

# 표정 강도에 강건한 얼굴 표정 인식

김진옥\*

요약

본 연구는 표정 인식을 개선하기 위한, 강도가 다른 표정을 인식하는 새로운 표정 인식 방법을 제안한다. 사람마다 다르게 나타나는 표정과 표정마다 다른 강도는 표정 인식을 저하에 지대한 영향을 미친다. 하지만 얼굴 표정의 다양한 강도를 처리하는 방법은 많이 제시되지 않고 있다. 본 연구에서는 표정 템플릿과 표정 강도 분포모델을 이용하여 다양한 얼굴 표정 강도를 인식하는 방법을 제시한다. 표정 템플릿과 표정 강도 분포모델은 얼굴의 특정 부위에 표시한 관심 점과 얼굴 특정 부위간의 움직임이 다른 표정과 강도에 따라 어떻게 달라지는지 설명하여 표정 인식을 개선에 기여한다. 제안 방법은 정지 이미지뿐만 아니라 비디오시퀀스에서도 빠른 측정 과정을 통해 다양한 강도의 표정을 인식할 수 있는 장점이 있다. 실험 결과, 제안 연구가 특히 약한 강도의 표정에 대해 타 방법보다 높은 인식 결과를 보여 제안 방법이 다양한 강도의 표정 인식에 강건함을 알 수 있다.

키워드 : 얼굴 표정, 얼굴 표정 강도, 강도 분포 모델, 얼굴 표정 템플릿

## Robust Facial Expression-Recognition Against Various Expression Intensity

Kim, Jin Ok\*

ABSTRACT

This paper proposes an approach of a novel facial expression recognition to deal with different intensities to improve a performance of a facial expression recognition. Various expressions and intensities of each person make an affect to decrease the performance of the facial expression recognition. The effect of different intensities of facial expression has been seldom focused on. In this paper, a face expression template and an expression-intensity distribution model are introduced to recognize different facial expression intensities. These techniques, facial expression template and expression-intensity distribution model contribute to improve the performance of facial expression recognition by describing how the shift between multiple interest points in the vicinity of facial parts and facial parts varies for different facial expressions and its intensities. The proposed method has the distinct advantage that facial expression recognition with different intensities can be very easily performed with a simple calibration on video sequences as well as still images. Experimental results show a robustness that the method can recognize facial expression with weak intensities.

Keywords : Facial Expression, Facial Expression Intensity, Intensity Distribution Model, Facial Expression Template

### 1. 서론

사람의 행동을 인지하는 신호를 전달하는 표정은 복잡한 심리적 절차의 기능적 구조를 반영하여 상호간의 정보를 해석하는 단초를 제공한다. 얼굴 표정은 감정의 종류에 따라 다를 뿐 아니라 사람마다 표정으로 감정의 정도를 나타내는 방법이 다르다. 또한 동일한 얼굴 표정이라도 표정의 강도는 사람마다 다르다. 이런 문제점은 얼굴 표정 인식 시스템

의 신뢰도를 저하시키는 요인이 된다. 표정 강도가 다른면 동일한 표정에 대해서도 인식시스템이 다른 표정으로 인식하여 표정 인식 결과에 악영향을 미치게 된다.

대부분의 표정 인식 알고리즘의 대부분은 학습 대상 인물의 표정 강도를 다양하게 구성하는 대신 표정을 극도로 강조한 얼굴을 대상으로 인식시스템을 학습시키므로 이 인식 방법을 표정이 약한 인물에 적용할 경우 표정 인식 결과가 현저하게 약화된다. 물론 서로 다른 사람이나 동일한 인물의 표정을 인식할 때 사람마다 다른 표정 강도를 고려하여 표정 모델 학습 과정을 매번 새롭게 수행할 수도 있으나 비용상의 문제점이 발생하고 표정 인식 알고리즘의 실행시간이 많이 소요되는 단점을 보인다. 따라서 각각 다른 표정

\* 이 논문은 2008년도 대구한의대학교 기린연구비 지원에 의한 것임  
† 정 회 원 : 대구한의대학교 모바일콘텐츠학부 조교수  
논문접수 : 2009년 4월 24일  
수정일 : 1차 2009년 6월 13일, 2차 2009년 7월 15일  
심사완료 : 2009년 7월 17일

강도에 대해 일정범위를 동일하게 처리하는 표정 인식 방법이 필요하다.

본 연구는 얼굴 표정의 다양한 강도를 동일하게 인식하고 이를 통해 표정 인식률을 높이는 방법을 제시하고자 한다. 본 연구는 얼굴의 각 표정을 빠르고 효과적으로 인식하여 표정 인식률을 높이는데 그 목적이 있으며 특히 사람마다 다르게 나타나는 얼굴 표정 강도로 인한 표정 인식률 저하를 개선하는 데 연구의 의의를 둔다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구를 기술하며 3장에서는 본 논문이 제안하는 표정 템플릿과 표정 강도 분포 모델을 이용한 표정 인식 방법을 설명하고 4장에서는 제안 연구의 실험 결과를 제시한다. 5장에서는 연구의 결론을 정리한다.

