

스마트폰에서 웃음 치료를 위한 표정인식 애플리케이션 개발

강선경[†], 이옥걸^{**}, 송원창^{***}, 김영운^{****}, 정성태^{*****}

요 약

본 논문에서는 스마트폰에서 웃음 치료를 위한 표정인식 애플리케이션을 제안한다. 제안된 방법에서는 스마트폰의 전면 카메라 영상으로부터 AdaBoost 얼굴 검출 알고리즘을 이용하여 얼굴을 검출한다. 얼굴을 검출한 다음에는 얼굴 영상으로부터 입술 영역을 검출한다. 그 다음 프레임부터는 얼굴을 검출하지 않고 이전 프레임에서 검출된 입술영역을 3단계 블록 매칭 기법을 이용하여 추적한다. 카메라와 얼굴 사이의 거리에 따라 입술 영역의 크기가 달라지므로, 입술 영역을 구한 다음에는 고정된 크기로 정규화한다. 그리고 주변 조명 상태에 따라 영상이 달라지므로, 본 논문에서는 히스토그램 매칭과 좌우대칭을 결합하는 조명 정규화 알고리즘을 이용하여 조명 보정 전처리를 함으로써 조명에 의한 영향을 줄일 수 있도록 하였다. 그 다음에는 검출된 입술 영상에 주성분 분석을 적용하여 특징 벡터를 추출하고 다층퍼셉트론 인공신경망을 이용하여 실시간으로 웃음 표정을 인식한다. 스마트폰을 이용하여 실험한 결과, 제안된 방법은 초당 16.7프레임을 처리할 수 있어서 실시간으로 동작 가능하였고 인식을 실험에서도 기존의 조명 정규화 방법보다 개선된 성능을 보였다.

Development of Recognition Application of Facial Expression for Laughter Therapy on Smartphone

Sun-Kyung Kang[†], Yu-Jie Li^{**}, Won-Chang Song^{***},
Young-Un Kim^{****}, Sung-Tae Jung^{*****}

ABSTRACT

In this paper, we propose a recognition application of facial expression for laughter therapy on smartphone. It detects face region by using AdaBoost face detection algorithm from the front camera image of a smartphone. After detecting the face image, it detects the lip region from the detected face image. From the next frame, it doesn't detect the face image but tracks the lip region which were detected in the previous frame by using the three step block matching algorithm. The size of the detected lip image varies according to the distance between camera and user. So, it scales the detected lip image with a fixed size. After that, it minimizes the effect of illumination variation by applying the bilateral symmetry and histogram matching illumination normalization. After that, it computes lip eigen vector by using PCA(Principal Component Analysis) and recognizes laughter expression by using a multilayer perceptron artificial network. The experiment results show that the proposed method could deal with 16.7 frame/s and the proposed illumination normalization method could reduce the variations of illumination better than the existing methods for better recognition performance.

Key words: Facial Expression Recongnition(표정인식), Three Step Block Matching(3단계 블록 매칭), PCA(주성분 분석), Multilayer Perceptron Artificial Neural Network(다층 인공 신경망)

※ 교신저자(Corresponding Author): 정성태, 주소: 전북 익산시 신용동 344-2 원광대학교 공대 5123호(570-749), 전화: 063)850-6886, FAX: 063)856-8009, E-mail: stjung@wku.ac.kr

접수일: 2010년 11월 30일, 수정일: 2011년 1월 28일
완료일: 2011년 2월 24일

[†] 중신회원, 원광대학교 컴퓨터공학과
(E-mail: doctor10@wku.ac.kr)

^{**} 정회원, 원광대학교 컴퓨터공학과
(E-mail: liyujie0211@nate.com)

^{***} 정회원, 원광대학교 컴퓨터공학과
(E-mail: colorart@nate.com)

^{****} 준회원, 원광대학교 컴퓨터공학과
(E-mail: yukim@goodit.kr)

^{*****} 중신회원, 원광대학교 컴퓨터공학과

※ 본 연구는 2009학년도 원광대학교 교비 지원에 의해서 연구되었음.

1. 서 론

웃음을 통해 인간의 정신적, 육체적, 심리적, 사회적으로 건강을 만들 수 있어 웃음 치료는 사회적, 의료적으로 중요한 가치를 가진다고 할 수 있다. 그래서 PC 환경에서 사람에게 웃음을 훈련할 수 있는 웃음 표정 인식에 대한 연구가 진행되고 있다. 그런데 요즘에 u-healthcare에 발전에 따라서 언제 어디서나 치료를 받을 수 있는 연구가 활발하게 진행되고 있다. 본 논문에서는 u-healthcare를 위해서 스마트폰에서 웃음 치료를 할 수 있도록 웃음 표정 인식 애플리케이션을 개발하였다.

스마트폰은 범용 운영체제를 사용하고, 표준화된 개발 환경을 제공하여 개발자들이 자유롭게 애플리케이션을 개발할 수 있는 환경을 제공하고 있다. 따라서 일반 사용자들 간에 개방형 운영체제를 기반으로 다양한 애플리케이션의 공유가 가능하다. 최근 스마트폰은 폭 넓은 사용과 함께 이를 지원하기 위한 모바일 애플리케이션에 대한 연구가 급속히 증가하고 있다. 그런데 스마트폰의 CPU 성능이 아직은 PC에 비하여 뒤떨어지므로 PC환경에서 개발된 애플리케이션을 바로 스마트폰에 옮겨 사용하기는 어렵다.

