

감성 인터페이스

- 얼굴표정의 인식 및 합성 모형 -

정찬섭, 오경자, 이일병*, 변혜란*

연세대학교 심리학과, 연세대학교 컴퓨터과학과*

얼굴표정은 내적상태를 명백하게 반영하며 언어를 제외한 내적상태의 전달수단 중에서 가장 강력하고 복잡한 신호로 간주된다(Ekman, 1982; Rinn, 1991). 얼굴표정을 인식한다는 것은 단순히 얼굴에 나타나는 물리적 변화를 탐지하는 것이 아니라 그것에 내재되어 있는 내적 정서 및 행위 준비 상태(state of action readiness)를 파악하는 것이기 때문에 의미인식의 한 유형이다(Frijda, 1986). 언어가 의미 전달을 정확하게 할 수 있는 수단이라는 점에서 얼굴표정보다 내적상태를 정확하게 전달한다고 볼 수도 있지만 반드시 그렇지 않을 수도 있다. 언어는 의식의 통제하에 있어 의도에 따라 위장시키거나 왜곡시킬 가능성이 있기 때문이다. 나아가서 얼굴표정에 의해 전달되는 내적 상태는 문화적 보편성을 지니고 있어서 언어를 초월하여 인종 간 및 인종 내에서 일관성을 보인다(Ekman 등, 1987). 그럼에도 얼굴표정 인식 연구는 상징체계를 사용하여 복잡한 영상요소들을 추상화할 수 없다는 점 때문에 본격화되기 어려운 연구 영역 중의 하나였다. 그러나 최근 컴퓨터를 이용하여 대량의 정보를 저장 및 처리할 수 있게 되어 영상처리, 상징추출, 영상해석 등 표정연구의 여러 가지 기술적 난제들에 돌파구가 마련됨으로써 지난 몇 년 사이에 표정인식과 합성에 대한 연구에 큰 발전이 있었다.

최근에 들어 표정 연구가 활기를 띠고 있는 것은 그것이 학문적으로 중요한 것도 있겠지만 그보다도 감성공학과 영상산업 등의 분야에서 표정을 통한 감성 인식 및 표현의 실용적 가치에 대한 인식이 고조되고 있기 때문이다. 요즘 컴퓨터 및 고

속 통신 기술이 급속하게 발전함에 따라 멀티미디어 환경, 디지털 TV, 가상 생활 공간 등이 그에 편승하여 함께 발전하고 있다. 이러한 변화는 한편으로는 사람들에게 생활의 풍요로움을 더해 주지만 다른 한편으로는 새로운 생활 환경에 적응해야 하는 인지적 및 정서적 부담을 가중시키고 있다. 표정연구는 이러한 변화가 인간 중심적인 방향으로 전개되어 일의 효율성과 삶의 질을 향상시키도록 하는데 응용될 수 있다. 가상 인물을 이용하여 자연스럽게 감성적인 방식으로 기계와 정보를 교환할 수 있는 인터페이스 개발, 모델 기반 얼굴 변화 부호화 방식을 이용한 실시간 원격 화상 회의 실현, 가상 배우나 가수의 창조를 통한 오락 세계의 현실적 및 초현실적 연장 등이 그러한 예이다. 표정연구는 이 이외에도 사람, 제품, 환경에 대한 감성 반응 측정 시스템, 표정인식을 이용한 보안 시스템, 표정 연기 교육 프로그램, 대인 감수성 훈련 프로그램, 대인 정서 진단 도구 등의 개발에도 큰 응용 가치를 지닌다.

정보화 사회가 본격화되면서 표정의 응용에 대한 요구는 더욱 늘어 날 것으로 예상되며 표정 인식 및 표현 기술이 소프트웨어 산업의 핵심 요소로 떠오를 전망이다. 이러한 추세에 부응하여 과학기술처가 주관하는 G-7 감성공학 프로젝트는 “표정/제스처 측정기술 및 DB 개발”이라는 연구 과제를 추진하고 있는데 현재 2년차 연구 수행 단계에 있는 이 연구의 목적, 연구내용, 연구 진행 상황을 소개함으로써 표정연구의 기본 문제, 연구접근 방법, 연구 결과의 응용 가능성을 개관해보도록 하겠다. 그림 1은 이 연구과제의 개요를 나타낸 것이다. 이 그림에서 보듯이 이 과제의 연구 목표인 표정과 제스처를 인식하고 표현하는 모형과 그것의 기반이 되는 DB를 개발하기 위해서는 내적 감성 상태의 표상 체계를 만드는 작업, 얼굴표정 관련 정보를 체계적으로 정리한 DB 구축 작업, 이 두 작업 결과에 기초하여 표정을 인식하고 표현하는 작업이 수행되어야 한다. 이러한 연구 작업을 연구 문제별로 좀 더 세분하면 내적 상태의 구조 모형 개발, 기초 표정자료 수집 및 DB 구축, 얼굴 표정의 영상적 특징요소 규정 및 수리적 구조 파악, 표정-내적 상태 상관 구조 모형 개발, 표정 인식 및 합성의 전처리 작업을 위한 얼굴의 특징 요소 추출 및 인식, 표정 인식 모형 개발, 얼굴 및 표정 합성 작업의 일곱 부분으로 크게 나누어 볼 수 있는데 여기서는 이들 세부 연구 문제별로 연구 소개와 연구 문제에 대한 논의를 전개하기로 한다.

1. 내적상태의 구조 분석

얼굴표정에 관한 이론적 관심과 DB구축 작업은 정서(emotion)에 대한 심리학적 연구가 발단이 되었다. 얼굴표정은 관찰자에게 그것을 표현하는 사람의 내적상태를 추론할 수 있도록 해 주는 강력한 단서들을 포함하고 있기 때문에 관찰 대상

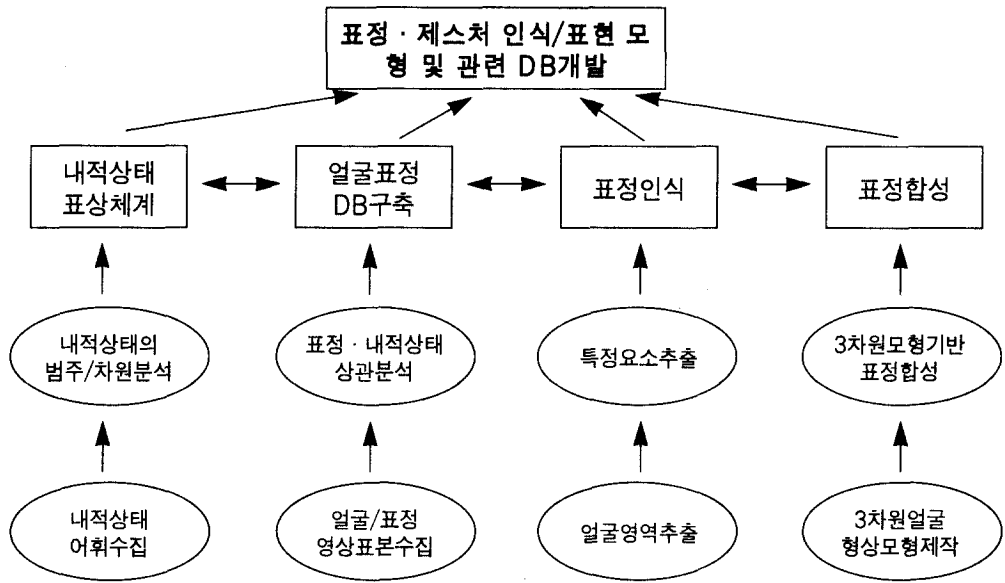


그림 1. '표정/제스처에 의한 감성측정 기술 및 DB개발' 연구 개념도

의 감성 반응, 특히 정서나 내적 행위준비 상태를 측정하는데 매우 유용한 자극 자료로 사용될 수 있다. 감성을 측정하기 위한 입력자료로서 얼굴표정이 유효한 자극이 되기 위해서는 표정의 분류 단계에서부터 그것의 배후 기체로서 체계화가 잘 된 내적상태의 모형이 뒷받침되어야만 한다. 내적상태의 구조를 규명하기 위한 두 가지 주요한 이론적 접근으로 기본정서모형과 차원모형이 있다.

기본정서모형에서는 진화론적, 생리적 근거를 갖는 문화보편적인 기본정서가 있으며 이 기본정서는 각기 독립적인 차원이라고 본다. 기본정서 이론을 주도해온 Ekman(1972)은 행복, 놀람, 공포, 분노, 혐오, 슬픔의 6가지를 기본정서로 보고 이러한 기본정서는 얼굴표정을 통해 잘 표현되며 문화적 보편성이 있어 문화나 언어, 연령층에 상관없이 쉽게 인식된다고 주장하였다. 이러한 전제위에서 Ekman(1978)은 각 기본정서의 얼굴표정을 구성하는 얼굴근육의 움직임을 분석하여 FACS를 구성함으로써 얼굴표정을 통한 정서연구를 주도해왔다. 그러나 이러한 기본정서모형의 주장에 많은 비판이 제기되어 왔다. 특히 기본정서모형의 주요한 이론적 근거가 되는 표정을 통한 정서인식의 문화적 보편성은 방법론상의 문제로 과대 추정되었을 가능성이 있음이 지적되었다. 기존 응답지의 보기에 다른 정서들을 추가하여 선택할 수 있게 했을 때는 인식률이 훨씬 떨어졌으며(Carroll & Russell, 1996), 표정보다 상황에 의존한 판단을 하는 경우가 많은 점, 연구에 이용된 대부분의 표정사진이 일상적인 표정이 아닌 배우의 과장된 표정이라는 점등

이 비판의 대상이 되고 있다(Russell, 1994).

기본정서 모형이 각 정서경험을 개별적으로 다루고자 하는 범주적 접근인 반면에 차원모형은 다양한 정서경험을 이들의 구조를 통해 포괄적으로 이해하려는 접근이다. 정서 경험을 몇 가지 하위차원으로 구조화하려는 시도는 Wundt(1897)이후로 심리학에서 비교적 일찍부터 이어져왔다. Russell(1978)은 각기 기존의 연구들을 종합하여 분석한 결과, 내적상태의 차원으로서 대개 쾌-불쾌(pleasant-unpleasant)와 각성-수면(arousal-sleep)의 두 차원은 비교적 일관되게 나타나고 있으며 3차원 이상은 통제소재나, 경험의 깊이, 주의정도 등 연구자마다 다양한 개념을 제안하고 있다(Averill, 1975; Russell & Bullock, 1986).

내적상태를 기술하는 어휘를 분석하여 내적체계의 차원을 규명하려고 시도한 국내 연구로 안신호 등(1993)의 연구가 있다. 이들은 선행 연구자들이 각성수준에 큰 부하를 갖는 '졸리운' 등의 단어를 근거없이 포함시킨 점을 비판하며 신체상태와 관련있는 단어들을 배제하고 국어사전을 이용하여 내적상태 기술 단어를 골라 압축하여 차원분석을 하였다. 연구 결과 내적상태의 제 1차원인 쾌-불쾌차원은 뚜렷하게 발견되었으나 각성수준은 발견되지 않았다고 보고하였다. 내적상태 관련 어휘에 관한 경험적 자료를 토대로 요인분석과 군집분석을 실시한 또 다른 국내연구(이만영 등, 1990)에서는 내적상태의 제 2차원으로서 각성-수면의 차원이 추출되기는 하지만 상대적으로 불안정하다고 결론짓고 이러한 차원을 내적상태의 주요 차원으로 포함시킬 것인가에 대해서 유보적인 입장을 취했다.

