

Ridge Regressive Bilinear Model을 이용한 조명 변화에 강인한 얼굴 인식 (Illumination Robust Face Recognition using Ridge Regressive Bilinear Models)

신 동 수 [†] 김 대 진 ^{**} 방 승 양 ^{***}
(Dong Su Shin) (Dai Jin Kim) (Sung Yang Bang)

요 약 얼굴 인식 시스템의 성능은 조명 변화로 인하여 발생하는 개인내 (intra-person) 차이가 개인 간 (inter-person) 의 차이보다 클 수 있기 때문에 조명 변화에 많은 영향을 받는다. 본 연구에서는 이러한 문제를 해결하기 위해서 대칭형 bilinear 모델을 이용하여 조명 요소와 신원 요소를 분리하는 방법을 제안한다. Bilinear 모델로 조명 요소와 신원 요소를 얻기 위한 translation 과정은 반복적 역행렬을 구하는 것이 요구되는데 입력 데이터에 따라 수렴하지 않는 경우가 발생할 수 있다. 이러한 문제를 완화하기 위해서 ridge regression 모델과 bilinear 모델을 결합한 ridge regressive bilinear 모델을 제안하였다. 제안된 모델은 조명 요소와 신원 요소의 분산을 적절히 줄여줌으로서 bilinear 모델에 안정성을 제공하며, 인식에 더 많은 고차원 요소 정보를 이용하게 함으로써 인식 성능을 높여 준다. 실험 결과에서 제안한 ridge regressive bilinear 모델이 bilinear 모델, 고유얼굴(eigenface) 방법, Quotient image 보다 좋은 인식 성능을 보여줌을 확인 할 수 있다.

키워드 : 얼굴 인식, 조명에 강인한 얼굴 인식, 접선형 모델, 능선 회귀 모형, 능선 회귀적 접선형 모델

Abstract The performance of face recognition is greatly affected by the illumination effect because intra-person variation under different lighting conditions can be much bigger than the inter-person variation. In this paper, we propose an illumination robust face recognition by separating identity factor and illumination factor using the symmetric bilinear models. The translation procedure in the bilinear model requires a repetitive computation of matrix inverse operation to reach the identity and illumination factors. Sometimes, this computation may result in a nonconvergent case when the observation has an noisy information. To alleviate this situation, we suggest a ridge regressive bilinear model that combines the ridge regression into the bilinear model. This combination provides some advantages: it makes the bilinear model more stable by shrinking the range of identity and illumination factors appropriately, and it improves the recognition performance by reducing the insignificant factors effectively. Experiment results show that the ridge regressive bilinear model outperforms significantly other existing methods such as the eigenface, quotient image, and the bilinear model in terms of the recognition rate under a variety of illuminations.

Key words : Face Recognition, Illumination Robust Face Recognition, Bilinear Model, Ridge Regression, Ridge Regressive Bilinear Model

1. 서 론

최근, 얼굴 인식은 생체인식분야에서 가장 활발한 연

구 분야 중의 하나이다. 1970년대 시작한 이후로, 많은 문제들이 해결되었으며, 인식 성능이 향상 되었다. 하지만, 조명에 강인한 얼굴 인식은 아직 남은 난제 중의 하나이다. 사람을 구분하는 기준으로서 다른 사람들과의 개인간(inter-person) 차이를 이용한다. 하지만, 조명이 변화에 따라 개인내(intra-person)의 차이가 개인간 차이보다 커질 수 있다. 이러한 문제를 해결하기 위해서 많은 접근 방법이 개발 되었다.

[†] 정 회 원 : LG전자 전자기술원

dsshin06@lge.com

^{**} 정 회 원 : 포항공과대학교 컴퓨터공학과 교수

dkim@postech.ac.kr

^{***} 종신회원 : 포항공과대학교 컴퓨터공학과 교수

sybang@postech.ac.kr

논문접수 : 2006년 2월 3일

심사완료 : 2006년 11월 20일

Riklin-Raviv[1]는 조명에 무관한 특성(feature)으로서 Quotient image 를 제안하였으며, 실험 결과에서 고유 얼굴(Eigenface) 보다 높은 인식 성능을 나타냄을 보였다. 그러나 Quotient image 는 입력 이미지에 그림자가 있는 경우 좋은 특성을 얻을 수 없게 된다. Zhou 와 Chellappa[2]는 Quotient image 에서 이상적인 클래스는 같은 모형을 지니고 있다는 가정으로 발생하는 제한 사항을 제거하기 위해서 알베도(albedo)와 표면 법선 벡터(surface normal)에 rank 제약을 부과하였다. Zhao[3]는 입력 이미지를 참조 이미지(reference image)와 비슷한 조명 환경의 이미지로 합성하는 Illumination ratio image를 제안하였다. Lui[4]는 참조 이미지와 가우시안 필터(Gaussian filter)를 거친 이미지를 이용하여 임의의 조명 환경에서 생성된 이미지를 정면 조명을 갖는 이미지로 변경하는 방법을 제안하였다. Shan 과 Gao[5]은 조명 효과를 제거하기 위해서 히스토그램 균일화(histogram equalization), 감마 보정(gamma intensity correction) 그리고 Quotient image 를 영역 기반 전략(region-based strategy)으로 결합한 방법을 제안 하였다. Belhumeur[6-8]는 다양한 조명 환경에서 사람의 얼굴을 인식하기 위한 생산적 외양 기반(generative appearance-based) 방법을 제안하였다. 이들은 모든 가능한 조명 환경에서 생성된 이미지들은 이미지 공간에서 convex cone을 이루며, 복원된 알베도와 형상(shape)은 새로운 조명 환경에서의 이미지를 합성하기 위한 생산적 모델(generative model)로 사용된다. 하지만, 이 방법은 반복적인 방법(iterative method)을 이용하기 때문에 인식에 많은 시간을 필요로 한다. Basri와 Jacobs[9,10]는 9차원 선형 공간으로 표현 가능한 모든 균등 반사 함수(lambertian reflectance function)를 정의했다. convex 모형을 가진 물체의 이미지들은 저차원 선형 공간으로 근사화 될 수 있으며, 이는 얼굴 인식 속도를 높일 수 있음을 의미한다. Wang [11]은 Ronen과 Jacobs의 이론에 근거하여 3차원 얼굴의 모형을 모델링 하는 방법을 제안하였다. 이 방법은 3차원 물체의 구조와 알베도를 복원하기 위해 다수의 이미지를 필요로 하는 광도 양안법(photometric stereo) 알고리즘과 연관성을 지니고 있다.

