

비디오속의 얼굴추적 및 PCA기반 얼굴포즈분류와 (2D)²PCA를 이용한 얼굴인식

Face Tracking and Recognition in Video with PCA-based Pose-Classification and (2D)²PCA recognition algorithm

김진율[†] · 김용석

Jin-Yul Kim[†] and Yong-Seok Kim

수원대학교 전자공학과

[†] Dept. of Electronic Engineering, University of Suwo

요 약

통상의 얼굴인식은 사람이 똑바로 카메라를 응시해야 하거나, 혹은 이동하는 통로의 정면과 같이 특정 얼굴포즈를 취득할 수 있는 위치에 카메라를 설치하는 등 통제적인 환경에서 이루어진다. 이러한 제약은 사람에게 불편을 초래하고 얼굴인식의 적용 범위를 제한하는 문제가 있다. 본 논문은 이러한 기존방식의 한계를 극복하기 위하여 대상이 특별한 제약 없이 자유롭게 움직이더라도 동영상 내에서 대상의 얼굴을 추적하고 얼굴인식을 하는 방법을 제안한다. 먼저 동영상 속의 얼굴은 IVT(Incremental Visual Tracking) 추적기를 사용하여 지속적으로 추적이 되며 이때 얼굴의 크기변화와 기울기가 보상이 되어 추출이 된다. 추출된 얼굴영상은 사람과 카메라의 각도를 특정각도로 제한하지 않았으므로 다양한 포즈를 가지게 되며 따라서 얼굴인식을 하기 위해서 포즈에 대한 관정이 선행되어야 한다. 본 논문에서는 PCA(Principal Component Analysis)기반의 얼굴포즈관정방법을 사용하여 추적기에서 추출된 이미지가 5개 포즈별 DB속의 학습된 포즈와 유사한 것으로 판정될 때만 얼굴인식을 수행하여 인식률을 높이는 방법을 제안하였다. 얼굴인식에서는 PCA, 2DPCA, (2D)²PCA의 인식알고리즘을 사용하여 얼굴인식률과 수행시간을 비교 제시하였다.

키워드: 얼굴추적, 얼굴포즈분류, 얼굴인식, PCA, (2D)PCA, (2D)²PCA

Abstract

In typical face recognition systems, the frontal view of face is preferred to reduce the complexity of the recognition. Thus individuals may be required to stare into the camera, or the camera should be located so that the frontal images are acquired easily. However these constraints severely restrict the adoption of face recognition to wide applications. To alleviate this problem, in this paper, we address the problem of tracking and recognizing faces in video captured with no environmental control. The face tracker extracts a sequence of the angle/size normalized face images using IVT (Incremental Visual Tracking) algorithm that is known to be robust to changes in appearance. Since no constraints have been imposed between the face direction and the video camera, there will be various poses in face images. Thus the pose is identified using a PCA (Principal Component Analysis)-based pose classifier, and only the pose-matched face images are used to identify person against the pre-built face DB with 5-poses. For face recognition, PCA, (2D)PCA, and (2D)²PCA algorithms have been tested to compute the recognition rate and the execution time.

Key Words : Face Tracker, Face Pose Classification, Face Recognition, PCA, (2D)PCA, (2D)²PCA

접수일자: 2013년 6월 28일

심사(수정)일자: 2013년 10월 4일

게재확정일자: 2013년 10월 16일

[†] Corresponding author

본 논문은 경기도의 경기도지역협력연구센터사업(GRRC 수원2013-B2, U-city 보안감시 기술협력센터)의 일환으로 수행하였음

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서 론

최근 통제되지 않은 자연스러운 환경에서 얻은 동영상 내의 얼굴을 추적하고 인식하는 필요성이 크게 증가하고 있다.[1][15]-[19] 하지만 기존의 얼굴인식은 고정된 카메라를 정면으로 응시해야만 하는 통제적인 방법이 널리 사용되고 있기 때문에 사람이 정면을 바라보지 않으면 인식이 불가능하다는 문제점을 지닌다. 통제되지 않은 시스템에서는 사람과 카메라 사이의 거리와 각도에 제약이 없기 때문에 획득된 동영상 속에서 추적 대상의 얼굴 크기와 각도 및 포즈는 매우 다양하게 된다. 얼굴인식을 위해서는 이러한 다양한 크기의 변화와 포즈의 변화에도 불구하고 얼굴의 위치를 지

속적으로 추적할 수 있어야 하며 정면이 아닌 포즈를 사용하여도 대상의 얼굴을 인식할 수 있어야 한다. 이를 위하여 다음과 같은 분야의 연구들이 뒷받침 되어져야 한다.

첫 번째는 동영상 속 대상의 움직임을 지속적으로 추적하는 추적기이다. 추적기는 윤곽정보를 기반으로 대상을 추적하는 ACM(Active Contour Model)[2]과 몇 번의 반복을 통해 대상과 가장 유사한 색상 분포를 갖는 영역을 추적하는 MS(Mean Shift)[3], 가중치가 부여된 다수의 샘플을 이용하여 대상의 컬러 분포가 유사한 영역의 확률을 측정하여 추적하는 색상기반 PF(Particle Filter)[4] 등이 가장 대표적인 추적알고리즘이다. 하지만 대상의 색상 분포, 윤곽의 특징 중 하나만 사용하는 추적시스템은 추적대상의 포즈변화에 따라 추적 성능이 크게 바뀌게 되는 문제점들이 존재한다. 이러한 문제를 해결하기 위해서는 대상의 특징을 지속적으로 업데이트하여, 대상의 포즈변화에도 강인하게 추적이 가능한 IVT(Incremental Visual Tracking) 알고리즘[5]과 같은 추적알고리즘이 사용되어야 한다. 두 번째로 대상이 누구인지 인식하기 위한 얼굴인식이다. 얼굴인식에는 대상의 특징을 추출하여 그것의 주성분을 분석하는 PCA(Principal Component Analysis)[6]기법과 특징들의 클래스간의 분산과 클래스 내의 분산의 비율을 최대화하는 방식인 LDA(Linear Discriminant Analysis)[7], 특징들 사이의 여백을 최대화 하는 방법인 SVM[8], 2차원의 이미지의 특징을 행으로 차원을 줄이는 2DPCA[9], 2차원의 이미지의 특징을 행과 열로 차원을 줄이는 $(2D)^2$ PCA[9] 등이 많이 사용되고 있다. 최근에는 동영상 내 얼굴움직임의 dynamics를 학습하기 위하여 ARMA(Auto Regressive Moving Average) 모델[16]과 HMM(Hidden Markov Model)을 사용[17]하기도 한다. 마지막으로 통제되지 않은 환경에서는 기존 연구와 달리 사람이 카메라를 정면으로 응시하지 않으므로 얼굴 포즈에 대한 추정 방법[10]도 별도로 요구된다. 또한 추정된 대상의 포즈별 얼굴이미지를 얼굴인식에 사용하기 위해서는 각 포즈별로 DB를 구성하고, 구성된 DB별로 얼굴인식이 실행되어야 한다[11].

