

얼굴 표정의 제시 유형과 제시 영역에 따른 정서 인식 효과*

Effects of the facial expression presenting types and facial areas on the emotional recognition

이정현** · 조경자****† · 박수진*** · 한광희** · 김혜리***

Jung-Hun Lee** · Kyung-Ja Cho****† · Soo-Jin Park*** · Kwang-Hee Han** · Hei-Rhee Ghim***

연세대학교 심리학과**

Dept. of Psychology, Yonsei University

충북대학교 심리학과***

Dept. of Psychology, Chungbuk University

Abstract : The aim of the experimental studies described in this paper is to investigate the effects of the face/eye/mouth areas using dynamic facial expressions and static facial expressions on emotional recognition. Using seven-seconds-displays, experiment 1 for basic emotions and experiment 2 for complex emotions are executed. The results of two experiments supported that the effects of dynamic facial expressions are higher than static one on emotional recognition and indicated the higher emotional recognition effects of eye area on dynamic images than mouth area. These results suggest that dynamic properties should be considered in emotional study with facial expressions for not only basic emotions but also complex emotions. However, we should consider the properties of emotion because each emotion did not show the effects of dynamic image equally. Furthermore, this study let us know which facial area shows emotional states more correctly is according to the feature emotion.

Key words : Facial expressions, facial areas, emotion, dynamic facial expression

요약 : 본 연구에서는 동영상 자극과 정지 영상 자극을 사용하여 얼굴 표정의 영역(얼굴 전체/눈 영역/입 영역)에 따른 정서 상태 전달 효과를 알아보려고 하였다. 동영상 자극은 7초 동안 제시되었으며, 실험 1에서는 12개의 기본 정서에 대한 얼굴 표정 제시 유형과 제시 영역에 따른 정서 인식 효과를, 실험 2에서는 12개의 복합 정서에 대한 얼굴 표정 제시 유형과 제시 영역에 따른 정서 인식 효과를 살펴보았다. 실험 결과, 동영상 조건이 정지 영상 조건보다 더 높은 정서 인식 효과를 보였으며, 입 영역과 비교하였을 때 동영상에서의 눈 영역이 정지 영상보다 더 큰 효과를 보여 눈의 움직임이 정서 인식에 중요할 것임을 시사하였다. 이는 기본 정서 뿐 아니라 복합

* 이 논문은 2005년도 학술진흥재단에 의하여 연구되었음(KRF-2005-079-HM0004).

† 교신저자 : 조경자(충북대학교 심리학과)

E-mail : chokj@yonsei.ac.kr

TEL/FAX : 043-261-0598

정서에서도 어느 정도 관찰될 수 있는 결과였다. 그럼에도 불구하고 정서의 종류에 따라 동영상의 효과가 달라질 수 있기 때문에 개별 정서별 분석이 필요하며, 또한, 얼굴의 특정 영역에 따라서도 상대적으로 잘 나타나는 정서 특성이 다를 수 있음을 시사해 준다.

주제어 : 얼굴 표정, 얼굴 영역, 정서, 동영상

1. 서론

인터페이스 상에서의 에이전트나 휴머노이드(humanoid) 로봇의 기능 문제뿐만 아니라 사용자와의 상호작용이 중요시되면서 정서 상태를 측정하고 표현하는 기술을 위한 연구의 필요성이 증가하고 있다.

인간의 정서를 파악하는 방법으로는 음조 및 표현 등 음성 언어를 이용한 것, 생리적 측정치를 사용하는 것, 그리고 얼굴 표정을 이용하는 것 등이 있다. 언어적 표현의 경우 문화적 차이가 있을 수 있으며, 생리적 측정법은 정서 차원(쾌-불쾌, 각성-이완) 상의 급격한 변화를 알려줄 뿐 정확하고 미묘한 정서 상태에 대한 정보를 제공하는 데에는 어려움이 있다. 이에 반해 얼굴 표정은 언어를 제외하고 인간이 가지고 있는 가장 강력한 정서 전달 수단 중 하나이며[13], 심지어 의식의 통제 하에 있어서 위장하거나 왜곡시킬 가능성이 높은 언어에 비해 진짜 정서 상태를 전달할 수 있다[2]. 또한 서로 다른 문화권에 대한 연구[14]에 의하면, ‘행복한’, ‘슬픈’, ‘분노한’, ‘힘오히는’, ‘공포스러운’, ‘놀란’ 과 같은 여섯 가지 기본 표정이 범문화적으로 인식되는 것으로 나타나, 얼굴 표정은 인종이나 언어를 초월한 문화적 보편성도 가지고 있는 것으로 알려져 있다. 정서 인식의 발달적 차원에 대한 연구에 의하면 8세 아동도 이미 성인의 정서 인식 수준과 차이를 보이지 않아 8세 이전에 얼굴 표정에서 정서를 판단하는 능력이 발달되는 것으로 보고되고 있다[3]. 이렇듯 문화적 보편성과 선천성이 입증된 얼굴 표정은 정서를 연구하는 데 좋은 지표가 될 수 있을 것이다.

본 연구에서는 정서 연구에서의 얼굴 표정의 중요성에 기초하여, 얼굴 표정의 제시 유형(동영상, 정지

영상)과 제시 영역(얼굴 전체, 눈 영역, 입 영역)에 따른 정서 인식 효과의 정도를 알아보고자 하였다. 연구 1에서는 얼굴 표정의 제시 유형과 제시 영역에 따른 정서 인식 효과를 비교하였는데, 제시 영역 중 눈과 입 영역의 정서 인식 효과를 쾌-불쾌, 각성-이완의 두 가지 정서 차원으로 나누어 정서 차원 상에서 결과를 비교, 분석하였다. 연구 1을 기반으로, 연구 2에서는 기본 정서 외에 인지상태를 표현하는 것을 포함하는 복합 정서가 얼굴 표정을 통해 얼마나 잘 파악되는지 살펴보았다.

1.1 얼굴 표정에서의 정서

기본 정서(basic emotions)는 모든 문화에서 나타나며, 그와 연관된 얼굴 표정의 특징 또한 보편적이다[25]. 행동생물학자(ethologist)들은 기본 정서의 몇몇은 초기 유아기에 나타난다고 하였으며[11], 영장류 동물학자(primatologist)들은 인간과 인간이 아닌 영장류의 기본 정서에 대한 얼굴 표정이 상응한다고 주장하였다[7]. 각각의 기본 정서는 구분되는 얼굴 표정뿐만 아니라, 구분되는 생리적인 반응도 가지고 있다[19].

