

논문 2005-42SP-6-7

DCT/LDA 기반 얼굴 인식에 관한 연구

(A Study on Face Recognition using DCT/LDA)

김 형 준*, 정 병 희**, 김 회 율***

(Hyoung-Joon Kim, Byunghee Jung, and Whoi-Yul Kim)

요 약

본 논문에서는 입력된 얼굴 영상으로부터 구한 DCT 계수에 대해 LDA를 적용하는 DCT/LDA를 이용한 얼굴 인식 방법을 제안한다. 제안된 방법은 적은 수의 DCT 계수를 이용하여 입력 영상을 저차원으로 표현함으로써 특징 공간의 차수보다 트레이닝 데이터의 수가 적은 경우 발생하는 LDA의 SSS 문제를 해결한다. DCT는 기저 벡터가 일정하며 PCA와 유사한 에너지 압축 효율을 가지기 때문에 제안된 방법은 기존의 PCA/LDA 방법보다 학습 속도는 빠르면서 실제 얼굴인식 시스템에 적용이 가능한 정도의 얼굴 인식율을 기대할 수 있다. 실험을 통해 제안된 방법이 PCA/LDA 방법과 유사한 얼굴 인식 성능을 보이면서 약 13,000배 빠르게 학습되는 것을 확인하였고, 기존의 Block-DCT/LDA 방법과 유사하거나 향상된 인식 결과를 확인하였다.

Abstract

This paper proposes a method to recognize a face using DCT/LDA where LDA is applied to DCT coefficients of an input face image. In the proposed method, SSS problem of LDA due to less number of training data than the size of feature space can be avoided by expressing an input image in low dimensional space using DCT coefficients. In terms of the recognition rate, both the proposed method and the PCA/LDA method have shown almost equal performance while the training time of the proposed method is much shorter than the other. This is because DCT has the fixed number of basis vectors while the property of energy compaction rate is similar to that of PCA. Although depending on the number of coefficients employed for the recognition, the experimental results show that the performance of the proposed method in terms of recognition rate is very comparable to PCA/LDA method and other DCT/LDA methods, and it can be trained 13,000 times faster than PCA/LDA method.

Keywords: 얼굴 인식, face recognition, discrete cosine transform(DCT), principal component analysis(PCA), linear discriminant analysis(LDA)

I. 서 론

최근의 얼굴 인식에 관한 연구는 얼굴 영역 영상을 그대로 사용하는 외관 기반 얼굴 인식 (appearance-based face recognition) 방법을 많이 사용하고 있다 [1][2][3]. 즉, 카메라로 획득된 영상에서 얼굴 영역에 대한 밝기 값은 $h \times w$ 크기의 행렬로 표현이 가능하고, 이를 $h \times w$ 길이를 갖는 벡터로 변환하여 특징값으로 사용한다. 그러나 $h \times w$ 길이를 갖는 특징 벡터를 직접 사용하

기에는 그 크기가 매우 크기 때문에 일반적으로 PCA(Principal Component Analysis)나 LDA(Linear Discriminant Analysis)와 같은 방법을 이용해서 특징 공간 축약 (feature space reduction) 과정을 거치게 된다.

PCA는 트레이닝 데이터로부터 데이터를 잘 표현할 수 있는 가장 중요한 축, 즉 기저 벡터(basis vector)들을 찾아 효율적으로 데이터의 차원을 줄일 수 있는 장점이 있다. 그러나 입력 영상의 크기가 커짐에 따라 기저 벡터를 구하는 계산량이 매우 높은 단점이 있다. LDA는 클래스 내부의 분산은 줄이고 클래스 간의 분산은 증가시키는 방향으로 입력 특징값을 변환시키는 변환 행렬을 구하는 방법이며, 트레이닝 데이터의 수가 증가할수록 PCA 방법 보다 인식률이 좋은 것으로 알려

* 학생회원, *** 정회원, 한양대학교 전자통신컴퓨터공학부 (Hanyang University)

** 정회원, KBS 방송기술연구팀 (Korean Broadcasting System)

접수일자: 2005년4월7일, 수정완료일: 2005년9월22일

져 있다^[3]. 그러나 트레이닝 데이터의 수 N 이 특징 공간의 차원 t 보다 적은 경우 SSS 문제(Small Sample Size problem)가 발생하여 변환 행렬을 구할 수 없게 되며, 얼굴 인식과 같은 실제 경우에 있어서 N 보다 t 가 매우 높은 경우가 일반적이다. 예를 들어, 46×56 크기의 얼굴 영상을 사용하여 인식하는 경우 $t=2,576$ 이 되고 LDA를 적용하기 위해서는 그 이상의 트레이닝 데이터가 필요하지만 이를 충족시키기는 매우 어렵다. 이를 해결하기 위해 일반적으로 PCA를 통해 입력된 영상을 g 차원의 중간 특징 공간(intermediate feature space)으로 변환한 후 LDA를 적용하여 최종 f 차원의 특징 공간으로 표현하는 방법을 사용한다 (단, $f < g < t$)^{[3][4][5]}.

