

# 프레임차 에너지의 전위차를 이용한 영역 기반의 비디오 객체 추출

정회원 꺾 종 인\*, 김 남 철\*\*

## Region-Based Video Object Extraction Using Potential of Frame-Difference Energies

Jong-In Kwak\*, and Nam-Chul Kim\*\* *Regular Members*

### 요 약

본 논문에서는 프레임차 에너지의 전위차를 이용한 비용으로 비디오 객체를 추출하는 영역 기반 분할 알고리즘을 제안한다. 첫 번째 단계에서는 파티션의 영역 내에 비디오 객체의 윤곽이 포함되지 않도록 공간적인 밝기 값을 이용하여 동질한 영역들로 세밀하게 분할을 한다. 이렇게 세밀하게 분할된 파티션은 두 번째 단계인 시공간 분할의 초기 파티션이 된다. 시공간 분할에서는 각각의 인접한 영역들의 동질성 비용을 두 영역의 프레임차 에너지 중에서 작은 에너지를 가지는 영역의 프레임차 에너지와 두 영역에 의하여 만들어지는 윤곽상에 존재하는 프레임차 에너지로 계산한다. 다음에 동질성의 비용이 최소인 두 영역을 찾아서 병합하고, 파티션을 새롭게 갱신한다. 이러한 반복적 병합은 프레임차 에너지의 전위차가 큰 윤곽들이 남을 때까지 수행한다. 마지막으로 후처리 단계에서는 객체 내부의 윤곽들을 제거하여 비디오 객체를 추출한다.

### ABSTRACT

This paper proposes a region-based segmentation algorithm for extracting a video object by using the cost of potential of frame-difference energies. In the first step of a region-based segmentation using spatial intensity, each frame is segmented into a partition of homogeneous regions finely so that each region does not contain the contour of a video object. The fine partition is used as an initial partition for the second step of spatio-temporal segmentation. In spatio-temporal segmentation, the homogeneity cost for each pair of adjacent regions is computed which reflects the potential between the frame-difference energy on the common contour and the frame-difference energy of the lower potential region of the two. The pair of adjacent regions whose cost is minimal then is searched. The two regions of minimum cost are merged, which result in updating the partition. The merging is recursively performed until only the contours remain which have frame difference energies of high potential. In the final step of post-processing, the video object is extracted removing the contours inside the object.

### 1. 서 론

입력 영상을 동일한 특성을 갖는 영역들로 나누는 영상분할(segmentation)<sup>[1,2]</sup>은 지금까지 영상분석, 객체인식(object-recognition) 등의 컴퓨터 비전을 위한 기본 과정으로 사용되어 왔다. 최근에는 비디오

프레임들로부터 배경과 비디오 객체들로 구분하여 객체기반(object-based)으로 부호화하는 MPEG-4<sup>[3,4]</sup>에서 필수적인 전처리 과정으로 그 중요성이 강조되고 있다. 이에 따라 비디오 객체 추출을 목적으로 하는 여러 가지 방법들이 연구되었는데, 이들 중 많은 방법들에서 초기 조건을 만족하도록 세밀하게

\* 경북대학교 전자전기공학부(jikwak@vcl.knu.ac.kr),  
논문번호 : K01043-0127, 접수일자 : 2001년 1월 27일

\*\* 경북대학교 전자전기공학부(nckim@ee.knu.ac.kr)

분할한 영역들에 대해서 동질성의 비용(homogeneity cost)을 최소로 하는 인접하는 두 영역을 찾아서 병합하고 이를 반복하여 객체를 추출한다<sup>5,9)</sup>.

Kim 등의 방법<sup>7)</sup>에서는 먼저 현재 프레임을 모폴로지 와 워터셰드 알고리즘으로 분할하여 초기 분할 영역을 구하고, 이전 프레임과 현재 프레임간의 밝기 값의 변화가 큰 변화 영역(changed area)을 구한다. 다음, 초기 분할 영역 중 변화 영역에 포함되는 각 영역의 특징값을 평균 밝기 값과 프레임차 에너지의 가중합으로 구한다. 이어서, 동질성의 비용으로 인접하는 두 영역사이의 특징값의 차로 구하고, 이 동질성의 비용을 최소로 하는 인접한 두 영역을 차례로 병합해 간다. 그러나, 여기서는 동질성의 비용 계산에 필요한 가중치의 선정이 분할 결과에 큰 영향을 미칠 수 있는데, 이러한 가중치를 어떻게 설정하는가에 대한 것은 언급되어 있지 않다.

동질성의 비용을 Kim 등의 방법과 같이 구하는 또 다른 방법으로, Wang<sup>8)</sup>의 방법에서는 각 프레임을 워터셰드 알고리즘으로 분할하여 초기 분할 영역을 구하고, 분할된 각 영역에 대한 특징값을 움직임 보상 오차(motion compensation error)로 구한다. 이하 동질성의 비용과 병합단계는 Kim 등의 방법에서와 동일하게 수행한다. 이 방법에서는 분할 결과에 영향을 주는 여러 개의 역치를 실험적으로 결정하여 사용하고 있고, 또한 객체를 움직임 정보만을 이용하여 분할하는 데는 어려움이 있다.

