

감정 자세 인식을 위한 자세특징과 감정에측 모델

Posture features and emotion predictive models for affective postures recognition

김진옥*
Kim Jin Ok

요 약

감정 컴퓨팅의 대표적 연구 주제는 기계가 사람의 감정을 인식하여 이에 적절히 대응하는 것이다. 감정 인식 연구에서는 얼굴과 목소리 단서를 이용하여 감정을 포착하는데 집중했으며 최근에 와서 행동자세를 주요 수단으로 이용하고 있다. 본 연구의 목적은 감정 표현에서 중요한 역할을 담당하는 자세 특징을 포착하고 확인하여 감정을 판별하는 것이다. 이를 위해 먼저 자세포착시스템으로 다양한 감정 자세를 수집하여 감정별 특징을 공간적 특징으로 설명한다. 그리고 동작을 취하는 행위자가 의도하는 감정과 관찰자가 인지하는 감정 간에 통계적으로 의미 있는 상관관계가 있음을 표준통계기술을 통해 확인한다. 6가지 주요 감정을 판별하기 위해 판별 분석법을 이용하여 감정 자세 예측 모델을 구축하고 자세 특징을 측정한다. 제안 특징과 모델의 평가는 행위자-관찰자 감정 자세 집단의 상관관계를 이용하여 수행한다. 정량적 실험 결과는 제안된 자세 특징으로 감정을 잘 판별하며 감정 예측 모델이 잘 수행됨을 보여준다.

ABSTRACT

Main researching issue in affective computing is to give a machine the ability to recognize the emotion of a person and to react to it properly. Efforts in that direction have mainly focused on facial and oral cues to get emotions. Postures have been recently considered as well. This paper aims to discriminate emotions posture by identifying and measuring the saliency of posture features that play a role in affective expression. To do so, affective postures from human subjects are first collected using a motion capture system, then emotional features in posture are described with spatial ones. Through standard statistical techniques, we verified that there is a statistically significant correlation between the emotion intended by the acting subjects, and the emotion perceived by the observers. Discriminant Analysis are used to build affective posture predictive models and to measure the saliency of the proposed set of posture features in discriminating between 6 basic emotional states. The evaluation of proposed features and models are performed using a correlation between actor-observer's postures set. Quantitative experimental results show that proposed set of features discriminates well between emotions, and also that built predictive models perform well.

☞ keyword : Affective Computing(감정컴퓨팅), Posture recognition(자세인식), Emotion Recognition(감정인식), Discriminant Analysis(판별인식), Motion Capture System(동작포착시스템), Posture features(자세특징)

1. 서 론

사람들은 커뮤니케이션 과정에서 행동이나 신체 언어로 다른 사람의 느낌과 의도를 이해한다. 또한 행동이나 신체 언어를 통해 상황이나 사건에 반응하는 방법으로 감정을 드러내는데 감정이란 슬픔과 같은 기분을 느끼거나 회상과 같은 어떤 정신적

상태에 있을 때를 의미한다[1]. 정확하게 감정에 대한 정의는 존재하지 않지만 감정은 생각하는 과정과 유사하며 사람 지능의 일부라는 주장이 지배적이다[2,3].

유비쿼터스적으로 비디오 데이터가 넘치고 영상 감시 장치, HCI, 게임, 동작 기반의 제어, 스마트 환경에서의 이벤트 검출과 같이 다양한 실세계 응용 기술 분야에서 컴퓨터시스템과 사람간의 자연스러운 상호작용을 통해 지능화된 시스템을 구축하려는 요구가 증가하면서 사람의 감정을 자동적으로

* 종신회원 : 대구한의대학교 국제문화정보대학
모바일콘텐츠학부 교수 bit@dhu.ac.kr
[2011/08/04 투고 - 2011/08/09 심사 - 2011/11/03 심사완료]

분석하고 이해하려는 수요가 늘고 있다.

사람이 컴퓨터 시스템과 상호작용하려면 사람간의 의사소통처럼 지능형 시스템도 사람의 느낌을 이해하고 적절하게 반응할 것을 기대하므로 사람의 의도, 마음 상태, 감정을 다양한 방법으로 예측하여 이해하는 기능을 갖춰야 한다. 컴퓨터 시스템이 사람의 감정에 자연스럽게 반응하는 지능형 감정컴퓨팅에서 가장 주목하는 연구는 기계가 상호작용하는 다른 에이전트 또는 사용자의 감정 상태를 인식하는 능력의 개발이다. 시스템이 이해하는 의사소통 방법은 주로 비언어적 신호로서, 특히 얼굴과 몸짓의 의미 있는 변화를 사회적 지능의 인식 수단으로 삼는다[4].

이를 위해 지능형 시스템은 시각적 센서를 통해 상대의 감정을 이해하도록 설계되어 얼굴 표정과 자세를 특정 감정 범위에 연관시켜 해당 감정을 분류하는 방법으로 감정 인식을 구현한다. 하지만 얼굴 표정을 통해 감정을 인식하는 연구들은 인위적으로 지은 표정을 감정과 연관시키므로 감정의 즉시성과 신뢰성은 떨어진다. 또한 얼굴은 여러 가지 기능에 연관되어 있고 인식한 얼굴 표정은 나타낼 수 있는 표정의 작은 부분만을 표현하므로 감정 상태를 인식하는 대표적 수단인 얼굴 표정의 유효성을 보완할 필요가 있다.

