

주요성분분석과 고정점 알고리즘 독립성분분석에 의한 얼굴인식 Face Recognition by Using Principal Component Anaysis and Fixed-Point Independent Component Analysis

조용현*
Yong-Hyun Cho*

< Abstract >

This paper presents a hybrid method for recognizing the faces by using principal component analysis(PCA) and fixed-point independent component analysis(FP-ICA). PCA is used to whiten the data, which reduces the effects of second-order statistics to the nonlinearities. FP-ICA is applied to extract the statistically independent features of face image. The proposed method has been applied to the problems for recognizing the 20 face images(10 persons * 2 scenes) of 324*243 pixels from Yale face database. The 3 distances such as city-block, Euclidean, negative angle are used as measures when match the probe images to the nearest gallery images. The experimental results show that the proposed method has a superior recognition performances(speed, rate). The negative angle has been relatively achieved more an accurate similarity than city-block or Euclidean.

Key Words : *Face Recognition, Principal Component Analysis, Independent Component Analysis, Fixed-point Algorithm, Feature Extractions*

1. 서 론

얼굴인식 시스템은 인터넷 사용자인증, 지능형 DVR, 출입통제, ATM 사용자인증, 그리고 공항검색 등의 다양한 시스템으로 이용될 수 있다[1,2]. 얼굴인식을 위해서는 얼굴검출, 특징추출, 분류로 이루어진다. 얼굴검출은 주로 훈련

된 얼굴 자료를 이용하거나 미리 결정된 기하학적 얼굴 모델을 영상에 나타난 얼굴과 일치시킴으로써 얻으며, 영상으로부터 직접 얼굴 각 부분의 모양을 추출하여 검출하기도 한다. 특징추출은 인식성능에 가장 많은 영향을 미치는 매우 중요한 과정이며, 여기에는 holistic 기법, 해석적 기법, 이들 기법들을 조합한 기법들이 있다 [1]. 마지막으로 분류는 추출된 특징들을 바탕

* 정회원, 대구가톨릭대학교 컴퓨터정보통신, 교수, 工博
712-702 경북 경산시 하양읍 금락리 330

* Prof. School of Compter & Information Comm., CUD.
E-mail : yhcho@cuth.cataegu.ac.kr

으로 얼굴을 식별하는 단계이다. 최근 생체신호를 이용한 인증기술이나 유비쿼터스 컴퓨팅을 위한 지능형 요소기술로 얼굴인식에 대한 연구가 국내외적으로 활발히 이루어지고 있다[1-3].

일반적으로 얼굴인식 방법에는 크게 특징을 기반으로 한 방법과 영상을 기반으로 하는 방법이 있다[1,3]. 특징을 기반으로 하는 방법은 얼굴의 특징과 관련된 정보들을 이용하는 방법이며, 영상을 기반으로 하는 방법은 얼굴 영상 전체를 대상으로 특징을 계산하는 방법이다. 최근 얼굴인식을 위한 방법으로 영상을 기반으로 하는 방법이 널리 이용되고 있다.

특징기반 얼굴인식 방법에는 얼굴의 주요 부분인 눈, 코, 입 등의 여러 특징을 추출하여 인식에 이용하는 지역적인 특징추출 기법, 색깔과 형태로 구성된 2차원의 템플릿을 이용하는 템플릿 패턴 정합 방법, 특정한 점에서의 방향과 변화에 따른 특성을 벡터화한 Gabor jet을 특징으로 사용하는 Gabor 필터를 이용한 방식들이 있다[1-3]. 이들 방법들에서는 얼굴을 표현을 위한 차원감소나 수행시간의 절약, 포즈나 크기 등에 강건한 특성이 있으나 인식에 필요한 정보 손실이나 템플릿 구성에 따른 공간 소모, 변형의 정도에 따라 인식성능이 달라지는 제약이 있다.

