

A Study on Appearance-Based Facial Expression Recognition Using Active Shape Model

Dong-Ju Kim[†] · Jeong-Hoon Shin^{**}

ABSTRACT

This paper introduces an appearance-based facial expression recognition method using ASM landmarks which is used to acquire a detailed face region. In particular, EHMM-based algorithm and SVM classifier with histogram feature are employed to appearance-based facial expression recognition, and performance evaluation of proposed method was performed with CK and JAFFE facial expression database. In addition, performance comparison was achieved through comparison with distance-based face normalization method and a geometric feature-based facial expression approach which employed geometrical features of ASM landmarks and SVM algorithm. As a result, the proposed method using ASM-based face normalization showed performance improvements of 6.39% and 7.98% compared to previous distance-based face normalization method for CK database and JAFFE database, respectively. Also, the proposed method showed higher performance compared to geometric feature-based facial expression approach, and we confirmed an effectiveness of proposed method.

Keywords : Active Shape Model, Face Normalization, Facial Expression Recognition

Active Shape Model을 이용한 외형기반 얼굴표정인식에 관한 연구

김 동 주[†] · 신 정 훈^{**}

요 약

본 논문에서는 ASM(Active Shape Model) 특징점(Landmark)을 이용하여 정밀한 얼굴영역을 획득하고, 외형기반 접근법으로 표정을 인식하는 방법에 대하여 제안한다. 외형기반 표정인식은 EHMM(Embedded Hidden Markov Model) 및 이진패턴 히스토그램 특징과 SVM(Support Vector Machine)을 사용하는 알고리즘으로 구성되며, 제안 방법의 성능평가는 공인 CK 데이터베이스와 JAFFE 데이터베이스를 이용하여 수행되었다. 더불어, 성능비교는 기존의 눈 거리 기반의 얼굴 정규화 방법과 비교를 통하여 수행되었고, 또한 ASM 전체 특징점 및 변형된 특징을 SVM으로 인식하는 기하학적 표정인식 방법론과 성능비교를 수행하였다. 실험 결과, 제안 방법은 거리기반 얼굴정규화 영상을 사용한 방법보다 CK 데이터베이스 및 JAFFE 데이터베이스 경우, 최대 6.39%와 7.98%의 성능향상을 보였다. 또한, 제안 방법은 기하학적 특징점을 사용한 방법보다 높은 인식 성능을 보였으며, 이로부터 제안하는 표정인식 방법의 효용성을 확인하였다.

키워드 : ASM, 얼굴정규화, 얼굴표정인식

1. 서 론

일반적으로 얼굴 표정인식은 얼굴검출, 특징추출, 그리고 표정분류의 단계로 수행된다. 표정인식의 첫 단계인 얼굴영역 검출 단계는 얼굴의 크기가 개인별, 표정별로 달라지므

로, 정밀한 얼굴영역 검출 과정은 신뢰성 높은 표정인식 시스템 설계를 위하여 중요한 단계이다. 특징추출 단계는 표정인식에 중요한 역할을 차지하며, 크게 외형기반 특징추출 방법과 기하학적인 특징추출 방법으로 구분할 수 있다. 기하학적인 특징추출 방법은 얼굴의 입, 눈, 볼 등의 위치 정보를 특징으로 이용하는 것으로 ASM[1]과 AAM(Active Appearance Model)[2]과 같은 대표적인 얼굴 특징점 추출 방식이 있다. 이러한 기하학적인 특징을 사용하는 표정인식 방법은 찾아진 특징점의 위치에 따라 인식 결과가 결정되므로, 정확하고 신뢰성 높은 특징점 추출이 요구된다. 하지만 특징점의 정확한 위치를 찾는 것은 어렵고 시간 또한 많이

※ 본 결과물은 2014년도 대구가톨릭대학교 학술연구비 지원에 의한 것임.
† 정 회 원 : 동서대학교 컴퓨터공학부 조교수
** 정 회 원 : 대구가톨릭대학교 IT공학부 부교수
Manuscript Received : April 29, 2015
First Revision : July 22, 2015
Accepted : September 9, 2015
* Corresponding Author : Jeong-Hoon Shin(only4you@cu.ac.kr)

소요되는 단점을 갖는다. 반면, 외형기반 특징추출 방법은 표정의 특징 변화를 얼굴 영상에 다양한 필터를 적용하여 직접 추출하는 것으로, 대표적인 방법으로는 가버 웨이블릿 (Gabor Wavelet)[3, 4], LBP(Local Binary Pattern)[5] 및 LDP(Local Directional Pattern)[6] 등을 이용한 방식이 있다. 일반적으로 이와 같은 외형기반 특징을 사용하는 표정 인식 방법은 기하학적 특징을 사용하는 표정인식 방법보다 인식 성능측면에서 우수한 결과를 보인다고 알려져 있다. 그러나, 영상의 텍스처 정보를 사용하므로 조명, 포즈변화 및 가려짐 등의 변화에 취약한 단점을 갖는다. 이에 본 논문에서는 조명 및 포즈 변화가 거의 없는 환경에서 수집된 얼굴 표정 데이터베이스를 이용하여 표정인식 시스템을 구성하였다. 마지막으로, 표정을 분류하기 위한 인식 알고리즘은 추출된 표정 특징의 변화된 패턴을 이용하여 표정을 분류하며, 대표적인 알고리즘으로는 SVM과 EHMM[7] 방식이 있다.

본 논문에서는 ASM 기반의 기하학적 특징점 추출을 통하여 정밀한 얼굴영역을 획득하고, 획득한 얼굴영상을 외형 특징기반의 접근법을 통하여 표정을 인식하는 방법을 제안한다. 본 논문과 유사하게 6개 표정에 대한 각각의 ASM 모델을 사용하여 얼굴 정규화를 수행하고, 16×16 크기의 윈도우에 대한 LBP 특징을 EHMM 알고리즘으로 표정 인식을 수행한 기존 연구가 있다[8]. 그러나, 본 논문에서는 한 개의 ASM 모델로 7개 표정에 대한 얼굴 특징점을 추출하여 얼굴 정규화를 수행하는 과정을 구체적으로 제시하였으며, 특징 추출을 위한 다양한 윈도우 크기에서의 실험을 통하여 2D-DCT 특징이 LBP 특징보다 우수하다는 점을 입증했다는 점에서 기존 논문[8]과 차별성을 갖는다.

