

# 얼굴인식을 위한 거리척도학습 방법 비교

뱃수리수브다<sup>†</sup>, 고재필<sup>\*\*</sup>

## 요 약

얼굴인식과 같이 클래스의 수가 변하는 분류 문제에는 학습이 필요하지 않은  $k$ -최근접이웃 분류기가 적합하다. 최근 학습 데이터의 분포를 반영하여 거리 척도를 학습하는 방법은  $k$ -최근접이웃 분류기의 획기적 성능향상을 보고하였다. 거리척도학습 방법은 적용 분야에 따라 성능 개선 정도가 다르다. 본 논문에서는 얼굴인식에 대하여 주요 거리척도학습 방법의 성능을 비교한다. 공개 얼굴 데이터베이스에 대한 실험 결과는 성능 및 계산시간 측면에서 주성분 분석 기반의 마하라노비스 거리척도가 얼굴인식 문제에서는 여전히 좋은 선택이 될 수 있음을 보여준다.

## A Comparison of Distance Metric Learning Methods for Face Recognition

Batsuri Suvdaa<sup>†</sup>, Jaepil Ko<sup>\*\*</sup>

## ABSTRACT

The  $k$ -Nearest Neighbor classifier that does not require a training phase is appropriate for a variable number of classes problem like face recognition. Recently distance metric learning methods that is trained with a given data set have reported the significant improvement of the  $k$ NN classifier. However, the performance of a distance metric learning method is variable for each application. In this paper, we focus on the face recognition and compare the performance of the state-of-the-art distance metric learning methods. Our experimental results on the public face databases demonstrate that the Mahalanobis distance metric based on PCA is still competitive with respect to both performance and time complexity in face recognition.

**Key words:** Distance Metric Learning(거리척도학습), Mahalanobis Distance(마하라노비스거리),  $k$ -Nearest Neighbor Classifier( $k$ NN 분류기), Face Recognition(얼굴인식)

## 1. 서 론

얼굴인식은 분류할 클래스가 많고 학습 단계에 주어지는 샘플도 해당 클래스의 일부에 지나지 않으며 샘플의 차원도 높다. 이러한 문제의 어려움으로 인해 지지벡터머신(support vector machine)[1]과 같은

복잡한 분류기가 사용되고 있다[2]. 이진 분류기인 SVM을 얼굴인식과 같은 다중 분류 문제에 적용하기 위해서는 다수의 이진 분류기를 구성하고 각각을 학습해야 한다[3]. 클래스의 수가 변하면 분류기 구성과 학습 과정을 다시 수행해야 한다.

얼굴인식은 클래스 수가 변하는 문제이므로 학습

\* 교신저자(Corresponding Author): 고재필, 주소: 경북 구미시 양호동 1번지, 금오공과대학교 디지털관 136호(730-701), 전화: (054)478-7529, FAX: (054)478-7539, E-mail: nonezero@kumoh.ac.kr

접수일: 2011년 4월 5일, 수정일: 2011년 4월 27일  
완료일: 2011년 5월 12일

<sup>†</sup> 준회원, 금오공과대학교 컴퓨터공학과  
(E-mail: suvda@hotmail.com)

<sup>\*\*</sup> 정회원, 금오공과대학교 컴퓨터공학과

\* 이 논문은 2008년도 교육인적자원부 학술연구조성사업비로 한국학술진흥재단의 지원을 받아 연구되었음(KRF-2008-331-D00537).

단계가 없는  $k$ -최근접이웃 (kNN) 분류기가 적합하다. kNN 분류기는 단순하면서도 복잡한 결정 경계를 만들 수 있으며, 클래스의 수가 많은 경우 테스트 시간을 줄이기 위한 트리 기반의 알고리즘도 존재한다[4].

kNN 분류기는 주어진 샘플을 학습 집합 중에서  $k$ 개의 가까운 이웃을 뽑아 이들의 다수 클래스로 분류한다. 주어진 샘플의 이웃 샘플은 이들의 유클리드 거리에 달려있다. 유클리드 거리는 주어진 학습 집합의 분포를 반영하지 못한다. 최근 이를 반영하여 거리 척도를 학습하는 방법이 kNN 분류기의 획기적인 성능 향상을 보고하고 있다[5-9]. 국내에서는 지문인식에 적용한 사례가 발표되었다[10].

거리척도학습 방법은 거리척도 관련 비용 함수의 정의와 최적화 문제를 푸는 방법에 따라 구분된다. 또한, 같은 방법이라도 적용 분야에 따라 성능 개선 정도에는 차이가 난다[7].

