

영상의 1차 모멘트와 고유벡터를 이용한 얼굴인식

조 용 현[†]

요 약

본 논문에서는 영상의 1차 모멘트와 고유벡터를 이용한 효율적인 얼굴인식 방법을 제안하였다. 여기서 1차 모멘트는 입력되는 얼굴영상의 중심좌표를 계산하는 것으로 이는 영상의 중심이동에 따른 전처리로 인식에 불필요한 배경을 배제시킴으로써 인식성능을 개선하기 위함이다. 고유벡터는 얼굴의 특징인 기저영상으로 주요성분분석을 이용하여 추출하였다. 이는 2차의 통계성을 고려한 중복신호의 제거로 인식성능을 개선하기 위함이다. 제안된 방법을 각각 320*243 픽셀의 60개(15명*4장) 얼굴영상에 적용하여 city-block, Euclidean, 그리고 negative angle의 3가지 거리 척도를 분류척도로 이용하여 실험하였다. 실험결과, 중심이동의 제안된 방법은 전처리과정을 거치지 않는 기존방법보다 45개의 시험영상에서 평균적으로 약 1.6배 정도의 우수한 인식률과 약 5.9배 정도의 정확한 분류가 가능함을 확인하였다. 특히 city-block이 Euclidean이나 negative angle의 거리척도보다 상대적으로 정확하게 분류함을 알 수 있다.

Face Recognition Using First Moment of Image and Eigenvectors

Yong-Hyun Cho[†]

ABSTRACT

This paper presents an efficient face recognition method using both first moment of image and eigenvector. First moment is a method for finding centroid of image, which is applied to exclude the needless backgrounds in the face recognitions by shifting to the centroid of face image. Eigenvector which are the basis images as face features, is extracted by principal component analysis(PCA). This is to improve the recognition performance by excluding the redundancy considering to second-order statistics of face image. The proposed methods has been applied to the problem for recognizing the 60 face images(15 persons * 4 scenes) of 320*243 pixels. The 3 distances such as city-block, Euclidean, negative angle are used as measures when match the probe images to the nearest gallery images. In case of the 45 face images, the experimental results show that the recognition rate of the proposed methods is about 1.6 times and its the classification is about 5.6 times higher than conventional PCA without preprocessing. The city-block has been relatively achieved more an accurate classification than Euclidean or negative angle.

Key words: Face Recognition(얼굴인식), First Moment(1차 모멘트), Centroid Shift(중심이동), Eigenvector(고유벡터), Basis Image(기저영상), Principal Component Analysis(주요성분분석)

1. 서 론

얼굴인식은 임의의 입력 얼굴영상에 대하여 데이

터베이스 내에 저장된 가장 유사한 얼굴로 인식시키는 것이다. 하지만 동일한 사람의 얼굴이라도 얼굴 표정, 포즈, 크기의 변화, 조명의 변화, 안경착용, 화

※ 교신저자(Corresponding Author) : 조용현, 주소 : 경북 경산시 하양읍 금락리 330(712-702), 전화 : 053)850-2747, FAX : 053)850-2704, E-mail : yhcho@cu.ac.kr

접수일 : 2005년 9월 9일, 완료일 : 2005년 11월 3일
[†]정회원, 대구가톨릭대학교 공과대학 컴퓨터·정보통신공학부

장, 노화, 장식품 등 여러 가지 왜곡 요인에 의해 획득된 얼굴영상의 정보가 크게 변한다[1-4]. 따라서 얼굴영상 자체의 방대한 정보량과 이에 따른 많은 계산량은 얼굴인식의 과제라 할 수 있다. 이처럼 많은 왜곡 요인들에 영향을 받지 않는 우수한 인식율과 실생활에 적용할 수 있는 계산 속도를 가지기 위한 여러 가지 방법들이 연구되고 있다[5-8].

일반적으로 얼굴인식에는 특징을 기반으로 한 방법과 영상을 기반으로 하는 방법이 있다[4]. 특징을 기반으로 하는 방법은 얼굴의 특징과 관련된 정보들을 이용하는 방법이며, 영상을 기반으로 하는 방법은 얼굴영상 전체를 대상으로 특징을 계산하는 방법이다. 최근 얼굴인식을 위한 방법으로 영상을 기반으로 하는 방법이 널리 이용되고 있다.

특징기반 얼굴인식 방법에는 얼굴의 주요 부분인 눈, 코, 입 등의 여러 특징을 추출하여 인식에 이용하는 지역적인 특징추출 기법, 색깔과 형태로 구성된 2차원의 템플릿을 이용하는 템플릿 패턴정합 방법, 특정한 점에서의 방향과 변화에 따른 특성을 벡터화한 Gabor jet을 특징으로 사용하는 Gabor 필터를 이용한 방법들이 있다[4]. 이들 방법들에서는 얼굴을 표현하기 위한 차원감소나 수행시간의 절약, 포즈나 크기 등에 강건한 특성이 있으나 인식에 필요한 정보 손실이나 템플릿 구성에 따른 공간 소모, 변형의 정도에 따라 인식성능이 크게 달라지는 제약이 있다.

