

논문 2014-51-11-24

얼굴정렬과 AdaBoost를 이용한 얼굴 표정 인식

(Facial Expression Recognition
using Face Alignment and AdaBoost)

정 경 중*, 최 재 식*, 장 길 진**

(Kyungjoong Jeong, Jaesik Choi, and Gil-Jin Jang[©])

요 약

본 논문에서는 얼굴영상에 나타난 사람의 표정을 인식하기 위해 얼굴검출, 얼굴정렬, 얼굴단위 추출, 그리고 AdaBoost를 이용한 학습 방법과 효과적인 인식방법을 제안한다. 입력영상에서 얼굴 영역을 찾기 위해서 얼굴검출을 수행하고, 검출된 얼굴 영상에 대하여 학습된 얼굴모델과 정렬(Face Alignment)을 수행한 후, 얼굴의 표정을 나타내는 단위요소(Facial Units)들을 추출한다. 본 논문에서 제안하는 얼굴 단위요소들을 표정을 표현하기 위한 기본적인 액션유닛(AU, Action Units)의 하위집합으로 눈썹, 눈, 코, 입 부분으로 나뉘지며, 이러한 액션유닛에 대하여 AdaBoost 학습을 수행하여 표정을 인식한다. 얼굴유닛은 얼굴표정을 더욱 효율적으로 표현할 수 있고 학습 및 테스트에서 동작하는 시간을 줄여주기 때문에 실시간 응용분야에 적용하기 적합하다. 실험결과, 제안하는 표정인식 시스템은 실시간 환경에서 90% 이상의 우수한 성능을 보여준다.

Abstract

This paper suggests a facial expression recognition system using face detection, face alignment, facial unit extraction, and training and testing algorithms based on AdaBoost classifiers. First, we find face region by a face detector. From the results, face alignment algorithm extracts feature points. The facial units are from a subset of action units generated by combining the obtained feature points. The facial units are generally more effective for smaller-sized databases, and are able to represent the facial expressions more efficiently and reduce the computation time, and hence can be applied to real-time scenarios. Experimental results in real scenarios showed that the proposed system has an excellent performance over 90% recognition rates.

Keywords : Facial expression recognition, face detection, face alignment, AdaBoost, action units.

* 정회원, 울산과학기술대학교 전기전자컴퓨터공학부
(School of Electrical and Computer Engineering,
Ulsan National Institute of Science and
Technology)

** 정회원, 경북대학교 전자공학부
(School of Electronics Engineering, Kyungpook
National University)

© Corresponding Author(E-mail: gjang@knu.ac.kr)

※ 이 논문은 2014년도 정부 교육부의 재원으로 한국
연구재단의 지원을 받아 수행된 기초연구사업임
(No. NRF-2010-0025642, No. NRF- 2014R1A1A1002662).
접수일자: 2014년08월22일, 수정일자: 2014년10월01일
게재확정: 2014년10월30일

I. 서 론

최근 스마트폰의 급속한 발전으로 인하여 디지털 이
미지 산업이 급성장 하며, 많은 산업분야에서 컴퓨터
비전의 기술이 요구되고 있다. 그 중에서 얼굴분야(얼
굴 검출, 얼굴 인식, 표정 인식 등)는 보안 및 인증, 스
마트 TV, CCTV 등 많은 산업분야에서 기술이 요구되
고 있고 더 나아가 사람과 로봇의 상호작용을 위한 기
반이 되는 기술로써 자리 잡고 있다.

표정인식 기술은 다양하고 도전적인 문제(조명, 포즈,

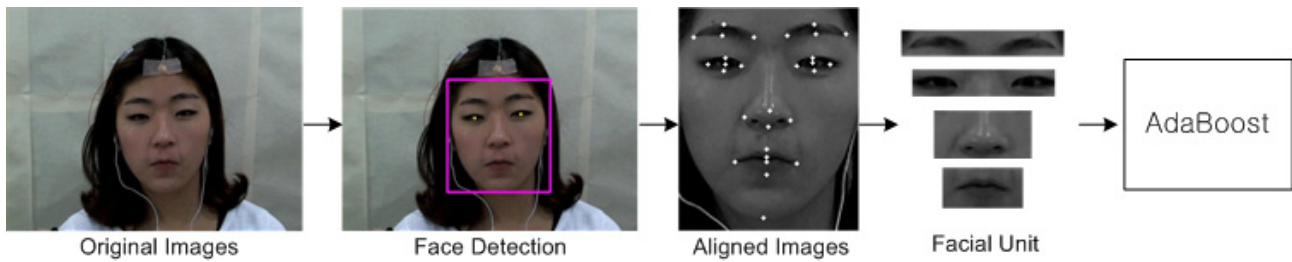


그림 1. 표정인식 학습을 위한 프레임워크

Fig. 1. Facial expression training framework.

