

색상 정보를 이용한 반자동 영상분할 기법

김민호*, 최재각**, 호요성*

*광주과학기술원 정보통신공학과, **경일대학교 제어계측공학과

Semi-Automatic Segmentation based on Color Information

Minho Kim*, Jae Gark Choi** and Yo-Sung Ho*

*Kwangju Institute of Science and Technology (K-JIST)

**Dept. of Control and Instrumentation Eng., Kyungil University

E-mail: mhkim@gogh.kjist.ac.kr

Abstract - This paper describes a new semi-automatic segmentation algorithm based on color information. Semi-automatic segmentation mainly consists of intra-frame segmentation and inter-frame segmentation. While intra-frame segmentation extracts video objects of interest from boundary information provided by the user and intensity information of the image, inter-frame segmentation partitions the image into the video objects and background by tracking the motion of video objects. For inter-frame segmentation, color information (Y, Cb and Cr) of the current frame can be used efficiently in order to find the exact boundary of the video objects. In this paper, we propose a new region growing algorithm which can maximize the ability of region differentiation, while preserving features of each color component.

용하는 기법들은 비디오 객체의 기하학적 특성을 제대로 다루지 못하지만, 형태학적 영상분할 기법은 기하학적 특성의 효율적인 이용이 가능하다. 형태학적 영상분할 기법은 형태학적 도구를 이용하여 영상을 단순화시킨 다음, Watershed 알고리즘으로 영상을 분할한다 [3, 4]. 그러나 형태학적 영상분할은 영상의 공간적 정보만을 이용하고 있을 뿐만 아니라, 분할된 영역 자체로는 하나의 비디오 객체를 나타내지 못하는 객체 대응의 문제점이 있다.

지금까지의 영상분할 기법들은 시간축 정보와 공간적 정보중에서 어느 하나만을 이용하므로 비디오 객체의 경계를 정확히 찾아내지 못하고, 심지어는 비디오 객체와 배경이 병합되는 결과를 얻을 수도 있다. 이러한 이유로 최근 시간축 정보와 공간적 정보를 함께 이용하여 좀 더 나은 분할 결과를 얻으려는 방법들이 시도되었다 [6]. 하지만 두 가지 정보 모두를 이용하는 방법에서조차도 이러한 문제를 완전히 해결하지 못했다. 왜냐하면 비디오 객체의 정의가 본질적으로 주관적인데다가, 자동 영상분할 기법은 단지 영상의 밝기 정보만을 이용하기 때문이다.

1. 서론

MPEG-4 동영상 부호화 표준은 영상을 객체 단위로 부호화한다. 따라서 객체기반 부호화를 위해서는 동영상 비디오의 각 화면을 VOP(Video Object Plane)의 형태로 표현해야 하고, VOP를 얻기 위해서는 영상분할이 이루어져야 한다. 비디오 객체를 분할하기 위해 몇몇 자동 영상분할 기술들이 개발되었는데, 가장 대표적인 방법들로서 변화검출(Change Detection) 기법, 움직임 예측 기반 영상분할 기법, 형태학적 영상분할 기법, 시공간 정보를 이용하는 영상분할 기법 등이 있다.

본 논문에서는 이러한 문제점들을 해결하기 위해 반자동 영상분할 기법을 이용하였다 [7]. 반자동 영상분할 기법은 첫 프레임에 대해서 사용자가 비디오 객체의 경계정보를 자동 영상분할 알고리즘에 제공하여 원하는 비디오 객체를 정확히 분할한다. 이때 사용자에 의해 제공된 정보를 효율적으로 이용하기 위해 이중 레이블 구조를 이용한다 [8]. 그 다음 프레임들에 대해서는 이전 프레임에서 분할된 비디오 객체의 움직임을 추적하여 영상분할을 반복한다 [7].