위에서 설명한 방법들은 특정 모델을 이용하여 조명에 무관한 얼굴 인식을 위해서 이미지의 합성에 의한 방법에 기초를 두고 있다. 이러한 접근 방법은 많은 연산을 필요로 하고 모델을 벗어난 조명 상황의 경우 인식 성능이 낮아지게 된다.

Tenenbaum과 Freeman[12]은 관측 데이터 개체(observation)를 스타일(style)과 콘텐츠(content) 같은 두 개의 독립된 요소로 분리하는 bilinear 모델을 제안

하였다. 그는 논문에서 기대치-최대화(Expectation-Maximization : EM) 방식이나 특이값 분해(Singular Value Decomposition : SVD)를 이용한 이요인(two-factors) 모델 학습과 모델 정합(fitting)에 대한 전체적인 프레임워크를 제공하였다. 그는 모음 구분, 새로운 문자에 대한 외삽(extrapolation), 입력 이미지를 새로운 조명하의 얼굴로 변환 및 인식에 대한 실험 결과를 보여 주었다.

임의의 조명 환경에서 생성된 이미지에 대하여 신원을 인식할 때, 본 논문에서는 조명 요소와 신원 요소를 얻어낸다. 이 과정에서 신원과 조명 요소에 대한 해를 얻기 위해서 반복적인 역연산(inverse operation)이 필요하다. 하지만, bilinear 모델의 입력 이미지가 많은 노이즈 정보를 포함하고 있을 때에는 올바른 수렴을 보장하기 힘들다. 이러한 문제를 완화하기 위해서 본 논문에서는 bilinear 모델에 ridge regression을 결합한 ridge regressive bilinear 모델을 제안한다. ridge regression [13-15]은 모델의 분산을 줄이기 위해서 패널티를 부과한다. 제안된 ridge regressive bilinear 모델은 조명 요소와 신원 요소의 분산을 적절히 줄여줌으로서 bilinear 모델에 안정성을 제공하며, 인식에 더 많은 고차원 요소 정보를 이용하게 함으로써 인식 성능을 높여 준다.

본 논문의 구성은 다음과 같다. 2장에서는 기존 bilinear 모델과 bilinear 모델의 학습 방법에 대한 설명을 한다. 3장에서는 기존 regression 모델에 대한 설명을 한다. 2장과 3장에서는 설명된 모델들을 토대로 하여 4장에서 두 모델을 결합하여, 이 논문에서 제안하고자 하는 ridge regressive bilinear 모델에 대한 설명을 한다. 5장에서는 실험과 결과에 대한 분석을 포함하고 있으며, 마지막으로 결론이 뒤 따른다.

2. Bilinear models

Bilinear 모델[12]은 관측 데이터 개체를 스타일과 콘텐츠 두 개의 요소로 분리하는 이요인 모델이다. Bilinear 모델은 하나의 요소가 고정되었을 때 선형 모델과 동일하게 동작한다. Bilinear 모델은 대칭형(symmetric) 모델과 비대칭형(asymmetric) 모델로 분류된다. 예를 들어, 문자를 볼 때, 문자를 폰트와 의미로 분리하여 인식한다. 이 때 폰트와 의미는 문자를 나타내는 독립적인 요소이다. 이 논문에서 신원 요소를 콘텐츠 요소로, 조명 요소를 스타일 요소로 정의한다.

대칭형 모델에서 bilinear 모델은 상호 작용 행렬(interaction matrix)을 이용하여 스타일과 콘텐츠 요소와 상호 작용을 한다. 상호 작용 행렬은 스타일과 콘텐츠에 독립적이다. 비대칭형 모델은 두 가지 중 하나의 요소를 알고 있다고 가정 한다. 따라서 모델은 선형 모

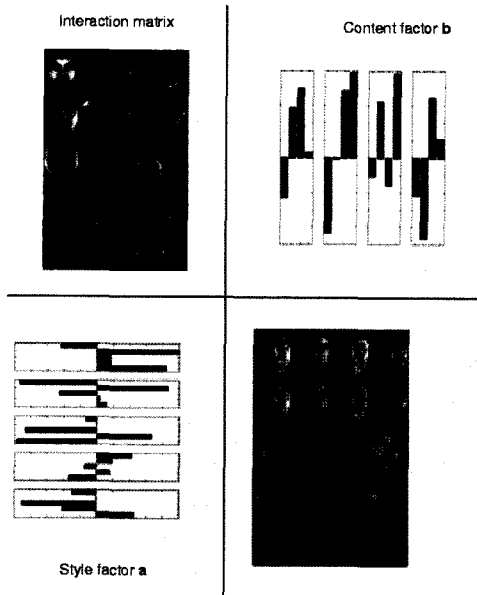


그림 1 bilinear 모델

델로서 동작하게 된다. 그림 1은 bilinear 모델에서 관측 데이터 개체가 두 요소의 상호 작용으로 표현되는 모습을 도해한 것이다.

