

에지 및 적응적 임계값을 이용한 움직이는 물체 및 정적 물체의 분할

손재식, 김주영, 이승익, 김덕규
경북대학교 전자공학과

Moving and Non-Moving Objects Segmentation Using Edge and Adaptive Thresholding

Jae-Sik Sohn, Ju-Young Kim, Seung-Ik Lee, Duk-Gyoo Kim
School of Electronic and Electrical Engineering
Kyungpook National University
E-mail : sjs25@palgong.knu.ac.kr

Abstract

움직이는 물체의 자동 분할은 컴퓨터 비전의 여러 응용분야에서 중요한 문제로 대두되고 있다. 본 논문에서는 감시 시스템에서 에지와 적응적 임계값을 이용한 효과적인 자동 움직임 분할 방법을 제안하였다. 먼저 연속 영상에서 현재 영상과 배경 영상과의 차를 얻어서 그 히스토그램을 만든다. 이 때, 앞에서 얻은 히스토그램은 영상 잡음의 평균이 0 인 가우시안 분포를 가진다고 가정한다. 그리고, 이 히스토그램을 이용하여 영상 잡음의 분산을 찾는다. 이 분산값을 이용하여 적응적 임계값과 움직임 영역창을 결정한다. 적응적 임계값에 의한 결과 영상에서 움직이는 물체를 분할하기 위해 본 논문에서는 움직임 영역창을 이용하는 방법을 제안하였다. 이 움직임 영역창에 의해 더욱 효과적인 움직임 분할이 이루어진다. 또, 잡음의 제거를 위해 수학적 모폴로지(mathematical morphology)와 화소의 연결성이 이용된다.

I. 서론

컴퓨터 사용의 증가와 하드웨어 성능의 향상으로 컴퓨터 비전 기술을 실생활에 응용하려는 시도가 여러 분야에서 활발히 진행되고 있다. 특히 비디오 영상에서

물체의 움직임 탐지 및 분할은 침입자 감시 시스템, 교통 감시 시스템 등 비디오 기반의 응용프로그램에서 중요한 역할을 하고 있다. 인간의 시각 시스템이 쉽게 움직이는 물체를 구별할 수 있음에도 불구하고, 이제까지 연구되어 온 방법들은 쉽게 움직임을 구별할 수 없었다. 이 때문에 자동 비디오 분할은 영상 처리 분야에서 가장 관심 있는 문제중의 하나로 알려져 있다. 따라서, 영상 처리와 컴퓨터 비전 분야에 종사하는 많은 연구원들과 엔지니어들이 이 문제에 관심을 기울이고 있다.

비디오 감시 시스템에서 가장 중요한 요소들 중의 하나가 연속 영상에서 움직이는 물체를 분할하고, 그 움직임을 예측하는 움직임 해석 부분이다. 게다가, 정확한 움직임의 예측은 움직이는 물체의 정확한 분할을 필요로 한다. 수년동안 움직이는 물체의 분할을 위한 하나의 방법으로 배경 영상이 이용되었다. 이 방법은 변화가 없는 배경 영상을 먼저 획득한 후, 그 배경 영상과 현재 다른 화소들을 움직이는 물체로 간주하여 분할하는 방법이다. 이 방법은 비디오 감시 시스템, 고속도로 감시 시스템 등에 사용 되어지고 있다. 또한 연속 영상에서의 움직임 분할을 위한 접근으로 다른 몇몇 방법들이 제시되어 왔다. 현저한 움직임 해석에 의한 분할[1], 시간적 집성을 하는 변화 감지 마스크 의한 분할[2], 화소 기반의 움직임 분할[3] 등이 있으며 그 외에 다른 여러 가지가 제시되어 왔다. 그러나, 이

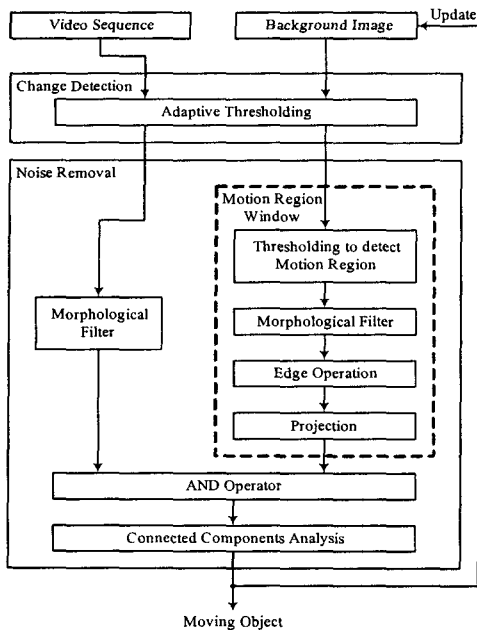


그림 1. 제안된 방법의 블록 선도

작업들은 다중 움직임 물체들, 불충분한 조명 상태, 빠른 변화, 다중 교합, 복잡한 환경 등의 문제점들을 가지고 있다[4]. 본 논문에서는 위에서 언급된 문제들 중에서 다중 움직임 물체에 대한 문제의 해결과 함께 본 논문에서 제안된 움직임 영역창을 이용하여 효과적으로 움직임 물체의 분할을 제안한다.