비디오 속의 얼굴인식을 위한 기존 관련 연구들에는 다음과 같이 다양한 방법들이 사용되었다. PCA와 LDA와 같은 선형변환은 널리 사용되는 차원축소방법으로 정지영상 및 동영상을 이용한 얼굴인식 알고리즘에 많이 사용되어져 왔다. Zhao[15]는 개선된 PCA와 LDA를 결합하여 비디오 동영상 내에서 한 장의 사진으로 주어진 얼굴을 찾아내는 방법을 제시하였다. 각 프레임마다 AdaBoost를 사용하여 얼굴을 검출한 후 주어진 사진과 비교하여 사용하여 얼굴을 찾는다. Aggarwal[16]은 ARMA 모델을 이용하여 포즈변화와 얼굴의 움직임에 따른 모습의 변화를 선형 dynamics 시스템으로 모델링하였다. Gallery 비디오시퀀스와 test 비디오시퀀스를 각각 ARMA 모델을 추정한 후 두 파라미터간의 부공간 각도를 측정하여 유사도를 계산하였다. Liu[17]는 비디오 시퀀스에서 얼굴을 인식하기 위하여 HMM 기법을 사용하였다. 훈련단계에서 각 대상의 비디오 시퀀스에서 추출한 얼굴의 공간적(spatial) 특징과 시간적(temporal) dynamics가 HMM로 학습된다. 인식단계에서는 테스트 비디오에서 추출한 얼굴이미지 시퀀스를 학습되어져 있는 각 대상의 HMM 정보와 비교하고 유사도가 가장 높은 시퀀스를 골라내어 얼굴을 인식한다. 그러나 [17]에서는 얼굴 추적을 수행하지 않고 수작업으로 얼굴영상을 추출해 사용하였다. ARMA와 HMM을 사용한 방법은 공간적-시간적 특징

을 동시에 사용한다는 장점을 가지지만 인식과정에서 긴 테스트 비디오 시퀀스를 요구한다는 점이 단점이다. 한편, 얼굴영상의 집합은 포즈, 조명등의 변화로 인해 본질적으로 비선형공간을 이룬다. Wang[18]은 영상집합을 하나의 선형공간으로 표현하는 대신 전체적으로는 비선형이지만, 국지적으로 선형인 부공간(local linear subspace)들의 집합, 즉 소위 "manifold"로 표현하였다. 얼굴 인식은 gallery manifold와 test manifold 간의 거리를 계산하여 유사도가 높은 얼굴 시퀀스를 찾아냄으로써 얼굴인식을 수행한다. [18]에서 Adaboost를 사용한 얼굴검출이 사용되었으며 얼굴추적기법은 사용하지 않았다. 포즈변화와 조명변화에 따른 성능저하를 극복하기 위하여 3D 얼굴 모델을 사용하기도 한다. [19]에서는 3D scanner를 사용하여 3D 얼굴모델을 미리 구축한 후, 입력 비디오로부터 추출된 얼굴의 조명상태와 포즈를 추정하여 3D 얼굴모델로부터 해당되는 2D 영상의 투영(projection) 영상을 얻고 이를 입력된 얼굴과 비교하여 얼굴 인식을 수행하였다. 그러나 3D영상을 구축하는 노력이 많이 들고 실제 얼굴과 투영된 얼굴영상과의 차이로 인하여 인식성능의 저하가 발생한다.

본 논문은 기존 정지영상이 아닌 얼굴추적의 결과로 획득한 얼굴영상 시퀀스로부터 포즈를 구분하고 사람을 인식하는 시스템의 구현에 관한 것이다. 기존의 연구에서는 얼굴추적의 어려움을 피하기 위해 수작업[17]을 사용하거나 얼굴추적기 대신 계산량이 더 많이 필요한 얼굴 검출기법을 사용[18][1]하기도 하였다. 본 연구에서는 얼굴추적기를 얼굴영상 DB 구축과 테스트 과정 모두에 적용하였다. 얼굴추적에는 IVT 추적기[5]를 사용하였다. IVT 얼굴추적기는 추적 대상의 얼굴 특징을 지속적으로 업데이트하기 때문에 주위 환경변화와 대상의 포즈변화에 강인한 특성을 가진다. 다음에, 추적기로부터 추출된 얼굴 이미지의 포즈를 분류하고 분류된 포즈 내의 DB와 비교하여 얼굴인식을 수행한다. 이때 추출된 이미지의 포즈가 포즈별 DB에 학습된 얼굴포즈와 유사할 때만 추출된 얼굴이미지를 저장하고 PCA기반 얼굴인식 알고리즘을 사용하여 얼굴인식을 수행한다. PCA기반 인식은 HMM기법과 달리 적은 수의 프레임으로도 얼굴인식을 수행할 수 있는 장점을 가진다. PCA의 성능저하를 극복하기 위하여 2DPCA와 $(2D)^2$ PCA를 사용하여 얼굴인식을 시도하였다. $(2D)^2$ PCA는 인식성능의 개선과 함께 기존 인식 방법들에 비해 매우 빠른 수행시간을 보인다.

본 논문의 구성은 다음과 같다. 2장에서는 실험에서 사용한 얼굴추적기인 IVT에 관하여 설명하였고, 3장에서는 PCA, 2DPCA, $(2D)^2$ PCA의 얼굴인식 알고리즘에 대해 기술하였다. 4장에서는 얼굴추적기에서 추출된 이미지를 사용하여 대상의 포즈를 판별하고, DB와 유사한 이미지만을 추출하는 방법에 대하여 설명하였다. 마지막으로 5장에서는 앞서 설명한 알고리즘을 구현하여 동영상에서 얻은 이미지를 사용한 얼굴인식 실험결과를 제시하였다. 실험은 본 실험실에서 구축한 동영상 DB와 Honda/UCSD 동영상 DB[20]를 이용하여 각각 이루어졌다.