영상 기반 방법은 입력 얼굴영상에서 직접 얼굴의 코딩을 얻어내는 외형을 기반으로 하며, 여기에는 입력 영상내의 상관관계를 제거하여 영상의 진폭스펙트럼을 획득하는 주요성분분석(principal component analysis : PCA)법, 선형 식별 분석(linear discriminant analysis : LDA)법, support vector machine(SVM) 방법 등이 있다[1,3,4]. 이들 방법들 역시 데이터의 차원 감소나 특징분석, 얼굴영상의 클래스들 간 최적구별, 처리시간 감소 등의 장점이 있으나 여기에도 정보의 손실, 계산의 복잡성, 실시간 인식 등의 제약은 여전히 남아 있다. 현재 가장 널리 이용되고 있는 PCA법은 얼굴 영상을 표현하는 고유 벡터를 추출하여 벡터 공간으로 투영시킨 가중치를 이용하는 방식으로, 전체의 얼굴 영상들을 적은 수의 벡터를 이용하여 표현할 수 있으며 빠르고, 계산이 간단하며, 점진적 변화에

덜 민감하다[3]. 또한 데이터 정보의 손실이 거의 없이 데이터 차원을 줄여 데이터 패턴을 파악할 수 있으며, 전체적인 영상의 특성을 파악할 수 있다는 장점을 가진다. 하지만 PCA를 이용할 경우, 눈, 코, 입 등과 같은 세부적인 표현의 제약이 있고, 표정이나 빛의 밝기 변화 및 회전 등의 왜곡에 민감하기 때문에 이에 대한 처리가 요구된다.

본 논문에서는 얼굴영상의 국부적인 특징을 추출함으로써 위상 스펙트럼까지도 고려하기 위하여 PCA와 고정점(fixed-point : FP) 알고리즘 독립성분분석(independent component analysis : ICA)[5-8]을 조합한 얼굴인식에 대해서 제안한다. 여기서 PCA는 2차 통계성에 따른 상관성을 사전에 고려됨으로써 좀 더 효과적으로 ICA를 수행하기 위함이고, 고정점 알고리즘은 빠른 ICA를 위함이다. 제안된 기법을 Yale 얼굴영상 데이터베이스[9]로부터 선택된 20개의 324*243 픽셀의 영상을 대상으로 시뮬레이션하고 기저영상의 개수에 따른 압축성과 city-block, Euclidean, 그리고 negative angle(cosine)의 거리척도에 따른 분류성능을 비교·검토하였다.

2. 고정점 알고리즘의 독립성분분석

ICA는 은닉신호의 분리(blind source separation : BSS)나 특징추출에 널리 이용되고 있다[5-8].

먼저 BSS를 위한 ICA는 m개의 은닉된 입력 신호 \mathbf{s} 로부터 선형적으로 혼합된 n개의 신호 \mathbf{x} 가 알려져 있을 때, 혼합된 신호로부터 역으로 m개의 독립인 입력신호를 찾는 기법이다. 하지만 입력신호들을 혼합하는데 이용된 혼합행렬 \mathbf{A} 는 알려져 있지 않고, 혼합과정에서 잡음 \mathbf{n} 가 추가 될 수 있으며, 그 관계는 다음과 같다.

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{n} = \sum_{i=1}^m s(i)\mathbf{a}(i) + \mathbf{n} \quad (1)$$

여기서 \mathbf{n} 는 보통 입력신호와 구별되지 않기 때문에 생략할 수도 있으며, $\mathbf{A}=[\mathbf{a}(1), \mathbf{a}(2), \dots, \mathbf{a}(m)]$ 으로 $\mathbf{a}(i)$ 는 ICA의 기저벡터이다. 혼합행

렬 \mathbf{A} 와 역혼합행렬 \mathbf{W} 에 대하여 살펴보면 다음 그림 1과 같은 구성도로 나타낼 수 있다. 결과적으로 ICA는 입력신호와 출력신호가 일치하도록 하는 즉, $\mathbf{y} = \mathbf{W}\mathbf{x} = \mathbf{W}\mathbf{A}\mathbf{s}$ 에서 $\mathbf{W}=\mathbf{A}^{-1}$ 의 관계를 만족토록 하는 것이다.

