

표정-내적상태 상관모형을 위한 얼굴 DB 구축

반세범, 한재현*, 정찬섭

연세대학교 심리학과, * 연세대학교 인지과학 협동과정

서울특별시 서대문구 신촌동 134, (우) 120-749

stuntman@psylab.yonsei.ac.kr

Construction of facial database for a facial expression-inner states correlation model

Sebum Ban, Jaehyun Han*, Chansup Chung

Department of Psychology, * Graduate Program of Cognitive Science at Yonsei University

120-749, Shinchon Dong 134, Seodaemun Gu, Seoul, KOREA

요약¹⁾

내적 감성상태와 표정을 연결하는 상관모형을 개발하기 위한 기초자료로서 광범위한 얼굴표정자료를 수집하였다. 다양한 얼굴유형의 표정자료를 DB화하기 위해서는 궁극적으로 50명 이상의 인물표정 자료집이 필요하다고 보고, 우선 영화배우 10명분의 표정자료를 수집하였다. 이를 위해 표정연기가 자연스럽게 평가된 영화배우 10명을 선정하고, 이들이 출연한 50여편의 영화비디오에서 표정자료를 수집하였다. 한사람의 인물당 50개의 표정을 한 세트로 하여 총 10명분 500가지의 표정자료수집을 완료하였다. 각각의 표정은 시작-상승-절정-하강-종료의 연속 5단계를 비율적으로 표집하여 동적정보를 포함하고 있다. 이들 자료들은 등장인물의 출연 맥락이나 상황별로 분류하여 표정관련 맥락단서를 기록하였고, 후속연구를 통해 각 표정에 대응되는 내적상태를 기록할 수 있도록 DB화 하였다. 이와 같이 구조화된 자료들은 표정인식과 합성의 입출력 모형을 위한 기초자료로 활용될뿐 아니라, 한국인의 얼굴 프로토타입 DB 구축 및 향후 표정관련 연구의 타당성을 검증할 수 있는 Benchmarking 정보를 제공할 수 있다.

서론

얼굴표정의 표준화된 원자료 및 한국인의 얼굴 프로토타입 DB를 구축하기 위해서는 먼저 다양하고 자연스러운 인물표정의 자료수집이 필요하다. 본 연구에서는 이들 표정자료를 인물의 내적상태와 관련지어 분류하고, 이를 표정인식 및 합성에 활용할 수 있도록 구조화된 DB를 구축하기 위한 기초자료를 수집하였다.

(1) 자연스럽게 다양한 표정자료의 수집방법

기존의 얼굴표정 연구에서 사용된 표정자료들은 대부분 Ekman(1972, 1978)의 기쁨, 슬픔, 놀람, 공포, 혐오, 분노 등의 6가지 기본정서와 관련된 정보만 담고 있기 때문에 인물의 다양한 내적상태를 제대로 반영하지 못하고 있다. 본 연구에서는 이러한 문제를 해결하기 위하여 이 6가지 정서뿐만 아니라 다양한 내적상태를 반영하는 얼굴표정 자료를 수집하기로 하였다. 이전의 연구는 대부분 특정인물이 연구자의 주문에 의해 지어내는 포즈된 얼굴표정을 사용하였다. 그러나 잘 훈련받지 못한 사람들로 하여금 특정한 내적상태에 해당되는 표정을 지어내게 하는 것은 무척 어려운 일이며, 특히 슬픔이나 공포와 같은 정서는 자연스럽게 표현하기가 무척 힘들기 때문에(Essa, I. A. & Pentland, A. P., 1995) 이런 방법을 사용하는 경우 수집된 자료의 신뢰도와 타당도에 문제가 있을 수 있다.

따라서 배우나 사람들로 하여금 특정한 표정을 짓도록 하는 것은 다음과 같은 문제점이 있다.

- ① 특정 표정자세를 취한 표정의 경우, 실험자의 지시에 맞게 적절한 표정을 지었는지 평가하기 힘들다.
- ② 연구자가 지시한 표정내용이 특정모형이나 이론에 근거하여 편파되었을 가능성이 있다.
- ③ 소수의 poser를 사용하면 얼굴모양 및 표정의 개인차를 반영하기 힘들고 실생활의 얼굴표정과 거리가 멀어지게 된다.

이러한 문제를 극복하기 위하여 본 연구에서는 비디오와 신문/잡지등의 대중매체를 통해 다양하고 광범위한 표정자료 수집을 하였고, 이 방법은 다음과 같은 장점이 있다.

