

# 윤곽선 재조정을 통한 의미 있는 객체 추적 알고리즘

## A Semantic Video Object Tracking Algorithm Using Contour Refinement

임정은, 이재연, 박현상, 나종범

한국과학기술원 전기 및 전자 공학과 영상시스템 연구실

Jungeun Lim, Jaeyoun Yi, Hyun Sang Park, and Jong Beom Ra

Dept. of Electrical Engineering, KAIST

E-mail: jbra@ee.kaist.ac.kr

### 요약

본 논문에서는 동영상에서 의미 있는 객체 영역을 추출하기 위해, 사람이 첫 장에서 관심 있는 객체를 표시하고, 그 다음 프레임부터는 사람의 도움 없이 객체를 추적하는 반자동 (semi-automatic) 방식의 객체 추적 알고리즘에 대해 개선된 알고리즘을 제안하고 이를 구현하였다. 제안한 객체 추적 알고리즘은 이전 프레임의 객체의 움직임 정보를 이용하여 현재 프레임에서 대략적인 객체의 위치를 찾은 후, 윤곽선의 불확실도를 조사하고, 윤곽선을 재조정하여 정확한 객체의 위치를 찾는다. 제안한 알고리즘은 다양한 영상에서 만족할 만한 결과를 얻었다.

### 1. 서론

컴퓨터와 신호처리 기술이 발달함에 따라 영상과 음향의 응용 범위는 통신과 전송 뿐만 아니라 멀티미디어 분야까지 확장하게 되었다. 영상과 음향 압축 기법인 MPEG-4에서는 VOP(Video Object Plane)이라는 새로운 개념을 도입하였다 [1]. 이 개념은 동영상을 독립적인 객체 영상들의 집합으로 간주하여 영상을 각각의 객체들로 분리하여 처리함을 의미한다. 그리고, 이 객체를 VO(Video Object)라고 한다.

이와 같은 이유에서 비디오 영상 내에서 의미 있는 객체를 VOP로 분리하는 영상 분할 기법의 개발이 소개, 연구되고 있다.

기존에 소개된 영상 분할 (image segmentation) 기법은 균질한 (homogeneous) 영역을 추출하여 압축하거나 [2,3], 동영상의 공간(spatial) 정보와 시간(temporal) 정보를 함께 이용하여 영상을 분할하는 방법 등이었다 [4]. 그러나, 두 가지 방법 모두 여러 색상과 질감, 그리고 움직임 정보를 함께 갖는 비디오 객체를 자동적으로 추출하기에는 아직 많은 어려움을 가지고 있다.

따라서 최근 의미 있는 영상을 추출하기 위해 반자동 방법이 소개되었다. 이 방법은 비디오의 첫 프레임의 영상 분할은 사람의 지식에 기반을 하여 수행하고, 그 다음 프레임부터는 사람의 도움 없이 자동으로 객체를 추적한다. 간혹 추적 결과가 어긋났다고 사람이 판단하는 경우는 잘못된 프레임에서 다시 사람의 지식에 기반하여 영상 분할을 한 후, 다시 자동으로 객체를 추적하도록 한다 [5].

이 논문에서는 동영상에서 의미 있는 객체를 추출하기 위해 이전 프레임의 추적 결과로부터 움직임 정보를 이용하여 현재 프레임의 대략적인 객체의 윤곽선을 구하고, 현재 프레임의 공간정보 즉, 밝기나 색상 등을 이용하여 앞

서 구한 윤곽선을 재조정하는 방법을 제안한다. 이 논문의 구성은 다음과 같다. 먼저 2장에서 제안하는 객체 추적 알고리즘에 관하여 설명하고, 3장에서 실험 결과를 보이고, 4장에서 결론을 맺는다.

## 2. 본문

### 2.1 전체구조

그림 1은 제안하는 객체 추적 알고리즘의 전체적인 구조를 보여 주고 있다. 제안하는 구조는 크게 3 단계로 이루어져 있다. 첫 단계에서 매개변수 움직임 모델을 이용하여 이전 프레임의 결과 객체의 윤곽선(contour)을 현재 프레임에 투영하고, 두 번째 단계에서 투영된 객체의 윤곽선의 부근 영역에서 투영의 결과가 불분명한 영역을 추출한다. 투영의 결과가 불분명한 영역은 움직임 보상 오차(motion estimation error)와 투영된 객체의 윤곽선에 인접한 배경과 객체의 밝기 비교 두 과정을 통해 추출한다 그리고, 마지막 단계에서 투영이 불분명한 영역과 그 부근을 다시 재조정함으로써

