

PCA와 개선된 k-Nearest Neighbor를 이용한 모델 기반형 물체 인식

정 병 수[†] · 김 병 기^{††}

요 약

주성분 분석법을 사용한 물체 인식 기술은 영상의 조명 변화가 있을 때 인식률이 떨어지는 경향이 있다. 본 논문에서는 실험영상이 학습영상에 대해 조명의 차이가 있는 경우에도 데이터베이스안의 물체인지 가려내는 새로운 PCA 분석방법을 사용한 물체 인식 기술을 제안하는데 그 목적이 있다. 그리고 개선된 k-nearest neighbor를 이용하여 물체 인식률을 향상 시켰다. 본 논문에서 제안된 물체 인식 알고리즘은 히스토그램 이퀄라이제이션과 미디언 필터를 이용하여 영상을 전처리하고 그것을 학습시켜서 물체 공간을 생성한다. 이때 히스토그램 이퀄라이제이션을 사용하여 히스토그램을 펼침으로써 조명 변화에 영향을 감소시키는 결과를 내었고, 이것은 기본적인 주성분 분석방법과 휘도치 정규화를 한 방법 등과 비교해 본 결과 조명 변화의 영향을 최소화하여 좋은 인식률을 유지할 수 있었다. 그리고 모델 영상내의 각각의 물체의 대표 값을 만든다. 그런 후 테스트영상을 물체 공간에 투영 시켜서 나온 성분과 대표 값의 거리를 비교하여 인식하게 된다. 기존의 방식으로는 거리 계산 오차가 많기 때문에 본 논문에서는 개선된 k-Nearest Neighbor를 이용하여 몇 개의 연속적인 입력영상에 대해 각 각의 모델 영상들을 인식의 단위로 이용하였다.

키워드 : 주성분 분석, 물체인식, 개선된 k-Nearest Neighbor, 히스토그램 이퀄라이제이션

Model-Based Object Recognition using PCA & Improved k-Nearest Neighbor

Byeong-Soo Jung[†] · Byung-Gi Kim^{††}

ABSTRACT

Object recognition techniques using principal component analysis are disposed to be decreased recognition rate when lighting change of image happens. The purpose of this thesis is to propose an object recognition technique using new PCA analysis method that discriminates an object in database even in the case that the variation of illumination in training images exists. And the object recognition algorithm proposed here represents more enhanced recognition rate using improved k-Nearest Neighbor. In this thesis, we proposed an object recognition algorithm, which creates object space by pre-processing and being learned image using histogram equalization and median filter. By spreading histogram of test image using histogram equalization, the effect to change of illumination is reduced. This method is stronger to change of illumination than basic PCA method and normalization, and almost removes effect of illumination, therefore almost maintains constant good recognition rate. And, it compares ingredient projected test image into object space with distance of representative value and recognizes after representative value of each object in model image is made. Each model images is used in recognition unit about some continual input image using improved k-Nearest Neighbor in this thesis because existing method have many errors about distance calculation.

Key Words : PCA, Object Recognition, Improved k-Nearest Neighbor, Histogram Equalization

1. 서 론

물체 인식 문제는 영상으로부터의 물체검출, 물체의 특징 추출과 물체 인식이라는 세부 주제로 다시 나뉘어 연구되어왔다[1]. 1990년대 이후의 물체 인식분야는 크게 PCA를

이용한 방법론, 동적 링크 구조를 이용한 방법론, 신경회로망을 이용한 방법론 등으로 나뉘어서 독립적으로 연구되고 있다[2]. PCA를 이용한 방법론은 물체 인식 기술에서 가장 효율적인 방법으로 입증되어져 왔다[3, 4]. 본래 PCA를 이용한 방법론은 데이터베이스상의 물체로 가장 이 물체를 잘 표현하는 기저벡터(eigenvector)를 만들어 물체 인식에 필요한 영상을 이 기저벡터에 투사시켜 이때 나온 인자들과 원래 각 물체의 대표 인자 값과 거리 비교로 물체를

† 정 회 원 : 남부대학교 디지털정보학과 교수
 †† 종신회원 : 전남대학교 전자컴퓨터정보통신공학 교수
 논문접수 : 2005년 9월 22일, 심사완료 : 2005년 12월 21일

인식하는 방법이다. 그러나 이 방법은 조명이 변하면 투사된 인자 값들도 크게 변한다. 따라서 인식률이 떨어지고 속도에 문제점이 있다[5, 6]. 본 연구에서 제안하고자 하는 물체 인식 알고리즘은 위의 알고리즘을 기초로 하여 조명 변화의 영향을 최소화하기 위해서 히스토그램 이퀄라이제이션과 미디언 필터를 이용하여 영상을 전처리하고 픽셀 값들의 분포를 비슷하게 하여 히스토그램 분포를 유사하게 만들었다. 그리하여, 기존의 PCA를 이용한 물체 인식 기술을 측면에서 비교하였다. 또한 투영된 입력 영상이 모델 영상과의 매칭의 경우 여러 포즈들이 투영되어 있는 물체 영상 공간에서 실제 물체 영상은 매칭에 성공했을지라도 다른 물체 영상으로 인식하는 잘못된 매칭이 발생한다[7]. 이러한 문제를 해결하기 위해 단일 물체 영상 단위의 매칭이 아닌 몇 개의 연속적인 입력 영상에 대해 각각 몇 개의 모델 영상들을 인식의 단위로 이용하는 개선된 k-Nearest Neighbor 방법을 이용하였다. 본 논문에서는 물체에 대해 5°씩 회전을 시켜서 물체 영상을(72개 영상) 찍은 후에 그 사이에 들어온 입력 영상에 대해서는 그 주변의 학습 영상들을 이용해서 그룹단위로 인식을 수행한다.

2. 모델 기반형 물체 인식

본 논문에서 제안한 모델 기반형 물체 인식의 주요 단계는 첫 번째로는 물체에 대해 5°씩 회전을 시켜서 물체 영상을 획득 단계이고, 두 번째는 획득한 물체 영상의 크기를 100×100으로 축소하는 크기 정규화 단계이다. 세 번째로는 정규화 영상의 히스토그램 이퀄라이제이션을 이용하여 밝기 변화에 강한 정규화 단계이고, 네 번째는 밝기 정규화된 영상을 잡음 제거와 영상의 위치를 중앙에 위치하기 위한 중심 이동 단계이다. 다섯 번째로는 앞의 단계를 통하여 얻어진 영상 집합을 주성분 분석을 이용한 고유 공간을 구하는 물체 영상 공간으로 구성하는 단계이며, 여섯 번째로는 물체를 인식하고 회전 정보를 추정하는 단계로 이루어진다.

