

상향식 영상분할 구조에서의 초기 영상분할을 위한 효율적인 마커 추출 알고리즘

박현상, 나종범

한국과학기술원 전기 및 전자공학과

대전시 유성구 구성동 373-1

Email: hspark@isscosmos.kaist.ac.kr, jbra@ee.kaist.ac.kr

EFFICIENT MARKER EXTRACTION ALGORITHM FOR INITIAL SEGMENTATION IN A BOTTOM-UP IMAGE SEGMENTATION SCHEME

Hyun Sang Park and Jong Beom Ra

Department of Electrical Engineering, Korea Advanced Institute of Science and Technology
373-1 Kusongdong Yusonggu, Taejon

ABSTRACT

In this paper, we propose an efficient marker extraction algorithm for initial image segmentation in a bottom-up segmentation scheme. The proposed algorithm generates dense markers in visually complex areas and coarse markers in visually uniform areas, which conforms to the human perceptual system. Experimental results show that the proposed method achieves better subjective quality for fine initial image segmentation.

1. 서론

영상분할은 균질한 특성을 가지는 영역 별로 구분하고, 각 영역 사이의 경계선을 결정하는 과정으로 정의된다. 영상분할은 컴퓨터 비전 분야에서 논의가 시작되었으며 [1]. 최근 초저전송률 환경에서의 영상 부호화에 대한 요구가 증가함에 따라서 [2], 영역기반 부호화의 핵심적인 기능으로 부각되고 있다.

영상분할은 영상 내의 객체가 가지는 실제 형태에 따라서 분할하는 것이 가장 바람직하다. 특히, 컴퓨터 비전의 문제에서는 영상분할은 토대로 영상의 내용을 기계적으로 판단해야 하기 때문에, 작은 영역이거나, 유사한 질감을 가지는 인접한 영역이더라도 정확히 분할해야 할 당위성이 있게 된다.

그러나, 영상 부호화의 입장에서, 영상분할로부터 영상의 내용을 판단하고 이해하는 주체가 인간이기 때문에 영상분할에 대한 요구사항이 컴퓨터 비전에서와 같이 엄격하지는 않다. 그대신, 유사한 질감을 가지는 인접 영역들을 하나의 영역으로 표현함으로써, 윤곽선에 대한 효율적인 부호화를 기하는 것이 더욱 중요한 문제가 된다.

이와 같이, 영상분할은 적용되는 시스템의 관점에 따라서, 상이한 요구조건을 가진다. 최근, MPEG-4 [3]에서는 영상 내의 서로 다른 의미 있는 객체 (semantic object)를

독립적으로 편집하거나 부호화하기 위해서 VOP (Video Object Plane)라는 개념을 도입했다. 이는 객체지향 부호화에 대한 본격적인 논의를 의미하며, 컴퓨터 비전의 문제인 객체분할 (object segmentation)과, 객체 단위의 영상부호화를 결합하는 계기가 되었다. 따라서, 영상분할은 객체의 대략적인 형태를 요구하던 관대한 조건으로부터, 정확한 형태를 찾아야 하는 엄격한 조건을 수용할 필요가 있게 되었다.

다양한 방법들이 영상분할을 위해서 제안되었으나 [4], 형태론적 분할에 많은 관심이 집중되고 있다. 이는 형태론적 도구들이 형태, 크기, 연결성 등과 같은 객체 지향적인 특성을 다루기 때문에, 영상분할과 같이 기하학적인 신호처리에 아주 적합한 성질을 가지고 있기 때문이다 [5]. 형태론적 영상분할은 다른 영상분할 방법들과 마찬가지로 영상단순화, 마커추출, 경계결정과정과 같이 3 단계로 구성되어 있다 [6][7]. 그러나, 이와 같은 과정을 한번만 실행함으로써, 원하는 결과를 얻을 수는 없기 때문에, 하향식이나, 상향식의 계층적인 구조가 일반적으로 사용된다 [7].

