

고정점 알고리즘의 ICA-basis에 의한 얼굴인식

조용현, 홍성준*

*대구가톨릭대학교 컴퓨터정보통신공학부

e-mail : yhcho@cu.ac.kr

Face Recognition by Using ICA-basis of Fixed-Point Algorithm

Yong-Hyun Cho, Seong-Jun Hong*

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

요 약

본 논문에서는 고정점 알고리즘의 독립성분분석을 이용한 얼굴인식을 제안하였다. 여기서 고정점 알고리즘은 뉴턴법에 기초한 것으로 빠른 인식을 위함이고, 독립성분분석의 이용은 통계적으로 독립인 기저영상을 효과적으로 추출하기 위함이다. 제안된 기법을 Yale 얼굴영상 데이터베이스로부터 선택된 20개의 320*243 픽셀의 영상을 대상으로 시뮬레이션한 결과, 주요성분의 개수에 따른 압축성능과 city-block, Euclidian, 그리고 negative angle의 거리척도에 따른 분류에서 우수한 인식성능이 있음을 확인할 수 있었다.

1. 서론

정보의 디지털화에 따른 이를 위한 보호와 보안 요구가 증가되면서 생체정보의 활용에 대한 연구들이 활발하게 진행되고 있다. 그 중에서도 얼굴인식 분야는 기존의 지문이나 홍채 등을 이용한 기법들보다 사용자와 친화적이면서도 데이터 획득의 용이성으로 많은 연구들이 이루어지고 있다[1-3]. 하지만 입력 얼굴로부터 해당 얼굴을 출력하는 얼굴인식에서는 많은 제약들이 있다. 즉, 3차원의 원 얼굴을 2차원으로 투영시킨 얼굴을 이용함으로써 깊이 정보, 크기, 회전등의 인식에 중요한 다양한 정보들의 손실, 패턴으로서 얼굴이 가지는 복잡성, 조명이나 배경 등의 환경의 복잡성, 인경이나 표정의 변화 등이 인식성능을 좌우하는 요인들이다.

일반적으로 얼굴인식에는 크게 특징을 기반으로 한 방법과 영상을 기반으로 하는 방법이 있다[1,2]. 특징을 기반으로 하는 방법은 얼굴특징과 관련된 정보들을 이용하는 방법이며, 영상을 기반으로 한 방법은 얼굴영상 전체를 대상으로 특징을 계산하는 방법이다. 특징 기반 방법에는 입력 영상으로부터 얼굴의 주요 부분인 눈, 코, 입 등의 여러 특징점을 추출하는 지역 특징 추출기법, 얼굴영상의 경계선들을 모아 만든 경계선 지도기법, Gabor 필터를 이용한 기법 등이 있다. 영상 기반 방법은 입력 얼굴영상에서 직접 얼굴의 코딩을 얻어내는 외형을 기반으로 하며, 여기에는 입력 영상내의 상관관계를 제거하여

영상의 진폭 스펙트럼을 획득하는 주요성분분석(principal component analysis : PCA)과 PCA를 일반화하여 영상의 위상 스펙트럼까지도 고려하는 독립성분분석(independent component analysis : ICA), 신경망, support vector machine(SVM) 등이 있다[3-8]. 얼굴인식 방법으로 영상 기반 방법이 최근 주로 이용되고 있으며 이에 대한 연구들이 활발하게 진행되고 있다.

한편 얼굴인식에서 PCA와 ICA는 중복신호 제거(redundancy reduction)의 효과가 있어 뇌의 시각 기능을 잘 반영하며 타 방법에 비해 높은 인식률을 얻을 수 있는 장점을 갖고 있다. 특히 PCA는 조밀한 코딩(compact coding)을 유도하여 중복 신호를 제거한다. 결국 얼굴영상 자체에서 불필요한 정보를 조밀하게 바꾸어 차원을 줄여 인식과정이 간단해질 뿐만 아니라 높은 인식률을 가지고 있어 많이 사용되고 있다. ICA는 PCA와 달리 드문 코딩(sparse coding)에 의하여 중복 신호를 제거한다. 즉, PCA가 신호의 차원 자체를 줄여 조밀한 코딩을 유도하는데 비해 ICA는 출력 신호의 값이 드물게 나타나는 방법으로 중복 신호를 제거하며 이는 뇌의 기능 모델에 좀 더 관련이 있다[5-8]. 따라서 물체인식을 위한 특징추출에 ICA를 이용하면 PCA와 달리 국부적인 특징(local feature)을 추출할 수 있어 PCA 보다 높은 인식률을 얻을 수 있다. 하지만 ICA에서는 데이터의 통계적 독립성을 판정하기 위해 확률밀도함수와 역행렬을 구한다든지 고차원의 cumulant나 moment를 구하는 등 복잡한 계산이 요구되어 PCA보다 계산이 복잡하여 추출시간이 오래 걸리는 제약이 있다.