인지신경과학의 연구[5]에서 제안한, 얼굴 인지와 상관관계가 있는 뇌의 일부는 신체 구성의 평가에도 연관되어 있다는 내용과 감정을 드러내는 신체 자세를 이용하여 감정 상태를 인식하는 실제적 연구[6-8]가 컴퓨팅시스템 분야에서 그다지 이루어지지 않고 있다는 점[9]에 주목하여 본 연구는 얼굴 표정을 통한 감정 인식 결과를 개선하고 지능형시스템에서 더 정확하게 사용자의 감정을 인식하기 위한 방편으로 신체 자세를 통해 감정 상태를 효과적으로 인식하는 방법을 제시하여 감정컴퓨팅 구현에 기여한다.

본 연구의 2장에서는 신체 감정자세 관련한 관련 연구들의 동향을 서술하고 3장에서는 실험 대상 집단 간의 감정 관련 상관성과 인식 일치성을 평가하는 방법을 제시한 후 실험 대상 집단의 자세가

감정을 드러낼 때의 자세 특징과 감정판별 모델을 설정한다. 4장에서는 제안 특징 집단의 판별력을 정량적으로 측정하고 평가한 내용을 기술한다. 마지막으로 5장에서는 결론을 맺는다.

2. 관련 연구

심리학과 인지과학에서는 감정 인식에 대한 연구를 다수 진행하면서 기쁨, 분노, 슬픔, 놀람, 혐오, 두려움과 같이 일반적으로 드러나며 인식 가능한 여섯 가지 기본 감정 표정을 정의하고 있다[10-13]. 그 결과 사람들이 드러내는 대표적 감정을 얼굴의 모든 동작을 설명하는 규칙 집단에 기반하여 인식할 수 있다는 연구 결과를 FACS(Facial Action Coding System)로 도출하였다[14].

얼굴에서 감정을 포착하는 대신 자세를 통해 감정 요인을 포착하고 이를 애니메이션 캐릭터 합성에 응용하는 감정 자세 모델링이 최근 다양하게 등장했지만 실질적으로 사람의 자세를 대상으로 직접적으로 감정 인식을 구현하지 못하는 이유는 자세 특징에 대한 체계적 연구가 부족하고 수행해야 하는 동작이 아주 다양하기 때문이다. 컴퓨팅 분야에서 감정관련 연구가 주로 표정을 가진 컴퓨터 캐릭터를 창조하는데 그 초점을 맞추는 동안 자세 인식 연구는 신호 언어, 사람 검출, 성별 인식을 주로 다루 왔다.

Neagle은 전문가가 발레를 가르치는데 사용가능한 시각화 시스템을 만들기 위해 감정 동작에 대한 특징분석을 시도하였다. 가상 발레 댄서의 시각적 외형자세와 사실화에 초점을 맞춰 사람 관찰자가 각 댄스 자세를 보고 감정을 판단할 때 꽤 높은 정확도를 보였다[15]. 다만 댄스 자세가 행동자세 연구의 단초를 제공했으나 댄스 동작은 그 자체가 과장된 자세이며 일상생활에서 자연스럽게 수행되는 동작은 아니라는 문제가 대두된다. Pollick의 연구에서는 간단한 팔 움직임은 감정적 의미로 해석하는 동작 특징에 관한 실험 결과를 제시하였다[16]. 팔 동작 속도를 판별 특징으로, 느린 움직임은 슬픔과 피곤함과 연관되어 있고 빠른 움직임은 흥분

상태로 인식하였다. 댄스를 통해 전달되는 감정의 자동 인식을 향한 노력은 Woo의 연구[17]에서도 계속되었다. 이 연구에서는 Laban의 댄스 관련 연구[18]에서 촉발된, 신체가 점유하는 공간의 양, 동작의 가속과 시간 등 3가지 요인으로 댄스 자세를 모델링하여 1초 동안의 댄스 동작을 자세 특징에 따라 신경망을 통해 자연스러움, 행복, 풍요함, 고독함, 생기 넘침, 날카로움, 엄숙함 등 7가지 감정 카테고리와 연관되어 제시하였다. 하지만 더 나은 인식 결과를 위해서는 1초보다 긴 동작 시퀀스를 분석하는 것이 필요하다.

교육학과 신경학에서는 자세를 관측하여 사람의 상태를 추론한 연구[19]를 제안하였으며 특히 Mota와 Picard는 e-러닝 시나리오에 감정 자세를 집목한 연구를 시도했다[20]. 이 연구에서는 의자의 등받이와 앉는 자리에 장착한 압력 센서를 통해 자세 정보를 수집하고 신경망이 센서 정보를 분석하여 학생 상체의 방향을 감지한 후 은닉마코프모델을 통해 자세 시퀀스를 각 감정 확률에 연관지었다.

Laban의 연구와 생물학적 동작 인식 연구에서 시초한, 정적 특징 기반의 감정적 신체 자세를 분류하는 점진적 카테고리화 모델에서는 전면, 후면, 수직 차원으로 사지에서 몸통까지 신체의 전체 시각에서 자세를 가장 잘 표현한 순간을 묘사하는 방법을 제시했다. 특히Beintema이 제시한 연구[21]에서는 사람이 동작을 인식할 때 동작 분석보다는 관절 위치에서 구한 연속적 자세 정보의 형태 분석으로 사람의 동작을 인식하는 방법을 제시했다. 즉 정적 자세의 관절 위치변화로 동작을 분석한 것이다. 움직임 정보는 주로 잡음이 심한 배경에서 사람의 신체를 검출할 때 발생하는 문제를 해결하기 위해 이용한다.

본 연구에서는 시스템이 자세의 정적 묘사만을 이용하여 분노, 행복, 슬픔, 기쁨 등의 감정을 두드러진 자세 특징을 통해 정확하게 분류하는 방법을 제시한다. 이를 위해 움직임 포착 데이터를 구성하여 한 집단은 지정된 감정을 표현한 자세를 수행한 사람들로 구성하고 다른 집단은 자세에서 해당 감정을 식별하는 사람들로 구성한다.

