Facial Emotional Expression Recognition with Soft Computing Techniques

2004년 10월 30일

변증남, 김대진, 이상완
한국과학기술원 (KAIST)
인간친화 복지 로봇시스템 연구센터 (HWRS-ERC)
djkim@ctrsys.kaist.ac.kr
bien@kaist.edu

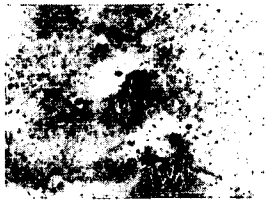
Contents

I. Introduction

FACE (顔)

What does it mean to you?

Faces



Mars Face

Monkey

Human

Monalisa

Paper Money (KRW)

[REF] http://images.google.co.kr/images?q=face&hl=ko&lr=&ie=UTF-8&newwindow=1&start=40&sa=N

Narcissism

[REF] http://blog.naver.com/fzoi74

Magic Mirror (from Snow White)

Snow White-The Movie, Walt Disney

[REF] http://www.animationusa.com/picts/wdpict/wkqst.gif

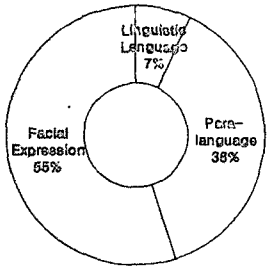
Princess Syndrom

[REF] http://blog.naver.com/sysisnom

Importance of Facial Expression

acial expression is very important and fundamental component in human-human interaction and/or human-computer interaction.

ommunication in natural way



[REF] A. Mehrabian, Communication without words, In Joseph A. DeVito, editor, Communication: Concepts and Process, pp 106-114, Prentice-Hall, Inc., 1971

From engineering point of view,

Case I. erving a Drink 1/3)

<Serving by Human Caregiver>

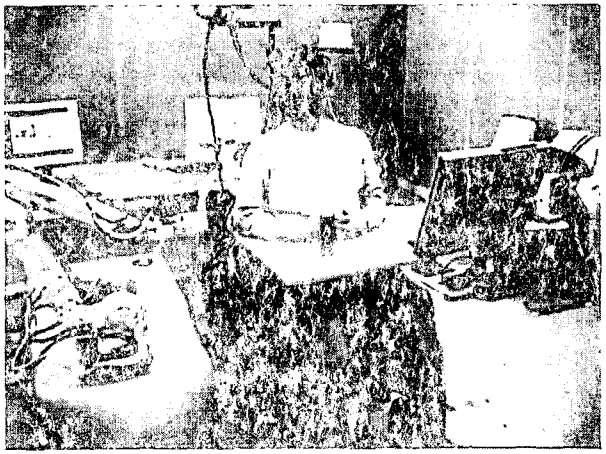
From engineering point of view,

Case I. erving a Drink 2/3)

<Advanced Serving is possible by
1) Head Shaking, and
2) Mouth Openning/Closing>

From engineering point of view,

Case I. erving a Drink 3/3)



From engineering point of view,

Case II. Service Robotic System

DOUMI Robot, HWRS-ERC, KAIST

PDA-based Facial Expression Recognition, HWRS-ERC, KAIST

From facial expression recognition,
Emotional/mental/physical/mental status monitoring is possible!

From engineering point of view,

Case III. Face-to-Face Interaction

Kismet

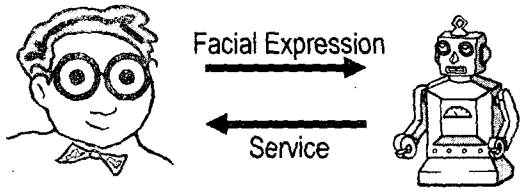
Regulating Interaction Intensity:
Face stimulus (human)

Cynthia Breazeal (Ferrell)
Brian Scassellati

MIT Artificial Intelligence Lab

[REF] http://www.ai.mit.edu/projects/sociable/videos.html

etter ervice via Facial Expressions



acial Expressions

1. Reading the intention of the user
2. Estimating the emotional/physical state of the user
3. Providing with emotional comfort for the user

⇒ In view of service provider (e.g. service robotic system), more intelligent and reliable services can be generated.

II. Human Face: Overview

Human Face (1/3)

hat is it? Anatomically speaking,

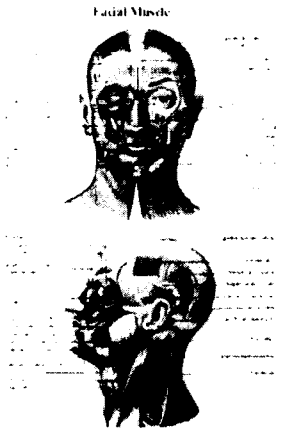
he most complex and expressive part (44 Muscles + 23 Bones)

pto 55,000 facial expressions!

haracteristics

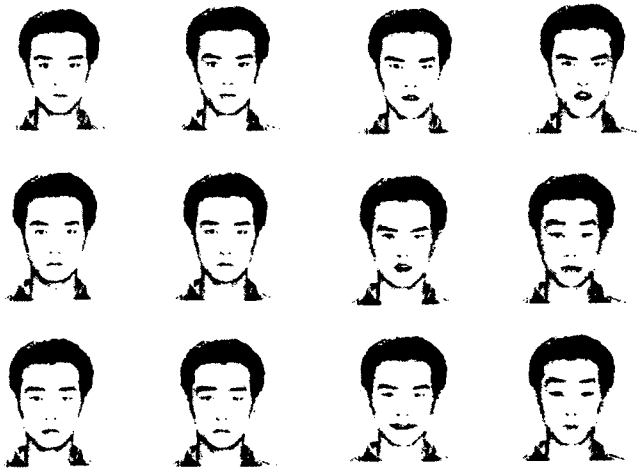
ue to its complex structure, it is very hard to model it in (classical) mathematical way.

he personality of each individual plays a key role to variate facial expressions.



[REF] G.-T. Park, "A Study on Extraction of Emotion from Facial Image using Soft Computing Techniques," Ph.D. Thesis, KAIST, 1999.

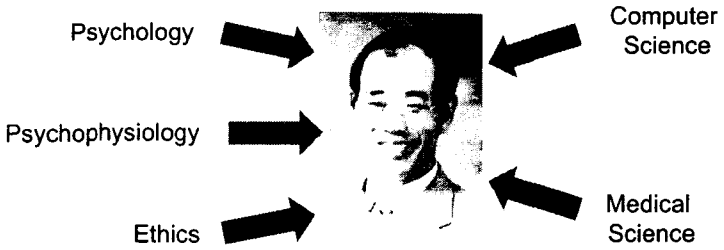
Human Face (2/3)



Artificially generated 12 face models (Max. 30)

[REF] C.-S. Choi, The Final Report on 'An Automatic Animation System for the Korean Sign Language,' Myongji Univ., Korea, 2003.

Human Face (3/3)



rom face, we, engineers, can

dentify who he/she is

stimate his/her emotional/physical/mental status

nderstand the meaning of expressions (intentions)

Engineering & Scientific Approaches for Face

ace Detection

ace Tracking

ace Identification (or Recognition)

acial Expression Recognition

acial Expression Generation (or Rendering)

ace Synthesis

ace Caricature

tc.

Face Detection

Where is the face?

Robitcs Institute, CMU

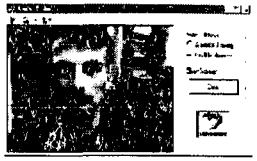
Face Group, Microsoft Research

[REF] http://www.ri.cmu.edu/projects/project_416.html
[REF] http://research.microsoft.com/~szli/Demos/default.asp

Face Tracking

Find & Track the Face!

Robotics Institute, CMU



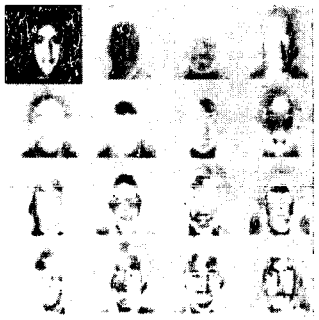
Nouse, Institute of Information Technology

R&CV Lab, KAIST

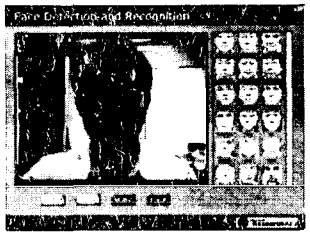
[REF] http://www.ri.cmu.edu/projects/project_448.html

Face Identification (or Recognition)

Who is he/she?



Media Lab, MIT



Face Group, Microsoft Research

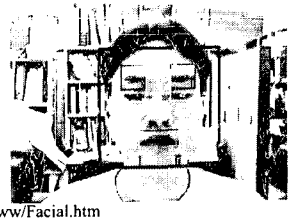
[REF] http://vismod.media.mit.edu/vismod/demos/facerec/basic.html
[REF] http://research.microsoft.com/~szli/Demos/default.asp

Facial Expression Recognition

What is the Meaning of Current Facial Expression?

