

계층적인 메쉬 구조를 이용한 영상분할 방법

임 동근, 호요성

광주과학기술원 정보통신공학과
500-712 광주광역시 북구 오룡동 1번지

Image Segmentation Using Hierarchical Meshes

Dong-Keun Lim and Yo-Sung Ho

Kwangju Institute of Science and Technology
1 Oryong-Dong Puk-Gu, Kwangju, 500-712, Korea
Email : {dklim,hoyo}@kjist.ac.kr

ABSTRACT

The object boundary of an image plays an important role for image interpretation. In this paper, we introduce a concept of hierarchical mesh-based image segmentation for finding object boundaries. In each hierarchical layer, we employ neighborhood searching and boundary tracking methods to refine the initial boundary estimate. We also apply a local region growing method to define closed contours. Experimental results indicate that reliable segmentation of objects can be accomplished by the proposed low complexity technique.

1. 서론

영상은 여러 개의 물체들로 구성되어 있기 때문에 각 물체의 경계선은 영상을 표현하는데 매우 중요한 정보이다. 물체의 경계를 찾는 영역분할 방법은 영상신호 처리에서 매우 중요한 기술중의 하나이다. 지금까지 영상분할을 위한 여러 가지 알고리즘들이 제안되었으며, 정적인 배경에 말하는 사람이 있는 화면으로 구성된 영상시퀀스를 가진 영상회의와 같은 분야에 응용하려는 시도가 늘고 있다[1-2]. 또한, MPEG-4 국제표준은 영상 자체를 여러 개의 물체로 구분하여 각각을 개별적으로 부호화하고 조합하는 방법을 다루고 있는데, 이를 제대로 사용하기 위해서는 영상분할이 먼저 이루어져야 한다.

지금까지 개발된 대표적인 영상분할 방법에는 경계값 분할(thresholding), 영역 확장(region growing), 분할 및 병합(split-and-merge), watershed, 에지 기반(edge-based) 영역분할 등

이 있으며, 각 방법들은 나름대로의 특성과 장점을 가지고 있다.

일반적으로 영역분할은 휘도(luminance)값에 기반하여 영상을 유사한(homogeneous) 영역들로 분할한다. watershed와 영역확장 방법은 이 목적을 위해 가장 널리 사용되는 방법이다. 그렇지만 이 방법들은 영역 단순화(region simplification)와 같은 처리를 수행해야 하므로 많은 계산량이 필요하다.

이 문제를 해결하기 위하여 본 논문에서는 계층적인 메쉬 구조(hierarchical meshes)를 사용한 영역 분할 방법을 제안하였다. 이것을 이용하면 원하는 정밀도를 가진 계층에서 실행을 중지할 수 있으며, 계층이 진행됨에 따라 물체의 경계선이 점점 세밀하게 보정(refine)된다. 또한, 물체 경계선 근처의 영역만을 후보로 삼아 계층적 영상분할이 진행되므로 영상 전체에 대해 적용되는 다른 방법들보다 계산량이 상당히 감소한다. 제안한 방법은 신뢰성있는 연결된 닫힌 경계선(closed contour)을 탐색하면서 영역을 분할할 수 있으며, 계산의 복잡도가 작다.

2. 새로운 영역분할 방법

본 논문에서는 영상시스템을 2 단계로 분류하였다. 1 단계에서는 단순한 하위 레벨의 영상처리 기법을 사용하여 영상의 특성을 추출하고, 2 단계에서는 1 단계의 결과에 대해 계층적인 구조를 사용하는 상위 레벨의 영상처리 기법을 적용한다.

제안한 영역분할 방법은 피라미드 데이터

구조를 가진 계층적인 메쉬 분류에 기반한다. 계층적인 접근 방법의 특징은 거친(coarse) 레벨에서 만들어진 대략적인 분류 결과를 점점 보정(refine)해 나가는 데 있다. 또한 피라미드 구조는 영상의 관심있는 부분에 초점을 집중시킬 수 있다. 본 논문에서는 물체의 경계선에 주로 관심을 두었다. 따라서 제안한 알고리즘은 top-down 분류에서 모든 피라미드 데이터의 일부만을 다루기 때문에 계산량을 효과적으로 줄일 수 있다[1,4].