3. 연구 내용

3.1 감정 자세 특징

본 연구는 감정이 어떻게 신체 자세를 통해 표현되는지 알아보기 위해 의사소통 수단으로서의 신체 주요 자세를 정립한다. 선행 연구[22]를 통해 사용자 행동 자세를 분석한 결과 사람들이 주로 상체 부위와 손의 동작으로 감정을 드러낸다는 것을 바탕으로 (표 1) (a)와 같이 신체 자세를 분류한다. (b)는 (a)의 신체자세를 통해 6가지 대표 감정을 드러낸 것을 분류한 결과이다. (c)의 놀람 자세는 (b)의 놀람 신체자세 특징을 드러낸다.

선행 연구는 2D 신체자세를 기반으로 한 것이므로 본 연구에서는 실제적인 감정 자세를 3D로 포착하기 위해 움직임 포착시스템을 이용, 20대 남녀 13명을 대상으로 감정 자세를 포착한다. 자세를 취한 사람을 행위자라 하고 행위자가 취한 감정 자세를 평가하여 인식하는 사람은 관찰자로 한다. 실험 행위자는 기쁨, 슬픔, 두려움, 분노, 혐오, 놀람을 표현하는 제자리 자세를 수행한다. 행위자는 동작을 취할 때 자연스럽게 감정을 표현하도록 하며 감정 동작을 표현하는 동안 다른 사람과 접촉하지 않는다. 행위자는 (그림 1)의 신체 자세를 고려하여 관절과 주요 신체 부위에 32개의 마커가 달린 수트를 입는다. 각 감정 자세는 8개 카메라로 포착하고 3D 공간에서 32개 마커의 위치를 설명하는 연속적 프레임으로 나타내 총 109개의 자세를 수집한 후 자세가 가장 잘 표현된 순간의 프레임을 선택한다. 각 마커를 3개 직각 평면에 투영하여 신체의 부피와 방향을 나타내고 신체의 옆쪽, 앞쪽, 수직 범위 최대치를 측정한다. 자세의 전면 모습을 시뮬레이션 하기 위해 초기에는 자세를 회전시킨다. 이런 방법으로 포착한 자세 특징과 선행 연구의 감정 자세 정보를 이용하여 (표 2)와 같은 24개 감정 결정 자세 특징을 설정한다. 각 특징은 행위자의 신체 최대 범위를 고려하여 신체 구조에 따라 정규화한다. 예를 들어 오른팔의 측면 쪽 벌린 거리는 팔의 최대 측면 확대 범위와 X축 방향을 따라 오른손과

(표 1) 신체 주요 자세

신체 자세			
1	몸을 뻗은 상태	11	으쓱하기
2	몸을 오무린 상태	12	만지기/신체 부위 껌치기/몸, 어깨주위로 팔 두르기
3	움직이기-몸앞으로 굽히기	13	손을 머리 뒤로 하기
4	움직이기-뒤로 굽히기	14	손 껌치기/ 주먹 불끈 쥐기
5	움직이기-방향 바꾸기/왼쪽, 오른쪽으로 움직이기	15	손바닥 위로 하기
6	머리 굽히기	16	손바닥 아래로 하기
7	머리 옆으로 기울이기	17	손을 허리에 놓기
8	어깨 떨어뜨리기	18	손가락으로 가리키기
9	몸에서 팔을 올리거나 뻗치기	19	두 손을 같이 누르기
10	팔짱끼기		
20	a	신체부위 만지기(머리, 얼굴, 뺨, 입, 이마, 귀, 턱)	
	b	두손으로 얼굴 감싸기	
	c	두손으로 뺨을 감싸기	
	d	두손으로 입을 감싸기	
	e	한손으로 이마 만지기	
	f	한손으로 귀 만지기	
	g	한손으로 턱 만지기	

(a) 2D 신체 주요 자세

감정	관련 신체자세
기쁨	1,9
슬픔	2,3,6,8,12,20b
놀람	4,20c, 20d
두려움	2,4,12,14,20b, 20d
분노	1,14,16,17,18
혐오	4,5

(b) 감정 관련 신체 자세



(c) 놀람 자세의 예

(표 2) 감정 결정 자세 특징

표시	특징벡터	자세 특징	표시	특징벡터	자세 특징
B1	방향 _{xy}	뒤쪽 머리-앞쪽 머리 축	B2	방향 _{xy}	뒤쪽 머리-앞쪽 머리 축
B3	방향 _{yz}	오른손-오른어깨	B4	거리 _z	왼손-왼어깨
B5	거리 _z	오른손-오른어깨	B6	거리 _y	왼손-왼어깨
B7	거리 _y	오른손-왼어깨	B8	거리 _x	왼손-오른어깨
B9	거리 _x	오른손-오른팔꿈치	B10	거리 _x	왼손-왼팔꿈치
B11	거리 _x	오른팔꿈치-왼어깨	B12	거리 _x	오른팔꿈치-왼어깨
B13	거리 _z	오른손-오른팔꿈치	B14	거리 _z	왼손-왼팔꿈치
B15	거리 _y	오른손-오른팔꿈치	B16	거리 _y	왼손-왼팔꿈치
B17	거리 _y	오른팔꿈치-오른어깨	B18	거리 _y	왼팔꿈치-왼어깨
B19	거리 _z	오른팔꿈치-오른어깨	B20	거리 _z	왼팔꿈치-왼어깨
B21	방향 _{xy}	어깨 축	B22	방향 _{xz}	어깨 축
B23	방향 _{xy}	발꿈치 축	B24	3D-거리	오른발꿈치-왼발꿈치

기쁨	행위자		놀람	행위자	
	관찰자			관찰자	
	특이 자세			특이 자세	
두려움	행위자		슬픔	행위자	
	관찰자			관찰자	
	특이 자세			특이 자세	
혐오	행위자		분노	행위자	
	관찰자			관찰자	
	특이 자세			특이 자세	

(그림 1) 감정 판별을 위한 자세 집단

원어깨 간의 거리 비율로 계산한다.

