

# 생리적 내재반응 및 얼굴표정 간 확률 관계 모델 기반의 감정인식 시스템에 관한 연구

## A Study on Emotion Recognition Systems based on the Probabilistic Relational Model Between Facial Expressions and Physiological Responses

고 광 은, 심 귀 보\*  
(Kwang-Eun Ko<sup>1</sup> and Kwee-Bo Sim<sup>1</sup>)

<sup>1</sup>School of Electrical and Electronics Engineering, Chung-Ang University

**Abstract:** The current vision-based approaches for emotion recognition, such as facial expression analysis, have many technical limitations in real circumstances, and are not suitable for applications that use them solely in practical environments. In this paper, we propose an approach for emotion recognition by combining extrinsic representations and intrinsic activities among the natural responses of humans which are given specific stimuli for inducing emotional states. The intrinsic activities can be used to compensate the uncertainty of extrinsic representations of emotional states. This combination is done by using PRMs (Probabilistic Relational Models) which are extended version of Bayesian networks and are learned by greedy-search algorithms and expectation-maximization algorithms. Previous research of facial expression-related extrinsic emotion features and physiological signal-based intrinsic emotion features are combined into the attributes of the PRMs in the emotion recognition domain. The maximum likelihood estimation with the given dependency structure and estimated parameter set is used to classify the label of the target emotional states.

**Keywords:** emotion recognition, facial expression, physiological responses, probabilistic relational model, feature fusion

### 1. 서론

최근의 인간의 감정을 인식하는 연구는 인공지능, 인간-컴퓨터 인터페이스 등을 망라하여 심리학, 정신의학, 뇌 의 과학 등의 다(多)학제 기반 융합을 통함으로써 HCI 연구의 새로운 방향성을 제시하고 있다. 인간의 감정은 외부와 소통하는 과정에서 가장 자연스럽게 발생하는 상태적 변화의 한 종류로 다양한 정의가 가능하다. 그 중에서도 HCI 분야에서는 신경생리학적 활동, 운동적 표현, 주관적 느낌과 행동경향 및 인지과정을 거쳐 표면적으로 드러나는 정보들의 집합체로 감정을 정의하고 있으며, 다양한 방식의 측정 및 분석이 이루어져 왔다. 감정을 시간 축 상에서의 유지 특성에 대해 분류할 때 가장 널리 쓰이는 정의는 초(second) 단위 구간에서 유지되는 표현형(expression)으로서 시계열 해석의 대상으로 여겨져 왔다[1]. 감정에 대한 대표적인 표현형의 정보로 음성, 표정, 행동 등이 있다. 각각에 대하여 대표적인 연구로부터 주목할 만한 여러 가지 성과들이 도출되었지만, 그 중에서도 자동화 된 시스템을 기반으로 감정인식을 구현하기 위해 가장 보편화된 방식은 얼굴 표정 해석에 의한 방법이다. 표정을 기반으로 감정을 인식하기 위

해 선행되어야 하는 연구로는 컴퓨터 비전 및 영상처리 기술, 패턴인식 기반 기계학습 혹은 추계학적(stochastic) 방법론 등이 있다. 이러한 선행 기술들은 얼굴 영역 검출, 표정 특징 계산, 표정 상태 분류로 구성되는 체계화된 분석 과정에 각각 적용되어왔다. 각 단계 별로 살펴보면, 안면 영역 부분 특징 혹은 전 영역을 기반으로 배경과 대상을 분리하는 얼굴 영역 검출은 외부 광원, 배경 효과 등 노이즈나 자세 다변화에 강인하기 위한 영상처리 알고리즘의 개발에 초점을 맞춰 연구되어져 왔다[2,13]. 그리고 검출된 얼굴 영역에서 표정의 주요 정보를 내포한 특징을 검출하기 위해 정적 특징, 동적 특징으로 구분하거나 Gabor wavelets, Point distribution model과 같은 전체적(holistic) 방법과 intensity profile, principal component analysis 등의 국소적 방법 등이 대표적으로 활용되어져왔다[2]. 이러한 방법들은 최근까지도 사용되고 있으며, 기존의 2차원 상의 특징을 대신하여 3차원 공간상의 깊이 정보를 함께 고려한 형태로 진화된 사례도 있다. 또한, 검출된 표정 특징 정보에 대한 분류를 위해 사용되는 패턴인식을 정의함에 있어 다양한 분야에 특성화된 정의가 있지만, 본 논문에서는 다차원 공간 내에서 확률 분포를 추정하고 공간을 카테고리 혹은 클래스 영역으로 나누는 문제에 초점을 맞추었다[3]. 이에 해당하는 방법들로는 support vector machine, hidden Markov models, recurrent neural network 등이 있다[2,4-6]. 경우에 따라서 추계학적 접근으로는 FACS (Facial Action Coding System) 등의 매뉴얼에 의해 수집된 통계적 데이터를 이용하여 선형 또는 비선형 회귀(regression) 방법을 사용

\* 책임저자(Corresponding Author)

Manuscript received March 25, 2013 / revised April 10, 2013 / accepted April 29, 2013

고광은, 심귀보: 중앙대학교 전자전기공학부

(kkeun@cau.ac.kr/kbsim@cau.ac.kr)

※ 본 논문은 한국연구재단 중견연구자지원사업(No. 2012-0008726)에서 지원하여 연구하였음.

하기도 한다. 이러한 일련의 표정 기반 감정인식 연구가 다양한 연구자 및 기관에서 수행되어져 왔으나 대부분의 기존 연구는 의식적으로 조절 가능한 얼굴 표정근에서 드러나는 표면적인 감정 상태만을 고려하기 때문에 내재적 감정 상태에 대한 불확정성에 따른 사항을 고려해야한다. 이는 표정이 감정 정보를 유지하는 간격이 짧기 때문에 외란이나 측정 대상자의 정서적 영향에 의해 원신호에 대한 훼손 또는 왜곡을 대비 해야한다. 이에 대한 해결책으로 음성이나 생리적 반응을 측정하여 활용하기도 한다.