2.1 Bilinear model 학습

이 절에서는 대칭 모델과 비대칭 모델의 학습 방법에 대해서 설명한다. 대칭형 모델은 다음과 같이 표현된다.

$$y = \sum_{i=1}^I \sum_{j=1}^J w_{ij} a_i b_j \quad (1)$$

y 는 관측 데이터 개체를 나타내는 벡터이고, w_{ij} 는 스타일 요소인 a 와 콘텐츠 요소인 b 와 상호 작용하는 기저(basis) 벡터이다. 학습에 사용되는 이미지의 크기가 N×M이라 할 때, w_{ij} 와 y 는 K(=N×M)차원의 벡터이며, a_i 와 b_j 는 계수이다.

대칭형 모델을 사용하기 위해서, 학습을 통하여 스타일 요소와 콘텐츠 요소가 상호 작용하기 위하여 필요한 상호 작용 행렬 W 를 얻어야 한다. S×C 개의 학습 이미지를 이용하여 관측 데이터 행렬 Y 를 생성한다. S 는 스타일의 개수로서 본 논문에서는 조명의 개수이며, C 는 콘텐츠의 개수로서 학습에 사용된 사람 수를 나타낸다.

$$Y = \begin{pmatrix} y_{11} & \dots & y_{1C} \\ \vdots & \ddots & \vdots \\ y_{S1} & \dots & y_{SC} \end{pmatrix}, \quad Y^{VT} = \begin{pmatrix} y_{11} & \dots & y_{1S} \\ \vdots & \ddots & \vdots \\ y_{C1} & \dots & y_{CS} \end{pmatrix} \quad (2)$$

어깨 글자 VT 는 벡터 전치(vector transpose)를 나타내며, 각 원소 y_{ij} 는 K 차원의 관측 데이터의 벡터이다. 관측 데이터 행렬 Y 의 크기는 SK×C 이다. 그림 2 는 크기가 K=N×M 픽셀인 이미지 S×C 개를 이용하여

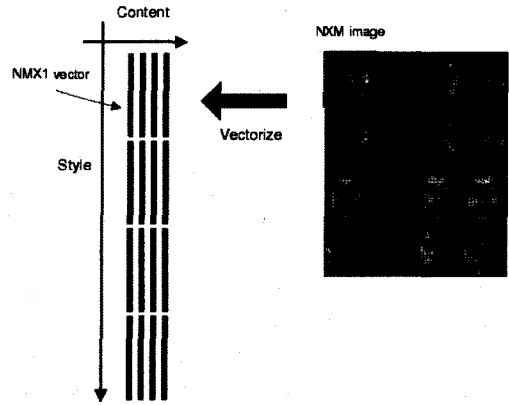


그림 2 관측 데이터 행렬 Y 의 생성 방법

크기가 SK×C 인 관측 데이터 행렬을 생성하는 방법을 도식화 한 것이다.

대칭형 모델은 식 (3)과 같이 간단한 형태로 표현된다.

$$Y = (W^{VT}A)^{VT}B \quad \text{또는} \quad Y^{VT} = (WB)^{VT}A \quad (3)$$

A 와 B 는 스타일과 콘텐츠 요소들이 쌓아져 생성된 I×S, J×C 크기의 행렬이다.

$$A = (a_1, \dots, a_S), \quad B = (b_1, \dots, b_C)$$

$$W = \begin{pmatrix} w_{11} & \dots & w_{1J} \\ \vdots & \ddots & \vdots \\ w_{I1} & \dots & w_{IJ} \end{pmatrix}$$

모델 파라미터 A, B 와 상호 작용 행렬 W 를 얻는 전 과정은 알고리즘 I 과 같다.

알고리즘 I : compute_model_parameter(input:Y; output:A, B, W)

1. Y=USV^T를 이용하여, B 를 V^T 로 초기화.
2. (YB^T)^{VT} = USV^T를 이용하여, V^T 의 처음 I 개의 행으로 B 갱신.
3. (Y^{VT}A^T)^{VT} = USV^T를 이용하여, V^T 의 처음 J 개의 행으로 A 갱신.
4. A 와 B 가 수렴할 때까지 2번과 3번 과정을 반복.
5. W = ((YB^T)^{VT}A^T)^{VT}.

비대칭형 모델은 다음과 같이 표현된다.

$$y = \sum_{j=1}^J \left(\sum_{i=1}^S w_{ij} a_i \right) b_j = W^s b \quad (4)$$

W^s 는 특정 스타일에 대한 기저이고, b 는 콘텐츠 파라미터 요소를 나타낸다. 비대칭형 모델은 식 (5)로 표현 가능하다.

$$Y = WB \quad (5)$$

W와 B의 크기는 각각 $SK \times J$ 이고 $J \times C$ 이다. 비대칭형 모델에서 B와 W는 Y를 SVD로 분해하여 얻어진다. SVD를 이용하여, $Y = USV^T$ 으로 분해되고, W는 US의 처음 J개의 열을 이용하여 생성하고, B는 V^T 의 처음 J개의 행을 이용하여 생성한다.

2.2 Bilinear model translation

Bilinear 모델을 이용하여, 임의의 조명 환경에서 얻어진 사람의 얼굴 이미지를 새로운 특정 조명 환경의 이미지로 변환 가능하다. 본 논문에서는 인식을 위해 대칭형 bilinear 모델을 학습시켜, 상호 작용 행렬 W를 얻고, 이를 이용하여 임의의 조명 환경에서 얻어진 임의의 사람의 테스트 이미지로부터 스타일 요소 a와 콘텐츠 요소 b를 얻는다. 이 과정에서 반복적인 유사 역행렬 연산이 필요하다. †는 유사 역행렬 연산을 나타낸다. 알고리즘 II는 새로운 테스트 이미지 y에서 스타일과 콘텐츠 요소 a와 b를 얻는 과정을 기술한다.

