

논문 2011-48SP-5-10

상관관계에 기반한 가려진 얼굴 영상 검출 및 복원

(Detection and Recovery of Occluded Face Images Based on Correlation)

이지은*, 곽노준**

(Ji-eun Lee and Nojun Kwak)

요약

본 논문에서는 화소들 간의 상관관계를 이용하여 가려진 얼굴 영상을 검출하고 복원하는 방법을 제안한다. 본 논문의 학습 단계에서는 기존에 이용된 주성분 분석법(PCA)의 변환 행렬 대신 상관계수를 계산하고, 테스트 단계에서는 학습 단계에서 구한 상관계수를 이용하여 가려진 얼굴 영역 검출 과정과 복원 과정을 수행한다. 검출된 영상과 복원된 영상은 실험을 통해 기존 방법과 비교한다. 실험 결과, 상관관계 방법에 의해 검출된 영상은 기존 주성분 분석법을 이용한 방법보다 가려진 얼굴 영역 및 주변 영역의 잡음이 적음을 확인하였다. 또한 복원된 얼굴 영상에서는 영상의 뭉개지는 현상이 줄어들었으며, 복원된 얼굴 영상의 가려진 부분과 가려지지 않은 부분과의 경계가 보다 매끄럽게 연결되는 것을 확인하였다.

Abstract

In this paper, we propose a method to detect and recover the occluded parts of face images using the correlation between pairs of pixels. In a training stage, correlation coefficients between every pairs of pixels are calculated using the occlusion-free face images. Once a new occluded face image is shown, the occluded area is detected and recovered using the correlation coefficients obtained in the training stage. We compare the performance of the proposed method with the conventional method based on PCA. The results show that the proposed method detects and recovers occluded area with much smaller noises than the conventional PCA based method. Moreover, recovered images by the proposed method were more smooth with reduced blurring effect.

Keywords: 상관관계, 가려진 얼굴 영상

I. 서론

전 세계적으로 감시카메라는 산업분야뿐만 아니라 사람이 많이 모이는 장소인 시가지나 편의점, 은행 등 곳곳에 설치되어 있다. 이렇게 우리 주변에 감시카메라가 많이 존재하고 있는 이유는 감시카메라가 적은 인원으로 넓은 장소 또는 여러 장소를 한 눈에 감시할 수 있고, 녹화 저장된 영상을 통해 문제 발생 시의 상황을

다시 볼 수 있어서 매우 편리하기 때문이다.

그러나 감시카메라가 항상 모든 상황에서 유용한 것은 아니다. 만일 범죄자들이 선글라스나 마스크 등으로 얼굴을 가린 채 감시카메라 영상에 잡힌다면, 실제 문제가 발생했을 시에도 그 범죄자의 신원을 확인할 수 있는 중요한 특징들을 잡아낼 수가 없을 가능성이 크다. 이는 곧, 최근 응용 범위가 확대되고 있는 얼굴 인식을 이용하여 범죄자의 신원을 알아내기가 어렵다는 뜻이다. 그러므로 더욱 정확히 신원을 확인하기 위해 얼굴 인식 단계 이전에 영상의 가려진 부분을 검출하고 검출된 영역을 복원하는 것이 영상 처리 분야에서 시급히 해결해야 할 중요한 과제이다.

얼굴 영상에서의 가려진 부분을 검출하고 복원하기 위한 대표적인 방법으로는 주성분 분석법(principal

* 학생회원, ** 정회원, 아주대학교 전자공학과
(Division of Electrical and Computer Engineering, Ajou University)

※ 이 연구는 한국연구재단 지원(KPF-2011-0005324)과 아주대학교 교내연구비 지원(20094720)에 의해 이루어졌음.

접수일자: 2010년12월17일, 수정완료일: 2011년6월7일

component analysis, PCA)을 응용하는 방법이 있다. 특히 주성분 분석법을 이용하여 얼굴 영상에서 안경을 제거하는 방법^[1~2], 또는 얼굴에 존재하는 잡음을 제거하는 방법^[3] 등이 제시되어 널리 이용되고 있다. 또한 기존 주성분 분석법을 발전시켜 영상의 가려진 부분을 더 빠르게 복원하는 Fast recursive PCA^[4], 확률론적 주성분 분석법 (probabilistic PCA)^[5]을 이용하는 Tipping과 Bishop의 연구 등 영상의 가려진 부분에 대한 연구가 다양한 방법^[6~7]으로 진행되고 있다.

이 중 본 논문에서 제안하는 방법과 주로 비교할 주성분 분석법을 이용하는 기본적인 가려진 부분 검출 및 복원 방법은 다음과 같다. 먼저 얼굴의 특정 부분이 가려진 영상을 가려지지 않은 원본 학습 영상들로부터 얻은 고유값(eigenvalue)들로 이루어진 주성분 분석법 계수를 이용하여 재구성(reconstruction)한 다음, 재구성된 영상과 가려진 영상과의 화소값의 차이가 일정 경계값을 넘으면 해당 화소들이 가려진 화소들이라 간주한다. 이러한 방식으로 영상에서 가려졌다고 판단되는 화소들을 선택한 후, 이들 화소들로 이루어진 영역을 재구성된 영상의 화소들로 채운다. 이렇게 재구성된 영상을 앞의 주성분 분석법 계수를 이용하여 또 다시 재구성한다. 최종적으로 복원된 영상은 위의 주성분 분석법을 이용하여 가려진 부분 검출과 복원하는 과정을 주어진 조건이 수렴할 때까지 계속 반복하여 얻어낼 수 있다.^[1~2]

이와 달리, 본 논문에서는 상관계수 값들을 이용하여 영상의 가려진 부분을 검출하고, 검출된 영역을 복원하는 과정을 제시한다. 제시된 방법에서는 우선 가려지지 않은 학습 영상들을 이용하여 각 화소들 사이의 상관계수를 구한 후, 얼굴의 특정 부분이 가려진 영상의 화소들의 값을 앞에서 구한 상관계수를 이용하여 예측한다. 이를 위해서는 먼저 얼굴에서 가려진 부분의 위치를 알아내야 한다. 만약 한 화소의 값은 모른다고 가정하고, 나머지 화소들의 값은 알고 있다고 가정하면, 우리가 모르는 화소의 값은 나머지 화소들 중 해당 화소와 상관관계가 높은 몇 개의 화소들의 값을 가중합하여 예측할 수 있다. 이 과정을 각각의 모든 화소들에 대하여 적용하면, 하나의 영상의 모든 화소들의 값을 예측된 값으로 대체할 수 있다. 이렇게 상관계수에 의해 재구성된 영상과 가려진 영상과의 화소값의 차이가 일정 경계값보다 크면 가려진 부분이라 여긴다. 이러한 방법을 통해 영상의 가려진 부분을 찾았다면, 이제는 그 부분