표정의 다양성, 학습 데이터 구성)로 인하여 해결해야 할 난제가 많다. Ekman and Friesen은 자세한 얼굴표정을 인식하기 위하여 Facial Action Coding System (FACS)을 제안하였다.^[1] FACS는 44개의 액션유닛(Action Units)으로 구성되어 있으며, 얼굴의 미세한 표정 변화를 나타내기 위해서 유닛들의 조합을 사용하였다. 액션 유닛은 얼굴의 미세한 표정변화를 나타내기 위한 단위로 정의되었지만 최근의 기계학습 기반 자동인식을 위해서는 각 표정 유닛의 데이터베이스를 구성해야 하기 때문에 많은 학습자료가 요구된다.

Y. Tian 등은 특징 기반의 액션유닛 인식 시스템을 제안하였다.^[2] 그들은 FACS를 개선하여 사람이 표정을 지을 때 액션 유닛이 독립적으로 나타나지 않고 조합으로 나타나기 때문에 기존의 액션 유닛을 조합하고 신경망을 이용하여 학습하는 새로운 방법을 제안하였다. 또한 표정의 정확한 기하학적 모델을 구성하여 특징으로 사용하였다. H. Sadeghi 등은 기하학적 평준화 기법,^[3] Local Binary Patterns (LBP),^[4] Support Vector Machine (SVM)^[5,6] 학습기를 이용하여 표정인식 연구를 수행하였다. 그들은 얼굴 정렬 방법을 사용하지 않고 기하학적 방법으로 얼굴을 정렬하였다. 기존 방법과 유사한 연구를 수행하였으나 성능은 90% 이상의 결과를 보였다. Kotsia 등은 3가지(Gabor, DNMF, and Geometric displacement) 방법을 JAFFE와 Cohn-Kanade 데이터베이스를 이용하여 테스트하고 각각 91.6%, 86.7%, 91.4%의 성능을 보였다.^[7]

사람의 얼굴모양과 크기는 다양하기 때문에 기하학적으로 얼굴을 모델링하는 방식은 오차가 크다. 그렇기 때문에 기하학적 방법에서 얼굴의 특징 점을 찾는 방향으로 연구가 진행되었다. 얼굴에서의 특징 점은 눈썹, 눈, 코, 입, 턱선 등의 중요 포인트를 가리키며, 사람이 직접 손으로 좌표 포인트를 찍은 학습자료가 요구된다.

얼굴의 특징 점을 찾기 위해서 가장 많이 사용되는 방법은 ASM (Active Shape Model),^[12-13] AAM (Active Appearance Model)^[14] 방법들이다. 하지만 학습데이터에 종속적인 특성을 가지고, 학습과 인식 시간이 실시간 어플리케이션에 적용하기에 힘들다는 단점이 있다.

기존의 연구를 통하여 얼굴의 정확한 특징 점 추출(Face Alignment)은 표정인식에서 중요한 역할을 한다. 또한 얼굴의 표정을 나타내는 특징의 정의도 중요하다. 본 논문에서는 표정인식을 위해 다양한 기반 기술 연구와 특징 개발 및 표정을 인식하기 위한 프레임워크를 제안한다. 그림 1에 묘사된 전체 프레임워크는 다음과 같이 구성된다. 얼굴 이미지에서 얼굴을 찾기 위해서 얼굴검출을 수행하고, 검출된 얼굴 이미지에서 특징점을 찾기 위해서 얼굴정렬 알고리즘을 수행한 후, 얼굴의 표정을 잘 표현하기 위하여 얼굴 유닛들(Facial Units)을 추출한 후, AdaBoost^[8-10] 알고리즘을 이용하여 얼굴단위를 학습한다. 얼굴검출은 LBP 특징벡터와 AdaBoost 학습으로 만들어졌다. LBP 특징은 전체적인 조명 변화에 강인하여 얼굴 모델을 학습할 때 얼굴의 특징을 잘 표현한다. 얼굴정렬 알고리즘은 X. Cao 등이 제안한 boosted regression 방법을 사용하였으며,^[11] 이는 다양한 얼굴에 강인한 특성을 가지고 있고, 실시간에 적용할 수 있을 만큼 속도가 빠르다.

본 논문에서 제안한 얼굴유닛(Facial Units)은 얼굴의 표정을 나타내기 위한 기본적인 유닛으로 눈썹, 눈, 코, 입 부분으로 나누어진다. Ekman and Friesen이 제안한 액션 유닛은 얼굴의 상위, 하위 등 얼굴의 모양을 여러 부분으로 나누어 다양한(눈 사이의 거리, 눈 코 입의 위치 등) 얼굴의 모양을 반영하기가 힘들다는 단점이 있다.^[1] 얼굴의 정확한 특징점을 얼굴정렬 알고리즘으로 찾아낼 수 있으면 얼굴의 지역적인 영역을 특징으로 선택하는 것이 더욱 강인한 특징을 만들어낼 수 있다. 최

종적으로 추출된 각각의 얼굴유닛을 AdaBoost 알고리즘을 이용하여 학습하게 된다.