한국어의 내적상태 구조에서는 어떤 차원이 관찰되는가를 알아보기 위해서는 특정 내적상태를 포괄적으로 반영하는 어휘목록을 선정하는 것이 핵심적인 문제이다. 성격이나 감각, 기분과 구분되는 내적상태를 미리 규정하는 것은 어려운 일일 뿐만 아니라 인위적일 수 있으므로 내적상태라는 개념으로 열어두고 어휘목록을 선정하였다. 또한 차원이 얼마나 안정적으로 관찰되는 것인가를 보기 위해서는 몇 가지의 다른 어휘 목록을 사용한 차원 분석 결과들을 비교할 필요가 있다고 보았다. 표정인식은 타인의 내적상태에 대한 인식 과정과 관련된 것이며 자신의 내적상태 경험에 대한 인식과는 다를 수 있다고 가정하고 두 가지 과정을 나누어 살펴보았다.

(1) 내적상태 기술 관련 어휘의 조사 및 기저 차원 분석차원

자신의 내적상태를 인식하는 경우에 그 내적상태들이 어떻게 구성되었는가를 알아보기 위하여 관련 어휘목록을 선정하고 그 어휘들의 구조를 분석해보았다. 내적상태와 관련된 어휘들을 성격이나 신체상태와 구별하여 추출할 수 있도록 하기 위해서 추출 방법을 다양화하고 내적상태에 대한 어휘임을 피험자들이 숙지하도록 하였다.

1) 방법

가. 내적상태 기술 어휘 조사 및 핵심 어휘 목록 선정

내적체계를 잘 반영하는 어휘는, 많은 사람들이 내적상태의 단어로써 쉽게 떠올리며, 내적체계의 여러 영역을 골고루 포함하는 대표성을 지녔을 것으로 가정하고, 다음의 단계를 거쳐 어휘 22개를 최종선정하였다.

제1단계: 1차 어휘목록 작성

학부생 200명과 30-40대 피험자 150명을 대상으로, 성격이나 신체용어와 구분되는 정서개념을 자연스럽게 유도하기 위하여 먼저 성격, 신체상태를 포함한 몇가지 범주의 예를 들어준 다음, 마음상태의 범주에 해당되는 단어들을 자유로이 연상하게 하고, 빈도순으로 100개의 어휘를 선정하였다.

제2단계: 유사성 평정을 통한 유사단어군 묶기

어휘목록에 매우 유사한 단어들이 많았기 때문에, 유사한 단어쌍을 묶는 과제를 시행하였다. 대학원생 20명에게 100개중 30개의 유사한 단어쌍을 고르되, 유사성 정도에 따라 가중치를 두게하여 유사성 점수를 계산하였다.

제3단계: 유사단어군별 연상빈도를 기준으로 최종 내적상태 어휘목록 선정

유사한 단어쌍 평정실험 결과를 이용하여 유사한 단어들을 묶고 그 빈도를 합산하였다. 대표단어는 그 중 가장 빈도가 높은 단어로 하였다. 이렇게 해서 다시 빈도순으로 목록을 마련하였는데, 이 경우 빈도는 단순히 자주 사용하는 의미만이 아니라, 여러개념을 대표하는 속성으로 해석될 수 있다.

나. 차원 분석

평정자 연세대학교 심리학 교양과목 수강생 161명이 참가하였다.

도구 및 절차 선정된 22개 단어를 두 단어씩 짝을 지어 모두 231개의 단어쌍에 대해 유사성을 7점 척도에 평정하게 하였다.

2) 결과

유사성 평정자료를 다차원척도법(Multi-Dimensional Scaling: MDS)으로 분석할 결과, 첫번째 차원이 63.3%, 두번째 차원은 25%로, 두 차원이 전체 변량의 88%를 설명하였다. 1차원은 “기쁨-즐거움-행복-만족-.....,-싫어함-슬픔-절망”의 순서를 보여 쾌-불쾌 차원으로 해석되었고, 2차원에서는 “부끄러움-긴장-놀

라움-.....-편안함-지루함-나른함”의 순서를 보여 각성수준으로 해석될 수 있었다. 세번째 차원은 전체변량이 5.9%로 설명변량이 작았고 명확한 해석도 어려웠다. 2차원상에 배열된 22개 단어가 그림 2(a)에 제시되어 있다.

(2) 표정 기술 어휘의 차원 분석

자신의 내적상태를 인식하는 경우와 다른 사람의 표정을 보고 그 사람의 내적상태를 인식하는 경우를 분리하여 두 경우에 내적상태의 구조가 다른가를 알아보았다. 그를 위해 얼굴표정을 제시하고 그에 해당되는 내적상태를 기술하도록 하여 표정으로부터 추출된 내적상태 어휘목록을 수집하였다. 수집된 어휘를 차원분석하여 구조를 살펴보았다.

1) 방법

가. 표정인식 관련 어휘목록 선정

피험자 연세 대학교 심리학 교양과목 수강학부생 21명이 참가하였다.

도구 및 절차 자극은 한국 배우(이경영, 심혜진, 이미연)의 표정 사진 105장으로, 피험자에게 얼굴표정을 한 장씩 제시하고, 현재 그 배우의 마음상태를 추론하여 형용사로 기술하게 하였다. 이 과제에서 피험자들이 표정을 보고 기술한 마음상태관련 형용사들을 빈도순으로 정리하여 가장 자주 언급된 어휘 22개를 선정하였다. 이중 9개는 연구1의 목록과 중복되었다.

나. 차원 분석

피험자 연세대학교 심리학 교양과목 수강생 140명이 참가하였다.

도구 및 절차 연구1과 동일한 방식으로 7점 척도의 유사성 평정과제를 시행하였다.

2) 결과

MDS 분석 결과 두 차원이 전체 변량의 95%를 설명하였는데 첫 번째 차원이 89%, 두 번째 차원은 6%로 대부분의 변량이 1차원에 의해 설명되는 것으로 나타났다. 차원은 연구 1과 마찬가지로 제 1차원이 쾌-불쾌이고 제 2차원이 각성-수면으로 해석되는데 별 무리가 없었다. 제 3차원은 1% 이하로 설명력이 매우 작았다. 2차원 평면 상에 배열된 22개 단어가 그림 2(b)에 제시되어 있다.

본적이고 대표적인 단어들일수록 두 차원의 구조로 더 잘 설명될 수 있는 가능성을 시사한다.

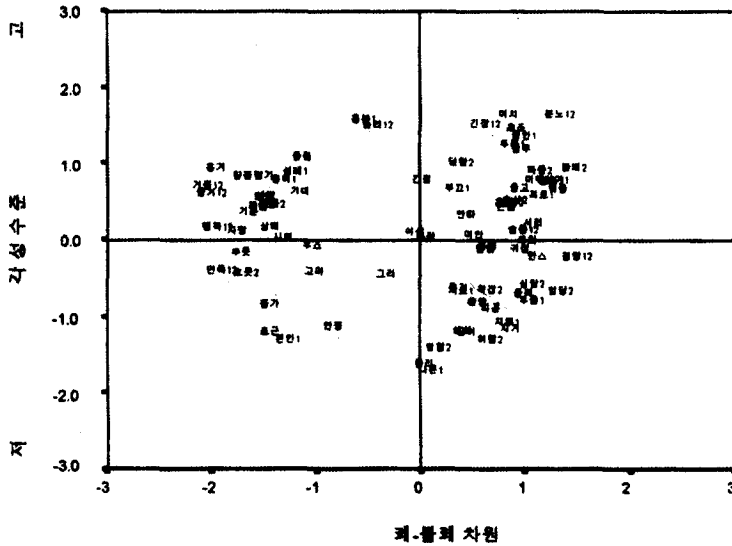


그림 3. 83개 기초 어휘목록의 차원분석 결과

2. 기초 표정 자료 수집 및 DB구축

얼굴표정 연구를 위해서는 얼굴표정의 표준화 된 원자료 및 한국인의 얼굴 프로토타입 DB를 개발하는 것이 필수적이며 이를 위해서는 먼저 다양하고 자연스러운 인물 표정의 자료 수집이 필요하다. DB화 된 표정 자료들은 인물의 내적상태와 관련지어 분류될 수 있으며 표정인식 및 합성에 활용될 수 있다.

MS ACCESS 프로그램을 사용하여 DB를 구축하였다. DB의 구조는 다음과 같은 항목들로 구성하였다.

(1) 대상인물과 성별

국내 영화배우 중 표정연기가 자연스럽고 표정 자료를 충분히 얻을 수 있을 만큼 출연 작품이 많은 인물을 사전 조사하여 남자배우 5명과 여자배우 5명 총 10명의 배우를 대상인물로 선정하였다. DB에는 이들을 명시하였다.

(2) 실제연령과 극중연령

대상인물의 영화 출연 당시의 실제 나이와 극중에서의 나이를 10세를 단위로 하여 어느 범주에 속하는가를 표시하였다.

(3) 영화제목과 장르

표정자료가 표집된 영화를 밝혀두었으며 어떠한 장르에 포함되는가를 밝혀놓았다.

(4) 인물위치

표정이 어느 위치와 방향 또는 각도에서 촬영된 것인가를 15도를 단위로 분류해 두었다.

(5) 인물번호/ 표정번호/ 표정단계

DB 내 표정 사진들의 대상인물의 분류에 용이하도록 고유 번호를 부여하였으며 수집된 각 표정도 일련 번호를 갖도록 하였다. DB 내의 한 표정은 시작-상승-절정-하강-종료의 다섯 단계로 구분하여 표집하였으므로 각 단계에도 시작에서 종료까지 순서대로 일련 번호가 붙여졌다.

(6) 표정관련 내적상태

표정DB를 자극으로 한 연구에서 도출될 수 있는 결과들을 기록할 수 있도록 구성하였다. 표정과 내적상태와의 관련을 평정을 통해 추론하도록 하고 그로부터 각 표정에 대해 내적상태의 범주와 차원값을 명기할 수 있도록 하였다.

(7) 원영상과 얼굴추출영상

원영상은 200×200 화소의 크기로 인물표정을 포착하였으며 얼굴추출영상은 얼굴부분만을 타원형으로 추출하여 입력하였다.

(8) 표집사건 맥락

표집된 표정 자료의 맥락적 정보를 이용할 수 있도록 하기 위해서 해당 표정 장면의 대사나 극중 상황에 대해 기록하였다.

위와 같은 구조를 가진 DB를 근거로 하여 먼저 내적상태 공간의 구성 차원을 추출하고 규정된 내적상태 차원과 얼굴표정의 관계를 규명해 보았다. 얼굴표정과 내적상태의 상관 모형은 얼굴인식의 연구에 기초 자료 및 검증 자료로 사용되었으며 얼굴표정으로부터 표정을 짓는 사람의 내적상태를 추론하는 함수적 관계를 밝히는 것은 역방향으로 얼굴표정의 합성에 이용될 수 있었다.

The image shows a software interface for a facial expression database. It includes several input fields for facial features and a large circular area for a face image. The fields are labeled as follows:

- 이름 (Name): [Redacted]
- 성별 (Gender): 여 (Female)
- 나이 (Age): 20
- 눈높이 (Eye Height): [Redacted]
- 눈폭 (Eye Width): 8, 50, 3
- 코 (Nose): [Redacted]
- 입 (Mouth): [Redacted]
- 턱 (Chin): 2
- 눈노 (Eye No.): 6
- 경도 (Degree): 15도
- 경도 (Degree): 6

At the bottom, there is a question: "대체 닥터인은 뭐하는 사람이야?" and a date: "97/07/27".