Robotics Institute, CMU

MIS Lab, ATR

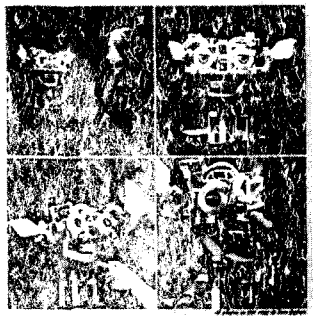


BSCL, KAIST

[REF] http://www-2.cs.cmu.edu/afs/cs/project/face/www/Facial.htm
[REF] http://www.mis.atr.co.jp/~mlyons/facial_expression.html
[REF] http://www.ee.kaist.ac.kr/lab/bscl/

Facial Expression Generation

Generate Facial Expressions!



KISMET, MIT

Facerobot, Tokyo Univ. of Science

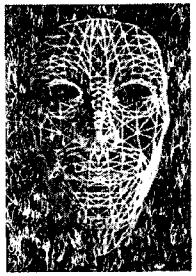
ULKNI, KAIST

[REF] http://www.samogden.com/Kismet.html

KFIS 추계학술대회 2004

Face Synthesis

Synthesize the Face!



K. Waters, HP

V. Blanz, Max-Planck Institute

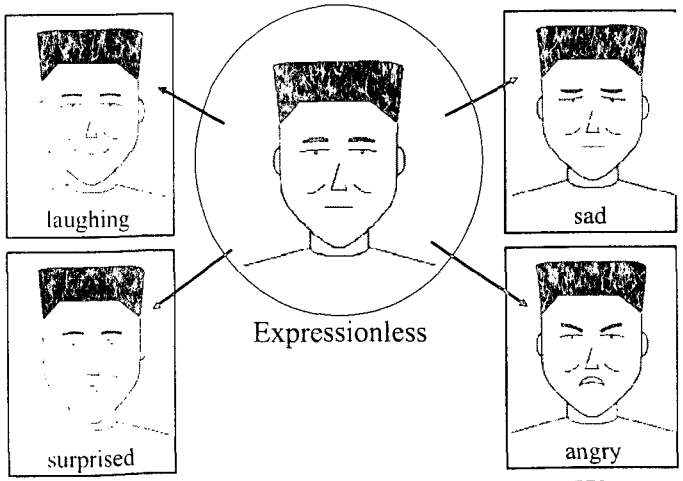
Movie Star, William Holden

[REF] http://www.crl.hpl.hp.com/projects/facial/facialdoc.html
[REF] http://www.kyb.tuebingen.mpg.de/bu/people/volker/
[REF] http://24.69.228.59/web_pics/images/users/larry/movie_stars/moviestars/william_holden.jpg

KFIS 추계학술대회 2004

Face Caricature

Express the Model impressions!



Prof. Onisawa, Univ. of Tsukuba

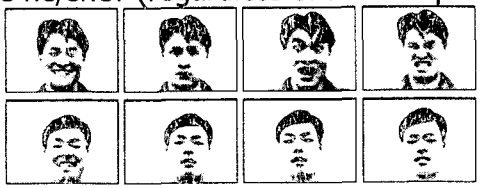
[REF] www.brain.riken.go.jp/labs/ lbis/member/iwas/thesis/FUZZ-IEEE01.ppt

KFIS 추계학술대회 2004

Face Recognition vs. Facial Expression Recognition

Face Recognition:

Who is he/she? (regardless of facial expressions)

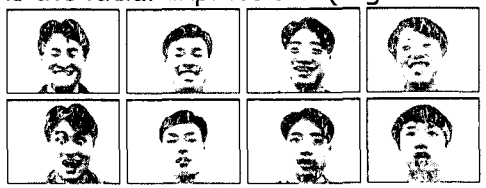


Mr. A

Mr. B

Facial Expression Recognition:

What is the facial expression? (regardless of his/her ID)

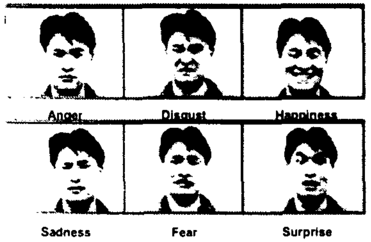


Happy

Surprise

KFIS 추계학술대회 2004

Target of Facial Expression Recognition: Six universal facial expressions



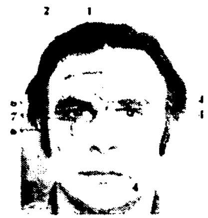
Emotion	Observed facial cues
Surprise	brows raised (curved and high) skin below brow stretched horizontal wrinkles across forehead eyelids opened and more of the white of the eye is visible jaw drops open without tension or stretching of the mouth
Fear	brows raised and drawn together forehead wrinkles drawn to the center upper eyelid is raised and lower eyelid is drawn up mouth is open lips are slightly tense or stretched and drawn back
Disgust	upper lip is raised lower lip is raised and pushed up to upper lip or is lowered nose is wrinkled cheeks are raised lines below the lower lid, lid is pushed up but not tense brows are lowered, lowering the upper lid
Anger	brows lowered and drawn together vertical lines appear between brows lower lid is tensed and may or may not be raised upper lid is tense and may or may not be lowered due to brows' action eyes have a hard stare and may have a bulging appearance lips are either pressed firmly together with corners straight or down or open, tensed in a squarish shape nostrils may be dilated (could occur in sadness too) [illegible] only if registered in all three facial areas
Happiness	corners of lips are drawn back and up mouth may or may not be parted with teeth exposed or not a wrinkle runs down from the nose to the outer edge beyond lip corners cheeks are raised lower eyelid shows wrinkles below it, and may be raised but not tense crow's-feet wrinkles go outward from the outer corners of the eyes
Sadness	inner corners of eyebrows are drawn up skin below the eyebrow is triangulated, with inner corner up upper lid inner corner is raised corners of the lips are drawn or lip is trembling

[REF] Y. Yacoob et al., "Recognizing human facial expression," TR-706, CVL-CAR, Univ. of Maryland, 1994.

KFIS 추계학술대회 2004

Well-known Psychological Work on Facial Expressions: FACS(Facial Action Coding System)

- eveloped by Ekman and Friesen in 1978
- rovides an objective description of facial signals in terms of component motions, or acial actions
- 6 Action Units(AUs)
- pplications
  - ehavioral science, medicine,
  - human-computer interactions
- ifficulty in coding
  - t takes over 100 hours of training to achieve minimal competency on FACS.

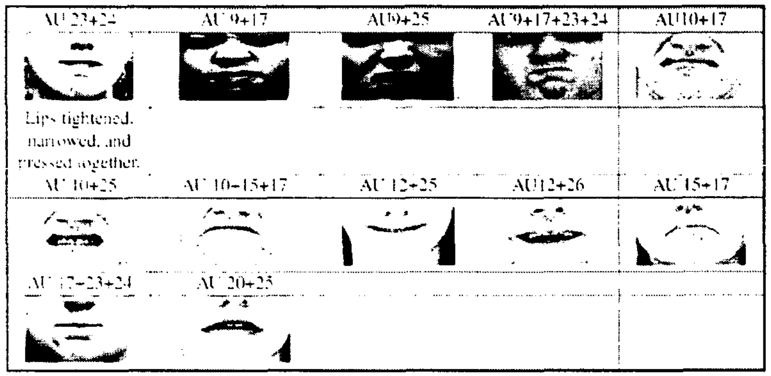


[REF] P. Ekman et al., "Facial Action Coding System," HIL, 1978.

KFIS 추계학술대회 2004

Example of FACS

- U(Action Unit) in Lower Face
  - ombinations of AUs

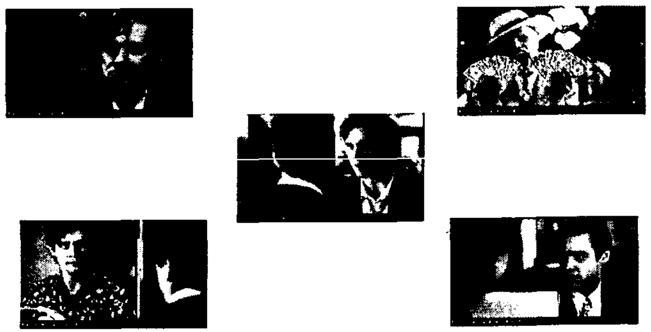


[REF] Y. Tian et al., "Recognizing action units for facial expression analysis," IEEE PAMI 23(2): 97-115, 2001.