Zhong 등의 방법<sup>9)</sup>에서는 첫 번째 프레임에서 사용자 개입으로 만들어진 대략적인 후보 객체 윤곽을 에너지를 최소화하는 방향으로 윤곽을 이동시키는 활동 윤곽 알고리즘(active contour algorithm : snakes)<sup>10)</sup>으로 객체 윤곽을 찾고, 객체 윤곽 내부를 칼라, 에지 및 움직임을 융합한 동질성의 비용으로 객체 분할을 한다. 두 번째 프레임부터는 이전 프레임에서 분할된 객체 내부의 영역을 현재 프레임으로 affine 모션 모델을 이용하여 영역별 움직임 투영(projection)<sup>11)</sup>을 한다. 이어서 움직임 투영한 모든 영역을 포함하는 확장된 경계 상자(expanded boundary box)를 만들고, 확장된 경계 상자 내의 영역들을 칼라, 에지 및 움직임을 융합한 동질성의 비용으로 객체의 윤곽을 추출한다. 여기서는 확장된 경계 박스의 경계에 걸쳐지는 영역들을 객체영역과 배경영역으로 나누기 위하여 역치를 사용하고 있는데, 역치를 어떻게 결정하는가에 대한 것은 언급이 되어 있지 않다.

지금까지 살펴본 바로 비디오 객체를 효과적으로

추출하기 위해서는 비디오 객체의 윤곽 특성을 잘 표현할 수 있는 동질성의 비용을 사용하되, 비용 내에 실험적으로 결정되어야 하는 역치 또는 파라미터의 수는 줄일 필요가 있다. 본 논문에서는 비디오 객체의 윤곽을 추출하기 위하여 비디오 객체의 윤곽을 특징짓는 프레임차 에너지의 전위차로 정의된 동질성의 비용으로 비디오 객체를 분할하는 영역 기반 자동 객체 분할 알고리즘을 제안한다. 제안한 알고리즘에서는 공간 분할, 시공간 분할, 후처리 과정으로 되어 있으며, 비디오 객체의 윤곽을 추출하기 위하여 공간 분할에서는 영역들이 비디오 객체의 윤곽을 포함하지 않도록 밝기 값으로 세밀하게 영역 기반 영상 분할을 한다. 이어서 공간 분할의 결과를 입력으로 하여 시공간 분할에서는 윤곽상에 존재하는 프레임차 에너지와 윤곽을 이루는 두 영역내의 프레임차 에너지 중에서 작은 에너지를 갖는 영역의 프레임차 에너지의 비율로 정의된 동질성의 비용으로 반복적 병합을 통해 비디오 객체를 추출한다. 실험에서는 제안한 알고리즘을 사용하여 Suzie와 Miss America 연속 영상에 대하여 비디오 객체를 추출한 결과를 제시한다.

## II. 제안한 비디오 객체 추출 방법

영상 분할은 입력영상  $I$ 를 동질성 및 연결성을 만족하는 영역들로 나누는 것으로, 분할된 결과는 파티션(partition)  $P=\{R_1, R_2, \dots, R_N\}$ 로 나타낼 수 있다. 여기서  $R_i \in P$ 인  $R_i$ 들은 동질성 및 연결성을 만족하고 분할된 영상  $P$ 는  $I$ 의 파티션이므로,  $R_i \cap R_j = \emptyset$  (단,  $i \neq j$ )와  $\bigcup_x R_x = I$  를 만족한다<sup>12)</sup>.

이때 동질성은 영역내의 모든 화소들이 공통의 성질을 갖고 있는 것이고, 연결성은 영역내 임의의 두 화소사이에 연결된 경로가 존재하는 것이다. 비디오 객체 분할은 움직임이나 밝기 값에 대한 동질성과 연결성을 만족하도록 연속 영상을 배경과 객체들로 분할하는 것이다. 최종적인 분할 결과를 얻는 여러 가지 방법중의 하나로 반복적 병합법<sup>13-17)</sup>이 있는데, 이 방법에서는 매 단계마다 최소의 비용을 갖는 두 영역을 찾아서 병합을 하고 파티션을 갱신해 나가면서 분할하고자 하는 단계까지 반복적으로 영역을 병합해 간다. 본 논문에서는 이러한 반복적 병합법을 사용하여 비디오 객체를 추출하고자 한다.

