

# MPEG-4 기반의 능동윤곽모델을 이용한 스테레오 영상에서의 객체분할에 관한 연구

김신형\*, 전병태\*\*, 박두영\*, 장중환\*

\*배재대학교 정보통신공학과

\*\*한국전자통신연구원

zeros@mail.pcu.ac.kr, chunbt@etri.re.kr, dypark@mail.pcu.ac.kr, jangjw@mail.pcu.ac.kr

## A Study on the Object Segmentation Using Active Contour Model based MPEG-4

Shin-Hyoung Kim\*, Byung-Tea Chun\*\*, Doo-Yeong Park\*, Jong-Whan Jang\*

\*Dept. of Information and Communication Engineering, PaiChai University

\*\*Electronics and Telecommunications Research Institute

### 요 약

본 논문에서는 능동윤곽모델(active contour model)로 잘 알려져 있는 스네이크(snake) 알고리즘을 스테레오영상에 적용하여 좌·우 영상의 disparity 정보를 이용, 객체의 경계선을 찾는 알고리즘을 제안한다. 스네이크는 객체의 경계를 얻기 위해 에지정보를 사용하는데 실제 이미지에서 객체의 경계가 아닌 인접한 주위의 강한 에지(edge)에 대해서도 영향을 받게 되는 문제가 있다. 이러한 문제를 해결하기 위해 스테레오영상의 disparity 정보를 이용하여 이를 개선하고 disparity 측정에 사용되는 블록매칭(block matching)방법을 스네이크 알고리즘에 적용시켰다.

### 1. 서론

영상분할(segmentation)은 컴퓨터 비전과 같은 분야에서 영상내의 객체를 해석하고 응용하기 위해 연구되어 왔다. 특히, 멀티미디어 표준안으로 제정된 MPEG-4에선 VOP(Video Object Plane) 개념을 소개하고 내용기반 부호화방식(content based coding)을 채택하였다. 내용기반 부호화방식은 영상을 영상 내 객체단위로 부호화 하여 압축률을 높이고 사용자 접근(user interactivity)을 가능하게 하여 내용기반 색인(content based indexing) 및 검색 그리고 영상 전송 등 다양한 응용이 가능하게 되는 장점이 있다. 이러한 내용기반 부호화방식에 있어서 가장 중요한 전처리 작업이 객체를 분할하는 영상분할 과정이다. 최근 인간의 다양한 욕구를 충족시키기에는 2차원 영상이 한계가 있기 때문에 2차원 영상에 입체감을 더한 3차원 영상의 정보 처리에 관한 연구가 다양한 분야에서 진행되고 있다. 그래픽스와 실사(real image)를 합성하여 표현하는 분야에서도 영상분할은 필수 과정이라 할 수 있다.

영상분할에 관한 연구는 화소들간의 유사성을 측정하는 임계치 방법(thresholding), 영역 성장(region growing), 영역 분할(region splitting), 영역 병합(region merging)<sup>[1][2]</sup>이 있으며 영역성장 방식의 일종으로 영역간의 경계를 찾는 분수령 알고리즘(watershed algorithm)<sup>[3][4]</sup>과 정의된 에너지 함수

를 최소화하는 과정을 통해 객체의 경계(boundary)를 추출하는 능동윤곽모델(active contour model)<sup>[5][6][7]</sup> 등이 제안되었다. 분수령 알고리즘은 영역간 경계면이 끊어지지 않고 하나의 폐곡선으로 객체를 분할할 수 있는 장점이 있지만 영상의 과분할(over segmentation)로 인한 영역병합과 같은 후처리를 해주어야 하는 단점이 있고, 능동윤곽모델은 객체에 인접한 주위의 강한 에지에 영향을 받는 문제가 있다.

본 논문에서는 스테레오 영상에서 좌·우 영상의 상대적 변이인 disparity와 능동윤곽모델로 잘 알려진 스네이크 알고리즘을 이용해 객체의 경계를 검출하는 방법을 제시한다.

