

# 얼굴표정 DB 및 사용자 인터페이스 개발<sup>1)</sup>

한재현\*, 문수종\*, 김진관\*, 김영아\*, 홍상욱\*, 심연숙, 반세범\*, 변혜란<sup>†</sup>, 오경자\*, 정찬섭\*  
연세대학교 대학원 인지과학 협동과정, 심리학과\*, 컴퓨터과학과<sup>†</sup>

## A DB for facial expression and its user-interface

Jaehyun Han, SooJong Moon\*, Jinkwan Kim\*, Youngah Kim\*,  
Sang Wook Hong\*, Youn-Sook Shim, Sebum Ban\*, Hye-Ran Byun<sup>†</sup>,  
Kyung Ja Oh\*, Chan-Sup Chung\*  
Cognitive Science Program, Department of Psychology\*,  
Department of Computer Science<sup>†</sup>, Yonsei University

얼굴 및 얼굴표정 연구의 기초 자료를 제공하고 실제 표정을 디자인하는 작업의 지침으로 사용되도록 하기 위하여 대규모의 표정 DB를 구축하였다. 이 DB 내에는 여러 가지 방법으로 수집된 배우 24명의 자연스럽고 다양한 표정 영상자료 약 1,500장이 저장되어 있다. 수집된 표정자료 각각에 대하여 내적상태의 범주모형과 차원모형을 모두 고려하여 다수의 사람들이 반응한 내적상태 평정 정보를 포함하도록 하였으며 사진별로 평정의 일치율을 기록함으로써 자료 이용에 참고할 수 있도록 하였다. 표정인식 및 합성 시스템에 사용될 수 있도록 각 표정자료들을 한국인 표준형 상모형에 정합하였을 때 측정된 MPEG-4 FAP 기준 39개 꼭지점들(vertices)의 좌표값들 및 표정추출의 맥락정보를 저장하였다. 실제 DB를 사용할 사람들이 가진 한정된 정보로써 전체 DB의 영상자료들을 용이하게 검색할 수 있도록 사용자 인터페이스를 개발하였다.

### 1. 얼굴표정 DB 구축

#### 1) DB구축 목적 및 개요

최근 영상산업과 감성과학 분야를 중심으로 얼굴 표정의 인식 및 합성기술에 대한 관심이 고조되면서 이러한 기술 구현 작업의 핵심 기반이 되는 표준 표정 자료집의 개발에 대한 요구가 증가하고 있다. 이와 같은 표준화된 원자료 및 한국인의 얼굴 프로토타입 DB를 구축하기 위하여 반세범, 한재현과 정찬섭(1997)은 자연스럽고 다양한 표정 자료를 수집

하였다. 그들은 방법상으로 저렴하고 용이할 뿐 아니라 다양한 자료의 수집이 가능하며 자연스러운 표정들을 표집할 수 있다는 장점에 초점을 두고 신문과 잡지 및 영화비디오와 비디오CD에서 표정을 추출하였다. 수집된 배우 10명에 대한 50장씩의 표정들을 대상으로 기초적 표정 인식 및 합성의 자료로서 사용가능한가의 여부를 확인하였다. 그 결과 비디오 및 CD에서 추출된 사진들은 해상도가 낮아 화질이 떨어지고 대상인물이 정면이 아닌 기울어지거나 돌아간 경우가 많아 자동 인식 및 합성을 위한 자료로서는 부적합하다는 것을 발견하였다. 일차적으로 수집된 자료에 대한 이 같은 한계점을 보완하

1) 본 연구는 과학기술부 감성공학기술개발사업 17-02-A-10의 연구비 지원을 받아 수행되었습니다.  
e-mail : jaehyun@psylab.yonsei.ac.kr

기 위하여 실제 연기를 통한 표정 자료를 수집하였다. 현재 컴퓨터를 이용한 표정인식 및 합성 연구 분야에서는 대부분 연구자 자신이나 일반인들의 표정자료를 이용하고 있다. 그러나 Essa와 Pentland(1995)가 밝혔듯이 훈련받지 못한 사람들에게 특정한 내적상태에 해당되는 표정을 짓도록 하는 것은 무척 어려운 일이므로 전문 연극배우와 연기 전공 학생들을 고용하여 표정을 촬영함으로써 이러한 문제점을 해결하도록 하였다.