변화 검출 기법은 연속하는 두 영상의 차 영상을 이용한 통계적 가설검증 기법이다. 이 기법은 Global Thresholding 방법을 이용하고 있기 때문에 본질적으로 국부 정보를 이용하지 못한다. 이를 보완하기 위해 MRF(Markov Random Field)를 이용한 방법이 제안되었다 [1]. 하지만 이 방법은 드러난 배경(Uncovered Background)과 덮여질 배경(Background to be Covered)을 구별하지 못한다. 움직임 예측 기반 영상분할 기법은 화소 밝기의 시공간 미분값을 이용한 Mapping Parameter의 예측을 통하여 이러한 문제를 해결할 수 있다 [2].

비디오 객체의 움직임을 추적하여 영상분할을 수행하는 과정중 움직임 추적에서 오류가 발생할 수 있는데, 이러한 오류를 보정하기 위해 현재 프레임의 색상 정보(Y, Cb, Cr)를 이용하여 비디오 객체의 경계선을 조율한다. 경계 조율은 계층적 큐를 이용한 영역성장 알고리즘을 사용하는데, 기존에 제안된 방법[7, 8]에서는 색상 정보를 효율적으로 이용하지 못했다. 따라서 본 논문에서는 색상 정보를 효율적으로 이용하여 비디오 객체의 경계선을 정확하게 찾아내는 새로운 영역성장 알고리즘을 제안한다.

위와 같이 연속되는 두 영상의 차이값만을 이

본 논문에서 제안한 새로운 영역성장 알고리즘

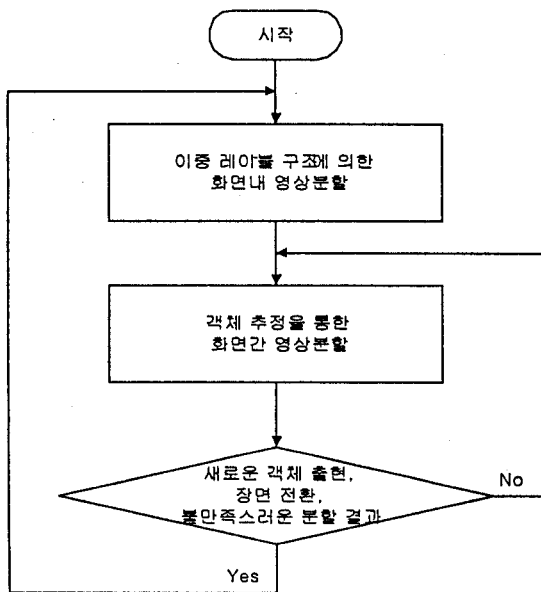
을 이용하는 반자동 영상분할 알고리즘은 기존의 알고리즘에서 색상 정보의 비효율적인 이용으로 인해 비디오 객체 사이의 경계가 무너지는 현상을 방지하여 더 좋은 분할 결과를 얻을 수 있다.

2. 반자동 영상분할 알고리즘

<그림 1>에서 보는 바와 같이 반자동 영상분할은 화면내 영상분할과 화면간 영상분할의 두 부분으로 이루어져 있다. 첫번째 프레임에 대해, 화면내 영상분할을 통해 관심있는 비디오 객체를 배경으로부터 분리한다. 이때 사용자의 조력을 통해 경계 정보가 화면내 영상분할 알고리즘에 제공된다. 다음으로 후속되는 프레임들에 대해서

- (a) 관심있는 새로운 비디오 객체 출현,
- (b) 장면 전환 (Scene Change),
- (c) 불만족스러운 분할 결과

가 발생하기 전까지는 화면간 영상분할을 반복한다. 만약 이러한 사건들이 발생한다면 화면내 영상분할을 다시 시작한다.



<그림 1> 반자동 영상분할 알고리즘

3. 화면내 영상분할

화면내 영상분할은 사용자의 개입을 통해 주어진 인지 정보와 영상에서 주어진 밝기 정보를 이용하여 영상을 분할한다. 화면내 영상분할은 Global Labeling, Local Labeling, 전체 레이블 마스크와 국부 레이블 마스크의 융합으로 구성된다.

