

분류된 영역 병합에 의한 객체 원형을 보존하는 영상 분할

박 현상, 나 종범

한국과학기술원 전기 및 전자공학과

대전시 유성구 구성동 373-1

E-mail: hspark@isscosmos.kaist.ac.kr, jbra@ee.kaist.ac.kr

Image segmentation preserving semantic object contours by classified region merging

Hyun Sang Park and Jong Beom Ra

Dept. of Electrical Engineering, Korea Advanced Institute of Science and Technology

373-1 Kusongdong Yusonggu, Taejon

Abstract

Since the region segmentation at high resolution contains most of viable semantic object contours in an image, the bottom-up approach for image segmentation is appropriate for the application such as MPEG-4, which needs to preserve semantic object contours. However, the conventional region merging methods, that follow the region segmentation, have poor performance in keeping low-contrast semantic object contours. In this paper, we propose an image segmentation algorithm based on classified region merging. The algorithm pre-segments an image with a large number of small regions, and also classifies it into several classes having similar gradient characteristics. Then regions only in the same class are merged according to the boundary weakness or statistical similarity. The simulation result shows that the proposed image segmentation preserves semantic object contours very well even with a small number of regions.

1. 서론

초저전송률을 지향하는 영상 압축 표준으로 새롭게 제안된 MPEG-4 [1]에서는 이전의 표준에서 다루기 않았던 VO (Video Object)라는 새로운 개념을 도입했다. VO란 독립적인 편집이 가능한 객관적인 영상 객체를 지칭하며, VO의 윤곽선을 고려한 움직임 보상에 의해서 고압축률에서도 시각적으로 우수한 화질이 가능하게 되었다.

VO의 도입은 객체의 형태를 보존하는 영상 분할을 필요로 한다. 동영상의 경우에는 시공간적 상관도를 이용하여 VO를 추출하거나 추적하기 위한 다양한 방법들 [2-3]이 제안되었으나, 동영상의 최초 영상에 해당하는 정지 영상에 대해서는 이에 대한 연구보다는 영상

압축에 적합한 영상 분할에 주된 관심이 집중되어 왔다 [4].

객체의 형태를 보존하기 위한 영상 분할은, 영역 분할 (region segmentation)과 영역 병합 (region merging)으로 구성되어 있는 상향 (bottom-up) 방식이 높은 가능성을 가진다. 이는 높은 해상도에서의 영역 분할이 비록 과분할된 결과를 낳더라도, 영상 내의 중요한 구성 요소에 대한 윤곽선의 대부분을 포함함에 따라 형태를 보존하기 위한 방법에서는 타당한 시발점이 되기 때문이다. 영역 병합은 과분할된 결과로부터 유사한 특성을 가지는 작은 영역들을 병합해서 목적에 일치하는 더 큰 영역으로 재구성하는 과정이다.

그러나 기존에 제안된 영역 병합 방법들은 병합하려는 영역의 통계적인 유사성에 기반을 두거나 [5-6], 통계적인 모델과의 차이값 [7]을 토대로 병합을 수행했으나, 과분할된 각 영역으로부터 얻어진 통계치의 신뢰도가 낮고 지역적인 정보만을 사용하기 때문에 통계적인 성능 개선에는 유리하더라도, 광역 정보의 결핍으로 인해서 객체의 실제적인 형태에 대한 손실을 가져오게 되었다.

따라서 객체 형태에 대한 광역 정보를 제공할 수 있는 선지식이 필요하게 되었다. 하향 (top-down) 방식에서는 상위 계층에서의 분할 결과를 하위 계층에서 사용함으로써, 광역 정보를 제공하지만, 이 경우 저대조로 구분된 평탄 영역의 분리가 어렵다는 문제를 가지고 있다.

본 논문에서는 gradient 영상을 이용하여, 분할된 영역을 균질도에 따라서 여러 개의 class로 분류하여, 이 class 정보를 영역 병합에 필요한 광역 정보로 사용한다.

병합은 높은 균질도를 가지는 영역부터 수행되며, 높은 유사도를 가지거나 낮은 경계선 경도를 가진, 같은 class에 속하는 인접 영역 중에서, 병합 후에 가장 낮은

분산을 제공하는 영역들을 병합시킨다.

본 논문은 다음과 같이 구성되어 있다. 2 장에서는 제안한 방식을 구성하는 각 구성 요소의 기능을 설명하고, 3 장에서 대표적인 동영상에 대한 실험 결과를 기술한다. 4 장에서는 결론과 함께 앞으로 다루어야 할 과제에 대해서 언급하겠다.

2. 본론

제안한 영상 분할 방식은 그림 1 과 같이 영역 분할 (region segmentation)과 영역 병합 (region merging)으로 구성되어 있다.

