

정서 재인 방법 고찰을 통한 통합적 모델 모색에 관한 연구*

Toward an integrated model of emotion recognition methods based on reviews of previous work

박미숙** · 박지은** · 손진훈**†

Mi-Sook Park** · Ji-Eun Park** · Jin-Hun Sohn**†

충남대학교 사회과학대학 심리학과/뇌과학 연구소**

Department of Psychology/Brain Research Institute, Chungnam National University**

Abstract

Current researches on emotion detection classify emotions by using the information from facial, vocal, and bodily expressions, or physiological responses. This study was to review three representative emotion recognition methods, which were based on psychological theory of emotion. Firstly, literature review on the emotion recognition methods based on facial expressions was done. These studies were supported by Darwin's theory. Secondly, review on the emotion recognition methods based on changes in physiology was conducted. These researches were relied on James' theory. Lastly, a review on the emotion recognition was conducted on the basis of multimodality(i.e., combination of signals from face, dialogue, posture, or peripheral nervous system). These studies were supported by both Darwin's and James' theories. In each part, research findings was examined as well as theoretical backgrounds which each method was relied on. This review proposed a need for an integrated model of emotion recognition methods to evolve the way of emotion recognition. The integrated model suggests that emotion recognition methods are needed to include other physiological signals such as brain responses or face temperature. Also, the integrated model proposed that emotion recognition methods are needed to be based on multidimensional model and take consideration of cognitive appraisal factors during emotional experience.

Keywords : Emotion recognition, facial expression, physiology, multimodality

요약

정서 컴퓨팅 분야는 인간과 컴퓨터 간 상호작용을 효과적이게 하기 위하여 사용자의 정서를 재인하는 컴퓨터 시스템을 개발했다. 본 연구의 목적은 심리학적 이론에 기반한 정서 재인 연구들을 고찰하고 보다 진보된 정서 재인 방법을 제안하고자 하였다. 본 연구의 본론에서는 심리학적 이론에 근거한 대표적인 정서 재인 방법을 고찰하였다. 첫째, Darwin 이론에 근거한 얼굴 표정 기반 정서 재인 방법을 고찰하였다. 둘째, James 이론에 근거한 생리신호기반 정서 재인 방법을 고찰하였다. 셋째, 앞의 두 이론에 통합적으로 근거한 다중 모달리티 기반 정서 재인 방법을 고찰하였다. 세 가지 방법들은 이론적 배경과 현재까지 이루어진 연구 결과들을 중심으로 고찰되었다. 결론에서는 선행 연구의 한계점을 종합하여 보다 진보된 정서 재인 방법을 제안하였다. 본 연구는 첫째, 현재 사용되고 있는 제한된 생리 신호에 대한 대안으로 다양한 생리 반응(예., 뇌 활동, 얼굴

* 이 논문은 2010년도 정부(교육과학기술부)의 재원으로 휴먼인지환경사업본부-신기술융합형 성장동력사업의 지원을 받아 수행된 연구임(No. 2010K001129)

† 교신저자 : 손진훈(충남대학교 사회과학대학 심리학과/뇌과학 연구소)

E-mail : jhsohn@cnu.ac.kr

TEL : 042-821-6369

FAX : 042-821-8875

온도 등)을 정서 재인에 포함하도록 제안하였다. 둘째, 애매한 정서를 구분할 수 있도록 정서의 차원 개념에 기반한 정서 재인 방법을 제안하였다. 셋째, 정서 유발에 영향을 미치는 인지적 요소를 정서 재인에 포함하도록 제안하였다. 본 연구에서 제안한 정서 재인 방법은 다양한 생리 신호를 포함하고, 정서의 차원적 개념에 기반하며, 인지적 요소를 고려한 통합적인 모델이다.

주제어 : 정서 재인, 얼굴 표정, 생리 신호, 다중모달리티

1. 서론

정서 컴퓨팅(Affective Computing; AC) 분야의 연구들은 인간과 컴퓨터가 효과적으로 의사소통할 수 있도록 하기 위하여 노력해 왔다. 예를 들면, 사용자의 정서를 재인하고 정서적으로 적절하게 반응하는 컴퓨터 시스템들이 정서 컴퓨팅 연구자들에 의해 개발되어 왔다. 사용자와 상호 작용하는 동안 사용자의 정서 상태를 자동적으로 인식하고 사용자에게 반응하는 시스템은 인간과 컴퓨터의 상호 작용의 질을 높여준다. 결과적으로, 사용자는 컴퓨터 인터페이스를 더 즐길 수 있고 유용하게 사용하게 된다(Calvo & D'Mello, 2010).

정서 컴퓨팅 분야에서 정서 인식은 중요하다. 만약 인터페이스가 사용자의 정서를 파악하지 못하면 사용자에게 적절하게 반응할 수 없기 때문이다. 그러나, 정서 인식은 어려운 문제이다. 그 이유는 정서가 경계가 애매하고 표현이나 경험에서 상당한 개인 차이가 존재하는 구성개념이기 때문이다.

정서 인식 연구에 종사하고 있는 공학자들은 정서를 활용한 컴퓨터 인터페이스를 개발하는데 있어 나타나는 여러 기술적인 문제들을 해결하는데 노력을 집중해 왔다. 반면, 이들은 정서의 이론적 배경에는 관심을 기울이고 있지 않다(Calvo & D'Mello, 2010). 그러나, 정서를 활용한 유용한 인터페이스를 개발하기 위해서는 정서의 과학적 이론과 실용적 공학이 잘 조화를 이루어야 한다.

본 연구의 목적은 대표적인 정서 재인 방법을 고찰하고, 이에 근거한 진보된 정서 재인 방법을 제안하는 것이다. 본 연구는 심리학적 이론에 근거하여 공학적 연구를 진행한 대표적인 정서 재인 방법을 다음과 같은 순서 테로 소개하고자 한다: 1) Darwin 이론에 근거한 얼굴 표정 기반 정서 재인 방법 2) James 이론에 근거한 생리신호 기반 정서 인식 방법 3) Darwin과 James의 이론에 통합적으로 근거한 다중 모달리티 기

반 정서 재인 방법. 본 연구에서는 위의 세 가지 정서 재인 방법들을 이론적 배경과 연구 결과를 중심으로 고찰할 것이다. 최종적으로, 본 연구에서는 위의 정서 재인 연구들이 가지는 한계점에 대한 부분적인 대안으로 진보된 정서 재인 방법을 제안하고자 한다. 본 연구에서 제안하는 정서 재인 방법은 다양한 생리 반응(예., 뇌 활동, 얼굴 온도 등)을 포함하고, 정서의 차원 개념에 기반하며, 정서의 인지적 요인을 고려한 통합적인 모델이다.

2. 얼굴 표정에 기반한 정서 인식 방법

현재의 정서 인식 기법들은 정서 경험동안 나타나는 얼굴 표정, 자세, 몸짓, 생리적 변화, 그리고 언어에 기반하여 정서를 구분하고 있다(Calvo & D'Mello, 2010; Zeng et al., 2009). 정서에서 '표현'을 강조한 Darwin의 이론에 의해 고무되어, 대다수 정서 인식 연구들은 얼굴 표정에서 기본 정서를 탐지하는데 초점을 맞추고 있다. 본 장에서는 얼굴 표정에 기반한 정서 인식 방법을 이론적 배경과 연구 결과를 중심으로 개관하고자 한다.

2.1. 이론적 배경

얼굴 표정에 기반하여 정서를 탐색하고 있는 연구들은 Darwin의 기본 정서적 관점에 근거한다. Darwin (2002)은 인간의 얼굴 표정과 몸짓은 동물과 유사하다는 것을 발견하고, 정서 경험의 행동적 결과물은 진화과정의 산물이라고 결론지었다. Darwin의 정서 이론은 여섯 가지 정서를 나타내는 얼굴 표정은 진화의 과정에서 발전했으며, 그 결과 이 정서들은 모든 문화에 걸쳐 보편적으로 존재한다고 주장한다. 이러한 여섯 정서는 타고난 것이며 정서 표현이 범문화적이고

보편적이라는 관점에서 “기본” 정서로 일컬어진다 (Ekman, 1971; Ekman & Friesen, 2003; Tomkins, 1962; Izard, 1971; Izard, 1994). 기본 정서적 입장을 지지하는 대부분의 연구자들은 분노, 혐오, 공포, 기쁨, 슬픔 그리고 놀람 여섯 가지 정서를 기본 정서로 정의한다. 그리고, 이들은 각 정서에 따른 독특한 얼굴 표정이 존재한다고 가정한다.

Darwin 이론의 중요한 측면은 정서 표현인데(예를 들면, 혐오스러워하는 얼굴) 그는 이를 “유용한 관련 습관(serviceable associated habits)”이라고 명명했다. 이러한 습관은 처음에는 정서 표현을 위해서 진화한 것이 아니라, 인간의 본능적인 행동과 관련있었다. 즉, 혐오스러워하는 얼굴은 처음에는 개인에게 혐오스러운 맛을 유발하는 음식을 거부하는 것과 관련 있었다. 그러나, 결국에는 그 음식이 없는 경우에도 혐오스러운 표정은 ‘혐오’를 유발하는 사건에 대한 정서 표현 수단으로 발전하였다.