2.1 물체 영상 집합의 정규화 과정

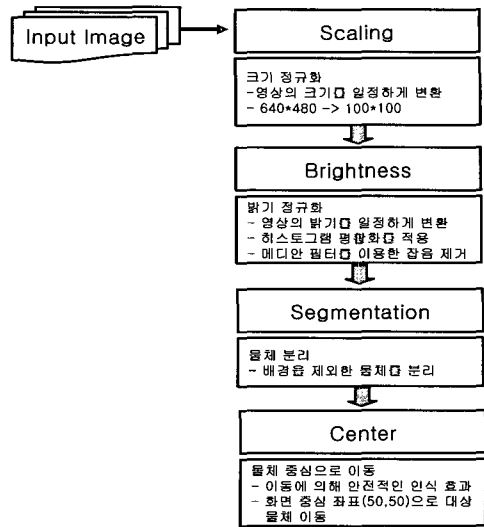
물체 영상 집합의 정규화는 카메라로부터 연속적으로 얻어진 영상들을 모델구성에 필요한 전처리 과정으로 시각 기반의 물체 인식에 필수적인 처리 과정이다. 특히 일반 환경에서 연속적 영상들은 각각 서로 다른 많은 잡음들을 포함하기 때문에 정규화 처리에 의해 인식률이 크게 좌우 될 수 있다. (그림 1)은 전처리에 대한 전체 흐름도이다.

2.1.1 물체 영상 집합의 크기 정규화

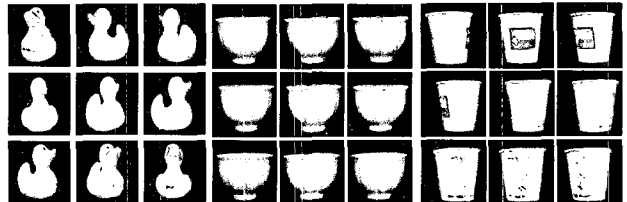
먼저 카메라로부터 촬영하여 (그림 2)와 같은 영상을 얻는다. 이러한 영상의 집합을 \hat{X} 라 하고 식(1)과 같이 표현한다.

$$\hat{X} = [\hat{X}_1, \hat{X}_2, \dots, \hat{X}_N]^T \quad (1)$$

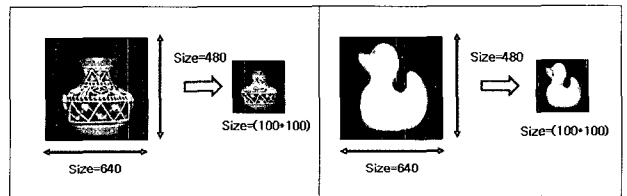
여기서 N은 물체 영상의 개수이다.



(그림 1) 전처리 전체 개략도



(그림 2) 카메라로부터 획득한 일부 영상 집합



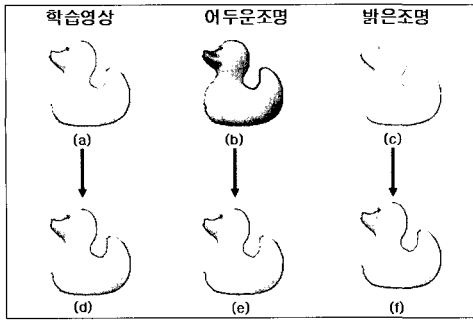
(그림 3) 원 영상과 축소된 영상

일반적으로 640 x 480의 큰 영상이므로 계산 량을 줄이기 위해 영상을 일정한 크기로 정규화 한다. 본 논문에서는 영상의 크기를 100 x 100으로 정규화 시켰다.

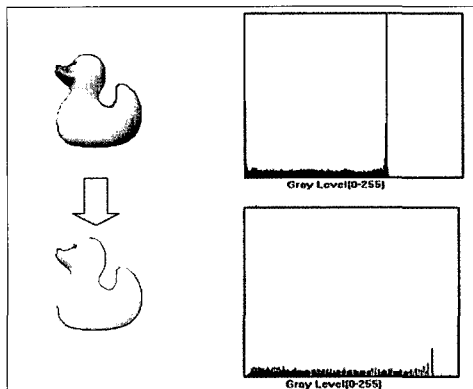
여기에서는 (그림 3)에서 보는 바와 같이 선형보간법을 이용하면 영상의 픽셀의 누설 없이 자연스럽게 축소되는 것을 알 수 있다.

2.1.2 정규화 영상의 밝기 정규화

PCA는 물체 인식 기술에서 가장 믿을 수 있는 방법이나 픽셀 값을 비교하기 때문에 조명의 변화에 따라 정확성이 떨어지는 한계점이 있다[8]. 이 한계점을 해결하기 위해서 히스토그램 이퀄라이제이션을 사용하여 해결한다. 물체 인식에서 히스토그램 이퀄라이제이션을 쓰는 이유는 이것을 실행하면 영상의 히스토그램이 펼쳐질 뿐 아니라 영상의 대조를 높여준다. 그래서 영상의 조명이 밝거나 어두우면 영상은 영상의 휘도치의 분포는 한곳에 모일 것이다. 그 경우 히스토그램 이퀄라이제이션의 사용으로 히스토그램을 펼쳐



(그림 4) 조명이 다른 영상을 히스토그램 이퀄라이제이션 후의 영상



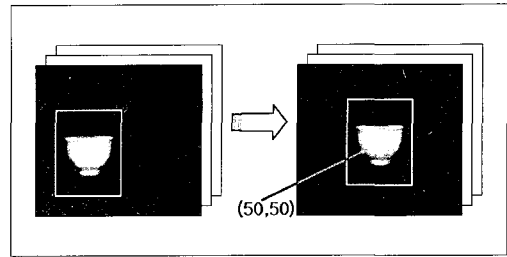
(그림 5) (그림 4)의 어두운 영상의 이퀄라이제이션 후 영상의 데이터

서 인식해야 할 물체의 특징점 찾기에 도움을 주게 된다. 특별한 조명을 사용하지 않는 경우 각 입력 영상은 조명의 밝기에 따라 인식에 큰 차이를 나타낸다. 이를 해결하기 위한 방법으로 본 논문에서는 히스토그램 이퀄라이제이션을 이용하여 밝기 변화에 강한 물체 인식을 수행한다. (그림 4)는 본 논문에서 제안한 방법을 적용하여 밝기의 변화에서도 일정한 인식을 행할 수 있다는 것을 보여준다. (그림 5)는 (그림 4)의 어두운 영상을 이퀄라이제이션 한 후 영상의 데이터를 보여준다.