Fig. 1은 입력영상으로부터 얼굴검출과 ASM 특징점을 기반으로 얼굴 정규화를 수행하고, 외형 특징기반의 표정인식 알고리즘으로 표정분류를 수행하는 제안 방법의 블록 다이어그램을 나타낸다. 일반적으로, 기존의 외형기반 표정인식 방법은 대부분 얼굴의 두 눈을 검출하고, 눈 중심점 사이의 거리를 기반으로 얼굴영역을 획득한 후, 표정인식을 수행한다[4, 6, 7]. 그러나 사람의 얼굴영역은 개인별, 그리고 표정별로 상이하게 달라지므로, 표정인식 성능 개선을 위해서는 보다 정밀한 얼굴영역 획득이 요구된다. 이에 본 논문에서는 정밀한 얼굴영역을 ASM 특징점을 이용하여 획득하고, 외형기반의 분류 알고리즘으로 표정을 인식하는 방법을 제안한다. 제안하는 방법은 ASM 특징점 자체를 표정인식에 이용하는 방법[9]과 다른 접근법으로, 특히 얼굴 내부의 ASM 특징점 추출이 정확하게 이루어지지 않아도 얼굴 외곽의 특징점은 상대적으로 안정하게 검출되므로, 부정확한 얼굴영역 검출에 의한 인식 성능 저하를 극복할 수 있는 장점을 갖는다. 본 논문에서의 표정인식 방법은 ASM 얼굴 정규화 영상을 기반으로 이진패턴 히스토그램 특징과 SVM 분류기를 이용하는 방법[5, 6]과 2D-DCT(Two-Dimensional Discrete Cosine Transform) 특징과 EHMM 알고리즘을 이용하는 방법[7] 등의 외형기반 알고리즘을 통하여 표정분류

를 수행한다. 더불어, 제안방법의 성능평가는 얼굴 정규화에 있어 기존의 두 눈의 거리에 기반한 얼굴영역 획득 방법과 비교하였으며, 인식 방법론에 있어서는 얼굴 특징점을 SVM으로 인식하는 기하학적 특징기반 표정인식 방법론[9]과 LBP 특징과 EHMM 알고리즘으로 표정인식을 수행하는 방법론[8]과 성능 비교를 수행하였다.

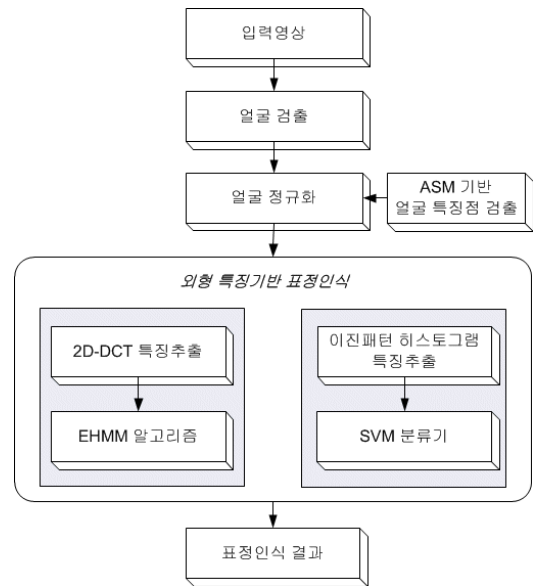


Fig. 1. Block Diagram of Proposed Facial Expression Recognition System

2. 얼굴 영역 획득

2.1 거리 기반 얼굴 정규화

대부분의 기존 표정인식 시스템은 얼굴영역을 검출하기 위하여 하 라이크(Haar-like) 특징기반의 아다부스트(AdaBoost) 알고리즘[10]으로 얼굴 영역과 눈 영역을 검출하고, Fig. 2와 같이 두 눈의 중심점 위치 및 거리 정보를 이용하여 얼굴 기우림 보상과 정규화 과정을 통하여 얼굴영역을 획득한다 [4, 6]. 일례로, 두 눈 사이의 거리가 D일 때, 눈 중심점을 기준으로 위로는 0.7D, 아래 방향으로 2.0D, 좌우 방향으로 0.5D의 비율로 얼굴영역은 획득된다.

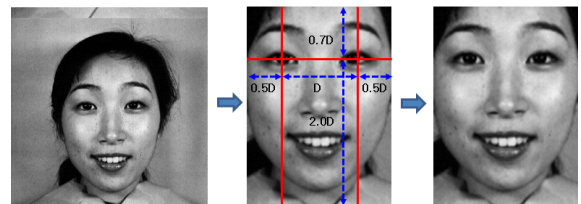


Fig. 2. Distance-Based Face Normalization

그러나 개인별 또는 다양한 표정마다 거리기반 방법에 의한 얼굴영역 획득은 다소 일괄적이지 못한 얼굴영상이 얻어

진다. Fig. 3은 두 눈의 거리 정보를 이용하여 획득한 얼굴 영상의 일례로써, Fig. 3의 (a)는 개인별에 따른 정규화 영상을, (b)는 한 사람에 대한 표정별 정규화 영상을 보여준다. Fig. 3으로부터 거리기반에 의하여 획득한 얼굴영상이 사람마다 그리고 표정마다 관심 얼굴영역이 다소 상이하다는 것을 알 수 있다.