본 논문에서 제안한 알고리즘의 주된 아이디어는 경계선 검출(boundary detection)에 사용되는 divide-and-conquer 알고리즘과 매우 유사하게 계층적인 메쉬 구조를 만든다. 각 계층에서 물체 경계선을 포함하는 메쉬는 다음 단계에서 보다 정밀한 메쉬 구조를 얻기 위한 후보가 된다. 이 과정에서 물체 경계선을 포함하는 크기가 작아진 보정된 메쉬 구조가 만들어진다. 여러 가지 다른 크기의 계층적인 메쉬 구조는 1단계 영상처리 결과인 에지 정보를 사용하여 구하며, 메쉬는 각 후보 블록의 무게 중심(centroid)을 연결하여 만든다. 이렇게 만들어진 메쉬 구조의 최외각선만을 좌수법(left-hand rule)과 우수법(right-hand rule)을 선택적으로 사용하여 탐색(tracking)하고, 이 결과는 다음 계층의 메쉬를 찾는 후보로 사용된다. 이와 같이 계층이 진행됨에 따라 물체 경계선 근처에만 후보가 분포하게 되고 그 갯수는 점점 줄어들어 계산량이 감소된다. 이와 같은 방법으로 최종적으로 화소 단위의 경계선까지 찾아낸 후 경계선 내부를 채우고, 이를 마스크로 사용하여 영역분할을 수행한다. 그림 1은 제안한 방법의 전체 흐름도이다.

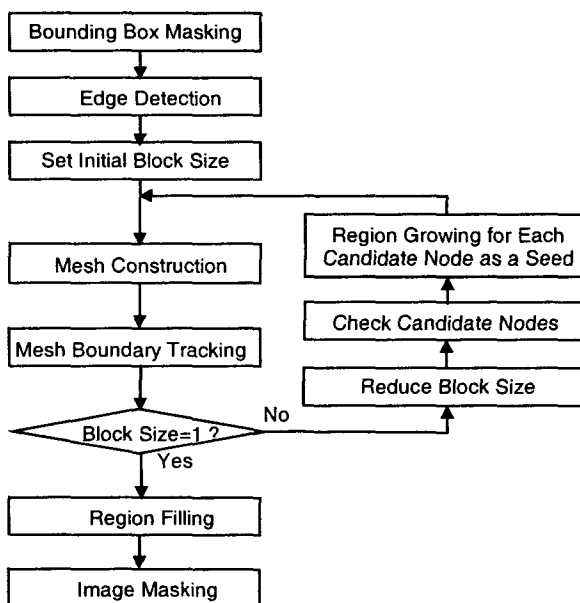


그림 1: 계층적 메쉬 구조에 기반한 영상분할

여기서 각 계층에서 후보 메쉬들은 좀 더 작은 메쉬들로 분리된다. 우리는 메쉬경계 추적방법과 메쉬경계 연결방법을 적용하였다. 이 과정은 원하는 정밀도에 이를 때까지 반복적으로 수행된다. 각 계층에서는 이전 계층의 개략적인 경계선에 대하여 지역적인 탐색이 수행된다.

결과적인 물체의 경계는 정밀하지 않기 때문에 경계보정 작업이 추가적으로 필요하다. 이를 위해 이전 계층의 경계선 위의 점을 시작점(seed)으로 사용하여, 그 주변에 대하여 영역확장 방법을 사용한다. 영역분할 마스크는 보정된 물체 경계선의 내부 영역을 채워서 구한다.

2.1 물체 경계의 구획

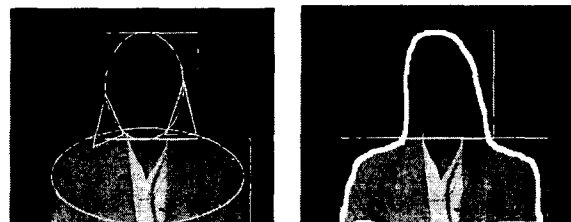
영상 자체에는 여러 가지 물체가 존재하며 원하는 물체가 어느 것인지를 자동적으로 찾는 것은 매우 어려운 일이다. 따라서, 자동 영역분할은 물체가 명확히 정의될 수 있는 일부 응용에만 사용될 수 있다. 일반적인 응용에서는 사용자가 원하는 물체의 경계 영역을 지적하여 정의할 수 있다.

