논문 2006-43SP-4-6

# NMF와 LDA 혼합 특징추출을 이용한 해마 학습기반 RFID 생체 인증 시스템에 관한 연구

## ( A Study on the RFID Biometrics System Based on Hippocampal Learning Algorithm Using NMF and LDA Mixture Feature Extraction )

오 선 문*, 강 대 성*

( Sun-Moon Oh and Dae-Seong Kang )

요 약

최근 각종 온라인 상거래 및 개인 신분카드 이용이 늘어나면서 개인 인증의 중요성이 부각되고 있다. RFID(Radio Frequency Identification) tag가 내장된 개인 신분 카드가 점차 증가하고 있지만, 본인의 인증을 할 수 있는 방법이 미비하기 때문에, 자동화 할 수 있는 대책이 시급하다. RFID tag는 현재 메모리 용량이 매우 작기 때문에, 개인의 생체정보를 저장하기 위해서는 효율적인 특징추출 방법이 필요하며, 저장된 특징들을 비교하기 위해서는 새로운 인식방법이 필요하다. 본 논문에서는 인간의 인지학적인 두뇌 원리인 해마 신경망을 공학적으로 모델링하여 얼굴 영상의 특징 벡터들을 고속 학습하고, 각 영상의 최적의 특징을 구성할 수 있는 해마 신경망 모델링 알고리즘을 이용한 개인생체 인증 시스템에 관한 연구를 수행하였다. 시스템은 크게 NMF(Non-negative Matrix Factorization)와 LDA(Linear Discriminants Analysis) 혼합 알고리즘을 이용한 특징 추출 부분과 해마신경망을 모델링하고 인식 성능을 실험하는 것으로 구성 되어 있다. 제안한 시스템의 성능을 평가하기 위하여 실험은 표정변화와 포즈변화가 포함된 이미지를 각각 구분하여 인식률을 확인하였다. 실험 결과, 본 논문에서 제안하는 특징 추출 방법과 학습 방법을 다른 방법들과 비교하였을 때, 학습시간비용과 인식률에서 우수함을 확인하였다.

Abstract

Recently, the important of a personal identification is increasing according to expansion using each on-line commercial transaction and personal ID-card. Although a personal ID-card embedded RFID(Radio Frequency Identification) tag is gradually increased, the way for a person's identification is deficiency. So we need automatic methods. Because RFID tag is very small storage capacity of memory, it needs effective feature extraction method to store personal biometrics information. We need new recognition method to compare each feature. In this paper, we studied the face verification system using Hippocampal neuron modeling algorithm which can remodel the hippocampal neuron as a principle of a man's brain in engineering, then it can learn the feature vector of the face images very fast and construct the optimized feature each image. The system is composed of two parts mainly. One is feature extraction using NMF(Non-negative Matrix Factorization) and LDA(Linear Discriminants Analysis) mixture algorithm and the other is hippocampal neuron modeling and recognition simulation experiments confirm the each recognition rate, that are face changes, pose changes and low-level quality image. The results of experiments, we can compare a feature extraction and learning method proposed in this paper of any other methods, and we can confirm that the proposed method is superior to the existing method.

**Keyword** : Hippocampus, Brain Modeling, Face recognition, NMF(Non-negative Matrix Factorization), RFID

## I. 서 론

오늘날 우리나라에서 개발되고 있는 생체 인식 시스템은 여러 가지 문제점을 지니고 있다. 일부 업체를 제외한 대다수 업체들이 소니, 아이덴틱스 등 해외 업체

* 정회원, 동아대학교 전자공학과
(Dept. of Electronic Eng., Dong-A University)
※ "본 결과물은 정보통신부의 2005년도 정보통신기초기술연구지원사업(정보통신연구진흥원)으로 수행한 연구결과입니다."
접수일자: 2005년9월5일, 수정완료일: 2006년6월19일

의 모듈 및 핵심 기술 수입에 의존하기 때문에 해마다 비싼 로열티를 지불하고 있는 실정이며, 아직까지 제품 가격이 너무 비싸다는 게 전문가들의 지적이다. 또한 여러 성능 개선 노력에도 불구하고 상용화된 생체인식 보안시스템이 사람을 잘못 인식하는 오인식률과 등록된 사람을 등록되지 않은 것으로 판정하는 오거부율이 높게 나타나고 있다는 점이다. 이러한 여러 문제점들을 해결하고 생체인식이 편리하고도 완벽한 보안을 보증하는 방법으로 정착되기 위해서 지문 데이터를 비롯한 여러 생체인식의 표준화가 필요하다. 개인 생체 정보가 저장된 RFID를 통한 생체인증 시스템을 도입하면 보안에 취약했던 모든 분야로 무궁무진한 진출이 가능하다. 비용이 저렴하고 설치도 간단하여 RFID리더기와 생체정보 입력장치만 갖춰진다면 어디서든지 쉽게 구축이 가능하다. 현재 수많은 현금지급기기에 설치가능하며 CCTV로 의존했던 빈약한 보안이 지문이라는 확실한 증거물로 범인색출이 가능하여 매년 약 400건의 금융사고가 줄어들 수 있는 범죄예방 효과도 예상된다. 또한 출입통제시스템에도 도입 가능하여 회사의 출입통제와 근태관리, 관공서, 개인가정에 보안을 책임질 수 있다.

