

# 계층적 능동형태 모델을 이용한 비정형 객체의 움직임 예측형 실시간 추적

## (Hierarchical Active Shape Model-based Motion Estimation for Real-time Tracking of Non-rigid Object)

강진영\*, 이성원\*\*, 신정호\*\*, 백준기\*\*

(Jinyoung Kang, Seongwon Lee, Jeongho Shin, and Joonki Paik)

### 요약

본 논문에서는 비정형 객체를 능동형태 모델을 사용하여 실시간으로 추적하기 위한 방법을 제시하였다. 객체를 추적 할 때, 가려진 부분의 윤곽을 추정해 낼 수 있는 능동형태 모델을 사용하였으며, 비디오의 각 프레임에서 처리과정의 시간을 줄이기 위해서 영상을 계층적으로 분리하여 실시간 처리를 가능하게 하였다. 또한 다음 입력영상의 초기 윤곽을 효율적으로 찾기 위해서 칼만필터(Kalman filter)를 사용하여 특징점을 예측하였고, 블록 정합(block matching) 기법을 추가하여 예측 안정성을 향상시켰다. 비 계층적 방법, 비 예측 방법 등과 비교 실험을 통해서 제안된 계층적, 예측형 방식이 수렴속도 증가와 모델링의 정확도에서 모두 개선된 효과를 얻을 수 있음을 확인하였다.

### Abstract

In this paper we proposed a hierarchical ASM for real-time tracking of non-rigid objects. For tracking an object we used ASM for estimating object contour possibly with occlusion. Moreover, to reduce the processing time we used hierarchical approach for real-time tracking. In the next frame we estimated the initial feature point by using Kalman filter. We also added block matching algorithm for increasing accuracy of the estimation. The proposed hierarchical, prediction-based approach was proven to out perform the exiting non-hierarchical, non-prediction methods.

**Keywords :** Active Shape Model, Block matching, Kalman filter, Hierarchical, Non-rigid object tracking

### I. 서론

영상 합성, 애니메이션을 위한 모션캡처, 보안 감시 시스템, HCI(human computer interaction), 등 많은 분야에서 비디오의 움직이는 객체를 인식하고 추적하는 기술들이 널리 응용되고 있다. 기존의 모델 기반 컴퓨터 비전 기법은 정형(rigid) 객체를 대상으로 하여 그 성능을 인정받았으나<sup>[1]</sup>, 움직이는 사람과 같이 연속되는 프레임에서 추적하는 물체의 형태가 변하는 경우에는

비정형(non-rigid) 모델을 이용하는 기술이 필요하다<sup>[2,3,4]</sup>. 비정형 모델 혹은 변형가능(deformable) 모델을 이용한 추출 및 추적 시스템은 객체의 변형에 강하기 때문에 가림현상(occlusion) 등을 극복할 수 있다.

초창기에 의료영상 분야에 성공적으로 사용되었던 비정형 모델의 일종인 능동 형태 모델(active shape model, ASM)은 객체 외형의 정보를 학습시키고 이 정보를 바탕으로 객체의 평균 형태에 변형을 가하여 새로운 영상 내에서 객체를 찾는 방법이다<sup>[5,6]</sup>. 하지만 매 프레임마다 객체의 윤곽을 모델과 정합시키기 위하여 반복적인 계산이 필요하고, 다음 프레임의 객체의 윤곽의 이동정보를 예측하기 어렵기 때문에 ASM을 상용화된 비디오 시스템에 사용하기에는 실시간 구현이 어렵다는

\* 학생회원, \*\* 정회원, 중앙대학교  
(Chung-Ang University)

※ 이 연구는 과학기술부 국가지정연구실 사업과 교육부 BK21 사업 지원으로 수행되었음

접수일자: 2004년7월8일, 수정완료일: 2004년7월24일

단점을 가지고 있다.

본 논문에서는 ASM 기반 객체추적 기법의 뛰어난 성능을 유지하면서, 동시에 실시간 구현이 가능한 시스템을 구현하기 위해서 계층구조와 예측기법들을 사용한 새로운 ASM 기반 추적기술을 제안하였다. 본 제안에서는, 계층적 영상을 사용함으로써 전체적인 정보량을 줄여 매 프레임 마다 반복적인 계산량을 줄이고, Kalman filter<sup>[7]</sup>, 블록정합 등을 이용하여 전 프레임의 객체의 윤곽정보를 이용하여 다음 프레임의 윤곽정보의 초기 값을 예측함으로써 계산량의 감소에도 불구하고 물체추적의 정확성을 높일 수 있게 하였다.

본 논문의 II장에서는 ASM의 기본적 개념 설명, III장에서는 제안된 계층적 기법의 소개, IV장은 제안된 예측기법에 대해서 설명을 하고, V장에서 실험결과를, 마지막으로 VI장에서 결론을 내리도록 할 것이다.

## II. 능동형태 모델(ASM)

연속적인 비디오 영상에서 비정형 객체를 추적하기 위해서는 기본적으로 객체의 형태와 위치를 추출하는 작업이 필요하다. 대부분의 2차원의 변형 가능한 모델은 객체를 표현하기 위해 경계의 정보를 이용한다. 변형 가능한 모델을 추적함에 있어, 추적하려는 객체에 대한 사전정보가 적용 가능한 조건에서는 정확성이나 효율성면에서 ASM은 최적의 접근방법 중 하나이다.

ASM은 찾고자 하는 객체 외곽 형태의 선형정보(a priori knowledge)를 이용한다. 따라서 불완전하거나 잡음이 많이 들어간 영상에서도 강건하게 객체의 외곽을 정합하여 추출할 수 있다. 이 모델은 훈련 집합(training set)을 필요로 하는데, 사용자가 직접 자동 혹은 수동으로 훈련 영상으로부터 객체의 외곽에 기준점들을 표시하여 여러 가지 형태 모델을 구축해야 한다. 기준점은 또한 객체의 외곽 특징을 잘 나타낼 수 있어야 한다. 예를 들면, 눈의 코너점(corner point)등이 좋은 기준점이 될 수 있다.

구체적으로 본 논문의 ASM 알고리즘은 크게 네 가지 단계로 나눌 수 있다. 즉 (i) 기준점(landmark point)의 할당 (ii) PCA(principal component analysis), (iii) 지역구조(local structure) 모델링, 그리고 (iv) 모델 정합 등으로 구성된다.