실험 관찰자는 20대 남녀 30명으로 구성되어 행위자의 감정 자세를 보고 6개 감정 중 하나를 선택하도록 한다. 분노, 슬픔, 기쁨, 두려움, 혐오, 놀람의 6개 감정을 여러 다른 강도로 표현한 후 세 가지 요소를 갖춘 감정 자세 쌍을 구성한다. 각 정적 자세 P_i 는 (X_i, A_i, O_i) 으로 구성하여 24개 특징 벡터 $X_i = x_{i1}, \dots, x_{i24}$ 와 행위자가 의도하는 감정 A_i , 관찰자가 가장 빈번하게 분류하는 감정 O_i 를 형성한다.

(그림 1)은 행위자가 감정 자세를 취하고 그 감정을 관찰자가 인식하면서 자세특징의 차이를 측

정하기 위해 감정을 결정하는 특징의 주요 성분을 반영하여 생성한 아바타 자세로 감정별로 80%의 자세 변동율을 보인다. PCA 적용 결과 사람들이 다양한 관절 구성 상태로 감정을 드러내지만 주요 성분의 몇가지 조합으로 자세를 설명할 수 있기 때문에 24개 특징으로 감정 자세의 가능한 변화를 포착할 수 있다. 자세 집단의 규모가 클수록 더 정확한 결론을 도출할 수 있다.

신체 자세의 감정 표현력을 확인하기 위해 (표 3)과 같이 행위자와 관찰자간의 감정 분류 일치/불일치 정도를 측정한다. (표 3)에서 행은 행위자가 자세를 통해 의도하는 감정이고 열은 관찰자가 분류한 감정을 의미한다.

(표 3) 행위자와 관찰자간의 감정 일치성

행위자 \ 관찰자	감정						
	분노	두려움	기쁨	슬픔	혐오	놀람	총
분노	18	2	4	1	2	1	28
두려움	2	23	4	1	1	2	33
기쁨	2	3	18	1	1	2	27
슬픔	2	1	1	21	3	1	29
혐오	2	1	1	3	18	1	26
놀람	1	2	2	1	1	20	27
총	27	32	30	28	26	27	170

대각선 셀의 값은 행위자와 관찰자간의 감정 일치치를 의미하고 다른 셀의 값은 불일치를 나타낸다. 예를 들어 첫 번째 분노 셀의 18은 총 28개의 분노 자세 중 행위자와 관찰자 둘다 분노라고 분류한 것이다. 각 감정에 대해 행위자와 관찰자간 감정 일치성을 계산한 결과 전체적으로 행위자와 관찰자는 69%의 분류 일치성을 보인다.

행위자와 관찰자, 두 범주의 불연속 변수 간 상관관계를 측정하기 위해 관찰자의 분류 결과는 행위자가 의도한 감정과는 독립적이라는 귀무가설을 설정하고 χ^2 (카이스퀘어) 테스트를 적용한다. 그 결과 자유도 25일 때 χ^2 의 총합 125.59는 유의 수준 $\alpha = 0.001$ 에서 기각역 51.618보다 훨씬 높아 귀무가설이 참일 확률은 0.001보다 훨씬 작으므로 귀무가설은 기각되며 두 범주 간에 의미 있는 관계가 성립함을 알 수 있다.

이 관계를 더 상세히 설명하기 위해 (그림 1)의 두 가지 집단을 구성하여 Set_A 는 행위자 집단으로, Set_B 는 관찰자 집단으로 설정하고 각각의 집단을 6개의 감정 자세를 취한 아바타로 나타낸다. 각 집단은 감정 강도를 달리하여 자세를 취하고 A_i 와 O_i 별로 강도가 다른 자세를 묶어 구축한다. 각 감정 자세에는 PCA(Principal Component Analysis)를 적용하여 같은 감정을 드러낸 자세지만 완전히 다른 형태의 자세나 반복되는 자세를 구분하여 잉여성과 특이성을 제거한다.

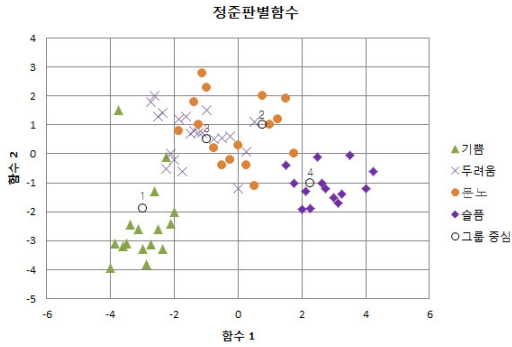
감정집단 내 잉여성을 확인하기 위해 각 자세

P_i 를 구성하는 모든 X_i 벡터에 PCA를 수행하여 반복되는 성분은 제거함으로써 감정 자세를 적은 수의 주요성분으로 표현한다. 감정 자세 PCA의 첫 번째 4개 주요성분이 약 80%의 변동성으로 자세를 결정한다. 슬픔의 경우에는 2개의 주요성분이 80%의 변동성으로 자세를 결정한다. 가장 높은 순서의 주요성분은 거의 자세에 포함되지 않아 특이자세로 분리되어 대표 자세로 간주되지 않는다.

