

# 잡음제거와 주요성분분석에 의한 효과적인 얼굴인식

## An Efficient Face Recognition by Denoising and Principal Component Analysis

조용현, 홍성준\*

대구가톨릭대학교 컴퓨터정보통신공학부

Yong-Hyun Cho, Seung-Jun Hong

School of Computer and Information Comm. Eng., Catholic Univ. of Daegu

E-mail : yhcho@cu.ac.kr

### 요 약

본 논문에서는 잡음제거와 주요성분분석을 이용한 효과적인 얼굴인식 기법을 제안하였다. 여  
기서 잡음제거는 필터링과 1차 모멘트 평형이동을 조합하여 영상의 특징정보와 관계가 없는  
배경을 제거함을 위한 것이고, 주요성분분석은 얼굴영상의 주요성분인 2차원의 중복성분이 제  
거된 특징을 효과적으로 추출하기 위함이다. 제안된 기법을 768\*576 픽셀 크기를 갖는 24개의  
AR 얼굴영상을 대상으로 시뮬레이션한 결과, 제안된 얼굴인식이 잡음제거를 하지 않은 기존의  
얼굴인식에 비해 주요성분의 개수에 따른 압축성능, 특징추출 시간, 그리고 city-block,  
Euclidean, negative angle(cosine)의 거리척도에 따른 인식에 있어서 보다 우수한 성능이 있  
음을 확인할 수 있었다.

### 1. 서론

정보의 보호나 보안을 위한 방안으로 생체정보의 활  
용에 대한 연구들이 활발하게 진행되고 있다. 그 중에서  
도 얼굴인식 분야는 지문이나 홍채 등을 이용한 기법들  
보다 사용자와 친화적이면서도 데이터 획득의 용이성으  
로 많은 연구들이 이루어지고 있다[1,2]. 하지만 입력 영  
상으로부터 해당 얼굴을 출력하는 얼굴인식에서는 조명이나  
배경 등의 환경의 복잡성과 안경이나 표정의 변화  
등에 따라 성능이 좌우되는 제약이 있다[3-6].

일반적으로 얼굴인식에는 특징기반 방법과 영상기반  
방법이 있다[1,2]. 특징기반 방법은 얼굴특징과 관련된  
정보들을 이용하는 방법으로 지역 특징 추출기법, 경계  
선 지도기법, Garbor 필터를 이용한 기법 등이 있다. 영  
상 기반 방법은 입력 얼굴영상에서 직접 얼굴의 코딩을  
얻어내는 외형을 기반으로 하며, 여기에는 입력 영상내  
의 상관관계를 제거하여 영상의 진폭 스펙트럼을 획득  
하는 주요성분분석(principal component analysis :  
PCA)과 PCA를 일반화하여 영상의 위상 스펙트럼까지  
도 고려하는 독립성분분석(independent component  
analysis : ICA), support vector machine(SVM) 등이  
있다[2-5]. 영상기반 방법이 주로 얼굴인식을 위해 이  
용되고 있으며 이에 대한 연구들이 활발하게 진행되고  
있다.

한편 입력으로부터 자동으로 출력패턴을 생성하는 기

법인 PCA와 ICA는 얼굴인식에 널리 이용되고 있다  
[2-5]. 이 방법들은 중복신호 제거의 효과가 있어 뇌의  
시각 기능을 잘 반영하며 타 방법에 비해 높은 인식률  
을 얻을 수 있는 장점을 갖고 있다. PCA가 신호의 차  
원 자체를 줄여 조밀한 코딩을 유도하는데 비해 ICA는  
출력 신호의 값을 드물게 나타내는 방법으로 중복 신호  
를 제거하며 이는 뇌의 기능 모델에 좀 더 관련이 있다.  
얼굴인식에 ICA를 적용하면 PCA와 달리 국부적인 특  
징을 추출할 수 있으며 PCA 보다 높은 인식률을 얻을  
수 있다고 알려져 있다[4]. 하지만 얼굴 인식에서 반드시  
ICA가 PCA보다 우수한 인식성능을 가지는 것은 아  
니며, 계산이 복잡한 제약도 있다.