일부 연구자들은 Darwin의 진화 이론을 다른 정서 표현까지 확장시켰다. Frijda(1987)는 정서 자극에 직면했을 때 특정한 방식으로 행동하는 “행동 경향성” 또는 준비 상태를 연구했다. 이러한 행동 경향성은 주변 환경에서 발견되는 문제를 해결해야 하는 우리의 욕구와 관련이 있다. 예를 들면, 행동 경향성에서 “접근”은 “원하는” 것을 소비할 수 있도록 하며, 일반적으로 “욕망”이라고 불리는 정서와 관련된다. 한편, “회피”의 목적은 자신을 보호하는 것이며 종종 “공포”라 불리는 정서와 연관된다.

‘정서 표현’을 강조한 Darwin의 기본 정서적 관점에 의해 영향을 받아, 얼굴 표정에 기반하여 정서를 탐색하는 기법을 정서 인식 연구에서 상당히 일반적으로 사용되고 있다.

2.2. 얼굴 표정에 기반한 정서 인식 연구 개관

최근에 Zeng 등(2009)은 얼굴 표정에 기반한 정서 인식 연구 29개를 개관했다(표 1). Darwin의 기본 정서적 관점은 각각의 기본 정서와 관련된 특징적이고 원형적인(prototypical) 얼굴 표정이 있다고 가정하기 때문에(Ekman & Friesen, 2003; Ekman, 1992; Ekman, 1984), 정서를 얼굴 표정에 따라 분류하는 것은 가능하다. 표정은 정서를 경험할 때 짧은 순간 동안 나타나기 때문에, 정서를 탐색하는 것은 정서에 따른 원형

적인(prototypical) 얼굴 표정을 탐색하는 것이다. Ekman과 Friesen(1978)는 얼굴 움직임을 객관적으로 묘사할 수 있는 얼굴 움직임 코딩 체계(Facial Action Coding System; FACS)를 개발하였다. 이 방법은 얼굴의 독립적인 움직임을 확인하기 위하여 일련의 움직임 단위(action units; AUs)를 정의한다. 얼굴 근육 움직임 단위를 이용하여 정서를 구분하기 위한 방법으로는 EMFACS(Emotional FACS) 그리고 FACS-AID (FACS Affect Interpretation Database)와 같은 고차적인 코딩 체계가 고안되었다(Ekman et al., 2002).

정서 인식에서 얼굴 표정을 이용하는 대부분의 연구들은 2차원 얼굴 표정의 전면 이미지를 인식하는데 초점을 맞추고 있다. 또한, 이 연구들은 대부분 머리 움직임이 거의 없는 범위에서 얼굴 표정을 분석하는 것에 적합하다. 예외적으로, Pantic과 Patra (2006) 연구는 얼굴 표정의 측면 이미지를 분석에 사용하였다.

현재 얼굴 표정에 기반하여 정서를 분류하는 연구들은 기계 학습(machine learning) 기법을 이용하고 있다. 얼굴 표정에 기반한 기계 학습 연구들의 두 가지 주요 흐름을 살펴보면, 첫째는 얼굴 표정에서 정서를 인식하는 기법을 사용하거나, 둘째는 얼굴 움직임 단위(AUs)를 인식하는 기법을 사용한다(Cohn, 2006; Pantic & Bartlett, 2007). 그 구체적인 내용은 다음과 같다. 얼굴 표정에 기반한 정서 인식 연구들은 다양한 패턴 인식 방법들을 사용하고 있다. 이 기법들은 2차원 얼굴 표정 이미지에서 추출된 **얼굴 특징(facial features)**에 기반하고 있다. 일반적으로 추출된 얼굴 특징들은 얼굴 구성요소(눈, 입 등)의 모양 그리고 얼굴의 두드러진 부위(눈, 입의 가장자리 등)의 위치와 같은 **기하학적 특징**이나 주름, 둔덕, 그리고 고랑과 같은 얼굴의 조직(texture)을 대표하는 **외형적 특징**이다. 기하학적 특징에 기반한 방법의 전형적인 예는 Chang 등(2006)의 연구인데, 이들은 58 개의 얼굴 지형에 의해 정의된 모양 모델을 사용했으며, Pantic 등의 연구는 입, 눈, 눈썹, 코, 그리고 그 주변의 일련의 얼굴 특징점을 사용했다(Pantic & Bartlett, 2007; Pantic & Patras, 2006; Pantic & Rothkrantz, 2004a; Valstar et al., 2006; Valstar et al., 2007). 외형 특징 기반 방법을 사용한 전형적인 연구들은 다음과 같다 :

표 1. 얼굴 표정에 기반한 정서 인식 연구 요약(Zeng et al., 2009)

연구자	사용된 feature	Classifier	정서 분류	정확율 (%)
Ashraf et al. (2007)	AAM	SVM	2	Im: EER: 19 %
Bartlett et al. (2004)	Gabor wavelets	SVM+HMM	3 AUs	Im/Vi:75-98
Bartlett et al. (2005)	Gabor wavelets	Adaboost SVM	17 AUs	Im:93.4(CK+EH), 90.5 (RU)
Cohen et al. (2003)	12 motion units	Tree-augmented DBN, HMM	6	Im:66.53(CH), 73.22(CK) Vi:58.63(CH)
Cohn et al. (2004)	shape models, Gabor wavelets	LDC	3 AUs	Im:76(3-class)
El Kaliouby & Robinson (2004)	24 facial points	DBN	6	Vi:77.4
Fasel et al. (2004)	Gray-level intensity	NN	7	Im:38-68
Gunes & Piccardi (2005)	Shape features, optical flow	C4.5, Bayes-Net	8	Im:80-100(various fusion)
Ioannou et al. (2005)	FAPs	neurofuzzy network	3	Im: 78
Ji et al. (2006)	Shape features	DBN	2	Correlation coefficient: 95.3
Kapoor & Picard (2005)	Facial and head gesture	GP, SVM, HMM, NN	2	Vi:86
Kapoor et al. (2007)	Pixel difference of mouth region	Same as in [73]	2	Vi:79.17 Baseline: 58
Lee & Elgammal (2005)	Pixel intensity of face region	decomposable model	6	Vi: 39.58 Im: 61.85
Littlewort et al. (2007)	Gabor wavelets	Adaboost SVM	2	Vi: 72
Lucey et al. (2007)	AAM	SVM	AUs: CK: 15 OD: 4	Im: 95 (CK) with 16.66% FAR, 70.47 (OD)

연구자	사용된 feature	Classifier	정서 분류	정확율 (%)
Pantic & Patras (2006)	Facial profile points	Rule-based	27 AUs	Vi: 86.3
Pantic & Rothkrantz (2004a)	frontal and profile facial points	Rule-based	32 AUs	Im:86
Pantic & Rothkrantz (2004b)	frontal and profile facial points	Rule-based, case-based	9	Im:83
Sebe et al. (2004)	12 motion units	kNN	4	Im:93 (CK) 95 (SD)
Tong et al. (2007)	Gabor wavelets	Adaboost DBN	14 AUs	Vi:93.2 (OD), 93.3 (CK)
Valstar et al. (2004)	Motion history images	SNoW kNN	15 AUs	Vi:61(MMI) 68(CK)
Valstar et al. (2006)	8 facial points	gentle boost, SVM	2	Vi:90.7
Valstar et al. (2007)	20 facial points	Gentle SVM-sigmoid	2	Vi:94
Wang & Ahuja (2003)	Shape and gray-level texture	NN with HOSVD	7	Im: 84.58
Wang et al. (2006)	3D surface labels	LDA	6	Im: 83.6
Wen & Huang (2003)	Geometric, ratio-image	Exemplars with GMM	4	Im: 75.37
Whitehill & Omlin (2006)	Haar features	Adaboost	11 AUs	Im: 92.35
Yeasin et al. (2006)	Pixel intensity of face	kNN+HMM	6	Vi: 90.7 (CK) 72-82 (OD)
Zeng et al. (2006)	Texture with LPP	SVDD	2	Im:79(male), 87(female)

표 1의 약자에 대한 설명: Image/Video based, EER: equal error rate, FAR: false acceptance rate, and GP:Gaussian process
AAI, BU, CH, CK, FABO, MMI, RU, and SD are the database name.
EH: the Ekman-Hager database, OD: Other database, and ?: missing entry.

1. Bartlett 등(Bartlett et al., 2003; Bartlett et al., 2005; Bartlett et al., 2006; Littlewort et al., 2007)과 Guo와 Dyer(2005)는 Gabor wavelets 기법을 사용했으며,
2. Whitehill과 Omlin(2006)은 Haar 특징 방법을 사용했고,
3. Chang 등(2004)은 일직선으로 배열된 얼굴 외형의 여러 부분공간(subspace)에 기반한 방법을 사용했다.

