

---

# PCA 와 KNN를 이용한 3차원 물체인식

## Three Dimensional Object Recognition using PCA and KNN

---

이기준

광주보건대학 의약정보관리학과

Kee-Jun Lee(leekj@ghc.ac.kr)

---

### 요약

기존의 주성분 분석을 이용한 물체 인식 기술은 모델 영상내의 각각의 물체의 대표 값을 만든 후에 실험 영상을 물체 공간에 투영 시켜서 나온 성분과 대표 값의 거리를 비교하여 인식하게 된다. 그러나 단순히 기존의 방법인 point to point 방식인 단순 거리 계산은 오차가 많기 때문에 본 논문에서는 개선된 Class to Class 방식인 k-Nearest Neighbor을 이용하여 몇 개의 연속적인 입력영상에 대해 각각의 모델영상들을 인식의 단위로 이용하였다. 또한, 물체 인식을 하는데 있어 본 논문에서 제안한 주성분 분석법은 물체 영상 자체를 계산하여 인식하는 게 아니라 물체 영상 공간이라는 고유 공간을 구성한 후에 단지 기여도가 큰 5개의 벡터로만 인식을 수행하기 때문에 자원 축소의 효과까지 얻을 수 있었다.

■ 중심어 : | 영상 인식 | 주성분 분석 | 최 인접 보간법 |

### Abstract

Object recognition technologies using PCA(principal component analysis) recognize objects by deciding representative features of objects in the model image, extracting feature vectors from objects in a image and measuring the distance between them and object representation. Given frequent recognition problems associated with the use of point-to-point distance approach, this study adopted the k-nearest neighbor technique(class-to-class) in which a group of object models of the same class is used as recognition unit for the images in-putted on a continual input image. However, the robustness of recognition strategies using PCA depends on several factors, including illumination. When scene constancy is not secured due to varying illumination conditions, the learning performance the feature detector can be compromised, undermining the recognition quality. This paper proposes a new PCA recognition in which database of objects can be detected under different illuminations between input images and the model images.

■ keyword : | Mage Recognition | Principal Component Analysis | K-Nearest-Neighbor |

---

## 1. 서론

물체 인식은 컴퓨터 비전 분야에서 최근 가장 활발히 연구되고 있는 분야중의 하나이다[1-3]. 물체 인식

을 정의하는 데에도 여러 분류가 있을 수 있지만 간단하게 정의하자면 미리 데이터베이스를 가지고 있는 물체에 대해서 영상 안에 그 물체가 있는지 없는지를 밝

접수번호 : #090518-005

접수일자 : 2009년 05월 18일

심사완료일 : 2009년 06월 01일

교신저자 : 이기준 e-mail : leekj@ghc.ac.kr

혀내는 것이라 할 수 있다. 이러한 정의에 의해서 크게 두 가지로 구분될 수 있는데 한 가지는 배경에 클러터(clutter)가 거의 없다고 가정하고 획득되는 영상은 한 물체에 대해서만 찍히는 것이다. 그러므로 한 영상 안에 데이터베이스에 있는 물체가 존재하는 지만 알아내면 되는 것이다[2]. 다른 방법으로는 영상 안에 다양한 클러터가 존재해서 이러한 클러터에도 불구하고 데이터베이스에 있는 물체가 존재하고 있고 또한 그 부분이 영상의 어느 영역에 있는 지까지 밝혀내는 것이다. 본 논문에서는 테스트 영상이 하나의 물체만을 포함하고 있고 물체 영역 이외에는 특별한 클러터를 가지고 있지 않은 영상을 사용하였다[3]. 물체 인식은 그 대상 물체에 따라서 크게 2D 물체 인식과 3D 물체 인식으로 나눌 수 있다. 2D 물체 인식은 물체의 단편만이 인식이 가능하기 때문에 대부분의 평평한 면을 가지는 물체들이다[4][5]. 이에 반해 3차원 인식은 물체의 모든 view에 대해서 인식이 가능한 방법이다. 2차원 인식에 비해서 대상 물체의 제한이 없는 것이 다른 점이라고 할 수 있겠다. 모든 view에 대해서 인식이 가능해야 하기 때문에 일반적으로 2차원 물체 인식보다 잘 풀리지 않는 문제이며 다양한 방법으로 접근되고 있다[6-8].

