

<http://dx.doi.org/10.7236/IIBC.2014.14.3.139>

IIBC 2014-3-20

# Active Shape Model과 통계적 패턴인식기를 이용한 얼굴 영상 기반 감정인식

## Video-based Facial Emotion Recognition using Active Shape Models and Statistical Pattern Recognizers

장길진\*, 조아라\*\*, 박정식\*\*\*, 서용호\*\*\*\*

Gil-Jin Jang\*, Ahra Jo\*\*, Jeong-Sik Park\*\*\*, Yong-Ho Seo\*\*\*\*

**요약** 본 논문에서는 얼굴 영상으로부터 자동으로 사람의 감정을 인식하는 효과적인 방법을 제안한다. 얼굴 표정으로부터 감정을 파악하기 위해서는 카메라로부터 얼굴영상을 입력받고, ASM (active shape model)을 이용하여 얼굴의 영역 및 얼굴의 주요 특징점을 추출한다. 추출한 특징점으로부터 각 장면별로 49차의 크기 및 변이에 강인한 특징벡터를 추출한 후, 통계기반 패턴분류 방법을 사용하여 얼굴표정을 인식하였다. 사용된 패턴분류기는 Naive Bayes, 다중계층 신경회로망(MLP; multi-layer perceptron), 그리고 SVM (support vector machine)이며, 이중 SVM을 이용하였을 때 가장 높은 최종 성능을 얻을 수 있었으며, 6개의 감정분류에서 50.8%, 3개의 감정분류에서 78.0%의 인식결과를 보였다.

**Abstract** This paper proposes an efficient method for automatically distinguishing various facial expressions. To recognize the emotions from facial expressions, the facial images are obtained by digital cameras, and a number of feature points were extracted. The extracted feature points are then transformed to 49-dimensional feature vectors which are robust to scale and translational variations, and the facial emotions are recognized by statistical pattern classifiers such as Naive Bayes, MLP (multi-layer perceptron), and SVM (support vector machine). Based on the experimental results with 5-fold cross validation, SVM was the best among the classifiers, whose performance was obtained by 50.8% for 6 emotion classification, and 78.0% for 3 emotions.

**Key Words** : Emotion recognition, Facial expression, ASM (active shape model), Pattern recognition, SVM (support vector machine).

### 1. 서론

전자기술과 정보기술의 발전에 따라 스마트폰과 노트북 컴퓨터 등이 개인 기기의 사용이 확대되고 있다. 특히

2013년 CES (consumer electronics show)에서는 기기가 사용자의 의도를 파악하여 지시가 있기 전에 알아서 대응하는 ‘맥락 인식’ 기술들이 많이 소개되었다. 이를 위하여 기계가 사람의 요구를 미리 파악하고, 사람의 요구사

\*정회원, 경북대학교 전자공학부

\*\*준회원, 울산과학기술대학교 전기전자컴퓨터공학부

\*\*\*정회원, 영남대학교 정보통신공학과 (교신저자)

\*\*\*\*정회원, 목원대학교 지능로봇공학과

접수일자 2014년 5월 21일, 수정완료 2014년 6월 11일

게재확정일자 2014년 6월 13일

Received: 21 May, 2014 / Revised: 11 June, 2014

Accepted: 13 June, 2014

\*\*\*Corresponding Author: parkjs@yu.ac.kr

Dept. of Information and Communication Engineering, Yeungnam University, Korea

항 이상의 결과물을 제공하는 기술이 산업계에서 지속적으로 요구되고 있으며, 기술의 파급효과는 새로운 사업 기회와 시장이 될 것이 자명하다.<sup>[1-5]</sup>