하향식 영상분할 구조 (top-down segmentation scheme)는, 전체 영상을 하나의 영역으로 간주하면서 시작한다. 상위 계층에서는 영상분할을 간략하게 수행하고, 하위 계층에서는 비균질 영역 (inhomogeneous region) 위주로, 계층을 따라서 상세한 분할을 연속적으로 수행해 나간다. 반면, 상향식 영상분할 구조 (bottom-up segmentation scheme)는 과분할 (over-segmentation)된 영상분할로부터, 원하는 결과가 나올 때까지, 유사한 인접 영역들을 계속해서 병합해 나가는 것이다. 따라서, 하향식 구조는 영역을 분할하는 규칙 (splitting rule)을 결정하는 문제가, 상향식 구조에서는 정확한 초기 영상분할 (initial image segmentation)과 영역 병합 (region merging)이 중요한 문제가 된다. 그러나, 최종적인 결과가 정확한 객체 형태를 유지하기 위해서는 구조적 차이에 무관하게 상세한

영상분할 (fine image segmentation)을 수행해야 하는 과정이 반드시 필요하며, 상향식 구조에서의 초기 영상분할과 하향식 구조에서의 마지막 영상분할이 이에 해당한다.

본 논문에서는, 원영상이 가지고 있는 중요한 영역들을 최대한 유지하면서도, 지나친 과분할의 문제를 야기하지 않는 상세한 초기 영상분할 방법을 제안하고자 한다. 이는 상향식 영상분할 구조에서 가장 근본적인 단계로서의 지위를 가진다.

2. 초기 영상분할

2.1 일반적인 마커추출 기법

마커 (marker)란, 균질한 영역을 분할하기 위해서, 그 영역의 내부에서 구성된 연결된 영역으로 정의된다. 일반적으로, 형태론적 영상분할에서 마커추출과 영상분할은 동일한 문제와 간주될 수 있다.

형태론적 영상분할에서는 형태론적 필터를 이용해서 영상을 단순화한다. 형태론적 필터는 영상 내에서 형태소 (structuring element)보다 작은 영역을 제거하면서도 영역간의 경계를 열화시키지 않는다는 장점을 가지고 있다. 그래서, 단순화된 영상에서 형태소보다 넓은 면적을 가지는 모든 연결된 영역을 마커로 추출하는 기법 [5]이 널리 사용된다. 이를 CONV 방법으로 명명하였다.

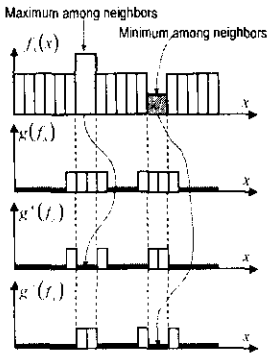


그림 1. 형태론적 경사연산자의 성능비교.

경사 영상 (gradient image)에서는 낮은 경사값으로 구성된 연결된 영역이나 극소 최저점을 마커로 추출할 수 있다. 그러나, 경사 연산자 (gradient operator)는 2차소 두께의 얇은 영역에 항상 높은 경사값을 주기 때문에, 이러한 영역이 가지는 시각적인 중요도에 상관없이 마커추출이 불가능해지며, 영역간의 경계가 완만할 경우에는 그 이상의 두께를 가지는 영역에서조차 마커를 추출하지 못하는 경우가 발생한다. 이를 해상도 손실문제 (resolution loss problem)라고 하며 [8], 정확한 영상분할을 어렵게 만드는 역할을 한다. 그림 1은 시각적으로 중요한 2개의 영역이 형태론적 경사연산자에 의해서 마커로 추출되지 못하는 경우를 도시하고 있다. 그림에서 g , g' , g'' 는 각각 형태론적 경사연산자, dilation에 의한 경사연산자, erosion에 의한 경사연산자를 나타낸다.

상세한 초기 영상분할을 위해서, 경사영상에 watershed 알고리즘을 직접 적용하거나 [9], 이에 따른 과분할을 방지하기 위해서, waterfall 알고리즘을 추가적으로 적용하기도 한다 [10][11]. 이러한 방법들은 CONV 방법에 비해서 과분할함에도 불구하고, 해상도 손실 문제로 인하여 많은 중요한 영역들이 유실되기 때문에, 낮은 주관적인 성능을 가진다. 이러한 경향은 3절에서 실험을 통해서 확인될 것이다.