영상 기반 방법은 입력 얼굴영상에서 직접 얼굴의 코딩을 얻어내는 외형을 기반으로 하며, 여기에는 입력 영상내의 상관관계를 제거하여 영상의 진폭스펙트럼을 획득하는 주요성분분석(principal component analysis : PCA), PCA를 일반화하여 영상의 위상 스펙트럼까지도 고려하는 독립성분분석(independent component analysis : ICA), 선형식별 분석(linear discriminant analysis : LDA), support vector machine(SVM) 방법 등이 있다[5-8]. 이들 방법들 역시 데이터의 차원 감소나 다차원의 특징분석, 얼굴영상의 클래스들 간 최적구별, 처리시간 감소 등의 특징을 가진다. 특히 입력패턴으로부터 자동으로 출력패턴을 생성하는 기법인 PCA와 ICA는 얼굴을 인식하는 방법으로 가장 널리 이용되고 있다. 이 방법들은 중복신호 제거의 효과가 있어 뇌의 시각 기능을 잘 반영하며 타 방법에 비해 높은 인식률을 얻을 수 있는 장점을 갖고 있다. PCA가 신호의 차원 자체를 줄여 조밀한 코

딩을 유도하는데 비해 ICA는 출력 신호의 값을 드물게 나타내는 방법으로 중복 신호를 제거한다[9-11]. 특히 PCA는 얼굴영상을 표현하는 고유벡터를 추출하여 벡터 공간으로 투영시킨 가중치를 이용하는 방식으로, 전체의 얼굴영상들을 적은 수의 벡터를 이용하여 표현할 수 있으며 빠르고, 계산이 간단하며, 점진적 변화에 덜 민감하다[7,8]. 또한 데이터 정보의 손실이 거의 없이 데이터 차원을 줄여 데이터 패턴을 파악할 수 있으며, 전체적인 영상의 특성을 파악할 수 있다는 장점을 가진다. 하지만 PCA를 이용할 경우, 눈, 코, 입 등과 같은 세부적인 표현의 제약이 있고, 표정이나 크기 및 위치 변화 등의 기하학적 왜곡에 민감하기 때문에 이에 대한 처리가 요구된다.

본 논문에서는 영상의 1차 모멘트[12]와 고유벡터를 이용한 효율적인 얼굴인식 방법을 제안한다. 여기서 1차 모멘트는 입력되는 얼굴영상의 중심좌표를 계산하는 것으로 이는 영상의 중심이동에 따른 기하학적 왜곡에 강건한 전처리로 인식에 불필요한 배경을 배제시켜 인식성능을 개선하기 위함이다. 고유벡터는 얼굴의 특징인 기저영상으로 주요성분분석을 이용하여 추출하였으며, 이는 2차의 통계성을 고려한 중복신호의 제거로 인식성능을 개선하기 위함이다. 제안된 방법을 Yale 얼굴영상 데이터베이스[13]로부터 320*243 픽셀의 선택된 60개(15명*4장) 영상을 대상으로 city-block, Euclidean, 그리고 negative angle의 3가지 거리를 분류척도로 이용하여 실험하였다. 실험결과, 중심이동의 전처리된 제안된 방법과 전처리를 거치지 않는 기존방법의 결과를 비교·검토하였다.

2. 영상의 중심이동과 고유영상의 추출

얼굴인식의 성능과 효율을 높이기 위한 영상의 전처리 기법에는 기하학정보들에 대한 정규화 기법과 화소들에 직접적인 효과를 갖는 필터링 기법들이 있다[5-8]. 일반적으로 영상은 배경과 의미 있는 형상으로 구분된다. 배경이 형상에 비해 비교적 변화가 적은 영상일 경우, 주어진 영상의 크기로부터 보다 작은 크기의 유효한 영상을 추출할 수 있다. 이렇게 추출된 영상을 입력영상으로 이용하면 보다 빠르고 정확한 인식이 가능하다. 한편 영상을 기반으로 하는 얼굴인식에서 영상의 크기나 위치 등의 변화는 고유영상을 이용한 기법에서 인식성능에 많은 영향을 미

친다[7].

유효영상의 중심을 계산하기 위하여 단면의 1차 모멘트[12]를 이용한다. 이는 토목공학에서 보의 처짐을 위해 많이 이용되는 면적의 중심을 찾는 기법으로 여기서는 중심이동에 의한 배경을 제거함으로써 영상의 차원을 감소시키기 위함이다. 특히 중심계산에 따른 차원의 감소는 필요한 얼굴영역만의 추출로 크기나 위치에 무관한 인식을 가능하게 한다. 또한 전처리된 얼굴의 고유영상을 추출하기 위해 주요성분분석을 이용한다.