사용자의 의도 및 요구를 나타내는 중요한 요소 중의 한 가지가 감정 상태로, 사용자의 감정의 파악은 입출력 인터페이스 디자인의 중요한 요소로 작용하게 되며, 단순히 사용자의 시간적인 사용 패턴에 따라 입출력 인터페이스를 조정하게 되면 감정 상태에 따라 달라지는 요구 사항을 제때에 적용하기 어렵게 된다. 기존의 심리학 분야에서는 사람의 감정의 상태를 6가지로 분류하였으며 많은 연구를 수행하였으나, 이를 공학적으로 해석하는 연구 결과는 따라가지 못하고 있다. 인간의 감정은 단순히 하나의 요소로 판단할 수 없으며, 이를 기존의 정량적인 분석법으로는 기술하지 못한다.<sup>[6-7]</sup>

사람의 얼굴 표정은 광학적 흐름 분석(optical flow analysis), 홀리스틱 분석(holistic analysis), 국부적인 표현(local representation) 방법 등으로 분석이 가능하고, 광학적 흐름 추정을 통하여 얼굴의 근육의 움직임이 자세히 연구되었다.<sup>[8-9]</sup> 또한 얼굴영상에서 추출한 표정과 음성신호, 뇌파 등의 생체신호를 조합하여 감정인식에 적용되기도 하였다. 이를 위하여 Haar-like 필터와 AdaBoost 알고리즘을 이용하여 얼굴영역을 추출하고, 눈과 코, 입 등의 감정이 나타나는 요소들의 위치를 추출하였다.<sup>[10]</sup> 또한 얼굴영상의 특징을 보다 잘 기술하기 위하여 Eigenface를 이용하여 최적의 감정이 표현된 얼굴을 찾아내었다.

본 연구에서는 사용자의 얼굴 영상을 이용하여 감정을 파악하는 기술을 개발하고, 이를 여러 가지 통계적 패턴분류 기술을 적용하여 최적의 방법을 찾아내고자 한다. 이를 위하여 다양한 종류의 통계기반 패턴분류기를 적용하였으며, 결과의 신뢰성을 높이기 위하여 5-fold cross-validation을 적용하여 자료의 선택에 관계없이 순수 패턴 분류기의 성능을 평가하였다. 평가 결과, SVM (support vector machine)이 감정의 변이에 관계없이 MLP (multi-layer perceptron)과 Naive Bayes 대비 좋은 성능을 보였다.

본 논문의 구성은 다음과 같다. 2장에서는 감정에 따른 표정의 변화를 나타낼 수 있는 특징을 ASM (active shape model)을 이용하여 추출하는 방법을 소개하고, 3장에서는 여러 가지 통계적인 패턴분류 방법을 기술한다. 4장에서는 제안한 감정분류 시스템의 성능평가 결과를

보이고, 5장에서 결론을 맺는다.

## II. 특징 추출 방법

감정의 변화에 의하여 영향을 받는 것은 얼굴을 구성하는 근육들의 움직임으로 나타나며, 이는 눈, 코, 입 등의 모양인 표정으로 나타난다. 이러한 변화를 모델링하기 위하여 본 연구에서는 얼굴의 형태와 모양을 기술하는 ASM (active shape model)을 사용하였다.<sup>[11-12]</sup> ASM은 눈, 코, 입, 눈썹 등의 위치를 사람에 의해 기재된 학습 영상으로부터 점 분포 모델(point distribution model, PDM)을 저장하고, 이를 입력 영상과의 최적의 정합을 구하여 입력 영상의 특징점들을 추출한다. ASM에 의해 구한 특징점의 예는 그림 1과 같다. 원래의 ASM은 얼굴 윤곽선 특징점까지 얻을 수 있지만 본 연구에서는 표정을 나타내는 얼굴의 구성요소에 대해서만 특징점을 사용하였다. 하나의 눈썹에 대하여 6개의 특징점, 눈에 대하여 5개, 코에서 10개, 그리고 입술에서는 13개의 특징점이 추출되며, 그리하여 총 45개의 특징점을 얻었다. 추출된 특징점들을 조합하여 최종적으로 영상의 크기와 위치에 영향을 정규화한 49개의 수치로 표정의 특징 기술자를 추출하였다.<sup>[13-15]</sup>