II장에서는 본 연구에 사용된 알고리즘을 기술하고, III장에서 실험 결과를 보이며 IV장에서 결론을 맺는다.

II. 본 론

1. 얼굴 검출(Face Detection)

얼굴 검출은 표정을 인식하기 위한 전처리 단계로서 학습 및 테스트 이미지가 주어졌을 경우 이미지에서 얼굴을 추출하는 단계이다. 얼굴을 검출하기 위하여 LBP^[4]와 AdaBoost^[10]을 사용하였다. LBP는 조명의 변화에 강인한 특징을 가진다. 그리고 AdaBoost는 weak classifier를 이용하여 strong classifier를 만드는 방법으로 얼굴 검출에 많이 이용되고 있다.^[8,9]

가. Local Binary Patterns (LBP)

LBP는 그림 2와 같이 center pixel(M)을 기준으로 8개의 이웃픽셀들과 비교하여 구성을 하고 그림 3과 같이 C1->C2->...->C8로 시계방향으로 M과 비교하여 M보다 크면 1, 그렇지 않으면 0 (Ci > M)으로 변환하여 저장한다.

LBP는 center pixel(M)을 기준으로 8개의 이웃픽셀들과 비교하여 구성되기 때문에 8bits로 표현되므로 [0-255]의 값으로 256가지로 표현되며, gray 값은 조명 변화에 민감하지만, LBP 값은 전체적인 조명 변화에 덜 민감한 특징을 보임을 알 수 있다.

C1	C2	C3
C8	M	C4
C7	C6	C5

그림 2. 임의의 3x3 영역의 픽셀 값
Fig. 2. The values of pixels in an arbitrary 3x3 region.

C8>M	C7>M	C6>M	C5>M	C4>M	C3>M	C2>M	C1>M
------	------	------	------	------	------	------	------

그림 3. LBP 변환
Fig. 3. LBP feature calculation.

나. LBP와 AdaBoost를 이용한 얼굴 검출기 학습

LBP는 입력영상에서 픽셀 값을 3×3 패치의 256가지 패턴들 중의 하나로 변환하는 것이다. 이는 $n \times n$ 크기의 학습 이미지에 대하여 상하좌우 외각영역을 제외한 내부 $(n-2) \times (n-2)$ 크기로 변환될 수 있음을 알 수 있다. 이때 AdaBoost는 $(n-2) \times (n-2)$ 크기의 LBP로 변환된 학습 이미지에서 얼굴과 비얼굴을 가장 잘 구분할 수 있는 위치와 해당 위치에서의 LBP의 패턴들의 가중치들을 weak classifier로 구성하고, 이러한 weak classifier들의 선형 조합으로 strong classifier로 구성하는 역할을 수행한다. AdaBoost 학습기를 학습하는 자세한 방법은 다음과 같다.^[10]

Step 1. 학습에 필요한 N_f 개의 얼굴 이미지와 N_{nf} 개의 비얼굴 이미지를 준비한다.

Step 2. 모든 학습(얼굴, 비얼굴) 이미지를 LBP변환을 통하여 LBP이미지($\Gamma_m^f, \Gamma_n^{nf}$)로 변환한다. 여기서 $m = 1, \dots, N_f$ 이고 $n = 1, \dots, N_{nf}$ 이다.

Step 3. 반복문의 인덱스($t = 1$)와 각 이미지의 가중치($D_t^f(m) = 0.5/N_f, D_t^{nf}(n) = 0.5/N_{nf}$)를 초기화 한다.

Step 4. 얼굴과 비얼굴 이미지의 pixel 위치에서 weighted kernels indices의 테이블을 수식(1), (2)을 이용하여 생성한다.

$$g_t^f(p, \gamma) = \sum_{m,p,\gamma} D_t^f(m) F(\Gamma_m^f(p) = \gamma) \quad (1)$$

$$g_t^{nf}(p, \gamma) = \sum_{n,p,\gamma} D_t^{nf}(n) F(\Gamma_n^{nf}(p) = \gamma) \quad (2)$$

여기서 $\gamma = 0, \dots, 255$ 값을 가지며 LBP이미지의 index를 나타내고, p는 pixel 위치를 나타내고, F()는 argument가 참이면 1, 거짓이면 0을 가지는 지시함수이다.

Step 5. 각각의 룩업테이블의 에러를 계산한다.

$$\delta_t(p) = \sum_{\gamma} \min \{g_t^f(p, \gamma), g_t^{nf}(p, \gamma)\} \quad (3)$$

Step 6. 반복문 t에서 에러가 가장 작은 위치 P_t 을 선택한다.

$$P_t \begin{cases} p | \delta_t(p) = \min_{p \in S} \{ \delta_t(p) \} \text{ if } |S_t| < N_p \\ p | \delta_t(p) = \min_{p \in S_t} \{ \delta_t(p) \} \text{ otherwise} \end{cases} \quad (4)$$

여기서 N_p 는 선택할 수 있는 위치의 최댓값이고, S_t 는 이전에 이미 선택된 위치들의 집합이다. 따라서 $S_t = S_i \cup p_t$ 가 된다.

