

레벨셋 방법을 이용한 비디오 객체 추출

이광연, 김성대

한국과학기술원 전자전산학과.

전화:042-869-8030 / 팩스: 042-869-8570

Video Object Extraction using Level Set Method

Kwang-Yeon Rhee, Seong-Dae Kim

EECS, Korea Advanced Institute of Science and Technology,

email : rheeky@sdvision.kaist.ac.kr

요약

비디오객체 추출 기법은 MPEG-4 및 MPEG-7의 응용을 목표로 최근 활발하게 연구되고 있다. 이들 연구는 객체 추출의 전체적인 구조와 정확한 윤곽선 검출 알고리즘의 개발에 초점을 맞추고 있으며 제한적인 조건하에서 만족할 만한 성능을 내고 있다. 그러나, 카메라 움직임, 객체의 빠른 움직임, 비강체 운동 등 보다 일반적인 상황에서는 객체 추출의 안정성이 떨어진다. 본 논문에서는 객체추출의 안정성을 높이기 위해 칼라, 움직임 정보 등의 특징정보(feature)가 균일한 영역으로 사전분할하고, 분할된 균일영역을 추적하는 알고리즘을 제안한다. 추적된 균일 영역간의 경계는 각 영역의 통계적 분포와 영역경계의 윤곽선으로 정의된 에너지를 레벨셋 방법으로 최소화함으로써 조정된다.

I. 서론

MPEG-4에서 비디오객체(video object)는 자연영상의 의미 있는 영역을 표현하는 시공간데이터(spatio-temporal data)로 정의된다. 이러한 비디오 객체는 MPEG-4의 영역기반 부호화(region-based coding)와 MPEG-7을 비롯한 내용기반검색(content-based retrieval) 시스템에서 기본단위로 취급된다. 이런 이유로 비디오객체 추출을 위한 연구는 최근 MPEG-4 및 MPEG-7의 응용을 목표로 활발하게 진행되고 있다.

최근의 비디오객체 추출 알고리즘[7]은 입력이 영상열(image sequence)이라는 특성에 착안해 주로 다음과 같은 3단계 과정으로 구성된다. 즉, (1)영상열의 첫 번째 프레임에서 추출하려는 영역 설정(object definition) (2)설정된 영역 추적(object tracking) (3)추적된 영역간의 경계 미세조정(region refinement)

영역설정에는 배경은 정지해 있고 추출하려는 영역이 움직이는 물체라는 가정이 있다면 움직임 해석(motion analysis)을 이용해서 자동화하는 것이 가능하나[8], 일반적인 경우 사용자의 개입에 의해 정의하는 것이 보통이다.[6]

영역추적은 설정된 영역 전체의 움직임이 균일(homogeneous)하다고 가정하면 움직임 파라미터 추정만으로 간단하게 해결될 수 있으나[6], 그러한 가정이 성립되지 않을 경우 영역을 움직임이 균일한 영역으로 사전 분할해서 추적하는 것이 타당하다[1][2].

영역추적 후 영역경계의 미세 조정은 active contour[9]을 이용하거나[2][3], watershed 등의 영역 확장 방법을 이용한다[1].

본 논문에서는 영역추적과 영역경계의 미세조정 문

제를 해결하는 알고리즘을 제안한다. 즉, 설정된 영역을 여러 개의 균일한 영역(homogeneous regions)으로 사전분할하고, 분할된 각각의 영역을 추적한다. 이 때, 각각의 균일 영역은 균일성의 지표(homogeneity measure)가 되는 특징정보가 정규분포를 따른다고 가정된다. 추적된 영역간의 경계조정은 각 균일 영역간의 통계적 분포와 영역경계의 윤곽선으로부터 정의된 에너지 함수를 최소화함으로써 수행된다. 에너지 함수의 최소화는 레벨셋 방법(level set method)[4]을 이용한다.