본 논문에서는 영상의 특징정보와 거리가 먼 배경잡  
음을 제거하기 위한 방법으로써 Median 필터링[6]과 1  
차 모멘트 평형이동[7]을 조합한 전처리와 주요성분분  
석을 이용한 효과적인 얼굴인식 기법을 제안한다. 여기  
서 Median 필터링과 1차 모멘트는 영상의 배경잡음 제  
거와 배경을 제거하여 유효영상만을 추출하기 위함이고,  
주요성분분석은 얼굴영상의 주요성분인 2차원 통계  
성의 중복성분이 제거된 특징을 추출하기 위함이다. 제  
안된 기법을 AR 얼굴영상 데이터베이스[8]로부터 선택  
된 24개의 768\*576 픽셀의 영상을 대상으로 시뮬레이션  
하고 주요성분의 개수에 따른 압축성능과 city-block,  
Euclidean, 그리고 negative angle의 거리척도에 따른 인

식을 위한 분류성능을 비교·검토하였다.

## 2. 잡음제거와 주요성분분석

### 2.1 잡음제거

얼굴인식은 크게 얼굴영상 획득과정, 전처리 과정, 주요성분 추출과정, 인식을 위한 식별과정으로 이루어진다 [1,2]. 전처리 과정은 인식성능과 효율을 높이기 위한 과정으로 일반적인 기하학 정보들에 대한 정규화 방식과 화소들에 직접적인 효과를 갖는 여러 가지 필터링 기법들이 있다[4,5]. 이들 중 가장 기본적이면서도 인식성능에 큰 영향을 미치는 기법은 기하학적 정보를 이용한 것이다. 이는 인식을 위한 특징추출 등의 계산 부하를 줄여 실시간 처리를 가능하게 한다.

영상은 보통 배경과 의미 있는 형상으로 구분할 수 있다. 배경이 형상에 비해 비교적 변화가 적을 경우, 주어진 영상의 크기로부터 보다 작은 크기의 유효한 영상을 추출할 수 있다. 본 논문에서는 Median 필터링[6]을 통한 1단계 배경잡음의 제거에 이어 차 영상의 1차 모멘트 평행이동[7]을 사용하여 유효영상만을 추출하는 전처리 기법을 제안한다.

일반적으로 Median 필터는 영상신호에 적용적인 잡음제거 방법으로 비선형의 특성을 가지며, 영상의 특징을 가장 잘 나타내는 경계면을 그대로 보존하는 특성을 가진다. 이러한 속성을 가지는 Median 필터로 본 연구에서는 먼저 얼굴영상의 배경에 포함된 잡음을 제거하였다.

한편 입력되는 얼굴영상에서 배경을 제외한 유효영상만을 얻기 위하여 1차 모멘트 평행기법을 이용한다. 즉, 인식을 위해 입력되는 일반적인 얼굴영상을 x축 또는 y축으로 평행이동 시키기 위하여 차 영상의 1차 모멘트 평행기법을 사용한다. 1차 모멘트를 구하는 과정은 먼저 Median 필터링을 통한 1단계 배경잡음 제거과정을 거쳐 P(n) 화소로 구성된 원 영상으로부터 차 영상(DeltaImage)을 생성한다. 여기서 n번째 좌표의 차 영상은  $\Delta Image(n) = P(n) - P(n-1)$ 로 계산하며, y행 화소 좌표로부터 x열 화소 좌표 방향으로 계산된다. 구해진 차 영상으로부터 유효한 영상의 화소만을 생성하기 위하여 논리필터링을 수행하는데, 이때 논리필터링은 차 영상에 대한 화소의 변화 평균값을 구하고, 그 값의 5% 값을 유효한 변화의 판별을 위한 임계값으로 설정하였다. 차 영상의 값이 임계값보다 크거나 같으면 유효한 화소로 1의 값을 갖도록 하며, 그렇지 않으면 0의 값을 갖도록 함으로써 균일한 형상의 영상을 생성한다. 논리필터링된 영상은 다음의 1차 모멘트 평행으로 중심이동을 시킨다.