그림 4. 표정DB의 구조

3. 얼굴표정의 영상적 특징요소 규정 및 수리적 구조 파악

얼굴표정에 대한 연구에서 연구의 직접적 대상인 얼굴표정의 영상적 특징 요소를 규정하는 작업이 요구되는 것은 당연하며 이를 위해서는 눈, 코, 입 등과 같은 얼굴의 특징요소들의 모양을 정량화하고 표준화할 필요가 있다(Benson, 1995). 이러한 정량화 및 표준화 작업에 얼굴 폭, 눈 간의 거리 등 얼굴 내의 변화 차원들로 구성된 공간인 얼굴 공간(face space)개념이 사용될 수 있다(Valentine 등, 1995).

표정을 만들어내는 얼굴 근육들의 수축과 이완은 근육과 연결된 피부와 조직들을 움직이도록 할 뿐만 아니라 얼굴표면에 주름과 선, 골을 만들고 얼굴 표면의 특징요소들을 다양하게 움직이도록 한다(Rinn, 1991). 또한 얼굴표정은 시선의 방

향, 머리의 움직임, 홍조, 안색의 변화, 땀의 발생, 동공의 확장 같은 자동적 반응들도 수반한다(Oster 등, 1989). 얼굴표정의 연구에 있어 얼굴표정에 수반되는 이와 같은 복합적인 요소들을 모두 포함시키기는 매우 어렵다. 실제로 사람들이 얼굴표정으로부터 내적상태를 추론하는 경우에도 관찰되는 얼굴표정의 모든 요소들에 동일한 중요성을 부여한다고 보기는 힘드므로 복합적인 얼굴표정을 핵심적인 요소들로 나타낼 수 있어야 한다. 얼굴표정을 결정짓는 표정 요소에 대해서는 얼굴의 개별적 특징요소들을 강조한 입장과 특징요소들 간의 조합 형태에 초점을 둔 입장이 있다.

눈썹, 눈, 코, 입 등은 얼굴의 특징을 가장 잘 파악할 수 있도록 해 주는 얼굴 표면의 표지물(landmark)이다. 이러한 표지물들은 얼굴을 3차원 표면으로 보았을 때 색과 밝기가 가장 민감하게 변하는 곳이며 얼굴 근육의 움직임이 가장 잘 반영되는 곳이다. 이러한 얼굴 표면의 특징요소들이 얼굴표정을 결정짓는다는 입장은 특징요소들 간의 상대적 중요성에 대해서도 동시에 연구하고 있다. 왜냐하면 모든 특징요소들이 같은 중요성을 가지고 그로 인해서 특징요소들 간에 우선 순위가 존재하지 않는다면 그 요소들을 모두 고려해야만 할 것이며 이러한 과정은 얼굴표정의 연구에 표정을 이루는 모든 요인들을 모두 포함시키는 것과 유사하게 비효율적이며 비경제적이기 때문이다.

표정이 인식될 때 각 특징요소들이 중요성에 따라 인식의 우선 순위를 가지는가를 알아보기 위한 실험들은 얼굴 표면의 특징요소들이 얼굴표정을 결정짓는다는 전제를 기저에 가지고 있다. Boucher 와 Ekman(1975)은 공포와 슬픔은 눈과 눈가풀로부터, 기쁨과 혐오는 뺨과 입으로부터 파악됨을 밝혀내었다. 이러한 연구 결과는 얼굴표정을 통해 내적상태를 추론하는 경우에 추론의 단서가 되는 부분은 파악된 정서에 따라 다를 것이라는 가설을 입증하였다. Fraser, Craig 와 Parker(1990)도 놀란 얼굴, 기쁜 얼굴 그리고 슬픈 얼굴을 인식하는데는 입이 가장 우세한 특징임을 밝히고 있으며 그 밖의 얼굴표정 특징요소들 간에도 중요성에 따른 위계적 순위가 있음을 밝혀내었다. 일부의 연구들은 근전도(electromyography) 측정을 통하여 특정한 내적상태에서는 얼굴의 특정 부분이 민감하게 반응함을 알아내었다. 눈썹, 눈, 그리고 뺨 부분의 근육들이 내적상태의 정서가(valence)와 강도(intensity)에 따라 다르게 반응하였는데 이 결과는 표정 인식에서 얼굴표정 특징요소들 간에 우선 순위가 있음을 밝힌 이전의 연구 결과들과 일치하는 것이다(Cacioppo 등, 1986; MaCanne 과 Anderson, 1987).

얼굴표정을 인식할 때의 상대적 중요성과 그에 따른 우선 순위에 대한 연구들은 눈과 입이 우세한 요소들임을 일관되게 밝혀주고 있다. 그러나 이와 같은 표정인식에 대한 연구들은 내적상태가 범주적으로 지각된다는 전제 하에 내적상태와 표정을 연관시켰으므로 표정을 통해서 연속적인 내적상태의 변화를 인식하는 과정을

설명할 수 없으며 특징요소들의 변화 강도는 무시된다는 제한점을 가진다. 그럼에도 표정 특징요소들이 상대적 중요성과 그에 따른 우선 순위를 가진다는 사실은 표정으로부터 내적상태를 추론하는 과정을 효율적으로 단순화시켜줄 수 있다는 점에서 의미가 있다.

얼굴표정 특징요소들이 얼굴과 표정인식에 결정적이라는 주장에 대해서, 이러한 특징요소들 자체보다는 특징요소들의 구성과 배치에 대한 정보가 내적상태에 대해 더 큰 예측력을 갖는다는 입장이 있다. 특징요소에 초점을 두는 입장과 특징요소들의 관계를 강조하는 입장이 상호 배타적인 관계는 아니며 따라서 대부분의 연구자들은 특징요소들 자체와 특징요소들의 관계 모두를 중요시해왔다. 대표적으로 Haig(1985)는 얼굴의 특징요소들을 인식하는 순서의 차이에 의해서라기 보다 요소들의 조합에 의해서 얼굴이 인식된다고 보았다. Young, Hellawell 과 Hay(1987)도 여러 사람의 얼굴 특징요소들을 임의로 조합한 합성 사진들을 자극으로 한 실험 결과를 토대로 얼굴인식에서 특징요소들의 상대적 배치에 대한 정보(configurational information)가 중요함을 밝혀내었다. Diamond 와 Carey(1986)는 상하가 바뀐 얼굴이 제시되었을 때 얼굴인식율이 급격히 떨어지는 것은 특징요소들 간의 관계로 얼굴이 구별되기 때문임을 밝혀내었다.

표정을 결정짓는 요인으로서 얼굴 표면의 특징요소들 또는 그들 간의 관계와 조합이 초점이 되고 있다. 두 견해가 상호 대립적이지는 않으며 어느 경우이든지 얼굴표정을 결정짓는 주요 요소들은 얼굴공간을 구성하는 차원들로 작용한다고 볼 수 있다. 여기서는 얼굴공간의 차원으로서 Pilowsky 등(1986)이 얼굴표정을 계량화하기 위해서 고안해 낸 얼굴의 열 두 차원을 사용하였다. 이 열 두 차원은 고정된 얼굴의 점들 사이, 즉 얼굴의 표지물들 간의 거리로 나타내지며 따라서 얼굴 근육들의 움직임의 정도를 나타낸다고 볼 수 있다. 이 열 두 차원에 의해서 얼굴표정의 주된 변화가 반영될 수 있다는 가정 하에 얼굴표정 공간을 구성하는 영상적 특징요소를 규정하고 그 배후의 수리적 구조를 파악해보았다.

1) 방법

가. 얼굴표정의 차원값 측정

측정차원. 얼굴표정을 계량화하기 위한 Pilowsky 등(1986)의 얼굴의 열 두 차원은 그림 5와 같이 얼굴의 각 특징요소들과 근육들의 움직임을 반영하도록 고안되어 있다. 이 차원들은 열 두 쌍의 얼굴의 특징요소들 간 거리로 구성되어 있다.

측정대상 및 절차. 얼굴표정 차원에 대한 값들을 알아내기 위한 측정대상 사진은 얼굴표정DB의 사진 중자연스러운 얼굴이면서 정면얼굴이라는 조건을 만족시키는 사진 150장을 이용하였다. 각 차원값들의 상대적인 비와 차이가 중요하였으

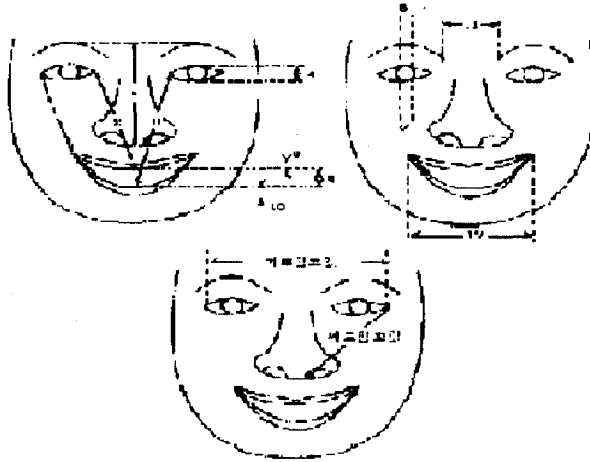


그림 5. Pilowsky 등(1986)이 제시한 얼굴의 열 두 차원 및 참조값:
 세로요인(1, 2, 3, 4, 5, 10), 입너비요인(8, 12), 눈열림요인(6, 7), 입열림요인(9,11)

므로 측정을 위해 사진들을 확대하였다. 크기 600×600 화소로 확대된 사진에서 마우스로 각 차원들의 끝점을 클릭하면 그 점들 간의 거리가 계산되도록 하여 열 두 얼굴 차원들의 값과 두 가지의 참조값 등 열 네 가지의 값을 얻었다. 왼쪽과 오른쪽의 길이로 나타나는 차원은 두 거리의 평균값을 사용하였으며 가로 길이에 대해서는 양쪽 눈꼬리 사이의 거리가, 세로 길이에 대해서는 눈꼬리에서 코끝까지의 거리가 참조값으로 이용되어 얼굴의 움직임이나 사진 크기에서 발생하는 차이를 보정하였다.

나. 표정변화의 수리적 배후구조 파악을 위한 주성분 분석

얼굴표정 변화를 결정하는 영상적 특징요소들 배후의 수리적 구조를 파악하려는 목적과 얼굴표정 공간을 구성하는 차원으로 사용된 Pilowsky 등(1986)의 열 두 차원 간의 독립성을 검증하려는 목적으로 얼굴표정에 대한 열 두 차원의 측정값들에 대해 주성분분석을 실시하였다.

2) 결과

분석 결과 네 가지의 요인이 열 두 차원 전체 변량의 71.0%를 설명하는 것으로 나타났다. 제 1요인은 전체 변량의 36.7%를 설명하였으며 세로요인으로 명명하였다(차원 1,2,3,4,5,10포함). 입너비요인이 전체 변량의 14.1%를 설명하였으며 차원 8, 차원 12가 포함되었다. 눈열림요인이 전체 변량의 10.5%를 설명하였으며 눈동자의 아래 폭, 윗 폭이 포함되었다. 입열림요인이 전체 변량의 9.6%를 설명하

여 제 4요인으로 추출되었으며 윗입술에서 아랫입술까지의 거리와 양눈썹사이의 거리가 포함되었다.