본 논문에서 제안된 알고리즘은 개념적으로 [1][2]와 유사하다. 그러나, 각 균일 영역에 대한 통계적 분포를 고려하고 이로부터 구성된 에너지 함수를 레벨셋 방법을 이용하여 해결한다는 점에서 기존의 방법에 비해 차이점을 갖는다. 한편, [5]에서도 레벨셋 방법을 적용하여 물체를 추적하고 있으나 본 논문과 달리 균일영역으로 사전분할하지 않고 설정된 영역전체를 추적하며 또한 추출하려는 영역이 정지된 배경 위에서 움직이는 물체라는 가정을 하고 있는 등 일반적인 비디오객체 추적에 적용하기에는 부적합하다.

II. 레벨셋 방법(level set method)

레벨셋 방법[4]은 active contour의 구현방법 중 하나라고 할 수 있으며, 유한개의 제어점(control point)에 기반한 구현방법[10]에서 발생하는 제어점 개수결정의 문제와 위상변화(topology change)의 문제를 자연스럽게 해결해주는 방법이다. 본절에서는 지면 관계상 자세한 수식은 생략하고 개념위주로 레벨셋 방법을 설명한다.

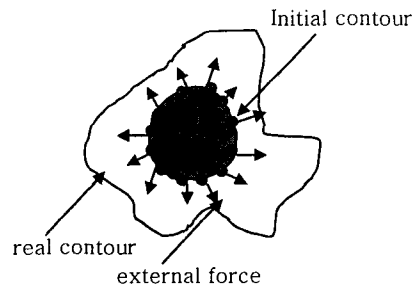


그림 1. 유한개의 제어점에 기반한 active contour 구현방법

그림 1을 보면 각각의 제어점에 영상으로부터 정의

된 힘(force)이 작용하고 있다. 초기에 주어진 제어점이 실제 윤곽선의 내부에 있으므로 각 제어점에는 바깥쪽 방향의 힘이 작용하게 된다. 이러한 힘은 보통 윤곽선의 smoothness와 영상의 공간방향 변화율(spatial gradient)로부터 정의된 에너지를 최소화하도록 계산된다. 이 때, 유한개의 제어점을 이용하는 경우 실제 윤곽선에 근사하려면 제어점의 개수가 윤곽선의 세부변화를 표현할 수 있도록 충분히 많아야 한다. 즉, 제어점의 개수가 많을수록 실제 윤곽선에 더욱 근사하게 된다. 그런데, 제어점의 개수가 증가하면 그림 2와 같은 문제가 발생한다.

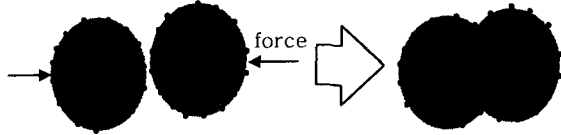


그림 2. 위상 변화(topology change)

그림 2에서는 2개의 떨어져 있던 원이 외부의 힘에 의해 서로 만나게 되는 경우를 보여준다. 이 때, 두 원이 만나는 부분에서의 제어점의 처리가 문제가 된다. 즉, 두 원이 만나는 부분의 제어점은 제거되어야 한다. 그러나 기존의 유한개의 제어점에 기반한 구현방법은 이 문제를 고려하고 있지 않다. 한편, 제어점의 개수가 증가할수록 그림 2와 같이 위상변화가 더욱 자주 발생하게 된다. 윤곽선 전개(curve evolution)에 따른 제어점 처리 방법의 어려움은 유한개의 제어점에 기반한 active contour 구현방법의 가장 큰 문제점 중에 하나이다.

