

표정별 가버 웨이블릿 주성분특징을 이용한 실시간 표정 인식 시스템

Real-time Recognition System of Facial Expressions Using Principal Component of Gabor-wavelet Features

윤현섭 · 한영준 · 한현수*

Hyunsup Yoon, Youngjoon Han and HERNSEO HAHN*

* 송실대학교 전자공학과

요 약

표정은 인간의 감정을 전달할 수 있는 중요한 수단으로 표정 인식은 감정상태를 알아낼 수 있는 효과적인 방법중 하나이다. 일반적인 표정 인식 시스템은 얼굴 표정을 표현하는 특징점을 찾고, 물리적인 해석 없이 특징을 추출한다. 하지만 특징점 추출은 많은 시간이 소요될 뿐 아니라 특징점의 정확한 위치를 추정하기 어렵다. 그리고 표정 인식 시스템을 실시간 임베디드 시스템에서 구현하기 위해서는 알고리즘을 간략화하고 자원 사용량을 줄일 필요가 있다. 본 논문에서 제안하는 실시간 표정 인식 시스템은 격자점 위치에서 얻어진 가버 웨이블릿(Gabor wavelet) 특징 기반 표정 공간을 설정하고, 각 표정 공간에서 얻어진 주성분을 신경망 분류기를 이용하여 얼굴 표정을 분류한다. 제안하는 실시간 표정 인식 시스템은 화남, 행복, 평온, 슬픔 그리고 놀람의 5가지 표정이 인식 가능하며, 다양한 실험에서 평균 10.25ms의 수행시간, 그리고 87%~93%의 인식 성능을 보였다.

키워드 : 표정 인식, 표정특징추출, 가버 웨이블릿, 주성분분석

Abstract

Human emotion can be reflected by their facial expressions. So, it is one of good ways to understand people's emotions by recognizing their facial expressions. General recognition system of facial expressions had selected interesting points, and then only extracted features without analyzing physical meanings. They takes a long time to find interesting points, and it is hard to estimate accurate positions of these feature points. And in order to implement a recognition system of facial expressions on real-time embedded system, it is needed to simplify the algorithm and reduce the using resources. In this paper, we propose a real-time recognition algorithm of facial expressions that project the grid points on an expression space based on Gabor wavelet feature. Facial expression is simply described by feature vectors on the expression space, and is classified by an neural network with its resources dramatically reduced. The proposed system deals 5 expressions: anger, happiness, neutral, sadness, and surprise. In experiment, average execution time is 10.251 ms and recognition rate is measured as 87~93%.

Key Words : Facial Expression Recognition, Expression Feature, Gabor Wavelet, PCA

1. 서 론

인간과 로봇간의 의사소통은 최근 급성장하고 있는 개인용 서비스 로봇시장에서 지능형 로봇을 위한 필수적인 요소 중 하나이다. 기존 인간의 명령에만 따르던 로봇은 단방향적이라는 한계가 있었기 때문에 지능형 로봇은 좀 더 자연스럽게, 인간 친화적으로 의사소통할 수 있는 특성이 필요하게 되었다. 감정인식은 인간과 로봇간의 의사소통 방법 중 하나로 얼굴영상의 표정 인식을 통한 감정인식 뿐만 아니라, 음성신호를 이용한 감정인식, 생체신호를 이용한 감정인식 방법들이 연구되고 있다. 이중 음성신호를 이용한 감정인식의 경우는 감정 상태에 따른 단어, 음성신호의 톤

(Tone) 그리고 피치(Pitch)등 음성신호의 특성을 이용한 방법들이 연구되고 있다[1]. 그러나 음성신호의 경우 개개인의 특성, 사투리에 따라 특성이 달라지기 때문에 일반적인 방법을 사용하여 감정인식을 하기에는 어렵다[2]. 그리고 생체신호를 이용한 감정인식의 경우는 대표적으로 얼굴에 EMG(Electromyography)센서를 부착하고 표정에 따른 특정 근육의 움직임을 찾고 표정을 구분하는 연구가 진행되었다[3]. 하지만 얼굴에 센서를 부착함에 따라 특정 얼굴 근육을 움직이기 어렵거나 표정이 부자연스러워지는 것이 단점으로 지적되었다[4]. 이에 반해 얼굴 영상을 이용한 감정인식방법은 비전센서를 이용하여 얻어진 얼굴 영상으로부터 표정에 영향을 주는 근육의 움직임 특성을 찾고 표정을 분류하는 방법을 사용하여 앞서 설명된 방법들에 비해 인간 친화적이고 보다 정확한 감정인식 결과를 얻을 수 있다.

