

---

# 고차 국소 자동 상관 특징 정보를 이용한 외관 기반 객체 인식

강명아\*

Appearance-based Object Recognition Using Higher Order Local Auto Correlation  
Feature Information

Myung-A Kang\*

## 요 약

본 논문에서는 고차 상관 특징 정보와 주성분 분석을 결합하여 차원을 낮추면서도 객체 인식을 유지하고, 고유 공간 구성 시간을 현저하게 줄이는 알고리즘에 대해 기술한다. 제안된 방법은 기존의 기하학적 정보를 이용하거나 스테레오 영상을 이용하는 방법에 비해 많은 계산량이 요구되지 않기 때문에 실시간 시스템 구축에 매우 적합하다는 것이 실험을 통하여 증명되었다. 또한 인식률을 향상시키기 위해 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 입력영상을 인식의 단위로 이용하여 인식 오차를 줄일 수 있었다.

## ABSTRACT

This paper describes the algorithm that lowers the dimension, maintains the object recognition and significantly reduces the eigenspace configuration time by combining the higher correlation feature information and Principle Component Analysis. Since the suggested method doesn't require a lot of computation than the method using existing geometric information or stereo image, the fact that it is very suitable for building the real-time system has been proved through the experiment. In addition, since the existing point to point method which is a simple distance calculation has many errors, in this paper to improve recognition rate the recognition error could be reduced by using several successive input images as a unit of recognition with K-Nearest Neighbor which is the improved Class to Class method.

## 키워드

고차 국소 자동 상관 특징, 고유 공간, 주성분 분석

## Key word

HLLAF(Higher order local auto correlation features), Eigen space, PCA(Principal Component Analysis)

## I. 서 론

산업이 급속하게 발전하는 환경에서 자동화된 객체 인식 시스템의 개발은 생산 라인, 의료 및 군사 등 여러 분야에서 노동 인력을 대체할 수 있는 중요한 위치를 차지하고 있다. 특히 부품조립 및 검사, 군사장비 등과 같은 3차원 객체를 자동적으로 인식하기 위한 연구가 컴퓨터 비전(Computer Vision) 또는 로봇 비전(Robot Vision)이라는 분야로 분리되어 컴퓨터에 감지된 객체가 무엇 인지를 판단하고, 판단 결과를 이용하여 적절한 임무를 수행할 수 있는 능력을 제공해 줌으로써 사람이 할 수 있는 시각 능력을 컴퓨터에 부여해 주고자 하는 많은 연구가 진행되고 있다[1].

일반적으로 컴퓨터를 이용하여 객체를 인식하기 위해 객체에 대한 많은 정보를 컴퓨터에 저장시켜 이를 기반으로 객체를 인식하는 방법을 이용한다. 이는 모델을 기반으로 한 인식방법으로서 인식하고자 하는 모델의 물리적 특징과 그들의 기하학적 연관 관계를 표현하는 특징을 추출하여 데이터베이스에 저장하고, 입력 영상으로부터 추출된 특징들과 모델특징들을 비교하여 영상 내에 존재하는 객체의 종류와 위치 및 자세 등을 인식해 내는 방법이다[2][3]. 객체 인식은 그 대상 객체에 따라서 크게 2D 객체 인식과 3D 객체 인식으로 나눌 수 있다. 2D 객체 인식은 객체의 단편만이 인식이 가능하기 때문에 대부분의 평평한 면을 가지는 객체들이다[4][5]. 이에 반해 3차원 인식은 객체의 모든 view에 대해서 인식이 가능한 방법이다. 2차원 인식에 비해서 대상 객체의 제한이 없는 것이 다른 점이라고 할 수 있겠다. 모든 view에 대해서 인식이 가능해야 하기 때문에 일반적으로 2차원 객체 인식보다 잘 풀리지 않는 문제이며 다양한 방법으로 접근되고 있다[6].