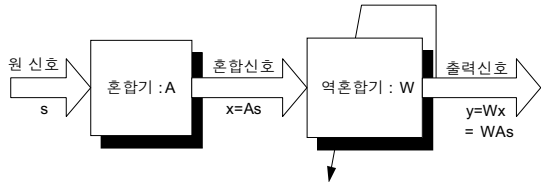


그림 1. ICA를 위한 구성도

한편 ICA는 영상의 특징을 추출하는 것에도 이용된다. 이때 원 신호 \mathbf{s} 는 영상패치 \mathbf{x} 를 형성하기 위한 혼합행렬 \mathbf{A} 의 가중치를 나타낸다. 결국 \mathbf{s} 의 선형합성에 의해 \mathbf{x} 는 형성되므로 \mathbf{s} 는 영상들의 근거들(causes)이다. 그림 2는 선형 영상합성 모델을 나타낸 것이다. 여기서 \mathbf{s} 의 열들은 근거들을 나타내며, \mathbf{A} 의 행들은 기저벡터(basis vector)로 구성된 기저함수이고, 그리고 \mathbf{x} 의 열들은 관찰된 영상패치들이다. 일반적으로 1차원 신호에서는 기저벡터이며, 영상과 같은 2차원 신호에서는 기저영상 또는 특징영상이다. 따라서 영상패치 \mathbf{x} 로부터 ICA를 이용하여 \mathbf{W} 를 구하면 $\mathbf{W} = \mathbf{A}^{-1}$ 의 관계에서 \mathbf{W} 의 행은 \mathbf{A} 의 열로 이들은 기저벡터인 영상의 특징이다.

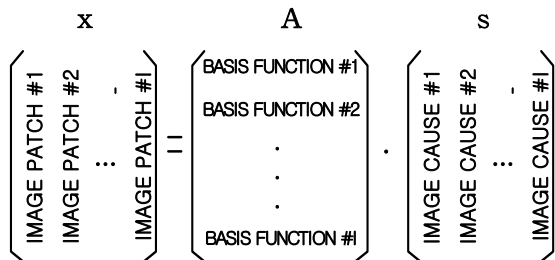


그림 2. 선형 영상합성 모델

ICA를 좀 더 효과적으로 수행하기 위해 신호를 대상으로 영평균과 whitening의 전처리 과정을 수행한다[5-8]. 이는 신호의 1차와 2차의 통계성을 고려한 정규화로 학습 알고리즘을 간략화 하기 위함이다. 신호의 영평균은 신호벡터

\mathbf{x} 에서 평균값 \mathbf{x}^* 를 뺀 차이이며, 영평균된 새로운 신호벡터 \mathbf{x} 는 다음과 같다.

$$\mathbf{x} = \mathbf{x} - \mathbf{x}^* \tag{2}$$

whitening은 데이터 벡터 \mathbf{x} 의 공분산행렬이 단위행렬값을 갖도록 함으로써 구할 수 있다. 즉, $E\{\mathbf{x}\mathbf{x}^T\}=\mathbf{I}$ 가 되도록 한다. whitening된 벡터의 성분들은 상호간의 상관성이 줄어들어 비상관성이 더 강한 독립조건을 위해 요구되는 전처리 기법이다. 본 연구에서는 whitening의 기법으로 PCA를 이용한다. 이는 PCA가 데이터 내의 2차원의 통계적 속성을 잘 반영하며, 표준 패키지들이 많이 개발되어 있기 때문이다. PCA whitening 행렬 \mathbf{V} 는 다음과 같다.

$$\mathbf{V} = \mathbf{D}^{(-1/2)}\mathbf{E}^T \tag{3}$$

여기서 \mathbf{D} 는 고유값 λ 로 구성되는 대각행렬이며, \mathbf{E} 는 고유벡터 \mathbf{c} 로 구성되는 직교행렬이다. 여기서 PCA는 ICA의 계산을 간략화하기 위해 신호의 차원인 주요성분의 개수를 조정한다. 즉 학습을 위한 얼굴영상의 속성을 가장 잘 반영하는 기저영상 개수를 결정한다.