얼굴 표정은 기본 정서 이상의 감정도 전달하고, 다른 사람의 행동에 대한 공감과 이해를 보여줌으로써 의사소통을 더 풍부하게 한다[12]. 얼굴 표정은 ‘당황스러운’, ‘생각하는’, ‘흥미있는’과 같은 복합 정서(complex mental states)라 언급되는, 인지적 정서 상태(cognitive mental states)를 전달하기도 한다[3]. 여기서 복합 정서란 정서적 요인과 인지적 상태가 통합된 것을 지칭하는 것으로, 기본 정서와 달리 인지적 평가를 수반하지 않고서는 경험될 수 없는 것

이다. 이러한 인지적 정서 상태는 일상의 상호작용에서 기본 정서보다 더 빈번하게 나타난다[28].

Premack과 Woodruff[26]는 “복합 정서가 행동에 의해 추론될 수 있는 것으로 관찰 불가능하다”고 하였으나, Baron-cohen 등[3]의 연구에서는 이와 다른 결과가 보고되었다. 그들은 아홉 개의 복합 정서(정멸하는, 인지하는, 위협하는, 후회하는, 죄책감을 갖는, 걱정하는, 불신하는, 복수심 있는, 음모를 꾸미는)를 포함한 11개의 정서 어휘에 맞는 17세기의 화가 Diego Da Velazquez의 작품들과 현대 영국 화가인 David Hockney의 작품들을 선정해, 영국, 스페인 그리고 일본 사람들을 대상으로 그림 속의 정서를 얼마나 잘 파악하는지를 실험하였다. 그 결과, 복합 정서를 나타내는 얼굴 표정을 인식하는 데 있어 문화적, 시대적 보편성이 존재함을 확인할 수 있다.

타인의 행동을 이해하고 예측하는 데 있어 복합 정서가 더 중요하고 보편성을 가정할 수 있음에도 불구하고, 얼굴 표정에 대한 연구는 거의 대부분 기본 정서의 발현에 대해서만 연구되어 온 경향이 있다. 이는 기본 정서가 정지 영상이나 적은 수의 프레임만으로도 확보될 수 있는 반면에 복합 정서는 연속적인 얼굴과 머리의 영상에 의해 신뢰롭게 인지될 수 있기[16] 때문일 것이다.

1.2 얼굴 표정의 제시 유형

얼굴 표정에 대한 연구들은 주로 정지 영상으로 이루어진 자극을 사용하여 이뤄졌다. 그러나 실제로 얼굴 표정을 통한 정서 인식은 연속적인 표정 변화 속에서 이루어진다. 얼굴 표정은 짧은 시간 동안 지속되는 얼굴 근육 활동의 급격한 변화에 의해 야기되는 복잡한 신호로[27], 눈동자나 얼굴 근육의 미세한 움직임은 얼굴 표정을 통해 정서를 인식하는 데 중요한 단서를 제공해 줄 수 있다. 얼굴 표정과 같은 신체 반응의 순간적 패턴은 정서를 구별하는 데 있어 유용한 단서가 될 수 있으며[32], 그 역동적 변화는 다른 사람들의 정서 상태의 순간적인 변화를 알

아채는 데 가치 있는 자료가 될 수 있다[30]. 뿐만 아니라, 우리는 일상생활의 의사소통에서 이러한 얼굴 움직임의 역동적 패턴에 의해 얼굴 표정, 시선, 말하기에 대한 중요한 정보를 이끌어 낼 수 있다[22]. 근육은 스스로 역동적 움직임을 가지기 때문에, 우리의 지각이 그 역동적 속성을 무시하고 정지된 패턴에만 주의를 준다면 오히려 이상할 것이다[22]. 따라서 얼굴 표정의 정지된 패턴을 사용한 기존 연구들은 표정 연구에서 제한된 성공을 보였을 가능성이 있다[5]. 이렇듯 정지 영상은 얼굴 표정 연구에서 단서를 제한하고 부자연스러운 자극이 될 수 있다.

Sato와 Yoshikawa[31]는 동영상 자극을 사용하여 얼굴 표정 흉내내기에 관한 연구를 수행하였다. 그들은 거울 뉴런(mirror neuron)에 관한 신경과학적 증거를 토대로, 실제 얼굴 표정으로도 흉내내기가 나타나는지 알아보았다. 그들은 분노와 행복 두 가지 표정을 동영상과 정지 영상으로 제시하여, 동영상에서는 얼굴 표정 흉내내기가 자발적으로 일어남을 확인하였다. 이 결과는 실생활에서 얼굴 표정 처리의 심리학적 기제를 확인하는 데 있어 동영상 자극 사용의 중요성을 보여주는 것으로 해석될 수 있을 것이다.

Kamachi 등[22]은 얼굴 표정의 역동적 속성을 증명하기 위해, 동영상 얼굴 표정 자극을 제시하여 실험을 수행하였다. 첫 번째 실험에서는 네 가지 정서(행복한, 슬픈, 놀란, 분노한)를 세 가지 속도(빠른, 중간, 느린)로 제시하여, 참가자가 정서를 얼마나 정확하게 파악해내는지 알아보았다. 실험 결과, ‘슬픈’은 느린 속도 조건에서, ‘행복한’과 ‘놀란’은 빠른 속도 조건에서, ‘분노한’은 중간 속도 조건에서 가장 잘 인식되었다. 이 결과가 전체 제시 시간에 의한 것인지 아니면 자극의 연속성 때문인지를 알아보기 위해 두 번째 실험에서는 정지 영상과 동영상 자극을 비교하였는데, 그 결과 자극 제시 시간의 효과가 동영상에서만 나타나 실험 1의 결과가 제시 시간 때문이 아니라 연속성의 속도에 의한 것으로 결론지을 수 있었다.

이외에도 동영상의 정서 인식 효과를 보인 몇몇 연구들이 있었다. 정지 영상에 의해 얼굴 표정을 범주화하기에는 어려움을 느끼는 실인증(agnosia) 환자들은 움직이는 광점 제시 조건(dynamic point-light display)에서는 상대적으로 더 나은 판단을 하였다[21]. 10가지 정서에 대해 합성된 얼굴 이미지를 이용하여 동영상과 정지 영상의 정서 인식 정확성을 목표 어휘 선택 점수로 측정하여 비교한 연구에서는 동영상에서의 정확성이 유의미하게 높았으며, 특히 선택 어휘들 중 유사한 의미의 혼동을 야기하는 어휘가 존재할 경우 동영상의 효과가 더 잘 나타나기도 하였다[33]. Lemay, Kirouac, 그리고 Lacouture [23]의 연구에서는 정적인 얼굴 표정보다 동적인 얼굴 표정에서 공포와 놀람이 유의미하게 더 잘 구별된다고 하였다.

1.3 얼굴 표정의 제시 영역

사람들이 얼굴 표정을 통해 상대방의 정서를 파악함에 있어서, 얼굴의 어느 영역이 더 많은 영향을 미치는가에 대해 많은 연구들이 행해져 왔다. 그 중에서도 얼굴의 구성요소 중 가장 많은 움직임을 가지는 눈 영역과 입 영역의 정서 인식 효과를 비교하는 연구가 많이 이루어졌다.