우선 Global Labeling에서는 사용자에게 의한 인지 정보를 GUI (Graphic User Interface)와 같은 도구를 이용하여 알고리즘에 부가한다. 좀 더 구체적으로 설명하면, GUI 도구를 이용하여 우리가 찾고자 하는 비디오 객체의 실제 경계 주위에 불확정 영역을 설정하여 전체 레이블 마스크 (Globally Labeled Mask)를 얻는다. 이렇게 얻어진 전체 레이블 마스크는 배경, 비디오 객체, 불확정 영역으로 이루어진다. Local Labeling에서는 영상의 밝기 정보를 이용하여 국부 레이블 마스크 (Locally Labeled Mask)를 얻는다. 이때 형태학적 영상분할

기법을 이용하며, 이것은 일반적으로 영상 단순화 (Simplification) 과정과 Watershed 알고리즘으로 이루어져 있다.

형태학적 영상분할 기법은 영상을 비디오 객체나 배경과 같은 의미있는 객체와는 관계없이 단지 밝기 정보가 비슷한 여러 개의 작은 영역들로 나눈다. 따라서 Local Labeling에서 나오는 결과를 국부 레이블 마스크라 하며, 마지막으로 전체 레이블 마스크와 국부 레이블 마스크를 융합하여 최종 마스크를 얻는다 [8].

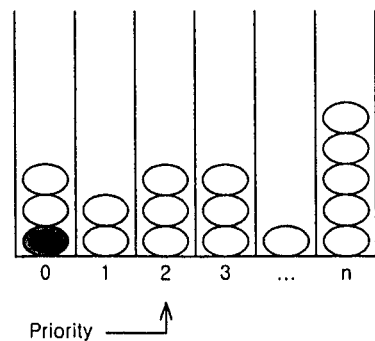
4. 화면간 영상분할

화면간 영상분할은 비디오 객체의 움직임을 추적하여 영상을 분할한다. 이를 위해 우선 비디오 객체 각각의 움직임을 예측한다. 이때 비디오 객체의 움직임을 나타내기 위해 Affine 6 파라미터 모델을 사용한다. 다음으로 예측된 움직임 변수를 이용하여 이전 프레임의 비디오 객체를 각각 현재의 프레임에 투영(Projection)시킨다. 하지만 움직임 예측시 잡음으로 인해 얻어진 변수가 정확하지 않을 수 있고, 비디오 객체를 투영할 때 다중으로 맵핑이 되거나 전혀 맵핑이 되지 않는 영역들이 발생할 수도 있다. 이러한 오류를 보정하기 위해 비디오 객체의 투영의 다음 단계로써 경계 조율을 행한다.

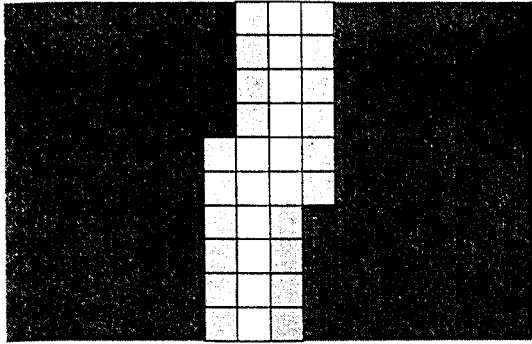
경계 조율에서는 우선 앞에서 언급된 다중으로 맵핑되거나 전혀 맵핑이 되지 않는 영역을 불확정 영역으로 설정한다. 그리고 정확하지 않은 움직임 변수로 인해 각 비디오 객체들이 만나는 경계 부분은 오류를 내포할 수 있기 때문에, 이 영역 또한 불확정 영역으로 설정한다. 이 불확정 영역은 영역 결정이 유보된 화소들의 집합을 의미하게 된다. 이렇게 설정된 불확정 영역의 화소들을 계층적 큐 (Queue)를 이용한 영역성장 알고리즘을 이용하여 특성이 비슷한 비디오 객체에 할당을 시킨다. 이 과정을 통해 비디오 객체의 정확한 경계를 찾을 수 있게 된다 [7].