컴퓨터 비전을 이용한 표정 인식은 이용되는 특징에 따

접수일자 : 2009년 9월 1일

완료일자 : 2009년 11월 27일

라 크게 영상기반(image base)의 표정 인식과 모델기반(model base)의 표정 인식의 2가지 접근 방법으로 나누어 설명할 수 있다[5,6]. 우선 영상기반의 표정 인식은 얼굴 영상으로부터 표정을 잘 표현할 수 있는 이미지 화소값에 기반을 둔 특징들을 이용하게 된다. Padgett[7]는 얼굴영상으로부터 눈, 코, 입영역에 해당하는 부분을 분리하고 각 세부 영상에 PCA(Principal Component Analysis)를 적용해 얻은 특징을 이용하였다. Geng[8]는 PCA를 통해 얻은 결과에 LDA(Linear Discriminant Analysis)방법을 응용하는 연구를 진행하였다. 그리고 Matthew[9]는 얼굴영상에서 가버 웨이블릿(Gabor wavelets) 특징을 추출하고 여기에 PCA를 적용하여 얻은 결과들을 표정 인식의 특징으로 사용하였다. 마지막으로 Kim[10]은 확장된 사각특징을 표정 인식 시스템에 적용하여 표정 인식의 특징으로 사용하였다.

다음으로 모델기반의 특징을 사용한 표정 인식 방법은 얼굴에서 움직일 수 있는 액션 유닛(Action Unit, AU)에 기초한 특징점의 위치나 움직임에 기반한다. Yacoob[11]의 광학적 흐름(Optical Flow)분석방법은 찾은 얼굴근육 움직임을 표정 인식에 특징으로 사용하였다. Bouchra[12]는 AAM(Active Appearance Model)을 적용하여 특징점을 찾고, 미리 정의된 얼굴의 특징점간의 거리관계를 특징으로 하여 표정 인식에 적용하였다. 그리고 Cheon[13]은 AAM으로 찾아진 얼굴 특징점의 움직임을 특징으로 하여 표정 인식에 적용하였다. 마지막으로 Gokturk[14]는 회전에 강인하게 이용할 수 있는 3차원 얼굴모델의 구성, 트래킹을 통하여 움직임을 측정하고 표정 인식을 위한 특징으로 활용하였다.

기존의 표정 인식을 위한 방법들은 표정이 잘 나타나는 특징점의 위치를 찾고 특징점의 화소, 위치정보들을 특징벡터로 이용하였다. 따라서 이러한 표정 인식 시스템은 찾아진 특징점의 위치에 따라 인식 결과가 결정된다. 그리고 특징점의 정확한 위치를 찾는 것은 어렵고 시간 또한 많이 소요되는 단점을 지닌다.

본 논문에서는 빠르게 변화하는 표정의 변화까지 인식을 할 수 있는 실시간 표정 인식 시스템 구현을 목적으로 연구를 진행하였다. 따라서 본 연구는 위에서 단점으로 지적한 특징점의 정확한 위치를 찾는 과정을 간략화 하고 자원 사용량을 줄이기 위한 내용이 연구의 중심이 된다. 이를 위한 알고리즘의 기본적인 개념은 각 표정별로 입력 얼굴 영상의 격자점 위치에서 얻어진 가버 웨이블릿 특징을 사용하여 주성분을 표현할 수 있는 표정 공간을 구성하고, 각각의 표정 공간에서 얻어진 주성분 가버 웨이블릿 특징을 신경망에 이용하여 분류하는 방법을 제안하려 한다. 따라서 본 논문에서는 미리 구해진 표정 공간을 사용함에 따라 특징점의 정확한 위치를 추정하는 과정을 간략화 하고, 낮은 차원으로 구성된 표정공간의 주성분만을 이용하기 때문에 리소스의 사용 또한 줄여 실시간으로 표정을 인식하는 방법을 제안하게 된다.

논문의 구성은 2장에서 구현된 표정 인식 시스템의 구조를 설명한다. 다음 3장에서는 전처리 과정을 포함한 표정별 주성분 특징 추출방법에 대해 설명하며, 4장에서는 분류기로 이용할 신경망 구성에 대해 설명한다. 그리고 5장에서는 제안한 시스템의 수행시간 측정실험과 인식을 실험결과를 비교하여 제안하는 시스템의 실시간성 및 인식률의 우수성을 검증하였다. 마지막으로 6장에서는 본 연구의 결론을 제시한다.