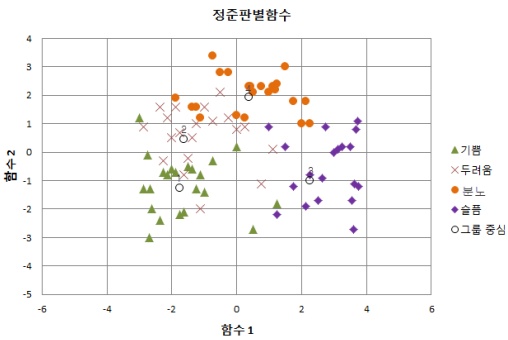
감정 집단에서 특이한 자세를 제거하기 전 PCA 결과와 행위자/관찰자의 일치성을 (표 4)와 같이 측정한다. 첫째 칸은 첫 번째 4개 주요성분에 대한 자세 일치/불일치를 나타낸다. 두 번째 줄의 셀은 특이 자세에 대한 일치/불일치를 나타낸다. PCA 결과와 행위자-관찰자간의 일치성은 독립적이라는 귀무가설을 설정하고 X_2 테스트를 적용한 결과 자유도 1에서 $X_2 = 19.43$ 는 유의 수준 $\alpha = 0.001$ 에서 기각역 10.827보다 더 높아 귀무가설은 기각되어 행위자 의도 감정과 관찰자 평가 결과는 서로 관계가 있음을 확인한다.

(표 4) 행위자/관찰자 일치성

PCA	행위자-관찰자		
	일치	불일치	전체
수용자세	92	44	136
특이자세	15	19	34
전체	107	63	170



(a) 행위자 집단



(b) 관찰자 집단

(그림 2) 2개의 판별함수에 의해 투영된 감정 자세

3.2 감정 자세 예측 모델

자세로부터 서로 다른 감정 자세를 판별할 수 있는 특징을 확인하기 위해 판별 분석(DA, Discriminant Analysis)법을 적용한다. 판별 분석은 두 개 이상의 범주나 그룹을 판별하기 위해 사용하는 통계적 기술로 요인 분석과 유사하게 판별함수그룹 즉 가장 치된 연속변수의 조합을 이용하여 그룹 판별을 수행한다.

$$D_k = d_{k1}x_1 + d_{k2}x_2 + \dots + d_{k24}x_{24} \quad (1)$$

x_i 는 자세 $k = 1, \dots, N-1$ 를 설명하는 24차원 벡터의 성분이며 N 은 감정 그룹 수이다. 판별상관계수 d_{kj} 는 그룹 내 차이에 대한 그룹 간 차이를 최대화하기 위해 이용한다. 판별함수를 이용, 감정 그

룹 간 분리를 최대화하는 축을 규정한 다차원 판별 공간에 자세를 연관 짓는 것이 가능하므로 각 판별 함수를 x_{ij} 특징 값에 적용하여 자세 P_i 를 공간에 매핑한다.

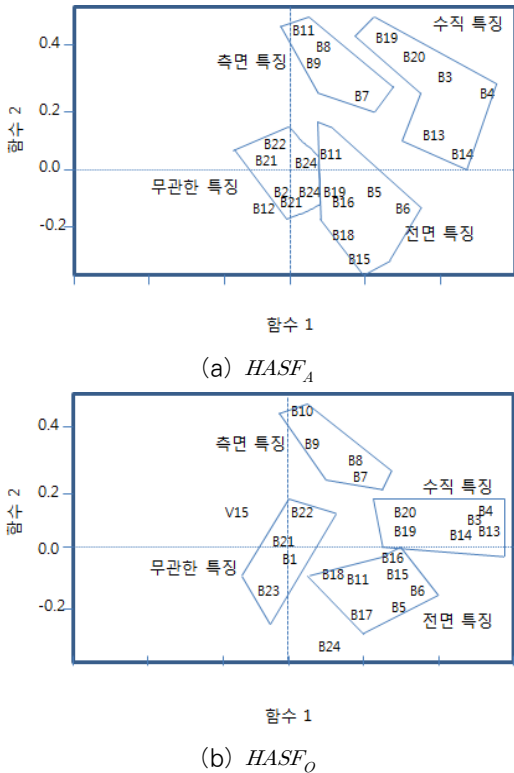
다양한 감정들이 어떻게 나타나는지를 이해하기 위해 감정 그룹을 여러 조합의 모델로 구성한다. 구성된 조합은 기쁨-분노-슬픔-두려움(HASF), 기쁨-분노-슬픔-혐오(HASD), 기쁨-분노-두려움-놀람(HAFS), 분노-슬픔-두려움-놀람(ASFS), 기쁨-슬픔-두려움-혐오(HSFD)이다.

판별분석 방법은 특이자세에 아주 민감하므로 PCA로 검출한 15개 특이자세를 Set_A , Set_O 에서 제외하고 적용한다. (그림 2) (a)는 Set_A 의 $HASF_A$ 모델을 판별함수로 정의한 2D공간이다. 세 개 판별함수에 대한 누적된 변수 값은 첫 번째 두개 함수 내에 변화량의 80%가 표시되었음을 보여준다. 2D공간에 4개 감정이 분리되어 잘 나타나 있고 두려움 그룹이 부분적으로 화남 그룹과 겹쳐있다. (그림 2) (b)는 Set_O 의 $HASF_O$ 모델에 대한 2D 판별공간이 구축된 결과이다. 4개 감정 자세 그룹이 잘 분리되어 있으며 부분적으로 두려움 자세가 기쁨 자세 그룹과 겹쳐 있다.

두 집단의 감정 간에 겹쳐지는 양상이 다르게 나타나는 것은 행위자가 정적 자세보다는 여러 움직임 동작으로 감정을 표현하는 동안 관찰자는 실험에서 요구하는 정적 자세 특징만으로 자세를 평가하기 때문이다. 하지만 두 공간에서 두 개의 판별 함수가 추정된 판별 규칙은 유사함을 알 수 있다. 두 공간에서 함수 1은 기쁨과 두려움에서 분노와 슬픔 자세를 판별하고 함수 2는 슬픔에서 분노를, 두려움에서 기쁨을 판별한다.