### 2.1 제안한 비디오 객체 추출 알고리즘의 개요

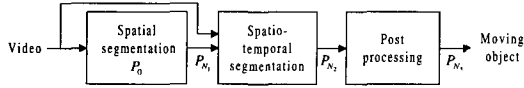


그림 1. 비디오 객체 추출을 위한 전체 구성도

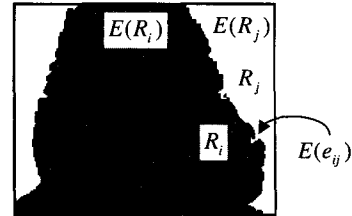
제안한 알고리즘의 구성은 그림 1에서처럼 공간 분할, 시공간 분할, 후처리 과정으로 이루어진다. 공간 분할에서는 현재 프레임의 초기 파티션  $P_0$ 로부터 비디오 객체의 윤곽이 파티션  $P_{N_1}$ 의 원소인 영역 내에 포함하지 않도록 세밀하게 분할한 파티션  $P_{N_1}$ 을 얻는다. 시공간 분할에서는 세밀하게 분할된 파티션  $P_{N_1}$ 을 입력으로 하여 비디오 객체 윤곽을 찾기 위하여 현재와 이전 프레임으로부터 시간적 정보(temporal information)인 프레임차를 구하고, 영역간 프레임차 에너지의 전위차를 최소화 하는 비용으로 반복적 병합을 통해서 프레임차 에너지의 전위차가 큰 객체 내부의 윤곽과 비디오 객체와 배경의 경계인 윤곽을 갖는 파티션  $P_{N_2}$ 를 얻는다. 후 처리에서는 영역의 크기를 이용한 비용을 최소화하는 반복적 병합으로 파티션  $P_{N_2}$ 에서 프레임차 전력이 큰 객체 내부의 윤곽을 제거하고, 비디오 객체 윤곽으로 객체와 배경이 분리된 파티션  $P_{N_3}$ 를 얻는다. 본 논문에서는 편의상 배경의 움직임이 없고 움직임 객체가 한 개인 경우만을 고려하는데, 이때  $P_{N_3}$ 은 움직임 객체와 배경의 두 영역으로 구성되어야 하므로 영역의 크기는  $|P_{N_3}|=2$ 가 된다.

### 2.2 공간 분할

공간 분할에서는 비디오 객체의 윤곽이 영역 내에 포함되지 않도록 세밀하게 분할하기 위하여 윌콕 기반 비용을 최소화 하는 반복적 병합법<sup>[17]</sup>을 선택하여 사용한다. 이 영상 분할법은 두 영역간의 이질성을 윌콕적 관점에서 고려한 Lagrange 비용함수를 정의하고, 각 단계마다 윌콕 기반 비용이 최소가 되는 가지의 인접 영역쌍을 병합함으로써 분할영상을 얻는 방법이다. 이때 윌콕 기반 비용  $\Delta J(e)$ 은  $t$ 단계에서 각 가지  $e$ 에 대하여 다음과 같이 정의된다.

$$\Delta J(e) = \Delta D(e) + \lambda_{t+1} \Delta R(e) \quad (1)$$

여기서  $\Delta D(e)$ ,  $\Delta R(e)$ 는 임의의 가지  $e$ 에 연결된 두 영역을 병합하면서 발생하는 왜곡(distortion)의 변화량 및 윌(rate)의 변화량을 나타내고,  $\lambda_{t+1}$ 는



- : High frame-difference energy
- : Low frame-difference energy

그림 2. 임의의 한 파티션내의 인접한 두 영역이 만드는 윤곽상에서의 프레임차 에너지의 전위차

$t+1$ 단계에서의 Lagrange 상수이다. 입력 영상을 이러한 윌콕 기반 비용을 이용한 반복적 병합법으로 세밀하게 분할함으로써 객체의 윤곽을 포함하지 않는 영역들로 분할되어 시공간 분할의 입력으로 적합한 결과를 얻을 수 있다.

### 2.3 시공간 분할

시공간 분할 과정을 설명하기에 앞서, 윤곽을 이루는 두 영역간의 프레임차 에너지의 전위차가 크고, 또한 윤곽상의 프레임차 에너지가 큰 비디오 객체의 윤곽을 검출하는 시공간 분할의 기본 원리를 그림 2를 통하여 살펴보기로 하자. 그림 2는 이전과 현재 프레임간의 프레임차를 공간 평균한 SAFD (spatial averaging of frame difference)를 구하고, 시공간 분할 과정내의 임의의 한 파티션에 속하는 각 영역에 대하여 프레임차 에너지를 구한 영상을 보여 준다. 진한 부분은 프레임 에너지가 큰 것을 나타내고, 연한 부분은 프레임차 에너지가 작은 것을 나타낸다.  $E(R_i)$ 와  $E(R_j)$ 는 임의의 영역  $R_i$ 와  $R_j$ 에 존재하는 프레임차 에너지를 나타내고,  $E(e_{ij})$ 는 두 영역에 의하여 만들어지는 윤곽  $e_{ij}$  상에 존재하는 프레임차 에너지를 나타낸다. 그림 2에서는  $E(R_i)$ 가 크고  $E(R_j)$ 가 작기 때문에 두 영역의 프레임차 에너지의 전위차가 크고, 또한  $E(e_{ij})$ 도 크기 때문에 진하게 표시된 윤곽은 비디오 객체의 윤곽이 된다. 객체 내부에 존재하는 영역들 사이의 프레임차 에너지의 전위차는 각 영역의 프레임차 에너지가 비슷하기 때문에 작고, 윤곽상의 에너지도 두 영역의 프레임차 에너지와 비슷하다.

