

감성ICT 교육을 위한 얼굴감성 인식 시스템에 관한 연구

A Study on the System of Facial Expression Recognition for Emotional Information and Communication Technology Teaching

송 은 지*

Eun Jee Song

요 약

최근 정보기술을 이용하여 인간의 감정을 인식하고 소통할 수 있는 ICT(Information and Communication Technology)의 연구가 증가하고 있다. 예를 들어 상대방의 마음을 읽기 위해서 그 사람과의 관계를 형성하고 활동을 해야만 하는 시대에서 사회의 디지털화로 그 경험이 디지털화 되어가며 마인드를 리딩 할 수 있는 디지털 기기들이 출현하고 있다. 즉, 인간만이 예측할 수 있었던 감정을 디지털 기기가 대신해 줄 수 있게 된 것이다. 얼굴에서의 감정인식은 현재 연구되어지는 여러 가지 감정인식 중에서 효율적이고 자연스러운 휴먼 인터페이스로 기대되고 있다.

본 논문에서는 감성 ICT에 대한 고찰을 하고 그 사례로서 얼굴감정 인식 시스템에 대한 메카니즘을 살펴보고자 한다.

Key Words : Emotion, ICT(Information and Communication Technology), Facial Expression Recognition

ABSTRACT

Recently, the research on ICT (Information and Communication Technology), which cognizes and communicates human's emotion through information technology, is increasing. For instance, there are researches on phones and services that perceive users' emotions through detecting people's voices, facial emotions, and biometric data. In short, emotions which were used to be predicted only by humans are, now, predicted by digital equipment instead. Among many ICT researches, research on emotion recognition from face is fully expected as the most effective and natural human interface.

This paper studies about sensitivity ICT and examines mechanism of facial expression recognition system as an example of sensitivity ICT.

* 남서울대학교 컴퓨터학과 (sej@nsu.ac.kr)

제1저자 (First Author) : 송은지

교신저자 : 송은지

접수일자 : 2012년 12월 6일

수정일자 : 2012년 12월 20일

확정일자 : 2012년 12월 24일

I. 서론

최근 감성적 제품이 소비를 자극하는 시대가 도래함에 따라 선진기업은 성숙기 ICT(Information and Communication Technology)산업 돌파구로 감정 전달 기술에 주목하여 시장 차별화에 주력하고 있으며 성능 및 가격 위주의 시장 전략에서 사용자 편의성과 만족도를 극대화시키는 소비자 감성 지향성 산업으로 재편을 꾀하고 있다[1].

최근 감성ICT의 발달로 상대방의 마음을 읽기 위해서 그 사람과의 관계를 형성하고 활동을 해야만 하는 시대에서 사회의 디지털화로 그 경험이 디지털화 되어가며 마인드를 리딩 할 수 있는 디지털 기기들이 출현하고 있다. 즉, 인간만이 예측할 수 있었던 감정을 디지털 기기가 대신해 줄 수 있게 된 것이다.

그 중에서 사람의 음성과 얼굴표정 및 생체데이터를 감지하여 사용자의 감정을 인식하는 휴대폰 및 관련서비스 기술에 관한 연구가 증가하고 있다. 얼굴에서의 감성인식은 현재 연구되어지는 여러 가지 감정인식 중에서 효율적이고 자연스러운 휴먼 인터페이스로 기대되고 있다. 심리학 분야에서도 얼굴 분석과 인식에 대한 연구가 수년간 이루어졌으며 심리학자인 Ekman과 Friesen의 연구에 따르면 사람의 여섯 가지 감정인 기쁨, 슬픔, 화남, 놀람, 혐오, 공포는 각각의 문화에 영향을 받지 않고 공통으로 인식되는 기본 감정으로 분류하였다[2]. 이러한 감정은 얼굴 표정을 비롯하여 뇌파, 맥박, 체온 등 생체 데이터로부터도 인식한다. 감정을 인식하는 휴대폰과 관련된 특허출원이 지속적인 증가세가 이어지고 있으며 이는 IT산업의 전체적인 방향이 PC중심에서 네트워크 중심을 거쳐 고객 중심으로 가고 있음에 따라, 고객의 행동은 물론 감정, 기호 등을 종합적으로 파악하여 맞춤형 서비스를 제공하는 것이 중요해지기 때문인 것으로 보인다.