2.1 영상의 중심이동

입력되는 일반적인 얼굴영상을 X축 또는 Y축으로 평행 이동시켜 이루어지는 중심이동은 차 영상의 1차 모멘트에 의해 유효영상의 중심을 계산함으로써 가능하다. 이때 1차 모멘트를 구하는 과정은 먼저 원 영상 OrgImage로부터 차 영상 DeltaImage를 생성한다. 여기서 n번째 좌표의 차 영상 $\Delta I(n) = P_i(n) - P_i(n-1)$ 로 계산되며, Y행 화소좌표로부터 X열 화소좌표 방향으로 계산을 수행한다. 여기서 $P_i(n)$ 은 i번째 행의 n번째 열 픽셀값이다. 그림 1은 입력되는 $m \times n$ 픽셀의 원 영상과 계산된 차 영상 DeltaImage를 각각 나타낸 것이다.

다음은 계산된 차 영상으로부터 유효한 영상의 화소만을 생성하기 위하여 논리필터링을 수행한다. 이를 위해 차 영상에 대한 평균 화소값을 구하고, 그 값을 유효한 변화의 판별을 위한 값으로 이용한다. 이때 논리필터링은 차 영상의 값이 판별값보다 크거나 같으면 유효한 화소로 1의 값을 갖도록 하며, 그렇지 않으면 0의 값을 갖도록 함으로써 균일한 형상의 영상을 생성한다. 그림 2는 논리필터링 후 1차 모멘트를 구하기 위한 임의의 영상단면을 나타낸 것이다.

$P_i(0)$...	$P_i(n-1)$	$P_i(n)$	0	$P_i(n) - P_i(n-1)$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$P_m(0)$...	$P_m(n-1)$	$P_m(n)$	0	$P_m(n) - P_m(n-1)$

그림 1. $m \times n$ 픽셀의 원 영상과 차 영상

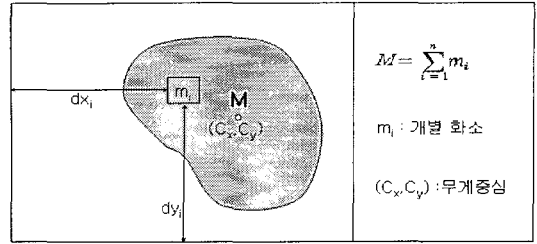


그림 2. 1차 모멘트를 위한 영상단면

여기서 1차 모멘트 평형은 각 개별 화소에 대한 모멘트 값의 합이 영상 전체의 중심거리와 영상 전체의 화소 값의 곱과 동일함을 의미한다[12]. 이때 중심계산을 위한 1차 모멘트 평형을 수식으로 표현하면 다음의 식 (1)과 같다.

$$M \times C_x = \sum_{i=1}^n (m_i \times dx_i), \quad M \times C_y = \sum_{i=1}^n (m_i \times dy_i) \quad (1)$$

여기서 m_i 는 i번째 픽셀의 논리값으로 유효한 화소이면 1, 그렇지 않으면 0이며, M 은 유효영상의 논리값에 대한 합이다. 이때 유효영상의 중심좌표 (C_x, C_y) 는 식 (2)와 같이 계산된다.

$$C_x = \frac{\sum_{i=1}^n dx_i}{M}, \quad C_y = \frac{\sum_{i=1}^n dy_i}{M} \quad (2)$$

따라서 1차 모멘트 평형으로부터 계산된 유효영상의 좌표 (C_x, C_y) 를 중심으로 보다 작은 영상을 선택하여 입력영상으로 이용한다. 즉, 임의로 입력되는 원 영상을 구해진 중심좌표로 이동시켜 유효한 영상을 선택하면 인식에 불필요한 배경을 배제시킬 수 있어 차원이 감소된 영상을 얻을 수 있다. 이렇게 하면 얼굴영상의 특징인 기저영상을 추출을 위해 이용되는 PCA의 입력영상 데이터를 줄일 수 있어 인식시간과 인식률을 개선시킬 수 있다.

2.2 고유영상 추출

통계학, 신호처리, 그리고 신경망 등에서 적당한 변환에 의한 알맞은 데이터의 표현을 찾기 위해 선형변환이 이용된다. 그 중에서 PCA는 데이터의 차원을 감소시키는 선형변환 기법으로 널리 이용되고 있는 2차원 방법이다[9-11]. PCA는 입력 데이터의 공분산에 의해 결정되는 정규직교기저로 표현되며, 가우

스 변수의 모든 정보는 공분산행렬 내에 포함되기 때문에 가우스 분포의 입력 데이터에 대해서는 적당한 것이다. 일반적으로 영상인식에서 영상을 기반으로 하는 방법은 영상 내에 포함된 특징추출에 따라 인식성능은 매우 달라진다. 특히 영상 내 각 픽셀사이의 2차원 또는 그 이상의 고차원 기저영상은 주요한 특징이며, 이러한 기저영상의 추출에 PCA와 ICA가 널리 이용되고 있다[8].