그림 1은 임의의 영상으로부터 1차 모멘트 평행을 설명한 것이다. 1차 모멘트 평행에 의해 각 개별 화소에 대한 모멘트 값의 합은 영상 전체의 중심거리와 영상 전체의 화소 값의 곱과 같다.

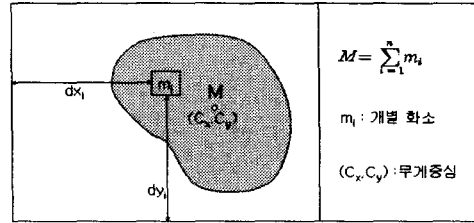


그림 1. 1차 모멘트 평형

그림 1로부터 유도되는 1차 모멘트 평행법칙을 이용한 중심계산식은 식 (1)과 같다.

$$M \times C_x = \sum_{i=1}^n (m_i \times dx_i)$$

$$M \times C_y = \sum_{i=1}^n (m_i \times dy_i) \quad (1)$$

여기서  $m_i$  값은  $i$ 번째 픽셀의 논리값으로 유효한 화소일 경우 1이며,  $M$ 은 유효영상의 논리값에 대한 합계값이다. 이때 유효영상의 중심좌표  $(C_x, C_y)$ 는 식 (2)와 같이 계산될 수 있다.

$$C_x = \frac{\sum_{i=1}^n dx_i}{M}, \quad C_y = \frac{\sum_{i=1}^n dy_i}{M} \quad (2)$$

따라서 임의의 위치에 입력되는 영상을 식 (2)에 의해서 얻어진 중심좌표로 이동시켜 유효한 영상을 선택하면 원영상보다 차원이 감소된 영상을 얻을 수 있다. 이렇게 함으로써 영상의 특징추출을 위해 이용되는 PCA의 입력데이터를 줄일 수 있다.

### 2.2 주요성분분석

주요성분분석은 데이터의 차원을 감소시키는 선형 변환 기법으로 널리 이용되고 있는 2차원 방법이다[2-4]. 이는 가우스 분포의 입력데이터의 특징을 추출하는 기법으로 데이터 내에 포함된 정보를 추출하고 압축하여 통계적 규칙들을 찾아내는 것이다. 결국 대용량의 입력 데이터를 통계적 독립인 특징들의 집합으로 변환시키는 것이며, n차원 입력공간의 데이터를 m차원 출력공간의 데이터로 투영시키는 것이다. 여기서 m (n이면 입력 데이터 벡터가 가지는 대부분의 내부정보를 유지면서도 차원의 감소가 가능하게 된다).

자기상관행렬  $R_{xx} = \langle \mathbf{x}\mathbf{x}^T \rangle$ 를 가진 평균이 영인 입력벡터  $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ 에 대해서 생각해 보자. 여기서 T는 전치를 나타내며,  $\langle \cdot \rangle$ 는 기대치를 나타낸다. 또한  $\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2, \dots, \hat{\mathbf{w}}_m$ 이  $R_{xx}$ 의 고유벡터와 직교되는 연결가중치 벡터라 할 때,  $\hat{\mathbf{w}}_1 = [\hat{w}_{11}, \hat{w}_{12}, \dots, \hat{w}_{1n}]^T$ 는 가장 큰 고유치  $\lambda_1$ 과 일치하며,  $\hat{\mathbf{w}}_2 = [\hat{w}_{21}, \hat{w}_{22}, \dots, \hat{w}_{2n}]^T$ 는 두 번째로 큰 고유치  $\lambda_2$ , 그리고  $\hat{\mathbf{w}}_n = [\hat{w}_{n1}, \hat{w}_{n2}, \dots, \hat{w}_{nn}]^T$ 는 가장 작은 고유치  $\lambda_n$ 과 각각 일치한다. 이상의 관계를 행렬방정식으로 나타내면 식 (3)과 같다.

$$R_{xx}\hat{w}_j = \lambda_j \hat{w}_j, (j=1,2,\dots, n) \quad (3)$$

여기서  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$  이다. 주어진 입력 벡터  $x$ 를 위한 첫 번째  $m$ 개의 주요 특징을 나타내는 고유벡터  $y$ 는 다음의 선형변환식 (4)로 나타낼 수 있다.