공학부분에서 많이 응용되고 있는 신경망은 뇌의 학습-기억 활동을 모태로 다양한 분야에 응용되고 있다. 인식률의 향상과 학습 속도를 개선하기 위해서는 기존 신경망의 구조적인 문제를 해결하는 방향으로 연구되어야 한다. 최근 신경생물학 모델들은 인간의 뇌에서 신피질(neocortex)의 표현의 구성에서 해마(Hippocampus)의 역할을 중요시 하고 있다. 공학부분에서 많이 응용되고 있는 신경망은 뇌의 학습-기억 활동을 모태로 다양한 분야에 응용되고 있다. 본 연구에서는 학습 시간과 인식률 문제점들을 해결하기 위하여 인간의 인지학적인 두뇌 원리인 대뇌피질과 해마 신경망을 공학적으로 모델링하여 개인 생체 정보의 특징 벡터들을 고속 학습하고, 각 영상의 최적의 특징을 구성할 수 있는 해마 신경망 모델링 알고리즘을 개발하고, 얼굴 생체 정보를 RFID를 이용하여 인증하는 시스템에 적용하고자 한다.

## Ⅱ. 특징벡터 계산

### 1. NMF

입력 얼굴 영상 및 학습을 위한 얼굴 영상은 모두 전처리 과정을 거친 후 사용된다. 전처리 이후의 영상은 NMF 학습 과정으로 학습된 basis를 이용한 얼굴 표현기를 통해 해당 얼굴 코딩으로 변환된다. NMF는 음수가 아닌 변수들로 행렬 형태의 데이터를 분해하고 이를 통해 데이터와 변수들의 관계를 두 개의 행렬로 나타낸다. 이 두 개의 행렬을 통해 데이터와 변수들의 관계를 단순하게 표현하여 새로운 해석을 가능하게 한다[1]. 특히, NMF는 음수가 아닌 가중치만을 가지므로 베이시스의 합으로만 결과를 나타나게 된다. 이러한 특징은 감소된 공간의 축이 실제적인 의미를 갖게 하고 전체적인 결과는 부분적인 특징을 합한 것이 된다. 예를 들어, 얼굴 영상을 분석하는 데 있어서 음수가 아닌 변수들의 사용은 입, 코, 눈 등의 부분 영상의 합으로 전체 영상을 나타내도록 만든다.

NMF, PCA 모두 X 행렬을 근사적인 행렬 W(베이시스 영상들)와 H(계수)의 곱으로 분해(factorization)하는 방법을 적용한다. 따라서 X 행렬이 m × n (m은 변수의 수, n은 데이터의 수)일 때 W와 H는 각각 (m×r), (r×n) (여기서 r은 변수와 데이터가 사상되는 차원)으로 분해된다. 따라서 데이터와 변수의 관계를 설명함에 있어 감소 차원 공간이 도입된다.

NMF의 베이시스와 계수는 각각 희소 행렬이며, 계수에 있어서 비음수 값을 갖도록 벡터 분해함으로써, 베이시스 영상들이 눈, 코, 입과 같이 얼굴의 부분적인 영상의 형태로 나타나게 되고, 나아가 베이시스 영상의 더하는 결합만이 존재한다. 여기서 각각의 베이시스는 다른 형태와 위치를 가지며, 모든 베이시스가 적어도 한 영상에 사용된다. 반면에 PCA의 베이시스와 계수는 모든 부호의 값을 가질 수 있기 때문에, 베이시스 영상 간의 공제(subtraction)가 존재하며, 베이시스 영상은 시각적으로 뚜렷한 형상을 가지고 있지 않다. 각각의 베이시스에 대한 계수도 전체적으로 분표되어있다.

NMF는 희소 선형 코딩(sparse linear coding)으로써, 수식 (1)과 같이 입력 영상 행렬 X를 입력 영상 보다 낮은 차원을 가진 두개의 행렬 W, H로 분해하여 차원을 축소한다. 그림 1에서처럼 X가 n×m이면 W, H는 각각 n×r, r×m이 되며, W의 r개의 열은 베이시스 영상이고, H의 각 열은 입력 영상 행렬 X에 대응하는 계수 값으로 구성된다.

$$X \approx WH \tag{1}$$

WH가 X에 가장 가까워지도록 W와 H 행렬을 반복적으로 업데이트하며, 업데이트의 방향은 목적함수(Objective function) F의 지역 최대값(local maximum)이 되는 방향이다. 이 목적 함수는 최대 우도(Maximum likelihood)에 로그를 취한 값이다. 여기서 목적함수는 수식 (2)가 된다.

$$F = \sum_{i,k} [X_{ik} \log (WH)_{ik} - (WH)_{ik}] \tag{2}$$

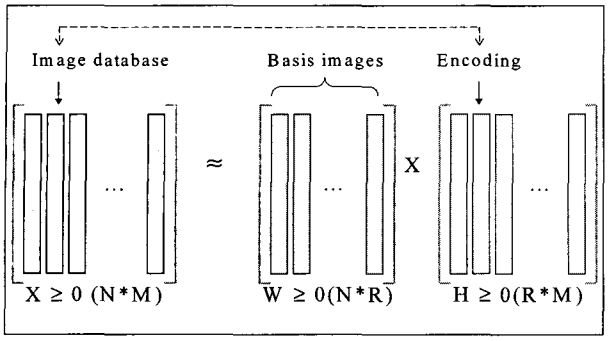



그림 1. NMF의 기본 원리
Fig. 1. A basic principle of NMF.

따라서 수식 (2)의 값이 지역 최대값이 되도록 W, H 각각을 수식 (3), 수식 (4)로 반복 수행한다[2].

$$W_{ij} \leftarrow W_{ij} \frac{\sum_k H_{jk} X_{ik} / (WH)_{ik}}{\sum_v H_{jv}} \tag{3}$$

$$H_{jk} \leftarrow H_{jk} \frac{\sum_i W_{ij} X_{ik} / (WH)_{ik}}{\sum_u W_{uj}} \tag{4}$$

특히 NMF는 PCA와 달리 W, H가 비음수 값만을 가지므로, 베이시스 영상간의 공제가 없는 베이시스 영상과 각각 대응하는 계수 곱의 선형결합으로 입력 영상을 표현한다.