이상의 원리에 근거한 시공간 분할 과정을 간단히 설명하면 다음과 같다. 먼저 이전 프레임과 현재 프레임으로부터 SAFD를 구한 다음, 움직임 객체의

움직임 벡터를 구하고 노출 배경 (uncovered background)의 SAFD를 제거한다. 마지막으로 윤곽상의 프레임차 에너지와 영역내의 프레임차 에너지를 이용한 비용으로 반복적 병합을 통해 객체 분할을 수행한다.

(1) SAFD

SAFD는 다음과 같이 구한다.

$$SAFD(x, y) = \frac{1}{N_W} \sum_{(x-k, y-d) \in W} |f_i(x-k, y-d) - f_{i-n}(x-k, y-d)| \quad (2)$$

여기서  $SAFD(x, y)$ 는 이전과 현재 프레임의 절대 프레임차에 대하여 창  $W$  내에서 평균을 취한 것이고,  $N_W$ 는 창  $W$ 내의 모든 화소 수이다. 이렇게 함으로써 잡음을 줄이고, 분할을 할 때 객체 내부에 발생하는 소영역을 줄이는 효과도 있다.

(2) 움직임 벡터와 노출 배경 영역

비디오 객체의 움직임 벡터가 큰 경우에는 프레임차 에너지가 큰 노출 배경(uncovered background) 영역이 비디오 객체에 포함되는데 이를 해결하기 위하여 노출 배경 영역의 SAFD를 제거해야 한다. 이를 위하여 먼저 SAFD의 이진화 마스크를 구하고, 현재 프레임과 이전 프레임에 마스크를 씌운다. 그리고, 현재와 이전 프레임에 대하여 움직임 물체의 움직임 벡터를 구하는데, 알고리즘은 MAD를 이용하고, 탐색 범위(search range)는 두 영상을 정확히 겹쳤을 때를 기준으로  $(\pm \Delta x, \pm \Delta y)$ 로 한다. 그림 3은 이렇게 구한 움직임 벡터가 큰 경우에 노출 배경 영역에 존재하는 SAFD를 없앤 예이다. 회색 영역과 흰색영역이 SAFD이고, 여기서 회색영역이 제거된 노출 배경 영역에 존재하는 SAFD이다. 움직임 객체의 움직임 벡터가 큰 경우에 노출 배경 영역의 SAFD를 이렇게 제거함으로써 노출 배경 영역이 객체에 포함되는 것을 막을 수 있다.



그림 3. 노출 배경 영역에 존재하는 프레임차를 없앤 예 (Suzie 연속 영상, frame#49, frame#50)

(3) 프레임차 에너지의 전위차를 이용한 반복적 병합법

앞서 구한 SAFD를 이용하여 공간 분할로 세밀하게 분할된 파티션  $P_{N_i}$ 의 윤곽상과 영역 내에 존재하는 SAFD로 구한 윤곽과 영역의 프레임차 에너지를 이용한 비용을 최소화 하는 반복적 병합을 통해 객체 분할을 수행한다. 여기서 비용은 윤곽상에 존재하는 화소에 대하여 공간 평균을 취한 프레임차를 윤곽 화소의 개수로 나눈 윤곽상의 프레임차 에너지와 영역 내에 존재하는 화소에 대하여 공간 평균을 취한 프레임차를 영역 화소의 개수로 나눈 영역내의 프레임차 에너지의 비율로 구성된다.

윤곽상의 프레임차 에너지와 영역내의 프레임차 에너지는 아래와 같이 각각 구한다.

$$E(e) = \frac{1}{N_C} \sum SAFD_C(x, y) \quad (3)$$

여기서  $N_C$ 는 윤곽상에 각각 존재하는 픽셀 수이고,  $SAFD_C(x, y)$ 는 윤곽상의 화소에 존재하는 절대 프레임차에 공간 평균을 취한 것이다.

$$E(R) = \frac{1}{N_R} \sum SAFD_R(x, y) \quad (4)$$

여기서  $N_R$ 은 영역  $R$ 내에 존재하는 픽셀 수이고,  $SAFD_R(x, y)$ 영역  $R$ 내의 화소에 존재하는 절대 프레임차에 공간 평균을 취한 것이다.