1) 본 연구는 과학기술처 G7 감성공학과제 17-01-09의 연구비 지원을 받아 수행되었습니다.

- ① 저렴하고 용이한 방법으로 다양한 표정자료가 수집가능하다.
- ② 특정표정에 해당하는 실제상황이 있거나 그와 유사한 script가 있으므로 자연스러운 표정의 수집이 가능하다.

(2) 표정인식에서 동적정보의 중요성

얼굴의 표정변화를 보고 인물의 내적상태를 추론하는데에는 각 표정별로 민감한 부위가 다르다. 이것은 각 표정별로 얼굴특징요소의 중요도가 다르게 나타난다는 것을 암시한다. 그러나 표정인식에 있어서 특징요소 그 자체보다는 얼굴표면의 동적인 운동정보가 더 중요하다(Bassili,1979). 얼굴의 특징요소를 잘 알아보지 못하게 검게 칠한후 무선적으로 하얀 점을 찍고 표정을 지을 때, 얼굴표면의 운동정보만으로도 정확한 표정인식이 가능하였다.

대부분의 표정연구에서 사용된 정치화상 표정자료는 특정 내적상태의 절정부분을 표집한 것인데, 이러한 일부분의 정보만으로는 표정의 사실감이 부족하며, 동적정보가 포함된 표정자료가 더욱 자연스럽다는 것을 알 수 있다.

본 연구에서는 이와같은 점을 고려하여 하나의 연속적인 표정을 시작-상승-절정-하강-종료의 5단계에 맞추어 비율적으로 표집하였다.

(3) 표정과 내적상태의 연관

얼굴 표정을 내적상태에 연관짓기 위해서는 먼저 잘 정의된 내적상태 기술체계가 필요하다. 일반적인 정서와 내적상태 연구에서는 기본정서를 가정하는 경향이 있고, 이들은 정서를 독립적이고 불연속적인 요소로 간주하는 범주모형을 지지하고 있다(Ekman, P., 1972; Izard, C. E., 1982). 이러한 범주모형에 따르면 얼굴표정 역시 몇가지 기본정서로만 구분될 뿐이다.

그러나 얼굴이라는 공간에 대응되는 내적상태는 차원모형(Osgood, C. E., 1966 ; Schlosberg, H., 1952)으로 표상하는 것이 유용하다. 얼굴표정은 다양한 특징요소와 근육운동의 조합으로 기술될 수 있는데, 이러한 표정변화는 몇가지 내적상태의 차원상에서 연속적인 변화량의 조합으로 나타낼 수 있기 때문이다.

특히 Russell(1980)의 원형모형에서는 '쾌-불쾌'와 '각성-수면'의 2차원 공간상에 다양한 내적상태를 배열하여 놓고, 이를 그림1과 같이 얼굴표정의 변화에 따른 배열로 대응시켜 놓았다(Russell, J. A., 1997).

얼굴표정과 내적상태를 이와같은 방식으로 대응시켜 놓으면 범주모형에서 생략된 부분을 보완해 놓을 수 있다. 이에따라 특정한 얼굴표정을 입력하면 그에 해당되는 내적상태를 출력할 수 있는 표정-내적상태 상관모형을 개발하고, 표정인식 및 합성 시스템으로



그림 1. Russell(1980)의 원형모형 내적상태 공간에 배열된 얼굴표정의 변화. 가로축이 쾌-불쾌, 세로축이 각성-수면의 차원에 대응되고, 중앙은 중립표정에 대응된다.

구현할 수 있게되는 것이다.

본 연구에서는 이러한 점을 고려하여, 수집된 표정자료와 내적상태를 연관시킬 수 있도록 DB 구조를 결정하였으며, 후속연구에서 표정-내적상태 평정절차를 거친 후 상관모형을 개발할 예정이다.

자료수집 방법

(1) 표정자료 수집의 원칙설정

자연스럽고 다양한 표정자료를 수집하기 위해 다음과 같은 자료수집원칙을 설정하였다.

가) 표집대상 미디어의 유형

① 신문/잡지

사건/인터뷰 기사와 광고 등과 함께, 얼굴 또는 스냅 사진이 많이 실려있고 양질의 화질로 되어있는 매체를 중심으로 표집한다.

