

얼굴 인식을 위한 연립 대각화와 국부 선형 임베딩

김은솔¹, 노영균², 장병탁¹

¹서울대학교 컴퓨터공학부

²한국과학기술원 전산학과

eskim@bi.snu.ac.kr, yungkyun.noh@gmail.com, btzhang@bi.snu.ac.kr

Locally Linear Embedding with Simultaneous Diagonalization for Face Recognition

Eun-Sol Kim⁰¹, Yung-Kyun Noh², Byoung-Tak Zhang¹

¹School of Computer Science and Engineering, Seoul National University

²Department of Computer Science, Korea Advanced Institute of Science and Technology

요 약

Locally Linear Embedding(LLE)는 manifold learning 알고리즘 중 하나로 고차원 공간에 있는 데이터들의 inner product 정보를 이용하여 embedding 하는 방법이다. 고차원 공간에서 같은 평면에 있는 데이터들은 inner product 값이 크기 때문에 low-dimensional 공간에서도 가깝게 위치하도록 embedding 되는 반면 수직으로 위치한 평면에 있는 데이터들은 inner product 값이 0이 되기 때문에 서로 떨어진 위치에 embedding 된다. 한편, 한 사람의 얼굴에 다양한 각도에서 조명을 비추면서 촬영한 이미지들은 low-dimensional linear subspace를 구성한다는 연구 결과가 있다. 이에 본 논문에서는 다른 평면에 위치하는 데이터들을 자연스럽게 분류할 수 있는 LLE 알고리즘을 얼굴 이미지에 사용하여, 효과적으로 얼굴 인식 문제를 해결할 수 있는 방법을 제안한다. LLE가 다른 평면에 위치하는 데이터들을 구분하는 성질을 극대화 하기 위하여, identity가 다른 데이터들이 서로 수직으로 위치하도록 하는 연립 대각화(Simultaneous Diagonalization, SD) 방법을 적용하였다. 실험 결과, 연립 대각화를 적용하고 LLE를 적용하면 identity가 다른 데이터들이 겹치지 않고 뚜렷하게 구분되는 효과가 있음을 확인하였다.

1. 서 론

Non-linear embedding 방법 중 하나인 Locally Linear Embedding (LLE)[1]는 high-dimensional 공간에 있는 데이터들 간의 inner product 정보를 사용한다. Inner product는 두 벡터가 같은 방향으로 있을 때에 최대 값을 가지고 수직으로 있을 때에 0이 되는 성질이 있으므로, LLE가 embedding을 위해 사용하는 정보는 high-dimensional 공간에 있는 데이터들이 얼마나 유사한 방향으로 위치하고 있는지, 같은 평면에 있는지에 대한 정보임을 알 수 있다. 반면 수직으로 위치하는 데이터들의 정보는 이용하지 않는다. 우리는 LLE가 수직으로 위치하는 데이터들 사이의 관계를 고려하지 않는다는 점에 주목하여 얼굴 데이터를 분리할 수 있는 방법을 제안하고자 한다.

얼굴 이미지에는 identity 이외에 illumination, pose, expression과 같은 외부 요인이 포함되어 있을 수 있는데, 이 중 illumination의 변화는 간단한 선형모델로 표현될 수 있다. Object (얼굴 이미지에서는 얼굴에 해당함)가 Lambertian 일 때에, object에 다양한 각도에서 빛을 비추면서 촬영한 이미지들이 low-

dimensional linear subspace를 구성한다는 특징이 잘 알려져 있다. [2,3] 본 논문에서는, 얼굴 이미지에 내재되어 있는 illumination 변화를 linear subspace로 표현할 수 있다는 사실을 기반으로 얼굴 인식의 성능을 크게 높일 수 있는 방법을 제시한다.

앞서 논의 하였듯이 LLE는 같은 평면에 있는 데이터들 사이의 관계를 고려하는 반면 수직으로 위치하는 데이터들 사이의 관계는 고려하지 않는 방법이다. 데이터들이 같은 평면에 있다면 low-dimensional 공간에서 가깝게 위치하도록 embedding 하고, 수직으로 위치하는 데이터들은 서로 떨어지도록 embedding 한다. 반면 얼굴 이미지는 illumination의 변화에 따라 linear subspace를 형성하므로, identity마다 하나의 평면을 이루고 있다. 본 논문에서는 두 가지 성질을 결합하여, 얼굴 이미지를 효과적으로 분류할 수 있는 LLE 방법을 고안하였다.

