

# Fuzzy C-means와 CONDENSATION을 이용한 객체 검출 및 추적 시스템

(An Object Detection and Tracking System using Fuzzy  
C-means and CONDENSATION)

김종호\*, 김상균\*\*, 황구선\*\*\*, 안상호\*\*\*\*, 강병두\*\*\*\*\*

(Jong-Ho Kim, Sang-Kyoon Kim, Goo-Seun Hang, Sang-Ho Ahn, and Byoung-Doo Kang)

**요약** 동영상에서의 움직이는 객체 검출과 추적은 객체 식별, 상황인식, 지능형 영상 감시 시스템 등 많은 시각 기반 응용 시스템에서 기본적인 전처리 작업이다. 본 논문에서는 배경과 조명이 실시간으로 변화하는 상황에서 움직이는 객체를 빠르고 정확하게 검출하고 움직이는 객체가 다른 물체에 가려지는 경우에도 강인하게 객체를 추적하는 방법을 제안한다. 객체의 효과적인 검출을 위해서 효과적인 고유 공간과 Fuzzy C-means(FCM)를 결합하여 사용하고 검출된 객체를 강인하게 추적하기 위해 Conditional Density Propagation (CONDENSATION) 알고리즘을 사용한다. 먼저 Principal Component Analysis(PCA)를 이용하여 배경 영상에서 수집한 학습 데이터를 주성분(Principal component)으로 선형변환 한다. 주성분들의 고유 특성에 대한 해석을 통하여 객체와 배경에 대하여 판별 능력이 우수한 주성분을 선별하여 고유 배경을 구성한다. 다음으로 이전단계에서 구성된 고유 벡터와 입력 영상을 결합한 연산 결과를 FCM의 입력 값으로 사용해서 객체를 검출한다. 최종적으로 검출된 객체의 좌표를 CONDENSATION의 입력으로 사용해서 객체를 추적한다. 고정된 카메라에서 조명변화와 배경변화에 적용 가능한 시스템을 구현하기 위해 고정된 카메라에서 움직이는 다양한 객체가 포함된 영상을 수집하여 학습데이터로 구성하여 사용하였다. 실험 결과에 따르면 제안하는 방법이 조명변화와 배경변화 그리고 객체의 부분적 움직임에 모두 강인하게 객체를 검출하고 다른 물체나 배경에 의해 객체가 일부 가려지더라도 객체를 추적함을 보여준다.

**핵심주제어** : 객체 검출, 객체 추적, 고유배경, 주성분 분석, 클러스터링, 퍼지씨먼즈

**Abstract** Detecting a moving object from videos and tracking it are basic and necessary preprocessing steps in many video systems like object recognition, context aware, and intelligent visual surveillance. In this paper, we propose a method that is able to detect a moving object quickly and accurately in a condition that background and light change in a real time. Furthermore, our system detects strongly an object in a condition that the target object is covered with other objects. For effective detection, effective Eigen-space and FCM are combined and employed, and a CONDENSATION algorithm is used to trace a detected object strongly. First, training data collected from a background image are linear-transformed using Principal Component Analysis (PCA). Second, an Eigen-background is organized from selected principal components having excellent discrimination ability on an object and a

\* 인제대학교 컴퓨터공학부, 제1저자

\*\* 인제대학교 컴퓨터공학부, 교신저자

\*\*\* 진코퍼레이션 비전팀

\*\*\*\* 인제대학교 전자지능로봇공학과

\*\*\*\*\* 전자 부품연구원 지능로보틱스 연구센터

background. Next, an object is detected with FCM that uses a convolution result of the Eigen-vector of previous steps and the input image. Finally, an object is tracked by using coordinates of an detected object as an input value of condensation algorithm. Images including various moving objects in a same time are collected and used as training data to realize our system that is able to be adapted to change of light and background in a fixed camera. The result of test shows that the proposed method detects an object strongly in a condition having a change of light and a background, and partial movement of an object.

**Key Words** : Object Detection, Object Tracking, Eigen-background, PCA, Clustering, FCM

## 1. 서 론

사람과 컴퓨터의 상호작용을 위해 시작된 컴퓨터 비전 기술은 사용자 인터페이스 설계, 로봇학습, 감시 시스템 등의 많은 분야에서 응용되고 있다. 그중에서도 배경과 객체를 정확하고 효과적으로 분리해 내는 객체 검출 기술과 검출된 객체를 추적하는 기술은 기본적인 매우 중요한 기술이다. 물체를 정확히 분리해내지 못한다면 다음 단계인 로봇 비전, 증강현실, 동작 인식 등과 같은 시각기반 시스템에서 정확한 행동 분석과 동작인식을 할수 없고, 객체를 정확하게 추적하지 못한다면 객체의 폐색이 발생되었을 때 연속된 객체 정보를 획득할 수 없다.

그러므로 많은 연구자들에 의해 배경으로부터 객체를 분리하고 추적하는 연구가 진행되고 있다. 객체에서 배경을 분리하는 대표적인 방법으로는 이전프레임이나 미리 정의된 배경과 현재 프레임과의 차이를 이용하는 방법, 비디오의 압축 정보를 이용하는 방법, 객체의 움직임을 이용하는 방법, visual attention을 이용하는 방법이 있다.

프레임과의 차이를 이용하는 차분 영상의 경우 구현이 쉽다는 장점이 있지만 객체의 움직임이 없거나 객체의 일부분이 움직일 경우 객체 전체를 검출하지 못하는 단점이 있다. 압축 정보를 이용하는 방법은 미리 정의된 배경을 요구하지 않고 처리속도가 빠르다는 장점이 있지만 압축방법에 따라 다른 구현 방법을 사용해야 하는 단점이 있다.

객체의 움직임 정보를 이용하는 방법은 처리속도가 빠르다는 장점이 있지만 객체의 정확한 형태정보를 추출할수 없고, 빛의 변화에 민감한 단점이 있다. visual attention을 이용하는 방법은 처리속도가 빠른

장점이 있지만 의미있는 객체를 추출하기 힘들고 정확한 형태정보를 추출하기 힘든 단점이 있다.

위 단점들을 극복하기 위한 방법으로 고유공간(Eigen-space) 모델을 이용한 방법이 있다[1,2]. 이 방법은 영상에서 배경의 정보를 분석하고 고유배경(Eigen-background)을 구성하여 배경으로 사용한다. 고유배경은 배경과 객체를 구분한다. 그리고 움직임이 없는 객체도 검출이 가능하게 된다. 고유공간모델을 개선한 방법으로 고유배경을 갱신(renewal)하여 배경을 재구성하는 방법이 있다[3]. 이 방법은 새로운 입력 영상을 학습데이터로 추가하여 고유배경을 갱신하는 방법이다. 그러나 고유공간을 사용하는 방법은 배경에 대한 잡음이 객체로 잘못 검출되는 경우가 생긴다. 따라서 객체 전체를 검출하지 못하고 움직임이 있는 객체의 일부분만이 검출되는 문제를 해결해야 한다. 뿐만 아니라 인위적으로 정의한 한 프레임을 배경으로 사용하는 문제를 해결하고 배경변화에 적응적이며 잡음에 강인한 객체를 검출하는 방법이 필요하다.