본 논문에서는 스마트폰을 이용하여 무표정과 웃음 표정을 실시간으로 인식하는 애플리케이션을 개발하는 것에 목적을 두었으며, 속도 개선을 위해 여러 가지 방법을 사용하였다. 표정 인식을 위한 특징을 얼굴 전체 영상으로부터 구하지 않고 입술 영역만을 사용하여 속도를 개선하였다. 카메라 영상으로부터 입술 영역의 영상을 얻기 위해서는 먼저 얼굴 영역을 검출하였다. 얼굴 검출단계에서는 뛰어난 성능으로 인해 최근 여러 분야에서 널리 응용되고 있는 AdaBoost 얼굴 검출 알고리즘을 이용하여 얼굴을 정확하게 검출하였다. 얼굴을 검출한 다음에는 얼굴 영상으로부터 입술 영역을 검출한다.

그런데 AdaBoost 얼굴 검출 알고리즘은 PC 환경에서는 실시간으로 동작하는데 아무런 문제가 없지만, 스마트폰 환경에서는 실시간으로 동작하지 못하는 것으로 실험을 통하여 확인할 수 있었다. 따라서 본 논문에서는 모든 프레임에서 얼굴 검출을 하지 않고 한번 얼굴이 검출되면 얼굴 영역에서 입술 영역을 검출하고 그 다음 프레임부터는 입술 영역을 추적하는 방법을 사용하여 속도를 향상시킬 수 있도록

하였다. 입술 영역 추적을 위해서는 3단계 블록 매칭 기법을 이용하여 이전 프레임에서 검출된 입술영역과 가장 유사한 값을 가지는 영역을 현재 프레임에서 탐색하는 방법을 사용하였다.

입술 영역을 검출한 다음에는 크기 정규화와 조명 정규화 작업을 수행한다. 카메라와 얼굴 사이의 거리에 따라 입술 영역의 크기가 달라지므로, 입술 영역을 구한 다음에는 고정된 크기로 정규화할 필요가 있다. 또한 주변 조명 상태에 따라 영상이 달라지므로, 입술 영상에 대한 좌우대칭 평균화와 히스토그램 매칭을 결합하는 조명 정규화 방법을 적용하여 영상의 조명 변화를 줄일 수 있도록 하였다. 그 다음에는 정규화된 입술 영상에 주성분 분석을 적용하여 특징 벡터를 구하고, 다층 퍼셉트론 인공신경망을 통해서 웃음 표정을 인식하였다. 그림 1에는 본 논문에서 제안된 스마트폰 기반 웃음 표정 인식 애플리케이션 구성도가 나타나 있다.

2. 관련 연구

기존의 얼굴 검출 알고리즘은 크게 지식 기반 방법, 특징 기반 방법, 템플릿 방법, 외형 기반 방법으로

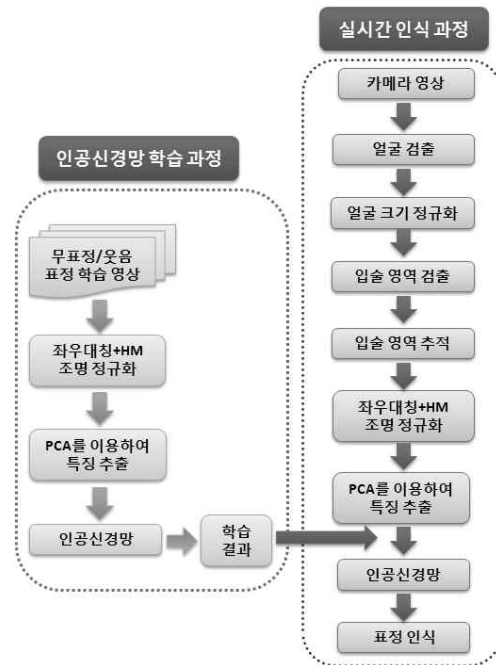


그림 1. 제안된 애플리케이션 구성도

나눌 수 있다. 그 중에서 외형 기반 방법은 얼굴의 특징을 이용하여 얼굴 이미지와 비 얼굴 이미지에 대해서 먼저 학습을 시켜서 얼굴과 비 얼굴 영상을 정확하게 인식하는 분류기를 생성한 다음에 생성된 분류기를 이용하여 검출할 이미지를 전체적으로 스캐닝해서 얼굴 검출하는 방법으로서, 최근에 널리 연구되고 있는 방법이다. 분류기를 생성하는 방법은 여러 가지가 있는데, 선형 부분 공간 방법(linear subspace methods), 인공신경망(neural network), 서포트 벡터 머신(support vector machine: SVM), 그리고 AdaBoost Boosting(AdaBoost)[1,2,3]알고리즘 등이 있다. 특히 Viola와 Jones 제안된 AdaBoost 얼굴 검출 알고리즘은 지금까지 다른 방법에 비해 뛰어난 성능으로 인해 최근 여러 분야에서 널리 응용되고 있다. 그래서 본 논문에서 기존의 AdaBoost 얼굴 검출 알고리즘을 적용하여 얼굴을 검출하였다.