본 논문에서는 이러한 3차원 물체 인식을 하기위해 주성분 분석이라는 방법을 이용하여 물체에 대한 5°씩 회전을 시켜서 영상을 찍은 후에 그 사이에 들어온 입력 영상에 대해서는 그 주변의 트레이닝 영상들을 이용해서 그룹단위로 인식을 수행한다. 주성분 분석을 이용한 인식은 픽셀 값들의 분포를 이용하기 때문에 물체의 회전(rotation), 평행 이동(translation), 크기 변화(scale), 조명변화에 대해서 매우 민감하다. 그러므로 많은 트레이닝 이미지가 필요하게 되고 영상의 크기를 모두 일정하게 맞추는 정규화 과정이 필요하게 된다[9].

모델 영상내의 각각의 물체의 대표 값을 만든 후에 실험 영상을 물체 공간에 투영 시켜서 나온 성분과 대표 값의 거리를 비교하는 기존의 방법인 point to point 방식은 단순 거리 계산을 계산하기 때문에 오차가 많아 실시간 인식 시 인식 저하를 발생 시킨다[10]. 본 논문에서는 개선된 Class to Class 방식인 K-Nearest Neighbor을 이용하여 몇 개의 연속적인 입력영상에 대

해 각각의 모델영상들을 인식의 단위를 그룹으로 이용하여 인식 오차를 줄일 수 있었다.

## II. 전체 시스템 구성도

본 논문에서는 실시간으로 들어오는 영상으로부터 크기와 밝기를 정규화 한 후 이 영상을 주성분 분석법을 가하여 물체를 인식한다. 또한 물체 고유공간을 구성하여 기존의 단순한 거리 계산이 아닌 개선된 방법인 K-Nearest Neighbor을 이용하여 인식을 수행한다.

제안된 알고리즘의 전체 구성도를 [그림 1]에 나타내었다.

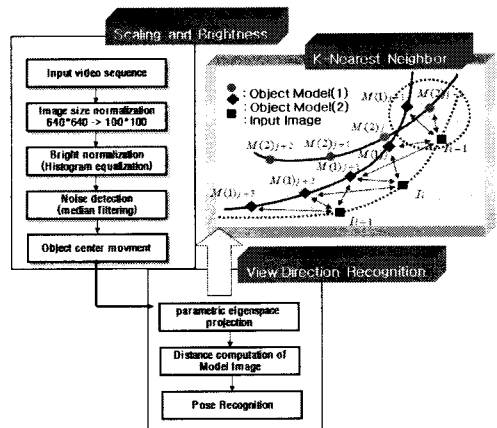


그림 1. 알고리즘 전체 구성도

## III. 물체 영상의 정규화

본 논문에서는 히스토그램 평활화를 이용하여 조명변화를 정규화 한 다음 잡음 제거를 위해 미디언 필터를 사용함으로써 일정한 영상을 얻을 수 있었다. 그림 2는 히스토그램 평활화의 효과를 나타낸 것이다. (a)(b)(c)는 원영상과 밝은 조명아래서의 영상 그리고 어두운 조명아래서의 영상이다. (d)(e)(f)는 각 영상의 히스토그램 평활화를 수행 한 후의 모습이다. 처리 후의 영상은 모두 비슷하게 나타남을 볼 수 있다.

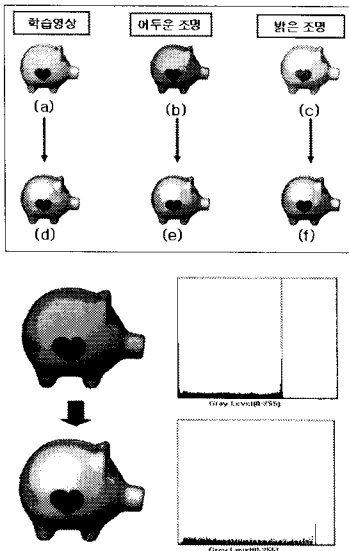


그림 2. 조명의 세기가 다른 영상과 히스토그램 평활화를 적용한 후의 영상

평활화를 적용한 결과 영상에서 노이즈로 인해 물체 형상이 매끄럽지 않은 현상이 생긴다. 이를 해결하기 위해 식(1)을 이용하여 [그림 3]처럼 물체의 노이즈를 제거하여 개선된 물체 형상을 얻을 수 있었다

$$med(x_n) = \begin{cases} x_{n-1}, & n = 2v+1 \\ \frac{1}{2}(x_n + x_{n-1}), & n = 2v \end{cases} \quad \text{식(1)}$$