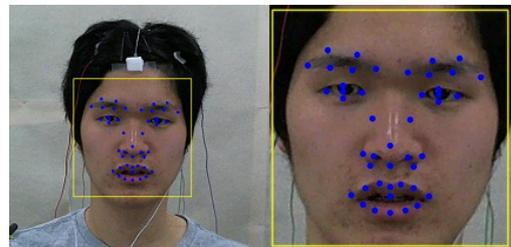


그림 1. ASM 특징점 추출 예

Fig. 1. Examples of extracted feature points by Active Shape models

## III. 패턴인식 방법

본 연구에서는 추출된 49차 벡터를 이용하여 통계적인 확률모델을 학습하였고, 이를 이용하여 감정들을 자동으로 분류하고 인식하는 시스템을 구축하였다. 사용된 패턴분류기는 각 특징들의 독립성을 가정하는 Naive

Bayes 분류기, MLP (multi-layer perceptron), 그리고 SVM (support vector machine)을 사용하였다.

### 1. Naive Bayes 분류기

Naive Bayes, 혹은 Naive Bayesian 분류기는 가장 기본적인 패턴 분류 방법의 하나로, 현재 시스템의 기본이 되는 성능을 얻고자 할 때 주로 사용된다. 이 분류기는 다차원 특징벡터에 대하여 모든 차원이 서로 독립적이라는 매우 강한 가정에 기반하며, 각 차원은 또한 일반적으로 1차원 가우시안 확률 분포임을 가정하여 전체 학습자료에 대하여 단지 평균과 분산만을 추정하기 때문에 그 학습과정이 매우 간단하다.

주어진 분류  $c$  에 대하여 입력벡터  $\mathbf{x}$  가 주어졌을 때 후험확률(a posteriori probability)는 다음의 식 (1)과 같이 표현된다.<sup>[16]</sup>

$$p(c|\mathbf{x}) = p(c)p(\mathbf{x}|c) \propto p(c) \prod_{i=1}^D p(x_i|c) \quad (1)$$

각각의 차원의 우도(likelihood)는 정규분포로 가정되며, 식 (2)와 같이 계산된다.

$$p(x_i|c) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{x_i - \mu_i}{2\sigma_i^2}\right] \quad (2)$$

학습과정은 차원  $i$  에 대하여 주어진 학습자료의 평균  $\mu_i$  와 분산  $\sigma_i^2$  만이 필요하기 때문에 학습이 매우 간단하다. 하지만 식 (1)에서 모든 차원이 독립적이라는 가정은 해당되지 않는 경우 인식성능이 떨어지게 된다. 하지만 구현이 매우 간단하고 성능에 영향을 주는 요소가 적기 때문에 본 연구에서는 감정인식의 baseline 성능을 구하기 위하여 사용하였다.

### 2. MLP (multi-layer perceptron)

신경회로망(artificial neural network)의 대표적인 구현방법으로, 선형분류기인 perceptron을 복층구조로 쌓아 올린 패턴분류기이다.<sup>[17]</sup> 역전파(backpropagation) 알고리즘으로 학습되며,<sup>[18]</sup> 본 연구에서는 한 단계의 숨겨진 단계, 즉 두 층의 perceptron 구조를 적용하였고, 패턴분류에 일반적으로 사용되는 시그모이드(sigmoid) 함수를

최종 결과를 얻는 데 사용하였다. 이를 이용하여 일반적인 비선형 패턴분류가 가능하였으며, 49차의 입력 차원에 대해서도 문제없이 학습과 인식을 수행할 수 있었다.

### 3. SVM (support vector machine)

학습자료의 수가 충분하지 않을 경우, 각 층별 선형분류기를 차원들의 총 조합으로 학습하여야 하는 MLP의 정확도는 매우 떨어지게 되어 성능저하 요인이 된다. SVM은 적은 수의 지지벡터(support vector)를 최적화 방법으로 구한 후, 이를 이용하여 입력 공간을 두 개로 나누게 된다. 임의의 비선형 공간을 나누기 위하여 본 연구에서는 Gaussian 함수와 유사한 형태를 가지는 radial-basis function kernel을 사용하였으며,<sup>[19]</sup> 이에 따르는 분류되는 패턴은 그림 2에서 예를 들었다.