$$y = \hat{W}x \quad (4)$$

여기서  $\hat{W} = [\hat{w}_1, \hat{w}_2, \dots, \hat{w}_m]^T \in R^{m \times n}$  이며, 이 식에서 연결가중치행렬  $\hat{W}$ 의 행은 가장 큰 고유치와 일치하는 상관행렬  $R_{xx}$ 의 고유벡터임을 의미한다. 다시 말하면, 입력데이터 공간에서  $m$ 차원의 주요특징을 나타내는 부공간은  $R_{xx}$ 의  $m$ 개 주요 고유벡터에 의해 구성된 부공간으로 정의된다. 결국 PCA는  $\langle \|\hat{w}_j^T x\|^2 \rangle$ 가 최대인 고유벡터  $\hat{w}_1, \hat{w}_2, \dots, \hat{w}_m$ 의 방향을 찾는 것이다.

그림 2는 본 논문에서 제안된 Median 필터링과 1차 모멘트 평형이동이 조합된 전처리 과정과 PCA를 이용한 학습과 시험의 얼굴 인식과정을 나타낸 것이다. 여기서 영 평균(zero-mean)은 입력되는 얼굴영상 데이터의 실제 값에서 평균값을 뺀 차 영상 데이터를 만드는 것이다. 이는 영상내의 1차적 통계성을 고려하여 계산을 간단하게 하기 위함이다. 고유치와 고유벡터의 계산은 계산의 복잡성을 줄이기 위하여 snapshot 방법을 이용하였다. 또한 정규화된 영상데이터를 내림차순으로 나열한 후 얼굴의 특징벡터인 고유벡터의 개수를 정하여 고유벡터를 추출한다. 이때 계산을 간단하게 하기 위하여 영상의 차원을 감소시킨다. 한편 시험얼굴의 인식을 위해 시험영상으로부터 학습영상의 평균값을 뺀 차 영상인 영 평균 영상을 구하고, 마찬가지로 이들로부터 고유벡터를 추출한다. 마지막으로 정합을 위해 학습과 시험 얼굴영상 각각의 고유벡터 상호간을 city-block, Euclidean, 그리고 negative angle(cosine) 각각의 거리 척도를 이용하여 유사성을 판단한다. 결국 시험얼굴에 대한 유사성이 가장 많은 즉 거리가 가장 가까운 학습 영상이 추출되어 인식되게 된다.



그림 2 제안된 얼굴인식과정

### 3. 시뮬레이션 결과 및 고찰

제안된 기법의 성능을 평가하기 위해 768\*576 픽셀의 24개(12명\*2개) AR 얼굴영상[8]을 대상으로 실험하였다. 여기서 12개의 영상은 학습영상으로 나머지 12개는 salt & pepper 형태의 랜덤잡음[6] 20%와 표정이 다른 시험 영상이다. 특징추출에 대한 복원성능을 평가하기 위해 원 영상의 픽셀 값  $x_i$ 와 복원된 영상의 픽셀 값  $y_i$  사이의 절대평균합오차 amse(absolute mean sum error)  $= (1/N) \sum_{i=1}^N |x_i - y_i|$ 를 정의하였으며, 여기서 N은 전체 픽셀수이다. 실험은 펜티엄 IV-2.2G 컴퓨터상에서 Matlab 5.1로 구현하였다.

그림 3과 4는 각각 인식을 위해 학습에 이용된 보통의 표정을 가진 학습 얼굴영상들과 salt & pepper 형태의 잡음이 반영된 다양한 표정의 시험 얼굴영상들이다. 그림 5는 Median 필터링과 차 영상의 1차 모멘트 평형이동을 통하여 영상의 크기가 384\*288 픽셀로 감소되고 중심이동의 전처리된 학습 얼굴영상들이다.



