

GrabCut 을 이용한 배경 분리 알고리즘의 정확도 개선

이상훈, *김기백, **조남익
 서울대학교, *숭실대학교, **서울대학교
 tkd1088@ispl.snu.ac.kr, *imkgb27@ssu.ac.kr, **nicho@snu.ac.kr

Improvement of Background Subtraction Algorithm using GrabCut

Sang-Hoon Lee *Gibak Kim, **Nam Ik Cho
 Seoul National University, *Soongsil University, **Seoul National University

요 약

본 논문에서는 기존의 배경 분리 알고리즘 결과에 GrabCut 알고리즘을 도입하여 보다 정확한 배경 분리를 수행하고자 한다. 기존의 알고리즘은 동영상의 프레임 간 정보만을 이용하여 배경 확률 모델을 만들고 배경과 전경을 분리한다. 제안하는 알고리즘에서는 먼저 프레임 간의 정보를 이용하여 간단하게 배경과 전경을 분리하는 기존의 배경 분리 알고리즘을 적용한다. 분리된 결과의 정확도를 향상시키기 위해 프레임 내의 정보를 이용하는 GrabCut 알고리즘을 적용한다. 즉, 본 연구에서는 동영상의 프레임 간 정보와 프레임 내 정보를 모두 이용하여 배경과 전경을 분리하고자 한다. 실험결과에서 Change Detection Workshop dataset 에 포함된 몇 가지 영상에 대해 실험 한 후 결과 영상 비교 및 F-measure 를 통해 개선된 결과를 확인할 수 있다.

1. 서론

최근 들어 동영상 촬영 기술과 전송 기술의 발전에 따라 동영상 자료가 많은 분야에서 다양한 용도로 이용되고 있다. 이러한 추세에 따라 동영상 자료들을 자동으로 분석하고 처리하는 기술들이 요구되고 있다. 동영상 처리 기술의 예로는 동영상에 등장하는 물체 인식, 등장 물체의 동작 분석, 교통 상황 모니터링, 동영상 내용에 따른 압축 등이 있다.

배경 분리 기법은 완벽히 혹은 어느 정도 고정된 카메라에 의해 촬영된 동영상에 대해, 배경과 전경을 구분하여 배경을 제외한 전경만을 검출해 내는 기법으로 다양한 동영상 처리 기술의 전처리 과정으로 많이 이용된다. 배경 분리 기법을 적용하여 전경 만을 분리해 낸 후 동영상 처리 기술 적용하면, 좀 더 향상된 결과를 얻을 수 있을 뿐만 아니라 처리 시간을 절약할 수 있다.

배경 분리 기법은 주로 프레임 내의 모든 픽셀에 대해 개별적으로 수행되며 크게 두 가지 과정으로 이루어져 있다. 첫 번째는 모델링 과정이고, 두 번째는 결정 과정이다. 모델링 과정에서는 여러 프레임 동안의 정보를 축적하여 픽셀 마다 각각의 배경 확률 모델을 만든다. 이 모델은 픽셀 값에 따라 배경일 확률을 나타낸다. 결정 과정에서는 한 픽셀의 픽셀 값과 해당 픽셀의 배경 확률 모델을 비교한다. 이러한 과정을 거쳐 해당 픽셀이 배경일 확률을 계산하고, 배경인지 전경인지 결정하게 된다. 이 두 과정을 모든 픽셀에 대해 수행하고 나면 최종 전경 지도를 얻을 수 있다. 최근에는 확률 모델을 만들 때 매번 새로 하지 않고, 기존의 모델을 업데이트 하는 방식을 사용하기 때문에 두 과정의 구분이 뚜렷하지 않은 경우가 많다.