### 1. 기준점(landmark point) 할당

주어진 입력영상으로부터 적당한 기준점을 객체의 윤

곽선에 할당한다. 이 작업은 주로 수동으로 이루어진다. 바람직한 기준점은 한 영상에서 기준점으로 선택된 특징점이 다음 영상에서도 특징점으로 연결되어 있어야 한다. 2차원 영상에서 우리는 n 개의 기준점을 2n차원의 벡터를 사용해서 다음과 같이 표현할 수 있다.

$$x = [x_1, \dots, x_n, y_1, \dots, y_n]^T. \quad (1)$$

본 논문에서 적용될 실험에 사용하는 기준점의 수는 n=42로 했다. 기준점을 자동적으로 결정하는 방법은 Tian 등<sup>[8]</sup>에 의해 논의되었다. 기준점의 역할은 모델의 윤곽 형태를 조절하는 것이다. 즉 초기값으로 설정된 기준점은 계속 유지되는 것이 아니라, 각각의 기준점에서 경계의 수직방향에 따라 추적하는 객체의 실제 윤곽과의 편차를 최소화하는 방향으로 갱신된다.

### 2. PCA(Principal Component Analysis)

n 개의 기준점의 집합은 객체의 형태를 표현한다. 훈련 집합에서 각각의 형태는 2차원 공간에서 존재하지만, 우리는 PCA 기술을 기반으로 하여 좀 더 적은 수의 파라미터를 가지고 객체의 형태를 모델링 할 수 있다<sup>[11]</sup>. 훈련 집합에서 m개의 형태를 가지고 있다고 가정하자. 여기서  $x_i$ 는 각각의 형태에 따른 훈련 집합을 의미 한다 ( $i=1, \dots, m$ ).

단계별로 PCA 알고리즘을 살펴보면 다음과 같다.

1. 훈련 집합에서 m개의 표본 형태의 평균을 구한다.

$$\bar{x} = \frac{1}{m} \sum_{i=1}^m x_i \quad (2)$$

2. 훈련 집합으로부터 공분산 행렬(covariance matrix)을 구한다.

$$S = \frac{1}{m} \sum_{i=1}^m (x_i - \bar{x})(x_i - \bar{x})^T. \quad (3)$$

3. 공분산 행렬로부터 다음과 같은 값을 얻을 수 있다.

$$\Phi = [\phi_1 | \phi_2 | \dots | \phi_q], \quad (4)$$

여기서  $\phi_j$ (단,  $j=1, \dots, q$ )는 S에서 q개의 가장 큰 고유값(eigen value)에 해당하는 고유벡터(eigen vector)를 의미한다.

4. 주어진  $\Phi$ 와  $\bar{x}$ 로부터 각각의 형태는 다음과 같이 간략화 시킬 수 있다.

$$x_i \approx \bar{x} + \Phi b_i \quad (5)$$

여기서

$$b_i = \Phi^T(x_i - \bar{x}). \tag{6}$$

벡터  $b$ 는 변형 가능한 모델 파라미터의 집합을 정의한다.  $b$ 의 요소가 변하면서 우리는 형태를 변화시킬 수 있다. PCA 알고리즘의 3번째 단계에서  $q$ 값은 데이터의 전체적인 분산값에 98%에 해당되는 모델형태의 개수를 나타낸다.

### 3. 지역구조 모델링 (local structure modeling)

ASM을 사용하여 입력영상의 형태를 해석하기 위해서는 영상과 모델을 최적으로 정합시키는 기준점의 집합을 찾아야 한다. 즉 각 기준점에 해당하는 점들에서 물체와 수직이 되는 방향으로 각각의 밝기 값들을 구하여 비교하는 것이다.

아래의 그림 1은 이러한 방법을 표기한 것이다.

### 4. 모델 정합(model fitting)

우리는 훈련 집합에 의한 모델의 형태를 실제 비디오 영상의 객체와 정합하여 최적의 위치와 형태를 갖는 파라미터를 얻어야한다. 이것은 모델 좌표프레임의 형태인  $x$ 와 영상 좌표프레임의 형태인  $y$ 를 정합하기위해 다음과 같은 여러 함수를 최소화함으로써 이루어질 수 있다.

$$E = (y - Mx)^T W (y - Mx), \tag{7}$$

여기서  $W$ 는 대각행렬이며, 그 대각성분은 각각 기준점에 대한 가중치를 나타낸다. 그리고  $M$ 은 회전성분, 전이성분  $t$ , 스케일링 성분  $s$ 로 이루어진 기하학적인 변환을 나타내는 행렬이다. 가중치 요소는 객체의 윤곽을

따라 선택된 이전의 기준점과 갱신된 기준점 사이 값의 변위에 관련이 있다. 변위가 크면 클수록 행렬에 해당되는 가중요소들은 작아지고, 반면에 변위가 작아질수록 가중요소들은 커진다. 하나의 점을  $[x_0, y_0]^T$ 으로 표현한다면, 기하학적인 변환행렬은 다음과 같이 정의 될 수 있다.

$$M \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = s \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}. \tag{8}$$

기하학적인 파라미터의 집합( $\theta, t, s$ )이 결정되면, 모델좌표 프레임으로 투영된  $y$ 값은 다음과 같이 주어진다.

$$x_p = M^{-1}y. \tag{9}$$

결국 모델의 파라미터는 다음과 같이 갱신된다.

$$b = \Phi^T(x_p - \bar{x}). \tag{10}$$

윤곽을 따라 위 단계들의 결과로써 최적의 기준점의 변위를 얻을 수 있다. 최적으로 갱신된 기준점의 조합은 영상좌표 프레임  $y$ 에 새로운 형태를 만들어낸다. 새로운 형태는 식 (7)을 사용하여 가장 근접한 형태를 찾는 데 이용될 수 있다.  $M$ 을 통해서 최적의 위치를 도출한 뒤, 이 새로운 형태는 주어진 훈련 집합에 주요 요소에 포함된  $\Phi$ 로 투영될 수 있다. 이 과정은 모델파라미터  $b$ 를 갱신한다. 그 결과로 주요 요소에 해당되는 비슷한 한 편차만이 모델 파라미터에 영향을 줄 수 있다. 모델 파라미터를 계산한 후에  $x$ 에 의해 주어지는 새로운 형태는 식 (5)에 의해 구할 수 있고, 이 새로운 형태는 식 (7)에 의해 다음 반복과정을 위해 사용된다. 이러한 과정은 식(7)이 최소값을 가질때까지의 반복을 통해서 최종적인 형태가  $x$ 로 얻어진다.

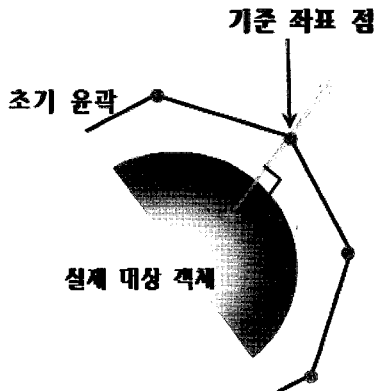


그림 1. 지역구조 모델링  
Fig. 1. Local Structure Modeling.