2. 얼굴추적기(Face tracker)

2.1 얼굴추적기(Face tracker)

비디오 영상속에서 물체를 추적하기 위해서 PF(Particle Filter) 기법이 널리 사용되어 왔다. 일반적인 PF의 구성은

다음의 4단계의 과정을 가진다. 먼저 이전 과정에서 얻은 샘플의 가중치에 비례하는 확률로 새로운 샘플을 선택하고 (Selection), 선택된 샘플을 새로운 위치로 전파시킨다 (Propagation). 그리고 각 샘플에서 유사도를 측정하여 가중치를 부여하고(Observation), 가중치가 부여된 샘플의 위치의 평균으로 객체를 추정한다(Estimation). 그리고 위의 과정을 반복하여 대상을 추적한다.

색상기반 PF 추적 방법[4]은 각 샘플 위치에서 추적대상과의 색분포의 유사도를 계산하여 대상이 존재하는 확률을 측정하고 대상을 추적하는 대표적인 추적알고리즘이다. 그러나 얼굴추적에서와 같이 추적대상인 얼굴의 포즈가 변하는 경우 포즈변화에 따라 추적대상의 모양과 색분포가 변화하게 되므로 추적성능이 떨어지는 문제점이 존재하게 된다. IVT[5]는 추적대상의 모양이 변화하는 경우에 성능이 저하되는 문제를 개선할 수 있다. IVT에서는 추적대상의 이전 영상들로부터 PCA를 사용하여 특징을 추출하고 이 특징을 PF 기법을 이용하여 추적함으로써 모양이 변하는 대상물체에 효과적으로 대응할 수 있다.

IVT에서는 다음과 같은 베이저안 확률을 계산하여 대상을 추적한다:

$$P(u_t|I_{0:t}) \propto P(I_t|u_t) \int P(u_t|u_{t-1})P(u_{t-1}|I_{0:t-1})du_{t-1} \quad (1)$$

여기서 u_t 와 I_t 는 각각 시간 t 에서의 대상의 상태와 입력 되는 이미지 프레임을 말한다. 그리고 초기의 u_0 는 알고 있다고 가정한다. 샘플의 상태 u_t 의 매개변수에는 대상의 중심좌표, 참조영상 대비 크기(즉, 스케일), 수평축에 대한 회전각도(roll) 등이 포함되어 있다. IVT에서는 통상의 얼굴 추적 PF 기법에서와 마찬가지로 u_{t-1} 에서 u_t 로 샘플을 전파시킬 때 복잡한 전파모델(dynamic model)을 사용하는 대신 다음과 같이 간단한 가우시안 분포모델을 사용하였다. 즉,

$$P(u_t|u_{t-1}) = N(u_t; u_{t-1}, \Sigma) \quad (2)$$

여기서 Σ 는 알맞게 선택된 대각선 공분산(diagonal covariance) 행렬이다.

IVT에서는 Observation단계에서 한 샘플의 상태 u_t 로부터 얻어지는 이미지 I_t 의 유사도, 즉, $P(I_t|u_t)$ 를 과거의 추적대상영상들 $I_{0:t-1}$ 에 PCA를 수행하여 얻은 평균(μ)과 고유벡터(B)로 표현되는 부공간(subspace)으로부터 현재의 영상 이미지 I_t 를 얻을 수 있는 확률을 계산함으로써 구한다[5]. 이 확률은 결국 영상 I_t 와 부공간내의 중심점까지의 거리에 반비례하므로 다음과 같이 2가지 거리를 계산하여 얻을 수 있다. 즉, I_t 를 부공간으로 투영할 때 I_t 로부터 투영된 점까지의 거리와 이 투영된 점에서 부공간 내의 중심점(평균 μ 로 주어지는)까지의 거리에 근거하여 유사도를 계산하였다. [14]에도 비슷한 방법이 제시된 바 있다.

한편, 새로운 영상프레임 I_t 이 추가될 때마다 과거의 추적 대상영상들 $I_{0:t-1}$ 에 근거해서 얻은 PCA 부공간의 평균(μ)과 고유벡터(B)는 새롭게 갱신되어야 한다. 그러나 매 프레임마다 새로운 PCA를 수행하면 계산량이 너무 많아져 실시간 처리가 어려우므로 [5]에서는 ISVD(Incremental SVD)[13]

를 사용하여 재귀적인 방법으로 새로운 평균(μ')과 고유벡터(B')를 갱신하는 방법을 제시하여 계산량을 낮추었다. 이러한 결과로 IVT는 추적대상의 모습이 조금씩 바뀌는 경우에도 적은 계산량으로 우수한 추적성능을 보인다.

본 논문에서는 얼굴추적기 구현에 [5]에서 제시된 얼굴추적 방법을 사용하였다.

3. 얼굴인식 알고리즘

3.1 PCA 알고리즘

얼굴추적기에서 추출된 이미지는 2차원의 영상 이미지이고, 고차원의 데이터를 가지고 있기 때문에 학습 성능 및 속도를 향상시키기 위해서 대표적인 차원축소 알고리즘인 주성분 분석법(PCA)을 사용하여 차원을 축소한다.[6] PCA는 고차원의 특징벡터를 저차원의 특징벡터로 정보의 손실이 최소화되도록 차원을 축소시키는 대표적인 알고리즘이다.

PCA알고리즘을 이용한 얼굴인식은 얼굴영상의 크기가 $N \times N$ 이고, 인식후보 얼굴영상의 개수가 M 개 일 때 각 후보 얼굴영상을 M 개의 $N^2 \times 1$ 행벡터로 인식후보의 얼굴벡터 집합 $S = [X_1, X_2, X_3, \dots, X_M]$ 을 구성한다. 인식 후보들의 얼굴벡터 집합을 사용하여 평균 영상을 구할 수 있다. 평균 얼굴 벡터는 다음과 같이 정의 한다.

$$\bar{X} = \frac{1}{M} \sum_{n=1}^M X_n \quad (3)$$

공분산 행렬을 구하기 위해 학습영상 벡터와 평균 얼굴 영상 벡터의 차 영상 벡터 $\phi_i = X_i - \bar{X}$ 를 구한다. 차 영상 벡터에 대한 공분산 행렬 C 는 다음과 같다.