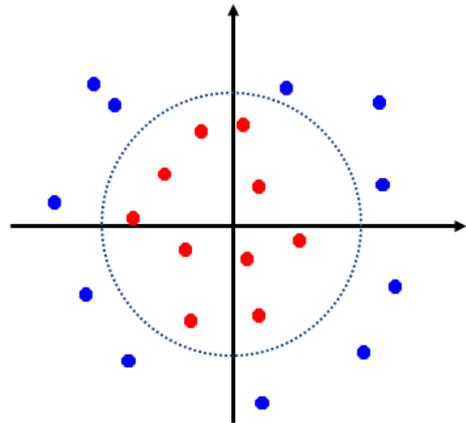


그림 2. Radial-basis kernel function을 이용한 SVM 패턴분류 결과의 예

Fig. 2. Example of pattern classification result using radial-basis function kernels in SVM

## IV. 실험 및 결과

### 1. 자료의 수집 및 학습-인식 부분집합 설계

제안된 시스템의 유효성을 검증하기 위하여 본 연구에서는 실험실에서 얻어진 정지영상에 대하여 감정인식 실험을 수행하였다. 실험에 사용된 감정의 개수는 총 6가지로 분노(Anger), 혐오(Disgust), 놀람(Surprise), 공포(Fear), 기쁨(Joy), 슬픔(Sadness)이며 일반적인 심리실

험에서 분류방법을 따랐다. 총 153명의 피험자에 대하여 517개의 감정이 뚜렷이 나타나는 장면을 추출하였으며, 각 감정별 학습자료의 수는 표 1과 같다.

**표 1. 각 감정별 장면 샘플들의 수 및 피실험자의 수**  
Table 1. Numbers of image samples and test subjects for individual emotion classes

	분노	혐오	놀람	공포	기쁨	슬픔	합계
장면수 (개)	32	150	119	15	178	23	517
피실험자수 (명)	19	77	78	8	113	12	153

**2. 5-fold Cross Validation 실험 설계**

감정의 추출은 피실험자에게 각 감정을 유발하는 영상물을 보여주고, 자연스럽게 감정이 유발되는 순간을 사람이 직접 추출하는 방식을 사용하였다. 인위적으로 감정을 나타내지 않았기 때문에 각 감정별 장면 샘플들의 수가 달랐고, 공포와 슬픔은 얼굴 표정이 나타나지 않는 경우가 많아 실험에 사용할 수 있는 샘플들의 수가 매우 부족했다.

이 문제를 해결하기 위해서 본 연구에서는 전체 자료를 5개의 부분집합으로 나누고, 이 중 4개를 학습에 사용하고 1개를 평가에 사용하였다. 이 경우 5개의 서로 다른 조합이 가능하며 모든 자료를 이용하여 평가를 할 수 있는 장점이 있다. 이러한 실험 설계 방식을 5-fold cross validation 이라고 한다.<sup>[20-22]</sup>

표 1의 장면 샘플들을 5개의 부분집합으로 나누었으며, 표 2는 나눈 부분집합들을 샘플들의 수로 표현한 것이고, 표 3은 각 부분집합의 샘플들을 추출한 피실험자의 수를 나타낸다. 왼쪽 열의 “CV(숫자)”은 5가지 cross validation 부분집합을 의미한다. 표 3에서 하나의 행의 피실험자의 수를 더하면 맨 오른쪽의 합과 같다. 실제 환경에서는 모델 학습에 참여한 피실험자와 인식 실험에 참여하는 피실험자들이 다를 것이기 때문에, 이를 반영하기 위하여 cross validation을 위한 부분집합들을 피실험자들의 겹침이 없도록 나누었다. 각 부분집합들의 최소 샘플들의 수는 공포, 슬픔, 분노에서 3개였으며, 최소 피실험자의 수는 슬픔에서 1명이었다.