2.1.3 물체 영상의 잡음 제거 및 중심 이동

밝기 정규화 된 영상에서도 노이즈가 발생하기 때문에 노이즈를 감소하기 위해 미디언 필터 방법을 이용하고, 영상의 위치가 항상 중앙에 위치해야만 좀 더 나은 인식을 가질 수 있기에 물체의 중심이동 방법에 대하여 기술을 한다. 본 논문에서 히스토그램 이퀄라이제이션을 수행한 후에 약간의 임펄스 잡음이 포함되어 있기에 이를 해결하기 위해 효과적인 방법으로 밝혀진 미디언 필터를 이용하여 잡음을 제거하였다. 영상에서 물체만을 세그멘테이션 한 영상을 얻기 위해선 식(2)에 표현한 것처럼 각 입력 영상을 미리 촬영한 배경 영상과의 차를 이용한다.

$$X(i, j) = \sum_{i=0}^N \sum_{j=0}^M [Img_i(i, j) - Img_k(i, j)] \quad (2)$$

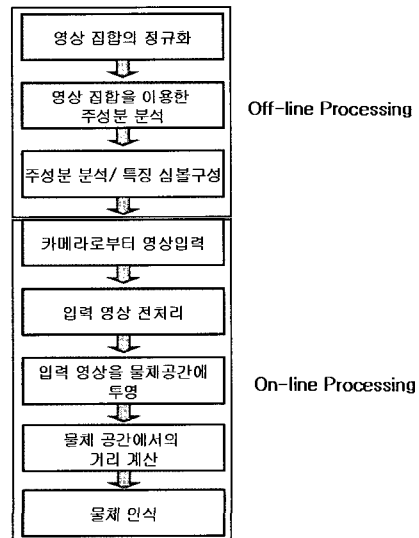


(그림 6) 물체 영상의 중심 이동

이때, $Img_i(i, j)$ 는 입력 영상이고, $Img_k(i, j)$ 는 배경 영상이며 $X(i, j)$ 는 입력 영상에서 배경 영상을 뺀 값으로 이루어진 영상이다. 이렇게 구해진 차 영상을 안정적인 인식 영상으로 사용하기 위해 (그림 6)과 같은 방법으로 물체를 100×100의 화면 중심(50, 50)으로 이동시킨다.

2.2 정규화 영상 집합을 이용한 물체 영상 공간 구성

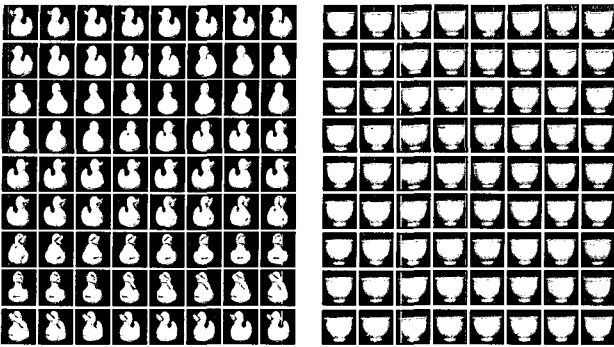
앞의 절에서 설명한 방법을 통하여 얻어진 영상들을 이용하여 물체의 전체적인 외관 특징을 표현할 수 있는 저차원 벡터 공간인 파라메트릭 물체 공간을 생성한다. 이 공간은 주성분 분석법이라는 통계적 방법에 의해 만들어진다. 아래 (그림 7)은 물체 영상 공간에서 전체 매칭과정을 나타낸 것이다.



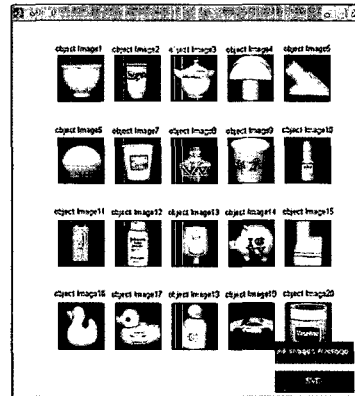
(그림 7) 물체 영상 공간을 이용한 전체 매칭 과정

2.2.1 주성분 분석법을 이용한 물체 영상의 고유 공간 구성

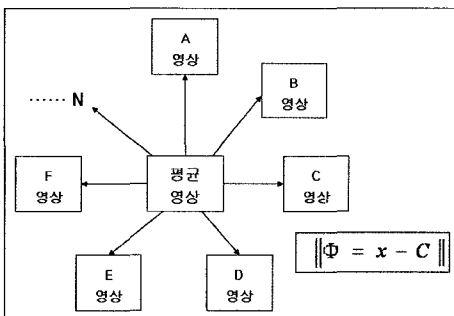
카메라를 이용하여 물체의 방향이 틀어지는 경우를 같은 형상으로 인식하기에는 많은 어려움이 있다. 따라서 본 논문에서는 물체의 방향성이 틀어지는 경우 (그림 8)과 같은 물체 영상으로 인식하기 위해 주성분 분석을 이용한 물체 방향성 문제의 해결 방법을 제안한다. 한 물체 영상 당 5°씩 회전을 해서 물체 영상을 습득하였다. 즉 한 물체는 72개의 이미지를 갖게 된다. 이 방법은 고유벡터와 고유치를 계산하여 물체 영상의 평균 모델을 구하여 이용한다. 고유 벡터



(그림 8) 물체의 방향이 틀어지는 경우의 영상



(그림 10) 20가지의 물체 영상



(그림 9) 입력 영상들과 평균 영상의 거리 관계

를 계산하기 위해서는 먼저 아래 (그림 9)에서처럼 하나의 물체 영상은 5° 틀어진 영상들의 집합이므로 72개의 이미지라 할 수 있는 모든 물체 영상들과의 차를 구한다. 그리고 (그림 10)에서 나타난 평균 영상 C와 새로운 영상 집합 X를 식(3), 식(4)와 같이 나타낸다.

식(3)을 이용하여 (그림 10)의 20개의 물체를 (그림 11)과 같이 평균 영상 C를 구성하였다. 한 물체 당 72개의 이미지를 갖고 있고, 물체 영상의 크기는 100 x 100 이다.

$$C = \left(\frac{1}{N}\right) \sum_{i=1}^N x_i \quad (3)$$

여기서 N은 물체 영상의 개수이며 x_i 는 입력 영상이다.

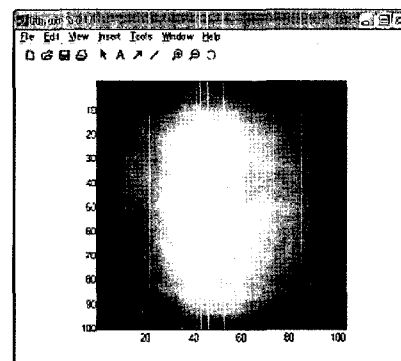
$$X = \{x_{1,1}^{(1)} - c, x_{2,1}^{(1)} - c, \dots, x_{R,1}^{(1)} - c, \dots, x_{R,L}^{(p)} - c\} \quad (4)$$

고유 공간을 구하기 위해서는 M x N의 크기를 지닌 영상 집합 X를 식(5)과 같이 계산하고 만족하는 고유벡터를 구하면 된다. 즉, 공분산 행렬 Q에 대한 고유치 λ 와 고유벡터 e를 구한다.

$$Q = \nabla XX^T \quad (5)$$

여기서 M은 한 영상의 $\lambda_i e_i = Qe_i$ 픽셀 수이고 N은 전체 영상의 개수를 나타내는 정수이다.