본 논문에서는 선행연구 [4]와 유사하게 얼굴 표정에 기반한 단일 신호에서의 감정 상태 전달과정에서 발생할 수 있는 원신호의 훼손 또는 왜곡을 보상하기 위해 자율신경계 및 불수의근 등에 의해 종속된 생리적 반응을 측정, 분석하고자 한다. 이러한 생리적 반응 기반으로 하는 내재 감정 특징을 정의하여 얼굴 표정에서 드러나는 표면 감정 특징과 감정인식 도메인 상에서 관계적 모델을 구축하였다. 이러한 관계 모델은 개별적인 특징 클래스 간의 확률 분포에 기초함으로써 감정 특징 간의 융합을 가능하게 하며, 별도의 분류기나 인식 과정이 없이도 확률 관계 모델의 출력을 통해 해당 시점에서 주어진 표면 및 내적 감정 특징에 대응하여 가장 높은 확률 분포로 가질 것으로 예측되는 상태를 결과로 도출할 수 있다. 이러한 확률 관계 모델의 출력력이 감정 상태에 클래스 라벨링 값으로 도출되어 얼굴 표정 기반 미세 감정 상태에 대한 인식을 수행하게 된다.

본 논문의 구성은 다음과 같다. II 장에서는 표면 감정 특징과 내재 감정 특징을 각각 얼굴 표정 기반 영상 데이터와 생리적 반응 신호들을 통해 검출하기 위해 적용된 방법론 및 검출 과정을 다룬다. III 장에서는 표면 감정 특징과 내재 감정 특징의 융합을 위해 활용하기 위한 확률 관계 모델에 대한 설명과 감정인식 도메인 상에 적용하기 위한 설정을 다룬다. IV 장에서는 감정인식 도메인에서의 종속 구조 및 파라미터에 대한 학습 과정을 논한다. 본 논문에서 구축한 확률 관계 모델을 기반으로 융합 감정 특징에 대응하는 감정 상태를 분류한 결과를 통해 제안하는 감정인식 방법의 성능을 평가한다. 마지막 V 장에서는 본 논문에서 제안하는 방법에 대한 최종 결론과 더불어 향후 보완이 필요한 사항을 소개한다.

## II. 관련 연구

본 논문에서 목표로 하는 감정인식 시스템의 입력에 해당하는 특징은 얼굴 표정에서의 표면 감정 특징과 생리적 반응에서의 내재 감정 특징으로 구성된다. 서론에서 언급한 바와 같이, 각각의 특징을 검출하기 위하여 보편적인 영상 처리 및 신호처리 알고리즘을 활용하였다. 본 장에서는 내재, 외재 감정 특징에 대한 검출 과정을 소개하고자 한다. 또한, 입력된 특징이 시스템 내에서 관계적 모델에 의해 융합되는 과정을 수행할 수 있도록 하는 확률 관계 모델에 대해서도 소개한다.

### 1. 표면 감정 특징: 얼굴 표정

표정을 기반으로 얼굴 영상의 특정 영역들에 분포된 landmark 집합, 즉 AU (Action Unit)의 연속된 궤적을 측정

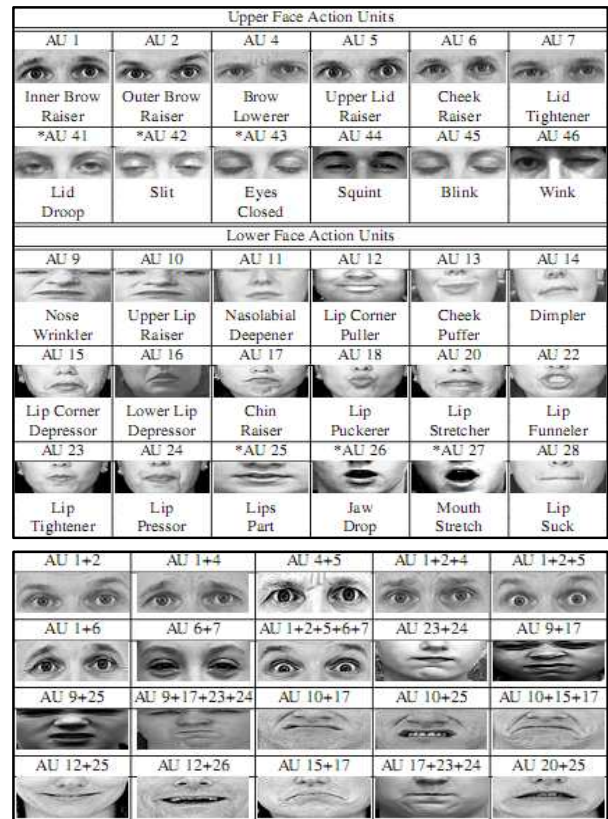


그림 1. 얼굴 표정 관련 (상) 단일 AU, (하) AU의 조합 [3].

Fig. 1. Facial expression-related (UP) single AU and (DOWN) combination of AUs [3].

하고 이에 기반한 감정 상태 별 확률 분포를 정의함으로써 표면 감정 특징을 검출할 수 있다. 얼굴 영역에 분산된 landmark의 집합을 AU 및 AU의 조합으로 정의하고 이러한 AU 영역의 움직임과 그에 따른 감정상태 간의 관계의 규칙을 명세화한 것이 P. Ekman이 개발한 FACS이다[10]. 일반적으로 사람 대 사람의 감정 상태나 그 변화를 해석할 때, FACS를 기반으로 하여 표정의 세부적인 의미를 분류하는 것이 가능하다[10]. FACS는 능숙한 사용을 위하여 사용자에 대한 반복적 훈련 및 학습이 요구되는데, 이를 자동화된 시스템으로 대체하고자 하는 많은 연구들이 진행된 바 있다. 이와 같이 각각의 AU는 안면근육 움직임을 특정 표정 데이터로 매칭한 결과인데, 그림 1과 같이 단일 AU와 AU 조합에 의해 표정을 구성하는 일부분들을 표현된다.