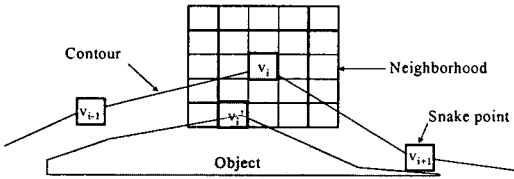
논문의 구성은 다음과 같다. 2장에서는 스네이크 알고리즘에 대해 설명하고, 3장에서는 스테레오 영상에서 disparity 측정을 위한 블록매칭에 대하여, 4장에서는 제안한 알고리즘에 대해 설명하였다. 그리고 실험결과와 5장에, 결론은 6장에 기술하였다.

### 2. 스네이크 알고리즘

스네이크 알고리즘은 에너지함수를 최소화하는 방향으로 스네이크 포인트(snake point)들이 반복적으로 움직여가며 객체의 경계를 찾아내는 방법이다. Williams과 Shah<sup>[6]</sup>는 각 스네이크 포인트마다 가중

치를 두어 안정적이고 빠르게 단일 객체의 경계를 찾는 greedy snake를 제안하였다.

$i=0, \dots, N-1$  일 때  $v_i=(x_i, y_i)$ 가 윤곽선(contour) 상의 스네이크 포인트들이라고 하면  $x_i, y_i$ 는  $i$  번째 스네이크 포인트의  $x, y$  좌표를 나타낸다.  $i$  번째 스네이크 포인트  $v_i$ 는 이전 스네이크 포인트인  $v_{i-1}$ 와 이후 스네이크 포인트인  $v_{i+1}$ 를 이용하여 이웃하는 포인트들 중 에너지가 가장 최소가 되는  $v_i$  포인트로 이동하게 된다.



[그림 1] 스네이크 포인트의 이동 예

스네이크 포인트들을 움직이게 하는 에너지 함수는 식 (1)과 같이 크게 내부에너지(internal energy)와 영상 에너지(image energy) 2개 함수로 나눌 수 있고, 각 함수는 식 (2)와 (3)처럼 3개의 항으로 나타낸다. 매개변수인  $\alpha, \beta$  그리고  $\gamma$ 는 3개의 항에 대한 에너지의 관계를 균형 있게 해주는 역할을 한다.

$$E_{snake}(v_i) = E_{internal}(v_i) + E_{image}(v_i) \quad (1)$$

$$E_{internal}(v_i) = \alpha \cdot E_{continuity}(v_i) + \beta \cdot E_{curvature}(v_i) \quad (2)$$

$$E_{image}(v_i) = \gamma \cdot E_{edge}(v_i) \quad (3)$$

각 항에 대해 알아보면  $E_{continuity}(v_i)$ 는 스네이크 포인트들 간의 사이 거리를 균일하게 하는 에너지로 식 (4)와 같이 스네이크 포인트들의 평균거리  $\bar{d}$ 와 두 스네이크 포인트간에 거리의 차로 유도된다.

$$E_{continuity}(v_i) \approx \bar{d} - |v_i - v_{i-1}|^2 \quad (4)$$

$$|v_i - v_{i-1}|^2 = (x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 \quad (5)$$

$E_{curvature}(v_i)$ 는 스네이크의 굽어짐(rigidity)을 제어하는 에너지로 스네이크 포인트를 윤곽선 안쪽으로 끌어당기게 된다. 식은 다음과 같이 현재 스네이크 포인트를 기준으로 이전 스네이크 포인트와 이후 스네이크 포인트를 이용하여 유도된다.

$$E_{curvature}(v_i) \approx |v_{i-1} - 2v_i + v_{i+1}|^2 = (x_{i-1} - 2x_i + x_{i+1})^2 + (y_{i-1} - 2y_i + y_{i+1})^2 \quad (6)$$

$E_{image}(v_i)$ 는 에지 강도(edge intensity)와 같은 영상 에너지로 나타낸다.

$$E_{edge}(v_i) = -|\nabla I(v_i)| = -|\nabla I(x_i, y_i)|^2 \quad (7)$$

각 에너지 값들은 이웃한 포인트들 중 가장 큰 에너지 값으로 나누어 정규화 함으로써 0에서부터 1의 값을 갖는다.

여기서 영상 에너지인  $E_{image}(v_i)$ 는 에지의 강도가 높은 값을 갖는 객체의 경계선에서 스네이크 포인트의 움직임을 멈추게 하는 힘을 갖고 있다. 그러나 실제 이미지에서 객체의 경계가 아닌 주위의 강한 에지에 대해서도 영향을 받게 되는 문제가 있는데 스테레오 영상에서 이를 해결하기 위한 방법으로 disparity 정보를 이용하였다.