본 논문에서는 특이치 분해를 이용하여 고유 공간을 구하



(그림 11) 물체 영상의 평균 영상

였다. 특이치 분해를 이용하여 식(6)에서처럼 행렬이 분해되는데, 영상 집합 X의 공분산 행렬에 대한 고유 벡터를 쉽게 얻을 수 있다.

$$[X] = [V] \cdot \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \cdot [U^T] \quad (6)$$

행렬은 식(6)와 같이 분해가 가능하며 식(7)의 표현을 이용하면 원 식은 식(8)과 같이 나타낼 수 있다.

$$\begin{aligned} X &= \sum_{i=1}^r \sigma_i V_i U_i^T \\ V &= [v_1, \dots, v_r] \\ U &= [u_1, \dots, u_r] \\ \Sigma &= \text{diag}(\sigma_1, \dots, \sigma_r) \end{aligned} \quad (7)$$

여기서 행렬 X의 특이치(singular value)는 $\sigma_i = \sqrt{\lambda_i}$, $i = 1, \dots$ 로 정의된다.

$$X = V \Sigma U^T \quad (8)$$

여기서 Σ 는 대각행렬이다. V 행렬과 U 행렬은 각각 직교 행렬이기 때문에 식(9)에서 나타난 것처럼 표현된다.

$$[V^T] \cdot [V] = [U^T] \cdot [U] = [I] \quad (9)$$

이런 과정으로 얻어진 행렬 중 공분산의 고유 벡터로 이용되는 행렬은 공분산 행렬 X와 크기가 일치하는 U 행렬이다. 특이치 분해 과정에서 나온 고유벡터를 고유치가 큰 순서대로 재구성한다. 각 고유벡터가 지닌 고유치의 크기는 그 고유벡터의 중요도를 의미하므로, 그 고유 공간을 규정하는 중요 고유벡터를 식(10)을 이용하여 선택한다. 따라서 모든 고유벡터를 고유 공간 구성에 이용하지 않고 많은 영상을 대표할 수 있는 주성분의 벡터만을 이용할 수 있다.

$$\frac{\sum_{i=1}^k \lambda_i}{N} \geq T_1 \quad (10)$$

여기서 T_1 은 고유벡터의 개수를 조정하는 문턱치이며, 인식과 물체 영상 평가 시 이용되는 고유벡터 $\{e_i | i = 1, 2, 3, \dots, k\}$ 는 저차원 공간을 구성하기 위해 ($k \ll N$)을 만족하며 실험에서는 $k=3$ 을 이용하였다. 이렇게 구성된 공간을 앞으로 물체 영상 공간으로 표현한다.

2.2.2 물체 영상 공간에서 상관관계와 거리

앞 절에서 얻어진 물체 영상 공간에서 평균 영상 C에서 뺀 영상 집합 x 를 모두 $f_j = [e_1, e_2, e_3, \dots, e_k]^T (x_n - C)$ 을 이용하여 투영시킨다. 투영시킨 결과는 이산적인 점들로 표현되며, 이들 각 점은 물체 영상을 의미한다. 연속적인 점들은 서로 연관성이 많기 때문에 물체 영상 공간으로 투영시킨 결과는 서로 깊은 상관관계를 가진다.

$m(m_1, m_2, \dots, m_n)$ 와 같이 각 물체들은 서로 관계있는 연속성을 가진 점들의 집합으로 나타나게 된다.

물체 영상 공간을 계산하기 위해 이용된 영상 x_1 과 x_2 를 물체 영상 공간에 투영시켜 얻은 점들이 각각 m_1 과 m_2 라면, 이 두 점 사이의 거리와 두 영상간의 상관관계는 서로 밀접하다. 즉, 거리가 가까울수록 두 영상은 닮은 영상이 된다.

공간상에 임의의 한 점을 m_1 이라하면 원래의 물체 영상은 식(11)과 같이 복원 될 수 있다.

$$x_1 = \sum_{i=1}^N m_{1i} e_i + C \quad (11)$$

여기서 C는 전체 영상 집합의 평균 영상이고 e_i 는 전체 고유벡터를 나타낸다. k개의 물체 영상 공간에 투영한 물체 영상은 식(12)으로 표현된다.

$$x_1 \approx \sum_{i=1}^N m_{1i} e_i + C \quad (12)$$

두 물체 영상간의 유사도는 식(13)과 같은 템플레이트 매칭으로 측정된다.

$$\|x_1 - x_2\|^2 = (x_1 - x_2)^T (x_1 - x_2) = 2 - 2x_1^T x_2 \quad (13)$$

위 식을 식(12)를 이용하여 나타내면 식(14)과 같다.

$$\|x_1 - x_2\|^2 = \left\| \sum_{i=1}^k m_{1i} e_i - \sum_{i=1}^k m_{2i} e_i \right\|^2 \quad (14)$$

다시 식(14)을 풀면 식(15)로 나타낼 수 있다.

$$\left| \sum_{i=1}^k m_{1i} e_i - \sum_{i=1}^k m_{2i} e_i \right|^2 = \left| \sum_{i=1}^k (m_{1i} - m_{2i}) e_i \right|^2 = \sum_{i=1}^k \sum_{j=1}^k e_i^T e_j \times (m_{1i} - m_{2j})^2 = |m_1 - m_2|^2 \quad (15)$$

여기서 $i=j$ 일 때 $e_i^T e_j = 1$ 이고, 그렇지 않으면 0이 된다. 식(14)과 식(15)으로부터 다음과 같이 $\|x_1 - x_2\|^2 \approx \|m_1 - m_2\|^2$ 을 얻을 수 있다[12, 13].

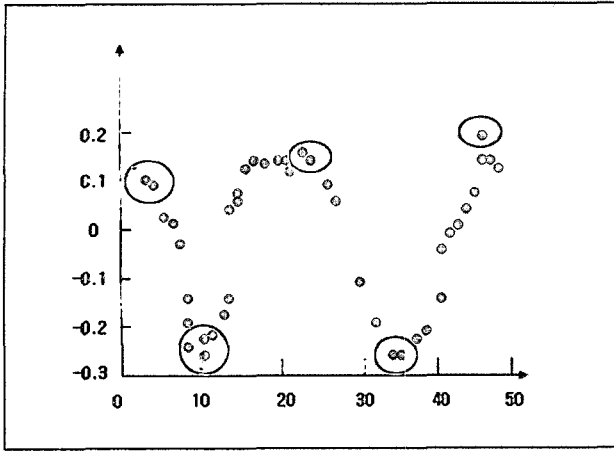
따라서 두 점간의 관계는 두 물체 영상간의 관계와 매우 유사하다는 것을 알 수 있다. 고유 공간에 투영된 점이 가까울수록 영상들의 높은 상관관계를 가진다는 것이 증명되었다.