객체를 잘 검출하더라도 검출 속도가 느릴 경우 객체를 연속적으로 검출할 수 없다. 이러한 문제점을 해결하기 위해 검출된 객체를 추적하는 기술은 필수적이다. 검출된 객체를 추적하기 위한 여러 가지 기법들로는 Optical Flow, Mean-shift, Kalman Filter, CONDENSATION 등이 있다.

Optical-Flow를 이용하는 방법은 입력 영상 프레임의 시간적 벡터와 프레임내의 화소들의 위치변화에 대한 속도를 공간적 벡터로 표현하여 물체의 이동을 인식하고 추적한다. 그러나 추적하는 객체의 이동량이 많은 경우 화소 단위로 계산하기 때문에 전후 프레임의 시공간 벡터를 구하기 위한 계산량이 많아지는 단점이 있다[4].

Mean-Shift를 이용한 방법은 단일 후보를 이용하기 때문에 구현이 간단하고 계산량이 적은 장점이 있다[5]. 그러나 지역 모드로 수렴하는 특성으로 인해 객체의 폐색이 발생할 경우 추적 성공률이 떨어지는 단점이 있다.

Kalman Filter를 이용하는 방법은 백색 가우시안 노이즈를 가진 선형동적 시스템에서 외부에서 관측된 값과 측정된 예러들로부터 다음 상태를 예측할 수 있는 방법을 제공한다. 또한 순차적이면서 재귀적인 알고리즘을 이용하여 적은 연산만으로도 쉽게 구현할 수 있다[6]. 하지만 추적방향이 선형이라는 단점을 가지고 있다.

CONDENSATION[7] 알고리즘을 이용한 방법은 요인 샘플링 기법에 기반한 확률적인 객체 추적 기법으로 복잡한 영상에서 특정 객체를 추적하는데 사용된다.

본 논문에서는 객체 검출과 추적 시스템들의 문제점을 해결하기 위해 영상에서 효과적인 정보를 얻을 수 있는 개선된 고유 배경을 구성하여 강인한 객체를 검출하고 CONDENSATION 알고리즘을 이용해서 검출된 객체를 추적하는 시스템을 제안한다. 먼저 일상 생활에서 사람의 이동과 배경 내에서 의자와 같은 객체의 움직임 정보가 있는 다양한 영상을 사용하여 배경의 변화에도 적응적이도록 학습데이터를 구성한다. PCA[8]를 이용하여 배경의 정보를 분석하고 고유 배경을 구성한다.

일반적인 설명비율을 사용하는 방법으로는 그림자와 같은 잡음의 오검출이 심하다. 제안하는 시스템은 PCA를 통해 구해지는 주성분을 클러스터링 알고리즘을 이용하여 해석하고 신체를 효과적으로 검출하는 고유값을 선별한다.

검출방법으로는 사용자의 인위적인 목표값 설정이 필요하지 않고 비교사 학습방법인 클러스터링을 사용한다.

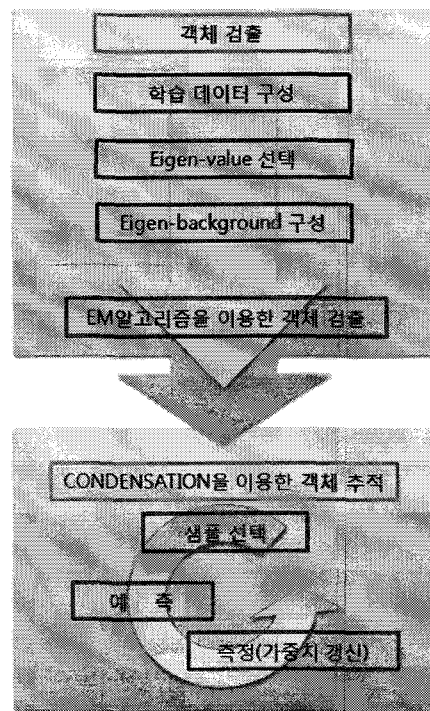
클러스터링 알고리즘으로는 FCM[9]을 이용한다. 선별된 고유벡터를 입력 영상의 가중치로 사용하고 FCM를 통해 클러스터링 하여 잡음에 강인한 객체를 검출한다. 하지만 본 논문에서 제안하는 검출 방법론의 경우 검출율은 뛰어나지만 많은 연산량으로 인해 검출속도가 크게 떨어지는 문제점을 가지고 있다. 그래서 본 논문에서는 CONDENSATION 알고리즘을 이용해서 검출된 객체를 추적한다. CONDENSATION 알고리즘을 이용해서 검출된 객체를 추적함으로써 검출기가 가지는 검출속도 문제를 해결하고 다른 물체

나 배경에 의해 폐색이 발생하는 경우에도 안정적으로 객체를 추적할 수 있다.

## 2. 객체 검출 및 추적 시스템

### 2.1 시스템 구조

본 논문에서 제안하는 시스템의 전체 구조는 <그림 1>과 같다. 먼저 다양한 배경 영상들을 학습 데이터로 하여 PCA의 입력으로 사용한다. PCA를 통해 학습 데이터 집합의 정보를 해석하고 배경과 움직이는 객체를 잘 구분하는 고유값을 선별한다. 선별된 값을 통해 고유 배경을 구성하고 클러스터링의 입력 차원으로 사용한다. 입력 영상과 구성된 고유 배경을 픽셀 단위로 각각 곱하고 이 결과를 FCM을 통해 클러스터링하여 영상 내에서 움직이는 객체를 검출한다. 다음으로 CONDENSATION 알고리즘을 이용해서 검출된 객체를 추적한다. <그림 1>은 학습 데이터 구성, 주성분 분석 및 선별, 그리고 검출 및 추적에 대한 전체 과정을 나타낸다.



