

# FP-ICA의 인수부호에 의한 얼굴인식

조용현, 홍성준\*

\*대구가톨릭대학교 컴퓨터정보통신공학부

e-mail : yhcho@cu.ac.kr

## Face Recognition by Using Factorial Face Code of FP-ICA

Yong-Hyun Cho, Seong-Jun Hong\*

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

### 요 약

본 논문에서는 고정점 알고리즘의 독립성분분석을 이용하여 얼굴영상의 인수부호를 찾아 얼굴을 인식하는 기법을 제안하였다. 여기서 고정점 알고리즘은 뉴우턴법에 기초한 것으로 빠른 특징추출을 위함이고, 독립성분분석의 이용은 통계적으로 독립인 계수로 구성된 인수부호를 효과적으로 추출하기 위함이다. 제안된 기법을 Yale 얼굴영상 데이터베이스로부터 선택된 20개의 324\*243 픽셀의 영상을 대상으로 시뮬레이션한 결과, 기저영상의 개수에 따른 압축성능과 L1- 및 L2-norm의 거리척도에 따른 분류에서 우수한 인식성능이 있음을 확인할 수 있었다.

### 1. 서론

얼굴인식 시스템은 인터넷 사용자인증, 지능형 DVR, 출입통제, ATM 사용자인증, 그리고 공항검색 등의 다양한 시스템으로 이용될 수 있다[1-3]. 얼굴인식을 위해서는 얼굴검출, 특징추출, 분류로 이루어진다. 여기서 얼굴영역의 검출은 주로 훈련된 얼굴 자료를 이용하거나 미리 결정된 기하학적 얼굴 모델을 영상에 나타난 얼굴과 일치시킴으로써 얻으며, 영상으로부터 직접 얼굴 각 부분의 모양을 추출하여 검출하기도 한다. 일반적으로 얼굴검출에서 이용되는 영상은 전면 얼굴영상이다. 하지만 카메라 등으로부터 입력되는 얼굴은 크기나 방향 및 zoom 등과 같은 많은 제약들이 있다. 특징추출은 인식성능에 가장 많은 영향을 미치는 매우 중요한 과정이며, 여기에는 holistic 기법, 해석적 기법, 이들 두 기법들을 조합한 기법들이 있다 [2,3]. 마지막으로 분류는 추출된 특징들을 바탕으로 얼굴을 식별하는 단계이다. 여기에는 전문가의 규칙이나 움직임 파라미터를 이용하는 규칙기반과 판별함수나 시공간의 움직임 에너지 형판을 이용하는 형판기반 기법들이 있다. 최근 생체신호를 이용한 인증기술이나 유비쿼터스 컴퓨팅을 위한 지능형 요소기술로 얼굴인식에 대한 연구가 국내외적으로 활발히 이루어지고 있다.

최근 얼굴에 따라 감정을 알아내어 인간의 의지나 감정을 인지하는 보다 인간친화적인 감성컴퓨터를 개발하려는 연구들이 활발하게 이루어지고 있다[1-3]. 얼굴인식을 위해 이용되는 방법에는 주요성분분석(principal component analysis : PCA) 나 local feature analysis(LFA) 등의

holistic 분석, Gabor 필터를 이용하는 local filter, 변화량을 취해서 특징점을 찾거나 템플릿을 이용하는 human subject 등이 연구되었다[3-7]. 하지만 알려진 가장 우수성의 기법은 holistic 분석으로 이는 PCA에 기반을 둔 방법들이다. 그러나 이들 방법들은 대부분 얼굴영상 내의 2차원의 통계적 정보만을 이용함으로써 대상이 되는 얼굴 환경에 따라서는 인식성능이 떨어지는 제약이 있다. 최근 이러한 제약을 극복하기 위해서 고차원의 통계적 정보를 이용하는 독립성분분석(independent component analysis : ICA)를 이용한 방법이 제안되었다[5-8]. 이들 연구들에서도 기존에 알려진 입력력 사이의 상호정보를 최대화하기 위한 수치적 방법들을 이용함으로써 문제에 따라서는 비현실적인 계산시간과 분석성능으로 ICA 고유의 제약을 그대로 가지고 있다. 따라서 보다 정확한 얼굴 특징의 추출을 위한 새로운 기법의 개발은 인식의 성능을 더욱 더 개선시키기 위한 절실한 과제이다.