## 2. 관련 연구

표정 강도 관련 연구들은 주로 표정 강도의 변화를 측정하고자 하였다. 표정을 결정하는 얼굴 움직임은 강도에 따라 다양하게 나타난다. FACS 코딩에서는 움직임 단위(AU, Action Unit)의 움직임 즉 표정 강도 변화를 설명하는데 여러 단계로 강도의 범위를 나누어 나타내는 새로운 연구를 진행했다[1]. 표정 움직임 강도 변화를 나타내기 위해 관련된 AU를 묶어 설명한 연구에서는 눈 분위에는 AU 41, AU42 또는 AU45를 묶어 눈을 살짝 내려 뜬 상태에서 감은 눈까지의 움직임 강도 변화를 나타냈다[1]. 얼굴의 관심 영역에서 표정의 강도 변화를 표현하거나 AU와 표정에서 강도를 정량화하는 방법을 제안한 연구도 있다[2-5]. 이 연구들은 얼굴 표정 강도를 서로 다른 얼굴 움직임간의 차이를 판별하기 위한 목적으로 적용했다. Bartlett[6]은 얼굴 표정의 강도가 체계적으로 변하는 것을 측정하는 알고리즘을 제안했으나 각 강도 변화 단계에서 개별적으로 그 결과 값을 나타내지는 못했다. Tian[7]은 가버 특징과 인공신경망을 이용하여 표정 강도 변화를 판별했으나 이 역시 표정 강도의 변화 정도를 다룬 것이다. 이 외, 얼굴 움직임 강도 측정에 관한 연구를 통해 얼굴 움직임의 여러 모습에서 나타내는 강도 변화를 자동으로 인식하려는 노력은 얼굴 표정 합성 분야에 많이 적용되고 있다.

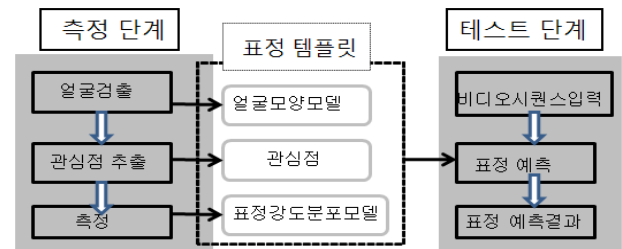
얼굴 표정의 강도가 사람에 따라 다양하게 나타나기 때문에 표정 인식시스템에서는 대표 표정의 기본 모델을 어떻게 설정하고 학습시키느냐에 따라 인식 결과에 많은 차이가 나타난다. 그래서 Wang[8]은 여러 인물의 표정 강도를 학습하는 방법 대신 스테레오 카메라와 3D 디지털타저를 이용하여 특정 인물의 얼굴 표정 모델을 생성하여 여러 강도의 표정을 동일하게 인식했으나 이를 비디오 시퀀스에는 적용하지 못했다. Dornakia[9]는 얼굴 표정 부위에 표시한 여러 개의 제어 점을 직접 조절하여 각 인물에 맞는 일반적인 얼굴 표정 모델을 생성함으로써 다양한 강도의 표정 인식을 시도했으나 이 방법은 구축비용이 너무 비싼 단점이 있다.

Lucy[10]은 얼굴 표정 모델을 설정할 때 얼굴 움직임을 이용한 구조화 기술을 적용하여 이미지에서 사용자의 얼굴 모델을 재구축하는 방법을 제안했으나 재구축된 얼굴 표정 모델이 오히려 표정 인식률을 저하하는 결과를 가져왔다.

본 연구는 얼굴 표정 모델 설정시 시간과 노력을 줄이면서 표정 템플릿을 빠르게 구축하여 비디오시퀀스 및 정지 영상에 나타난 다양한 강도의 표정을 동일한 표정으로 인식하도록 한다.

## 3. 제안 연구(3장의 각 단락을 재 구성하고 연구의 독창성을 설명하기 위해 내용을 전반적으로 편집함)

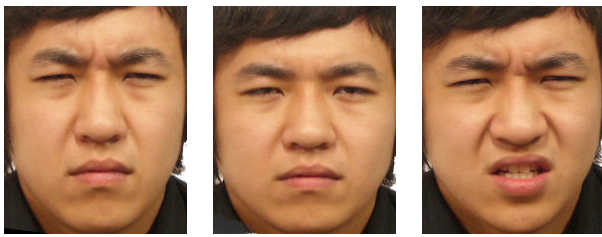
제안 연구의 구조는 그림 1과 같이 측정 단계와 테스트 단계로 나뉜다. 먼저 측정단계에서는 학습 대상 인물의 얼굴을 검출하고 얼굴 특징 부위에 관심 점을 설정하여 표정 템플릿을 구축한다. 이 표정 템플릿의 관심 점 위치 좌표의 강도는 표정 변화에 따라 얼굴 특징 부위가 움직이면서 같이 변한다. 그 표정 변화 과정을 표정 강도 분포 모델로 설정한다. 테스트 단계에서는 표정 템플릿의 표정 강도 분포 모델과 표정 예측 대상의 표정을 비교하여 표정을 예측한다. 표정강도 분포 모델은 서로 다른 표정은 물론 다양한 표정강도라도 허용할 수 있는 범위 이내라면 같은 표정으로 예측하는데 이용한다.