기존의 표정 인식 연구 방법은 크게 광학적 흐름 분석(Optic Flow Analysis)방법[4], 국부적인 표현 방법[5], 홀리스틱 분석(Holistic Analysis)방법[6] 등이 있다. 그 중에서 홀리스틱 분석 방법은 영상 전체에 대한 분석으로 얼굴 영상에 대해 통계적으로 학습 시키고 인식하는 방법이다. 홀리스틱 방법에서는 차원을 축소하여 계산 량을 줄일 수 있고 영상 전체의 특징을 이용함으로써 표정 인식의 정확도를 향상 시킨다는 장점을 가지고 있기 때문에 본 논문에서 홀리스틱 분석 방법을 이용하여 웃음 표정 인식하는 방법을 사용하였다.

그러나 홀리스틱 분석 방법을 사용할 경우, 입력 영상에 대해 조명 변화에 따라 인식률에 영향을 미칠 수 있기 때문에 조명에 따른 전처리 과정으로 영상에 대한 정규화 과정이 필요하다. 따라서 본 논문에서는 기존의 조명 표준화[7], 백색화(whitening)[8], 히스토그램 매칭[9] 등의 방법보다 개선된 방법으로 좌우 대칭 평균화과 히스토그램 매칭을 결합하는 방법을 제안한다.

3. 제안한 표정인식 애플리케이션

3.1 입술 영역 검출

스마트폰 성능의 한계성에 인하여, 본 논문에서는 웃음 표정을 인식할 때 전체 얼굴 영상을 이용하지 않고 입술 영역을 찾은 후에 입술 영역만 이용하여

웃음 표정을 인식하였다. 사람이 웃음 표정을 지을 때 주로 입술 부분의 모양이 변화되므로 입술 영역의 영상만으로도 인식이 가능한 것으로 분석된다.

3.1.1 AdaBoost 알고리즘 이용하여 얼굴 영역 검출

본 논문에서 스마트폰의 전면 카메라를 사용하여 영상을 획득하였다. 윈도우즈 모바일 운영체제를 사용한 스마트폰을 이용하였고, DirectShow 라이브러리를 이용하여 영상을 획득하였으며, 카메라가 제공하는 여러 영상 형식 중에서 320×240 크기의 YUV12 형식의 영상을 사용하였다. 획득된 영상에 대해 바로 입술 영역을 검출하지 않고 먼저 얼굴 영역을 찾은 다음에 얼굴의 기하학적 특징을 이용하여 입술을 검출하였다.

본 논문에서는 Adaboost 얼굴 검출 알고리즘을 이용하여 얼굴을 검출하였다. 이 알고리즘은 Haar-like 특징을 적분 이미지를 통해서 빠른 속도로 계산하고, AdaBoost 알고리즘과 캐스케이드 구조를 이용하여 얼굴을 검출하는 방법이다. 즉 이 얼굴 검출 방법은 그림 2와 같이 크게 세 단계로 구성한다.

첫 번째, Haar-like 특징을 이용하여 얼굴을 표시하고, 적분 이미지를 이용하여 Haar-like 특징 값을 빠르게 계산한다.

두 번째, AdaBoost 알고리즘을 통해서 얼굴 특징을 제일 잘 구별하는 Harr-like특징 즉 약한 분류기들을 선택한 후에 선택된 가중치를 가진 약한 분류기

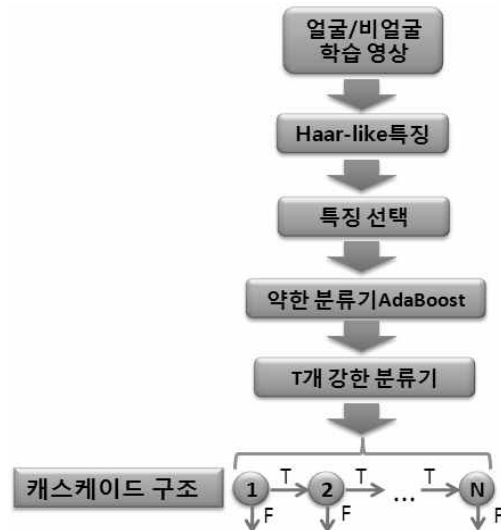


그림 2. AdaBoost 얼굴 검출 알고리즘

들을 결합하여 강한 분류기를 생성한다.

세 번째, 학습을 통해서 생성된 여러 개 강한 분류기를 캐스케이드 구조로 만들어 검출 속도와 정확도를 향상시켰다.

그림 3은 본 논문에서 AdaBoost 얼굴 검출 알고리즘을 이용하여 검출된 얼굴 영상의 예를 보여 주고 있다.

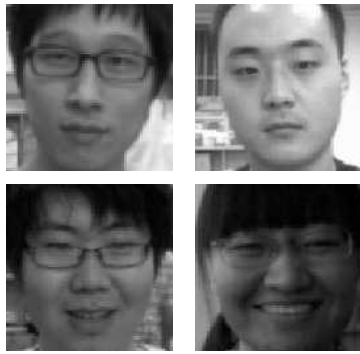


그림 3. 일반적인 DRM 흐름도

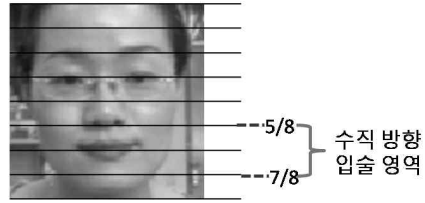
3.1.2 입술 영역 검출

본 논문에서는 얼굴 기하학 분포 특징과 그레이 픽셀의 분포 특징을 이용하여 입술 영역을 검출하였다. 사람의 입술은 얼굴의 하단 부분에 있으므로, 얼굴영역 하단 부분에 입술 후보 영역을 설정하였다. 입술의 중간에 있는 줄은 주변에 비하여 더 어두운 특징을 가지고 있으므로 명암도를 이용하여 입술 영역을 검출한다.