을 새롭게 예측된 화소값으로 채워 넣어 복원하여야 한다. 검출 부분에서 이미 모든 화소값들을 예측하여 채워 넣었지만, 상관관계가 높은 몇 개의 화소들 중 가려진 영역의 화소들이 포함되었을 수 있으므로 예측된 값의 오차가 클 수 있다. 그러므로 이미 알고 있는 영상의 가려진 영역과 가려진 영역 외의 화소의 값들을 이용할 수 있다. 가려진 영역의 하나의 화소값은 영상의 가려지지 않은 화소들 중에서 해당 화소와 상관관계가 높은 몇 개의 화소값들의 가중합에 의해 구해진다. 이를 가려진 영역의 모든 화소들에 적용하여 가려진 영역의 모든 화소값들에 예측된 값들을 채워 넣는다. 얼굴 영상의 가려진 영역의 모든 화소값들이 예측된 값들로 채워진 영상이 상관관계를 이용하여 최종적으로 복원된 영상이다.

본 논문은 영상의 가려진 부분을 검출하고 복원하는 문제에 있어서, 기존의 주성분 분석법이 아닌 상관관계를 이용한 방법을 통해 해결한다. 이 논문의 각 장의 내용은 다음과 같다. 제 II장에서는 주성분 분석법을 이용하는 기본적인 가려진 부분 검출 및 복원 방법에 대하여 살펴보고, 제 III장에서는 본 논문에서 제안한 상관관계를 이용하는 가려진 부분 검출 및 복원 방법에 대해 자세히 살펴본다. 제 IV장에서는 실험을 통해 기존 주성분 분석법과 제안한 상관관계를 이용한 방법을 비교 분석하고, 마지막으로 제 V장에서는 결론을 제시한다.

II. PCA를 이용한 가려진 얼굴 영역 검출 및 복원

1. 학습

학습 단계에서 최종적으로 구하고자 하는 것은 축소하고자 하는 차원 수만큼의 고유벡터로 구성된 변환 행렬이다.

화소의 개수가 d 개인 N 개의 학습 얼굴 영상을 $X_{tr} = [x_1, x_2, \dots, x_N] \in R^{d \times N}$ 이라 표현한다. 여기서 각각의 영상 x_n 은 학습을 위해 하나의 열벡터로 표현해 준다. 그 다음 전체 얼굴 영상들의 평균 영상을 구하고, 각 얼굴 영상들로부터 평균을 뺀다. 위의 내용을 식으로 표현하면 다음과 같다.

$$\bar{x} = \frac{1}{N} \sum_{n=1}^N x_n \quad (1)$$

$$\Phi_n = x_n - \bar{x}, \quad n = 1, 2, \dots, N \quad (2)$$

각 얼굴 영상들로부터 평균을 뺀 영상들의 공분산 행렬(covariance matrix)을 구한다.

$$C = \frac{1}{N} \sum_{n=1}^N \Phi_n \Phi_n^T \quad (3)$$

공분산 행렬에 대한 고유분석을 행하여 d' 개의 가장 큰 고유값 $\lambda_1, \dots, \lambda_{d'}$ 들을 선택한다. 선택된 고유값과 관련된 고유벡터를 구하고 변환 행렬 $W \in R^{d' \times d}$ 를 다음과 같이 만든다.

$$W = [w_1, w_2, \dots, w_{d'}] \quad (4)$$

여기에서 w_n 은 n 번째 고유값 λ_n 에 해당하는 고유벡터로 이러한 고유벡터는 원래의 차원 수인 d 개만큼 얻을 수 있는데, 변환 행렬 W 를 구할 때는 이들 중 축소하고자 하는 차원 수인 d' 차원만큼만 취하도록 한다.

2. 가려진 얼굴 영상의 검출 및 복원

이 단계에서는 테스트 얼굴 영상 $X_i = [x_1, x_2, \dots, x_L]$ 가 입력되면 이를 재구성하여 얼굴 영상의 가려진 부분을 검출하고, 검출된 부분은 재구성한 영상을 이용하여 복원한다. 본 논문에서 테스트 얼굴 영상들은 임의의 너비와 높이를 가진 사각형으로 얼굴의 일부분을 가린 영상을 의미한다.

각 테스트 얼굴 영상들은 앞의 학습 단계에서 구한 평균 영상으로 정규화시킨 다음, 학습 단계에서 마지막으로 도출해 낸 고유벡터에 대하여 가중치를 구하고, 이를 이용하여 각 테스트 얼굴 영상들을 고유벡터의 가중합의 형태로 재구성한다.

$$u_i = w_i^T(x - \bar{x}), \quad i = 1, 2, \dots, d' \quad (5)$$

$$x' = \bar{x} + w_1 u_1 + \dots + w_{d'} u_{d'} = \bar{x} + \sum_{i=1}^{d'} w_i u_i \quad (6)$$

여기서 x 는 테스트 얼굴 영상이고, x' 는 변환 행렬에 의해 재구성된 새로운 얼굴 영상이다. 이렇게 재구성된 얼굴 영상 x' 와 테스트 얼굴 영상 x 의 각 화소 사이의 차이값이 일정 경계값 ϵ 을 넘느냐 안 넘느냐에 따라 가려진 영역인지 아닌지를 판별하게 된다.

이러한 방법으로 얼굴 영상에서 가려진 영역을 판별

한 후, x' 이라는 새로운 영상에 재구성된 얼굴 영상과 테스트 얼굴 영상으로 채운다. 즉, 가려진 부분이라 판단된 화소값에는 앞에서 재구성된 얼굴 영상의 화소 x'_i 값으로 채우고, 가려지지 않은 부분이라 판단된 화소값에는 원래 입력으로 들어온 테스트 얼굴 영상의 화소 x_i 값으로 채운다.

가려진 부분 검출과 복원 과정을 간단한 수식으로 나타내면 각각 식 (7) 및 (8)과 같다.

$$\begin{aligned} |x'_i - x_i| \leq \epsilon &\Rightarrow i \in O^c \\ |x'_i - x_i| > \epsilon &\Rightarrow i \in O \end{aligned} \quad (7)$$

$$x_i'' = \begin{cases} x_i & \text{if } i \in O^c \\ x'_i & \text{if } i \in O \end{cases} \quad (8)$$

여기서 O^c 는 가려지지 않은 영역을 뜻하고, O 는 가려진 영역을 뜻한다. 이렇게 새롭게 구성된 영상 x'' 를 x 로 여기고 위의 과정을 $x = x''$ 이 될 때까지 계속 반복한다. 모든 반복 과정을 거치고 기준치가 수렴했을 때 생성된 영상을 최종적으로 복원된 영상이라 한다.