2. 표정 인식 시스템 구조

제안하는 실시간 표정 인식 시스템의 구조는 그림 1과 같다. 표정의 특징을 추출하기 위해서 먼저 입력영상으로부터 얼굴영역의 검출 및 정규화하는 전처리 과정이 필요하다. 다음으로 정규화 과정을 마친 영상에서 가버 웨이블릿 특징을 추출하기 위해서 미리 정의된 격자점(165개)의 위치 정보와 가버 웨이블릿 커널(18개)들을 이용하여 특징을 추출하게 된다. 추출된 가버 웨이블릿 특징 벡터(2970차원)는 차원이 클 뿐 아니라 중복되는 불필요한 성분을 많이 포함하고 있기 때문에 각 표정별 데이터베이스에 의해 만들어진 주성분 기저(80개)들을 이용하여 표정별 주성분 가버 웨이블릿 특징을 추출한다. 마지막으로 신경망으로 구성된 분류기를 이용하여 화남, 행복, 평온, 슬픔 그리고 놀람의 5가지 표정을 구분하고 그 결과를 각 표정을 갖는 아바타 영상을 이용하여 표현한다.

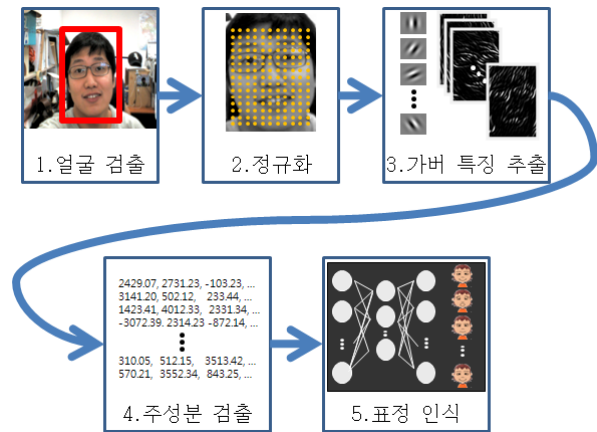


그림 1. 제안하는 표정 인식 시스템의 흐름도
Fig. 1. Flow chart of proposed FER system

3. 표정별 가버 웨이블릿 주성분 추출

3.1 전처리 과정

얼굴 표정의 특징을 추출하기 위한 사전 과정은 입력 영상으로부터 얼굴 영상을 검출하는 것이다. 본 표정 인식 시스템에서는 얼굴 검출을 위해 입력영상의 밝기값 정보에 기반을 둔 사각특징과 Adaboost 학습 알고리즘을 이용한 Viola와 Jones의 얼굴 검출 알고리즘[15]을 사용하였다. Adaboost[16]는 Freund가 제안한 학습 알고리즘으로 여러 개의 약분류기들로부터 강분류기를 구성하는 방법 중 하나이다. Viola와 Jones는 Adaboost 학습 알고리즘을 얼굴 검출에 적용하였고, 적분영상을 이용하여 효율적으로 추출된 사각특징과 cascade 구조를 사용하여 빠르고 높은 검출률을 보이는 얼굴 검출 시스템을 구현하였다.

얼굴 영역을 검출 후, 다음 과정으로 정확한 위치에서 얼굴 표정의 특징을 추출하기 위해 얼굴 영상의 크기를 정규화 하는 과정을 거치게 된다. 본 표정 인식 시스템에서는 얼굴 영역의 격자점 위치에서 추출한 특징을 이용하기 때문에 정확한 격자점을 구성하기 위해 크기 및 위치 정규화 과정이 필요하다.

기존의 얼굴 표정 특징은 AU(Action Unit)에 기반을 둔

위치에서 추출되었다. AU는 Ekman과 Friesen이 제안한 FACS(Face Action Coding System)에서 얼굴 표정의 변화를 일으키는 요소로 이들 AU들의 조합을 통해 전형적인 표정을 기술하였다[17]. 표정 인식 시스템을 위해서는 자동적으로 정확한 AU들의 위치를 찾는 것이 필요하다. 하지만 AU들의 위치를 찾는 것은 시간이 오래 걸릴 뿐 아니라 정확한 위치를 찾기 쉽지 않다.

그래서 제안하는 알고리즘은 격자점위치에서 특징을 추출하는 방법을 사용한다. 따라서 별도의 특징점 위치를 찾는 과정을 생략할 수 있고 수행 시간을 줄인다.