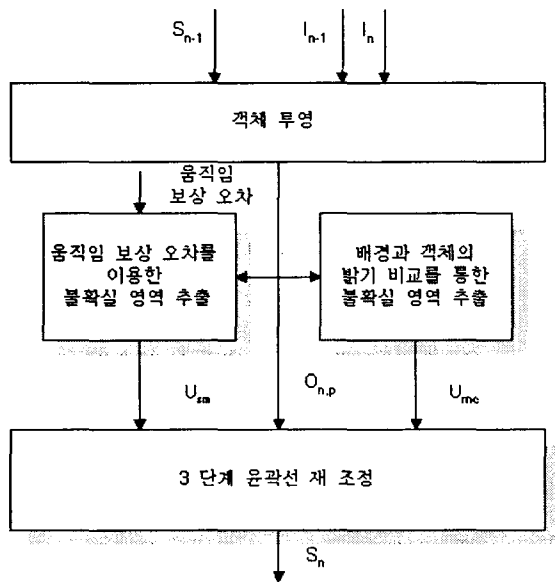


그림 1. 객체 추적 알고리즘의 전체 구조

써 정확한 객체의 윤곽선을 구한다. 윤곽선 재조정은 3 단계로 이루어져 있다.

### 2.2 객체 윤곽선 투영

객체 윤곽선 투영(temporal projection)은 움직임 모델(motion model)을 이용하여 이전 프레임 결과를 현재 프레임으로 투영하는 과정이다. 이전 프레임에서 현재 프레임으로 객체의 이동은 매개변수 움직임 모델(parametric motion model)을 이용하여 기술할 수 있다. 매개변수 움직임 모델은 물체를 강체(rigid object)라고 고려하는데, 동영상에서 연속하는 프레임 사이의 시간 간격은 매우 짧으므로 물체의 모양은 크게 변화하지 않을 것이다. 그러므로, 움직임이 다양한 객체라 할지라도 충분히 하나의 움직임 모델로 표현할 수 있을 것이며, 대략적인 객체의 현재프레임에서 위치를 알 수 있다. 특히, 이 논문에서는 6개의 매개변수를 갖는 affine 움직임을 이용하였다.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_4 & a_5 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} a_3 \\ a_6 \end{bmatrix}, \quad (x', y') \in O_{n-1} \quad (1)$$

Affine 움직임 모델은 식 (1)과 같다. 여기에서  $(x, y)$ 는 현재 프레임의 화소의 위치  $(x', y')$ 는 이전 프레임의 화소의 위치를 나타낸다. 그리고,  $O_{n-1}$ 는 이전 프레임의 결과  $S_{n-1}$ 에서 객체를 의미한다.

대체로 모델의 매개변수는 실수이므로, 이전 프레임의 각 화소에 대해 현재 프레임의 대응하는 화소의 좌표는 실수 값을 갖게 된다. 그래서, 이 논문에서는 그림 2에서 설명하는 바와 같이 이전 프레임의 객체  $O_{n-1}$ 의 윤곽선을 구성하고 있는 각 화소를 affine 움직임을 이용하여 현재 프레임으로 옮기고 이 실수 좌표를 둘러싸고 있는 4 화소를 모두 객체의 윤곽선으로 결정하였다. 그리고, 그 결과가  $O_{n,p}$ 이다.