**표 2. Cross-validation을 위한 부분집합별 장면 샘플들의 수**  
Table 2. Numbers of image samples for individual cross-validation subsets

	분노	혐오	놀람	공포	기쁨	슬픔	합계
CV1	4	36	27	3	37	5	112
CV2	3	23	27	3	34	3	93
CV3	9	34	28	3	37	3	114
CV4	11	25	19	3	39	5	102
CV5	5	32	18	3	31	7	96
합계	32	150	119	15	178	23	517

**표 3. Cross-validation을 위한 부분집합별 피실험자들의 수**  
Table 3. Numbers of test subjects for individual cross-validation subsets

	분노	혐오	놀람	공포	기쁨	슬픔	합계
CV1	2	21	16	2	9	3	53
CV2	2	13	13	2	14	1	45
CV3	6	20	18	2	10	2	58
CV4	6	15	12	2	10	3	48
CV5	3	19	12	2	9	1	46
합계	19	88	71	10	52	10	153

각각의 Naive Bayes, MLP, SVM 패턴분류기는 4개의 subset으로 학습하고 1개의 subset으로 평가 실험을 수행하였다. 모든 장면은 5번의 평가에서 각각 1번씩 평가 자료에 포함되고, 4번의 학습자료에 포함되게 되며, 학습 자료에는 평가자료의 피실험자가 포함되지 않는다. 따라서, 총 5번의 학습-인식 평가실험이 가능했고, 이의 평균치로 최종성능을 평가하였다. 최종 cross-validation 설계방식과 샘플들의 수는 표 4에 정리하였다.

**표 4. 5-fold cross-validation 설계 및 학습 및 평가 샘플들의 수**

**Table 4. Design of 5-fold cross-validation and numbers of training and test samples for validation cases**

CV 부분집합 할당 (학습)-평가	학습 샘플들의 수 (장면수)	평가 샘플들의 수 (장면수)
(2345)-1	405	112
(1345)-2	424	93
(1245)-3	403	114
(1235)-4	415	102
(1234)-5	421	96

**3. 반복 샘플링을 이용한 학습자료 균등화**

표 1에 나온 바와 같이 장면의 수는 감정의 종류 및 특성에 따라 다르게 수집되었다. 이는 자연스러운 감정 발현을 위한 자료수집으로, 실제 다른 감정들의 발현 빈도

를 어느 정도 반영하는 수치이다. 하지만 일반적인 패턴 분류 성능은 학습 샘플의 수가 많은 감정으로 모델이 학습되는 경향이 있기 때문에 감정 장면의 수가 불균일할 경우, 각 감정별 성능차가 발생할 수 있다. 이를 해결하기 위하여 학습자료가 부족한 감정에 대하여, 샘플링을 통해 학습자료를 인위적으로 생성하였다.

표 5. 다양한 학습자료 할당 방법 및 장면 샘플들의 수  
 Table 5. Various training data assignments and their numbers of image samples

학습자료 할당	분노	혐오	놀람	공포	기쁨	슬픔	합계
2345	28	114	92	12	141	18	405
1345	29	127	92	12	144	20	424
1245	23	116	91	12	141	20	403
1235	21	125	100	12	139	18	415
1234	27	118	101	12	147	16	421

표 6. 반복 샘플링을 통한 감정별 성능 변화  
 Table 6. Performance comparison of single sampling and multiple sampling

	분노	혐오	놀람	공포	기쁨	슬픔	평균
1번 샘플링	0.0%	78.7%	73.9%	0.0%	81.5%	60.9%	42.9%
반복 샘플링	34.4%	59.3%	55.5%	20.0%	74.7%	60.9%	50.8%

예를 들면, 표 5의 학습자료 2345에서 장면샘플들의 최대 개수는 141개인 기쁨이며, 장면샘플들의 최소 개수는 공포로 12개이다. 이 숫자를 맞추어 주기 위해 공포의 샘플을 11번 반복 샘플링하고, 작은 크기의 Gaussian noise를 첨가하여 학습자료의 수를 균등하게 맞추어 주었다. 표 6은 개수를 균등하게 맞추어 주었을 때의 성능 변화를 나타낸다. 패턴 분류기는 SVM을 사용하였다. 반복 샘플링을 사용한 경우, 각 감정별 성능은 분노가 0%에서 34.4%, 공포가 0%에서 20.0%로 향상되어 적은 샘플수에서도 적절한 성능을 보장하였다. 하지만 좋은 성능을 보였던 혐오와 놀람의 성능은 크게 떨어져, 전체 평균 성능은 크게 차이가 없었다. 따라서 반복 샘플링은 응용 분야에 따라 전체 성능이 중요하지, 개별 성능이 중요하지 판단하여 적용이 고려되어야 한다.