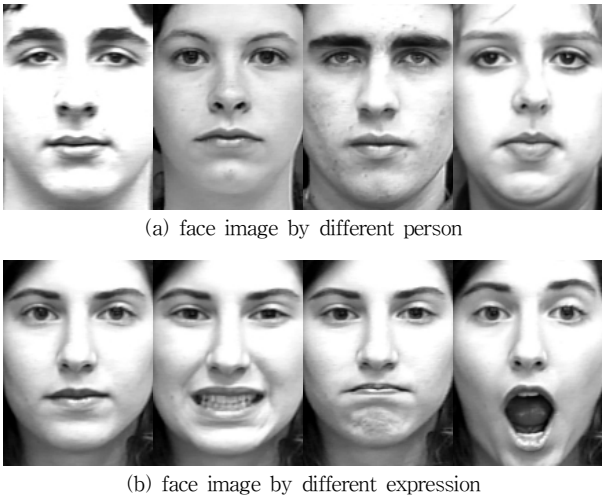


Fig. 3. Sample Images Via Distance-Based Face Normalization

2.2 ASM 기반 얼굴 정규화

이에 본 논문에서는 기존 방법보다 정밀하고 일괄성 높은 관심 얼굴영역을 획득하여 표정인식 성능을 개선하고자, ASM 특징점을 이용한 얼굴 정규화 방법을 제안한다. 일반적으로, ASM은 통계적인 모델을 이용하여 특징점을 추출하는 방법으로 특징점 검출 속도가 빠르고, 형태 모델의 여러 가지 변형을 잘 표현할 수 있으며, 조명변화에 덜 민감한 방법으로 알려져 있다[1]. ASM에서 형태 모델을 만드는 학습 과정은 학습 데이터로부터 라벨링된 특징점들을 이용하여 평균 형태를 구한다. 이러한 형태 모델은 Equation (1)으로 표현된다.

$$x \approx \bar{x} + Pb \quad (1)$$

여기서 \bar{x} 는 평균 형태 벡터이고, P 는 형태의 공분산에 대한 고유벡터 열로 구성된 행렬이며, b 는 형태 파라미터를 원소로 하여 구성된 벡터이다. 학습된 형태 모델을 이용하여 특징점을 추출하는 과정은 초기의 모양을 형태 모델과 입력 영상을 이용하여 조정한다. 다음으로 b 를 구하여 제한 조건 범위 밖에 있으면 재조정 한 후, b 의 모든 값들이 제한조건 범위를 만족할 때까지 반복하여 형태 파라미터 b 를 계산하여 입력영상에 대한 형태를 구한다.

본 논문에서는 MUCT(Milborrow University of Cape Town) 얼굴 데이터베이스[11]의 76개 특징점에 대한 좌표와 얼굴영상을 이용하여 ASM 모델을 생성하여, 이를 얼굴 정규화 과정에 이용하였다. Fig. 4는 ASM 모델 생성에 이용된 76개의 ASM 특징점 정보를 나타내며, Fig. 5는 ASM을 이용하

여 검출한 얼굴 특징점을 보여준다. Equation (2)와 같이 얼굴의 ASM 특징점을 α_i 라 할 때, 얼굴영역은 Equation (3)의 좌표로 결정되는 영역을 정규화한다. 여기서 $\alpha_i(x)$ 와 $\alpha_i(y)$ 는 각각 i 번째 특징점의 x 와 y 의 좌표를 의미한다.

$$\alpha_i = (x_i, y_i), \quad i = 1, 2, \dots, 76 \quad (2)$$

$$\begin{aligned} &[\text{left}, \text{top}, \text{right}, \text{bottom}] = \\ &[\alpha_{22}(x), \alpha_{24}(y) - (\alpha_{26}(y) - \alpha_{24}(y)), \\ &\alpha_{16}(x), (\alpha_6(y) + \alpha_{10}(y))/2] \end{aligned} \quad (3)$$

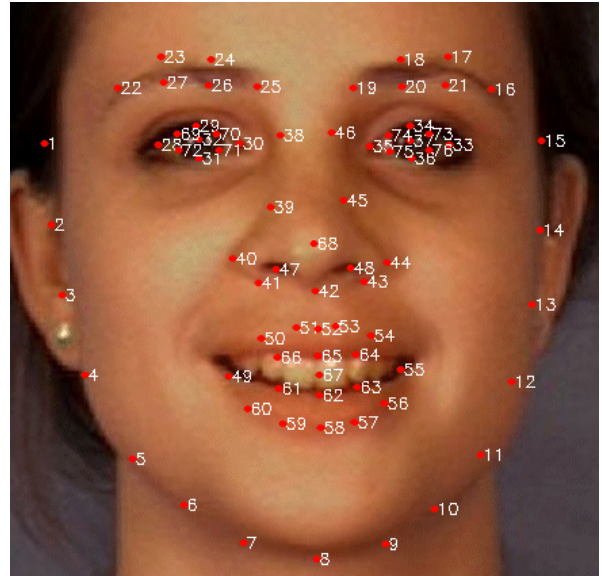


Fig. 4. ASM landmark information

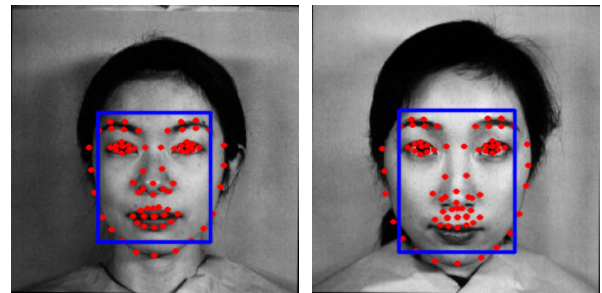


Fig. 5. ASM-based face normalization

특히, ASM을 이용한 얼굴정규화 방법은 얼굴 내부의 ASM 특징점 추출이 정확하게 이루어지지 않아도 얼굴 외곽의 특징점은 상대적으로 안정하게 검출되므로, 부정확한 내부 특징점 검출이 표정인식 성능 저하에 영향을 주지 않는 장점을 갖는다. Fig. 6은 ASM 기반 얼굴 정규화 방법을 이용한 개인별, 표정별 예시 영상을 보여준다. Fig. 6의 ASM 기반 정규화 방법은 거리기반의 정규화 방법과 비교하여 개인별, 표정별 영상에서 보다 일관성 있는 정밀한 얼굴영역이 획득되는 것을 확인할 수 있다.