KFIS 추계학술대회 2004

cf. other Facial Expressions

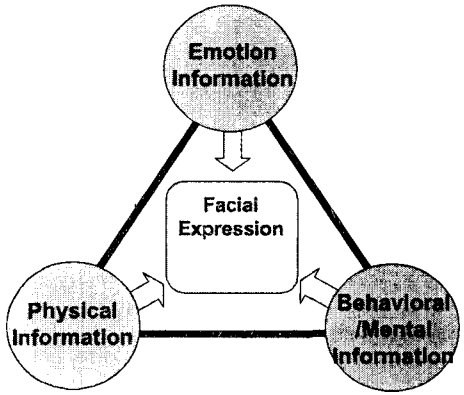
quinting, Empathy, Displeasure, Recall, Worry, Pouting, Sobbing, Weeping, Grief, Tenderness, …



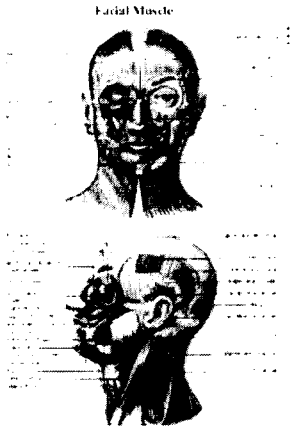
From the film "The Mask", 1994.

KFIS 추계학술대회 2004

Facial Expression as a Combination of Three Kinds of Information



Facial Expression as Human Biosign



Is he happy?

If yes, how much?

How do you know?

Isn he surprise?

[REF] A. Silverstein, "Human Anatomy and Physiology", John Wiley & Sons, 2nd edition, 1983.

Facial Expression as Human Biosign: What it is difficult to extract some information?

haracteristics of Human Biosign

- igh-dimensionality
- mbiguity
- nconsistency / Locality
- usceptibility to noise
- ubjectivity
- •
- •
- •

⇒ Mathematical Modeling is difficult !

OBSERVATION

by L. Zadeh

- *Machines are driven by measurements*
- *Humans are driven by perceptions*
- *To enable a machine to mimic the remarkable human capability to perform a wide variety of physical and mental tasks using perception-based information, it it necessary to have a means of converting measurements into perceptions*

LAZ 2-24-03

Soft Computing Techniques
= Computational Intelligence Technique

hat? (L. Zadeh)

set of approaches of computing which parallels the remarkable ability of the human mind to reason and learn in an environment of **uncertainty, imprecision and partial truth.**

uzzy Logic
rtificial Neural Networks
enetic Algorithm
ough Set Theory

uiding principle

xploit the tolerance for imprecision, uncertainty and partial truth to achieve tractability, robustness and low solution cost.

[REF] http://www.iiasa.ac.at/Research/DAS/dmsd/abo1/tsld022.htm

Human Biosign vs. Soft Computing (CI) Techniques

Characteristics	Capabilities	Techniques
Ambiguity, Susceptibility	Linguistic modeling	Fuzzy Inference Logic
Ambiguity, Inconsistency, Susceptibility	Generalization by Learning	Neural Network
High-dimensionality, Inconsistency	Optimization	Genetic Algorithm (Evolutionary Computation)
High-dimensionality	Rule reduction	Rough Set

III. Facial Expression Recognition:
Brief Survey for conventional (non-SCT) methods and SCT-based methods

*SCT: Soft Computing Techniques

Well-known non-SCT Method for Facial Expression Recognition (~1999)

Correlation-based optic flow

ocal velocity extraction

lassification Procedure

earest neighbor & Template matching

erformance

uccess rate: 53.1%~85.6%

[REF] G.Donato, M.S.Bartlett, J.C.Hager, P.Ekman and T.J.Sejnowski, lassifying Facial Actions," IEEE Tr. on PAMI, vol.21, no.10, oct.1999

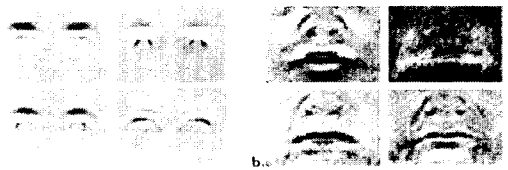
Principal Component Analysis: igenActions"

nsupervised learning method based on the second-order dependencies among the pixels.

pplications

- ecognizing facial identity
- lassifying gender
- ecognizing facial expressions

uccess rate: 79.3%



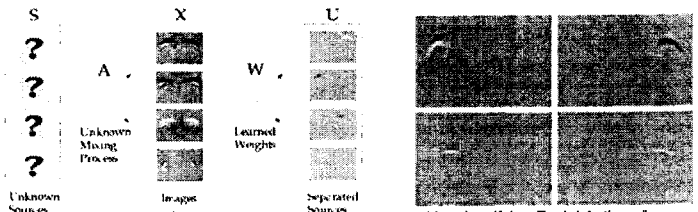
[REF] G.Donato, M.S.Bartlett, J.C.Hager, P.Ekman and T.J.Sejnowski, lassifying Facial Actions," IEEE Tr. on PAMI, vol.21, no.10, oct.1999

Independent Component Analysis

nsupervised learning algorithm derived from the principle of optimal information transfer between neurons.

aximizes the mutual information between the input and the output of a nonlinear transfer function.

uccess rate: 95.5%



[REF] G.Donato, M.S.Bartlett, J.C.Hager, P.Ekman and T.J.Sejnowski, lassifying Facial Actions," IEEE Tr. on PAMI, vol.21, no.10, oct.1999

Facial Expression Recognition using Continuous Dynamic Programming (2001)

nput Feature

acial Motion Graph (FMG)

⇒ **Too Time-consuming Process**

< Facial Motion Graph : FMG >

lassification

ontinuous Dynamic Programming

$$d(t,\tau) = \| G_A(t) - G_B(\lambda) \|$$

lassification Result

- 4 Facial Expressions (Angry, Disgust, Happiness, Sad)
- 15 training data, 21 test data (including some training data)
- Average Rate: 95%

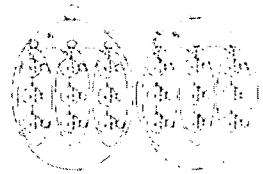
⇒ **Too Small Dataset**

[REF] Haihong Zhang and Yan Guo, "Facial Expression Recognition Using Continuous Dynamic Programming," IEEE ICCV Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, 2001.

Facial Expression Recognition using Pseudo 3-D Hidden Markov Models (2002)

seudo 3-D Hidden Markov Models

- Extension of 1-D HMM paradigm
- Concept of yperstates'
- Triple embeded Viterbi Algorithm



< Pseudo 3-D HMM >

lassification Result

- Varying number of Gaussian mixtures
- 75.0% (1 mixture, rank 1)
- 95.8% (4 mixtures, rank 2)

⇒ **Too Time-consuming Process**

< Equivalent 1-D HMM >

[REF] Stefan Muller et al., "Facial Expression Recognition Using Pseudo 3-D Hidden Markov Models," 16th International Conference on PatternRecognition (ICPR'02),2002.

Summary for Facial Expression Recognition with non-SCT methods

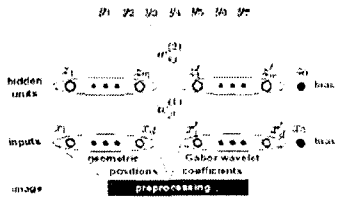
	Method (Author)	# of Expressions	Accuracy (%)
Non-SCT	Optical Flow Histogram (Dai)	7	-
	Continuous Dynamic Programming (Zhang)	4	95
	HMM chains (Hoey)	5	80-92
	HMM using MPEG4 FAP (Pardas)	6	83.6
	Pseudo 3-HMM (Muller)	6	75-95.8
	State-based Dynamic Model (Bourel)	6	72
	Tree-Augmented-Na e Bayes:TAN (Cohen)	6	65.5-83.6
	Linear Programming with Gabor Filter (Guo)	6	91
	Line-based Disparity Measure (Gao)	3	86.6

87.4% for 5.4 expressions

ue to geometrical relations, it often takes too long time to calculate the necessary features.
n-line learning is almost impossible.
ot robust to variation of the input data

Comparison Between Geometry-Based and Gabor-Wavelets-Based Facial Expression Recognition using Multi-Layer Perceptron (1998)

- Input Features
  - Geometric position: 34 fiducial points (manually selected)
  - Set of Gabor Wavelets : 3 freq. × 6 theta (fixed) = 18 kernels
- 2-layer perceptron / 7 expressions
- Learning by fast error-backpropagation
- Classification Result
  - 90.1% (7 expressions)
  - 92.3% (excluding 'Fear')



⇒ **Manually Selected Features**

[REF] Zhengyou Zhang, Michael Lyons, Michael Schuster and Shigeru Akamatsu. 1998. Comparison Between Geometry-Based and Gabor-Wavelets-Based Facial Expression Recognition Using Multi-Layer Perceptron. Proceedings, Third IEEE International Conference on Automatic Faceand Gesture Recognition. pp. 454-459.