## 2. LDA

LDA[3]는 PCA와 함께 데이터 분류와 차원 축소를 위하여 널리 알려진 기술 중의 하나이다. LDA 방법은 분산 벡터를 구성하는 과정에서 within-class는 최소화하고, between-class는 최대화시키는 방향으로 선형 변환을 시킨 후 분산 벡터를 학습하는 것으로 특징 공간에서 특징 벡터들의 사영들 중에서 클래스 분리를 최대화하는 선형 부공간을 찾아내는 방법을 말한다[4]. LDA의 목적은 변환 후 서로 다른 클래스에 속하는 영상간의 분산은 최대로 하고 같은 class의 영상간의 분산은 최소가 되도록 하는 변환을 구해내는 것이다[5]. 따라서 전체 얼굴 영상 벡터의 분산에 대하여 고유값 분석을 행하는 Eigenface 방법보다 인식률이 우수하다.

## 3. NMF와 LDA를 혼합한 특징 계산 방법

얼굴 표현 기법으로 PCA[3]를 적용하면 클래스와의 관계를 고려하지 못하고 전체데이터의 분산을 최대로 하는 방향으로만 선형변환하기 때문에 클래스간의 겹침이 발생할 수 있다는 단점이 있다. 이를 해결할 수 있는 LDA 방법에서는 within-class 분산 벡터와 between-class 분산 벡터를 사용하여, 클래스간의 데이터 분산은 최대가 되고, 각 클래스내의 데이터 분산은 최소가 되는 방향으로 선형 투사함으로써 보다 정확한 분류가 가능하게 된다. 알고리즘의 구현 시 within-class 분산 벡터의 역행렬 계산이 필요하게 되는데, 학습데이터의 수에 비하여 영상의 화소 개수가 많은 경우 LDA를 그대로 적용하기는 곤란하다. 본 연구에서는 전체영상의 차원을 줄이기 위해서 NMF를 먼저 적용하고 차원이 줄여진 데이터의 분류를 위하여 LDA를 사용한다.

먼저 본 논문에서 소개한 NMF 기법을 이용하여 비음수 값을 갖는 고유 벡터 행렬을 구한다. 여기서 R의 차원은 (N+M)R<NM의 범위 안에서 N=64 x 64, M=400 인 경우 R=360로 정했다. LDA는 NMF의 고유벡터 행렬을 사용하여 기저벡터를 구한 뒤 최종적으로 그것에 대한 고유 벡터 행렬을 구한다. LDA 고유 벡터의 전치행렬과 NMF 고유 벡터 전치행렬 그리고 입력 영상 행렬을 곱하면 실제 영상에 대한 가중치 벡터 값을 가지게 된다. 여기서 가중치 벡터 값이 입력 영상에 대한 고유값이 된다.

C개의 클래스를 $w_1, w_2, ..., w_c$, $i$번째 클래스에 속한 얼굴 영상의 개수를 $n_i$, NMF를 적용하기 전과 후의 $i$번째 클래스 평균을 각각 $\mu_i$와 $\widetilde{\mu_i}$라고 하면, $Y_{ij} = W_{NMF}^T X_{ij}$이고, $\mu_i = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij}$이므로 $\widetilde{\mu_i} = \frac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij} = \frac{1}{n_i}\sum_{j=1}^{n_i} W_{NMF}^T X_{ij} = W_{NMF}^T \mu_i$가 된다. 따라서 $Y_{ij}$의 within-class 분산 행렬 $S_W'$와 between-class 분산 행렬 $S_B'$는 수식 (5)과 같다. 여기서 $X_{ij}$와 $Y_{ij}$는 $W_{NMF}$에 의한 선형 변환 이전 및 이후의 $i$번째 클래스에 속한 $j$번째 얼굴 벡터를 나타낸다.

$$S_W' = W_{NMF}^T \; S_W \; W_{NMF} \;,$$
$$S_B' = W_{NMF}^T \; S_B \; W_{NMF} \tag{5}$$

수식 (6)을 적용하여 선형 변환 행렬 $W_{LDA}$을 계산하게 된다. 여기서 $W_{LDA}$의 각 열은 $(S_W)^{-1}S_B$의 가장 큰 $(C-1)$개의 고유치에 해당하는 고유벡터로서 구한다.

$$W_{LDA} = \underset{W}{argmax} \left( \frac{W^T S_B W}{W^T S_W W} \right) \tag{6}$$

결과적으로 얼굴 영상 벡터 $X_{ij}$를 $Z_{ij} = W_{LDA}^{T} W_{NMF}^{T} X_{ij}$에 의해 선형 변환된 $Z_{ij}$로 표현되어 미리 구하여진 벡터들과의 비교에 의하여 얼굴 영상을 인식하게 된다.

## Ⅲ. 해마신경망 모델과 구현

### 1. 해마신경망 학습 알고리즘 구현

해마의 기능 중 가장 중요하고 강력한 것은 단기기억에서 장기기억으로의 확장에 있다. 해마의 역할은 방금 저장된 정보를 기억하고 필요한 정보와 그렇지 못한 정보를 분류하는 것이다. 필요한 정보의 경우, 해당 뉴런에 저장해두고 동일한 패턴이 입력되면 인식하게 되는 것이다. 단기 기억과 장기 기억으로 나누는 방법에 있어서 호감도 측정은 통계적 패턴 인식에 사용되는 방법을 사용하면 통계적으로 많이 나타나는 값들에 대해서 호감도를 높게 부여하는 것이다. 이러한 특징을 해마 신경망에도 도입한다면, 입력되는 중간에 입력받는 값들의 가중치를 사용하여 단기 기억을 할 것인지 장기 기억을 할 건인지를 결정할 수 있다. 이러한 결정 값을 해마기억 장소 분류 기준 값으로 사용할 수 있다. 전체 데이터 개수 대비 해당뉴런의 누적 개수가 문턱치를 만족하면 장기 기억 장소로 저장시키도록 한다[6].

각 블록의 기능 설명은 다음과 같다.