본 논문에서는 입력 영상을 전처리 과정을 통해 실루엣 영상으로 변환한 다음 고차 국소 상관 특징(HLAF : Higher order Local Autocorrelation Features)을 통하여 추출된 특징을 주성분 분석법(PCA : Principal Component Analysis)이라는 통계적인 수법에 의해 객체 외관 특징들을 표현할 수 있는 저 차원 벡터 공간, 즉 파라메트릭 고유 공간으로 투영된다. 각 객체들은 이 공간 내에서 순차적으로 연속된 점의 궤적으로 표현되

고, 미리 학습된 모델 궤적과 입력 영상의 궤적을 비교함으로써 객체인식이 이루어지게 된다. 본 논문에서는 고유 공간의 구성 속도를 향상시키기 위해서 실루엣 영상과 HLAF를 이용한 특징을 사용하였다. 또한 인식과정에서 기존의 방법인 Point to Point 방식은 투영된 입력 영상이 모델 영상과의 매칭의 경우 여러 객체들이 투영되어 있는 객체 공간에서 실제 객체 영상이 매칭에 성공했는지라도 다른 객체 영상으로 인식하는 잘못된 매칭이 빈번하게 발생한다. 그래서 본 논문에서는 인식을 향상시키기 위해 개선된 Class to Class 방식인 K-Nearest Neighbor를 이용하여 몇 개의 연속적인 입력 영상을 인식의 단위로 이용하여 인식 오차를 줄일 수 있었다.

## II 전처리

### 2.1. 배경제거

카메라를 통하여 얻은 영상 시퀀스는 단순 배경에서 취득한 것이며, 일반 환경에서 취득한 영상에는 객체 인식에 필요 없는 많은 오브젝트(배경)이 포함되어 있다. 그러나 객체 인식에 필요한 것은 객체 영역(전경)이므로 우선 배경과 신체 영역을 분리하는 작업이 필요하고 이를 위해서는 먼저 배경 모델을 생성해야 한다. 그러나 조명의 밝기가 일정하지 않고 수시로 변하기 때문에 같은 카메라로 일정 시간 동안 똑같은 배경을 촬영한다고 할 지라도 모두 동일하지 않아 안정적인 배경 모델을 얻는 데 어려움이 따른다.

본 논문에서는 조명 변화로 인한 배경의 밝기 변화를 측정하여 시간 요소( $t$ )를 고려해서 일정 시간  $T_i$  동안 배경 영상  $I_t$ 를 취득한 다음 영상영역  $R$  내에 있는 각 픽셀( $x$ )들을 분석하여 조명이 가장 밝았을 때의 화소값  $P_{\max}(x)$ 와 가장 어두울 때의 화소  $P_{\min}(x)$ 을 얻는다. 이 두 화소값의 차이  $D(x)$ 는 조명의 변화로 나타날 수 있는 밝기의 임계치이다. 이 3가지의 요소를 이용해 배경 모델(Background Model)을 구성한다. 이와 같은 내용은 식(1-4)와 같다[7].

$$BM = \{P_{\max}(x), P_{\min}(x), D(x)\}_{x \in R} \quad (1)$$

$$P_{\max}(x) = \text{Max} I_t(x), (1 \leq t \leq T_i) \quad (2)$$

$$P_{\min}(x) = \text{Min} I_t(x), (1 \leq t \leq T_i) \quad (3)$$

$$D(x) = P_{\max}(x) - P_{\min}(x) \quad (4)$$

일단 배경 모델이 만들어지면, 이진 영상  $B(x)$ 는 식(5)에서 보여주는 것처럼 입력 영상  $I(x)$ 가 가장 밝은 화소값  $P_{\max}(x)$ 와 가장 어두운 화소값  $P_{\min}(x)$ 의 차분 연산을 통해 얻은 차이 값이 임계치  $D(x)$ 보다 크면 최대값(255)을 그렇지 않으면 최소값(0)의 화소값을 갖게 한다.

$$B(x) = \begin{cases} 255 & \text{if } |P_{\max}(x) - I(x)| \text{ or } |P_{\min}(x) - I(x)| > D(x) \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

식(5)는 조명으로 인해 생길 수 있는 밝기 차이는 무시하고 객체의 포즈 변화의 움직임의 차이를 갖는 영역만 분리하는 기준이 된다. 식(2)의 결과로 얻어진 이진 영상에는 배경 모델에서 설정한 밝기 값의 임계치를 벗어나는 조명의 변화로 인해 배경임에도 불구하고 전경 영역으로 분리되어 1픽셀의 작은 점들이 포함될 수 있다. 따라서 이 잡음을 없애기 위해 모폴로지(Morphology) 연산을 통해서 잡음을 제거하였다. 한 번의 침식연산을 수행하고 이때 객체의 줄어드는 현상이 발생하기에 다시 팽창 연산을 사용하여 복원하였다.