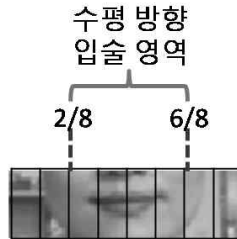
입술 후보 영역은 입술 수직 방향으로 대부분 그림 4(a)와 같이 전체 얼굴 영상의 5/8부터 7/8까지 부분을 차지하고, 수평 방향으로 대부분 그림 4(b)와 같이 전체 얼굴 영상의 2/8부터 6/8까지 부분을 차지한다. 그래서 얼굴의 기하학 분포 특징에 근거하여 그림 4(c)와 같이 입술 후보 영역을 설정한다.

입술 후보 영역을 설정한 다음 후보 영역에서 한 줄씩마다 픽셀 값의 합을 계산하여 제일 어두운 줄을 검출한다. 따라서 검출된 줄의 중심은 그림 5와 같이 $(x_{lip_{center}}, y_{lip_{center}})$ 을 정의할 수 있다. $x_{lip_{center}}$ 는 영상의 수평 중심 위치로 설정하였다.

입술 중심을 검출한 다음에는 다음과 같은 영역을 입술 영역으로 설정하였다. 여기에서 H 는 검출된 얼굴 영역의 높이이고 W 는 검출된 얼굴 영역의 넓



(a) 입술 수직 방향 영역



(b) 입술 수평 방향 영역



(c) 입술 후보 영역

그림 4. 입술 후보 영역 검출 과정

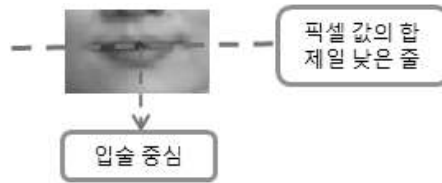


그림 5. 입술 중심 검출

이다.

$$y_{min} = y_{lip_{center}} - H/8$$

$$y_{max} = y_{lip_{center}} + H/8$$

$$x_{min} = x_{lip_{center}} - W/4$$

$$x_{max} = x_{lip_{center}} + W/4$$

그림 7은 위에서 설명한 입술 검출 알고리즘을 이용하여 다양한 환경에서 입술 영역을 검출하는 화면

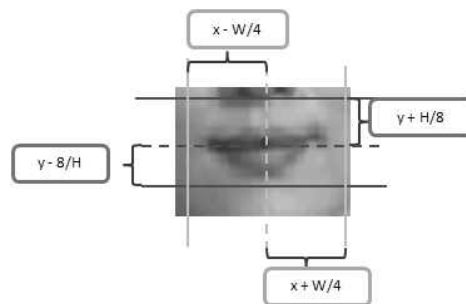


그림 6. 입술 영역 검출

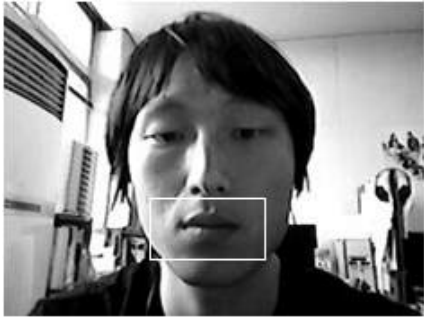


그림 7. 입술 영역 검출 화면

이다. 구현된 애플리케이션은 그림과 같이 입술의 위치를 표시 해주고 있다.

3.2 입술 영역 추적

교통 시스템에서 차에 대한 추적, 보안 시스템에서 사람에 대한 추적 등 사람 또는 객체 추적에 대한 연구는 활발히 진행되고 있다. 추적 방법 중에서 블록 매칭에 의한 추적 방법이 있는데, 블록 매칭 추적 알고리즘으로는 3단계 탐색(Three-step Search : TSS)[10], Two Dimensional Logarithmic search (TDL)[11], Binary Search(BS)[12], Cross Search Alogrithm(CSA) [13], Hierarchical Search Block Matching Algorithm[14] 등이 있다. 그 중에서 3단계 탐색 알고리즘은 1981년에 Koga 등이 제안한 탐색 방법이다. 이 방법은 간단하고 견고하면서, 결과는 최적 궤적과 가까이하기 때문에 움직임 추적 분야에서 널리 응용되고 있다. 본 논문에서는 3단계 탐색 알고리즘을 이용하여 입술 영역에 대해 추적하였다.

본 논문에서는 이전 프레임에서 검출된 입술 영역과 가장 유사한 영역을 현재 프레임에서 탐색하는데, 탐색 영역은 이전 위치를 중심으로 15×15 영역이다. 이전 프레임의 입술 영상과 가장 비슷한 영역을 찾기 위하여, 각 후보 위치 (i, j) 에 대하여 후보 영상과 이전 입술 영상 사이의 픽셀 값의 절대 차이의 합(SAD: Sum of Absolute Difference)을 구하여 이 값이 가장 작은 위치를 현재 프레임의 입술 영상으로 결정하게 된다.

$$SAD(i, j) = \sum_{y=0}^{M-1} \sum_{x=0}^{N-1} |A(bx+x, by+y) - B(bx+x+i, by+y+j)|$$

여기에서, $-7 \leq i, j \leq 7$ 이고, $A(bx+x, by+y)$ 는 이

전 프레임에서 입술 영상의 픽셀값을 나타내고 $B(bx+x+i, by+y+j)$ 는 현재 프레임에서의 픽셀값이다. (bx, by) 는 이전 프레임 영상에서 입술 영역의 시작 위치를 나타낸다. M 과 N 은 추적하는 입술 영역의 높이와 넓이를 나타낸다.