감정 간의 판별을 위해 감정의 두드러진 특징을 구분하기 위해서는 각 자세 특징과 판별함수간의 상관관계 계수의 정준구조행렬을 계산한다. (그림 3)은 $HASF_A$ 와 $HASF_O$ 에 대한 구조계수로 $HASF_A$ 가 $HASF_O$ 보다 더 많은 양의 계수를 보인다.

하지만 두 경우에서 모두, 신체의 수직 방향 확대를 나타낸 특징 $B3, B4, B13, B14, B19, B20$



(그림 3) $HASF_A$ 와 $HASF_O$ 의 구조 상관관계

이 두 함수와 양의 상관관계가 있다. 이것은 슬픔 집단과 기쁨 집단이 서로 잘 분리됨을 의미한다. 두 번째 두드러진 자세 특징은 $B7, B8, B9, B10$

으로 이 특징들은 신체의 측면방향으로 확대되는 것들이다. 감정 자세는 신체의 전면방향 확대 특징에는 가중치가 없어 보인다. 어떤 특징은 두 개 판별함수에 대해 0에 가까운 상관관계를 보인다. 이 특징들은 감정 간의 판별에는 상관없으므로 제외할 수 있다. (표 5)에 나타난 것처럼 Set_A, Set_O 는 제외 가능한 특징의 대부분이 비슷하다. $HASF$ 모델의 경우 네 가지 감정들은 많은 자세 특징을 공유해야 한다는 점을 강조하면서도 특징들의 대부분이 판별에서 제외 가능하다. 하지만 일반적으로 판별결과는 제안된 특징이 감정 자세를 판별하는 두드러진 특징이며 행위자가 의도하는 표현과 관찰자 평가간의 관련성이 통계적으로 의미가 있음을 보여준다.

4. 평가실험

제안된 특징을 이용해 어떻게 다른 감정적 자세를 잘 판별하는지에 대한 정량적 평가 측정이 필요하다. 이를 위해 판별 분석을 이용하여 두 가지 집단(행위자와 관찰자)을 대상으로 예측 감정이 맞는지 평가한다. 두 가지 $HASF$ 모델에 대한 혼합행렬 결과는 (표 6)과 같다.

(표 6)의 행은 행위자와 관찰자가 자세로 나타내거나 자세를 보고 결정한 감정이다(\bar{E}). 열은

(표 5) 모델별 제외 가능 자세 특징

모델	집단	제외가능 신체 특징
HASF	행위자	$B1, B17, B21, B22, B23, B24$
	관찰자	$B1, B21, B22, B23$
HASD	행위자	$B1, B21, B22, B23, B24$
	관찰자	$B1, B18, B21, B22, B23$
HAFS	행위자	$B1, B5, B17, B21, B22, B23$
	관찰자	$B1, B5, B6, B11, B15, B16, B17, B18, B19, B20, B21, B22$
ASFS	행위자	$B1, B8, B9, B21, B23, B23$
	관찰자	$B1, B7, B8, B21, B22, B23$
HSFD	행위자	$B1, B21, B22, B23, B24$
	관찰자	$B1, B8, B12, B21, B22, B23$

(표 6) HASF 모델 예측 감정

(a) 행위자가 의도한 감정 대 $HASF_A$ 의 예측 감정

행위자의 의도감정	분노	두려움	기쁨	슬픔	혐오	놀람	총
분노	20	1	1	0	1	0	23
두려움	1	26	3	0	1	0	31
기쁨	1	1	22	0	0	1	25
슬픔	0	0	0	23	1	1	25
혐오	1	0	1	1	22	0	25
놀람	0	0	1	1	0	24	26
총	23	28	28	25	25	26	155

(b) 관찰자가 평가한 감정 대 $HASF_O$ 의 예측 감정

관찰자의 평가감정	분노	두려움	기쁨	슬픔	혐오	놀람	총
분노	19	0	1	2	1	1	24
두려움	1	25	1	1	0	1	29
기쁨	1	0	22	0	1	1	25
슬픔	0	1	1	23	1	0	26
혐오	1	1	1	0	22	0	25
놀람	0	0	1	0	1	24	26
총	22	27	27	26	26	27	155

$HASF_A$ 와 $HASF_O$ 모델이 각각 예측한 감정이다 (\bar{E}_p). 표의 셀 값 $c[i,j]$ 에서 i 번째 행은 감정 자세에 대해 행위자와 관찰자가 분류한 자세 개수를, j 번째 열은 해당 $HSAF$ 모델이 분류한 자세 개수를 나타낸다. 자세집단의 판별력을 특징짓기 위해 전달된 정보를 계산한다. 정보 T 는 각 혼합 행렬 표에 식 (2)를 적용하여 결정한다.

$$T(\bar{E}_l, \bar{E}_p) = H(\bar{E}_p) + H(\bar{E}_l) - H(\bar{E}_l, \bar{E}_p) \quad (2)$$