Step 7. 반복문 t 와 픽셀의 위치 P_t 에서의 weak classifier 록업테이블을 생성한다.

$$w_t(\gamma) = \begin{cases} 0 & \text{if } g_m^f(p_t, \gamma) < g_n^f(p_t, \gamma) \\ 1 & \text{otherwise} \end{cases} \quad (5)$$

Step 8. 반복문 t 에서 에러 값으로부터 α_t 값을 얻는다.

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \delta_t}{\delta_t} \right) \quad (6)$$

Step 9. 픽셀 위치 P_t 의 classifier를 업데이트 한다.

$$h_{p_t}(\gamma) = h_{p_t}(\gamma) + \alpha_t w_t(\gamma) \quad (7)$$

Step 10. 각 이미지의 가중치 분포를 업데이트 하고 normalize 한다.

$$D_{t+1}^f(m) = D_t^f(m) \times \begin{cases} e^{-\alpha_t} & \text{if } w_t(\Gamma_m^f(p)) = 1 \\ e^{\alpha_t} & \text{otherwise} \end{cases} \quad (8)$$

$$D_{t+1}^{nf}(n) = D_t^{nf}(n) \times \begin{cases} e^{-\alpha_t} & \text{if } w_t(\Gamma_n^{nf}(p)) = 0 \\ e^{\alpha_t} & \text{otherwise} \end{cases} \quad (9)$$

Step 11. 만약 종료조건을 만족시키지 못하면 반복문 t 를 1증가시키고 Step 4로 돌아간다. 종료조건은 validation dataset을 이용하여 학습하기 이전에 정해놓은 FAR(False Acceptance Rate)와 FRR(False Rejection Rate)의 기준을 만족하는 것으로 한다.

Step 12. 종료조건을 만족하면 최종 strong classifier 를 만든다.

$$H_i(\Gamma) = \sum_{p \in S_i} h_p(\Gamma(p)) \quad (10)$$

2. 얼굴정렬(Face Alignment)

얼굴을 검출한 후 표정을 위한 학습 데이터는 다양한 얼굴의 크기, 위치, 포즈를 가지기 때문에 추출된 얼굴 이미지를 그대로 이용하여 학습하는 것은 성능을 떨어뜨리는 요인이 된다. 그래서 정확한 특징점을 찾는 방법이 요구된다. X. Cao^[11] 등은 boosted regression 방법을 사용하여 face alignment를 수행하였다. 주어진 얼굴 특징점이 실제 얼굴 특징점에 가장 가깝게 목표 함수를 추론하는 방법이다. 얼굴에 있는 전체 특징점을 vectorial regression을 이용하여 각각의 점이 아닌 전체 점을 한 번에 학습하였고, 학습 방법으로는 random fern을 사용하였다. 또한, shape indexed feature를 사용하여 사람의 얼굴 크기나 모양에 상관없이 특징점을 찾을 수 있는 장점이 있다.

가. Shape regression을 이용한 얼굴정렬

얼굴 이미지가 주어졌을 때, 얼굴 정렬의 최종 목표는 초기 shape이 실제 shape과 가장 근접되게 만드는 것이다. 그래서 얼굴정렬을 위해서 boosted regression 방법을 사용하였다. Boosted regression은 초기위치와 실제위치의 경로 T를 연결하기 위하여 weak regressors($R^1, \dots, R^t, \dots, R^T$)을 더해주는 방법을 사용한다. 얼굴 이미지 I 와 얼굴 shape의 초기 값 S_0 가 주어졌을 때, 각각의 regressor는 shape 증가분 δS 을 이미지의 특징으로부터 계산하고, 얼굴의 shape을 업데이트 한다.^[11]

$$S^t = S^{t-1} + R^t(I, S^{t-1}), \quad t = 1, \dots, T \quad (11)$$

식(11)에서 볼 수 있듯이, t^{th} weak regressor R^t 는 이전 shape S^{t-1} 을 새로운 shape S^t 로 갱신한다.

나. Shape indexed feature

일반적인 컴퓨터 비전 어플리케이션에서 이미지를 처리할 때 이미지의 중앙은 기하학적으로 이미지의 넓이와 높이의 반 지점을 사용한다. X. Cao^[11]에 나와 있는 그림처럼 사람의 얼굴은 인종, 성별, 나이 등에 따라서 얼굴의 모양이 다르게 나타나기 때문에 이미지의 중앙을 기준으로 얼굴의 특징점의 좌표를 설정하면 얼굴 모양의 다양함 때문에 얼굴의 특징분포가 사람마다 조금씩 다른 위치에 존재하게 되고 일반화하기 어려운 문

제가 있다. 이런 문제를 해결하기 위하여 shape indexed feature를 사용한다. 글로벌 좌표를 사용하는 방법(이미지의 중앙을 기준)보다 특징점을 기반(로컬한 특징점을 기준좌표로 설정)으로 같은 거리에 떨어져 있는 이미지의 pixel값이 더욱 유사하기 때문에 random fern을 사용할 때 인종, 성별, 나이 등에 받는 영향이 적어진다.