#### 4. 각 패턴분류기 별 감정 분류 성능 비교

각 감정별 분류 성능을 증시하여, MLP (multi-layer perceptron)과 SVM (support vector machine)의 학습에는 반복 샘플링을 사용하였다. 하지만 Naive Bayes는 차

원별 평균과 표준편차만이 요구되기 때문에 반복 샘플링을 하더라도 성능의 차이가 없어 사용하지 않았다. 전체 6개 감정들에 대해 실험 이후, 성능이 떨어지는 공포, 분노, 슬픔을 하나씩 제외해 가면서 성능을 평가하였다. 각 감정별 자세한 평가결과는 표 7에서 정리하여 나타내었다.

평균 성능으로 볼 때, Naive Bayes 패턴분류기는 6개 감정에서 41.7%, 5개에서 49.3%, 4개에서 64.0%, 그리고 혐오, 놀람, 기쁨 세가지에서는 65.7%로 감정의 개수를 줄여감에 따라 그 성능이 계속적으로 향상되었다. MLP는 감정의 개수를 줄여감에 따라 50.9%, 54.5%, 67.9%, 76.3%로 역시 성능이 향상되었으며, SVM은 50.8%, 58.3%, 73.9%, 78.0%로 패턴분류기 중 가장 좋은 성능을 보였다.

표 7. 각 패턴분류기별 성능평가 결과  
 Table 7. Performance comparison of various pattern classification methods

(a) Naive Bayes 결과

(a) Results of Naive Bayes

분노	혐오	놀람	공포	기쁨	슬픔	평균
3.1%	50.7%	50.4%	20.0%	56.2%	69.6%	41.7%
3.1%	60.7%	53.8%	X	59.6%	69.6%	49.3%
X	61.3%	63.9%	X	61.2%	69.6%	64.0%
X	71.3%	63.9%	X	61.8%	X	65.7%

(b) MLP 결과

(b) Results of MLP

분노	혐오	놀람	공포	기쁨	슬픔	평균
25.0%	68.0%	63.9%	13.3%	74.2%	60.9%	50.9%
18.8%	68.7%	67.2%	X	70.2%	47.8%	54.5%
X	80.0%	68.1%	X	75.8%	47.8%	67.9%
X	78.0%	75.6%	X	75.3%	X	76.3%

(c) SVM 결과

(c) Results of SVM

분노	혐오	놀람	공포	기쁨	슬픔	평균
34.4%	59.3%	55.5%	20.0%	74.7%	60.9%	50.8%
25.0%	65.3%	63.9%	X	76.4%	60.9%	58.3%
X	80.7%	74.8%	X	79.2%	60.9%	73.9%
X	77.3%	76.5%	X	80.3%	X	78.0%

## V. 결론

본 논문에서는 표정으로 발현된 감정을 얼굴영상을 이용하여 자동으로 인식하는 방법을 제안하였다. 얼굴의

각 요소는 ASM (active shape model)을 사용하였으며, 특징점들의 상대적인 위치로부터 총 49개의 입력벡터를 구성하였고, Naive Bayes, MLP (multi-layer perceptron), SVM (support vector machine)을 사용하여 감정패턴을 자동으로 분류하였다. 성능평가는 5개의 부분집합으로 학습-평가를 반복하는 5-fold cross validation을 수행하였으며, 적은 자료에서도 높은 수준의 신뢰성을 보일 수 있었다.