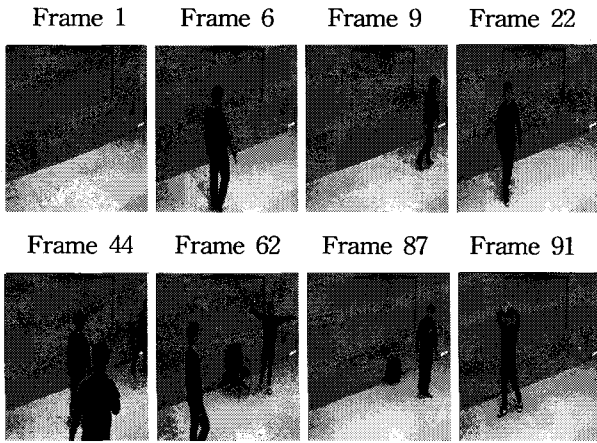
<그림 1> 객체 검출 및 추적 시스템 전체 구성

## 2.2 객체 검출

### 2.2.1 고유배경

본 논문에서는 영상에서 배경과 움직이는 객체를 구분하기 위해서 PCA를 이용한 고유배경을 구성한다.

고유배경은 움직임이 있는 객체를 효율적으로 검출할 수 있고, 인위적으로 정의한 프레임을 배경으로 사용하는 경우 매번 관리자에 의한 배경 설정과 갱신을 해야 하는 문제를 해결할 수 있다. 또한 움직임이 없을 때 객체를 찾지 못하는 문제도 보완한다.



<그림 2> 본 논문에 사용된 학습 데이터의 일부

먼저, 고유배경을 구성하기 위하여 학습 데이터로 사용될 배경의 샘플을 획득한다. 자연 영상에서 배경을 정의하는데 있어서 인위적으로 위치가 변경된 후 움직임이 없는 의자와 같은 가구는 배경에 포함되어야 하고 지나가는 행인과 같은 경우에는 움직이는 객체로 인식되어야 한다. 따라서 배경 변화에 따른 검출율을 높이기 위하여 <그림 2>와 같이 고정된 시점에서 움직이는 객체가 포함된 영상을 학습데이터로 사용하였다.

두 번째로 PCA를 통해 학습데이터를 분석하고 배경의 주성분을 추출하여 고유배경을 생성한다. 학습데이터를 PCA의 입력으로 사용하기 위하여 2차원 영상을 일차원 열벡터  $I_j$ 로 구성한다.

$$I_j = \begin{bmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{Nj} \end{bmatrix}, \quad (1 \leq j \leq M) \quad (1)$$

$$S = [I_1, I_2, \dots, I_M] = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1M} \\ x_{21} & x_{22} & \dots & x_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \dots & x_{NM} \end{bmatrix} \quad (2)$$

$S$ 는 학습데이터 전체이고,  $j$ 는 전체 학습 영상 집합에서 이미지의 인덱스이다.  $M$ 은 학습 데이터의 개수이고,  $N$ 은 특징의 수이다. 본 논문에서는  $240 \times 320$  영상을 사용하고 있으므로 총 특징의 개수  $N$ 은 76,800개이다. 학습 데이터에 대한 평균은 식 (3)과 같다.

$$\Psi_i = \frac{1}{M} \sum_{j=1}^M x_{ij}, \quad (1 \leq i \leq N, 1 \leq j \leq M) \quad (3)$$

$$\Psi = [\Psi_1, \Psi_2, \dots, \Psi_N]^T$$

$x_{ij}$ 는  $I_j$ 에서  $i$ 번째의 픽셀 값이다. 본 논문에서는 100개의 학습데이터를 사용한다. 편차는 행렬식 (4)를 통해 식 (5)로 표현된다.

$$\Phi_j = I_j - \Psi \quad (4)$$

$$A = \begin{bmatrix} \Phi_{11} & \Phi_{12} & \dots & \Phi_{1M} \\ \Phi_{21} & \Phi_{22} & \dots & \Phi_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{N1} & \Phi_{N2} & \dots & \Phi_{NM} \end{bmatrix} = [\Phi_1, \Phi_2, \dots, \Phi_M] \quad (5)$$

고유값과 고유벡터를 계산하기 위하여 공분산 행렬을 만든다. 공분산 행렬  $C$ 는 식 (6)을 통해 구해진다.

$$C = \frac{1}{M} \sum_{j=1}^M \Phi_j \Phi_j^T \quad (6)$$

$$= AA^T$$

공분산 행렬  $C$ 는  $N \times N$  차원의 행렬인데,  $240 \times 320$  영상에서 이 행렬의 특징 차원의 수  $N$ 은 76,800이다.  $76,800 \times 76,800$  차원의 공분산 행렬을 이용하여 76,800개의 고유값과 이에 따른 고유벡터를 계산하기란 현실적으로 어렵다. 따라서  $A^T A$ 에 대한 고유벡터  $v_j$ 를 고려하여  $M \times M$  차원에서 고유값과 고유벡터를 구한다.

$$AA^T v_j = \lambda_j v_j \quad (7)$$

양변에  $A^T$ 를 곱하면

$$A^T A A^T v_j = \lambda_j (A^T v_j) \quad (8)$$

$A^T v_j$ 는  $A^T A$ 에 대한 고유 벡터  $v'_j$ 가 된다. 이 행렬은  $M \times M$  행렬이다.

$$A^T v_j = v'_j, \quad v_j = A v'_j \quad (9)$$

최종적으로 식(9)를 통해  $M \times N$  행렬의 고유 벡터  $v_j$ 를 구할 수 있다.

공분산 행렬  $C$ 로부터  $M$ 개의 고유값  $\lambda_j$ 와 각 고유값에 대응하는  $N$ 개의 고유벡터  $v_j$ 를 계산한다.

$$B_j = \frac{1}{M} \sum_{j=1}^M v_j \phi_j^T \quad (j = 1, 2, \dots, M) \quad (10)$$

$B_j$ 은 고유값  $j$ 번째에 해당하는 고유배경이고  $v_j$ 은  $j$ 번째 고유값에 해당하는 고유벡터이다. 고유값  $j$ 번째에 해당하는 고유벡터  $v_j$ 와 편차벡터  $\phi^T$ 를 통하여  $B_j$ 을 구한다.

### 2.2.2 일반적인 고유배경 방법

일반적인 고유배경 방식은 고유값을 내림차순으로 정렬한 후 식(11)을 만족하는 고유값만을 선택하여 사용한다.

$$\frac{\sum_{j=1}^m \lambda_j}{\sum_{j=1}^T \lambda_j} \geq th (= 0.9) \quad (11)$$

$T$ 는 고유값 전체의 개수이고, 경계값을  $th(0.9)$ 로 설정하여 설명비율이 90%까지 반영되는 고유값  $m$ 개를 사용하여 고유배경을 만든다. 가장 높은 설명비율을 가지는 고유값은 객체와 배경을 구분하기에 적합하지 않고 잡음 성분을 많이 가지고 있다.  $th$ 를 0.9이상으로 설정을 하더라도 첫 번째 고유값은 포함되기 때문에 객체와 배경의 구분이 명확하지 않다. 고유배경  $B$ 는 식 (12)를 이용하여 생성한다.

$$B = \frac{1}{m} \sum_{j=1}^m B_{l_j}, \quad (1 \leq l \leq n, 1 \leq j \leq m) \quad (12)$$

$$D_j = |I_j * B| > t \quad (13)$$

입력영상  $I$ 에서 식 (13)을 통하여 경계값  $t$ 보다 크면 객체이고 그렇지 않으면 배경을 나타낸다. 고유배경을 구성하기 위하여 PCA를 통해 학습데이터의 성분을 분석한다.