### III. 계층적 능동형태 모델

ASM의 계층적인 접근은 여러 가지 형태와 크기의 객체를 다루는 데 있어서 필수적이다<sup>[9]</sup>. 계층적 모델은 비 계층적 모델의 방법에 비해 계산량이나 성능면에 있어 상당히 효과적이고 우월하다. Baumberg는 Kalman filter가 형태를 정합하는데 있어서의 성능 향상을 위해 계층적 스네이크 모델에 어떻게 적용될 수 있는지 연구했다<sup>[10]</sup>. 구체적으로는 계층적으로 표본화의 정도를 작은 값에서 큰 값으로 변화시켜가면서 기준점

의 수를 변화시켰다. 본 논문에서는 입력 영상의 정보와 각 기준점간의 상관관계만을 고려하는 스네이크 알고리즘 대신에 사전 학습을 통해 가려진 부분에도 강건한 ASM방법을 사용하였다. Cootes는 자신의 ASM을 계층적으로 확장하였다<sup>[11]</sup>. 그가 제안한 방식은 일반적인 계층적 방식과 마찬가지로 저해상도에서 해를 구하고 고해상도로 옮겨간 다음 그 해를 수정한다.

본 논문에서는 Baumberg의 형태정합에 쓰인 Kalman filter 이룬 접근방법과 Cootes의 계층적인 접근방법을 물체 추적에 적용하여 계층적 예측형 ASM 알고리즘을 제안하였다.

1. 계층적 접근

실시간으로 비정형 객체를 추적하기 위해서는 초기 입력 영상에서 다음 영상이 들어오기까지의 시간동안에 각 영상에서 모두 처리가 이루어져야한다. 따라서 각각의 시간 간격의 영상에서 ASM의 처리 시간을 줄이기 위해서 피라미드 모델을 사용하였다. 영상 분석에 있어서 피라미드 모델을 적용하는 방법은 윤곽을 검출하기 위해서 Tanimoto와 Pavlidis에 의해 처음 소개되었다. 피라미드 모델의 가장 중요한 성질은 비-피라미드 모델의 방법에 비해 계산량이나 성능면에 있어 효과적이라는 것이다. 칼라스테레오 영상을 가지고 한 실험에서도 스테레오 영상을 정합을 하는 데 있어 비계층적인 방법보다는 계층적인 방법이 일반적으로 좀 더 정확한 것임을 증명했다. 게다가 계산량에 있어서도 현저히 감소됨을 보였다<sup>[8]</sup>.

아래의 그림 2는 본 논문에서 사용된 각 단계별 영상의 크기를 나타낸 것이다.

IV. 예측 방법

ASM 수행이 이루어지기 전에 초기 윤곽형태를 이동시켜주는 것은 정확도 향상에 중요하다. 본 논문에서는

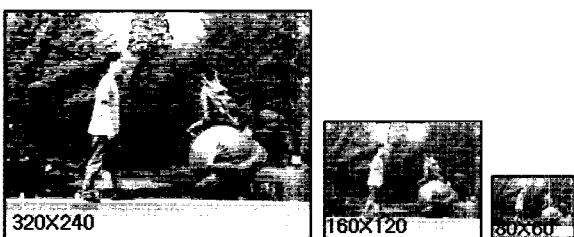


그림 2. 계층적 영상의 단계  
Fig. 2. Hierarchical images.

연속되는 영상에서 전 프레임의 각 기준점에서 다음 프레임에서의 움직임을 추정하여 초기 윤곽형태를 이동시키는 기법을 사용한다.

움직임 추정 기법은 객체의 특징 정보를 검출하고 추적하는 방법에 따라 여러 가지 방법들이 있다<sup>[12,13]</sup>. 본 논문에서는 Kalman filter와, 블록 정합 기법을 이용한 예측 방법을 사용 하였다.

1. 칼만 필터(Kalman filter)

칼만 필터란 잡음에 의해 간섭 받는 선형 동적 시스템(linear dynamic system)에서 상태 벡터  $x_k$ 의 최적의 추정치  $\hat{x}_k$ 를 구하기 위한 재귀(recursive) 알고리즘이다.

여기서 k는 시간을 나타낸다. 칼만 필터링은 크게 예측(prediction), 측정(measurement), 수정(update)의 세 단계로 나누어진다.

예측단계에서는 상태 벡터  $\hat{x}_k^-$ 과 상태 추정 오차의 공분산 행렬  $P_k^-$ 의 예측이 이루어진다. (-)는 아직 측정을 통한 수정 단계가 이루어지지 않았음을 나타낸다. 아래 그림 3은 칼만 필터의 상태도를 나타낸 것이다.

예측의 단계가 끝나고 나면 측정 단계에서 칼만 이득,  $K_k$ 을 계산한다. 이 칼만 이득의 값과 예측단계의 상태벡터  $\hat{x}_k^-$ 으로 상태 벡터  $\hat{x}_k$ 를 수정한다. 이때  $z_k$ 는 측정 벡터, H는 상태 벡터와 측정 벡터 사이의 관계를 나타내는 관측 행렬(observation matrix)를 나타낸다. 또한 다음 단계의 예측단계를 위한 오차 공분산 행렬도 수정하게 된다. 이러한 과정이 순환함으로써 칼만 필터가 완성되게 된다.

본 논문에서는 미리 찾아진 기준점들을 측정벡터로

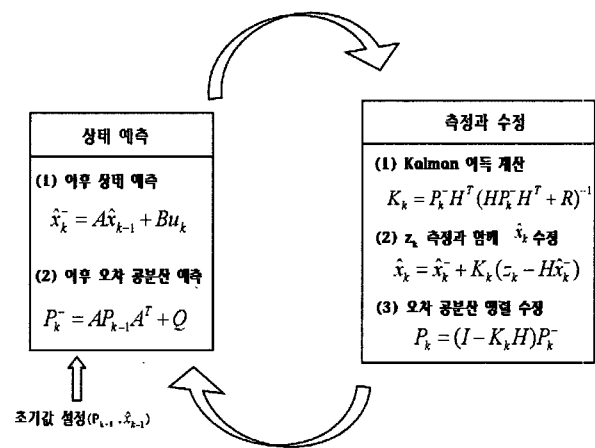


그림 3. 칼만 필터 상태도  
Fig. 3. Kalman filter state diagram.

하고 다음 프레임에서 찾아야할 기준점들을 상태벡터로 놓고 칼만 필터 예측 알고리즘을 사용하여 윤곽을 형성한다.