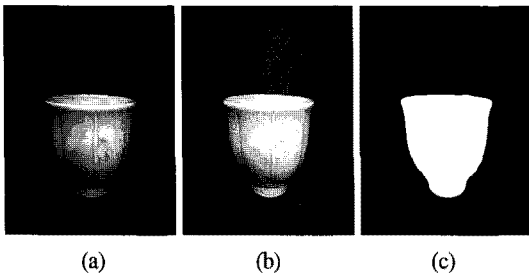


그림 1. (a)최소 밝기의 영상 (b)최대 밝기의 영상 (c)입력영상을 배경 모델 파라미터를 이용해 추출한 실루엣 영상

Fig. 1. (a) The minimum brightness of the image (b) The maximum brightness of the image (c) Silhouette image extracted using the background model parameters of input image

## 2.2. 고차 국소 자동 상관 특징 생성(HLAF)

앞의 절에서 설명한 방법을 통하여 얻어진 정규화 된 영상들을 이용하여 고유공간을 구성하게 되면 객체 이미지의 크기가  $320 * 240$  이기에 76,800차원이라는 계산량이 필요하게 된다. 본 논문에서는 이 차원의 크기를 줄이기 위해 고차 국소 자동 상관 계수를 이용하여 차원의 크기 즉 특징 벡터의 크기를 25개로 줄이는 방법을 제안한다. 고차 국소 자동 상관 특징(Higher order local autocorrelation features)들은 이동 불변의 함수로서 알려져 있다. 이러한 자동 상관 특징을 더 확장 시킨 함수가 고차 국소 자동 상관 함수이다. 이미지의 영역에  $P$ 가 표시되어 있을 때  $N$ 차원의 자동 상관 특징은  $N$ 을  $a_1, \dots, a_N$ 으로 대치시킨다면 다음 식(6)과 같이 정의할 수 있다.

$$x_i^N(a_1, \dots, a_N) = \int_P f(\tau + a_1) \dots f(\tau + a_N) d\tau \quad (6)$$

식(6)에서  $f(\tau)$ 는  $\tau$ 상에서 그레이 레벨을 표현한 것이다. 큰 영역을 지니고 있는  $P$ 를 조합하여 얻은 자동 상관 특징은 그 계수의 수가 실질적인 응용을 하기에 너무 많기 때문에 일정할 과정을 거쳐 이를 감소시켜야 한다. 따라서 먼저  $N$ 차원의 범위를 2차원 정도로 제한한다. ( $N=0,1,2$ ) 만일 0차원 자동 상관 특징은 영역  $P$ 의 그레이 레벨의 평균을 의미한다.

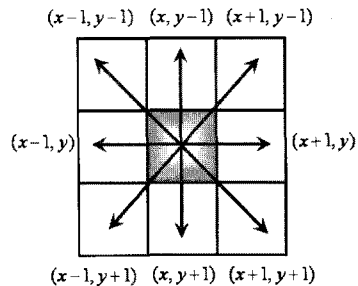


그림 2. 3x3 국소 마스크 패턴 Fig. 2. 3x3 local mask pattern

그림 2와 같이 3x3의 국소 마스크 패턴(local mask patterns)를 지정된 범위 내에 대치시키게 된다. 이 국소

마스크 패턴 중 가운데 픽셀에 대한 값을 참조하게 되며, 패턴의 수는 25개로 제한한다. 이 25개의 패턴을 지나고 있는 국소 마스크 패턴을 순차적으로 이미지 영상에 위치시킴으로서 25개의 특징 벡터를 계산할 수 있다. 그림 3은 25개의 패턴들을 보여주고 있으며, 여기서 심볼 \*는 “don't care”를 의미한다.