- 내후각뇌피질(EC) : 해마와 신피질 사이에서 인터페이스를 구성하고 해마 신경망 모델의 입력과 출력을 나타낸다.
- 치아이랑(DG) : 내후각뇌피질에 직접 연결되는 치아이랑 구조는 그 모델에 동일한 객체의 다양한 특징 성분들을 과거 입력 패턴에서 동일성을 판단하여 특징을 단순화 시킨다. 패턴 평균값에서 편차율 범위의 임계수치를 넘으면 1, 그렇지 않으면 -1로 하여 특징의 동일성을 이진화 시킨다.
- CA3 : 어떤 사건을 기억해낼 때에는 연상 작용을 반복함으로써 보다 나은 결과를 얻을 수 있으며, 이러한 순환 개념을 도입한 것이 순환 연상 메모리이다. CA3 영역은 치아이랑에 연결되고 이 후자 구조로부터 단순화된 정보를 받는다. CA3 영역에서는 자기 연상을 하며, 노이즈가 있는 특징 정보를 더욱 분별력 있게 정돈하는 역할을 한다. 과거에 학습된 특징들과 차이가 큰 특징이 입력 될 경우에는 새로운 패턴으로 분류한다.

## Ⅳ. RFID 시스템

### 1. 전체 시스템

본 시스템은 크게 네 개의 블록으로 나누어 볼 수 있다. 첫 번째는 CCD 카메라를 통한 영상 입력 부 이고, 두 번째는 RF-Reader와 RF-Tag로 구성되는 RF시스템이다. 세 번째는 얼굴특징 및 개인 신상정보를 저장하고 있는 데이터베이스이고, 마지막으로 네 번째는 본 시스템의 핵심 엔진이라고 할 수 있는 응용 어플리케이션 부 이다. 본 논문에서 구현한 시스템의 동작과정 블록도는 그림 3과 같다.

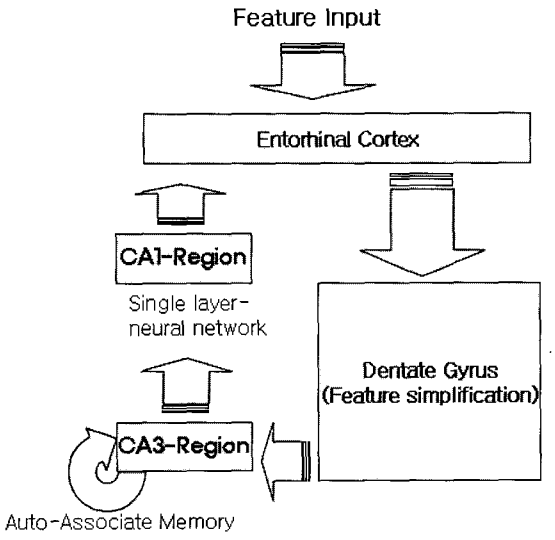



그림 2. 해마신경망 모델
Fig. 2. Hippocampal Neural Network model.

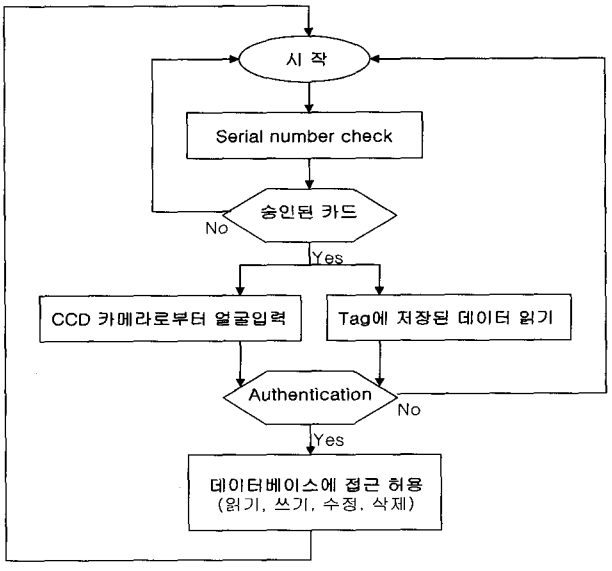



그림 3. 전체 시스템 블록도
Fig. 3. The block-diagram of RFID biometric system

본 시스템에서는 Client-Server 환경을 지원함으로서 인터넷 환경에서의 사용자 인증이 가능하도록 설계하였다. 이러한 인터넷 환경에서의 사용자인증은 보안의 문제로 인해 실제 사용자의 인증에 필요한 생체데이터가 전달되는 것은 아니고 Client PC에서 사용자 인증의 결과를 이미 도출에 놓은 상태에서 Server에 있는 데이터베이스에 저장되어 있는 사용자의 기본 정보를 주고받는 것이다. 이러한 과정은 Stand alone 시스템에서의 일반 사용에는 문제가 없으나 시스템의 장기간 사용 시 발생하는 사용자의 갱신과 여러 시스템에서의 사용자 구분환경을 구축하는데 목적을 가지고 있다. 만약 이러한 방식의 설계가 아니면 사용자에게 태그를 발급한 순간부터는 더 이상의 관리가 어렵게 된다. 그리고 많은 시스템을 관리하는 Server로서의 입장을 생각해서 사용자 인증시의 프로세싱 최대 부하는 Client로 하였고, 네트워크 환경에서 Server에 걸리는 부하는 최소로 줄이기 위하여 사용자의 기본정보만 전달하여 확인하는 방식을 사용하였다.