따라서 국내 IT관련 학과에서도 감성ICT에 대한 교육이 필요하다.

본 논문에서는 감성ICT에 대한 고찰을 하고 이해를 돕기 위해 그 사례로서 사람의 감정을 읽을 수 있는 얼굴감정 인식시스템의 메커니즘을 살펴본다.

II. 관련 연구

영국 왕립 예술학교의 인터랙션 디자인(interaction design)학과 졸업생인 버나드 호펜가트너가 진행했던 'Belief Systems' 프로젝트에는 고객이 매장에 들어섰을 때 고객의 얼굴만 읽는 스캐너가 있다. 스캐너 밑에 있는 스크린을 통해서 사고 싶은 물건의 이미지들이 순차적으로 지나가고 스캐너는 고객의 미묘한 얼굴 표정을 인식해 그 사람이 가장 마음에 들었던 물건을 자동적으로 결정해주는 시스템이다.

이것은 디지털기기의 마인드 리딩이 가능함과 동시에 새로운 소비패턴을 재창조하는 사용자 경험을 보여주고 있는 예이다.

지금까지 ICT가 국민의 삶의 방식은 크게 개선시켰으나 ICT활용에 따른 정보이용 편의성과 삶의 질 보장 측면에서는 미흡했다는 인식이 확산되고 있으며 국내 ICT산업의 제2의 스마트폰 위기 도래에 따라 최근 감성 ICT의 혁신적 기술 확보와 생태계 구축 등 선제적 대응체계 마련의 필요성이 제기 되고 있다. 80년대는 생산력이 경쟁력 관건이었고 90년대는 신기술 확보가 시장 지배력을 좌우했으며 2000년대는 기술과 감성이 제품 구매력을 결정하는 시대로 변화하고 있는 것이다.

감성ICT산업은 일상생활에서 제공되는 제품과 서비스에 인간의 감성을 자동 인지하고 사용자 상황에 맞게 감성정보를 처리하여 사용자 감성 맞춤형 서비스를 제공하는 감성 ICT기술이 융합된 산업으로 정의 되고 있다. 이에 따라 휴먼 인터페이스에 관한 연구에서 인간의 감정을 어떻게 인식할 수 있느냐가 새로운 문제로 부각되고 있다[1].

1970년대 심리학자 Kendon[3]은 사람들이 함께 고개를 끄덕거리는 등의 바디랭귀지가 대화를 긍정적으로 느끼는 만족도와 상관관계가 있다는 것을 발표했다. 그 후 많은 사회심리학자들과 뇌 과학자들에 의해 이러한 Nonverbal communication 에 대한 연구가 진행되었다. Nonverbal communication 이란 구두로 전달되는 것이 아니고 그 외의 경로를 통해서 전달되는 의사소통의 요소인 시선, 고개를 끄덕이는 등의 머리움직임, 얼굴표정, 손동작, 어깨를 으쓱거리는 몸동작 등을 말한다.

심리학자들 사이에서 진행되어온 Nonverbal communication 에 대한 연구는 최근 공학자들 사이에서도 관심을 끌며 새로운 연구 분야를 열어가고 있다.

2004년 스텐포드 대학교의 Dept. of Communication 에서는 Jeremy[4]가 그의 연구에서 사람의 Nonverbal communication을 따라하는 가상 에이전트를 구현해서 그 에이전트의 행동이 사람들이 그 에이전트와 인터랙션에서 느끼는 만족도에 어떠한 영향을 미치는지 발표했다. MIT 미디어랩의 Human Dynamics 그룹 연구를 이끌고 있는 Sandly A. Pentland 교수는 사람들 사이의 대화를 분석하는 Sociometric Badge[5]를 개발하여 심리학 실험을 분석하고 있다. 또한 MIT 미디어랩의 Affective Computing 그룹의 Rosalind W. Picard 교수는 사람들이 표현하는 감정을 분석하는 기술을 개발 중이며 특히 얼굴표정을 통해 드러나는 감정을 읽는 소프트웨어를 개발하고 있다.