얼굴 이미지들은 identity에 따라 평면을 이루는 성질이 있기 때문에 LLE를 이용하여 자연스럽게 identity를 분류할 수 있지만, 서로 다른 identity에 해당하는 데이터들이 가깝게 embedding되어 분류 성능이 떨어지는 문제가 있다. 이는 identity 들이

만드는 평면이 항상 수직으로 위치하지 않기 때문이다. 이에 본 논문에서는 identity들이 만드는 평면들이 항상 수직으로 위치하도록 변형한 뒤 LLE를 적용하는 방법을 제안한다. Identity들이 만드는 평면이 수직으로 위치하도록 하기 위하여 연립 대각화 (Simultaneous Diagonalization, SD) 방법을 적용하였다.

실험결과, Synthetic 데이터와 실제 데이터 셋에 실험해 본 결과 SD를 적용한 후 LLE를 수행하면 identity 별로 데이터가 잘 구분되어 embedding 되는 것을 확인하였다.

2. LLE (Locally Linear Embedding)

LLE는 고차원 공간에서 인접해 있는 데이터들 사이의 선형적 구조를 보존하면서 embedding한다. 차원이 D 인 고차원 공간에 데이터가 N 개 있고, 데이터 x_i 와 인접한 $k(k \ll N)$ 개의 데이터들을 $x_j(1 \leq j \leq k)$ 라고 하자. LLE는 x_j 들의 선형 결합으로 x_i 를 재생성 할 때, 그 차이를 가장 줄일 수 있는 선형 결합 계수(가중치)들을 찾고, 이 선형 결합을 차원이 $d(d \ll D)$ 인 낮은 차원의 공간에서도 유지할 수 있는 점 Y_i 를 찾는다. 이 때 Y_i 의 좌표가 회전이나 이동과 같은 연산에 영향을 받지 않게 하기 위하여 가중치의 합이 1이 되도록 한다.

W_{ij} 를 구하는 방법은 다음과 같다.

x_i, x_j 는 각각 크기가 D 인 column 벡터, X_j 를 k 개의 인접한 데이터들로 이루어진 $D \times k$ 행렬, W_{ij} 를 $D \times N$ 행렬이라고 하자. 이때 고차원 공간에서의 재생성 에러는 다음과 같이 나타낼 수 있다.

$$E = \sum_i \left\| x_i - \sum_{j=1}^k W_{ij}^T \cdot x_j \right\|^2 = \sum_i \|x_i - W_{ij}^T \cdot X_j\|^2$$

이때 E 를 최소화하는 W_{ij} 는 다음과 같이 전개하여 구할 수 있다.

$$\frac{\partial E}{\partial W} = 2 \sum_i (x_i - W_{ij}^T \cdot X_j) \cdot X_j^T = 0, \quad s.t \sum_j W_{ij} = 1$$

$$x_i X_j = W_{ij}^T X_j X_j^T \\ \therefore W_{ij} = (X_j^T X_j)^{-1} X_j^T x_i$$

합이 1인 조건이 없다면 계산한 W_{ij} 는 x_i 를 X_j 들이 span 하는 평면에 정사영시키는 projection matrix임을 알 수 있다. 이와 비교하여 합이 1인 조건이 추가된다면 x_i 를 X_j 들이 이루는 convex hull과 가장 가까운 곳으로 mapping을 하게 된다. (그림 1)

또, W_{ij} 행렬을 보면 x_i 와 인접한 데이터들 중에서, x_i 와 수직으로 위치하는 데이터들에 대한 가중치가 0이 되는 것을 알 수 있다. 즉, embedding을 할 때 x_i 와 인접해 있지만 수직으로 위치한 데이터들과 x_i 사이의 상관 관계는 고려되지 않는다. 본 논문에서는 LLE의 이러한 특성을 이용하여 얼굴 이미지 분류 문제를 개선하고자 한다.