본 논문에서는 얼굴인식에 대하여 최신 거리척도 학습 방법의 성능을 비교한다. 본 논문의 실험 결과는 성능 및 계산시간 측면에서 주성분 분석 기반의 마하라노비스(Mahalanobis) 거리척도가 얼굴인식에서 여전히 좋은 선택이 될 수 있음을 보여준다.

본 논문의 구성은 다음과 같다. 2장에서는 거리척도학습 방법에 관하여 거리척도학습과 대표적 방법을 차례대로 설명한다. 3장에서는 공개 얼굴 데이터 베이스에 대한 성능평가실험 결과를 보이며, 4장에서는 실험결과에 대해 토의한다.

## 2. 거리척도학습 방법

### 2.1 거리척도학습

거리척도학습은 크게 교사 및 비교사 거리척도 학습 방법으로 구분된다[11]. 본 논문에서는 클래스를 고려한 교사 거리척도학습 방법을 주로 다룬다. 교사 거리척도학습의 목적은 같은 클래스에 속한 샘플들의 쌍과 다른 클래스에 속한 샘플들의 쌍에 대하여, 전자는 가깝게 후자는 멀게 측정되도록 거리 척도를 학습하는 것이다.

차원이  $N$ 인 벡터  $x \in R^N$ 들 간의 거리척도로 널리 사용되는 유클리드 거리를 위한 함수  $d$ 는 다음과 같이 정의된다.

$$d(x_1, x_2) = \|x_1 - x_2\|_2^2 = (x_1 - x_2)^T (x_1 - x_2) \quad (1)$$

유클리드 거리는 단위행렬  $\Sigma = I$  을 도입하여 다음과 같이 표현할 수 있다.

$$d(x_1, x_2) = \|x_1 - x_2\|_2^2 = (x_1 - x_2)^T \Sigma (x_1 - x_2) \quad (2)$$

단위행렬은 대각이 1이고 주변이 모두 0이므로, 유클리드 거리는 변수들의 분산이 모두 같고 변수들 간의 상관관계가 없는 경우에 적합하다. 변수들 간의 분산과 공분산을 고려하여 거리척도를 정규화 시키면 다음과 같다. 이를 마하라노비스 거리라 한다.

$$d_M(x_1, x_2) = (x_1 - x_2)^T \Sigma^{-1} (x_1 - x_2) \\ \Sigma = \Sigma_{ij} (x_i - \bar{x})(x_j - \bar{x})^T \quad (3)$$

여기서  $\bar{x}$ 는 입력 벡터들의 샘플 평균을 나타낸다. 마하라노비스 거리를 양의준정부호(positive semi-definite) 행렬  $A$ 로 일반화시키면 거리 척도를 위한 함수는 다음과 같다.

$$d_A(x_1, x_2) = (x_1 - x_2)^T A (x_1 - x_2) \quad (4)$$

만약  $A$ 가 양의준정부호행렬이 아니면  $d_A(x_1, x_2) \leq 0$ 이 될 수 있으므로 거리 척도로 적합하지 않다. 거리척도학습은 거리척도학습 목적에 부합하는 행렬  $A$ 를 구하는 것이다. 한편,  $A$ 는  $L^T L$ 로 분해되어 다음과 같이 전개된다.

$$d_A(x_1, x_2) = (x_1 - x_2)^T A (x_1 - x_2) = (x_1 - x_2)^T L^T L (x_1 - x_2) \\ = (Lx_1 - Lx_2)^T (Lx_1 - Lx_2) = \|Lx_1 - Lx_2\|_2^2 \quad (5)$$

따라서 거리척도학습은 변환공간에서의 유클리드 거리가 입력공간에서 의미거리가 되도록 하는 변환  $L$ 을 구하는 것과 같다. 이러한 이유로 다양체학습(manifold learning) 방법은 거리척도학습으로 분류될 수 있다.

### 2.2 비교사 학습 방법

차원축소에 효과적인 주성분 분석도 거리척도 학습 방법으로 간주될 수 있다. 주성분 분석은 학습 집합  $\{x_i | x_i \in R^N\}_{i=1}^n$ 을 분산이 최대가 되는 특정 공간으로 투영하는 선형 변환  $L: x \rightarrow y$ 을 구한다. 주성분 분석에서 선형 변환  $L$ 은 다음과 같이 제약이 있는 최적화 문제로 계산된다.