2. 블록 정합(Block matching)

연속된 영상에서 움직이는 물체는 영상간의 화소값을 변화시키며, 이 화소값 변화는 물체의 이동값을 나타내는 수학적 모델의 계수를 추정하는데 사용된다. 이때 물체의 이동을 나타내는 수학적 모델로 움직임 보상 부호화에서 사용되고 있는 블록정합 알고리즘은 주로 상관관계를 이용한다.

일반적으로 움직이는 물체에는 회전 운동 또는 크기의 확대나 축소가 있을 수 있으나 블록정합 알고리즘은 시간 경과에 따른 물체의 변형, 즉 이전 프레임에서의 물체는 현재 프레임에서 변형이 많이 생기지 않고 물체의 움직임은 단지 그 물체를 구성하는 블록들의 움직임으로 근사 시킬 수 있다는 전제하에 사용되는 움직임 추정 방식으로 주로 동영상 혹은 스테레오 영상의 압축에 많이 응용되고 있다<sup>[14]</sup>. 블록정합은 탐색 영역 결정 방식에 따라 전역 탐색(full search), 3단계(three step search)방식, 병렬 계층 1차원 탐색(parallel hierarchical 1-D search)방식 및 계층적 움직임 추정(hierarchical motion estimation)방식 등이 있다. 이러한 방식들은 움직임 벡터를 찾는 방식에서 차이가 있을 뿐 이전 프레임의 가장 적합한 블록을 찾는다는 점에서 동일하다<sup>[15]</sup>.

이전 프레임에서 가장 적합한 블록을 찾는 기준은 아래의 식(11)의 비용함수(cost function)로 표현할 수 있다.

$$E(i, j) = \frac{1}{MN} \sum_{k,l=0}^{M-1,N-1} (c - b)^2, \tag{11}$$

$$c = C(x + k, y + l),$$

$$b = B(x + i + k, y + j + l),$$

$$-p \leq i, j \leq p,$$

여기서  $p$ 는 탐색영역을 나타내고,  $C(x + k, y + l)$ 는 현재 프레임의 매크로 블록내의 화소를 나타낸다.

$B(x + i + k, y + j + l)$ 는 이전 프레임의 정합블록내의 화소이다. 따라서 식(11)로부터 비용함수가 최소값을 갖는  $(i, j)$ 를 움직임 벡터로 정의한다.

본 논문에서는 전 프레임의 윤곽의 각 기준점들을 다음 프레임에서 블록정합을 이용하여 초기 윤곽 위치의 정확성을 높였다.

V. 실험결과

제안된 계층적 ASM과 그를 이용한 실시간 추적을 실험하기 위해서 SONY DCR-TRV900 3CCD 카메라로 촬영된 320\*240, raw gray영상을 사용하였다.

그림 4는 실험에 이용된 영상의 첫 번째 프레임을 나타낸다. 그림 5(a)는 초기 영상의 초기 윤곽을 나타낸 것이며, 그림 5(b)는 이 영상에서 반복수행 없이 한번의 ASM과정을 통해 얻은 결과 영상이다.

1. 움직임 예측이 포함 되지 않는 경우

다음 프레임의 움직임 예측 없이 이전 프레임의 결과 윤곽을 다음 프레임의 초기 윤곽으로 하고 ASM을 수행한 결과 이다. 영상에서 각 프레임별로 큰 움직임이 존재하지 않기 때문에 첫 번째 프레임에서 바로 다섯 번째 프레임의 영상의 윤곽을 찾는 것을 실험하였다.

그림 6(a)(b)(c)는 영상의 해상도에 따라 320\*240, 160\*120, 80\*60의 영상에서 움직임 예측 없이 전 프레임의 ASM수행 결과의 윤곽을 그대로 다섯 번째 프레임의 초기 윤곽으로 하고 ASM을 수행한 결과 영상이다.



그림 4. 실험영상의 첫 프레임  
Fig. 4. The first frame of the test sequence.

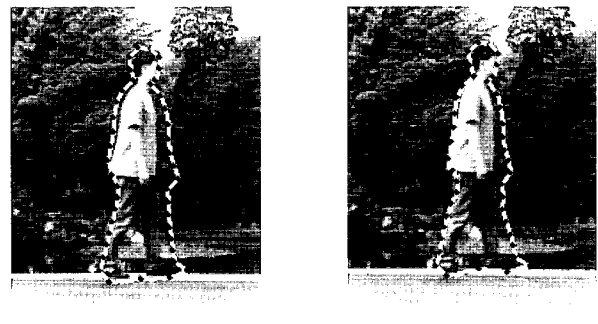


그림 5. (a)초기윤곽 과 (b)ASM 1회 실행결과  
Fig. 5. (a)Initial and (b)the modeling result one-time ASM.

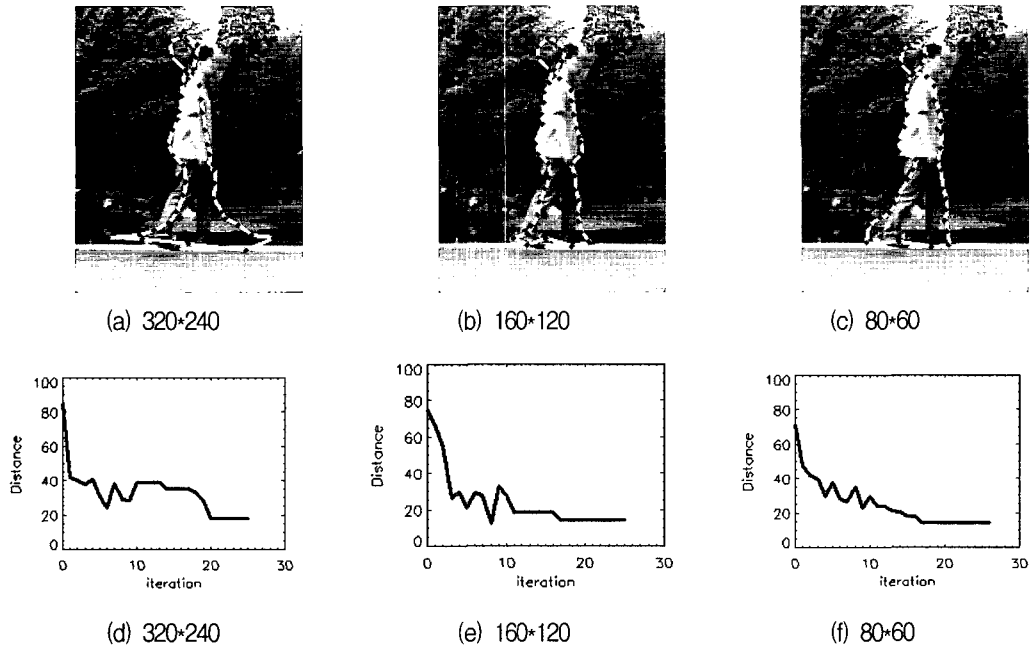


그림 6. 움직임 예측을 수행하지 않은 경우 결과 영상 윤곽((a),(b),(c)), 각 반복 수행 당 거리차 그래프((d),(e),(f))  
 Fig. 6. Resulting shape using each resolution((a),(b),(c)) and the distance graph without motion prediction.