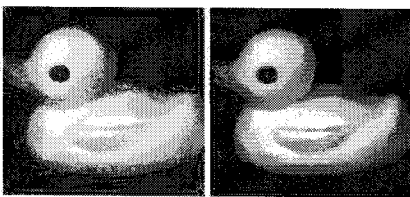


그림 3. 물체 영역 잡음 제거 전, 후 영상

#### IV. 주성분 분석을 이용한 물체인식

단일 카메라를 이용하여 물체의 방향이 틀어지는 경우를 같은 형상으로 인식하기에는 많은 어려움이 있다. 본 논문에서는 물체의 방향성이 틀어지는 경우[그림 5]에도 같은 형상으로 인식하기 위해 주성분 분석을 이용한 물체 방향성 문제의 해결 방법을 제시한다. 3장에서

설명한 방법을 통하여 얻어진 영상을 이용하여 물체형상의 전체적인 외관 특징을 표현할 수 있는 저차원 벡터 공간, 즉 파라메트릭 공간을 생성한다. 이 공간은 주성분 분석법이라는 통계적 방법에 의해 만들어진다.

주성분 분석이란 각 물체 형상 영상의 데이터들의 공간적 위치 값이 주로 각 영상에서 어디에 분포하는가를 계산하여 확률 빈도가 높은 벡터 값들을 고유치 값에 비례하여 재구성하는 방법이다. 따라서 이 방법은 고유 벡터(eigenvector)와 고유치(eigenvalue)를 계산하여, 물체 형상의 전체의 평균 모델을 구하여 이용한다. [그림 4]는 물체인식에 사용한 물체들이고 [그림 5]는 한 가지 물체를 5°씩 틀어서 획득한 전체 영상이다.(1개의 포즈 = 72 image).

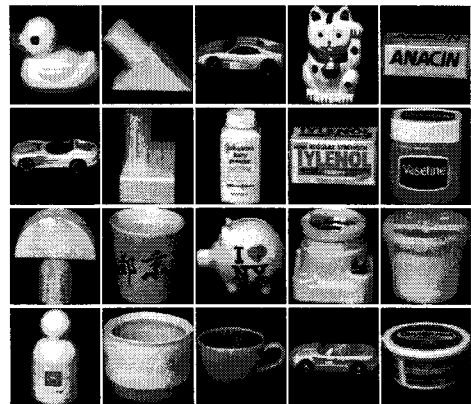


그림 4. 물체 인식에 사용한 물체 영상

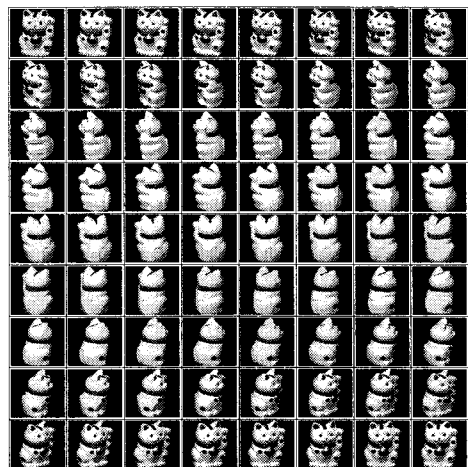


그림 5. 카메라로부터 물체를 5° 씩 틀어서 획득한 일부 영상

### 1. 주성분 분석을 이용한 고유공간 구성

고유벡터를 계산하기 위해서는 먼저 모든 영상의 평균 데이터를 구하여 각 영상들의 데이터의 차를 구한다. 평균 데이터 C와 새로운 영상 데이터 집합 X를 식(2)와 식(3)과 같이 나타낸다.

$$C = (1/N) \sum_{i=1}^N x_i \quad \text{식(2)}$$

$$X = \{x_{1,1}^{(1)} - c, x_{2,1}^{(1)} - c, \dots, x_{R,1}^{(1)} - c, \dots, x_{R,L}^{(P)} - c\} \quad \text{식(3)}$$

고유공간을 구하기 위해서는 M \* N의 크기를 지닌 영상 집합 X를 식(4)과 같이 계산하고 식(5)을 만족하는 고유벡터를 구하면 된다.