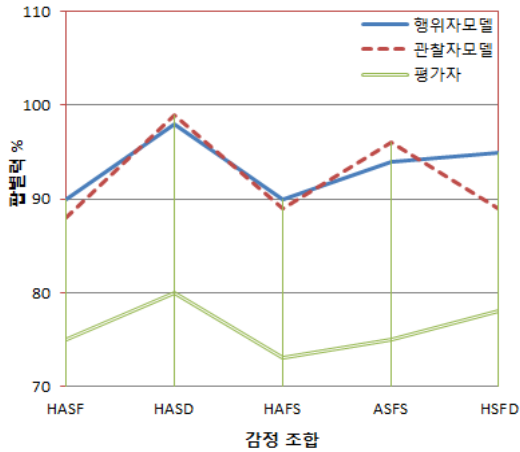
여기서 $H(\bar{E}_l) = -\sum_i p_i \log_2 p_i$,

$$H(\bar{E}_p) = -\sum_j p_j \log_2 p_j,$$

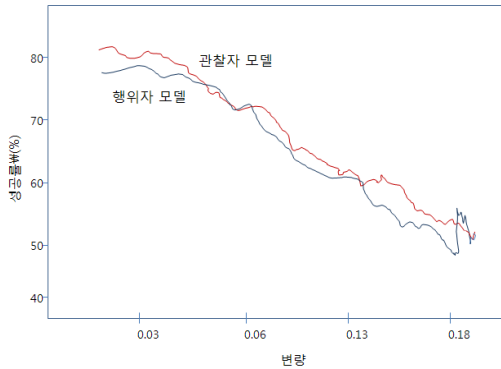
$$H(\bar{E}_l, \bar{E}_p) = -\sum_i \sum_j p_{ij} \log_2 p_{ij} \quad \text{이다.}$$

p_i 는 i 번째 감정 전달을 의도한 자세 확률이고 p_j 는 j 번째 감정 전달을 모델이 예측한 확률이다. p_{ij} 는 i 번째 감정전달을 의도한 자세가 대신 j 번째 감정을 전달하는 것으로 예측한 조합가능성이다. 판별력은 $T(\bar{E}_l, \bar{E}_p)$ (전송된 정보)와 $H(\bar{E}_l)$ (전송될 수 있는 최대 가능 정보량)간의 비율로 나타낸다. 즉 의도한 감정과 예측한 감정이 일치할 때 판별력은 최대가 된다. 각 혼합 행렬표에 대해 계산한 판별력 결과는 (그림 4)와 같다.

(그림 4) (a)는 행위자와 관찰자 집단의 5가지 다른 혼합 감정 모델의 판별력 값을 나타낸다. 행위자집단과 관찰자집단에 예측 모델의 판별력은 86% 이상의 높은 값을 보인다. 가장 낮은 판별력 곡선은 별도의 평가자에게 (그림 1)의 아바타 감정 자세를 측정된 것으로 평가자에게 얼마만큼의 정보가 전달되는 지를 나타낸다. 이 값은 행위자, 관찰자



(a) 판별력



(b) HASF 모델 잡음테스트
(그림 4) 판별력 분석과 잡음 테스트

두개 집단을 평가하기 위한 참조 값으로 이용한다. 결과적으로 행위자 모델과 관찰자 모델을 보고 평가한 제안 모델의 판별력 값이 아바타를 시각적으로 보고 감정을 판별한 사람의 판별력 값보다 높다. 따라서 제안 연구에서 사용된 자세 특징들이 자세의 감정 상태를 잘 포착함을 알 수 있다.

마지막으로 실제 테스트집단이 없는 상태에서 훈련 집단에 백색 잡음을 추가해 구축한 새로운 데이터 집단을 이용하여 판별 함수의 정확도를 테스트했다. (그림 4) (b)는 추가한 잡음의 양 대비 정확하게 분류된 자세의 백분율을 보여준다. 두 집단에 대한 판별함수 인식 결과는 원 훈련집단의 변동량이 10% 일 때까지는 인식이 잘 수행됨을 보여준다.

그 이상의 변동량에 대해서는 인식 결과가 계속 낮아져 20%에 이르면 인식할 수 없게 된다.

5. 결 론

사람의 감정 인식 연구는 사람과 컴퓨터간의 상호작용과 같은 자연스러운 사람 대 기계 커뮤니케이션 등에 적용할 수 있으며 특히 대상의 감정 상태에 따라 효과적인 교육을 제공하려는 교육학 분야에서도 응용 사례를 보이고 있다. 유비쿼터스 환경에서 대상의 감정을 자연스럽게 포착하고 이에 대응하여 적절한 서비스를 제공하는 감정 인식 컴퓨팅 시스템이 주로 집중해 온 사람의 얼굴 표정을 통한 감정 인식 외 자세와 행동에서 드러나는 특징을 감정과 연관시키면 더 효과적인 감정인식이 가능해진다. 제안 연구에서는 신뢰성 높은 자세 특징 검출과 해석을 통해 사람과 컴퓨터간의 상호작용을 위한 더 효과적이고 개선된 컴퓨팅시스템을 설계하는 데 기여한다.

본 연구에서는 신체 자세의 감정표현이 어떻게 수행되는지와 주요 감정별로 자세가 어떻게 나타나는지 확인하고 이 자세들을 결정하는 주요 신체 특징을 포착한다. 두드러진 주요 신체 특징을 설정하고 이를 이용해 감정별 자세들을 판별하기 위해 표준통계 기법을 이용한다. 실험 결과 자세는 감정 커뮤니케이션의 강력한 채널임을 확인하였다. 관찰자들의 50% 이상은 미리 정의된 콘텍스트가 없는 상황에서 행위자의 자세에서 의도하는 감정 메시지를 인식할 수 있었다. 판별 분석을 통해 본 연구에서 제시한 감정 결정 특징을 이용한 판별력 지수가 일반적 자세를 통해 계산한 판별력 지수보다 상회하면서 본 연구의 24개 자세 특징으로 주요 감정 그룹을 더 잘 판별함을 알 수 있으며 감정 예측 모델을 통해 높은 감정 인식률을 보였다.

하지만 더 많은 자세 특징들이 주요 감정 간의 정확한 판별을 위해 필요하다. 오분류된 자세를 분석해 보면 벌리거나 오무린 손 모양 등 손의 형태와 위치가 감정을 드러내는 신체 특징에서 중요한 역할을 하고 있음을 알 수 있으며 신체의 움직임

방향을 더 정확하게 설명하는 것이 감정 분류 시 추가되어야 한다. 여자와 남자의 신체 감정 표현력 역시 상당히 다른 양상을 보이고 있다.