## 2. RFID 시스템

RFID 시스템은 기본적으로 그림 4와 같이 태그, 리더, 데이터베이스(응용 어플리케이션)로 구성된다. RFID 시스템에서 사용하는 주파수의 경우 크게 세 가지 정도로 나눌 수 있다. 각각의 용도에 따라 100-500㎑, 10-15㎒, 850-950㎒, 2.5-5.8㎓ 정도로 나눌 수 있는데 그 중 일찍부터 관심을 끌었던 주파수는 125㎑, 13.56㎒, 2.45㎓이다. 13.56㎒의 경우 주요 RFID 제조화사들이 자사의 태그에 이를 채택하고 있고 ISO 15693에 지정되어 있는 등, 전 세계적으로 표준화가 진행 중에 있다. 주파수의 특성에 따라 태그의 데이터를 읽어들이는 속도나 거리 그리고 가격의 차이가 발생하며 주파수가 높을수록 이 세 가지 특성이 모두 높아진다[7]. 각각의 기능은 다음과 같다.

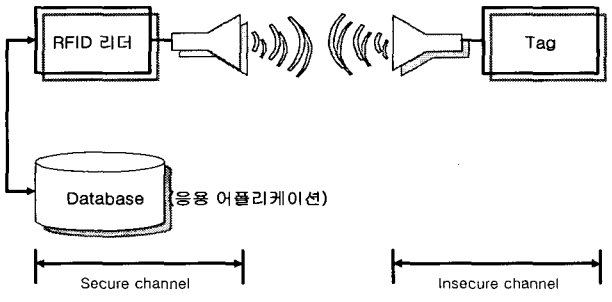



그림 4. RFID 시스템
Fig. 4. The RFID system.

### 가. 태그(Tag)

리더의 요청에 응답하는 트랜스폰더(Transponder)로써 태그는 각각의 고유한 시리얼 번호를 저장하고 있다. 태그는 실제 물체에 부착하는 것으로 칩과 안테나를 가지고 있다. 적게는 64비트에서 많게는 8천 비트 정도까지 정보를 담을 수 있다. 본 논문에서는 카드 타입에 부착된 것을 사용하였고, 좀 더 구체적으로는 1Kbyte(1024×8)의 용량을 가지는 ISO 14443-A 형인 Phillips사의 Mifare 태그 IC를 사용하였다. 태그의 EEPROM은 16개의 섹터로 이루어져 있으며, 각 섹터는 16bytes 단위의 4개의 블록으로 구성되어 있다. 섹터 0번의 블록 0번은 IC 생산자 데이터가 포함되어 있다. 각각의 섹터는 16byte 단위의 블록 4개로 구성되며, 4번째 블록은 데이터 보호를 위한 키 값을 저장하게 된다.

### 나. 리더(Reader)

태그에 정보를 요청하고 수신한 데이터를 판독하고 태그를 인식하는 트랜시버(Transceiver). 리더는 태그에게 RF 신호를 통해 태그의 전원을 공급하는 역할을 한다. 본 논문에서는 13.56㎒ 대역을 사용하는 GEMPLUS사의 GemEasyLink 680SP 를 사용하였다.

# V. 얼굴 인식 실험

## 1. 실험 자료

### (1) ORL 데이터베이스

ORL 데이터베이스[8]는 Cambridge의 Olivetti Research Laboratory에서 획득한 얼굴 영상들로써(그림 5 참조), 40명에 대하여 일정한 어두운 배경에서 얼굴 표정과 안경 착용 유무에 따라 각 10장씩 구성되어 있다. 특히 ORL 데이터베이스는 안경·수염 등이 있는 얼굴들이 많이 분포 되어 있고 좌우 회전을 통해 변형된 얼굴 영상도 다소 존재한다.

그림 5. ORL Database의 예
Fig. 5. Some examples of ORL database.

(2) CMU AMP Face EXpression Database

표정변화 실험을 위하여 CMU AMP Face EXpression Database 얼굴 DB를 사용하였다. Carnegie Mellon University의 Advanced Multimedia Processing Lab[9]에서 만든 얼굴 DB로 동일한 조명조건에서 정면 얼굴의 다양한 표정을 지원한다. 그림 6은 CMU 얼굴 DB의 일부 얼굴 영상의 일부를 보여준다.

그림 6. CMU AMP Face EXpression Database의 예
Fig. 6. Some examples of CMU AMP Face EXpression Database.

(3) DAUFace 데이터베이스

제안한 시스템의 인식 성능 비교를 위해 자체 제작한 DAUFace 데이터베이스는 그림 7과 같이 화상 통신용 저가의 CCD 카메라를 가지고 일정한 환경에서 다양한 포즈와 표정 변화로 검출된 얼굴 영상이다. 각 얼굴 영상은 8-bit gray level의 64×64 크기로 되어있으며, 총 400개(20명×20개)로 구성되어 있다.

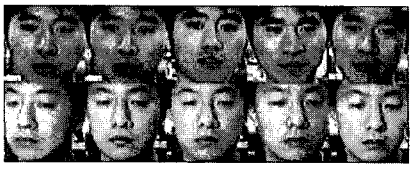


그림 7. DAUFace Database의 예
Fig. 7. Some examples of DAUFace database.

## 2. 실험 방법

얼굴을 표현한 후에는 특정 방법에 의하여 표현된 입력 얼굴 영상을 기존에 표현된 영상과의 비교 과정을 거치게 되는데 이것이 분류 과정이다. 이 분류 과정을 통하여 입력 얼굴 영상이 어느 얼굴의 범주에 속하는가를 결정하며 이것이 바로 얼굴 인식 시스템의 출력이 된다. 얼굴을 분류하는 방법은 Euclidean 거리 측정에 의한 특징 벡터의 거리를 측정하는 Nearest Neighbor 방법으로 구하거나, BP, SVM 등 각종 신경망과 HMM 같은 분류기를 이용하여 구할 수 있다. 그러므로 본 논문의 실험에서는 해마신경망 알고리즘의 성능을 평가하기 위하여 주로 위와 같은 학습 방법들과 비교하는 방식으로 수행하였다. 앞 절에서 소개한 Database들을 사용하여 해마신경망 학습 모델의 얼굴 인식 실험을 수행하였는데, 크게 얼굴 포즈변화와 표정변화 그리고 자체 제작한 저화질 영상에서 포즈 및 표정변화로 나누어 얼굴 인식 실험을 수행하였다.

