

MPEG-4 부호화를 위한 반자동 영상분할

김진철, 김제환, 하중수, *김영로, 고성제
고려대학교 전자공학, *(주)삼성전자

전화 : 02-3290-4719 / 핸드폰 : 011-415-7729

Semiautomatic segmentation for MPEG-4 encoding

Jin-Cheol Kim, Jae-Hwan Kim, Jong-Soo Ha, *Young-Ro Kim, Sung-Jea Ko

Dept. of Electronics Engineering, Korea University

*Samsung Electronics Inc.

E-mail : arias711@dali.korea.ac.kr

Abstract

In this paper, We propose a new semiautomatic segmentation method using spatio-temporal similarity. In the proposed scheme, segmentation is performed using gradual region merging and bi-directional spatio-temporal refinement. Simulation results show the efficiency of the proposed method in semantic object extraction.

I. 서론

MPEG-4는 기존의 비디오 압축 표준인 MPEG-1/2에 없었던 객체 기반 부호화 개념과 이에 따른 VOP(video object plane)라는 새로운 개념을 도입하여 복호화 및 부호화를 VOP 단위로 가능하게 함으로써 향후 멀티미디어 환경에서 객체별로의 접근이 가능하도록 지원하는 것을 목표로 하고 있다[1]. 따라서, 영상 분할은 MPEG-4 표준의 직접적인 표준화 대상은 아니지만 객체기반 부호화에서 절대적으로 필요한 전처리 과정이다.

일반적으로 영상에서 사람이 추출하고자 하는 객체를 설정하는 것은 매우 주관적이므로 사용자의 개입이 없는 자동 영상분할 알고리즘을 이용하여 정확한 경계의 객체를 추출하는 것은 현실적으로 매우 어렵다. 따라서 사람이 직접 원하는 객체를 지정해주는 반자동 영상분할 알고리즘을 이용하는 것이 효과적이다. 반자동 영상분할 기법에는 공간적 정보를 이용한 화소기반 기법, 예지기반 기법, 영역기반 기법 등이 있고 시간적

정보를 이용한 기법에는 통계적 방법에 의한 기법, 움직임 정보를 이용한 기법등이 있다[2],[3]. 반자동 영상 분할을 할 때 공간적 정보만을 사용하면 객체와 배경의 화소값이 비슷한 경우에 에러를 크게 발생시키고, 시간적 정보만을 사용하면 객체의 움직임이 작을 때 에러를 크게 발생시킨다.

따라서, 본 논문은 공간적 유사성에 기반한 점진적 영역 병합 알고리즘과 양방향 시공간적 동질성에 기반하여 정확한 객체의 경계를 얻을 수 있는 알고리즘을 제안하였고 이를 토대로 향상된 반자동 영상분할 결과를 얻고자 하였다.

II. 반자동 영상분할 알고리즘

임의의 영상에서 사람이 객체를 추출하고자 할 때 사전 정보 없이 의미 있는 객체를 추출한다는 것은 만족스러운 결과를 얻기가 어렵다. 따라서 사람의 도움이 개입된 반자동 영상분할이 의미있는 객체를 추출하는데 효과적이다. 그림 1에서 보여주듯이 반자동 영상 분할은 크게 프레임내(intraframe) 영상분할과 프레임간(interframe) 영상분할의 두 단계로 구성된다.

먼저 프레임내 영상 분할에서는 GUI(graphic user interface) 환경에서 사람의 도움으로 얻어진 대략적인 내부경계와 외부경계 사이에서 객체를 추출한다. 결과가 만족스럽지 못하다면 다시 사람의 도움을 얻어 프레임내 영상분할을 하고 결과가 만족스럽다면 프레임간 영상분할을 한다. 프레임간 영상분할은 프레임내 영상분할에서 추출된 객체를 추적하는 단계이다.

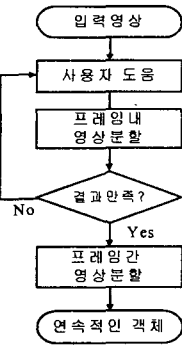


그림 1. 반자동 영상분할의 구조도

2.1 프레임내 영상분할

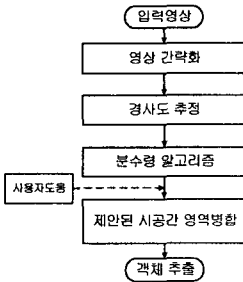


그림 2. 프레임내 영상분할의 구조도

그림 2는 프레임내 영상분할의 구조도를 보여주고 있다. 먼저 입력 영상이 들어오면 영상을 간략화 한다. 영상을 간략화 하는 것은 영상에서 잡음을 제거하고 불필요한 정보들을 없애기 위해서 이다. 본 논문에서는 형태학적 재생 여과기와 3x3 구조적 요소를 사용하여 영상을 간략화 하였다[4].