$$C = \frac{1}{M} \sum_{n=1}^M \phi_n \phi_n^T = \Omega \Omega^T \quad (4)$$

이 결과 $\Omega = [\phi_1, \phi_2, \phi_3, \dots, \phi_M]$ 는 $N^2 \times M$ 행렬이 되며, 공분산 행렬은 $N^2 \times N^2$ 의 차원을 갖는다. 공분산 행렬로부터 M 개의 고유값 λ_i 와 이에 대응하는 고유벡터 u_i 를 계산한다. 여기서 고유값에서 얻어진 고유벡터를 $N \times N$ 로 재배열하면 사람 모양과 닮은 고유얼굴을 얻을 수 있다. 고유얼굴은 다음과 같은 식으로 표현되며, 고유얼굴이 완성되면 얼굴인식을 위한 학습단계가 끝나게 된다.

$$u_l = \frac{1}{M} \sum_{k=1}^M \lambda_k \phi_k \quad l = 1, 2, \dots, M \quad (5)$$

새로운 인식 후보(A)가 들어오게 되었을 때, 인식 후보들의 평균벡터(\bar{X})를 사용하여 차벡터를 만들고, 그것을 학습단계에서 만든 고유얼굴에 투영하여 ($\omega_k = u_k^T(A - \bar{X})$) 각 인식후보들과의 유클리디언 거리가 가장 작게 나오는 사람을 인식 대상으로 판정한다.

본 논문에서는 PCA를 이용한 얼굴인식과 얼굴인식 이전에 필요한 포즈 분류에서 사용하였다. 포즈분류 과정에서는 포즈별로 PCA 공간을 구축하고 테스트 이미지의 유사성을 비교하여 얼굴포즈를 분류한다.

3.2 행 방향으로 차원을 줄인 2DPCA 알고리즘

PCA 얼굴인식은 2차원 영상을 1차원 벡터로 표현 하였다. 하지만 이미지는 본질적으로 행렬로 되어있다. 2DPCA 알고리즘[9]은 이미지를 1차원 벡터로 변환하여 차원을 축소하지 않고, 2차원의 이미지 그대로 행 방향으로 차원을 축소하는 방식이다. M개의 학습이미지가 있을 때, i번째 학습이미지는 $X_i^{m_1 \times m_2}$ ($i=1,2,\dots,M$)로 나타낼 수 있다. M개의 학습이미지에 대한 평균 이미지는 $\bar{X}^{m_1 \times m_2}$ 로 나타낸다. 이때 학습이미지들에 대한 공분산 행렬은 다음과 같이 나타낼 수 있다.

$$C = \frac{1}{M} \sum_{i=1}^M (X_i - \bar{X})(X_i - \bar{X})^T \quad (6)$$

여기서 공분산 행렬 C를 이용하여 행 방향으로 $X^{m_1 \times d_1}$ 만큼 차원을 축소한 특징행렬 U를 구한다. d_1 는 행렬 U에서 d_1 개의 큰 특징행렬 만큼 차원을 줄인 것을 말한다. 따라서 2DPCA의 특징행렬은 $m_1 \times d_1$ 의 크기를 갖는다. 이렇게 구해진 행렬 U는 새로운 인식후보 A가 주어질 때 $Y = U^T(A - \bar{X})$ 로 투영 된다. 2DPCA는 기존의 PCA보다 더 작은 공분산의 크기를 가지기 때문에 계산 속도가 빠르고 인식률이 PCA보다 더 좋은 장점을 가진다.

3.3 열 방향으로 차원을 줄인 2DPCA 알고리즘

열 방향으로 차원을 줄인 2DPCA[9]알고리즘은 행 방향 2DPCA와 같이 2차원 이미지를 1차원으로 변환하여 차원을 축소하지 않고, 2차원의 이미지 그대로 차원을 축소하는 방식이다. 하지만 앞서 설명한 2DPCA에서 행 방향으로 차원을 축소하였다면, 이번에는 열 방향으로 차원을 축소한다. 이때 학습이미지들에 대한 공분산은 다음과 같은 식으로 표현된다.

$$C = \frac{1}{M} \sum_{i=1}^M (X_i - \bar{X})^T (X_i - \bar{X}) \quad (7)$$

여기서 공분산 행렬 C를 이용하여 행 방향으로 $X^{d_2 \times m_2}$ 만큼 차원을 축소한 특징행렬 V를 구한다. d_2 는 행렬 V에서 d_2 개의 큰 특징행렬 만큼 차원을 줄인 것을 말한다. 따라서 열 방향으로 차원을 줄인 2DPCA의 특징행렬은 $d_2 \times m_2$ 의 크기를 갖는다. 이렇게 구해진 행렬 V는 $Y = (A - \bar{X})V$ 로 투영 된다.

3.4 (2D)²PCA 알고리즘

(2D)²PCA 알고리즘[9]은 행 방향으로 차원을 축소한 2DPCA와 열 방향으로 차원을 축소한 2DPCA를 결합한 방식으로 2차원 이미지의 특징을 양쪽 방향에서 차원을 축소하는 방법을 말한다. 이렇게 생성된 2개의 특징행렬들을 이용하여 새로운 인식후보들이 들어왔을 때 다음과 같은 식으로 표현 할 수 있다.

$$Y = U^T(A - \bar{X})V \quad (8)$$

앞서 설명한 행 방향으로 차원을 줄인 2DPCA와 열 방

향으로 차원을 줄인 2DPCA의 특징행렬 U와 V는 행 방향으로 $m_1 \times d_1$ 만큼 차원이 줄어들고, 열 방향으로 $d_2 \times m_2$ 만큼 차원이 줄어든다. 따라서 Y의 행렬의 크기는 $d_1 \times d_2$ 의 크기를 갖는다. 이렇게 행과 열의 특징 벡터 중에서 가장 큰 몇 개의 주성분만을 추출하여 차원을 줄임으로써 2DPCA와 같은 인식률을 가지며 계산하는 차원이 줄어서 연산의 속도가 감소하는 장점이 있다.

4. 얼굴 포즈 분류

4.1 Multi-Space PCA를 이용한 포즈 분류

얼굴추적기에서 얻어지는 얼굴 영상 시퀀스를 그대로 얼굴인식에 사용할 경우 포즈 변화에 따른 얼굴 이미지의 차이를 고려하지 않으므로 성능의 저하가 불가피하다. 얼굴 인식 성능의 향상을 위해서는 추적기에서 추출된 얼굴을 포즈별로 분류하고, 분류된 포즈 안에서 얼굴인식을 해야 한다. 얼굴 포즈를 분류하기 위해서는 먼저 다른 사람들의 같은 포즈들을 그림 1과 같이 yaw 각도($\pm 90^\circ, \pm 45^\circ, 0^\circ$)에 따라 분류한다. 분류된 이미지로 DB를 구성하고 각 포즈별로 PCA를 수행하여 포즈별 고유 얼굴벡터를 갖는 multi-space PCA 공간을 구성한다.