정서 인식에서 눈과 입의 차이가 없다는 연구 결과가 있는가 하면[8, 18], 반대로 정서를 가장 잘 구분할 수 있는 얼굴 영역이 있다는 주장도 있었다[10, 29]. 이에 반해 Ekman, Friesen, 그리고 Tomkins[15]는 이런 주장들과 달리 정서에 따라 해당 정서가 잘 나타나는 얼굴 영역이 다르다고 하였다. 이들은 ‘혐오’, ‘공포’, ‘슬픔’, ‘행복’, ‘분노’, ‘놀람’의 여섯 가지 정서에 대해 얼굴 영역을 ‘뺨-입’, ‘눈-눈꺼풀’, ‘눈썹-이마’와 같이 세 영역으로 나누어, 참가자가 얼굴 부분 사진만으로 정서를 얼마나 잘 파악해내는지 살펴보았다. 실험 결과, ‘혐오’는 ‘뺨-입’ 영역에서, ‘공포’는 ‘눈-눈꺼풀’ 영역에서, ‘슬픔’은 ‘눈썹-이마’와 ‘눈-눈꺼풀’ 영역, ‘행복’은 ‘뺨-입’과 ‘눈-눈꺼풀’ 영역, ‘분

노’는 ‘뺨-입’과 ‘눈썹-이마’ 영역이라는 두 영역에서, 마지막으로 ‘놀람’은 세 영역 모두에서 대등하게 잘 나타났다. 이밖에도 정서에 따라 잘 나타나는 얼굴 영역이 다르다는 연구들이 있는데, Hanawalt[20] 연구에 의하면 행복한 표현에서는 입이 더 중요하고, 두려움이나 놀람에서는 눈이 더 중요하였다. Boucher와 Ekman[6] 연구에 의하면 공포와 슬픔은 눈과 눈꺼풀로부터, 기쁨과 혐오는 뺨과 입으로부터 더 잘 인식되며, Fraser, Craig, 그리고 Parker[17] 연구에 의하면 놀란 얼굴, 기쁜 얼굴, 그리고 슬픈 얼굴은 입이 가장 잘 나타내었다.

그런데 이 연구들은 모두 기본 정서를 통해 이루어진 연구 결과로, 인간의 인지상태를 포함하는 복합 정서를 살펴보지 않았다. Cunningham, Kleiner, Wallraven, 그리고 Bühlhoff[9]는 기본 정서와 복합 정서를 모두 포함한 아홉 가지 정서(동의하는, 동의하지 않는, 무력한, 행복한, 놀란, 당황한, 혐오스러운, 슬픈, 생각하는)에 대해 이 정서들이 정확하게 인식되기 위해 어떤 얼굴 영역들이 움직일 필요가 있는지 연구하였다. 그들은 눈만 움직이는 조건, 눈과 눈썹이 움직이는 조건, 입만 움직이는 조건, 그리고 세 영역이 모두 움직이는 조건과 머리의 움직임이 있는 조건, 이렇게 총 다섯 가지 조건에서 정서가 얼마나 잘 나타나는지 실험 참가자로 하여금 7점 척도로 평정하도록 하였다. 실험 결과, 대부분의 정서는 한 가지 얼굴 영역으로 그 의미가 전달되었다. ‘동의하는’, ‘동의하지 않는’, ‘무력한’은 머리 움직임만으로도 그 효과가 나타났으며, ‘행복한’, ‘놀란’은 입의 움직임, 그리고 ‘당황한’은 눈썹의 움직임, ‘생각하는’은 눈의 움직임, ‘슬픔’과 ‘혐오’는 네 가지 얼굴 영역 모두를 요구하였다.

Baron-Cohen, Wheelwright, 그리고 Jolliffe[4]는 여섯 개의 기본 정서(행복한, 슬픈, 화난, 두려운, 놀란, 혐오하는, 걱정하는)와 아홉 개의 복합 정서(교활한, 죄책감이 드는, 생각하는, 감동하는, 어리둥절한, 희롱하는, 지루한, 흥미로운, 거만한)를 연기한 배우의 사진을 얼굴 전체, 눈 영역, 입 영역으로 나

누어 실험 참가자가 얼마나 정확하게 정서를 인식하는지 알아보았다. 그 결과, 실험 참가자들은 기본 정서에서는 얼굴 전체 영역을 통해 정서를 가장 잘 파악했으며, 눈과 입 영역에서는 인식의 차이가 나타나지 않았다. 그러나 복합 정서를 파악하는 데 있어서는 눈이 얼굴 전체만큼의 효과를 가지고 있는 반면, 입은 그만큼의 효과를 나타내지 못해 복합 정서에서는 눈이 중요하다는 결론을 내릴 수 있었다.

2. 실험 1: 기본 정서에서 얼굴 표정의 제시 유형과 제시 영역에 따른 정서 인식 효과

실험 1에서는 12가지의 기본 정서를 나타내는 얼굴 표정을 사용하여 제시 유형과 제시 영역에 따라 정서 인식의 정확도가 어떻게 달라지는지 알아보았다. 독립변인은 제시 유형(동영상, 정지 영상)과 제시 영역(얼굴 전체, 눈 영역, 입 영역)이었고, 종속변인은 정서 인식의 정확도였다. 얼굴 표정 제시 유형 조건은 실험 참가자 간, 제시 영역 조건은 실험 참가자 내 설계로, 전체 실험은 혼합요인설계로 구성되었다.

2.1 실험 참가자

연세대학교에서 심리학 관련 과목을 수강하는 56명의 학생들이 수업의 일부로 실험에 참여하였으며, 이들은 무선적으로 각 실험 조건에 배정되었다. 참가자 중 29명은 동영상 조건에, 나머지 27명은 정지 영상 조건에 배정되었다.

2.2 실험 자극

예비 실험을 통해 남녀 배우 각 여섯 명으로 된 DB에서 표정 표현이 정확한 것으로 조사된 배우 네 명(남녀 각 두 명)의 동영상과 정지 영상을 추출하였다. 이들 동영상을 Adobe Premiere 6.0을 사용하여 얼굴 전체, 눈 영역, 입 영역의 세 가지 자극으로 만

들었다. 얼굴색에 따른 영향을 배제하기 위해 영상 자극은 모두 흑백으로 전환되었다(그림 1).