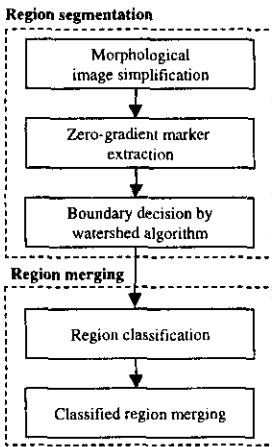


Fig. 1. Overall procedure

수리형태론을 이용한 영역 분할은 다음과 같이 3 단계로 구성된다. 1 단계: 형태론적 필터 (morphological filter)를 이용하여 영상을 단순화 한다. 2 단계: 단순화된 영상으로부터 균질한 영역을 대표할 수 있는 마커 (marker)를 추출한다. 이는 다음 절에 설명되어 있다. 3 단계: watershed 알고리즘을 적용하여, 균질한 각 영역 간의 경계를 추출한다.

영역 병합은 2 단계로 구성되어 있다. 먼저 과분할된 영역을 균질도에 따라서 여러 개의 class로 분류하고, 마지막 단계에서는 같은 class를 가지는 영역들에 대해서 병합을 수행한다.

2.1. Zero-gradient 마커 추출에 의한 영역 분할

일반적인 마커 추출법은 평탄 영역 내에서는 많은 수의 마커를, 비평탄 영역 내에서는 적은 수의 마커를 추출하지만, 이는 인간의 시각적인 특성에는 반하는 것이다. 따라서 비평탄 영역에서는 많은 수의 마커를 추출하고, 평탄 영역에서는 적은 수의 마커를 추출하기 위하여 그림 2와 같은 마커 추출 알고리즘을 제안한다.

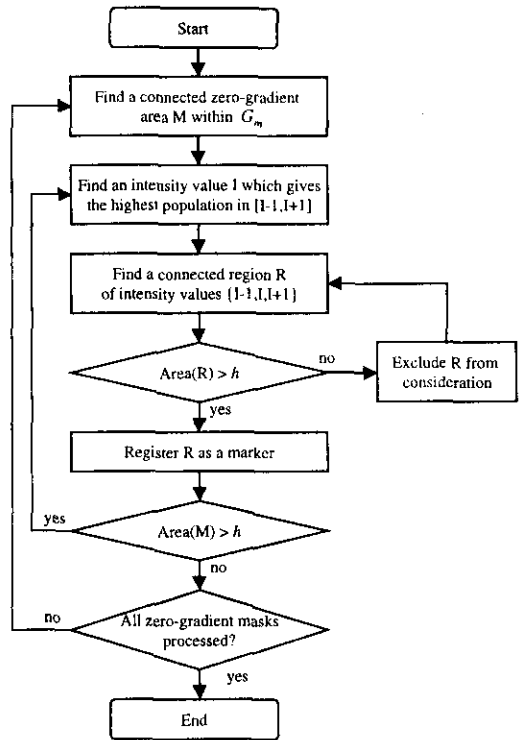


Fig. 2. The zero-gradient marker extraction algorithm

먼저 마커가 발생할 수 있는 기초 영역을 식 (1)에 정의된 G_m 영상이 '0'을 가지는 위치만으로 제한한다. 식에서 f_s , δ_1 , ϵ_1 은 각각 단순화된 영상, 크기 1의 dilation, 크기 1의 erosion을 나타낸다.

$$G_m = \text{Min}(\delta_1(f_s) - f_s, f_s - \epsilon_1(f_s)) \quad (1)$$

그림에서 크기 기준 h 는 형태론적 필터의 크기에 의해서 결정된다. 그리고 단순화된 영상의 밝기값을 사용해서 마커를 추출한다.

이 방법에 의해서 추출된 마커들은 매우 유사한 밝기값을 가지는 화소들로 이루어진 영역이 된다. 마커 내에서 좁은 범위의 밝기값 변화를 허용함으로써, 균질한 영역 내에서 적은 수의 마커를 추출하는 것이 가능해진다. 또한 크기 기준 h 를 낮춤으로써 비균질한 영역에 대한 마커의 수를 증가시키는 것이 가능하다.

2.2. 영역 분류 (region classification)

영역 분류는 화소 단위의 이진 분류가 선행된다. 식 (2)의 형태론적 gradient 영상에서, σ_{G_m} 보다 낮은 gradient를 가지는 모든 화소는 평탄 영역으로, 그렇지 않은 영역은 비평탄 영역으로 나눈다. σ_{G_m} 은 G_m 영상의 분산을 나타낸다.

$$g = \delta_1(f_s) - \varepsilon_1(f_s). \quad (2)$$

평탄 영역을 구성하는 화소들은 비평탄 영역의 화소들에 의해서 서로 구분된다. 이때 크기가 작은 연결된 평탄 영역과 비평탄 영역에 해당하는 모든 화소의 class 를 0으로 결정하고, 평탄 영역의 화소들로 구성된 연결된 영역들에 서로 다른 class 를 할당함으로써, 화소 단위의 분류를 마친다.