또한 자연적으로 발생하는 감정들의 샘플 개수를 균일하게 맞추어 주기 위한 반복 샘플링 방법을 제안하였고, 전체적인 성능은 유지하면서도 각 감정별 성능을 균일하게 맞추었다.

본 시스템은 스마트폰 등에 적용되어 사용자의 감정 상태에 따른 맞춤형 서비스를 가능하게 하여, 새로운 시장을 창출하게 만들 수 있으며, 탑재 기기의 부가 가치를 높일 수 있을 것으로 보인다. 현재 소형 디지털 기기에 적합한 얼굴 특징점 추출 방법의 연구 및 인식기의 최적화에 대하여 연구가 진행중에 있다.

## References

- [1] C. H. Ahn, J. H. Choi, S. J. Yang, W. T. Lim, and J.H. Cha, "Trends of Emotional Information Service," *Electronics and Telecommunications Trends*, Vol. 27, No. 6, December 2012.
- [2] J. S. Park and J. W. Lee, "Trends of Emotional User Experience," *Electronics and Telecommunications Trends*, Vol. 26, No. 5, October 2011.
- [3] WonBum Yoon, HeuiSeok Lim, and SungHyun Yoon, "Smart Phone Real Time Q&A System," in *Proc.the KAIS Fall Conference*, Vol. 1, pp. 12-15, November 2010.
- [4] Yeong-Wook Yang, Jae-Choon Jo, and Heui-Seok Lim, "The Smart Phone Application Implement for Cognitive Ability Training," in *Proc.the KAIS Fall Conference*, Vol. 1, pp. 43-46, November 2010.
- [5] Cheol-Ho Yoon, "Usability Study of Different Types of Mobile Phones Using KLM Model," *Journal of the Korea Academia-Industrial cooperation Society* Vol.14, No.5, pp. 2081-2085, 2013.
- [6] Ji-Soo O, Jeong-Jin Kang, Myung-Jae Lim, and Ki-Young Lee, "Design and Implementation of an Emotion Recognition System using Physiological Signal," *The Journal of The Institute of Internet, Broadcasting and Communication (JIIBC)*, Vol. 10, No. 1, pp.57-62, February 2010.
- [7] Yeon-Soo Lee and Young B. Park, "An acoustic study of feeling information extracting method," *The Journal of The Institute of Internet, Broadcasting and Communication (JIIBC)*, Vol. 10, No. 1, pp.51-55, February 2010.
- [8] Hyunsup Yoon, Youngjoon Han, and Hernsoo Hahn, "Facial Expression Recognition based on Expression Space Feature," in *Proc. Workshop on Speech Communication and Signal Processing*, pp. 395-396, August 2009.
- [9] Tae-Hun Yoo, Jong-Yong Lee, Jin-Soo Kim, and Sang-Hun Lee, "A Study on Detecting Salient Region using Frequency-Luminance of image," in *Proc.the KAIS Spring Conference*, Vol. 2, pp. 486-489, May 2012
- [10] Paul Viola and Michael J. Jones, "Robust Real-Time Face Detection," *International Journal of Computer Vision*, Vol. 57, Issue 2, pp 137-154, May 2004.
- [11] T.F. Cootes and C.J. Taylor and D.H. Cooper and J. Graham, "Active shape models - their training and application," *Computer Vision and Image Understanding*, Vol. 61, pp. 38-59, 1995.
- [12] Byungtae Ahn, Yudeog Han, and In So Kweon, "Real-Time Facial Landmarks Tracking using Active Shape Model and LK Optical Flow", in *Proc. International Conference on Ubiquitous Robots and Ambient Intelligence (URAI 2012)*, pp. 541-543, November 2012.
- [13] G. Antonio, M. Sorci, M. Bierlaire, and J. Thiran, "Discrete choice models for static facial expression recognition," in *Proc. International Conference on Advanced Concepts for Intelligent Vision Systems*, Vol. 4179, pp. 710-721, 2006.