III. 상관관계를 이용한 가려진 얼굴 영역 검출 및 복원

1. 상관계수

각 화소를 하나의 확률변수라 가정하자. 상관계수는 두 확률변수 사이에 얼마나 연관성이 존재하는지를 나타내는 지표이다. 상관계수는 공분산을 표준편차로 나눠서 표준화 시킨 것으로, -1에서 1까지의 값을 가진다. 상관계수는 1 또는 -1값에 가까울수록 두 확률변수 간의 관련성이 크다는 뜻이고, 0에 가까울수록 관련성이 없다는 뜻이다.

$$\rho_{ij} = \frac{C_{ij}}{\sigma_i \sigma_j} = \frac{E[(x_i - \mu_i)(x_j - \mu_j)]}{\sqrt{E[(x_i - \mu_i)^2]} \sqrt{E[(x_j - \mu_j)^2]}} \quad (9)$$

여기서 x_i 는 i 번째 확률변수이고, x_j 는 j 번째 확률변수를 의미한다. C_{ij} 는 두 확률변수의 공분산을 나타내고, σ_i, σ_j 는 각각의 확률변수의 표준편차를 나타내며 $\mu_i = E[x_i], \mu_j = E[x_j]$ 를 의미한다.

영상에서는 각각의 화소를 하나의 확률변수로 생각할 수 있으므로 한 장의 영상은 하나의 d 차원 확률벡터의 실현으로 생각할 수 있고, 각 화소들 사이의 연관도

는 위 식 (9)에서 정의한 상관계수를 통해 유추할 수 있다. 가려진 영상의 경우 가려진 부분의 화소값을 해당 화소와 연관도가 높은 가려지지 않은 부분의 화소값들로부터 예측할 수 있다는 것이 본 논문의 기본 아이디어이다.

N 장의 학습 얼굴 영상이 주어졌다고 가정하면, 실제 상황에서 식 (9)의 두 확률변수의 기댓값 $E(\cdot)$ 을 알 수 없으므로 N 개의 샘플을 이용한 상관계수의 예측치를 다음과 같이 구할 수 있다.

$$\hat{\rho}_{ij} = \frac{A}{B} \tag{10}$$

$$A = \sum_{n=1}^N (x_{in} - \bar{x}_i)(x_{jn} - \bar{x}_j)$$

$$B = \sqrt{\sum_{n=1}^N (x_{in} - \bar{x}_i)^2} \sqrt{\sum_{n=1}^N (x_{jn} - \bar{x}_j)^2}$$

여기서 x_{in} 은 n 번째 학습 영상에서의 i 번째 화소값이고, \bar{x}_i , \bar{x}_j 는 i 번째, j 번째 화소값의 샘플 평균을 나타낸다.

실제 얼굴 영상에서 얼굴의 특정 4부분, 즉 눈과 코, 입, 볼 부분에 대하여 각각의 상관계수 도를 그림 1을 통해 살펴보겠다.

그림 1에서는 X로 표시된 화소의 상관계수를 나타내고 있다. 그림 1의 상관계수도에서 하얀색으로 표시된 부분은 상관계수의 절대값이 높은 부분이고, 검정색으로 표시된 부분은 상관계수의 절대값이 낮은 부분이다. 여기서 눈에 대한 상관계수도 살펴보면, 왼쪽 눈동자의 한 화소와 연관도가 높은 화소들, 즉 왼쪽 눈동자 주

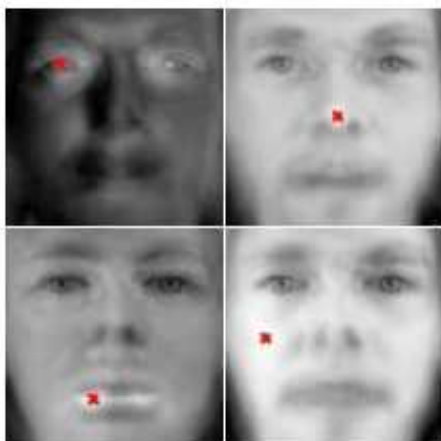


그림 1. 상관계수도 (왼쪽부터 눈, 코, 입, 볼)
Fig. 1. Correlation coefficient map.
(eye, nose, mouth, and cheek from left)

변과 오른쪽 눈동자 주변이 밝게 표시되어 있는 것을 볼 수 있다. 이렇게 상관계수를 이용하면, 가까운 화소들의 정보뿐만 아니라 그 화소와 멀리 떨어져 있는 화소들의 정보도 이용하여 더 좋은 결과를 얻을 수 있을 것이라 예상된다.

2. 결합적(Jointly) 가우시안 분포

얼굴 영상의 x_i 의 화소값을 모른다고 가정하고, x_j 의 화소값은 알고 있다고 가정하자. x_i 의 화소값은 해당 화소와 연관도가 높은 여러 화소들의 집합 $P_i = \{j | x_i \text{와 상관관계가 높은 } m \text{개의 } x_j \text{의 index}\}$ 를 이용해서 구할 수 있는데, 이 때 우리는 화소 i 와 j 가 결합적 가우시안 분포를 따른다고 가정한다.

$|P_i| = 1$ 일 때, x_j 번째 화소값이 v 라고 가정하고 x_i 의 조건부 확률을 식으로 표현하면 다음과 같다.

$$P(x_i | x_j = v) = \frac{P(x_i, x_j = v)}{P(x_j = v)} \tag{11}$$

$P(x_j = v)$ 는 상수이기 때문에 $P(x_i | x_j = v)$ 는 $P(x_i, x_j = v)$ 에 비례한다고 볼 수 있다. 그러므로 화소 i 와 j 가 결합적 가우시안 분포를 따른다는 가정 하에 $P(x_i, x_j = v)$ 에 대하여 정리하면,

$$P(x_i, x_j = v) \propto \exp\left(-\frac{1}{2} \begin{pmatrix} x_i \\ v \end{pmatrix}^T \Sigma^{-1} \begin{pmatrix} x_i \\ v \end{pmatrix}\right)$$

$$= \exp\left(-\frac{1}{2} \frac{1}{(1 - \rho_{ij}^2)\sigma_i^2\sigma_j^2} \left(x_i - \frac{\sigma_i}{\sigma_j} \rho_{ij} v\right)^2\right) \tag{12}$$

여기서 Σ^{-1} 는 x_i 와 x_j 의 공분산 행렬 Σ 의 역행렬로 다음과 같이 정리될 수 있다.