본 논문에서는 기존의 배경 분리 기법에서 프레임 간

정보만 이용하고, 프레임 내 정보를 사용하지 않는다는 점에 착안하여, 프레임 내 정보를 이용하여 기존 배경 분리 기법의 정확도를 개선하는 알고리즘을 제안한다. 기존의 알고리즘의 결과를 살펴보면 실제 배경인 영역을 전경으로 인식하는 경우가 있다. 또, 전경인 영역을 완벽하게 찾지 못하여, 전경과 배경의 경계를 뚜렷하게 잡아내지 못하는 경우가 많다. 제안하는 알고리즘에서는 이러한 점들을 프레임 내 정보를 이용하여 개선하고자 한다. 먼저 최신 배경 분리 알고리즘인 PBAS[1] 알고리즘을 적용하여 전경 영역과 배경 영역을 분리한 후, 그 결과를 이미지 분리 알고리즘인 GrabCut[2]의 입력으로 사용하여 배경과 전경을 좀 더 정확하게 분리하고자 한다.

본 논문은 다음과 같이 구성되어 있다. 2 절에서는 배경 분리 기법에 관련된 연구에 대해 소개한다. 3 절에서는 기존의 PBAS 알고리즘으로 전경과 배경을 분리해 내는 방법을 설명하고, 4 절에서는 3 절에서의 결과를 입력으로 GrabCut 을 적용하여 그 결과를 개선하는 방법을 소개한다. 5 절에서는 기존의 PBAS 알고리즘 결과와 GrabCut 을 적용한 결과의 비교를 통해서 성능향상을 확인한다. 마지막으로 6 절에서는 본 논문에 대한 결론을 맺는다.

2. 관련 연구

대부분의 배경 분리 기법은 픽셀단위로 독립적으로 수행된다. 위치가 \vec{x} 인 어떤 한 픽셀의 시간 t 에서의 픽셀 값을 $\vec{I}(\vec{x}, t)$ 라 하면 그 픽셀의 전경 여부를 나타내는 값

$F_{\vec{x},t}$ 는 다음과 같이 계산될 수 있다.

$$F(\vec{x},t) = \begin{cases} 1 & p(\vec{I}(\vec{x},t) | BG) < threshold \\ 0 & otherwise \end{cases} \quad (1)$$

여기에서, $p(\vec{I}(\vec{x},t) | BG)$ 는 위치가 \vec{x} 인 픽셀의 배경 확률 모델에 픽셀 값 $\vec{I}(\vec{x},t)$ 를 대입하여 배경일 확률을 구한 것 이고, 배경 확률 모델은 $\vec{I}(\vec{x},1), \dots, \vec{I}(\vec{x},t-1)$ 를 바탕으로 만들어진다. 해당 픽셀의 배경 확률이 일정 수준 보다 낮으면 전경, 높으면 전경으로 판단한다. 프레임의 구성하는 모든 픽셀 에 대해 위 과정을 적용하면 최종 전경 지도를 얻을 수 있다.

배경 분리 기법은 위의 배경 확률 모델을 만들 때 파라미터를 이용하는 방법과 이용하지 않는 방법, 이렇게 크게 두 가지 방법으로 나뉜다. 먼저, 파라미터를 이용하는 방법은 평균과 분산 등을 이용해 모델을 나타내는 방법으로 Gaussian model 을 이용하는 것이 대표적이다. Gaussian model 을 이용한 방법으로는 [3-6]이 있다. 두 번째 방법은 파라미터를 이용하지 않고 모델을 만드는 방법으로서 배경으로 판단된 픽셀 값들을 저장해 두었다가 배경 모델로 이용한다. [7]에서는 정해진 개수의 픽셀 값을 누적해서 배경 모델로 저장한다. 그 후, 새 픽셀 값과 배경 모델에 포함된 픽셀 값들 사이의 거리를 계산하여 배경 여부를 결정한다.

최근에는 앞에서 소개한 방법에 기초하여 모델링 과정이나 결정 과정을 향상시킨 알고리즘 또는 새로운 아이디어를 이용한 알고리즘들이 제안되고 있다. [8]에서는 GMM 을 이용하되 추가적으로 Flux Tensor 를 이용해 프레임간 움직임 정보를 이용한다. [9]는 기존의 알고리즘과는 다른 방식으로 모델 대신 분류기를 이용하여 배경을 분리한다. [1]의 경우 파라미터를 이용하지 않고 배경 확률 모델을 만드는 방법으로서, 모델링 과정에서 배경 모델로 가지고 있는 픽셀 값들을 새 픽셀 값으로 갱신할 확률, 결정 과정에서 배경 여부를 판단할 때의 기준 등을 영상의 상황에 맞게 조절하는 방법을 제안하고 있다.