PCA는 입력데이터의 특징을 추출하는 기법으로 데이터 내에 포함된 정보를 추출하고 압축하여 통계적 규칙들을 찾아내는 것이다. 이는 대용량의 입력데이터를 통계적 독립인 특징들의 집합으로 변환시키는 것이며, n차원 입력공간의 데이터를 k차원 출력공간의 데이터로 투영시키는 것이다. 여기서 $k \ll n$ 이면 입력데이터 벡터가 가지는 대부분의 내부정보를 유지하면서도 차원의 감소가 가능하게 된다.

자기상관행렬 (autocorrelation matrix) $R_{xx} = \langle xx^T \rangle$ 를 가진 평균이 영인 입력벡터 $x = [x_1, x_2, \dots, x_n]^T$ 에 대해서 생각해 보자. 여기서 T는 전치를 나타내며, $\langle \cdot \rangle$ 는 기대치(expected value)를 나타낸다. 또한 $\hat{w}_1, \hat{w}_2, \dots, \hat{w}_n$ 이 R_{xx} 의 고유벡터(eigen vector)와 직교되는 연결가중치 벡터라 할 때, $\hat{w}_1 = [\hat{w}_{11}, \hat{w}_{12}, \dots, \hat{w}_{1n}]^T$ 는 가장 큰 고유값(eigen value) λ_1 과 일치하며, $\hat{w}_2 = [\hat{w}_{21}, \hat{w}_{22}, \dots, \hat{w}_{2n}]^T$ 는 두 번째로 큰 고유값 λ_2 , 그리고 $\hat{w}_n = [\hat{w}_{n1}, \hat{w}_{n2}, \dots, \hat{w}_{nn}]^T$ 는 가장 작은 고유값 λ_n 과 각각 일치한다. 이상의 관계를 행렬방정식으로 나타내면 식 (3)과 같다.

$$R_{xx}\hat{w}_j = \lambda_j\hat{w}_j, \quad (j=1,2,\dots, n) \quad (3)$$

여기서 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ 이다. 주어진 입력벡터 x 를 위한 첫 번째 k개의 주요 특징을 나타내는 고유벡터 y 는 다음의 선형변환식 (4)로 나타낼 수 있다.

$$y = \hat{W}x \quad (4)$$

여기서 $\hat{W} = [\hat{w}_1, \hat{w}_2, \dots, \hat{w}_k]^T \in R_{k \times n}$ 이며, 이 식에서 연결가중치행렬 \hat{W} 의 행은 가장 큰 고유값과 일치하는 상관행렬 R_{xx} 의 고유벡터임을 의미한다. 다시 말하면, 입력데이터 공간에서 k차원의 주요특징을 나타내는 부공간은 R_{xx} 의 k개 주요 고유벡터에 의해 구성된 부공간으로 정의된다. 결국 PCA는 $\langle \|\hat{w}_j^T x\|_2^2 \rangle$ 가 최대인 고유벡터 $\hat{w}_1, \hat{w}_2, \dots, \hat{w}_k$ 의 방향을

찾는 것이다. 일반적으로 얻어지는 고유값은 크기에 따라 정렬하고 고유벡터도 해당 고유값의 위치대로 정렬한다. PCA에서 순서대로 정렬된 고유값의 뒤쪽은 0에 가까운 값을 가지게 되어 이를 삭제할 수 있다. 이는 고유벡터의 작은 값들을 고려하지 않음으로써 입력 데이터의 차원을 줄이기 위함이다.

입력되는 데이터의 주요특징들을 추출하기 위한 수치적 기법에 의한 PCA들은 먼저 자기상관행렬 R_{xx} 를 계산한 다음 그것과 일치하는 고유벡터를 계산한다. 그러나 입력데이터 벡터가 대규모의 차원을 가진다면, 상관행렬은 대단히 크게 되어 요구되는 고유벡터를 찾는 것은 매우 복잡한 일이다. 이를 해결하기 위한 대안으로 공분산 행렬의 차원을 감소시켜 계산량을 줄이는 수치적 방법이나 상관행렬의 추정을 위한 계산이 요구되지 않는 단층구조의 신경망을 이용한 방법도 있다[10].

한편 PCA를 좀 더 효과적으로 수행하기 위해 신호를 대상으로 영 평균(zero-mean)을 수행한다[9,10]. 이는 신호의 1차적 통계성을 고려한 정규화로 PCA를 간략화 하기 위함이다. 신호의 영 평균은 신호벡터 x 에서 평균값 x^* 를 뺀 차이이며, 영 평균된 새로운 신호벡터 x 는 다음 식 (5)와 같다.

$$x = x - x^* \quad (5)$$

따라서 중심이동의 전처리된 얼굴영상을 대상으로 PCA를 수행하면 2차의 고유영상을 추출할 수 있다. 중심이동과 PCA를 조합한 제안된 얼굴인식 과정을 요약하면 그림 3과 같다. 여기서는 고유값과 고유벡터의 계산은 복잡성을 줄이기 위하여 snapshot 방법을 이용하였다[9]. 학습과정은 먼저, 1차 모멘트 평형