② 영화/비디오

시중에 유통되고 있는 방화 비디오와 TV 연속극의 코믹, 애정, 액션, 역사, 기타 장르에서 동일 비율로 인물표정을 골고루 표집한다. 코믹이나 액션장르와 같은 영화에서는 등장인물의 표정이 과장되기 쉽기 때문에 여러 가지 다양한 장르에 걸쳐 표집하여야 인물표정을 골고루 수집할 수 있다.

나) 표정유형별 표집비율

중립얼굴과 표정이 뚜렷한 얼굴을 1:20의 비율로 표

집한다. 중립얼굴은 얼굴모양의 프로토타입 작성에 활용될 수 있으며, 얼굴표정의 절대영점과 같은 기준점 역할을 한다. 여기서 중립얼굴이라 함은 모든 얼굴근육이 이완된 상태에서 많이 벗어나지 않은 얼굴로서, 내적상태를 추론하기 힘든 증명사진과 같은 '무표정함' 얼굴을 말한다.

다) 표정파일의 크기와 형식 및 유형

① 크기(size)와 형식

표정인식과 합성, 그리고 표정-내적상태 상관연구에 활용될 수 있는 적절한 자료크기와 함께 저장용량등을 고려한다. 표정한 얼굴은 200x200 pixel 크기의 24비트 칼라 비트맵 파일로 저장한다.

② 표정자료의 유형

하나의 표정을 5단계의 연속적인 동적정보가 담긴 화상으로 수집한다. 즉, 표정변화가 일어나는 연속동작을 시작-상승-절정-하강-종료의 5단계에 맞추어 비율적으로 포착한 사진을 저장한다.

라) 대상인물의 위치

상반신 이상의 정면 및 45도 측면사진, 85도 측면사진을 5 : 3 : 1의 비율로 표집하며, 좌우 5도 각도 범위의 편차를 허용한다. 측면사진은 얼굴의 3차원 입체구조를 표현할 때 활용될 수 있다. 원본 사진

에서 얼굴부분만 따로 뽑아낸 사진을 저장하여 얼굴 인식 자료로 활용한다.

마) 대상인물의 특성

① 연령층

어린이와 노년층을 제외한 일반 연령층의 인물로서 10대-20대, 30-40대, 50대 이상의 세 연령군으로 나누어 표집하고, 남녀비율은 가급적 비슷하게 한다.

② 영화배우의 경우 한사람의 얼굴을 계속 추적하여 50가지의 표정을 한세트로 하여 총 10명분의 인물표정 자료를 수집한다.

③ 대사내용과 상황맥락을 근거로하여 해당표정에 대한 내적상태 추론단서를 함께 입력한다.

(2) 비디오 편집을 통한 표정자료 수집

앞에서 설정한 자료수집 원칙에 따라 국내 영화배우 중 표정연기가 자연스럽게 표정자료를 충분히 얻을 수 있을만큼 출연작품이 많은 대상인물을 사전 조사하여 남자배우 5명과 여자배우 5명, 총 10명의 배우가 선정되었다. 이들 영화배우가 출연한 방화비디오의 목록을 조사하여 총 50여 편의 영화비디오를 선정하고 영상편집을 하였다. 영상편집장치로는 Video-blaster 영상보드와 Bravado 영상보드가 사용되었고, 편집 프