본 논문에서는 입력 얼굴 영상 및 학습을 위한 얼굴 영상은 영상 내의 얼굴의 크기와 위치를 맞추는 전처리 과정을 통하여 얼굴 인식이 제대로 수행되게 하였다. 얼굴 인식에서 가장 중요한 특징이라 할 수 있는 눈의 위치를 맞추도록 하였다. 얼굴 영상의 각 눈의 가운데 두 점을 영상 내 지정 위치에 오도록 2D 유사 변환(affine transformation)을 수행하였다. 그 뒤 배경의 영향을 없애기 위하여 윈도우 영역을 정의하여 얼굴을 잘라내었다. 영상을 축소하는 down sampling 과정에서 경계선들의 정보를 유지하기 위해 잘라낸 얼굴 영상에 번짐 효과(blurring)를 주고 64x64 픽셀 크기로 down sampling을 수행하였다. 마지막으로 영상을 gray level로 변환하였다.

전처리 이후의 영상은 NMF와 LDA 과정으로 학습된 basis를 이용하여 해당 얼굴 코딩으로 변환된다. 이 basis를 이용하여 코딩으로 변환하는 과정은 원래의 입력 공간의 subspace로 변환하는 과정으로 볼 수 있으며, 이것은 해마신경망 학습 알고리즘의 입력으로 사용되고, 입력 얼굴 코딩을 학습된 얼굴 코딩들과 비교하여 가장 비슷한 얼굴을 찾아낸다.

## 3. 실험 결과

해마신경망 학습 알고리즘의 인식률 비교를 위하여, leaving one out 방법을 사용하였는데, leaving one out은 n 개의 표본이 있을 때, n-1 개의 표본으로 학습 하고 나머지 하나로 테스트하는 절차를 n번 반복하는 것으로써, 표본 내의 모든 케이스가 테스트에 사용되어 신뢰성 있는 결과를 획득할 수 있으나 연산 비용이 비싸다는 단점이 있다.

(1) basis 개수 변화에 따른 실험 결과

NMF와 LDA의 얼굴 인식 성능 분석을 위하여 ORL, CMU, DAUFace 데이터베이스를 이용하여, basis 수를 변화해 가며 얼굴 인식 실험을 수행하였다. 본 실험은 NMF와 LDA를 혼용한 basis 수의 변화에 따른 인식 성능을 비교하여, 해당 기법의 최적의 basis 수를 찾는다. 32, 64, 96, 128 개로 basis 수를 변화해가며 유클라

디안 방법으로 인식률이 가장 우수한 것을 선택하도록 한다. 유클라디안 방법을 선택한 이유는 학습 알고리즘의 성능에 의존하지 않고 특징 벡터의 자체 분리도를 평가하기 위해서이다. 표 1을 보면 공통적으로 basis의 수가 32, 64, 128개일 때 보다는 96개일 때 인식 성능이 더 좋은 것을 알 수 있다. 표 1을 보면 NMF와 LDA를 혼용하는 것만으로도 좋은 성능을 보이고 있지만, 해마 신경망 학습 알고리즘의 성능을 평가하기 위하여 다른 학습 알고리즘과 비교를 통하여 성능을 평가하도록 한다. 학습 알고리즘의 입력으로써 basis를 최적의 개수인 96개로 정하도록 한다.

클래스 분리도를 정량적으로 측정하기 위하여 NMF와 LDA를 혼용한 방법과 LDA 각각에 대하여 $\frac{W^T S_B W}{W^T S_W W}$ 값을 구하였다. ORL 데이터베이스와 CMU 데이터베이스, DAUFace 데이터베이스에 대한 클래스 분리도 값을 표 2에 나타내었다.

표 2의 값이 클수록 클래스 내의 분산 값은 작고, 다른 클래스 간의 분산 값이 크기 때문에 각각의 얼굴 영상 클래스를 보다 잘 분리할 수 있다. 표 2를 살펴보면, 제안한 방법이 LDA보다 높음에 따라 변별력이 증가하였음을 알 수 있다. 이처럼 NMF와 LDA의 혼용은 NMF 알고리즘을 이용하여 차원을 효과적으로 축소하고, 이러한 축소 특징 벡터 공간에 LDA를 적용함으로 클래스 분리도가 상당히 향상되어, 인식률 크게 향상시킬 수 있었다.

표 1. Basis 수 변화에 따른 인식 결과(%)
Table 1. Recognition result with change of basis number(%).

| DB / basis | ORL | CMU AMP | DAUFace |
|---|---|---|---|
| 32개 | 84.03 | 88.34 | 87.95 |
| 64개 | 86.35 | 89.36 | 88.35 |
| 96개 | 91.34 | 92.62 | 91.78 |
| 128개 | 87.56 | 90.66 | 90.24 |

표 2. NMF와 LDA 혼용의 클래스 분리도 측정
Table 2. The class separability measurement of mixed method of NMF and LDA .

| Database | LDA | NMF+LDA |
|---|---|---|
| ORL database | $5.85\times10^{17}$ | $8.91\times10^{18}$ |
| CMU databas | $8.72\times10^{17}$ | $9.43\times10^{18}$ |
| DAUFace database | $7.54\times10^{17}$ | $8.79\times10^{18}$ |

(2) CA3의 연상 메모리 성능 확인

해마에서 CA3영역의 연상 기억 성능을 평가하기 위해서 CMU AMP 데이터베이스 전체를 인식을 위한 training set으로 하여 실험을 하였다. 그림 8은 대표 얼굴의 Edge 영상이며, 그림 9에서 표정이 다른 얼굴영상이 자기연상 수행 후에 대표 Edge영상으로 고쳐지는 것을 확인할 수 있다. 본 결과는 임의의 패턴이 CA3 연상과정으로 결과로 학습과정에 참여 했던 패턴으로 연상이 정확히 되는 것을 보여준다.