본 논문에서는 BosphorusDB를 기반으로 기본 6정서 = {화남, 혐오, 공포, 기쁨, 놀라움, 슬픔}를 나타내는 얼굴 영상에 분포된 22개의 2D landmark들을 활용하였다[8,9]. 본 연구의 선행연구를 기반으로 구축한 그림 2의 안면 메쉬 모델 [11]에서 사전에 정의된 AU 영역들 가운데, BosphorusDB의 22개 landmark와 본 논문에서 추가로 정의한 2개의 landmark를 이용하여 표정의 상태와 변화에 따른 AU의 움직임을 관측하였다. 그림 2에서 AU영역에 분포된 24개 landmark를 도시하였으며 각각의 landmark 인덱스에 대한 명세는 표 1로 정의하였다. 이를 기반으로 AU에 대한 landmark의 조합을 구성함으로써 AU 조합의 움직임 데이터

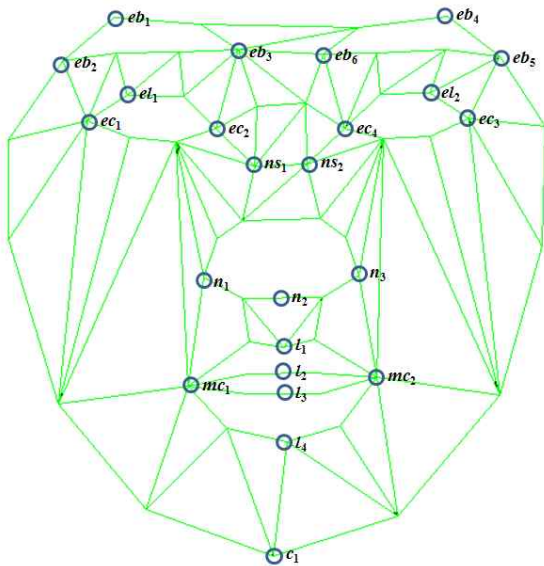


그림 2. 얼굴 표정 관련 AU 영역에서의 24개 landmark 분포.  
Fig. 2. Facial expression-related 24 landmark distribution in the AU regions.

표 1. AU 영역에서 분포된 landmark index의 명세.

Table 1. Labels of landmark index distributed in the AU regions.

Landmark Labels	Index	Landmark Labels	Index
'Outer left eyebrow'	eb <sub>2</sub>	'Nose saddle left'	ns <sub>1</sub>
'Middle left eyebrow'	eb <sub>1</sub>	'Nose saddle right'	ns <sub>2</sub>
'Inner left eyebrow'	eb <sub>3</sub>	'Left nose peak'	n <sub>1</sub>
'Inner right eyebrow'	eb <sub>6</sub>	'Nose tip'	n <sub>2</sub>
'Middle right eyebrow'	eb <sub>4</sub>	'Right nose peak'	n <sub>3</sub>
'Outer right eyebrow'	eb <sub>5</sub>	'Left mouth corner'	mc <sub>1</sub>
'Left eyelid'	el <sub>1</sub>	'Right mouth corner'	mc <sub>2</sub>
'Right eyelid'	el <sub>2</sub>	'Upper lip outer middle'	l <sub>1</sub>
'Outer left eye corner'	ec <sub>1</sub>	'Upper lip inner middle'	l <sub>2</sub>
'Inner left eye corner'	ec <sub>2</sub>	'Lower lip inner middle'	l <sub>3</sub>
'Inner right eye corner'	ec <sub>4</sub>	'Lower lip outer middle'	l <sub>4</sub>
'Outer right eye corner'	ec <sub>3</sub>	'Chin middle'	c <sub>1</sub>

시퀀스를 구축할 수 있다. 이때의 AU에 대한 landmark의 조합은 FACS을 참조하였다[10]. 또한, 본 논문에서는 AU 단위의 움직임 방향을 2차원 축으로만 구성하였고 움직임이 각각의  $x, y$ 축에 대한 편미분이 양의 값일 때 유한 차분 간격  $h$ 를 1, 음의 값일 때 -1이라고 정하였다.

영상 시퀀스가 초당 10 프레임을 가정한다면 landmark 기반 AU 특징들의 수치에 대한 미분의 유한차분간격  $h$ 를 0.1로 정해지고, 이때의 표정 변화에 대응하는 특징벡터를 다음 표 2와 같이 정의할 수 있다. 인식하고자 하는 감정을 은닉 상태로 정의할 때, 관측 상태는 AU 또는 AU 조합에 해당하는 landmark의 움직임, 즉 표 2의 특징 파라미터들이고, 각 표정에서의 AU 또는 AU 조합에 대한 사전 확률 분포를 알 수 있다면, 주어진 관측 상태에 대한 은닉 상태의 확률을 추론할 수 있다.

표 2. AU에 대응하는 표면 특징 파라미터 집합.

Table 2. The set of external feature parameters correspond to AUs.

	External Feature Parameters	associated AUs
$X_1$	$\{\nabla_y d(eb_3, ec_2)\}, \{\nabla_y d(eb_6, ec_4)\}$	$AU_1$
$X_2$	$\{\nabla_y d(eb_2, ec_1)\}, \{\nabla_y d(eb_5, ec_3)\}$	$AU_2$
$X_3$	$\{\nabla_\theta \angle (eb_2, eb_3, ec_2)\}, \{\nabla_\theta \angle (eb_6, eb_5, ec_3)\}$	$AU_4$
$X_4$	$\{\nabla_y d(el_1, ec_2)\}, \{\nabla_y d(el_2, ec_4)\}$	$AU_5$
$X_5$	$\{\nabla_\theta \angle (ec_1, el_1, ec_2)\}, \{\nabla_\theta \angle (ec_4, el_2, ec_3)\}$	$AU_6, AU_7$
$X_6$	$\{\nabla_\theta \angle (eb_3, ec_2, ns_1)\}, \{\nabla_\theta \angle (eb_6, ec_4, ns_2)\}$	$AU_9$
$X_7$	$\{\nabla_y d(n_2, l_2)\}$	$AU_{10}$
$X_8$	$\{\nabla_x d(n_1, mc_1)\}, \{\nabla_x d(n_3, mc_2)\}$	$AU_{12}$
$X_9$	$\{\nabla_y d(n_1, mc_1)\}, \{\nabla_y d(n_3, mc_2)\}$	$AU_{15}$
$X_{10}$	$\{\nabla_\theta \angle (mc_1, c_1, mc_2)\}$	$AU_{17}$
$X_{11}$	$\{\nabla_\theta \angle (mc_1, l_2, mc_2)\}, \{\nabla_\theta \angle (mc_1, l_3, mc_2)\}$	$AU_{23}$
$X_{12}$	$\{\nabla_y d(l_1, l_4)\}, \{\nabla_y d(l_1, c_1)\}$	$AU_{24}$
$X_{13}$	$\{\nabla_y d(l_2, l_3)\}, \{\nabla_x d(mc_1, mc_2)\}$	$AU_{20}, AU_{25}$
$X_{14}$	$\{\nabla_y d(l_2, l_4)\}, \{\nabla_y d(l_2, c_1)\}$	$AU_{26}$