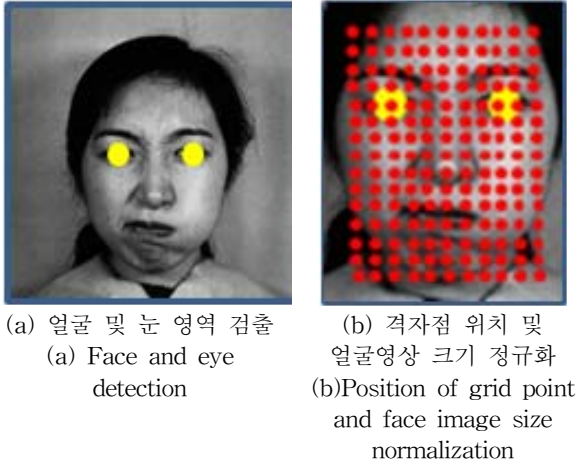


그림 2. 얼굴 영상 정규화
Fig. 2. face image normalization

격자점에 맞는 크기 및 위치 정규화를 하기 위한 첫 번째 과정으로 검출된 얼굴 영역으로부터 밝기값 정보에 기반을 둔 사각특징을 이용하여 두 눈의 위치를 찾는다.

다음 과정으로 두 눈의 위치가 찾아지면 두 눈 사이 거리에 기반을 두어 정해진 격자의 개수가 들어갈 수 있도록 격자의 간격을 결정하게 된다. 다음으로 격자의 간격이 결정되면 찾아진 눈이 특정 격자점에 위치할 수 있도록 영상을 이동하여 위치를 정규화하고, 영상이 일정한 크기를 갖으며 11x15크기를 갖는 격자가 구성될 수 있도록 얼굴영상의 크기를 조절하여 얼굴 영상의 크기가 정규화된 영상을 얻게 된다.

3.2 가버 웨이블릿 특징

가버 웨이블릿 특징은 가버 웨이블릿 커널과 입력영상의 컨벌루션을 통해 얻어진 계수로써 인간의 시각적 인식 체계를 잘 표현하고 있으며, 복잡한 질감에 대한 정보를 효과적으로 찾아낼 수 있다. 이는 가버 웨이블릿 특징이 특정 주파수, 특정 방향 그리고 공간적 지역성에 반응하는 특징을 이용한 것으로 가버 웨이블릿 커널의 변수를 조절함으로써 다양한 가버 웨이블릿 특징값을 얻어낼 수 있다. 그리고 가버 웨이블릿 특징은 이미지 화소기반의 방법과 비교하여 조명, 자세, 크기, 그리고 회전등 환경변화에 더 강인하기 때문에 얼굴 표정의 특징을 강인하게 추출할 수 있다[18]. 본 표정 인식 시스템에서 사용한 2차원 가버 웨이블릿 커널은 식 (1)과 같이 표현된다.

$$g(x,y;\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \psi\right) \quad (1)$$

여기서 x,y 는 가버 웨이블릿 커널에서 중심점을 원점으로 하는 위치를 나타내며, $x' = x \cos \theta + y \sin \theta$, $y' = -x \sin \theta + y \cos \theta$ 를 나타내고, θ 는 웨이블릿의 방향을 나타내고, λ 는 웨이블릿의 파장($\frac{1}{\text{주파수}}$)을 나타내며, ψ 는 웨이블릿의 이동된 정도, 그리고 σ 는 가우시안커널의 크기를 나타내게 된다.

본 시스템에서는 가버 웨이블릿 커널의 $\theta \in \{0, 30, 60, 90, 120, 150\}$ 와 $\lambda \in \{6, 8, 16\}$ 의 조합하여 총 18개의 가버 웨이블릿 커널을 만들었다. 그림 3은 만들어진 가버커널의 집합을 보여준다.

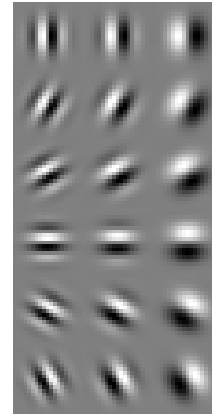


그림 3. 얼굴 표정 인식 시스템에 적용된 가버 웨이블릿 커널 영상
Fig. 3. The Gabor wavelet kernel image which is applied FER system

다음으로 가버 웨이블릿 특징은 위에서 만든 가버 웨이블릿 커널과 입력영상과의 컨벌루션을 통해 얻어질 수 있다. 본 표정 인식 시스템에서는 격자위치에서 얻어진 가버 웨이블릿 특징을 사용하므로 입력 영상의 모든 격자점에서 가버 웨이블릿 커널과 곱을 수행한다.

$$G(j,k) = \sum_{x,y} I_j(x,y) \times g_k(x,y) \quad (2)$$

식 (2)에서 $G(j,k)$ 는 j 번째 격자점 위치에서의 k 번째 웨이블릿 커널을 곱했을 때의 가버 웨이블릿 특징을 나타내며, x,y 는 각 가버 웨이블릿 커널의 너비, 높이에 해당하는 영역을 나타낸다.