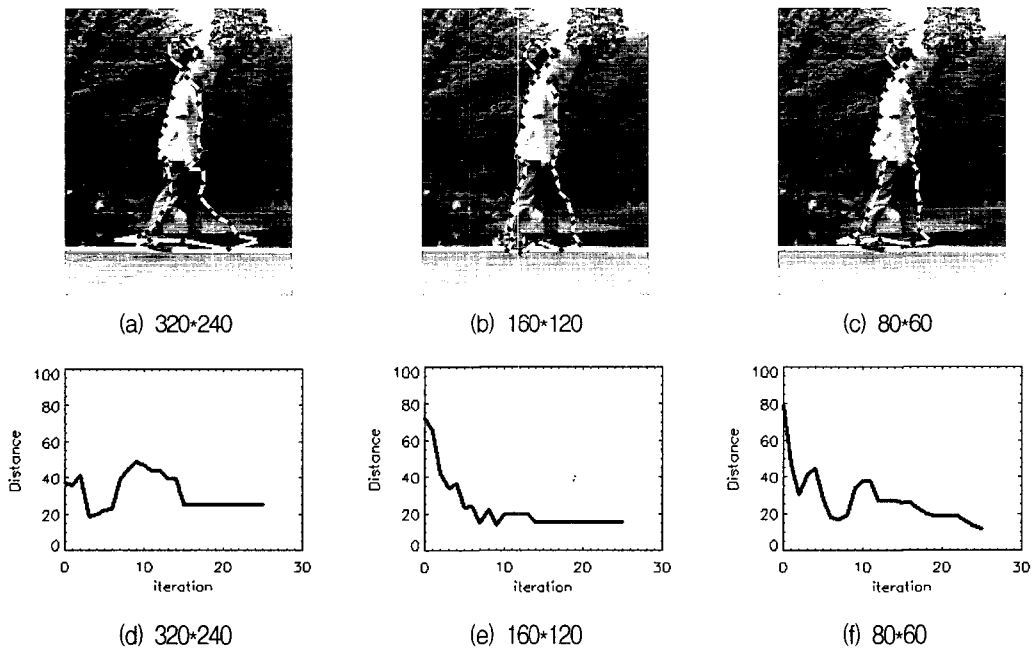


그림 7. 모든 기준점의 중심점을 Kalman filter로 예측한 각 단계별 영상의 결과윤곽((a),(b),(c)), 각 반복 수행 당 거리차 그래프((d),(e),(f))

Fig. 7. In each resolution (a)resulting Kalman prediction of the center point of total landmark points and (b)the distance graph.

그림 6(d)(e)(f)는 각 해상도에서 반복 수행을 거치면서 각 반복에 따라 윤곽의 기준점들의 유클리드거리 변화를 그래프로 나타낸 것이다. 모델이 수렴하기 까지 20회 이상의 반복이 수행이 되어야 한다. 또한 320\*240크기의 영상에서는 실험한 윤곽이 머리 부분과 발 부분에서 실제 윤곽과 크게 어긋나는 결과를 보이고 있다.

제안된 계층적 모델에서 실험한 결과가 원영상의 실

험 결과보다 윤곽의 안정성, 즉 기준점들의 거리 변화가 가장 작아지는 반복횟수가 줄어들며, 결과의 윤곽 또한 훨씬 정확하다.

2. 움직임 예측을 수행한 경우

이 절에서는 움직임 예측을 사용하여 연속되는 프레임의 초기값을 갱신한 실험한 결과를 제시한다. 먼저, 윤

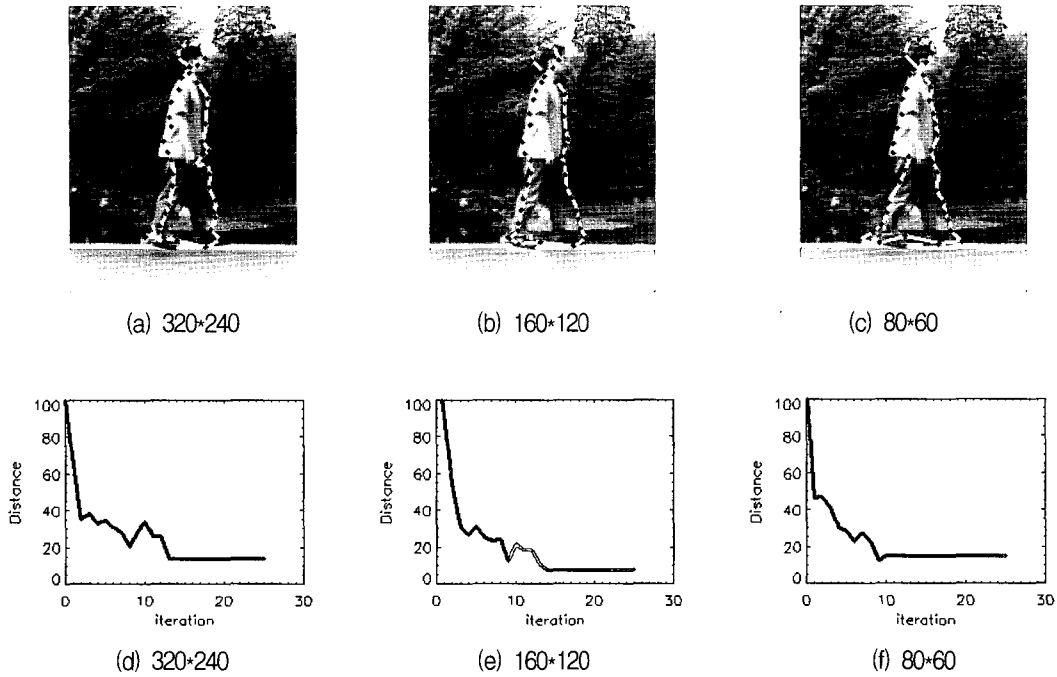


그림 8. 상하좌우 네 점의 중심점을 Kalman filter로 예측한 각 단계별 영상의 결과윤곽((a),(b),(c)), 각 반복 수행 당 거리차 그래프((d),(e),(f))  
 Fig. 8. In each resolution (a)resulting shape (b)the distance graph with Kalman prediction of the center point of left, right, up and bottom.