### 3. Disparity 측정

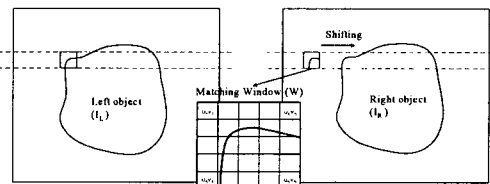
#### 3.1 블록매칭 (Block Matching)

스테레오 좌·우 영상을 이용하면 영상의 3차원 정보인 깊이(depth) 정보를 얻을 수 있고 이를 얻기 위한 전처리로 disparity 측정을 하게 된다. 이때 블록매칭과 같은 방법을 사용할 수 있는데 방법에는 SAD (Sum of Absolute Differences), ZSAD (Zero mean Sum of Absolute Differences), SSD (Sum of Squared Differences), ZSSD (Zero mean Sum of Squared Differences) 등이 사용된다. 여기에선 매칭의 정확도 및 계산 속도를 고려하여 ZSAD 블록매칭 방법을 사용하였다. ZSAD는 식 (8)과 같이 나타낼 수 있다.

$$ZSAD(v_i) = \quad (8)$$

$$\text{MIN}_{(dx, dy)} \left\{ \sum_{(u,v) \in W} |(I_L(u, v) - \bar{I}_L) - (I_R(u+dx, v+dy) - \bar{I}_R)| \right\}$$

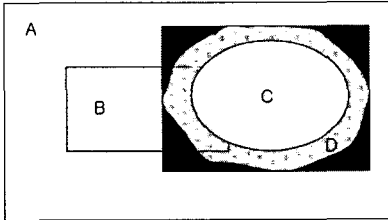
$I_L, I_R$ 는 각각 좌·우 이미지,  $W$ 는 매칭 윈도우,  $u, v$ 는 매칭윈도우 내의 좌표,  $\bar{I}_L, \bar{I}_R$ 은 이미지 평균값을 나타낸다. 블록매칭은 [그림 2]와 같이 좌·우 영상이 수평이라고 할 때 매칭 윈도우 사이즈만큼의 왼쪽 영상을 오른쪽 영상에서 수평으로 이동하며 상관성(correlation)이 가장 높은 곳에서의  $d_x$ 를 찾아 disparity를 예측할 수 있다. 단, 매칭윈도우를 수직으로는 이동하지 않기 때문에  $d_y = 0$ 이다.



[그림 2] 블록매칭을 이용한 disparity 측정

### 3.2 Disparity Map

스테레오 영상에서 측정된 disparity는 카메라에서 멀리 떨어진 객체에선 작게, 카메라와 가까운 객체에선 크게 나타나는데 이 차이를 레벨로 나누어 표현한 것이 disparity map(변이지도)이다. 여기서 disparity와 영상 객체간의 관계를 [그림 3]과 같이 정의하였다.

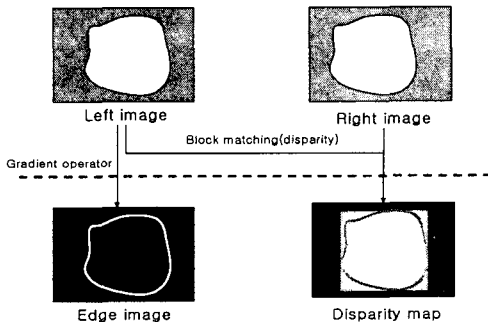


[그림 3] 배경 및 객체의 disparity 맵

A영역은 배경이며 disparity가 다른 영역에 비해 작고, B 영역은 분할하고자 하는 객체에 인접한 또 다른 객체이고 disparity는 배경보다는 크고 관심객체인 C보다는 작다. C 영역은 분할하고자 하는 관심 객체이며 disparity가 크다. D 영역은 관심객체인 C는 아니지만 disparity 측정 결과 C와 같은 영역으로 오측정된 구간이다. 오측정된 구간은 매칭 윈도우 사이즈에 따라 그 넓이가 다를 수 있다. 매칭 윈도우의 크기가 너무 작으면 광도변화를 탐지하지 못해 disparity 측정의 신뢰도가 떨어지고 너무 크게 되면 D 영역과 같이 disparity가 다른 지역까지 섞이게 된다.