알고리즘 II : compute_style_content_BM(input:W,y; output:a,b)

1. B의 평균값으로 b를 초기화 한다.
2. $a = ((Wb)^{VT})^\dagger y$ 를 이용하여, 스타일 요소 a를 갱신한다.
3. $b = ((W^T a)^{VT})^\dagger y$ 를 이용하여, 콘텐츠 요소 b를 갱신한다.
4. a와 b가 수렴할 때까지, 2번 과정과 3번 과정을 반복한다.

비대칭형 bilinear 모델의 경우, y의 스타일을 알고 있다고 가정하기 때문에 콘텐츠 요소는 한 번에 계산된다.

$$b = (W_i)^\dagger y.$$

3. Regression model

Regression 분석은 독립 변수와 종속 변수 사이의 관계를 분석하는 전통적인 테크닉으로 독립 변수들의 선형 조합으로서 종속 변수를 표현한다.

3.1 최소 자승법

Regression 모델은 식 (6)으로 표현된다.

$$\hat{b} = \arg \min_b \left(\sum_{i=1}^J (y_i - b_0 - \sum_{j=1}^J x_{ij} b_j)^2 \right). \quad (6)$$

Regression의 목적은 잔류 에러(residual error)를 최소화하는 \hat{b} 를 찾는 것이다. 식 (6)은 식 (7)과 같은 간단한 형태로 표현 가능하다.

$$y = Xb + e. \quad (7)$$

y, X, b, 그리고 e는 관측 데이터 벡터, 의존 변수, 모델 파라미터, 그리고 에러 벡터를 의미한다. e의 평균은 0이고, 분산은 σ^2 이다. 종속 변수 행렬 X가 완전 계수(full rank)이면, 최소 자승법(least square estimation)을 이용하여 \hat{b} 을 얻을 수 있다.

$$\hat{b} = (X^T X)^{-1} X^T y. \quad (8)$$

\hat{b} 의 분산은 식 (9)와 같이 표현 할 수 있다.

$$\begin{aligned} \text{Var}(\hat{b}) &= E[(\hat{b} - b)(\hat{b} - b)^T] \\ &= E[(X^T X)^{-1} X^T e e^T X (X^T X)^{-1}] \\ &= (X^T X)^{-1} X^T E[ee^T] X (X^T X)^{-1} \\ &= \sigma^2 (X^T X)^{-1} \\ &= \sigma^2 (VDU^T UDV^T)^{-1} \\ &= \sigma^2 (VD^2 V^T)^{-1} \\ &= \sigma^2 \sum_{j=1}^J v_j \frac{1}{d_j} v_j^T. \end{aligned} \quad (9)$$

$X = UDV^T$ 이고, v_j 는 행렬 V의 j번째 열인 고유벡터(eigen vector)이고, d_j 는 v_j 에 해당하는 고유치(eigen value)이다. 데이터 사이에 강한 상관관계가 존재하면, 공선성(collinearity)이 발생한다. 이는 고유치가 작은 고유 벡터를 만든다. 그리고 이것은 모델의 분산을 크게 만들어, 모델을 불안정하게 만든다. 이러한 문제를 해결하기 위해서 ridge regression, partial least square 그리고 principal component regression 같은 편향 추정법(biased method)이 개발 되었다.

3.2 Ridge regression

Ridge regression은 모델의 분산을 줄이기 위해서 계수(coefficient)에 페널티를 부과한다.

$$\hat{b} = \arg \min_b \left(\sum_{i=1}^J (y_i - b_0 - \sum_{j=1}^J x_{ij} b_j)^2 + \lambda \sum_{j=1}^J b_j^2 \right). \quad (10)$$

식 (10)은 식 (11)와 같은 간단한 매트릭스 형태로 표현될 수 있다.

$$RSS(\lambda) = (y - Xb)^T (y - Xb) + \lambda b^T b. \quad (11)$$

해는 식 (12)를 통하여 얻을 수 있다.

$$\hat{b} = (X^T X + \lambda I)^{-1} X^T y. \quad (12)$$

Ridge regression으로 얻은 \hat{b} 의 분산은 식 (12)와 같이 계산될 수 있다.

$$\begin{aligned} \text{Var}(\hat{b}) &= E[(\hat{b} - E(\hat{b}))(\hat{b} - E(\hat{b}))^T] \\ &= E[(X^T X + \lambda I)^{-1} X^T e e^T X (X^T X + \lambda I)^{-1}] \\ &= (X^T X + \lambda I)^{-1} X^T E[ee^T] X (X^T X + \lambda I)^{-1} \\ &= \sigma^2 (X^T X + \lambda I)^{-1} X^T X (X^T X + \lambda I)^{-1} \\ &= \sigma^2 \sum_{j=1}^J v_j \left(\frac{d_j}{(d_j + \lambda)^2} \right) v_j^T. \end{aligned} \quad (13)$$

Ridge regression에서는 $X^T X$ 의 대각 원소(diagonal element)에 작은 상수 λ 를 더해지게 된다. 이것은 파라미터 벡터 \hat{b} 의 분산을 줄이고, regression 모델을 더욱 안정하게 만들어 준다.