4. 표정-내적상태 상관구조 모형

얼굴표정은 영상을 통해서 나타나고 내적상태는 의미를 기반으로 하므로 얼굴표정을 내적상태에 직접 연결시킨다는 것은 불가능하며 얼굴표정과 내적상태의 직접 연결을 가능하도록 만들어주는 기저의 표상구조 간의 연결 가능성을 살펴보아야 한다. 이러한 목적을 위하여 먼저 얼굴표정-내적상태 간 함수 관계의 성립 조건을 검증하고 그 결과를 바탕으로 내적상태와 얼굴표정의 함수를 탐색해보고자 하였다.

(1) 얼굴표정-내적상태 간 함수 관계 성립 조건 검증

내적상태의 구조에 대해서는 연구 1, 2, 3의 결과로부터 확인된 세 차원이 구성하는 공간모형을 사용할 수 있다. 그러나 얼굴표정에 대해서는 함수관계 추정의 토대가 될만한 공간모형이 없기 때문에 내적상태 공간의 요소들에 유일하게 대응되는 얼굴표정들이 있다는 것을 확인함으로써 함수의 존재 가능성을 간접적으로 검증할 수 있었다. 내적상태의 특성에서 기인하는 무선적인 잡음(random noise)을 배제하기 위해서 통계적으로 유의미한 차이를 갖지만 그 거리가 최소인 신호 대 잡음(signal to noise)의 비가 큰 요소들을 추출하였다. 내적상태 공간의 구별되는 두 점이 얼굴표정 공간 내의 두 점으로 대응된다는 사실은 내적상태와 얼굴표정을 각각 공간모형으로 가정하는 것이 타당하다는 것과 내적상태로부터 얼굴표정으로의 함수가 존재할 가능성을 증명하는 것이다.

1) 방법

가. 내적상태기술단어들의 차원 평정

평정자 연세대학교 재학 중인 남녀 학부생 91명이 평정자로 조사에 참가하였다.

평정대상단어 김영아, 김진관, 박수경, 오경자(1997)가 연세대학교에 재학 중인 학부생 200명과 30,40대 피험자 150명을 대상으로 실험을 통하여 수집한 내적상태기술단어 83개를 평정대상단어로 선정하였다.

평정차원 기존의 연구들이 제시하는 쾌-불쾌와 각성-수면의 차원을 평정차원으로 사용하였다. 평정대상단어 83개 각각을 두 차원에 대해 7점 척도상에 평정하도

록 하였다.

평정절차 평정은 설문지를 통하여 실시하였다. 내적상태 기술 단어들은 임의의 순서로 섞여 있었으며 설문지의 왼쪽에 단어를, 오른쪽에 7점 척도를 주어 평정하도록 하였다.

나. 유의미한 최소한의 거리에 있는 단어 쌍 추출

두 차원에 대해 평정된 83개 단어들의 평균평정값들 중 최소한의 유의미한 거리에 있는 단어 쌍을 찾기 위해서 평균평정값들을 크기 순서대로 배열하였다. 인접한 두 단어들로부터 점차 단계를 늘려가면서 단어들 간의 거리를 조사하였다. 두 단어가 유의한 차이가 있다고 할 수 있기 위해서 초과해야 할 임계치는 Tukey의 HSD로 결정하였다. 단, HSD에서의 MS오차는 83개 문항에 대한 91명의 평정값에 대한 전체변량(pooled-variance)을 사용하였고 단어간의 거리가 뚜렷하게 떨어진 쌍을 찾기 위해서 유의도 수준 .01에서 검증하였다.

다. 추출된 단어에 대응되는 얼굴표정의 결정

응답자 연세대학교에 재학 중인 남녀 학부생 51명이 분류응답에 참가하였다.

분류대상사진 최소한의 유의미한 거리에 있는 단어 쌍들에 대응하는 얼굴표정을 결정하기 위해서 반세범, 한재현, 정찬섭(1997)이 개발한 얼굴표정DB를 이용하였다. DB의 총 2,500장의 사진들 중 위치와 해상도를 고려하여 150장을 선별하였다.

분류의 범주 유목 선정 분류의 보기는 최소 유의미거리에 있는 단어 두 쌍씩 총 8개였다. 쾌-불쾌차원에서의 졸리움, 흥분, 부러움, 놀라움, 각성-수면차원에서의 나른함, 허탈함, 졸리움, 흥가분함의 7개의 단어와 '해당없음'을 포함하여 8개의 항목이 보기로 주어졌다.

분류절차 각 응답자는 개별적으로 각 사진 속의 표정을 8개 분류유목 중 하나에 분류하였다. 마우스를 이용하여 얼굴표정 사진 대상인물의 내적상태를 가장 잘 기술하는 단어를 보기들 중에서 선택하여 클릭하도록 하였다.

라. 대응되는 얼굴표정 간의 거리 계산

평정자 연세대학교에 재학 중인 남녀 학부생 65명이 평정자로 참가하였다.

평정대상사진 쌍으로 추출된 단어에 대응되는 얼굴표정사진들을 대상으로 하였다.

평정차원 평정의 척도는 쾌-불쾌차원에서는 1점이 '쾌', 7점이 '불쾌'로, 각

성-수면차원에서는 1점이 '각성된', 7점이 '나른하고 풀어진'으로 제시되었다.

평정절차 평정이 시작되면 쾌-불쾌차원에 대해 평정해야 할 얼굴표정사진 7장이 무선적으로 먼저 제시된 후에 각성-수면 차원에 대해 평정할 사진 5장이 다시 무선적으로 제시되었다. 제시된 사진 아래에 1에서 7까지의 숫자가 쓰여있는 버튼들 중 화면에 보여지는 대상인물의 내적상태를 추론하여 7점척도 상의 위치를 판단하여 클릭하도록 하였다. 두 단어 간의 거리는 평정값들에 대해 t검증을 실시하였다.

2) 결과

가. 내적상태기술단어들의 차원 평정

쾌-불쾌차원 평정대상단어 83개에 대해서 91명의 평정자들이 1점에서 7점까지의 7점척도에 평정한 평정값의 총 평균은 4.25, 표준편차는 1.83으로 불쾌 쪽으로 치우쳐있었다.

각성-수면차원 평정대상단어 83개에 대해 91명의 평정자들이 1점에서 7점까지의 7점척도에 평정한 평정값의 총 평균은 3.53, 표준편차는 1.73으로 각성 쪽으로 약간 치우쳐있었다.

나. 유의미한 최소한의 거리에 있는 단어 쌍 추출

쾌-불쾌차원 평정평균값의 크기 순 배열에서 네 단계간에서 임계치를 초과하는 단어 쌍인 졸리움(4.22)-흥분(3.13) 쌍과 다섯 단계간에서 유의미한 차이를 보이는 쌍들 중 그 거리가 최소인 부러움(4.55)-놀라움(3.60) 쌍이 추출되었다.

각성-수면차원 평정평균값의 크기 순 배열에서 아홉 단계간에서의 나른함-허탈 쌍, 열 단계간에서의 나른함-공허함 쌍, 열 한 단계간에서의 졸리움-홀가분함 쌍이 유의미한 차이를 보였다. 나른함-허탈과 나른함-공허함이 매우 유사하였으므로 각성-수면 차원에서는 나른함(6.02)-허탈(5.00) 쌍과 졸리움(5.78)-홀가분함(4.58) 쌍을 추출하였다.

다. 추출된 단어에 대응되는 얼굴표정의 결정 및 대응되는 얼굴표정 간의 거리 계산

얼굴표정사진 150장에 대해서 51명의 응답자가 8개 응답 범주 유목에 분류시킨 결과를 토대로 각 유목에서 가장 높은 빈도를 보인 사진을 그에 대응되는 사진으로 결정하였다. 추출된 단어에 대응되는 얼굴표정 및 그들의 함수적 대응관계는 그림 4에 나타나 있다.



졸리움 (82.4%)



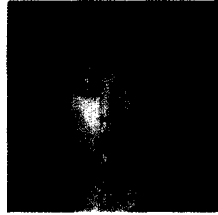
홍분 1 (82.4%)
 $t=6.90^{***}$



홍분 2 (82.4%)
 $t=4.98^{***}$



부러움 (41.2%)



놀라움1 (96.1%)
 $t=10.39^{***}$



놀라움2(96.1%)
 $t=12.69^{***}$

(a) 꽤-불쾌차원



놀라움3(96.1%)
 $t=11.06^{***}$



나른함 (33.3%)



허탈 2 (64.7%)
 $t=-3.21^{**}$



허탈 1 (64.7%)
 $t=2.44^*$



졸리움 (82.4%)



홀가분함(64.7%)
 $t=-.29$

(b) 각성-수면차원

그림 4. 최소유의이거리에 있는 단어 쌍에 대응되는 사진 및 함수적 대응조건 검증 결과

(2) 표정-내적상태의 함수적 관계 모형

타당성이 검증된 두 공간모형을 바탕으로 얼굴표정으로부터 내적상태의 추론 과정을 함수적으로 표현하는 것이 가능함이 확인되었다. 얼굴표정 공간은 얼굴표정을 결정하는 특징요소들의 모양 변화를 반영해야 하므로 이러한 특징요소들이 n 개가 있다면 이 값을 나타내는 n 항의 벡터를 구성할 수 있다. 또한 내적상태 공간에서는 내적상태의 주요 변화를 포착할 수 있는 r 차원을 추출할 수 있다. 내적상태 공간에서 얼굴표정 공간으로의 함수가 있으므로 두 공간 내의 차원값을 측정하고 그를 바탕으로 함수 관계를 규명하였다.

1) 방법

가. 얼굴차원값 측정

(1)에서 측정된 얼굴표정의 열 두 차원값을 사용하였다.

나. 얼굴표정의 내적상태 차원값 평정

평정자 연세대학교에 재학 중인 남녀 학부생 111명이 평정자로 참가하였다.

평정차원 대상인물의 내적상태를 추론하여 주어진 차원에 대해 7점 척도로 평정하도록 하였다. 내적상태의 차원은 쾌-불쾌, 각성-수면차원, 외부-내부지향차원을 제시하였다.

평정절차 평정대상사진 150장을 30장씩 다섯 묶음으로 나눈 뒤 한 평정자가 그 중 두 묶음만을 평정하도록 하였다. 따라서 한 평정자가 평정한 사진은 총 180장 ($30\text{장} \times 3\text{차원} \times 2\text{묶음}$)이었다.

다. 얼굴표정 차원값과 내적상태 차원값 간의 관계 구조 분석

얼굴표정 차원값과 내적상태 차원값의 관계를 조사하기 위해 중다회귀분석(Multiple Regression Analysis), 정준상관분석(Canonical Correlation Analysis)을 실시하였다.

2) 결과

가. 얼굴표정의 내적상태 차원값 평정

내적상태의 세 차원에 대해서 150장의 얼굴표정사진을 평정하도록 한 결과 쾌-불쾌차원의 평균과 표준편차는 $4.34(\pm 0.95)$, 각성-수면차원은 $3.35(\pm 1.20)$, 외

부지향-내부지향차원은 3.67(±1.45)였다. 조사대상사진 150장은 볼쾌 쪽, 각성 쪽, 외부지향 쪽으로 약간 치우쳐 있으며 표준편차가 1점 내외로 비교적 전 척도 차원에 걸쳐 분포되어 있었다.