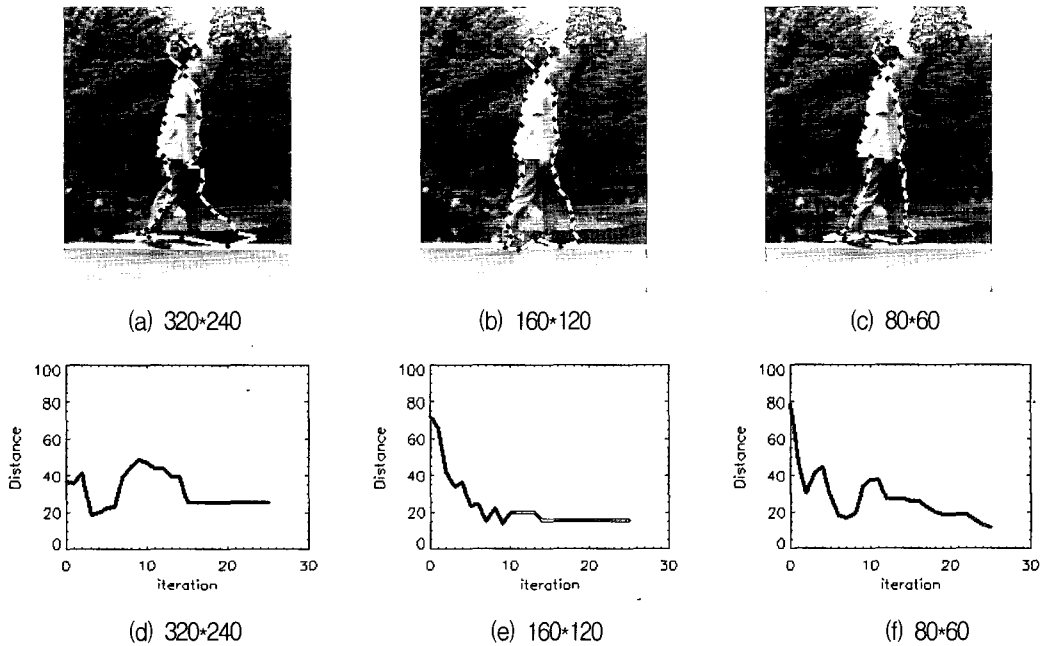


그림 9. 모든 기준점을 Kalman filter로 예측한 각 단계별 영상의 결과윤곽((a),(b),(c)), 각 반복 수행 당 거리차 그래프 ((d),(e),(f))  
 Fig. 9. In each resolution (a)Resulting shape and (b)the distance graph with Kalman prediction of the total landmark points.

곽의 전체 기준점의 평균점을 구하고 이점에 해당되는 연속된 프레임의 대응점을 Kalman filter를 사용하여 예측하였다. 그리고 이 예측된 점과 전 프레임의 평균점과의 거리를 구하고, 이 거리를 모든 기준점들에 적용시켜

서 다음 프레임의 초기 윤곽으로 정한다.

그림 7 (a)(b)(c)는 영상의 해상도에 따라 320\*240, 160\*120, 80\*60의 영상에서 실험한 결과 영상이고, 그림 7(d)(e)(f)는 각 반복 회수 당 기준점들의 유클리드 거

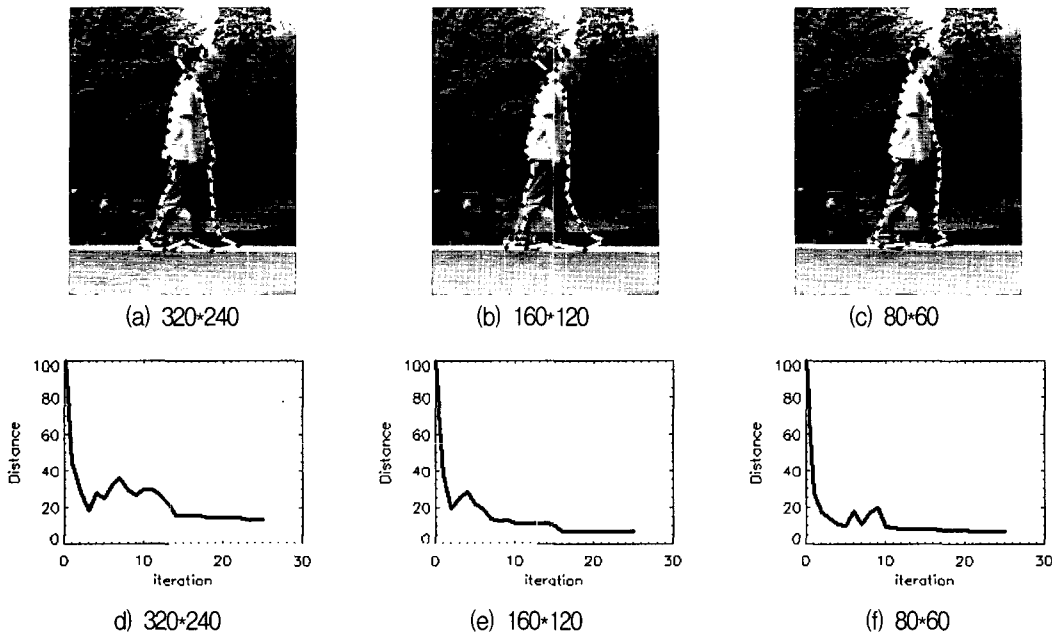


그림 10. 모든 기준점을 블록정합으로 예측한 각 단계별 영상의 결과윤곽((a),(b),(c)), 각 반복 수행 당 거리차 그래프 ((d),(e),(f))

Fig. 10. In each resolution (a)Resulting shape and (b)the distance graph with block matching prediction of the total landmark points.



그림 11. 초기 영상과 윤곽  
Fig. 11. Initial image and shape.

리를 그래프로 나타낸 것이다. 그림 8은 윤곽의 전체 기준점에서 상하좌우의 가장 큰 좌표 네 개를 추출하여 이 점들의 평균을 구하고 그림 7과 마찬가지로 각 해상도에서 실험한 결과와 그래프이다. 그림 9는 모든 윤곽의 기준점을 Kalman filter로 다음 프레임의 움직임을 예측하여 이 예측된 점들을 초기 윤곽으로 정하고 그림 7과 마찬가지로 각 해상도에서 ASM을 수행한 결과 영상과 거리 차를 나타낸 그래프이다. 그림 10은 모든 기준점들을 이용하여 다음 프레임(5th 프레임)의 블록정합을 이용하여 다음 프레임의 초기 윤곽을 구하고 ASM을 반복하여 수행한 후의 결과 윤곽과 그래프이다.