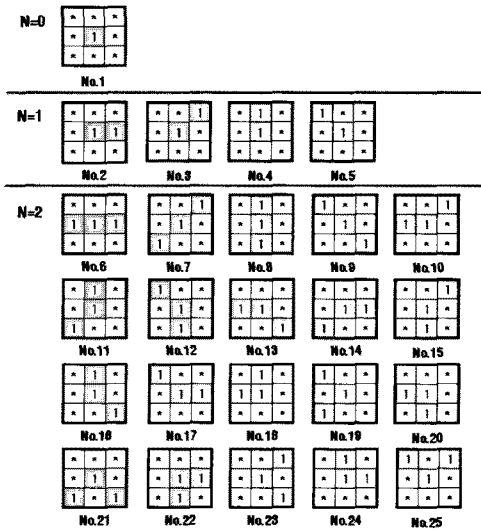


그림 3. 25개의 3×3 국소 마스크 패턴  
Fig. 3. 25 3×3 local mask patterns

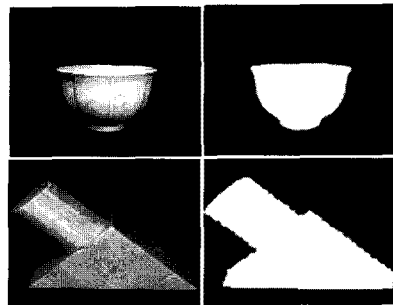
고차 국소 자동 상관 특징은 이미지 영역  $P$ 를 25개의 3×3 마스크 패턴을 이용해 탐색하여 각각의 마스크 패턴의 참조 픽셀 값의 합을 통하여 총 25개의 특징 벡터를 계산하게 된다. 이러한 특징 벡터는 명백하게 이동불변의 특징을 가지고 있다. 그림 3에서 국소 마스크 패턴을 사용하여 얻은 특징 벡터  $f^v$ 는 식(7)로 정의된 고차 국소 자동 특징의 특징 벡터를 가지게 된다.

$$f^v = f_1, \dots, f_{25} \quad (7)$$

$I_{x,y}$ 를 3×3 마스크 패턴을 영상의  $x, y$ 좌표를 탐색하였을 때의 특징 벡터라 한다면,  $I_{x,y}$ 의 0차원 자동 상관 함수를  $f_1$ 로 표현할 수 있다. 그러므로 각각의 특징 벡터  $f_i (i = 1, 2, 3, \dots, 25)$ 는 식(8)로 정의할 수 있다.

$$\begin{aligned} f_1 &= \sum_x \sum_y I(x,y) \\ f_2 &= \sum_x \sum_y I(x,y)I(x+1,y) \\ &\vdots \\ f_5 &= \sum_x \sum_y I(x,y)I(x-1,y-1) \\ f_6 &= \sum_x \sum_y I(x,y)I(x+1,y)I(x-1,y) \\ f_7 &= \sum_x \sum_y I(x,y)I(x+1,y-1)I(x-1,y+1) \\ &\vdots \\ f_{14} &= \sum_x \sum_y I(x,y)I(x-1,y+1)I(x+1,y) \\ f_{15} &= \sum_x \sum_y I(x,y)I(x,y+1)I(x+1,y-1) \\ &\vdots \\ f_{25} &= \sum_x \sum_y I(x,y)I(x-1,y-1)I(x+1,y-1) \end{aligned} \quad (8)$$

주어진 입력 영상에서 정규화된 객체 영상은 고차 국소 상관 특징을 이용하여 객체영상이 담고 있는 25개의 특징 벡터를 생성한다. 객체 추출 알고리즘을 통하여 얻은 객체 영상은 3×3 마스크 패턴 탐색을 통하여 특징 벡터가 생성되어진다. 검출된 영상은 이진영상으로 마스크 패턴에 대한 픽셀의 총합으로 결정되어진다. 그림 4는 2개의 객체의 고차 국소 상관 특징을 추출한 데이터이다.