그림 3. 제안된 얼굴영상 인식과정

을 이용하여 획득된 학습영상의 중심좌표를 계산하여 얼굴의 중심을 찾는다. 다음은 찾아진 중심좌표를 기준으로 차원이 감소된 유효영상을 획득하는 전처리 과정을 수행한다. 전처리된 영상을 대상으로 영평균과 공분산행렬을 각각 계산하여 고유값과 고유벡터를 계산한다. 한편 영상의 차원을 감소시켜 계산을 간단하게 하기 위해 고유값을 정규화하고 내림차순으로 나열하여 정해진 개수만큼의 고유벡터를 추출한다. 추출된 고유벡터는 얼굴의 특징벡터로 2차원의 속성을 가지는 기저영상이며 시험영상의 판별을 위해 저장한다. 한편 식별과정에서도 시험얼굴을 획득한 후 전처리하여 입력영상보다 작은 유효영상을 추출한다. 다음은 유효영상으로부터 학습영상의 평균값을 뺀 차 영상인 영 평균을 구하고, 마찬가지로 저장된 고유벡터를 이용하여 시험영상의 기저영상을 계산한 다음, 저장된 기저영상과 비교 판별하여 얼굴을 인식하게 된다. 이때 유사성 판별을 위한 거리척도로 city-block, Euclidean, negative angle을 각각 이용하였다. 결국 시험얼굴에 대해 유사성이 가장 많은 즉 거리가 가장 가까운 학습영상이 추출되어 인식되게 된다.

3. 실험 및 결과분석

중심이동과 PCA에 의한 제안된 얼굴인식 방법의 성능을 평가하기 위해 320*243 픽셀의 60개(15명*4장) Yale 얼굴 데이터베이스[13] 영상을 대상으로 실험하였다. 여기서 학습 얼굴영상은 15명의 각 표준영상이고, 시험 얼굴영상은 표준 얼굴영상에서 위치, 표정, 안경착용의 변화를 준 영상으로 각 3장씩 총 45개(15명*3장)이다. 특징추출에 의한 복원성능을 평가하기 위해 원 영상의 픽셀값 x_i 와 복원된 영상의 픽셀값 y_i 사이의 절대평균합오차 AMSE(absolute mean sum error) = $(1/N) \sum |x_i - y_i|$ 를 정의하였으며, 여기서 N 은 얼굴영상의 전체 픽셀수이다. 실험은 펜티엄 IV -2.2G 컴퓨터에서 Matlab 5.1로 구현하였다.

그림 4는 15명의 얼굴영상들로 기저영상의 데이터베이스 구성을 위해 학습에 이용된 표준 얼굴들이다. 그림 5는 15명의 학습자에 대한 표준 얼굴영상을 대상으로 다양한 변화를 준 시험 얼굴영상이며, 전체 45개 중에서 12개(4명*3장)의 얼굴영상들만 나타낸

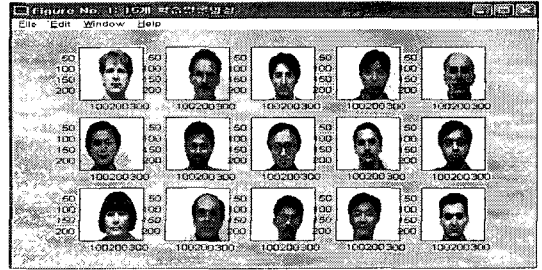


그림 4. 15명의 표준 학습 얼굴영상

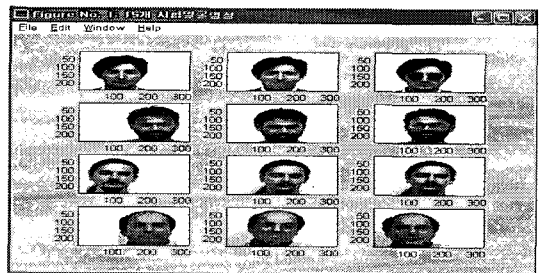


그림 5. 4명의 시험 얼굴영상

것이다. 여기서 보면 각 학습자별로 첫째와 둘째 열의 영상들은 위치와 표정의 다양한 변화가 있으며, 셋째 열의 영상들은 안경착용뿐만 아니라 위치의 변화된 것이다.