$$\Sigma^{-1} = \frac{1}{(1 - \rho_{ij}^2)\sigma_i^2\sigma_j^2} \begin{bmatrix} \sigma_i^2 & -\rho_{ij}\sigma_i\sigma_j \\ -\rho_{ij}\sigma_i\sigma_j & \sigma_j^2 \end{bmatrix} \tag{13}$$

식 (12)를 통해 x_i 의 조건부 기댓값과 분산값을 아래와 같이 정리할 수 있다.

$$E[x_i | x_j = v] = \frac{\sigma_i}{\sigma_j} \rho_{ij} v \tag{14}$$

$$Var[x_i | x_j = v] = (1 - \rho_{ij}^2)\sigma_i^2 \tag{15}$$

$|P_i| = 2$ 인 경우에 x_j 값이 v , x_k 값이 w 라고 가정할 때, 조건부 기댓값 $E[x_i|x_j = v, x_k = w]$ 을 정확한 식으로 표현하기 어려우므로 $E[x_i|x_j = v]$ 와 $E[x_i|x_k = w]$ 를 가중합하여 다음과 같이 그 예측치를 구해 보도록 한다.

$$E[x_i|x_j = v, x_k = w] \simeq \alpha E[x_i|x_j = v] + \beta E[x_i|x_k = w] \quad (16)$$

여기에서 가중치 α 와 β 가 $\alpha + \beta = 1$ 이라는 조건을 만족하도록 해야 한다. 가중치 α 는 x_j 값에 따른 x_i 값의 추정치의 신뢰도라 말할 수 있으므로 식 (15)에 나타난 x_i 의 조건부 분산이 작으면 작을수록 신뢰도가 높으며 따라서 가중치도 크게 되어야 한다. 따라서 본 논문에서는 가중치를 1/표준편차에 비례하는 값으로 정한다. 즉 $|P_i| = 2$ 인 경우에 (16)의 식에서 $\alpha \propto 1/\sqrt{1-\rho_{ij}}$, $\beta \propto 1/\sqrt{1-\rho_{ik}}$ 로 놓고 $\alpha + \beta = 1$ 를 만족하도록 α, β 값을 정한다.

$|P_i| > 2$ 의 경우에도 이를 확장하여 각각의 화소에 가중치를 두도록 한다. x_i 를 추정하는 문제에서 만일 x_i 와 x_j 의 상관관계가 높아 신뢰도가 높으면 가중치를 크게 주고, 신뢰도가 낮으면 가중치를 작게 준다. 즉, 하나의 화소값을 추정할 때 그 화소와 상관도가 높은 10개의 화소들을 이용하여 원하는 화소값을 추정한다고 할 때, 10개의 화소들 중 상관도가 최댓값인 화소값을 가장 많이 이용하고 상관도가 최솟값인 화소값을 가장 적게 이용하여 해당 화소값을 결정한다.

3. 가려진 얼굴 영상의 검출 및 복원

본 논문에서는 위에서 설명된 화소 간의 상관관계를 이용하여 가려진 얼굴 영상에서 가려진 부분을 검출하고 이를 복원하는 방법을 제시한다.

이를 위해 먼저 가려지지 않은 학습 얼굴 영상들 X_{tr} 의 모든 화소에 대하여 상관계수를 구한다. 위의 식 (10)을 보면 알 수 있듯이, 먼저 학습 얼굴 영상들의 평균 영상과 표준편차 영상을 구하고, 각 얼굴 영상들에서 평균 영상과의 차영상들을 구한다. 이렇게 구해진 차영상들의 각 화소에 대해 상관계수를 구하고, 상관계수 값들 중 큰 값들의 일부만을 저장한다.

각 테스트 얼굴 영상들의 앞에서 구한 평균 영상으로 정규화시킨 다음, 정규화된 각 테스트 얼굴 영상들의 첫 번째 화소값을 모른다고 가정하자. 첫 번째 화소에 대해 해당 화소와 연관도가 높은 여러 화소들의 화소값

을 이용하여 해당 화소값을 예측할 수 있다. 이 때, 해당 화소값은 상관계수 값이 일정값을 넘는 여러 화소값들의 가중합에 의해 구한다. $|P_i| = 2$ 일 때에 해당 화소값 x_i 를 구하는 방법을 (14) 내지 (16)에 나타내었다. 이를 보다 일반화하여 $|P_i| = m$ 인 경우로 확장하여 식으로 표현하면 다음과 같다.

$$w_j = \frac{1}{\sqrt{1-\rho_{ij}^2}} \quad (17)$$

$$\begin{aligned} z_i &= w_1 \frac{\sigma_i}{\sigma_1} \rho_{i1} x_1 + \dots + w_m \frac{\sigma_i}{\sigma_m} \rho_{im} x_m \\ &= \sum_{j=1}^m w_j \frac{\sigma_i}{\sigma_j} \rho_{ij} x_j \end{aligned} \quad (18)$$

$$x_i' = \bar{x} + \frac{z_i}{\sum_{j=1}^m w_j} \quad (19)$$

여기서 w_j 는 i 번 째 화소와 j 번 째 화소 사이의 상관관계에 대한 가중치, 즉 앞 절에서 언급한 신뢰도를 의미한다. $x_j (j = 1, \dots, m)$ 는 i 번 째 화소와 연관도가 높은 화소들로, 그 화소들의 값은 우리가 알고 있다고 여긴다.

즉, 간단히 그림으로 나타내어 보면 아래 그림 2와 같다. 그림 2를 살펴보면, □로 나타낸 것은 우리가 화소값을 알지 못해 값을 추정해야 하는 i 번 째 화소를

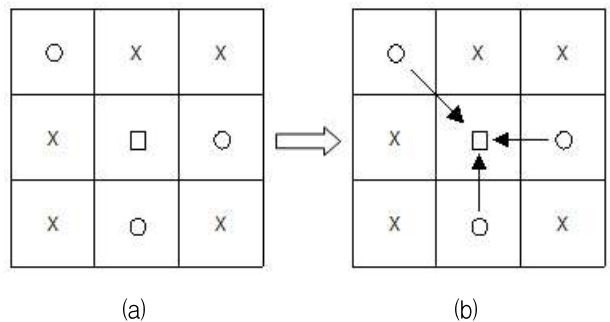


그림 2. 얼굴 영상 부분을 간단하게 나타낸 그림(a)과 연관도가 높은 화소들이 x_i 에 영향을 주는 그림(b)

Fig. 2. The image simply represented the part of a face image (a) and the image x_i is influenced by the pixels which have with correlation with x_i (b).

나타낸 것이고, \bigcirc 는 i 번 째 화소와 연관도가 높은 화소들을 그림으로 도시화한 것이다. 결론적으로 \square 의 화소값은 연관도가 높은 각각의 \bigcirc 의 화소값들의 가중합에 의해 결정된다.