## 2.3 클러스터링(Clustering)

일반적인 고유배경을 생성하는 방법은 설명 비율이 높은 값을 가지는 고유값들을 선택하여 학습 데이터를 해석한다. 고유값의 설명 비율이 높은 값은 학습 데이터에서 많은 부분의 정보를 가지고 해석한다. 그러므로 형광등과 같은 미세한 조명에 반응하는 배경의 픽셀변화, 그림자 등에 크게 반응하는 특징을 가진다. 이런 성분들을 가지고 있으므로 배경의 질감으로 인한 잡음과 객체에 의해 발생한 그림자를 객체로 잘못 검출하는 것을 볼 수 있다.

일반적인 고유배경 문제를 해결하기 위해서 클러스터링을 사용한다. 제안하는 방법은 고유값의 설명 비율이 큰 것을 기준으로 선택하지 않고 배경과 객체를 잘 구분하는 성분을 가지고 있는 고유 값을 선별한다. 클러스터링 과정을 통해 고유 값을 선별하고 이 고유값에 해당하는 벡터들로 고유배경을 구성한다. 클러스터링은 데이터를 배경과 객체를 나타내는 성분으로 분석하여 나눌 수 있다. 클러스터링 방법으로는 Nearest Neighbor(이하 NN), K-Means, FCM 등이 있다.

### 2.3.1 FCM(Fuzzy-C-Means)

FCM알고리즘은 퍼지 집합에 근거하여 클러스터링 방법이다[9]. 이 알고리즘은 k-means와 유사한데 입력 데이터가 각 군집에 속할 확률과 거리를 근거로 클러스터링 방법이다. 본 논문에서는 클러스터링 방법으로 기준이 되는 데이터의 선별이 필요 없고, 초기데이터의 영향이 적은 FCM을 사용한다.

입력 영상  $I$ 와 식(9)의 고유벡터  $v$ 를 사용하여 FCM의 입력으로 사용될 데이터  $E$ 를 식(14)와 같이

생성한다.

$$E_s = v_s \cdot I_{diag}, \quad I_{diag} = \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_N \end{bmatrix} \quad (14)$$

행렬 연산을 위하여 I를 대각행렬  $I_{diag}$ 으로 변환한다. S는 선별된 고유값의 인덱스이다. 본 논문에서는 3개의 고유 값에 해당하는 고유벡터를 선별하여 사용한다. 식(15)을 최소화에 기반을 두어 클러스터링 한다.

$$J_r = \sum_{p=1}^P \sum_{q=1}^Q u_{pq}^r \|e_p - c_q\|^2 \quad (15)$$

여기서  $e_p$ 는 클러스터링 하고자하는 입력 데이터이고, P는 그 입력데이터의 총 개수, Q는 cluster의 개수이다.  $u_{pq}$ 는  $e_p$ 가 군집 q에 속하는 확률을 나타낸다.  $c_q$ 는 q번째 군집의 중심이고,  $\|\cdot\|$ 는 norm이다.

$$u_{pq} = \frac{1}{\sum_{q=1}^Q \left( \frac{\|e_p - c_z\|}{\|e_p - c_q\|} \right)^{\frac{2}{m-1}}}, \quad (z = 1, 2, \dots, Q) \quad (16)$$

$$c_q = \frac{\sum_{p=1}^P u_{pq}^m \cdot e_p}{\sum_{p=1}^P u_{pq}^m} \quad (17)$$

식(16)과 (17)를 통해  $u_{pq}$ 와  $c_q$ 을 갱신하면서 식(18)이 만족할 때까지 반복한다.

$$\|u_{pq}^{(t+1)} - u_{pq}^{(t)}\| < \delta \quad (18)$$

t는 반복 횟수이며,  $\delta$ 는 0과 1 사이의 확률 값이다. 입력 데이터  $e_p$ 와 클러스터 중심  $c_q$ 사이 거리 정보에 확률  $u_{pq}$ 을 곱하여 가중치를 준다. 즉  $J_r$ 은 거리와 확률에 의해서 구해지는 특징이다. 이 과정을 통해 거리가 최소화된  $J_r$ 을 찾아 클러스터링을 하여 배경과 객체를 분류할 수 있다.

### 2.3.2 클러스터링을 이용한 객체 검출

클러스터링의 입력으로 사용될 특징 값 F는 식(19)을 이용하여 구성한다.

$$F = I * B \quad (19)$$

입력 영상 I와 형성된 고유 배경 B의 결합연산 결과 F를 클러스터링의 입력으로 사용한다. 선별된 고유 값을 입력영상에 적용하여 구해진 특징 값들을 클러스터링 한다. 입력 영상에서 각각의 특징들은 선별된 고유값에 대응하는 새로운 특징 값으로 구성된다. 이 새로운 특징 값은 군집에 의해 클러스터링 되고 객체와 배경으로 각각 나누어진다.

### 2.4 객체 추적

본 논문에서는 이전 단계에서 검출된 객체를 CONDENSATION 알고리즘[7]을 이용해서 추적한다. CONDENSATION 알고리즘은 통계적인 방법에 기초하여 추적 대상의 존재 확률이 높은 곳을 대상이 있는 곳으로 지정하여 추적을 하는 방식이다. Kalman Filter와 유사한 방법을 이용하지만 Kalman Filter가 가우시안 확률 분포에 기반을 두고 있는 반면에 CONDENSATION은 좀 더 일반적인 상황을 가정한 멀티 모달 확률 분포에 기반을 두고 있어서 복잡한 배경에서도 높은 추적성공률과 빠른 처리속도를 보인다.

CONDENSATION 알고리즘은 크게 “Factored Sampling”과 “동적 모델의 학습”에 기초를 두고 있다. Factored sampling은 샘플링 된 상태에 확률 값을 부여하는 방법을 말하고 동적 모델의 학습은 샘플링 된 샘플들의 집합이 전파되어가는 양식을 미리 학습에 의해 정해 두는 것을 의미한다.