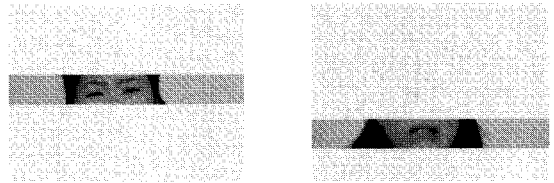


그림 1. 실험 자극의 예 '고통스러운' 표정. A는 얼굴 전체, B는 눈 영역, C는 입 영역

2.3 절차

실험 참가자들에게 주어진 과제는 동영상 혹은 정지 영상 자극 아래 제시된 12개의 기본 정서 어휘 중 자극 속 인물의 정서를 가장 잘 표현하는 어휘를 선택하는 것이었다(그림 2). 12개의 기본 정서 어휘는 김혜리 등[1]에서 추출한 것이었다.

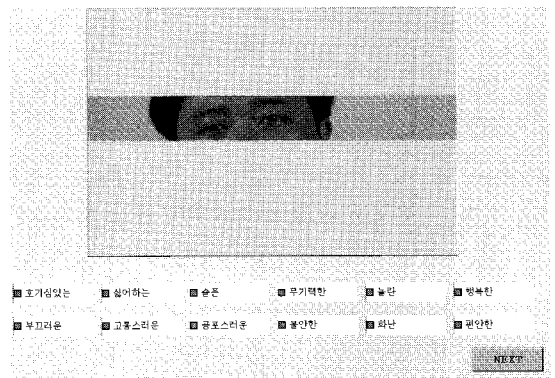


그림 2. 실험 화면의 예

어휘 추출 과정은 예비 조사를 통해 이뤄졌으며 김혜리 등의 연구에서 사용한 17개의 어휘를 유사어

에 대한 전문가 평정을 통해 묶어 정리한 것이었다. 실험에 사용된 어휘는 다음과 같았다: 고통스러운, 공포스러운, 놀란, 멍한, 부끄러운, 불안한, 슬픈, 싫어하는, 편안한, 행복한, 호기심있는, 화난.

실험은 얼굴 전체 영역에 의해 얼굴 부분 영역을 추측하는 것을 배제하기 위해 눈 영역과 입 영역을 우선적으로 제시하는 블럭과 얼굴 전체 영역만을 제시하는 블럭으로 구성되었다. 총 시행 수는 ‘배우(네명)×정서(12가지)×얼굴 영역(세 조건)’의 총 144회로, 동영상과 정지 영상은 모두 각 7초씩 제시되었다. 자극의 제시 시간은 예비 실험을 통해 표정이 80% 이상 정확하게 판단되는 지점을 선정한 것이다. 영상 자극이 나타났다 사라지고 나면 참가자는 마우스를 이용하여 답을 선택하였다. 참가자들의 응답 자료는 영상 속 배우가 표현한 정서와 일치하면 ‘1’로, 일치하지 않으면 ‘0’으로 기록되는 로그파일로 남았으며, 이 자료를 이용하여 정서 판단의 정확도를 분석하였다.

2.4 결과 및 논의

2.4.1 제시 유형과 제시 영역에 따른 정서 인식 정확도

얼굴 표정으로 정서를 인식하는 데 있어서, 자극의 제시 유형과 제시 영역에 따른 차이가 나타날 것이라는 가설이 실험 1을 통해 확인되었다. 동영상 조건이 정지 영상 조건보다 세 가지 얼굴 제시 영역 모두에서 정서 인식의 정확성이 더 높게 나타났으며, 제시 영역 내에서는 동영상과 정지 영상 모두 얼굴 전체→눈→입의 순서로 정서 인식의 정확성이 낮았다.

표 1. 기본 정서의 정서 인식 정확도

제시 영역	제시 유형	평균	표준편차
얼굴 전체	동영상	35.71	4.26
	정지 영상	31.30	3.10
눈 영역	동영상	25.90	3.42
	정지 영상	20.78	3.92
입 영역	동영상	20.10	3.96
	정지 영상	18.52	3.48

표 1에 얼굴 표정 제시 유형과 제시 영역에 따른 정서 인식 정확도 점수의 평균과 표준편차가 제시되어 있다. 제시 유형에 따른 각 제시 영역의 총점은 48점(정서 12가지×배우 4명)으로 동영상 조건에서의 평균은 얼굴 전체는 약 36점, 눈 영역 조건은 약 26점, 입 영역 조건은 약 20점이었으며, 정지 영상 조건에서의 평균은 얼굴 전체가 약 31점, 눈 영역 조건이 21점, 입 영역 조건이 19점이었다.

혼합 이요인 변량분석에 의한 통계 분석 결과 제시 영역에 따른 주효과($F(2, 108)=312.204, p<.001$)와, 제시 유형에 따른 주효과($F(1, 54)=25.986, p<.001$)가 각각 통계적으로 유의하였다. 제시 유형과 제시 영역 간의 상호작용 효과 또한 통계적으로 유의하였다; $F(2, 108)=5.113, p<.05$. 세 가지 제시 영역 중 눈 영역이 동영상과 정지 영상 간 차이가 가장 크게 나타났으며, 동영상에서 눈의 효과가 정지 영상보다 더 높은 것으로 나타났다(그림 3). 이는 제시 영역 중 눈이 눈동자와 같은 요소로 인해 입보다 상대적으로 많은 움직임을 가지고 있어, 동영상 조건에서 정서 전달에 있어 더 많은 효과를 갖는 것으로 해석될 수 있을 것이다.

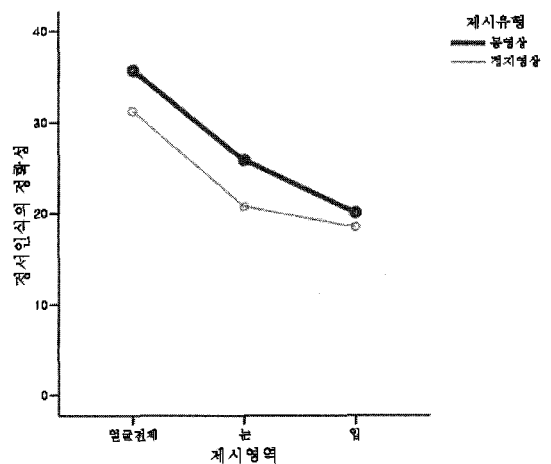


그림 3. 기본 정서의 제시 유형과 제시 영역에 따른 정서 인식 정확도

2.4.2 각 기본 정서별 제시 유형에 따른 정서 인식 정확도

전체 기본 정서의 제시 유형에 따른 정서 인식 효과의 차이가 통계적으로 유의하였다. 이에 기본 정서 각각의 제시 유형에 따른 정서 인식 정확도의 차이를 알아보았다.