5. 계층적 큐를 이용한 영역성장

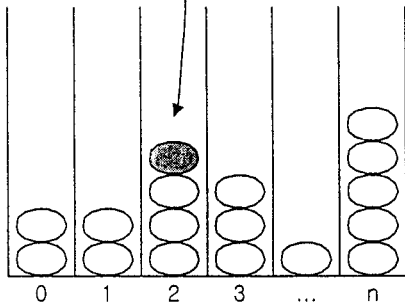
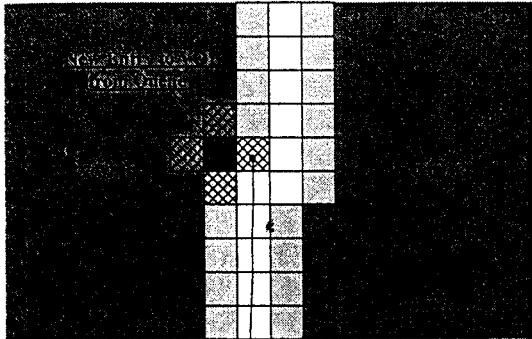
본 논문에서 사용된 영역성장 알고리즘은 계층적 큐를 이용하여 불확정 영역의 화소들을 해당 비디오 객체에 할당시킨다. 이때 계층적 큐는 화소의 인접성(proximity) 및 특성의 유사성(similarity)을 동시에 이용하여 불확정 영역의 화소들을 해당 비디오 객체에 할당하는 것을 가능하도록 한다. <그림 2>에서 계층적 큐를 보여주고 있다.



<그림 2> 계층적 큐



<그림 3> 영역성장 알고리즘의 초기화



Priority ↑

<그림 4> 영역의 성장과정

영역성장 알고리즘은 초기화 과정과 범람과정의 2 단계로 이루어진다. 초기화 과정에서는 <그림 3>와 같이 비디오 객체에 인접한 화소들을 유사성 정도에 따라 계층적 큐중 하나의 큐에 넣는다.

범람 과정은 화소를 포함하고 있는 큐 중에서 우선 순위가 가장 높은 큐에서 화소 (<그림 2>에서 회색 화소)를 추출하여, 특성이 가장 비슷한, 즉, 아래에서 정의된 AbsDiff가 가장 작은 비디오 객체에 그 화소를 할당시킨다. 다음으로 현재 처리된 화소의 인접 화소중에서 아직 큐에 들어가지 않은 불확정 영역의 화소를 큐에 집어넣는다. 이 과정을 <그림 4>에서 보여주고 있다.

우선 순위가 높은 큐에 있는 화소들이 먼저 영역 결정에 이용되기 때문에, 비디오 객체에 인접한 화소를 계층적 큐에 넣을 때 사용되는 우선 순위 (Priority)를 [7]에서는 다음과 같이 정의하였다.

<식 1>
 $Priority = \min(AbsDiff(n))$
 $AbsDiff(n) = |CurrAvg - LAvg(n)|$

$CurrAvg = \alpha * Crnt_Y + \beta * Crnt_Cb + \gamma * Crnt_Cr$
 $LAvg(n) = \alpha * LAvg_Y(n) + \beta * LAvg_Cb(n) + \gamma * LAvg_Cr(n)$
 $\alpha, \beta, \gamma: \text{Weighting Factor, } \alpha + \beta + \gamma = 1$

하지만 위의 방법은 단순히 평균만을 취함으로써 색상 정보 개개의 특성을 살리지 못하고 있다. 이러한 문제점은 실험 결과에서도 알 수 있다.

본 논문에서는 색상 정보의 개개의 특성을 살려 영역 사이의 분별 능력을 강화할 수 있는 다음과 같은 계층값을 제안한다.