다음은 이전 단계에서 간략화된 영상으로부터 공간 경사도 영상을 구한다. 경사도 영상에서 분할된 영역은 분수령 결과 영상에서 경계로 나타나며 경사도의 최소값은 분할 결과 영상에서 한 영역으로 나타나게 된다. 따라서 공간 경사도는 영상을 영역 분할하는 분수령 알고리즘의 입력값으로 사용할 수 있다. 본 논문에서는 Sobel 연산자를 이용해 경사도 영상을 구했다.

분수령 알고리즘(watershed)은 영상의 밝기 값 분포를 지형학적인 구조로 보고 지형학적인 구조에서 영역들의 경계가 되는 분수령을 찾는 알고리즘이다. 분수령 알고리즘은 그림 3과 같이 집수 지역들(catchment basis)로 불리는 경사도의 극부 최저값들을 분할될 영역의 중심으로 가정한다. 그 점으로부터 시작하여 같은 성질의 영역들을 공통 영역으로 할당함으로써 영

역을 분할한다[5]. 전 단계의 경사도 영상에서, 경사도 값이 큰 영역은 분수령 영상에서 경계로 나타나고, 상대적으로 작은 값을 갖는 영역들은 한 영역으로 병합된다. 분수령 알고리즘의 구현 방법으로 침식 기법[6]과 비내림 기법[5]이 있는데 본 논문에서는 비내림 기법의 개념을 이용하여 알고리즘을 구현하였다.

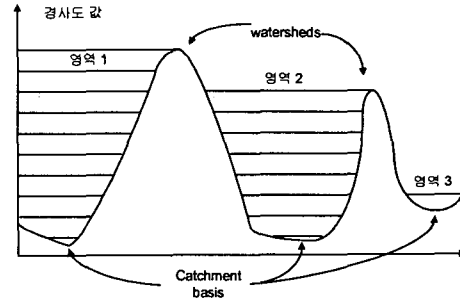


그림 3. 지형학적 분수령 알고리즘

지금 까지 설명한 분수령 알고리즘의 결과영상만으로는 사용자가 원하는 의미있는 객체를 추출하지 못하므로 제안하는 시공간 영역 병합과정과 양방향 시공간 경계 개선 기법을 이용해 프레임내 영상분할 결과를 얻는다.

2.2 프레임간 영상분할

프레임간 영상분할은 두 번째 단계인 추적 알고리즘이다. 이러한 객체 추적을 위해서 흑백이나 칼라의 동질성을 기준으로 영역들을 추적하는 방법, 움직임 정보의 동질성을 기준으로 추적하는 방법, 또는 시간적 공간적 정보를 조합한 기준으로 추적하는 방법이 연구되었다[7],[8]. 이러한 연구에서 사용한 추적 알고리즘은 순방향 추적기술을 사용하는 것이었다. 이 기술은 이전 프레임의 영역이나 객체를 움직임 정보에 의해 현재 프레임에 투사해서 현재 프레임의 객체를 판단하는 것으로 투사한 후 투사된 영역이나 객체들을 조합하거나 후처리해서 추적하는 것이다. 그러나 순방향 추적의 단점은 움직임 정보에서 폐색영역(occlusion) 문제와 투사된 객체의 경계를 처리해 주는데 어려움이 있고 결과로 나온 객체의 경계에 찌그러짐이나 빈 공간의 영역이 나올 수 있다는 것이다.

최근 들어 역방향 추적기법들이 객체 추적의 대안적인 방법으로써 제안되고 있다[9]. 역방향 추적 기법의 주요 장점은 객체 경계를 잘 찾기 위해서 각 프레임마다 공간적 영상분할을 실행함으로써 객체 경계를 다시 후처리하여 조정할 필요가 없다는 점과 형태가 변

MPEG-4 부호화를 위한 반자동 영상분할

하는 객체(non-rigid object)를 잘 추적하는데 있다. 따라서 본 논문에서는 역방향 추적 기법을 사용하여 프레임간 영상분할을 수행하였다.

III. 제안하는 영상분할 알고리즘

3.1 점진적인 영역병합

본 논문에서 제안하는 영역 병합 알고리즘은 사용자가 그려주는 밴드 내에서의 작은 영역들을 전경과 배경으로 구분하여 병합한다. 그림 4(a)에서 분수령 영역과 사용자가 조작한 경계를 각각 볼 수 있다. 이 때 ROI(s) (regions of interest)로 표현되는 부분이 사용자가 조작한 경계를 나타낸다. 그림 4(b)에서 보듯이, ROI s 안팎의 경계 B_{in} , B_{out} 사이의 분수령 영역들은 영역들 사이의 평균 절대값 차이가 문턱값 보다 작을 때 전경과 배경으로 각각 병합된다. 이 과정은 ROI s 안에서 점진적으로 이루어지며 문턱값은 점진적으로 증가시키며 병합한다.

$$|M_{R_i} - M_{R_j}| < T \quad (1)$$

식 (1)이 만족하면 R_i 와 R_j 는 병합한다. 여기서 R 은 영역을, M 은 평균, T 는 문턱값을 나타낸다. 그림 4(a)에서 ROI s를 제외한 나머지 영역들을 RON s (regions of non-interest)라고 한다.