CONV 방법은 단순화된 영상에서, 형태소보다 작은 영역은 큰 평탄 영역 사이를 잇는 전이영역이라는 가정을 가진다. 그러나, 이 가정을 잘못된 것이다. 형태론적 필터는 평탄영역에서 돌출된 잡음은 쉽게 제거하지만, 비평탄 영역에서는 잡음과 영역을 구분할 수 없기 때문에, 예측할 수 없는 형태로 잡음을 제거하거나, 아예 잡음을 제거하지 못하는 경우가 발생한다. 이가 가장 기본적인 형태론적 도구인 dilation이나 erosion은 형태소로 정의된 영역 안에서 최대값이나 최소값을 찾는 연산자라는 사실에서도 유추할 수 있다. 따라서, CONV 방법은 실제 객체에 해당하는 복잡한 영역에서는 저분할되는 양상을, 단순한 영역에 대해서는 과분할되는 양상을 보여주며, 이는 인간의 시각특성과 상반되는 것이다.

2.2 마커군 기반 마커추출 기법

본 논문에서 기술하는 마커추출 기법은 ZGA (Zero-Gradient marker extraction Algorithm) [12]를 개선한 것이다.

CONV 방법은 전체 영상 영역에서 마커를 추출하지만, 본 논문에서는 마커가 발생할 수 있는 영역을 제한하고자 한다. 단순화된 영상 f 에는 많은 평탄영역이 존재하지만, 그 중에서 마커로 추출할, 중요한 영역들은 다음과 같다.

- 형태소보다 넓은 영역.
- 형태소보다 작더라도 주변보다 밝거나, 어두운 영역.

전자는 일반적인 마커추출 기준이고, 후자는 시각적인 중요도에 기인한 기준이다. 형태론적 경사연산자는 이러한 영역들을 추출하기 적합한 성질을 가지고 있다. 그림 1에서처럼, dilation (또는 erosion)을 이용한 경사연산자는 주변보다 밝은 (또는 어두운) 영역 전체에 대해서 0인 경사값을 가진다. 따라서, 다음과 같은 연산자에 의해서 0의 값을 가지는 영역은 시각적으로 높은 중요도를 가지는 것으로 간주할 수 있게 된다.

$$g'''(f_s) = \min(g'(f_s), g(f_s)); \quad (1)$$

마커군 (marker cluster)은 $g'''(f_s)$ 에서 0으로만 구성된 연결된 영역을 나타낸다. 이때, 마커추출을 마커군 내부로 제한할 수 있다. 일반적으로는 하나의 밝기값만을 가지는 영역을 마커로 고려함으로써, 단순한 영역에서 지나친 과분할을 야기한다. 이를 피하기 위해서, 본 논문에서는 다음 알고리즘에 의해서 다양한 밝기값으로 구성된 마커를 추출한다.

1. $g'''(f_s)$ 에서 마커군 M 을 추출한다.

2. 마커군을 대표하는 밝기값 I_p 를 식(2)와 같이 결정한다. 식에서, $p_M(i)$ 는 마커군 M 에서, 밝기값 i 를 가지는 화소의 수를 나타내고, 합의 범위를 지정하는 파라미터는 $k = \max(1, \lceil \sigma_M \rceil)$ 와 같다.

$$I_p = \arg \max_i \sum_{j=i-k}^k p_M(i+j), 0 \leq i \leq 255. \quad (2)$$

3. $\{I_p - k, \dots, I_p + k\}$ 로 구성된 연결된 영역 R 을 마커군에서 찾는다. 마커군 M 은 R 에 해당하는 영역을 제거함으로써 갱신된다.
4. 영역 R 은 h 보다 넓은 영역일 때, 마커로 등록되며, 그렇지 않으면, 단계 3.으로 돌아간다.
5. 마커군 M 이 h 보다 넓은 영역이면, 단계 2로 돌아간다.
6. 모든 마커군에 대해서 마커추출이 이루어지지 않았으면, 단계 1로 돌아간다.