일원변량분석을 실시한 결과, 총 일곱 가지 정서에서 동영상의 정서 인식 정확도가 정지 영상보다 통계적으로 유의하게 높은 것으로 나타났다; 화난($F(1, 54)=6.951, p<.05$), 공포스러운($F(1, 54)=15.523, p<.001$), 무기력한($F(1, 54)=4.859, p<.05$), 부끄러운($F(1, 54)=68.157, p<.001$), 불안한($F(1, 54)=26.555, p<.001$), 슬픈($F(1, 54)=9.201, p<.005$), 호기심 있는($F(1, 54)=20.331, p<.001$). 그러나 '싫어하는', '고통스러운', '행복한', '편안한' 같은 정서들은 제시 유형 간 차이가 통계적으로 유의하지 않았다. 후자의 정서들은 얼굴 표정의 움직임 없이 순간의 표정만으로도 충분히 그 정서가 전달되는 특성이 있기 때문인 것 같다.

한편, '놀란' 표정에서는 정지 영상에서의 정확도가 동영상에서의 정확도보다 통계적으로 유의하게 높은 것으로 나타났다; $F(1, 54)=45.065, p<.05$. '놀란'은 높은 각성 값을 가진 정서로 정지 영상에 의해서도 충분한 정서 인식 효과가 있으며, 특성상 순간적으로 나타나는 것으로 지속시간이 짧다. 때문에 7초의 동영상으로 제시한 경우 도리어 다른 정서 판단이 개입되어 정서 판단에 혼란이 야기되었을 가능성이 있다.

2.4.3 정서 차원별 눈 영역과 입 영역의 정서 인식 정확도

정서 인식 정확도가 더 높게 나타난 동영상 조건에서, 눈 영역과 입 영역의 정서 인식 정확도의 차이를 쾌-불쾌, 각성-이완의 정서 차원에서 분석해 보았다. 그림 4의 정서 차원 분포도는 김혜리 등[1]에서 보고된 것이다.

반복 측정 설계 방법에 의한 통계 분석 결과, 정서

영역의 주효과($F(3, 84)=5.906, p<.005$)와 제시 영역(눈 영역, 입 영역) 간의 주효과($F(1, 28)=70.763, p<.001$) 그리고 정서 차원 영역과 제시 영역 간의 상호작용 효과($F(3, 84)=5.808, p<.005$)가 각각 통계적으로 유의하였다. 눈 영역에 대한 정서 인식 정확도는 쾌-불쾌 차원을 따라 다르게 나타났는데, 쾌 차원보다 불쾌 차원에서 더 높은 정서 인식 정확도를 보였다. 입 영역에 대해서는 각성-쾌 차원에서 가장 높은 정확도가 보인 반면, 이완-쾌 차원에서는 가

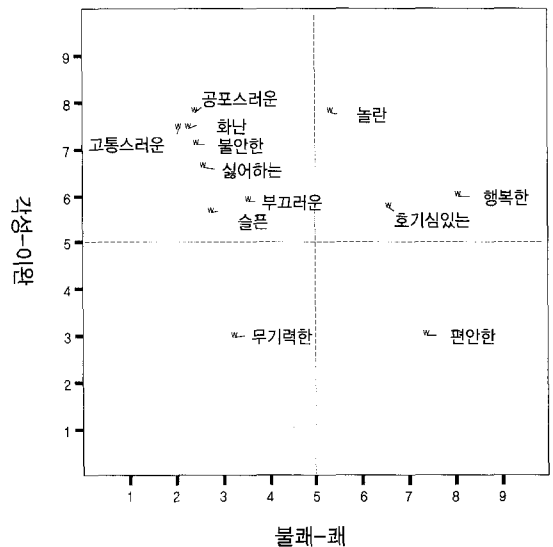


그림 4. 정서 차원 상의 기본 정서 분포

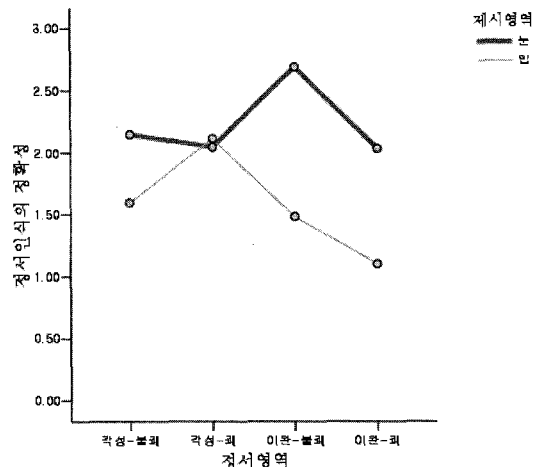


그림 5. 기본 정서에서 정서 차원 공간별 눈 영역과 입 영역에서의 정서 인식 정확도

장 낮은 정확도가 보였다(그림 5). 후자의 결과가 나온 것은 이완-쾌에 속하는 어휘가 ‘편안한’ 하나였기 때문일 가능성이 있다. 해당 어휘의 특성상 중립적 표정에 가깝기 때문에 입 모양의 변화가 크지 않아서 낮은 수행점수를 보였을 가능성이 있다.

3. 실험 2: 복합 정서에서 얼굴 표정의 제시 유형과 제시 영역에 따른 정서 인식 효과

실험 2에서는 기본 정서를 넘어서 인지상태를 포함하는 12가지의 복합 정서에서 얼굴 표정의 제시 유형과 제시 영역에 따라 사람들의 정서 인식의 정확도가 어떻게 달라지는지 알아보았다. 독립변인은 제시 유형(동영상, 정지 영상)과 제시 영역(얼굴 전체, 눈 영역, 입 영역)이었고, 종속변인은 정서 인식의 정확도였다. 얼굴 표정 제시 유형 조건은 실험 참가자간, 제시 영역 조건은 실험 참가자내 설계로, 전체 실험은 혼합요인설계로 구성되었다.

3.1 실험 참가자

연세대학교에서 심리학 관련 과목을 수강하는 63명의 학생들이 수업의 일부로 실험에 참여하였으며, 이들은 무선적으로 각 실험 조건에 배정되었다. 참가자 63명 중 32명은 동영상 조건에, 나머지 31명은 정지 영상 조건에 배정되었다.

3.2 실험 자극

실험 1과 마찬가지로 네 명의 배우의 복합 정서에 해당하는 동영상과 정지 영상 각각을 Adobe Premiere 6.0을 사용하여 얼굴 전체, 눈 영역, 입 영역의 세 가지 자극으로 만들었다. 얼굴색에 따른 영향을 배제시키기 위해 자극은 모두 흑백으로 전환되었다.