본 논문에서는 입력된 얼굴 영상으로부터 구한 DCT(Discrete Cosine Transform) 계수에 대해 LDA를 적용하는 DCT/LDA를 이용한 얼굴 인식 방법을 제안한다. DCT는 기저 벡터가 고정되어 있으면서 특히 자연 영상에 대해 PCA와 유사한 에너지 압축 효율을 갖기 때문에 별도의 학습 과정 없이 적은 수의 DCT 계수로 영상을 잘 표현할 수 있다. 따라서 PCA/LDA 방법의 학습 과정에서 PCA가 대부분의 계산량을 차지하는 것을 고려할 때, DCT/LDA 방법은 PCA/LDA 방법 보다 빠른 학습 속도와 유사한 인식 성능을 기대할 수 있다. 이 방법은 이미 우리가 제안한 얼굴 구성 요소 기반 DCT/LDA 얼굴 인식 방법에서 사용되었으나^{[6][7]}, 얼굴 구성 요소를 사용하지 않고 전체 얼굴에 대해서 DCT/LDA를 적용한 것에 대한 비교 분석이 없었기 때문에 본 논문에서는 기존의 PCA/LDA 방법 및 DCT를 이용하는 방법들과 성능 비교를 하고자 한다.

본 논문의 구성은 다음과 같다. II장에서 관련 연구를 소개하고 III장에서 DCT/LDA를 이용한 얼굴 인식 방법을 설명한다. IV장에서 AR 데이터 셋과 ORL 데이터 셋에 대한 얼굴 인식 실험 결과를 보이고, V장에서 결론을 맺는다.

II. 관련 연구

PCA와 LDA를 혼합한 PCA/LDA 방법은 일반적으로 우수한 인식 성능을 보여주고 있다. 그러나 PCA 기저 벡터를 구하는 과정에서 연산량이 많이 요구 되어 학습 시간이 오래 걸리는 단점이 있다. 따라서 학습 시간을 단축하기 위해 영상에 대해 PCA와 유사한 에너지 압축 효율을 갖는 것으로 알려진 2D DCT(2 Dimensional Discrete Cosine Transform)을 이용하여 PCA를 대체한 여러 가지 DCT/LDA 방법이 제안되었

다^{[8][9][10][11]}.

S. Annadurai는 입력된 얼굴 영상에서 중첩되지 않도록 8×8 block DCT를 이동하면서 수행하고 각 블록에서 선택된 $n \times n$ 개의 DCT 특징값들을 1차원 벡터로 표현한 후 이에 대해 LDA를 적용하는 Block-DCT/LDA 방법을 제안하였다^[8]. 이 때, 8×8 block DCT를 수행하기 위해 입력 영상을 8의 배수가 되도록 64×64 크기로 줄이는 과정을 거치고, 각 블록에서 선택하는 특징값의 수는 1×1 , 2×2 , 3×3 의 세 가지 중에서 선택할 수 있다. 만약 1×1 개의 DCT 특징값을 선택하는 경우, 이는 8×8 블록의 평균값을 의미하며 따라서 영상을 단순히 $1/8$ 로 줄이는 것과 같다. 그렇지 않고, 2×2 혹은 3×3 개의 DCT 특징값을 선택하는 경우, 각각 256개와 576개의 DCT 특징값으로 영상을 표현하고 이에 대해 LDA를 수행하게 되는데, 만약 트레이닝 데이터의 수가 이들보다 적게 되면 SSS 문제에 직면할 수 있는 가능성이 있다.

X.-Y. Jing은 Block-DCT/LDA 방법과 다르게 영상 전체에 대해 2D DCT를 수행하는 Discriminant DCT/LDA 방법을 제안하였으며, 2-D Separability Judgment 방법을 통해 DCT 계수들 중에서 변별력이 높은 것을 선택하여 LDA를 적용하였다^[9]. 2-D Separability Judgment 방법은 각 주파수 대역별 DCT 계수들에 대해 클래스 내의 분산은 작고 다른 클래스들과의 분산은 큰 주파수 대역을 선택하는 것으로 [6]과 [7]에서 제안된 FSWBV(Feature Selection method using Within and Between class Variance)와 유사하다. 다만, FSWBV는 DCT 계수 각각에 대해 변별력을 보는 반면, 2-D Separability Judgment 방법은 주파수 대역별로 변별력을 계산하는 차이점이 있다.

M. J. Er이 제안한 DCT/LDA/RBF 방법은 DCT/LDA 방법에 분류기(classifier)로서 RBF(Radial Basis Function) 신경망을 적용하였다^[10]. 또한, LDA 전에 DCT 특징값을 이용하여 클러스터링을 수행하였는데, 이 과정에서 두 얼굴 영상이 동일 인물이라도 서로 추출된 DCT 특징값이 다를 경우 다른 클래스(인물)로 간주하고 LDA를 수행한다. 이러한 클러스터링 과정을 통해 LDA가 더 잘 동작하도록 하였으며, 최종적으로 RBF 신경망을 통해 입력 영상을 인식함으로써 우수한 인식 성능을 보여주었다.

A. Pnevmatikakis가 제안한 DCT/PCA/LDA 방법은 DCT 특징값으로부터 PCA 기저 벡터를 구하고 구해진 PCA 기저 벡터에 DCT 특징값을 투영한 후 이에 대해 LDA를 수행한다^[11]. 여기서, DCT 특징값 추출과정에서

는 8×8 block DCT를 절반이 중첩되게 이동하면서 사용하였고 각 블록에서 DCT-mod2 방법^[12]을 통해 6+n 개의 DCT 계수를 선택한다. 단, $0 \leq n \leq 63$ 이며, [11]에서는 $n=26$ 을 사용하였다. 이 경우, 각 블록 당 29개의 DCT 특징값을 사용하고, 92×112 영상에 대해 총 14,500개(=900(블록 수)×29(블록 당 DCT 특징값 수))의 DCT 특징값이 생성된다. 따라서 원래 영상의 크기보다 추출된 DCT 특징값의 수가 더 많아지는 결과를 초래하며 LDA를 바로 적용할 수 없게 된다. 이 문제를 해결하기 위해 PCA를 통해 차수를 줄이고 LDA를 수행하게 된다.