위의 알고리즘에서 h 는 마커의 크기에 대한 기준이다. 이는 단순화된 영상에서 의미있는 영역이 가지는 최소크기를 기준으로 삼는 것이 타당하다. 형태론적 필터로 이진영상 (binary image)을 단순화할 경우에는, 형태소보다 작은 영역들은 모두 제거되므로, 형태소의 크기는 적합한 기준이 된다. 그러나, 그레이 레벨을 가지는 일반적인 영상에서는 단순화 후에 1화소 크기의 영역까지도 남아있게 된다. 따라서, 형태소의 크기별 기준으로 삼는 것은 저분할을 야기할 개연성을 가지게 된다. 본 논문에서는 식(3)과 같이, 형태소의 한 면의 길이별 크기 기준으로 삼는다.

$$h = 2N + 1. \quad (3)$$

여기서, N 은 형태소의 크기별 나타내는 파라미터이고, 형태소의 넓이는 $(2N + 1)^2$ 가 된다.

3. 실험결과

상세한 초기영상분할을 위해서, 크기 1의 open-close by reconstruction 필터를 사용해서, 가장 낮은 수준의 단순화를 수행한다. 마커추출은 제안한 방법에 의해서, 크기 기준 $h=3$ 을 가지고 수행한다. 추출된 마커에 대해서 영역성장의 기능을 수행하는 변형된 watershed 알고리즘 [7]을 적용해서 영역간의 경계를 결정한다.

그림 2(a)는 실험에 사용한 QCIF 영상을 나타낸다. 비교를 위해서, 3 종류의 다른 영상분할 결과를 같이 도시한다. 그림 2(b)는 경사영상에 직접 watershed 알고리즘 [9]을 적용한 결과이고, 이 결과에 다시 waterfall 알고리즘 [10]을 적용한 결과가 그림 2(c)에 해당한다. 앞 절에서 언급한 바와 같이 경사영상을 토대로 한 영상분할은 중요한 부분이 유실되는 단점을 보이는 것을 알 수 있다. 그림 2(d)는 CONV 방법 [5]에 의한 영상분할 결과이다.

제안한 방법에 의해서, *Miss America*, *Akiyo* 영상은 다른 방법들에 비해서 적은 수의 영역으로 분할되었으며, 주관적 화질의 향상을 보여준다. 특히, *Miss America*의

머리와, *Akiyo*의 오른쪽 귀절이와 같이 시각적으로 중요한 부분들이 제안한 방법에 의해서는 표현됨에도 불구하고, 더 적은 수의 영역으로 분할된다는 사실은 상당히 고무적인 일이다.

Carphone 영상에서는 제안한 방법이 CONV 방법에 비해서 과분할되지만, 이는 제안한 방법이 과분할되기 보다는 CONV 방법이 저분할된 것으로 해석할 수 있다. Watershed 알고리즘의 적용결과는 영상이 얼마나 많은 상세영역으로 구성되어 있는지를 알려준다는 관점에서, 다른 두 영상에서는 CONV 방법이 watershed 적용결과와 1/2 정도의 영역을 가짐에 비해서, 이 경우에는 1/3로 줄어든다는 사실은 CONV 방법에 의한 저분할 현상을 보여주는 것으로 해석할 수 있게 한다.

4. 결론

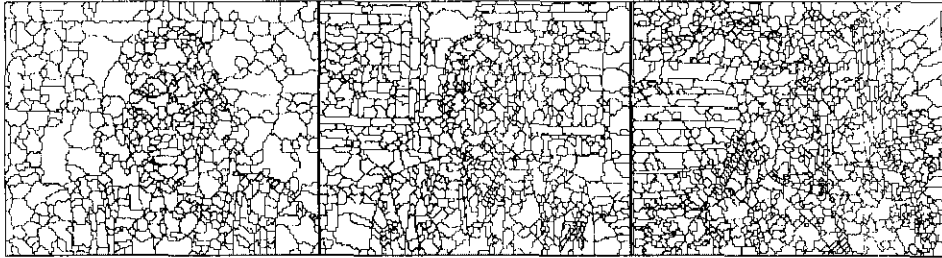
본 논문에서는 상세한 영상분할에 적합한 효율적인 마커추출 기법에 대해서 기술하였다. 제안한 방법은 복잡한 영역에 대해서 많은 마커를, 단순한 영역에서 적은 마커를 추출함으로써, 인간의 시각체계를 따르기 때문에, 다른 방법에 비해서 우수한 주관적 성능을 가진다. 따라서, 제안한 방법은 상향식 영상분할 구조에서 초기 영상분할로서 아주 적합하다.