단순한 표정자료들의 나열만으로는 실제 얼굴 및 표정 연구를 위한 자료로서의 가치가 부족하므로 수집된 모든 표정 영상자료들이 내적상태 정보를 포함하도록 하였다. 이를 위하여 내적상태 구조에 대한 차원모형과 범주모형을 모두 고려한 평정체계를 개발하여 대규모의 평정실험을 실시, 내적상태 평정자료를 수집하였다. 또한 인식 및 합성을 위한 기초 자료로서 한국인 형상모형(정찬섭, 오경자, 이일병과 변혜란, 1998)에 정합하였을 때 MPEG-4 FAP에서 정의한 39개 꼭지점(vertices)의 좌표값을 결정, 저장해두었다. 다양한 방법으로 수집된 이러한 정보들은 사용자들이 쉽게 알아볼 수 있는 형태로 정리되었다.

## 2) 자료수집 절차 및 방법

### (1) 표정자료 수집

#### 가) TV드라마 배우의 표정 표집

여러 사람의 자연스럽게 다양한 표정을 수집할 목적으로 TV드라마 영상자료로부터 표정을 표집하였다. 영화에서 TV드라마로 표집 대상 미디어를 변경한 것은 영화에서의 일차 표집 결과 영화라는 장르의 특성상 과장된 표정이 많았고 특정 내적상태의 표정(예를 들면, 분노 표정 또는 우스워하는 표정)에 편중되는 경향이 있었기 때문이었다. 사전 조사 결과 TV드라마에 나타나는 표정들이 영화에서보다 일상생활에서 관찰할 수 있는 표정들을 더욱 가깝게 묘사하는 것으로 밝혀졌다.

#### ① 표집 대상 미디어의 유형 및 대상인물

TV드라마 - 일차 표정 자료 수집 과정에서는 10명의 영화배우(남 : 5명, 여 : 5명)에 대한 표정 영

상자료 500장을 수집하였다. 이에 포함되지 않은 대상인물을 선택하여 TV드라마로부터 8명 배우(남 : 3명, 여 : 5명)에 대한 표정 영상자료 400여장을 표집하였다. 총 18명 배우의 표정자료 900여장을 표집하였다.

#### ② 표정 영상자료의 파일 크기 및 형식

이미 수집된 영상자료들과 일관되도록 동일한 크기와 형식으로 저장하였다. 연구에의 활용과 저장 용량을 모두 고려하여 표정 자료의 크기 및 형식은 200×200 화소 크기의 24비트 칼라 비트맵 파일을 사용하였다.

#### ③ 표집 도구

일차 표집과정과 동일한 도구를 사용하였다. 영상편집장치로는 Video-blaster 영상보드와 Bravado 영상보드를 사용하였고 편집 프로그램으로는 Adobe Premier 5.0과 Xing MPEG player 3.12를 사용하였다. 이들 장치를 이용하여 비디오화면을 컴퓨터에 avi 포맷 동화상으로 저장한 후 대상인물의 표정자료를 프레임단위로 분석하여 적절한 자료를 표집하였다.

#### 나) 연극배우의 표정 실연 촬영

영상물에서 추출한 표정자료들의 화질 문제와 대상 인물의 얼굴이 기울어지거나 돌아간 문제를 보완하기 위하여 연극배우 및 연기 전공 학생들의 표정 실연 자료를 촬영하였다.

#### ① 촬영 대상 배우

훈련받지 못한 사람들에게 특정 내적상태에 해당되는 표정을 짓도록 하는 것은 매우 어려운 일이므로 전문 남, 여 연극배우 각 1명과 연기전공 남, 여 학생 각 2명씩 총 6명을 촬영 대상 배우로 선정하였다.