Recognizing Upper/Lower Face Action Units for Facial Expression Analysis(2000)

- Feature Extraction
  - Permanent features : Eye / Brow / Cheek / Lips
  - Transient feature : Furrow (edge & thresholding by heuristics)

<CMU - Prof. Kanade's LAB>
http://www-2.cs.cmu.edu/~face/

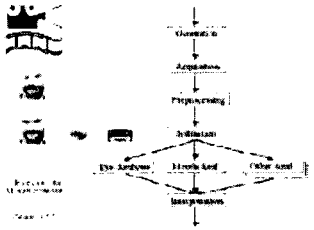
- Recognition
  - Neural Network Model (1 hidden)
  - Average Recognition Rate:
    93.7%(Upper)
    96.7%(Lower)

⇒ **Just for AU, not Facial Expression**

[REF] Ying-li Tian, Takeo Kanade and Jeffrey F.Cohn. 2000. Recognizing Upper Face Action Units for Facial Expression Analysis. Computer Vision and Pattern Recognition (CVPR'00)
[REF] Ying-li Tian, Takeo Kanade and Jeffrey F.Cohn. 2000. Recognizing Lower Face Action Units for Facial Expression Analysis. Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition (FG'00)

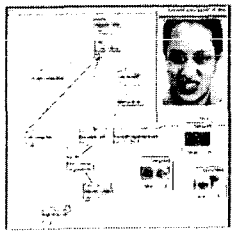
Expert System for Automatic Analysis of Facial Expressions (2000)

- Facial Feature Detection
  - Single detector mode
    - Selectable/Connectable by the User
  - Multi-detector mode
    - Each and every facial feature detector integrated into the workbench is invoked automatically.



- Classification Result
  - 6 Expressions
  - Average Rate: 86.3%

⇒ **Not User-friendly Interface**



[REF] M. Pantic, L.J.M. Rothkrantz. 2000. Expert system for automatic analysis of facial expressions. Image and Vision Computing 18 (2000) 881-905.

Comparison between SCT methods and non-SCT methods in Facial Expression Recognition

	Method (Author)	# of Expressions	Accuracy (%)	
Non-SCT	Optical Flow Histogram (Dai)	7	-	
	Continuous Dynamic Programming (Zhang)	4	95	
	HMM chains (Hoey)	5	80-92	
	HMM using MPEG4 FAP (Pardas)	6	83.6	
	Pseudo 3-HMM (Muller)	6	75-95.8	87.4% for 5.4 expressions
	State-based Dynamic Model (Bourel)	6	72	
	Tree-Augmented-Na e Bayes:TAN (Cohen)	6	65.5-83.6	
	Linear Programming with Gabor Filter (Guo)	6	91	
	Line-based Disparity Measure (Gao)	3	86.6	
SCT	MLP (Christine)	6	-	
	Geometry + Gabor-wavelet MLP (Zhang)	7	90.1-92.3	
	Expert System (Pantic)	6	86.3	86.8% for 6.4 expressions
	MLP & RBFN (Gargesha)	6	65-73	
	Convoultional NN (Fasel)	7	87.1-91.4	
	Upper/Lower Facial Action Unit NN (Tian)	- (AU-based)	93.7-96.7	
	PCA-based SVM (Kapoor)	- (AU-based)	62.5-81.2	

Discussions

omparing with other non-SCT, SCT enable us to deal with various kinds of uncertainty and complex aspects due to;

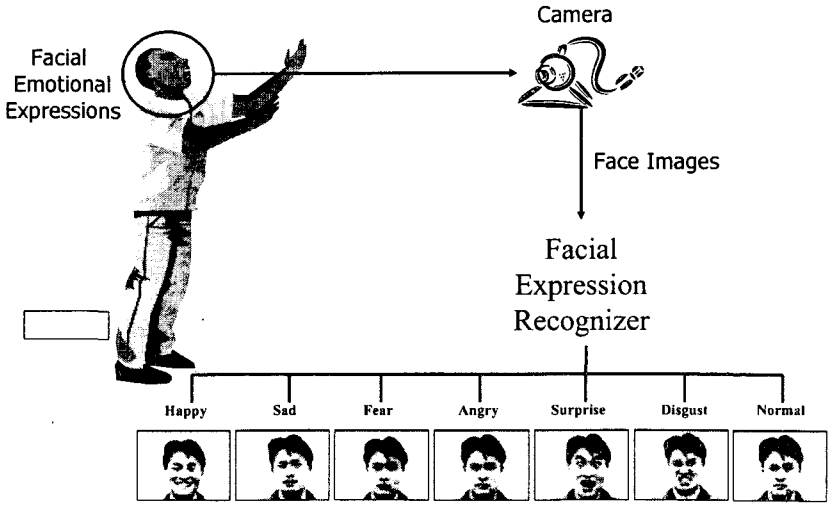
1. **Robustness to the noisy input data**
2. **On-Line Learning Capability**

owever, shortcomings of previous SCT-based works are still remained such as;

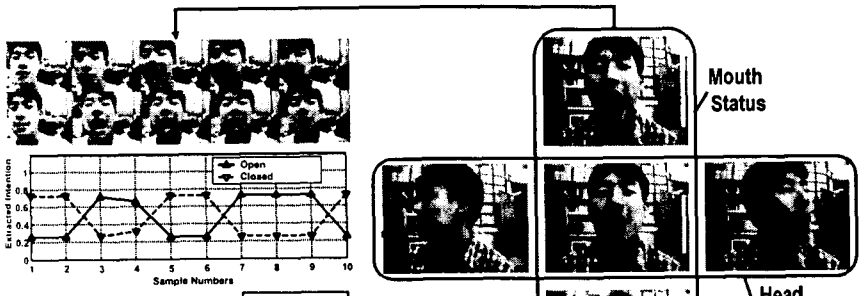
1. The previous works **didn effectively use the descriptive (linguistic) results given by Ekman et al.**
2. The previous works can deal with **the individual characteristics for each subject.**
3. Stick to one specific method of SCT, not **Toolbox-based Approach (fusion of various techniques)**

IV. Our Experiences on Facial Expression Recognition with Soft Computing Techniques

Assumption: Recognition is limited to one of 7 facial emotional expressions.



Intention Reading Result (1)



	Outputs: Open	Outputs: Closed
Inputs: Open	30/35	5/0
Inputs: Closed	1/0	74/75

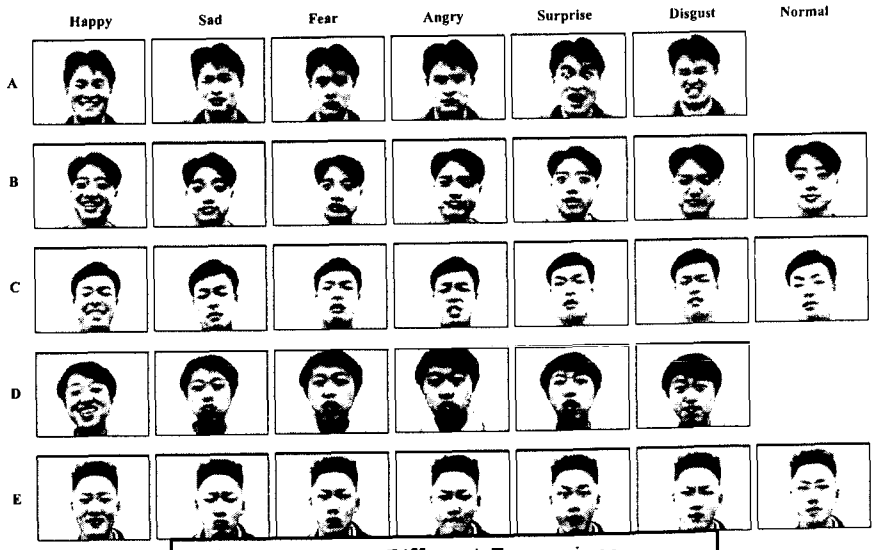
Success Rate: 92.2 %

Intention Reading Result (2)

IV-C. Facial Emotional Expression Recognition System

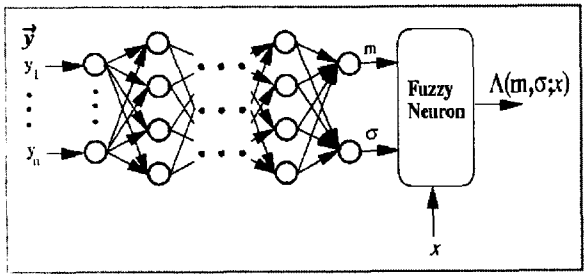
Personalized Classifier using Fuzzy Neural Networks (D.-J. Kim, 2004)

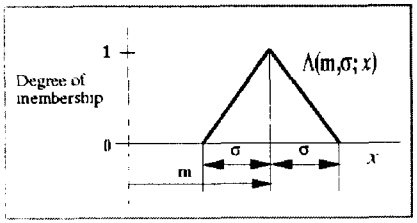
D.-J. Kim and Z. Bien, "A 'Personalized' Facial Expression Recognition with Fuzzy Similarity Measure and Novel Feature Selection Method," Proc. of FUZZ-IEEE2004, 2004.