한편 ICA를 위한 FP 알고리즘은 신경망이 가지는 병렬성과 분산성, 그리고 더 작은 메모리 요구 등의 제약을 해결하기 위해 제안된 알고리즘이다[5-8]. FP 알고리즘은 엔트로피 최적화 방법으로부터 유도된 기법으로 지금까지 알려진 기법 중 가장 빠른 학습속도를 가진다. 이는 신호 내에 포함된 상호정보를 최소화함으로써 ICA를 수행하는 기법이다.

데이터 벡터 \mathbf{x} 의 상관행렬 $E\{\mathbf{x}\mathbf{x}^T\}=\mathbf{I}$ 로 whitening되어 있다고 가정할 때, 근사화된 반복기법의 역혼합행렬 \mathbf{W} 를 구하는 뉴우턴(Newton)법은 다음과 같다.

$$\begin{aligned} \mathbf{W}^+ &= \mathbf{W} - [E\{\mathbf{x}\mathbf{g}(\mathbf{W}^T\mathbf{x})\} - \beta\mathbf{W}] / [E\{\mathbf{g}'(\mathbf{W}^T\mathbf{x})\} - \beta] \\ \mathbf{W}^* &= \mathbf{W}^+ / \|\mathbf{W}^+\| \end{aligned} \tag{4}$$

여기서 \mathbf{W}^* 는 \mathbf{W} 의 새롭게 경신된 값이고, $\beta = E\{\mathbf{W}^T\mathbf{x}\mathbf{g}(\mathbf{W}^T\mathbf{x})\}$ 이다. 결국 식 (4)는 뉴우턴법

에 기초를 둔 ICA를 위한 FP 알고리즘이다. 또한 식 (4)의 첫째식 양쪽에 $\beta - E\{g'(W^T \mathbf{x})\}$ 를 곱해 구해지는 더욱 간단해진 뉴우턴법의 FP 알고리즘은 다음과 같다.

$$\begin{aligned} W^+ &= E\{xg(W^T \mathbf{x})\} - E\{g'(W^T \mathbf{x})\}W \\ W^* &= W^+ / \|W^+\| \end{aligned} \quad (5)$$

위의 과정에서 벡터 \mathbf{x} 는 whitening된 신호벡터이며, ε 은 1보다 적은 임의의 값으로 실험에서는 0.0001로 하였다. 특히 경신의 식에서 $g(\cdot)$ 는 비선형 함수이며 일반적으로 $(\cdot)^3$ 과 $\tanh(\cdot)$ 의 함수값을 가진다. 본 연구에서는 $\tanh(\cdot)$ 함수를 이용하였다.

따라서 PCA와 FP 알고리즘 ICA에 의한 얼굴영상의 인식 과정을 요약하면 다음의 그림 3과 같다. 여기서 식별과 인식 단계는 시험 얼굴영상과 학습 얼굴영상을 비교하여 인식하는 단계이다. 이때 다양한 거리척도들이 이용될 수 있다.

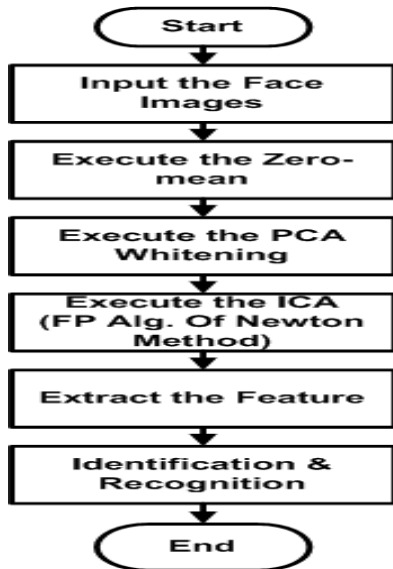


그림 3. 제안된 기법에 의한 얼굴인식

3. 시뮬레이션 결과 및 고찰

Yale 얼굴영상 데이터베이스[9]로부터 선택된 20개의 324*243 픽셀의 영상을 대상으로 PCA와 FP-ICA의 제안된 조합형 기법으로 실

험하였다. 얼굴영상의 복원성능을 평가하기 위해 원 얼굴의 픽셀값 x_i 와 복원된 얼굴의 픽셀값 y_i 사이의 절대평균합오차 AMSE(absolute mean sum error) $= (1/N^2) \sum |x_i - y_i|$ 를 정의하였으며, 여기서 N^2 은 전체 픽셀수이다. 실험은 펜티엄 IV-2.2G 컴퓨터상에서 Matlab 5.1로 실행하였다.