II. 움직이는 물체의 분할 과정

움직이는 물체를 분할하는 목적은 연속하는 영상 중 현재 프레임에서 움직임이 있는 부분과 정지해 있는 부분을 분할하는 것이다. 앞서서도 언급했듯이, 움직이는 물체의 분할은 배경 영상과 연속 영상들 사이의 차를 감지하는 기술과 결합될 수 있다. 본 논문은 이 차를 이용하여 움직이는 물체를 분할한다.

2.1 변화 감지

배경 영상에서 변화를 감지하는 것은 연속 영상에서 움직이는 물체의 예측을 제공한다. 이 변화를 감지하기 위해, 본 논문은 연속 영상에서의 현재 프레임과 배경 영상간의 차를 취한다. 즉, 이 차가 0 인 부분은 배경 영역이 되고 차가 0 보다 크거나 작은 부분은 변

화가 있는 영역으로 간주한다.

아래 (1)식은 변화 감지 마스크(CDM)를 나타낸 것이다.

$$CDM(x, y) = \begin{cases} 1, & \text{if } |I_x(x, y) - B(x, y)| > T \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

여기서, $I_x(x, y)$ 와 $B(x, y)$ 는 각각 그레이 스케일(gray scale)을 화소값으로 갖는 연속 영상에서의 현재 프레임과 배경 영상을 나타낸다. T 는 임계값이다.

여기서 문제는 잡음이나 움직이는 물체가 아닌 인자들에 의해 현재 프레임과 배경 영상간의 강도 분산(intensity variation)이 발생한다는 것이다. 이 문제의 해결 방법은 잡음에 의한 강도 분산을 없애기 위해 임계값을 적절히 결정하는 것이다. 본 논문은 적응적 임계값을 사용하여 이 문제를 해결하였다.

2.1.1 적응적 임계값

적응적 임계값은 Habili 등[5]이 제시한 방법을 이용한다. 이 방법은 적당한 오경보 확률(false alarm probability)에 기반한다. 본 절에서는 임계값을 결정하는 방법을 알아본다.

$p(d|B)$ 와 $p(d|F)$ 는 현재 프레임과 배경 영상간의 차를 나타내는 히스토그램에서 각각 배경과 움직이는 물체들의 상태 확률 밀도 함수를 나타낸다. 오경보 확률과 누락경보 확률(miss alarm probability)은 다음과 같다.

$$P_{false} = \sum_{d \in D_F} P(d|B) \quad (2)$$

$$P_{miss} = \sum_{d \in D_B} P(d|F) \quad (3)$$

P_{false} 는 히스토그램에서 그 값이 배경일 때 움직이는 물체로 오인될 확률이고, P_{miss} 은 그 값이 움직이는 물체일 때 배경으로 간주될 확률이다. 이상적으로 P_{miss} 와 P_{false} 이 최소값을 가져야 된다. 그러나, 이 확률들은 상호 대립적인 객체들이다.

실험적인 결과들을 보면, 변화 감지 마스크에서 소금과 후추 잡음(salt and pepper noise)의 양이 최소가 될 때 견고한 움직임 분할 마스크가 얻어진다. 그러므로, P_{false} 가 최소가 되거나, 또는 배경을 감지하는 확률이 최대가 되면 변화 감지 마스크에서 소금과 후추 잡음이 최소가 된다. 이럴 경우 오직 배경 모집단(background population)만으로 적당한 임계값을 찾을 수 있게 된다. 우리는 히스토그램에서 배경의

값들이 가우시안 분포를 가지며 평균이 0 이라고 가정한다. 따라서 $p(d|B)$ 는 다음과 같이 표현될 수 있다.

$$p(d|B) = \frac{1}{2\pi\sigma_B} e^{-\frac{1}{2}\left(\frac{d}{\sigma_B}\right)^2} \quad (4)$$

식 (4)에서 d 는 히스토그램의 값을 의미한다.