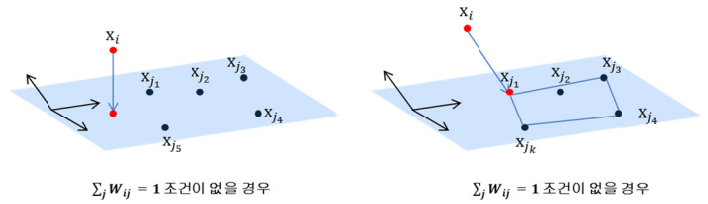


그림 1. 가중치의 합이 1인 조건 유무에 따른 차이

3. 연립 대각화 (Simultaneous Diagonalization)

2장에서 살펴보았듯이, LLE는 high-dimensional 공간에서 수직으로 위치하지 않는 데이터들의 관계를 유지하여 low-dimensional 공간에서도 가깝게 위치하도록 embedding을 하는 반면, 인접한 데이터들 중에서 수직으로 위치하는 데이터들은 떨어뜨려 위치하도록 한다. 즉, 서로 다른 클래스에 속한 데이터들이 high-dimensional 공간에서 수직으로 위치하고 있지 않다면 low-dimensional 공간에서 겹쳐서 위치하게 되는 문제가 생기는데, 이를 반대로 생각하면 클래스가 다른 데이터들이 high-dimensional 공간에서 수직으로 위치하도록 바꾼 후에 LLE를 적용하면 embedding의 성능을 향상시킬 수 있다.

클래스가 다른 데이터들이 서로 수직으로 위치하도록 하기 위하여 연립 대각화 (Simultaneous Diagonalization, SD) 방법을 고안하였다. (그림 2) 연립 대각화는 다음과 같은 방법으로 이루어진다.

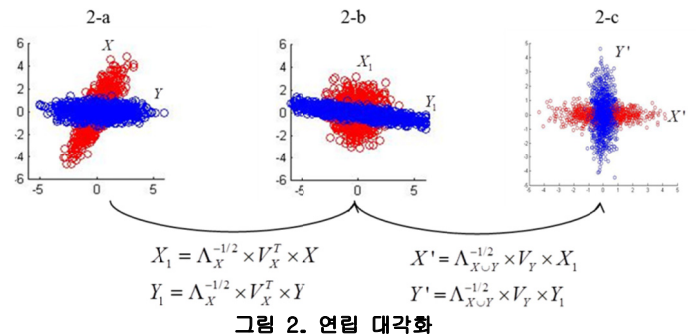


그림 2. 연립 대각화

데이터가 두 개의 클래스 X, Y 로 이루어 있다고 하자 (그림 2-a). 먼저, 클래스 X 만을 whitening 시키는 transformation을 클래스 X, Y 에 동시에 적용하면 변형된 X_1 의 covariance matrix가 identity 행렬이 된다(그림 2-b). 그 다음, X_1, Y_1 전체를 whitening 시켜서 모든 방향의 variance가 같아지도록 하면 그림 2-c와 같이 두 클래스가 수직으로 위치하게 된다. 연립 대각화 방법에 대한 자세한 내용은 표 1에 정리되어 있다.

4. 실험 및 논의

실험을 위하여 2개의 클래스로 구성된 synthetic 데이터와 Yale Face Database B [4, 5]를 사용하였다.

표 1. 연립 대각화 과정

$X = \{X_i\}_{i=1..N}, Y = \{Y_i\}_{i=1..N}$ Σ_x : covariance matrix of X Λ_x : diagonal matrix, diagonal elements are the corresponding eigenvalues of Σ_x V_x : square matrix, i^{th} column is the eigenvector of Σ_x $X_1 = \Lambda_x^{-1/2} \times V_x \times X$ $Y_1 = \Lambda_x^{-1/2} \times V_x \times Y$ $X' = \Lambda_y^{-1/2} \times V_y \times X_1$ $Y' = \Lambda_y^{-1/2} \times V_y \times Y_1$

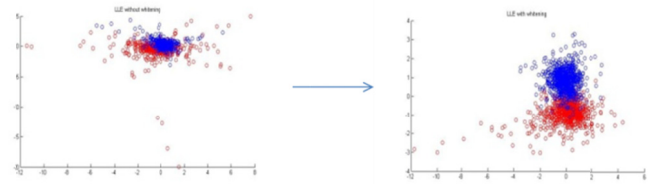
Synthetic 데이터 셋은 두 가지 방법으로 생성되었다. 첫 번째 synthetic 데이터는 20-dimensional data이고, 그 중 10차원은 intrinsic dimension으로서 가우시안 분포를 따르고, 나머지 10차원은 noise dimension이다. 두 번째 데이터는 50-dimensional data이고 10차원이 intrinsic dimension으로, 가우시안 분포를 따른다. 이 두 가지 데이터에 SD를 적용시킨 후에 LLE를 수행한 결과가 그림 3에 정리되어있다. 그림 3의 왼쪽 그래프들은 SD를 적용하지 않고 LLE를 수행한 결과인데, 같은 클래스에 속하는 데이터들이 인접하여 embedding이 되긴 하지만, 두 개의 데이터가 겹쳐서 위치하고 있다. 반면 SD를 적용한 후 LLE를 적용한 결과를 보면 서로 다른 클래스에 속하는 데이터들이 따로 떨어져서 위치하고 있는 것을 확인할 수 있다.