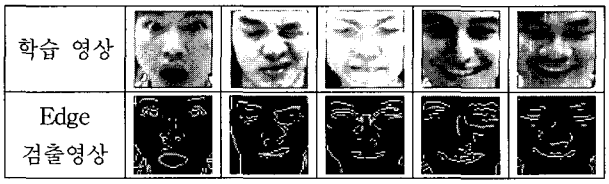



그림 8. 각 얼굴의 고유 Edge 검출 영상
Fig. 8. Edge Detection Images of each face images.

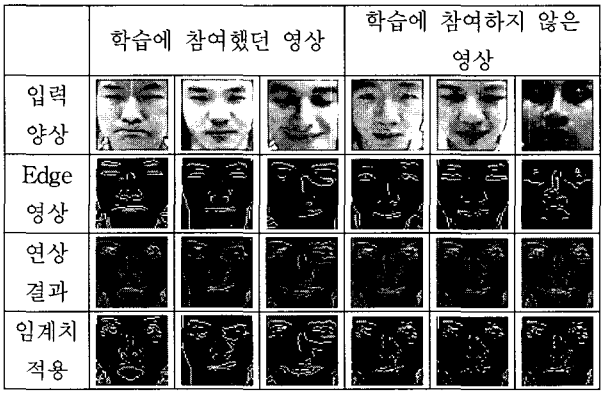



그림 9. 해마의 CA3의 연상기억모델 적용결과
Fig. 9. Result of Auto-associate of CA3 in Hippocampus.

(3) 포즈변화에 따른 실험 결과

현재 SVM이나 HMM등과 같이 성능 좋은 분류기들을 응용한 연구들이 많지만, 이들 학습알고리즘 또한 영상전체를 입력 패턴으로 사용하진 않기 때문에, 계산된 입력 패턴이 주어져야 한다. 100% 분별력이 있는 패턴이 존재한다면 학습이 필요한 분류기 또한 필요하지 않겠지만, 현실에서는 그렇지 못하므로, 학습 알고리즘이 꾸준히 연구되어 오고 있으며, 그와 동시에 분별력이 우수한 특징패턴을 구성하는 것이 중요하다고 할 수 있다.

포즈변화에 따른 얼굴 인식 성능을 비교하기 위하여 ORL 데이터베이스의 40명의 영상에서 각각 10가지의 포즈를 해마신경망 학습 알고리즘에 이용하고, 학습에 이용하지 않은 나머지 1개의 영상들로 인식 성능을 측

표 3. 학습 알고리즘과 특징 계산 방법에 따른 포즈 인식률 비교(%)

Table 3. Comparisons of pose recognition rates with learning algorithms and feature caculation method(%).

| 학습방법 / 특징계산방법 | Nearest Neighbor | BP | SVM | HMM | Hippocampus |
|---|---|---|---|---|---|
| NMF | 87.85 | 90.66 | 90.89 | 91.96 | 92.42 |
| LDA | 89.04 | 90.74 | 91.85 | 92.34 | 93.03 |
| PCA | 88.65 | 90.15 | 91.51 | 92.09 | 92.15 |
| NMF+LDA | 91.34 | 93.21 | 93.88 | 94.21 | 97.94 |

정하였다. 먼저 NMF와 LDA를 혼용한 특징 벡터 사용 방법의 인식 성능을 확인하기 위하여, NMF, LDA, PCA를 단독으로 사용하여 계산한 특징 벡터들과 비교 하였다. 그 결과는 표 3과 같다.

표 3에서와 같이 NMF와 LDA를 혼용한 것은 Nearest Neighbor 방법에서도 91.34%의 가장 높은 인식률을 보이고 있고, 다른 학습 방법들에서도 단독으로 사용될 때보다 인식률을 향상시키는 것을 보면, NMF와 LDA를 혼용한 방법으로써 계산한 특징 벡터는 각종 학습 알고리즘의 좋은 학습 패턴으로 이용되어 질 수 있다는 것을 알 수 있다.

(4) 표정변화에 따른 실험 결과

표정변화에 따른 얼굴 인식 성능을 비교하기 위하여 CMU 데이터베이스의 13명의 영상에서 각각 50가지의 포즈를 해마신경망 학습 알고리즘에 이용하고, 학습에 이용하지 않은 나머지 영상들로 인식 성능을 측정하였다. 먼저 NMF와 LDA를 혼용한 특징 벡터 사용방법의 인식 성능을 확인하기 위하여, NMF, LDA, PCA를 단독으로 사용하여 계산한 특징 벡터들과 비교하였다. 그 결과는 표 4와 같다.

표 4. 학습 알고리즘과 인식률 비교(%)

Table 4. Comparisons of recognition rates with learning algorithms(%).

| 학습방법 / 특징 계산방법 | Nearest Neighbor | BP | SVM | HMM | Hippocampus |
|---|---|---|---|---|---|
| NMF | 89.00 | 92.56 | 92.68 | 92.63 | 92.42 |
| LDA | 91.04 | 93.65 | 93.85 | 94.25 | 95.03 |
| PCA | 91.56 | 93.81 | 93.51 | 93.88 | 95.65 |
| NMF+LDA | 93.55 | 95.36 | 96.45 | 96.12 | 99.10 |

표 5. DAUFace 데이터베이스를 이용한 알고리즘별 얼굴 인식 실험 결과(%)

Table 5. The simulation result of facial recognition using DAUFace database.

| 학습방법 / 특징계산방법 | Nearest Neighbor | BP | SVM | HMM | Hippocampus |
|---|---|---|---|---|---|
| NMF | 86.63 | 88.15 | 89.51 | 90.09 | 91.15 |
| LDA | 88.31 | 89.65 | 90.87 | 91.21 | 92.45 |
| PCA | 87.00 | 89.47 | 90.86 | 91.96 | 91.09 |
| NMF+LDA | 90.84 | 92.20 | 92.11 | 94.34 | 96.53 |

(5) 자체 데이터베이스에서의 실험 결과

저화질 화상캠으로 획득한 얼굴 영상으로 구축한 DAUFace 데이터베이스를 이용하여 앞의 실험들과 같은 방법으로 성능을 확인 해 보았다. DAUFace 데이터베이스는 실시간 시스템에 적용할 때를 고려하여 비슷한 표정으로 약간의 포즈 변화를 주도록 하였다. 인식 결과는 표 5와 같다.