영역의 class 는 영역을 구성하는 화소들이 가지는 class 중에서 가장 큰 class 를 택함으로써 영역 단위의 분류를 수행한다. 그림 4의 3 번째 행은 여러 영상에서 영역 분류 후에 얻어진 class 경계선을 나타낸다.

2.3. 분류된 영역 병합 (classified region merging)

그림 3은 제안한 영역 병합 방법의 순서도이다. 각 class 는 독립적으로 처리되며, 한 class 내에서는 Dev 보다 낮은 분산을 가지는 영역을 중심으로 병합을 수행한다. 임의의 영역 R 에 대한 분산 σ_R 과 평균 μ_R 은 원 영상에서 계산한다.

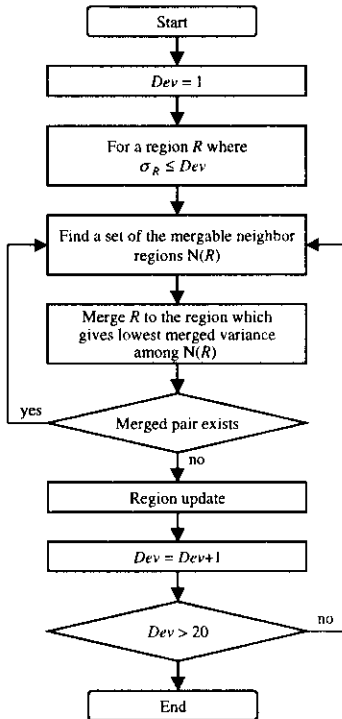


Fig. 3. Classified region merging algorithm

현재 병합을 수행하고자 하는 영역 R 에 대해서, 같은 class 를 가지는 인접한 영역 중에서 낮은 경계선 경도를 가지거나, 유사한 통계 특성을 가지는 영역을 모아서 병합이 가능한 후보 영역 집합 $N(R)$ 을 다음과 같

이 구성한다.

영역 R 과 인접한 영역 R_n 을 동시에 접하는 경계 영역 화소의 50% 이상이 식 (3.1)의 경계선 경도 문턱치보다 작은 gradient 를 가지거나, 영역 R_n 이 원영상에서 가지는 평균 밝기값이 식 (3.2)를 만족하면, 병합 후보 영역으로 삼는다. 식에서 σ_C 와 μ_C 는 식 (1)의 G_m 영상에서 계산한 class C의 분산과 평균을 나타낸다.

$$\text{Boundary weakness: } \text{Max}([\sigma_C], 2). \quad (3.1)$$

$$\text{Region similarity: } |\mu_R - \mu_{R_n}| < \mu_C. \quad (3.2)$$

결정된 병합 후보 집합 $N(R)$ 에서 식 (4)와 같이, 병합 후에 가장 낮은 분산을 가질 수 있는 이웃 영역 R_n^* 으로 병합을 수행한다.

$$n^* = \underset{n}{\text{argmin}} \sigma_{R+R_n}, \text{ where } R_n \in N(R). \quad (4)$$

3. 모의실험

모의 실험의 환경을 다음과 같다. 영상 단순화는 크기 1을 가지는 open-close by reconstruction 필터에 의해서 수행하였고, 그림 2의 마커 추출 알고리즘에서 크기 조건 $h=3$ 으로 마커를 추출했다. 영역 경계 결정은 변형된 watershed 알고리즘 [4]을 사용했다. 이어서 분할된 결과에 제안한 영역 병합을 적용했다.

그림 4의 각 열은 QCIF 형식을 가지는 “Claire”, “Miss America”, “Mother & daughter”, “Carphone” 표준 비디오 영상의 최초 프레임에 대한 영상 분할 결과를 나타낸다. 하나의 열은 원영상, zero-gradient 마커에 의한 영역 분할, class 경계 영상, 영역 병합 결과 등을 연속적으로 보여준다.

최종적으로 분할된 영역의 수는 “Claire”는 67 개, “Miss America”는 125 개, “Mother & daughter”는 261 개, “Carphone”은 229 개였다.

4. 결론

분할된 영상은 객체의 원형이 보존됨을 보여준다. 특히 기존의 방법에서는 배경과 객체 간의 저대조 때문에 구분이 어려웠던 “Miss America”의 경우에서도 시각적으로 우수한 분할 결과를 얻을 수 있었다.