이들 방법 외에 LDA는 사용하지 않지만 DCT를 사용하는 얼굴 인식 방법으로 DCT 방법^[13]과 DCT/PCA 방법^[14], DCT/HMM(Hidden Markov Model) 방법^[15] 등이 제안되었다.

본 논문에서 제안하는 DCT/LDA 방법은 S. Annadurai이 제안한 Block-DCT/LDA 방법과 X.-Y. Jing이 제안한 Discriminant DCT/LDA 방법과 유사한 특성을 갖는다. 그러나 영상 전체에 대해 DCT를 수행하는 점에서 Block-DCT/LDA 방법과 다르다. 또한 Discriminant DCT/LDA 방법과 비교할 때, 단순히 저주파 대역의 DCT 계수를 사용하는 점에서 다르지만, 우리가 이전에 제안했던 FSWBV 방법^{[6][7]}을 함께 사용한다면 DCT/LDA 방법은 Discriminant DCT/LDA 방법과 거의 유사한 학습 및 인식 과정을 가지며, 인식에 있어서도 비슷한 성능을 보일 것으로 기대된다.

III. DCT/LDA 기반 얼굴 인식

DCT/LDA를 이용한 얼굴 인식 방법은 그림 1과 같이 입력된 얼굴 전체 영상에 대해 2D DCT를 취한 후 DCT 계수에 대해 LDA를 수행함으로써 얼굴에 대한 특징값을 정의한다. DCT는 영상 신호에 대해 PCA와 유사한 에너지 압축 효율을 갖는다^[16]. 즉, PCA와 같이 DCT 또한 적은 수의 특징값으로 영상을 표현하는데 적합하며, 이를 통해 LDA의 SSS 문제를 해결할 수 있다. 또한, PCA/LDA 기반 얼굴 인식 방법의 학습 과정에서 대부분의 계산량이 PCA 기저 벡터를 구하는데 소요되는 점을 고려할 때, DCT는 기저 벡터가 고정되어 있으므로 DCT를 활용한 LDA 방법은 학습에 필요한 시간을 크게 줄일 수 있다.

DCT/LDA 방법의 학습 과정은 다음과 같다. 트레이닝 영상에 대해 조명의 영향을 줄이기 위한 전처리로 영상의 평균은 0이고 분산은 1이 되도록 ZMST(Zero

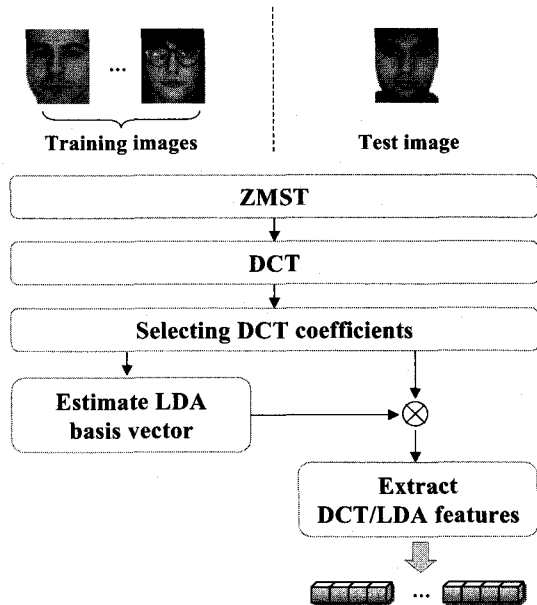


그림 1. DCT/LDA 특징값 추출 과정
Fig. 1. The process extracting DCT/LDA features.

Mean Unit Variance)^[17]를 수행한다. 다음으로 2D DCT를 수행하고^[16], DCT 계수들 중에서 DC 값을 제외한 저주파 영역의 계수들을 선택한다. 적은 수의 DCT 계수로 영상을 효율적으로 표현하기 위해 저주파 대역에 대해 그림 2와 같은 zigzag 순서로 DCT 계수를 선택하여 이를 DCT 특징값으로 사용한다. 이때 첫 번째 DCT 계수인 DC 값은 영상의 평균값으로서 조명의 영향을 많이 받고, 또한 조명의 영향을 줄이기 위한 전처리로 ZMST 과정을 거치면 DC 값은 항상 0이 되므로 특징값으로 사용하지 않는다. 이렇게 추출된 DCT 특징값에 대해 LDA를 학습하여 LDA의 기저 벡터를 추출한다^{[6][7]}.

인식 과정에서는 학습 과정과 같이 인식하고자 하는 얼굴 영상에 대해 ZMST를 수행하고 DCT 특징값을 추출한다. 추출된 DCT 특징값을 학습 과정에서 구한 LDA 기저 벡터에 투영시켜서 DCT/LDA 특징값을 계산한다. 추출된 DCT/LDA 특징값과 DB에 등록된 얼굴 영상의 특징값에 대해 유사도를 계산하고 가장 가까운

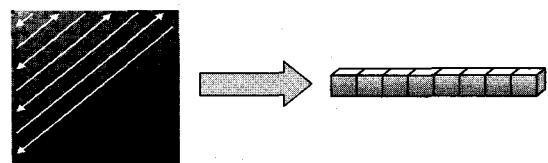


그림 2. Zigzag 순서에 의한 저주파 대역의 DCT 계수 선택
Fig. 2. Selecting DCT coefficients by zigzag scan in low frequency.