### 4. 영상 에너지 함수

스테레오 영상에서 얻을 수 있는 특징 중 disparity 정보를 얻기 위한 블록매칭 방법을 스네이크 알고리즘의 영상 에너지에 사용하는 알고리즘을 소개한다.



[그림 4] 스테레오 영상정보

제안된 방법에서 사용되어지는 영상은 스테레오 좌·우 영상을 이용하여 얻을 수 있는 disparity 영상과 이진 에지영상을 사용한다. 그러나 이진 에지영상을 사용할 경우 경계선이 오목한 부분에서 에너지 부족현상이 일어나 경계선을 잘 찾지 못하기 때문에 이진 에지영상에 유클리드 거리 변환(Euclidean Distance Transform : EDT)<sup>[9]</sup>을 적용하였다. 유클리드 거리 변환은 그레이영상 또는 이진 에지영상의 강도(intensity)를 확산(diffusion)시켜 오목면에 대한 에너지 부족현상을 해결할 수 있다.

Disparity 정보를 스네이크 에너지 항에 적용하기 위해서 선택된 관심 객체의 disparity를 기준으로 정규화(normalize) 시켰다. 이는 에너지 함수를 최소화하는 관심객체의 disparity까지 포인트를 이동시킨다. 정규화 함수를 ND(Normalized Disparity)라고 할 때 다음과 같이 나타낼 수 있다.

$$ND(v_i) = (ZSAD(v_i) - interested\ disparity) / Max \quad (9)$$

*interested disparity*는 관심객체의 disparity이고 *Max*는 disparity map에서의 최대값이다.

식 (9)를 식 (3)에 적용하여 스네이크 알고리즘에서의 영상 에너지함수를 다음과 같이 표현하였다.

$$E_{image}(v_i) = \lambda \cdot ND(v_i) + (1 - \lambda) \cdot \gamma \cdot E_{edge}(v_i) \quad (10)$$

여기에서  $\lambda$ 는 상수로서 관심객체의 disparity 영역에서는 0으로 설정하여 영상의 에지 정보를 가지고 객체의 경계를 찾을 수 있게 하고, 그 외의 disparity 영역에서는 1로 설정하여 에지 정보는 사용하지 않고 정규화된 disparity 정보를 이용하여 관심객체의 disparity 영역까지 유도한다.

if  $(x_i, y_i) \in \{CUD\}$

$$\lambda = 0$$

else if  $(x_i, y_i) \in \{AUB\}$

$$\lambda = 1$$

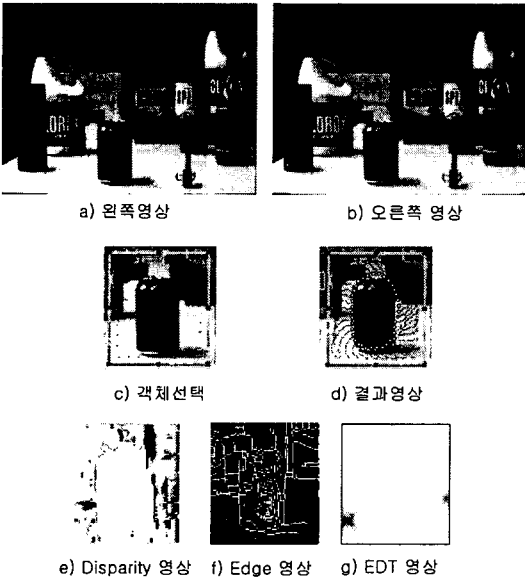
전체 스네이크 에너지함수 식은 다음과 같다.

$$E_{snake}(v_i) = \alpha \cdot E_{continuity}(v_i) + \beta \cdot E_{curvature}(v_i) + \lambda \cdot ND(v_i) + (1 - \lambda) \cdot \gamma \cdot E_{edge}(v_i) \quad (11)$$