그림 3. 학습 얼굴영상

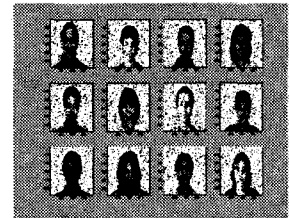


그림 4. 시험 얼굴영상

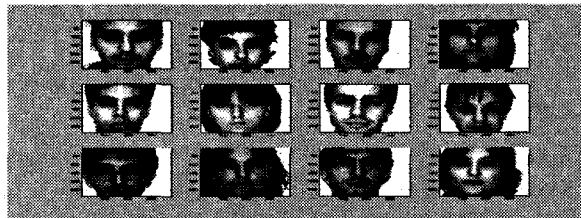


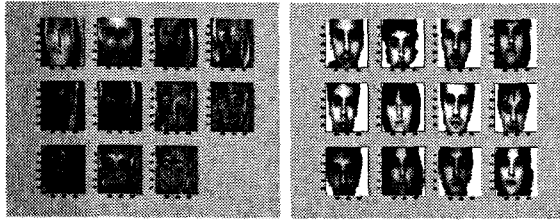
그림 5 전처리된 학습 얼굴영상

표 1은 1차 모멘트 평형이동에 의해 전처리된 영상과 그렇지 않은 영상을 대상으로 기저벡터의 개수에 따른 복원성능을 절대평균합오차 amse로 나타낸 것이다. 여기서는 2가지 경우의 영상 모두에서 기저벡터인 기저영상의 수가 증가함에 따라 amse로 표시되는 복원성능은 개선되다 다시 저하됨을 알 수 있다. 이는 기저영상의 개수인 특징벡터 수의 지나친 증가는 오히려 과학습을 일으켜 복원성능이 저하되기 때문으로 추측된다. 본 연구에서는 기저영상의 수를 복원성능이 가장 우수한 11개일 때를 기준으로 실험하였다.

표 1. 기저영상의 개수에 따른 절대평균합오차 amse

기저영상 수	1	2	3	4	5	6	7	8	9	10	11	12
amse (전처리)	2.928e-012	2.300e-012	2.094e-012	1.813e-012	1.547e-012	1.336e-012	1.076e-012	0.822e-012	0.571e-012	0.327e-012	1.423e-012	0.713e-012
amse (비전처리)	1.913e-012	1.690e-012	1.470e-012	1.325e-012	1.147e-012	0.968e-012	0.825e-012	0.659e-012	0.427e-012	0.275e-012	1.851e-012	0.446e-012

그림 6은 가장 우수한 복원성능을 가진 전처리된 11개의 기저영상과 그로부터 복원된 12개의 학습얼굴들을 각각 나타낸 것이다. 기저영상은 12개의 학습 얼굴영상으로부터 11개의 특징영상을 추출한 것으로 차원감소에 따른 데이터의 압축효과를 갖는다. 육안으로 볼 때, 그림 6(a)의 기저영상으로부터 복원된 그림 6(b)의 복원된 학습 얼굴영상은 그림 5의 전처리된 원 얼굴영상 매우 유사함을 알 수 있다. 이때 12개의 영상 모두에 대해 복원오차는 표 1에서 나타난 1.4283e-012 정도로 매우 작은 값을 가진다.



(a) 11개의 기저영상 (b) 12개의 복원영상  
그림 6. 기저영상과 복원된 학습 얼굴영상

표 2는 Median 필터링과 1차 모멘트 평형으로 전처리된 시험영상과 그렇지 않은 시험영상 각 12개에 대한 인식처리 결과를 나타낸 것이다. 표에서 전처리된 영상은 그렇지 않은 영상에 비해 특징추출에 이용되는 영상의 픽셀수가 약 75% 줄었으며, 그에 따라 CPU 시간으로 나타낸 특징 추출시간도 약 5배 정도 개선됨을 알 수 있다. 분류척도로서 city-block, Euclidean, negative angle의 거리를 각각 이용하였을 때 전처리 영상은 그렇지 않은 영상에 비해 city-block과 negative angle은 약 86%, Euclidean은 약 75%정도의 픽셀수의 감소가 있더라도 동일한 인식률을 가짐을 알 수 있다. 이는 Median 필터링과 1차 모멘트 평형이동의 전처리를 수행함으로써 작은 크기의 유효영상 추출이 가능하며, 이에 따라 상대적으로 보다 빠르고 정확한 영상의 특징추출과 인식이 가능함을 의미한다.