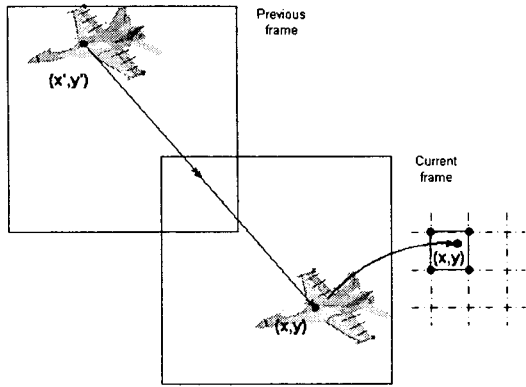


그림 2. 객체 윤곽선 투영

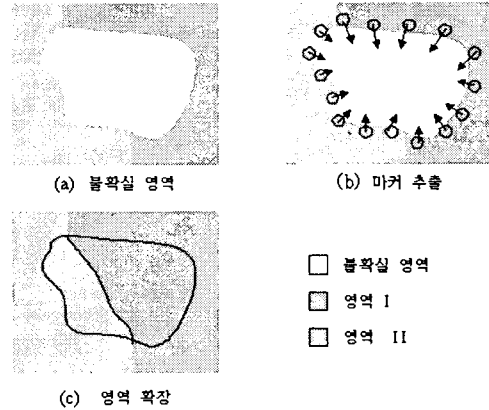


그림 3. 영역 확장

### 2.3 영역 확장 (Region extension)

영상 분할에서 영상은 여러 작은 영역들로 분할되어 표현된다. 이 분할된 각 영역은 각기 다른 값을 지정함으로써 다른 영역임을 표시할 수 있다. 그림 3(a)는 영역 분할이 결정된 '영역 I, II'와 아직 결정되지 않은 "불확실 영역 (uncertain region)"을 표시하고 있다. 그리고, 불확실 영역 내 화소들은 '영역 I, II'를 불확실 영역으로 확장하여 영역 결정을 할 수 있다. 이 때, 영역 확장은 **watershed** 알고리즘을 이용한다. 영역 확장은 화소의 밝기 값의 차이가 작을수록, 더 가까운 화소 쪽으로 확장이 되어야 하므로, **watershed** 알고리즘을 변형하여 그림 3(b)에서 설명하는 바와 같이 불확실 영역을 둘러싸고 있는 모든 화소들을 마커로 선택하고, '영역 I'에 포함되는 화소부터 확장된 불확실 영역 내의 모든 화소는 '영역 I'으로 결정함으로써 영역 확장을 실행할 수 있다. 그림 3(c)는 영역 확장의 결과로 불확실 영역 내의 화소들을 영역 결정한 그림이다.

### 2.4 투영이 불분명한 영역 추출

이 단계는 마지막 윤곽선 재조정 단계에서 재조정할 영역을 추출하는 단계이다. 곧 2.2의 결과에서 윤곽선 투영 결과가 어긋났다고 고려되는 부분은 재조정의 정도를 높게 하고, 반대

로 윤곽선 투영이 비교적 옳게 된 부분은 윤곽선 재조정을 낮은 단계로 하도록 한다. 이 과정은 움직임 보상 오차를 이용한 단계와 투영된 객체의 윤곽선에 인접한 배경과 객체의 화소 밝기 비교의 두 단계로 구성되어 있으며, 투영이 불분명한 화소들은 두 단계에서 구한 영역을 모두 포함한다.

#### 2.4.1 움직임 보상 오차

2.2에서 매개 변수 움직임 모델을 이용하여 대략적인 추적 결과  $O_{n,p}$ 를 구할 수 있었다.  $I_n$ 에서  $O_{n,p}$ 를 제외한 부분을  $B_{n,p}$ 라 하면 이 두 영역  $O_{n,p}$ 와  $B_{n,p}$ 에 대해서 각각 이전 프레임  $I_{n-1}$ 로 매개 변수 움직임 모델을 구할 수 있다. 그리고, 이 두 움직임 모델을 이용하면 식 (2)와 같이 현재  $n$ 번째 프레임의 움직임 보상 오차 영상  $e_n$ 을 얻을 수 있다.

$$e_n(x, y) = \min_j |I_n(x, y) - I_{n-1}(x'_j, y'_j)| \quad (2)$$

$$\text{where } (x'_j, y'_j) \in R(x', y'), (x, y) \in I_n$$

여기에서  $(x', y')$ 는 현재 프레임의 화소  $(x, y)$ 가 매개 변수 움직임 모델로 이동한 이전 프레임의 실수형 화소이며,  $R(x', y')$ 는  $(x', y')$ 를 둘러싸고 있는 4개의 정수형 화소를 뜻한다.

움직임 보상 오차 영상  $e_n$ 을 이용하여 윤곽선 근처의 화소에 대해서 투영이 불분명한

화소를 추출할 수 있다. 특히, 화소  $(x, y)$ 에서 움직임 추정이 옳지 않아 움직임 보상 오차가 크거나 혹은 객체의 움직임에 따라서, 새로 나타난 배경 부근의 화소를 추출할 수 있을 것이다. 그림 4(a)는  $O_{n,p}$ 와  $B_{n,p}$ 의 움직임 모델에 대한 보상 오차 영상의 이진(binary) 영상이다. 움직임 보상 오차를 구하는 데에 있어, 영상의 잡음에 의한 화소 선택이 될 수 있기 때문에 오차가 문턱치(threshold)이상인 화소만을 선택한다. 그리고, 이진 오차 영상에 대해서 윤곽선에서 일정한 폭 내의 오차가 큰 화소만을 투영이 불분명한 영역으로 고려한다. 그 결과가 그림 4(b)인  $U_{mc}$ 이다.