식별을 위해 신경망을 이용하는 비전 시스템에 대한 연구도 있으나 대부분의 연구들에서는 특징추출 등을 위한 기반기술의 개발이 부족하다. 하지만 지금까지의 대부분의 연구들에서 특징추출은 기존의 PCA를 이용한 기법들을 주로 활용하고 있는 실정이다. 특히 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 논문에서는 뉴우턴(Newton)법[8]에 기반한 고정점 알

고리점(fixed point : FP)의 독립성분분석(FP-ICA)을 이용하여 얼굴 내의 통계적 독립인 인수를 찾아 얼굴을 인식하는 기법을 제안한다. 여기서 뉴턴법의 FP 알고리즘은 엔트로피 최적화로부터 유도된 빠른 특징추출의 속성을 가지며, 독립성분분석의 이용은 통계적으로 독립인 계수로 구성된 인수부호를 효과적으로 추출하기 위함이다. 제안된 기법을 Yale 얼굴영상 데이터베이스[9]로부터 선택된 20개의 324\*243 픽셀의 영상을 대상으로 시뮬레이션하고 기저영상의 개수에 따른 압축성과 L1- 및 L2-norm의 거리척도에 따른 분류성능을 비교·검토한다.

2. 뉴턴법에 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 **s**로부터 선형적으로 혼합된 n개의 신호 **x**가 알려져 있을 때, 혼합된 신호로부터 역으로 m개의 독립인 입력신호를 찾는 기법이다. 하지만 입력신호들을 혼합하는데 이용된 혼합행렬 **A**는 알려져 있지 않고, 혼합과정에서 잡음 **η**가 추가 될 수 있으며, 그 관계는 다음과 같다.

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

여기서 **η**는 보통 입력신호와 구별되지 않기 때문에 생략할 수도 있으며, **A**=[**a**(1), **a**(2), ..., **a**(m)]으로 **a**(i)는 ICA의 기저벡터이다. 혼합행렬 **A**와 역혼합행렬 **W**는 **y** = **Wx** = **WAs**에서 **W=A**<sup>-1</sup>의 관계를 만족토록 하는 것이다.

ICA는 영상의 특징을 추출하는 것에도 이용된다. 이때 원신호 **s**는 영상패치 **x**를 형성하기 위한 혼합행렬 **A**의 가중치를 나타낸다. 결국 **s**의 선형합성에 의해 **x**는 형성되므로 **s**는 영상들의 근거들(causes)이다. 그림 1은 선형 영상합성 모델을 나타낸 것이다. 여기서 **s**의 열들은 근거들을 나타내며, **A**의 행들은 기저벡터로 구성된 기저함수이고, 그리고 **x**의 열들은 관찰된 영상패치들이다. 따라서 영상패치 **x**로부터 ICA를 이용하여 **W**를 구하면 **W = A**<sup>-1</sup>의 관계에서 **W**의 행은 **A**의 열로 이들은 기저벡터인 영상의 특징이다.

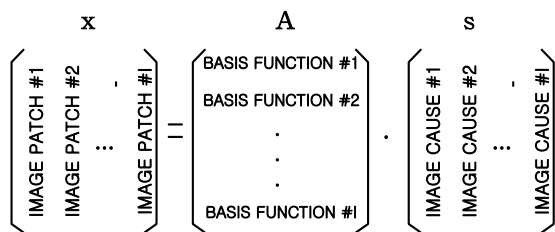


그림 1. 선형 영상합성 모델

한편 영상 내에 존재하는 통계적으로 독립인 계수로 구성된 인수부호를 효과적으로 추출하기 위해서 이용되는 ICA의 구성은 그림 2와 같다.

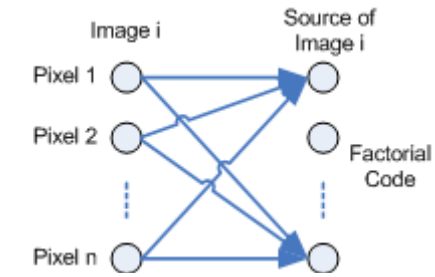


그림 2. 인수부호 추출을 위한 ICA 구조

여기에서 행은 다른 픽셀을 나타내고, 열은 다른 영상을 나타내는 데이터 행렬을 구성한다.