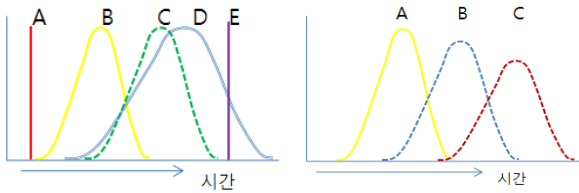
(그림 1) 연구 구조

### 3.1 표정 강도

얼굴 표정은 (그림 2)(a)과 같이 개인의 표정 특성, 감정의 정도에 따라 다양하게 나타난다. 이를 반영하여 FACS의 2002년 버전은 이전 FACS 버전과 비교하여 AU(Action Unit)를 서로 다른 표정 강도로 구분지어 표정 범위를 정한 내용을 포함하고 있다[1]. 이는 표정 강도가 한 가지 형태로만 나타나지 않는다는 것을 고려한 것으로 FACS는 얼굴 표정 강도를 (그림 2)(b)와 같이 A부터 E까지 다섯 단계로 나누어 A는 최소 외형 변화를, E는 최대 외형 변화를 나타낸다. 하지만 FACS의 얼굴 표정 강도를 5단계로 나누어 이를 시간의 변화에 따라 증가하는 강도의 분포도를 구한 결과 중간 단계의 얼굴 표정 강도 간에는 심하게 겹치는 결과를 보이며 특히 D 단계의 표정이 많은 범위를 차지하고 있음을 알 수 있다. 이는 FACS 강도 코드가 상당히 주관적이



(a) 혐오 표정의 서로 다른 표정 강도



(b) FACS의 강도 분포도

(c) Reilly의 강도 분포도

(그림 2) 얼굴 표정의 강도

어서 서로 다른 표정 강도를 정확하게 구분하지 못하기 때문이다. 이를 해결하기 위해 Reilly[11]는 AU를 저, 중, 고 강도에 해당하는 3개 단계로 단순하게 나눈 (그림 2)(c)의 방법으로 표정을 분류한다.

얼굴 표정 인식률은 이와 같은 얼굴 표정의 다양한 표정 강도 때문에 낮아질 수밖에 없다. 따라서 표정이 변하거나 표정 강도가 다를 경우에도 해당 표정을 제대로 인식할 수 있는 표정 인식 시스템을 구현해야 하며 이를 위해 본 연구에서는 기본 표정을 설정한 표정 템플릿을 구축하고 측정 대상 얼굴의 표정과 표정 템플릿의 표정의 유사한 정도를 판단하여 얼굴 표정 인식을 처리하는 방법을 제안한다.

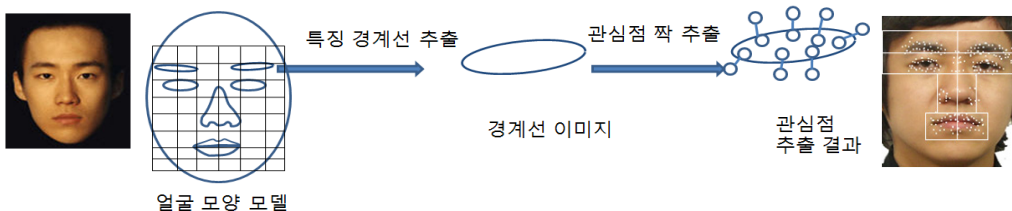
3.2 표정 템플릿과 표정 강도 분포 모델

제안 연구의 얼굴 표정 인식은 얼굴 표정 템플릿  $M = \{S, Z, T\}$ 을 이용한다. 표정 템플릿의  $S$ 는 얼굴 모양 모델,  $Z$ 는 관심 점 집합,  $T$ 는 표정강도 분포 모델을 각각 나타낸다. 표정 템플릿은 얼굴 특징 경계선을 추출하여 얼굴 모양 모델을 구축한 후 표정을 결정하는 얼굴 부위에 표시한 관심 점들이 표정이 변화함에 따라 얼마나 강도가 달라지는 지 측정하여 표정강도 분포 모델을 만든다. 표정강도 분포 모델은 표정을 결정하는 얼굴의 특징 부위에서 관심 점 위치의 강도가 서로 다른 표정에 대해 어떻게 달라지는 지 설명하고 표정 강도의 범위를 예측하여 다양한 강도의 표정을 단일 표정으로 이해하도록 한다.

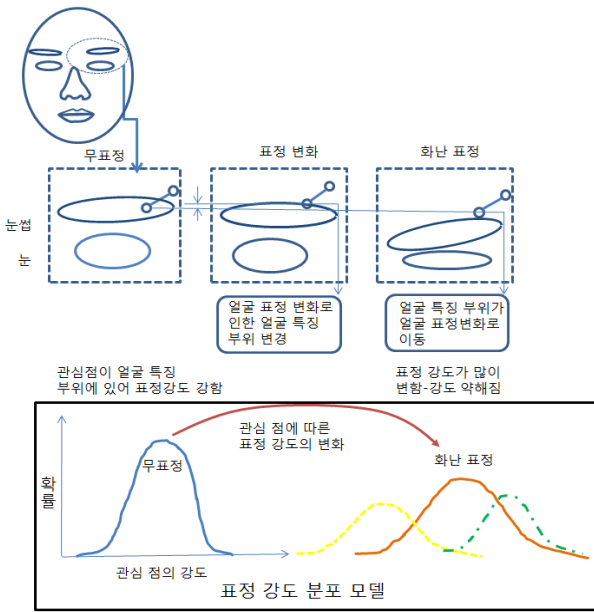
얼굴 모양 모델  $S$ 는 (그림 2)와 같이 한국인의 얼굴 평균 모델[12]을 기반으로 얼굴 표정을 만드는 특징 부위의 윤곽선을 검출하여 이 특징 부위에 관심 점 위치를 지정하여 만든다. 얼굴과 얼굴 특징 부위 윤곽선으로 구성된 얼굴 모양 모델  $S$ 는 각 표정에 대해 얼굴 정면 상태의 일반적 모델을 적용한다.