Obj	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	2769	2757	2748	2757	2748	2748	2742	2754	2742	2753	2754	2739	654	633	624	654	684	681	624	660	627	745	648	795	666
2	1983	1962	1953	1968	1956	1929	1856	1929	1923	1935	1921	1944	1914	1903	1920	1932	1944	1935	1914	1935	1893	1956	1926	1956	1935

그림 4. 2개의 객체의 고차 국소 상관 특징데이터  
Fig. 4. Higher local correlation feature data of 2 objects

### III. 객체 특징 생성과 인식

단일 카메라를 이용하여 객체의 방향이 회전하는 경우 같은 형상으로 인식하기에는 많은 어려움이 있다. 본 논문에서는 객체의 방향성이 회전하는 경우에도 같은 형상으로 인식하기 위해 주성분 분석을 이용한 객체 방향성 문제의 해결 방법을 제시한다. 회전하는 객체를 같은 현상으로 인식하기 위해 객체를 5° 회전하여 영상을 획득한 후 목차 2에서 설명한 방법을 통하여 얻어진 고차 상관 특징을 이용하여 객체 모델 데이터베이스를 그림 5와 같이 구성하였다.

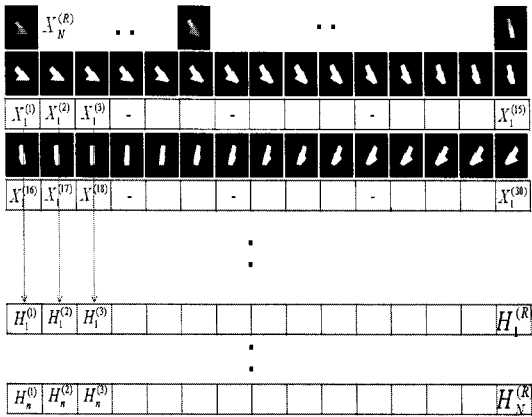


그림 5. 2진 영상에서의 고차 상관 특징 계수 추출 및 DB구성

Fig. 5. Extraction of HLAf in binary image and DB configuration

#### 3.1. 주성분 분석을 이용한 공간 생성

주성분 분석법은 고차원의 입력 데이터 집합을 저차원의 의미 있는 데이터 집합으로 줄일 수 있다. 객체 영상 데이터의 경우 한 개의 객체가 회전 동작을 구성하는 프레임(frame)의 수가 많고 특징을 추출하기 비교적 어렵기 때문에 빠른 인식속도와 효과적인 특징 추출이 가능한 방법을 적용해야 한다. 따라서 3장에서 추출한 연속적인 회전 동작의 고차 상관 특징을 이용하여 저차원 벡터로 표현하는 방법에 대해 기술 한다.

앞 절에서 구한 특징 벡터를  $x$ 라 하고, 식(9)와 같이 표현 할 수 있다. 이 벡터의 고유공간을 계산하기 위해서는 먼저 모든 특징 벡터의 평균 벡터를 구하여 각 특징 벡터의 차를 구해야한다. 평균 벡터  $c$ 와 새로운 특징 집

합  $X$ 는 식(10-11)과 같다.

$$x = [x_1, x_2, x_3, \dots, x_N]^T \quad (9)$$

$$c = (1/N) \sum_{i=1}^N x_i \quad (10)$$

$$X = [x_1 - c, x_2 - c, \dots, x_N - c]^T \quad (11)$$

그런 다음, 식(12)를 만족하는 고유벡터를 구하기 위해 공분산 행렬  $Q$ 에 대한 고유치  $\lambda$ 와 고유벡터  $e$ 를 구하면 된다.

$$Q \triangleq XX^T, \lambda_i e_i = Q e_i \quad (12)$$

고유 벡터를 구하기 위해 고유치 분해를 사용하지 않고 특이치 분해(Singular Value Decomposition)이용함으로써 특징 집합  $X$ 의 공분산 행렬에 대한 고유벡터를 쉽게 얻을 수 있었다. 이렇게 얻어진 고유공간에 평균 벡터  $c$ 를 뺀 특징 집합  $X$ 를 식(13)을 이용하여 투영시켰다.