그림 2와 같이 여러 개의 사각형 상자, 타원, 사용자가 지정한 관심 영역(ROI, region of interest) 표시선을 사용하여 경계를 구획지을 수 있다. 또한 이것들을 결합하여 사용할 수도 있다.



(a) 여러 개의 사각형

(b) 사각형과 타원



(c) 사각형, 타원, 다각형

(d) 사용자 지정 곡선

그림 2: 관심영역의 경계

2.2 경계선 추적(Boundary Tracking)

우리는 보통 이차원이나 삼차원의 영상신

호를 모델링하여 메쉬 구조로 표현할 수 있다 [2,8]. 메쉬들은 일반적으로 길이가 긴 물체의 경계선을 따라서 분포한다. 이 경계선은 물체라는 의미를 제공하기 때문에 계층적인 구조에서 중요한 사항이다. 즉, 길이가 긴 경계선은 물체를 의미한다고 할 수 있다. 물체 경계선은 에지선들의 연결에 의해 구해진다. 개별적인 지역 에지선(local edge, edge segment)을 따라서 물체 경계선이 구해지도록 단계별로 여러 계층의 영역분할이 이루어진다.

후보 경계선의 위치를 찾은 후, 물체 경계에 대한 정밀도를 보정해야 한다. 그림 3은 계층적인 메쉬의 구성 방법과 경계선 연결 방법을 설명하고 있다. 계층 k에서 에지들은 3개의 사분면에 존재한다. 이 영역들은 계층 k+1에서의 후보 위치가 된다. 비슷한 방법으로, 계층 k+1에서 에지들은 7개의 영역 안에 위치하며, 이 영역들은 계층 k+2에서의 후보 위치가 된다.

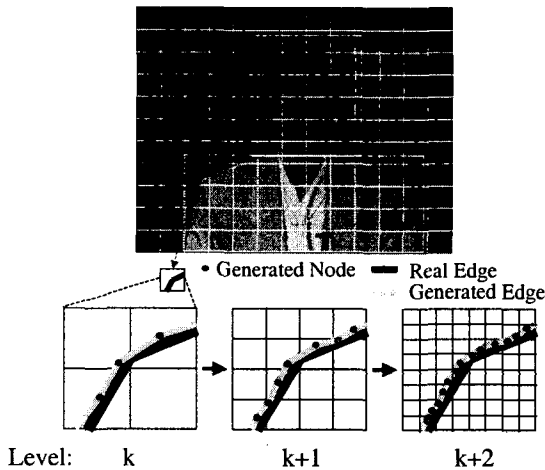


그림 3: 서로 다른 레벨에서 물체의 위치

그림 4는 복잡한 물체의 경계선을 추적하는 예를 보여준다. 마치 복잡한 해안선과 같은 그림인데, 계층이 진행됨에 따라서 물체의 경계선이 점진적으로 보정되고 있다.

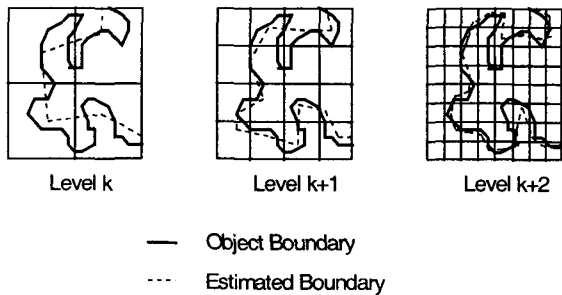


그림 4: 경계선 보정

메쉬 구조는 다음 식으로 정의되는 각 블록의 무게중심(centroid) (C_x, C_y)를 연결하여 구한다.