이와 같이 AU나 AU의 조합에 따라 감정 상태에 대응하는 고유의 움직임을 표면 감정 특징화 할 수 있으며, 이에 대한 확률분포를 구함으로써, 표정에 해당하는 감정 상태에 대한 클래스 라벨링을 할 수 있다. 그렇지만 표면 감정 특징만을 이용해 표정으로 나타나는 감정 상태가 피험자의 자연적인 반응인지, 인위적인 반응인지 여부에 대한 결정을 내리기에 다소 미흡하다. 본 논문에서는 이를 보완하고자 자율신경계에 의해 발생하는 생리적 반응을 감정 상태에 관련된 특징으로 활용하여 생리적 반응들과의 감정 상태에 대한 관계 모델을 구축하고자 한다.

2. 내재 감정 특징: 생리적 반응

본 논문에서의 생리적 반응들의 수집과정은 본 연구실에서 선행한 생리적 신호 대응 감정인식 연구에서의 측정실험과정을 따른다[12]. 공포, 기쁨, 혐오, 평활의 4가지 감정을 유발하기 위하여, 70회의 시도가 1 세션을 구성하도록 설계된 측정실험 패러다임을 통해 4명의 피험자에 대해 각각 10 세션에 해당하는 데이터를 수집한 바 있다. 이렇게 측정된 생리적 반응에 대한 실험 데이터는 biofeedback 2000 x-pert MULTI 모듈에 의해 측정되는 SC (Skin Conductance), ST (Skin Temperature), BVP (Blood Volume Pulse)로 구성되어 있다.

첫 번째 생리적 반응으로 SC는 피부 상에 발생하는 전기적 포텐셜을 저장함으로써 측정되며 일반적으로 구형파 형태를 가진다. 본 연구실에서 선행된 연구에서 측정된 신호의 측정 범위는 0에서 50 $\mu$ S(siemens)이고, 시간 해상도는 0.001 $\mu$ S/sec이다. SC는 스트레스 반응의 척도 중 응답속도가 가장 빠른 신호 중 하나이며 자율신경계활동에 대한 측정에 쓰이는 비침습적 방법 중 가장 강인하다. 주로 피험자에 대한 외적 또는 내적인 환경의 변화가 SC의 변화를 유발하는데, 일반적인 신호처리 기법들이 이러한 변화를 측정하기 위하여 활용된다. 일반적으로 SC 특징 검출은 로우패스필터에 의한 원 신호 평탄화(smoothed) 후, 상승 에지에

서의 급강하 구간에서의 신호에 대한 미분을 통해 계산된 임계치를 기반으로 환경의 변화 유무를 판단할 수 있는 지표를 얻는 과정을 통한다. SC를 이용한 감정인식을 위해서는 신호의 내적 상태가 얼마만큼 변화되었는가를 정량화된 수치로 나타내거나 SC의 임계치(threshold) 구간의 폭과 감정 상태와의 연관성을 판단함으로써 가능하다.

두 번째로 BVP는 피부표면상에 근사한 혈액의 흐름에 대한 평균을 측정하며 값의 범위는 0~100%이고 시간해상도는 매 0.25%마다이다. 이 파라미터는 30~200 bpm (beats per minute)을 가지며 1 bpm의 해상도를 가진다. 이때 샘플링비율은 20Hz로 초당 10 데이터 비율로 100의 시정수를 가지도록 설정했다. BVP는 가장 대표적인 심장 활동 신호로서 흥분상태와 이완상태에 따라 변화의 폭이 종속된다는 특성을 보인다. 이를 감정인식에 적용하기 위해 원신호의 상승 에지를 측정하기 위한 임계치를 계산해야하는데, IBI (Inter-Beat Interval) 방법은 상승 에지 간 시간구간을 IBI로 계산하고 IBI의 역으로써 해당 구간에서의 심박으로 기록하는 방식으로 임계치를 설정한다. 시간에 따른 SC와 BVP의 측정치는 그림 3과 같이 측정된 바 있다[12].

기존의 생리적 반응 기반 감정인식 관련 선행 연구에서는 ST도 고려 대상이었다[12]. ST는 초당 4가지 수치의 데이터 비율을 디지털 형태로 멀티모듈 내에서 처리 및 전달되고, 이를 처리한 결과가 10-40 °C 범위 내에서 온도의 범위에서 0.01°C마다 시간해상도를 가지는 값으로 표현되었다. 이러한 ST의 특성은 감정 상태의 변화에 따른 신호의 반응속도가 상대적으로 SC나 BVP보다 늦기 때문에 표면 감정 특징과 융합에 적용하는 단계에 있어 시간적 동기화가 어렵기 때문에 본 논문에서는 활용을 배제하였다.

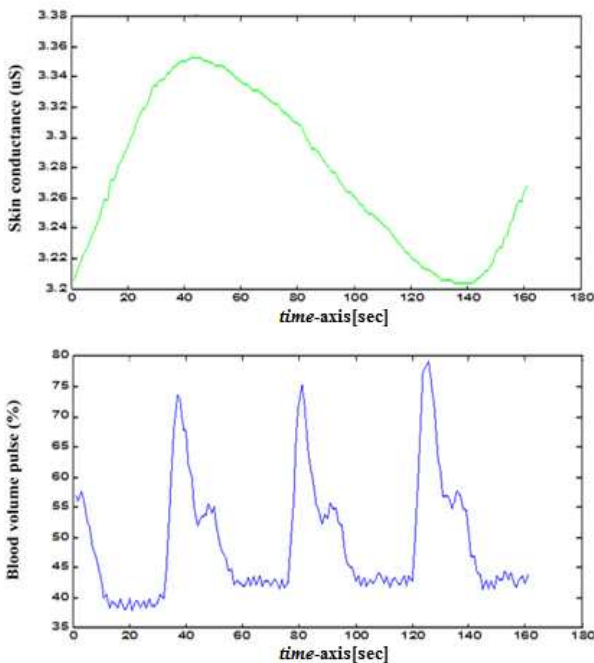


그림 3. SC, BVP 측정치 예시 (공포 감정 자극 부여).