$$\max_L \text{trace}(L^T CL) \quad \text{s.t. } LL^T = I \quad (6)$$

여기서  $C_x = \Sigma_i^n (x_i - \bar{x})^T (x_i - \bar{x}) / n$ 는 학습 집합의 공분산 행렬이다. 변환  $L$ 은  $C_x$ 의 고유벡터를 행으로 구성한다. 이때, 고유벡터의 배치 순서는 고유값의 크기( $\lambda_1 \gg \lambda_2 \gg \dots \gg \lambda_n$ )에 따르며 상위  $K(\ll N)$ 개의 고유벡터만을 사용하는 경우 차원축소에 효과가 있다. 또한,  $L$ 을 통해 변환된 특징 공간에서의 각 변수  $y_1, y_2, \dots, y_K$ 들의 상관계수가 0이 되는 특징이 있다. 이러한 특징으로 인해  $k$ NN 분류기는 입력 벡터  $x$ 들 간의 유클리드 거리 대신 주성분 분석을 통해 변환된 벡터  $y$ 들 간의 유클리드 거리를 활용하여 간단하게 성능 향상을 달성한다.

$$d_{EucPCA}(x_1, x_2) \cong \|y_1 - y_2\|_2^2 = (y_1 - y_2)^T (y_1 - y_2) \quad (7)$$

한편 특징 공간에서의 샘플 평균을  $\bar{y}$ 라고 할 때 공분산 행렬  $C_y = \Sigma_i^n (y_i - \bar{y})^T (y_i - \bar{y}) / n$ 의 대각은  $C_x$ 의 고유값이고 그 이외는 모두 0이 된다. 따라서 마하라노비스 거리는 다음과 같이 고유값의 역수로 간단하게 계산된다.

$$d_{MaharPCA}(x_1, x_2) \cong (y_1 - y_2)^T C_y^{-1} (y_1 - y_2) \\ = \sum_{k=1}^K \frac{1}{\lambda_k} (y_k^{(1)} - y_k^{(2)})^2 \quad (8)$$

주성분 분석은 학습 집합의 지역적 분포를 고려하지 않고 클래스를 구분하지 않으므로 전역적 비교사 방법으로 분류된다[11].

### 2.3 교사 학습 방법

#### 2.3.1 전역적 교사 학습 방법

교사학습으로서의 거리척도학습 방법은 거리척도에 관한 목적함수를 볼록 최적화(convex optimization) 문제로 간주하여  $k$ -평균 군집화 성능을 개선한 방법에서 시작되었다[5]. 이들이 제시한 목적함수와 제약은 다음과 같다.

$$\min_A \sum_{(x_i, x_j) \in S} d_A(x_i, x_j) \\ s.t. A \geq 0, \sum_{(x_i, x_j) \in D} d_A(x_i, x_j) \geq 1 \quad (9)$$

여기서  $S$ 는 같은 클래스에 속한 모든 샘플 쌍에 대한 집합이고,  $D$ 는 서로 다른 클래스에 속한 모든 샘플 쌍에 대한 집합을 나타낸다. 모든 샘플 쌍을 고려하므로 이 방법은 전역적 방법으로 분류된다. 전역

적 방법은 각 클래스가 단일 모달(modal) 임을 가정하므로 다중 모달인 경우에는 효과가 없다. 이러한 문제는 샘플 쌍을 구성할 때 지역성을 고려하는 지역적 방법으로 개선될 수 있다.

#### 2.3.2 지역적 교사 학습 방법

지역적 교사학습방법으로는 이웃 샘플을 확률적으로 선택하는 방법을 제안한 이웃성분분석(neighbor component analysis)[6]과 목적함수에 SVM의 마진 개념을 도입한 넓은마진최근접이웃(large margin nearest neighbor)[7]이 대표적이다.

##### 1) 이웃성분분석(NCA)

일반적으로  $k$ NN 분류기의 성능 평가를 위한 테스트 에러는 leave-one-out(LOO) 교차검증으로 계산한다. NCA는 LOO 에러가 최소가 되도록 하는 거리척도를 구한다. 이때, 이웃 샘플을 확률적으로 선택하는 방법[12]을 도입하여 미분이 가능한 LOO 에러 관련 식을 유도하였다. 특징 공간에서 샘플  $x_i, x_j$ 가 이웃일 확률  $p_{ij}$ 는 다음과 같이 정의된다.

$$p_{ij} = \frac{\exp(-\|Lx_i - Lx_j\|_2^2)}{\sum_{k \neq i} \exp(-\|Lx_i - Lx_k\|_2^2)}, \quad p_{ii} = 0 \quad (10)$$

식 (10)에서 주목해야 할 것은 명시적으로  $k$ 개의 이웃을 결정해야 하는 기준과는 달리 이웃의 개수가  $L$ 에 의해 간접적으로 자연스럽게 결정된다는 점이다. 식 (10)으로부터 주어진 샘플  $x_i$ 가 올바르게 분류될 확률  $p_i = \sum_{j \in C_i} p_{ij}$ 가 유도된다. 여기서  $C_i = \{j \mid c_j = c_i\}$ 는 샘플  $x_i$ 와 같은 클래스에 속한 샘플  $x_j$  집합을 나타낸다. 이제 NCA에 의한 선형 변환  $L$ 은 올바르게 분류될 샘플의 평균 개수를 최대화 하는 문제로 표현된다.