그러면 비디오 객체 추출을 위한 동질성의 비용은 (3)식과 (4)식을 이용하여 다음과 같이 정의한다.

$$AJ(e) = c(\{R_i, R_j\}) = \begin{cases} \frac{E(e)}{\max[\min\{E(R_i), E(R_j)\}, C]} \cdot \begin{cases} ((E(e) > E(R_i)) \text{ or } (E(e) > E(R_j))) \\ \end{cases} \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

여기서  $C$ 는 분모가 0이 되는 것을 방지하기 위한 상수이다. 영역을 병합하기 위한 동질성의 비용인 (5)식은 (3)식인 윤곽상의 프레임차 에너지를 분자로 하고, 분모를 윤곽을 이루는 두 영역 각각의 프레임차 에너지 중에서 작은 것으로 한다. 이렇게 함으로써 객체 윤곽을 이루는 영역들간의 동질성 비용을 크게 할 수 있으며, 분할이 진행될 때 객체 윤곽이 남을 수 있다. 배경 영역들을 합칠 경우 윤곽상의 프레임차 에너지가 없고, 또한 영역내의 프레임차 에너지가 없는 경우이므로 상수  $C$ 가 분모가 되어 비용이 0되어 먼저 병합이 된다. 객체 내부의

영역들을 합칠 경우 두 영역의 프레임차 에너지와 윤곽상의 프레임차 에너지가 비슷하기 때문에 두 영역 중에서 작은 것을 취하여 윤곽상의 프레임차 전력을 나누어주면, 비용이 1에 가까워지게 된다. 이 경우에 객체 내부 영역을 먼저 병합하기 위하여 비용을 0으로 한다. 두 영역 프레임차 에너지 중에서 작은 것이 0이 되거나 상당히 작은 경우에 C를 사용한다. 이상의 절차로 비디오 객체의 윤곽과 프레임차 에너지의 전위차가 큰 객체 내부에 윤곽을 가진 파티션  $P_{N_i}$ 를 얻는다.

(4) 후처리

시공간 분할의 결과인 파티션  $P_{N_i}$ 를 보면, 프레임차 에너지를 이용한 비용 때문에  $P_{N_i}$ 의 결과에서 볼 수 있었던 프레임차 전력이 없는 배경 영역들이 병합이 됨을 알 수 있다. 그래서 후처리 과정에서는 비용을 임의의 인접하는 두 영역의 크기 즉, 화소수의 합으로 정의하고 비용을 최소화하는 반복적 병합으로 움직임 객체의 윤곽을 추출한다. 잡음 때문에 생기는 배경 영역에 존재하는 작은 고립 영역은 배경 영역에 합치고, 객체 내부에 존재하는 작은 영역들은 객체에 합쳐서 움직임 객체의 윤곽을 얻기 위하여 인접하는 두 영역의 크기의 합이 최소인 영역들을 병합한다. 분할에 사용되는 비용을 아래와 같이 표현할 수 있다.

$$\Delta f(e) = c(\{R_i, R_j\}) = |R_i| + |R_j| \quad (6)$$

여기서,  $|R_i|$ 은 영역  $R$ 의 화소 수이다. 이렇게 비디오 객체의 윤곽을 검출하여 움직임 객체를 추출한 결과를 공간 분할한 결과인 파티션  $P_{N_i}$ 과 연결하여 정보를 저장한다.

현재 프레임에 현재 프레임에 대한 움직임 객체를 추출하고 과거에 저장한 정보와 비교하여 이전 프레임에는 움직임 정보가 있어서 객체에 포함되었다가 현재 프레임에는 움직임 정보가 없어서 빠진 부분이 있으면 저장된 이전의 정보를 이용하여 빠진 영역을 객체에 추가하여 완벽한 움직임 객체를 추출한 파티션  $P_{N_i}$ 를 얻는다.

IV. 실험 결과 및 고찰

제한한 알고리즘의 성능을 평가하기 위하여 176 X144 QCIF(Quarter Common Intermediate Format)의 Suzie와 Miss America 연속 영상을 시험

연속 영상으로 사용하였다. 그림 4의 (a)는 Suzie 연속 영상의 첫 번째 프레임이고, (b)는 Miss America 연속 영상의 첫 번째 프레임이다.

그림 5는 Suzie 연속 영상 중에서 61번째와 63번째 프레임을 이용하여 제안한 알고리즘으로 비디오 객체를 추출한 결과를 보여 준다. (a)는 Suzie 연속 영상의 61번째 프레임 즉, 이전 프레임이며, (b)는 Suzie 연속 영상의 63번째 프레임 즉, 현재 프레임이다. 결과 (c)는 현재 프레임에 대하여 객체와 배경의 경계가 영역에 포함되지 않도록 세밀하게 6000개의 영역으로 분할한 것이다. (d)는 세밀하게

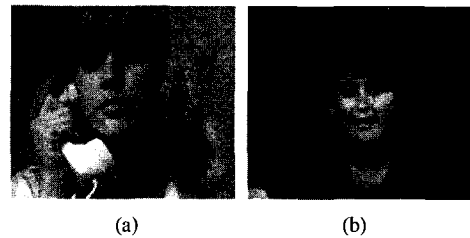


그림 4. 실험에 사용된 연속 영상의 첫 번째 프레임 : (a) Suzie(프레임#0), (b) Miss America(프레임#0)