한편 그림 6은 그림 5의 첫째 행 첫째 열의 시험얼굴을 대상으로 판별값으로 평균화소값의 0.25 정도로 하여 논리 필터링한 후, 1차 모멘트에 의해서 계산된 중심좌표를 함께 나타낸 영상과 구해진 중심좌표로 얼굴영상을 중심이동 시킨 다음, 77760 (320*243) 픽셀의 영상을 28032(192*146) 픽셀로 축소시킨 전처리된 얼굴영상이다. 여기서 좌측 영상은 각 픽셀에 대해 차 영상 $\Delta Image(n) = P_i(n) - P_i(n-1)$ 을 구한 후, 그에 대한 화소의 평균값을 기준으로 설정된 판별값을 이용하여 논리 필터링한 것으로 얼굴영상의 경계영역이 추출됨을 알 수 있다. 우측 영상은

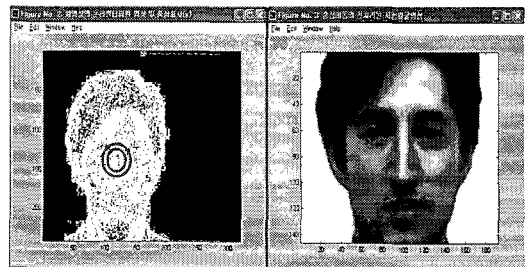


그림 6. 논리필터링 및 전처리된 영상

구해진 중심좌표로 얼굴영상을 중심이동 시킨 후, 차원을 감소시킨 전처리된 시험영상이며, 원 영상에 비해 약 2.8배정도 차원이 감소된 영상으로 실험에서는 이 영상을 이용한다. 이는 입력된 영상에서 인식에 불필요한 배경을 제거하여 입력데이터의 차원을 감소시켜 인식성능을 개선하기 위함이다.

그림 7은 위치가 중심으로부터 벗어난 위의 시험영상을 대상으로 제안된 PCA 방법을 이용한 후 city-block, Euclidean, negative angle의 3가지 거리척도에 따라 식별된 결과를 나타낸 것이다. 그림 7(a)는 시험얼굴이고, 그림 7(b), (c), (d)는 각각 city-block, Euclidean, negative angle의 거리척도에 따라 시험얼굴과의 거리에 따라 식별된 영상을 가까운 순서대로 위에서부터 3개씩만을 나타낸 것이다. 3가지 거리척도 모두 육안으로 보면, 첫 번째 나타난 영상은 시험얼굴과 동일함을 알 수 있다. 이는 3가지 거리척도 모두에서 시험영상이 정확하게 인식됨을 나타낸 것이다.

표 1은 1차 모멘트에 의한 중심이동의 전처리된 영상과 그렇지 않은 영상을 대상으로 기저영상의 개수에 따른 복원성능을 알아보기 위해 절대평균합오차 AMSE를 나타낸 것이다. 여기서 2가지 방법 모두는 기저영상의 수가 증가함에 따라 AMSE로 표시되는 복원성능은 개선되다 다시 저하됨을 알 수 있다. 이는 기저영상인 특징벡터의 개수가 증가함에 따라 복원성능도 개선되지만 지나치게 증가하면 오히려 과학습을 일으켜 복원성능이 저하되는 것으로 추측된다. 한편 전처리된 영상은 그렇지 않은 영상에 비해 AMSE가 큼을 알 수 있다. 이는 전처리에 의해 유효얼굴의 일부가 배경과 함께 제거됨으로써 복원에 필요한 정보가 손실되기 때문이다. 전처리나 비전처리 모두에서 기저영상의 수가 14개 일 때 가장 우수한 복원성능이 있음을 알 수 있어 본 논문에서는 이를 기준으로 실험하였다.

그림 8은 전처리된 영상을 대상으로 PCA에서 가장 낮은 AMSE를 가진 14개의 기저영상으로부터 복원된 15개의 학습얼굴들을 각각 나타낸 것이다. 기저영상은



그림 7. 거리척도에 따라 식별된 결과영상

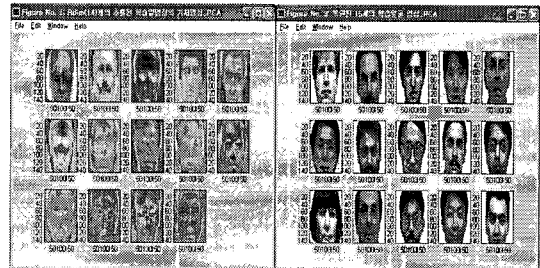


그림 8. 14개의 기저영상과 15개의 복원된 학습 얼굴영상

학습 얼굴영상으로부터 추출된 특징영상으로 차원감소에 따른 데이터의 압축효과를 갖는다. 육안으로 볼 때, 기저영상으로부터 복원된 학습 얼굴영상은 전처리된 원 얼굴영상과 매우 일치함을 알 수 있다. 이때 15개의 복원영상 모두에 대한 오차는 표 1에서 나타난 것처럼 약 1.8332×10^{-12} 정도의 매우 작은 값을 가진다.