$$\max_L \sum_i p_i = \sum_i \sum_{j \in C_i} p_{ij} \quad (11)$$

이 식은  $L$ 에 대해 미분 가능하므로 변화율 기반 알고리즘으로 풀 수 있다. NCA는 데이터 분포에 대한 어떠한 가정도 하지 않기 때문에 샘플 분포가 멀티 모달인 경우에도 적용할 수 있다. 그러나 목적함수가 볼록 최적화 문제는 아니므로 지역 해에 빠질 수 있다.

##### 2) 넓은마진최근접이웃(LMNN)

LMNN은 마진개념을 도입하여 주어진 샘플에 대

한  $k$ 개의 인접이웃에 대하여 같은 클래스에 속한 이웃은 보다 근접하게 되고, 그렇지 않고 다른 클래스에 속한 이웃은 멀어지도록 손실함수를 다음과 같이 정의하였다.

$$\begin{aligned} \epsilon(L) &= \epsilon_{pull}(L) + \epsilon_{push}(L), \\ \epsilon_{pull}(L) &= \sum_{i,j \in C_i} \|I(x_i - x_j)\|_2^2, \\ \epsilon_{push}(L) &= \sum_{i,j \in C_i, i \neq j} (1 - y_{ij}) \\ &\quad [1 + \|I(x_i - x_j)\|_2^2 - \|I(x_i - x_i)\|_2^2]_+ \end{aligned} \quad (12)$$

여기서  $[z]_+ = \max(z, 0)$ 이고  $y_{ij}$ 는 샘플  $x_i$ 와  $x_j$ 의 클래스가 동일한 경우 1로 그렇지 않은 경우 0로 설정한다. 수식 (12)는  $L$ 에 대해 볼록이 아니므로  $L$ 대신  $A = L^T L$ 를 사용하여 볼록 최적화 문제를 수립한다.

$$\begin{aligned} \min_A \quad & \sum_{i,j \in C_i} (x_i - x_j)^T A (x_i - x_j) + \\ & \sum_{i,j \in C_i, i \neq j} (1 - y_{ij}) \\ & [1 + (x_i - x_j)^T A (x_i - x_j) - (x_i - x_i)^T A (x_i - x_i)]_+ \\ \text{s.t.} \quad & A \geq 0 \end{aligned} \quad (13)$$

LMNN을 제안한 논문에서는 다양한 데이터에 대한 실험결과를 제시하였으며 다중 클래스 SVM과 대등한 우수한 성능을 보였다. 또한 이 방법은 커널을 도입하여 비선형 방법으로 확장되었다[13]. 그러나 본 방법은 특징공간에서가 아니라 입력공간에서 이웃의 개수를 명시적으로 정의해 주어야 하는 문제를 안고 있다.

### 2.3.3 에너지 함수 기반 방법

에너지 기반 방법[8]은 얼굴검증에 적용된 방법으로 기하학적인 변형에 강인하도록 설계된 컨벌루션(convolution) 신경망을 이용하여 입력공간을 특징공간으로 변환한다. 이때 사용된 변환을  $G_W$ 라 하면, 입력벡터  $x_i, x_j$ 에 대한 에너지 함수는 다음과 같이 정의된다.

$$E_W = \|G_W(x_i) - G_W(x_j)\|_2^2 \quad (14)$$

같은 클래스에 속한 모든 샘플 쌍에 대한 집합을  $S$ , 서로 다른 클래스에 속한 모든 샘플 쌍에 대한 집합을  $D$ 라고 하고,  $S$ 와  $D$ 의 원소에 대한 각각의 손실함수를  $L_S, L_D$ 라 하자. 그러면 에너지 함수를 기

반으로 하는 손실함수는 다음과 같이 정의될 수 있다. 컨벌루션 신경망은 이를 최소화하도록 학습한다.

$$L(W, x_i, x_j) = (1 - Y)L_S(E_W(x_i, x_j)) + YL_D(E_W(x_i, x_j)) \quad (15)$$

여기서  $Y$ 는  $(x_i, x_j)$  쌍이  $S$ 의 원소일 때는 1,  $D$ 의 원소일 때는 0으로 하여 주어진 쌍에 대응되는 손실함수만 고려되도록 한다. 또한  $S$ 의 원소에 대해서는 거리가 짧게 나오는 것이 손실이 작고,  $D$ 의 원소에 대해서는 거리가 멀게 나오는 것이 손실이 작기 때문에 손실함수는  $E_W$ 에 대해  $L_S$ 와  $L_D$ 는 각각 증가함수와 감소함수가 되도록 한다.