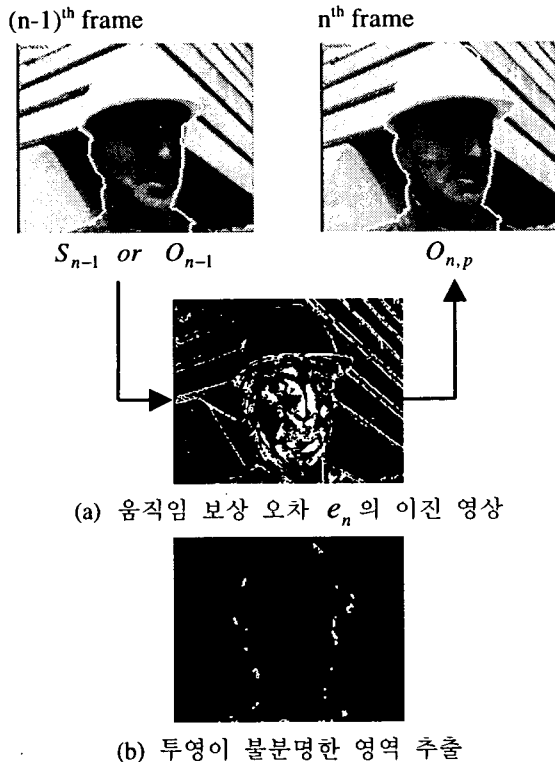
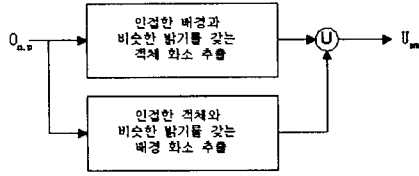


그림 4. 움직임보상오차를 이용한 불분명영역 추출

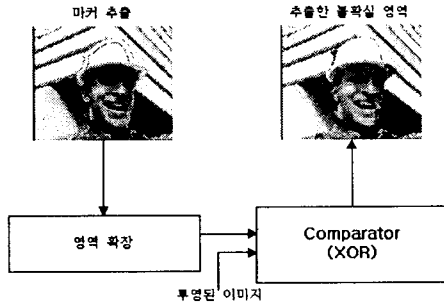
#### 2.4.2 인접한 배경과 객체의 화소밝기 비교

만약 배경과 객체가 확실한 밝기 차이가 있다고 가정한다면, 객체의 투영이 잘못된 경우에는 투영된 객체의 윤곽선에 인접한 객체와 배경의 밝기는 비슷할 것이다. 이 과정에서는

$O_{n,p}$ 의 윤곽선 부근에서 밝기가 비슷한 배경과 객체의 영역을 추출한다. 그림 5는 이 과정을 설명하고 있다. 그림 5(a)는 이 과정이 첫째, 배경과 비슷한 밝기를 가지는 객체 영역 추출과 둘째, 객체와 비슷한 밝기를 가진 배경 영역 추출의 두 단계로 분류하여 실행함을 보여주고 있다. 그리고, 각 추출 단계는 2.3에서 설명한 영역 확장을 이용한다. 그림 5(b)는 이 과정을 더 자세히 설명하고 있다. 그림 5(b)는 배경과 비슷한 밝기를 가지는 객체 영역 추출 단계를 설명한 것이다. 우선 배경과 객체를 의미하는 화소들을 마커(marker)로 추출한다. 그림에서 볼 수 있는 것과 같이  $O_{n,p}$ 의 윤곽선에 접한 배경의 화소들을 배경을 대표하는 마커,  $O_{n,p}$  객체 내부의 화소들을 객체를 나타내는 마커로 선택한다. 그리고, 두 마커 사이의 영역을 불확실 영역으로 고려하고 하여 영역 확장을 한다. 이 때, watershed 알고리즘의 우선 순위(priority)는 마커와의 밝기 차이이며, 밝기 차이가 작을수록 거리가 가까울수록 더 빨리 영역 결정이 이루어진다. 또, 이 경우 모든 불확실 영역의 화소가 결정될 때까지 영역 확장을 하는 것이 아니라, 우선 순위의 제한을 줌으로써 비슷한 밝기를 가지는 화소에 대해서만 영역을 결정하도록 한다. 이렇게 얻어진 결과와 투영된 결과  $O_{n,p}$ 와 비교하여 다르게 결정된 화소, 즉 투영의 결과에서는 배경이었으나 객체와 비슷한 밝기를 가진 화소는 투영이 불분명한 화소로 추출한다. 배경 영역 역시 비슷한 방법으로 객체를 나타내는 마커를 투영된 객체  $O_{n,p}$ 의 윤곽선에서, 배경을 나타내는 마커는 배경 내부의 화소로 고려함으로써, 객체와 비슷한 밝기를 가지는 화소를 추출할 수 있다. 그리고, 두 단계에서 추출된 화소들의 합집합은 윤곽선 부근의 비슷한 밝기를 가지는 배경과 객체 화소들인 투영이 불분명한 영역  $U_{sm}$ 이다.