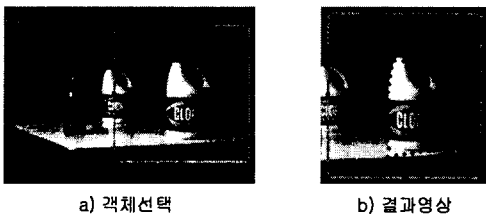
### 5. 실험 결과

본 논문에서 제안한 객체분할 방법은 윈도우 환경에서 Visual C++ 6.0을 이용하여 구현을 하였고, 사용된 영상은 320×240 크기의 스테레오 영상을 이용하였다. 초기 스네이크 포인터는 사용자 입력을 통해 관심객체를 포함한 사각형의 외곽선에 일정한 간

격으로 스네이크 포인트를 설정하였고, 블록매칭에 사용한 윈도우  $W$ 의 사이즈는 신뢰도를 고려해  $9 \times 9$  사이즈를 사용하였다. 또한, 관심객체의 disparity 측정은 전처리 과정에서 블록매칭 방법으로 구하였다. 스네이크 알고리즘에 적용한 매개변수는 실험에 의해  $\alpha=1.0$ ,  $\beta=0.6$  그리고  $\gamma=1.2$ 로 설정하였다. 실험 결과 영상은 [그림 5]에 보였다.



[그림 5] 실험 결과 I



[그림 6] 실험 결과 II

사용자는 [그림 5]의 c)와 같이 스테레오 영상에서 좌측영상에 대해 분할하고자하는 객체를 포함하는 사각형의 영역을 선택한다. 이때 사각형 안쪽에 균일한 간격으로 초기 스네이크 포인트들이 설정되고 알고리즘 실행 후 d)와 같은 결과영상을 얻을 수 있다. 영상 에너지에 사용되는 영상정보는 e), f) 그리고 이진 에지영상을 유클리드 거리 변환한 g) 영상이 쓰였다. [그림 6]은 다른 영상에 같은 실험을 한 결과를 보여주고 있다.

실험결과 관심객체의 주위에 있는 또 다른 객체 및 에지의 영향을 받지 않고 대체로 좋은 결과를 보였으나 스테레오 영상에서 발생하는 폐색(occlusion)

부분으로 인해 경계선을 벗어나거나 오 검출하는 경우가 있었다.

## 6. 결론

본 논문은 MPEG-4 기반의 스네이크 알고리즘을 사용하여 스테레오 영상에서 객체를 분할하는 방법을 제시하였다. 관심 객체 주위의 강한 에지 성분으로 인한 문제점을 개선하기 위해 disparity 정보를 스네이크 알고리즘에 적용하는 방법을 사용하였다. 이는 현장감을 주는 3차원 객체기반의 여러 응용분야에 적용할 수 있을 것이다. 그러나 객체 주위의 폐색부분에 대한 disparity 오측정을 후처리 해주는 부분과 블록매칭과 같은 방법에서 연산속도의 개선 및 신뢰성 높은 disparity 측정방법이 필요하겠다.

## 참고문헌

- [1] R. C. Gonzalez and R. E. Woods, Digital Image Processing, Addison Wesley Publishing company, 1992.
- [2] R. M. Haralick and L. G. Shapiro, "SURVEY: Image Segmentation Techniques," Computer Vision Graphics and Image Processing, vol. 29, 1985.
- [3] S. Beucher and C. Lantuejoul, "Use of Watershed in Contour Detection", Proceedings of the International Workshop on Image Processing, CCETT/IRISA, Rennes, France, 1979.
- [4] L. Vincent and P. Soille, "Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 13, pp. 583-598, 1991.
- [5] M. Kass, A. Witkin, and D. Terzopoulos, "Snake: Active Contour Models," Int'l J. Computer Vision, vol. 1, no. 4, pp. 321-331, 1987.
- [6] D. J. Williams and M. Shah, "A Fast Algorithm for Active Contours And Curvature Estimation," Computer Vision, Graphics, and Image Processing, vol. 55, pp. 14-26, 1992.
- [7] K. M. Lam, and H. Yan, "Fast Greedy Algorithm for Active Contours," Electron Lett., vol. 30, pp. 21-23, 1994.
- [8] A. Redert, E. Hendriks and J. Biemond, "Correspondence Estimation in Image Pairs," IEEE Signal Processing Mag., vol. 16, no. 3, pp. 29-46, May 1999.
- [9] L. D. Cohen and I. Cohen, "Finite-element methods for active contour models and balloons for 2-D and 3-D images," IEEE Trans. Pattern Anal. Machine Intell., vol. 15, pp. 1131-1147, Nov. 1993.