그림 4는 정규화된 얼굴영상에 가버 웨이블릿 커널을 컨벌루션한 결과를 보여주며, 본 시스템에서는 전체 이미지가 아닌 격자점에서만 가버 웨이블릿 커널을 곱하여 얻은 값을 특징벡터로 이용하게 된다.

3.3 표정별 주성분 공간 추출

격자점에서 추출된 가버 웨이블릿 특징을 살펴보면 165개의 격자점에 각각 18개의 가버 웨이블릿 커널을 적용함에 따라 총 2970차원의 특징벡터가 생성된다. 이 특징벡터는 얼굴 표정에 상관없는 조명이나 얼굴 자세에 따른 많은 잡음이 포함되어 있기에 때문에 얼굴 표정을 구분하기에 민감하고, 또 고차원의 데이터를 사용함에 따라 복잡하고 처리 시간 또한 오래 걸리는 단점이 있다. 주성분 분석(Principal Component Analysis; PCA)은 이러한 고차원 데이터를 분

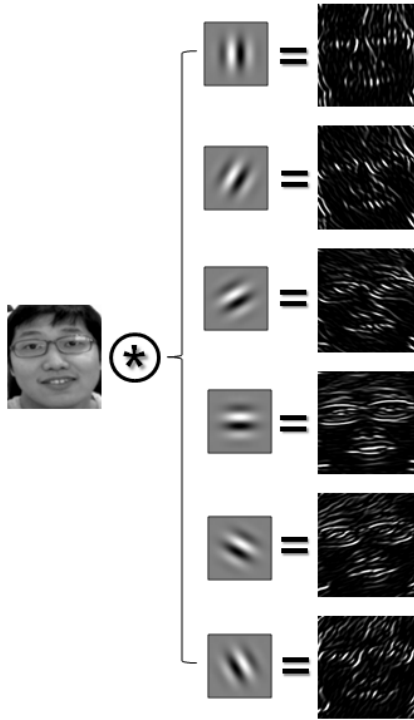


그림 4. 추출된 가버 웨이블릿 특징
Fig. 4. Extracted Gabor wavelet feature

석하는 방법 가운데 하나로 원하는 만큼의 축소된 주성분 데이터로 재구성할 수 있고, 고차원 데이터의 패턴인식에 사용되는 통계적 기법이다[19]. 주성분 분석은 데이터를 새로운 좌표축으로 투영하였을 때 최대 분산이 나올 수 있는 축, 즉 새로운 공간의 기저벡터들을 찾고 새로운 좌표공간으로 선형변환 하여 얻은 데이터를 주성분 데이터로써 사용하는 방법이다. 주성분 분석을 위해서는 통계적 정보인 평균, 공분산행렬, 그리고 고유벡터, 고유값등을 이용하게 된다. 첫 번째로 m 차원 공간의 데이터가 n 개 존재한다면 각 차원 데이터의 평균($\bar{X}_i = \frac{1}{n} \sum_{j=1}^n X_{ij}$)을 구한다. 여기서 i 는 데이터의 차원을, j 는 i 차원내 j 번째 데이터를 의미한다. 그리고 각 차원의 데이터에서 각 차원의 평균을 빼서 평균이 0인 벡터들로 구성된 데이터($X'_{ij} = X_{ij} - \bar{X}_i$)를 구할 수 있다. 다음으로 이 평균이 0인 데이터들을 이용하여 공분산 행렬($C_{xy} = \frac{1}{n-1} \sum_{j=1}^n (X'_{xj})(X'_{yj})$)을 계산 할 수 있다. 여기서 x 와 y 는 공분산 행렬에서 성분들의 위치를, X' 는 평균이 0인 행렬을 의미한다. 그리고 대수적 방법을 이용하여 앞서 구해진 공분산 행렬의 고유값 $\lambda_1, \lambda_2, \dots, \lambda_m$ (이때 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$)와 이에 대응하는 고유벡터 v_1, v_2, \dots, v_m 을 구하게 된다. 이때 고유값의 크기순으로 정렬된 각 고유벡터는 주성분을 공간을 구성하는 기저벡터가 된다. k 개의 고유벡터를 선택하여 얻은 주성분 행렬($K = [v_1, v_2, \dots, v_k]$)을 구성할 수 있고, 이 주성분 공간에서 표현된 k 차원의 데이터를 얻을 수 있다. 이때 입력데이터에서 편차의 비율을 $p\%$ 만큼 유지하기 위해서는 다음 식 (3)을 만족하는 벡터의 수 k 를 선택할 수 있다.

$$\sum_{t=1}^k \lambda_t \geq \frac{p}{100} \times \sum_{t=1}^m \lambda_t \quad (3)$$