(a) 화소 밝기 비교를 통한 불분명한 영역 추출



(b) 인접한 배경과 비슷한 밝기를 갖는 객체 화소 추출의 예

그림 5. 화소 밝기 비교를 통한 불분명한 영역 추출

## 2.5 영역 확장을 이용한 객체의 윤곽선 결정

앞 2.3 과 2.4 에서 설명한 투영이 불분명한 영역  $U_{proj}$  과 그 부근의 화소들은 영역 확장을 이용한 3 단계 윤곽선 결정을 통해 최종 결정을 내린다. 윤곽선 결정은 컬러 RGB 값을 이용하였다. 그림 6 은 제안한 알고리즘을 설명한다.

첫 단계는 수리 형태학(**mathematical morphology**)을 이용하여  $U_{proj}$  을 일정한 폭만큼 확장하고, 이 영역을 불확실 영역으로 고려하고 영역확장을 통해 결정한다. 그리고, 다음 두 단계에서는 이전 단계의 영역확장을 통한 결과  $S_{n,i}$  와 투영 결과  $O_{n,p}$  를 비교하여 서로 다르게 결정된 화소에 대해서 영역 확장을 통하여 재결정을 한다. 만약 객체의 윤곽선 부근에서 배경이나 객체가 복잡한 가장자리(**edge**)를 가지고 있거나 객체의 색상이 배경과 비슷하면 단 한번의 윤곽선 결정은 잘못된 결과를 낳을 수 있다. 앞서 구한 투영 결과  $O_{n,p}$  는 윤곽선 부분에서는 정확한 결과는 아니나, 객체의 모

양은 유지한다고 볼 수 있다. 이 점을 이용하여, 잘못 결정될 수 있는 영역을 여러 단계를 뒤서 다시 고려함으로써, 기존의 방법에 비해 정확한 객체의 윤곽선을 얻을 수 있는 가능성이 증가할 수 있다. 또, 단계가 진행될 때마다, 수리형태학의 형태소(**structure element**)의 크기를 감소시켜서 불확실 영역의 크기를 점차 줄여나간다. 이와 같이 여러 단계로 윤곽선 재조정을 함으로써 배경이 복잡한 텍스처(**texture**)인 경우 결과 윤곽선이 잘못된 가장자리로 향하는 것을 막을 수 있다.

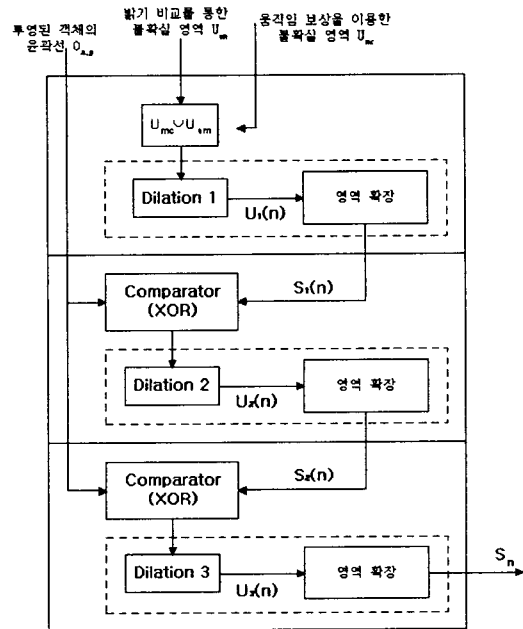


그림 6. 윤곽선 재조정

## 3. 모의 실험

이 논문에서 제안한 객체 추적 알고리즘의 실험을 위해서, MPEG-2 및 MPEG-4 의 표준화 과정에 사용된 테스트 영상 중 'Miss America', 'Akiyo', 'Foreman', 'Mother and daughter', 'Clair' 영상을 사용하였다. 각 실험 영상은 모두 176×144 크기의 QCIF 30Hz 영상을 사용했으며, 실험 영상의 intra 프레임의 영상 분할 결과는 사람에 의해서 직접 주어진 것으로 사