모든 후보 위치에서 SAD를 계산하면 $15 \times 15 \times M \times N$ 번의 뺄셈 연산을 필요로 하게 되어 많은 연산을 해야 한다. 이에 반하여 3단계 탐색 알고리즘은 대략적인 탐색으로부터 정확한 탐색까지 단계적으로 탐색해나가는 방법이다. 각 단계에 9개의 후보 위치에서 SAD를 계산하여 총 $27 \times M \times N$ 번의 연산을 함으로써 모든 위치에서 SAD를 계산하는 방법보다 8배 정도 연산량을 줄일 수 있다. 연산량이 줄어드는 대신에 정확도는 모든 후보 위치를 탐색하는 것에 비하여 떨어질 수 있지만 실험결과 입술영역 추적에는 문제가 없는 것으로 분석된다.

3단계 탐색 알고리즘의 첫 번째 단계에서는 그림 8에서 1번으로 표시되어 있는 9개의 위치에서 SAD 계산을 수행하고 제일 작은 후보 위치를 선택한다. 두 번째 단계에서는 첫 번째 단계에서 SAD 값이 제일 작았던 후보 위치를 중심으로 9개 후보 위치에 대하여 SAD를 계산한다. 그림 8에서는 좌측 하단의 후보 위치가 첫 번째 단계에서 SAD 값이 제일 작았고 그 점을 중심으로 2번으로 표시되어 있는 후보 위치를 설정하고 SAD 값을 계산하는 것이 나타나 있다. 세 번째 단계에서는 두 번째 단계에서 SAD 값이 제일 작았던 후보 위치를 중심으로 9개 후보 위치에 대하여 SAD를 계산한다. 그림 8에서는 중앙 하단의 후보 위치가 두 번째 단계에서 SAD 값이 제일 작았고 그 점을 중심으로 3번으로 표시되어 있는 후보 위치를 설정하고 SAD 값을 계산하는 것이 나타나 있다. 그림에 나타나 있듯이 단계별로 후보 위치를 정하는 거리가 4, 2, 1로 줄어드는 것을 알 수 있다.

3.3 입술 영상 정규화

일반적으로 영상을 인식하는 시스템은 전처리 단계, 특징 추출 단계, 인식 단계 세 단계로 나눈다. 영상을 획득한 후에 우선 영상에 대한 전처리를 실행하고, 다음에 특징을 추출한 후에 마지막으로 추출된 특징을 인식하는 과정으로 진행한다. 검출된 입술 영상을 이용하여 표정을 인식하기 전에 보다 좋은 결과

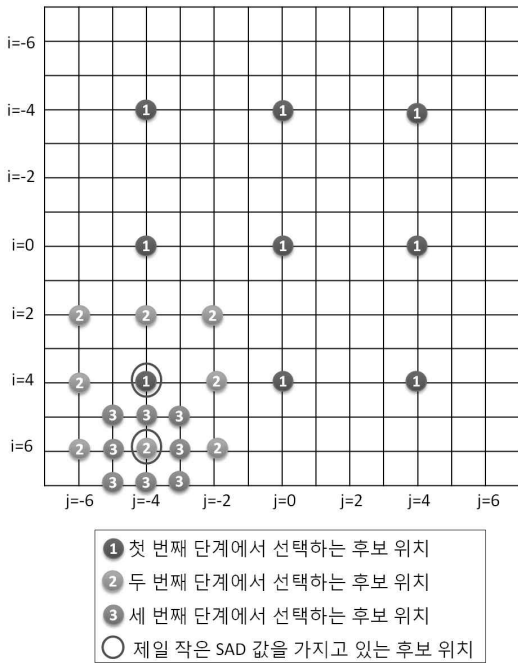


그림 8. 3단계 탐색 알고리즘 적용 예

를 얻기 위해 입술 영상 전처리 과정을 거쳐야 된다. 영상에 대한 전처리 과정은 표정 인식기 성능에 많은 영향을 미칠 수 있다. 전처리 과정으로는 크기 정규화와 조명 정규화를 수행하였다. 휴대폰과 얼굴 사이의 거리가 때에 따라 달라질 수 있으므로 크기 정규화를 수행하였다. 또한 스마트폰을 사용하는 조명 환경은 다양하기 때문에 검출된 입술 영상에 대해 조명 정규화 전처리과정을 실행하였다.

본 논문에서는 입술 영상의 크기를 64×32 크기로 일정하게 정규화 하였다. 이렇게 함으로써 학습과 인식에 사용되는 영상에 있어서 크기 차이로 인한 문제를 제거하였다.

다양한 환경에서 21명 실험자에 대하여 제안된 입술 영역 검출 및 추적 알고리즘을 적용하여 획득된 입술 영상 예가 그림 9에 나타나 있다. 그림 9(a)는 검출된 무표정 입술 영상들이고, 그림 9(b)는 검출된 웃음 표정 입술 영상들이다.