나. 얼굴표정 차원값과 내적상태 차원값 간의 관계 구조 분석

① 중다회귀분석 결과

얼굴표정의 열 두 차원이 내적상태의 세 차원 상에서의 변화를 설명하는 정도를 알기 위해서 중다회귀분석을 실시한 결과 내적상태의 쾌-볼쾌차원은 21.5%, 각성-수면차원은 40.4%, 외부지향-내부지향차원은 44.5% 정도로 얼굴표정차원에 의해 설명되었다. 또한 이 결과는 내적상태의 쾌-볼쾌차원과 각성-수면차원에서의 평정값을 예측하도록 하는 얼굴 차원들 간에는 큰 차이가 있음을 보여준다. 내적상태의 쾌-볼쾌차원은 얼굴표정차원 중 주로 입의 가로의 길이에 의해서 설명된 반면 각성-수면차원은 눈과 입의 세로의 길이에 의해서 설명되었다. 제 3차원인 외부지향-내부지향차원의 경우에는 제 2차원인 각성-수면차원과 크게 차이가 없었다.

표 1. 얼굴표정차원들이 내적상태의 세 차원을 예언하는 정도에 대한 중다회귀분석 결과

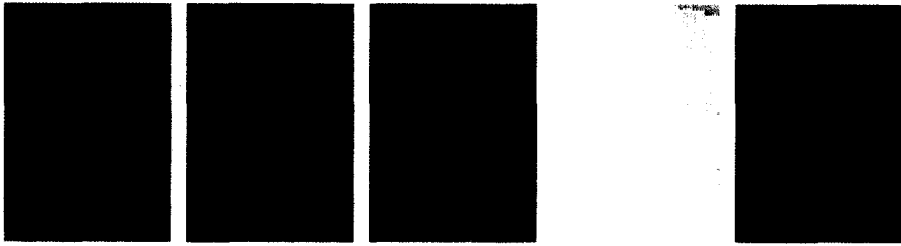
내적상태차원	R ²	얼굴차원	Beta	t	p		
쾌-볼쾌 F(1,148)=40.644, p<.001		12	-.464	-6.375	.000		
		각성-수면	.404	4	-.462	-6.264	.000
		F(5,144)=19.501, p<.001	9	-.331	-5.060	.000	
			5	.323	4.394	.000	
			6	.315	4.857	.000	
11	.211	3.237	.001				
외부-내부지향 F(5,144)=23.066, p<.001		4	-.535	-7.234	.000		
		9	-.387	-6.094	.000		
		5	.360	5.012	.000		
		6	.273	4.301	.000		
		7	-.153	-2.279	.024		



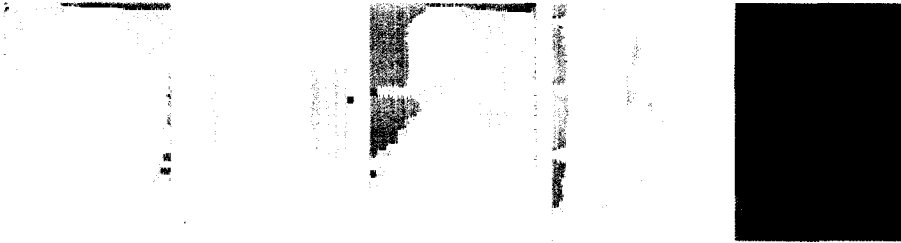
② 정준상관분석 결과

얼굴표정의 열 두 차원이 내적상태를 설명하는 경향을 더 자세히 알아보기 위해서 얼굴표정과 내적상태의 상관이 최대가 되도록 얼굴 차원들과 내적상태 차원들을 모두 조합하여 표현하는 정준상관분석을 실시하였다. 정준상관분석 결과로 얻은 정준함수들에서 얼굴표정의 열 두 차원이 내적상태 변량의 53.4%, 42.1%, 13.6%를 설명하였다. 이들의 설명력은 회귀분석을 통해 얻어진 설명력보다 크지만 유사한 결과를 보여주었다.

세로요인



입너비요인



눈열림요인



입열림요인

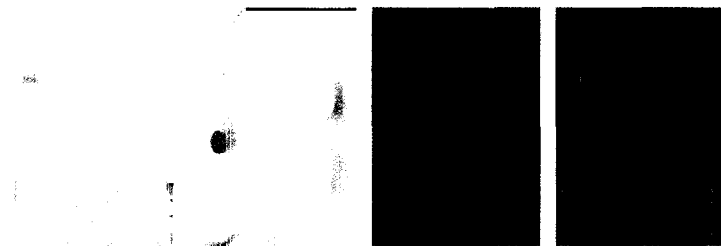


그림 5. 주성분분석결과 도출된 요인들에 대한 요인점수의 배열에서 등간으로 추출한 사진

5. 얼굴의 특징요소 추출 및 인식 : 표정인식 및 합성의 전처리 작업

얼굴 및 요소 추출을 위한 특징 연구 자체도 중요하지만 연구되어진 특징을 영상 데이터로부터 효과적으로 획득하기 위한 영상처리 기법에 대한 연구도 중요하다. 이들 기법은 얼굴 및 요소 추출을 위한 효율적인 특징들이 선정되는 과정에서 추가적으로 포함되어 개발되어야 할 과제이나 이러한 특징들에 대한 연구와 병렬적으로 수행되고 있으므로 먼저 범용적으로 사용할 수 있는 영상처리 기법들의 특징에 대해 연구하고 시스템에 적용하여보았다.

1) 방법

가. 얼굴 영역 추출

비교적 복잡한 배경하에서 얼굴 크기 변화 및 위치에 무관한 시스템을 구하였다. 이를 위하여, 얼굴의 특징을 가장 잘 표현할 수 있도록 통계적인 분석을 통해 형판을 만들고 이를 영상 내에서 다양한 크기로 변화시키면서 탐색하도록 하였다.

통계적인 분석을 통한 형판 모델링은 얼굴 인식 분야에서 가장 주목받고 있는 기법중의 하나인 주성분 분석법(PCA)이용하였으며, 모자의 영상(Huang, 1994)을 사용하여 얼굴 크기 정규화 및 주성분 분석법으로 인한 시간 복잡도를 줄이도록 하였다.

얼굴 탐색을 위한 주성분 분석 얼굴 영상에 대한 주성분 분석은 얼굴 영상에 존재하는 상대 정보를 추출하여 영상 공간을 얼굴 공간으로 축소하여 얼굴을 특징화하는 것으로 이는 수학적으로 얼굴 영상 집합의 공변량 매트릭스로부터 고유벡터(eigenvector)를 찾는 것이다.

이 벡터들은 영상으로 표현했을때, 얼굴과 유사한 모습을 띄게 되어 고유얼굴(eigenface)(Turk 와 Pentland, 1991)이라고 불리운다. 얼굴 영상 집합내의 변화를 가장 잘 나타내는 값이 큰 고유값에 해당하는 고유얼굴들을 이용하여 얼굴공간을 나타낼 수 있다.



학습영상

Eigen Faces

그림 7. 학습영상 및 Eigen Faces

얼굴을 탐색하기 위한 고유얼굴 학습 얼굴 영상들을 $T_1, T_2, T_3, \dots, T_N$ 라

하면, 이들 영상 집합의 평균은 다음과 같이 정의할 수 있다 -
$$\psi = \frac{1}{M} \sum_{n=1}^M T_n$$

다시, 평균 영상으로 정규화된 각 얼굴 영상들로 부터 구한 공변량 매트릭스의 고유벡터에 대한 매트릭스를 u 라고 하면, 새로운 입력 얼굴 영상(T)을 다음 과정을 통해 얼굴 공간으로 투영(projection)시킬 수 있다 - 이를 통하여 새로운 얼굴 영상은 길이가 n 로 표현 된다. 이때, 평균으로 정규화한 입력 영상과 얼굴 공간에 투영된 영상은 다음 수식을 통해 표현되며, 아래 그림을 통해 비교해 볼 수 있다.

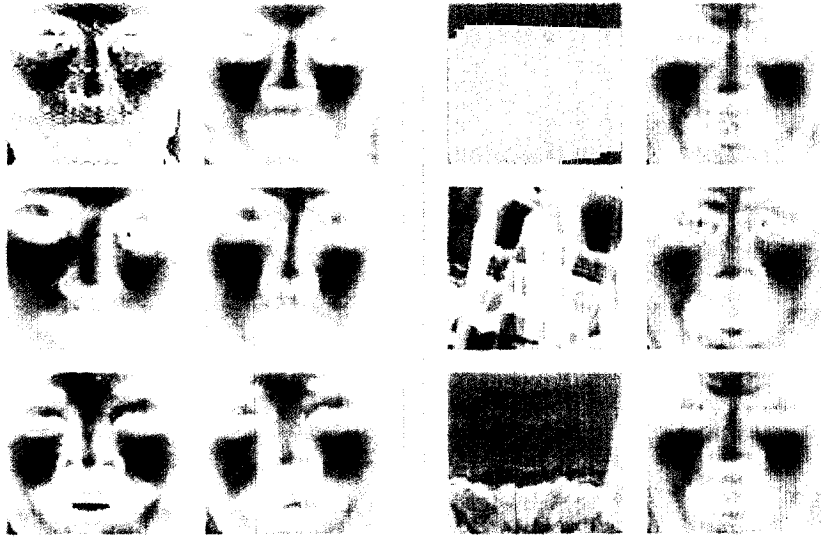


그림 11. 얼굴프로젝션 및 배경프로젝션

그림 11에서 보는 것과 같이 얼굴이 입력되었을 경우 투영된 영상은 이와 비슷한 얼굴이 되고 배경에 해당하는 얼굴이 아닌 영상이 입력되면 입력된 영상과 상이한 모습이 된다. 본 시스템에서는 이러한 특성을 이용하여 얼굴을 탐색하였다.

나. 얼굴 요소 추출

그레이 얼굴 영상에서 얼굴의 각 요소의 특징을 찾아내는 방법으로는 엣지 정보, 그레이 스케일 정보를 이용하는 방법이 있다. 본 연구에서는 각 얼굴 요소별로 구분할 수 있는 많은 정보를 가지고 있는 엣지 정보와 그레이 스케일 정보를 같이 병합하여 이용하였다.

얼굴 요소(눈, 입)의 경계영역 추출 얼굴 각 요소의 엣지를 추출하는 방법으로는 성능이 우수하다고 알려진 소벨(Sobel)연산자를 이용하였다. 이렇게 추출된 엣지를 프로젝션 기법을 이용하여 그레이 스케일을 나타내는 히스토그램을 그린다. 이 히스토그램을 얼굴의 지형적 특징을 고려하여 분석하여 눈, 입이 존재할 후보영역을 찾는다. 이렇게 찾아진 후보영역에 템플릿 매칭 또는 각 얼굴 요소의 부분 특징을 이용한 분석 방법을 이용하여 눈, 입의 경계영역을 찾았다.

한편 영상의 조명에 의한 영향을 감소시키고 얼굴 각 요소의 효과적인 추출을 위하여, 영상을 보강하는 전처리 방법으로 히스토그램 평활화(Histogram

Equalization), Contrast Strecth, Sharpening등의 다양한 영상 처리 기법을 적용하여 최적 성능을 내는 방법을 선택하여 입력영상을 보강하였다.

얼굴 인식 분야에서 주로 쓰이는 대표적인 방법에는 템플릿 매칭(Template Matching), 지형적 특징(Geometric Feature)를 이용한 방법이 있는데, 본 연구에서는 이 두가지를 결합하여 먼저 지형적 특징을 이용하여 얼굴 요소의 후보 영역의 위치를 찾아낸 후, 여기에 다시 템플릿 매칭을 사용하여 눈과 입의 경계영역을 추출하는 방법을 사용하였다.