표 5의 결과를 보면 실시간 시스템의 적용에 학습 속도가 빠르고, 인식률이 우수한 해마신경망 학습 알고리즘을 적용할 경우 효과가 좋을 것으로 판단된다.

## VI. 결 론

NMF와 LDA는 모두 비교사 학습 기법으로 NMF는 조밀한 코딩, LDA는 드문 코딩에 의하여 인식에 있어 필요한 중복 신호 제거 효과를 갖기 때문에 얼굴 영상의 특징 벡터로써 많이 연구되어 왔다. 본 논문에서는 이 두 기법을 혼용하고, 해마신경망 학습 알고리즘을 이용한 얼굴 인식 모델을 설계하고 얼굴 인식 실험을 수행하여 인식 특징 및 성능을 분석하였고, 다른 학습 방법들과의 비교 실험을 통하여 보다 면밀한 분석을 수행하였다. basis 수를 조절해가며 얼굴 인식을 수행한 실험에서 96개의 basis 갯수가 가장 성능이 좋은 것으로 보아 basis수가 너무 적거나 많을 경우 성능에 영향을 준다는 사실도 알 수 있었다. 그리고 얼굴 인식 실험을 통하여 학습방법들을 비교했을 때, 포즈와 표정에 따른 얼굴 인식 실험에서 제안한 알고리즘이 97% 이상의 인식 성능을 나타내는 등, 소규모 인식에서 다른 학습 방법보다 높은 인식 성능을 보였다. 또한 계산량이 NMF와 LDA를 혼용한 방법이 훨씬 적으므로 단독으로 사용하였을 때 인식률 측면에서 성능이 좋은 LDA보다가 오히려 효과적인 결과를 얻었다. 그리고 3~5 번 째

답까지를 포함한 인식에서 해마신경망 학습 알고리즘이 다른 학습 방법보다 높은 인식률을 얻는 현상이 발견되었고, 이러한 점으로 인하여 해마신경망 학습 알고리즘이 우수하다라는 결론을 얻을 수 있었다.

본 논문에서 강조하고 싶은 내용은 얼굴 영상의 인식을 위한 특별한 시스템에 관한 설계방향이다. 기존의 시스템의 경우 조명의 변화나 얼굴의 가려짐, 그리고 표정 변화를 갖는 얼굴 영상을 인식하기 위해서 NMF, PCA, LDA등을 개별적으로 적용하였기 때문에 우수한 인식률을 기대하기 어려웠다. 본 논문에서는 정규화된 영상에 NMF와 LDA를 혼용하여 차원을 축소한 후 해마신경망 학습 알고리즘을 사용하여 학습한 후 인식을 가능하게 하였으며, 분석된 결과를 종합해 볼 때, 얼굴 인식에 있어서 NMF와 LDA를 혼용하는 방법이 단독으로 사용될 때 보다 더 나은 기법으로 볼 수 있다. 그리고 단순한 벡터 거리 비교 방법 보다는 학습 알고리즘을 이용하였을 때, 인식 속도나 성능 면에서 우수하다고 할 수 있겠다.

## 참 고 문 헌


[1] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization", Nature, vol.401, pp.788-791, 1999.

[2] Daniel D. Lee, H. Sebastian Seung, "Algorithms for Non-negative Matrix Factorization", NIPS, 2001

[3] R.Duda, P.Hart, D.Stork: "Pattern Classification - Second Edition", John Wiley&Sons, 2001.

[4] K. Etemad and R. Chellappa, "Discriminant Analysis For Recognition of Human Face Images" JI Optical Society of America Vol.14 aug. 1997. p1724-1733

[5] P. N. Belhumeour, J. P. Hespanha, D. K. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", IEEE Trans. On PAMI, 19(7):711-720, 1997.

[6] Stephen A. Weis, "Security and Privacy in Radio-Frequency Identification Devices",MIT, May 2003.

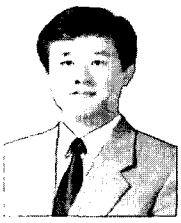
[7] Qin YL, McNaught on BL, Skaggs WE., Barnes CA(1997) "Memory reprocessing in corticocortical and hippocampocortical neuronal ensembles", Phil Trans R Soc Lond B Biol Sci 352: 1525-1533.

[8] Homepage of the AT&T Laboratories at Cambridge : http://www.uk.research.att.com/facedatabase.html/

[9] Homepage of the Advanced Multimedia Processing Lab in the Carnegie Mellon University : http://amp.ece.cmu.edu/


## 저 자 소 개

오 선 문(정회원)
2004년 동명정보대학교 컴퓨터공학과 학사 졸업.
2006년 동아대학교 전자공학과 석사 졸업.
<주관심분야 : 영상처리, 패턴인식, 신호처리>



강 대 성(정회원)
1984년 경북대학교 전자공학과 학사 졸업.
1991년 Texas A&M 대학교 전자공학과 석사 졸업.
1994년 Texas A&M 대학교 전자공학과 박사 졸업.
1995년~현재 동아대학교 전자컴퓨터공학부 부교수.
<주관심분야 : 영상처리, 비디오 색인 및 검색, 영상 압축>