Fig. 3. Example of measurement of SC, BVP (providing fear emotion-related stimuli).

SC와 BVP로 구성된 생리적 반응에서 내재 감정 특징을 검출하기 위해 각각의 신호에 대하여 전절에서 시간 축 수치 미분을 위해 설정된 유한차분간격  $h = 0.1\text{sec}$ 에 따른 신호 샘플링을 선행한다. 구형과 특성을 보이는 SC의 경우 수치미분에 의한 시간 축 도함수의 상승 구간 간격의 평균  $m_{SC}$ 와 표준 편차  $std_{SC}$ 로 구성된 쌍을 내재 특징으로 활용하고 BVP의 경우, 최대 크기 값에 해당하는  $peak_{BVP}$ 와 최대의 파워스펙트럼을 가지는 신호 주파수 대역  $freq_{BVP}$ 의 쌍을 내재 특징으로 활용한다. 이렇게 구해진 내재 감정 특징 벡터는 확률관계모델로 적용하여 표면 감정 특징과의 융합 및 본 논문에서 제안하는 감정 상태 간 종속 관계에 대한 확률 추론을 수행하기 위해 활용된다.

### III. 감정인식 도메인에서의 확률 관계 모델 구축

표정과 생리적 반응 각각의 측정 과정에서 개별적 측정 결과에 대한 인식을 수행한 여러 기존 연구들의 (얼굴표정 감정인식[5,6,8,9], 생리적 반응 감정인식[4,12]) 수행된 연구는 제한된 환경 하에서 일정 수준의 결과를 보여주지만 실제 환경에서는 인상적이지 못한 결과를 보인다. 훈련 데이터베이스에 학습되지 않은 외란이나 unlabeled 특성도 있겠지만, 가장 큰 원인은 주어진 환경에 반응하여 드러나는 감정의 자연적인(natural), 무의식적(unconscious), 내재적(inherent) 특성에 의한 것으로 보인다. 이러한 가설 하에 실생활 가운데 보다 정확하게 감정인식을 수행하려면 주변 상황정보에 대한 인식이(context awareness) 선행되어야 하는데, 이를 고려한 시스템 구축 시 하드웨어나 소프트웨어적인 복잡도가 커지는 문제가 있다. 본 장에서는 이러한 상황정보를 대체할 수 있는 대안을 소개하고자 한다.

본 장에서는 개별 감정 특징들을 PRM으로 적용하여 감정 상태 클래스 라벨링에 대한 입력으로 활용하는 과정을 논한다. 이를 위해, 확률 관계 모델의 구조 및 파라미터에 대한 학습과 학습된 확률 관계 모델을 기반으로 하는 감정 인식 과정에 대하여 설명한다.

#### 1. 확률 관계 모델(probabilistic relational models)

확률관계모델(PRM)은 목표 도메인을 가정하고 도메인 내에 존재하는 클래스들로 구성된 관계 로직 스키마의 모든 인스턴스에 대한 확률 분포를 구한 결과이다[7]. 이 모델의 스키마는 목표 도메인 내에 존재하는 클래스와 클래스의 속성들(attributes), 그리고 각 속성의 값으로 구성되며 속성 간의 종속과 클래스 간의 종속을 표현하는 관계 골격(relational skeleton,  $\sigma$ )을 기반으로 모델이 구축된다. 임의의 클래스  $x$ 와 속성  $\alpha$ 의 모든 가능한 집합을 각각  $X$ ,  $A(x)$ 로 표현하면 속성 간의 종속 관계를 "Parent"로 표현하여  $x.\alpha \leftarrow PA(x,\alpha)$ ,  $x \in X$ ,  $\alpha \in A(x)$ 로 정의 가능하다. 클래스 내의 속성 중에 참조 슬롯(reference slot)이라하는 속성인 동시에 하위 클래스의 역할을 하는 존재가 있으며, 이는 주어진 클래스 내의 관계 로직을 정의함에 있어 계층적 관계에 대한 지원을 가능하게 하고, 조건부 결합확률 분포 (CPD: Conditional Probability Distribution)로  $P(x,\alpha|PA(x,\alpha))$ 가 주로 쓰인다. 관계 골격  $\sigma$ , 종속 구조  $\psi$ , 도메인 내  $X.A$ 와  $PA(X.A)$ 의 값, 즉 파라미터 집합  $\theta_\psi$ 가 주어졌을 때 임의

의 인스턴스  $I$ 에 대한 CPD를 다음 식 (1)과 같다.

$$P(I | \sigma, \Psi, \theta_\Psi) = \prod_{x \in \sigma} \prod_{\alpha \in A(x)} P(I_{x,\alpha} | I_{PA(x,\alpha)}) \quad (1)$$

이러한 PRM에 대한 파라미터 및 구조 학습은 베이지안 네트워크 (Bayesian network, BNT) 의 학습과정과 유사하다. PRM의 파라미터 학습은 목표 도메인 내  $i$ 개의 클래스가 존재하고 각각의 클래스  $X_i$ 에 대한 객체  $x$ 를  $k$ 개 생성할 경우, 식 (1)에 의한 CPD를 만족시키는  $\theta_\Psi$ 를 구하는 과정으로 BNT에서와 마찬가지로 MLE (Maximum Likelihood Estimation) 문제로 풀 수 있다. PRM의 적정 파라미터 집합에 대한 ML  $l$ 은 다음과 같이 구할 수 있다.

$$l(\theta_\Psi | I, \sigma, \Psi) = \log P(I | \sigma, \Psi, \theta_\Psi) = \sum_{X_i} \sum_{\alpha \in A(X_i)} \left[ \sum_{x_k} \log P(I_{x,\alpha} | PA(I_{x,\alpha})) \right] \quad (2)$$

PRM의 구조학습은 BNT의 종속 모델(dependency model) 구축 과정에 해당하며 학습데이터베이스로부터의 속성 간의 확률 모델을 우선 구축해야 한다. 이러한 모델의 구축은 목표 도메인 상의 클래스 및 속성 간의 무수히 많은 가능한 종속 구조 중에 학습 데이터베이스에 대응하는 가설 공간(hypothesis space)에서 스코어링 함수(scoring function)를 기반으로 적절한 탐색 알고리즘을 적용함으로써 가능해진다. 일반적으로 Greedy-search 알고리즘 등이 활용 된다[7].