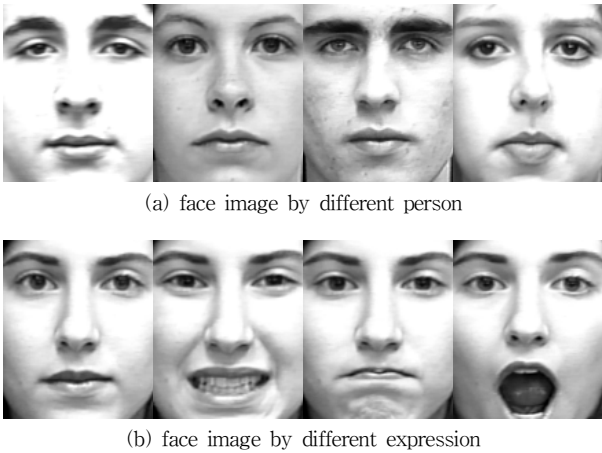


Fig. 6. Sample Images Via ASM-Based Face Normalization

3. 표정 인식

본 논문에서는 ASM 특징점을 이용하여 정밀한 얼굴영역을 추출하고, 대표적인 외형기반 표정분류 방법인 EHMM 알고리즘 및 이진패턴 특징기반 접근법을 사용하여 표정인식을 수행한다. 더불어, 제안 방법의 성능비교를 위하여 ASM의 특징점 자체를 이용하는 기존의 기하학적 특징기반 표정인식 접근법에 대하여 설명한다.

3.1 ASM 기반 얼굴 정규화

본 논문에서는 얼굴 정규화 영상을 입력으로 외형기반 접근법 중에 하나인 2D-DCT 특징과 EHMM을 이용하여 표정인식을 수행하였다[7]. 표정인식 수행 과정은 처음으로 입력 영상을 다수의 블록으로 나뉘고 각 블록마다 2D-DCT 특징 추출한 후, EHMM 알고리즘의 입력 관측 벡터로 사용한다. EHMM은 일차원의 HMM을 2차원 구조로 나타내기 위하여 일반화한 방법으로, 슈퍼 스테이트(super-states)와 삽입 스테이트(embedded-states)의 집합으로 구성된다. 여기에서 각각의 super-state는 하나의 일차원 HMM을 포함하며, 관측확률이 없고 상태전이 확률만을 갖는다. 표정 모델 학습은 처음으로 2D-DCT 관측벡터를 추출하고, 추출한 관측벡터를 세로축과 가로축 방향으로 균일하게 분할하여 EHMM 모델 파라미터를 초기화한다. 초기화된 모델 파라미터들은 이중 삽입 비터비 세그멘테이션(doubly embedded Viterbi segmentation) 알고리즘과 세그멘탈 K-평균(segmental K-means) 알고리즘을 이용하여 재 추정되고, 모델 파라미터들이 수렴할 때까지 재 추정 과정을 반복한다. 인식 과정에서는 입력 영상에 대한 관측벡터와 이중 삽입 비터비(doubly embedded Viterbi) 알고리즘을 이용하여 각 표정별 모델에 대한 확률값을 계산하고, 가장 높은 확률값을 갖는 모델을 선택하여 표정인식을 수행한다.

3.2 히스토그램 특징 기반 표정인식

더불어, 본 논문에서는 외형기반 접근법 중에 하나인 이

진패턴 히스토그램 특징과 SVM을 이용하여 얼굴 정규화 영상을 입력으로 표정인식을 수행하였다. 이진패턴 히스토그램 특징으로는 최근 조명에 강인한 변환기법으로 소개된 LBP[5]와 LDP[6]을 사용하였다. 이와 같은 이진패턴 기반의 일반적인 표정인식 접근 방법은 Fig. 7과 같은 과정으로 수행된다. 처음으로 얼굴 영상은 LBP 또는 LDP 연산을 이용하여 이진패턴으로 변환되고 구분된 몇 개의 영상 블록에 대한 히스토그램을 추출하여 이들을 연결한 벡터를 표정을 구분하기 위한 특징으로 사용한다. 다음으로, 추출한 히스토그램 특징 벡터는 SVM 분류기를 통하여 표정을 인식한다.

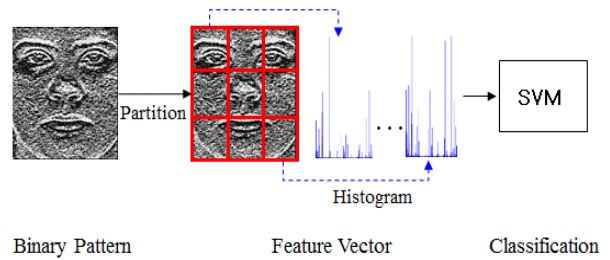


Fig. 7. Facial Expression Recognition Approach Based on Histogram Feature and SVM

3.3 얼굴 특징점 기반 표정인식

제안 방법의 성능비교를 위하여 얼굴 영상의 픽셀 레벨 정보를 이용하는 외형기반 접근법과 다르게, 본 절에서는 ASM 특징점 정보와 기하학적 얼굴 특징 정보를 추출하고, 이를 SVM 분류기로 표정인식을 수행하는 기하학적 접근법에 대하여 소개한다[9]. 기하학적 특징으로는 ASM 전체 특징점, slope 특징, angle 특징 등을 사용하였으며, 특징 추출 과정은 다음과 같다. Equation (2)로 표현되는 얼굴의 특징점이 주어졌을 때, 코의 중심점 특징인 α_{68} 를 중심으로 Equation (5)로 표현되는 좌표 변환을 수행하여 β 를 생성하고, 다음으로 스케일 정규화를 Equation (6)으로 수행하여 정규화된 특징점 벡터를 생성한다.

$$origin = \alpha_{68} = (x_{68}, y_{68}) \tag{4}$$

$$\beta_i = \alpha_i - origin, \quad i = 1, 2, \dots, 76 \tag{5}$$

$$\hat{\beta}_i = \beta_i / (\beta_{46} - \beta_{38}), \quad i = 1, 2, \dots, 76 \tag{6}$$