기존의 조명 정규화 방법은 표준화, 백색화, 히스토그램 매칭 등 영상을 전역적으로 보정하기 때문에 영상이 전체적으로 어둡거나 또는 밝은 영상에 대해서 좋은 조명 정규화 효과를 나타내지만 조명이 좌측 혹은 우측에서 비추질 때 영상이 영역별로 차이가



(a) 무표정 입술 영상



(b) 웃음 표정 입술 영상

그림 9. 검출된 입술 영상 예

생겨서 조명 정규화 효과가 크지 않다는 것으로 분석되었다. 이를 보완하기 위해서 본 논문에서 사람의 입술 영상의 좌우 대칭성을 근거하여 좌우대칭 평균화와 히스토그램 매칭을 결합하는 방법을 제안한다.

제안된 방법은 검출된 입술 영역에 대하여 먼저 좌우대칭 평균화를 실행해서 국부적인 조명 왜곡을 개선하였고, 다음에 좌우 명암이 고르게 분포된 입술 영상에 대하여 히스토그램 매칭을 실행하여 전체적으로 조명을 개선하였다. 이 방법은 부분적 조명 정규화를 수행한 후에 전체 조명 정규화를 수행하는 방식을 통해서 전체적으로 어둡거나 밝은 영상뿐만 아니라 부분적으로 왜곡되는 현상을 개선하였다.

그림 10에서는 조명 정규화 방법에 따른 실행 결과 예가 나타나 있다. 제안된 방법이 조명 변화가 심한 영상에 대해서도 기존 방법에 비하여 안정된 결과를 생성하는 것을 알 수 있다.

3.4 입술 영상을 이용한 표정인식

입술 영역을 검출하고 정규화한 다음에 본 논문에서는 차원을 축소 방법인 주성분분석(PCA)을 이용하여 입술 영상의 특징 벡터를 추출하였다.

그림 11은 PCA알고리즘을 통해서 구한 고유 벡터 영상, 즉 고유 입술 영상들이다.



그림 10. 조명 정규화 방법에 따른 결과 영상

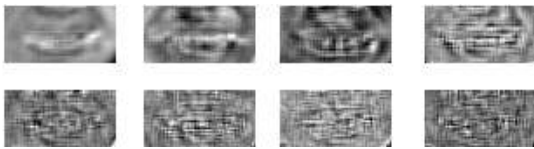


그림 11. PCA로부터 구한 고유 입술 영상들

입술 영상의 특징 벡터를 추출한 다음에 본 논문에서 인공신경망을 이용하여 표정을 인식하였다. 인공신경망은 무감독 학습과 감독 학습으로 나눌 수 있다. 감독 학습 알고리즘은 그림 12와 같이 입력된 데이터의 출력 결과가 의도한 결과와 일치하는지 알려주는 교사 신호를 필요하므로 인위적으로 제어할 수 있고, 학습된 시스템의 신뢰도가 높기 때문에 많이 사용되고 있다.

그래서 본 논문에서 대표적인 감독 학습 알고리즘인 그림 13과 같은 3계층 역전파 인공신경망[12]을 이용하여 표정을 인식하였다. 본 논문에서는 무표정

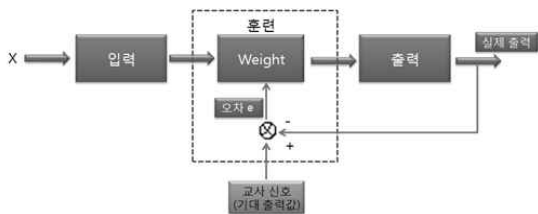


그림 12. 감독 인공신경망

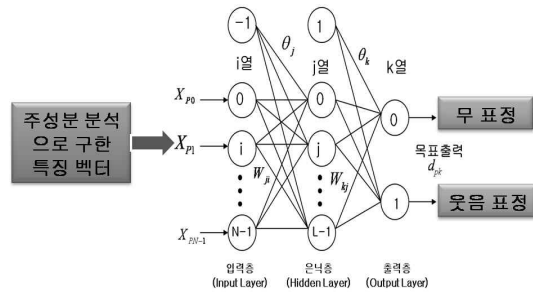


그림 13. 3계층 역전파 인공신경망

과 웃음 표정을 분류하기 위한 인공신경망에서 PCA를 통해서 구한 특징 벡터를 입력 노드로 사용하고, 출력 노드는 2개(무표정은 0로 표시하고 웃음 표정은 1로 표시한다)로, 은닉 계층 노드는 입력 층 노드와 같은 계수로 설정하였다.

4. 실험결과 및 고찰

스마트폰에서 실시간으로 웃음 표정 인식 애플리케이션을 개발하기 위해 본 논문에서 삼성 OMNIA II를 이용하여 구현하였다. 실험 환경은 표 1과 같다.