표 2는 기저영상을 14개로 하였을 때, 학습 얼굴영상 15개와 시험 얼굴영상 45개를 대상으로 거리척도에 따른 인식결과를 각각 나타낸 것이다. 기존의 방법은 중심이동의 전처리되지 않은 얼굴영상을 대상으로 PCA를 행한 것이고, 제안된 방법은 중심이동의 전처리된 얼굴영상을 대상으로 PCA를 행한 것이다. 여기서 학습영상과 시험영상 사이의 유사성을 측정하기 위한 척도로 city-block, Euclidean, negative angle의 3가지 거리척도를 이용하였다. 또한 오차합은 올바르게 인식된 영상을 대상으로 각 기저영상을 대상으로 거리척도에 의한 식별오차의 절대합이다. negative angle의 거리척도에서는 학습과 시험의 기

표 1. 기저영상 수의 변화에 따른 복원성능

기저영상 수		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
절대 평균	전처리	34.4	30.3	26.3	23.9	20.6	17.9	15.7	13.7	11.9	10.0	8.2	6.0	3.5	4.7972	7.8
	비전처리	669	704	257	364	327	434	805	059	225	564	426	039	814	e-013	548
합오차, AMSE	전처리	31.4	24.7	22.8	20.3	18.5	16.3	13.7	12.0	9.4	8.0	6.2	4.2	2.6	1.8332	8.4
	비전처리	933	839	722	563	235	711	413	371	489	330	562	301	225	e-012	693

표 2. 기존방법과 제안된 방법에 의한 얼굴인식 결과

거리척도	입력영상	기존 방법(PCA)			제안된 방법(중심이동+PCA)		
		인식률(%)	오차합	오인식영상	인식률(%)	오차합	오인식영상
City-Block (L1-norm)	학습얼굴	100	8.13e-09	-	100	1.08e-09	-
	시험얼굴	55.56	2.36e+05	(위치변화15개, 표정변화4개, 안경유무1개)	88.89	3.01e+05	(위치변화3개, 표정변화1개, 안경유무1개)
Euclidean (L2-norm)	학습얼굴	100	3.47e-09	-	100	4.66e-010	-
	시험얼굴	55.56	7.97e+04	(위치변화15개, 표정변화4개, 안경유무1개)	88.89	1.00e+05	(위치변화3개, 표정변화1개, 안경유무1개)
Negative Angle (Cosine)	학습얼굴	100	-15.00	-	100	-15.00	-
	시험얼굴	55.56	-24.14	(위치변화15개, 표정변화4개, 안경유무1개)	88.89	-38.68	(위치변화3개, 표정변화1개, 안경유무1개)

저영상이 완전히 일치되는 경우 -1.0의 값을 갖는다. 표에서 보면 2가지 방법에서 이미 학습된 얼굴영상에 대한 인식에서는 3가지 척도 모두에서 100%의 인식률을 보이나, 시험 얼굴영상에 대해서는 기존방법과 제안된 방법은 약 55.56%와 약 88.89%의 인식률을 각각 나타낸다. 따라서 중심이동의 제안된 방법은 기존방법에 비해 평균 약 1.6배 정도 개선된 인식률을 가진다. 15개의 학습영상에 대하여 제안된 방법은 기존방법에 비해 city-block과 Euclidean 거리에서의 오차합은 평균적으로 약 7.53배와 약 7.45배 정도의 작은 값을 가지며, negative angle에서는 동일한 값을 가진다. 또한 45개의 시험영상에 대하여 제안된 방법은 기존방법에 비해 city-block, Euclidean, negative angle의 3가지 거리척도 각각에 대한 오차합에서 평균적으로 약 1.23배, 약 1.59배, 약 1.01배 정도의 작은 값을 가진다. 이는 중심이동의 제안된 방법은 그렇지 않은 기존방법에 비해 인식을 위한 식별과정에서 좀 더 정확하게 얼굴을 분류함을 보여준 것이다. 결국 인식률과 오차합의 개선에서 보면 제안된 1차 모멘트의 중심이동에 따른 전처리 과정의 우수성을 재확인할 수 있다. 한편 45개의 시험얼굴에서 기존방법은 city-block, Euclidean, negative angle의 3가지 거리척도 모두에서 각각 20개 얼굴이 오인식되었으며, 여기서 위치변화에 따른 시험얼굴 15개, 표정변화에 따른 시험얼굴 4개, 안경착용 유무에 따른 시험얼굴 1개가 각각 오인식되었다. 특히 기존방법에서는 학습된 영상에 비해 좌우로 위치가 이동된 시험영상 모두는 각기 다른 학습영상으로 오인

식되었다. 이는 표정변화나 안경착용 여부에 비해 위치변화가 기존방법의 얼굴인식에서 더욱 큰 영향을 미침을 알 수 있다. 하지만 제안된 방법에서는 3가지 거리척도에서 각각 5개 얼굴이 오인식되었으며, 위치변화에 따라 3개, 표정변화와 안경착용 여부에 따라 각 1개의 시험얼굴이 오인식되었다. 여기서도 인식률은 다른 변화에 비해 위치변화가 가장 크게 좌우됨을 알 수 있다. 특히 제안된 방법에서 오인식 원인을 살펴보면, 5개 시험얼굴 모두에는 조명에 따른 배경잡음이 첨가되었고, 이는 1차 모멘트의 논리필터링에서 유효화소 값의 판별에 영향을 미쳐 중심좌표가 정확히 구해지지 않았기 때문이다. 한편 이들 오인식 시험얼굴을 대상으로 유효영상의 판별값을 바꾸어 실험한 결과 모두 올바르게 인식되었으며, 단순히 위치만 좌우로 이동된 배경잡음이 없는 얼굴영상들도 모두 올바르게 인식되었다. 이는 제안된 전처리가 위치나 표정 등의 변화뿐만 아니라 조명 등에 의하여 발생하는 잡음에도 강건한 성능이 있음을 보여준 것이다. 실험의 결과에서 학습영상과 시험영상의 유사성 판별을 위한 거리척도 중에서 city-block이 다른 2가지 거리척도에 비해 학습 및 시험 얼굴영상에 대한 인식률이 높음도 알 수 있다. 따라서 중심이동의 전처리 과정을 수행하면 인식에 불필요한 정보를 제거할 수 있어 인식성능이 개선됨을 확인할 수 있다.