Different Persons, Different Expressions
→ Difficult to model it in the unified viewpoint!

Structure of Fuzzy Observer

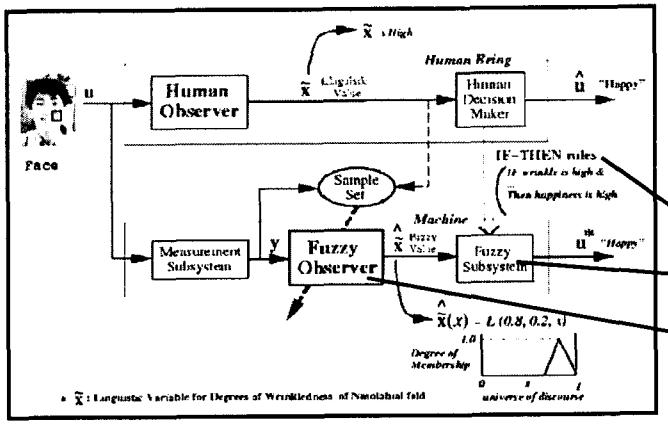




$\hat{\tilde{x}}(x) = \Lambda(m,\sigma\,;x)$

$\hat{\tilde{x}}(x)$: estimated feature vector in linguistic form

Human Observer vs. Fuzzy Observer



Soft computing techniques

Fuzzy Rule Set

pper face

Var	Ner	Sur	Fear	Dis	Ang	Hap	Sad
u1	M	H	H	L	L		M
u2	H	H	L	H	L		L
u3	L	L	H	L	L		H

- Ex) If $u_1 = H$ & $u_2 = H$ & $u_3 = L$ Then Surprise is *H*
  Otherwise Surprise is *L*
- No rule in Happiness

iddle face

Var	Ner	Sur	Fear	Dis	Ang	Hap	Sad
m1	M	H	M	L	L	L	L
m2	M	M	M	H	M	H	M
m3	M	M	M	M	M	H	M
m4	M	M	M	M	H	M	L

- Ex) If $m_1 = L$ & $m_2 = H$ & $m_3 = H$ & $m_4 = M$ Then Happiness is *H*
  Otherwise Happiness is *L*

ower face

Var	Ner	Sur	Fear		Dis	
l1	L	H	H	M	H	L
l2	M/L	M	M	M	M	M/L
l3	M	M	H	H	M	M
l4	L	H/M	H	H	L	L
l5	M	H	M	M	M	M
l6	M	M	H	H	M	VH
l7	L	L	L	L	L	H

Var	Ang		Hap			Sad
l1	L	VH	L	M	H/VH	L
l2	L	H	M/L	M	M/L	L
l3	M	M	VH	M	M	L
l4	L	L	L	M	M	L
l5	M	H	M	M	H	M
l6	M	M	M	M	M	M
l7	L	L	L	L	L	L

■ Ex)
If $l_1 = L$ & $l_2 = M$ & $l_3 = VH$ & $l_4 = L$ & $l_5 = M$ & $l_6 = M$ & $l_7 = L$ Then Happiness is *M*
If $l_1 = L$ & $l_2 = L$ & $l_3 = VH$ & $l_4 = L$ & $l_5 = M$ & $l_6 = M$ & $l_7 = L$ Then Happiness is *M*
Otherwise degree of Happiness is *L*

Automatic Rule Reduction

using Rough Set Theory

4 input linguistic variables

n each facial region, rules are constructed as follows :

x3x2 = 18 (upper face)

x2x2x3 = 36 (middle face)

x3x4x3x2x4x2 = 2,304 (lower face)

o. of rules in an integrated rule form

ogical AND of all facial regions :

8x36x2,304 = 1,492,992! Too many rules!

- **That is, Ekman original linguistic description requires too big memories!**

o. of rules in the decomposed rule set

ogical OR of all facial regions :

8+36+2,304 = 2,358!

Key Observation for Personalization

he number of efficient features and their values depend on person to person.

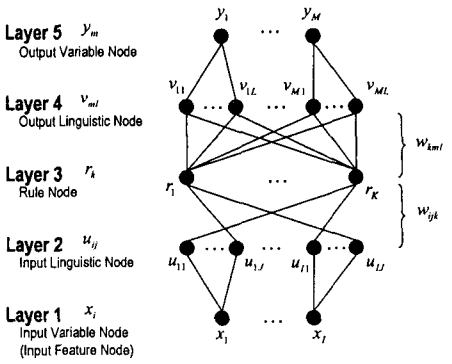
olution

- *Measure the importance of each feature to the output*
- *Select the efficient feature according to the above measured values*

Connection Degree & Histogram

for Measure the Importance of A Feature

in FNN structure



connection degree

$$F(u_{ij}, v_{ml}) = \frac{1}{(J)^{I-1}} \sum_{k=1}^{K} h(u_{ij}, v_{ml}, r_k)$$

$$h(u_{ij}, v_{ml}, r_k) = w_{ijk} \times w_{kml}$$

istogram of connection degree

$$G(i,j) = \frac{1}{M} \left| \sum_{m=1}^{M} \sum_{l=1}^{L} g(F(u_{ij}, v_{ml})) F(u_{ij}, v_{ml}) \right|$$

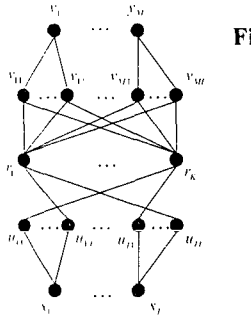
$$H(i,j,m) = \left| \sum_{l=1}^{L} g(F(u_{ij}, v_{ml})) F(u_{ij}, v_{ml}) \right|$$

$$g(F(u_{ij}, v_{ml})) = 2\frac{l-1}{L-1} - 1$$

$G(i,j)$ Effect of u_{ij} to the whole outputs

$H(i,j,m)$ Effect of u_{ij} to the specific output y_m

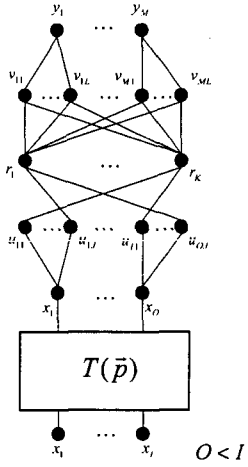
Problem Formulation of Feature Selection/Modification



Find $T(\vec{p}) \in \Re^{O \times I}$ **s.t.** $E' < E$

$$E = \frac{1}{MS} \sum_{s=1}^{S} \sum_{m=1}^{M} \left\| \vec{y}_s^m - C(\vec{x}_s^m, \vec{p}) \right\|$$

$$E' = \frac{1}{MS} \sum_{s=1}^{S} \sum_{m=1}^{M} \left\| \vec{y}_s^m - C(\vec{x}_s^{m\prime}, \vec{p}') \right\|$$



Feature Selection by Determining $T(\vec{p})$

$O < I$

Feature Selection/Modification Procedure

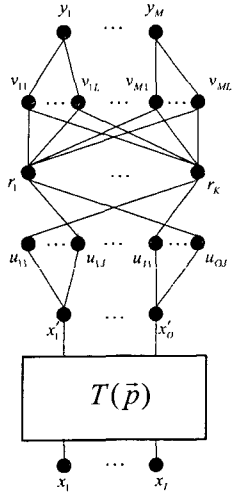
for Personalized Classifier

```
procedure FeatureSelection{G(i,j), H(i,j,m), x_i, u_ij, y_m}
begin {FeatureSelection}
Phase 1
  for i = 1,...,I,
    index(i) = 1, C = 0
    for j = 1,...,J,
      if G(i,j) < γ1 G(i,j)^max,
        D = 0
        for m = 1,...,M,
          if H(i,j,m) ≤ γ2 H(i,j,m)^max and e_m ≥ γ3,
            delete (or prune) node u_ij, D = 1
        C = C + D
    if C == J,
      delete (or prune) node x_i, index(i) = 0
Phase 2
  for o = 1,...,O,
    for i = 1,...,I,
      t_oi(p) = 0
      if index(i) == 1
        t_oi(p) = 1
end {FeatureSelection}
```