표 1. 웃음 표정 인식 애플리케이션 실험 환경

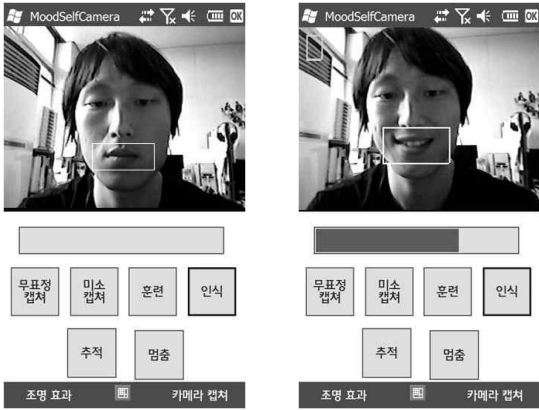
구 분	환 경	비 고
운영체제	Windows Mobile 6.5	
CPU	800MHz	
RAM	512MB RAM	
프로그램 개발도구	MS Visual C++ 2008	
입력 장치	500만 화소 AF 카메라	

본 논문에서 설계된 인식 인터페이스 화면은 그림 14와 같다. 그림에는 개발된 웃음 표정 인식 애플리케이션을 실시간으로 실험한 결과가 나타나 있다. 그림 14(a)에 보여주는 것과 같이 무표정일 때 표시를 안 해 주고, 그림 14(b)는 실시간으로 웃는 표정을 인식하여 웃음 표정 인식 결과를 그래프로 표시하여 나타내고 있다.

제안된 방법의 수행속도에 대한 실험 결과가 표 2에 나타나 있다. 표에 나타나 있는 바와 같이, PC에서 개발된 애플리케이션의 속도에 비해서 많이 떨어지지만 제안된 방법에서는 입술 추적과 고정소수점 연산을 같이 이용함으로써 스마트폰에서도 실시간

표 2 제안된 방법의 수행 속도

구 분	스마트폰			PC
	추적기능+ 고정소수점 연산	추적기능+ 부동소수점 연산	추적기능 사용하지않음+ 부동소수점 연산	추적기능 사용하지 않음+ 부동소수점 연산
속도(프레임/초)	16.7	9.2	1.7	44.4



(a) 무표정 인식 화면 (b) 웃음 표정 인식 화면

그림 14. 표정 인식 인터페이스

동작을 할 수 있었다.

본 논문에서 제안된 표정 인식 애플리케이션의 성능을 검증하기 위해서 그림 15와 같이 다양한 조명 환경(좌측 광원, 우측 광원, 평균 광원 등)에서 21명 실험자들을 촬영한 영상으로부터 입술 영상을 검출하여 실험을 해 보았다. 각 사람에 대해 무표정, 웃음 표정 각 10장씩 총 420장을 검출하였다. 그 중에 입의 17명 실험자의 입술 영상은 학습 영상으로 사용하고, 나머지 4명 실험자의 입술 영상은 테스트 영상으로 사용하였다. 본 논문에서 모두 15개 그룹을 만들어서

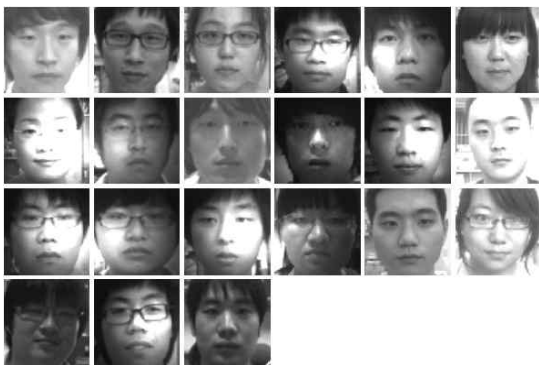


그림 15. 실험자 얼굴 영상

이용하여 실험해 보았다.

표정 인식 결과는 특징 벡터의 크기에 따라 달라질 수 있는데, 본 실험의 경우에는 특징 벡터의 크기가 60일 때에 가장 좋은 성능을 보였다. 특징 벡터 60개일 때 조명 정규화 방법에 따른 15번째 실험의 평균적인 표정 인식 결과가 그림 16에 나타나 있다.

이 실험 결과를 통해서 조명 변화가 인식률에 영향을 미칠 수 있다는 것을 알 수 있다. 그리고 본 논문에서 제안된 좌우대칭 평균화와 히스토그램 매칭을 결합하는 조명 정규화 방법은 다른 방법보다 인식률이 높은 98.75%의 인식률을 얻을 수 있었다. 제안된 방법은 다른 조명 정규화 방법보다 조명 변화에 대해 영향을 더 효과적으로 줄여 인식률을 향상시킬 수 있다는 것을 알 수 있었다.

표정 인식률

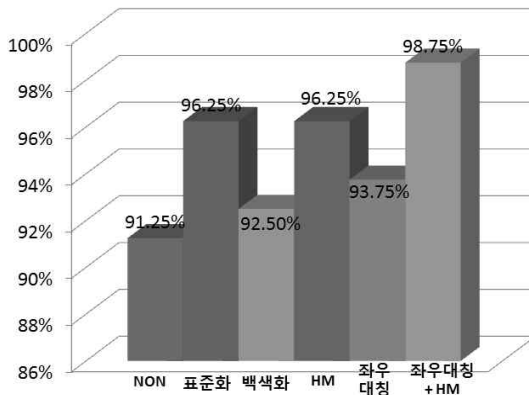


그림 16. 조명 정규화 방법에 따른 표정 인식률

5. 결 론

본 논문에서는 스마트폰에서 웃음을 훈련할 수 있는 웃음 표정 인식 애플리케이션을 개발하기 위해서 Haar-like 특징을 이용하여 얼굴 영역을 검출한 다음에 얼굴 영상의 기하학특징과 픽셀 분포 특징을 이용하여 입술 영역을 검출하였다. 입술 영역을 검출한

다음에 3단계 탐색 추적 방법을 이용하여 검출된 입술 영역에 대해서 추적을 실행하는 방식을 통해서 속도 문제를 개선하였다. 그리고 좌우대칭 평균화과 히스토그램 매칭을 결합하는 방법을 이용하여 조명 정규화를 수행함으로써 조명 변화에 대해 영향을 효과적으로 줄여서 인식률을 향상시켰다.