$$c_x^{i,j} = \frac{\sum_{k=0}^{BS-1BS-1} \sum_{l=0}^{BS-1BS-1} k \cdot I(i \cdot BS + k, j \cdot BS + l)}{\sum_{k=0}^{BS-1BS-1} \sum_{l=0}^{BS-1BS-1} I(i \cdot BS + k, j \cdot BS + l)} \quad (1)$$

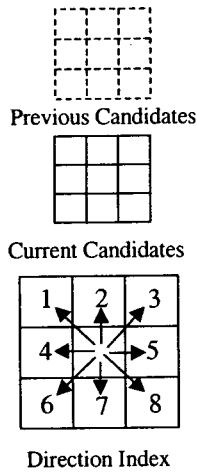
$$c_y^{i,j} = \frac{\sum_{k=0}^{BS-1BS-1} \sum_{l=0}^{BS-1BS-1} l \cdot I(i \cdot BS + k, j \cdot BS + l)}{\sum_{k=0}^{BS-1BS-1} \sum_{l=0}^{BS-1BS-1} I(i \cdot BS + k, j \cdot BS + l)} \quad (2)$$

여기서 (i,j)는 각 블록의 위치를 나타내는 인덱스(index)이다. BS는 블록 크기, I(.)는 해당 위치의 에지값을 나타낸다. 에지값을 사용하여 구한 무게중심이 영상의 휘도값을 사용하여 구한 무게중심보다 좀 더 물체의 경계에 잘 접근하였다.

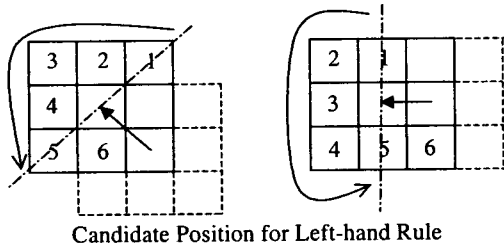
정확한 에지 경계 추적을 위해 좌수법과 우수법을 사용하였다. 이때 다음과 같은 방법으로 효과적으로 물체 경계를 찾을 수 있다.

- 영상의 왼쪽 위에서 시작하여 에지를 가진 블록이 처음 나타날 때까지 차례로 주사(raster scan)한다. 처음 찾은 위치는 경계선 추적을 위한 시작 위치가 된다. 이 위치에서 처음에는 좌수법을 사용한 경계선 추적이 사용된다.
- 후보 위치의 우선순위(priority)에 따라 에지를 가진 블록을 찾기 위해 주변의 3x3 블록을 탐색한다. 이 과정이 그림 5와 그림 6에 설명되어 있다. 만약 완전히 닫힌(closed) 경계선을 찾기 전에 영상의 경계에 이르면, 이전의 시작 위치로 되돌아와서 반대 방향으로 다른 탐색 방법인 우수법을 적용한다. 이 과정은 모든 물체의 경계를 검사할 때까지 반복적으로 수행된다.
- 에지를 가진 것으로 판명된 에지 블록의 위치에서, 에지 경계를 따르는 이전 단계의 움직임 방향을 저장하기 위하여 DIR이라는 변수를 사용한다. 다음 경계선 위치는 그림 6의 우선순위표에 따른 순서로 탐색된다. 이렇게 구해진 위치는 다음 위치의 DIR이 된다.
- 에지 블록의 크기를 줄이기 위해 물체의 경계에서 지역적인 영역확장이 이루어진다.

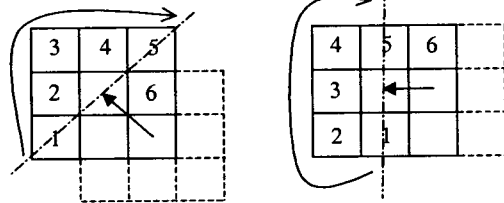
e) 위의 과정은 화소단위 해상도의 물체 경계선에 이를 때까지 계속해서 반복된다.