#### ② 촬영 대상 표정

기존의 얼굴표정 연구에 사용된 표정자료들이 대부분 Ekman(1972, 1978)의 여섯 가지 기본 정서와 관련된 정보만을 담고 있었기 때문에 표정과 관련된 다양한 내적상태를 제대로 반영하지 못하고 있다. 이를 보완하기 위한 방법으로서 김영아, 김진관, 박수경, 오경자와 정찬섭(1998)의 연구 결과를 적용, 내적상태를 쾌-불쾌, 각성-수면의 두 차원이 구성하

는 구조로 가정하고 각 사분면 상에 고루 분포하는 83개의 내적상태 기술 어휘를 선택하였다. 내적상태 기술 단어 83개의 캐-블래, 각성-수면 두 차원 상에서의 분포는 그림 1에 제시되어 있다.

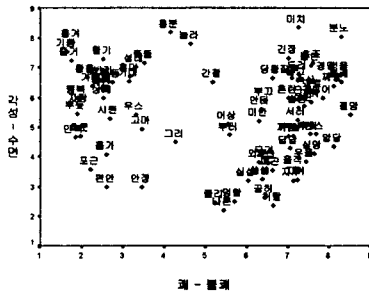


그림 1. 내적상태 기술단어 83개의 캐-블래, 각성-수면 차원 상 분포

### ③ 촬영 영상 자료의 파일 크기 및 형식

추후 영상 분석에 필요한 화질을 얻기 위하여 실연을 촬영한 영상자료는 640×480 화소 크기의 24비트 칼라 비트맵 파일로 저장하였다.

### ④ 촬영 도구

인위적인 상황에서 표정을 지어내는데 에서 비롯되는 부자연스러움을 최소화하고 표정을 짓는 배우들이 자신의 표정을 관찰하면서 확인할 수 있도록 Lyons와 Akamatsu(1998)가 사용한 것과 유사한 장비를 사용하였다. 이 장비는 앞면에 일방향거울이 달린 어둡상자인데 배우가 거울을 보면서 자연스럽게 지시된 표정을 짓는 동안 어둡상자 속에 있는 카메라로 그 표정을 촬영하도록 되어 있다. (그림 2 참조) 카메라는 SONY의 디지털비디오카메라를 사용하였으며 촬영된 영상은 전용 보드인 WVBK 2000과 전용 캡처 프로그램인 DV shef 1.0을 이용하여 편집하였다.

### ⑤ 촬영절차

표정을 촬영할 배우에게는 83개 내적상태를 표현하는 상황 스크립트를 촬영 일주일 전에 주고 충분히 숙지하도록 하였다. 배우가 디지털 카메라가 든 상자 앞에서 일방향거울을 본 상태로 촬영 준비를 마치면 83개 내적상태를 표현하는 스크립트를 하나씩 주어 그에 해당되는 표정을 짓도록 하였다. 실연하려는 표정의 내적상태에 몰입할 수 있는 충분



그림 2. 연극배우 실연의 촬영 장면

한 시간을 준 후 배우의 표정연기를 촬영하였다. 시간 제한은 두지 않았으며 배우가 스스로의 표정을 보면서 만족할 때까지 반복 촬영하였다. 주어진 83가지의 표정은 시간 간격을 두고 2-3일에 걸쳐 촬영하였다. 디지털 비디오 카메라로는 중립(neutral) 얼굴에서 표정의 정점(apex), 다시 중립얼굴에 이르는 전 과정을 촬영하여 필요한 경우에 연속된 프레임으로 사용할 수 있도록 하였다.

### (2) 내적상태 평정 자료 수집

다양한 방법으로 수집된 표정자료들이 얼굴 연구 및 얼굴표정 연구를 위한 의미 있는 자료로 사용될 수 있기 위해서는 대부분의 사람들이 그 표정들을 어떤 내적상태로 인식하는가에 대한 정보가 필요하다. 이와 같은 내적상태 정보를 수집하기 위하여 대규모의 학부생들로 하여금 표정자료들을 평정하도록 하였다.