다음으로, 정규화된 특징벡터 $\hat{\beta}_i$ 로터 눈과 입에 대한 slope 특징과 angle 특징을 추출한다. 눈과 입에 대한 slope 특징은 각각 Equation (7)과 (8)로 계산되며, $S = [s_1, s_2, \dots, s_{10}]$ 를 slope 특징벡터로 사용한다. 또한 눈과 입에 대한 angle 특징은 각각 Equation (9)와 (10)으로 계산되며, $A = [a_1, a_2, \dots, a_{12}]$ 를 angle 특징벡터로 사용한다. 여기서 $\|*\|$ 는 norm 연산자를 의미한다.

$$\begin{aligned} s_1 &= (\hat{\beta}_{30} - \hat{\beta}_{29}) & s_2 &= (\hat{\beta}_{30} - \hat{\beta}_{31}) & s_3 &= (\hat{\beta}_{30} - \hat{\beta}_{28}) \\ s_4 &= (\hat{\beta}_{29} - \hat{\beta}_{28}) & s_5 &= (\hat{\beta}_{31} - \hat{\beta}_{28}) \end{aligned} \quad (7)$$

$$\begin{aligned} s_6 &= (\hat{\beta}_{55} - \hat{\beta}_{52}) & s_7 &= (\hat{\beta}_{55} - \hat{\beta}_{58}) & s_8 &= (\hat{\beta}_{55} - \hat{\beta}_{49}) \\ s_9 &= (\hat{\beta}_{52} - \hat{\beta}_{49}) & s_{10} &= (\hat{\beta}_{58} - \hat{\beta}_{49}) \end{aligned} \quad (8)$$

$$\begin{aligned} a_1 &= \arccos((s_1 \cdot s_3) / (\|s_1\| \cdot \|s_3\|)) \\ a_2 &= \arccos((s_2 \cdot s_3) / (\|s_2\| \cdot \|s_3\|)) \\ a_3 &= \arccos((s_4 \cdot s_3) / (\|s_4\| \cdot \|s_3\|)) \\ a_4 &= \arccos((s_3 \cdot s_5) / (\|s_3\| \cdot \|s_5\|)) \\ a_5 &= \arccos((s_1 \cdot s_4) / (\|s_1\| \cdot \|s_4\|)) \\ a_6 &= \arccos((s_2 \cdot s_5) / (\|s_2\| \cdot \|s_5\|)) \end{aligned} \quad (9)$$

$$\begin{aligned} a_7 &= \arccos((s_6 \cdot s_8) / (\|s_6\| \cdot \|s_8\|)) \\ a_8 &= \arccos((s_7 \cdot s_8) / (\|s_7\| \cdot \|s_8\|)) \\ a_9 &= \arccos((s_9 \cdot s_8) / (\|s_9\| \cdot \|s_8\|)) \\ a_{10} &= \arccos((s_{10} \cdot s_8) / (\|s_{10}\| \cdot \|s_8\|)) \\ a_{11} &= \arccos((s_6 \cdot s_9) / (\|s_6\| \cdot \|s_9\|)) \\ a_{12} &= \arccos((s_7 \cdot s_{10}) / (\|s_7\| \cdot \|s_{10}\|)) \end{aligned} \quad (10)$$

4. 실험 및 결과

제안하는 표정인식 방법의 성능평가는 CK(Cohn-Kanade AU-Coded Facial Expression) 데이터베이스[12]와 JAFFE (Japanese Female Facial Expression) 데이터베이스[13]를 이용하여 수행하였다. CK 표정 데이터베이스는 100명의 학생에 대한 7가지(무표정, 기쁨, 화남, 놀람, 공포, 슬픔, 혐오) 표정을 갖는 이미지 시퀀스로 구성되어 있으며, 이 중에서 320개의 시퀀스를 실험에 사용하였다. JAFFE 표정 데이터베이스는 213장의 일본인 여성 모델 10명으로 구성되어 있으며, 한 사람에 대하여 7가지(무표정, 기쁨, 화남, 놀람, 공포, 슬픔, 혐오) 표정 이미지로 구성되어 있다. 각 데이터베이스의 영상은 기존의 거리기반 얼굴 정규화 방법과 ASM 특징점 기반 얼굴 정규화 방법을 적용하여, 각각 얼굴 영역을 획득하고 이를 표정인식 성능평가에 이용하였다. 표정분류 접근법으로는 이진패턴 히스토그램 특징과 SVM을 이용한 방법과 2D-DCT 특징과 EHMM을 이용한 방법 및 ASM 특징점 자체를 특징으로 사용하는 기하학적 특징기반 접근법을 이용하여 성능평가를 수행하였다. 더불어, 본 논문에서는 표정인식 성능평가를 위한 학습 및 테스트 데이터셋을 랜덤 샘플링을 통하여 구분하는 기존 연구[5, 6]와는 다르게, 개인별 표정 영상을 순차적으로 선택하여 구성하였다. 이러한 방법을 통하여 본 실험에서는 각 데이터베이스에 대한 데이터 셋을 5개의 그룹으로 나누고, 한 개 그룹을 테스트에, 나머지 4개 그룹을 학습에 사용하는 5-fold 교차검증을 수행하여 결과를 도출하였다[7]. 단, 기존 연구[5, 6]는 데이터 셋을 랜덤 샘플링을 통하여 선택하므로, 학습과 테스트 데이터 셋에 동일한 사람에 대한 표정영상이 포함되

는 경우가 많아 인식률이 높게 나타난다. 즉, CK 및 JAFFE 데이터베이스는 한 사람에 대하여 3장의 표정 영상을 포함하고 있어, 표정인식 결과가 개인의 얼굴 영상에 의해 영향을 받는다. 이러한 영향을 줄이기 위하여, 본 논문에서는 개인별 표정 영상을 순차적으로 사용하여 학습과 테스트 셋에 동일인에 대한 영상이 겹쳐지는 확률을 낮춤으로써 표정인식 실험의 공정성을 추구하였다.