향후 감정 인식 시스템은 실시간 멀티모달 시스템으로 진보하여 신뢰도 높은 감정 인식시스템을 구현하도록 한다. 이를 위해 실시간 환경에서 다양한 자세를 통해 정확한 감정 분류를 시도하고 그 결과를 평가하기 위해 얼굴 표정과 자세를 동시에 포착하고 더 복잡한 손의 형태를 보여주는 대상의 대형 데이터베이스를 구축해야 하고 시스템에서 인식하는 감정상태도 기본 감정 6가지 외 다양하게 확장하도록 한다.

참 고 문 헌

- [1] R. W. Picard, *Affective Computing*, MIT Press, 2000. M. Pantic, *Toward an Affect-Sensitive Multimodal Human-Computer Interaction*, Proc. of the IEEE, Vol. 91, No. 9, 2003.
- [2] H. M. Paterson, F. E. Pollick, A. J. Sanford, *The role of velocity in affect discrimination*. In *Proceedings of the Twenty-third Annual Conference of the Cognitive Science Society*, 2000.
- [3] Salovey, J.D. Mayer, *Emotional Intelligence, Imagination, Cognition and Personality*, Vol. 9, No. 3, pp. 185-211, 1990.
- [4] A. Vinciarelli, M. Pantic, H. Bourlard, A. Pentland, *Social Signal Processing : State-of-the Art and Future Perspectives of an Emerging Domain*, *Proceedings of the ACM International Conference on Multimedia*, pp. 1061-1070, 2008.
- [5] B. de Gelder, j. Snyder, G. Gerard, N. Hadjikhani, *Fear fosters fight: A mechanism for fear contagion when perceiving emotion expressed by a whole body*, *Proceedings of the National Academy of Science*, Vol. 101, No. 47, pp. 16701-16706, 2003.
- [6] M. Coulson, *Attributing emotion to static body postures : Recognition accuracy, Confusions and viewpoint dependence*, *Journal of Nonverbal Behavior*, Vol. 28, No. 2, pp. 117-139, 2004.
- [7] P. Fagerberg, A. Stahl, K. Hook, *Designing gestures for affective input: an analysis of shape, effort and valence*, *Proceedings of the National Academy of Science*, Vol. 102, No. 45, pp. 16518-16523, 2005.
- [8] E. Back, T. R. Jordan, S. M. Thomas, *The recognition of mental states from dynamic and static facial expressions*, *Visual Cognition*, Vol. 17, No. 8 pp. 1271-1286, 2009.
- [9] Z. Zeng, M. Pantic, G. Roisman, T. Huang, *A survey of affect recognition methods: Audio, visual, and spontaneous expressions*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 31, No. 1, pp. 39-48, 2009.
- [10] M. Minsky, *The Society of Mind*, Simaon and Shuster, New York, 1985.
- [11] M. Minsky, *The Eemotion Machine: Commonsense Thinking, Artificial Intelligence and the Future of the Human Mind*, Simon and Schuster, New York, 2006.
- [12] P. Ekman, *An argument for basic emotions*. *Cognition and Emotion*, pp. 169-2000, 1992.
- [13] J. A. Panksepp, *A critical role for effective neuron science in resolving what is basic about basic emotion*. *Psychological Review*, pp. 554-560, 1992.
- [14] S. Kimihiro, N. Kenichi, *Useful information for face perception is described with FACS*. *Journal of Nonverbal Behavior*, Vol. 27, pp. 43-55, 2003.
- [15] R. J. Neagle, K. C. Ng, R. A. Ruddle, *Studying the fidelity requirements for a virtual ballet dancer*, In *Proceedings of Vision, Video and Graphics Conference (VVG2003)*, pp.

- 181-188, 2003.
- [16] F. E. Pollick, H. Patterson, A. Bruderlin, A. J. Sanford, Perceiving affect from arm movement, *Cognition*, Vol. 82, pp.51-61, 2001.
- [17] W. Woo, J. Park, Y. Iwadate, Emotion analysis from dance performance using time-delay neural networks, *Proceedings of the JCIS-CVPRIP'00*, Vol. 2, pp. 374-377, 2000.
- [18] R. von Laban, *The Mastery of Movement*, Plays Inc., 1971.
- [19] M. M. Coulson, Attributing emotion to static body postures: recognition accuracy, confusions and viewpoint dependence, *Journal of Nonverbal Behavior*, Vol. 28, No. 2, pp. 117-139, 2004.
- [20] S. Mota, R. Picard, Automated posture analysis for detecting Learner's Interest Level, *Proceedings of the IEEE Workshop on Computer Vision and Pattern Recognition for Human Computer Interaction*, pp. 49-50, 2003.
- [21] J. A. Beintema, M. Lappe, Perception of biological motion without local image motion. *Proceedings of the National Academy of Sciences*, Vol. 99, pp. 5661-5663, 2002.
- [22] 김진옥, 사용자 행동 자세를 이용한 시각계 기반의 감정 인식 연구, *한국정보처리학회 논문지 B*, 제 18-B권, 5호, pp. 1-10, 2011.

● 저 자 소 개 ●

김진옥 (Kim Jin Ok)

1989년 성균관대학교 졸업(학사)

1998년 성균관대학교 대학원 정보통신공학과 졸업(석사)

2002년 성균관대학교 대학원 전기전자 및 컴퓨터공학과 졸업(박사)

2004년~현재 대구한의대학교 국제문화정보대학 모바일콘텐츠학부 교수

관심분야 : 멀티미디어공학, 패턴 인식, 영상처리, 유비쿼터스 컴퓨팅

E-mail : bit@du.ac.kr