식 (5)는 어떤 임의의 값 $-T$ 와 T 사이를 샘플링해서 나타난 결과 집합을 추정 매개 변수로 나타낸 것이다.

$$\hat{P}(d|B, T) = \frac{1}{\sqrt{2\pi}\hat{\sigma}_B(T)} e^{-\frac{1}{2}\left(\frac{d}{\hat{\sigma}_B(T)}\right)^2} \quad (5)$$

그리고, 이 때 주어진 T 값에 대해 평균 절대 오차 (MAE)를 구하면 식(6)과 같이 된다.

$$MAE(T) = \frac{1}{n} \sum_d |\hat{P}_B(T) \cdot (\hat{p}(d|B, T) - h(d))| \quad (6)$$

(6)식에서 n 은 현재 프레임과 배경 영상간의 차를 나타내는 히스토그램에서 차값들의 총개수를 나타낸다. 이때, 평균 절대 오차가 가장 낮을 때의 T 값이 히스토그램 값과 가우시안 형태가 가장 일치하는 것으로 보고 가우시안 잡음 분산의 추정을 할 수 있다. $h(d)$ 는 히스토그램을 나타낸다. P_{false} 는 다시 Q 함수를 이용하여 계산될 수 있는데 Q 함수는 식(7)과 같이 정의될 수 있다.

$$Q(z) = \frac{1}{2\pi} \int_z^{\infty} e^{-\frac{1}{2}u^2} du \quad (7)$$

$Q(z)$ 는 분산이 1 이고 평균이 0 인 가우시안 밀도 함수의 꼬리에 해당하는 부분의 확률이다. 그러므로 P_{false} 는 다시 식(8)로 정의 될 수 있다.

$$P_{false} = 2Q\left(\frac{\tau}{\sigma_B}\right) \quad (8)$$

표 I 에서는 P_{false} 의 예와 임계값을 보여주고 있다. 본 논문에서는 실험 결과를 토대로 $3\sigma_B$ 를 임계값으로 정하였다. 본 논문에서는 이 임계값을 적응적 임계값이라 한다.

표 I. 임계값 τ 에 따른 P_{false} 값의 예

τ	P_{false}
σ_B	0.32
$2\sigma_B$	0.045
$3\sigma_B$	0.003

2.2 움직임 영역창과 잡음 제거

본 절에서는 움직임 영역창을 이용하여 잡음 제거 및 움직임 분할을 효과적으로 하는 과정을 설명한다. 이 움직임 영역창을 결정하기 위해 앞 절에서 구한 σ_B 값을 이용한다. $n\sigma_B$ 는 n 의 값이 증가할수록 P_{false} 의 값은 감소한다. 그러나, 반대로 n 이 증가하면 P_{miss} 의 값은 증가하게 된다. 이론적으로 P_{false} 와 P_{miss} 가 최소가 되는 것이 이상적이지만, P_{false} 가 작아지면 작아질수록 영상 잡음이 줄어들어서 움직임 영역창을 찾기 쉬워진다. 이것을 이용하여 $(n-1)\sigma_B$ 의 값일 때의 결과 영상과 $n\sigma_B$ 의 값을 가질 때의 결과 영상간의 차이 개수가 임계값보다 작아질 때 즉, $(n-1)\sigma_B$ 에 의한 결과 영상과 $n\sigma_B$ 에 의한 결과 영상이 변화가 거의 없을 때, $n\sigma_B$ 의 값에 의해 움직임 영역창이 구해진다. $n\sigma_B$ 의 값이 적용된 결과 영상에서 에지 성분을 추출하여 수평과 수직 축으로 투영한 후 그 투영된 값이 있는 존재하는 부분을 이루는 창을 만든다. 이를 움직임 영역창이라 한다. 이 움직임 영역창을 앞 절에서 얻은 적응적 임계값에 의한 결과 영상에 적용시켜서 효과적인 움직임 물체의 분할이 이루어진다. 움직임 영역창 안의 잡음은 모폴로지 필터인 침식과 팽창 필터를 사용하여 제거되고 제거되지 않은 나머지 잡음은 다시 화소의 연결성을 보고 연결성이 작은 물체를 잡음으로 간주하여 제거한다.

III. 실험 결과

본 논문에서 제안한 자동 움직임 분할 방법은 비디오 감시 시스템에서의 움직임 분할을 위한 방법이다. 실험에 사용된 연속 영상은 휘도 성분을 갖는 CIF

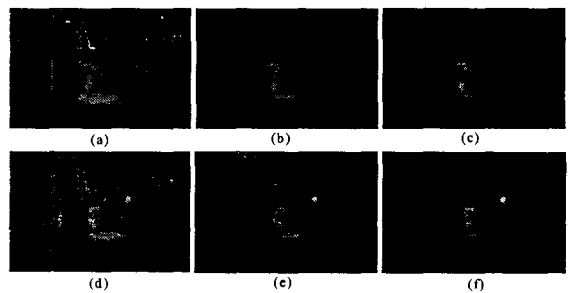


그림 2. $n\sigma_B$ 의 값에 의한 결과 영상의 예 (a)와 (d)는 $n = 3$ 일 때, (b)와 (e)는 $n = 8$, (c)와 (f)는 $n = 14$ 일 때의 결과 영상