거리를 가지는 얼굴을 입력 영상의 얼굴로 인식한다. 본 논문에서는 유클리디언 거리(Euclidean distance)를 이용하여 유사도를 계산하였다.

IV. 실험 결과

제안된 DCT/LDA 기반 얼굴 인식 방법의 성능을 평가하기 위해 AR 데이터 셋과 ORL 데이터 셋을 이용하였고, Block-DCT/LDA와 PCA/LDA 기반 얼굴 인식 방법과 그 성능을 비교하였다. 입력 얼굴 영상과 DB에 들어있는 얼굴 영상과의 유사도는 유클리디언 거리를 이용하고, nearest-neighbor를 통해 입력 얼굴 영상을 인식하였다. 모든 실험은 Intel P-IV 2.5GHz CPU와 1Gbyte 메모리를 갖는 PC에서 Visual C++ 6.0을 이용하여 이루어졌고, PCA 및 LDA에 사용되는 고유 분석(Eigen analysis)을 위해 OpenCV^[18]를 이용하였다.

1. 실험 데이터 셋

AR 데이터 셋은 126명에 대해 그림 3과 같이 2주의 시간 간격을 두고 촬영된 26장의 정면 얼굴 영상으로 구성되어 있다^[19]. 본 논문에서는 임의로 선택된 100명(남자 50명, 여자 50명)에 대해 그림 3의 (a)~(g)와 같이 얼굴의 가려진 부분이 없는 첫 주기에 촬영된 7장의 영상을 이용하여 총 700장의 얼굴 영상으로 실험 데이터 셋을 구성하였다.

AR 데이터 셋의 각 영상은 768×576 크기를 갖는 컬러 영상이므로 이를 인식에 적합한 얼굴 영상을 만들기 위해 정규화 과정이 필요하다. 정규화 과정은 원본 영상에서 수작업에 의해 찾아진 두 눈의 위치를 그림 4와 같이 각각 (12, 18)과 (33, 18)에 위치하도록 원본 영상을 변형한다. 이때 정규화된 영상은 46×56의 크기를 가지며, R, G, B 컬러 채널의 평균값으로 표현된 그레이

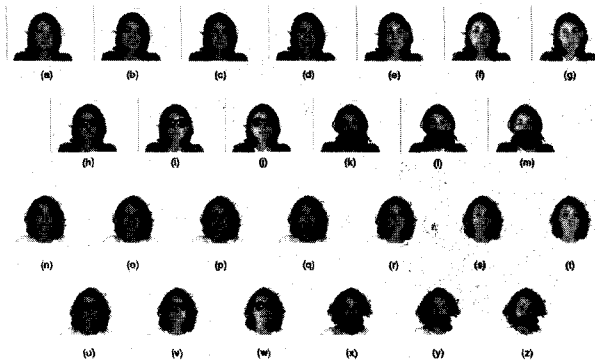


그림 3. AR 데이터 셋 ([3]에서 인용)
Fig. 3. AR face dataset (cited from [3]).

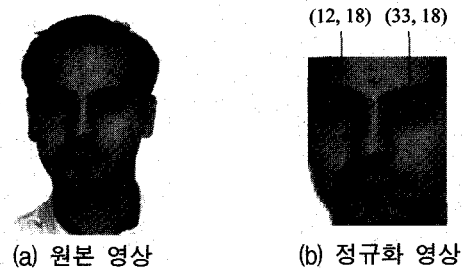


그림 4. 얼굴 영상 정규화
Fig. 4. Face normalization.



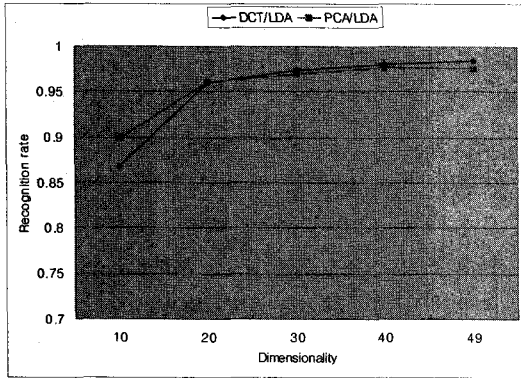
그림 5. ORL 데이터 셋
Fig. 5. ORL face dataset.

영상이다.

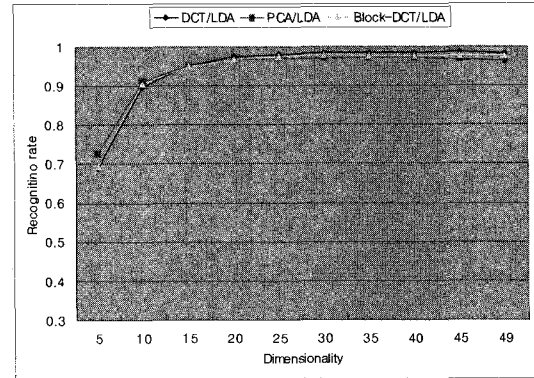
ORL 데이터 셋은 40명의 인물로 구성되어 있고 각 인물들은 그림 5와 같이 10장의 얼굴 영상을 갖는다^[20]. 이때 모든 영상은 92×112 크기의 그레이 영상이다. AR 데이터 셋에 비해 ORL 데이터 셋은 얼굴 영역만을 포함하도록 적당히 정규화가 이루어져 있으므로, 400장(=40명×10장)의 얼굴 영상에 대해 별도의 정규화 과정을 갖지 않고 원본 영상을 그대로 실험에 사용하였다.