실험결과를 보면 제안된 계층적 예측형 알고리즘이 비 계층적 비 예측 알고리즘 보다 반복의 횟수가 작고 윤곽의 정확도가 높은 것을 알 수 있다.

위의 실험들은 첫 번째 프레임에서 곧 바로 다섯 번째 프레임의 영상의 윤곽을 찾는 것이다. 그래서 유클리드

거리가 최소가 되는 때까지는 최소 5회 이상의 반복이 필요함을 알 수 있다.

표 1은 각 실험에 적용된 방법에 따라서 비교 한 것이다. 이 표를 보면 제안된 계층적 예측형 방법이 반복 횟수와 최소 값에 도달 했을 때 각 기준점들이 움직인 유클리드 거리가 작음을 알 수 있다. 형태의 안정성 또한 높아진다.

그래서 그림 12는 블록 정합을 이용하여 각 프레임마다 연속으로 추적되는 영상을 몇 개의 프레임만을 추출한 것과 최초로 가장 적은 거리차를 보일 때의 반복회수 그래프이다. 처음 최소 값으로 떨어진 후 반복을 중단하고 다음 프레임으로 넘어가게 된다. 실시간 추적을 위해서는 반복되는 회수는 프레임 별로 같아야한다. 따라서 반복되는 회수는 3으로 제한하였다. 표 1에서는 최소의 반복 회수는 5이지만, 이 실험은 첫 프레임에서 다섯 번째 프레임임을 감안 하여야 한다.

그림 11은 추적할 영상의 첫 프레임에서 초기 윤곽을 나타내고 있다.

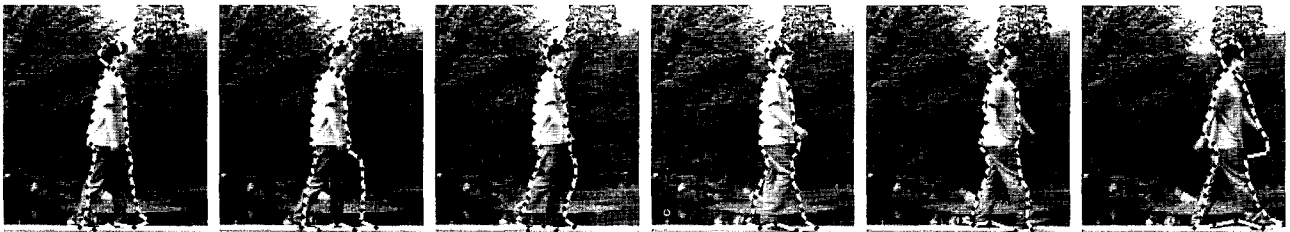
### VI. 결 론

본 논문에서는 ASM을 이용하여 비정형 객체를 실시간으로 추적하는 방법을 제시하였다. 초기 영상에서 ASM 수행을 거친 후 물체의 윤곽을 찾고 이를 바탕으로

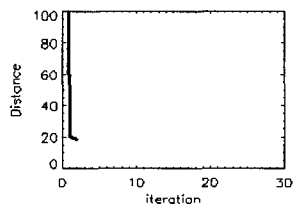


표 1. 각 적용이론에 따른 비교  
Table 1. The comparison by each method.

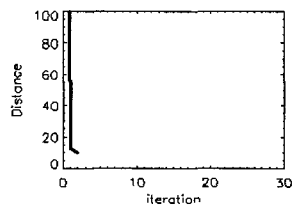
움직임 예측 유무		적용이론	해상도 (pixel)	첫 번째 최소 거리까지 반복 횟수	유클리드 거리 (pixel)	형태의 정확도 (주관적 판단)		
움직임 예측 없을 경우		전 프레임의 ASM 결과윤곽을 초기 윤곽으로 정합	320*240	20	20	머리, 다리, 몸의 일부분 실제 물체와 오차 큼		
			160*120	17	15	몸 윤곽만 정확		
			80*60	17	15			
제안된 방식	전체 기준점의 평균	Kalman filter 예측	320*240	15	25	머리, 다리가 실제 물체와 오차 큼		
			160*120	9	15	머리·다리 일부 오차		
			80*60	7	15	머리의 일부분 오차		
	네 점의 평균점		320*240	13	13	① 사람의 뒤쪽 일부 오차		
			160*120	13	8	'①' 보다 오차가 적음		
			80*60	9	8	머리 일부 약간의 오차		
	모든 기준점		320*240	13	25	머리, 다리 큰 오차		
				160*120	9	15	머리일부, 다리 일부 오차	
				80*60	7	15	머리 일부 오차	
			모든 기준점	블록 정합	320*240	14	15	머리일부 오차
					160*120	10	5	머리 일부 오차
					80*60	5	5	거의 정확함



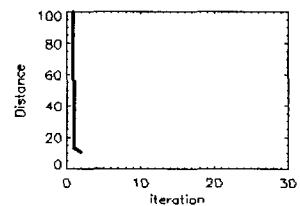
(a) 3<sup>th</sup> frame      (b) 6<sup>th</sup> frame      (c) 9<sup>th</sup> frame      (d) 16<sup>th</sup> frame      (e) 21<sup>th</sup> frame      (f) 25<sup>th</sup> frame



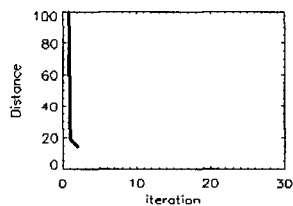
(g) 3<sup>th</sup> frame



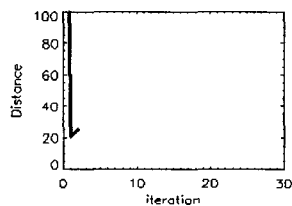
(h) 6<sup>th</sup> frame



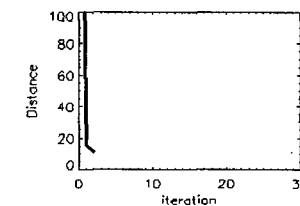
(i) 9<sup>th</sup> frame



(j) 16<sup>th</sup> frame



(k) 21<sup>th</sup> frame



(l) 25<sup>th</sup> frame

그림 12. 각 프레임에서 추적되는 윤곽((a)~(f))과 거리차 그래프((g)~(l))

Fig. 12. In each frame (a)~(f)The resulting tracking shape and (g)~(l)the distance graph.