(a) 방향 인덱스



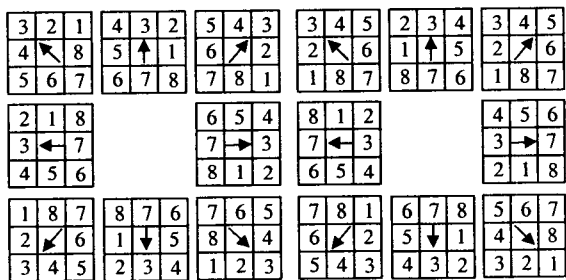
Candidate Position for Left-hand Rule



Candidate Position for Right-hand Rule

(b) 후보 위치에서의 탐색

그림 5: 경계선 추적



(a) 좌수법

(b) 우수법

그림 6: 경계 추적을 위한 우선순위

그림 7은 경계선 추적의 예를 보여주고 있다. 표 1에는 이때의 에지 위치와 방향(DIR)이 정리되어 있다. 처음에는 좌수법으로 시작하였지만, No. 18에서 영상의 경계에 이르러서,

다음 단계에 시작 위치로 돌아와서 반대 방향으로 우수법이 적용된다. No. 13에서 No. 16은 결가지라는 것을 알 수 있으며, 닫힌 영역분할 마스크를 얻기 위해 제거된다.

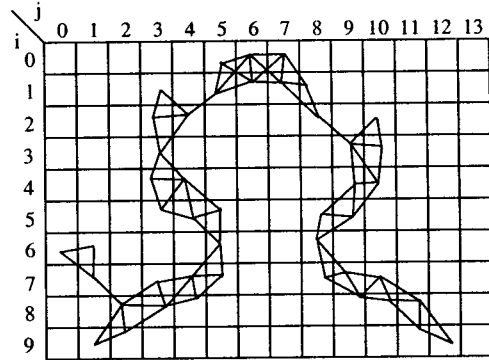


그림 7: 경계선 추적의 예

표 1: 방향(DIR)과 위치값(Pos.)

No	Pos	DIR	Scan	No	Pos	DIR	Scan
0	(0,5)	7	left	17	(8,2)	8	left
1	(1,5)	6	left	18	(9,1)	6	left
2	(2,4)	1	left	19	(0,5)	5	right
3	(1,3)	7	left	20	(0,6)	5	right
4	(2,3)	7	left	21	(0,7)	8	right
5	(3,3)	7	left	22	(1,8)	7	right
6	(4,3)	7	left	23	(2,8)	8	right
7	(5,3)	7	left	24	(3,9)	3	right
8	(5,4)	5	left	25	(2,10)	7	right
9	(6,5)	8	left	26	(3,10)	7	right
10	(7,4)	6	left	27	(4,10)	6	right
11	(7,3)	4	left	28	(5,9)	6	right
12	(8,2)	6	left	29	(6,8)	8	right
13	(7,1)	1	left	30	(7,9)	5	right
14	(6,1)	2	left	31	(7,10)	8	right
15	(6,0)	4	left	32	(8,11)	8	right
16	(7,1)	8	Left	33	(9,12)		end

2.3 에지 검출과 영역 확장

물체들의 경계 위치에 있는 화소들은 에지 검출기의 출력에서 지역적인 최대값을 나타낸다. 이러한 에지들은 영상의 특성을 잘 설명할 수 있기 때문에 영상분할에서 매우 중요한 정보이다. 영상분할의 목적은 경계에서 휘도값의 변화가 일어나는 에지들을 연결하여 닫힌 영역을 찾아내는 것이다.

영역확장은 영상분할을 위한 단순한 방법이다. 만약 두 개의 인접한 화소의 휘도값의 차이가 어떤 문턱치(threshold)보다 작으면 두 화소는 결합되어 한 영역에 포함된다. 본 논문에서는 물체를 이루는 조각들(segments)을 결합시키기

위해 화소 휘도값 차이 대신에 각 조각의 통계적 특성인 평균값을 사용하였다.