에너지 기반 방법은 변화율 기반의 학습방법인 컨벌루션 신경망을 이용하여 기하학적 변형에 강인한 비선형 방법으로, 볼록 최적화에 기반을 두지 않으므로 학습 데이터의 수에 제약받지 않는다.

## 3. 실험결과

얼굴인식 성능평가는 2장에서 설명한 마하라노비스 거리최도학습 방법과 에너지 기반 방법을 대상으로 하며, 수식 (7)의 유클리드 거리 기반 방법을 성능평가의 기준으로 삼는다. 실험에 사용한 데이터는 얼굴인식에서 널리 사용되는 얼굴 데이터베이스인 ORL[14], PIE[15], 확장 Yale B[16]이다. 얼굴영상은 차원이 크기 때문에 주성분 분석을 적용하여 입력벡터의 차원을 200차원으로 줄였으며, LMNN에서는 이웃의 개수  $k$ 를 3으로 설정하여 실험을 수행하였다. 본 논문에서 제시한 실험결과는 입력벡터의 차원을 30차원 이내로 한정한다. 우리의 초기 실험에서 50차원 이상의 경우 MahaNCA의 성능이 급격히 떨어졌으며 나머지 방법은 200차원까지 인식을 상승이 거의 없었기 때문이다.

단순하고 일관성 있는 표기를 위해 주성분 분석 기반의 유클리드 거리는 EucPCA, 주성분 분석 기반의 마하라노비스 거리는 MahaPCA, LMNN 기반의 마하라노비스 거리는 MahaLMNN, NCA 기반의 마하라노비스 거리는 MahaNCA, 그리고 컨벌루션 신경망 기반의 에너지 함수 방법은 Energy라고 표기한다.

### 3.1 데이터베이스

ORL 데이터베이스는 40명에 대해 다양한 변화를

담고 있는 10장씩의 얼굴영상으로 총 400장으로 구성되어 있다. 얼굴영역만 담고 있는 각 영상의 크기는 112×92이다. PIE 데이터베이스로부터 우리는 15명에 대해 총 1,782장의 얼굴영상을 사용한다. 각 영상은 포즈, 조명, 표정 변화를 갖고 있다. 원 영상의 크기는 400×600인데 얼굴영역만 300×300으로 자른 후 다시 보간법을 적용하여 50×50으로 크기 조절을 하였다. 확장 YaleB 데이터베이스로부터 우리는 38명에 대한 1,368장의 얼굴영상을 사용한다. 한 명당 9개의 포즈, 4방향의 얇은 조명변화를 갖고 있다. 원 영상으로부터 얼굴영역만 54×54 크기로 잘라내어 사용한다. 표 1에 실험에 사용한 데이터베이스를 요약하였다. 모든 데이터베이스는 주성분 분석을 통해 차원을 줄였으며 실험에서는 30차원까지의 결과를 보인다. 별도의 전처리는 적용하지 않았다.

3.2 얼굴인식 오류 비교 및 분석

각 데이터베이스는 학습과 테스트를 위해 무작위로 7:3 비율로 나누었고 이를 10번 수행하여 평균오

류와 표준편차를 계산하였다. 전체 실험결과는 표 2에 제시한다. 사용한 차원에 따른 각 방법의 최저 에러를 밑줄로 표시하였다.

표 2에 따르면 성능 평가의 기준이 되는 EucPCA보다 나머지 거리척도학습 방법의 성능이 모두 높은 것을 확인할 수 있다. 각 방법의 최저 오류를 비교해보면, ORL에 대해서 MahaLMNN 및 Energy 방법이, PIE에 대해서는 MahaPCA가, YaleB에 대해서는 MahaNCA가 EucPCA에 비해 2배 이상의 성능향상을 보였다.

전반적으로 PIE 및 YaleB에 대한 성능은 ORL에 비해 현저히 낮게 나왔다. 이는 해당 데이터베이스가 상당한 포즈변화와 조명변화를 갖고 있기 때문으로 보인다. 거리척도학습 방법이 기존의 방법에 비해 상대적으로 성능 향상을 보이거나 여전히 포즈나 조명에는 취약함을 나타낸다.