본 알고리즘은 추적을 위한 입자 필터의 3가지 갱신 단계로 선택, 예측, 측정의 3단계로 구성된다.

단계 1, 2, 3은 시간  $t_{k-1}$ 에서의 샘플 집합  $\{s_{k-1}^{(n)}, \pi_{k-1}^{(n)}, c_{k-1}^{(n)}, n = 1, \dots, N\}$ 로부터 시간  $t_k$ 에서의 샘플 집합  $\{s_k^{(n)}, \pi_k^{(n)}, c_k^{(n)}, n = 1, \dots, N\}$ 을 구하는 단계이다.

먼저 선택단계에서는 임의의 수  $r \in [0, 1]$ 을 균일 확률 분포로부터 만들고,  $c_{k-1}^{(j)} \geq r$ 을 만족하는 가장

작은  $j$ 를 찾는다. 다음으로  $s_k^{(n)} = s_{k-1}^{(n)}$ 으로 놓는다. 여기서  $s_k^{(n)}$ 는 시간  $t_k$ 에서의  $N$ 개의 샘플 중  $n$ 번째 샘플을 나타낸다.

두 번째로 예측단계에서는 각각의 샘플  $s_k^{(n)}$ 을 아래와 같이 표본화하여 예측한다.

$$p(X_k | X_{k-1} = s_k^{(n)}) \quad (20)$$

세 번째로 측정단계로써 측정된 특징량  $Z_k$ 를 이용하여 아래와 같이 새로운 위치를 측정하고 가중치를 갱신한다.

$$\pi_k^{(n)} = p(Z_k | X_k = s_k^{(n)}) \quad (21)$$

여기서  $Z_k$ 는 시간  $t_k$ 에서 관찰된 측정치,  $X_k$ 는 시간  $t_k$ 에서 표본화 된 상태를 나타내고  $\pi_k^{(n)}$ 는 샘플  $s_k^{(n)}$ 의 확률을 나타낸다.

다음 가중치들을  $\sum_n \pi_k^{(n)} = 1$ 이 되도록 정규화하고 누적 확률과 함께  $(s_k^{(n)}, \pi_k^{(n)}, c_k^{(n)})$ 으로 저장한다.

$$\begin{aligned} c_k^{(0)} &= 0, \\ c_k^{(n)} &= c_k^{(n-1)} + \pi_k^{(n)} \end{aligned} \quad (22)$$

여기서  $c_k^{(n)}$ 는 샘플  $s_k^{(n)}$ 의 누적 확률을 나타낸다. 이렇게 한 시간 스텝에서의 샘플 집합이 구해지면 현재 상태를 추정하고 추적할 수 있다.

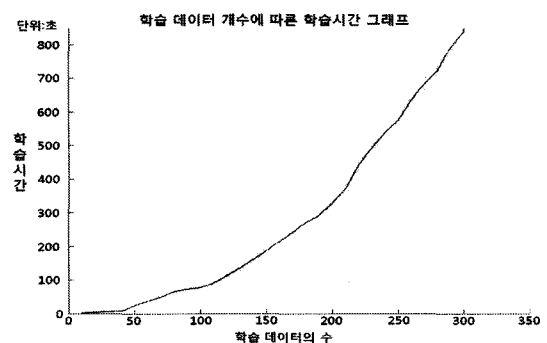
### 3. 실험 및 결과 분석

제안하는 방법은 구성요소와 조명의 다양한 변화하에서 실내와 실외에서 HVR-2030 webcam을 이용하여 활용한 IJUdataset과 비전 그룹에서 사용하는 pets 2000, 2001, 2004를 이용해서 실험하였다. 실험을 위해서 3.4GHz 인텔cpu, 2GB RAM, 윈도우즈 운영체제를 사용하였다.

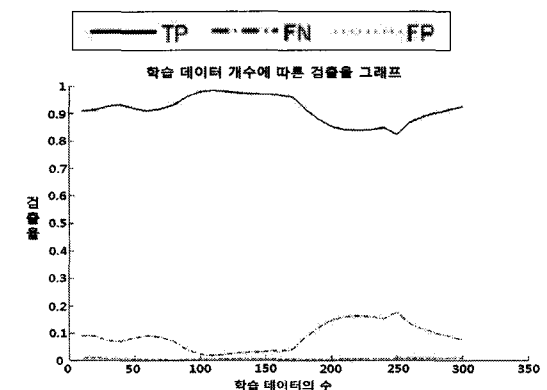
### 3.1 학습데이터 개수에 따른 학습 속도와 검출을 실험

<그림 3>은 학습데이터 개수에 따른 학습 속도와 검출율을 나타낸 그래프이다.

<그림 3(a)>에서 보는 바와 같이 학습 데이터 개수가 증가하면 공분산 행렬이 커져서 학습시간이 급격히 증가한다. 그리고 학습데이터의 개수가 300개가 넘을 경우 더 이상 학습이 되지 않았다. 그리고 검출율의 경우 학습데이터의 개수가 증가하면 대체로 검출율도 증가하나 100개 일 때 가장 좋은 결과를 나타내고 200개 이상일 경우 오히려 검출율이 떨어졌다. 그래서 본 논문에서는 검출율이 가장 높고 학습시간이 적게 걸리는 100개를 이용해서 데이터를 학습하고 객체를 검출했다.