$$f_j = [e_1, e_2, e_3, \dots, e_k]^T (x_n - c) \quad (13)$$

이와 같이 얻어진 저 차원 벡터 공간을 고유공간이라고 부른다. 앞서 설명한 바와 같이 주성분 분석은 몇 개의 주성분 벡터를 유도하여 이를 통해 차원의 축소와 자료의 요약을 주목적으로 하고 있다. 따라서 전체 객체 영상을 적절히 설명하기 위하여 보유해야 할 주성분의 수를 결정해야 한다. 식(14)는 우리가 필요한 개수를 선택하는 방법이다.

$$\left( \sum_{i=1}^k \lambda_i / p \right) \times 100 \geq 80 \quad (14)$$

여기서  $\lambda_i$ 는  $i$ 번째 고유값,  $p$ 는 전체 고유값 개수라고 할 때 전체 분산 중 주성분  $C_i$ 가 설명할 수 있는 비율은  $\lambda_i/p$ 이다. 즉  $k$ 개 주성분들이 설명할 수 있는 누적 비율이 80%이상일 때의 개수를 선택하게 된다. 그림 6은

실험에 사용한 객체들의 특징 집합으로부터 구한 고유치 개수에 따른 주성분의 누적기여도를 보여주고 있다.

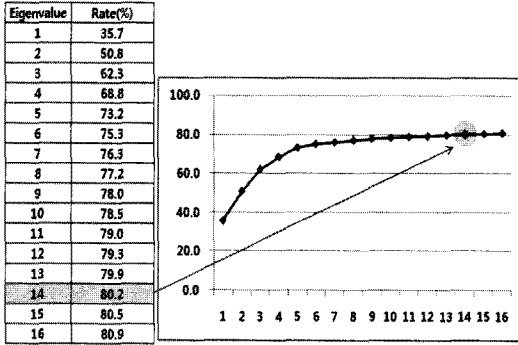


그림 6. 고유치의 개수에 따른 누적 기여도  
Fig. 6. Cumulative contribution depending on the number of eigenvalue

각각의 6개의 실루엣 객체 포즈에 대한 고유공간내의 분포는 그림 7과 같다.

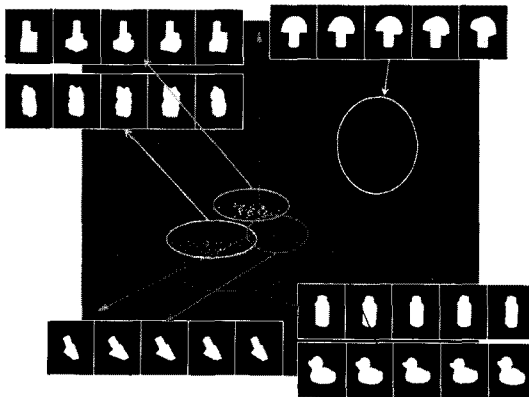


그림 7. 6개의 객체의 영상 시퀀스를 객체공간으로 투영한 결과

Fig. 7. The result of projecting 6 object image sequences to object space

3.2. 개선된 K-Nearest Neighbor를 이용한 거리 평가와 객체 인식

투영된 입력 영상이 모델 영상과의 매칭의 경우 (Point to Point) 여러 동작들이 투영되어 있는 객체 영상 공간에서 실제 객체 영상은 매칭에 성공했을지라도 다른 객체 영상으로 인식하는 잘못된 매칭이 발생

하였다. 이러한 문제를 해결하기 위해 단일 객체 영상 단위의 매칭이 아닌 몇 개의 연속적인 입력 영상에 대해 각각 몇 개의 모델 영상들을 인식의 단위로 이용하였다(Class to Class). 매칭 알고리즘은 식(15), 식(16)에서 나타낸 것처럼 K-Nearest Neighbor 방법을 이용하였다. 여기서  $\arg S(M_j) = j$ 는 모델의 번호를 구하는 연산자이다.

$$w = \frac{Max(\arg S(M_j)) - Min(\arg S(M_j))}{d(k-1)} \quad (15)$$

$$\frac{\sum \sum w(I_i - M_j)}{k} \quad (16)$$

위의 식(16)에 의해 구해진 값을 이용하여 모델 영상과 입력 영상의 인식을 결정한다. 여기서는 누적비율이 80% 이상이 되는  $k=14$ 를 이용하였다. 그림 8은 공간상에 투영된 연속적인 입력 영상들과 모델 영상들의 위 방법을 이용한 매칭 방법을 나타낸 것이다.