얼굴인식을 기반으로 하는 감정인식 연구에는 광학적 흐름분석(optic flow analysis), 홀리스틱 분석(holistic analysis),국부적인 표현(local representation) 등이 있다. 과학적 흐름 분석에는 Lien[6]이 얼굴감정인식을 수행하기 위해 광학적 흐름 추정을 통한 얼굴의 모션 분석을 하였다. 또한, 홀리스틱 분석은 얼굴 전체에 대한 분석으로 얼굴 영상들의 통계학으로부터 학습된 데이터 구동커널을 사용한다. 대표적인 방법으로 PCA(principal component analysis)[7], LFA(Local Feature Analysis)[8],LDA(Linear Discriminant Analysis)[9] 등이 연구되어지고 있으며 이와 같은 방법들은 영상집합의 2차 중속성에 근거하지만 고차 중속성에는 민감한 단점이 있다.

III. 얼굴감정 인식 시스템

복잡하고 다양한 환경에서의 강인한 얼굴 표정 인식 연구는 위에서 살펴본 바와 같이 과거로부터 지금까지 활발하게 진행되고 있다.

최근의 이미지 분석 기술과 패턴분류의 발전은 자동 얼굴 인식과 얼굴 표정 인식의 가능성을 열어주었다. 하지만 얼굴 표정 인식은 세 가지 면에서 어려움이 있다. 먼저 얼굴 이미지나

이미지 시퀀스(image sequence)에서 얼굴을 검출하여야 하고 검출한 얼굴에서 얼굴 표정 데이터를 추출해야 하며 추출한 데이터를 가지고 얼굴 표정 분류를 해야 한다. 이런 어려움들 때문에 얼굴 표정 인식에 관련된 연구는 입력 데이터에 대해서 사전 처리(pre-processing)와 특징점 추출(feature extraction), 분류(classification),사후 처리(post-processing) 등의 단계를 거쳐 얼굴 표정을 구분하게 된다.

얼굴표정을 인식하는 과정은 표1에 나타나 있다. 먼저 얼굴 영역을 추출하고 이로부터 다시 매칭 특징을 뽑아낸 후 데이터베이스와의 비교를 통해 얼굴을 인식한다. 미리 얼굴 영상과 신상 데이터베이스를 구축한 후 카메라를 통해 들어오는 영상을 분석하여 데이터베이스의 자료와 비교해 얼굴을 인식한다. 얼굴 인식 후 얼굴표정을 구분하는 특징점을 추출하여 표정을 인식한다[10].

<표 1> 얼굴감정인식 과정

처리 과정	제공 정보
전처리 정보	Gray Image → Binary Image → Difference Image → Morphology Image
얼굴 인식	AdaBoost 알고리즘
특징점 추출	점(코너)기반 접근으로 각 특징별 코너 추출.
표정 인식	코너의 방향성과 각도, 거리(움직임의 양), 면적 등의 정보로 표정인식을 가능하도록 함.

얼굴을 인식하는 방법 중에서도 특징점 기반 방식으로 추출하는 방법은 얼굴인식에서 사용하는 전통적인 방법으로 눈, 코, 입 등 얼굴의 특징을 나타낼 수 있는 곳에 점을 찍고 이 점들 사이의 관계를 이용해 얼굴을 구분하는 방법이다. 일반적으로 많이 사용하는 것이 눈 사이의 거리, 눈썹의 꺾어진 각도, 코의 길이, 입의 크기 등이다.

얼굴인식단계로서 전 처리(preprocessing)과정은 캡처된 이미지를 인식하기 위한 준비 단계이다. 얼굴 색상 분포는 얼굴인식이나 영상인식에서 비슷하게 GSCD(General Skin Color Distribution)로 쓰기도 한다. 캡처된 이미지를 보면 피부색에 대한 색상 발생 빈도가 특정 영

역을 중심으로 밀집되어 있다. 입력된 컬러 영상에서 얼굴의 피부 색상만을 추출하기 위해 얼굴 피부 색상과 비슷한 값을 가지는 픽셀의 색상 값은 높은 확률 값을 갖게 하고 그 이외의 색상의 영역은 낮은 확률 값을 갖게 하는 방법이다. 영상처리의 가장 기본적인 기법 중 하나인 이진화(Binary) 기법은 지문인식 기술, 문자인식 기술에서 특정 정보를 추출하기 위한 전 처리 과정으로 영상을 향상시키기 위해 사용된다. 임계값이란 경계 값이라고 할 수 있으며 특정 정보 추출을 위해 영상의 특징을 나눌 때 기준으로 잡는 값을 말한다.