4. Ridge regressive bilinear model

이 장에서는 ridge regression 과 bilinear 모델을 결합한 ridge regressive bilinear 모델을 설명한다. Ridge regression 의 특성을 bilinear 모델에 추가하기 위해서, bilinear 모델의 목적함수(objective function)를 수정하였다.

$$E(a,b) = \sum_{k=1}^K \left(y_k - \sum_{i=1}^I \sum_{j=1}^J w_{ijk} a_i b_j \right)^2 + \lambda \sum_{i=1}^I \sum_{j=1}^J (a_i b_j)^2. \quad (14)$$

식 (14)는 식 (15)와 같이 간단한 행렬 형태로 표현할 수 있다.

$$E(a,b) = (y - (Wb)^{VT}a)^T (y - (Wb)^{VT}a) + \lambda a^T a b^T b. \quad (15)$$

E 의 최소 값을 얻기 위해서 목적함수 E 를 모델 파라미터 벡터 a 와 b 로 미분하면, 스타일과 콘텐츠 요소를 얻을 수 있다.

$$a = \left(((Wb)^{VT})^T ((Wb)^{VT}) + \lambda b^T b \right)^{-1} ((Wb)^{VT})^T y, \\ b = \left((W^{VT}a)^T (W^{VT}a) + \lambda a^T a \right)^{-1} (W^{VT}a)^T y. \quad (16)$$

알고리즘 III 은 ridge regressive bilinear 모델에서 입력 이미지 y 를 이용하여, 스타일과 콘텐츠 요소를 얻는 과정을 기술한다.

알고리즘 III : compute_style_content_RRBM(input:W,y; output:a,b)

1. B 의 평균값으로 b 를 초기화 한다.
2. $a = \left(((Wb)^{VT})^T ((Wb)^{VT}) + \lambda b^T b \right)^{-1} ((Wb)^{VT})^T y$ 를 이용하여, a 를 갱신한다.
3. $b = \left((W^{VT}a)^T (W^{VT}a) + \lambda a^T a \right)^{-1} (W^{VT}a)^T y$ 를 이용하여, b 를 갱신한다.
4. a 와 b 가 수렴할 때까지, 2번 과정과 3번 과정을 반복한다.

비대칭형 bilinear 모델에서는 입력 이미지 y 의 스타일을 사전에 알기 때문에, 반복적인 연산을 사용할 필요 없이 한 번에 결과를 얻을 수 있다. 관측 데이터 벡터의 스타일 기저가 W_i 라면, 콘텐츠 요소는 식 (17)을 통하여 얻을 수 있다.

$$b = (W^T W_i + \lambda I)^{-1} W_i^T y. \quad (17)$$

실제 구현에서는 고유치가 커지는 부작용을 피하기 위해서 λ 보다 큰 고유벡터의 고유치는 그대로 보존하고, λ 보다 작은 고유 벡터의 고유치는 λ 로 변환해 주었다.

5. 실험 및 분석

ridge regressive bilinear model은 대칭과 비대칭 모델에 사용할 수 있지만, 실험에서는 대칭형 ridge regressive bilinear model을 사용하였다. 비대칭 모델의 경우에는 하나의 요소를 알고 있어야 사용 가능하지만, 실제 환경에서 조명의 상태를 아는 경우가 매우 드물고, 또한 다른 방법들과의 성능 비교에 따른 형평성을 고려하여 대칭형 모델을 이용하였다.

실험에는 CMU PIE 데이터베이스[16]를 사용 하였다. 데이터베이스는 21개의 서로 다른 조명 이미지를 갖고 있는 68명으로 구성 되어 있다. 그림 3은 CMU PIE 데이터베이스에 포함된 샘플 이미지들을 나타내었다. 실험을 위하여 얼굴 이미지에서 두 눈의 중점과 거리를 기준으로 얼굴을 수작업으로 이미지에서 잘라 낸 후에 이미지의 크기를 56×46으로 변환 하였다. 이미지에서 머리 모양이나 이미지의 백그라운드 같은 노이즈 정보를 제거하기 위해서 마스크를 사용했다.

실험에서 PIE 데이터베이스의 이미지를 학습용 데이터 셋(set)과 테스트용 데이터 셋으로 나누었다. 테스트 셋은 다시 probe 데이터 셋과 gallery 데이터 셋으로

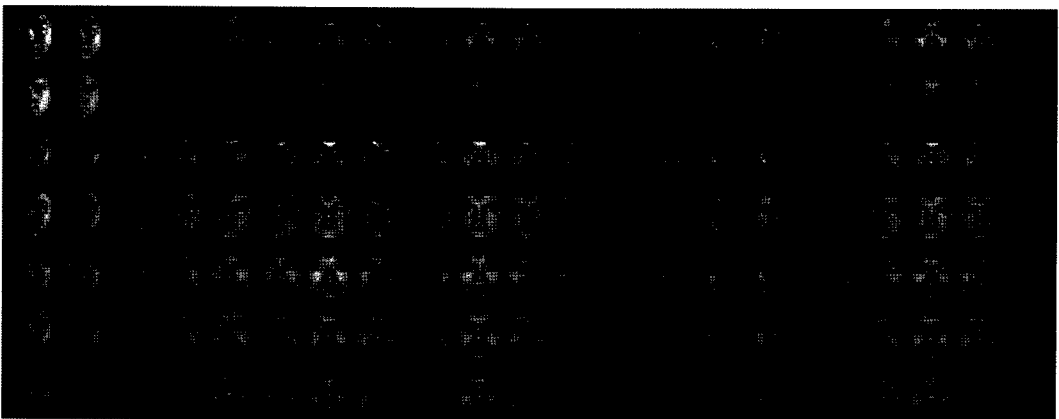


그림 3 PIE 데이터베이스의 샘플 이미지

나누었다. 테스트 데이터 중에서 정면 조명을 가진 이미지를 이용하여 gallery 데이터 셋으로 사용하고, 나머지 이미지들을 이용하여 probe 데이터 셋을 구성하였다.