차원변화에 따른 오류를 보면, 차원이 증가함에 따라 MahaNCA를 제외하고 대체로 오류가 줄어드는 것을 확인할 수 있다. MahaNCA는 가장 낮은 오

표 1. 실험데이터 요약

데이터베이스	클래스 수	변화	영상 크기	사용차원	영상 수
ORL	40	다양한 변화	112×92	1-35	400
PIE	15	포즈, 조명, 표정 변화	50×50	1-35	1,782
YaleB	38	포즈, 조명 변화	54×54	1-35	1,368

표 2. 얼굴인식 오류 비교

얼굴영상	차원	EucPCA	MahaLMNN	MahaNCA	MahaPCA	Energy
ORL	5	22.6±2.4	21.9±3.9	9.9±3.0	21.4±3.6	27.1±3.0
	10	11.9±2.0	8.3±1.7	4.4±2.4	7.2±1.9	6.5±2.5
	15	8.7±1.1	6.1±1.6	5.2±2.9	7.5±1.7	4.5±1.1
	20	8.2±1.6	5.0±1.5	9.8±4.0	6.2±1.6	4.1±1.1
	25	7.5±1.9	3.2±1.2	13.5±3.9	6.4±2.4	3.2±1.3
	30	6.2±2.0	3.0±1.2	15.3±4.4	6.4±2.3	3.0±1.1
PIE	5	47.7±19.0	49.9±15.2	37.3±18.3	47.3±17.6	60.4±7.0
	10	37.4±18.2	37.2±14.7	36.4±18.7	32.3±13.4	44.6±10.3
	15	33.8±17.3	31.4±12.8	27.7±12.8	22.8±8.8	34.0±11.1
	20	32.6±16.8	27.8±11.7	27.2±11.6	19.2±7.8	30.5±10.8
	25	31.0±16.2	25.4±11.6	26.8±10.6	16.6±7.4	26.3±12.2
	30	29.3±15.5	23.7±11.8	26.1±10.4	16.4±7.8	26.6±11.4
YaleB	5	91.6±1.0	92.1±2.0	42.1±8.7	91.2±1.9	90.7±1.7
	10	83.6±2.6	76.4±8.4	30.4±8.0	70.6±5.4	74.2±10.0
	15	75.4±2.9	63.2±12.0	26.6±9.9	43.5±4.6	61.6±12.6
	20	64.8±3.9	55.8±13.3	23.5±9.7	28.8±7.2	55.4±13.6
	25	57.6±4.7	43.3±13.7	36.1±7.8	33.0±10.0	42.6±14.3
	30	53.2±5.6	41.5±17.7	48.3±13.2	33.0±13.0	40.9±17.6

류의 차원이 타 방법과는 차이를 보여 차원 선정에 주의가 요구된다. MahaNCA는 볼록 최적화 문제가 아니므로 실험에 따라서는 지역 해에 빠질 수 있다. 이를 반영하듯 ORL에서는 타 방법에 비해 표준편차가 높게 나타났다.

실험결과에서 특이한 점은 최저 오류 측면에서 MahaPCA가 클래스를 고려하지 않는 비교사 거리척도학습 방법임에도 타 방법과 경쟁할 만한 오류를 보이고 있다는 점이다. 이러한 경향은 전체적으로 성능이 가장 낮은 25차원에 대한 그림 1의 그래프에서 더욱 명확히 드러난다. 그림 1에서 MahaPCA는 PIE 및 YaleB에서 가장 낮은 오류를 보여준다.

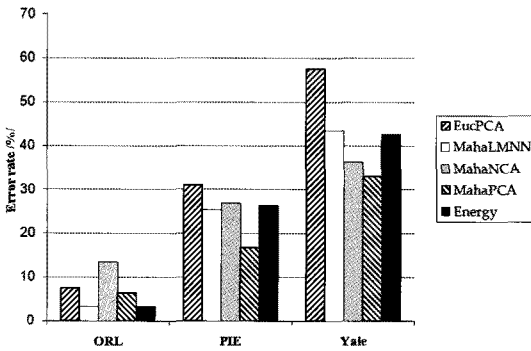


그림 1. 25차원 데이터에 대한 얼굴인식 오류 비교

### 3.3 최저 오류에 대한 비교

우리의 실험에서는 얼굴 데이터베이스별로 가장 낮은 오류를 달성한 거리척도학습 방법이 다르게 나왔다. 표 1에 따르면 ORL에서는 MahaLMNN과 Energy, PIE에서는 MahaPCA, YaleB에서는 MahaNCA가 각각 최저 오류를 보였다. 본 절에서는 각 거리척도학습 방법별 최저 오류에 대한 상관관계와 최저 오류에 대한 종합적인 그래프를 통해 이들의 우위를 비교한다. 각 거리척도학습 방법 간의 상관관계는 표 3에 제시한다.