배경 라벨링은 캡처된 이미지에서 얼굴과 배경을 분리하는 것이다. 얼굴 영역 내에서 얼굴의 특징점을 찾기 위해 얼굴 성분, 즉 눈과 눈썹, 입과 코 후보 영역을 찾아야 한다. 먼저 얼굴이 아닌 배경이 되는 부분을 라벨링 알고리즘을 이용해 찾는다. 다음으로 눈과 눈썹을 찾기 위해 앞에서 설명한 눈과 눈썹 색상을 이용해 가우시안 변환한 이미지와 원본 입력 영상을 그레이 스케일한 이미지를 이용해 연산을 한다. 이로써 우리는 눈과 눈썹 영역을 추출할 수 있다. 입술 영역은 앞서 추출한 눈과 눈썹 영역에 기반한 얼굴성분 요소들의 지정학적 위치 관계를 이용해 입술 후보 영역을 추출한다.

얼굴감성인식 시스템의 설계단계 내용은 다음과 같다.

1. 영상 획득 : 카메라는 초당 30frame을 지원하고 디지털 컬러 이미지는 3색 컬러 평면으로 구성한다. 알고리즘에서 물체구분과 그림자 영향을 제거하기 위해 RGB(Red Green Blue)를 사용한다.

2. 전처리과정 ①Gray :이진 영상보다는 더 밝은데, 각 화소의 밝기가 여러 단계로 보통 흑백 사진이 이에 해당되며 밝기의 단계는 검정색에서 시작해서 중간에 회색이 있고 마지막에 흰색으로 끝난다. ②이진화 : 임의의 경계 값(Threshold)을 기준으로 하여 어두우면(작으면) 전부 0으로, 밝으면(크면) 전부 255로 설정한다. 이진화영상은 0과 1, 즉 완전한 흑과 백으로만 밝기를 표현한다.③차영상 : 두 영상간의 차(빼기) 연산 결과를 취하는 영상. 3채널 영상 → 1채널 영상 → 차 연산 처리한다. ④모폴로지 : 기본적으로 팽창(Dilation)과 침식(Erosion). 통상적으로 영상 내에서 잡음 제거, 또는 구성요

소들의 결합 또는 분리 등의 폭넓은 분야에서 사용된다.

3. 배경제거: 배경영상과 입력되는 영상간의 픽셀 비교를 통해 전경영역 추출을 하여 프로세싱에서 기준 프레임과 현재 프레임을 비교하여 배경을 제거한다.

4. 관심 영역 지정: Adaboost 알고리즘을 이용하여 입력된 영상에서 얼굴 영역을 사각형 혹은 다각형으로 지정하고 개체의 군집 특성을 표본 샘플에 의한 크기로 환산하여 군집의 크기를 계산한다.

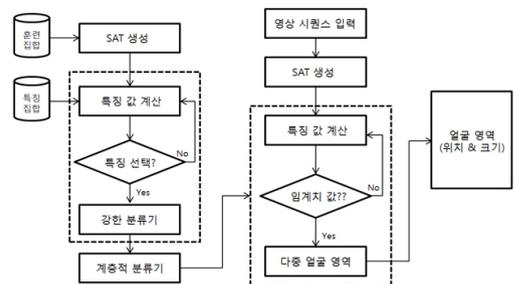
5.움직임 영역 추출: 전경영역은 입력영상과 배경영상의 픽셀 유사도가 높을 경우 전경 영역으로 추출되지 못하는 경우가 있어 구멍이나 영역 끊김 현상 발생하는데 성능을 높이기 위해 Adaboost 사용한다.

6. 얼굴 추적: 영상 내 움직이는 물체의 경우 각 프레임 당 움직이는 거리 및 크기가 크게 변하지 않고 빠른 처리 속도 가능하며 얼굴이 겹쳐지는 경우의 구분을 위해 처음 두 객체 이상의 크기 정보를 가지고 있다가 급격하게 커지는 경우와 얼굴의 개수가 줄어드는 경우의 정보를 합하여 두 개 이상의 객체가 있다고 판단한다.