일반적으로 ICA를 좀 더 효과적으로 수행하기 위해 혼합신호를 대상으로 영평균과 whitening의 전처리 과정을 각각 수행한다[5-8]. 이는 신호의 1차와 2차의 통계성을 고려한 정규화로 학습 알고리즘을 간략화 하기 위함이다. 먼저, 신호의 영평균은 신호벡터 **x**에서 평균값 **x\***를 뺀 차로 다음과 같다.

$$\mathbf{x} = \mathbf{x} - \mathbf{x}^* \quad (2)$$

또한 whitening은 데이터 벡터 **x**의 공분산행렬이 단위행렬값을 갖도록 함으로써 구할 수 있다. 즉, E{**xx**<sup>T</sup>}=**I** 가 되도록 하기위해 PCA가 널리 사용되며, 이때 PCA whitening 행렬 **V**는 다음과 같다.

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

여기서 **D**는 고유값 λ로 구성되는 대각행렬이며, **E**는 고유벡터 **c**로 구성되는 직교행렬이다. 특히 여기에서는 ICA의 계산을 간략화하기 위해 신호의 차원인 주요성분의 개수를 조정한다.

FP 알고리즘은 신경망이 가지는 병렬성과 분산성, 그리고 더 작은 메모리 요구 등의 제약을 해결하기 위해 제안된 ICA 알고리즘이다[8]. FP 알고리즘은 알고리즘의 한 단계에 많은 수의 데이터들이 이용되는 batch mode로 계산을 수행하며, 신경망으로 수행될 수 있는 알고리즘이다. 데이터 벡터 **x**의 상관행렬 E{**xx**<sup>T</sup>}=**I**로 whitening되어 있다고 가정할 때, 근사화된 반복기법의 역혼합행렬 **W**를 구하는 뉴턴법은 다음과 같다.

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

여기서 **W\***는 **W**의 새롭게 경신된 값이고, β=E{**W**<sup>T</sup>**xg**(**W**<sup>T</sup>**x**)}이다. 결국 식 (4)는 뉴턴법에 기초를 둔 ICA를 위한 FP 알고리즘이다. 또한 식 (4)의 첫째식 양쪽에 β - E{g'(**W**<sup>T</sup>**x**)}를 곱해 구해지는 더욱 간단해진 뉴턴법의 FP 알고리즘은 다음과 같다.

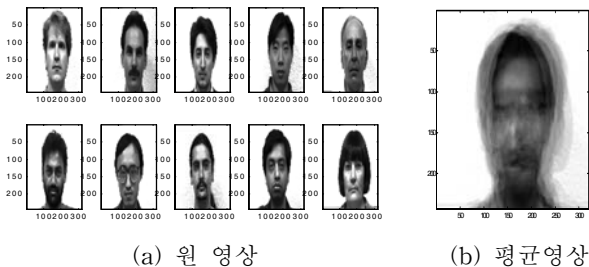
$$\mathbf{W}^* = E(\mathbf{xg}(\mathbf{W}^T\mathbf{x})) - E(g'(\mathbf{W}^T\mathbf{x}))\mathbf{W}, \mathbf{W}^* = \mathbf{W}/\|\mathbf{W}\| \quad (5)$$

위의 과정에서 벡터  $\mathbf{x}$ 는 whitening된 신호벡터이며,  $\varepsilon$ 은 1보다 작은 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.

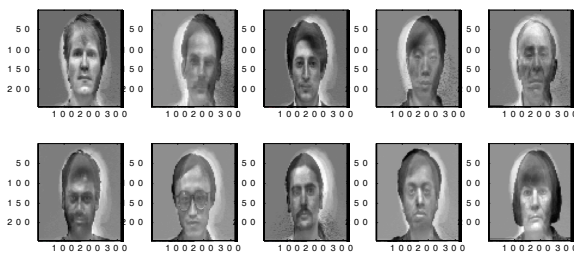
### 3. 시뮬레이션 결과 및 고찰

Yale 얼굴영상 데이터베이스[9]로부터 선택된 20개의 324\*243 픽셀의 영상을 대상으로 제안된 FP 알고리즘의 ICA로 실험하였다. 얼굴영상의 복원성능을 평가하기 위해 원 얼굴의 픽셀값  $x_i$ 와 복원된 얼굴의 픽셀값  $y_i$  사이의 절대평균합오차 AMSE(absolute mean sum error)  $= (1/N) \sum_{i=1}^N |x_i - y_i|$ 를 정의하였으며, 여기서 N은 전체 픽셀수이다. 실험은 펜티엄 IV-2.2G 컴퓨터상에서 Matlab 5.1로 실행하였다.