본 시스템에서는 PCA방법을 활용하여 각 표정별 주성분 공간을 구성하고 얻어진 주성분 데이터를 최종 표정 특징으로 사용하였다. 표정별 주성분 공간을 구성하기 위해 5가지 표정에 대해서 각 표정별로 이전 과정에서 생성한 가버 웨이블릿 특징 벡터를 입력으로 하여 주성분 공간을 생성한다. 본 시스템에서는 2970차원의 각 표정별 입력데이터를 30개씩 사용하여 각 표정별 입력데이터 집합을 구성하였다. 그리고 입력데이터에서 편차의 비율을 95%만큼 유지하였을 때 각 표정별 16개의 주성분 벡터를 구하여 5가지 표정에 대하여 총 80개의 주성분 벡터로 구성된 주성분 행렬을 구성할 수 있다.

4. 표정 인식을 위한 신경망 구성

본 표정 인식 시스템에서는 표정 특징을 이용한 표정 인식을 위해 1개의 은닉층으로 구성된 다층 신경망을 사용하였다.

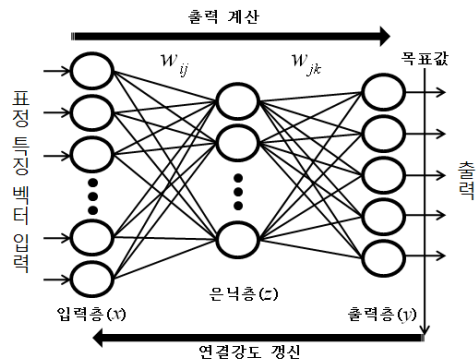


그림 5. 신경망 구조
Fig. 5. Structure of neural network

입력 노드의 수는 표정 특징의 차원수와 같이 80개로 구성되었으며, 은닉층은 40개의 노드로 구성, 그리고 출력 노드는 구분된 5가지 표정 표현하기 위해 5개의 노드로 구성되었다. 그리고 w 는 노드사이에 연결 강도를 의미하며, 최종 출력은 연결 강도(w)와 입력값(x) 그리고 활성화함수(f)에 의해 결정된다.

$$y_k = f\left(\sum_{j=1}^{40} f\left(\sum_{i=1}^{80} x_i w_{ij}\right) w_{jk}\right) \quad (4)$$

$$f(N) = \frac{1}{1 + e^{-N}} \quad (5)$$

위 결과값 식 (4)를 보면 x_i 는 i 번째 입력벡터값을, w_{ij} 는 i 번째 입력노드와 j 번째 은닉노드사이의 연결 강도를, 그리고 w_{jk} 는 j 번째 은닉노드와 k 번째 출력노드사이의 연결 강도를 의미한다. 활성화함수로 적용된 식 (5) $f(N)$ 은 비선형 단극성 단조증가 함수이고 0에서 1사이 출력을 갖는 시그모이드 함수를 사용하였다. 그리고 연결강도 w 는 순방향 구조의 다층 신경망에서 학습에 사용할 수 있는 역전파 알

고리즘(Backpropagation)을 이용해 계산하였다.

5. 실험 결과 및 고찰

제안하는 얼굴 표정 인식 시스템은 C++언어로 구현하였으며, 2.4GHz의 CPU기반의 PC에서 실험되어졌다.

실험에 사용된 얼굴 영상 데이터베이스는 CAS-PEAL (Chinese Academy Sciences-Pose, Expression, Accessories, and Lighting) DB[20], JAFFE(Japanese Female Facial Expression) DB[21], PF01(Postech Faces' 01) DB[22], 그리고 자체 제작된 SSU FE(SoongSil University Face Expression) DB로 구성되었다.

CAS-PEAL DB에서는 중국인을 모델로 하여 얼굴의 자세, 표정, 장신구, 그리고 조명 변화를 준 데이터베이스로 이중 표정에 관련된 영상만을 사용하였다. JAFFE는 일본 여성을 모델로 하여 7가지 표정에 대한 영상을 포함하고 있다. PF01 DB의 경우는 아시아인을 모델로 하여 조명, 자세, 표정 변화를 포함하고 있으며 이중 4가지 표정 변화에 대한 영상을 사용 하였다. 마지막으로 SSU FE DB는 한국인을 대상으로 7가지 표정에 대한 영상을 포함하고 있다. 실험을 위해 구성된 데이터 베이스 영상은 각 표정별로 200장씩, 5가지 표정에 대해 총 1000장의 표정 영상으로 구성하였다. 아래 그림 6은 사용된 데이터 베이스 영상의 예를 보여준다.