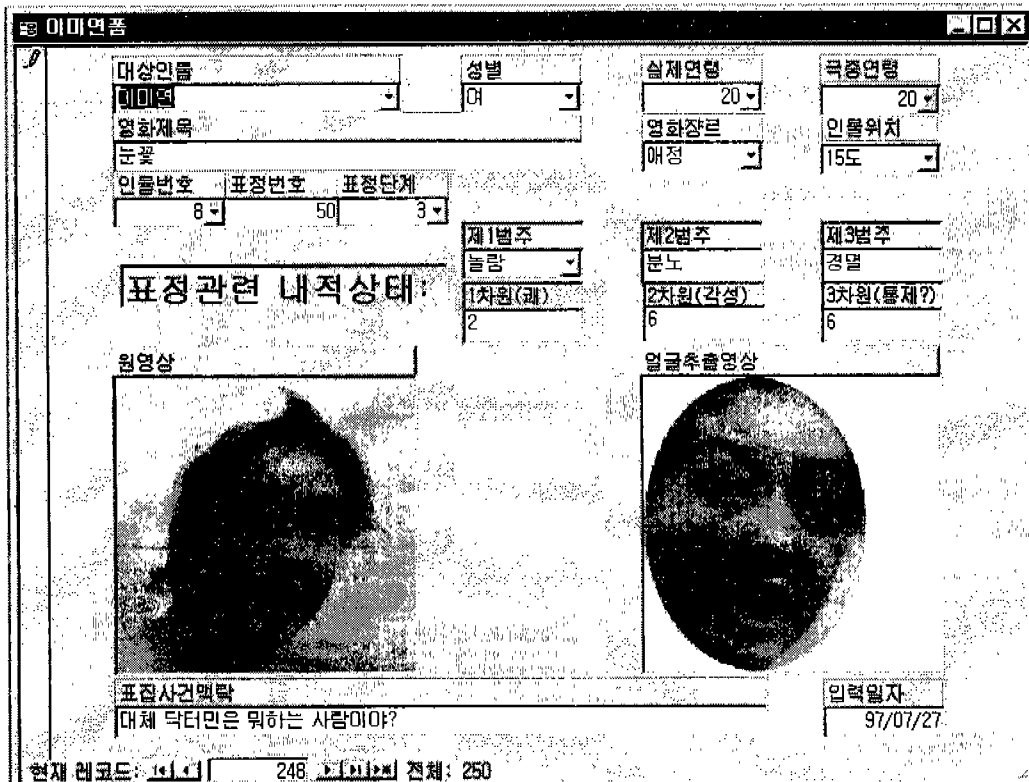


그림 2. 표정-내적상태 상관모형을 위한 얼굴 DB 구조

로그그램으로는 Adobe Premier 4.0과 Xing MPEG player 3.12가 사용되었다. 이들 장치를 이용하여 비디오화면을 컴퓨터에 avi 포맷 동화상으로 저장한 후, 등장인물의 표정자료를 frame 단위로 분석하여 적절한 자료를 화상데이터로 수집하였다. 수집된 자료는 MS ACCSES '97을 이용하여 DB화하였다.

결과

(1) DB 구조의 결정과 구축

MS ACCESS 프로그램을 사용하여 그림2와 같이 DB를 구축하였다.

이 DB구조에 나타나 있는 항목을 설명하면 다음과 같다.

가) 대상인물과 성별

국내 유통되고 있는 방화 비디오에 출연중인 영화배우 중 표정연기가 자연스럽고 출연작품이 충분히 많은 남자배우 다섯명과 여자배우 다섯명을 선정하였다. 남자배우로는 박중훈, 안성기, 이경영, 이영하, 한석규가 선정되었고, 여자배우로는 강수연, 심혜진, 이미연, 최명길, 최진실이 선정되었다.



그림 3. DB의 대상인물

나) 실제연령과 극중연령

실제연령은 출연 영화배우의 실제나이를 말하고, 극중연령은 극중에서 어느 연령대에 해당하는 인물로 연기하는가를 말한다. 이들 배우들은 대부분이 20-30대에 속하며 극중에서도 비교적 동일연령대의 배역을 연기하였다.

다) 영화제목과 장르

표집된 표정자료가 어느 영화에서 나온것인지 알 수 있도록 영화제목을 기록하였다. 영화장르는 코믹, 애정, 역사, 액션, 기타로 구분하였다.

라) 인물위치

인물의 표정이 어느 위치와 방향/각도에서 촬영되었는지 알 수 있도록 정면(0도)에서 90도 측면까지 기록되

게 하였다.

마) 인물번호/표정번호/표정단계

10명의 대상인물 중 몇번째 인물인지 인덱스 번호를 주었으며, 50가지 표정 중 몇번째 표정인지, 그리고 그 표정 중에서 몇번째 연속 단계의 표정인지를 기록하게 하였다.



그림 4. DB의 표정단계

바) 표정관련 내적상태

표집된 표정이 어떤 내적상태와 관련이 있는지 일반인들로 하여금 추론평정하게 하여 해당 내적상태에 대한 정보를 기록할 수 있게 하였다. 이러한 내적상태 정보는 크게 두가지 형태로 기록되는데, 범주정보는 기쁨, 슬픔 등과 같이 언어적으로 구분되는 불연속적 정보를 말하며, 차원정보는 쾌-불쾌의 연속적인 차원에서 해당표정이 어디쯤에 위치하는지 숫자로 나타내는 정보를 말한다.