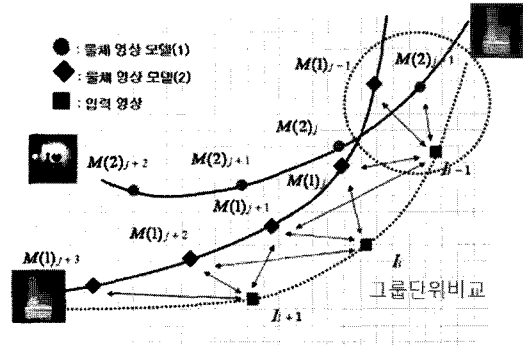


그림 8. K-Nearest Neighbor를 이용한 객체 영상 인식  
Fig. 8. Object image recognition using K-Nearest Neighbor

그림 8에서 나타나듯이 입력 영상과 모델 영상 간 가장 가까운 거리를 가지는 영상일지라도 실제 다른 객체 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 객체 영상 전체인 궤적간의 일치성까지 알 수 있으며, 이를 이용하여 얻어진 결과는 객체 영상 전체를 평가하는 데도 이용할 수 있었다.

#### IV. 객체 인식에 대한 실험결과

실험에 이용된 객체의 수는 25개를 사용하였고, 한 객체는 그림 9에서처럼 5°씩 회전한 영상을 카메라로 촬영하여 획득한 것이다.

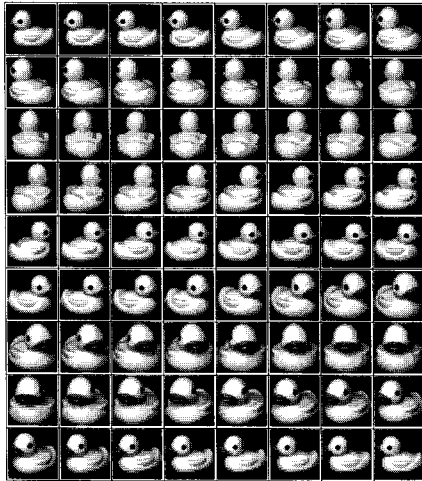


그림 9. 오리 객체의 회전 모델 집합 영상  
Fig. 9. Set image of rotation model of a duck object

각 객체들이 0° ~ 360°씩 회전한 객체의 집합을 한 영상의 객체라고 한다. 주성분 분석을 이용한 인식에 가장 문제점은 고유공간을 구성하는 시간과 객체 인식을 수행하는 속도가 빠르지 않기에 실시간 적용에는 무리가 있었다. 그래서 객체 인식에 대한 고유공간을 구성하기 위해 640 \* 480으로 촬영한 객체 영상은 크기 정규화를 거쳐 320 \* 240영상으로 변환하였다. 그리고 본 연구에서는 객체의 모든 영상의 데이터를 사용하지 않고 실루엣 영상을 생성하였다. 객체의 실루엣 영상만으로도 객체 특징을 상당 부분 가지고 있기에 이 특징을 HLAF의 입력데이터로 사용하였다. 또한 객체 실루엣 영상이 320 \* 240 = 76,800개의 데이터를 가지고 있는데 데이터를 HLAF를 이용하여 25개 특징 데이터만 추출하게 된다. 즉 전처리 과정에서 76,800개의 데이터를 25개로 축소시키게 되는 것이다. 표 1은 주성분 분석의 고유 공간 구성 시간과 기존의 주성분 분석과 제안한 방법의 인식률을 비교 나타내었다. 그 결과 인식률의 큰 변화 추이 없이 상당한 속도 개선의 효과를 만들었다.