** 본 논문은 대구광역시 우수 연구센터 지원사업(2004)으로 수행되었음.

본 논문에서는 뉴우턴(Newton)법[8]에 기반한 고정점 알고리즘(fixed point : FP)의 독립성분분석을 이용한 얼굴인식을 제안한다. 여기서 뉴우턴법의 FP 알고리즘은 엔트로피 최적화로부터 유도된 빠른 특징추출의 속성을 가지며, 독립성분분석은 복원성능을 개선시키기 위함이다. 제안된 기법을 Yale 얼굴영상 데이터베이스[9]로부터 선택된 20개의 320*243 픽셀의 영상을 대상으로 시뮬레이션하고 주요성분의 개수에 따른 압축성능과 city-block, Euclidian, 그리고 negative angle(cosine)의 거리척도에 따른 인식을 위한 분류성능을 비교·검토한다.

2. 뉴우턴법에 기반한 고정점 알고리즘의 ICA

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

먼저 BSS를 위한 ICA는 m개의 은닉된 입력신호 s로부터 선형적으로 혼합된 n개의 신호 x가 알려져 있을 때, 혼합된 신호로부터 역으로 m개의 독립된 입력신호를 찾는 기법이다. 하지만 입력신호들을 혼합하는데 이용된 혼합행렬 A는 알려져 있지 않고, 혼합과정에서 잡음 η가 추가 될 수 있으며, 그 관계는 다음과 같다.

$$x = As + \eta = \sum_{i=1}^m s(i)a(i) + \eta \tag{1}$$

여기서 η는 보통 입력신호와 구별되지 않기 때문에 생략할 수도 있으며, A=[a(1), a(2), ..., a(m)]으로 a(i)는 ICA의 기저벡터이다. 혼합행렬 A와 역혼합행렬 W에 대하여 살펴보면 다음 그림 1과 같은 구성도로 나타낼 수 있다. 결과적으로 ICA는 입력신호와 출력신호가 일치하도록 하는 즉, y = Wx = WAs에서 W=A⁻¹의 관계를 만족토록 하는 것이다.

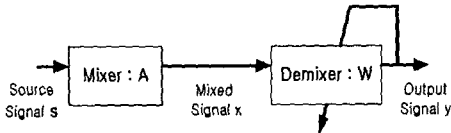


그림 1. ICA를 위한 구성도

한편 ICA는 영상의 특징을 추출하는 것에도 이용된다. 이때 원신호 s는 영상패치 x를 형성하기 위한 혼합행렬 A의 가중치를 나타낸다. 결국 s의 선형합성에 의해 x는 형성되므로 s는 영상들의 근거들(causes)이다. 그림 2는 선형 영상합성 모델을 나타낸 것이다. 여기서 s의 열들은 근거들을 나타내며, A의 행들은 기저벡터로 구성된 기저함수이고, 그리고 x의 열들은 관찰된 영상패치들이다. 따라서 영상패치 x로부터 ICA를 이용하여 W를 구하면 W = A⁻¹의 관계에서 W의 행은 A의 열로 이들은 기저벡터인 영상의 특징이다.

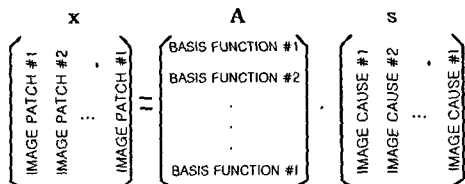


그림 2. 선형 영상합성 모델

일반적으로 ICA를 좀 더 효과적으로 수행하기 위해 혼합 신호를 대상으로 영평균과 whitening의 전처리 과정을 수행한다[5-8]. 이는 신호의 1차와 2차의 통계성을 고려한 정규화로 학습 알고리즘을 간략화 하기 위함이다. 신호의 영평균은 신호벡터 x에서 평균값 x'를 뺀 것이며, 영평균된 새로운 신호벡터 x는 다음과 같다.

$$x = x - x' \tag{2}$$

한편, whitening은 데이터 벡터 x의 공분산행렬이 단위행렬값을 갖도록 함으로써 구할 수 있다. 즉, E(xx^T)=I 가 되도록 한다. whitening된 벡터의 성분들은 상호간의 상관성이 줄어들어 비상관성이 더 강한 독립조건을 위해 요구되는 전처리 기법이다. 이러한 whitening의 기법으로 PCA가 널리 사용되며, 이때 PCA whitening 행렬 V는 다음과 같다.