표 3. 최저 오류에 대한 거리척도학습 방법 간의 상관관계

방 법	EucPCA	MahaLMNN	MahaNCA	MahaPCA	Energy
EucPCA	1	-	-	-	-
MahaLMNN	0.325	1	-	-	-
MahaNCA	0.336	0.547	1	-	-
MahaPCA	0.257	0.356	0.866	1	-
Energy	0.192	0.588	0.452	0.416	1

표 3에서 상관계수가 0.5 이상인 경우만 보면, MahaNCA와 MahaPCA가 가장 유사하고 그 다음은 MahaLMNN과 Energy, 그리고 MahaNCA와 MahaLMNN이 차례대로 그 뒤를 따른다.

그림 2는 거리척도학습 방법별로 최저 오류에 대한 최대값, 최소값, 중앙값, 평균값을 동시에 표시한 다. 그림에서 막대의 상단은 최대값, 하단은 최소값, 가로선은 중앙값, 십자원은 평균값을 나타낸다.

그림 2에 따르면 최대 및 최소의 좁은 범위와 낮은 평균값 측면에서 MahaNCA 및 MahaPCA가 가장 우수함을 알 수 있다. 이 둘만 비교하면 최소값 측면에서는 MahaNCA, 중앙값 측면에서는 MahaPCA가 우수하다고 볼 수 있다. 표 3에서 MahaPCA와 Energy 및 MahaLMNN의 상관계수가 0.5이하이므로 MahaPCA의 성능은 이들에 비해 두드러진다고 볼 수 있다. 이는 유의수준 95%의 t-검증을 통해서도 확인된다.

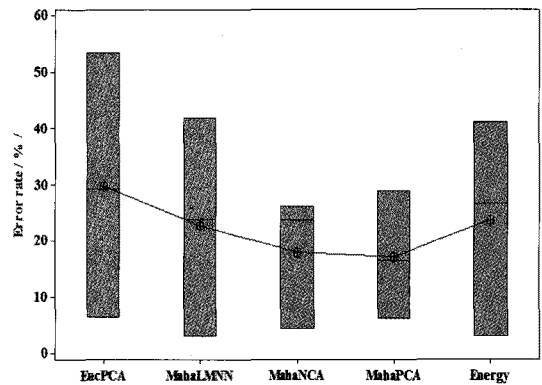


그림 2. 최저 오류에 대한 거리척도학습 방법의 오류 비교

### 3.4 학습시간 비교

얼굴영상의 차원을 30으로 고정하였을 때 각 방법의 학습시간을 비교하였다. 표 4는 학습시간 측정결과를 보여준다. 학습에 사용한 얼굴영상 수는 전체

표 4. 거리척도학습 방법의 학습시간 비교

방 법	ORL(280장)	PIE (1,260장)	Yale B (956장)
EucPCA	0.015±0.003	0.20±0.42	0.13±0.012
MahaLMNN	4.014±1.215	28.48±6.37	12.30±2.86
MahaNCA	3.8±0.324	38.80±0.632	21.41±4.42
MahaPCA	0.015±0.003	0.20±0.42	0.13±0.012
Energy	0.25±0.05	0.51±0.63	0.31±0.25

영상의 70%이다.

학습시간은 EucPCA(MahaPCA)가 가장 작고 그 뒤를 Energy가 이었다. 지역적 방법으로 분류되는 MahaLMNN 및 MahaNCA는 이들에 비해 상대적으로 학습시간이 오래 걸렸다. MahaLMNN은 볼록 최적화 해법에 기반을 두기 때문에 학습 샘플의 수가 늘어남에 따라 처리 시간이 급격하게 늘어났다. 이웃의 개수를 높이면 처리시간은 더욱 늘어날 수 있다. MahaNCA는 모든 가능한 쌍에 대하여 수식 (11)을 계산해야 함에 따라 MahaLMNN과 마찬가지로 학습 샘플의 수가 늘어남에 따라 처리 시간이 급증한다. 이에 반해 신경망 기반의 Energy 방법은 학습 샘플의 수에 따른 처리시간의 차이가 거의 나지 않는다. 표 4의 실험결과는 각 방법론의 이러한 특징을 뒷받침한다.

표 4의 수치에 따르면 MahaPCA는 타 방법에 비해 상대적으로 월등히 빠른 학습시간을 보인다. 예를 들어, ORL에서 MahaNCA에 약 250배, MahaLMNN에 약 260배, Energy에 약 16배 빠르게 나타났다.

#### 4. 결 론

본 논문에서는 거리척도학습 방법을 교사/비교사, 전역적/지역적, 볼록최적화/변화율기반최적화 등으로 구분하여 대표적인 방법론을 설명하였다. 거리척도함수의 성능향상을 확인하기 위해 주성분 분석을 통한 유클리드 거리를 성능평가 기준으로 삼았다.