2.2.3 물체 영상 공간에서 심볼을 이용한 모델 구성

이 절에서는 물체 영상 공간에서 여러 변화에 대해 안정된 특징을 가지는 모델을 선정하는 방법에 대하여 설명한다. 일반적으로 시각적 학습 방법을 이용한 인식은 모든 외관들과 일치하는 영상들은 모델화되어 등록되었다. 그리고 입력 영상은 등록된 모델 영상들과 비교하면서 인식을 위한 매칭이 수행된다. 효과적인 매칭을 수행하기 위해서는 매칭 횟수와 인식 속도를 개선시키는 것이 필요하다.

여기서는 각 포즈들의 모델들에 대한 특징들을 구하여 모든 모델들을 비교하지 않고 특징이 되는 일부의 특징 심볼들과 비교하여 인식률의 저하 없이 매칭 속도를 증가시켰다. (그림 12)는 e_3 에서 시간 축에 대한 물체 영상 궤적을 나타낸 것이다. 여기서 e_3 는 3차원 물체 공간 영상의 고유치이다. 이 궤적으로부터 반복된 곡선이 비슷한 형태를 가지며 국소적인 최대 값과 최소 값으로 표현되는 것을 알 수 있다. 하지만 물체 영상의 큰 변화에 대해 모든 차원에서 물체 영상을 정확하게 표현하는 곡선을 나타내지 않기 때문에 모델 영상을 가장 잘 표현할 수 있는 차원으로 새로 구성한다. 선택된 차원에 의해 구해진 모델 영상들 중 국소적으로 최대 값과 최소 값으로 나타난 모델 $S(m)$ 을 물체 영상 특징 심볼로 식(16)과 같이 구성한다.

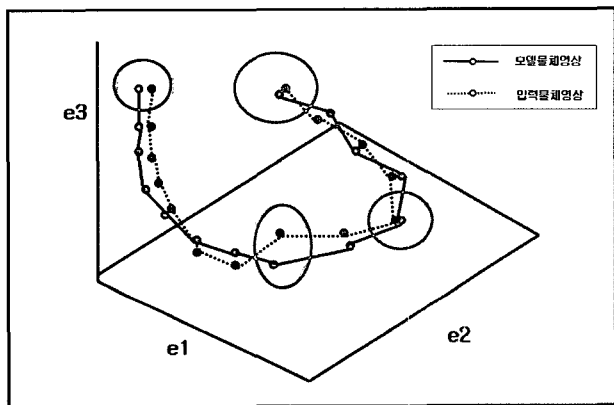
$$S(m) \equiv [s_{m,1}, s_{m,2}, s_{m,3}, \dots, s_{m,n}] \quad (16)$$



(그림 12) e_3 에서 시간 축에 대한 물체 영상 모델 심볼

2.3 물체 인식과 회전 정보 추정

연속적으로 입력되는 영상은 앞의 절에서 구성된 물체 특징 심볼들에 의해 인식된다. 물체 영상 공간에서 입력 영상들은 어떤 궤적을 그리게 되는데 이 궤적은 연속적인 입력 심볼로 표현된다. 이런 심볼들을 모델 심볼에서 물체 특징 심볼을 구한 방법으로 물체 입력 특징 중 첫 번째 심볼이 이 물체 특징 심볼내에 존재하면 심볼 매칭은 순차적으로 수행되게 된다. 연속적 입력 심볼이 완벽하게 모델 특징 심볼과 일치하면 일치된 모델 물체 영상으로 인식하게 된다. 아래 (그림 13)은 3차원 물체 공간상에서 연속적인 입력 물체 영상과 모델 물체 영상간의 매핑을 나타내고 있다. 여기서 점선으로 표현된 심볼들은 순차적으로 들어오는 입력 물체 영상이고 직선은 모델 물체 영상의 심볼들이다.



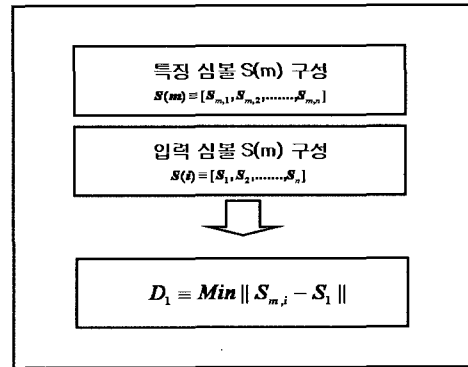
(그림 13) 물체 영상 공간에서 입력 영상과 모델 영상의 매핑

2.3.1 입력 영상과 모델 영상에 대한 매칭 방법

정규화된 영상들이 물체 영상 공간에 투영되어 인식을 위한 모델들로 정해지면 실제 인식에 필요한 처리는 매우 간단하다. 먼저 평균 영상에서 입력 영상 y 를 뺀 다음 고유 공간에 식(17)와 같이 투영하면 된다.

$$z_n = [e_1, e_2, e_3, \dots, e_k]^T (y_n - c) \quad (17)$$

구해진 z_n 물체 영상 공간상에서 점들로 표현되는데 이러한 점들을 물체 특징 심볼을 구한 것과 같이 입력 특징 심볼을 구성하게 된다. 여기서 고유벡터 $\{e_i | i = 1, 2, 3, \dots, k\}$ 는 저차원 공간을 구성하기 위해 ($k \ll N$)을 만족한다. 구성된 입력 특징 심볼들은 모델 특징 심볼들과 비교하여 매칭이 이루어진다. (그림 15)는 매칭 과정을 나타낸다. 만약 (그림 14)의 D_1 이 미리 정해진 임계치보다 적으면 입력 특징 심볼은 비교된 모델 특징 심볼 물체 영상으로 인식하게 된다. 물체 인식은 입력 영상의 잡음이나 시스템의 착오를 감안하여 임계값을 정해서 각 모델들과의 거리 중 (그림 14)에서 나타낸 최소거리를 구하여 이루어진다.



(그림 14) 물체 영상 공간에서 매칭 과정

2.3.2 개선된 k-Neast Neighbor 방법을 이용한 물체 인식

앞의 절에서처럼 투영된 입력 영상이 모델 영상과의 매칭의 경우(Point to Point) 여러 포즈들이 투영되어 있는 물체 영상 공간에서 실제 물체 영상은 매칭에 성공했는지라도 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).