$G(i,j)^{\max} = \max_{i,j} G(i,j)$

$H(i,j,m)^{\max} = \max_{i,j,m} H(i,j,m)$

$\gamma_1, \gamma_2, \gamma_3$: user-defined parameters



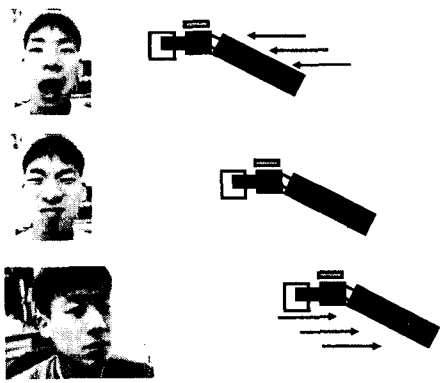
Intention Reading by Vision Sensor

erving a beverage to the user

outh Openning (Approach)

outh Closing (Stop)

ead Shaking (To keep away)



Fuzzy Rule Base

ositive/Negative intention reading

Mouth Openness **Positive/Negative Intention**

IF (Premise)	THEN (Consequent)
Mouth Openness is High	Positive is High (which means that Negative is Low)
Mouth Openness is Low (or Head Shaking)	Positive is Low (which means that Negative is High)

Image Features (1)

: Mouth Openness

eometrical Features (Global)

n area ratio of the lip part to the face

height ratio of the lip part to the face

G (Gabor-Gaussian) Feature (Local)

abor filter response of the lip part

$$f_G = \frac{\sum_{i=1}^{H-1}\{w_G(i)\cdot dy_{proj}(i)\}}{\sum_{i=1}^{H-1} w_G(i)}$$

$w_G(i)$: Gaussian weights

$dy_{proj}(i)$: Absolute value of the derivative of the projected value $y_{proj}(i)$

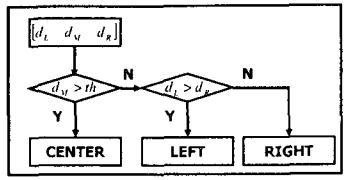
H : Height of a Gabor-filtered image

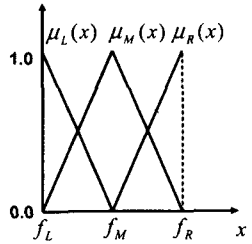
... W.-K. Song, J. S. Han and Z. Zenn Bien, "Soft Computing based Intention Reading Techniques as a Means of Human-robot ... for Human Centered System," Journal of Soft Computing 7, pp.160-166, 2003.

Image Features (2)

:Head Shaking

acial Aspect (R./L./C.)

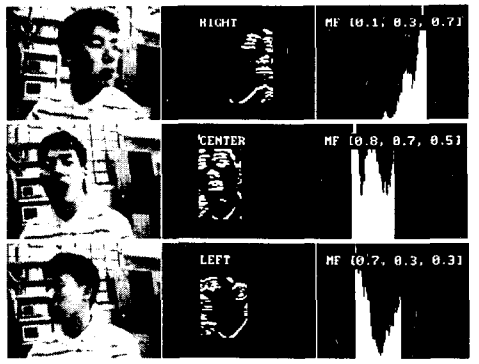




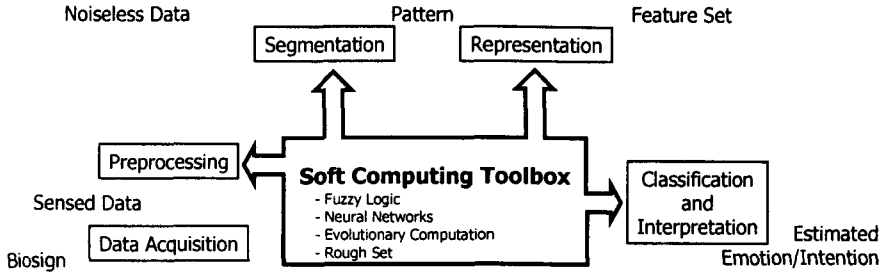
$$d_L = \sum_{x=f_L}^{f_M} \bar{I}_{xproj}(x)\cdot\mu_L(x)$$

$$d_M = \sum_{x=f_L}^{f_R} \bar{I}_{xproj}(x)\cdot\mu_M(x)$$

$$d_R = \sum_{x=f_M}^{f_R} \bar{I}_{xproj}(x)\cdot\mu_R(x)$$



General architecture of soft computing-based human biosign recognition system: Toolbox-based Approach



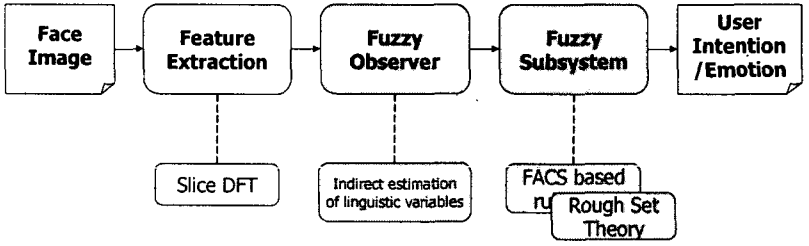
[REF] Z. Zenn Bien, Jung-Bae Kim, Dae-Jin Kim, Jeong-Su Han and Jun-Hyeong Do "Soft computing-based Emotion/Intention Reading for Service Robot", Submitted to AFSS 2002 (International Conference on Fuzzy Systems), pp.121-128, India, Feb. 3-6, 2002.

IV-A. Facial Emotional Expression Recognition System

Fuzzy Observer-based Approach (G.-T. Park, 2000)

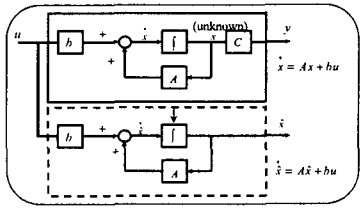
G.-T. Park and Z. Bien, "Neural Network-based fuzzy observer with application to facial analysis," Pattern Recognition Letters, vol. 21, pp. 93-105, Feb., 2000.

Overall Structure of the System

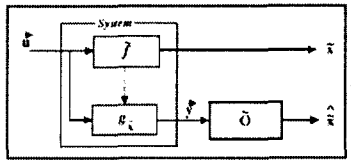


Design of Fuzzy Observer

onventional state observer



ndirect estimation of a linguistic variable from conventionally measured data



$\tilde{x} = \tilde{f}(\vec{u})$

$\vec{y} = g_{\tilde{x}}(\vec{u})$

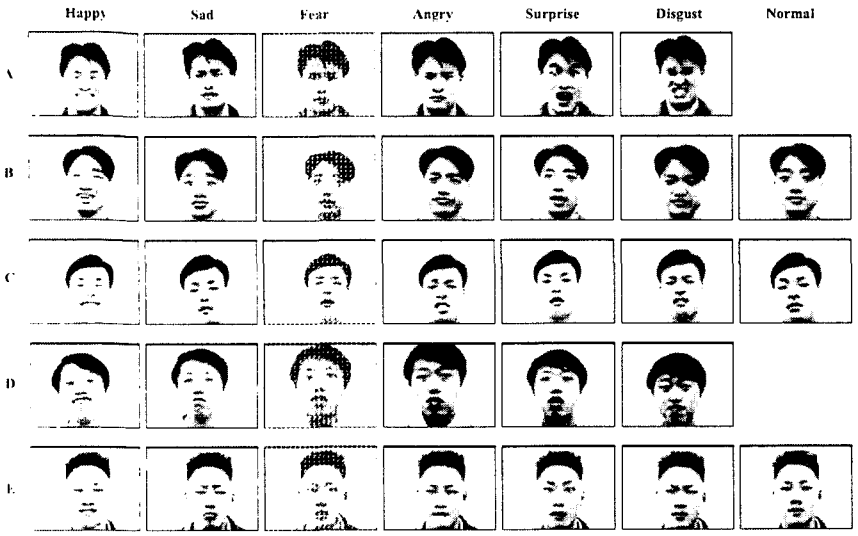
$\hat{\tilde{x}} = \tilde{O}(\vec{y})$

$\vec{u}$: input feature vector in numerical form

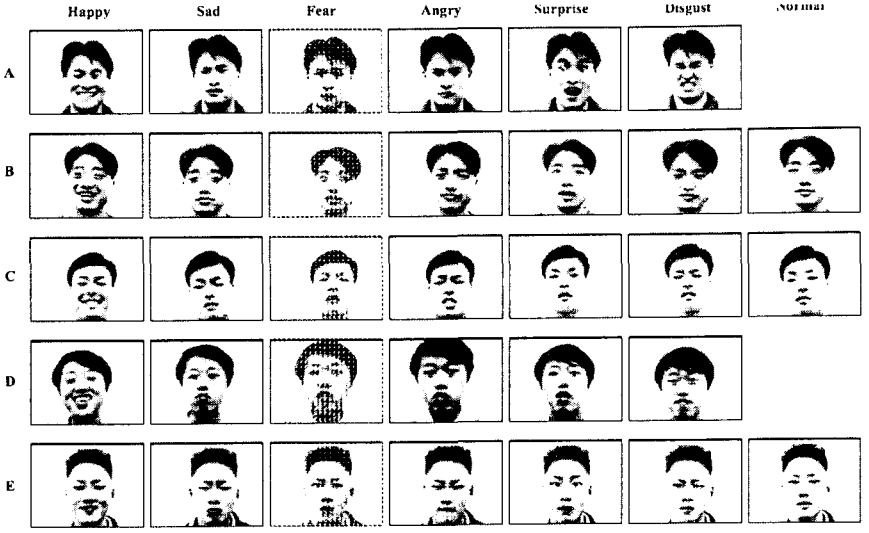
$\tilde{x}$: feature vector in linguistic form