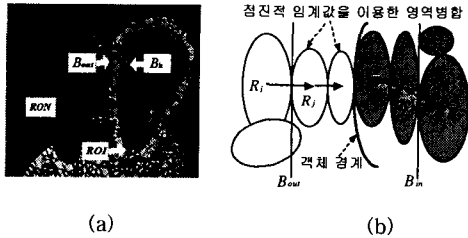


그림 4. 제안된 점진적 영역 병합. (a) 분수령 영역들과 사용자 조작 경계. (b) 영역병합.

3.2. 양방향 시공간적 경계 개선

그림 5에서 제안된 양방향 시공간적 경계 개선 기법을 확인할 수 있다. 먼저 현재와 이전 프레임 사이의 광학흐름(optical flow) 추정을 한다[10]. 다음 각 RON s 마다 영역 어파인(affine) 모델링을 통해 어파인 모션 파라미터(motion parameter)를 구한다. 이 어파인 모션 파라미터는 R_E (점진적 영역 병합을 통해 얻어지는 객체 경계에 인접한 분수령 영역)에 가장 가

까이 위치한 RON s의 어파인 모션 파라미터 R_A 와 R_B 를 말하며 다음과 같이 이용된다. 모션 파라미터 $A=\{a_1, a_2, \dots, a_5, a_6\}$ $B=\{b_1, b_2, \dots, b_5, b_6\}$ 는 어파인 변환에서 구하고자 하는 값에 대한 파라미터로 작용한다.

$$\begin{bmatrix} x'_A \\ y'_A \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_4 & a_5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_3 \\ a_6 \end{bmatrix},$$

$$\begin{bmatrix} x'_B \\ y'_B \end{bmatrix} = \begin{bmatrix} b_1 & b_2 \\ b_4 & b_5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_3 \\ b_6 \end{bmatrix} \quad (2)$$

식 (2)에서 (x'_A, y'_A) 와 (x'_B, y'_B) 는 현재 프레임의 위치 (x, y) 로부터 어파인 변환된 이전 프레임의 위치들이다. 이처럼 프레임의 위치를 구한 다음에는 A 와 B 를 이용해 이전과 현재 프레임 사이의 전방으로 치환된 프레임 차이값(forward displaced frame differences) DFD^2_{Af} 와 DFD^2_{Bf} 를 각각 계산한다. 수식은 아래 식 (3),(4)와 같다.

$$DFD^2_{Af} = \frac{1}{N} \sum_{(x,y) \in R_E} [I^t(x,y) - I^{t-1}(x'_A, y'_A)]^2 \quad (3)$$

$$DFD^2_{Bf} = \frac{1}{N} \sum_{(x,y) \in R_E} [I^t(x,y) - I^{t-1}(x'_B, y'_B)]^2 \quad (4)$$

N 은 영역 R_E 의 화소수이며, $I^t(x, y)$ 는 현재 프레임의 위치 (x, y) 에서의 화소값이다. $I^{t-1}(x'_A, y'_A)$ 는 이전 프레임의 위치 (x'_A, y'_A) 에서의 화소값이다. 움직임 추정에서 찾기 힘든 폐쇄영역에서 생기는 에러를 감소시키기 위해 현재와 다음 프레임 사이에서 역방향 DFD인 DFD^2_{Ab} 와 DFD^2_{Bb} 를 계산한다.

$$\min(DFD^2_{Bf}, DFD^2_{Bb}) > \min(DFD^2_{Af}, DFD^2_{Ab}) \quad (5)$$

최종적으로 식 (5)를 만족한다면 영역 R_E 는 객체로 병합된다.