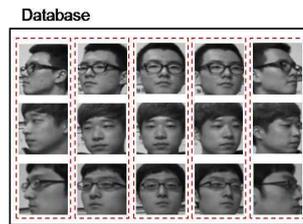


그림 1. Multi-Space PCA
Fig. 1. Multi-Space PCA

다음으로, 테스트하려는 얼굴 이미지를 각 PCA 공간으로 투영하여 그 거리를 계산하고 최소 거리를 갖는 포즈로 분류한다. 이때, 각 포즈별 고유 얼굴벡터의 모든 성분을 사용하여 포즈를 분류 하려 한다면, 포즈별 고유벡터 안에서 잘못된 값으로 매칭 되어 포즈가 분류 되지 않는다[12]. 따라서 포즈를 분류하기 위해서는, 각 포즈별 PCA공간이 가지고 있는 가장 큰 몇 개의 주성분을 추출하고, 추출된 주성분을 테스트 이미지에 투영하여 해당 포즈별 PCA 공간과의 유클리디언 거리가 최소일 때 포즈로 추정한다. 이상을 정리하면, 포즈를 분류하는 과정은 다음과 같이 진행된다.

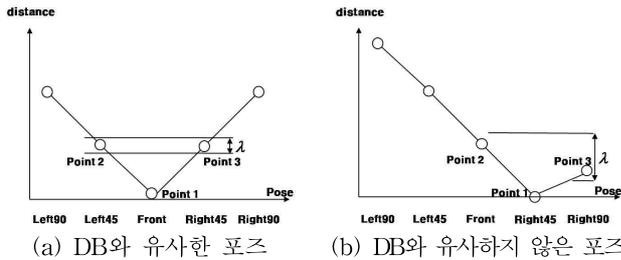
포즈를 분류하는 과정

- 1단계 : 각 포즈별로 PCA알고리즘을 실행하여 고유 얼굴벡터를 생성.
- 2단계 : 각 포즈의 특징을 나타내는 큰 주성분을 추출.
- 3단계 : 인식후보의 특징을 각 포즈별 주성분으로 투영하여 가까운 포즈를 찾음.

4.2 DB와 유사한 포즈를 갖는 이미지 추출

얼굴추적기에서 추출된 각 이미지를 앞서 설명한 4.1절의 방법으로 포즈를 판별하면 5개의 포즈 중 하나로 포즈를

분류할 수 있다. 하지만 분류된 포즈별 DB내에서 그대로 얼굴인식을 실행하게 된다면 DB속의 학습된 이미지와 판별된 이미지의 포즈 각도 차이가 존재하므로 얼굴 인식이 하락한다. 이러한 문제점을 해결하기 위해 추출된 이미지가 포즈별 DB와 유사한 포즈 일 때만 얼굴인식에 사용하도록 하여 동영상 속에서의 얼굴인식의 성능을 개선한다.



(a) DB와 유사한 포즈 (b) DB와 유사하지 않은 포즈
그림 2. 인접 포즈간의 거리 값

Fig. 2. Value of Distance between Adjacent Poses

대상이 어떤 포즈를 가질 때 다른 포즈들에 대한 거리의 차이는 2차 곡선 모양이 되어 추정된 포즈에서 인접한 다른 포즈까지의 거리가 같다고 알려져 있다[12]. 예를 들어 현재 포즈가 그림 2(a)와 같이 정면(Front)이라고 하자. 이때 다음으로 가까운 포즈인 좌(Right45), 우(Left45)의 두 포즈간의 거리값의 차이는 거의 없게 된다. 따라서 좌(Right45), 우(Left45) 포즈와의 거리값을 각각 계산하여 그 차이가 거의 없을 때, 현재 이미지가 DB와 가장 유사한 이미지임을 알 수 있다. 이렇게 계산된 Left45(Point2), Right45(Point3)의 거리값 차이를 λ 로 두고, 그 차이가 일정한 범위(0과 가까운)안에 들어올 때 이미지는 DB속의 이미지와 유사한 포즈임을 나타내므로 이 얼굴이미지를 얼굴인식에 사용한다. 반면, 그림 2(b)에서는 yaw 각도가 변하여 얼굴이 한 포즈에서 다음 포즈로 넘어 가는 과정에서 두 포인트의 거리 차 λ 가 증가하는 것을 볼 수 있다.

5. 제안된 알고리즘의 구현 및 실험결과

5.1 제안된 알고리즘의 구현

본 실험에서는 앞서 설명한 2장의 얼굴추적기와 3장의 얼굴인식 알고리즘, 4장의 얼굴포즈 판별 방법을 구현하여 동영상 속에서 대상의 포즈 및 얼굴을 인식하는 시스템을 구성하였다. 그림 3은 제안한 알고리즘의 흐름도이다. 제안한 시스템은 크게 세 부분으로 나누어져있다. 첫 번째로 대상을 추적하기 위한 얼굴 추적기이고, 두 번째로 추적된 대상이 바라보는 포즈를 추정하는 포즈판별부이다. 마지막으로 판별된 포즈 안에서 PCA, 2DPCA, (2D)²PCA 등의 얼굴인식 알고리즘을 사용한 얼굴인식부이다. 실험에서 사용된 DB는 12명을 학습하였고, 5개의 포즈에 한 사람당 10장씩의 이미지를 사용하여 총 600장의 이미지를 학습에 사용하였고, 얼굴이미지의 크기는 90 × 90을 사용하였다. 그리고 실험에 사용된 동영상은 초당 30프레임의 속도로 촬영되었고, 실험 대상인 사람당 평균 542장의 프레임에 가지는 동영상을 실험에 사용하였다. 알고리즘은 Matlab으로 구현되었고 실험에서 얼굴의 pitch 방향은 고려하지 않았다.