에지 검출과 영역 확장은 영역 경계에서 스텝 에지가 있고 영역 내에서는 평활화(smoothing)된 휘도값을 가진다는 가정하에서는 동일한 과정이다. 두 과정 모두 물체 경계선을 찾거나 균일한 영역을 구한다[1,7]. 균일성(homogeneity)은 색상(color), 텍스춰(texture), 움직임(motion), 깊이(depth) 등을 통해 측정할 수 있는데, 여기서는 휘도값의 변화를 사용하였다. 혼성 연결(hybrid link)을 사용한 영역 확장에서 에지 검출기는 각 화소에 대해 적용하여 에지(edge) 혹은 비에지(non-edge)로 구분한다. 주변에 있는 화소중 에지에 속하지 않는 화소들은 영상 조각들(segments)로 결합된다. 물체 조각들은 비에지 화소들을 결합하여 이루어진다. 잡음이나 복잡한 영상들에 대한 에지 검출은 잘못된 물체 조각을 만들기 때문에 단하지 않은 경계선이 구해질 수도 있다.

통계적 연결(statistical link)을 사용한 영역 확장은 주변 영상 조각들의 의미있는 통계적 특성과 결합하려는 화소값을 비교한다. 통계적 특성의 한 예는 사각형 블록의 평균 화소값이다. 두 개의 인접한 영상 조각들은 경계선을 따라서 균일한 값을 가지면 결합된다. 이 방법은 항상 연결된 닫힌 구조를 가지지만, 계산량이 많아진다.

앞의 두 방법을 결합시킨 방법은 개별적인 방법을 적용했을 때 발생하는 문제들을 상호 보완하게 된다. 그렇지만 개별적인 방법이 가지는 장점을 완전하게 가지지는 못하기 때문에 지역적으로만 사용된다. 이 방법은 일반적으로 연결된 에지 구조를 가지게 된다. 다음 계층에서 블록의 크기를 줄일 때, 블록이 뚜렷한 에지를 가지는지 여부를 검사해야 한다. 이를 위하여 각 블록의 평균 에지 값을 문턱치와 비교한다.

에지 검출을 위해 일반적으로 많이 사용되는 Sobel 검출기를 사용하였다. 여기서, 수직 방향의 1차 도함수(first-order derivative)를 구하기 위한 연산자 h_1 과 수평 방향의 1차 도함수를 구하기 위한 연산자 h_2 는 다음과 같다.

$$h_1 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}, h_2 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

지역적인 영역 확장의 결과는 다음 계층의 물체 경계의 추적을 위한 초기 후보값이 된다.

2.4 영역 채움 (Region Filling)

영역분할 마스크를 찾기 위하여 물체 경계

선의 정보를 이용하여 내부 영역을 채운다. 이를 위하여 일반적인 다각형 스캔변환(polygonal scan conversion) 알고리즘을 사용한다[9]. 이 알고리즘은 볼록 다각형(convex)과 오목 다각형(concave)을 모두 다루는데 유용하다. 그림 8은 영역 채움 과정을 보여준다. 이 과정은 다음과 같이 3 단계로 수행된다.

- 다각형(polygon)의 에지와 스캔 라인(scan line)의 교차점을 찾는다.
- x 축 값이 작은 순서부터 정렬한다.
- 두 개의 교차점 사이에 있는 다각형의 내부의 모든 화소를 경계선의 화소값으로 채운다. 이때 영역의 내부에 있는 화소인지를 검사하기 위해 홀수-패리티 규칙(odd parity rule)을 사용한다. 패리티는 최초에는 짝수로 지정하고 교차점을 만날 때마다 패리티를 바꾼다. 패리티가 홀수이면 해당 화소를 채우고, 아니면 채우지 않는다.

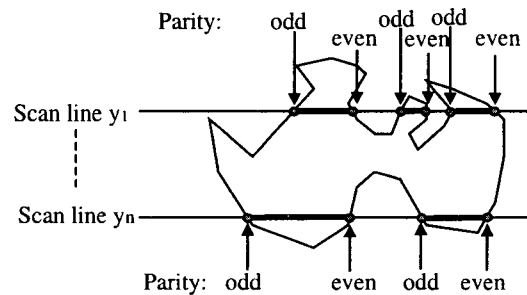


그림 8: 영역 채움

2.5 영상 마스크 (Image Masking)

영상의 물체를 얻기 위해, 원래 영상과 최종 결과인 영상분할 마스크와 AND 연산을 수행한다.