몇몇 연구에서 제안된 것처럼(Pantic & Patras, 2006) 얼굴 표정을 이용한 정서 인식에서 기하학적 그리고 외형 특징 기법을 사용하는 것이 최선의 방법으로 여겨진다. 기하학적인 그리고 외형 특징에 기반한 두 기법을 혼합적으로 사용한 연구도 있다. Tian 등(2005)은 얼굴 구성 요소의 모양과 웃을 때 생기는 눈가 주름과 같은 일시적으로 나타나는 얼굴 특징들을 사용하였다. 그리고 Zhang과 Ji(2005)는 눈, 눈썹, 그리고 입술 주변에 26 개의 점과 Tian 등(2005)에 의해 제안된 일시적인 얼굴 특징을 동시에 사용하였다. 또 다른 예는 Lucey 등(2007)에 의해 제안된 방법으로, 얼굴의 외형적 특징과 얼굴 표정의 모양을 획득하기 위하여 활성화된 외형 모델(Active Appearance Model; AAM)을 사용하였다.

얼굴 표정을 이용한 29개의 정서 인식 연구를 개관한 Zeng 등(2009)의 메타분석에 나타난 얼굴 표정기반 정서 인식 연구들의 한계점은 다음과 같다. 첫째, 거의 모든 연구는 6가지 기본 정서를 구분하는데 중점을 두고 있다. 즉, 기본 정서 이외에 우리가 실제 생활에서 경험하는 다른 정서의 탐색은 간과되고 있다. 대부분의 연구들이 기본 정서의 인식에 집중하고 있는 이유는 기본정서가 가진 보편적 속성 때문이다. 또한, 기본 정서를 나타내는 얼굴 표정 데이터는 가용성이 높아 이 데이터를 이용하여 시스템을 훈련시키고(training) 테스트하는 것이 용이하기 때문이기도 하다(예., Kanade et al., 2000). 그러나, 정서 인식 방법을 실제 생활에 효과적으로 응용하기 위해서는 기본 정서 이외의 정서 상태를 탐색하려는 노력이 필요하다. 일부 연구에서는 기본 정서가 아닌 정서 상태, 예를 들면 피로한(fatigue)(Gu & Ji, 2004; Ji et al., 2006), 동의하는(agreeing), 집중하는(concentrated), 흥미있는(interested), 생각하는(thinking), 혼동스러운(confused), 좌절스러운(frustrated)과 같은 정신 상태를 탐색하려고 시도가 진

행되고 있다(El Kaliouby & Robinson, 2004; Kapoor et al., 2007; Kapoor & Picard, 2005; Teeters et al., 2006; Yeasin et al., 2006).

얼굴 표정에 기반한 정서 인식 연구들의 두 번째 한계점은, 약 절반 가량의 연구들이 인위적으로 연기한 얼굴 표정을 분석에 사용하고 있다는 점이다. 인위적으로 지은 표정은 일상의 자연스러운 표정에 비해 상당히 과장되어 있다(Tian et al., 2005). 그러므로, 연기한 표정을 사용하여 개발된 정서 인식 시스템을 실생활에 적용하는 것은 한계가 있다. 최근에 몇몇 연구들은 자연스러운 얼굴 표정을 사용하고 있는데, 자연스러운 얼굴 표정에서 정서를 탐색하거나(Bartlett et al., 2003; Bartlett et al., 2005; Bartlett et al., 2006; Cohn et al., 2004; Cohn & Schmidt, 2004; Ioannou et al., 2005; Lucey et al., 2007; Sebe et al., 2004; Valstar et al., 2006; Zeng et al., 2006; Littlewort et al., 2007; Ashraf, 2007; Valstar et al., 2007), 얼굴 근육의 움직임을 파악하려고 시도하고 있는 연구들이 이에 속한다(Bartlett et al., 2003; Bartlett et al., 2005; Bartlett et al., 2006; Cohn et al., 2004; Cohn & Schmidt, 2004; Valstar et al., 2006; Valstar et al., 2007). 자연스러운 얼굴 표정과 억지스러운 표정 사이에는 중요한 차이점이 존재한다(Pantic & Patras, 2006; Afzal & Robinson, 2009; Cohn, & Schmidt, 2004; Ekman et al., 1990). 그러므로, 인위적인 표정에 근거한 연구에서 나온 통찰은 실제 상황에서 일반화되기 어렵다. 결론적으로, 정서 인식 연구를 실제 생활에 적용하기 위해서는 사용자와 컴퓨터가 상호 작용하는 맥락 속에서 자연스러운 얼굴 표정을 탐색하려는 노력이 필요하다.

얼굴 표정에 기반한 정서 인식 연구들의 세 번째 한계점은 몇몇 연구들만이 삼차원 얼굴 이미지를 사용하고 있다는 점이다. 예를 들면, Huang과 동료들은 3차원 이미지에서 추출된 얼굴 특징을 정서 구분에 사용하였다(예., Cohen et al., 2003; Sebe et al., 2004; Wen & Huang, 2003; Zeng et al., 2006). Cohn 등(2004)은 정서를 탐색하기 위하여 3차원 이미지 상에서 눈썹 움직임 단위와 머리 움직임 분석에 초점을 맞추었다. Chang 등(2005) 그리고 Wang 등(2006)과 Yin 등(2006)의 연구도 얼굴 표정 인식을 위하여 3차원 이미지를 사용한 전형적인 예이다. 3차원 얼굴 모델에 기반한 정서 인식 기법의 이점은 얼굴을 촬영한 면(전면 혹은 측면)에 관계없이 얼굴 표정 인식이 가능하다는 점이다. 따라서, 2차원 이미지를 주로 사용하고 있는

연구들은 실생활 환경에서 자연스러운 얼굴 표정을 인식하는데 제약이 있다.

마지막으로, 추가적인 몇 가지 한계점들은 다음과 같다. 먼저, Zeng 등(2009)이 개관한 29개 시스템 중 6개만이 실시간으로 작동한다. 다음으로, 대다수의 정서 인식 연구들은 자연스러운 비디오 장면이 아닌 비디오에서 임의로 선정된 일부 구간을 정서 분석에 사용하고 있다. 또한, 많은 연구에서 머리 움직임, 시선 방향, 몸짓 단서를 분석에서 제외하고 있다. 행동 단서의 정확한 판단을 위해서는 얼굴 표정과 함께 몸 움직임을 고려하는 것이 필요하다. 그러나 최근, 몇 편의 상대적으로 적은 연구들에서 인식율을 향상시키기 위하여 얼굴 표정과 함께 머리 움직임 정보(예., Cohn et al, 2004; Ji et al., 2006; Zhang & Ji, 2005; Zhu, & Ji, 2006; Valstar et al., 2007), 제스처 정보(예., Balomenos et al., 2005; Gunes & Piccardi, 2005a, 2005b; Valstar et al., 2007), 몸의 자세 정보(예., Kapoor et al., 2007; Kapoor & Picard, 2005; Gunes & M. Piccardi, 2005)를 조합하여 사용하고 있다. 정서 인식 연구들의 실용화를 고려한다면 이상에서 언급된 한계점들은 개선되어야 할 중요한 사항들이다.

3. 생리신호에 기반한 정서 인식 방법

정서에서 ‘생리적 변화’를 강조한 James의 이론의 영향을 받아, 일부 정서 인식 연구들은 정서에 따른 생리적 반응을 확인하는데 초점을 맞추고 있다. 본 장에서는 생리신호에 기반한 정서 인식 방법을 이론적 배경과 연구 결과를 중심으로 개관하고자 한다.

3.1. 이론적 배경

정서 경험에서 ‘정서 표현’을 강조한 Darwin의 이론에 비해 James의 이론은 ‘생리적 변화(physiological changes)’를 강조했다. James(1884)는 자극에 대한 반응으로 나타나는 ‘생리적 변화’를 지각하는 것 자체가 정서라고 해석했다. James와 동시대 인물이며 심리생리 연구에서 선구자인 Lange는 몇몇 일반 정서(예., 슬픔, 기쁨, 분노, 공포)에 수반되는 생리적 반응을 연구했다. James와 마찬가지로, Lange 도 정서적인 경험은

말초 신경계에서 생리적 반응으로 표현된다는 사실을 강조했다.

후에, 이 두 사람의 이론은 “James-Lange 이론”으로 발전하였다. 이 이론은 자율 신경계(Autonomic Nervous System; ANS)의 일부인 교감신경계(Sympathetic Nervous System; SNS)에서 나타나는 “변화”로 정서를 설명하고자 하였다. 정서 경험에서 생리적 변화를 지지하는 대부분의 연구자들은 각 정서에 따른 전형적인 생리 반응이 존재한다고 가정한다.

정서의 ‘생리적 변화’를 강조한 James의 관점에 의해 영향을 받은 일부 정서 인식 연구들은 각 정서와 관련된 생리적 변화 유형을 분석하여 정서를 탐색하고자 하였다.