엣지를 추출하는 방법에서도 각 요소별 특성을 잘 반영할 수 있는 방법을 모색한 결과 수직, 수평으로 분리하여 엣지를 찾는 것이 좋은 결과를 얻을 수 있음을 알 수 있었다. 따라서 각 요소별로 분리된 연산자로 엣지를 추출하고 이를 프로젝션 히스토그램 분석하여 각 요소의 후보 영역을 찾도록 하였다.

이렇게 찾아진 후보 영역에 템플릿 매칭을 이용하여 각 요소를 추출해 보았고, 또 다른 방법으로는 각 요소의 특징을 기반으로 한 지식을 도출하여 이를 이용하여서도 추출해 볼 수 있도록 연구하였다. 지형적 특징을 고려한 프로젝션 히스토그램을 통한 분석과 템플릿 매칭 또는 각 요소의 부분적 특징을 병합하여 사용함으로써 각 요소 추출의 시간을 줄일 수 있었고 효율적인 성능을 보일 수 있었다.

얼굴 요소(눈, 입)의 윤곽선 추출 위에서 언급한 방법을 사용하여 찾아진 얼굴 요소의 경계영역에 대하여 각 얼굴요소의 특징점을 찾을 수 있도록 눈과 입의 윤곽선을 추출할 수 있는 연구를 하였다. 얼굴 영상에 관련된 윤곽선 연구에는 주로 가변 템플릿(Deformable Template) 방법(Yuille, 1992)과 스네이크(Snakes:Active Contour Model)를 이용한 연구가 있다. 기존의 연구들에서는 눈과 입의 경우에는 주로 가변 템플릿 방법을 이용하였고 눈썹, 얼굴 윤곽선 추출에 스네이크를 이용하였다(Huang, 1992).

여기서는 스네이크를 이용하여 눈과 입의 윤곽선 추출 연구를 하였다. 이 방법은 연속성, 곡률, 엣지값으로 이루어진 에너지함수에 대하여 이를 최소화하는 방향으로 수렴해가면서 영상의 윤곽선을 추출해 가는 방법이다(Kass, 1988). 이 방법은 초기값에 영향을 많이 받는 단점이 있지만, 얼굴요소 영역 추출에서 정확한 결과를 구할 수 있다면 이점은 보완되리라 본다. 또한 적당한 파라미터의 결정과 에너지 함수의 조정에 의해 비교적 정확한 눈,입의 윤곽선 추출 결과를 보일 수 있다.

2) 결과

가. 얼굴 영역 추출

얼굴 탐색 결과를 보여 준다. 디지털 카메라 및 CCD카메라를 통하여 얻은 영상 및 영화상에서 비교적 정면 영상만을 선택하여 뽑은 영상들을 대상으로 실험을 해 보았다. 명암 및 크기에 무관하게 얼굴 영역을 잘 찾는 것을 볼 수 있다. 그러나, 모델링된 형편이 다양한 얼굴을 모두 표현하기에는 부족할 수 밖에 없기 때문에, 사용자가 시스템에 호의적인 제한된 상황의 응용분야에 적용해 볼 수 있을 것이다.

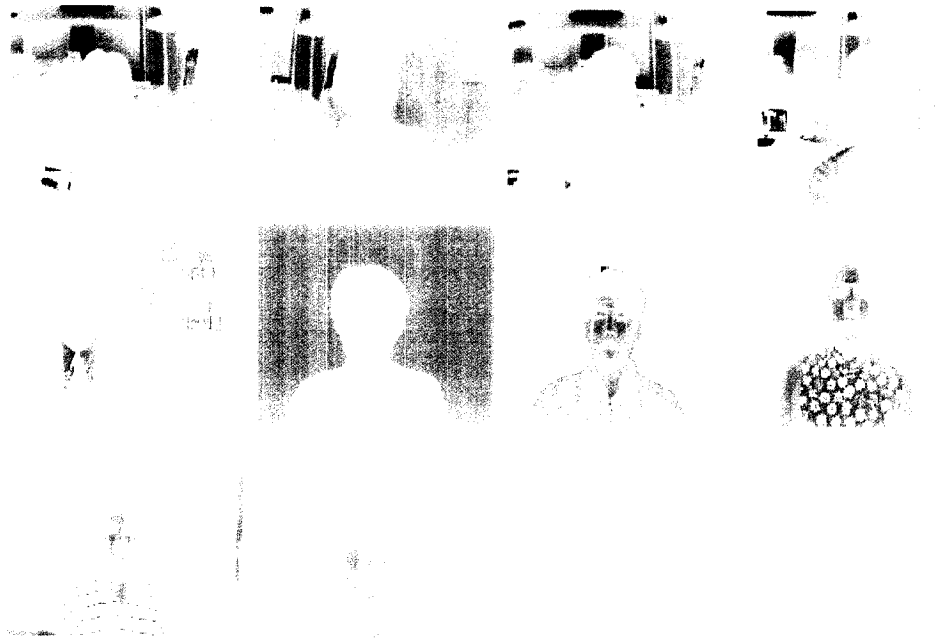


그림 9. 얼굴영역 추출 결과

나. 얼굴 요소 추출

얼굴 요소(눈, 입)의 경계영역 추출 그림 10과 같이 무표정한 얼굴과 약간의 표정이 있는 얼굴에서도 비교적 눈, 입의 경계영역을 잘 추출하고 있다.

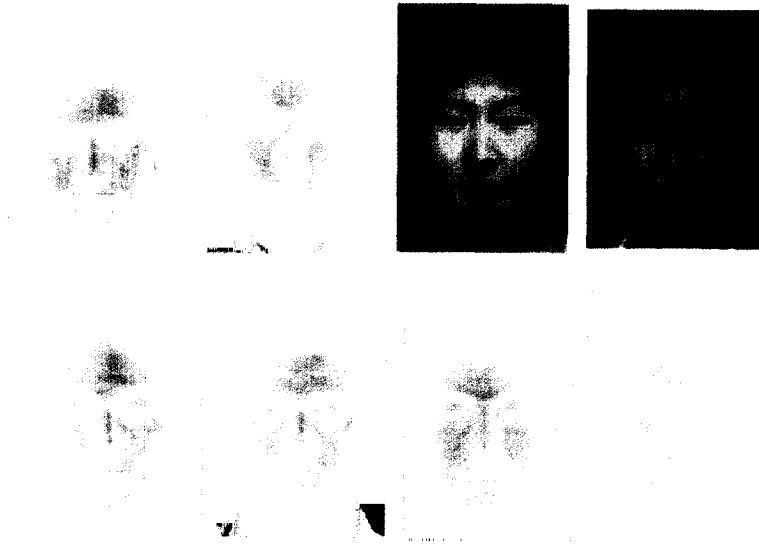


그림 10. 얼굴요소의 경계영역 추출

얼굴 요소(눈, 입)의 윤곽선 추출 앞부분에서 서술한 얼굴 및 얼굴요소 경계 영역 추출에 의한 후보영역에 대하여 윤곽선을 추출한 결과는 그림 14와 같다.

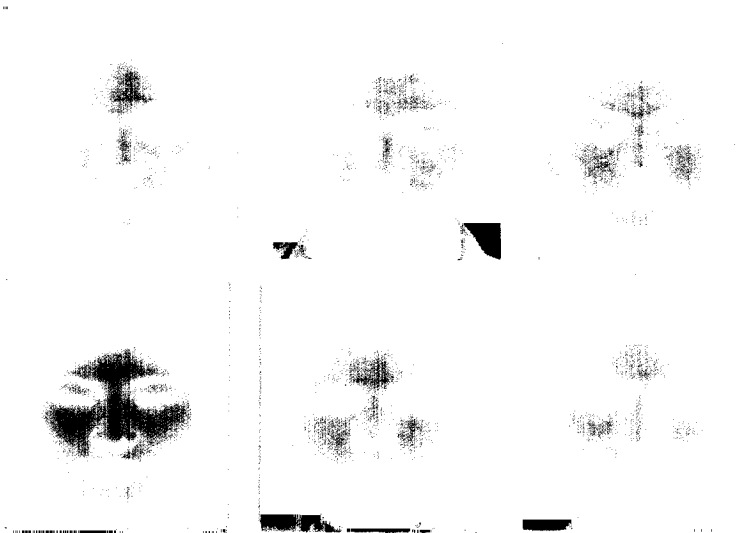


그림 11. 얼굴요소의 경계선 추출

6. 표정인식 모형

얼굴표정의 인식에 대한 연구는 얼굴 합성 및 얼굴인식의 연구와 연구 방법 및 기술을 대거 공유한다. 예를 들어, 얼굴을 인식하기 위해서는 무엇이 얼굴이며 얼굴의 주요 구성요소가 무엇인지 또 그 요소들의 시각적 및 공간적 특징이 무엇인지에 대해 알아야하는데 이러한 구성요소와 관련된 부호들을 표정인식과정과 공유하기 때문에 얼굴인식연구 결과의 상당부분을 표정인식연구에 활용할 수 있다. 또 표정인식을 위해서는 사람마다 각기 다른 얼굴 내, 외부의 모양을 표준화하여야 되는데 이를 위해서는 얼굴합성 연구 방법을 도입할 필요가 있다. 이러한 연구들이 모두 병렬적으로 수행되고 있고 관찰자가 다른 사람의 표정을 통해 내적상태를 추론하는 과정에 대해서도 연구되고 있으므로 그 일부를 대상으로 인식에의 응용 가능성을 검증해보았다.

1) 방법

내적상태와 얼굴표정의 함수 관계를 규명하기 위해 측정된 얼굴표정의 열 두 차원값들을 신경망의 입력층에 입력하였다. 신경망 훈련은 백프로퍼게이션 알고리즘을 사용하였으며 훈련을 위하여 100개의 얼굴표정을 사용하였다. 신경망에 의한 얼굴표정 인식의 정확성을 검증하기 위하여 훈련에 사용되지 않은 38개의 얼굴표정을 실험데이터로 사용하였다.