관심 점 집합은  $Z = \{z_i\}_{i=1}^{N_z}$ 로 얼굴 모양 모델  $S$ 에 표시된 여러 개의 고정된 점으로 구성한다. 변수  $z_i$ 는 학습 이미지에서  $i$ 번째 관심 점의 이미지 좌표를 나타내고  $N_z$ 는 관심 점의 개수이다. 관심 점은 (그림 2)와 같이 얼굴 모양 모델  $S$ 의 정면 얼굴 이미지에 설정하는데 얼굴 자세가 임의로 변할 수 있으므로 관심 점의 위치는 얼굴 모양 모델에 투영하여 자세에 따라 모양 모델을 회전하거나 위치를 옮긴 다음 이미지 단면에 2차원 위치로 투영하여 만든다. 관심 점을 정면 얼굴 이미지에 표시할 때 특징 부위의 윤곽선에서 서로 걸치도록 윤곽선 안쪽과 바깥쪽을 연결하는 쌍으로 표시한다. 관심 점은 눈썹, 눈, 코, 입 등 얼굴 표정을 결정하는 4개 얼굴 부위 영역에 설정한 개수만큼의 관심 점 쌍을 오름차순 형식으로 표시한다. 이때 관심 점의 각 쌍은 얼굴 특징 부위 경계에 걸쳐 있고 특징 부위 중심을 향하도록 한다. 표정 부위 간의 구분을 위해서 얼굴 이미지를 라플라시안 가우시안 필터[13]를 이용, 제로 크로싱 경계로 표시한다.

표정강도 분포 모델  $T$ 는 특징 부위에 위치한 관심 점의 강도가 표정에 따라 어떻게 달라지는지를 설명하여 특정 표정과 표정 강도를 나타내는 것으로 얼굴 모양 모델  $S$ 에 표시한 관심 점 집합의 좌표 값으로 만들어진다. 얼굴 표정 변화는 얼굴 특징 영역에서 강도 변화가 있음을 의미한다. 얼굴 표정에서 관심 점의 강도란 얼굴 표정을 결정하는 각 부위를 나타내는 관심 점의 주요 위치가 해당 영역에서 차지하는 픽셀 수이다. 관심 점 강도는 (그림 4)와 같이 표정에 따라 얼굴 부위에 변화가 일어나면서 같이 변한다. 예를 들어 무표정일 때는 관심 점이 얼굴 특징 부위와 겹쳐 있어서 관심 점 강도가 높으나 화난 표정을 지을 때는 표정 부위의 위치가 달라져 관심 점은 그대로 있는 상태에서 눈썹, 눈, 입 등 특징 부위의 위치가 변하므로 해당 관심 점의 강도는 약화된다. 즉 무표정일 때 눈썹 부위에 대한 표정 강도 분포 모델의 확률이 높고 화난 표정의 눈썹 표정 강도 분포 모델은 더 낮은 결과를 보인다. 화난 표정은 단일한 상태가 아닌 여러 상태로 나타나므로 표정 강도 분포 모델을 통해 어떤 표정이 어떤 강도로 나타나는 것을 확인할 수



(그림 3) 관심 점이 설정된 표정 템플릿



(그림 4) 표정에 따른 특징 부위 강도 변화와 표정 강도 분포 모델

있다.

표정은 특정 순간에 나타나는 것이 아니라 무표정에서 시작하여 표정이 나타나고 다시 무표정으로 되돌아오는 변환 과정으로 나타나며 각 표정은 이전 표정과 관계가 없이 독립적으로 발생하며 얼굴 특징 부위 역시 서로 다른 표정에 대해 각기 다르게 변화한다. 따라서 본 연구에서 얼굴 특징 부위의 관심 점 강도는 다른 관심 점으로부터 독립적이고 표정을 짓고 있는 동안 관심 점 강도의 변화는 정규 분포를 따른다고 가정한다.

표정 강도 분포 모델  $T$ 는 표정의 변화 정도가 불연속인 상태에서 각 표정의 분포를 나타내는 혼합분포도로 나타내어[11] 얼굴 표정 카테고리  $e \in \{1, \dots, N_e\}$ 의  $i$ 번째 관심 점의 평균  $\mu_i(e)$ 와 표준편차  $\sigma_i(e)$ 가 어떻게 변화하는지 설명하며 식 (1)과 같이 나타낼 수 있다.

$$\begin{aligned} T &= \{N_1, \dots, N_N\}, \\ N_i &= N(\mu_i(e), \sigma_i^2(e)), \\ \sigma_i(e) &= k\mu_i(e) \end{aligned} \quad (1)$$

여기서  $N(\mu, \sigma^2)$ 은 표정의 평균  $\mu$ 와 표준 편차  $\sigma$ 를 가진 정규 분포를 의미한다.  $e \in \{1, \dots, N_e\}$ 은 얼굴 표정을 나타낸다.  $N_e$ 는 대상 표정의 수를 나타내고  $e=0$ 는 무표정을,  $\mu_i(e)$ 와  $\sigma_i^2(e)$ 는 표정  $e$ 에서 관심 점  $i$ 의 강도 평균과 분산을 의미한다. 표준편차  $\sigma_i$ 는 비례 상수  $k$ 가 있는 평균  $\mu_i$ 에 비례한다고 가정한다. 이는 (그림 4)의 화난 표정의 표정 강도 분포 모델 예와 같이 표정의 분포도는 강한 표정, 중간 표정, 약한 표정에 따라 다르게 나타날 수 있으므로 표정 강도 분포 모델에 표준 편차 범위를 설정함으로써 그

범위 내의 표정은 모두 같은 표정으로 인식하여 표정 강도에 강건한 표정 인식을 위함이다.