그림 3은 인식에 이용된 보통의 표정을 가진 학습 얼굴 영상들이다. 그림 3(a)는 실험에 이용된 10인의 20개의 얼굴영상 중에서 각자 1개씩인 10개의 원 얼굴영상을 나타낸 것이며, 그림 3(b)는 이들의 평균 얼굴영상이고, 그림 3(c)는 원 얼굴영상에서 평균 얼굴영상을 뺀 영 평균된 얼굴영상들을 각각 나타낸 것이다.



(a) 원 영상 (b) 평균영상

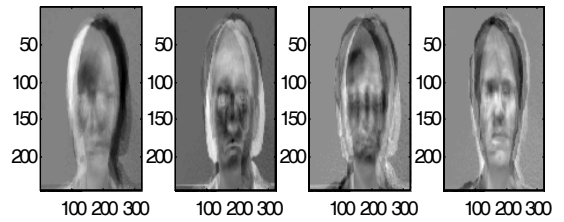


(c) 영 평균영상

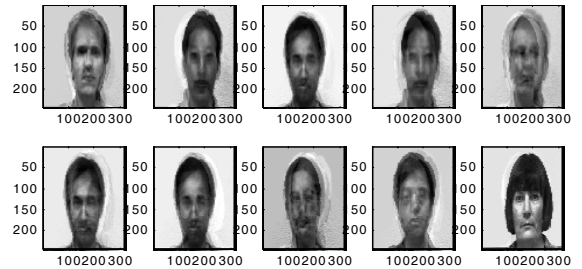
그림 3. 학습 얼굴영상

그림 4는 기저벡터에 의한 4개의 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 4(a)로부터 복원된 그림 4(b)의 얼굴영상에는 많은 잡영이 존재하여 원 얼굴영상과 다

소 차이가 있음을 알 수 있다. 이는 기저벡터의 수를 줄인 것에 대한 결과이다.



(a) 4개의 기저얼굴



(b) 10개의 복원얼굴

그림 4. 기저얼굴과 복원얼굴

한편 그림 5는 FP-ICA에 의한 인수부호를 이용한 기저벡터인 기저영상의 개수에 따른 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.

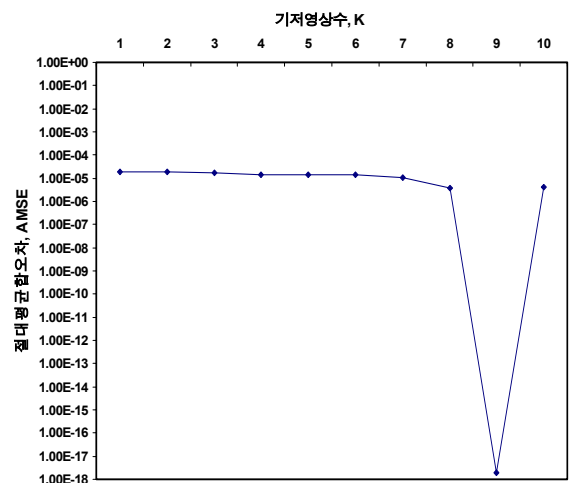
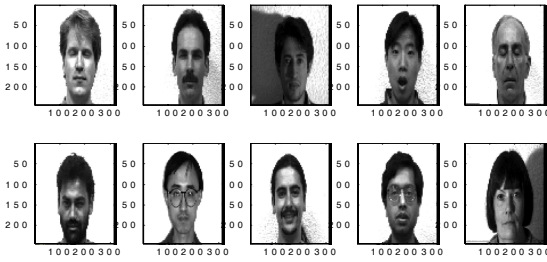