3.3 절차

실험 참가자들에게 주어진 과제는 동영상 혹은 정지 영상 자극 아래 제시된 12개의 복합 정서 어휘 중 자극 속 인물의 정서를 가장 잘 표현하는 어휘를 선택하는 것이었다. 실험 1과 같은 방식으로 예비 조사를 통해 김혜리 등[1]의 연구에서 사용된 15개의 복합 정서 어휘 중 12개의 어휘를 사용하였다; 간절함, 걱정하는, 냉담한, 뉘우치는, 당황스러운, 부러운, 안심한, 안타까운, 열광적인, 자신있는, 절망한, 흐뭇한. 복합 정서를 나타내는 영상과 어휘가 사용되었다는 점을 제외한 나머지 실험 절차는 실험 1과 동일하였다.

3.4 결과 및 논의

3.4.1 제시 유형과 제시 영역에 따른 정서 인식 정확도

실험 1과 마찬가지로 복합 정서를 사용한 실험 2에서도 자극의 제시 유형과 제시 영역에 따른 차이가 나타났다. 동영상 조건이 정지 영상 조건보다 세 가지 얼굴 제시 영역 모두에서 정서 인식의 정확도가 더 높게 나타났으며, 제시 영역 내에서는 동영상과 정지 영상 모두 얼굴 전체→눈→입의 순서로 정서 인식의 정확성이 낮았다. 표 2에 얼굴 표정 제시 유형과 제시 영역간 정서 인식 정확도의 평균과 표준편차가 제시되어 있다. 제시 유형에 따른 각 제시 영역의 총점은 48점(정서 12가지×배우 4명)으로 동영상 조건에서의 평균은 얼굴 전체는 약 24점, 눈 영역은 약 18점, 입 영역은 약 15점이었으며, 정지 영상 조건에서의 평균은 얼굴 전체가 약 23점, 눈 영역이 15점, 입 영역이 14점이었다.

혼합 이요인 변량분석을 실시한 결과, 제시 영역에 따른 주효과($F(2, 122)=156.483, p<.001$)와 제시 유형에 따른 주효과($F(1, 61)=5.294, p<.05$)가 각각 통계적으로 유의하였다. 제시 유형과 제시 영역간의 상호작용 효과는 통계적으로 유의하지 않았다; $F(2, 122)=.712$ n.s.

표 2. 복합 정서의 정서 인식 정확성

제시 영역	제시 유형	평균	표준편차
얼굴 전체	동영상	24.47	3.56
	정지 영상	23.16	3.80
눈 영역	동영상	17.50	3.63
	정지 영상	15.32	3.85
입 영역	동영상	15.16	2.89
	정지 영상	14.26	3.53

3.4.2 각 복합 정서별 제시 유형에 따른 정서 인식 정확도

실험 2의 전체 복합 정서의 제시 유형에 따른 정서 인식 효과의 차이가 통계적으로 유의한 것으로 나타나서, 복합 정서 각각의 제시 유형에 따른 정서 인식의 차이를 알아보았다.

일원변량분석 결과, 총 네 가지 정서에서 동영상의 정서 인식 정확도가 정지 영상보다 통계적으로 유의하게 높은 것으로 나타났다; 간절한($F(1, 61)=11.177, p<.005$), 뉘우치는($F(1, 61)=.6.679, p<.05$), 부러운($F(1, 61)=4.059, p<.05$), 안심한($F(1, 61)=9.948, p<.005$). ‘걱정하는’, ‘냉담한’, ‘당황스러운’, ‘안타까운’, ‘열광하는’, ‘자신있는’, ‘좌절하는’, ‘흐릿한’과 같은 여덟 가지 정서에서는 제시 유형 간 차이가 통계적으로 유의하지 않았다. 총 12개 복합 정서 중 3분의 2에 해당되는 여덟 개의 정서에서 제시 유형 간 차이가 나타나지 않았기 때문에, 본 연구로는 복합 정서인식에서 동영상이 더 효과가 있다고 단정짓기는 어렵다. 동영상 조건에서 전체 평균이 약 57점에 불과해 총점 144점의 반에도 미치지 못하는 결과가 나타났기 때문에 본 연구에서 사용한 복합 정서 과제가 너무 어려워서 각 정서별 제시 유형 간 차이를 보기 어려웠을 가능성이 있다. 또한, 제시 유형에 따른 통계적 차이를 보이지 않은 여덟 가지 복합 정서들의 경우, 특성상 눈이나 입과 같이 얼굴의 한 영역만으로 나타내기에는 어려움이 있을 가능성이 있고, 얼굴 내부 요소의 움직임보다는 고개의 움직임과 같은 다른 요소들이 보다 중요한 정서들일 가능성도 있다.

3.4.3 정서 차원별 눈 영역과 입 영역의 정서 인식 정확도

전체 제시 유형 분석에서 그 효과가 더 높게 나타난 동영상 조건에서, 눈 영역과 입 영역의 정서 인식 정확도를 왜-불왜, 각성-이완의 정서 차원에서 분석해 보았다. 그림 6의 정서 차원은 김혜리 등[1]에서 보고된 것이다.

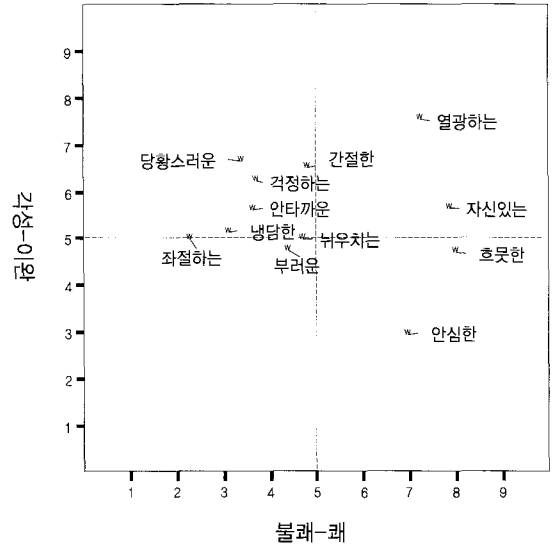


그림 6. 정서 차원 상의 복합 정서 분포

반복 측정 설계 방법에 의한 통계 분석 결과, 정서 영역의 주효과($F(3, 93)=24.959, p<.001$)와 제시 영역(눈 영역, 입 영역) 간의 주효과($F(1, 31)=4.317, p<.5$) 그리고 정서 차원 영역과 제시 영역 간의 상호작용 효과($F(3, 93)=5.214, p<.005$)가 각각 통계적으로 유의하였다. 전반적으로 왜 차원에서는 눈 영역보다 입 영역이 더 중요한 것으로 나타났으며, 이와 반대로 불쾌 차원에서는 눈 영역이 입 영역보다 더 중요한 것으로 나타났다(그림 7).