### 3.3 유사도와 사후 확률밀도를 이용한 표정 예측

표정 템플릿 및 표정 강도 분포 모델을 구축한 후에는 얼굴 표정을 예측하는 과정을 수행한다. 표정 예측 대상이 되는 비디오 시퀀스에는 여러 인물의 여러 표정이 혼합되어 분포되어 있다. 이 경우 비디오 내 인물 이미지의 얼굴 표정이 다양하게 섞여 있으므로 베이즈안추론의 일반적 방식에 따라 비디오 시퀀스에 사후 확률 밀도 함수를 적용하여 얼굴 표정을 예측한다.  $t$  시간 동안 얼굴 이미지의 비디오 시퀀스  $(I_{1:t})$ 는 식 (2)와 같이 여러 표정이 분포되어 반복적으로 나타나며  $t-1$ 은 이전 표정 상태를 의미한다.

$$P(h_t, e_t | I_{1:t}) = \alpha P(I_t | h_t, e_t) \int P(h_t | h_{t-1}) \sum_{e_{t-1}} P(e_t | e_{t-1}) P(h_{t-1}, e_{t-1} | I_{1:t-1}) dh_{t-1} \quad (2)$$

식(2)의  $P(I_t | h_t, e_t)$ 는 얼굴 이미지에서 머리 자세  $h$ 와 얼굴 표정  $e$ 에 대한 표정 유사도로서 사후 확률밀도 함수는 표정 유사도를 이용하여 결정된다.  $\alpha$ 는  $\alpha = 1/P(I_t)$ 로 정규화 상수이다. 얼굴의 자세  $h_t$ 와 표정 상태  $e_t$ 는 처음 상태에서 상태가 바뀌는 1차 마코프 프로세스를 따른다고 가정하고 얼굴 자세 상태  $h_t$ 는 이미지 평면에서 이차원으로 움직이며 이전 표정과 현재 상태 표정은  $P(e_t | e_{t-1})$ 로 동일하다고 본다.

사후 확률밀도를 결정하는 표정 유사도는 측정 대상 얼굴 표정과 표정 강도 분포 모델  $T$ 를 비교하여 표정의 유사 여부를 판단한다. 얼굴 표정  $e_t$ 에서  $i$  관심 점의 강도 변화는 식 (3)과 같이 정규 분포로 나타난다.

$$\frac{1}{\sqrt{2\pi}\sigma_i(e_t)} \exp\left[-\frac{1}{2}\rho\left(\frac{I_{i,t} - \mu_i(e_t)}{\sigma_i(e_t)}\right)\right] \quad (3)$$

$\mu_i(e_t)$ 는 얼굴 표정  $e_t$ 에서 관심 점  $i$ 의 강도에 대한 평균이고  $\sigma_i(e_t)$ 는 표준 편차를 의미하며  $I_{i,t}$ 는  $t$  시간에  $i$  좌표 점에서  $I_t$  이미지의 강도를 나타낸다. 얼굴 표정 부위에서 각 관심 점들의 관계는 독립적이기 때문에 식 (2)의 유사도  $P(I_t | h_t, e_t)$ 는 식 (4)와 같이 예측한 표정의 관심 점 강도를 곱해서 관심 점 범위를 확대함으로써 표정 유사도를 계산하는데 유리하도록 한다.

$$P(I_t | h_t, e_t) = \prod_{i=1}^N P(I_{i,t} | h_t, e_{i,t}) \quad (4)$$

정규 분포 형태로 관심 점이 강조된 인식 대상 표정과 표정 강도 분포 모델  $T$ 의 차이를 기반으로 식 (5)과 같이 표정 유사도를 계산하여 이를 통해 해당 표정의 유사 정도를 판단한다. 식 (5)에서  $I_i$ 는 얼굴 이미지에서  $i$ 번째 관심 점



의 강도이고  $T_i(e)$ 는 표정  $e$ 에서의  $i$ 번째 관심 점에 대한 표정 강도 분포 모델이다.

$$P(I_i|h_i, e_i) = \frac{1}{\sqrt{2\pi\sigma_i(e_i)}} \exp\left[-\frac{1}{2}\rho(d(I_i(h), T_i(e)))\right] \quad (5)$$

식 (5)의 함수  $\rho(\cdot)$ 는 강건 함수인 Geman-McClure 함수[15],  $p(x, \sigma) = x^2/\sigma^2 + x^2$  (여기서  $\sigma$ 는 도함수 인수)를 이용하여 도함수를 계산한 것으로 이 함수는 모양 모델 오류 때문에 얼굴 위치가 크게 바뀌는 문제나 이미지 잡음과 같은 문제에 대해 강건하게 표정을 예측하도록 한다. 식(5)에서 인식 대상 표정과 표정 강도 분포 모델  $T$ 간의 거리  $d(\cdot, \cdot)$ 는 식 (6)으로 정의하여 관측한 표정의 관심 점과 표정 강도 분포 모델의 관심 점간의 거리가 일정 범위 이내에 있으면 이 두 표정은 유사한 표정이라고 판단한다.  $\Upsilon_i$ 는 표정 강도의 범위를 결정하는 적응 변수치이다.

$$d(I_i, T_i(e)) = \begin{cases} \frac{\Upsilon_i I_i - \mu_i(e)}{\sigma_i(e)}, & \text{범위 내일 경우} \\ d_0, & \text{범위 밖일 경우} \end{cases} \quad (6)$$

단, 비디오 시퀀스에서는 인물 자세 파라미터가 연속적이고 관심 점 분포도가 표정 부위 겹침 등의 요인 때문에 복잡하게 나타날 경우 표정 범위를 계산하기 어려우므로 사후 확률 밀도 함수를 파티클이라 불리는 표정 결정 가중치 샘플 집합의 근사치로 계산한다[14]. 파티클은 표정과 그에 대한 가중치로서  $l$ 번째 파티클의 상태와 가중치는  $[h_i^{(l)}, e_i^{(l)}]$