그림 4는 인식에 이용된 보통의 표정을 가진 학습 얼굴영상들이다. 그림 4(a)는 실험에 이용된 10인의 20개의 얼굴영상 중에서 각자 1개씩인 10개의 원 얼굴영상을 나타낸 것이며, 그림 4(b)는 이들의 평균 얼굴영상이고, 그림 4(c)는 원 얼굴영상에서 평균 얼굴영상을 뺀 영 평균된 얼굴영상들을 각각 나타낸 것이다.

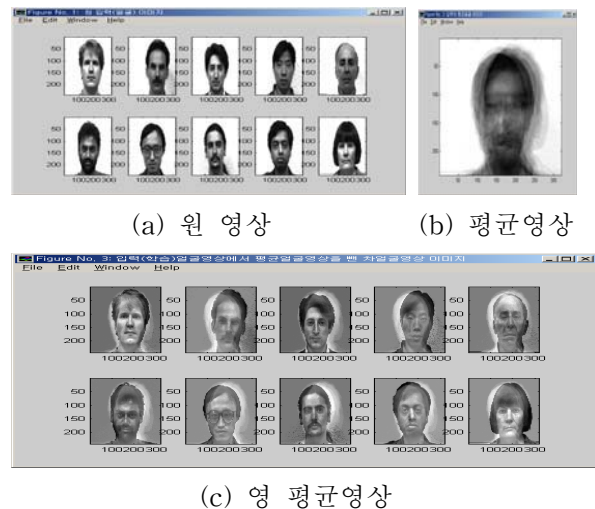
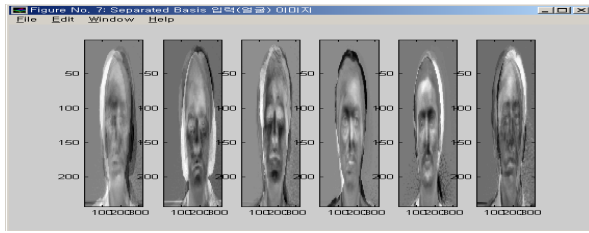
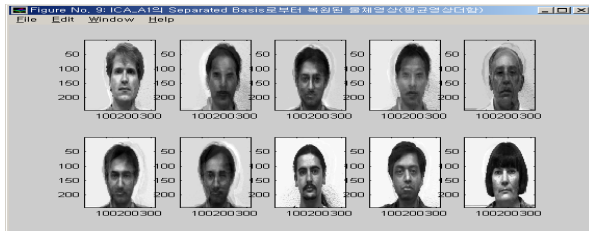


그림 4. 학습 얼굴영상

그림 5는 기저벡터에 의한 6개의 기저얼굴영상과 이들에 의해 복원된 10개의 얼굴영상들을 나타낸 것이다. 이는 10개의 얼굴영상 중에서 6개의 특징얼굴을 추출한 것으로 차원감소에 따른 데이터의 압축효과를 보여 준 것이다. 하지만 육안으로 볼 때, 그림 5(a)로부터 복원된 그림 5(b)의 얼굴영상에는 많은 잡영이 존재하여 원 얼굴영상과 다소 차이가 있음을 알 수 있다. 이는 기저벡터의 수를 줄인 것에 대한 결과이다.



(a) 6개의 기저얼굴



(b) 10개의 복원얼굴

그림 5. 기저얼굴과 복원얼굴

한편 그림 6은 기저얼굴 영상인 기저벡터의 개수에 따라 절대평균합오차를 나타낸 것으로 복원성능을 나타낸 것이다. 여기서는 기저얼굴 영상의 수가 증가함에 따라 절대평균합오차의 복원성능이 향상되다 다시 저하됨을 알 수 있다. 이는 기저얼굴의 개수인 특징벡터수의 지나친 증가는 오히려 과학습에 의한 복원성능의 저하를 초래함을 알 수 있다. 따라서 본 연구에서는 기저얼굴의 수를 복원성능이 가장 우수한 9개일 때를 기준으로 실험하였다.