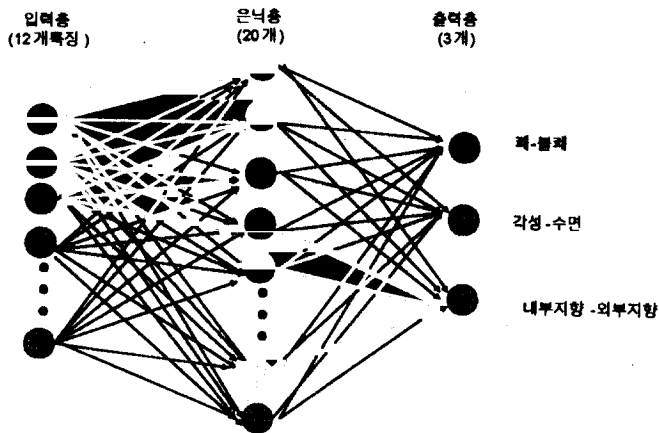
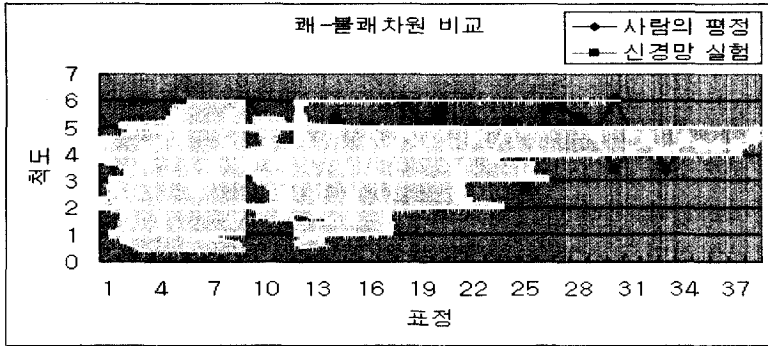
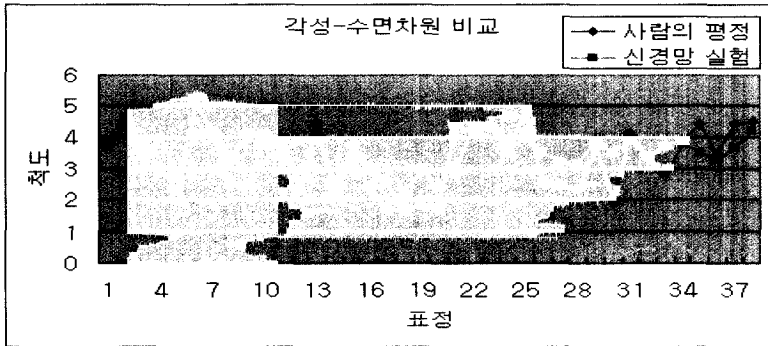


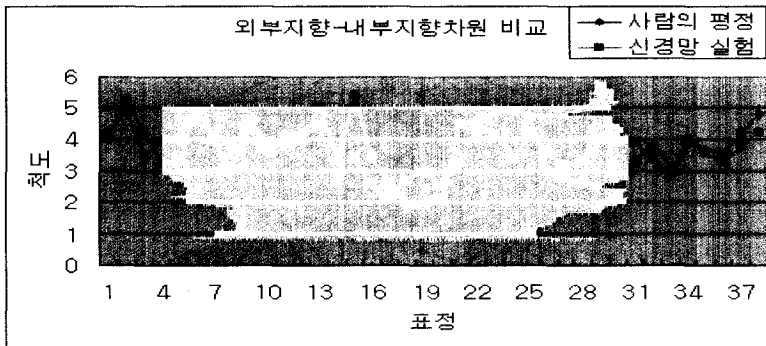
그림 12. 신경망 구조



(a) 쾌-불쾌차원



(b) 각성-수면차원



(c) 외부지향-내부지향차원

그림 13. 시험데이터 38개 표정에 대한 신경망의 인식 결과와 표정에 대한 사람들의 평균 평정 결과의 비교

2) 결과

신경망 훈련에 사용되지 않은 표정 데이터 38개를 사용하여 신경망에 의한 인식의 정확성을 검증하였다. 인식실험 결과 100개의 학습데이터로 훈련된 신경망은 패-블패차원에 대해 68%(26/38), 각성-수면차원에 대해서는 60%(23/38), 그리고 외부지향-내부지향차원에서는 76%(29/38)의 얼굴표정 인식율을 보였다. 인식 수준의 임계치는 사람의 평정결과의 80%~120%를 인식으로 간주하여 인식률을 측정하였다. 표정차원별로 사람의 평정과 신경망의 수행 결과는 그림으로 나타낼 수 있다.

7. 얼굴 및 표정합성

얼굴연구를 가장 먼저 체계적으로 시도한 사람은 Francis Galton(1878)이다. 그는 변이가 너무 심하지 않은 여러 얼굴들을 합성하는 방법을 개발하였는데 이 방법을 사용하면 친숙성과 현실성이 높은 가상의 얼굴이 출현한다는 것과 이렇게 만들어진 얼굴은 사람들이 잘 생긴 얼굴로 지각한다는 것을 발견하였다. Galton이 고안한 얼굴합성 원리는 최근 기술적인 보완을 거쳐 얼굴 미모 결정 요인 연구(Perrett, May 와 Yoshikawa, 1994)와 표준 얼굴 합성 연구(Craw, 1995) 등에 이용되고 있다. 다양한 얼굴들이 지닌 시각적 특징들을 수치로 전환하여 그것의 평균치를 환산하는 기법은 얼굴의 개인차를 통제하고 표정과 관련된 얼굴 내부요소들의 표준화 된 특징들을 포착할 수 있게 해 줌으로써 표정 분류 작업에서 생략될 수 없는 기술적 요소가 된다.

얼굴합성 및 표정합성은 최근에 멀티미디어를 이용하여 사용자에게 가장 친근한 휴먼-머신 인터페이스를 개발하려는 연구에 이용되고 있다. 휴먼-머신 인터페이스는 커뮤니케이션 시스템이나 컴퓨터의 모니터를 통하여 자연스러운 얼굴 및 표정, 제스처, 음성 등을 제공함으로써 사용자와 컴퓨터와의 상호작용을 가능하게 해준다. 특히 얼굴은, 신빙성의 문제가 제기되는 언어에 비하여 인간의 감정과 정서 등의 내적상태를 가장 명백하게 나타내 준다는 점에서 얼굴 표정을 이용한 휴먼-머신 인터페이스 구현 연구의 핵심이 되고 있다. 휴먼 인터페이스의 구현에 있어서 심리학적인 연구를 배경으로, 얼굴 표정에 대하여 감정 상태를 파악한다면, 더욱 효과적인 휴먼 인터페이스를 구현할 수 있다.

1) 방법

가. 얼굴의 3차원 모델링

얼굴합성 단계는 입력된 얼굴영상에 대한 3차원 모델링 단계와 얼굴합성 단계로 나누어진다. 얼굴합성 기법은, 3차원 모델링 단계의 결과로 얻어진 얼굴 형상과 얼굴 텍스처에 대한 정보를 이용하여, 형상 합성 과정과 텍스처 합성 과정으로 이루어진다.

합성을 위한 얼굴의 3차원 모델링은 메쉬 형태로 구성된 표준형상모델을 이용하여 구성한다. 얼굴영상이란 3차원 얼굴이 2차원 화면에 보여지는 결과이므로, 얼굴영상의 합성을 위해서는 영상 속의 얼굴에 대한 충실한 3차원 모델이 필요하다. 이러한 방법은 입력된 2차원 얼굴 영상에 대하여 몇 개의 특징점들이 주어지면, 이 점들을 중심으로 표준형상모델이 대상 인물에 정합되고 각 개인 얼굴의 3차원 형상모델을 얻어내는 것이다. 최근 많이 연구되고 있는 영상기법 중 하나인 모델 기반 영상 부호화 방식의 요소 기술인 3차원 모델링을 통하여 보다 효과적인 모델링을 할 수 있다.

나. 얼굴 합성

Galton의 연구에 착안점을 두어 표준 프로토타입 구성을 위한 얼굴합성 기법을 개발하였다. 표준 프로토타입 구성을 위한 얼굴 합성은 얼굴의 형상 합성과 얼굴의 텍스처에 대한 합성으로 분류하여 이루어진다.

얼굴의 형상 합성 얼굴의 3차원 모델링을 통하여 얻어진 형상의 3차원 좌표값을 이용하여, 얼굴형상을 나타내는 벡터를 X 라 정의할 수 있다. 식 ②는 n 개 특징점의 3차원 좌표로 이루어진 벡터 X 를 의미한다.

$$X = \{x_1, y_1, z_1, x_2, \dots, y_n, z_n\}^T \quad - \text{공식 ②}$$

벡터 X 를 이용하여 합성할 얼굴 영상간의 평균 형상벡터를 구할 수 있다.

$$X = \frac{1}{m} \sum_{i=1}^m X_i \quad - \text{공식 ③}$$

m 은 합성할 영상의 수를 의미한다.

공식 ③에서 구하여진 평균 형상벡터를 이용하여 합성하고자 하는 원영상의 텍스처를 평균형상의 얼굴에 입혀 같은 형상을 갖는 합성할 영상들을 얻는다.

얼굴의 텍스처 입히기(rendering) 얼굴의 형상합성에 의해 구해진 평균 형상에 원영상의 텍스처를 입혀 평균 형상을 갖는 개인의 얼굴을 먼저 구한다. 그러면 합성할 각 영상들은 모두 같은 형상을 갖는 다른 텍스처의 얼굴 영상이 된다. 이에 대하여 텍스처에 대한 합성은 합성할 각 얼굴 영상의 가중치를 조정함으로써 합성된 결과를 얻을 수 있다. 가중치의 선형합으로써 최종의 합성된 얼굴영상을 얻을 수 있다.

다. 표정 합성

FACS를 바탕으로 하여, 3차원 모델링 결과 얻어진 형상모델을 표정에 따라 변형시킨 후, 텍스처 매핑함으로써 표정을 합성하였다. 형상모델을 변형시키는 규칙에 FACS를 적용하였으며, 얼굴의 윗 부분에 해당되는 AU는 눈썹과 관련된 "Inner brow raiser (AU1)", "Outer brow raiser (AU2)", "Brow lowerer (AU4)" 등의 3가지 규칙이 적용되었다. 그리고 얼굴의 아랫 부분은 입술과 턱 부분으로 나뉘어 많은 AU 들로 구성된다. 입술 부분에 대해서는 위/아래, 수평, 사선, 그리고 임의로 움직이는 5가지 형식의 규칙이 적용되며, 턱부분은 귀 아래쪽 점을 중심으로 움직이는 규칙이 적용되었다.

2) 결과

가. 얼굴의 3차원 모델링

그림 17은 디지털 카메라로 입력된 2차원 영상을 3차원 모델링한 결과를 보여주고 있다.

A. 입력영상

B. 3차원 모델링 영상

그림 14. 입력영상 및 3차원 모델링영상

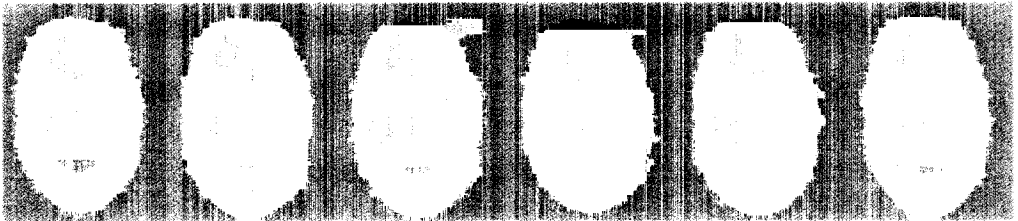
나. 얼굴 합성

합성이 이루어질 입력 영상들에 대해 평균 형상을 구한 후 합성할 얼굴의 텍스처를 입힌 결과는 그림 18의 양 끝 사진으로 제시되어 있다. 왼쪽 사진의 가중치를 감소시켜 가면서 두 사진을 합성한 결과가 그림 18의 사진들이다.

그림 15. 평균형상모델에 따른 합성할 영상(위쪽 양 끝 사진)과
왼쪽 사진의 가중치 변화에 따른 합성 영상
(100%, 75%, 50%, 25%, 0%)

다. 표정 합성

FACS를 적용시켜 표정을 합성한 결과가 그림 19에 제시되어 있다. Happiness는 AU 5, 12, 26의 조합으로 Anger는 AU 4, 5, 7, 10, 25, 26의 조합으로 합성되었다. Surprise는 AU 1, 2, 5, 26의 변형으로 Disgust는 AU 4, 9, 17의 변형으로 합성되었다. Sadness는 AU 1, 4, 5를 변형시켜 표현하였으며 Fear는 AU 1, 2, 4, 5, 7, 20, 25, 26을 변형시켜 표현하였다.