실험은 학습, 등록, 그리고 테스트 세 개의 단계로 구성되어 있다. 학습 단계에서 Y를 이용하여 행렬 W를 구한다. 등록 단계에서는, 알고리즘 II나 알고리즘 III을 이용하여 gallery 이미지들에 대한 스타일과 콘텐츠 요소를 얻고, 이를 데이터베이스에 등록한다. 테스트 단계에서는 각각의 probe 이미지에 대하여 스타일과 콘텐츠 요소를 얻고, 이를 등록 단계에서 저장한 gallery 이미지들의 콘텐츠 요소와 비교하여 probe 데이터와 가장 유클리드 거리(Euclidean distance)가 가장 작은 gallery 이미지를 probe 이미지와 신원이 동일한 사람으로 인식하게 하였다. 그림 4에 얼굴 인식을 위해 사용된 전체적인 구조가 도식화 하였다. 모든 실험은 트레이닝 셋과 테스트 셋의 조합을 변경하며 10번 실행하였고, 결과 값으로 평균값을 이용하였다.

5.1 안정성

두 bilinear 모델의 분산을 이용하여 안정성을 비교하였다. 그림 5는 λ의 값을 변경하면서 얻은 콘텐츠 요소

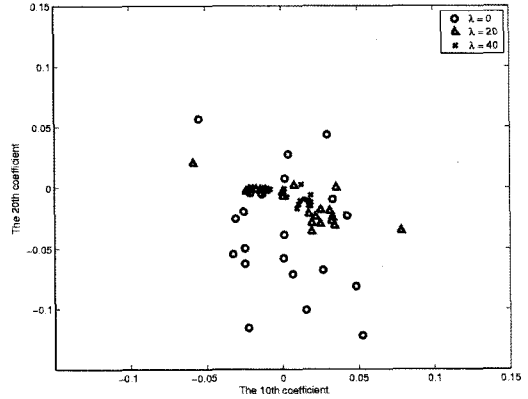


그림 5 λ의 크기에 따른 콘텐츠 요소의 분산

들의 분포를 보여준다. λ가 커지면 ridge regressive bilinear 모델은 고유 값이 작은 성분에 더 큰 영향을 미치며 해당하는 계수의 크기를 축소시킨다. 이를 도식하기 위해서 10번째, 20번째 성분(component) 사용하였다. 그림 5에서 λ가 증가함에 따라 콘텐츠 요소의 크기와 그것의 분산이 줄어들게 됨을 확인 할 수 있다. 이것

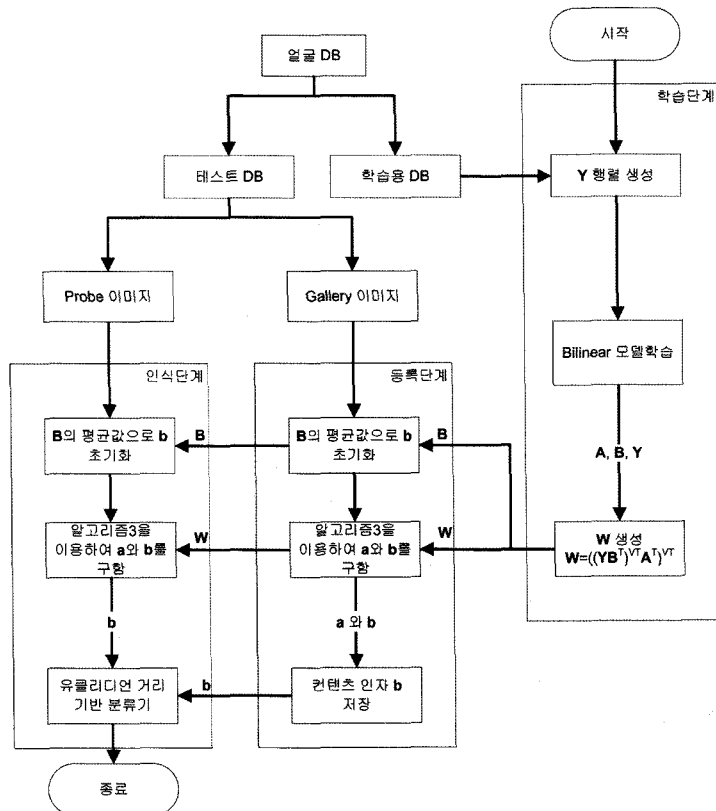


그림 4 전체적 실험 구조

은 bilinear 모델 보다 ridge regressive bilinear 모델이 더 안정적인 이유를 설명해 준다. 그림 5에서 λ 가 0인 경우, ridge regressive bilinear 모델은 bilinear 모델과 동일하다.

5.2 얼굴 인식

첫째, 여러 λ 값을 이용하여 인식 실험을 하였고, 최적의 λ 값을 결정 하였다. 콘텐츠 요소의 성분 개수를 20, 25, 30개로 변경하며 에러율을 계산하였다. 이를 도식화 한 것이 그림 6이다. 주어진 데이터 셋에 대하여 λ 값이 12일 때 최적임을 확인 할 수 있다. 만약 λ 가 일정 수준의 크기를 넘으면 콘텐츠 요소가 너무 많이 줄어들게 되고, 유용한 정보를 잃어버리게 된다. 이것은 λ 값이 일정 수준 이상으로 커지면 인식 성능이 낮아지는 이유를 설명해 준다.

둘째, 학습에 사용한 사람 수를 변경하며 두 bilinear 모델의 인식 성능을 비교해 보았다. 실험에서 N 명의 사람을 무작위로 선택하여 트레이닝 셋으로 사용 하였고, 나머지 68-N 명의 사람을 테스트 셋으로 사용 하였다. (68-N)×1 개의 정면 조명 이미지를 gallery 이미지로 사용하고 나머지 (68-N)×20 이미지를 probe 이미지로 사용하였다. 콘텐츠 요소에 사용된 성분의 개수를 20개로 고정 하였다. 그림 7에 두 bilinear 모델의 평균 에러율을 도식화 하였다. 이때의 λ 값은 12로 설정하였다. 그림 7을 통하여 (1) 두 bilinear 모델의 평균 에러율은 트레이닝에 사용한 사람 수가 많을수록 더 낮아지며, (2) ridge regressive bilinear 모델이 더 낮은 에러율을 보여주는 것과 (3) ridge regressive bilinear 모델의 에러율의 분산이 더 작은 것을 확인할 수 있다.