로 Kalman filter와 블록정합에 의해서 다음 프레임의 움직임을 예측하여 초기 윤곽을 찾은 후 ASM 수행과정을 거친다. 본 논문에서는 ASM 수행의 반복 횟수와 결과 윤곽 추출을 통해서 이루어진다. 계층적 영상의 이용이 물체를 실시간으로 추적 하는데 있어서 반복회수를 줄이고, 정확도를 향상 시킬 수 있다는 것을 보였다. 예측하는 방법에서 모든 기준점의 평균점, 상하좌우 네 점의 평균점을 각각 Kalman filter로 다음 점을 예측하여 그 이동 거리만큼을 모든 기준점들에 적용 시켜서 이를 ASM을 수행하는 것 보다는 모든 점들을 Kalman filter로 예측하여 초기 윤곽을 나타내어 ASM을 수행하는 것이 수렴정도를 나타내는 유클리드의 거리 변화가 적게 나타나고, 수렴되기까지 반복 횟수도 적은 결과를 얻을 수 있다. 또한 각 점의 위치에 의해 예측되는 Kalman filter 알고리즘을 사용하여 ASM을 수행하는 것보다 영상의 정보를 분석하여 예측하는 블록정합 방법의 결과가 영상의 윤곽의 정확성과 수렴되는 정도가 반복횟수로 인해 결과가 더 정확하다는 것을 확인하였다. 또한 제안된 계층적으로 예측된 방법이 반복되는 횟수를 줄인다는 것을 보였다. 이로써 실시간으로 ASM을 이용하여 물체를 효과적으로 추적해 낼 수 있다.

실험 영상에서 움직임이 많거나 보다 복잡한 영상에 적용하기 위해서는 훈련 집합의 다양성이 존재하여야 한다. 움직임이 많다는 것은 다양한 자세를 가지고 있다고 볼 수 있기 때문이다. 또한 사용자의 의도나 혹은 더 많은 영상의 정보를 획득하여, 예를 들어 칼라의 변화나 조명의 변화를 감지 할 때 물체의 윤곽을 저장하고 좀더 자세한 윤곽의 정보가 필요할 때 적응적으로 계층적 모델을 적용하는 연구가 진행되어야 할 것이다.

## 참고 문헌

- [1] Y. Boykov, D.P. Huttenlocher, "Adaptive Bayesian recognition in tracking rigid objects," *Computer Vision and Pattern Recognition*, pp. 697-704, Vol.2 2000.
- [2] I. Haritaoglu, D. Hartwood, and L. S. Davis, "Real-time surveillance of people and their activities," *IEEE Trans. PAMI*, Vol. 22, pp.809 -830, 2000.
- [3] S. J. McKenna, Y. Raja, S. Gong, "Tracking colour objects using adaptive mixture models," *Image and Vision Computing*, vol. 17, pp.225-231, 1999.
- [4] R. Plankers, P. Fua, "Tracking and modeling people in video sequences," *Comp. Vision and Image Understanding*, vol. 81, pp. 285-302, 2001.
- [5] T. J. Cootes, C. J. Taylor, D. H. Cooper, and J. Gragam, "Training models of shape form sets of examples," *In British Machine Vision Conference*, pp.9-18, September 1992.
- [6] T. F. Cootes, A. Hill, C. J. Taylor C.J., and J. Haslam, "The use of active shape models for locating structures in medical images," *Information Processing in Medical imaging*, pp.33-47, 1993.
- [7] G. Welch, G. Bishop, "An Introduction to the Kalman Filter," *Technical Report, Department of Comp. Sc. and Engg., Univ. of North Carolina at Chapel Hill*, 2002.
- [8] S. Tanimoto and T. Pavlidis, "A hierarchical data structure for picture processing," *Comput. Graphics Image Process.* vol. 4, pp. 104-119, 1996.
- [9] A. Koschan, S. Kang, J. Paik, B. Abidi, and M. Abidi, "Color active shape models for tracking non-rigid objects," *Pattern Recognition Letters*, Vol. 24, no. 11, pp. 1751-1765, July 2003.
- [10] A. Baumberg, "Hierarchical shape fitting using an iterated linear filter," *Image and Vision Computing*, vol. 16, pp. 329-335, 1996.
- [11] T. F. Cootes, C. J. Taylor, and A. Lanitis, "Active shape models: evaluation of a multi-resolution method for improving image search," *Proc. British Machine Vision Conference*, pp. 327-336, 1994.
- [12] S. M. Smith, "Reviews of Optical Flow, Motion Segmentation, Edge Finding and Corner Finding," *Technical Report, Dept. of Clinical Neurology, Oxford University*, 1997.
- [13] S. Araki, T. Matsuoka, H. Takemura, and N. Yokoya, "Real-time Tracking of Multiple Moving Objects in Moving Camera Image Sequence," *IEICE Trans. Inf. & Syst.* Vol. E83-D, No7, 2000.
- [14] Chin-Chen Chang, Lin-Li Chen, Tung-Shou Chen, "An improvement of bottom-up variable-sized block matching technique for video compression," *IEEE Transactions on Consumer Electronics*, Vol.44, No. 4, pp.1234 - 1242, 1998.
- [15] A. Murat Tekalp, "Digital Video Processing," Prentice hall signal processing series, 1995.

— 저 자 소 개 —



**강 진 영**(학생회원)  
 2003년 울산대학교 전자공학과  
 학사졸업.  
 2004년 현재 중앙대학교 첨단  
 영상대학원 영상공학과  
 석사과정.

<주관심분야: 모델기반 객체추적, 영상압축>



**이 성 원**(정회원)  
 1988년 서울대학교 제어계측  
 공학과 학사졸업.  
 1990년 서울대학교 제어계측  
 공학과 석사졸업  
 2003년 University of Southern  
 California 전기공학과  
 박사졸업

2004년 현재 중앙대학교 첨단영상대학원  
영상공학과 연구교수.

<주관심분야: 미디어프로세서 및 SOC설계, 멀티  
미디어 신호처리, Power-Aware Computing>



**신 정 호**(정회원)  
 1994년 중앙대학교 전자공학과  
 학사졸업.  
 1998년 중앙대학교 전자공학과  
 석사졸업.  
 2001년 중앙대학교  
 첨단영상대학원  
 영상공학과 박사졸업

2004년 현재 중앙대학교 첨단영상대학원  
영상공학과 연구교수

<주관심분야: 영상처리, 영상복원, 해상도개선, 영  
상융합, 컴퓨터 비전>



**백 준 기**(정회원)  
 1984년 서울대학교 제어계측  
 공학과 학사졸업.  
 1987년 노스웨스턴대학교 전기 및  
 컴퓨터 공학과 석사졸업.  
 1990년 노스웨스턴대학교 전기 및  
 컴퓨터 공학과 박사졸업.

2004년 현재 중앙대학교 첨단영상대학원  
영상공학과 교수

<주관심분야: 영상복원, 신호처리, 반도체>