4. 결 론

본 논문에서는 영상의 1차 모멘트와 고유벡터를

이용한 얼굴인식 기법을 제안하였다. 여기서 1차 모멘트는 입력되는 얼굴영상의 중심좌표를 계산하여 중심이동하는 전처리로 인식에 불필요한 배경을 배제시킴으로써 인식성능을 개선하기 위함이다. 또한 고유벡터는 얼굴의 특징인 기저영상으로 주요성분 분석을 이용하여 추출하였으며, 이는 중복성분의 제거함으로써 인식성능을 개선하기 위함이다.

제안된 방법을 각각 320*243 픽셀의 60개(15명*4장) 얼굴영상에 적용하여 city-block, Euclidean, 그리고 negative angle의 3가지 거리 척도를 분류척도로 이용하여 실험하였다. 실험결과, 중심이동의 제안된 방법은 전처리과정을 거치지 않는 기존방법보다 학습영상에서 인식률은 동일하나 평균적으로 약 7.5배 정도 정확한 분류성능이 있으며, 시험영상에서는 평균적으로 약 1.6배 정도의 인식률과 약 5.9배 정도의 정확한 분류성능이 있음을 확인하였다. 또한 city-block이 Euclidean이나 negative angle의 거리 척도보다 작은 오차함을 가져 상대적으로 정확한 분류성능이 있음도 알 수 있었다.

향후 제안된 방법을 좀 더 큰 규모의 인식문제와 다양한 분야에 적용하는 연구와 각도 등의 변화에도 견실한 인식성능을 가지는 전처리 기법에 대한 연구가 지속적으로 이루어져야 할 것이다.

참 고 문 헌

- [1] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," *Proceedings of IEEE*, Vol. 83, No. 5, pp. 705-741, May 1995.
- [2] M. H. Yang, D. Kriegman, and N. Ahuja, "Detecting Faces in Images: A Survey," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 1, pp. 64-58, Jan. 2002.
- [3] S. H. Jeng, H. Y. M. Liao, C. C. Han, M. Y. Chern, and Y. T. Liu, "Facial Feature Detection Using Geometrical Face Model: An Efficient Approach," *Pattern Recognition*, Vol. 31, No. 3, pp. 273-282, 1998.
- [4] R. Brunelli and T. Poggio, "Face Recognition : Feature Versus Templates," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 10, pp. 1042-1052, Jan. 1993.
- [5] 이형찬, "기하학적 영상처리를 이용한 얼굴인식 시스템," 멀티미디어학회논문지, 제 6권, 제 7호, pp. 1139-1149, 2003.
- [6] 지형근, 이경희, 정용화, "SVM을 이용한 얼굴 검출 성능향상 방법," 정보처리학회논문지, 제 11-B권 제1호, pp. 13-20, 2004.
- [7] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86, 1991
- [8] B. A. Draper, K. Baek, M. S. Barlett, and J. R. Beveridge, "Recognizing Faces with PCA and ICA," *Computer Vision and Imaging Understanding 91*, pp. 115-137, 2003.
- [9] N. Kambhatla and T. K. Leen, *Dimension Reduction by Local PCA*, Neural Computation 9, pp. 1493-1516, 1997.
- [10] K. I. Diamantaras and S. Y. Kung, *Principal Component Neural Networks : Theory and Applications, Adaptive and Learning Systems for Signal Processing, Communication, and Control*, John Wiley & Sons. Inc., New York, 1996.
- [11] A. Hyvaerinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, Inc., New York, 2001.
- [12] 김상철역, 재료역학, 청문출판사, 서울, 1992.
- [13] "Yale Face Databases," <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>.



조 용 현

1979년 경북대학교 공과대학 전자공학과(공학사)
 1981년 경북대학교 대학원 전자공학과(공학석사)
 1993년 경북대학교 대학원 전자공학과(공학박사)
 1983년~1984년 삼성전자(주) 사원
 1984년~191984년~1987년 한국전자통신연구소 연구원
 1987년~1997년 영남이공대학 전자과 교수
 1997년~현재 대구가톨릭대학교 컴퓨터정보통신공학부 교수
 관심분야 : 신경망, 병렬분산처리, 신호처리, 영상인식, 교환기 등