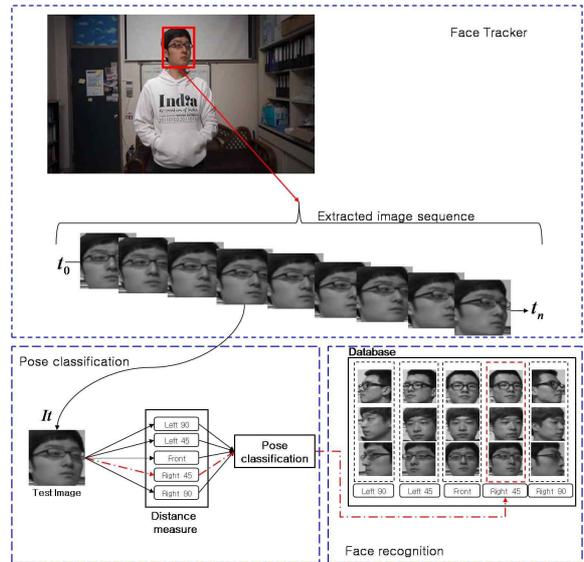


그림 3. 제안된 알고리즘의 흐름도
Fig. 3. Flow of the Proposed Algorithm

그림 3에서 $t_0 \sim t_n$ 는 2.1절의 IVT를 사용하여 동영상 속에서 추출한 얼굴이미지 시퀀스이다. 이렇게 얻어진 각 얼굴이미지는 앞서 설명한 4장과 같이 포즈 판정용 DB구축 과정에서 미리 구해놓은 주요 주성분을 이용하여 투영되고 가장 작은 거리 값을 가지는 포즈를 현재 포즈로 추정한다. 이때 추출되는 얼굴 이미지 시퀀스에는 동영상의 특성상 급격한 포즈변화가 존재하지 않는다. 따라서 포즈 분류에 그림 4와 같은 포즈별 가중치를 추가로 사용하였다.

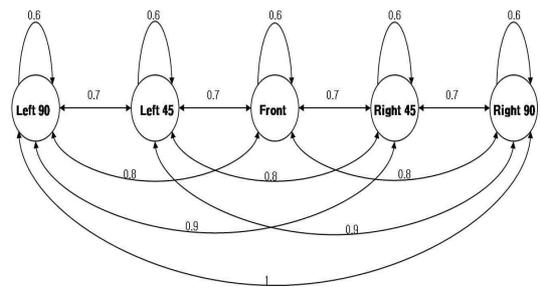


그림 4. 분류된 포즈에 사용한 가중치
Fig. 4. Weights for Classified Poses

예를 들어 현재 포즈가 Front라고 가정한다면 추정된 포즈에 가장 작은 가중치를 주고, 그 다음으로 가까운 포즈인 Left45, Right45에 두 번째로 작은 가중치를 주는 방식이다. 여기서 포즈를 추정할 때 가장 작은 값을 가지는 포즈를 현재 포즈로 추정하기 때문에 현재의 포즈와 거리가 멀수록 큰 가중치를 준다. 이렇게 하면, 동영상 속에서 추정된 포즈가 급격히 다른 포즈로 잘못 투영 되어 포즈 판별의 정확성을 저하 시키는 것을 방지 할 수 있다. 다음으로 추출된 이미지의 포즈를 추정하였다면, 추정된 포즈 안의 DB를 사용하여 3장의 얼굴인식 알고리즘을 사용하여 얼굴인식하게 된다.

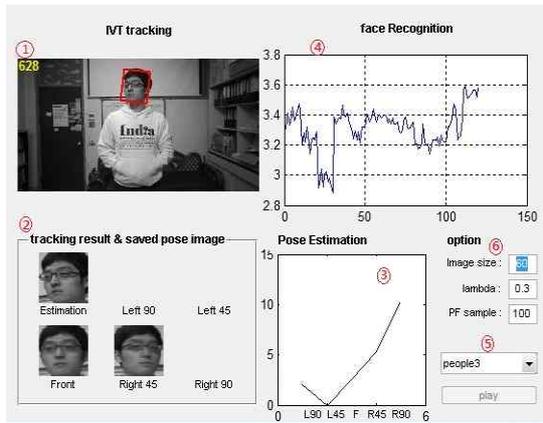


그림 5. 제안된 시뮬레이터
Fig. 5. Proposed Simulator

그림 5은 제안된 알고리즘을 바탕으로 만들어진 시뮬레이터이다. 그림 5에서 표시된 ①번은 추적기이다. ②번은 추적기에서 추정된 이미지와 추정된 이미지 중에서 DB와 유사한 포즈의 이미지일 때, 각 포즈별로 저장되는 이미지를 나타낸 것이다. 또한 ③번은 현재 추적중인 대상이 어느 포즈를 바라보고 있는가에 대한 포즈의 정보를 나타낸다. 그리고 ④번은 각 포즈별로 저장된 이미지를 사용하여 얼굴인식을 한 결과를 반영한 그래프이다. 여기서 그래프에 나타난 얼굴인식 알고리즘은 PCA, 2DPCA, (2D)²PCA를 선택하여 사용할 수 있다. ⑤번은 시뮬레이터에서 작동하는 동영상을 선택한다. 마지막으로 ⑥번은 IVT추적기에서 사용되는 샘플 개수와 얼굴인식에 사용되는 이미지의 크기를 선택하는 설정이다.

아래 5.2절에는 제안된 알고리즘의 성능을 검증하기 위하여 자체로 구축한 DB를 사용하여 정지영상에서의 얼굴인식을 포함하여 동영상 속에서 추출된 이미지를 사용한 여러 가지 실험조건에 따른 얼굴인식 실험결과를 제시하였다. 그리고 5.3절에서는 기존 결과와의 성능 비교를 위하여 Honda/UCSD DB[20]를 사용한 실험결과를 제시하였다.

5.2 제안된 알고리즘의 실험결과

동영상 속의 얼굴인식 실험에 앞서 정지영상에서 얼굴인식을 한 결과(실험 I)를 표 1에 나타내었다. 정지영상의 얼굴인식은 PCA 알고리즘을 이용하여 인식률을 계산하였고,

실험에 사용한 테스트 이미지들은 카메라를 정면으로 바라보는 동영상에서 대상의 얼굴을 추출하여 얼굴인식에 사용하였다.

실험 II는 동영상 속의 얼굴인식을 위한 실험으로 추적기에서 추출되는 모든 이미지를 4장에서 설명한 방법으로 포즈를 분류하여 PCA 얼굴인식을 수행한 결과이다. 이때 인식률은 정지영상에서 얼굴 인식한 결과보다 인식률이 10%정도 하락 되는 것을 볼 수 있다. 그 이유는 동영상에서 추정된 대상의 포즈가 5개의 포즈 중 한곳으로 투영될 때 DB속의 학습된 이미지와 포즈 차이가 존재하기 때문이다.