따라서 제안한 방식의 영상 분할은 객체의 윤곽선의 보존이 필요한 영상 분할과, 높은 주관적인 화질이 필요한 영상 압축에 적합하다. 동영상 압축에서는 분할된 경계선이 실제 객체의 윤곽선과 일치하기 때문에 높은 시간적 상관도로 가지며, 이로 인하여 효율적인 윤곽선 부호화가 가능하다.

분할된 영역의 수는 비록 적지는 않지만, 과분할이

남아 있는 대부분의 영역은 평탄한 영역으로 둘러 쌓인 비평탄 영역을 구성하고 있기 때문에 효율적인 추가적 영역 감소가 가능해진다.

참고문헌

[1] ISO/IEC JTC1/SC 29/WG 11/1469, "MPEG-4 video verification model V.5.0," Nov. 1996.
 [2] J. G. Choi, S. W. Lee, and S. D. Kim, "Spatio-temporal segmentation using a joint similarity measure," *IEEE Trans. Circ. and Sys. for Video Tech.*, vol. 7, no. 2, pp. 279-286, Apr. 1997.
 [3] J. Yi, H. S. Park, and J. B. Ra, "A hierarchical semantic video object tracking algorithm using mathematical morphology," *Proc. of Intern. Workshop on Advanced Image Tech.*, pp. 29-34, Jan. 1998.
 [4] P. Salembier, "Morphological multiscale segmentation for image coding," *Signal Processing*, vol. 38, pp. 359-386, 1994.
 [5] L. Vincent and P. Soille, "Watersheds in digital space: An efficient algorithm based on immersion simulations," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583-598, June 1991.
 [6] D. Cortez, P. Nunes, and F. Pereira, "Image segmentation towards new image representation methods," *Signal Processing: Image Communication*, vol. 6, pp. 485-496, June 1995.
 [7] K. W. Song, Y. S. Park, K. P. Han, Y. S. Kim, and Y. H. Ha, "Image segmentation using contrast and contour simplification for very low bit rate coding," *Proc. of Picture Coding Symposium*, pp. 189-193, Sep. 1997.

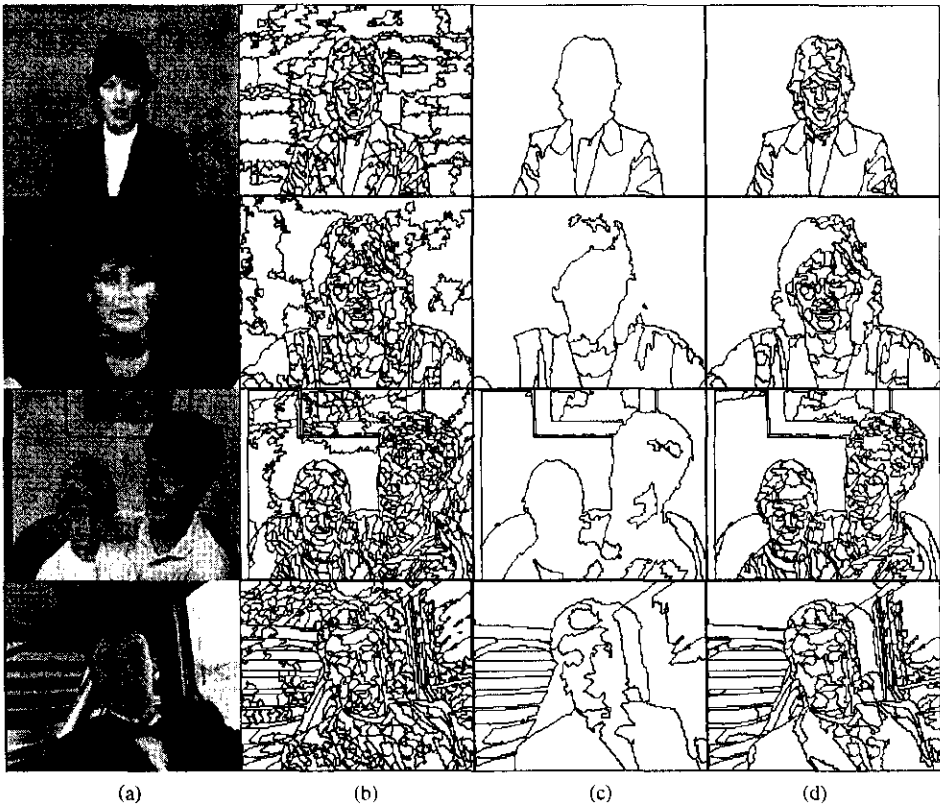


Fig. 4. Segmentation results; (a) original image, (b) initial segmentation, (c) classification, and (d) final merged segmentation result using (b) and (c).