다. Random fern을 이용한 학습 방법

Regressor r을 학습하기 위하여 random fern 알고리즘을 사용한다. Fern은 F개의 features와 thresholds를 이용하여 feature 공간을 2^F 개의 bin으로 나눈다. 각각의 bin b는 bin으로 나뉘지는 학습샘플의 에러를 최소화 하는 regression 출력 δS_b 와 관련되어 있다.

$$\delta S_b = \operatorname{argmin}_{\delta S} \| \hat{S}_i - (S_i + \delta S) \| \quad (12)$$

식(12)의 해는 shape 차이 값의 평균이다.

$$\delta S_b = \frac{\sum_{i \in \Omega_b} (\hat{S}_i - S_i)}{\Omega_b} \quad (13)$$

이미지를 이용하여 random fern을 적용하여 학습하는 방법은 첫 번째 노드에서 얼굴이미지의 2개의 쌍의 점을 랜덤하게 추출하여 추출된 점의 pixel값을 비교하여 random하게 선택한 threshold값의 크기와 비교를 한다. 두 번째 노드에서는 다른 2개의 쌍의 점을 추출하여 크기를 비교한다. 계속해서 5개의 점의 쌍을 비교하여 5번째 노드까지의 threshold값과 비교하여 최종 종단 노드에 도달하게 된다. 종단 노드에 도달한 이미지를 이용하여 delta-shape을 구하고 종단 노드의 사이즈로 나누어 줌으로서 실제 delta-shape을 구하게 된다. 모든 학습데이터를 random fern을 이용하여 종단 노드로 구분시켜주면 전체 학습데이터를 이용하여 delta-shape의 크기를 구할 수 있다.

3. 표정인식(Facial Expression Recognition)

표정 인식을 위하여 얼굴 검출, 얼굴 정렬 기능을 수행하였다. 얼굴이 검출되면 여기에 얼굴 정렬 기능을 수행하여 중요 특징점을 추출한 후 그 특징점을 이용하여 표정 인식을 수행하게 된다. 표정인식을 위하여 선택된 감정은 무표정(Neutral), 기쁨(Happy), 화남

(Angry), 놀람(Surprise) 4가지 감정이다.

가. 얼굴유닛(Facial Units)

액션유닛을 사용하면 얼굴의 좀 더 많은 표정을 분류할 수 있지만, 여기에 사용된 기하학적 특징은 인종, 성별, 나이에 따른 다양한 얼굴의 변화를 반영하기 힘들어 인식률이 떨어질 수 있다. 또한 액션유닛의 조합은 7000가지 정도가 되기 때문에 각 유닛을 학습하기 위해서는 현실적으로 방대한 학습 데이터베이스가 필요하게 된다. 방대한 학습데이터와 기하학적 특징의 단점을 보완하기 위하여 본 연구에서는 얼굴유닛을 제안하였다. 얼굴 유닛은 표정을 인식하기 위한 단위 유닛으로서, 얼굴에서 4가지 부분(눈썹, 눈, 코, 입)을 추출하였다. 일반적으로 화가 난 얼굴에서는 미간이 찌푸러지며 놀란 얼굴에서는 입이 크게 벌어지며 눈이 커지는 현상을 볼 수 있다. 이러한 얼굴유닛의 조합을 이용하면 정확하게 표정을 인식 할 수 있다. 그리고 얼굴 유닛을 학습하기 위해서 4가지 단위 유닛만을 학습하면 되기 때문에 각 유닛의 데이터만을 수집하면 되고 학습데이터 수가 줄어든다. 표 1에서 자세한 얼굴유닛의 특징을 볼 수 있다.

표 1. 얼굴유닛: 얼굴에서 각 감정에서 보이는 파트를 구분하여 표정을 나타내기 위한 유닛의 세트

Table 1. Facial units: Combination of facial units for emotion recognition.