<식 2>
 $AbsDiff(n) = \max(|Diff_Y(n)|, |Diff_Cb(n)|, |Diff_Cr(n)|)$
 $Diff_Y(n) = Crnt_Y - LAvg_Y(n)$
 $Diff_Cb(n) = Crnt_Cb - LAvg_Cb(n)$
 $Diff_Cr(n) = Crnt_Cr - LAvg_Cr(n)$

<식 2>의 유사성 척도는 우선 색상 정보의 개개의 특성을 살릴 수 있으며, 색상 정보 각각의 차이값의 최대값을 취함으로써 영역 사이의 차별 능력도 강화할 수 있다. 다시 말해서, <식 2>의 척도는 밝기성분(Y)의 차이가 작지만 색상 성분(Cb 또는 Cr)의 차이가 명확한 곳이나, 색상 성분의 차이는 작으나 밝기 성분의 차이가 명확한 곳도 정확히 구별할 수 있다. 비디오 객체에 인접하는 불확정 영역의 화소를 계층적 큐에 넣을 때의 Priority와 큐에서 추출한 화소를 영역에 할당할 때 이 값을 사용함으로써 비디오 객체의 정확한 경계를 찾을 수 있다.

6. 실험 결과

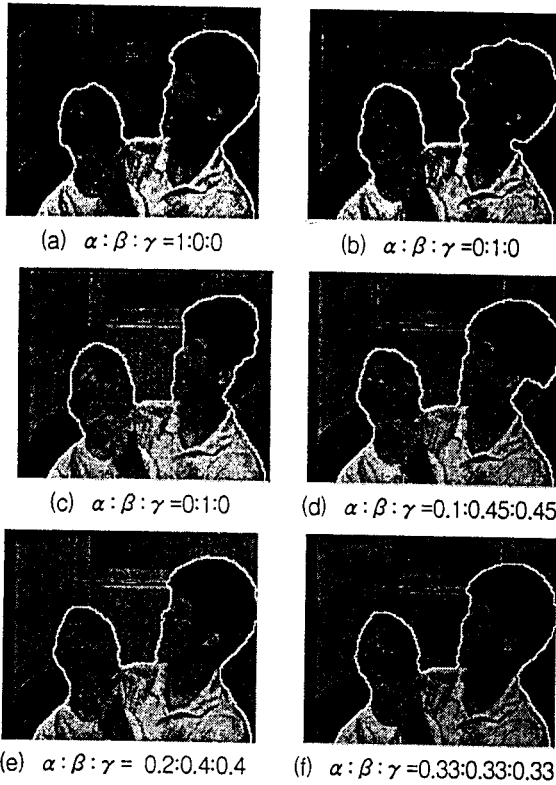
본 논문에서 제안된 계층값을 이용한 반자동 영상분할 기법의 성능을 알아보기 위해 176x144 크기의 QCIF 형태의 Mother & Daughter 영상을 사용하여 실험하였다. <그림 5>는 실험에서 사용된 Mother & Daughter의 원영상을 보여주고 있다.



<그림 5> 원영상

제안된 계층값의 성능을 알아보기 위해서 우선 반자동 영상분할 알고리즘의 첫 번째 단계인 화면 내 영상분할 알고리즘을 통해 원하는 비디오 객체를 정의한다. 앞에서 언급했듯이, 일단 비디오 객체가 정의된 후에는 화면간 영상분할을 반복하게 된다.

<그림 6>는 <식 1>의 성능을 알아보기 위해 α, β, γ 를 변화시켜가며 얻은 실험 결과이다. 우선 Y만을 이용한 경우, 아이의 왼쪽 얼굴과 머리 윗부분의 경계가 무너진 것을 볼 수 있다. 이에 반해 Cb나 Cr만을 사용한 경우, 아이의 얼굴 부분의



<그림 6> <식 1>을 이용한 결과



<그림 7> <식 2>을 이용한 결과

경계가 정확히 보존된 반면, 엄마의 얼굴 부분의 경계가 무너진 것을 알 수 있다.