위에서 첫 번째 화소값을 모른다고 가정했다. 이제 정규화된 각 테스트 얼굴 영상들의 각각의 화소값을 모른다고 가정해 보자. 테스트 얼굴 영상들의 각각의 화소값은 연관도가 높은 화소값들의 가중합에 의해 구해진 것으로 간주한다. 이 때, 테스트 얼굴 영상 데이터들의 각각의 실제 화소값 x_i 와 상관계수를 이용하여 예측된 화소값 x_i' 와의 차이값이 일정 경계값 ϵ 보다 크면 그 부분은 가려진 부분이라 여긴다. 이를 도식화하면 다음과 같다.

$$\begin{aligned} |x_i' - x_i| < \epsilon &\Rightarrow i \in O^c \\ &\geq \epsilon \Rightarrow i \in O \end{aligned} \quad (20)$$

위의 식 (20)에 의해 각 테스트 얼굴 영상의 특정 화소가 가려진 부분인지 아닌지를 판단한다. 이로써 얼굴 영상의 가려진 부분을 검출할 수 있다.

일단 (20)의 식에 의해 전체 영상 중 가려졌다고 판단되는 부분을 검출한 후에는 가려졌다고 판단되는 화소들을 복원하는 작업이 필요하다. 이를 위해서 각각의 가려졌다고 판단되는 화소들을 제외한 나머지 부분, 즉 우리가 화소값을 알고 있다고 가정한 부분만을 이용하여 가려진 화소와 연관도가 높은 화소들을 찾는다. 그 후 가려졌다고 판단되는 화소는 식 (17)과, (18), (19)를 이용하여 예측값을 구한다. 예측된 화소값들은 식 (21)의 아래 조건을 만족할 경우에 한해서 기존 화소값들과 대체된다.

$$x_i'' = \begin{cases} x_i & \text{if } i \in O^c \\ x_i' & \text{if } i \in O \end{cases} \quad (21)$$

만약 우리가 화소값을 알고 있다고 가정한 부분이 우리가 값을 예측하고자 하는 화소의 주변이고 그 부분도 가려진 영역이라면, 가려진 화소는 식 (20)의 조건을 만족하지 못하므로 다시 가려진 영역이라 판단된다. 반대로 우리가 화소값을 알고 있다고 가정한 부분이 가려지지 않은 영역에 속한다면, 예측된 화소값은 식 (20)의 조건을 만족하여 우리가 모른다고 가정한 화소값은 예측된 값으로 채워진다. 그러므로 우리가 알고 있다고 가정한 화소들이 가려진 영역에 속한 경우에는 먼저 앞에서 설명한 후자의 경우에 속하는 화소들부터 예측된

화소값으로 채워진 다음, 예측된 값들에 의해 유추된다. 결국 최종적으로 복원된 얼굴 영상은 영상의 가려진 영역의 각 화소값들이 x_i' 로 대체될 때까지 위의 과정을 반복하여 마지막으로 얻어진 영상을 말한다.

IV. 실험

1. BioID 데이터에 대한 실험

실험에 사용한 얼굴 영상은 BioID 데이터로 총 1521장의 그레이 영상이다. 각 영상은 23명의 정면 얼굴로 구성되어 있다.^[8] 각각의 영상의 크기는 100×100 화소의 크기로 이루어져 있으며, 각 얼굴 영상은 두 눈의 중심을 기점으로 고정되어 있다. 이 데이터는 현실 조건에 맞춰 촬영된 것으로 각 영상마다 매우 큰 변화를 가지는 조명 환경 상태에서 획득되었다.

주성분 분석법을 이용한 실험의 학습 단계에서 얼굴 영상 1521장 중 1000장을 임의로 선택하여 학습 영상으로 사용하였고, 나머지 521장을 테스트 영상으로 사용하였다. 상관관계를 이용한 실험의 학습 단계에서도 주성분 분석법을 이용한 실험에서와 같이 1521장 중 임의의 1000장을 학습 영상으로 사용하였고, 나머지 521장을 테스트 영상으로 사용하였다.

실험을 위해 테스트 얼굴 영상에 임의의 너비와 높이를 가진 직사각형을 임의의 위치에 배열시켜 얼굴 영상



그림 3. BioID 데이터 얼굴 영상의 부분을 가린 영상
Fig. 3. Occluded face images for BioID data.

의 부분이 가려지게끔 만들었다. 영상의 부분이 가려진 결과는 그림 3에서 보인다.

가. 주성분 분석법을 이용한 실험 결과

학습 단계에서 10000차원인 1000장의 학습 얼굴 영상을 주성분 분석법을 이용하여 50차원으로 축소시키고, $W \in R^{50 \times 10000}$ 인 변환 행렬을 구한다. 테스트 얼굴 영상의 가려진 얼굴 부분을 식 (6)을 이용하여 재구성한다. 이렇게 재구성된 얼굴 영상과 테스트 얼굴 영상과의 차이값이 실험을 통해 정한 일정 경계치보다 작으면 가려지지 않은 부분이라 여기고, 크면 가려진 부분이라 여긴다. 그림 4는 주성분 분석법을 이용한 방법에 의해 가려진 부분이 검출된 영상들이다. 여기서 경계치를 넘는 부분은 1의 값으로, 넘지 않는 부분은 0의 값으로 표현하여 도시화하였다.

그림 4를 보면, 얼굴의 가려진 부분을 찾았지만 얼굴 영역과 그 주변 영역에 많은 잡음이 발생했다. 이제 앞의 식 (8)에서 언급했듯이, x_i'' 가 얼굴 영상의 가려진 부분의 화소면 재구성된 영상의 화소값으로 채우고, 가려지지 않은 부분의 화소면 테스트 얼굴 영상의 가려지지 않은 부분의 화소값으로 채운다. 새로 생성된 영상 x'' 를 다시 테스트 얼굴 영상으로 가정하고, 모든 화소값들이 가려지지 않은 부분이라 간주될 때까지 위의 과

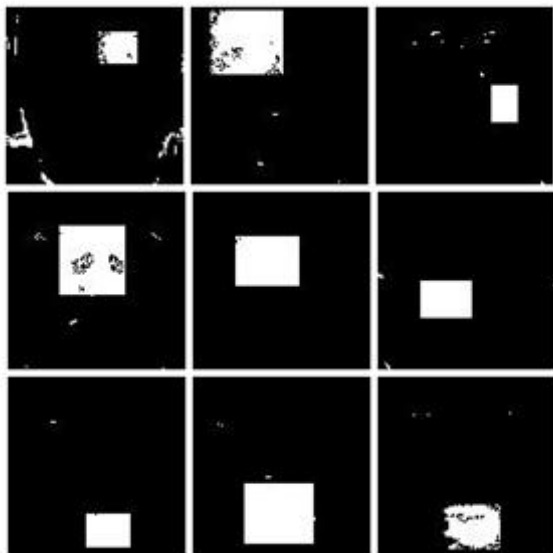


그림 4. BioID 데이터의 가려진 얼굴 영상과 재구성된 얼굴 영상과의 차영상

Fig. 4. Differential images between occluded images and reconstructed images for BioID data.