ORL, PIE, YaleB 얼굴 데이터베이스에 대한 우리의 실험결과는 최고 성능의 거리척도학습 방법이 유클리드 거리에 비해 kNN 분류기의 성능을 오류 측면에서 두 배 가량 향상시키는 것을 확인하였다. 그러나 포즈와 조명의 변화가 큰 데이터베이스인 PIE 및 YaleB에서는 거리척도학습 방법이 절대적인 수준에서는 여전히 포즈나 조명에는 취약함을 보여주었다.

거리척도학습 방법별로 최저 오류를 비교한 결과

MahaPCA가 MahaNCA와 동등한 수준의 최저오류를 보였다. 학습시간 측면에서는 MahaPCA가 타 방법에 비해 월등히 빠른 학습시간을 보였다. 결론적으로, 우리의 실험결과는 성능 및 학습시간 측면에서 MahaPCA가 얼굴인식에서는 여전히 경쟁력 있음을 보여주었다. 본 방법은 개체기반 인식방법[17]의 성능향에 적용될 수 있다.

#### 참 고 문 헌

[1] B. Schölkopf and A. J. Smola, "Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond, MIT Press, 2002.

[2] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Computing Surveys*, Vol.35, No.4, pp.399-458, 2003.

[3] E. Allwein, R. Schapire, Y. Singer, and P. Kaelbling, "Reducing Multiclass to Binary: A Unifying Approach for Margin Classifiers," *Journal of Machine Learning Research*, Vol.1, No.1, pp.113-141, 2000.

[4] J. Friedman, J. Bentley, and R. Finkel, "An algorithm for finding best matches in logarithmic," *ACM Trans. on Mathematical Software*, Vol.3, No.3, pp.209-226, 1997.

[5] E. Xing, A. Y. Ng, M. I. Jordan, and S. Russell, "Distance Metric Learning, with Application to Clustering with Side-Information," *Advances in Neural Information Processing Systems*, Vol.14, pp.521-528, 2002.

[6] J. Goldberger, S. Roweis, G. Hinton, and R. Salakhutdinov, "Neighbourhood Components Analysis," *Advances in Neural Information*

- Processing Systems*, Vol.17, pp.513-520, 2005.
- [7] K. Q. Weinberger and L. K. Saul, "Distance Metric Learning for Large Margin Nearest Neighbor Classification," *Journal of Machine Learning Research*, Vol.10, pp.207-244, 2009.
- [8] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a Similiarty Metric Discriminatively, with Application to Face Verification," Proc. of Int'l Conf. on Computer Vision and Pattern Recognition, pp.349-356, 2005.
- [9] J. Davis, B. Kulis, S. Sra, and I. Dhillon, "Information-Theoretic Metric Learning," Proc. of Int'l Conf. on Machine Learning, Vol.227, pp.209-216, 2007.
- [10] D. Jang and C. Yoo, "Fingerprint Matching Based on Distance Metric Learning," Proc. of Int'l Conf. on Acoustics, Speech and Signal Processing, pp.1529-1532, 2009.
- [11] L. Yang, R. Jin, L. Mummert, R. Sukthankar, A. Goode, B. Zheng, S. Hoi, and M. Satyanarayanan, "A Boosting Framework for Visuality-Preserving Distance Metric Learning and Its Application to Medical Image Retrieval," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.32, No.1. pp.30-44, 2008.
- [12] G. Hinton and S. T. Roweis, "Stochastic Neighbor Embedding," *Advances in Neural Information Processing System*, Vol.15, pp.833-840, 2002.
- [13] L. Torresani and K. C. Lee, "Large Margin Component Analysis," *Advances in Neural Information Processing Systems*, Vol.19, pp.1385-1392, 2007.
- [14] ORL face database: <http://www.uk.research.att.com/facedatabase.html>.
- [15] PIE face database: <http://vasc.ri.cmu.edu/idb/html/face>.
- [16] Yale Face Database B: <http://cvc.yale.edu/projects/yalefaceB/yalefacesB.html>.
- [17] D. Song and C. Lee, "Instance Based Learning Revisited: Feature Weighting and Its Applications," *Journal of Korea Multimedia Society*, Vol. 9, No. 6, pp.762-772, 2006.



**박수리수브다**

2004년 6월 국립중앙대학교 전자공학과 석사  
 2007년 9월~현재 금오공과대학교 컴퓨터공학과 박사과정  
 관심분야: 패턴인식, 영상처리



**고재필**

2004년 2월 연세대학교 컴퓨터과학과 박사  
 2004년 3월 현재 금오공과대학교 컴퓨터공학과 부교수  
 관심분야: 패턴인식, 영상처리