향후 연구 방안으로는 무표정과 웃음 표정의 구분 뿐 만 아니라 웃음의 강약 정도를 측정할 수 있는 애플리케이션을 개발할 필요가 있는 것으로 사료된다.

참 고 문 헌

- [1] Lienhart R, and Maydt J, "An extended Set of Harr-Like Features for Rapid Object Detection," *Proceedings of International Conference on Image Processing*, Vol.1, pp. I-900-I-903, 2002.
- [2] Viola P and Jones M, "Rapid Object Detection Using a Boosted Cascade of Simple Features," *Computer Vision and Pattern Recognition*, CVPR 2001. Vol.1, pp. I-511-I-518, 2001.
- [3] 이우주, 김진철, 이배호, "AdaBoost 알고리즘을 이용한 실시간 얼굴 검출 및 추적," 한국멀티미디어학회논문지, Vol.9, No.10, pp. 1266-1275, 2006.
- [4] J. Lien, T. Kanade and C. Li, "Detection, Tracking, and Classification of Action Units in Facial Expression," *Journal of Robotics and Autonomous Systems*, Vol.31, No.3, pp. 131-146, 2000
- [5] C. Padgett and G. Cottrell, "Representing Face Images for Emotion Classification," *Advances in Neural Information Processing Systems*. Vol. 9, MIT Press. 1997.
- [6] 김호덕, 양현창, 박창현, 심귀보, "PCA을 이용한 얼굴 표정의 감정인식 방법," Proceedings of KFIS Autumn Conference, 제16권, 제2호, pp. 772-776, 2006.
- [7] 신영숙 "PCA 표정을 이용한 강인한 얼굴 표정 인식," 한국인지과학회 논문지, 제16권, 제4호, pp. 323-331, 2005.
- [8] 한수정, 광근창, 고현주, 김승석, 전명근, "ICA-factorial 표현법을 이용한 얼굴감정인식," 한국 퍼지 및 지능시스템학회 논문지, 제13권, 제3호, pp. 371-376, 2003.
- [9] 이옥걸, 강선경, 김용운, 정성태, "웃음 치료 훈련을 위한 웃음 표정 인식 시스템 개발," 한국 컴퓨터 정보학회 논문지, 제15권, 제4호, pp. 47-55, 2010.
- [10] T. Koga, K. Linuma, A. Hirano, Y. Iijima, and T. Ishiguro, "Motion-Compensated Interframe Coding for Video Conference," in *Proc. NTC 81*, pp. C961-965, 1981.
- [11] J. R. Jain and A. K. Jain, "Displacement Measurement and Its Application in Interframe image Coding," *IEEE Trans. on Communications*, Vol.COM-33, pp. 1011-1014, 1985.
- [12] Th. Zahariadis and D. Kalivas, "A Spiral Search Algorithm For Fast Estimation Of Block Motion Vectors," *Signal Processing VIII, theories and applications. Proceedings of the EUSIPCO 96.*, Vol.2, pp. 1079-1082.
- [13] M.Ghanbari, "The Cross Search Algorithm for Motion Estimation," *IEEE Transactions on Communications*, Vol.38, No.7 pp. 950-3, 1990.
- [14] Kwon Moon Nam, Joon-Seek Kim, and Rae-Hong Park, "A Fast Hierarchical Motion Vector Estimation Algorithm Using Mean Pyramid," *IEEE Transactions on Circuits and Systems for Video technology*, Vol.5, No.4, pp. 344-351, 1995.



강 선 경

2000년 2월 원광대학교 전기·전자공학부(공학사)
2004년 2월 원광대학교 정보·컴퓨터교육학과(교육학석사)
2010년 2월 원광대학교 컴퓨터공학과(공학박사)

관심분야: HCI, 영상처리, 패턴인식, 임베디드시스템



김 영 운

2003년 2월 원광대학교 컴퓨터·정보통신공학부(공학사)
2005년 8월 원광대학교 컴퓨터공학과(공학석사)
2006년 2월~현재 원광대학교 컴퓨터 공학과 박사과정

관심분야: 영상처리, 컴퓨터비전, 패턴인식, HCI, 컴퓨터 그래픽스



이 옥 걸

2006년 8월 원광대학교 전기전자 및 정보통신 학부(공학사)
2010년 8월 원광대학교 컴퓨터 공학과 석사

관심분야: 영상처리, 컴퓨터비전, 패턴인식, 컴퓨터그래픽스



정 성 태

1987년 2월 서울대학교 컴퓨터공학과(공학사)
1989년 2월 서울대학교 컴퓨터공학과(공학석사)
1994년 8월 서울대학교 컴퓨터공학과(공학박사)

1995년 3월~현재 원광대학교 전기전자 및 정보공학부 교수

관심분야: 영상인식, 영상처리, 컴퓨터 그래픽스



송 원 창

2010년 2월 원광대학교 전기전자 및 정보통신 학부(공학사)
2010년 3월~현재 원광대학교 컴퓨터 공학과 석사과정

관심분야: 영상처리, 컴퓨터비전, 패턴인식, 증강현실, 임베디드시스템