(a) 학습데이터 개수에 따른 학습시간

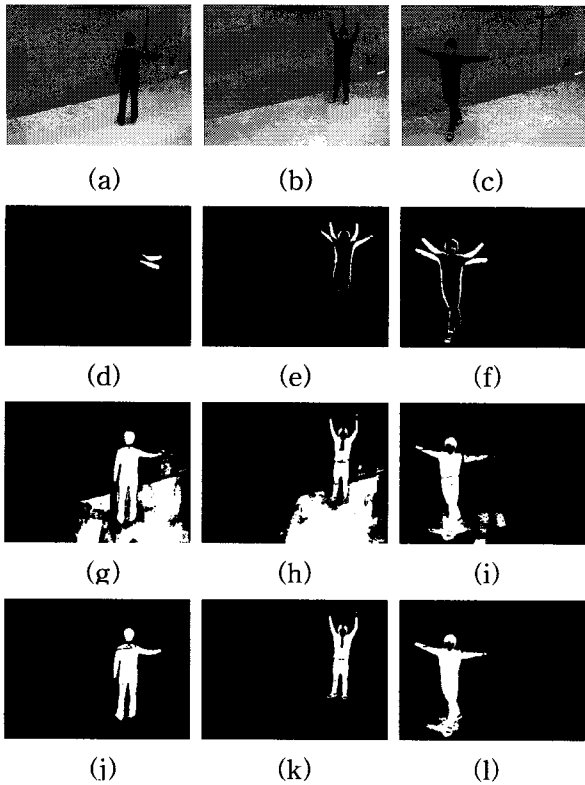


(b) 학습 데이터 개수에 따른 검출율

<그림 3> 학습 데이터 개수에 따른 학습 시간과 검출을 실험

### 3.2 기존 방법과의 비교 실험

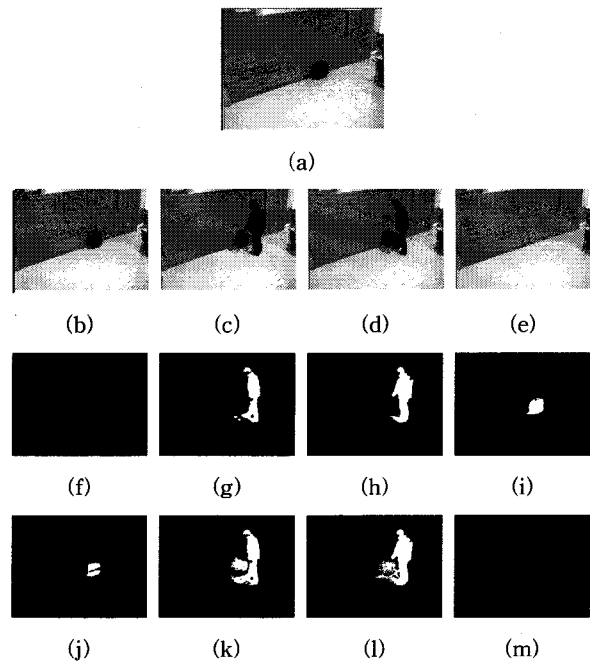
<그림 4>는 (a)~(c)와 같은 프레임이 입력되었을 때 차영상을 이용한 방법, 일반적인 고유배경 방법과 본 논문에서 제안하는 방법을 비교 실험한 결과이다. (d)~(f)는 전 프레임과의 비교를 통한 차이를 나타내는 차영상을 이용한 방법의 결과이다. 그림에서 보는 바와 같이 움직임이 없는 몸통 부분은 검출이 되지 않고 움직이는 손 부분만 검출됨을 알 수 있다. (g)~(i)는 일반적인 고유배경 방법을 이용한 검출 결과이다. 그림에서 보는 바와 같이 설명 비율이 높은 주성분들은 잡음과 그림자의 성분을 많이 가지고 있어 배경에서 잡음과 그림자가 오검출되고 있다. (j)~(l)은 본 논문에서 제안하는 클러스터링을 통해 개선된 고유배경을 사용한 방법이다. 그림에서 보는 바와 같이 움직임이 없는 부분도 정확하게 검출되고 배경과 객체를 잘 분류하는 주성분을 선별하기 때문에 잔상을 남기지 않고 잡음 및 그림자가 제거되어 강인하게 객체를 검출하고 있다.



<그림 4> (a)~(c) 입력 영상 (d)~(f) 차영상 방법 (g)~(i) 일반적인 고유배경 (j)~(l) 제안하는 방법

### 3.3 인위적으로 정의한 배경을 사용하는 방법과 비교 실험

본 논문에서는 연속적으로 입력되는 100개의 배경 영상을 이용해서 고유배경을 생성한다. 만약 <그림 5(a)>와 같이 인위적으로 정의한 한 프레임을 배경으로 사용하여 객체를 검출할 경우 초기에 입력되는 배경 영상에 따라 객체를 배경으로 배경을 객체로 오인하는 문제가 발생한다.



<그림 5> 인위적으로 정의한 프레임과 객체 영상이 포함된 프레임을 배경으로 사용했을 때 비교 실험

그림 5(f)~(i)는 인위적으로 정의한 프레임을 배경으로 사용했을 때의 결과이다. <그림 5(f)>에서 보는 바와 같이 가방을 배경으로 인식하고 <그림 5(i)>에서 보는 바와 같이 가방을 뺄 경우 가방이 있는 위치가 객체로 인식되는 문제가 발생한다. 하지만 전체 학습 데이터에서 객체가 포함된 영상을 40% 사용할 경우 <그림 5(j)~(m)>에서 보는 바와 같이 가방이 포함된 영상이 삽입될 경우 가방을 객체로 인식하고, 가방을 치울 경우 배경으로 정상적으로 인식한다.



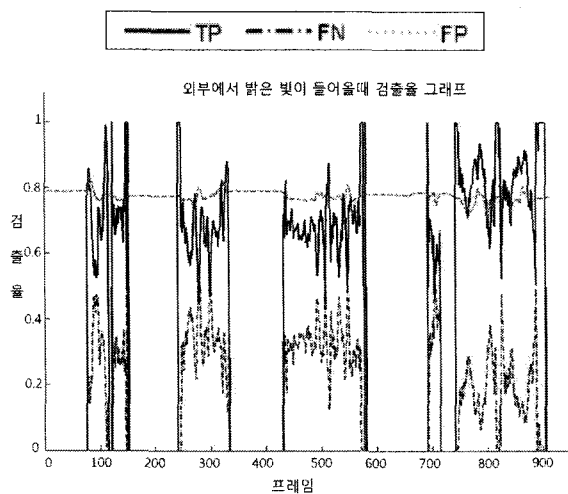
### 3.4. 다양한 조명하에서의 실험 결과

제안한 시스템의 성능을 분석하기 위하여 <그림 6>과 같이 일반적인 조명하에서와 어두운 조명, 그리고 외부에서 강한 빛이 들어왔을 때의 영상에 대해서 객체를 추출하였다.