3. PBAS 알고리즘

PBAS 알고리즘은 [1]에서 제안된 “The Pixel-Based Adaptive Segmenter”라는 배경 분리 알고리즘 이며, 알고리즘의 이름에서 알 수 있듯이 영상의 상황에 맞게 자체적으로 적용할 수 있는 능력을 갖고 있다. 배경 분리 알고리즘은 모델링 과정에서 매 프레임마다 모델 갱신 빈도에 따라 배경 모델을 갱신한다. 그리고, 결정 과정에서 갖고 있는 배경 모델과 새 픽셀 값을 비교하여 판단 기준에 따라 전경 여부를 판단한다. 그렇기 때문에 모델 갱신 빈도와 판단 기준은 배경 분리 알고리즘에서 매우 중요하다. 하지만, 이 두 값들은 영상이 안정적인지 불안정적인지, 해당 영역이 배경인지 전경인지에 따라 다르게 적용되어야 한다. 기존의 많은 배경 분리 알고리즘들은 모델 갱신 빈도와 판단 기준으로 일정한

값을 사용하기 때문에, 영상의 특징에 따라 성능 차이가 크게 나타날 수 있다. 반면에, PBAS 알고리즘은 영상의 상황에 맞게 모델 갱신 빈도와, 판단 기준을 조절하는 능력을 갖고 있기 때문에 다양한 상황에서도 좋은 성능을 나타낸다.

위치가 \vec{x} 인 픽셀의 배경 모델은 아래 식 (2)와 같이 N 개의 픽셀 값들을 누적한 집합으로 나타낸다.

$$B(\vec{x}) = \{B_1(\vec{x}), B_2(\vec{x}), \dots, B_N(\vec{x})\} \quad (2)$$

새로운 프레임의 픽셀 값 $\vec{I}(\vec{x},t)$ 에 의해 모델을 갱신할 때에는 N 개의 픽셀 값들 중 하나를 임의로 선택하여 $\vec{I}(\vec{x},t)$ 로 교체함으로써 갱신하게 된다. 이 때, 매 프레임마다 갱신하는 것이 아니라, 모델 갱신 빈도 $T(\vec{x})$ 에 따라 $\frac{1}{T(\vec{x})}$ 의 확률로 갱신 한다. 모델 갱신 빈도는

해당 픽셀이 배경인지 전경인지에 따라 조절되는데, 해당 픽셀이 배경이면 감소하고, 전경이면 증가한다. 즉, 해당 픽셀이 배경에 해당하면 갱신될 확률이 높아지고, 전경에 해당하면 갱신될 확률이 낮아지는 것이다.

앞에서 만든 배경 모델을 바탕으로 새로운 프레임의 픽셀 값 $\vec{I}(\vec{x},t)$ 이 전경인지 배경인지 판단하는 결정 과정은 다음 식 (3)과 같이 나타낸다.

$$F(\vec{x},t) = \begin{cases} 1 & \#\{\text{dist}(\vec{I}(\vec{x},t), B_n(\vec{x})) < R(\vec{x})\} < \#_{\min} \\ 0 & otherwise \end{cases} \quad (3)$$