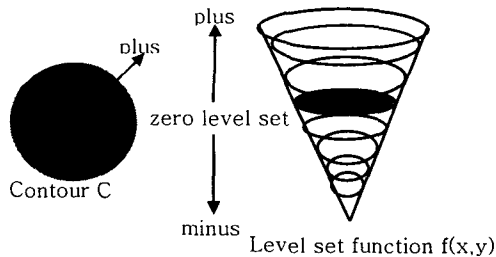


그림 3. 2차원 윤곽선과 level set간의 변환

레벨셋 방법은 위상변화의 문제를 해결하기 위해 그림 3과 같이, 2차원 윤곽선을 직접 다루지 않고 3차원의 레벨셋으로 변화해서 윤곽선의 변화를 처리한다. 레벨셋 함수 $f(x,y)$ 는 2차원의 윤곽선으로부터 식 1과 같이 정의된다.

$$f(x,y) = \begin{cases} +d(x,y,C) & \text{if } (x,y) \text{ is outside the contour } C \\ -d(x,y,C) & \text{if } (x,y) \text{ is inside the contour } C \end{cases}$$

where $d(x,y,C)$ is the distance from the contour C to (x,y)

식 1. level set function definition

즉, 윤곽선으로부터의 (x,y) 까지의 거리를 계산한 후, 윤곽선 내부이면 -, 외부이면 + 부호를 할당한다. 반대로 3차원 레벨셋 함수 $f(x,y)$ 로부터 2차원 윤곽선의 위

치는 $C = \{(x,y) | f(x,y) = 0\}$ 과 같이 정의된다.

그림 4는 레벨셋으로 변환하여 위상이 변화하는 상황에서의 처리를 보여준다. 3차원의 레벨셋을 다루게 되므로 직접 2차원 윤곽선을 다루는 경우처럼 제어점의 제거나 생성은 불필요하게 된다.

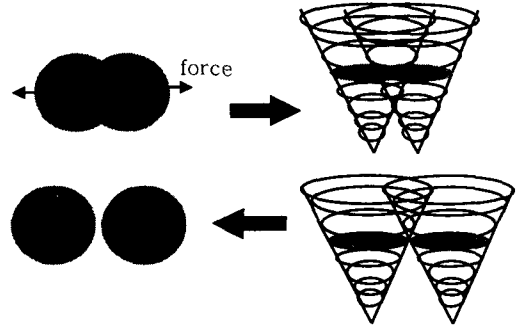


그림 4. level set 함수에 의한 위상 변화

한편 윤곽선에 작용하는 힘에 의한 레벨셋 함수 $f(x,y)$ 의 변화는 영상으로부터 정의된 에너지 함수를 최소화하도록 식 2와 같이 steepest descent rule에 의해 표현된다.

$$f_{new}(x,y) = f_{old}(x,y) - \Delta t \frac{df(x,y)}{dt}$$

식 2. level set function update eq.

III 제안된 알고리즘

제안된 알고리즘의 구성은 다음과 같다. 먼저 사용자 개입에 의해 주어진 영역을 watershed 알고리즘에 의해 사전 분할하고, 각 영역의 통계적 분포를 계산한다. 분할된 영역은 SSD(sum of squared difference)가 최소가 되도록 추적된다. 추적된 영역을 씨앗(seed)으로 하여 사전영역분할에서 계산된 각 균일영역의 통계적 분포와 윤곽선으로부터 정의된 에너지가 최소화되도록 최종 분할 영역의 경계선을 결정한다.

1. 영역 설정 및 사전 분할

추출하려는 영역 설정은 영상열의 첫번째 프레임에서 사용자 개입에 의해 비교적 정확하게 주어진다. 설정된 영역은 marker 기반의 watershed 알고리즘에 의해 사전분할(pre-segmentation)된다. 사전분할 알고리즘은 [11]과 유사하다. 즉, morphological filtering에 의해 영상을 단순화(smoothing)하고 이로부터 connected component labeling을 거쳐 marker를 찾는다. marker로부터 watershed 알고리즘에 의해 영역 확장하여 균일한 영역으로 분할한다. 분할된 영역은 수식 3과 같이 물체영역과 배경영역이 각각 N, M 개의 균일한 영역으로 표시된다.