표 1. PCA와 HLAF를 이용한 PCA의 처리시간 및 인식률  
Table 1. PCA processing time and recognition rate using PCA and HLAF

모델 영상 개수	PCA		PCA using HLAF	
	Training Time(sec)	Recognition Rate	Training Time(sec)	Recognition Rate
50	1.47	92.80%	0.015	87.50%
100	5.67	93.00%	0.016	89.00%
150	12.73	93.50%	0.017	89.50%
200	23.91	93.70%	0.017	89.70%
250	39.47	93.70%	0.031	89.50%
300	57.69	94.25%	0.047	90.00%
350	77.52	94.30%	0.062	89.00%
400	102.21	94.30%	0.094	90.50%
450	130.33	95.20%	0.125	90.50%
500	165.55	95.20%	0.150	90.70%
550	200.97	95.90%	0.195	91.50%
600	249.96	96.50%	0.214	91.50%
650	259.74	96.50%	0.236	91.00%
700	361.12	96.30%	0.287	91.90%
750	321.22	97.10%	0.338	92.00%
800	390.55	97.40%	0.387	92.00%
950	469.75	97.50%	0.452	92.50%
1,000	578.54	97.90%	0.512	92.50%
합계	3,448.4	95.28%	3.195	90.60%

그리고 표 2는 각 매칭 방법별 매칭 성공률에 대한 분석 결과를 나타내었다. 표 2에서 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 2. 매칭 방법에 따른 매칭 성공률  
Table 2. Successful Matching Rate According to the Matching Methods

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%

#### IV. 결론 및 향후계획

본 논문에서는 HLAF와 주성분 분석법을 결합하여 기존의 객체 영상을 사용하는 방법이 아닌 객체의 형태의 정보인 실루엣 영상을 HLAF의 입력으로 사용하여 얻어진 특징 데이터를 산출하여 이를 주성분 분석에 활용함으로써 차원을 낮추면서도 인식률을 90% 이상 유지하고, 고유 공간 구성 시간을 현저하게 줄이는 방법을 제시하였다. 제안된 방법은 기존의 기하학적 정보를 이용하거나 스테레오 영상을 이용하는 방법에 비해 많은 계산량이 요구되지 않기 때문에 실 세계에서 구현이 용이하고 실시간 시스템 구축에 매우 적합하다는 것이 실험을 통하여 증명되었다. 또한 2차원 객체 영상을 가지고 3차원 객체 인식을 수행하기 위하여 객체를 5°씩 회전하여 모델 영상을 구성함으로써 객체의 단면만을 인식하는 게 아니라 객체 영상의 단위인 그룹 단위로 인식을 수행하여 3차원 객체 인식도 가능하였다. 그러나 복잡한 배경에서는 객체 영역만을 분리하는데 어려움이 있었고 향후에는 이러한 문제점을 해결하여 보다 안정적인 객체 인식 알고리즘을 개발할 계획이다.

#### 참고문헌

[1] P. J. Besl and R. C. Jain, "Three-Dimensional Object Recognition," ACM computing surveys, vol. 17, no. 1, pp. 75-145, March 1985.

[2] Jong-Min Kim, and Maeng-Kyu 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, Oct. 2008.

[3] Rafael C. Gonzalez and Richard E. Woods, Digital Image Processing, Addison Wesley, 1992.

[4] Rafael C.Gonzalez, Richard E.Woods, "Digital Image Processing 2/E," Prentice Hall, pp.519-532, 2001.

[5] JY. Tak and E. Hwang, "Indexing and Matching Scheme for Recognizing 3D Objects from Single 2D Image," International Conference on Internet and Multimedia Systems and Applications (IMSA'09), pp.60-67, 2009.

[6] Ragael C. Gonzalez and Richard E. Woods "Digital Image Processing", Prentice Hall, 2002.

[7] Ismail Haritaoglu, David Harwood and Larry S. Davis, "W4: Who? When? Where? What? A Real-time System for Detecting and Tracking People," Third Face and Gesture Recognition Conference, pp.222-227, 1998.

[8] M. Saito and K. Kitaguchi, "Appearance based object pose estimation using regression models," SICE Annual Conference, pp.1926-1929, 2008

#### 저자소개

강명아(Myung-A Kang)



1992년 2월 : 광주대학교  
전자계산학과(이학사)

1995년 2월 : 조선대학교  
전산통계학과(이학석사)

1999년 2월 : 조선대학교 전산통계학과(이학박사)

2004년 9월 ~ 현재 : 광주대학교 컴퓨터공학과 조교수

※ 관심분야 : 영상처리, 패턴인식, 신경망