이와 같이 관계 골격을 기반으로 객체 간과 속성 간의 확률 분포만을 정의하면 되는 확률 관계 모델의 특성은 기존의 방법들에 비교했을 때 여러 가지 장점이 있다. 기존의 BNT 등의 확률 그래픽 모델들은 입력으로 적용되는 통계적 데이터가 flat-file (다른 파일, 즉 응용 도메인)으로 연결을 고려하지 않은 독립 형태)로 고정하여 그 활용 폭을 제한해야 했다. 그리고 BNT에서 클래스 간 의존 관계를 정의하기 때문에 클래스의 모든 인스턴스 간 확률 의존성 또한 전부 정량화되어야 하는 단점이 있다. 반면에 확률 관계 모델에서는 관계 골격에 기반한 객체 간이나 객체 내 속성 간 의존 관계에 대한 정량화만이 요구되므로 규모의 집적화가 가능하다. 또한 객체 간 관계가 도메인에 따라 유동적인 변화가 가능하기에 다양한 도메인 상에 활용 가능한 템플릿이라고 볼 수 있다.

2. 감정인식 도메인 확률 관계 모델 구조 학습

우리가 측정하고자 하는 감정의 유효한 시간 구간이 초 단위로 짧지만 연속된 감정 상태의 변화를 가정할 경우 감정 발생과정의 중첩으로 인해 실생활 조건에서는 정확한 결과가 나올 것이라고 기대하기 어렵다. 본 논문에서는 이러한 조건을 기반으로 특정 감정상태를 유발하는 자극에 대응하는 표정 및 생리적 반응을 동시에 측정하는 패러다임을 구상한다. 표정의 측정을 수행할 때 활용되는 패러다임은 공인된 자극 데이터베이스를 기반으로 하며 피험자는 자극에 대응하는 표정을 발생시키도록 요구받는다. 이 경우, 일정 비율로 의식적으로 자극에 부합되지 않는 표정을 발생시키도록 한다. 생리적 반응은 자율신경계를 통해 발생하는 신호로 사람의 인위적인 제어가 불가능하다. 즉 상황에 즉각적인 반응을 보인다고 할 수 있다. 따라서 연속적

표 3. 감정인식 도메인에서의 관계 골격 명세.

Table 3. Specification of relational skeleton in domain of emotion recognition.

Class	Emotion	Inherent Feature	Extrinsic Feature	AUs	CLMs
Attributes or Ref. Slot (type)	Inherent feature	$m_{SC}$ ; (numeric)	AUs	[ $au_1 \dots au_m$ ]; (single AU, nominal)	[ $X_1, \dots, X_{14}$ ]; (derivative & aggregation values, nominal)
		$std_{SC}$ ; (numeric)			
	Extrinsic feature	$peak_{BVP}$ ; (numeric)	CLMs	[ $cau_1 \dots cau_m$ ]; (comb.of AU, nominal)	
intensity; (numeric)	$freq_{BVP}$ ; (numeric)				

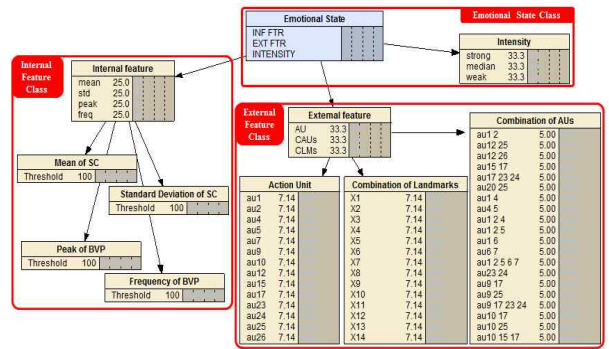


그림 4. 감정인식 도메인 관계 골격,  $\sigma_E$ .

Fig. 4. Relational skeleton  $\sigma_E$  of domain of emotion recognition.

자극을 순차적으로 부여하고 그에 대응하는 연속적인 표정 측정과 생리적 반응 측정을 동시에 진행하여 생리적 반응 지표에 따른 표정 변화의 영향력, 즉 종속성뿐만 아니라 감정 상태, 즉 클래스 객체 간의 의존성을 구조화 한다면 별도의 상황인식 과정에 대한 고려 없이 실제 환경에 유연하게 적용할 수 있는 템플릿이 될 수 있을 것이다. 이는 본 논문에서 적용하고자 하는 PRM의 특성에 잘 부합된다. 본 논문에서 목표로 하는 도메인이 감정인식이고, 해당 도메인 상에 구체화 가능한 객체를 기본 6정서로 가정한다. 이 때 관계 골격을 구성하는 클래스 및 속성에 대한 명세는 다음 표 3에 도시되어 있다.

이러한 명세를 바탕으로 감정인식 도메인에서의 관계 골격을 구성한다면 다음 그림 4와 같다. 이 때 감정 상태를 최상위 계층의 클래스로 보고, 각 클래스에 대한 속성 중 내재 감정 특징과 표면 감정 특징은 참조 슬롯으로 하위 계층 클래스를 생성하였다. 감정 상태 별로 가중치가 부여되는 내재 및 표면 감정 특징이 각각 다르므로 속성 간의 확률 관계 또한 특정되어야 한다.

그림 4의 관계 골격을 구성하는 노드 및 네트워크 연결선 (arc)의 구성과 종속 관계 간 확률 값의 분포를 계산하기 위하여 본 논문에서는 Norsys Corp.의 *Netica Application* 및 *Netica API*를 이용하였다.

3. 감정인식 도메인 확률 관계 모델 파라미터 학습

본 논문에서는 표정 인식을 위하여 105명의 피험자에 대한 기본 6정서를 표현하는 2차원 데이터 3,675개의 영상 데

표 4. 얼굴표정에서 AU 발생 확률 분포,  $P(AUs|x)$ .

Table 4. Probability distributions of the AU occurrences for each facial expression,  $P(AUs|x)$ .