분할된 각 영역은 특징정보(feature)에 대해 식 4와 같이 정규분포(Gaussian distribution)를 갖는다고 가정된다. 특징정보는 주어진 영상의 특성에 따라 칼라나 밝기값, 텍스처, 움직임 정보 등이 될 수 있다.

$$R = R_{obj} \cup R_{bak}, \quad R_{obj} \cap R_{bak} = \phi$$

$$R_{obj} = \bigcup_{i=1}^N R_i, \quad R_{bak} = \bigcup_{i=N+1}^{N+M} R_i,$$

$$R_i \cap R_j = \phi \quad \text{for } i \neq j$$

식 3. 영역분할

$$p(\mathbf{c}(x, y) | (x, y) \in R_i) \sim N(\boldsymbol{\mu}_i, \boldsymbol{\sigma}_i^2)$$

식 4. 각 영역의 특징정보 벡터 $\mathbf{c}(x, y)$ 는 정규분포를 갖는다고 가정된다.

2. 영역추적 및 씨앗(Seed)영역 검출

현재 프레임에서 사전 분할된 각 영역은 다음 프레임까지 영역기반 움직임 추정에 의해 추적된다. 추적된 영역은 다음 프레임에서 씨앗(seed) 영역이 되어 영역경계 조정을 위한 씨앗으로 사용된다. 이 때, 사전분할된 영역의 크기가 작으므로, 운동 모형은 2차원 병진운동 모형으로 충분하다. 물론, affine움직임 모형 등 복잡한 움직임 모형을 사용할 수도 있겠지만 실험결과 거의 차이가 없었다.

3. 영역경계 조정

3.1 에너지 정의

사전 분할된 각 영역의 특징정보(feature)의 확률값으로부터 전체 영역의 에너지를 정의한다. 본 논문에서는 [12]의 정의를 따른다. [12]에서는 MDL(minimum description length)를 기준으로 하여 식 5와 같이 다중밴드 영상에 적용 가능한 에너지 함수를 정의하고 있다. 에너지를 최소화하면 영역 R_i 의 특징정보 벡터 $\mathbf{c}(x, y)$ 의 확률을 최대로 하면서 동시에 영역의 경계 ∂R_i 의 길이를 최소로 하게 된다.

$$E(\{\partial R_i, \boldsymbol{\mu}_i, \boldsymbol{\sigma}_i^2\}) = \alpha \int \|\partial \hat{R}_i(p)\| d - \sum_{i=1}^{N+M} \sum_{(x,y) \in R_i} \log p(\mathbf{c}(x, y) | R_i)$$

식 5. MDL에 기반한 에너지 함수

3.2 에너지 최소화와 레벨셋 표현

에너지는 steepest descent 방법에 의해 최소화된다. 에너지를 최소화하기 위한, 영역경계 ∂R_i 위의 점 $v=(x, y)$ 의 시간에 대한 미분치는 Euler-Lagrange eq. 과 Green정리에 의해 식 6과 같이 유도된다[12].

κ_i 는 점 v 에서의 곡률(curvature)이고 \mathbf{n}_i 는 영역 R_i 의 바깥쪽 방향으로의 단위법선벡터(outward unit normal vector)이다. 식 6은 영역 R_i 와 R_j 가 인접해 있을 때 특징정보 벡터 $\mathbf{c}(x, y)$ 의 분포가 더욱 더 잘 어울리는 영역으로 점 v 가 이동함을 의미한다.

$$\frac{dv}{dt} = -\frac{\delta E}{\delta \mathbf{v}} = \left(-\alpha \kappa_{i(v)} + \log \left(\frac{p(\mathbf{c}(v) | R_i)}{p(\mathbf{c}(v) | R_j)} \right) \right) \mathbf{n}_{i(v)}$$

식 6. 영역경계의 시간에 대한 미분치

한편, 2차원 평면상의 영역경계의 변화가

$$\frac{dv}{dt} = \beta(x, y) \mathbf{n}_{i(v)}$$

이면 레벨셋 함수의 변화는 다음과 같다[4].

$$\frac{d\phi}{dt} = \beta(x, y) \|\nabla \phi\|$$

3.3 에너지 최소화에 의한 영역경계 조정

2절에서 영역추적으로 얻은 씨앗영역으로는 다음 프레임의 전체 영역을 표시할 수 없으며 씨앗영역끼리 서로 겹치기도 한다. 씨앗 영역은 식 7과 같이 표시된다.