$\vec{y}$: output vector in numerical form

$\hat{\tilde{x}}$: estimated feature vector in linguistic form



(conventional) 'Facial Expression Recognition'



(proposed) 'Personalized Facial Expression Recognition'

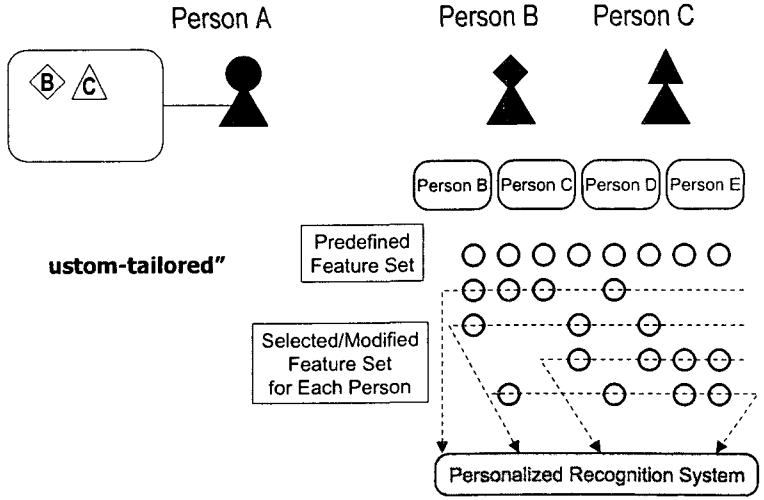
Personalized Expression
= Customized Expression

➡ Ownership is possible.

Related Work for Personalized Facial Expression Recognition:

J. F. Cohn, K. Schmidt, R. Gross and P. Ekman, "Individual Differences in Facial Expression: Stability over Time, Relation to Self-Reported Emotion, and Ability to Inform Person Identification," IEEE International Conference on Multimodal Interfaces (ICMI 2002), 2002.

How to Personalize?
by "Feature Selection/Modification"



Automatic Rule Reduction (cont)

using Rough Set Theory

Rule reduction for surprise in upper face

Conditional			Decision
u1	u2	u3	Surprise
H	H	L	H
H	H	H	L
H	L	H	L
H	L	L	L

18 rules

Rough Set Theory ⇩

Conditional			Decision
u1	u2	u3	Surprise
H	H	L	H
x	x	H	L
x	L	L	L

o. of rules after rule reduction using rough set theory: 221

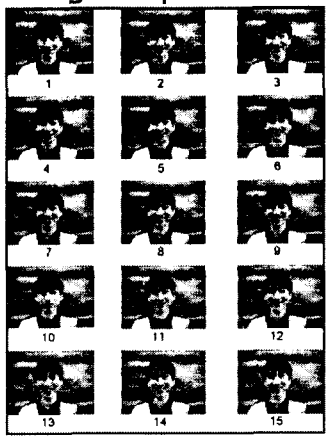
Emotion	no. of reduced rules	core	reducts
Surprise	6/7/16	all/all/all	all/all/all
Fear	6/7/16	all/all/all	all/all/all
Disgust	6/7/18	all/all/all	all/all/all
Anger	6/7/18	all/all/all	all/all/all
Happiness	x/7/35	x/all/all	x/all/all
Sadness	6/7/16	all/all/all	all/all/all
Neutral	7/7/16	all/all/all	all/all/all

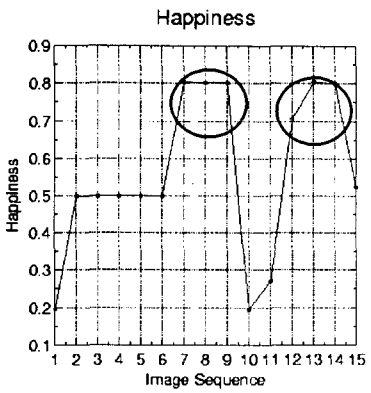
Total sum of rules = 221

6 rules

Result for Happy Face

mage sequence of 15 frames





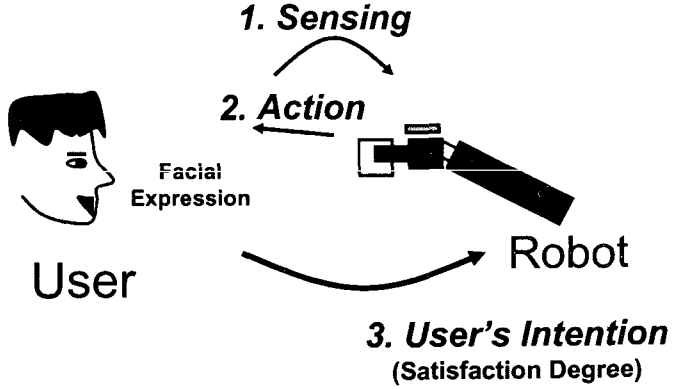
IV-B. Facial Emotional Expression Recognition System

Human-friendly Robot via Intention Reading (D.-J. Kim, 2002)

D.-J. Kim and Z. Bien, "Soft Computing-based Intention Reading through the User's Mouth for Human-friendly Human-Robot Interaction," Proc. of SCIS&ISIS2002, Tsukuba, Japan, 2002.

Scenario: Serving a Drink

or human-friendly interaction, reading of human spontaneous intention is essential!



Setup

Inputs & 7 Outputs

egree of eye openness, degree of NLR, degree of NLF, degree of mouth openness, distance between eyebrows and eyes

appy, Sad, Fear, Angry, Surprise, Disgust and Neutral expression

ata Set

XP-DS#1 & EXP-DS#2 are taken from the Ekman DB by statistical variations.

$$z_i = x_i + 2\sigma(e^{-(x_i-0.5)^2} - 0.5), \qquad i = 1,\ldots,5$$

$$\vec{x} = [x_1,\ldots,x_5]^T \in \Re^5$$

$$\vec{z} = [z_1,\ldots,z_5]^T \in \Re^5 \qquad \sigma = 0.05$$

SCL DB consists of the facial images of 10 persons captured by USB WebCam.

Data set	# of persons	# of expressions	# of images	# of training images	# of test images
EXP-DS#1	14	5-7	9,400	4,700	4,700
EXP-DS#2	22	5-7	15,000	7,500	7,500
BSCL DB	10	7	7,000	3,500	3,500

Numerical Result:

Comparison with other pattern classifiers

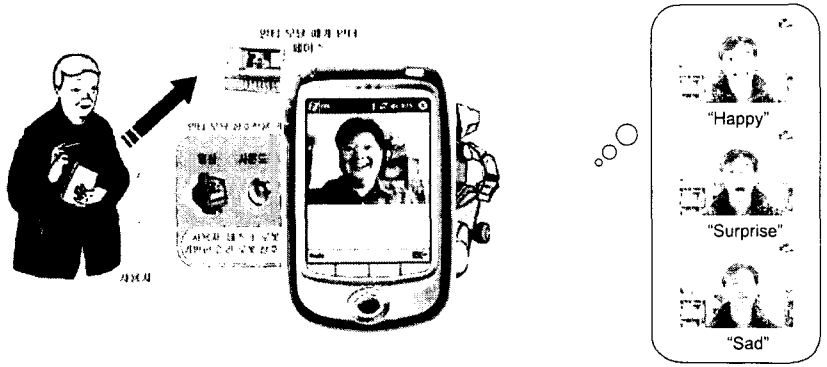
	Classification Rate (%)			
	NN-M	ANFIS	Lin FNN	Proposed
EXP-DS#1	73.1	52.9	35.4	94.3
EXP-DS#2	66.8	65.9	29.9	90.3
BSCL DB	26.3	43.2	16.3	78.0

IV-D. Facial Emotional Expression Recognition System

User Adaptive System
using Gabor Wavelet Neural Networks and Q-Learning
(S.-W. Lee, 2004)

S.-W. Lee, D.-J. Kim, K.-H. Park and Z. Bien, "Gabor Wavelet Neural Network-Based Facial Expression Recognition," Proc. of SCI2004, 2004.