더불어, Fig. 8은 ASM 얼굴 정규화 영상과 EHMM 알고리즘을 사용하여 구현한 본 논문의 표정인식 시스템을 나타낸다. Fig. 8에서 ASM 얼굴 특징점이 사용자의 얼굴에 검출되는 것과 7가지 표정에 대한 인식 확률 및 결과가 나타나는 것을 확인할 수 있다.

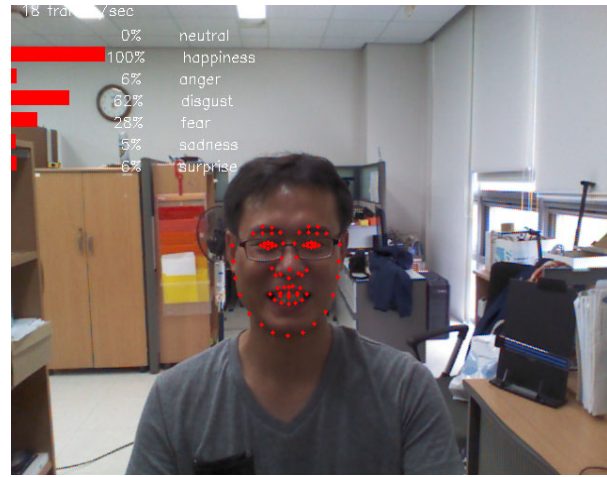


Fig. 8. Implemented Facial Expression Recognition System

CK 데이터베이스와 JAFFE 데이터베이스상에서 2D-DCT 특징과 EHMM 알고리즘을 이용한 외형기반 표정인식 방법의 실험 결과는 각각 Table 1과 Table 2와 같다. 본 논문에서는 2D-DCT 특징 추출을 위한 윈도우 크기를 8×8 , 16×16 , 24×24 , 32×32 로 변화하여 인식 성능을 조사하였으며, 5-fold 교차검증에 대한 인식률과 평균 인식률을 얼굴 정규화 방법에 따라 나타내었다. 실험 결과로부터 CK 데이터베이스와 JAFFE 데이터베이스 경우에 최대 인식률은 ASM 정규화 영상을 사용할 때, 각각 77.05%와 66.69%를 보였다. 또한 모든 윈도우 크기에 대하여, ASM 정규화 영상을 사용한 방법은 거리기반 정규화 영상을 사용한 경우보다 더 나은 인식률을 보였으며, 실제 CK 데이터베이스와 JAFFE 데이터베이스의 경우에 평균 6.39%와 4.18%의 인식 성능이 향상됨을 확인하였다. 또한, LBP 특징과 EHMM 알고리즘을 이용한 표정인식 방법[8]에 대한 CK 데이터베이스상에서의 결과를 Table 3에 JAFFE 데이터베이스상에서의 결과를 Table 4에 각각 나타내었다. 인용논문[8]과 유사하게 JAFFE 데이터베이스에서 16×16 윈도우를 사용할 경우, LBP 특징을 사용하는 방법은 2D-DCT 특징을 사용한 방법보다 0.48%의 더 좋은 결과를 보였지만, 다른 크기의 윈도

우를 사용한 경우에는 2D-DCT 방법이 더 높은 인식률을 보였다. 또한, CK 데이터베이스의 경우에도 2D-DCT 특징을 사용한 방법이 LBP 특징을 사용한 방법보다 우수한 인식 성능을 보임을 Table 1과 Table 3으로부터 알 수 있다.

Fig. 9는 위와 같은 표정인식 방법에서 오인식되는 JAFFE 데이터베이스의 얼굴 영상 예시를 보여준다. 실제 Fig. 9의 4장의 영상은 슬픔(sadness)의 표정으로 분류되나, 인식 결과는 처음의 두 영상은 화남(anger)으로, 다음 두 영상은 혐오(disgust)로 인식된다. 이러한 오인식 이유는 Fig. 9의 영상에서 확인 할 수 있듯이 사용자의 표정 분류에 대한 판단이 모호하여 오인식 결과를 나타낸 것으로 사료되며, 이는 표정별로 명확히 구분되는 많은 영상 데이터를 사용하여 표정 모델을 학습함으로써, 이와 같은 오인식 문제를 해결 할 수 있으리라 사료된다.

Table 1. Recognition Results of DCT-EHMM Approach for CK Database

	8×8		16×16		24×24		32×32	
	Eye	ASM	Eye	ASM	Eye	ASM	Eye	ASM
Set1	57.14%	76.59%	71.83%	80.56%	68.25%	79.37%	71.43%	76.98%
Set2	59.06%	73.62%	70.47%	78.35%	75.59%	77.95%	75.98%	75.20%
Set3	63.81%	66.54%	64.98%	75.49%	72.37%	79.77%	75.88%	79.38%
Set4	59.07%	68.34%	68.34%	74.52%	71.81%	74.13%	71.04%	72.20%
Set5	54.65%	66.28%	67.83%	68.99%	70.54%	74.03%	72.48%	72.09%
Average	58.74%	70.27%	68.69%	75.58%	71.71%	77.05%	73.36%	75.17%

Table 2. Recognition Results of DCT-EHMM Approach for JAFFE Database

	8×8		16×16		24×24		32×32	
	Eye	ASM	Eye	ASM	Eye	ASM	Eye	ASM
Set1	50.00%	50.00%	52.38%	52.38%	66.67%	54.76%	69.05%	73.81%
Set2	45.24%	59.52%	59.52%	64.29%	71.43%	85.71%	83.33%	85.71%
Set3	66.67%	64.29%	59.52%	66.67%	64.29%	61.90%	61.90%	57.14%
Set4	38.10%	50.00%	57.14%	54.76%	47.62%	57.14%	40.48%	52.38%
Set5	37.78%	53.33%	55.56%	64.44%	66.67%	64.44%	60.00%	64.44%
Average	47.55%	55.42%	56.82%	60.50%	63.33%	64.79%	62.95%	66.69%