그림 5. 주성분 분석법을 이용하여 BioID 데이터의 가려진 얼굴 영상 복원

Fig. 5. Recovery of occluded face images for BioID data using the conventional PCA based method.

정을 반복한다. 이렇게 최종적으로 복원된 영상은 그림 5와 같다.

나. 상관관계를 이용한 실험 결과

학습 단계에서 BioID 데이터 1521장 중 1000장의 학습 얼굴 영상을 식 (10)을 이용하여 각 화소에 대한 상관계수를 구하고, 그 중 값들이 큰 1000개의 상관계수들만 이용한다.

테스트 단계에서는 학습 단계에서 구해둔 상관계수 값들을 결합적 가우시안 방법을 이용하여 얼굴 영상의 가려진 부분을 찾아낸다. 본 연구에서는 1000개 중 큰 값부터 20개의 상관계수 값을 뽑아내어 그 상관계수 값이 0.95이상인 경우에만 결합적 가우시안 방법을 이용하였다. 만일 한 화소값을 모르는 상태에서 상관도가 0.95이상인 20개의 화소값을 알 때, 그 중 상관도가 큰 10개의 화소값에 각각 가중치를 주어 해당 화소값을 추정하도록 조건을 두었다. 예측된 화소값과 테스트 영상의 화소값의 차가 실험을 통해 정한 일정 경계값보다 크면 가려진 영역이라 간주하고, 그 외에는 가려지지 않은 영역이라 간주한다. 그림 6은 상관관계를 이용한 방법에 의해 가려진 영역이 검출된 얼굴 영상들이다. 여기서도 위에 언급한 것과 같이 가려진 영역은 1의 값으로, 가려지지 않은 영역은 0의 값으로 정하여 도시화

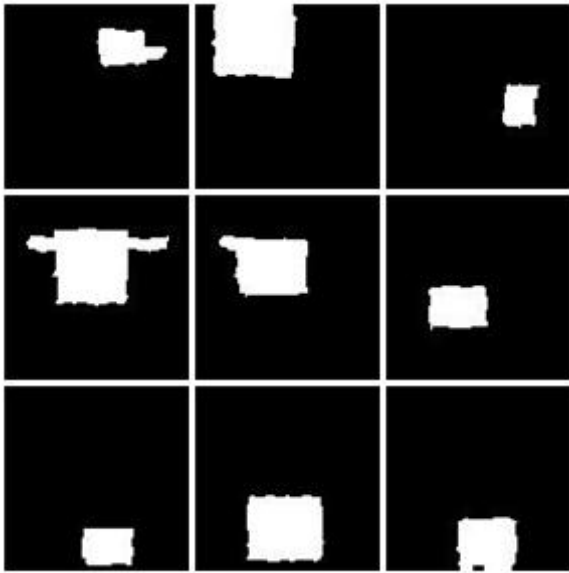


그림 6. 상관관계를 이용하여 BiID 데이터의 가려진 얼굴 영상 검출

Fig. 6. Detection of occluded face images for BiID data using a correlation based method.



그림 7 상관관계를 이용하여 BiID 데이터의 가려진 얼굴 영상 복원

Fig. 7. Recovery of occluded face images for BiID data using a correlation based method.

하였다.

그림 6을 보면, 실제 가려진 영역의 주변 영역에 작은 잡음이 발생하지만, 그림 4의 주성분 분석법을 이용한 방법보다 가려진 영역과 거리가 먼 영역에서 발생하는 잡음이 존재하지 않으며 가려진 영역 내에서의 잡음

이 발생하지 않는 것을 볼 수 있다.

그림 7의 복원된 영상은 육안으로 볼 때, 그림 5보다 가려진 영역과 가려지지 않은 영역의 연결이 매끄럽고 테스트 영상을 임의로 가리기 전의 기준 영상과 좀 더 비슷한 결과를 도출했다고 말할 수 있다.

다음 항에서는 육안의 차이를 수치적으로 알아보기 위해 절대 오차값과 최소 제곱 오차값을 구하여 비교하겠다.

다. 주성분 분석법을 이용한 실험 결과와 상관관계를 이용한 실험 결과와의 비교

상관관계를 이용하여 복원된 얼굴 영상들은 주성분 분석법을 이용하여 복원된 얼굴 영상들보다 뭉개짐 현상이 줄어들었고, 얼굴의 가려진 부분과 가려지지 않은 부분의 경계가 더 매끄럽게 연결되어 있는 것을 눈으로 확인하였다.

여기에서는 상관관계를 이용하여 복원된 얼굴 영상들과 주성분 분석법을 이용하여 복원된 얼굴 영상들이 가려지기 전의 영상들, 즉 기준 얼굴 영상들과의 오차값을 계산하여 결과를 수치적으로 확인한다.

먼저 주성분 분석법을 이용하여 최종적으로 복원된 영상을 x_i^{pca} 라 하고, 상관관계를 이용하여 최종적으로 복원된 영상을 x_i^{corr} 라고 하자. 또한, 테스트 얼굴 영상이 임의의 너비와 높이를 가지는 사각형에 의해 가려지기 전의 얼굴 영상, 기준 얼굴 영상을 x_i^{gt} 라고 하자. 이를 이용하여 절대 오차값과 최소 제곱 오차값에 대한 식을 정리한다.

$$error_{L1} = \frac{\sum_{i=1}^L |x_i^{gt} - x_i^{pca} \text{ or } x_i^{corr}|}{L} \quad (22)$$

$$error_{L2} = \frac{\sum_{i=1}^L \sqrt{(x_i^{gt} - x_i^{pca} \text{ or } x_i^{corr})^2}}{L} \quad (23)$$

표 1에 절대 오차값과 최소 제곱 오차값을 정리해 놓았다. 절대 오차값과 최소 제곱 오차값 둘 다 주성분 분석법을 이용한 방법보다 상관관계를 이용한 방법이 더 작은 것을 수치적으로 확인할 수 있다.

표 2에는 각 방법에 대한 반복 수행 횟수와 처리시간을 나타내었다. 이 표에서 상관관계를 이용한 방법이 주성분 분석법을 이용한 방법보다 반복 수행 횟수가 더

표 1. 기준 BioID 데이터와 복원된 BioID 데이터와의 절대 오차값과 최소 제곱 오차값 결과

Table 1. L1-norm error and L2-norm error between the ground truth BioID data and the reconstructed BioID data.