Emotion-Objects AU Comb.	$x$					
	Happy	Disgust	Angry	Surprise	Sad	Fear
1+2	0.533	0.048	0.247	0.773	0.264	0.659
1+4	0.298	0.912	0.327	0.180	0.928	0.729
4+5	0.152	0.592	0.981	0.345	0.429	0.332
1+2+4	0.341	0.733	0.969	0.155	0.871	0.839
1+2+5	0.610	0.029	0.021	0.952	0.011	0.723
1+6	0.078	0.666	0.007	0.564	0.866	0.802
6+7	0.931	0.001	0.001	0.021	0.512	0.032
1+2+5+6+7	0.602	0.202	0.003	0.989	0.031	0.362
23+24	0.066	0.603	0.692	0.002	0.856	0.039
9+17	0.013	0.595	0.801	0.001	0.562	0.009
9+25	0.002	0.698	0.793	0.001	0.447	0.008
9+17+23+24	0.001	0.761	0.908	0.001	0.301	0.002
10+17	0.021	0.927	0.635	0.001	0.633	0.001
10+25	0.220	0.860	0.992	0.002	0.237	0.011
10+15+17	0.031	0.298	0.621	0.003	0.902	0.007
12+25	0.899	0.001	0.004	0.011	0.008	0.132
12+26	0.808	0.001	0.001	0.135	0.002	0.291
15+17	0.011	0.218	0.132	0.028	0.997	0.001
17+23+24	0.068	0.594	0.753	0.001	0.571	0.001
20+25	0.005	0.802	0.368	0.007	0.219	0.521

이터를 보유한 BosphorusDB를 활용하였다[8]. 또한, 생리적 반응에 의한 감정인식을 위해 8명의 피험자에게서 수집된 [12]의 데이터베이스를 활용하였다. 각각의 학습 데이터베이스를 기반으로 EM (Expectation-Maximization) 알고리즘을 적용함으로써 기본 6정서의 객체에 대한 속성 및 클래스 간 종속 관계를 표현하는 PRM의 파라미터 집합을 구할 수 있다. 아래의 표 4는 파라미터 집합 중 감정상태 객체  $x$ 가 주어졌을 때, AU 조합이 발생될 확률 분포를 구한 것이다. 파라미터 집합은 [E\_state, Intensity, mean\_SC, std\_SC, peak\_BVP, freq\_BVP, AU, cAU, cLndmrk]로 구성되며 ASCII 텍스트 파일 형태로 저장되어 그림 4에 구성된 종속 구조 모델에 대한 학습을 수행하게 된다.

4. 확률 관계 모델 기반 감정인식 시뮬레이션 결과 분석

본 논문에서는 감정인식 도메인에서 존재하는 표면 감정 특징 관련 클래스와 내재 감정 특징 관련 클래스 간의 종속 구조 각 클래스를 구성하는 속성 간의 종속 구조에 대한 모델링을 수행했으며, 이를 근거하는 파라미터에 대한 학습을 통해 감정 상태에 대한 특징 벡터들을 융합하였다. 이 결과를 기반으로 감정상태 별 분류를 위해 표 4에서와 같이 기본 6정서 기반 감정객체  $x$ 에 대한 최우 추정 (MLE: Maximum Likelihood Estimation) 방법을 활용함으로써 감정 상태의 각 클래스 별 라벨링을 통한 감정인식이 가능하다. 이러한 과정은 특징 단계의 융합과 더불어 융합된 데이터의 은닉 상태에 대한 추론 결과가 일괄적으로 도출되므로 계산 과정에 큰 기여를 하게 된다. 본 논문에서는 주 감정 요소로 표정을, 보조 요소로 생리적 반응을 정

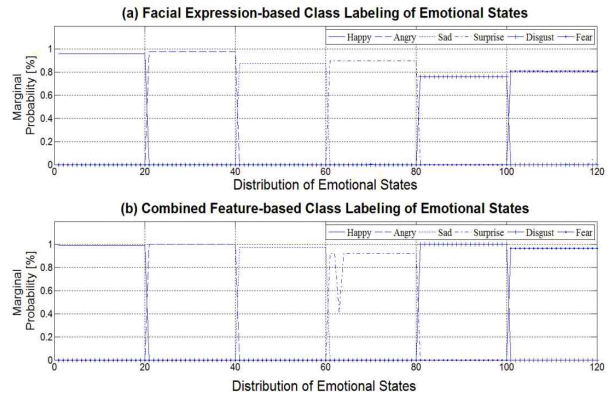


그림 5. 감정인식 시뮬레이션 결과 I. (a) 표정 기반, (b) 융합 감정특징 기반 감정 클래스 라벨링을 위한 경계확률분포.

Fig. 5. Emotion recognition simulation result I. Distribution of marginal probability for (a) facial expression-based class labeling, (b) combined feature-based class labeling.

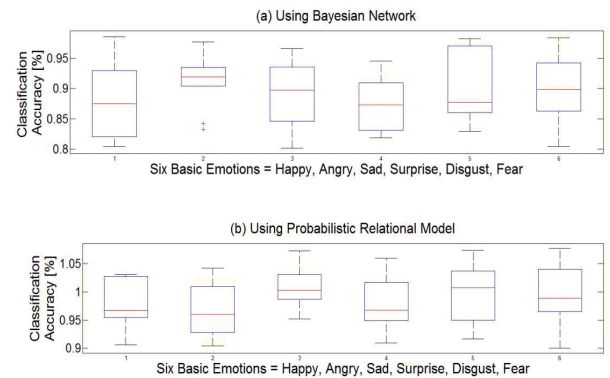


그림 6. 감정인식 시뮬레이션 결과 II. (a) BNT 기반, (b) PRM 기반 감정상태 분류 정확도.

Fig. 6. Emotion recognition simulation result II. (a) BNT-based, (b) PRM-based emotional states classification accuracies.

의하였으며 각각의 특징 간 확률 관계 모델에 기반한 감정 인식률을 추정하였다. 6개의 감정 클래스에 대하여 82개 요소 (내재+외재 특징)로 구성된 융합 감정 특징을 활용하는 제안 방법론의 성능 검증에 위하여 주 감정 요소만을 활용하는 경우와의 인식률 비교를 수행하여 그림 5를 도출하였다.

그림 5의 결과에서 표정만을 단독을 활용한 경우의 감정 상태 클래스 라벨링 결과에 대한 통계적 분석을 수행한 결과, 본 논문에서 제안하는 융합 감정 특징 기반의 경우가 우세한 것으로 보인다. 또한, EM 알고리즘에 의해 학습된 junction tree-BNT를 활용한 경우와의 인식률 및 소요 시간에 대한 비교분석을 그림 6을 통해 수행하였다.