$$R' = R'_{obj} \cup R'_{bak} \cup R'_\phi, \quad R'_{obj} \cap R'_{bak} \neq \phi$$

$$R'_{obj} = \bigcup_{i=1}^N R'_i, \quad R'_{bak} = \bigcup_{i=N+1}^{N+M} R'_i,$$

$$R'_i \cap R'_j \neq \phi \quad \text{for } i \neq j$$

식 7. 영역추적 후의 씨앗영역

R'_ϕ 는 전체 영역 R 에서 씨앗영역 R'_{obj}, R'_{bak} 을 제외한 영역이다. 영역경계조정은 3.1에서 정의된 에너지를 최소화하도록 식 6을 변형하여 사용한다. 즉, 식 8과 같이 영역경계상의 점 $v=(x, y)$ 가 어느 영역의 경계에 속하는지를 결정하여 그에 적합한 영역경계의 미분치를 결정한다.

$$\frac{dv}{dt} = \begin{cases} \left(-\alpha \kappa_i + \log \left(\frac{p(\mathbf{c}(x, y) | R_i)}{p(\mathbf{c}(x, y) | R_j)} \right) \right) \mathbf{n}_i & \text{if } R'_i \subset R'_{obj}, R'_j \subset R'_{bak} \\ \left(-\alpha \kappa_i + \log \left(\frac{p(\mathbf{c}(x, y) | R_i)}{\delta} \right) \right) \mathbf{n}_i & \text{if } R'_i \subset R'_{obj}, R'_j \subset R'_\phi \\ \left(-\alpha \kappa_i + \log \left(\frac{\delta}{p(\mathbf{c}(x, y) | R_j)} \right) \right) \mathbf{n}_i & \text{if } R'_i \subset R'_\phi, R'_j \subset R'_{bak} \\ 0 & \\ \left(-\alpha \kappa_i + \log \left(\frac{p(\mathbf{c}(x, y) | R_i)}{p(\mathbf{c}(x, y) | R_j)} \right) \right) \mathbf{n}_i & \text{if } R'_i, R'_j \subset R'_{obj} \text{ or } R'_i, R'_j \subset R'_{bak} \end{cases}$$

식 8. 씨앗영역경계의 시간에 대한 미분치

IV. 실험결과

MPEG영상열 중 빠른 비강체(non-rigid) 운동을 하는 stefan영상에 대해 제안된 알고리즘을 적용시켜 보았다. 특징 정보는 칼라와 움직임 추정 오차를 동시에 사용하였다. 최종 결과를 보면 경계선 부분이 다소 지저분하나 적절한 후처리를 하면 간단하게 해결될 수 있을 것이다. 어깨 부분이 제대로 검출되지 않았는데, 그 이유는 영역내의 칼라가 정규분포로 모델링할 수 없을 정도로 텍스처가 존재하기 때문이며 칼라의 평균치가 셔츠의 바탕색보다는 배경영역과 더 비슷하기 때문이다.

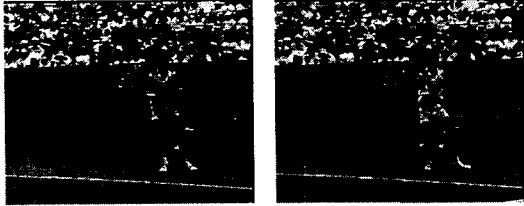
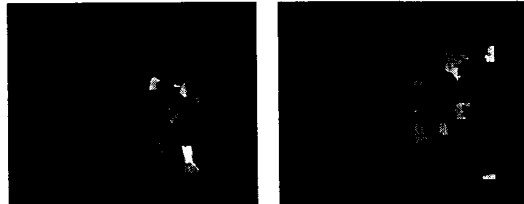