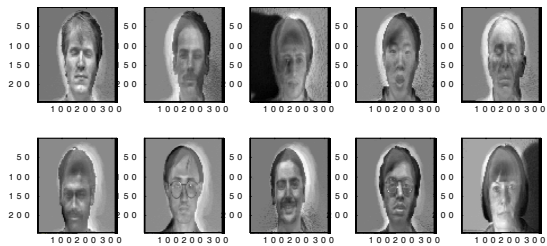
그림 5. 기저영상 개수에 따른 절대평균합오차

그림 6은 다양한 표정과 조명(놀람, 슬픔, 행복, 잠음, 윙크, 안경감, 중앙조명, 오른쪽조명, 왼쪽조명)의 변화에 따른 실험을 위해 이용된 10개의 시험 얼굴영상이다. 그림 6(a)는 원 시험 얼굴영상이며, 그림 6(b)는 원 시험영상에서 학습영

상의 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 1은 10개의 3(c)의 학습 얼굴영상과 그림 5(b)의 시험 얼굴영상을 각각 인식한 결과이다. 여기서는 이미 저장된 학습얼굴과 시험 얼굴사이의 유사성을 측정하기 위해 L1-norm과 L2-norm의 2가지 거리척도를 이용하였다. 2가지 척도에서 이미 학습된 얼굴영상과 새로운 시험 얼굴영상은 각각 100%와 90%의 인식률을 가짐을 알 수 있다. L1-norm과 L2-norm의 거리척도를 이용한 경우의 인식결과에 따른 오차값은 학습 얼굴영상은 각각 약  $10^{-12}$  -  $10^{-13}$ 정도의 우수한 인식성능을 가지나 상대적으로 시험 얼굴영상은 약 10 정도의 오차가 있음을 알 수 있다. 하지만 시험 얼굴영상은 이 정도의 오차를 가지지만 조명의 영향을 받는 3번 시험영상은 9번 학습영상으로 오인식된 경우를 제외하고는 모두 잘 인식되었다. 한편 L2-norm의 거리척도를 이용한 분류가 L1-norm에 비해 상대적으로 약 10배 정도 정확하게 분류되었다. 또한 시험의 결과에서는 영상의 조명등의 주변 환경이 인식에 큰 영향을 미침도 추측할 수 있다.

표 2. 10개의 학습과 시험 얼굴영상 각각에 대한 인식결과

거리척도	입력영상	인식률(%)	오차값	오인식 시험영상
L1-norm	학습얼굴	100	11.6634e-013	·
	시험얼굴	90	64.2417	3(9)
L2-norm	학습얼굴	100	44.7076e-014	·
	시험얼굴	90	27.151	3(9)

\* i(j) : i = 바르게 인식되어야 할 학습 얼굴번호,  
j = 잘못 인식된 학습 얼굴번호

이상의 실험결과에서 고정점 알고리즘의 독립성분석을 이용하여 얼굴영상의 인수부호를 찾아 인식하는 기법은 새로운 시험 얼굴영상에 대해 약 90%이상의 인식률을 가지며, 인식을 위한 유사성 측정에서 거리척도로 L2-norm이 L1-norm 보다 우수한 인식률이 있음을 알 수 있다.

#### 4. 결론

본 논문에서는 뉴우턴에 기반한 고정점 알고리즘의 독립성분석을 이용하여 얼굴 내의 통계적 독립인 인수를 찾아 얼굴을 인식하는 기법을 제안한다. 여기서 뉴우턴법의 FP 알고리즘은 엔트로피 최적화로부터 유도된 빠른 특징추출을 위함이며, 독립성분석의 이용은 통계적으로 독립인 개수로 구성된 인수부호를 효과적으로 추출하기 위함이다.

제안된 기법을 Yale 얼굴영상 데이터베이스로부터 선택된 20개의 324\*243 픽셀의 영상을 대상으로 시뮬레이션한 결과, 주요성분의 개수에 따른 압축성능과 L1-norm과 L2-norm의 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. S. Bartlett, "Face Image Analysis by Unsupervised Learning and Redundancy Reduction," Ph.D. Dissertation, Univ. California-San Diego, La Jolla, 1998
- [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>