그림 6에서는 기존의 BNT를 활용했을 경우의 감정 상태 분류결과와 비교하여 PRM을 적용한 제안 방법의 분류 결과가 우세한 정확도를 보이고 있다. 또한, 주어진 데이터셋과 동일한 조건 하에서 상대적으로 적은 평균 연산 시간, ET (eval. time)={BNT: 1.352 sec, PRM: 1.045sec}을 소모함으로써 제안하는 방법이 보다 효율적임을 나타낸다.

#### IV. 결론

본 논문에서는 얼굴 표정과 생리적 반응에 기반한 감정 인식을 위한 확률 관계 모델의 사용을 제안하였다. 개별적으로 검출된 표면 감정 특징과 내재 감정 특징 벡터로 구성되는 감정인식 도메인 상에서, 각 클래스 및 속성 간의 중속 구조에 기반한 관계 모델로 구축함으로써 특징 간의 융합을 도출하고 이때의 확률 모델을 통한 최우 추정을 통해 개별 감정 상태에 대한 확률 계산으로 감정 클래스 라벨링을 수행하였다. 관계 모델의 구축을 위한 구조 학습을 위해 일반적으로 가장 널리 활용되는 Hill-Climbing 알고리즘을 사용하였으며, 파라미터의 학습에는 영상 및 생리적 반응 DB를 기반으로 EM 알고리즘을 적용하였다. 기본 6경서 기반 감정 클래스에 대하여 본 논문에서 제안하는 PRM 기반 감정인식 결과는 표정 특징만을 활용한 경우와 junction-tree BNT 기반의 인식 결과와 비교를 통해 제안 방법이 보다 우수한 성능임을 보였다.

향후에는 개별적인 감정 특징 단계에서의 차원 축소와 더불어 감정 상태와 상관관계가 높은 특징을 검출하기 위해 기존의 사용한 방법을 보완할 예정이며, 실 환경 하에서의 감정인식을 가정한 별도의 실험 패러다임을 설계하고 이 때 수집된 학습 데이터베이스를 통해 제안 모델에 대한 학습 등을 수행하여 감정인식 성능을 향상시킬 예정이다.

#### REFERENCES

- [1] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. G. Taylor, "Emotion recognition in human-computer interaction," *Signal Processing Magazine, IEEE*, vol. 18, no. 1, pp. 32-80, 2001.
- [2] B. Fasel and J. Luetttin, "Automatic facial expression analysis: a survey," *Pattern Recognition*, vol. 36, no. 1, pp. 259-275, 2003.
- [3] K. Fukunaga *Introduction to Statistical Pattern Recognition*, 2Ed., Academic Press, p. 2, 1990.
- [4] J. N. Bailenson, E. D. Pontikakis, I. B. Mauss, J. J. Gross, M. E. Jabon, C. A. C. Hutcherson, C. Nass, and O. John, "Real-time classification of evoked emotions using facial feature tracking and physiological responses," *International Journal of Human-Computer Studies*, vol. 66, no. 5, pp. 303-317, 2008.
- [5] K. H. Joo, C. M. Geun, and K. K. Paliwal, "Face recognition using emotional face images and fuzzy fisher-face," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 15, no. 1, pp. 94-98, Jan. 2009.
- [6] Z. Zhang, "Feature-based facial expression recognition: Sensitivity analysis and experiments with a multilayer perceptron," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 13, no. 6, pp. 893-911, 1999.
- [7] N. Friedman, L. Getoor, D. Koller, and A. Pfeffer, "Learning probabilistic relational models," pp. 1300-1309.

- [8] A. Savran, B. Sankur, and M. Taha Bilge, "Regression-based intensity estimation of facial action units," *Image and Vision Computing*, vol. 30, no. 10, pp. 774-784, 2012.
- [9] A. Savran, B. Sankur, and M. Taha Bilge, "Comparative evaluation of 3D vs. 2D modality for automatic detection of facial action units," *Pattern Recognition*, vol. 45, no. 2, pp. 767-782, 2012.
- [10] P. Ekman and W. V. Friesen, "Facial action coding system: a technique for the measurement of facial movement," *Consulting Psychologists Press*, 1978.
- [11] K.-E. Ko and K.-B. Sim, "Study of Emotion Recognition based on Facial Image for Emotional Rehabilitation Biofeedback," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 16, no. 10, pp. 957-962, Oct. 2010.
- [12] M. Vanny, K.-E. Ko, S.-M. Park, and K.-B. Sim, "Physiological responses-based emotion recognition using multi-class SVM with RBF kernel," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 19, no. 4, pp. 1-8, Apr. 2013.
- [13] S.-J. Lee and S.-W. Kim, "Classifying scratch defects on billets using image processing and SVM," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 19, no. 3, pp. 177-280, Mar. 2013.



#### 고 광 은

2007년 중앙대학교 전자전기공학부(공학사). 2007년~현재 중앙대학교 대학원 전자전기공학부 석·박사 통합과정 수료. 관심분야는 Human-Computer Interaction, Brain-Computer Interface System, Intention Recognition, Neuro-Science, Neuro-Robotics, 소프트웨어 컴퓨팅 등.



#### 심 귀 보

1984년 중앙대학교 전자공학과(공학사). 1986년 중앙대학교 전자공학과(공학석사). 1990년 The University of Tokyo 전자공학과(공학박사). 1991년~현재 중앙대학교 전자전기공학부 교수. 2002년~현재 중앙대학교 중소기업산학협력센터 센터장. 2006년~2007년 한국지능시스템학회 회장. 2007년~현재 (사)한국산학연협회 서울지역협회 회장. 2009년~2010년 중앙대학교 중앙도서관장 및 박물관장. 2011년~현재 중앙대학교 스마트지능로봇연구센터 센터장. 관심분야는 인공지능, 뇌-컴퓨터 인터페이스, 의도 인식, 감정 인식, 유비쿼터스 지능형로봇, 지능시스템, 컴퓨테이션 인텔리전스, 지능형 홈 및 홈 네트워크, 유비쿼터스 컴퓨팅 및 센서 네트워크, 소프트웨어 컴퓨팅(신경망, 퍼지, 진화연산), 다개체 및 자율분산로봇시스템, 인공 번역시스템, 지능형 감시시스템 등. ICROS Fellow.