와  $w_i^{(l)}$ 로 나타내며  $\sum w_i^{(l)} = 1$ 을 만족한다. 가중치  $w_i^{(l)}$ 는 표정 유사도  $P(I_i|h_i^{(l)}, e_i^{(l)})$ 에 비례한다. 식 (7)과 같이 표정 유사도와 가중치를 기반으로 예측한 얼굴 표정  $\hat{e}_i$ 은 관측 대상 얼굴 표정  $P(e_i|I_{1:t})$ 에서 최대 확률로서 해당 표정으로 결정한다.

$$\hat{e}_i = \arg \max_{l \in e_i^{(l)} = e_i} \sum w_i^{(l)} e_i \quad (7)$$

### 4. 실험

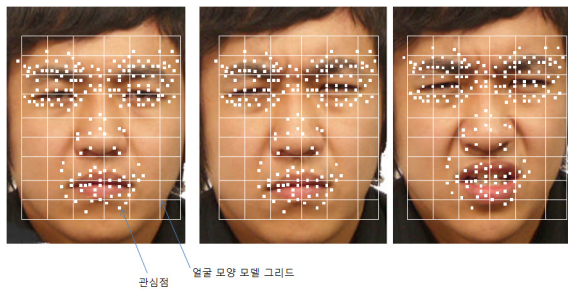
제안 연구는 개개인의 표정 강도가 보여주는 다양성으로 인한 차이를 극복하여 얼굴 인식률을 높이는데 초점을 두고 있다. 따라서 다양한 얼굴 표정 강도를 대상으로 표정 인식률을 비교하는 과정이 필요하다. 제안 방법을 평가하기 위해 다양한 표정의 비디오 시퀀스를 이용하여 실험했다. 실험 대상자의 머리는 2차원 평면상의 움직이는 범위 내에서 여러 가지 얼굴 표정을 짓는다.

(그림 5)와 같이 실험 대상 얼굴은 무표정, 혐오, 분노, 슬픔, 놀람, 기쁜 표정을 지을 때 강한 표정, 중간 표정, 약한 표정으로 3단계로 나누어 짓는다. 이에 따라  $N_e = 6$ 으로 설정한다. 9명의 실험 대상자들 연령대는 20대로 구성한다.  $512 \times 384$  크기의 비디오 시퀀스는 실험 대상자 당 15 fps로 찍는다. 처리시간은 3G RAM이 있는 3.73GHz 프로세스에서 80 ms/frame가 소요되었다.

실험 대상자는 머리는 고정된 채 모니터의 설명에 따라

표정	약한 표정			중간 표정			강한 표정		
놀람									
혐오									
기쁨									
슬픔									
화냄									

(그림 5) 실험 대상 표정



(그림 6) 표정 강도에 따른 관심 점 분포 결과

60프레임의 간격을 두고 60 프레임씩 한 번에 한 표정을 3번의 강도로 나눠 짓는다. 얼굴 표정 인식률은 지시가 표시된 후 각 표정당 최초 20프레임은 제거하고 계산한다. 표정 짓기에 대한 지시가 있는 후 그에 따라 얼굴 표정이 나타나는 시간차가 있기 때문이다. 프레임 기반의 얼굴 표정 인식을 위해 각 이미지의 LBP (Local Binary Pattern) 특징 추출에 59-bin LBP 오퍼레이터를 선택하고 얼굴 영역은 48개 영역(6×8)으로 나눈다[17].

(그림 6)은 얼굴 이미지에 표시된 점과 격자를 이용, 관심점 위치와 얼굴 모양 모델을 나타낸 후 표정 강도에 대한 관심 점의 분포 결과를 이용하여 표정을 예측하는 과정을 보여 준다. 이미지 프레임들은 약한 표정, 중간 표정, 강한 표정의 비디오 시퀀스에서 뽑은 것으로 단일 프레임으로 예측한 표정 결과는 확률로 나타난다.

<표 1>은 실험에서 얼굴 표정을 예측한 결과를 보여준다. 비디오 시퀀스에서 동일한 표정에 대한 저, 중, 고강도의 표정 프레임들을 추출한 얼굴 표정을 대상으로 실제 표정에 대해 본 연구를 통해 인식한 표정 결과를 확률로 나타낸 결과로써 프레임의 평균 얼굴 표정 인식률은 80.4%이며 슬픔을 슬픔으로 인식한 것은 70.3%, 슬픔을 혐오로 오인식한 것은 14.9%이다.

제안 방법을 <표 2>와 같이 비디오 시퀀스에 적용하여 표정을 분류한 결과 프레임을 이용하여 얼굴 표정을 인식한 결과보다 높은 86.5%의 표정 인식률을 나타냈다. 이는 실험 대상의 동일한 표정에 대해 표정 강도를 달리한 비디오 시퀀스를 대상으로 표정 인식을 시도한 것으로 비디오 시퀀스가 프레임보다는 동일한 표정에 대한 다양한 데이터를 다수 포함하고 있으므로 확률적으로 비디오 시퀀스의 표정 인식 신뢰도가 더 나은 결과를 보임을 알 수 있다.