또한 Cr나 Cr만 사용한 경우에서 α 의 값을 증가시켜 가며 실험을 한 결과, 엄마의 얼굴 부분의 경계가 점점 더 잘 보존되어가는 반면, 아이의 얼굴 부분의 경계가 점점 무너지는 결과를 보여주었다. 즉 α 를 증가시켜감에 따라 Y만 사용한 결과와 비슷해짐을 알 수 있다.

이 결과에서 알 수 있듯이, 각각의 성분만을 따로 사용하는 경우, 사용된 정보의 제한성 때문에 어머니나 아이의 얼굴부분의 경계 부분이 무너지는 결과를 초래했으며, 평균을 이용하는 경우 각 성분의 특성이 무시되는 것을 알 수 있다.

<그림 7>은 <식 2>를 사용하여 얻은 결과이다. <그림 7>에서 볼 수 있듯이, <식 1>을 이용한 결과와 달리 아이와 엄마의 얼굴 부분의 경계가 모두 보존되어 있음을 알 수 있다. 이것은 우선 각 성분 개개의 차이를 구한 후에 최대값을 선택했기 때문에 각 성분이 가지고 있는 영역 분별 능력을 보존할 뿐만 아니라, 극대화할 수 있다.

7. 결론

본 논문에서는 반자동 영상분할 알고리즘을 위해 색상 정보를 효율적으로 이용할 수 있는 영역성장 알고리즘을 제안하였다. 반자동 영상분할 알고리즘은 화면내 영상분할과 화면간 영상분할로 이루어진다. 화면내 영상분할에서는 사용자에게 의해 제공되는 경계 정보를 이용하여 원하는 비디오 객체를 정확히 분할하고, 화면간 영상분할에서는 이전 프레임의 분할결과와 움직임을 추적하여 영상을 분할한다. 움직임 추적 과정에서 오류가 발생할 수 있는데, 이 오류를 보정하기 위해 경계 조율을 수행한다. 경계 조율은 계층적 큐를 이용하는 영역성장 알고리즘을 이용한다. 본 논문에서는 각 색상 성분의 분별 능력을 보존하고 강화시킬 수 있는 계층값을 제안하였고, 실험을 통해 그 우수성을 확인하였다. 이와 같은 색상 정보를 효율적으로 이용하는 반자동 영상분할 알고리즘은 MPEG-4와 같은 객체 기반의 영상처리 분야에서 유용하게 이용될 수 있을 것이다.

8. 감사의 글

본 연구는 광주과학기술원 초고속광네트워크연구센터를 통한 한국과학재단 우수연구센터 지원금에 의한 것입니다.

참고 문헌

- [1] T. Aach and A. Kaup, "Statistical Model-based Change Detection in Moving Video", *Signal Processing*, 31, pp. 165-180, 1993.
- [2] M. Hötter and R. Thoma, "Image Segmentation Based on Object Oriented Mapping Parameter Estimation", *Signal Processing*, 15, pp. 315-334, 1988.
- [3] F. Meyer and S. Beucher, "Morphological Segmentation", *J. Visual Communication and Image Representation*, vol. 1, no. 1, pp. 21-46, September 1990.
- [4] P. Salembier, "Morphological Multiscale Segmentation for Image Coding", *Signal Processing*, 38, pp. 359-386, 1994.
- [5] 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, no. 6, June 1991.
- [6] J. G. Choi, S. W. Lee, and S. D. Kim, "Spatio-Temporal Video Segmentation Using a Joint Similarity Measure", *IEEE Trans. Circuits and Systems for Video Technology*, vol. 7, no. 2, April 1997.
- [7] J.G. Choi and et al., "A User-Assisted Segmentation Method for Video Object Plane Generation", *ITC-CSCC '98*, pp. 7-10, July 1998.
- [8] 김민호, 김대회, 호요성, "동영상 부호화를 위한 반자동 영상분할 기법", 제 11회 신호처리 합동학술대회, pp. 99-102, 1998.10.