<Happiness> <Anger> <Surprise> <Disgust> <Sadness> <Fear>

그림 16. FACS를 이용한 표정합성

맺는말

지금까지 각 연구 문제별로 연구 내용을 소개하였는데 연구과제 목표인 표정을 인식하고 표현 할 수 있는 시스템을 개발하기 위해서는 이들 각 세부 연구 과제들이 유기적으로 묶이어 하나의 모형으로 통합되어야 한다. 이 통합 모형은 얼굴 표정 자극을 카메라로 입력받아 기초 영상처리를 하는 모듈, 얼굴의 특징요소들을 추출하는 모듈, 확인된 얼굴 특징 요소들을 얼굴공간으로 매핑시키는 모듈, 표정 DB와 표정-내적 상태 상관 모형을 토대로 입력 표정에 해당되는 내적 상태를 추론하는 모듈, 특정 내적 상태에 해당하는 얼굴 표정을 합성하여 출력하는 모듈과 이 모듈들 사이를 연결하는 전환 알고리즘들로 구성될 것이다. 이러한 통합 모형이 개발되면 그림 20에 제시되어 있는 것과 같은 방법으로 모형의 신뢰성과 타당성을 점검하기 위한 전반적인 수행력 검사가 실시될 예정이다. 즉, 이 통합 모형에 임의의 얼굴 표정 영상 두 개를 입력하면 그 중 하나의 영상(#1)을 선택하여 일련의 영상 처리 작업을 거친 뒤 그에 해당되는 내적 상태를 인식하고 이 내적 상태에 해당되는 표정을 짓도록 나머지 영상(#2)을 변형시킬 것이다.

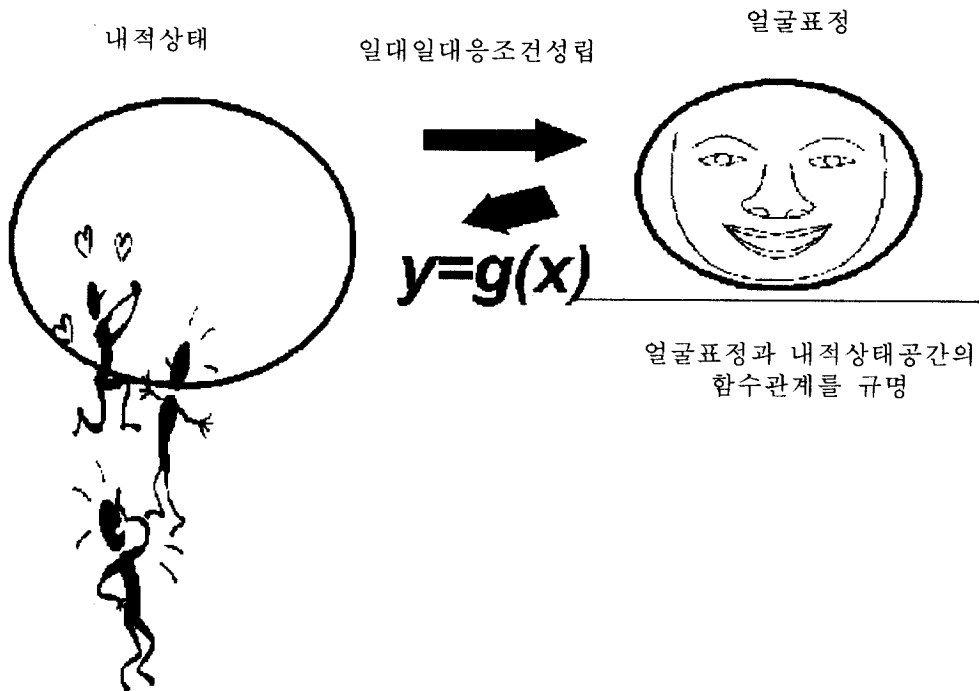


그림 20. 통합모형에 기초한 연구개발 기본 목표 개념도

대부분의 시각 문제들이 그렇듯 표정 인식과 합성의 문제도 그것을 정보처리 관점에서 풀어나가다가 보면 곳곳에서 생각하지 못한 어려운 복병들을 만나게 된다. 그림 17의 통합모형은 표정을 인식하고 합성하는 시스템의 핵심 부분일 뿐, 신뢰롭고 타당하며 유연한 시스템을 개발하려면 이것에 많은 작업 모듈들이 추가되어야 한다. 입력 영상의 화질을 최적화하는 문제, 여러 착시점에서 얻은 2D영상을 3D영상 모형으로 변환하는 문제, 표정이나 제스처의 동적 영상 정보를 이용하는 문제, 내적 상태를 신뢰롭고 타당하게 추론하기 위하여 얼굴표정 정보 이외의 상황 및 맥락, 목소리, 제스처 등 다양한 채널의 정보들을 고려하는 문제, 다양한 인물의 다양한 표정을 확보하여 편파되지 않은 표정 DB를 얻는 문제, 행동적 및 생리적 지표를 사용하여 모형의 신뢰도와 타당도를 검증하는 문제 등이 이러한 작업에 속한다. 다년차에 걸쳐서 수행될 이 연구과제에서는 이러한 문제들을 단계적으로 해결하여 연구 결과가 감성 인터페이스의 개발은 물론 다양한 분야에 응용될 수 있도록 할 예정이다.

참 고 문 헌

- 김영아, 김진관, 박수경, 오경자. (1997). 정서관련 어휘 분석을 통한 내적상태의 차원 연구. '97 한국감성과학회 연차학술대회 논문집, 209-214.
- 반세범, 한재현, 정찬섭. (1997). 표정-내적상태 상관모형을 위한 얼굴 DB 구축. '97한국감성과학회 연차학술대회 논문집, 215-219.
- 안신호, 이승혜, 권오식. (1993). 정서의 구조: 한국어 정서단어 분석. 한국심리학 회지: 사회, 7(1), 107-123.
- 이만영, 이흥철. (1990). 형용사 서술 의미의 구조에 관한 연구. 한국심리학회지: 실험 및 인지, 2, 118-138.
- Aizawa, K., Harashima, H., & Satio, T. (1989). Model-based analysis synthesis image coding system for a person's face. *Signal Process. Image com*, 1(2), 139-152.
- Averill, J. R. (1975). A semantic atlas of emotional concepts. *JSAS Catalogue of Selected Documents in Psychology*, 5, 330 (Ms. No. 421).
- Benson, P. J. (1995). Perspectives on face recognition. Directing research by exploiting emergent prototypes. In Valentine, T. (Ed.). *Cognitive and computational aspects of face recognition*. London: Routledge.
- Benson, P. J., & Perett, D. I. (1993). Extracting prototypical facial images from exemplars. *Perception*, 2, 257-262.
- Cacioppo, J. T., Petty, R. E., Losch, M. E., & Kim, H. S. (1986). Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology*, 50(2), 260-268.
- Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology*, 70(2), 205-218.
- Craw, I. (1995). A manifold model of face and object recognition. In Valentine, T. (Ed.). *Cognitive and computational aspects of facial recognition*. London: Routledge.
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, 115(2), 107-117.



- Ekman, P. (1972). Universal and cultural differences in facial expressions of emotions. In J. K. Cole (Ed.), Nebraska symposium on motivation: 1971 (pp. 207-283). Lincoln: University of Nebraska Press.
- 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.
- Ekman, P., & Friesen, W. V. (1975). Unmasking the face. Englewood Cliffs, N.J.: Prentice Hall.
- Ekman, P., & Friesen, W. V. (1978). Facial action coding system. Palo Alto, CA : Consulting Psychologist Press.
- Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., LeCompte, W. A., Pitcairn, T., Ricci-Bitti, P. E., Scherer, K., Tomika, M., & Tzavaras, A. (1987). Universals and cultural differences in judgments of facial expressions of emotion, *Journal of Personality and Social Psychology*, 53(4), 712-717.
- Fraser, I. H., Craig, G. L., & Parker, D. M. (1990). Reaction time measures of feature saliency in schematic faces. *Perception*, 19, 661-673.
- Frijda, N. H. (1986). The emotions. Cambridge, England: Cambridge University Press.
- Frijda, N. H. (1986). Facial expression processing. In H. D. Ellis, M. A. Jeeves FRSE, F. Newcombe & A. Young (Eds.), Aspects of face processing. Dordrecht: Martinus Nijhoff Publishers.
- Gosselin, P., Kirouac, G., & Dore, F. Y. (1995). Components and recognition of facial expression in the communication of emotion by actors. *Journal of Personality and Social Psychology*, 68(1), 83-96.
- Haig, N. D. (1985). How faces differ - a new comparative technique. *Perception*, 14, 601-615.
- Harashima, H., Aizawa, K., & Satio, T. (1989). Model-based analysis synthesis coding of videophone images. *Trans. on IEICEJ*, Eee-72(5), 452-459.
- Huang, C., & Cohen, C. (1992). Human facial feature extraction for face interpretation and recognition. *Pattern Recognition*, 25(12).

- Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: Active contour models. *International Journal of Computer Vision*.
- McCanne, T. R., & Anderson, J. A. (1987). Emotional responding following experimental manipulation of facial electromyographic activity. *Journal of Personality and Social Psychology*, 52(4), 759-768.
- Oster, H., Daily, L., & Goldenthal, P. (1989). Processing facial affect. In A. W. Young & H. D. Ellis (Eds.), *Handbook of research on face processing*. Amsterdam: Elsevier Science Publishers B. V.
- Pilowsky, I., Thornton, M., & Stokes, B. B. (1986). Towards the quantification of facial expressions with the use of a mathematic model of the face. In h. d. Ellis, M. A. Jeeves FRSE, F. Newcombe & A. Young (eds.), *Aspects of face processing*. Dordrecht: Martinus Nijhoff Publishers.
- Rinn, W. E. (1991). Neuropsychology of facial expression. In R. S. Feldman & B. Rim (Eds.) *Fundamentals of nonverbal behavior*. New York: Cambridge university press.
- Russell, J. A. (1978). Evidence of convergent validity on the dimensions of affect. *Journal of Personality and Social Psychology*, 30, 1152-1168.
- Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological bulletin*, 115(1), 120-141.
- Russell, J. A., & Bullock, M. (1986). Fuzzy concepts and the perception of emotion in facial expressions. *Social Cognition*, 4, 309-341.
- Terzopoulos, D., & Waters, K. (1993). Analysis and synthesis of facial image sequences using physical and anatomical model. *IEEE Trans. on PAML*, 15(6).
- Turk, M. A., & Pentland, A. P. (1991). Face recognition using eigenfaces. *Proc. Computer Vision and pattern recognition* 91, 586-591.
- Valentine, T., Chiroro, P., & Dixon, R. (1995). An account of the own-race bias and the contact hypothesis based on a 'face space' model of face recognition. In Valentine, T. (Ed.). *Cognitive and*



- computational aspects of face recognition. London: Routledge.
- Vetter, T., & Poggio, T. Linear object classes and images synthesis from a single example image. *IEEE Trans. on PAMI*, 19(7).
- Wundt, W.(1897). *Outlines of psychology* (trans. by G.H. judd). New york: Gustav E. strecht.
- Yang, G., & Huang, S. H. (1994). Human face detection in a complex background. *Pattern Recognition*, 27(1), 53-63.
- Young, A. W., Hellowell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, 16, 747-759.
- Yuille, A. L., Hallian, P. W., & Cohen, D. S. (1992). Feature extraction from faces using deformable templates. *International Journal of Computer Vision*.