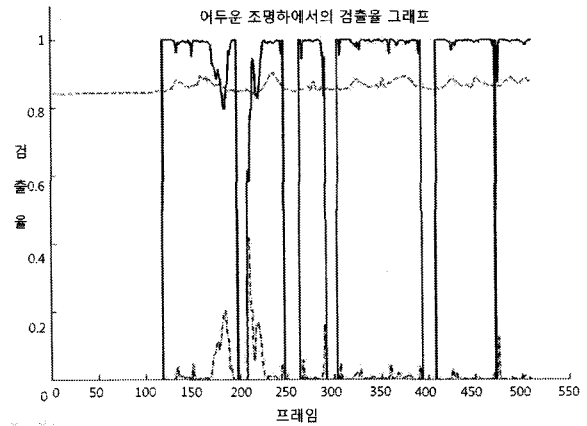


(a) 밝은 조명(b) 어두운 조명(c) 일반적인 조명  
<그림 6> 다양한 조명하에서의 검출을 실험

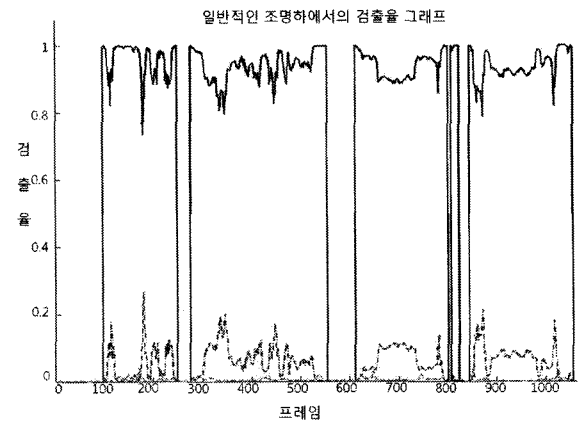
<그림 7(c)>의 그래프에서 보는 바와 같이 일반적인 조명하에서 객체를 객체로 정인식하는 TP, 배경을 객체로 오인식하는 FP, 객체를 배경으로 오인식하는 FN 모두 좋은 결과를 보임을 알수 있다. 하지만 <그림 7(a)>에서 보는 바와 같이 실내 환경에서 외부의 강한 빛이 들어왔을 때 그래프에서 보는바와 같이 TP가 크게 떨어지고 FN, FP가 크게 증가함을 알수 있다. 이는 외부에서 유입되는 강한 빛으로 인해 생긴 그림자를 객체로 오인식하고, 외부에서 유입되는 빛으로 인해 객체 부분의 명도값이 급격하게 증가해서 객체를 배경으로 오인식하기 때문이다. 그리고 <그림 7(b)>와 같이 조명이 어두운 환경에서는 TP와 FN은



(a) 외부에서 밝은 빛이 들어올 때 TP, FN, FP 그래프



(b) 어두운 조명하에서의 TP, FN, FP 그래프

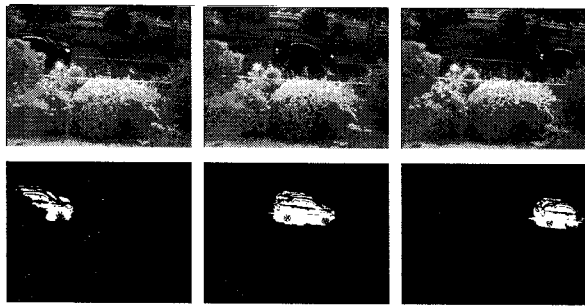


(c) 일반적인 조명하에서의 TP, FN, FP 그래프

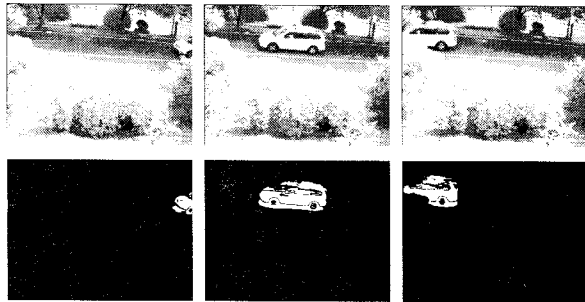
<그림 7> 다양한 조명하에서의 검출을 그래프

좋은 결과를 나타내나 객체를 배경으로 판단하는 FP가 높음을 알수 있다. 이는 어두운 조명으로 인해 객체의 명도값이 떨어져서 발생하는 현상이다.

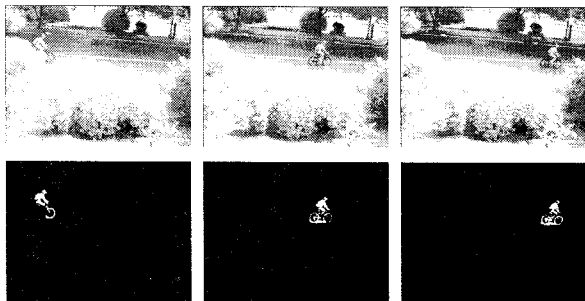
<그림 8>은 다양한 객체와 자연광의 변화 속에서의 검출 결과를 나타낸다. <그림 8(a)>의 경우 구름 낀 자연광상태에서 검출 결과를 나타낸다. 자연광 상태이고 아스팔트 색과 객체의 색이 유사함에도 불구하고 객체가 정확하게 검출된다. <그림 8(b)>의 경우 밝은 자연광으로 인해 차량과 도로의 색이 구분되지 않는 상황에서의 객체 검출 결과이다. <그림 8(c)>의 경우 반사광으로 인해 객체의 색과 배경의 색이 유사함에도 불구하고 객체를 정확하게 검출할 수 있다.



(a)



(b)



(c)

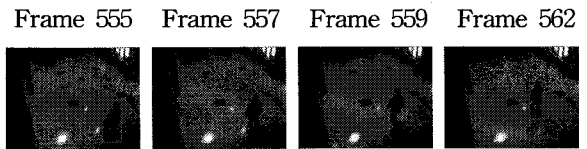
<그림 8> 다양한 객체와 자연광의 변화 속에서의 검출 결과

검출 실험 결과, 본 논문에서 제안한 개선된 고유배경은 조명과 환경에 의한 잡음, 빛에 의해 생기는 그림자, 움직임이 없는 객체 등을 효과적으로 배경과 객체로 분리하였다. 제안하는 방법은 차영상을 이용한 방법에서 나타나는 잔상이 생기는 문제점과 움직이지 않는 객체를 검출하지 못하는 문제를 해결하였다. 그리고 인위적인 배경으로 초기화 했을때의 문제점을 해결하였다.

그러나 조명이 급격히 변할 경우 검출율이 크게 저하되는 문제가 발생했다. 이를 해결하기 위해 조명의 급격한 변화시 영상의 명도값(intensity)을 일정하게 유지하도록 값을 조정해주는 연구가 필요하다.

### 3.5 CONDENSATION을 이용한 객체 추적

<그림 9>는 동일한 입력 영상에서 객체 검출 결과와 CONDENSATION을 이용한 객체 추적 결과를 보여준다. <그림 9(a)>는 본 논문에서 제안한 방법으로 객체를 검출한 결과이고, <그림 9(b)>는 검출된 객체를 CONDENSATION 알고리즘을 이용해서 추적한 결과이다. 그림에서 보는 바와 같이 객체를 555프레임에서 검출하고 CONDENSATION 알고리즘을 이용해서 추적한 결과 <그림 9(a)>에서 보는바와 같이 모든 프레임에서 객체를 정확하게 추적했다.