7. 얼굴 가정: 전경영역을 추출하지 못할 때의 문제를 해결하기 위한 알고리즘으로 AdaBoost 알고리즘을 이용했을 때 순간적으로 벗어난 경우에도 기존 프레임에서 얼굴이라고 인식된 경우 계속해서 인식하도록 가정한다.

8.이벤트 검출 :얼굴의 특징점을 추출하여 특징점이 관심영역의 지정된 위치를 통과하는 경우 검출되는 이벤트가 관심 영역 안에 1초 이상 표정이 기준과 부합되는 경우만 인식한다.

일반적으로 얼굴표정을 인식하는 시스템의 구성 및 설계순서는 다음의 그림 1,2와 같다 [10].



[그림 1] 얼굴인식 시스템 구성도



[그림 2] 영상처리 설계 순서

IV. 결론

정보시대에 살고 있는 현대인들은 정보와 소통의 기술 즉, ICT(Information and Communication technology)에 인식하고 있지 않은 사이에 많은 영향을 받고 있다. 기존 ICT 제품군에 인성적인 요소를 접목하는 감성ICT는 모바일 기기, PC, 디스플레이 등의 ICT 제품에 감성적 소통을 제공함으로써 사용자에게 새로운 경험과 가치를 부여하고 제품의 글로벌 경쟁력을 강화시키는 ICT+감성인지 융합산업으로 나타나고 있다.

감성적 제품의 소비를 자극하는 시대가 도래함에 따라 선진기업은 성숙기 ICT산업 돌파구로 감성전달 기술에 주목하여 시장 차별화에 주력하고 있으며 성능 및 가격 위주의 시장 전략에서 사용자 편의성과 만족도를 극대화시키는 소비자 감성 지향형 산업으로 재편을 꾀하고 있다.

이에 따라 감성 ICT에 관한 연구가 증가하고 있다. IT관련 학과에서도 관심을 갖고 학생들에게 교육해야 할 것이다.

본 연구에서는 감성 ICT에 대한 이해를 돕기 위한 고찰을 하고 그 사례로서 얼굴감정인식 시스템을 살펴보았다.

참고문헌

[1] 이해룡외, “감성UX 기술동향”, 전자통신동향분석 제26권 제5호, 2011.10
 [2] P. Ekman, W. Frisen, "Facial Action Coding

System : A Technique for the Measurement of Facial Movement", Consulting Psychologists Press, 1978.

[3] Kendon, A., "Movement coordination in social interactions", Acta. Psychological, Vol.32, 101-125, 1970.
 [4] Jeremy, N. Beilenson & Nick, Yee, Psychological Science, Vol.16. No.10, 814-819, 2004.
 [5] Kim, T. Chang, A. & Pantland, A. "Meeting Mediator: Enhancing Group Collaboration with Sociometric Feedback ", Proceedings of Conference on Computer Supported Collaborative Work, 457-466, 2008.
 [6] J. Lien, T. Kanade, C. Li, "Detection, tracking, and classification of action units in facial expression", J. of Robotics and Autonomous Systems, Vol. 31, No. 3, pp. 131-146, 2000.
 [7] M. Turk, A. Pentland, "Eigenfaces for recognition", J. of Cognitive Neuroscience, Vol. 3, No. 1, pp. 71-86, 1991.
 [8] P. Penev, J. Atick, "Local feature analysis: a general statistical theory for object representation", Network : Computation in Neural Systems, Vol. 7, pp. 477-500, 1996.
 [9] P. Belhumeur, J. Hespanha, D. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 711-720, 1997.
 [10] 장민식, "고객만족도 피드백을 위한 효율적인 얼굴감정 인식 시스템에 대한 연구", 정보보안 논문지 제12호 제2호, 2012.6

송은지 (Eun-Jee Song)

정희원



1984년 : 숙명여자대학교 수학과 (이학사)
 1988년 : 일본 나고야(名古屋) 국립대학 정보공학과(공학석사)

1991년 : 일본 나고야(名古屋) 국립대학 정보공학과 (공학박사)

1991년~1992년 : 일본 나고야(名古屋)국립대학 정보공학과 객원 연구원

2007년 : 오클랜드대학교 컴퓨터학과 방문교수

1996년~현 재 : 남서울대학교 컴퓨터학과 교수
 멀티미디어 기술사

관심분야 : 실천공학교수법, IT융합, 수치해석, 정보보호 등