2. AR 데이터 셋에 대한 얼굴 인식 실험

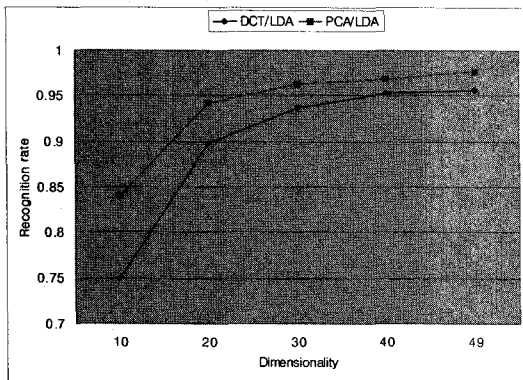
AR 데이터 셋에서 얼굴 인식 성능을 비교하기 위해 임의로 50명(350장)을 트레이닝 데이터로 선택하고 나머지 50명(350장)을 테스트 데이터로 사용하였으며, 트레이닝 데이터와 테스트 데이터를 무작위로 바꿔가면서 총 10가지의 트레이닝 데이터와 테스트 데이터 쌍을 생성하였다. 입력 영상과 DB에 들어있는 영상과의 유사도는 유클리디언 거리를 이용하고, nearest-neighbor를 통해 입력 영상을 인식하였다. 얼굴 인식 성능은 leave-one-out 방법^[21]을 이용하여 측정하였고, 10개 트레이닝 데이터와 테스트 데이터 쌍에 대한 평균 인식 성능을 그림 6에 나타내었다. 그림 6의 각 그래프는 사용된 DCT 및 PCA 특징값이 100개와 200개일 때 DCT/LDA 방법과 PCA/LDA 방법의 인식 성능을 보여 주며, x축은 인식에 사용된 LDA 특징값의 수를, y축은 인식률을 나타낸다.



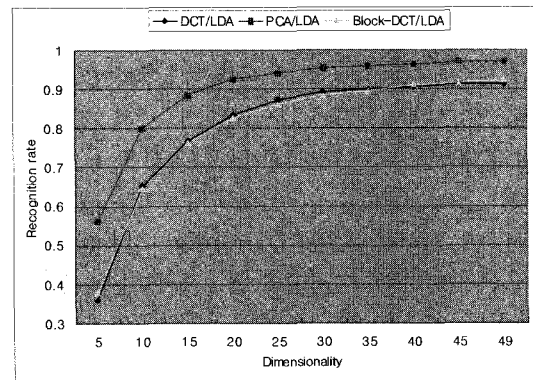
(a) $g=100$



(a) $g=64$



(b) $g=200$



(b) $g=256$

그림 6. DCT와 PCA 특징값을 (a) 100개, (b) 200 개 사용할 때 얼굴 인식 성능

Fig. 6. Face recognition rate when (a) 100, (b) 200 features of DCT and PCA are used.

그림 7. DCT와 PCA 특징값을 (a) 64개, (b) 256 개 사용할 때 얼굴 인식 성능

Fig. 7. Face recognition rate when (a) 64, (b) 256 features of DCT and PCA are used.

그림 6을 보면, DCT/LDA 방법과 PCA/LDA 방법 둘 다 사용된 LDA 특징값의 수가 증가함에 따라 얼굴 인식 성능이 증가함을 알 수 있으며, $g=100$ 일 때 두 방법은 유사한 성능을 보여주고 있다. 그러나 사용된 DCT 특징값과 PCA 특징값의 수가 100개에서 200개로 증가하더라도 성능 향상보다는 성능 저하를 보여주고 있으며, 특히 DCT/LDA 방법에서 성능 저하가 뚜렷하게 나타나고 있다. 이는 DCT 특징값 중에서 고주파 성분에 노이즈 성분이 많이 포함되어 있고 따라서 고주파 성분을 많이 사용함에 따라 성능이 저하되는 것으로 보이며, PCA/LDA 방법에서도 동일하게 해석할 수 있다. 이러한 현상에 대해 [9]와 [13]에서도 동일한 분석을 하고 있다.

그림 7은 Block-DCT/LDA 방법^[8]과 동일한 조건에서 얼굴 인식 성능을 비교하기 위해 $g=64$ 와 $g=256$ 일 때 얼굴 인식 성능을 보여준다. $g=64$ 일 때 Block-DCT/LDA 방법은 각 DCT 블록에서 1×1개의 DCT 특징값을 사용하고, $g=256$ 일 때 2×2개의 DCT 특징값을 사용한다. 각 DCT 블록에서 3×3개의 DCT 특

징값을 사용할 수도 있으나, 이렇게 되면 총 DCT 특징값의 수가 576개(=64(DCT 블록 수)×64(각 블록 당 DCT 특징값 수))가 되어 SSS 문제로 인해 LDA의 기저 벡터를 구하지 못하게 된다. 그림 7의 결과는 그림 6과 유사한 성능을 보여주고 있다. $g=64$ 일 때 세 방법 모두 유사한 성능을 보여주며, $g=256$ 일 때 PCA/LDA 방법에 비해 성능은 떨어지지만, Block-DCT/LDA 방법과 DCT/LDA 방법은 유사한 성능을 보여주고 있다.