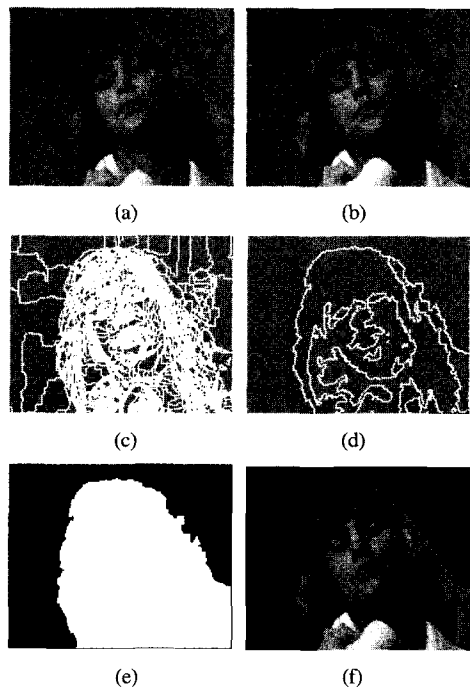


그림 5. Suzie 연속 영상의 61번째와 63번째 프레임과 제안한 알고리즘의 단계별 결과 : (a)이전 프레임(#61), (b)현재 프레임(#63), (c)현재 프레임에 대한 세밀한 분할의 결과 윤곽 영상, (d)움직임 객체 윤곽을 추출한 후의 윤곽 영상, (e)움직임 객체 마스크, (f)움직임 객체 영상

공간 분할된 윤곽 영상 (c)를 입력으로 하여 (5)식에 정의된 비용으로 반복적 병합을 통해 20개의 영역으로 분할한 결과이고, (5)식에서의 상수  $C$ 는  $10^{-6}$ 으로 사용하였고 모든 실험에서 동일하게 적용하였다. (e)는 영역이 크기를 비용으로 정의하여 반복적 병합으로 분할하는 후처리 과정에서 얻은 움직임 객체 마스크이며, (f)는 최종 결과인 움직임 객체 평면이다.

그림 6, 그림 7 및 그림 8은 Suzie 연속 영상에 대하여 비디오 객체 추출 결과를 과정별로 보여 준다. 그림 6은 세밀하게 분할된 영역들이 움직임 객체의 윤곽을 포함하지 않으면서 6000개의 영역으로 공간 분할한 결과를 보여 주고 있다. 그림 7은 그림 6의 결과를 입력으로 받아서 (5)식에서 정의한 프레임차 에너지의 전위차를 이용한 비용으로 움직임 객체의 윤곽을 포함하지 않으면서 20개의 영역으로 시공간 분할한 결과를 보여 주고 있다. 마지막으로 그림 8은 후처리 과정을 통하여 비디오 객체를 추출한 결과를 보여주는 움직임 객체 평면 영상들이다. 그림 6, 그림 7 및 그림 8을 통해서 제안한 알고리즘의 각 과정별로 정의한 비용으로 사람의 눈에 보이는 것과 대체로 일치하도록 객체와 배경의 경계를 찾아서 객체를 추출함을 볼 수 있다. 여기서 그림 6, 그림 7 및 그림 8의 49번째, 50번째, 51번째 프레임에서는 노출 배경 영역의 SAFD를 제거하여 비디오 객체를 추출한 결과를 보여 준다.

그림 9, 그림 10 및 그림 11은 Miss America 연속 영상에 대하여 비디오 객체 추출 결과를 과정별로 보여 준다. 그림 9는 세밀하게 분할된 영역들이 움직임 객체의 윤곽을 포함하지 않으면서 6000개의 영역으로 공간 분할한 결과를 보여 주고 있다. 그림 10은 그림 9의 결과를 입력으로 받아서 (5)식에서 정의한 프레임차 에너지의 전위차를 이용한 비용으로 움직임 객체의 윤곽을 포함하지 않으면서 70개의 영역으로 시공간 분할한 결과를 보여 주고 있다. 마지막으로 그림 11은 후처리 과정을 통하여 비디오 객체를 추출한 결과를 보여주는 움직임 객체 평면 영상들이다. 그림 9, 그림 10 및 그림 11을 통해서 제안한 알고리즘의 각 과정별로 정의한 비용으로 사람의 눈에 보이는 것과 대체로 일치하도록 객체와 배경의 경계를 찾아서 객체를 추출함을 볼 수 있다. 그림 10의 26번째와 28번째 프레임에서 보이는 결과와 같이 움직임 정보가 없어서 어깨선이 추출되지 않는 것을 볼 수 있다. 이러한 경우에 이전 프레임에서 추출한 비디오 객체의 정보를 이용하여 어깨선

이 보상된 비디오 객체를 추출한 결과를 그림 11의 26번째와 28번째 프레임에서 확인할 수 있다.