(a) 화남
(a) Anger



(b) 행복
(b) Happiness



(c) 평온
(c) Neutral



(d) 슬픔
(d) Sadness



(e) 놀람
(e) Surprise

그림 6. 얼굴 표정 데이터베이스 표본
Fig. 6. Facial Expression Sample DB

제안하는 얼굴 표정 인식 시스템의 성능을 평가하기 위해서 첫 번째로 표정 인식에 걸리는 수행시간을 측정하였다. 그리고 다음 평가 방법으로는 각 표정별 인식 결과를 바탕으로 인식률을 비교하였다.

첫 번째 실험은 표정 인식 알고리즘의 수행시간 측정 실험으로 전처리과정에서 얼굴 영역이 검출된 데이터베이스 영상으로 부터 눈 검출, 정규화, 얼굴 표정 특징 추출, 그리고 신경망을 이용한 표정 분류에 이르는 과정에서 소요되는 시간을 측정 하였다. 정규화된 얼굴 영상의 크기는 160x200을 갖으며 500개 영상에서 수행한 결과의 평균으로 수행시간을 계산하였다. 표 1은 표정 인식의 각 과정에서의 수행시간을 보여준다.

표 1. 얼굴 표정 인식의 단계별 수행 시간(단위:ms)
Table 1. Each Stage Execution time of FER(Unit:ms)

과정		시간
전처리	눈 검출	5.702
	정규화	0.010
특징 추출 및 구성		4.504
표정 분류		0.034
전체		10.251

수행시간 실험 결과에 따라 얼굴영역이 검출되었을 경우 평균 10.251ms에 표정 인식 과정이 수행되었고, 따라서 카메라에서 입력되는 영상에 대해 충분히 실시간으로 표정을 인식할 수 있는 것을 보여준다.

두 번째 실험은 표정 인식 알고리즘의 표정 인식률에 대한 실험으로 실험 환경은 다음과 같다. 첫 번째로 표정 공간을 구성하기 위해서 5가지 표정에 대해 각 표정당 30장의 영상을 이용하여 표정의 공간을 구성하였다. 그리고 각 표정당 100장의 영상을 신경망의 연결강도 학습을 위해 이용하였다. 마지막으로 나머지 표정당 100장씩의 이미지를 이

용하여 표정 인식을 실험에 사용하였다. 각 표정에 대한 표정 인식 결과는 표 2와 같다.

표 2. 표정 인식 결과
Table 2. Result of facial expression recognition

		인식 결과				
입 력		화남	행복	평온	슬픔	놀람
	화남	87	5	5	3	0
	행복	2	90	5	1	2
	평온	4	1	93	2	0
	슬픔	5	2	4	88	1
	놀람	2	1	3	2	92

5가지 표정에 대한 표정 인식 결과를 살펴보면 평균 90%의 인식률을 보였다. 평온은 93%로 가장 높은 인식률을 보인 반면 바른 표정으로 부터의 오인식을 또한 높게 나타났다. 표정의 특징이 잘 들어나는 놀람, 행복이 다음으로 높은 인식률을 보였으며 놀람 같은 경우는 다른 표정으로 부터의 오인식을 또한 낮게 나타나는 결과를 보였다. 다음으로 슬픔과 화남의 경우도 이전 표정들보다 약간 낮은 인식률을 보였으나 평균인식률에서 크게 벗어나지 않는 인식률을 결과를 볼 수 있었다.

6. 결 론

본 논문에서는 실시간으로 동작하는 표정 인식 시스템 구현을 목표로 연구를 진행하였다. 기존의 가버 웨이블릿 특징 기반의 얼굴 표정 인식 방법들은 얼굴의 주요 특징점을 찾아야 했기 때문에 특징점 위치를 찾는데 시간이 많이 걸리고, 특징점 위치에 따라 인식률의 결과가 예민해지는 단점을 가지고 있었다. 본 논문에서는 격자점에서 얻어진 가버 웨이블릿 기반의 표정별 공간에서 얻어진 특징과 신경망을 이용하여 빠르고 효과적인 표정 인식 시스템을 제안하였다. 표정 인식 단계의 수행시간 실험을 통하여 실시간으로 표정 인식 시스템의 수행을 보여주었다. 그리고 표정 인식을 실험을 통해 평균 90%의 높은 인식률을 결과를 얻을 수 있었고, AU에서 얻어진 특징과 인식률 비교를 통해 제안하는 가버 웨이블릿 주성분 특징의 우수성을 증명하였다.