Facial Units	Characteristic
FU1, FU2	정상 눈썹, 찡그린 눈썹
FU3, FU4	정상 눈, 큰 눈
FU5, FU6	정상 코, 주름 있는 코
FU7, FU8, FU9	정상 입, 벌린 입, 이가 보이는 입

나. 표정인식을 위한 AdaBoost 학습

본 논문에서 표정을 인식하기 위하여 사용된 학습 알고리즘은 AdaBoost이다. 미리 지정된 얼굴유닛을 이미지 데이터로부터 추출한 이후에 각 얼굴유닛을 학습 이미지로부터 모아서 표 1의 구성처럼 AdaBoost 학습을 수행하였다. (FU1, FU2), (FU3, FU4), (FU7, FU8, FU9)를 각 세트로 구성하여 각각의 세트를 AdaBoost로 학습한 다음, 테스트 이미지에서 나온 얼굴유닛을 조합하여 표정을 인식하는 방법이다. 표 2에 자세한 조

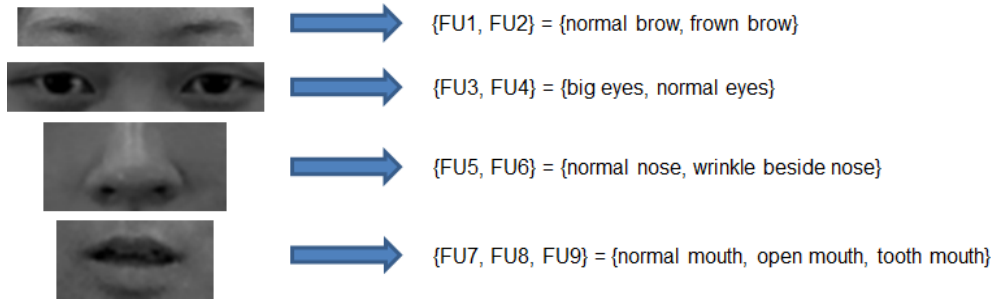


그림 4. 이미지를 통해서 보는 Facial Units
Fig. 4. Facial units from a facial image.

표 2. Facial Units의 조합
Table 2. Combinations of facial units.

	FU1	FU2	FU3	FU4	FU7	FU8	FU9
N	•		•		•		
A		•	•		•		
H	•		•				•
S	•			•		•	

합 방법을 나타내었다. 실제 제안한 얼굴유닛의 개수는 9가지이지만 코 부분의 특징이 정확하게 나타나지 않는 단점이 있어서 실험에는 코 부분을 제외한 7가지의 조합을 사용하였다.

그림 4는 이미지에서 얼굴유닛의 예제를 보여준다. 얼굴 정렬 알고리즘을 사용하여 특징점을 추출하면 각 특징 점을 이용하여 그림 4처럼 얼굴의 중요한 파트를 분리할 수 있다. 각 파트는 조명의 변화에 강한 특성을 가질 수 있게 LBP 특징으로 바꾸어서 학습한다.

III. 실험

본 논문에 구현된 시스템은 PC Windows7 환경에서 Visual C++와 OpenCV 2.4.6을 이용하여 구현하였고, 실험은 노트북에서 Logitech camera를 사용하여 테스트



그림 5. 학습에 사용된 샘플 사진
Fig. 5. Sample images for training.

트 하였다. 학습에 사용된 DB는 피실험자를 모집하여 각 감정별로 감정을 유도하여 영상을 제작한 후 이미지를 추출하여 수집하였으며 총 4300장의 이미지가 학습에 사용되었다. 그림 5는 학습에 사용된 이미지의 샘플을 보여준다.

1. Offline 실험

본 논문에서 offline 실험은 실시간으로 테스트를 진

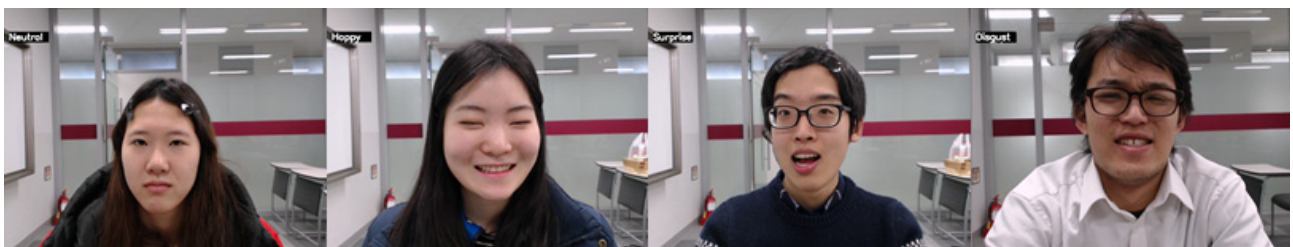


그림 6. Offline 실험 결과 샘플
Fig. 6. Image samples for offline experiments.