내적상태의 구조 모형은 크게 여섯 개에서 열 두 개 사이의 독립적이고 단극적(monopolar)인 요인들이 있다는 범주모형 또는 기본정서모형(예. Ekman, 1972)과 내적상태들은 독립적이지 않으며 체계적인 방식으로 관련되어 연속적인 차원들로 구성된다는 차원모형(예. Russell, 1980)으로 나뉜다. 이러한 두 모형의 입장을 모두 수용한다면 일상적으로 사람들을 범주모형에서와 같이 내적상태를 '기쁨', '슬픔' 등의 범주유목으로 판단하는 경향이 있지만 그 기저에는 심층구조로서 '캐-블래', '각성-수면'과 같은 차원으로 구성된 공간이 존재한다는 것을 가정하는 것이 타당하다. 이러한 타당성에 근거하여 다단계로

구성된 평정절차를 고안하였다. 평정은 한 장씩 제시되는 표정 영상자료를 본 후 쾌-불쾌, 각성-수면의 두 차원이 구성하는 평면상에서의 위치를 결정하면 그 점에서 가장 가까운 내적상태 기술 단어 다섯 개가 제시되고 그들 중 가장 적합한 단어를 선택함으로써 평정이 완료되도록 하였다. 동일한 표정자료에 대하여 다수의 사람들이 평정한 후에 이들의 평균 차원값 및 표준편차가 결정되고 그 인물의 내적상태를 가장 잘 기술하는 형용사 9개 및 각각에 대한 평정일치율이 계산되었다. 내적상태 구조의 범주모형과 차원모형을 모두 고려하도록 고안된 프로그램의 예는 그림 3에 제시하였다.

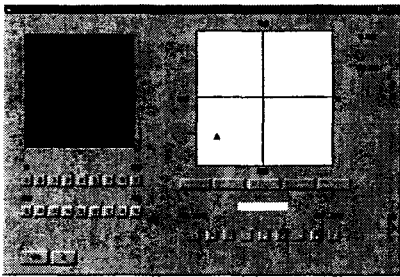


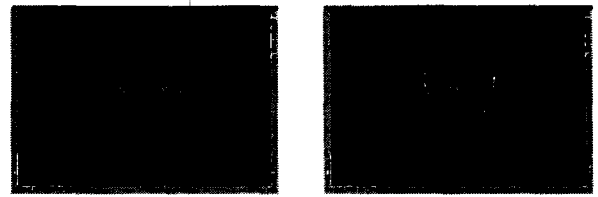
그림 3. 표정평정프로그램의 예

영화 및 TV드라마에서의 추출 방법과 연극배우 실연의 촬영 방법을 사용하여 수집된 표정영상자료 약 1,500장 각각에 30명 이상이 내적상태에 대하여 평정한 자료들을 바탕으로 내적상태에 대한 차원값 및 형용사 범주 정보를 포함하도록 하였다.

### (3) 형상모형 꼭지점(vertices)의 좌표값 수집

내적상태 정보가 포함된 표정자료 각각을 인식 및 합성 기술의 기초 자료로서 사용할 수 있도록 하기 위해서 각 자료들을 한국인 표준형상모형(정찬섭 등, 1998)에 정합, 모형의 정면을 구성하는 1,513개 꼭지점들의 좌표값들을 조사하였다. 정합한 모형은 3차원을 가정하고 있으나 이에 정합된 표정자료들은 2차원 정보만을 가지고 있었으므로 현재의 조사 결과로 추출된 3차원 좌표값(z)의 정확성에 대해서는 앞으로 더 연구해야 할 것이다. 그러나 1, 2차원 좌표값(x, y)은 인식 및 합성의 기초자료로 사용할 수 있다. 그림 4는 표정자료 표본을 한국인 표준형상

모형에 정합한 결과를 보여준다.