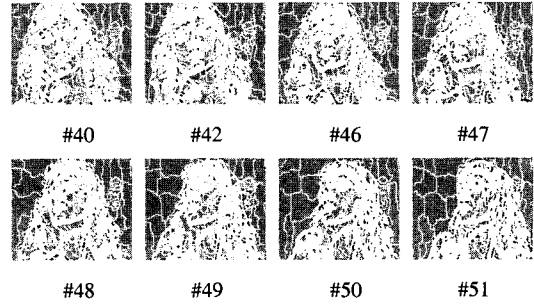


그림 6. Suzie 연속 영상에 대한 공간 분할 결과

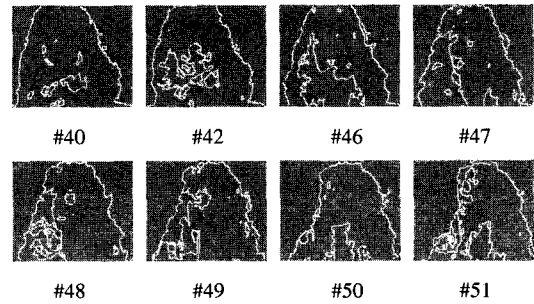


그림 7. Suzie 연속 영상에 대한 시공간 분할 결과

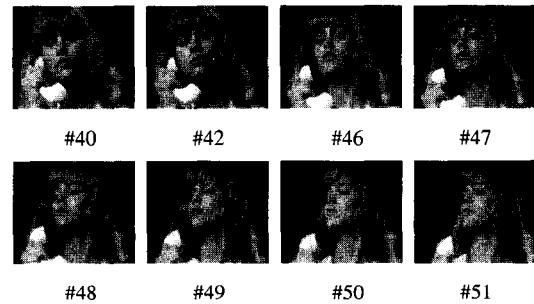


그림 8. Suzie 연속 영상에 대한 후처리 및 비디오 객체 추출 결과

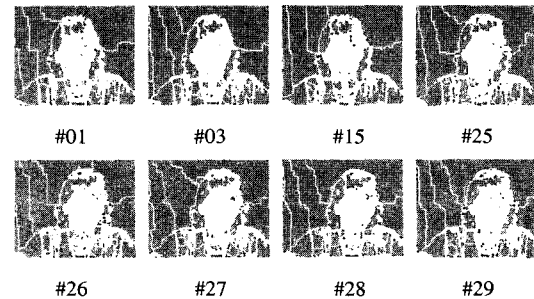


그림 9. Miss America 연속 영상에 대한 공간 분할 결과

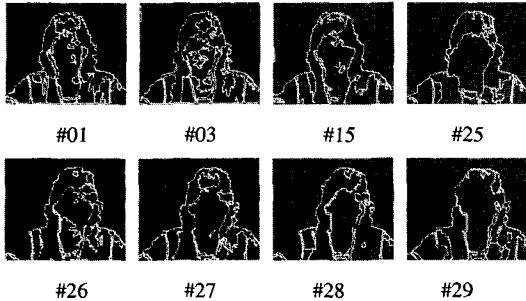


그림 10. Miss America 영속 영상에 대한 시공간 분할 결과

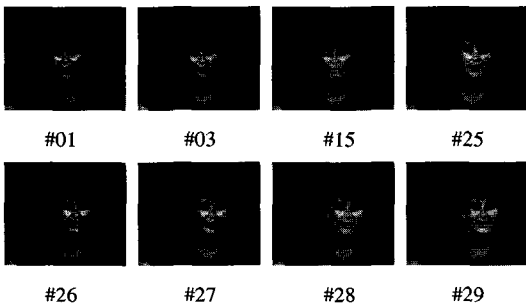


그림 11. Miss America 영속 영상에 대한 후처리 및 비디오 객체 추출 결과

## V. 결론

본 논문에서는 비디오 객체 윤곽을 잘 특징짓는 프레임차 에너지의 전위차로 비용을 정의하고, 이 비용을 최소화하는 반복적 병합법으로 비디오 객체 추출을 수행하였다. 제안한 알고리즘은 공간 분할에서 비디오 객체의 윤곽을 포함하지 않는 영역들로 입력 영상을 윌쇼프 기반의 비용을 최소화하는 반복적 병합법으로 세밀하게 분할하였고, 이 영상을 입력으로 하여 시공간 분할에서는 임의의 인접한 두 영역에 의해서 만들어지는 윤곽상에 존재하는 프레임차 에너지와 인접한 두 영역내의 프레임차 에너지 중에서 작은 에너지를 가지는 영역의 프레임차 에너지의 비로 표현된 비용을 최소화하는 반복적 병합법으로 비디오 객체를 추출하였다.

제안한 알고리즘의 비디오 객체 추출 성능을 평가하기 위하여 Suzie와 Miss America 연속 영상에 대하여 각각 실험하였다. 실험 결과, 공간 분할에서는 윌쇼프 기반 비용으로 세밀하게 분할함으로써 영역들이 비디오 객체의 윤곽을 포함하지 않으면서 분할됨을 볼 수 있었다. 세밀하게 분할된 영상을 입력으로 하는 시공간 분할에서는 객체의 윤곽을

잘 특징짓는 프레임차 에너지의 전위차로 추출함으로서 객체가 손상되는 것 없이 잘 분할됨을 확인할 수 있었다. 두 경우 모두 좋은 결과를 볼 수 있었으며 사람의 눈에 보이는 것과 대체로 일치하도록 객체와 배경의 경계를 찾아서 객체를 추출함을 볼 수 있었다. 앞으로 카메라의 움직임이 있는 경우 즉, 배경이 움직이는 경우와 복잡한 배경에 대하여 비디오 객체를 추출하는 방안과 실시간 처리를 위하여 수행시간을 단축시키는 노력이 있어야 한다.