<표 1> 프레임 기반 얼굴 표정 인식률

(단위:%)

표정	무표정	분노	슬픔	놀람	기쁨	혐오
무표정	91.0	0.1	3.2	0.4	0.4	0.3
분노	2.3	79.3	12.3	1.8	6.2	14.2
슬픔	11.2	10.6	70.3	8.3	0.7	14.9
놀람	0.3	0.1	0.7	90.0	0.5	0.8
기쁨	0.3	0.2	0.2	0.1	81.0	0.5
혐오	0.3	8.5	3.4	3.6	9.2	71.0

<표 2> 시퀀스 기반 얼굴 표정 인식률

(단위:%)

표정	무표정	분노	슬픔	놀람	기쁨	혐오
무표정	97.0	0.0	2.8	0.1	0.1	0.2
분노	1.9	84.1	12.3	1.0	4.7	13.5
슬픔	9.2	9.6	75.7	7.2	0.3	14.3
놀람	0.3	0.1	0.3	98.0	0.3	0.5
기쁨	0.1	0.0	0.0	0.0	88.3	0.1
혐오	0.1	7.1	3.0	2.3	7.5	76.3

<표 1><표 2>의 표정 인식 결과 놀람과 기쁨 표정은 표정을 결정하는 얼굴 부위의 변화가 많은 반면 슬픔, 혐오 표정을 나타내는 얼굴 표정 부위는 변화가 적어 인식률이 떨어짐을 확인할 수 있다.

표정에 대한 다양한 데이터를 포함하고 있는 비디오 시퀀스가 확률적으로 표정 인식 결과가 더 우수하지만 본 연구가 제시하는 표정 강도에 강건한 표정 인식 방법을 구체적으로 설명하기 위해 비디오 시퀀스의 강도별로 다른 표정을 프레임으로 포착하여 이를 대상으로 제안 방법과 SVM (Support Vector Machine)분류 방법[16]을 비교함으로써 표정 강도의 다양함에 대응하는 표정 인식 결과를 설명한다. 강도가 다른 얼굴 표정 프레임에 제안 방법과 SVM을 적용한 결과 다양한 표정 강도의 프레임을 섞어서 인식한 <표 1>에 비해 <표 3>의 인식률을 강한 표정을 가진 얼굴만을 대상으로 한 결과 <표 3>의 표정 인식률이 더 높다. 또한 제안 방법과 SVM 방법이 강한 표정에 대해서는 유사한 표정 인식률을 보였다. 그러나 <표 4><표 5>의 중간 표정 강도와 약한 표정 강도에 대해서는 제안 방법이 80.1%, 62.8%로 SVM의 75.3%, 47.3%에 비해 더 나은 인식률을 보임을 알 수 있다. 표정 강도가 강한 인물의 경우 연구 방법에 따른 표정 인식률 결과가 그다지 영향을 받지 않지만 약한 강도의 표정의 경우 제안 방법을 적용할 경우 SVM보다 높은 인식률을 보여 제안 방법이 표정 강도가 다양한 여러 사람의 표정 인식에 유리함을 알 수 있다.

<표 3> 강한 표정 인식률

(단위:%)

표정	평균	분노	슬픔	놀람	기쁨	혐오
제안방법	91.3	84.2	81.7	98.8	99.0	93.0
SVM	90.7	83.1	83.3	98.0	99.0	90.5

<표 4> 중간 표정 인식률

(단위:%)

표정	평균	분노	슬픔	놀람	기쁨	혐오
제안방법	80.1	70.7	62.0	94.6	94.5	78.5
SVM	75.3	66.1	54.5	90.0	88.6	77.6

<표 5> 약한 표정 인식률

(단위:%)

표정	평균	분노	슬픔	놀람	기쁨	혐오
제안방법	62.8	54.2	47.3	78.2	74.1	60.5
SVM	47.3	43.8	43.6	47.0	53.4	49.0

## 5. 결 론

본 연구에서는 강도가 다른 얼굴 표정을 예측하여 인식하는 새로운 방법을 제안한다. 표정 인식률을 결정하는 데에는 빛, 타 물체와의 겹침 등 외부적 원인 외에 사람마다 다른 표정의 강도가 인식 결과에 많은 영향을 미친다. 사람들의 표정은 표정 특성만큼 그 변화량이 다양함에도 현재까지의 많은 표정 인식 연구는 의도적으로 강하게 짓고 있는 표정을 대상으로 인식률을 제시하였다. 본 연구는 약한 표정, 중간 표정, 강한 표정 등 표정의 강도를 나누어 특히 약한 표정에 대해서도 강건한 인식 결과를 보일 수 있는 방법을 제시함으로써 인식률을 개선하고 있다.

이를 위해 본 연구에서는 얼굴 모양 모델을 설정하고 표정을 결정하는 얼굴 특징부위에 관심 점을 표시하여 표정 변화에 따른 관심 점과 얼굴 특징 부위간의 관계를 표정강도 분포모델로 나타낸 얼굴 템플릿을 구축함으로써 실제 다양한 강도의 표정을 보이는 인식 대상의 표정을 예측한다.

제안 연구는 무표정, 분노, 놀람, 슬픔, 기쁨, 혐오 등 6개의 얼굴 표정 범주의 3가지 강도표정에 대해 정면에서 본 표정을 개별 프레임에 대해서는 80.4%, 비디오 시퀀스에 대해서는 86.5%의 확률로 인식했다. 다양한 표정 강도에 무관하게 표정을 인식하는 제안 연구 방법은 특히 중간 표정과 약한 표정을 짓고 있는 인물들의 표정 인식에 강건한 인식 결과를 보였으며 강한 표정의 인식률은 타 연구와 유사한 인식결과를 보였지만 사람들의 표정이 일반적으로 아주 강하게 나타나지 않는다는 점을 감안할 때 제안 방법이 다양한 표정 강도에 더 강건함을 알 수 있다.