(a) 표정자료 표본

(b) 표준형상모형 정합결과

그림 4. 각 표정에 해당되는 구조를 파악하기 위해 표정 표본을 한국인 표준형상모형에 정합한 결과

### (4) 기타정보

수집된 자료들을 실제로 사용할 때에 참고하도록 하기 위하여 영상자료의 대상 인물에 대한 정보 및 표집 맥락 정보를 기록하였다. 대상인물의 이름, 나이, 성별을 기록하였고 영상물에서 추출된 자료들이 영화 또는 TV드라마의 어떤 상황맥락에서 나온 것인지 알 수 있도록 해당표정 장면의 대사나 극중 상황에 대한 정보를 기록하였다.

### 3) layout 결정 및 정보 수록

영상물로부터의 표집과 실연의 촬영을 통하여 수집된 표정자료들과 그 영상자료들에 대한 다단계의 평정절차를 거쳐 얻어진 내적상태 정보와 자동 인식 및 합성의 기초 자료로서 사용하기 위한 표준형상모형의 좌표값들 및 기타 관련 정보들을 쉽게 접근할 수 있도록 DB를 구조화하였다. 이러한 구조화 작업에서 사용상의 우선 순위와 정보의 중요도, 그리고 가시성을 고려하여 DB의 layout을 결정하였다. 최종 DB의 화면 구성은 그림 5에 제시되어 있다.

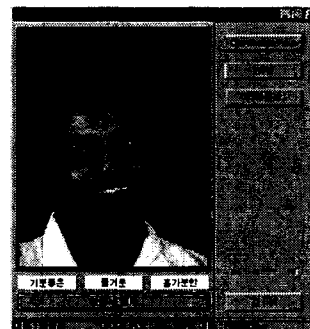


그림 5. 표정DB 화면

## 2. 사용자 인터페이스 개발

### 1) 인터페이스 개발 목적

실제 DB를 이용할 사람들이 쉽고 편리하게 알고자 하는 정보를 찾을 수 있으려면 그 과정을 도와줄 수 있는 검색 도구가 필요하다. 이를 위해서 이미 구축된 DB를 실제 이용할 사람들이 가진 정보와 찾고자 하는 정보를 중심으로 사용자 인터페이스를 구성하였다.

### 2) 인터페이스 설계 내용

표정 DB에 대한 사용자 중심의 인터페이스는 표정자료 검색의 기준, 검색 대상 배우, 그리고 검색 자료의 수를 선택, 지정할 수 있도록 설계하였으며 검색 결과로 제시된 자료들을 선별하여 저장하는 기능을 포함하도록 구성하였다.

#### 가) 검색 기준 선택

DB를 사용하여 표정자료 및 관련정보를 참고하고자 하는 사용자들의 대부분은 스크립트나 상황에서 비롯된 특정 내적상태나 정서에 대한 정보를 가지고 있게 된다. 그러므로 검색의 기초적 기준이 되는 옵션을 다음과 같이 세 가지로 설정하였다.

#### ① 내적상태 기술 단어(형용사 범주)로 검색

스크립트나 상황에서 비롯되는 내적상태 또는 정서는 내적상태를 기술하는 단어 즉, 형용사 형태로 정의될 수 있다. 형용사 범주로 표정자료들을 검색하는 경우에 사용자들은 찾고자하는 표정의 내적상태를 직접 타이핑하거나 또는 주어진 50개의 단어 중에서 선택할 수도 있도록 하였다. 이 단어들은 김영아 등(1998)의 연구에서 추출된 83개의 단어로부터 중복되거나 유사한 단어 또는 저빈도 단어를 제거한 결과로 추출되었다. 내적상태의 쾌-불쾌, 각성-수면의 두 차원 상에서의 83개의 단어 분포를 관찰한 후에 그 중 하나를 선택하는 경우에는 선택된 단어와 가장 의미적으로 유사한 단어로 연결되도록 설계하였다.

#### ② 내적상태의 차원값

내적상태 공간의 기저 차원인 쾌-불쾌, 각성-수면의 두 차원에서의 값으로 표현된 내적상태를 기준으로 그에 해당하는 표정들을 검색할 수 있도록 구

성하였다. 두 차원 상에서의 특정 점을 지적하고 검색하고자 하는 자료의 수를 지정하면 특정 점에서부터 가장 가까운 순서대로 표정 자료들을 찾을 수 있도록 설계하였다.