3.2. 생리신호에 기반한 정서 인식 연구 개관

생리적 상태를 측정하기 위하여 사용되는 대부분의 방법은 뇌, 심장, 근육, 그리고 피부에서 발생하는 전기 신호를 기록하는 것이다. 구체적으로는, 근육 활동을 측정하는 근전도(Electromyogram; EMG), 피부 표면의 땀선 활동을 측정하는 피부전기 활성(Electrodermal Activity; EDA), 심장 활성화를 측정하는 심전도(Electrocardiogram; EKG or ECG), 눈 움직임을 측정하는 전기안구도(Electrooculogram, EOG)이다. 또한, 뇌 활성화를 기록하는 뇌파전위기록(Electroencephalography; EEG) 그리고 더 최신의 뇌 영상 촬영 기법(functional Magnetic Resonance Imaging, fMRI 등)들도 포함된다.

생리 반응에 기반한 정서 인식 연구들을 살펴보기에 앞서 생리반응이 가지는 다음과 같은 특징을 살펴보는 것은 도움이 될 것이다(Andreassi, 2007).

자극-반응(SR) 특정성[Stimulus-response(SR) specificity] :

이것은 특정한 자극 상황에서 참가자는 특정한 생리적 반응 양식(patterns)을 경험한다는 이론이다. 많은 연구자들과 Ekman 등(1983)은 자율신경계의 특정성으로 인해 정서가 구분될 수 있다는 사실을 밝혀냈다.

개인-반응(IR) 특정성[Individual-response(IR) specificity] :

이것은 자극-반응 특정성을 보완하는 이론이지만 두드러진 차이가 있다. 자극-반응 특정성은 대부분 사람들의 반응 패턴이 유사하다고 가정하지만, 개인-반응

특정성은 여러 다른 자극에 대한 개인의 반응은 비교적 일관적으로 나타난다고 가정한다.

습관화 그리고 반동(rebound) : 이것은 자극에 의해 발생하는 두 가지 현상이다. 먼저, 습관화는 자극이 반복적으로 제시될 때, 생리적 반응성이 감소하는 현상이다. 반동은 자극이 제시되면 생리적 신호는 변하지만, 자극 제시 한참 후에는 생리적 반응이 원래 상태로 되돌아오는 현상이다(Andreassi, 2007).

이러한 생리 반응의 특성 중 개인-반응 특정성은 정서 인식에서 중요한 의의를 가진다. 왜냐하면 일부 연구자들은 개인-반응 특정성에 근거하여 정확한 정서 인식을 위해서는 각 피험자에게 맞게 조정된 모델을 사용해야 한다고 제안하기 때문이다. 연구자들은 이러한 제안을 실현할 수 있는 효과적인 방법을 찾고 있다. 예를 들면, Nasoz 등(2004)은 그들의 MAUI (Multimodal Affective User Interface) 시스템을 이용하여 생리적 신호를 정서에 매핑할 수 있는 여러 가지 알고리즘을 비교했다. 더 최근에, 이 연구자들은 사용자 의존적으로 생리 신호(심박율과 피부 전도)와 정서 차원[정서가(valence)와 각성(arousal)]을 매핑을 시켜주는 심리생리 정서맵(Psychophysiological Emotional Map; PPEM)(Villon & Lisetti, 2006)을 제안했다.

표 2에 생리신호 기반 정서 인식 연구들을 요약하여 제시하였다(Calvo & D'Mello, 2010). 표에는 각 연구에서 사용된 신호, 특징(feature) 추출 기법, 그리고 정서 모델(개별 정서 또는 차원 정서)이 포함되어 있다. 생리 신호에 기반한 정서 인식의 일부 연구들은 차원 모델을 채택하고 있지만, 대부분의 연구들은 개별 정서 모델에 기반하고 있다. 생리 신호에 기반하여 개별 정서를 구분하고 있는 연구들은 2~10 범주의 정서 분류를 포함한다. 예를 들면, Nasoz 등(2004)은 심박 및 피부 전도 관련 신호를 이용하여 6가지 정서(슬픔, 분노, 공포, 놀람, 좌절, 그리고 기쁨)를 분류하였다. Calvo 등(2009)은 심박, 피부 전도, 그리고 근육 관련 신호를 이용하여 중성, 분노, 증오, 슬픔, 정신적인 사랑, 낭만적 사랑, 기쁨, 그리고 숭배 등의 8가지 정서를 구분하였다.

표 2에 제시된 생리 신호 기반 정서 재인 연구들의 한계점을 정리하면 다음과 같다. 첫째, 연구들 중 일부만이 실시간 특징 추출을 목표로 하고 있다. 앞서 말했듯이, 실생활에서 적용을 고려할 때 실시간으로

작동하는 시스템이 유용하다. 둘째, 정서 인식을 위해 사용되는 생체 신호가 제한적이라는 점이다. 연구의 대부분은 정서 인식에서 피부 전도도와 심박율을 핵심적인 생리 지표로 사용하고 있다. 그러나, 핵심적으로 사용하고 있는 생리 지표인 피부 전도도와 심박율은 정서를 정의하는 유일한 지표는 아니다. 정서에 의해 조절되는 생리신호는 다양하기 때문에 피부 전도도와 심박율을 정서 탐색의 대표적 지표로 사용하는 것은 문제가 될 수 있다.

이러한 제한적인 생리 신호에 대한 대안으로 보다 다양한 생리 신호를 정서 재인에 포함하는 노력이 필요하다. 지난 20년 간 신경 과학은 정서 처리와 관련된 신경 기질을 규명하기 위한 새로운 기술과 기법들을 제안함으로써 정서 연구에 기여해 왔다. 이러한 노력의 일환으로 개별 정서와 관련된 특정적 신경 기질이 규명되었다(Damasio et al., 2000; for review Phan et al., 2002; Murphy et al., 2003). Phan 등(2002)은 정서와 관련된 뇌 기능 메타 연구에서 공포는 편도체와 슬픔은 뇌량밑 대상과 관련 있다고 보고하였다. Murphy 등(2003)은 그들의 메타 연구에서 공포는 편도체, 혐오는 뇌섬엽/창백핵, 그리고 분노는 외측 안와전두피질의 활성화와 연관되는 된다고 보고하였다. 이에 추가적으로, 최근의 연구는 기계 학습을 이용하여 뇌 신호에 기반한 특정 정서를 구분하는데 성공하였다(Sitaram et al., 2010). 구체적으로, Sitaram 등(2010)은 Support Vector Machine(SVM) 기법을 이용하여 fMRI 신호로부터 행복 또는 혐오와 같은 개인의 정서 상태를 분류하였다.

정서 인식 분야에서 fMRI 신호 이외에 다른 생체 정보를 사용하는 것도 가능하다. 최근에는 얼굴 온도에 기반하여 정서를 구분하려는 노력들이 시도되고 있다(Khan et al., 2006; Nakanishi & Imai-Matsumura, 2008; Sugimoto et al., 2000; Yoshitomi et al., 2002). Yoshitomi 등(2002)은 Neural Network(NN) 기법을 이용하여 얼굴 온도에 기반한 중립, 행복, 놀람, 그리고 슬픔 정서를 구분하였다. Khan 등(2006)은 주성분 분석(Principal Components Analysis; PCA) 기법을 이용하여 중립, 기쁨, 슬픔, 혐오, 그리고 공포 정서 등 5가지 정서를 분류하였다. 추가적으로, Sugimoto 등(2000)은 얼굴 표정과 얼굴 온도 정보를 통합적으로 사용하여 참가자가 비디오 게임을 하는 동안 경험하는 정서를 구분하였다. 이처럼 뇌 활성화 또는 얼굴 온도에 기반하여 정서를 분류하는 것이 가능하다는

점을 감안하면, 이들 생체 신호는 정서 재인에서 심박이나 피부 전도에 대한 새로운 대안을 제공할 수 있다. 뇌 반응 그리고 얼굴 온도 등 보다 다양한 생체 신호를 정서 재인에 포함한다면 정서 인식률을 향상시킬 수 있는 방안이 모색될 수 있을 것으로 본다.

셋째, 연구들이 사용하고 있는 정서 유발 방법은 대부분 자기 연출(self-enactment)이다. 이는 참여자에게 개인적인 정신적 심상을 제시하여 정서를 유발하는 하는 기법이다. 심리생리학자들은 정서를 유발하기 위해 사진, 비디오, 그리고 스크립트 등을 이용하고 있지만, 실제적인 정서를 유발하기 위해서는 피험자가 실제 정서 유발 상황을 경험할 수 있는 현실적인 시나리오[예., 상금을 받는 과제에 직접 참여하여 기쁨을 경험하는 것(Kleef et al., 2010)]를 사용해야 한다.