5. REFERENCES

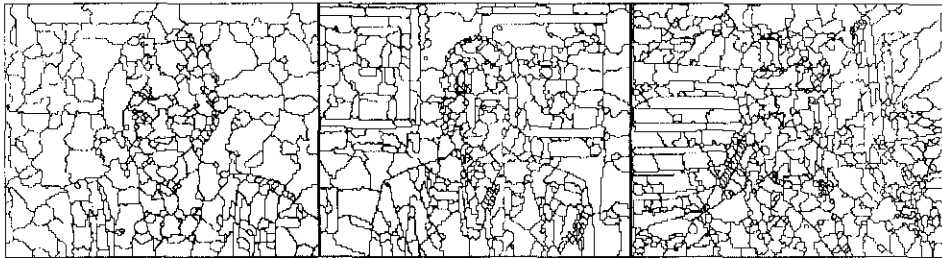
- [1] D.H. Ballard and C.M. Brown, *Computer Vision*, Prentice Hall, New Jersey, 1982
- [2] M. Kunt, A. Ikononopoulos, and M. Kocher, "Recent results in high compression image coding," *IEEE Trans. on Circ. and Syst.*, vol. 34, no. 11, pp. 1306-1336, Nov. 1987.
- [3] ISO/IEC JTC1/SC 29/WG 11/1469, "MPEG-4 video verification model V.5.0," Nov. 1996.
- [4] R.M. Haralick and L.G. Shapiro, "Image segmentation technique," *Computer Vision Graphics and Image Processing*, vol. 29, pp. 100-132, 1985.
- [5] P. Salembier and J. Serra, "Flat zones filtering, connected operators, and filters by reconstruction," *IEEE Trans. on Image Process.*, vol. 4, no. 8, pp. 1153-1160, Aug. 1995
- [6] F. Meyer and S. Beucher, "Morphological segmentation," *Journal of Visual Comm. and Image Representation*, vol. 1, no. 1, pp. 21-46, Sep. 1990.
- [7] P. Salembier, "Morphological multiscale segmentation for image coding," *Signal Processing*, vol. 38, pp. 359-386, 1994.
- [8] J. Crespo and J. Serra, "Morphological pyramids for image coding," *Proc. of SPIE Visual Comm. and Image Process.*, vol. 2727, pp. 159-170, 1996.
- [9] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," *IEEE Trans. on Patt. Anal. and Mach. Intell.*, vol. 13, no. 6, pp. 583-598, June 1991.
- [10] S. Beucher, "Watershed, hierarchical segmentation and waterfall algorithm," *Mathematical Morphology and its applications to Image Processing*, pp. 69-76, Kluwer Publishers, 1994.
- [11] L. Shafarenko, M. Petrou, and J. Kittler, "Automatic watershed segmentation of randomly textured color images," *IEEE Trans on Image Process.*, vol. 6, no. 11, pp. 1530-1544, Nov. 1997.
- [12] H.S. Park and J.B. Ra, "Zero-gradient marker extraction for morphological image segmentation," *Proc. of ITC-CSCC '98*, pp. 205-208, July 13-15, 1998.



(a) 원영상: Miss America, Akiyo, Carphone



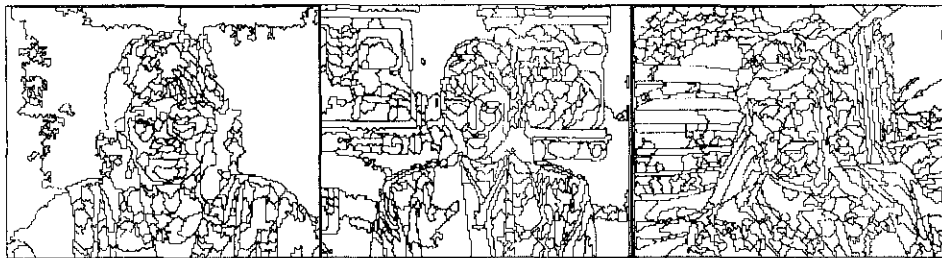
(b) Watershed 영상분할 (영역수): 620, 823, 1075.



(c) Waterfall 영상분할 (영역수): 332, 457, 630.



(d) CONV 방법 (영역수): 366, 436, 346.



(e) 제안한 방법 (영역수): 284, 400, 472.

그림 2. 다양한 영상분할 결과.