## 참고 문헌

- [1] R. C. Gonzalez and R. E. Woods, Digital Image Processing, Addison Wesley, 1993.
- [2] S. L. Horwitz and T. Pavlidis, "Picture segmentation by a tree traversal algorithm," Journal of ACM, vol. 23, no. 2, pp. 368-388, Apr. 1976.
- [3] ISO/IEC JTC1/SC29/WG11, N1992, MPEG-4 Video Verification Model Version 11.0, Mar. 1998.
- [4] ISO/IEC JTC1/SC29/WG11, MPEG-4 Visual Final Committee Draft, Dublin, Ireland, July 1998.
- [5] J. G. Choi, M. Kim, M. Ho Lee, and C. Ahn, "Automatic segmentation based on spatio-temporal information," ISO/IEC JTC1/SC29/WG11 MPEG97/m2091, Bristol, U.K., Apr. 1997.
- [6] J. G. Choi, S.-W. Lee, and S.-D. Kim, "Spatio-temporal video segmentation," IEEE Trans. Circuits Syst. Video Technol., vol. 7, pp. 279-286, Apr. 1997.
- [7] M. C. Kim, J. G. Choi, and Y. S. Ho, "A VOP generation tool: Automatic segmentation of moving objects in image sequences based on spatio-temporal information," IEEE Trans. Circuits Syst. Video Technol., vol. 9, no. 8 pp. 1216-1226, Dec. 1999.
- [8] D. Wang, "Unsupervised video segmentation based on watersheds and temporal tracking," IEEE Trans. Circuits Syst. Video Technol., vol. 8, no. 5 pp. 539-546, Sept. 1998.
- [9] D. Zhong and S.-F. Chang, "AMOS: An active system for MPEG-4 video object segmenta-

tion," in Proc. Int. Conf. Image Processing, Chicago, IL 1998.

[10] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes : Active contour models," Int. J. Computer Vision, vol. 1, no. 4, pp.321-331, 1988.

[11] D. Zong and S.-F. Chang, "Video object model and segmentation for content-based video indexing," ISCAS'97, HongKong, June 9-12, 1997.

[12] S. W. Zucker, "Region growing: Childhood and adolescence," Computer Vision Graphics and Image Processing, vol. 5, pp. 382-399, 1976.

[13] M. Kunt, M. Benard, and R. Leonardi, "Recent results in high- compression image coding, IEEE Trans. Circuits and Systems, vol. CAS -34, no. 11, pp. 1306-1336, Nov. 1987.

[14] M. Kocher and R. Leonardi, "Adaptive region growing technique using polynomial functions for image approximation," Signal Proc- essing, vol 11, pp. 47-60. 1986.

[15] O. J. Morris, M. de J. Lee, and A. G. Constantinides, "Graph theory for image analysis: An approach based on the shortest spanning tree," IEE Proc., vol. 133, pt. F, no. 2, pp. 146-152, Apr. 1986.

[16] S. H. Kwok and A. G. Constantinides, "A fast recursive shortest spanning tree for image segmentation and edge detection," IEEE Trans. on Image Processing, vol. 6, no. 2, pp. 328-332, Feb. 1997.

[17] C. W. Lim, S. C. Jun, and N. C. Kim, "Recursive merging for rate-distortion based image segmentation," IEEE Trans. Circuits Syst. Video Technol., vol. 10, no. 7 pp. 1121-1134, Oct. 2000.

곽 종 인(Jong-In Kwak)

정회원

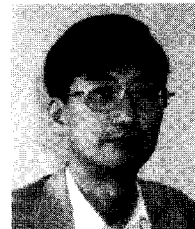


1994년 2월: 경북대학교  
전자공학과 (공학사)  
2001년 2월: 경북대학교  
전자공학과 (공학석사)  
2001년 3월~현재: 경북대학교  
전자공학과 박사과정

<주관심 분야> 영상처리, 영상분할, 영상압축, 컴퓨터 비전

김 남 철(Nam Chul Kim)

정회원



1978년 2월: 서울대학교  
전자공학과 (공학사)  
1980년 2월: 한국과학기술원  
전기 및 전자공학과  
(공학석사)  
1984년 2월: 한국과학기술원  
전기 및 전자공학과  
(공학박사)

1984년 3월~현재: 경북대학교 전자전기공학부 교수  
1990년 1월~1994년 12월: 생산기술연구원 HDTV  
신호처리분과위원  
1994년 1월~1996년 12월: 한국통신학회 편집위원  
1994년 1월~1996년 12월: 대한전자공학회 편집위원  
1996년 1월~현재: 한국방송공학회 편집위원, 이사  
1997년 1월~1998년 12월: 한국통신학회 신호처리  
연구회 위원장

<주관심 분야> 영상처리, 영상압축, 영상복원, 컴퓨터 비전