실험 III은 추적기에서 추출되는 모든 이미지를 사용하지 않고 DB속의 포즈와 가장 유사한 포즈일 때 얼굴추적기의 영상을 추출하여 그것을 이용하여 얼굴인식을 실행한 결과이다. 표 1에서 보이는 것처럼 이러한 방법으로 이미지를 추출하고 얼굴 인식을 한 경우, 정지영상에서 정면을 바라보고 얼굴인식을 실행한 경우와 비슷한 인식률이 나오는 것을 확인 할 수 있었다. 이것은 사람이 움직이는 동영상에서 DB와 유사한 이미지를 찾아내어 추출 한다면 정지영상의 얼굴인식과 비슷한 결과를 얻을 수 있다는 가능성을 확인한 실험이며, 동영상 속에서 얼굴인식에 사용하기 좋은 이미지를 추출하기 위한 실험이다.

실험 IV에서는 실험 III과 동일한 조건에서 2DPCA 알고리즘을 사용하여 얼굴인식을 하였다. 먼저 2DPCA에 사용한 특징행렬의 크기는 $m_1 \times d_1$ 의 크기를 갖는다. 여기서 실험에서 d_1 의 크기는 6을 사용하였다. 따라서 2DPCA의 특징 행렬의 크기는 90×6 의 크기를 갖는다. 그리고 특징 행렬의 행 방향으로 차원을 축소하고 얼굴인식 한 결과를 표 1에 나타내었다. 그리고 두 번째 실험인 DB와 유사한 이미지를 추출한 이미지에 기존의 얼굴인식 알고리즘보다 2DPCA를 사용한 얼굴인식이 약 3% 정도 인식률이 상승하는 것을 확인 할 수 있었다.

마지막으로 실험 V는 (2D)²PCA를 사용한 얼굴인식 실험이다. 먼저 $m_1 \times m_2$ 의 행렬에서 행 방향으로 가장 큰 성분인 d_1 개만큼 차원을 축소하였고, 열 방향으로 가장 큰 성분인 d_2 개만큼 차원을 축소하였다. 여기서 $m_1 \times m_2$ 의 크기는 이미지의 크기와 동일하고 특징벡터 d_1 과 d_2 는 각각 7, 6으로 차원을 축소하였다. 동영상 속에서 대상의 포즈를 추정하고, 포즈 별로 얼굴인식을 실행하기 위해서는 얼굴인식에 사용되는 차원이 크지 않아야 한다. 그래서 얼굴인식에 사용되는 차원의 크기를 60차원에서부터 점차 차원을 줄였

표 1. 각 실험에 대한 결과 비교
Table 1. Comparison of Experimental Results

Exp. No.	Conditions	Algorithm	No. Images Used / Total No. Images	Dimension	Processing Time(sec)	Recognition Rate	
I	Still-Images (Front)	PCA	-	60	-	91.0 %	
II	Images from video	All images	PCA	542 / 542	60	6.2	81.4 %
III		Images with similar poses in DB	PCA	55 / 542	60	0.63	91.3 %
IV			2DPCA	55 / 542	7×90	0.29	94.4 %
V			(2D) ² PCA	55 / 542	7×6	0.09	94.6 %

을 때, 7×6 의 차원이 인식시간 대비 인식률이 가장 좋게 나타나 실험에서는 d_1 과 d_2 의 차원의 크기를 각각 7, 6으로 사용하였다. 실험결과를 통해 (2D)²PCA는 행과 열의 두 방향으로 차원을 축소하고 얼굴인식을 실행하기 때문에 얼굴인식에 필요한 연산시간이 2DPCA보다도 약 1/3배 정도 감소하면서도 2DPCA의 인식률과 유사한 인식률을 얻을 수 있는 장점을 지님을 확인할 수 있었다.

5.3 Honda/UCSD Database에 적용한 실험결과

제안된 방법의 성능을 검증하기 위하여 국외에 널리 알려진 Honda/UCSD DB[20]를 이용하여 기존 연구결과와 비교하였다.

Honda DB는 20명에 대하여 총 60여개의 동영상 이 존재하는데 그중 학습에 20개의 동영상을 사용하고, 나머지 40개의 동영상을 테스트 동영상으로 사용한다. 먼저 앞서 실험한 알고리즘의 포즈정보를 이용하여 Honda/UCSD DB의 학습동영상에서 포즈를 자동으로 분류하고, 분류된 포즈를 기반으로 얼굴인식에 사용하기 위한 DB를 새로 구성하였다. 실험에서 사용된 DB는 20명을 학습하였고, 한 사람당 5개의 포즈, 또 포즈별로 5장씩의 이미지를 사용하여 총 500장의 이미지를 사용하였고, 사용된 얼굴 이미지의 크기는 90×90 이다. Honda/UCSD DB에 제안된 알고리즘을 적용한 결과를 표 2에 보였다.

실험VI, VII, VIII은 실험II,III, V와 각각 동일한 조건에서 수행한 실험이다. 표 2에서 Honda/UCSD DB는 pitch방향으로 급격한 움직임이 많기 때문에 인식률이 크게 저하된 것을 볼 수 있다. 실험IX은 동영상이 pitch방향으로 급격한 움직임을 가지는 문제를 해결하기 위하여 pitch 방향으로 임계값을 설정하여 같은 포즈라 하더라도 pitch가 임계값 밖이라고 판단되면 얼굴인식에 그 영상을 사용하지 않도록 하였다.

한편, 같은 Honda/UCSD DB를 사용한 기존 연구결과에 따르면 [16]에서 90%, [18]에서 96.9%의 인식률을 각각 얻었다. 본 논문에서 제안된 방법은 인식률에 있어서 큰 개선을 보이지 못하고 있으나 (2D)²PCA를 사용함으로써 연산 수행시간을 크게 줄일 수 있다.

6. 결론 및 향후 연구

본 논문은 하나의 카메라를 사용하여 움직이는 대상의 얼굴을 추적하고 추적된 결과로 획득한 얼굴 이미지를 사용하

여 대상의 포즈와 대상이 누구인지 찾아내는 시스템 구현에 관한 것이다. 추적기에서는 대상의 특징을 점차적으로 업데이트 하여 대상을 추적하는 IVT 알고리즘을 사용하였다. 사람의 포즈를 구분하기 위해서 각 포즈별 가장 큰 특징을 추출하여 그것을 추적기에서 얻어진 얼굴이미지에 투영하여 대상의 포즈를 판별하였고, 추출되는 이미지들 중에 DB와 유사한 이미지만을 찾아내어 얼굴인식에 사용하였다. 동영상에서 추출된 이미지의 포즈를 판별하고 판별된 포즈내의 DB에 PCA, 2DPCA, (2D)²PCA의 얼굴인식 알고리즘을 사용하여 얼굴 인식한 결과를 제시하였다.