즉 공분산 행렬 Q에 대한 고유치  $\lambda$ 와 고유벡터 e를 구한다.

$$Q = XX^T \quad \text{식(4)}$$

$$\lambda_i e_i = Qe_i \quad \text{식(5)}$$

이런 과정으로 얻어진 행렬 중 공분산의 고유 벡터로 이용되는 행렬은 공분산 행렬 X와 크기가 일치하는 U이다. 특이치 분해 과정에서 나온 고유벡터를 고유치가 큰 순서대로 재구성한다. 각 고유벡터가 지닌 고유치의 크기는 그 고유벡터의 중요도를 의미하므로 그 고유공간을 규정하는 중요 고유벡터를 식(6)을 이용하여 선택한다. 따라서 모든 고유벡터를 고유 공간 구성에 이용하지 않고 많은 영상을 대표할 수 있는 주성분의 벡터만을 이용할 수 있다.

$$\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^N \lambda_i} \geq T_1 \quad \text{식(6)}$$

여기서  $T_1$ 은 고유벡터의 개수를 조정하는 문턱 치

이며, 인식과 포즈 평가 시 이용되는 고유벡터는 저차원 공간을 구성하기 위해 실험에서는  $k = 5$ 를 이용했다. 이렇게 구성된 공간을 물체 고유 공간으로 표현한다.

### 2. 물체 영상 공간에서 상관관계와 거리

앞 절에서 얻어진 물체 고유 공간에서 평균 영상 C에서 뺀 영상 집합 x를 모두 식(7)을 이용하여 투영시킨다.

$$f_j = [e_1, e_2, e_3, \dots, e_k]^T (x_n - c) \quad \text{식(7)}$$

구해진  $f_j$  물체 고유 공간상에서 점들로 표현되는데 이러한 점들을 물체 포즈 특징 심볼을 구한 것과 같이 입력 특징 심볼을 구성하게 된다. 투영시킨 결과는 이산적인 점들로 표현되며, 이들 각 점은 입력 물체 하나 하나를 의미한다. 고유공간의 특성상 비슷한 특징 값을 가지는 벡터는 고유공간에서도 가까운 곳에 투영된다. 따라서 같은 물체의 경우 서로 비슷한 특징 벡터 값을 가지게 되고, 고유공간에서 비슷한 영역에 분포하게 된다. 6개의 물체포즈(회전한 물체)를 본 고유공간에 투영시킴으로서 모델 동작을 모델링 한다. 각 물체 포즈에 대한 고유공간내의 분포는 [그림 6]과 같다.

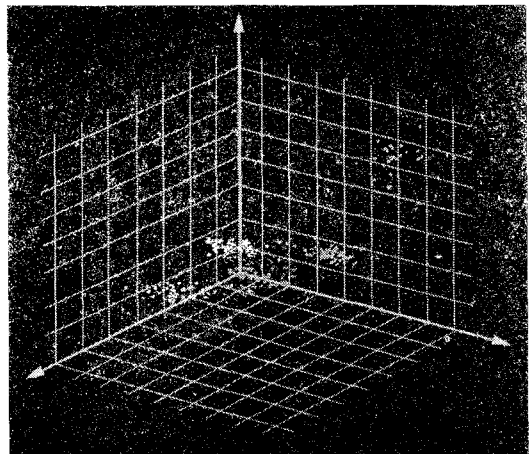


그림 6. 고유공간에서의 물체 분포

고유 공간에 투영된 점이 가까울수록 영상들의 높은 상관관계를 가진다.

### 3. 개선된 K-Nearest Neighbor를 이용한 거리 평가와 물체 인식

투영된 입력 영상이 모델 영상과의 매칭의 경우 (Point to Point) 여러 동작들이 투영되어 있는 물체 영상 공간에서 실제 물체 영상은 매칭에 성공했을지라도 다른 물체 영상으로 인식하는 잘못된 매칭이 발생하였다. 이러한 문제를 해결하기 위해 단일 물체 영상 단위의 매칭이 아닌 몇 개의 연속적인 입력 영상에 대해 각각 몇 개의 모델 영상들을 인식의 단위로 이용하였다 (Class to Class).

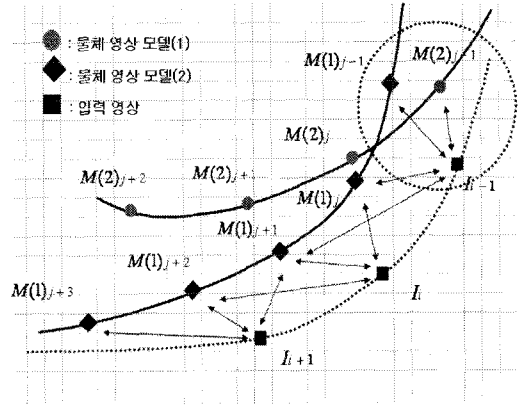


그림 7. 개선된 KNN을 이용한 물체 영상 인식

$$w = \frac{Max(arg S(M_j)) - Min(arg S(M_j))}{d(k-1)} \quad \text{식(8)}$$

매칭 알고리즘은 식(8), 식(9)에서 나타낸 것처럼 k-Nearest Neighbor 방법을 이용하였다. 여기서  $arg S(M_j) = j$ 는 모델의 번호를 구하는 연산자이다.