표 1은 DCT, PCA, LDA 특징값의 수를 10개 단위로 변화시키면서 인식률을 계산하고, 최대 성능을 보일 때의 인식률과 사용된 LDA 특징값의 수를 보여준다. 여기서 LDA 특징값의 수는 최대 49개(=50(클래스 수)-1)까지 실험하였고, Block-DCT/LDA 방법은 DCT 특징

표 1. AR 데이터 셋에서 얼굴 인식 성능
Table 1. Face recognition on AR face dataset.

Method	Feature dimension	Recognition rate(%)
DCT/LDA	40(70)	98.66
Block-DCT/LDA	30(64)	97.97
PCA/LDA	49(130)	97.83

값의 수가 64개 혹은 256개 두 경우에 대해서 실험하였다. 결과를 보면, DCT/LDA 방법이 98.66%로 가장 우수한 결과를 보이고, Block-DCT/LDA 방법과 PCA/LDA 방법이 각각 97.97%, 97.83%를 보여준다. 이때 DCT/LDA 방법과 Block-DCT/LDA 방법에서는 각각 70개와 64개의 DCT 특징값이 사용되었고, PCA/LDA 방법에서는 130개의 PCA 특징값이 사용되었다.

3. ORL 데이터 셋에 대한 얼굴 인식 실험

ORL 데이터 셋에서는 [9][13][14]와 동일한 조건에서 인식 성능을 비교하기 위해 각 인물의 첫 5장의 영상을 트레이닝 셋으로 사용하고, 나머지 5장의 영상을 테스트 셋으로 사용하였다. 즉, 트레이닝 셋과 테스트 셋으로 각각 200장(=40명×5장)의 얼굴 영상을 사용하였다.

그림 8은 LDA 특징값 수에 따른 DCT/LDA, PCA/LDA, Block-DCT/LDA 방법의 인식률을 보여준다. Block-DCT/LDA 방법의 경우 SSS 문제로 인해 각 DCT 블록 당 1×1개의 DCT 특징값만을 이용할 수 있었으며, 동일한 조건에서 성능 비교를 위해 DCT/LDA

방법과 PCA/LDA 방법에서 DCT와 PCA 특징값의 수를 64개로 고정하였다. 이와 같은 조건에서 PCA/LDA 방법이 가장 우수한 성능을 보여주면서 최대 96% 인식률을 보여주었다. 다음으로 Block-DCT/LDA 방법과 DCT/LDA 방법의 순으로 인식 성능을 보여주고 있다.

표 2는 AR 데이터 셋에 대한 실험과 유사하게, DCT, PCA, LDA 특징값의 수를 변화시키면서 ORL 데이터 셋에서 최대 성능을 보일 때의 인식률과 사용된 LDA 특징값의 수를 보여주며, 괄호 안의 숫자는 사용된 DCT 혹은 PCA 특징값의 수이다. 여기서 DCT/LDA/RBF, Discriminant DCT/LDA, DCT/PCA, DCT 방법은 각각 [10], [9], [14], [13]에서 결과를 인용하였고, Eigenface 방법은 [9]의 결과를 인용하였다.

표 2에서 인식률만을 고려하면, PCA/LDA 방법이 98%로 가장 우수하고 DCT/LDA/RBF 방법과 Discriminant DCT/LDA 방법이 다음을 잇고 있다. DCT/LDA/RBF 방법은 얼굴을 인식하는 과정에서 유클리디언 거리식을 쓰지 않고 RBF 신경망을 쓰고 있으며, Discriminant DCT/LDA 방법은 DCT 특징값에서 단순히 저주파 성분을 취하지 않고 보다 변별력이 높은 DCT 특징값을 선택해서 이에 대해 LDA를 적용하고 있다. 따라서 단순히 저주파 대역의 DCT 특징값을 사용하는 DCT/LDA 방법 보다 성능이 우수한 결과를 보여주고 있으며, 만약 DCT/LDA 방법에 [6]과 [7]에서 제안된 FSWBV 방법을 적용하게 되면 현재보다 많은 성능 향상을 보일 것으로 기대된다. Block-DCT/LDA 방법의 경우, 총 64개의 DCT 블록에서 각 DCT 블록 당 1×1개의 DCT 특징값을 선택하고 있다. 이 경우 선택된 DCT 특징값은 결국 각 블록의 평균값에 해당하고 원래 영상을 1/8로 축소하는 효과만을 갖는다. 따라서 현재로서는 Block-DCT/LDA 방법이 DCT/LDA 방법과 동일한 성능을 보여주고는 있으나, Block-DCT/LDA 방법에서는 FSWBV와 같은 변별력이 높은 DCT 특징값 선택과 같은 방법을 적용할 수 없으며 더 이상의 성능 향상은 기대할 수 없다.

다음으로 DCT/PCA 방법과 Eigenface 방법, DCT 방법의 순으로 성능이 나타나고 있는데, 에너지 압축 관점에서 보면 DCT와 PCA는 거의 유사한 성능을 보이는 것으로 알려져 있고, 따라서 세 방법 모두 유사한 성능을 보일 것으로 기대할 수 있다. 실제로 Eigenface 방법과 DCT 방법은 거의 동일한 성능을 보이고 있고, 두 가지를 함께 쓴 DCT/PCA 방법이 좀 더 좋은 성능을 보여주고는 있으나 사용되는 특징값의 수가 많다는 단점이 있다.