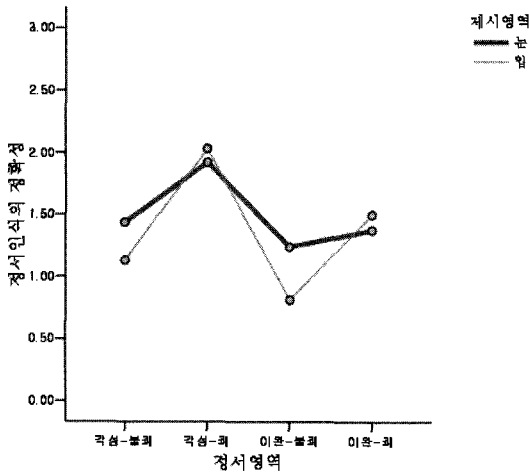


그림 7. 복합 정서에서 정서 차원 공간별 눈 영역과 입 영역에서의 정서 인식 정확도

4. 종합 논의

본 연구에서는 얼굴 표정에서 정서를 인식하는 데 있어 자극의 제시 유형과 제시 영역에 따른 효과를 검증하고, 눈 영역과 입 영역에 대한 정서 인식 정확도가 정서 차원에 따라 어떻게 달라지는지 알아보고자 하였다.

전반적으로 볼 때 동영상이 정지 영상에 비해 얼굴 전체, 눈 영역, 입 영역의 세 가지 제시 영역 모두에서 더 높은 정서 인식 정확도를 보이는 것으로 나타났다. 이는 Wehrle 등[33]의 동영상 효과에 대한 기존 연구를 지지하는 것으로, 동영상은 정서 판단을 하는 데 중요한 단서들을 더 제공하고 자극을 좀 더 자연스러운 표정으로 지각되게 할 가능성이 있다. 또한 동영상의 효과는 얼굴 전체뿐만 아니라, 눈 영역과 입 영역에서도 나타나 얼굴의 주요 영역의 부분 제시에서도 동영상이 정지 영상보다 더 유리한 것으로 나타났다. 이 결과는 얼굴의 전체적인 움직임뿐만 아니라 부분적인 움직임도 정서를 인식하는 데 중요한 단서가 됨을 시사한다.

쾌-불쾌, 각성-이완의 정서 차원 상에서의 눈 영역, 입 영역에 대한 정서 인식 정확도를 살펴본 결과, 쾌-불쾌 차원 상에서 서로 다른 양상을 보이는

것으로 나타났다. 눈 영역은 불쾌 차원에서 입 영역보다 상대적으로 높은 정서 인식 정확도를 보인 반면, 입 영역은 쾌 차원에서 눈 영역보다 나은 정확도를 보였다. 따라서 정서 전체에 대해 눈 영역이 입 영역보다 더 높은 정서 인식 정확도를 보인 것은 불쾌 차원 정서 자극이 쾌 차원 정서 자극보다 더 많았기 때문일 가능성이 있다. 제시 영역별 차이는 정서를 잘 나타내는 특정한 얼굴 영역이 있는 것이 아니라 정서에 따라 잘 표현되는 얼굴 영역이 다르다고 한 Ekman 등[15]의 주장을 지지한다.

12개의 기본 정서를 사용하여 이루어진 실험 1에서 제시 유형의 효과를 각 기본 정서별로 나누어보았을 때, ‘싫어하는’, ‘고통스러운’, ‘행복한’, ‘편안한’의 총 네 가지 정서에서는 제시 유형 간 차이가 나타나지 않았다. 오히려 ‘놀란’ 정서에서는 정지 영상이 더 높은 정확성을 보였는데, ‘놀란’의 경우 빠르고 짧게 나타나는 특성상 순간의 정지 영상만으로도 충분한 정서 정보를 제공하는 것으로 여겨진다. 이들 다섯 가지의 정서 중 정서 인식 정확도가 매우 낮았던 ‘편안한’을 제외한 나머지 정서들은 모두 각성 수준이 높은 정서라는 공통점을 갖는다. 이는 모든 종류의 정서를 인식하는 데 있어, 움직임이 정서 인식 효과를 높여주는 것은 아님을 시사하며, 각성 수준이 높은 일부 정서의 경우 정지 영상으로도 충분히 정서 전달을 할 수 있음을 의미한다.

실험 2는 복합 정서를 대상으로 실험 1과 같은 방식으로 이루어졌다. 실험 2를 통해서도 동영상의 정서 인식 효과가 지지되었으나, 이를 개별 정서별로 살펴보았을 때 ‘걱정하는’, ‘냉담한’, ‘당황스러운’, ‘안타까운’, ‘열광하는’, ‘자신있는’, ‘좌절하는’, ‘흐뭇한’과 같은 총 여덟 가지 정서에서는 제시 유형에 따라 통계적으로 유의한 차이가 나타나지 않았다. 얼굴 전체에 대한 동영상 조건만을 살펴봐도, 48점 만점에 실험 1의 기본 정서에서는 평균 정확도가 34인데 반해, 실험 2의 복합 정서의 경우엔 24에 불과했다. 이 결과는 복합 정서가 기본 정서와 달리 연속적인 영상에 의해 더 잘 인지될 수 있다는 El Kaliouby

와 Robinson[16]의 기존 연구에는 반하는 결과이다. Cunningham 등[4]의 연구에서는 ‘동의하는’이나 ‘동의하지 않는’과 같은 복합 정서를 인식하는 데 얼굴 표정뿐만 아니라, 머리의 움직임 또한 고려했다는 점을 감안하면, 본 연구에서 머리 움직임의 요소를 고려하지 않았기 때문에 이러한 결과가 나온 것일 수 있다. 또한, 4지선다 강제 선택에 의한 과제 방식이 너무 쉬울 것을 고려해 12개의 어휘를 모두 선택지로 준 것이 복합 정서 판단에서는 지나치게 어려운 과제였을 가능성이 있다.

복합 정서에서 제시 영역에 따른 실험 결과는 두 제시 유형 조건 모두에서 얼굴 전체→눈→입의 순서로 정서 인식의 정확도가 낮아져서 눈이 정서 판단에 중요함을 보여주었다. 그러나 얼굴 전체 조건에 비해서는 눈 영역을 통한 정서 인식 정확도가 떨어지는 것으로 나타나, 복합 정서를 나타내는 데 있어 눈이 얼굴 전체만큼의 효과를 가진다는 Baron-cohen 등[4]의 연구 결과를 지지하지는 않았다.

본 연구의 결과들을 종합해 볼 때, 얼굴 표정을 이용한 정서 연구는 동영상 자극을 사용하여 수행되었을 때 더 자연스럽게는 하지만, 이것이 모든 정서에 해당되는 것은 아니며, 특히 각성 수준이 높은 정서일수록 동영상에서 부가적인 이점이 나타나지 않을 수 있다고 결론지을 수 있다. 또한 눈 영역을 통한 정서 인식 정확도가 입보다 높고, 정서 차원 상에서 눈이 볼래 차원의 정서를 잘 드러내는 것으로 나온 결과는 사람들이 행복한 얼굴보다 화난 얼굴을 더 잘 인식하는 것이 생존과 관련된 이유가 있다는 Murphy와 Zajonc[24]의 기존 연구와도 잘 부합된다. 왜 차원에 비해 더 많이 분화된 볼래 차원의 정서를 잘 표현하는 눈 영역에 대한 정확 인식은 생존에 더 유리할 것이고, 이런 이유로 눈 영역을 통한 정서 인식이 입 영역보다 높아졌을 가능성이 있는 것이다.