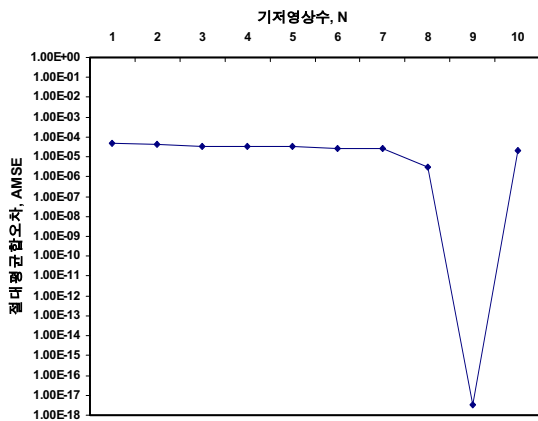
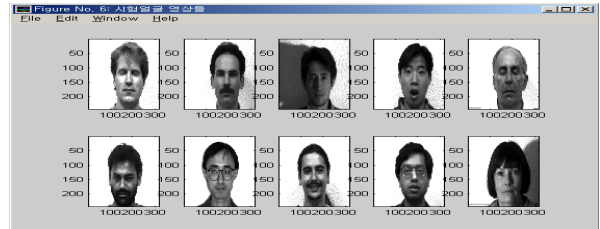


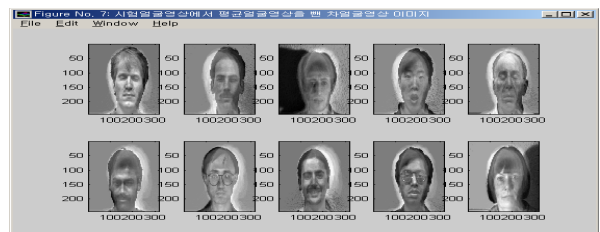
그림 6. 기저영상 개수에 따른 절대평균합오차

그림 7은 다양한 표정과 조명(놀람, 슬픔, 행복, 잠움, 윙크, 안경깸, 중앙조명, 오른쪽조명, 왼쪽조명)을 가진 시험을 위한 10개 얼굴영상이다. 그림 7(a)는 원 시험 얼굴영상이며, 그림 7(b)는 원

시험영상에서 학습영상의 평균영상을 뺀 차로 영 평균된 얼굴영상이다. 시험영상은 조명이나 표정 등에서 학습영상과 다르며, 시험영상의 고유벡터는 영 평균영상을 이용하여 얻어진다.



(a) 원 영상



(b) 영 평균영상

그림 7. 시험 얼굴영상

한편 표 1은 10개의 학습 얼굴영상을 대상으로 그림 3(c)의 학습 얼굴영상과 그림 5(b)의 다른 표정이나 조명을 가진 새로운 시험 얼굴영상을 인식한 결과이다. 여기서는 학습과 시험 얼굴영상 사이의 유사성을 측정하기 위해 city block(L1-norm), Euclidean(L2-norm), negative angle(cosine)의 3가지 거리척도를 이용하였다. 3가지 척도에서 이미 학습된 얼굴영상과 새로운 시험 얼굴영상은 각각 100%와 80%의 인식률을 가짐을 알 수 있다. city block과 Euclidean의 거리척도를 이용한 경우의 인식 결과에 따른 오차 값은 학습 얼굴영상은 거의 10^{-15} 이하의 아주 우수한 인식성능을 가지나 상대적으로 시험 얼굴영상은 10^2 정도의 오차가 있음을 알 수 있다. 하지만 시험 얼굴영상은 이 정도의 오차를 가지지만 조명과 안경을 낀 3번과 9번이 각각 학습 얼굴영상의 6번과 3번으로 오인식된 경우를 제외하고는 모두 잘 인식되었다. 한편 negative angle의 경우 오차 값은 학습 얼굴영상은 0.0이고, 시험 얼굴영상은 1.2352로 위의 2가지 거리척도보다 오차가 적음을 알 수 있다. 10개의 얼굴영상이 완벽하게 인식될 때

negative angle의 값이 -10.0을 가지게 되나 실험에서는 -8.8254이 되어 여기서의 오차 값은 그 차인 1.1746이 된다. 또한 negative angle에서는 다른 2가지 척도와 달리 3번과 9번이 각각 학습 얼굴영상의 9번과 1번으로 오인식된 경우를 제외하고는 모두 잘 인식되었다. 결국 시험의 결과에서는 영상의 조명이나 안경 등의 주변 환경이 인식에 큰 영향을 미침을 알 수 있다.