추가적으로 얼굴인식의 성능을 향상시키기 위하여 추출되는 이미지의 배경 성분을 ASM(Active Shape Model)등을 이용한 분할법을 적용하여 배경 성분을 제거하면 얼굴인식의 성능을 좀 더 향상시킬 수 있을 것으로 기대한다. 또한 테스트이미지의 밝기 변화를 고려한 연구와 안정적인 추적을 위하여 IVT추적기가 가지는 단점인 장시간 동안 추적하여 배경의 성분들이 계속 누적되어 추적에 실패 하는 부분에 대한 연구가 되어야 한다.

References

[1] H. K. Ekenel, J. Stallkamp, and R. Stiefelhagen, "A video-based door monitoring system using local appearance-based face models," *Computer Vision and Image Understanding*, Vol. 114, Issue 5, pp. 596-608, 2010.

[2] M. Kass, "Snake: Active contour model," *International Journal of Computer Vision*, Vol.1 pp. 312-331, 1988.

[3] D. Comaniciu, V. Ramesh, "Mean-shift and opimal prediction for efficient object tracking," *International Confèrence on Image Processing*, pp 70-73, 2000.

[4] K. Nummiaro, E. Koller-Meier and L.van Gool, "A color-based particle filter," *First International Workshop on Generative Model Based Vision*, pp. 53-60, June 2002.

[5] D. Ross, J. Lim, R.-S. Lin, and M.-H. Yang. "Incremental learning for robust visual tracking," *IJCV*, 2007.

[6] M. A. Turk, A. P. Pentland, "Face recognition using

표 2. Honda/UCSD DB에서 각 실험에 대한 결과 비교
Table 2. Comparison of Experimental Results in Honda/UCSD DB

Exp. No.	Conditions (Images from Video)	Algorithm	No. Images Used / Total No. Images	Dimension	Processing Time(sec)	Recogniti on Rate
VI	All images	PCA	386 / 386	60	6.2	65.2%
VII	Images with similar poses in DB	PCA	96 / 386	60	0.63	74.1%
VIII		(2D) ² PCA	96 / 386	7×6	0.09	87.6%
IX	Images with a large pitch variation rejected	(2D) ² PCA	72 / 386	7×6	0.09	92.3%
Results from previous researches			Aggarwal[16] 90%, Wang[18] 96.9%			

eigenface," *Proc. Computer Vision and Pattern Recognition*, pp 586 - 591, 1991.

[7] S. Mika, G. Ratsch, J. Weston, B. Scholkopf, and K. R. Mullers, "Fisher discriminant analysis with kernels," *Neural Networks for Signal Processing IX, Proceedings of the IEEE Signal Processing Society Workshop*, pp. 41-48, 1999.

[8] D. Anguita, S. Ridella, F. Riviaccio, R. Zunino, "Hyperparameter design criteria for support vector classifiers," *Neurocomputing*, Volume 55, Issues 1 - 2, Sep. 2003, pp. 109 - 134.

[9] Z. Daoqiang, Z. Zhi-Hua, "Two-directional two-dimensional PCA for efficient face representation and recognition," *Neurocomputing*, Vol 69, Issues 1 - 3, December 2005, Pages 224-231.

[10] E. Murphy-Chutorian, "Head Pose Estimation in Computer Vision: A Survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 31, pp 607-626, 2009.

[11] Y.-O. Kim, "A new 3D active camera system for robust face recognition by correcting pose variation," *ICCV 2004*, pp. 1485-1490.

[12] S. Srinivasan, "Head pose estimation using view based eigenspaces," *Proc. 16th International Conference on Pattern Recognition 2002*, Vol. 4, 302-305, 2002.

[13] A. Levey, and M. Lindenbaum, "Sequential Karhunen-Loeve basis extraction and its application to images," *Image Processing, IEEE Transactions on*, Vol. 9, Issue 8, 1371 - 1374, Aug 2000.

[14] B. Moghaddam, and A. Pentland, "Probabilistic visual learning for object detection," *Proc. IEEE International Conference on Computer Vision*, pp. 786 - 793, 1995.

[15] Q. Zhao, B. Liang, and F. Duan, "Combination of Improved PCA and LDA for Video-based Face Recognition." *Journal of Computational Information Systems*, 9.1, pp. 273-280, 2013.

[16] G. Aggarwal, A. K. R. Chowdhury, and R. Chellappa, "A system identification approach for video-based face recognition," *17th International Conference on Pattern Recognition, ICPR 2004*, Vol. 4. IEEE, 2004.

[17] X. Liu and T. Chen, "Video-based face recognition using adaptive hidden markov models," *Proc. Computer Vision and Pattern Recognition, 2003*, Vol. 1, pp. I-340.

[18] R. Wang, S. Shan, X. Chen, and W. Gao, "Manifold-manifold distance with application to face recognition based on image set," *Proc. Computer Vision and Pattern Recognition, 2008*, pp. 1-8.

[19] U. Park, H. Chen, and A. K. Jain, "3D model-assisted face recognition in video," *Proc. Computer and Robot Vision, 2005*, pp. 322-329.

[20] Honda UCSD Video Database, Available: <http://vision.ucsd.edu/content/honda-ucsd-video-database>, 2005, [Accessed: Oct 22, 2013]

저 자 소 개



김진율(Jin-Yul Kim)

1986년 : 서울대학교 전자공학과 학사
 1988년 : 한국과학기술원 전기 및 전자공학과 공학석사
 1993년 : 한국과학기술원 전기 및 전자공학과 공학박사
 1994년~현재 : 수원대학교 전자공학과 부교수

관심분야 : 영상 내 객체추적, 신호처리시스템 설계,
 Phone : +82-31-220-2323
 E-mail : jykim@suwon.ac.kr



김용석(Yong-Seok Kim)

2011년 : 수원대학교 전자공학과 학사
 2011년 ~ 현재 : 수원대학교 전자공학과 석사 수료

관심분야 : 객체추적, 얼굴인식,
 Phone : +82-31-220-2531
 E-mail : ameba413@nate.com