	절대 오차값	최소 제곱 오차값
주성분 분석법	17052.45	788.06
상관계수 방법	10500.65	532.22

표 2. BioID 데이터에 대한 반복 수행 횟수와 처리시간

Table 2. Iteration and processing time for BioID data.

		반복 수행 횟수		처리시간	
		평균	편차	평균	편차
주성분 분석법		3.1171	1.2262	0.0921	0.0509
상관계수 방법	검출	3	0	1.1728	0.4211
	복원	5.6334	3.8202		

많으며, 처리시간 또한 오래 걸리는 것을 볼 수 있다. 상관관계를 이용한 방법은 검출 과정에서 어느 부분이 가려진 영역인지 모르기 때문에 각각의 모든 화소에 대해 값을 각각 계산하여 예측한 다음, 차이값으로 가려진 영역을 찾아내기 때문이고, 복원 과정에서 가려진 영역의 각 화소에 대해 값을 각각 계산하여 예측한 다음, 가려진 영역의 모든 화소값의 차이값이 가려지지 않은 부분이라 여길 때까지 반복하기 때문이다.

즉, 주성분 분석법을 이용한 방법이 한 번에 모든 화소에 대해 계산하는 것에 반해, 상관관계를 이용한 방법은 각각의 화소별로 계산하기 때문에 알고리즘 처리시간이 오래 걸리는 것이다.

2. Yale 데이터에 대한 실험

Yale 데이터는 총 165장의 그레이 얼굴 영상으로 구성되어 있으며, 15명의 사람에 대해 각각 11장씩 구성되어 있다. 각각의 영상의 크기는 100×80 화소의 크기로 이루어져 있다. Yale 데이터는 빛의 방향과 사람의 표정 변화를 다양하게 설정하여 획득된 영상이다.

주성분 분석법을 이용한 실험과 상관관계를 이용한 실험의 학습 단계에서 총 165장 중 100장을 임의로 선택하여 학습 영상으로 사용하였고, 나머지 65장을 테스트 영상으로 사용하였다.

실험을 위해 테스트 얼굴 영상은 그림 8과 같이 입의 너비와 높이를 가진 직사각형에 의해 입의 얼굴 부분을 가려 놓았다.



그림 8. Yale 데이터 얼굴 영상의 부분을 가린 영상
Fig. 8. Occluded face images for yale data.

가. 주성분 분석법을 이용한 실험 결과

학습 단계에서 8000차원인 100장의 학습 얼굴 영상을 주성분 분석법을 이용하여 50차원으로 축소시키고, $W \in R^{50 \times 8000}$ 인 변환 행렬을 구했다. 그림 9는 BioID 데이터 실험 때와 같이 주성분 분석법을 이용한 방법에 의해 가려진 부분이 검출된 영상들이다.

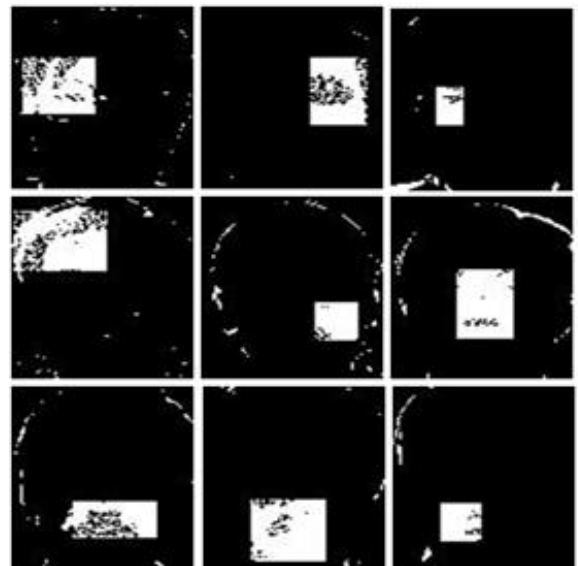


그림 9. Yale 데이터의 가려진 얼굴 영상과 재구성된 얼굴 영상과의 차 영상
Fig. 9. Differential images between occluded images and reconstructed images for yale data.



그림 10. 주성분 분석법을 이용하여 yale 데이터의 가려진 얼굴 영상 복원

Fig. 10. Recovery of occluded face images for yale data using the conventional PCA based method.

그림 9의 Yale 데이터도 마찬가지로 검출된 얼굴 영역과 주변 영역에 많은 잡음이 발생하는 것을 확인할 수 있고, 그림 10의 최종 복원된 얼굴 영상들에서 역시 뭉개짐 현상이 심하고 가려진 얼굴 부분과 가려지지 않은 부분의 불연속성이 강한 것을 볼 수 있다.

나. 상관관계를 이용한 실험 결과

Yale 데이터 165장 중 100장의 학습 얼굴 영상을 식 (10)을 이용하여 각 화소에 대한 상관계수를 구한다. 그 중 상관계수 값이 큰 1000개만을 추려 저장해 둔다.

우리가 알지 못하는 화소와 연관도가 높은 20개 중 10개의 화소값에 각각 가중치를 주어, 가중합에 의해 해당 화소값을 예측한다. 여기서 예측된 화소값과 테스트 영상의 화소값의 차가 실험을 통해 정한 일정 경계 값보다 크면 가려진 영역, 그 외에는 가려지지 않은 영역이라 간주한다. 그림 11은 상관관계를 이용한 방법에 의해 가려진 영역이 검출된 얼굴 영상들이다.

그림 11을 보면, 가려진 영역의 주변 영역으로 튀어나오는 잡음이 존재하지만 주성분 분석법을 이용한 방법보다 검출된 영상의 잡음이 적고 보다 정확하게 찾았다는 것을 확인할 수 있다.

그림 12는 최종 복원된 영상으로 조금 뭉개지는 경향이 보이나, 주성분 분석법에 의해 복원된 영상보다 뭉

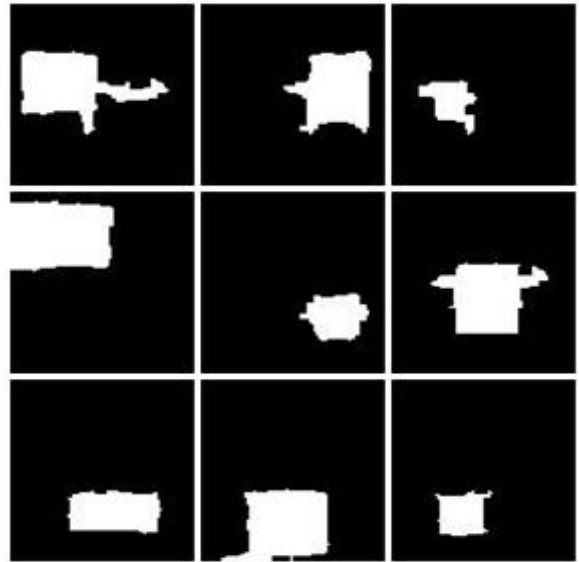


그림 11. 상관관계를 이용하여 yale 데이터의 가려진 얼굴 영상 검출

Fig. 11. Detection of occluded face images for yale data using a correlation based method.