용하였다. 그리고, 모든 실험은 0번 프레임을 **intra** 프레임으로 하였다. 그림 7은 제안한 알고리즘을 이용한 추적 결과를 보여주고 있다. 움직임 보상 오차: 투영된 객체 윤곽선부터 폭 2내의 화소만을 고려 하였다. 그리고, 밝기 비교 단계에서는: 밝기 차가 10이하인 화소만을 선택하였으며, 객체와 배경을 의미하는 마커추출은 각각 투영된 윤곽선에서 폭 4인 화소들로 추출하였다.

윤곽선 재조정 단계에서 수리형태학 불림 (**dilation**)의 형태소 크기는 첫 단계와 두 번째 단계는 2이고 마지막 단계는 1로 실행하였다.

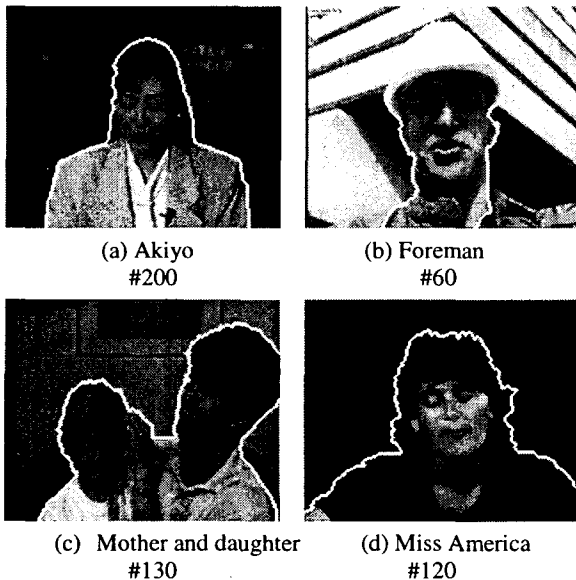


그림 7. 실험 결과

#### 4. 결론

이 논문은 비디오 영상에서 반자동으로 의미 있는 객체를 추적하는 방법을 제안했다. 제안한 알고리즘은 이전 프레임의 결과를 이용하여 현재프레임의 대략적인 결과를 얻고 그 결과를 수정함으로써, 정확한 결과를 얻고자 하였다. 그리고, 실험 결과에 의해서 제안한 알고리즘으로 의미 있는 객체의 추적이 비교적 잘 수행됨을 확인할 수 있었다. 그러나, 객체의 움직임이 심한 경우는 객체의 일부분을 잃는 경우가 있었고, 또 객체가 복잡하고 윤곽선 부근

에서 컬러 정보가 다양하거나 새로 객체가 나타나는 부분은 옳은 결과를 얻지 못했다. 앞으로 이 문제를 해결하기 위하여 객체 기반의 연구가 좀 더 보충되고, 윤곽선 재조정 단계에서 이전 프레임의 정보를 사용하는 기법이 더 연구된다면 좋은 성능의 객체 추적 알고리즘 구현이 가능할 것이다.

#### 5. 참고 문헌

- [1] ISO/IEC JTC1/SC29/WG11 N1730, "MPEG-4 Overview," Stockholm MPEG meeting, July 1997.
- [2] M. Kunt, A. Ikonomopoulos, and M. Kocher, "Second generation image-coding techniques," *Proc. IEEE*, vol. 73, no. 11, pp. 549-574, Apr. 1985
- [3] H. S. Park and J. B. Ra, "Efficient marker extraction algorithm for fine image segmentation," *IEE Electronic Letters*, vol. 34, no. 22, pp. 2107-2108, Oct. 1998.
- [4] J. G. Choi, S. Jee, and S.D.Kim, "Spatio-temporal video segmentation using a joint similarity measure," *IEEE Trans. Circuits Syst. Video Technol*, vol. 7, no. 2, Apr. 1997.
- [5] J. Yi, H.S. Park, and J.B. Ra, "A hierarchical semantic video object tracking algorithm using mathematical morphology," *Intern. Workshop on Advanced Image Technology*, vol. 1, pp. 29-34, Cheju, Korea, Jan. 14-15, 1998.

본 연구는 산자부 지원에 의해 수행되었음