두 번째로 Yale Face Database B 중에서 배경이 제거되고 얼굴 부분만 남긴 cropped image를 이용하였다. 이미지의 크기는 가로 192 pixel, 세로 168 pixel로 데이터의 dimension은 약 30,000이며, 사람마다 illumination이 다른 64장의 이미지가 있다. 이 중에서 두 사람의 이미지 셋을 뽑아 실험한 결과를 그림 4에 정리하였다.

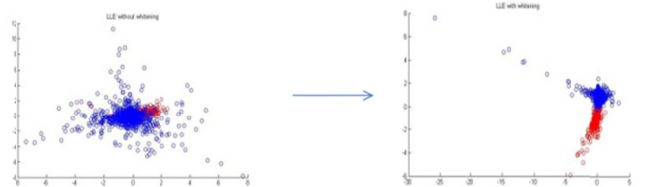
실험 결과를 보면, SD를 사용하여 두 개의 클래스가 low-dimensional 공간에서 서로 떨어져서 위치하도록 embedding 된 것을 볼 수 있다.

그림 3,4를 통해 SD를 이용하면 LLE의 성질을 극대화 하여, high-dimensional 공간에서 같은 평면에 있지 않는 데이터들을 분리할 수 있음을 확인하였다. 이 알고리즘을 얼굴 이미지에 적용하면, 같은 사람에 속하는 이미지들은 illumination의 변화가 만드는 linear subspace 위에 있기 때문에 identity를 기준으로 데이터를 분류할 수 있음을 확인하였다.

한편, 본 논문에서는 초기 실험 결과로 두 사람의 이미지만을 이용한 결과를 보였지만 추후에 여러 사람으로 이루어진 데이터 셋을 효율적으로 분류하기 위한 실험을 진행할 계획이다.



20-dimensional data를 이용한 실험 결과. (좌) LLE 결과 (우) SD를 이용한 LLE 결과



50-dimensional data를 이용한 실험 결과. (좌) LLE 결과 (우) SD를 이용한 LLE 결과

그림 3. 실험 결과

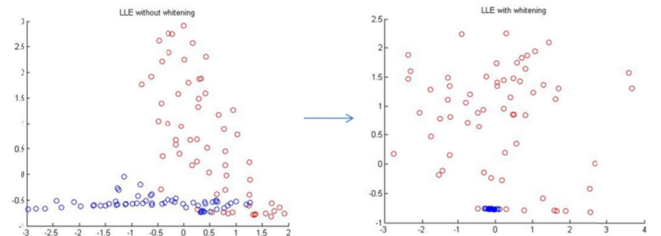


그림 4. Yale Face Database를 이용한 실험 결과. 2명의 이미지 (64개씩 128개)에 LLE를 적용한 결과(좌)와 SD 과정을 거친 후 LLE를 적용시킨 결과 (우)

감사의 글

이 논문은 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구이며(NRF-2010-0017734-Videome), 정부(산업통상자원부)의 재원으로 한국산업기술평가관리원 지원(KEIT-10035348-mLife, KEIT-10044009)을 일부 받았음.

참고문헌

[1] S. Roweis, L. Saul, Nonlinear dimensionality reduction by locally linear embedding. *Science* 290.5500 (2000): 2323-2326.
 [2] M. Turk, P. Alex. Face recognition using eigenfaces. *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on.* IEEE, 1991.
 [3] W. Zhao, R. Chellappa, P. J. Phillips, A. Rosenfeld. Face recognition: A literature survey. *Acm Computing Surveys (CSUR)* 35.4 (2003): 399-458.
 [4] A.S. Georghiades, P.N. Belhumeur, D. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 23.6 (2001): 643-660.
 [5] K. C. Lee, J. Ho, D. Kriegman. Acquiring linear subspaces for face recognition under variable lighting. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 27.5 (2005): 684-698.