그림 12. 상관관계를 이용하여 yale 데이터의 가려진 얼굴 영상 복원

Fig. 12. Recovery of occluded face images for yale data using a correlation based method.

개짐 현상이 적고 가려진 얼굴 부분과 가려지지 않은 부분과의 경계선 또한 매끄러운 것을 볼 수 있다.

다. 주성분 분석법을 이용한 실험 결과와 상관관계를 이용한 실험 결과와의 비교

앞의 BioID 데이터에 대한 실험에서 서술한 식 (22) 과 식 (23)를 이용하여 yale 데이터에 대한 절대 오차값과 최소 제곱 오차값을 구할 수 있다.

표 3에서 절대 오차값과 최소 제곱 오차값을 비교해보면, yale 데이터에 대한 오차값 역시, 상관계수를 이용한 방법이 주성분 분석법을 이용한 방법보다 오차가 더 작다는 것을 확인할 수 있다. 이는 곧, 상관관계를 이용하여 가려진 부분을 복원한 영상이 기준 얼굴 영상과 더 비슷하다는 것을 알 수 있다.

표 4는 각 방법에 대한 반복 수행 횟수와 처리시간을 나타낸 것으로, 상관관계를 이용한 방법이 주성분 분석법을 이용한 방법보다 약 2.6이상 많이 알고리즘 반복을 수행하였고, 그에 따른 처리시간 또한 더 오래 걸렸다. 그 이유로는 앞서 BioID 문제에서와 마찬가지로 상관관계를 이용한 방법이 검출 과정에서 각각의 모든 화소에 대해 값을 각각 계산하여 예측하고, 복원 과정에서 가려진 영역의 각 화소에 대해 값을 각각 계산하여 예측하기 때문이다.

표 3. 기준 yale 데이터와 복원된 yale 데이터와의 절대 오차값과 최소 오차값 결과
Table 3. L1-norm error and L2-norm error between the ground truth yale data and the reconstructed yale data.

	절대 오차값	최소 제곱 오차값
주성분 분석법	32389.53	1416.22
상관관계 방법	16142.30	840.42

표 4. Yale 데이터에 대한 반복과 처리시간
Table 4. Iteration and processing time for yale data.

		반복		처리시간	
		평균	편차	평균	편차
주성분 분석법		4.9077	1.2339	0.0168	0.0213
상관계수 방법	검출	2	0	1.4031	0.7492
	복원	5.6	3.4315		

V. 결 론

본 논문에서 제안한 상관관계를 이용한 방법에서 가려진 영역의 각 화소값들은 상관도가 높은 화소값들의 가중합에 의해 결정된다. 제안한 방법과 기존의 주성분 분석을 이용한 방법을 비교해 본 결과, 제안한 방법이 더 적은 수의 화소들의 정보를 이용하여 가려진 영역을 복원한다. 또한, 복원된 영상의 뭉개짐 현상이 적고 가

려진 얼굴 영역과 가려지지 않은 영역과의 경계선이 매끄럽게 연결된다. 두 방법에 대한 오차값 결과를 통해 제안한 방법이 기준 영상과의 차이가 더 적다는 것을 알 수 있다. 그러나 주성분 분석을 이용한 방법에 비해 처리속도가 비교적 느리므로, 이를 좀 더 보완하는 방법에 대한 연구가 필요하다.

참 고 문 헌

- [1] 오유화, 안상철, 김형곤, 김익재, 이성환, “반복적인 PCA 재구성을 이용한 얼굴 영상에서의 안경 제거,” 대한전자공학회논문지, 제41권 제3호, 35-49쪽, 2004년 5월
- [2] C. Wu, C. Liu, H. Y. Shum, Y. Q. Xy and Z. Zhang, “Automatic eyeglasses removal from face images,” Asian conference on Computer Vision, IEEE Trans. Pattern Analysis and Machine Intelligence., Vol. 26, no. 3, pp. 322-336, March 2004.
- [3] 박현, 문영식, “2차원 칼라 얼굴 영상에서 반복적인 PCA 재구성을 이용한 자동적인 잡음 제거,” 대한전자공학회논문지, 제43권 제2호, 62-71쪽, 2006년 3월
- [4] Z. M. Wang and J. H. Tao, “Reconstruction of partially occluded face by fast recursive PCA,” International Conference on Computational Intelligence and Security Workshops, pp. 304-307, Dec 2007.
- [5] S. xiaoFeng, T. Min, C. deren and T. RuoFeng, “A two step method to recover occluded part of face,” International Conference on Information Sciences and Interaction Sciences, pp. 169-173, June 2010.
- [6] B. W. Hwang and S. W. Lee, “Reconstruction of partially damaged face images based on a morphable face model.” IEEE Trans. Pattern analysis and Machine Intelligence, Vol. 25. no. 3. pp. 365-372, March 2003.
- [7] D. Lin and X. Tang, “Quality-driven face occlusion detection and recovery,” IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-7, June 2007.
- [8] BioID face database
<http://www.bioid.com/support/downloads/software/bioid-face-database.html>

— 저 자 소 개 —



이 지 은(학생회원)
 2009년 아주대학교 전자공학부
 학사 졸업.
 2011년 아주대학교 전자공학부
 석사 졸업.
 2011년~현재 아주대학교
 전자공학과 박사과정.

<주관심분야 : 컴퓨터 비전, 영상 처리>



곽 노 준(정회원)
 1997년 서울대학교 전기공학부
 졸업.
 1999년 서울대학교 전기 컴퓨터
 공학부 석사 졸업.
 2003년 서울대학교 전기 컴퓨터
 공학부 박사 졸업.

2003년 3월~2006년 8월 삼성전자 정보통신총괄
 통신연구소 표준연구팀 책임연구원.
 2006년 9월~2007년 2월 서울대학교 전기 컴퓨터
 공학부 BK조교수.
 2007년 3월~2011년 2월 아주대학교 전자공학부
 조교수.
 2011년 3월~현재 아주대학교 전자공학부
 부교수.

<주관심분야 : 컴퓨터 비전, 영상 처리>