참 고 문 헌

[1] 심귀보, 박창현, "음성인식으로 감정인식 요소 분석," *퍼지 및 지능 시스템학회 논문지*, 11권, 11호, pp.510-515, 2001.
 [2] 고현주, 이대중, 박장환, 전명근, "웨이블릿 변환을 이용한 음성에서의 감정 추출 및 인식 기법," *퍼지 및 지능 시스템학회 논문지*, 14권, 2호, pp.150-155, 2004.
 [3] Andreas Haag, Silke Goronzy, Peter Schaich and Jason Williams, "Emotion Recognition Using Bio-sensors: First Steps towards an Automatic System," *In Proceedings of LNCS*, Vol.3068, pp.36-48, 2004.

[4] Shishir Bashyala, Ganesh K. Venayagamoorthy, "Recognition of facial expressions using Gabor wavelets and learning vector quantization," *Engineering Applications of Artificial Intelligence*, Vol.21, No.7, pp.1056-1064, 2008.
 [5] Pantic M, Rothkrantz L.J.M, "Automatic analysis of facial expressions: the state of the art," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.22, No.12, pp.1424-1445, 2000.
 [6] B. Fasel, J. Luetttin, Beat Fasel, Juergen Luetttin, "Automatic Facial Expression Analysis: A Survey," *Pattern Recognition*, Vol.36, pp.259-275, 1999.
 [7] C. Padgett, G.W. Cottrell, "Representing face images for emotion classification," *Proc. Conf. Advances in Neural Information Processing Systems*, Vol.9, pp.894-900, 1997.
 [8] Geng Xue, Zhang Youwei, "Facial Expression Recognition Based on the Difference of Statistical Features," *International Conference on Signal Processing 2006*, Vol.3, pp.16-20, 2006.
 [9] Matthew N. Dailey, Garrison W. Cottrell, Curtis Padgett, and Ralph Adolphs, "EMPATH: A Neural Network that Categorizes Facial Expressions," *Journal of cognitive neuroscience*, Vol.14, pp.1158-1173, 2002.
 [10] 정성욱, 김도형, 안광호, 정명진, "실시간 얼굴 표정 인식을 위한 새로운 사각 특징 형태 선택기법", *제어자동화시스템공학회지*, 12권, 2호, pp.130-137, 2006.
 [11] Michael J. Black, Yaser Yacoob, "Recognizing Facial Expressions in Image Sequences Using Local Parameterized Models of Image Motion," *International Journal of Computer Vision*, Vol.25, No.1, pp.23-48, 2004.
 [12] Bouchra Abboud, Franck Davoine, Mo Dang, "Facial expression recognition and synthesis based on an appearance model," *Signal Processing: Image Communication*, Vol.19, No.8, pp.723-740, 2004.
 [13] Yeongjae Cheon, Daijin Kim, "A Natural Facial Expression Recognition Using Differential-AAM and k-NNS," *Pattern Recognition*, Vol.42, No.7, pp.1340-1350, 2008.
 [14] Gokturk. S.B, Bouguet. J.-Y, Tomasi. C, Girod. B, "Model-based face tracking for view-independent facial expression recognition," *IEEE International Conference on Automatic Face and Gesture Recognition*, pp.287-293, 2002.
 [15] Viola P., Jones M., "Rapid object detection using a boosted cascade of simple features," *Computer Vision and Pattern Recognition 2001*, Vol.1 pp.511-518, 2001.
 [16] Yoav Freund, Robert E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *In Computational Learning Theory: Eurocolt' 95*, pp.23-37, 1995.

- [17] P. Ekman, W. Friesen, "Facial Action Coding System: A Technique for the Measurement of Facial Movement," *Consulting Psychologists Press*, 1978.
 - [18] Jie Zou, Qiang Ji, G. Nagy, "A Comparative Study of Local Matching Approach for Face Recognition," *IEEE Transactions on Image Processing*, Vol.16, No.10, pp.2617-2628, 2007.
 - [19] I.T. Jollie, "Principal Components Analysis," *Springer-verlag New York*, 1986
 - [20] The Chinese Academy Sciences-Pose, Expression, Accessories, and Lighting(CAS-PEAL) Database <http://www.jdl.ac.cn/peal/index.html>
 - [21] The Japanese Female Facial Expression(JAFFE) Database <http://www.kasrl.org/jaffe.html>
 - [22] The Postech Faces' 01 Database <http://nova.postech.ac.kr/special/imdb/imdb.html>
-

저 자 소 개



윤현섭(Yoon Hyunsup)

2008년 : 숭실대학교 정보통신전자공학부
(공학사)

2008년~현재 : 숭실대 대학원 전자공학과
석사과정

관심분야 : 영상처리, 얼굴표정인식

Phone : 02-821-2050

E-mail : beat83@ssu.ac.kr



한영준(Han Youngjoon)

제 14권 7호(2004년 12월호) 참조



한헌수(Hahn Hernsoo)

제 13권 4호(2003년 8월호) 참조