향후 연구에서는 기존 연구[18, 19]를 바탕으로 실험 대상자의 연령대, 성별 분포도를 다양화시켜 강도 변화가 큰 얼굴 표정에 대해서도 신뢰도 높은 표정 인식결과를 보이도록 하며 얼굴 표정을 결정하는 부위를 표시하는 관심 점과 표정 부위를 효과적으로 추출함으로써 더 섬세하고 정확하게 인식률을 개선하는 방법에 대해 진행할 것이다. 또한 머리의 방향에 따라 얼굴의 각도가 달라질 때 정면이 아닌 상태에서 표정을 동시에 인식하는 방법과 얼굴 표정이 변할 때 어떤 과정을 통해 표정 전이가 진행되는 지에 대한 과정을 분석하여 이 내용을 얼굴 표정 합성연구에 적용하도록 한다.

## 참 고 문 헌

[1] Paul Ekman, Wallace Friesen, Joseph Hager, "Facial Action Coding System Manual," 2002.  
 [2] M. A. Sayette, J. F. Cohn, J. M. Wertz, M. Perrott, D. Parrott, "Psychometric Evaluation of the Facial Action Coding System for Assessing Spontaneous Expression," *Journal of Nonverbal Behavior*, vol. 25. no. 3, pp. 167-185, Springer, 2001.  
 [3] Mohammed Yeasin, Baptiste Bulot, Rajeev Sharma,

"Recognition of Facial Expressions and Measurement of Levels of Interest from Video," *IEEE Transactions on Multimedia*, Vol.8, No.3, 2006.  
 [4] S. Kimura, M. Yachida, "Facial Expression Recognition and its Degree Estimation," *International Conference on Computer Vision and Pattern Recognition*, pp.295-300, 1997.  
 [5] J. J. Lien, T. Kanade, J. Cohn, C. Li, "Subtly different facial expression recognition and expression intensity estimation," *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pp.853-859, 1998.  
 [6] M. Barelett, G. Littlewort, M. Franc, C. Ainscsek, I. Fasel, J. Movellan, "Automatic Recognition of Facial Actions in Spontaneous Expressions," *Journal of Multimedia*, Vol.1, No.6, pp.22-35, 2006.  
 [7] Y. L. Tian, T. Kanade, J. Cohn, "Eye-state Aciton Unit Detection by Gabor Wavelets," *Proceedings of International Conference on Multi-Model Interfaces*, pp.143-150, 2000.  
 [8] J. Wang, L. Yin, X. Wei, Y. Sun, "3D Facial Expression Recognition based on Primitive Surface Feature Distribution," *IEEE Computer Vision and Pattern Recognition*, pp.1399-1406, 2006.  
 [9] F. Drnaika, F. Davoine, "Simultaneous Facial Action Tracking and Expression Recognition in the presence of head motion," *Int. Journal of Computer Vision*, Vol.76, No.3, pp.257-281, 2008.  
 [10] S. Lucey, I. Mathews, C. Hu, Z. Ambadar, F. Torre, J. Cohn, "AAM Derived Face Representations for Robust Facial Action Recognition," *IEEE Int. Conf. Automatic Face and Gesture Recognition*, pp.155-160, 2006.  
 [11] Jane Reilly, John Ghent, John Mcdonald, "Investigating the Dynamics of Facial Expressions," *Lecture Notes in Computer Science*, Vol.4292, pp.334-343, 2006.  
 [12] <http://koreanface.culturecontent.com/>  
 [13] J. O. Kim, J. H. Chung, "On a Face Detection with an Adaptive Template Matching and an Efficient Cascaded Object Detection," *Lecture Notes in Computer Science*, Vol.3645, pp.414-422, 2005.  
 [14] M. Isard, A. Blake, "Condensation-conditional density propagation for visual tracking", *International Journal of Computer Vision*, Vol.29, No.1, pp.5-28, 1998  
 [15] R. Gonzalez, R. Woods, "Digital Image Processing", Addison-Wesley Publishing Company, pp.142, 1992.  
 [16] M. S. Bartlett, "Machine Learning Methods for Fully Automatic Recognition of Facial Expressions and Actions," *Proc. IEEE Conf. of SMC*, pp.592-597, 2004.  
 [17] M. Black, A. Rangarajan, "On the Unification of Line Processes, Outlier Rejection, and Robust Statistics with Applications in Early Vision," *International Journal of Computer Vision*, Vol.19, No.1, pp.57-92, 1996.

- [18] 김진옥, “표정 정규화를 통한 얼굴 인식율 개선”, 정보처리학회 논문지B, 제15-B권, pp.477-486, 2008.
- [19] 김진옥, “상황에 민감한 베이지안 분류기를 이용한 얼굴 표정 기반의 감정 인식”, 정보처리학회논문지B, 제13-B권, pp.653-662, 2006.



### 김진옥

e-mail : bit@dhu.ac.kr

1989년 성균관대학교 졸업(학사)

1998년 성균관대학교 졸업(공학석사)

2002년 성균관대학교 대학원 전기전자 및  
컴퓨터공학과 졸업(공학박사)

1992~1994년 현대전자산업(주)정보통신  
사업부 재직

1994~1999년 현대정보기술(주)인터넷사업부 재직

2004~현 재 대구한의대학교 모바일콘텐츠학부 조교수

관심분야: 멀티미디어공학, 패턴인식, 영상처리, 유비쿼터스  
컴퓨팅, 웹서비스 등