#### ③ 내적상태의 차원값 범위

각 차원 상에서 일정한 범위 내에 포함되어 있는 표정자료들을 모두 보고자 하는 경우에 내적상태 두 차원에 대한 9점 척도 중 최소값과 최대값을 지정함으로써 자료들을 찾을 수 있도록 하였다.

#### 나) 배우 선택

DB에는 다양한 배우들의 표정자료들이 저장되어 있다. 이들 중 사용자가 보고자하는 배우들만의 자료를 대상으로 검색할 수 있도록 대상 배우의 선택 기능을 포함시켰다. 남녀별로 구분하여 검색이 가능하며 각 배우별 검색도 가능하도록 구성하였다.

#### 다) 자료 수 선택

검색 결과로 몇 장의 자료를 찾을 것인가를 사용자가 결정할 수 있도록 설계하였다. 내적상태 차원값의 범위로 지정하는 경우에는 사진 수를 별도로 선택할 필요는 없다. 그러나 범주 또는 특정 차원값을 기준으로 검색하는 경우에는 찾고자 하는 사진 수를 입력함으로써 각각 지정된 범주에 대한 평정일 치율이 높은 자료부터 그리고 지정된 특정 차원값에 가까운 자료부터 순서적으로 정해진 사진 수만큼 제시되도록 하였다.

#### 라) 검색 결과

검색 기준과 검색 대상 배우 및 자료의 수를 지정한 후의 검색 결과 화면에는 한 번에 다섯 장의 사진을 볼 수 있도록 하였다. 제시된 사진들에 대해서는 각각의 자료 또는 모든 자료를 원하는 자리에 원하는 이름으로 저장할 수 있는 기능을 추가하였다. 뿐만 아니라 검색 결과로 제시된 사진들에 대한 다양한 정보를 다시 찾고자 하는 경우에 대비하여 원 DB화면으로 되돌아갈 수 있도록 설계하였다. 검색 기준 및 옵션 설정 화면은 그림 6에 제시하였다.

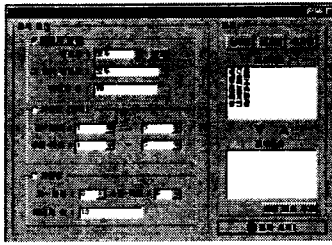


그림 6. 검색 옵션 및 기준 설정 화면

### 3. 후속 연구 방향

DB에 저장된 자료들을 실제 인식 또는 합성 시스템에 적용하여 정보와 구성의 유용성을 검증하고 이를 바탕으로 DB의 내용 및 구조를 지속적으로 개선, 향상시켜 나갈 것이다. 또한 화질이 떨어지는 표정자료들에 대한 화질개선 문제와 기울어지거나 돌아간 얼굴들의 보정 문제를 해결하여 수집된 자료들을 효과적이고 다양하게 사용할 수 있도록 할 계획이다.

### 참고문헌

- 김영아, 김진관, 박수경, 오경자, 정찬섭 (1998). 정서관련어휘분석을 통한 내적상태의 차원연구. *한국감성과학회지*, 1(1), 145-152.
- 반세범, 한재현, 정찬섭 (1997). 표정-내적상태 상관모형을 위한 얼굴 DB 구축. '97 한국감성과학회 연차학술대회 논문집, 215-219.
- 정찬섭, 오경자, 이일병, 변혜란. (1998). 표정/제스처에 의한 감성측정기술 및 DB개발. 과학기술부 감성공학기술개발사업 최종보고서, G - 17 - 01 - 09.
- 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., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto, CA : Consulting Psychologists Press.
- Essa, I. A., & Pentland, A. P. (1995). *Coding, analysis, interpretation, and recognition of facial*

- expressions*. MIT Media Laboratory Perceptual Computing Section Technical Report No. 325.
- Lyons, M., & Akamatsu, S. (1998). Coding facial expressions with Gabor wavelets. *Proceedings of the third international conference on automatic Face and Gesture Recognition*, 200-205.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161-1178.