연구들이 가지는 이러한 한계점과 더불어, 정서 인식에 생리신호를 사용하는 데는 몇 가지 도전들이 존재한다. 먼저, 생리신호를 측정하는 기법들은 침습적이고 움직임에 의해 발생하는 잡음에 상당히 민감하다는 것이다. 또한, 시간 해상도와 관련된 도전도 존재한다. 일반적으로 생리신호를 측정하는 표준 장비는 1,000Hz에서 신호를 기록할 수 있다. 그러나 신호의 실제 시간 해상도는 신호에 따른 특징(features)과 앞에서 언급한 자극-반응 양식에 따라 달라진다 (Andreassi, 2007). 결론적으로, 생리신호에 기반한 정서 인식 연구들을 실생활에 적용하기 위해서는 앞서 제시한 한계점들을 해결하려는 노력과 생리신호를 사용하는데 있어 나타나는 문제점들을 해결할 수 있는 진보된 데이터 획득 기법이 개발되어야 할 것이다.

4. 다중모달리티에 기반한 정서 인식 방법

최근에는 얼굴 표정이나 생리신호 만을 사용하는 단일모달리티 기법에 비해 여러 채널에서 획득한 정보를 통합적으로 사용하는 다중모달리티 기법이 정서 인식 분야에서 지지를 얻고 있다. 본 장에서는 다중모달리티에 기반한 정서 인식 방법을 이론적 배경과 연구 결과를 중심으로 개관하고자 한다.

표 2. 생리 신호에 기반한 정서 인식 연구 요약(Calvo & D'Mello, 2010)

연구자	사용된 신호	Classifier	정서 분류	정확율 (%)
Picard et al. (2001)	EKG, SC, EMG	Fisher Proj with SFFS/KNN	8	Best acc: 81.25
Wagner et al. (2005)	EKG, SC, EMG	PCA	4	80-90
Kim et al. (2004)	EKG, ST, SC	SVM	3 & 4	78 for 3 62 for 4
Nasoz et al. (2004)	SC, HR, ST	KNN, DFA, MBP	6+강도	71, 74, 83
Villon & Lisetti (2006)	HR+SC	신호 처리 방법	정서 차원	-
Calvo et al. (2009)	EKG, SC, EMG	SVM	8	68-80
Bailenson et al. (2008)	ECG, SC, + Face	SVM, Logistic Reg	2+강도	-
Liu et al. (2008)	ECG, SC, EMG, ST	SVM	3	83
Vyzas & Picard (1998)	EMG, BVP, GSR, RESP	SFFS/KNN	8	40-46
Haag et al. (2004)	EKG, EMG, SC, ST, BVP, RESP	MLP	정서 차원	10-20 bands: Valence (90-97), Arousal (63-90)
AlZoubi et al. (2009)	EEG	SVM	10	33-55
Heraz & Frasson (2007)	EEG	Nearest Neighbours, DT, Bagging	정서 차원 (세차원)	-

4.1. 이론적 배경

Darwin과 James의 이론에 의해 영향을 받아 ‘정서의 표현’과 ‘생리적 변화’를 지지하는 연구자들은 정서를 경험하는 동안 다양한 생리적 그리고 행동적 반응 체계가 활성화된다고 제안한다. 예를 들면, 분노

경험은 특정한 얼굴 표정, 목소리, 그리고 신체 표현을 통해서 명백히 드러나며, 심장 박동의 증가와 같은 생리적인 변화를 함께 수반한다.

4.2. 다중모달리티에 기반한 정서 인식 연구 개관

인간-컴퓨터 상호작용에서 다중모달리티를 이용한 기법의 이점이 널리 알려지고 있다. 이에 따라 다중모달리티 기법은 기존에 사용되던 단일모달리티 기법에 비해 정서 인식에서 더 진보된 방법으로 여겨지고 있다. 다중모달리티를 이용한 정서 인식에서 여러 채널에서 획득한 신호를 융합하는 방법은 다음과 같다(Sharma et al., 1998).

데이터 융합(Data Fusion)은 각 신호에서 획득한 원 데이터(raw data)에 대해 적용될 수 있는 방법이며, 신호가 동일한 시간 해상도를 가질 때만 적용될 수 있다. 일반적으로 이 기법은 동일한 측정 장치로부터 획득한 생리신호를 통합하는데 사용되지만 텍스트가 포함된 비디오 신호를 통합하는 데는 사용될 수 없다. 이 기법은 일반적으로 사용되지 않는 데 왜냐하면 센서의 배열 상태가 어긋나거나 기능 불량인 경우 센서에서 발생하는 잡음에 대해 민감한 방법이기 때문이다.

특징 융합(Feature Fusion)은 각 신호에서 추출된 일련의 특징들(feature)에 대해 적용될 수 있다. 이 접근법은 다중모달리티를 이용한 인간-컴퓨터 상호작용에서 흔히 사용되는 기법이다. 예를 들면, 이 기법은 정서 분류를 위해 개발된 Augsburg Biosignal Toolbox(Wagner et al., 2005)에서 사용된다. EKG나 EMG 등의 신호에서 우선적으로 사용되는 특징은 평균, 중앙치, 표준 편차, 최대값, 그리고 최소값이며, 신호에 따른 몇몇 특징적인 특징들이 함께 사용되기도 한다. 특징들은 각 신호에서 개별적으로 추출된 후에 여러 신호에 걸쳐 통합된다.

결정 융합(Decision Fusion)은 각 신호에 대한 정서 분류 결과를 통합하는 방법이다. 정서 상태는 처음에 각각의 신호에 의해 분류되며, 그 다음에 이 결과들은 여러 신호를 망라하여 전체적인 관점에서 통합된다. 이 기법이 다중모달리티를 이용한 인간-컴퓨터 상호

작용에서 가장 일반적으로 사용되는 접근법이다(Ekman et al., 1990; Russell, 2003).

다중모달리티를 이용한 연구의 이점에도 불구하고, 실제적으로는 소수의 연구만이 세 개 또는 그 이상의 모달리티를 정서 인식에 사용하고 있다. 표 3은 다중모달리티에 기반한 정서 인식 연구들을 요약하여 제시하고 있다. 먼저, Scherer와 Ellgring(2007)은 14개의 정서를 구분하기 위하여 얼굴 표정, 음성 특성, 그리고 몸 움직임(자세 그리고 몸짓)을 분석하였다. 더 최근에, Castellano 등(2008)은 8개의 정서를 구분하기 위하여 얼굴 특성, 언어 정보, 그리고 몸짓을 측정하였다. 두 연구 모두에서 단일모달리티에 비해 다중모달리티를 조합하여 사용한 경우에 정서 인식율은 증가하였다.

다중모달리티에 기반한 일부 연구에서는 문맥적 정보가 활용되고 있다. 먼저, Kapoor와 Picard(2005)는 아동의 흥미 수준을 추론하기 위하여 얼굴 움직임, 자세, 그리고 맥락적인 정보를 사용하는 시스템을 개발하였다. 이 연구에서 단일모달리티에 비해 다중모달리티를 조합한 기법은 인식율에서 큰 향상을 가져왔다. 후에, Kapoor 등(2007)은 아동이 ‘하노이 탑 과제’를 수행하는 동안 경험할 수 있는 좌절감을 예측하기 위하여 얼굴, 자세, 그리고 맥락적 정보와 함께 피부 전도도를 이용하는 시스템을 개발하였다. 이 시스템이 좌절감을 탐지하는 정확율은 79%로 나타났다. 최근의 연구에서, Arroyo 등(2009)은 학교 상황에서 학생의 자신감, 좌절, 흥분, 그리고 흥미 수준을 파악하기 위하여 문맥, 얼굴 특징, 자세, 피부 전도도, 그리고 마우스를 통해 측정되는 압력 정보 모두를 정서 인식에 사용하였다. 그 결과, 얼굴 특징과 문맥 정보를 조합하여 사용한 경우가 가장 최적임이 드러났다. 정서는 실제생활에서 사회적 기능을 가지는 문맥 의존적 특성이 있다는 점을 고려할 때 문맥 효과를 고려한 이러한 연구들은 정서 인식 분야에서 그 의의가 크다고 할 수 있다.

그러나, 이 연구들에도 몇 가지 한계점은 존재한다. 첫째, 다중모달리티 모델이 단일모달리티 모델과 비교하여 정서 인식율에서 향상을 가져왔는지를 평가하는 것은 어렵다. 이는 단일모달리티와 다중모달리티 모델에서 사용하는 특징(feature)의 수가 동일하지 않기 때문이다.

다음으로, 또한 이러한 연구들이 가정하고 있는 ‘정

서 **경험**은 신체적으로 **표현**된다’는 가정은 아마도 가장 보편적이고 잠재적으로 문제가 있는 가정일 것이다. 정서 인식 시스템은 화난 얼굴(**표현**)을 탐지했을 때 사용자가 화가났다(**경험**)고 추론한다. 그러나, 정서 표현으로써 대표적인 얼굴 표정은(Ekman, 1992) 많은 경우 본인이 보고한 정서 상태와 완전히 일치하지는 않는다(Fridlund et al., 1987; Ruch, 1995). 따라서, 정서 경험과 표현을 일대일로 대응시키는 것은 어렵다. 생리적이고 신체적인 신호는 인간의 정서 상태에 대해 중요한 정보를 제공해 주지만, 신체 신호가 정서의 지표로써 가지는 정보적 가치는 명확히 정의하기는 여전히 어려운 문제이다.