행하지 않고 사진을 촬영한 후 테스트 프로그램을 이용하여 성능을 측정하는 실험방식을 말한다. 테스트를 위해서 피실험자 5명 (남자3+여자2)을 모집하여 조명변화, 얼굴 각도 변화, 거리 변화를 적용하여 테스트를 진행하였다. 조명 변화를 위해서 실내, 실외 두 가지 환경에서, 각도 변화를 위해서 정면 얼굴의 각도가 +15°에서 -15°까지 변화한 상태에서, 거리 변화를 위해서 0.5m과 1m 두 가지 거리를 이용하여 테스트를 진행하였다. 총 실험 횟수는 조명별(2가지), 각도별(3가지), 거리별(2가지), 피실험자(5명)로 감정별로 60번 실험을 행하였다. 감정을 평가하는 방법으로는 감정을 유발하여 감정이 나타나는 5 frame을 추출하여 연속으로 3번 이상 정확한 감정이 나오면 맞는 것으로 평가하였다. 인식률은 (인식된 감정/전체 감정 평가 수)로 측정하였고 학습에는 5가지 감정을 사용하였으나 Angry와 Disgust의 감정이 유사하여 테스트에는 두 개의 감정을 통합하여 총 4가지 감정으로 테스트를 하였다.

표 3을 보면 Neutral의 감정은 실내, 실외 모두 100%의 성능으로 정확하게 평서 감정을 인지하는 것을 알 수 있다. 다른 감정들은 실험자들의 감정유발이 정확하지 않아서 표정을 정확하게 표현하지 못하는 경우에는 평서 감정으로 나타난다.

그림 6은 4가지 감정의 실험 결과 예제를 보여준다. 실험 결과는 offline에서 Neutral(100%), Happy(95.3%), Surprise(94.0%), Angry(81.3%)의 성능을 보였으며 평균 인식 시간은 30msec가 나왔다. 실내, 실외, 성별에 따른 성능차이는 미미했으며, 가까운 거리가 성능이 좋게 나왔다. 거리가 멀어지면 얼굴의 크기가 100×100

표 3. 실내, 실외에서의 offline 실험 결과
Table 3. Results of offline experiments in indoors and outdoors.

실내	N_out	H_out	S_out	A_out	Total
N_in	150	0	0	0	150
H_in	8	142	0	0	142
S_in	8	0	142	0	142
A_in	20	19	0	111	111

실외	N_out	H_out	S_out	A_out	Total
N_in	150	0	0	0	150
H_in	6	144	0	0	144
S_in	9	0	140	1	140
A_in	3	14	0	133	133

pixel 이하로 작아지며, 얼굴 검출의 성능에는 문제가 없었지만 얼굴 특징점 추출이 실패할 확률이 높아져서 거리를 1m로 제한하였다.

2. Online 실험

Online 실험은 스마트폰의 안드로이드 환경에 표정인식 시스템을 탑재하여 실시간으로 표정인식 테스트를 수행하는 방법을 말한다. Offline실험과 유사하게 4개의 감정(Neutral, Happy, Surprise, Angry)으로 테스트 하였다. 스마트폰에서 표정인식이 이루어질 경우 휴대가 편리한 이점이 있어서 좋지만 앱 프로그래밍을 해야 하는 번거로움과 스마트 폰의 하드웨어가 컴퓨터보다 사양이 낮아서 프로그램이 느려져 딜레이가 생길 수 있다는 단점이 있다. 하지만 휴대의 편리함은 단점을 극복하기에 충분히 효용의 가치가 있다.

Online 실험의 테스트를 위해서 피실험자 4명 (남자2+여자2)을 모집하여 스마트폰을 들고 손을 뻗을 수 있는 거리 내에서 감정을 유발하여 테스트를 하였다. Online 환경은 조명 변화, 손의 흔들림, 기기의 성능 등의 문제로 offline 실험보다 조건이 더욱 까다롭다. 총 실험 횟수는 피실험자(4명)로 감정별로 10번 실험을 행하였다. 감정을 평가하는 방법으로는 감정을 유발하여 감정이 나타나는 5 frame을 추출하여 연속으로 3번 이상 정확한 감정이 나오면 맞는 것으로 평가하였다. 인식률은 (인식된 감정/전체 감정 평가 수)로 측정하였고 학습에는 5가지 감정을 사용하였으나 Angry와 Disgust의 감정이 유사하여 테스트에는 두 개의 감정을 통합하여 총 4가지 감정으로 테스트를 하였다.

실험 결과는 표 4에 나타내었다. Online에서 Neutral (100%), Happy(90%), Surprise(90%), Angry(82.5%)의 성능을 보였다. Online 실험은 실시간으로 이루어지고 실험 환경이 offline보다 까다롭기 때문에 offline 실험보다 성능이 낮게 나왔다고 판단된다.

표 4. Online 실험 결과
Table 4. Results of online experiments.

	N_out	H_out	S_out	A_out	Total
N_in	40	0	0	0	40
H_in	3	36	1	0	36
S_in	4	0	36	0	36
A_in	7	0	0	33	33

IV. 결 론

본 논문에서는 얼굴 유닛(Facial Unit)과 AdaBoost를 이용한 표정인식 시스템을 제안하였다. 기존의 액션 유닛이 많은 DB를 요구하기 때문에 적은 수의 얼굴 유닛 집합을 제안하고 AdaBoost를 이용하여 학습하였다. 기존의 방법에 비해 더욱 간단하고 적은 수의 학습 데이터를 이용하여 학습이 가능하며 만족할 만한 성능을 얻었다. 이번 연구에서는 피실험자의 감정을 유발하여 학습데이터 및 테스트 데이터를 제작하였다. 더 도전적인 과제는 감정의 유발이 아닌 자연스런 감정을 인식하는 방법이다. 향후에는 이에 대해서 연구해 볼 계획이다.