표 1. 10개 학습과 시험얼굴 영상의 인식결과

거리척도	입력영상	인식률(%)	오차값	오인식영상
City Block (L1-norm)	학습얼굴	100	43.3764e-015	.
	시험얼굴	80	176.1504	3(6), 9(3)
Euclidean (L2-norm)	학습얼굴	100	28.0689e-015	.
	시험얼굴	80	79.8866	3(6), 9(3)
Negative Angle (Cosine)	학습얼굴	100	0.0	.
	시험얼굴	80	1.1746	3(9), 9(1)

* i(j) : i = 바로 인식되어야 할 학습 얼굴번호
j = 잘못 인식된 학습 얼굴번호

이상의 실험결과에서 PCA와 FP-ICA를 조합한 얼굴인식은 이미 학습된 얼굴은 100% 인식을 하며, 새로이 입력되는 시험 얼굴영상은 80%가 인식됨을 알 수 있다. 또한 분류를 위한 성능을 위한 유사성 측정을 위해 이용된 거리척도에서는 negative angle의 척도를 이용하는 것이 city-block이나 Euclidean보다 우수한 인식률이 있음을 알 수 있다.

4. 결론

본 논문에서는 주요성분분석과 고정점 알고리즘의 독립성분분석을 조합한 얼굴인식을 제안하였다. 여기서 PCA는 2차 통계성에 따른 상관성을 사전에 고려됨으로써 좀 더 효과적으로 ICA를 수행하기 위함이고, 고정점 알고리즘은 뉴우턴법에 기초한 것으로 빠른 특징추출성을 얻기 위함이며, ICA는 통계적으로 독립인 기저 영상을 효과적으로 추출하기 위함이다.

제안된 기법을 Yale 얼굴영상 데이터베이스로부터 선택된 20개의 324*243 픽셀의 영상을

대상으로 시뮬레이션한 결과, 학습된 얼굴과 시험얼굴 각각에서 100%와 80%의 인식률을 가지며, 기저영상의 개수에 따른 압축성능과 city-block, Euclidean, 그리고 negative angle의 거리척도에 따른 우수한 인식성능이 있음을 확인하였다. 또한 negative angle의 거리척도가 city-block과 Euclidean보다 우수한 분류성능이 있음도 확인하였다.

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

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) J. T. Tou and R. C. Gonzalez, 'Pattern Recognition Principles,' Addison-Wesley Pub., London, 1974
- 3) M. Turk and A. Pentland, "Eigenfaces for Recognition," Journal of Cognitive Neuroscience, Vol.3, No.1, pp.71-86, 1991
- 4) K. I. Diamantaras and S. Y. Kung, 'Principal Component Neural Networks: Theory and Applications, Adaptive and Learning Systems for Signal Processing, Communications, and Control,' John Wiley & Sons, Inc., 1996
- 5) T. W. Lee, Independent Component Analysis : Theory and Applications, Kluwer Academic Pub., Boston, 1998
- 6) J. Karhunen, "Neural Approaches to Independent Component Analysis and Source Separation", "4th European Symp., Artificial Neural Network, ESANN96, Burges, Belgium, pp. 249-266, Apr. 1996
- 7) S. Roberts and R. Everson, Independent Component Analysis : Principles and Practice, Cambridge Univ. Press, 2001
- 8) A. Hyvarinen, J. Karhunen, and E. Oja, Independent Component Analysis, John Wiley & Sons, Inc., 2001
- 9) "Yale Face Databases," <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>

(2004년 12월 4일 접수, 2005년 7월 20일 채택)