표 3. 다중모달리티에 기반한 정서 인식 연구 요약

연구자	사용된 신호	정서	정확율(%)
Scherer & Ellgring (2007)	face, voice, posture, and gesture	14	-21 facial features: 52.2 -16 acoustic parameters: 52.5 -37-channel model: 79 -10-channel model: 54
Castellano et al. (2008)	face, speech, contours, and gestures	8	-face: 48.3 -gestures: 67.1 -speech: 57.1 -multimodality: 78.3
Kapoor & Picard (2005)	face posture, and context	1	-upper face: 67 -lower face: 53 -contextual information: 57 -posture features: 82 -multimodality: 86
Kapoor et al. (2007)	the face, posture, context, +SC	1	79
Arroyo et al. (2009)	context, facial features, seat pressure, +SC	4	·face + context -confidence: 52 -interest: 29 -excited: 69 ·seat pressure + context -frustration: 46

5. 결론

본 연구의 2, 3, 4장에서는 얼굴 표정, 생리 신호, 그리고 다중모달리티에 기반한 정서 재인 방법을 각각

개관하였다. 각 장에서는 정서 재인 방법이 근거하고 있는 이론적 배경과 연구 결과, 그리고 한계점을 간단히 소개하였다. 결론에서는 앞서 기술한 연구 방법들이 가지는 한계점을 고려하여, 한계점을 부분적으로 극복할 수 있는 진보된 정서 재인 방법을 제안하고자 한다.

첫째로, 본 연구에서는 차원 정서 모델에 기반한 정서 재인 연구를 제안한다. 많은 정서 컴퓨팅 연구자들은 정서가 분노, 행복과 같은 개별 범주로 구분될 수 있다고 가정한다. 대부분의 경우 정서를 활용한 시스템을 실제로 응용하는데 있어 정서의 범주화가 요구된다. 그러나, **정서가 개별적인지 아니면 차원적인지**에 관해서는 여전히 논쟁이 끊이지 않고 있으며(Russell, 2003), 어떤 연구자들은 정서를 활용한 인터페이스를 구축하는데 차원적 접근이 더 적합하다고 주장한다(Barrett, 2006). 정서의 차원 이론가들은 정서는 개별적이기 보다는 정서가(valence) 그리고(또는) 각성(arousal) 차원으로 설명할 수 있는 연속적인 개념이라고 주장했다(Wundt, 1905; Russell, 1983). 정서 재인 연구에서 차원 개념을 채택한다면 구분하기 애매한 정서(예를 들면, 기쁨과 흥미 사이에 존재하는 정서)를 차원 상에서 설명할 수 있다. 결과적으로, 개별 범주 모델에 비해 차원 모델에서 보다 많은 정서의 설명이 가능하다. 최근에, Ciceri 등(2007)은 정서 재인 연구에서 차원 이론을 채택하였고, 정서가(valence), 각성(arousal), 새로움(novelty), 그리고 대처(coping) 차원 등 4차원으로 다양한 정서를 구분하고자 하였다. 정서 재인 연구에서 차원 개념을 채택한다면 개별 범주 모델에 비해 보다 많은 정서의 구분과 재인이 가능할 것이다.

다음으로 본 연구는 정서 재인에서 인지적 요소를 포함하도록 제안한다. 많은 정서 컴퓨팅 연구자들은 **정서**를 단순히 **수동적**이라고 가정한다. 이러한 가정을 지지하는 연구자들은 단순히 자극을 제시하면 정서 반응이 유발되는 것으로 간주한다. 이러한 가정은 정서 처리에서 인지적 효과를 무시하는 것이다. 정서는 단순히 자극에 의해 나타나는 수동적 상태가 아니라, 인지적 평가에 의해 조절되는 적극적 상태이다. 그러나, 현재 대부분의 정서 재인 연구들은 인지적 효과가 배제된 환경에서 수집된 데이터를 시스템 훈련과 테스트에 사용하고 있다. 정서 표현은 인지적 평가에 따라 다른 의미를 가질 수 있기 때문에, 인지적 요인을 고려하지 않은 데이터를 이용하여 정서 재인 시스템

을 훈련시킨다면 그 시스템은 모든 상황에서 일반화되기 어렵다. 따라서, 보다 진보된 정서 재인 방법은 인지적 요소를 포함할 필요가 있다. 맥락적 상황, 개인의 성격, 그리고 과거 경험 등은 정서 처리에 영향을 미치는 주요한 인지적 요소들이다. 맥락적 요소를 고려하여 과제의 길이, 하루 중 과제를 수행하는 시간, 과제의 복잡성 및 유형 등을 파악한다면 정서 상태를 추정하는 것이 보다 용이할 수 있다. 또한, 개인의 성격(외향적, 공격적, 강박적, 등) 그리고 개인의 과거 경험(과거의 실패 또는 성공 경험, 현재 과제와 관련된 정서 상태)을 정서 재인에서 고려한다면 보다 신뢰로운 정서 재인이 가능할 것으로 보인다(Hudlicka, 2002; Hudlicka, 2003).

현재의 정서 재인 방법에 대한 부분적인 대안으로 본 연구의 본문에서는 뇌 반응 또는 얼굴 온도 등의 생체 신호를 포함하는 것을 제안하였다. 결론에서는 보다 다양한 정서를 구분하기 위하여 정서의 차원 개념에 기반한 정서 재인 방법과 인지적 요인을 포함하는 대안을 제시하였다. 종합적으로, 본 연구에서 제안한 정서 재인 방법은 보다 다양한 생체 신호를 포함하고, 차원 개념에 기반하며, 인지적인 요인을 고려하는 통합적인 정서 재인 모델이다.

참고문헌

- Afzal, S. & Robinson, P. (2009). Natural Affect Data: Collection & Annotation in a Learning Context, *Proc. Third Int'l Conf. Affective Computing and Intelligent Interaction and Workshops*, 1-7.
- AlZoubi, O., Calvo, R. A., & Stevens, R. H. (2009). Classification of EEG for Emotion Recognition: An Adaptive Approach, *Proc. 22nd Australasian Joint Conf. Artificial Intelligence*, 52-61.
- Andreassi, J. L. (2007). *Human Behaviour and Physiological Response*. Taylor & Francis.
- Arroyo, I., Cooper, D. G., Bursleson, W., Woolf, B. P., Muldner, K., & Christopherson, R. (2009). Emotion Sensors Go to School. *Proc. 14th Conf. Artificial Intelligence in Education*, 17-24.
- Ashraf, A. B., Lucey, S., Cohn, J. F., Chen, T., Ambadar, Z., Prkachin, K., Solomon, P., & Theobald, B. J. (2007). The Painful Face: Pain Expression Recognition Using Active Appearance Models, *Proc. Ninth ACM Int'l Conf. Multimodal Interfaces (ICMI '07)*, 9-14.
- Bailenson, J. N., Pontikakis, E. D., Mauss, I. B., Gross, J. J., Jabon, M. E., Hutcherson, C. A. C., Nass, C., & John, O. (2008). Real-Time Classification of Evoked Emotions Using Facial Feature Tracking and Physiological Responses, *Int'l Journal of Human-Computer Studies*, 66, 303-317.
- Balomenos, T., Raouzaoui, A., Ioannou, S., Drosopoulos, A., Karpouzis, K., & Kollias, S. (2005). Emotion Analysis in Man-Machine Interaction Systems. *LNCS 3361*, 318-328.
- Barrett, L. (2006). Are Emotions Natural Kinds? *Perspectives on Psychological Science*, 1, 28-58.
- Bartlett, M.S., Littlewort, G., Braathen, P., Sejnowski, T. J., & Movellan, J. R. (2003). A Prototype for Automatic Recognition of Spontaneous Facial Actions, *Advances in Neural Information Processing Systems*, 15, 1271-1278.
- Bartlett, M.S., Littlewort, G., Frank, M. Lainscsek, C., Fasel, I., & Movellan, J. (2005). Recognizing Facial Expression: Machine Learning and Application to Spontaneous Behavior, *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition (CVPR '05)*, 568-573.
- Bartlett, M. S., Littlewort, G., Frank, M. G. Lainscsek, C., Fasel I., & Movellan., J. (2006). Fully Automatic Facial Action Recognition in Spontaneous Behavior, *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition (AFGR '06)*, 223-230.
- Calvo, R. A. & D'Mello, S. (2010). Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. *IEEE Transactions on Affective Computing*, 1(1), 18-37.
- Calvo, R. A., Brown, I., & Scheduling, S. (2009). Effect of Experimental Factors on the Recognition of Affective Mental States through Physiological Measures, *Proc. 22nd Australasian Joint Conf. Artificial Intelligence*.
- Castellano, G. Kessous, L., & Caridakis, G. (2008). *Emotion Recognition through Multiple Modalities: Face, Body Gesture, Speech, Affect and Emotion in Human-Computer Interaction* (pp. 92-103), Springer.