Table 3. Recognition Results of LBP-EHMM Approach for CK Database

	8×8		16×16		24×24		32×32	
	Eye	ASM	Eye	ASM	Eye	ASM	Eye	ASM
Set1	40.87%	41.27%	57.94%	68.65%	62.70%	78.17%	65.87%	72.22%
Set2	60.63%	59.06%	68.11%	73.23%	69.69%	69.69%	70.08%	71.26%
Set3	44.36%	54.86%	64.59%	66.54%	65.76%	69.65%	76.26%	72.76%
Set4	54.83%	44.40%	63.71%	65.25%	69.11%	68.73%	69.88%	71.43%
Set5	55.04%	51.16%	75.97%	74.81%	76.36%	78.68%	77.13%	73.26%
Average	51.14%	50.15%	66.06%	69.69%	68.72%	72.98%	71.84%	72.18%

Table 4. Recognition Results of LBP-EHMM Approach for JAFFE Database

	8×8		16×16		24×24		32×32	
	Eye	ASM	Eye	ASM	Eye	ASM	Eye	ASM
Set1	38.10%	42.86%	61.90%	57.14%	57.14%	57.14%	50.00%	54.76%
Set2	45.24%	52.38%	57.14%	66.67%	61.90%	61.90%	59.52%	61.90%
Set3	30.95%	42.86%	47.62%	50.00%	50.00%	59.52%	52.38%	50.00%
Set4	40.48%	40.48%	33.33%	66.67%	42.86%	57.14%	35.71%	64.29%
Set5	37.78%	48.89%	71.11%	64.44%	68.89%	71.11%	62.22%	57.78%
Average	38.51%	45.49%	54.22%	60.98%	56.15%	61.36%	51.96%	57.74%



Fig. 9. An Example of Misclassified Facial Expression Images

다음으로, 이진패턴 히스토그램 특징과 SVM 기반의 외형기반 접근법을 사용할 경우, CK 데이터베이스와 JAFFE 데이터베이스에서의 실험 결과를 각각 Table 5와 Table 6에 나타내었다. 표정인식 결과는 LBP와 LDP 연산에 따라 구분하였으며, 히스토그램 특징 추출에 사용한 서브 블록 개수에 따라 인식 결과를 구분하였다. 실험으로부터 ASM 정규화 영상을 사용할 경우, CK 데이터베이스에서 LBP는 77.63%, LDP는 76.83%의 최대 인식률을 보였으며, JAFFE 데이터베이스는 LBP가 69.74%, LDP는 63.86%의 최대 인식률을 보였다. 더불어, 실험 결과로부터 ASM 정규화 영상을 사용한 방법이 거리기반 정규화 영상을 사용한 경우보다 향상된 인식 성능을 보였으며, 이는 CK 데이터베이스에서 LBP는 평균 2.16%, LDP는 1.98%의 성능 향상과 JAFFE 데이터베이스에서는 LBP는 평균 7.98%, LDP는 5.18%의 향상된 인식 결과로부터 확인할 수 있다.

Table 5. Recognition Results of Histogram Feature-Based Approach for CK Database

	LBP(5×5)		LBP(7×6)		LDP(5×5)		LDP(7×6)	
	Eye	ASM	Eye	ASM	Eye	ASM	Eye	ASM
Set1	72.22%	69.84%	73.01%	76.58%	69.44%	69.04%	66.87%	73.41%
Set2	65.74%	75.59%	68.89%	73.22%	72.44%	72.04%	72.86%	71.65%
Set3	73.54%	80.93%	73.93%	78.98%	70.80%	78.98%	72.76%	75.48%
Set4	79.53%	76.06%	82.62%	78.76%	79.53%	78.37%	81.85%	81.46%
Set5	80.23%	84.10%	83.33%	80.62%	77.90%	81.39%	79.68%	82.17%
Average	74.25%	77.30%	76.35%	77.63%	74.02%	75.96%	74.80%	76.83%

Table 6. Recognition Results of Histogram Feature-Based Approach for JAFFE Database

	LBP(5×5)		LBP(7×6)		LDP(5×5)		LDP(7×6)	
	Eye	ASM	Eye	ASM	Eye	ASM	Eye	ASM
Set1	57.14%	57.14%	69.04%	66.66%	57.14%	54.76%	59.52%	64.28%
Set2	50.00%	69.04%	59.52%	73.80%	57.14%	71.42%	69.04%	73.80%
Set3	64.28%	73.80%	57.14%	61.90%	57.14%	64.28%	57.14%	50.00%
Set4	64.28%	64.28%	52.38%	61.90%	50.00%	66.66%	57.14%	61.90%
Set5	66.66%	75.55%	68.88%	84.44%	62.22%	62.22%	60.00%	68.88%
Average	60.47%	67.96%	61.39%	69.74%	56.72%	63.86%	60.56%	63.77%

본 논문에서는 ASM 특징점을 이용하여 얼굴영역을 정규화하고 이를 외형기반의 표정인식 알고리즘을 적용하는 접근법에 대하여 제안하였다. 우선, 위와 같은 제안 방법의 효용성을 평가는 거리기반의 얼굴 정규화 방법과 비교하여 수행되었고, 선행 실험으로부터 제안 방법이 표정인식 성능 개선에 효과적임을 확인하였다. 다음으로, 본 실험에서는 영상 자체를 표정인식에 이용하는 외형기반 접근법이 아닌, ASM 특징점 및 기하학적 얼굴 특징 정보를 표정인식에 적용하여 제안 방법의 성능비교를 수행하였다[9]. 기하학적 특징기반 표정인식은 76개의 ASM 특징점, Slope 특징 및 Angle 특징을 개별적으로 SVM으로 인식하는 방법으로, 5-fold 교차검증에 의한 실험결과를 Table 7에 나타내었다. 실험 결과, CK 데이터베이스에서 ASM 특징점은 68.66%, Slope 특징은 66.65%, Angle 특징은 64.93%의 인식률을 보였으며, JAFFE 데이터베이스에서는 ASM 특징점은 45.99%, Slope 특징은 45.51%, Angle 특징은 39.87%를 보였다. 이로부터, 모든 데이터베이스에 대하여 제안하는 외형기반의 표정인식 방법이 기하학적 특징기반의 표정인식 방법보다 인식 성능 측면에서 더 우수한 접근법임을 확인할 수 있었다.