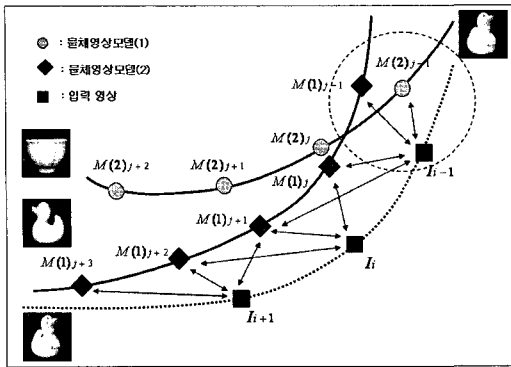
$$w = \frac{Max(arg S(M_j)) - Min(arg S(M_j))}{d(k-1)} \quad (18)$$

매칭 알고리즘은 식(18), 식(19)에서 나타낸 것처럼 k-Nearest Neighbor 방법을 이용하였다.

여기서 $arg S(M_j) = j$ 는 모델의 번호를 구하는 연산자이다.

$$\frac{\sum \sum w(I_i - M_j)}{k} \quad (19)$$

위의 식(19)에 의해 구해진 값을 이용하여 모델 영상과 입력 영상의 인식을 결정한다. $\sum \sum$ 는 공간상에 투영된 연속적인 입력 영상들과 모델 영상들의 거리를 의미하고, I_i 는 입력영상, M_j 는 모델영상을 나타낸다. 여기서는 $k=3$



(그림 15) 개선된 k-Nearest Neighbor 방법을 이용한 물체 영상 인식

을 이용하였다. (그림 15)는 공간상에 투영된 연속적인 입력 영상들과 모델 영상들은 이 방법을 이용한 매칭 방법을 나타낸 것이다. (그림 15)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 궤적간의 일치성까지 알 수 있으며, 이를 이용하여 얻어진 결과는 물체 영상 전체를 평가하는데도 이용할 수 있었다.

3. 실험 결과

본 논문에서는 주성분 분석과 개선된 k-Nearest Neighbor 매칭 방법을 사용해서 물체 인식 시스템을 VC++로 구현하였다. 조명의 변화에 영향을 받지 않도록 히스토그램 이퀄라이제이션을 적용하였으며, 계산량을 줄이기 위해서 선형 보간법으로 크기 정규화를, 물체의 잡음제거를 위해서 미디언 필터법을, 배경과 물체를 분리하기 위해서 차 영상 기법을 적용했으며, 또한, 투영된 입력 영상이 모델 영상과의 매칭의 경우 여러 포즈들이 투영되어 있는 물체 영상 공간에서 실제 물체 영상은 매칭에 성공했는지라도 다른 물체 영상으로 인식하는 잘못된 매칭이 발생할 수 있는데 이러한 문제를 해결하고, 물체 인식률을 개선하기 위해서 개선된 k-Nearest Neighbor 매칭 방법을 적용하였다. 모델 영상은 이미지의 크기는 100 x 100으로 사용하였다. 하나의 물체가 가지는 이미지의 개수는 5°씩 회전한 72장이며, 20개의 물체를 실험에 사용하였다. 고유 공간을 구성하는데 걸리는 시간은 대략 2분 40초가 걸렸고, 새로운 이미지가 들어와서 복원하는데 걸리는 시간은 평균 21.245 밀리 초가 걸렸다.

3.1 모델 기반형 물체인식 실험 결과

먼저 물체 인식 실험 이용된 물체 영상은 (그림 16)에서 처럼 물체를 5°씩 회전한 영상으로 카메라로 촬영하여 획득한 것이다. 각 물체들이 0°~360°씩 회전한 물체의 집합을 하나의 물체 영상이라 한다. 640 x 480으로 촬영한 물체 영



(그림 16) 찌잔 물체의 방향이 틀어지는 경우의 모델 영상

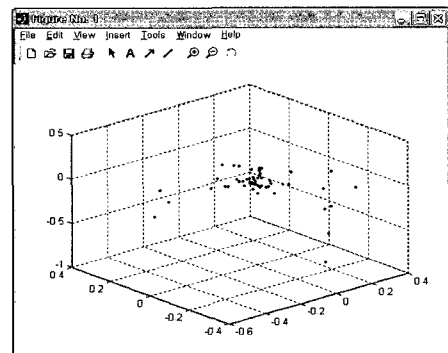
<표 1> 매칭 방법별 매칭 성공률

매칭 방법	입력 영상	매칭 실패	잘못된 매칭	매칭 성공
최소 거리 계산 (Point to Point)	모델로 구성된 경우	10.1%	11.3%	78.6%
	모델로 구성되지 않은 경우	16.8%	21.2%	62%
개선된 K-Nearest Neighbor 알고리즘 (Class to Class)	모델로 구성된 경우	6.1%	3.7%	90.2%
	모델로 구성되지 않은 경우	10.2%	12.8%	77%

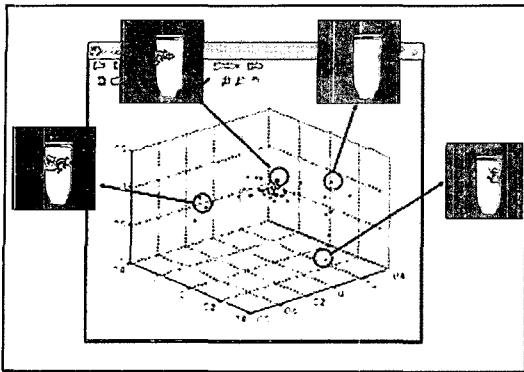
상은 크기 정규화를 거쳐 100x* 100 영상으로 크기를 축소 변환하였다. 물체 영상 집합의 고유벡터를 계산한 후 재구성된 물체 영상을 가장 잘 복원하는 16차원의 벡터를 선택하여 물체 영상 공간으로 이용하였다. 따라서 100 * 100 = 10000 차원의 이미지가 16차원으로 압축되는 효과도 거둘 수 있었으며, 실제 인식 처리 과정은 실시간 처리에도 적합하다는 것을 알 수 있었다.

<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 매우 높음을 알 수 있다. 특히 잘못된 매칭에 대해서는 많은 개선율을 나타내었다.

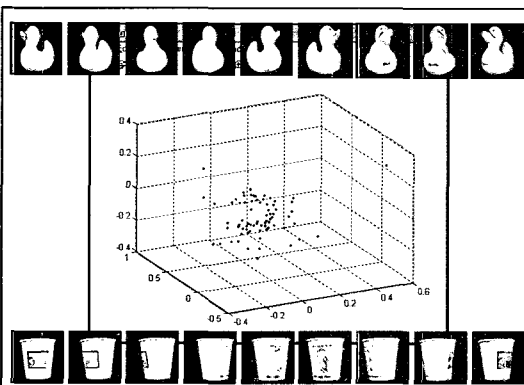
(그림 17)은 (그림 16)을 3차원 물체 영상 공간에 투영시킨 결과이다. (그림 18)은 물체 영상 공간에서 각 영상들이 매핑 되는 결과를 나타내었다.