- Chang, Y., Hu, C., Feris, R., & Turk, M. (2006). Manifold Based Analysis of Facial Expression. *Journal of Image and Vision Computing*, 24 (6), 605-614.
- Chang, Y., Hu, C., & Turk, M. (2004). Probabilistic Expression Analysis on Manifolds. *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition (CVPR '04)*, 2, 520-527.
- Chang, Y., Vieira, M., Turk, M., & Velho, L. (2005). Automatic 3D Facial Expression Analysis in Videos, *Proc. IEEE Int'l Workshop Analysis and Modeling of Faces and Gestures (AMFG '05)*, 3723, 293-307.
- Ciceri, M. R. & Balzarotti, S. (2007). Multidimensional Emotional Appraisal Semantic Space(MEAS): Evaluating HM Affective Interactions, *Knowledge-Based Intelligent Information and Engineering Systems*, 4693, 395-402.
- Cohen, L., Sebe, N., Garg, A., Chen, L., & Huang, T. (2003). *Facial Expression Recognition from Video Sequences: Temporal and Static Modeling*. *Computer Vision and Image Understanding*, 91(1-2), 160-187.
- Cohn, J. F. (2006). Foundations of Human Computing: Facial Expression and Emotion. *Proc. Eighth ACM Int'l Conf. Multimodal Interfaces (ICMI '06)*, 233-238.
- Cohn, J. F., Reed, L. I., Ambadar, Z., Xiao, J., & Moriyama, T. (2004). Automatic Analysis and Recognition of Brow Actions and Head Motion in Spontaneous Facial Behavior, *Proc. IEEE Int'l Conf. Systems, Man, and Cybernetics (SMC '04)*, 1, 610-616.
- Cohn, J. F. & Schmidt, K. L. (2004). The Timing of Facial Motion in Posed and Spontaneous Smiles. *Int'l Journal of Wavelets, Multiresolution and Information Processing*, 2, 1-12.
- Damasio A. R. et al. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience*, 3(10), 1049-1056.
- Darwin, C. (2002). *Expression of the Emotions in Man and Animals*. New York/Oxford:Oxford Univ. Press.
- Davidson, R. J., Ekman, P., Saron, C. D., Senulis, J. A., & Friesen, W. V. (1990). Approach-Withdrawal and Cerebral Asymmetry: Emotional Expression and Brain Physiology I. *Journal of Personality and Social Psychology*, 58, 330-341.
- Ekman, P. (1971). *Universals and Cultural Differences in Facial Expressions of Emotion*. Nebraska: Univ. of Nebraska Press.
- Ekman, P. (1984). *Expression and the Nature of Emotion, Approaches to Emotion*, In K. Scherer & P. Ekman (Eds.), 319-344, Erlbaum.
- Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, 6, 169-200.
- Ekman, P. & Friesen, W. V. (1978). *Facial Action Coding System: A Technique for the Measurement of Facial Movement: Investigator's Guide 2 Parts*. Consulting Psychologists Press.
- Ekman, P. & Friesen, W. V. (2003). *Unmasking the Face*. Malor Books.
- Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial Action Coding System. A Human Face*.
- Ekman, P., Levenson, R., & Friesen, W. (1983). Autonomic Nervous System Activity Distinguishes among Emotions, *Science*, 221, 1208-1210.
- Ekman, P., Friesen, W., & Davidson, R. (1990). The Duchenne Smile—Emotional Expression and Brain Physiology .2. *Journal of Personality and Social Psychology*, 58, 342-353.
- El Kaliouby, R. & Robinson, P. (2004). Real-Time Inference of Complex Mental States from Facial Expression and Head Gestures, *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition (CVPR '04)*, 3, 154.
- Fasel, B., Monay, F., & Gatica-Perez, D. (2004). Latent Semantic Analysis of Facial Action Codes for Automatic Facial Expression Recognition, *Proc. Sixth ACM Int'l Workshop Multimedia Information Retrieval (MIR '04)*, 181-188.
- Fridlund, A. J., Ekman, P., & Oster, H.. (1987). *Facial Expressions of Emotion, Nonverbal Behavior and Communication*, A. W. Siegman and S. Feldstein (Eds.), (pp. 143-223), Erlbaum.
- Frijda, N. (1987). Emotion, Cognitive Structure, and Action Tendency. *Cognition and Emotion*, 1, 115-143.
- Gu, H. & Ji, Q. (2004). An Automated Face Reader for Fatigue Detection, *Proc. IEEE Int'l Conf. Automatic*

- Face and Gesture Recognition (AFGR '04)*, 111-116.
- Gunes, H. & Piccardi, M. (2005a). Affect Recognition from Face and Body: Early Fusion versus Late Fusion. *Proc. IEEE Int'l Conf. Systems, Man, and Cybernetics (SMC '05)*, 3437-3443.
- Gunes, H. & Piccardi, M. (2005b). Fusing Face and Body Display for Bi-Modal Emotion Recognition: Single Frame Analysis and Multi-Frame Post-Integration. *Proc. First Int'l Conf. Affective Computing and Intelligent Interaction (ACII '05)*, 102-111.
- Guo, G. & Dyer, C. R. (2005). Learning from Examples in the Small Sample Case: Face Expression Recognition, *IEEE Transactions on Systems, Man, and Cybernetics Part B*, 35 (3), 477-488.
- Haag, A., Goronzy, S., Schaich, P., & Williams, J. (2004). *Emotion Recognition Using Bio-Sensors: First Steps towards an Automatic System. Affective Dialogue Systems*, 36-48, Springer.
- Heraz, A. & Frasson, C. (2007). Predicting the Three Major Dimensions of the Learner's Emotions from Brainwaves. World Academy of Science, *Engineering and Technology*, 25, 323-329.
- Hudlicka, E. (2002). Increasing SIA architecture realism by modeling and adapting to affect and personality. In Da utenhahn, K., Bond, A. H., Canamero, L., Edmonds, B. (Eds.), *Multiagent Systems, Artificial Societies, and Simulated Organizations*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Hudlicka, E. (2003). To feel or not to feel: The role of affect in human - computer interaction, *International Journal of Human-Computer Studies*, 59, 1-32.
- Ioannou, S., Raouzaoui, A., Tzouvaras, V., Mailis, T., Karpouzis, K., & Kollias, S. (2005). Emotion Recognition through Facial Expression Analysis Based on a Neurofuzzy Method, *Neural Networks*, 18, 423-435.
- Izard, C. (1971). *The Face of Emotion*. Appleton-Century-Crofts.
- Izard, C. (1994). Innate and Universal Facial Expressions: Evidence from Developmental and Cross-Cultural Research, *Psychological Bulletin*, 115, 288-299.
- James, W. (1884). What Is an Emotion? *Mind*, 9, 188-205.
- Ji, Q., Lan, P., & Looney, C. (2006). A Probabilistic Framework for Modeling and Real-Time Monitoring Human Fatigue, *IEEE Systems, Man, and Cybernetics Part A*, 36(5), 862-875.
- Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive Database for Facial Expression Analysis, *Proc. IEEE Int'l Conf. Face and Gesture Recognition (AFGR '00)*, 46-53.
- Kapoor, A. & Picard, R. W. (2005). Multimodal Affect Recognition in Learning Environments, *Proc. 13th Ann. ACM Int'l Conf. Multimedia*, 677-682.
- Kapoor, A., Burleson, B., & Picard, R. W. (2007). Automatic Prediction of Frustration. *Int'l Journal of Human-Computer Studies*, 65, 724-736.
- Khan, M. M., Ingleby, M., & Ward, R. D. (2006). Automated Facial Expression Classification and Affect Interpretation Using Infrared Measurement of Facial Skin Temperature Variations, *ACM Transactions on Autonomous and Adaptive Systems*, 1 (1), 91 - 113.
- Kim, K., Bang, S., & Kim, S. (2004). Emotion Recognition System Using Short-Term Monitoring of Physiological Signals. *Medical and Biological Engineering and Computing*, 42, 419-427.
- Kleef, G. A. V., De Dreu, C. K. W., & Manstead, A. S. R. (2010). An Interpersonal Approach to Emotion in Social Decision Making: The Emotions as Social Information Model, *Advances in Experimental Social Psychology*, 42, 45-96.
- Lee, C. & Elgammal, A. (2005). Facial Expression Analysis Using Nonlinear Decomposable Generative Models, *Proc. Second IEEE Int'l Workshop Analysis and Modeling of Faces and Gestures (AMFG)*.
- Littlewort, G. C., Bartlett, M. S., & Lee, K. (2007). Faces of Pain: Automated Measurement of Spontaneous Facial Expressions of Genuine and Posed Pain, *Proc. Ninth ACM Int'l Conf. Multimodal Interfaces (ICMI '07)*, 15-21.
- Liu, C., Conn, K., Sarkar, N., & Stone, W. (2008). Physiology-Based Affect Recognition for Computer-Assisted Intervention of Children with Autism Spectrum Disorder, *Int'l Journal of Human-Computer*