① 제1범주/제2범주/제3범주

내적상태의 범주로서 제1범주는 6가지 기본정서를 주었고, 제2범주와 제3범주는 내적상태 기술체계 연구에서 나온 표정어휘 범주를 배당하여 위계적인 구조를 만들었다.

② 제1차원/제2차원/제3차원

내적상태의 차원으로서 제1차원은 쾌-불쾌를, 제2차원은 각성-수면 차원을 주었고, 제3차원은 3차원은 본 연구와 병행하여 진행되고 있는 내적상태 기술체계 연구에서 나온 차원을 배당하기로 하였다.

사) 원영상과 얼굴추출영상

원영상은 200x200 화소크기로 인물표정을 포착한 화상을 말하며, 얼굴추출화상은 원영상중에서 얼굴부분만 타원형으로 잘라낸 화상을 말한다. 이중 얼굴추출영상은 얼굴표정 인식모형에 입출력 자료로서 제공된다.

아) 표집사건 맥락

표집된 인물표정 자료가 영화의 어떤 상황맥락에서 나온것인지 알 수 있도록 해당표정 장면의 대사나 극중

상황에 대한 정보를 기록하였다.

(2) 자료수집의 결과분석

가) 대상 인물당 비디오편수

대상인물 10명분의 표정자료를 얻기 위해 총 50편의 비디오를 사용하였고, 1명의 대상인물에게서 50가지의 표정자료를 구하기 위해서는 평균 5편의 비디오가 필요하였다.

나) 영화 장르별 표집

수집된 50편의 비디오 중 애정물이 제일 많았고 코믹물, 액션물, 기타, 역사물 등의 순서였다. 일반적으로 애정물의 경우 주인공들의 감정표현이 주요소재이기 때문에 표정자료 수집을 위한 비디오 장르 역시 애정물이 많았던 것으로 볼 수 있다. 장르별로 표현되는 표정의 종류가 차이가 있었는데, 코믹물이나 액션물의 표정연기는 과장되거나 몇 가지로 제한되는 경향이 있었다.

다) 대상인물 위치

인물의 위치별로 분석한 결과를 보면, 정면보다 15도 정도의 측면위치가 가장 많았고, 그다음으로 정면, 30도, 45도, 60도, 70도 90도의 순으로 나타났다. 이는 영화의 특성상 완전한 정면위치보다는 동적인 정보와 함께 측면위치의 인물위치 설정이 많기 때문인 것으로 볼 수 있다.

라) 표정변화의 다양성

대상인물당 50가지 표정을 한 세트로 10명분의 표정자료가 수집되었다. 이들 표정자료는 한 인물 내에서 일어날 수 있는 표정변화의 다양성을 반영하는 것으로 볼 수 있다. 이러한 표정자료를 그림5와 같이 패-블래 및 각성-수면의 차원으로 분류해 보면 얼마나 다양한 내적상태의 범위가 포함되는지 알 수 있다.

맺음말

본연구에서 수집된 자료는 후속연구에서 신경망 등을 사용한 표정인식 모형을 개발하기 위한 기초 화상 자료로 사용될 예정이기 때문에 영상처리를 위한 자료의 화질상태가 좋아야한다. 따라서 이미 입력된 자료를 점검하여 화질상태와 다양한 표정 등을 계속 보완해나갈 예정이다.



그림 5. 정서의 원형모형에 대응시킨 표정의 변화

참고 문헌

Bassili, J.N. (1979). Emotion recognition : The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology*, 37(11), 2049-2058. Cambridge university press.

Ekman, P. (Ed.). (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). *The Facial Action Coding system*. Palo Alto, CA: Consulting Psychological Press.

Essa, I. A. & Pentland, A.P. (1995). Coding, analysis, interpretation, and recognition of facial expressions. M.I.T. Media laboratory perceptual computing section technical report No.325

Izard, C. E.(1982). Measuring emotions in human development. In Izard, C. E. (Ed.), *Measuring Emotions in Infants and Children*, Cambridge:Cambridge University Press.

Osgood, C. E.(1966). Dimensionality of the semantic space for communication via facial expressions, *Scandinavian Journal of Psychology*, 7, 1-30.

Russell, J.A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161-1178.

Russell, J.A. (1997). A dimensional-contextual perspective. In Russell, J.A. & Fernández-Dols, J.M. (Eds.) *The psychology of facial expression*. New York, NY: Cambridge university press.

Schlosberg, H. (1952). The description of facial expressions in terms of two dimensions. *Journal of Experimental Psychology*, 44, 229-237.