후속 연구에서는 동영상의 시간에 대한 좀 더 체계적인 연구가 시도될 것이다. 제시 시간에 따른 정서 인식 변화를 기반으로 정서별 지속 패턴이나 지속 시간, 정서 유발까지의 시간 등 좀 더 다양한 요

소를 고려해 볼 것이다.

참고문헌

- [1] 김혜리 등 (2006). 마음이론에 대한 다학문적 연구: 진화적 관점에 기초한 발달심리, 정서심리, 정신병리 및 신경과학 연구. 학술진흥재단 기초학문 육성 지원인문 사회분야 심화 연구 중간보고서.
- [2] 한재현, 정찬섭 (1998). 얼굴 표정에 의한 내적상태 추정. 한국감성과학회지, 1(1), 41-58.
- [3] Baron-Cohen, S., Riviere, A., Cross, P., Fukushima, M., Bryant, C., Sotillo, et al (1996). Reading the mind in the face: A cross cultural and developmental study. Visual Cognition, 3(1), 39-59.
- [4] Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a “language of the eyes”? Evidence from normal adults and adults with autism or Asperger syndrome. Visual Cognition, 4(3), 311-331.
- [5] Berry, S. D. (1990). What can a moving face tell us? Journal of Personality and Social Psychology, 58(6), 1004-1014.
- [6] Boucher, J., & Ekman, P. (1975). Facial areas and emotional information. Journal of Communication, 25, 21-29.
- [7] Chevalier-Skolnikoff, S. (1973). Facial expression of emotion in nonhuman primates. In P. Ekman (Ed.), Darwin and facial expression, New York and London: Academic Press.
- [8] Coleman, J. (1949). Facial expressions of emotion, Psychological Monographs, 63, 1-296.
- [9] Cunningham, W. D., Kleiner, M., Wallraven, C., & Bühlhoff, H. H. (2004). The components of conversational facial expressions, ACM International Conference Proceeding, 73, 143-150.
- [10] Dunlap, K. (1927). The role of eye muscles and mouth muscles in the expression of the emotions. Genetic Psychology Monographs, 2, 197-233

- [11] Eibl-Eibesfeldt, I. (1973). The expressive behavior of the deaf and blind born. In M. von Cranach & I. Vine(Eds.), *Social communication and movement* (pp. 163-194). London: Academic Press.
- [12] Ekman, P. (1979). About brows: Emotional and conversational signals. In M. von Cranach, K. Foppa, W. Lepenies, & D. Ploog (Eds.), *Human ethology: Claims and limits of a new discipline* (pp. 169-202), Cambridge: Cambridge University Press.
- [13] Ekman, P. (1982). Methods for measuring facial action. In K. R. Scherer, & P. Ekman (Eds.), *Handbook of methods in nonverbal behavior research*. Cambridge: Cambridge University Press.
- [14] Ekman, P., & Friesen, W. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17, 124-129.
- [15] Ekman, P., Friesen, W., & Tomkins, S. S. (1971). Facial affect scoring technique: A first validity study. *Semiotica*, 3, 37-58
- [16] El Kaliouby, R., & Robinson, P. (2003). Real-time head gesture recognition in affective interfaces. *Proceedings of the 2003 IFIP TC13 International Conference on Human-Computer Interaction* (pp. 950-953), Zurich: IOS Press.
- [17] Fraser, I. H., Craig, G. L., & Parker, D. M. (1990). Reaction time measures of feature saliency in schematic faces. *Perception*, 19, 661-673.
- [18] Frois-Wittman, J. (1930). The judgment of facial expression. *Journal of Experimental Psychology*, 13, 113-151.
- [19] Griffiths, E. P. (2002). Basic emotions, complex emotions, Machiavellian emotions. In A. Hatzimoysis (Ed.), *Proceedings of the Royal Institute of Philosophy*, 1-43
- [20] Hanawalt, N. (1944). The role of the upper and lower parts of the face as the basis for judging facial expressions: II. In posed expressions and "candid camera" pictures. *Journal of General Psychology*, 31, 23-36.
- [21] Humphreys, G. W., Donnelly, N., & Riddoch, M. J. (1993). Expression is computed separately from facial identity, and it is computed separately for moving and static faces: neuropsychological evidence. *Neuropsychologia*, 31, 173-181.
- [22] Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., & Akamatsu, S. (2001). Dynamic properties influence the perception of facial expressions. *Perception*, 30, 875-887.
- [23] Lemay, G., Kirouac, G., & Lacouture, Y. (1995). Expressions faciales emotionnelles spontanees dynamiques et statiques: comparaison d'etudes de jugement categoriel et dimensionnel [Dynamic and static emotional facial expressions: A comparison of categorical and dimensional judgment studies]. *Canadian Journal of Behavioral Science*, 27, 125-139.
- [24] Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology*, 64, 723-739.
- [25] Ortony, A., & Turner, J. T. (1990). What's basic about basic emotions? *Psychological Review*, 97(3), 315-331
- [26] Premack, D., & Woodruff, F. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 4, 515-526.
- [27] Richardson, C., Bowers, D., Bauer, R., Heilman, K., & Leonard, C. (2000). Digitizing the moving face during dynamic displays of emotion. *Neuropsychologia*, 38, 1028-1039.
- [28] Rozin, P., & Cohen, B. A. (2003). High frequency of facial expressions corresponding to confusion, concentration, and worry in an analysis of naturally occurring facial expressions of americans,

- Emotion, 3(1), 68-75.
- [29] Ruckmick, C. A. (1921). A preliminary study of the emotions. *Psychology Monograph*, 30, 30-35.
- [30] Sato, W., & Yoshikawa, S. (2004). The dynamic aspects of emotional facial expressions. *Cognition & Emotion*, 18(5), 701-710.
- [31] Sato, W., & Yoshikawa, S. (2006). Spontaneous facial mimicry in response to dynamic facial expressions. *Cognition*, (in press).
- [32] Tomkins, S. S. (1982). Affect theory. In P. Ekman (Ed.), *Emotion in the human face* (pp. 353-395). New York: Cambridge University Press.
- [33] Wehrle, T., Kaiser, S., Schmidt, S., & Scherer, K. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology*, 78(1), 105-119.

원고접수 : 07.01.23

수정접수 : 07.03.20

게재확정 : 07.03.20