- [14] Y. Zhang and Q. Ji, "Facial expression understanding in image sequences using dynamic and active visual information fusion," in Proc. International Conference on Computer Vision, Vol. 2, pp. 1297-1304, 2003.
- [15] D. Bimler, and G. Paramei, "Facial-expression affective attributes and their configural correlates: components and categories", Spanish Journal of Psychology, Vol. 9, pp. 19-31, 2006.
- [16] Pedro Domingos and Michael Pazzani, "On the optimality of the simple Bayesian classifier under zero-one loss," Machine Learning Vol. 29, pp. 103-137, 1997.
- [17] Dong-Hyung Lee, Man-Mo Kang, Young-Kee Kim, and Soo-Dong Lee, "Sign Language recognition Using Sequential Ram-based Cumulative Neural Networks," Journal of IWIT, Vol.9, No.5, pp. 205-211, October 2009.
- [18] Keon-Jun Park, Yong-Kab Kim, Byun-Gon Kim, and Geun-Chang Hoang, "Design of Optimized Pattern Recognizer by Means of Fuzzy Neural Networks Based on Individual Input Space," Journal of IWIT, Vol.13, No.2, pp. 181-189, April 2013.
- [19] Kang-Yun Hee and Park-Young, "Design and Implementation of Document Classifier for IT documents based on SVM," Journal of IWIT, Vol.4, No.1, pp. 71-78, October 2004.
- [20] Ron Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in Proc. International Joint Conference on Artificial Intelligence, Vol. 2, No. 12, pp. 1137-1143, 1995.
- [21] Richard Picard and Dennis Cook, "Cross-Validation of Regression Models," Journal of the American Statistical Association, Vol. 79, No. 387, pp. 575-583, 1984.
- [22] Gyoo-Seok Choi, Jong-Jin Park, and Ha-Nam Nguyen, "Feature Selection Algorithm using Random Forest to Diagnose Cancer," International Journal of Advanced Smart Convergence (IJASC), Vol.1, No.1, pp.10-15, 2009.

## 저자 소개

### 장 길 진(정회원)



- 1997년 : KAIST 전산학과(공학사)
- 1999년 : KAIST 전자전산학과(공학석사)
- 2004년 : KAIST 전자전산학과(공학박사)
- 2004년~2006년 : 삼성종합기술원 전 문연구원
- 2006년~2007년 : 미국 Softmax, 연구원
- 2008년~2009년 : 미국 UCSD, 박사후연구원
- 2009년~2014년 : 울산과학기술대학교 조교수
- 2014년~현재 : 경북대학교 조교수
- 관심분야 : 음성신호처리, 패턴인식, 컴퓨터비전, 의료영상 처리

### 조 아 라(준회원)



- 2010년 : 안양대학교 디지털미디어공학과(공학사)
- 2010년~현재 : 울산과학기술대학교 석박사통합과정
- 관심분야 : 컴퓨터비전, 패턴인식, 휴먼-컴퓨터 인터랙션

### 박 정 식 (정회원)



- 2001년 : 아주대학교 정보및컴퓨터공학부(공학사)
- 2003년 : KAIST 전자전산학과(공학석사)
- 2010년 : KAIST 전산학과(공학박사)
- 2010년~2012년 : KAIST 전산학과 박사후연구원
- 2012년~2014년 : 목원대학교 지능로봇공학과 조교수
- 2014년~현재 : 영남대학교 정보통신공학과 조교수
- 관심분야 : 음성정보처리, 감정인식, 휴먼-로봇인터랙션

## 서 용 호 (정회원)



- 1999년 : KAIST 전산학과(공학사)
- 2001년 : KAIST 전자전산학과(공학 석사)
- 2007년 : KAIST 전자전산학과(공학 박사)
- 2007년 : 미국 MS 로보틱스그룹 인턴 연구원
- 2008년 : 미국 쉐일 과장
- 2010년~현재 : 목원대학교 지능로봇공학과 조교수
- 관심분야 : 지능로봇, 로봇비전, 휴머노이드, 임베디드시스템

※ 이 논문은 2013년도 정부(미래창조과학부)의 재원으로 휴먼인지환경사업본부-신기술융합형 성장동력사업의 지원을 받아 수행된 연구임 (No. 2013K000359).