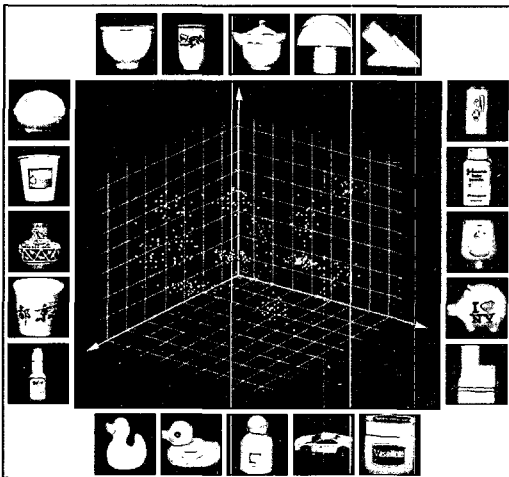
(그림 17) (그림 16)을 3차원 물체 영상 공간에서 투영



(그림 18) 3차원 물체 영상 공간에서 영상들의 매핑



(그림 19) 두 모델 입력 영상의 3차원 물체 영상 공간상의 매핑

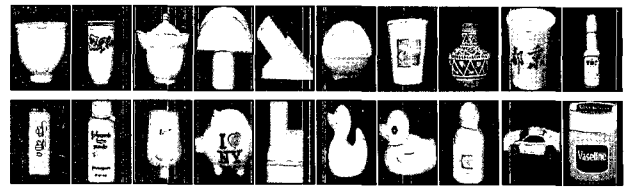


(그림 20) 고유 공간에서의 전체 물체 영상 분포

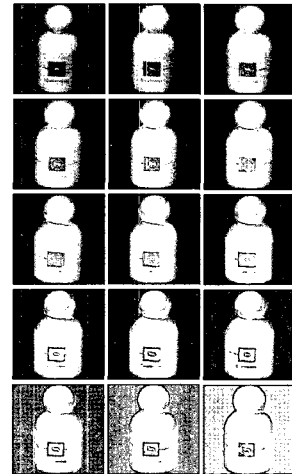
(그림 19)는 두 모델을 3차원 물체 영상 공간상에 매핑한 결과를 나타내었다. (그림 20)은 본 논문에서 사용한 물체 영상 모두 20가지를 고유 공간에서의 전체 물체 영상 분포를 나타낸다.

3.2 조명 변화에 따른 물체인식 실험 결과

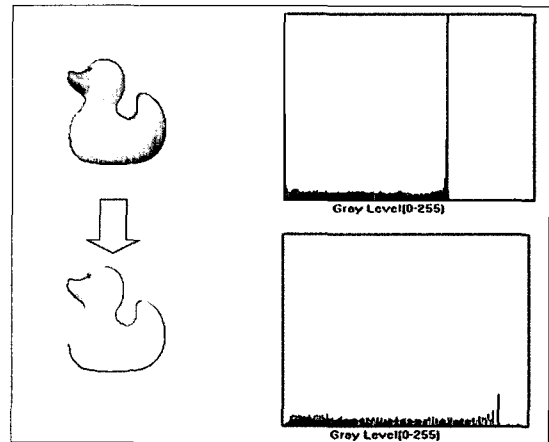
본 논문에서는 조명 변화에서도 인식률을 일정하게 유지하기 위해 히스토그램 이퀄라이제이션라는 기법을 사용하여 조명 변화에 따른 인식률 저하를 해결하였다.



(그림 21) 학습 영상의 예



(그림 22) 테스트 영상의 예



(그림 23) 이퀄라이제이션 수행 후 히스토그램 데이터

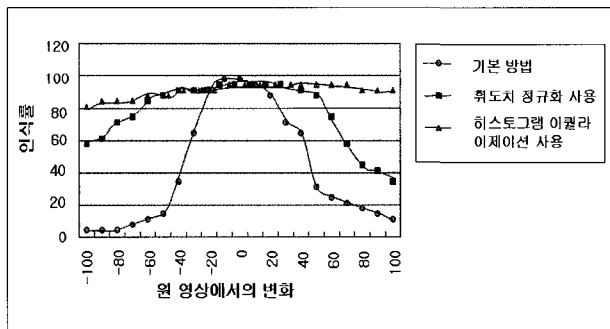
모델로 하는 영상은 물체 데이터베이스에서 100 * 100 크기의 그레이 스케일(Gray-Scale) 영상을 사용했다. 물체 영상의 데이터베이스는 총 20개의 물체로 구성 되어있다. (그림 21)은 실험에 사용한 학습 영상의 예이다. 테스트 영상은 모델 영상 각각에 픽셀 값을 약 -100에서 10의 차이로 +100까지의 변화가 있는 영상이다. 모델 영상 하나마다 조명의 변화가 있는 20개로 구성되어 총 1440개로 구성되어 있다.

(그림 22)는 테스트 영상의 예이다. 가장 중간에 있는 영상이 모델 영상과 조명의 변화가 없는 영상이고 왼쪽 위부터 차례로 어두운 조명부터 오른쪽 아래의 영상인 가장 밝은 조명의 영상을 나타낸 것이다.

(그림 23)은 어두운 영상을 히스토그램 이퀄라이제이션을 수행한 후 히스토그램 데이터의 변화를 나타낸다.

〈표 2〉 조명의 밝기 변화에 따른 인식을 비교

조명변화	기본적인 PCA	휘도치 정규화	히스토그램 이퀄라이제이션
-100	4.33%	56.00%	82.23%
-90	4.33%	60.33%	84.00%
-80	4.34%	70.67%	86.00%
-70	6.00%	76.00%	86.33%
-60	8.67%	82.67%	88.67%
-50	12.33%	87.67%	93.00%
-40	36.00%	91.33%	93.33%
-30	66.66%	91.33%	93.55%
-20	89.20%	91.67%	93.67%
-10	98.00%	92.33%	93.87%
0	97.25%	95.12%	94.67%
10	96.33%	91.33%	92.67%
20	91.67%	90.67%	92.67%
30	69.30%	89.87%	92.47%
40	63.67%	89.67%	93.55%
50	31.33%	84.33%	93.60%
60	26.00%	73.00%	91.00%
70	21.33%	57.66%	91.33%
80	18.00%	44.00%	91.20%
90	15.33%	38.00%	91.25%
100	13.67%	33.67%	90.77%
합계	41.61%	75.59%	90.94%



(그림 24) 조명 변화에 따른 인식을 비교

〈표 2〉는 실험을 통하여 조명의 변화에 따라 인식을 비교하여 나타낸 것이다. 각각 조명에 대해 아무런 전처리를 하지 않은 기본적인 PCA 방법, 휘도치 정규화를 사용한 방법, 그리고 본 논문에서 제안한 히스토그램 이퀄라이제이션을 사용해 전처리한 방법을 비교하였다.