REFERENCES

- [1] P. Ekman and W. V. Friesen, "The Facial Action Coding System: A Technique For The Measurement of Facial Movement," *Consulting Psychologists Press Inc.*, San Francisco, CA, 1978.
- [2] Y. Tian, T. Kanade, and J. Cohn, "Recognizing action units for facial expression analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 97 - 115, Feb. 2001.
- [3] H. Sadeghi, A. Raie, and M. Mohammadi, "Facial Expression Recognition using Geometric Normalization and Appearance Representation," *2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP)*, pp. 10-12, Sept. 2013
- [4] D.J. Kim, M.K. Sohn, and S.H. Lee, "A Study on Face Recognition Method based on Binary Pattern Image under Varying Lighting Condition," *The Institute of Electronics Engineers of Korea - Computer and Information*, Vol.49-CI, No.2, pp. 61-74, 2012.
- [5] E. Osuna, Support Vector Machines: Training and Applications, *PhD thesis*, MIT, EE/CS Dept., Cambridge, MA, 1998.
- [6] A. Mohan, C. Papageorgiou, and T. Poggio, "Example-based object detection in images by components," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(4):349-361, 2001.
- [7] I. Kotsia, I. Buciu and I. Pitas, "An Analysis of Facial Expression Recognition under Partial Facial Image Occlusion," *Image and Vision Computing*, vol. 26, no. 7, pp. 1052-1067, 2008.
- [8] J.H. Kim, K.H. Jang, J.H. Lee, and Y.S. Moon, "Multi-target Classification Method Based on Adaboost and Radial Basis Function," *The Institute of Electronics Engineers of Korea - Computer and Information*, Vol.47-CI, No.3, pp. 22-28, 2010.
- [9] P. Viola and M. Jones, "Fast and Robust Classification using Asymmetric Adaboost and a Detector Cascade," *Advances in Neural Information Processing System 14*, MIT Press, Cambridge, MA, 2002.
- [10] B. Froba and A. Ernst, "Face Detection with the Modified Census Transform," in *Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04)*, 91-96, 2004.
- [11] X. Cao, Y. Wei, F. Wen, and J. Sun, "Face alignment by explicit shape regression," in *Proceedings of the IEEE international Conference of Computer Vision and Pattern Recognition (CVPR)*, 2012.
- [12] S.J. Lim, Y.Y. Jeong, and Y.S. Ho, "Three-dimensional Model Generation for Active Shape Model Algorithm," *The Institute of Electronics Engineers of Korea - Signal Processing*, Vol.43, No.6, pp 28-35, November 2006.
- [13] J.Y. Kang, S.W. Lee, J.H. Shin, and J.K. Paik, "Hierarchical Active Shape Model-based Motion Estimation for Real-time Tracking of Non-rigid Object," *The Institute of Electronics Engineers of Korea - Signal Processing*, Vol.41, No.5, pp. 1-11, September 2004.
- [14] H.J. Park, J.H. Lee, T.H. Kim, and E.Y. Cha, "Multiple Active Appearance Model Based Facial Expression Recognition," in *Proceedings of the Institute of Electronics Engineers of Korea*, pp. 2138-2141, June 2010.

— 저 자 소 개 —



정 경 중(정회원)
2005년 울산대학교 전기전자
공학부 학사
2009년 POSTECH 정보통신학과
석사
2009년~2010년 삼성테크윈
연구원

2013년~현재 UNIST 전자공학과 박사과정
재학중
<주관심분야 : 이미지프로세싱, 컴퓨터비전, 기계
학습>



최 재 식(정회원)
2004년 서울대학교 컴퓨터공학과
학사
2012년 University of Illinois at
Urbana Champaign 전산
학과 박사

2013년 Lawrence Berkeley National Laboratory
박사후연구원
2013년~현재 울산과학기술대학교 전기전자컴퓨
터공학과 조교수
<주관심분야 : 인공지능, 기계학습, 컴퓨터비전,
로보틱스>



장 길 진(정회원)
1997년 KAIST 전산학과 학사
1999년 KAIST 전자전산학과
석사
2004년 KAIST 전자전산학과
박사
2004년~2006년 삼성종합기술원
전문연구원

2006년~2007년 Softmax, 연구원
2008년~2009년 University of California, San
Diego 박사후연구원
2009년~2014년 울산과학기술대학교 조교수
2014년~현재 경북대학교 조교수
<주관심분야 : 음성신호처리, 패턴인식, 컴퓨터비
전>