Table 7. Recognition Results of Geometric Feature-Based Approach

	CK			JAFFE		
	ASM	Slope	Angle	ASM	Slope	Angle
Set1	67.06%	68.65%	66.26%	30.95%	42.85%	42.85%
Set2	70.47%	68.11%	67.71%	50.00%	47.61%	38.09%
Set3	65.75%	64.98%	62.64%	47.61%	35.71%	26.19%
Set4	69.49%	64.47%	63.32%	54.76%	54.76%	50.00%
Set5	70.54%	67.05%	64.72%	46.66%	46.66%	42.22%
Average	68.66%	66.65%	64.93%	45.99%	45.51%	39.87%

5. 결론

본 논문에서는 ASM 특징점을 이용하여 얼굴 정규화를 수행하고, 이를 외형기반 접근법으로 표정 인식을 수행하는

방법에 대하여 제안하였다. ASM의 얼굴 특징점 자체를 표정분류에 이용하는 일반적인 기하학적 접근법과는 다르게, 본 논문에서는 ASM 기반 얼굴 정규화 영상을 외형기반 알고리즘을 이용하여 표정 분류를 수행함으로써, 표정인식 성능 향상을 도모하였다. 실제, CK 데이터베이스와 JAFFE 데이터베이스를 사용하여 실험한 결과, 제안 방법은 거리기반 정규화 영상을 이용한 기존 방법보다 최대, 6.39%와 7.98%의 성능 향상을 확인하였으며, 또한 기존의 기하학적 특징기반 표정인식 접근법보다 우수한 인식 성능을 보였으며, 이로부터 제안 방법의 효용성을 확인하였다. 향후, 조명 변화 환경에서 다양한 얼굴 포즈에 대한 얼굴 특징점 추출 정확도 향상 방법에 대한 연구와 이를 이용한 기하학적 특징기반 접근법을 외형 특징기반 표정인식 접근법과 결합하여 표정 인식 성능을 높이는 연구를 수행할 예정이다. 더불어, 표정 분류에 강인한 특징 기술자와 최근 기계학습 분야에서 각광을 받고 있는 깊은 학습(Deep Learning) 방법을 도입하여 표정 인식을 향상 위한 연구를 수행 할 계획이다.

References

- [1] T. F. Cootes, C. J. Taylor, D. H. Cooper and J. Graham, "Active shape models - their training and application," *Computer Vision and Image Understanding*, Vol.61, pp.38-59, 1995.
- [2] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.23, No.6, pp.681-685, 2001.
- [3] T. S. Lee, "Image representation using 2D Gabor wavelets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.18, No.10, pp.959-971, 1996.
- [4] Hyunsup Yoon, Youngjoon Han, and Hernsoo Hahn, "Real-time Recognition System of Facial Expressions Using Principal Component of Gabor-wavelet Features," *Journal of The Korean Institute of Intelligent Systems*, Vol.19, No.6, pp. 821-827, 2009.
- [5] C. Shan, S. Gong, and P. W. McOwan, "Facial Expression Recognition based on Local Binary Patterns: A Comprehensive Study," *Image and Vision Computing*, Vol.27, No.6, pp. 803-816, 2009.
- [6] T. Jabid, M. H. Kabir, and O. Chae, "Robust Facial Expression Recognition based on Local Directional Pattern," *ETRI Journal*, Vol.32, No.5, pp.784-794, 2010.
- [7] Dong-Ju Kim, Sang-Heon Lee, and Myoung-Kyu Sohn, "Feature Extraction Method of 2D-DCT for Facial Expression Recognition," *KIPS Transactions on Software and Data Engineering*, Vol.3, No.3, pp.135-138, 2014.
- [8] J. Cao and C. Tong, "Facial expression recognition based on LBP-EHMM," *2008 Congress on Image and Signal Processing*, pp.371-375, 2008.
- [9] X. Li, Q. Ruan, and Y. Ming, "A Remarkable Standard for Estimating the Performance of 3D Facial Expression

Features,” *Neurocomputing*, Vol.82, pp.99-108, 2012.

[10] P. Viola and M. J. Jones, “Robust real-time object detection,” *Technical Report Series, Compaq Cambridge research Laboratory*, CRL 2001/01, 2001.

[11] S. Milborrow, J. Morkel, and F. Nicolls, “The MUCT Landmarked Face Database,” in *Proc. Pattern Recognition Association of South Africa*, 2010.

[12] T. Kanade, J. Cohn, and Y. Tian, “Comprehensive Database for Facial Expression Analysis,” *IEEE International Conference Automatic Face Gesture Recognition*, pp.46-53, 2000.

[13] M. J. Lyons, J. Budynek, and S. Akamatsu, “Automatic Classification of Single Facial images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.21, No.12, pp.357-1362, 1999.



김 동 주

e-mail : kkb0320@gdsu.dongseo.ac.kr
 1998년 충북대학교 전파공학과(공학사)
 2000년 충북대학교 대학원 전파공학과
 (공학석사)
 2010년 성균관대학교 대학원 전기전자
 및 컴퓨터공학과(공학박사)

2011년~2015년 대구경북과학기술원 IT융합연구부 선임연구원
 2015년~현 재 동서대학교 컴퓨터공학부 조교수
 관심분야 : 얼굴인식, HCI, 영상신호처리, 패턴인식



신 정 훈

e-mail : only4you@cu.ac.kr
 1992년 성균관대학교 전자공학과(공학사)
 1994년 성균관대학교 대학원 전자공학과
 (공학석사)
 2005년 성균관대학교 대학원 전기전자
 및 컴퓨터공학과(공학박사)

1992년 SKC중앙연구소 연구원
 1994년 데이콤 종합연구소 주임연구원
 2006년~현 재 대구가톨릭대학교 IT공학부 부교수
 관심분야 : Human Computer Interaction, Brain Machine Interface