〈표 2〉와 (그림 24)를 통해 결과를 분석해 보면 기본 PCA방법은 학습영상의 조명 변화가 거의 없을 경우에만 좋은 결과를 내지만, 조명 변화에 따라 바로 인식이 크게 떨어짐을 알 수 있다. 그리고 휘도치 정규화를 사용한 방법도 기본 PCA방법 보다는 조명 변화에 강하게 나타나지만 심한 변화에 대해서는 인식이 크게 떨어짐을 알 수 있다. 하지만 본 논문에서 제안한 방법은 심한 조명 변화에 대해

서도 인식률의 변화 없이 거의 90%이상의 높은 인식률을 일정하게 가짐을 볼 수 있다.

그리고 학습 영상과 거의 조명 변화가 없는 상황에서는 기본적인 PCA방법, 휘도치 정규화를 사용한 방법, 히스토그램 이퀄라이제이션을 사용한 방법의 순서로 인식이 좋게 나타난다. 이것은 원래의 학습 영상이 각각 한 물체에 대해서는 같은 조명 아래서 찍힌 영상이지만, 물체에 따라서는 그 조명의 세기가 달라서 이 조명의 차이가 하나의 큰 물체 마다의 차이를 나타내는 특징으로 나타나기 때문이다. 그렇기 때문에 조명의 차이가 없을 경우 아무런 처리를 하지 않은 기본 PCA방법은 이 조명의 차이로써 물체를 인식하는데 하나의 중요한 특징을 더 가짐으로써 조금 더 좋은 결과를 가지지만 이것이 조명의 변화된 영상을 실험의 입력 영상으로 들어 왔을 경우 잘못된 특징 점으로 작용하여 더욱 나쁜 결과를 야기하게 된다. 이러한 이유로 히스토그램 이퀄라이제이션을 사용한 방법은 조명에 매우 강하기 때문에 조명의 변화에 거의 일정한 좋은 인식률을 가진다. 휘도치 정규화를 해도 어느 정도 조명에 강하나 심한 조명의 변화에 (그림 25)과 같이 영향을 받으므로 히스토그램 이퀄라이제이션을 사용한 방법에 비해 대체로 낮은 인식률을 나타내었다.

4. 결 론

본 논문에서는 실험영상이 학습영상에 대해 조명의 차이가 있는 경우에도 데이터베이스안의 물체인지 가려내는 새로운 PCA 분석방법과 개선된 k-Nearest Neighbor 매칭을 이용한 물체 인식 기술을 제안하였다. 본 논문에서는 히스토그램 이퀄라이제이션과 미디언 필터를 이용하여 영상을 전처리하고 그것을 학습시켜서 물체 공간을 생성한다. 이때 히스토그램 이퀄라이제이션을 사용하여 히스토그램을 펼침으로써 조명 변화에 영향을 크게 감소시키는 결과를 나았고, 이것은 기본적인 주성분 분석 방법과 휘도치 정규화를 한 방법 등과 비교해 본 결과 조명 변화에 영향을 최소화하여 90% 이상의 일정한 인식률을 유지할 수 있었다. 그리고 개선된 k-Nearest Neighbor 방법을 이용하여 물체 영상 공간에서 투영된 입력 영상이 모델 영상과의 매칭에서 다른 물체 영상으로 인식하는 잘못된 매칭 발생률을 7.0%이상 감소시킬 수 있었다.

참 고 문 헌

- [1] R. Chellappa, Charles L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces : A Survey," Proc. IEEE, Vol. 83, No.5, pp.704-740, 1995.
- [2] Ming-Hsuan Yang, David Kriegman, Narendra Ahuja, "Detecting Faces in Images : A Survey", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.24, No.1, pp. 1-38, 2002.
- [3] J. Daugman, "Face and gesture recognition: overview," IEEE Trans, on Pattern Analysis and Machine Intelligence, Vol.19,

No.7, pp.675-675, 1997.

- [4] P. C. Yuen, D. Q. Dai and G. C. Feng, "Wavelet -based PCA for Human Face Recognition," IEEE, 1998.
- [5] Sami Romdhani, "Face Recognition using Principal Components Analysis," 1998.
- [6] M. Kirby and L. Sirovich, "Application of Karhunen -Loeve procedure for characterization of human faces," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.12, pp.103-108, 1990.
- [7] Rafael C. Gonzalez and Richard E. Woods, "Digital Image Processing," Addison Wesley, 1992.
- [8] Randy Crane, "simplified approach to Image Processing," Prentice Hall, 1997.
- [9] Hiroshi Murase and Shree K. Nayar, "Visual Learning and Recognition 3-D object from appearance," international journal of Computer Vision, Vol.14, 1995.
- [10] Christopher Eveland, Kurt Konolige and Robert C. Bolles, "Background Modeling for Segmentation of Video-Rate Stereo Sequences," pp.266-271, 1998.
- [11] Gnanadesikan, R. "Methods for Statistical Data Analysis of Multivariate Observations. New York : Wiley," 1997.
- [12] Ragael C. Gonzalez and Richard E. Woods "Digital Image Processing," Prentice Hall, 2002.
- [13] H. WU, Q. Chen, and M. Yachida, "Facial Feature Extraction and Face Verification," IEEE Proc. ICPR, pp.484-488, 1996.
- [14] C. Katropoulos and I. Pitas, "Rule-Based Face Detection In Frontal Views," Proc. ICASSP'97, pp.2537-2540, 1997.



정 병 수

e-mail : bsjung@nambu.ac.kr

1986년 전남대학교 계산통계학과(이학사)

1988년 전남대학교 전산통계학과(이학석사)

2006년 전남대학교 전산통계학과(이학박사)

1991년~1999년 전남과학대학 전산정보

처리과 교수

2000년~현재 남부대학교 디지털정보학과 교수

관심분야: 소프트웨어공학, 디지털영상처리, 컴퓨터비전, 임베디드 시스템



김 병 기

e-mail : bgkim@chonnam.chonnam.ac.kr

1978년 전남대학교 수학교육과(이학사)

1980년 전남대학교 수학과(이학석사)

2002년 전북대학교 수학과(이학박사)

1981년~현재 전남대학교 전자컴퓨터

정보통신공학 교수

관심분야: 소프트웨어공학, 객체지향시스템, 디지털 영상처리, 컴퓨터비전