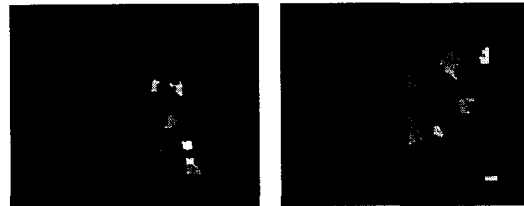
그림 5. (가)실험 영상 : Stefan 1st, 3th frame



(나) 좌: 1st frame, 사용자가 설정한 물체영역
우: 3th frame, 알고리즘 적용결과 추적해낸 영역



(다) 1st frame, Morphological 영역분할
좌: 물체영역, 우: 배경영역



(라) 영역추적으로 얻은 씨앗영역
좌: 물체영역, 우: 배경영역



(마) 레벨셀 표현에 의한 영역경계 조정 결과
좌: 물체영역, 우: 배경영역

이러한 문제는 텍스처 등 다른 특징정보를 추가로 사용하면 해결할 수 있을 것이다.

V. 결론 및 향후과제

사용자가 지정한 물체영역을 추적할 수 있는 알고리즘을 제안하였다. 빠른 비강체 운동에 적응할 수 있도록 영역을 미리 분할한 후 분할된 영역을 추적하여 씨앗영역으로 삼고, 다시 씨앗 영역들간의 경계를 조정하여 물체영역을 추출해내었다. 경계조정은 레벨셀 방법에 의해 에너지 함수의 최소화를 통해 구현했다. 앞으로는 보다 다양한 특징정보를 추가적으로 사용하여 검출의 안정성을 높일 계획이다.

참고문헌

- [1] F. Marques, J. Llach, "Tracking of generic objects for video object generation", International Conference on Image Processing, 1998. vol.3, pp 628-632
- [2] C. Toklu, A. M. Tekalp, A. T. Erdem, "Semi-automatic Video Layering using 2-D Mesh Tracking", Visual Communication and Image Processing, 1998, SPIE Vol. 3309, pp 1020-1031
- [3] S. Kruss, P. Kauff, "Fine Segmentation of image objects by means of active contour models using information derived from morphological transformations", Visual Communication and Image Processing, 1996, SPIE Vol. 2727, pp1164-1171
- [4] R. Mallad, J. A. Sethian, B. C. Vemuri, "Shape Modeling with Front Propagation: A Level Set Approach", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 17, No. 2, Feb. 1995, pp158-175
- [5] N. Paragios and R. Deriche, "Geodesic Active Contours and Level Sets for the Detection and Tracking of Moving Objects", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, pp. 1-15, Mar. 2000.
- [6] Chuang Gu, Ming-Chieh Lee, "Semiautomatic segmentation and tracking of semantic video objects", IEEE Transactions on Circuits and Systems for Video Technology, Vol: 8 5, Sept. 1998, pp 572-584
- [7] P.Salembier, F. Marques, "Region-based representations of image and video: segmentation tools for multimedia services", IEEE Transactions on Circuits and Systems for Video Technology, Vol: 9 8, Dec. 1999, pp 1147-1169
- [8] M Kim et al, "A VOP generation tool: automatic segmentation of moving objects in image sequences based on spatio-temporal information", IEEE Transactions on Circuits and Systems for Video Technology, Vol: 9 8, Dec. 1999, pp 1216-1226
- [9] M. Kass et al, "Snakes: Active Contour Models", International Journal of Computer Vision, 1988, pp321-331
- [10] D. Willams, M. Shah, "A Fast Algorithms for Active Contours and Curvature Estimation", CVGIP: Image Understanding, Vol 55, No 1, January, 1992 pp14~26
- [11] P.Salembier, M. Pardas, "Hierarchical Morphological Segmentation for Image Sequence Coding", IEEE Tran. Image Processing, Vol 3. No 5, Sep. 1994, pp 639-651
- [12] S.C. Zhu, A. Yuille, "Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 18, No 9, Sep. 1996, pp864-900
- [13] A. Yezzi et al, "A Geometric Snake Model for Segmentation of Medical Imagery", IEEE Trans. Medical Imaging, Vol 16, No 2, April 1997, pp 199-209