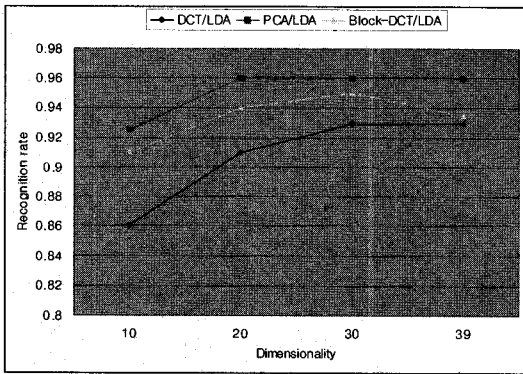


그림 8. DCT와 PCA 특징값을 64개 사용할 때 얼굴 인식 성능

Fig. 8. Face recognition rate when 64 features of DCT and PCA are used.

표 2. ORL 데이터 셋에서 얼굴 인식 성능
Table 2. Face recognition on ORL face dataset.

Method	Dimensionality	Recognition Rate
DCT/LDA	20(40)	95
PCA/LDA	20(90)	98
DCT/LDA/RBF ^[10]	30(55)	97.55
Discriminant DCT/LDA ^[9]	29(49)	97.5
Block-DCT/LDA ^[8]	30(64)	95
DCT/PCA ^[14]	110(192)	94.5
Eigenface ^[22]	74	91.11
DCT ^[13]	49	91

표 3. 학습 시간
Table 3. Training speed.

Method	Time(sec)
DCT/LDA	0.3939
Block-DCT/LDA	0.3294
PCA/LDA ¹	5393.203
PCA/LDA ²	16.1376

4. 학습 속도

학습 속도를 측정하기 위해 AR 데이터 셋에서 임의의 50명(350장)을 이용하여 학습 하였다. Block-DCT/LDA 방법과의 비교를 위해 DCT/LDA 방법과 PCA/LDA 방법에서 DCT와 PCA 특징값을 64개 사용하였고 이때 학습에 소요된 시간을 표 3에 나타내었다. 표 3에서 PCA/LDA¹은 PCA 기저 벡터를 구하기 위해 2,576×2,576 크기의 공분산 행렬에 대해 고유 분석을 한 경우이고, PCA/LDA²는 PCA 기저 벡터를 구할 때 특징 공간(feature space)의 차수 보다 트레이닝 데이터의 수가 적은 경우 적용할 수 있는 snap-shot-method [23]를 적용한 경우이다.

표 3에서 알 수 있듯이, DCT/LDA 방법과 Block-DCT/LDA 방법은 고정된 기저 벡터를 갖는 DCT를 이용하기 때문에 학습 속도가 매우 빠르게 이루어진다. 반면에 앞서 언급한 바와 같이 PCA/LDA 방법의 경우 PCA 기저 벡터를 구하기 위한 고유 분석을 하는 과정에서 매우 많은 계산량을 요구한다. 비록 350×350 크기의 공분산 행렬에 대해 고유 분석을 하는 snap-shot-method를 적용하더라도 DCT/LDA 방법에 비해 학습 시간이 오래 걸리게 된다.

V. 결 론

본 논문에서는 DCT/LDA를 이용한 얼굴 인식 방법을 제안하였다. 제안된 방법은 LDA의 SSS 문제를 해결하기 위해 입력 영상을 적은 수의 DCT 계수로 표현하고 이에 대해 LDA를 적용하였다. LDA의 SSS 문제를 해결하기 위한 기존의 PCA/LDA 방법은 트레이닝 데이터의 영상 크기가 커짐에 따라서 PCA 기저 벡터를 구하는 시간이 매우 오래 걸리는 단점이 있다. 그러나 DCT/LDA 방법은 기저 벡터는 고정되어 있으면서 PCA와 유사한 에너지 압축 효율을 갖는 DCT를 이용하기 때문에 학습 속도는 빠르면서 동일한 얼굴 인식 성능을 기대할 수 있다.

실험에서 AR 데이터 셋과 ORL 데이터 셋을 이용하여 기존의 PCA/LDA 방법을 비롯해서 DCT를 이용하

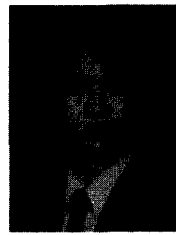
는 유사한 방법들에 대한 얼굴 인식 성과와 학습 속도를 비교하였다. 실험 결과 DCT/LDA 방법이 AR 데이터 셋에서 PCA/LDA 방법과 Block-DCT/LDA 방법 보다 우수한 성능을 보여주었고, ORL 데이터 셋에서는 PCA/LDA 방법 보다는 낮지만, Block-DCT/LDA 방법과 동일한 인식 성능을 보여주었다. 학습 속도에 있어서는 기대한 것과 같이 PCA/LDA 방법 보다 매우 빠른 속도로 학습이 가능하였다. 그러나 Block-DCT/LDA 방법은 8×8의 작은 block DCT를 수행하기 때문에 DCT/LDA 방법이 Block-DCT/LDA 방법 보다 학습 속도가 느렸으며, 얼굴 영상의 크기가 커질수록 DCT/LDA 방법이 더 느려질 것으로 예측된다.