$$\frac{\sum \sum w(I_i - M_j)}{k} \quad \text{식(9)}$$

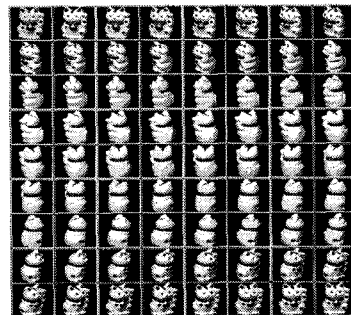
위의 식(7)에 의해 구해진 값을 이용하여 모델 영상과 입력 영상의 인식을 결정한다. 여기서는  $k=5$ 을 이용하였다. [그림 7]은 공간상에 투영된 연속적인 입력 영상들과 모델 영상들의 위 방법을 이용한 매칭 방법을 나타낸 것이다. [그림 7]에서 나타나듯이 입력 영상과 모델 영상 간 가장 가까운 거리를 가지는 영상일지라도 실제 다른 물체 영상일 경우가 있다는 것을 보여준다.

이러한 문제를 해결하기 위해 본 논문에서 제안한 방법은 두 점간의 거리로만 판단하지 않고 연속적인 몇 개의 영상 단위로 매칭을 수행함으로써 물체 영상 전체인 레적간의 일치성까지 알 수 있으며, 이를 이용하여 얻어진 결과는 물체 영상 전체를 평가하는 데도 이용할 수 있었다.

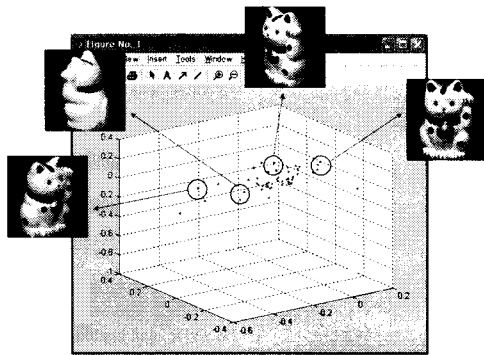
### V. 물체 인식에 대한 실험결과

실험에 이용된 물체 영상은 [그림 8](a)에서처럼 물체의 5°씩 회전한 영상을 카메라로 촬영하여 획득한 것이다. 각 물체들이 0° ~ 360°씩 회전한 물체의 집합을 한 물체 영상이라 한다. [그림 8](b)는 고유공간에서 고양이 물체가 매칭 되는 과정을 보여주고 있다.

물체 인식에 대한 실험은 640\*480으로 촬영한 물체 영상의 크기 정규화를 거쳐 100\*100 영상으로 변환하였다. 물체 영상 집합의 고유벡터를 계산한 후 재구성된 물체 영상을 가장 잘 복원하는 8차원의 벡터를 선택하여 물체 영상 공간으로 이용하였다. 따라서 100\*100 = 10,000 차원의 이미지가 8차원으로 압축되는 효과도 거둘 수 있었으며, 실제 인식 처리 과정은 실시간 처리에도 적합하다는 것을 알 수 있었다.



(a) 고양이 물체의 모델 영상



(b) 고유 공간에서의 고양이 물체의 매칭 과정

그림 8. 고양이 물체의 방향이 틀어지는 경우의 모델 영상과 매칭 영상

[표 1]은 각 매칭 방법별 매칭 성공률에 대한 분석 결과를 나타내었다. [표 1]에서 나타나듯이 개선된 k-Nearest Neighbor를 이용한 매칭 방법이 기존의 최소 거리 매칭 방법보다는 매칭 성공률이 높음을 알 수 있다. 특히 잘못된 매칭에 대해서는 많은 개선률을 나타내었다.