표 2. 시험영상의 인식 결과

구분	픽셀수	인식률(%)			추출시간 (Sec)
		city-block	Euclidean	negative angle	
전처리된 시험영상	768x576	91.7	91.7	91.7	3.1250
	690x518	91.7	91.7	91.7	2.1710
	614x460	91.7	91.7	91.7	1.5940
	538x402	91.7	91.7	91.7	1.2810
	460x346	91.7	91.7	83.3	0.8750
	384x288	91.7	91.7	83.3	0.6250
	306x230	83.3	83.3	83.3	0.3900
미전처리된 시험영상	768x576	83.3	91.7	83.3	3.1250

이상의 실험결과에서 제안된 Median 필터링과 1차 모멘트 평형을 통한 전처리와 PCA를 이용한 얼굴인식

기법은 보다 작은 크기의 유효영상 추출에 따른 빠른 특징추출이 가능하고, 학습 얼굴영상에 대해서는 100%의 인식률을 가지며, 잡음이 포함된 새로운 시험 얼굴영상에 대해서는 특징추출 픽셀수가 약 75% 감소되더라도 약 91.7% 이상의 높은 인식률이 있음을 확인하였다. 특히 얼굴영상의 인식을 위한 분류척도로 city-block과 Euclidean 거리를 이용하는 것이 우수한 인식 성능이 있음을 알 수 있다.

### 3. 결론

본 논문에서는 Median 필터와 차 영상의 1차 모멘트 평형이동을 통한 전처리와 수치기법의 주요성분분석을 이용한 얼굴인식을 제안하였다. 여기서 Median 필터는 배경잡음 제거와 1차 모멘트 평형이동은 보다 작은 크기의 유효영상을 추출하기 위함이고, 주요성분분석은 얼굴영상의 주요성분인 2차원 통계성의 중복성분이 제거된 특징을 추출하기 위함이다.

제안된 기법을 768\*576 픽셀 크기를 갖는 24개의 일반 얼굴영상을 대상으로 시뮬레이션 한 결과, 전처리를 하지 않은 단순한 PCA만에 의한 인식보다 빠른 특징추출이 가능하였고, 학습 얼굴영상의 경우 100%의 인식률을 보였으며, 시험 얼굴영상에 대해서는 특징추출 픽셀수가 약 75% 감소되더라도 약 91.7% 이상의 높은 인식률을 갖는 사실을 확인하였다. 특히 분류를 위한 거리척도로 city-block과 Euclidean 거리를 이용하는 것이 가장 우수한 인식률이 있음도 확인하였다.

향후 제안된 방법을 좀 더 큰 규모의 문제와 다양한 분야에의 응용에 대한 연구가 계속 진행되어야 할 것이다.

### 5. 참고문헌

- [1] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," Proceedings of IEEE., Vol.83, No.5, pp.705-741, May 1995
- [2] M. Turk and A. Pentland, "Eigenfaces for Recognition," Journal of Cognitive Neuroscience, Vol.3, No.1, pp.71-86, 1991
- [3]N. Kambhatla and T. K. Leen, 'Dimension Reduction by Local PCA,' Neural Computation 9, pp. 1493-1516, 1997
- [4]B. A. Draper, K. Baek, M. S. Barlett, and J. R. Beveridge, " Recognizing Faces with PCA and ICA," Computer Vision and Imaging Understanding 91, pp.115-137, 2003
- [5] 이대호, 박영태, "기하학적 특징에 기반한 순수 얼굴영역 검출기법," 정보과학회 논문지 : 소프트웨어 및 응용, 제 30권, 제 8호, pp.773-779, 2003
- [6] 유현중, 김태우 공역, MATLAB을 이용한 디지털 영상처리, ITC, 2004년 8월
- [7] 김상철역, '재료역학' 청문출판사, 1992년 4월
- [8] A.M. Martinez and R. Benavente, "The AR face database", [http://rv11.ecn.purdue.edu/~alex/alex\\_face\\_DB.html](http://rv11.ecn.purdue.edu/~alex/alex_face_DB.html), CVC Tech. Report #24, 1998.