참 고 문 헌

- [1] R. Gross, I. Matthews, and S. Baker, "Appearance-Based Face Recognition and Light-Fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, No. 4, pp. 449-465, April 2004.
- [2] J. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Face Recognition Using LDA-Based Algorithms," *IEEE Transactions on Neural Networks*, Vol. 14, No. 1, pp. 195-200, January 2003.
- [3] A. M. Martinez and A. C. Kak, "PCA versus LDA," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 2, pp. 228-233, February 2001.
- [4] W. Zhao, R. Chellappa, and A. Krishnaswamy, "Discriminant Analysis of Principal Components for Face Recognition," *Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 336-341, April 1998.
- [5] 장혜경, 오선문, 강대성, "PCA 기반 LDA 혼합 알고리즘을 이용한 실시간 얼굴인식 시스템 구현," *전자공학회논문지*, 제41권 SP편, 제4호, 45-50쪽, 2004년 7월
- [6] H.-J. Lee, H.-J. Kim, and W.-Y. Kim, "Face Recognition using Component-based DCT/LDA," 제2회 BERC 생체인식 워크샵, pp. 219-222, 2004년 2월
- [7] H.-J. Lee, H.-J. Kim, and W.-Y. Kim, "Face Recognition Using Component-based DCT/LDA," *International Workshop on Advanced Image Technology (IWAIT2005)*, pp. 25-30, January 2005.
- [8] S. Annadurai and A. Saradha, "Discrete Cosine Transform Based Face Recognition Using Linear Discriminant Analysis," *IJSIT Lecture Note of International Conference on Intelligent*

- Knowledge Systems, Vol. 1, No. 1, pp. 90-94, August 2004.
- [9] X.-Y. Jing and D. Zhang, "A Face and Palmprint Recognition Approach Based on Discriminant DCT Feature Extraction," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 34, No. 6, pp. 2405-2415, December 2004.
- [10] M. J. Er, W. Chen, and S. Wu, "High-Speed Face Recognition Based on Discrete Cosine Transform and RBF Neural Networks," *IEEE Transactions on Neural Networks*, Vol. 16, No. 3, pp. 679-691, May 2005.
- [11] A. Pnevmatikakis and L. Polymenakos, "Comparison of Eigenface-Based Feature Vectors under Different Impairments," 17th International Conference on Pattern Recognition (ICPR'04), Vol. 1, pp. 296-299, August 2004.
- [12] C. Sanderson, "Automatic Person Verification Using Speech and Face Information," Ph. D. Thesis, Griffith University, February 2003.
- [13] Z. M. Hafeed and M. D. Levine, "Face Recognition Using the Discrete Cosine Transform," *International Journal of Computer Vision*, Vol. 43, No. 3, pp. 167-188, 2001.
- [14] Z. Jianke, V. M. I, and M. P. Un, "Face Recognition Using 2D DCT with PCA," Chinese Conference on Biometric Recognition (Sinobiometrics'03), December 2003.
- [15] V. V. Kohir and U. B. Desai, "Face Recognition Using a DCT-HMM approach," IEEE Workshop on Applications of Computer Vision (WACV'98), pp. 226-231, 1998.
- [16] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, second ed., Prentice Hall, 2002.
- [17] J. Kittler, Y. P. Li, and J. Matas, "On Matching Scores for LDA-based Face Verification," The British Machine Vision Conference (BMVC2000), pp. 42-51, September 2000.
- [18] <http://www.intel.com/research/mrl/research/opencv/>
- [19] A.M. Martinez and R. Benavente, "The AR Face Database," CVC Technical Report 24, June 1998.
- [20] <http://www.uk.research.att.com/facedatabase.html>
- [21] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, second ed., Academic Press, 1990.
- [22] M. Turk and A. Pentland, "Eigenfaces for recognition," *International Journal of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86, 1991.
- [23] L. Sirivich and R. Everson, "Management and analysis of large scientific datasets," *International Journal of Supercomputer Applications*, Vol. 16, No. 1, pp. 50-68, 1992.

 저자 소개



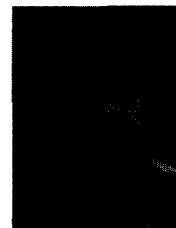
김형준(학생회원)

1999년 한양대학교 전자·전자통신 전파공학과 학사 졸업.

2001년 한양대학교 전자통신전파공학과 석사 졸업.

2001년~현재 한양대학교 전자통신컴퓨터공학부 박사과정

<주관심분야 : 영상처리, 얼굴 인식, 멀티미디어 검색, MPEG-7 등>



정병희(정회원)

1994년 이화여자대학교

전자계산학과 학사 졸업.

1996년 한국과학기술원 전산학과 석사 졸업.

1996년~현재 KBS 방송기술 연구팀 선임연구원.

2000년~현재 한국과학기술원 전자전산학과 (전산학전공) 박사과정.

<주관심분야 : 멀티미디어 검색, 비디오 요약, 멀티미디어 저작 환경 등>



김희울(정회원)

1980년 한양대학교

전자공학과 학사 졸업.

1983년 Pennsylvania State University 전기공학과 석사 졸업.

1989년 Purdue University 전기공학과 박사 졸업.

1989년~1994년 University of Texas 조교수.

1994년~현재 한양대학교 전자통신컴퓨터공학부 정교수.

<주관심분야 : 영상처리, 패턴 인식, 의료 영상, 머신비전, 멀티미디어 검색, MPEG-7 등>