3. 실험결과

본 논문에서 제안한 영상분할 알고리즘을 검증하기 위해 CIF (352x288) 크기를 가진 영상을 이용하여 실험을 수행하였다. 머리와 어깨 부분을 가진 CLARE와 AKIYO 영상을 사용하였다. 그림 9는 서로 다른 블록 크기를 가진 두 계층의 메쉬 구조를 보여준다.

첫번째 계층에서는 영상을 16x16 크기를 가진 블록으로 나누고, 관심있는 영역에서 메쉬를 생성한다. 메쉬들은 주로 물체의 에지를 따라서 분포한다. 경계선 추적 방법을 사용하여 각 계층에서 완전하게 닫히도록 연결된 경계선을 구한다. 다음 계층에서는 좀 더 작은 8x8 크기의 블록을 사용하여 물체 경계선을 보정한다. 여기서 적용되는 지역적인 영역 확장은 닫힌 형태의 경계선을 찾도록 하는데 도

움을 준다. 메쉬들이 주로 물체의 경계를 따라서 분포하기 때문에 다음 계층에서의 계산량을 줄이게 된다.



(a) BS=16

(b) BS=8

그림 9: 블록 크기가 다른 계층적 메쉬 구조

그림 10은 제안한 방법의 실험 결과의 한 예를 보여준다. 여기서 각 화면은 서로 다른 크기의 블록을 사용하고 있으며, 그림에 있는 사각형은 블록의 크기를 나타낸다. 다음 단계로 진행함에 따라 물체의 경계선이 보정됨을 알 수 있다.



(a) BS=16

(b) BS=8



(c) BS=4

(d) BS=2



(e) BS=1

(f) 분할 마스크

그림 10: 영상분할 결과

영역 분할된 영상은 원 영상과 영역 분할 마스크를 AND 연산하여 구한다.

4. 결론

본 논문에서는 계층적인 메쉬를 사용한 새

로운 영상분할 알고리즘을 제안하였다. 물체의 신뢰성있는 영역분할을 얻으면서 계산량을 줄이는 방법을 사용하였다. 구해진 결과는 영상의 에지 정보와 물체의 경계 정보를 유지하기 때문에 시각적으로 향상된 성능을 얻기 위한 객체기반 부호화(object-based coding)에 사용될 수 있다. 또한, 이 결과는 형상 부호화(shape coding)와 메쉬 구성을 위한 초기 경계선 정보로 사용할 수 있다.

감사의 글

본 연구는 광주과학기술원(K-JIST) 초고속 광네트워크(UFON) 연구센터를 통한 한국과학재단 우수연구센터(ERC) 지원금에 의한 것입니다.

참고문헌

- [1] L. Westberg, "Hierarchical contour-based segmentation of dynamic scenes," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, no. 9, pp. 946-952, September 1992.
- [2] Y. Altunbasak, "Object-scalable mesh-based coding of synthetic and natural image objects," *ICIP'97*, pp. 94-97, October 1997.
- [3] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. 13, pp. 583-598, June 1991.
- [4] L. Najman and M. Schmitt, "Geodesic saliency of watershed contours and hierarchical segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, pp. 1163-1173, December 1996.
- [5] S.A. Hojjatoleslami and J. Kittler, "Region growing: a new approach," *IEEE Trans. Image Processing*, vol. 7, no. 7, July 1998.
- [6] T. Pavlidis and Y.T. Liow, "Integrating region growing and edge detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 7, no.3, pp. 225-233, September 1992.
- [7] M. Tabb and N. Ahuja, "Multiscale image segmentation by integrated edge and region detection," *IEEE Trans. Image Processing*, vol. 6, pp. 642-655, May 1997.
- [8] O. Lee and Y. Wang, "Nonuniform image sampling and interpolation over deformed meshes and its hierarchical extension," *Visual Comm. And Image Proc. SPIE* vol. 2501, pp.389-400, May 1995.
- [9] J.D. Foley, et al., *Introduction to computer graphics*, Addison-Wesley, 1994.