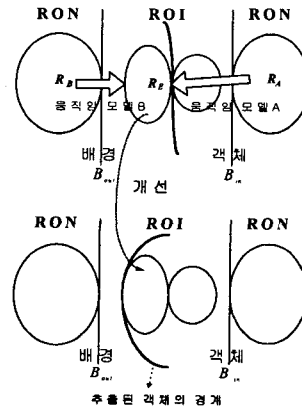


그림 5. 제안하는 양방향 경계 수정

IV. 실험 결과

그림 6(a)는 Motherdaughter, Flowergarden 영상의 점진적 영역 병합에 의한 결과 영상이며, 그림 6(b)는 점진적 영역 병합에 의한 결과영상에 양방향 시공간적 경계 개선 알고리즘을 적용해 얻은 결과영상이다. Motherdaughter 영상은 배경과 객체의 화소값 차이가 크고 객체의 움직임도 작기 때문에 공간적 정보인 점진적 영역 병합만을 사용하여도 만족할 만한 결과를 얻을 수 있었다. 그러나 Flowergarden 영상은 배경과 객체의 화소값이 비슷하고 객체의 움직임이 크기 때문에 점진적 영역 병합만으로는 만족할 만한 결과를 얻을 수 없었다. 그러나 시간적 정보인 양방향 시공간적 경계 개선을 통해 향상된 결과를 도출해 낼 수 있었다.

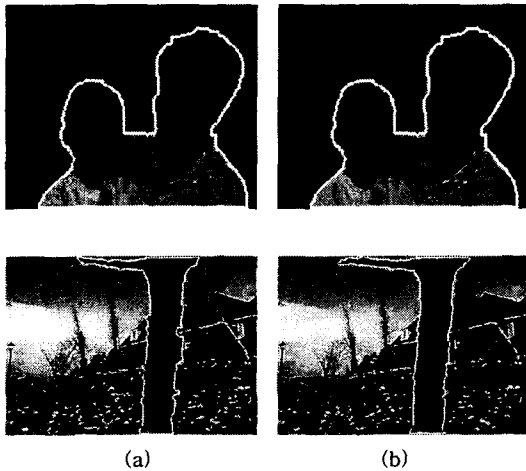


그림 6. Motherdaughter, Flowergarden 동영상의 영역 병합과 개선 결과. (a) 점진적 영역 병합, (b) 제안된 시공간적 경계 개선

V. 결론

본 논문에서는 기존의 반자동 영상분할의 성능을 향상시킬 수 있는 방안으로 프레임내 영상분할에서 시간적, 공간적 정보를 모두 사용하는 방법을 제안하였다. 제안한 알고리즘은 먼저 공간적 정보를 이용하여 점진적 영역 병합을 하고 시간적 정보를 이용하여 양방향 경계 개선을 하였다. 제안한 알고리즘을 이용한 실험 결과, 객체와 배경의 화소값이 비슷한 경우나 객체의 움직임이 없을 경우 제안한 기법이 기존의 기법보다 객체경계 검출에 있어 향상된 성능을 보였다. 특히 양방향 시공간적 경계 개선 기법은 객체의 움직임 때문

에 새로이 나타나는 영역으로 인한 오차를 감소 시켜 주었다. 실용적인 측면에서 영상분할을 사용하기 위해서는 처리시간, 의미있는 객체의 효과적인 추출에 대해 지속적인 연구가 진행되어야 한다.

참고문헌

- [1] ISO/IEC JTC1/SC29/WG11 m4478, "Text for VOP generation tools in MPEG-4 version 2 Visual Committee Draft," Seoul MPEG meeting, March 1999.
- [2] K. Haris, S. N. Efstratiadis, "Hybrid image segmentation using watersheds and fast region merging," *IEEE Trans. on Image Process.*, vol. 7, no. 12, pp. 1684-1699, Dec. 1998.
- [3] 김대회, 김민호, 호요성, "MPEG-4 동영상 부호화를 위한 영상 객체 추출 기법", *전자공학회지*, vol. 26, no. 7, pp. 42-51, 1999.
- [4] P. Salembier and M. Pardas, "Hierarchical morphological segmentation for image sequence coding," *IEEE Trans. on Image Process.*, vol. 3, no. 5, pp. 639-651, Sep. 1994
- [5] Luc Vicent, Pierre. Soille, "Watershed in digital space: An efficient algorithm based on immersion simulation," *IEEE Trans. on analysis and machine intelligence*, vol. 13, no. 6, pp. 583-598, June, 1991.
- [6] P. Salembier and M. Pardas, "Hierarchical morphological segmentation for image sequence coding," *IEEE Trans. on Image Process.*, vol. 3, no. 5, pp. 639-651, Sep. 1994
- [7] J. Wang and E. Adelson, "Representing moving images with layers," *IEEE Trans. on Processing*, vol. 3, no. 5, pp. 625-638, Sept. 1994.
- [8] M. J. Black, "Combining intensity and motion for incremental segmentation and tracking over long image sequences," *ECCV92*, pp. 485-493, Italy, May 1992.
- [9] C. Gu and M. Lee, "Semantic video object tracking using region-based classification," *ICIP 98*, pp. 628-632, 1998.
- [10] C. Kervrann and F. Heitz, "Statistical deformable model-based segmentation of image motion," *IEEE Trans. Image Processing*, vol. 8, pp. 583-588, no. 4, Apr. 1999.