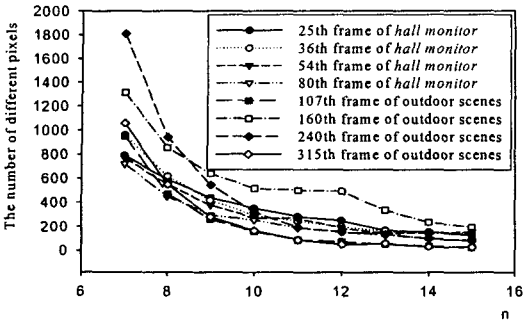


그림 3. $(n-1)\sigma_B$ 에 의한 결과 영상과 $n\sigma_B$ 에 의한 결과 영상간의 차이의 개수를 나타낸 예

형식의 영상들이다. 그림 2에서는 $n\sigma_B$ 에서 n 이 증가할수록 임계값이 높아지면서 잡음이 감소함을 보여준다. 이를 이용하여 움직이는 물체의 대략적인 위치인 움직임 영역창을 얻었다.

그림 3에서는 n 이 증가함에 따라 $(n-1)\sigma_B$ 과 $n\sigma_B$ 의 값이 적용된 결과 영상간 화소의 차이개수를 나타내고 있다. n 이 증가함에 따라 차이의 개수가 감소하는 것을 알 수 있다. 그리고, 감소하는 정도가 n 이 증가할수록 작아졌다. 이는 n 이 어떤 값 이상일 때, 잡음이 제거되어서 임계값에 의한 결과 영상이 영향을 거의 받지 않게 됨을 보여준다. 그림 4에서는 제안된 방법에 의해 자동 분할된 움직임 물체의 영상을 보여 주고 있다.

IV. 결론

본 논문에서는 감시 시스템에서 자동 움직임 분할하기 위한 방법을 제안하였다. 이를 위해 움직임 영역창을 적용시킴으로써 적응적 임계값만을 이용하는 움직임 분할보다 더욱 효과적으로 움직임 분할이 되는 방법을 제안하였다. 실험 결과 제안된 방법은 움직이는 물체가 비교적 잘 분할되는 것을 알 수 있었다. 향후 과제로서 더 정확한 분리와 다중 교합의 문제 등을 해결할 수 있는 방법을 제안할 것이다.

참고 문헌

[1] Irani M. and Peleg S., "Motion analysis for image enhancement: resolution, occlusion, and transparency," *Int. J. Visual Comm. Image Rep.* Vol.4 Issue.4,

pp.324-335, 1993.

[2] Stauffer C. and Grimson W. E. L., "Adaptive Background Mixture Models for Real-Time Tracking," in *Proceedings. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*. IEEE Comput. Soc. Part, Vol.2, 1999.

[3] Foresti, G. L., "A Real-Time System for Video Surveillance of unattended outdoor environments," *IEEE Trans. Circuits and Systems for Video Tech.* Vol.8 Issue.6, pp.697-704, Oct. 1998.

[4] Badenas J., Bober M., and Pla F., "Segmenting traffic scenes from gray level and motion information," *Pattern Anal. Appl.* 4, pp.28-38, 2001.

[5] N. Habili and A. R. Moini, "Automatic Temporal Segmentation for Content-Based Video Coding," *Tech. Report, Department of Electrical & Electronic Eng., Adelaide University*, November 2000.

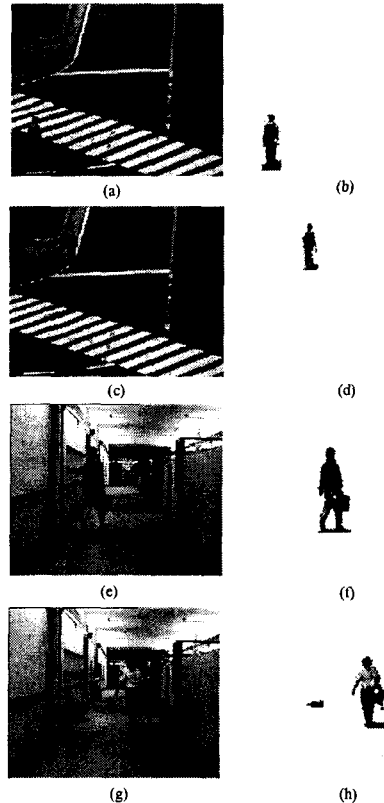


그림 4. 제안된 방법에 의한 움직임 물체 분할의 예