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

여기서 D는 고유값 λ로 구성되는 대각행렬이며, E는 고유벡터 c로 구성되는 직교행렬이다. 특히 여기에서는 ICA의 계산을 간략화하기 위해 신호의 차원인 주요성분의 개수를 조정한다.

독립성분분석을 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 ICA의 해를 구하는 기법이다.

FP 알고리즘은 알고리즘의 한 단계에 많은 수의 데이터들이 이용되는 batch mode로 계산을 수행하며, 신경망으로 수행될 수 있는 알고리즘이다. 데이터 벡터 x의 상관행렬 E(xx^T)=I로 whitening되어 있다고 가정할 때, 근사화된 반복기법의 역혼합행렬 W를 구하는 뉴우턴법은 다음과 같다.

$$W' = W - [E(xg(W^T x)) - \beta W] / [E(g'(W^T x)) - \beta], \tag{4}$$

$$W' = W' / \|W'\|$$

여기서 W'는 W의 새롭게 경신된 값이고, β = E(W^Txg(W^Tx))이다. 결국 식 (4)는 뉴우턴법에 기초를 둔 ICA를 위한 FP 알고리즘이다. 또한 식 (4)의 첫째식 양쪽에 β - E(g'(W^Tx))를 곱해 구해지는 더욱 간단해진 뉴우턴법의 FP 알고리즘은 다음과 같다.

$$W' = E(xg(W^T x)) - E(g'(W^T x))W, W' = W' / \|W'\| \tag{5}$$

위의 과정에서 벡터 x는 whitening된 신호벡터이며, ε은 1보다 적은 임의의 값으로 실험에서는 0.0001로 하였다. 특히 경신의 식에서 g(·)는 비선형 함수이며 일반적으로 (·)³과 tanh(·)의 함수값을 가진다. 본 연구에서는 tanh(·) 함수를 이용하였다.

따라서 뉴우턴법의 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.

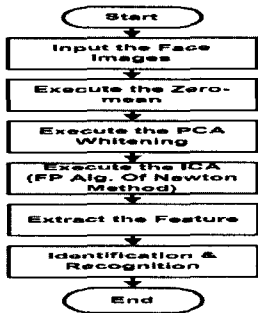


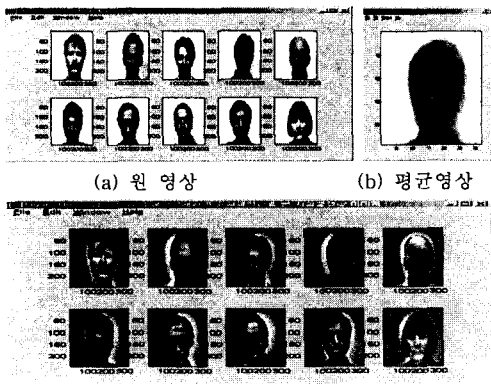
그림 3. 뉴우턴의 FP 알고리즘 ICA에 의한 얼굴인식과정

3. 시뮬레이션 결과 및 고찰

Yale 얼굴영상 데이터베이스로부터 선택된 20개의 320*243 픽셀의 영상을 대상으로 제안된 FP 알고리즘의 ICA-basis로 실험하였다. 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.

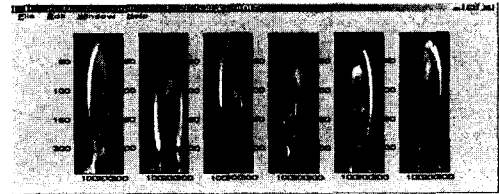
그림 4는 인식에 이용된 보통의 표정을 가진 학습 얼굴 영상들이다. 그림 4(a)는 실험에 이용된 10인의 20개의 얼굴영상 중에서 각자 1개씩인 10개의 원 얼굴영상을 나타낸 것이며, 그림 4(b)는 이들의 평균 얼굴영상이고, 그림 4(c)는 원 얼굴영상에서 평균 얼굴영상을 뺀 영평균된 얼굴영상들을 각각 나타낸 것이다.