셋째, 성분 개수를 변경하며 두 bilinear 모델의 성능을 비교하였다. 그림 8은 λ 값이 12일 때, 성분 개수를 12개에서 34개로 변경하며 얻은 평균 에러율을 표시하였다. 그림 8에서 ridge regressive bilinear 모델의 평

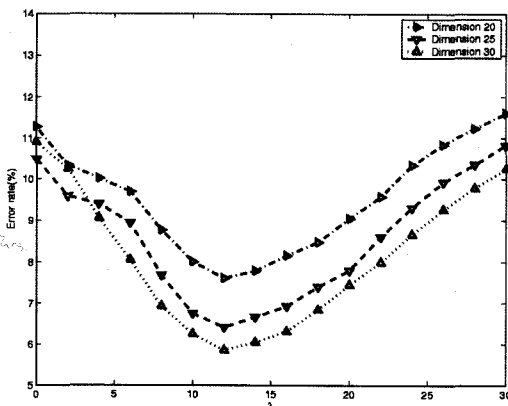


그림 6 λ 값의 변환에 따른 에러율 곡선

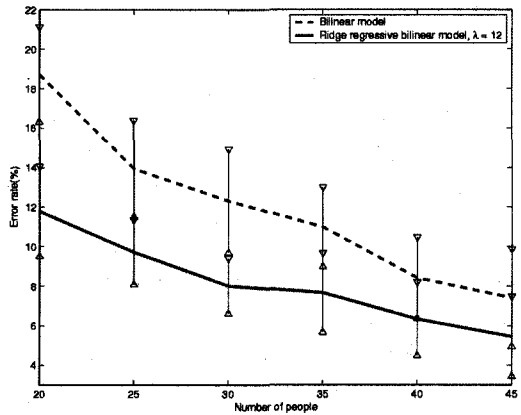


그림 7 트레이닝 사람 수에 따른 두 bilinear 모델의 에러율 곡선

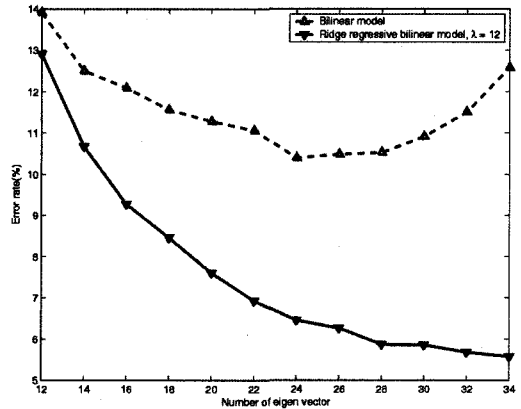


그림 8 콘텐츠 요소의 성분개수에 따른 두 bilinear 모델의 에러율 곡선

균 에러율은 컴포넌트의 개수가 증가함에 따라 지속적으로 낮아짐을 확인할 수 있다. 반면, bilinear 모델의 에러율은 점차 떨어지다가 컴포넌트의 개수가 일정 수준을 넘게 되면 에러율이 높아진다. 이것은 콘텐츠 요소의 성분 개수가 적절히 선택되지 못하면 bilinear 모델은 불안정해 질 수 있음을 의미한다.

마지막으로, bilinear 모델의 성능을 고유 얼굴, Quotient image[1]과 비교해 보았다. 각 실험에서 Quotient image를 제외한 모든 방법들은 동일한 테스트 셋과 트레이닝 셋을 사용하였다. 그림 9는 네 가지 방법들에 대한 성능을 도식화 한 그래프이다. 바의 중간은 평균 인식 성능 나타낸다. 그리고 각 바의 범위는 인식 성능의 최대값과 최소값을 나타낸다. 그림 8에서 (1) Quotient image는 낮은 인식 성능을 나타낸다. 왜냐하면 Quotient image는 그림자가 없는(non-shadow) 모델을 사용했기 때문이다. Quotient image를 생성할 때 입력 이

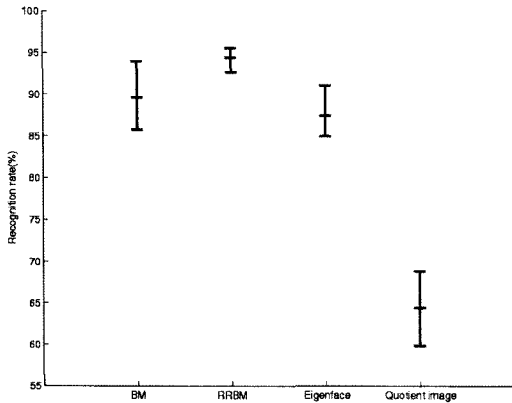


그림 9 네 가지 알고리즘의 인식 성능

미지를 합성된 이미지로 나누는데, 입력 이미지에서 그림자로 인해서 가려진 영역이 있는 경우에 이 영역의 값은 0이나 0에 가까운 값이 나오게 된다. 이런 영역이 인식 성능에 영향을 미친다. 따라서 그림자의 영향을 많이 받은 이미지일수록 인식 성능이 떨어지게 된다. 실제로 CMU DB에는 그림자의 영향을 받은 영상이 다수 포함되어 있으며, 이런 조명을 갖는 이미지들에 대해서 낮은 인식 성능이 나왔다. (2) Ridge regressive bilinear 모델은 가장 작은 인식 성능의 분산과 높은 성능을 보여 준다. 이것은 제한된 ridge regressive bilinear 모델이 강인하고 안정적인 인식 성능을 나타냄을 보여 준다. 표 1에는 네 가지 실험 방법에 대한 실험 파라미터와 인식 성능을 나타냈다.