표 1. 매칭 방법에 따른 매칭 성공률

매칭 방법	입력 영상	매칭 실패	잘못된 매칭	매칭 성공
기존의 거리계산 (Point to Point)	모델로 구성된 경우	10.5 %	11 %	78.5 %
	모델로 구성되지 않은 경우	15.8 %	20.2 %	62 %
개선된 k-Nearest Neighbor (Class to Class)	모델로 구성된 경우	6.1 %	3.7 %	90.2 %
	모델로 구성되지 않은 경우	13.2 %	16.8 %	70 %

## VI. 결 론

본 논문에서는 주성분 분석법을 이용하여 물체 영상 집합으로부터 고유벡터를 구하고 기여도가 큰 벡터만으로 구성된 고유 공간을 이용하여 물체를 인식하는 방법을 제안하였다. 제안된 방법은 물체의 구분뿐만 아니라 구체적인 정보인 물체의 크기나 빠르기의 정보를 얻어내는데 유용하다는 것을 알 수 있었다.

기존의 기하학적 정보를 이용하거나 스테레오 영상을 이용하는 방법에 비해 처리가 간단하고 안정된 결과를 얻을 수 있음이 실험을 통하여 증명되었다. 또한 2차원 물체 영상을 가지고 3차원 물체 인식을 수행하기 위하여 물체를 5°씩 회전하여 모델 영상을 구성함으로써 물체의 단면만을 인식하는 게 아니라 물체 영상의 단위인 그룹 단위로 인식을 수행하여 3차원 물체 인식도 가능하였다. 그러나 3차원 인식 수행 후 물체가 90° 되었을 경우 잘못된 물체로 인식하는 경우가 종종 발생하였고, 복잡한 배경에서는 물체 영역만을 분리하는데 어려움이 있었다. 향후에는 이러한 문제점을 해결하여 보다 안정적인 3차원 물체 인식 알고리즘을 개발할 계획이다.

## 참 고 문 헌

- [1] Hiroshi Murase and Shree K. Nayar, "Visual Learning and Recognition 3-D object from appearance", international journal of Computer Vision, Vo.14, 1995.
- [2] S. J. Dickinson, A. P. Pentland, and A. Rosenfeld, "3-D Shape Recovery Using Distributed Aspect Matching," IEEE trans. on PAMI Vol.14, No.2, pp.174-200, 1992(2).
- [3] P. L. Rosin and G. A. W. West, "Nonparametric Segmentation of Curves into Various Representations," IEEE trans. on PAMI, Vol. 17, No.12, pp.1140-1153, 1995(12).
- [4] Shigeyoshi Hiratsuka, Kohtarō Ohba, Hikaru Inooka, Shinya Kajikawa, and Kazuo Tanie, "Stable Gesture Verification in Eigen Space," LAPR Workshop on Machine Vision Application, 1998(11).
- [5] J. Weng, N. Ahuja, and T. S.Huang, "Learning recognition and segmentation of 3-D object from 2-D images," Proc. of Fourth Int'l Conf. on Computer Vision, pp.121-128, Belin, 1993(5).

- [6] Ragael C. Gonzalez and Richard E. Woods  
"Digital Image Processing," Prentice Hall, 2002.
- [7] V. Pavlovic, R. Sharama, and T. Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review," IEEE Trans. on PAMI, Vol.19, No.7 pp.677-695, 1997.
- [8] J. Segen and S. Kumar, "Shadow Gestures: 3D Hand Pose Estimation Using a Single Camera," CVPR99, Vol.1, pp.479-485, Fort Collins, Colorado, 1999(6).
- [9] J. M. Kim and M. K. Song, "Three Dimensional Gesture Recognition Using PCA of Stereo Images and Modified Matching Algorithm," IEEE Fuzzy Systems and Knowledge Discovery Vol.4, pp.116-120, 2008(10).
- [10] X. Liang, W. Zhihui, and W. Huizhong, "Robust Orientation Diffusion Via PCA Method and Application to Image Super-Resolution Reconstruction," Lecture Notes in Computer Science LNCS.4688 pp.726-735, 2007.
- [11] T. Jilin, T. Huang and T. Hai, "Face as mouse through visual face tracking," IEEE Computer and Robot Vision, 2005. Proceeding. The 2nd Canadian Conference on 9-11 May, 2005.
- [12] Y. Zhang and C. Liu, "Face recognition using kernel principal component analysis and genetic algorithms," Proceedings of the 2002 12th IEEE Workshop on Neural Networks for Signal Processing, pp.337-343, 2002.

## 저자 소개

이 기 준(Kee-Jun Lee)

정회원



- 1994년 2월 : 조선대학교 전산통계학과(이학사)
- 1997년 2월 : 조선대학교 전산통계학과(이학석사)
- 2001년 2월 : 조선대학교 전산통계학과(이학박사)

▪ 2002년 3월 ~ 현재 : 광주보건대학 의약정보관리과 부교수

<관심분야> : 영상처리, 내용기반검색, 멀티미디어