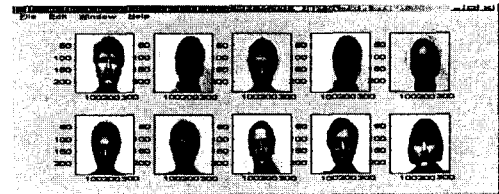
(c) 영 평균영상
그림 4. 학습 얼굴영상

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. 하

지만 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



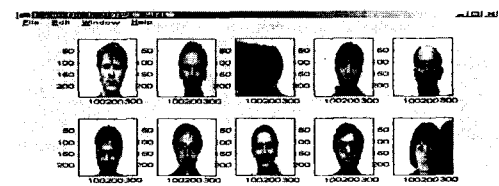
(b) 10개의 복원얼굴
그림 5. 기저얼굴과 복원얼굴

한편 표 1은 기저얼굴영상인 기저벡터의 개수에 따라 절대평균합오차를 나타낸 것으로 복원성능을 나타낸 것이다. 여기서는 기저얼굴의 수가 증가함에 따라 절대평균합오차의 복원성능이 향상되다 다시 저하됨을 알 수 있다. 이는 기저얼굴의 개수인 특징벡터수의 지나친 증가는 오히려 과학수에 의한 복원성능의 저하를 초래함을 알 수 있다. 따라서 본 연구에서는 기저얼굴의 수를 복원성능이 가장 우수한 9개일 때를 기준으로 실험하였다.

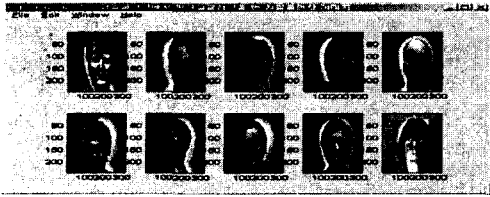
표 1. 기저영상의 개수에 따른 절대평균합오차

기저영상수	1	2	3	4	5	6	7	8	9	10
절대평균합오차	4.6376e-005	4.2520e-005	3.4990e-005	3.4040e-005	3.2310e-005	2.6480e-005	2.4930e-005	2.9460e-005	3.4450e-005	2.1740e-005

그림 6은 다양한 표정과 조명(놀람, 슬픔, 행복, 잠움, 윙크, 안경깎, 중앙조명, 오른쪽조명, 왼쪽조명)을 가진 시험을 위한 10개 얼굴영상이다. 그림 6(a)는 원 시험 얼굴영상이며, 그림 6(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 6. 시험 얼굴영상

한편 표 2는 10개의 학습 얼굴영상을 대상으로 그림 4(c)의 학습 얼굴영상과 그림 6(b)의 다른 표정이나 조명을 가진 새로운 시험 얼굴영상을 인식한 결과이다. 여기서는 학습과 시험 얼굴영상 사이의 유사성을 측정하기 위해 city block(L1-norm), Euclidian(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 block(L1-norm)와 Euclidian(L2-norm)의 거리척도를 이용한 경우의 인식결과에 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 주변환경이 인식에 큰 영향을 미침을 알 수 있다.

표 2. 10개의 학습과 시험 얼굴영상 각각에 대한 인식결과

거리척도	입력영상	인식률(%)	오차값	오인식 시험영상
City Block (L1-norm)	학습얼굴	100	43.3764e-015	.
	시험얼굴	80	176.1504	3(6), 9(3)
Euclidian (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 = 바르게 인식되어야 할 학습 얼굴번호,
j = 잘못 인식된 학습 얼굴번호

이상의 실험결과에서 자율학습의 주요성분분석을 이용한 얼굴인식은 새로운 시험 얼굴영상에 대해 약 80%이상의 인식률을 가지며, 인식을 위한 유사성 측정을 위한 거리척도에서는 negative angle의 척도를 이용하는 것이 우수한 인식률이 있음을 알 수 있다.

4. 결론

본 논문에서는 고정점 알고리즘의 독립성분분석을 이용한 물체영상의 특징추출을 제안하였다. 여기서 고정점 알고리즘은 뉴우턴법에 기초한 것으로 빠른 특징추출성능을 얻기 위함이고, 독립성분분석의 이용은 통계적으로 독립인 basis 영상을 효과적으로 추출하기 위함이다.

제안된 기법을 Yale 얼굴영상 데이터베이스로부터 선택된 20개의 320*243 픽셀의 영상을 대상으로 시뮬레이션한 결과, 주요성분의 개수에 따른 압축성능과 city-block, Euclidian, 그리고 negative angle의 거리척도에 따른 분류에서 우수한 인식성능이 있음을 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] M. Turk and A. Pentland, "Eigenfaces for Recognition," Journal of Cognitive Neuroscience, Vol.3, No.1, pp.71-86, 1991
- [3] J. T. Tou and R. C. Gonzalez, 'Pattern Recognition Principles,' Addison-Wesley Pub., London, 1974
- [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>