표 1 인식 성능 비교

알고리즘	성분의 개수	인식성능(%)	분산
RRBM($\lambda=12$)	34	94.4	0.8
Bilinear model	24	89.6	5.45
Eigenface	100	87.4	3.67
Quotient image	56×46	64.4	9.14

6. 결론

이 논문은 ridge regressive bilinear 모델을 이용한 조명에 강인한 얼굴 인식 방법을 제안한다. bilinear 모델은 조명 조건의 변화와 관계없이 좋은 콘텐츠 요소를 얻어낸다. 하지만, bilinear 모델은 공선성에 의한 작은 고유치를 가진 고유 벡터에 의해 올바른 수렴에 실패할 수 있다. 이러한 문제를 해결하기 위해서, 본 논문에서는 ridge regression과 bilinear 모델을 결합한 ridge regressive bilinear 모델을 제안하였다. 또한 몇 가지 실험으로 ridge regressive bilinear 모델이 다른 방법보다 성능과 안정성 면에서 우수함을 보였다. 그리고 사

효과에 따라 콘텐츠 요소의 범위가 제한되게 되고 이것이 성능 향상에 영향을 미치게 되었음을 알 수 있었다.

참고 문헌

- [1] Tammy Riklin-Raviv and Amnon Shashua, "The quotient image: class-based rerendering and recognition with varying illuminations," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, issue 2, pp. 129-139, 2001.
- [2] Shaohua Zhou and Rama Chellappa, "Rank constrained recognition under unknown illuminations," *Proceedings of Analysis and Modeling of Faces and Gestures*, pp. 11-18, 2003.
- [3] Jiali Zhao, Yong Su, Dejun Wang and Siwei Luo, "Illumination ratio image: synthesizing and recognition with varying illuminations," *Pattern Recognition Letters*, vol. 23, issue 15, pp. 2703-2710, 2003.
- [4] Dang-Hui Lui, Kin-Man Lam and Lan-Sun Shen, "Illumination invariant face recognition," *Pattern Recognition*, vol. 38, pp. 1705-1716, 2005.
- [5] Shiguang Shan, Wen Gao, Bo Cao and Debin Zhao, "Illumination normalization for robust face recognition against varying lighting conditions," *Proceedings of Analysis and Modeling of Faces and Gestures*, pp. 157, 2003.
- [6] Peter N. Belhumeur and David J. Kriegman, "What Is the Set of Images of an object under all possible illumination conditions?," *International Journal of Computer Vision*, vol. 28, num. 3, pp. 245-260, 1998.
- [7] Athinodoros S. Georghiades, Kriegman. D. J and Peter N. Belhumeur, "Illumination cones for recognition under variable lighting: Faces," *Computer Vision and Pattern Recognition*, p.52, 1998.
- [8] Athinodoros S. Georghiades, Kriegman. D. J and Peter N. Belhumeur, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, issue 6, pp. 643-660, 2001.
- [9] Ronen Basri and David Jacobs, "Lambertian reflectance and linear subspace," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, issue 2, pp. 218-233, 2003.
- [10] Ronen Basri and David Jacobs, "Photometric stereo with general, unknown lighting," *Computer Vision and Pattern Recognition*, vol. 2, pp. 374, 2001.
- [11] Haitao Wang, Stan Li, Yangsheng Wang and Weiwei Zhang, "Illumination modeling and normalization for face recognition", *Proceedings of Analysis and Modeling of Faces and Gestures*, pp. 104-111, 2003.
- [12] Joshua B. Tenenbaum, "Separating style and

- content with bilinear models," *Neural Computation*, vol. 12, pp. 1247-1283, 2000.
- [13] Tormod Næs and Bjørn-Helge Mevik, "Understanding the collinearity problem in regression and discriminant analysis," *Journal of Chemometrics*, vol. 15, issue 4, pp. 413-426, 2001.
- [14] Trevor Hastie, Robert Tibshirani and Jerome Friedman, "The Elements of Statistical Learning Data Mining, Inference and Prediction," Springer, pp. 59-64, 2001.
- [15] Maring Høy, Frank Westad and Harald Martens, "Combining bilinear modelling and ridge regression", *Journal of Chemometrics*, vol. 16, issue 6, pp. 313-318, 2002.
- [16] T. Sim, S. Baker and M. Bsat, "The CMU pose, illumination, and expression (PIE) database," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, issue 12, pp. 1615-1618, 2003.



신 동 수

2006년~현재 LG Electronics. 2004년~2006년 M.S., Computer Science & Engineering, POSTECH. 1999년~2004년 B.S., Computer Engineering, Dongguk University



김 대 진

1981년 2월 연세대학교 전자공학과(학사). 1984년 2월 KAIST 전기 및 전자공학과(석사). 1991년 8월 Syracuse University, Electrical and Computer Eng. (박사). 1984년 3월~1986년 12월 한국방송공사 기술연구소. 1992년 3월~1999년 6월 동아대학교 컴퓨터공학과 부교수. 1999년 7월~현재 포항공과대학교 컴퓨터공학과 부교수. 주 관심분야는 지능시스템, 멀티미디어 처리



방 승 양

1966년 일본 Kyoto대학 전기공학에서 학사. 1969년 서울대학교 전기공학에서 석사. 1974년 미국 University of Texas 전산학에서 박사를 받았음. 미국 Wayne State University, NCR, Bell 연구소 등에서 근무하다가 1981년 귀국. 한국전자기술연구소 시스템부 실장, 부장 역임, (주)유니온시스템 전무. 1986년부터 포항공대 컴퓨터공학과 교수. 현재 뇌연구센터 소장. 관심분야는 패턴인식, 신경회로망