- Studies*, 66, 662-677.
- Lucey, S. Ashraf, A. B., & Cohn, J. F. (2007). *Investigating Spontaneous Facial Action Recognition through AAM Representations of the Face, Face Recognition*. In K. Delac, and M. Grgic (Eds.), 275-286, I-Tech Education and Publishing.
- Murphy, F. C., Nimmo-Smith I., & Lawrence, A. D. (2003). Functional neuroanatomy of emotions: A meta-analysis, *Cognitive, Affective, & Behavioral Neuroscience*, 3(3), 207-233.
- Nakanishi, R. & Imai-Matsumura, K. (2008). Facial Skin Temperature Decreases in Infants with Joyful Expression. *Infant Behavior and Development*, 31, 137-144.
- Nasoz, F. Alvarez, K. Lisetti, C. L., & Finkelstein, N. (2004). Emotion Recognition from Physiological Signals Using Wireless Sensors for Presence Technologies, *Cognition, Technology and Work*, 6, 4-14.
- Pantic, M. & Bartlett, M. S. (2007). *Machine Analysis of Facial Expressions, Face Recognition*, In K. Delac and M. Grgic (Eds.), 377-416, I-Tech Education and Publishing.
- Pantic, M. & Patras, I. (2006). Dynamics of Facial Expression: Recognition of Facial Actions and Their Temporal Segments Form Face Profile Image Sequences, *IEEE Transactions on Systems, Man, and Cybernetics Part B*, 36 (2), 433-449.
- Pantic, M. & Rothkrantz, L. (2003). Toward an Affect-Sensitive Multimodal Human-Computer Interaction, *Proc. IEEE*, 91(9), 1370-1390.
- Pantic, M. & Rothkrantz, L. J. M. (2004a). Facial Action Recognition for Facial Expression Analysis from Static Face Images, *IEEE Transactions on Systems, Man, and Cybernetics Part B*, 34(3), 1449-1461.
- Pantic, M. & Rothkrantz, L. J. M. (2004b). Case-Based Reasoning for User-Profiled Recognition of Emotions from Face Images, *Proc. 12th ACM Int'l Conf. Multimedia (Multimedia '04)*, 391-394.
- Petrantonakis, P. C. & Hadjileontiadis, L. J. (2010). Emotion Recognition from EEG Using Higher Order Crossings. *IEEE Transactions on Information Technology in Biomedicine*, 14(2), 186-194.
- Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional Neuroanatomy of Emotion: A Meta-Analysis of Emotion Activation Studies in PET and fMRI, *NeuroImage* 16, 331-348.
- Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward Machine Emotional Intelligence: Analysis of Affective Physiological State, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10), 1175-1191.
- Ruch, W. (1995). Will the Real Relationship between Facial Expression and Affective Experience Please Stand Up—The Case of Exhilaration. *Cognition and Emotion*, 9, 33-58.
- Russell, J. A. (1983). Pancultural aspects of the human conceptual organization of emotions, *Journal of Personality and Social Psychology*, 45, 1281-1288.
- Russell, J. A. (2003). Core Affect and the Psychological Construction of Emotion. *Psychological Review*, 110, 145-172.
- Scherer, K. & Ellgring, H. (2007). Multimodal Expression of Emotion: Affect Programs or Componential Appraisal Patterns? *Emotion*, 7, 158-171.
- Sebe, N., Lew, M. S., Cohen, I., Sun, Y., Gevers, T., & Huang, T. S. (2004). Authentic Facial Expression Analysis, *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition (AFGR)*.
- Sharma, R., Pavlovic, V. I., & Huang, T. S. (1998). Toward Multimodal Human-Computer Interface. *Proc. IEEE*, 86(5), 853-869.
- Sitaram, R., Lee, S., Ruiz, S., Rana, M., Veit, R., & Birbaumer, N. (2010). Real-time support vector classification and feedback of multiple emotional brain states, *NeuroImage*, doi:10.1016/j.neuroimage.
- Sugimoto, Y., Yoshitomi, Y., & Tomit, S. (2000). A Method for Detecting Transitions of Emotional States using a Thermal Facial Image based on a Synthesis of Facial Expressions. *Robotics and Autonomous Systems*, 31, 147-160.
- Teeters, A., El Kaliouby, R., & Picard, R. W. (2006). Self-Cam: Feedback from What Would Be Your Social Partner, *Proc. ACM SIGGRAPH'06*, 138.
- Tian, Y. L., Kanade, T., & Cohn, J. F. (2005). Facial

- Expression Analysis, *Handbook of Face Recognition*. In S. Z. Li & A. K. Jain (Eds.), 247- 276, Springer.
- Tomkins, S. S. (1962). *Affect Imagery Consciousness: Volume I*. Springer.
- Tong, Y., Liao, W., & Ji, Q. (2007). Facial Action Unit Recognition by Exploiting Their Dynamics and Semantic Relationships, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(10), 1683-1699.
- Valstar, M.F. Gunes, H., & Pantic, M. (2007). How to Distinguish Posed from Spontaneous Smiles Using Geometric Features, *Proc. Ninth ACM Int'l Conf. Multimodal Interfaces (ICMI '07)*, 38-45.
- Valstar, M., Pantic, M., Ambadar, Z., & Cohn, J. F. (2006). Spontaneous versus Posed Facial Behavior: Automatic Analysis of Brow Actions, *Proc. Eight Int'l Conf. Multimodal Interface (ICMI '06)*, 162-170.
- Valstar, M., Pantic, M., & Patras, I. (2004). Motion History for Facial Action Detection from Face Video, *Proc. IEEE Int'l Conf. Systems, Man, and Cybernetics (SMC '04)*, 1, 635-640.
- Villon, O. & Lisetti, C. (2006). A User-Modeling Approach to Build User's Psycho-Physiological Maps of Emotions Using Bio-Sensors, *Proc. IEEE RO-MAN, 15th IEEE Int'l Symp. Robot and Human Interactive Comm., Session Emotional Cues in Human-Robot Interaction*, 269-276.
- Vyzas, E. & Picard, R. W. (1998). Affective Pattern Classification, *Proc. AAAI Fall Symp. Series: Emotional and Intelligent: The Tangled Knot of Cognition*, 176-182.
- Wagner, N. J., Kim, & Andre, E. (2005). From Physiological Signals to Emotions: Implementing and Comparing Selected Methods for Feature Extraction and Classification, *Proc. IEEE Int'l Conf. Multimedia and Expo*, 940-943.
- Wang, H. & Ahuja, N. (2003). Facial Expression Decomposition, *Proc. Ninth IEEE Int'l Conf. Computer Vision (ICCV '03)*, 958.
- Wang, J., Yin, L., Wei, X., & Sun, Y. (2006). 3D Facial Expression Recognition Based on Primitive Surface Feature Distribution, *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition(CVPR '06)*, 2, 1399-1406.
- Wen, Z. & Huang, T. S. (2003). Capturing Subtle Facial Motions in 3D Face Tracking, *Proc. Ninth IEEE Int'l Conf. Computer Vision (ICCV'03)*, 1343-1350.
- Whitehill, J. & Omlin, C. W. (2006). Haar Features for FACS AU Recognition, *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition (AFGR '06)*, 217-222.
- Wundt, W. (1905). *Fundamentals of Physiological Psychology(5th Ed.)*, Engelmann, Leipzig.
- Yeasin, M., Bullot, B., & Sharma, R. (2006). Recognition of Facial Expressions and Measurement of Levels of Interest from Video, *IEEE Transactions on Multimedia*, 8(3), 500-507.
- Yin, L., Wei, X., Sun, Y., Wang, J., & Rosato, M. J. (2006). A 3D Facial Expression Database for Facial Behavior Research, *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition (AFGR '06)*, 211-216.
- Yoshitomi, Y., Miyawaki, N, Tomita, S., & Kimura, S. (1997). Facial expression recognition using thermal image processing and neural network. *The 1997 6th IEEE International Workshop on Robot and Human Communication*, 1-6.
- Zeng, Z., Fu, Y., Roisman, G. I., Wen, Z., Hu, Y., & Huang, T. S. (2006). Spontaneous Emotional Facial Expression Detection, *J. Multimedia*, 1 (5), 1-8.
- Zeng, Z., Pantic, M., Roisman, G. I., & Huang, T. S. (2009). A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1), 39-58.
- Zhang, Y. & Ji, Q. (2005). Active and Dynamic Information Fusion for Facial Expression Understanding from Image Sequences, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(5), 699-714.
- Zhu, Z. & Ji, Q. Robust Real-Time Face Pose and Facial Expression Recovery. (2006). *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition(CVPR '06)*, 1, 681-688.

원고접수 : 11.02.11

수정접수 : 11.03.07

게재확정 : 11.03.08