Facial Expression Recognition (FER) for Human-Robot Interface

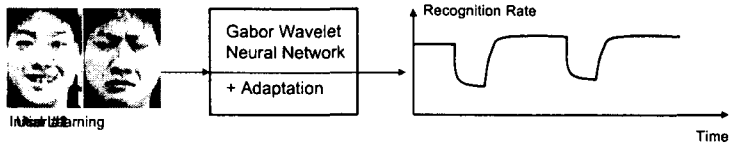


21C Frontier Project, evelopment of Vision Technology for Multi-Modal Intermediate Human-Robot Interface upported by Ministry of Science and Technology, Korea (2003.10.1~)

User Adaptation using Q-Learning and Fuzzy Neural Network Model

With these technique...

Facial Expression Recognition System



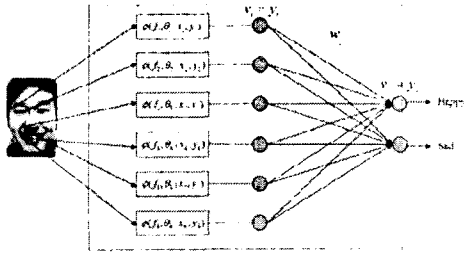
- A Classifier which has an adaptation capability even when another users try to use it.

➔ **GWNN**
+ Q-Learning
+ Fuzzy Neural Network Model

[REF] Yong Soo Kim, Chang Hyun Ham and Yong Sun Baek, "A Fuzzy Neural Network Model Solving the Underutilization Problem," Korea Fuzzy Logic and Intelligent Systems Society, Vol. 11, pp. 354-358.

Proposed Method: GWNN-based Classifier

Gabor Wavelet Neural Network (GWNN)

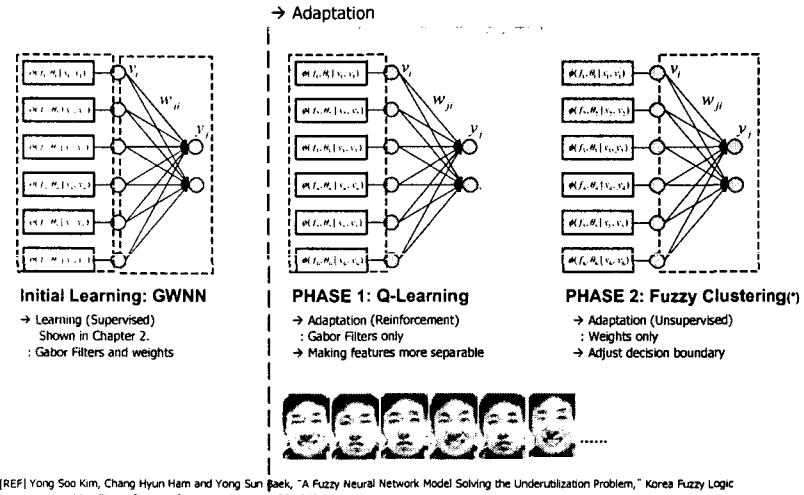


Merits

1. Integrating feature extraction and Classification process
2. Learning both weights and all Gabor Wavelet Filters

➔ **NO heuristics in feature extraction**

From Initial Learning to 2-Phases Adaptation



Initial Learning: GWNN

→ Learning (Supervised)
Shown in Chapter 2.
: Gabor Filters and weights

PHASE 1: Q-Learning

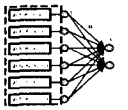
→ Adaptation (Reinforcement)
: Gabor Filters only
→ Making features more separable

PHASE 2: Fuzzy Clustering(*)

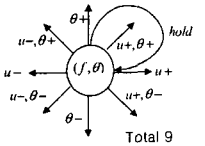
→ Adaptation (Unsupervised)
: Weights only
→ Adjust decision boundary

[REF] Yong Soo Kim, Chang Hyun Ham and Yong Sun Baek, "A Fuzzy Neural Network Model Solving the Underutilization Problem," Korea Fuzzy Logic and Intelligent Systems Society, Vol. 11, pp. 354-358.

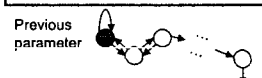
Adaptation Phase 1: Q-learning



State : parameter set
State Movement : Changing the parameter (f, θ)
Policy : When Unknown next input comes, parameters to be adjusted.
Action : Recognition using parameters of state
Reward : Reward from the user and the system itself.

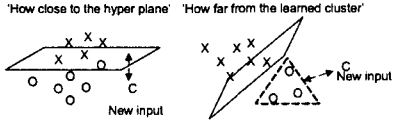


Total 9 actions
& 1020 states



Reward : $r = r_b + r_h |y_1 - y_2| - (|d_1 - y_1| + |d_2 - y_2|)$

'How close to the hyper plane' 'How far from the learned cluster'



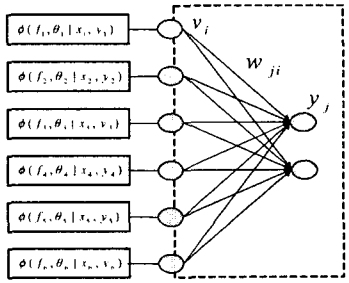
Adaptation Phase 2:

2-Layer Perceptron Model

Algorithm

1. Initialization: parameter setting
2. Competition: Input pattern is applied, competition among output neurons occurs in winner-take-all.
3. Vigilance Test
   : Distance btw the input pattern and candidate cluster.
   $c^{(w)} : \|x - v_i\| \le T$
4. Vigilance test → Fail
   : Current Input vector becomes New cluster (weights).
5. Vigilance test → Pass
   : Centroids (weights) are updated.

$$v_i^{(old)}(n+1) = v_i^{(old)}(n) + f(l)u_i^2(n)\Pi(x, v_i(t), T)(x - v_i^{(old)})$$

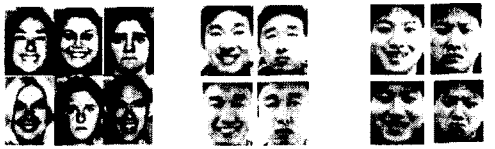


[REF] Yong Soo Kim, Chang Hyun Ham and Yong Sun Baek, "A Fuzzy Neural Network Model Solving the Underutilization Problem," Korea Fuzzy Logic and Intelligent Systems Society, Vol. 11, pp. 354-358.

Numerical Result:

Training DB & Test DB

	Training data	Test data	
	EKMAN DB	DB #1	DB #2
TOTAL TEST DATA	30 (15/15)	114 (53/61)	120 (60/60)
Specification	15 different people (To test generalization)	1 person (To test adaptation)	1 person (To test adaptation)



Numerical Result:

Adaptation Process is Effective to Improve the Classification Rate!

	Initial learning	Adaptation Process	
		Phase I	Phase II
USER#1 DB	100.0 %	(*)	(*)
USER#2 DB	87.5 %	89.15 %	93.35 %

(*) We can say that the learned system is favorably fit for current user, so the adaptation process is not necessary "at present".

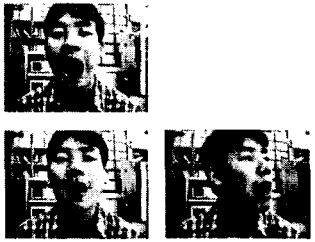
However, because the user's face changes, the recognition rate would fall down slowly, then the adaptation process will be activated.

V. Concluding Remarks

Summary of Our Works (1/2)

(1) Fuzzy Observer: Degree for happiness

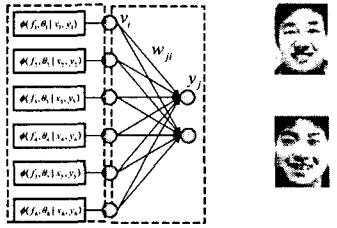
(2) Intention Reading by Robotic System: Mouth Openness/Head Shaking



Summary of Our Works (2/2)

(3) Personalized Classifier: Customized Classifier with Feature Selection

(4) Adaptive GWNN with Q-Learning: Adaptation for New User



Conclusions

acial expression is important information for engineering HiLS systems (service robots, intelligent sweet home and so on).

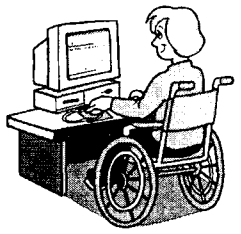
ue to its complexity, FER is not an easy task for the machine.

any FER methods with SCT emerged to show the effectiveness of SCT in FER task.

Further Works

ore effective & efficient adaptation and learning mechanism for facial expression recognition, applicable for many humans and/or for a life-long time.

mplementation in portable electronics such as PDA, tablet PC and so on for intelligent/human-friendly human-robot interaction system applications



Thank you

for your kind attention

Any Questions/Comments?

djkim@ctrsys.kaist.ac.kr
bien@kaist.edu

Thanks to:
Mr. Sang-Wan Lee
Mr. Young-Joon Kwon