(a)



(b)

<그림 9> 객체 검출과 추적

<그림 10>은 객체가 다른 사물에 가려졌을 때의 검출 실험결과이다. <그림 10(a)>에서 보는 바와 같이 검출기만 사용할 경우 사물에 의해 객체가 가려졌을 때 객체를 검출하지 못한다. 하지만 <그림 10(b)>에서 보는바와 같이 CONDENSATION 알고리즘을 이용해서 객체를 추적할 경우 객체가 다른 사물에 가려지더라도 객체가 추적할 수 있다.



(a)



(b)

<그림 10> 객체가 다른 사물에 가려졌을 때의 검출과 추적

#### 4. 결론 및 향후 연구

본 논문에서는 객체를 검출하고 추적하기 위해 개선된 고유배경을 구성하고, 클러스터링 알고리즘을 이용하여 배경변화에 강인하게 객체를 검출하고 CONDENSATION 알고리즘을 이용해서 검출된 객체를 추적하는 시스템을 제안하였다. 먼저 환경 변화에도 적응적으로 객체를 검출할 수 있도록 다양한 종류의 배경 영상을 학습 데이터로 구성하였다. PCA를 이용하여 학습 데이터를 분석하고 클러스터링 알고리즘을 이용하여 배경과 객체를 잘 분류하는 주성분을 선별하였다.

전통적인 주성분 분석방법인 설명비율에 의존하지 않고 배경과 객체를 잘 해석하는 주성분 공간을 선별하여 사용함으로써 복잡한 배경을 객체로 오인식하는 문제를 해결할 수 있었다. 선별된 주성분으로 고유 배경을 구성하고 클러스터링 알고리즘의 입력차원으로 사용하였다. 클러스터링 알고리즘으로는 FCM을 사용하였다. 고유 벡터와 입력 영상을 결합한 연산 결과를 FCM의 입력 값으로 사용함으로써 기존의 객체 검출 시스템이 적은 조명변화에도 민감하게 반응하여 영상 내에 잡음이 발생하는 오류를 줄일 수 있었고 객체의 움직임이 없거나 적은 경우 배경으로 오인식하는 문제도 해결할 수 있었다. 하지만 FCM을 사용함으로써 검출율은 높아지지만 많은 연산량으로 인해 처리속도가 떨어지는 단점이 있다. 이를 보완하기 위해서 검출된 객체를 CONDENSATION 알고리즘을 이용해서 추적한다.

제안한 시스템을 보완하기 위해서는 극심한 조명 변화에도 적응적으로 객체를 검출하기 위한 연구가 이루어져야 하고 배경을 업데이트할 때에 많은 연산 부하로 인해 즉각적인 배경 변화 대응이 어렵다는 문제점을 해결해야 할 것이다. 앞으로의 발전 방향으로 다중 카메라를 이용하여 더 많은 정보를 수집하고 이를 3차원으로 모델링하여 3차원 동작 인식 시스템을 개발할 계획이다.

#### 참 고 문 헌

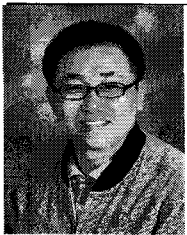
- [1] C.R. Wren, A. Azarbayejani, T. Darrll, and A.P. Pentland, "Pfinder: real-time tracking of the human body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.19, No.7, pp. 780-758, 1997.
- [2] N.M. Oliver, B. Rosario, and A. Pentland, "A Bayesian computer vision system for modeling human interactions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.22, No.8, pp. 831-843, 2000.
- [3] J. Yumel, J. Renno, D. Greenhill, J. Orwell, and G.A. Jones, "Adaptive Eigen-backgrounds for object detection," *International Conference on Image Preccessing*, 2004.
- [4] 김경규, 박경남, "광류를 이용한 적응적인 블록 정합 움직임 추정 기법," *한국산업정보학회 논문지*, 제 13권, 1호, pp57-67, 2008.
- [5] D. Comaniciu, V. Ramesh, and P. Meer, "Real Time Tracking of Non-Rigid Objects Using Mean Shift," *IEEE International Conference Computer Vision and Pattern Recognition*, VOL. 2, pp.142-149, 2000.
- [6] 성치영, 강병두, 김종호, 김상균, "효과적인 검출기와 칼만 필터를 이용한 강인한 얼굴 추적 시스템," *한국멀티미디어학회 논문지*, 제 10권, 제 1호, pp.26-35, 2007.
- [7] M. Isard and A. Blake, "CONDENSATION - Conditional Density Propagation for Visual Tracking," *Internal Journal of Computer Vision*, Vol. 29, pp5-28, 1998.
- [8] J.H. Kim, B.D. Kang, J.S. Eom, C.S. Kim, S.H. Ahn, B.J. Shin and S.K. Kim, "Real-Time Face Tracking System Using Adaptive Face Detector and Kalman Filter," *LNCS 4552, Human-computer Interaction, Part III, HCII 2007*, pp. 669-678, 2007.
- [9] M.S. Yang, K.L. Wu, J.N. Hsieh, et al. "Alpha-Cut Implemented Fuzzy Clustering Algorithms and Switching Regressions," *IEEE*



김 중 호 (Jong-Ho Kim)

- 정회원
- 인제대학교 전산학과 이학사
- 인제대학교 전산학과 전산학석사
- 인제대학교 전산학과 전산학박사

• 관심분야 : 패턴인식, 컴퓨터 비전



김 상 균 (Sang-Kyoon Kim)

- 정회원
- 경북대학교 통계학과 이학사
- 경북대학교 컴퓨터공학과 공학석사
- 경북대학교 컴퓨터공학과 공학박사

• 인제대학교 공과대학 컴퓨터공학부 교수

• 관심분야 : 패턴인식, 컴퓨터 비전



황 구 선 (Goo-Seun Hang)

- 정회원
- 인제대학교 컴퓨터공학부 공학사
- 인제대학교 전산학과 전산학석사
- 진코퍼레이션 비전팀 주임

• 관심분야 : 패턴인식, 컴퓨터 비전



안 상 호 (Sang-Ho Ahn)

- 정회원
- 경북대학교 전자공학과 공학사
- 경북대학교 전자공학과 공학석사
- 경북대학교 전자공학과 공학박사

• 인제대학교 공과대학 전자지능로봇공학과 교수

• 관심분야 : 영상공학, 로봇공학, IRCM, 적외선 영상처리



강 병 두 (Byoung-Doo Kang)

- 정회원
- 인제대학교 전산학과 이학사
- 인제대학교 전산학과 전산학석사
- 인제대학교 전산학과 전산학박사
- 전자부품연구원 지능로보틱스 연구센터 연구원
- 관심분야 : 패턴인식, 컴퓨터 비전, SLAM

논문 접수일 : 2011년 10월 20일

1차수정완료일 : 2011년 11월 15일

2차수정완료일 : 2011년 11월 25일

게재확정일 : 2011년 11월 29일