$\#\{\text{dist}(\vec{I}(\vec{x},t), B_n(\vec{x})) < R(\vec{x})\}$ 는 $n = 1, 2, \dots, N$ 에 대해 해당 조건을 만족하는 n 의 개수를 뜻한다. 다시 말해서, 총 N 개의 배경 모델 값들 과 현재 픽셀 값의 거리를 각각 계산하고, 그 거리가 판단 기준 보다 작은 것의 개수가 $\#_{\min}$ 이하이면 해당 픽셀을 전경으로 그 외의 경우는 배경으로 판단한다. 이 때, 판단 기준이 되는 $R(\vec{x})$ 는 해당 픽셀이 안정적인 영역에 해당하는지 불안정적인 영역에 해당하는지에 따라 조절된다. 해당 픽셀이 안정적인 영역인 경우 $R(\vec{x})$ 값은 감소하고, 불안정적인 영역인 경우 증가하게 된다. 안정적인 영역인 경우 판단 기준을 좀 더 엄격하게 적용하여야 하고, 불안정적인 영역의 경우 판단 기준을 좀 더 느슨하게 적용할 필요가 있기 때문이다.

4. GrabCut 을 통한 정확도 개선

3 절에서 설명한 [1]의 알고리즘을 적용하여 1 차적인 배경 분리를 수행하고 나면 그림 1 의 오른쪽 과 같은 전경 지도를 얻을 수 있다. 실제 전경 지도와 비교해 보면 전경을 정확하게 잡아내지 못하여 완벽한 실루엣을 얻지 못하는 경우가 발생하거나, 전경이 아닌 영역을 전경으로 판단하는 상황도 발생한다. 이러한 상황들은 배경 분리 후에 진행될 물체인식이나 움직임분석 등의 결과에 좋지 않은 영향을

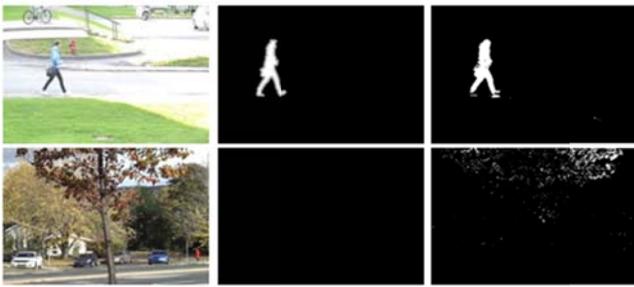


그림 1 입력 영상(왼쪽), 실제 전경 지도(가운데), PBAS 결과(오른쪽)

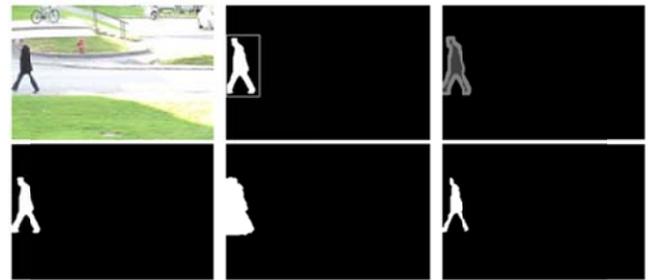


그림 3 입력 영상과 PBAS 결과(왼쪽), 전경을 포함하는 사각형과 그를 이용한 GrabCut 결과(가운데), 확실한 전경/배경 표시와 그를 이용한 GrabCut 결과(오른쪽)

반면에 오른쪽 위 영상처럼 확실히 배경인 영역(검은색), 확실히 전경인 영역(진한 회색) 그리고 애매한 영역(회색)을 정해주면 오른쪽 아래 영상과 같은 좋은 결과를 얻을 수 있다.

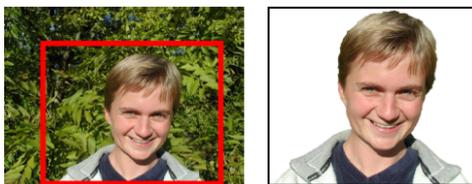


그림 2 입력 영상과 전경을 포함하는 사각형(왼쪽), GrabCut 결과(오른쪽)

미친다. 본 논문에서는 이러한 문제점을 프레임 내 정보를 사용하는 이미지 분리를 이용하여 개선하고자 한다. 즉, 배경 분리 알고리즘을 적용하여 1 차적으로 구한 전경 지도와 주변 픽셀 값들을 이용하여 전경을 한번 더 정확하게 분리한다.

[2]에서 제안된 GrabCut 은 그림 2 처럼, 정지영상에서 사용자가 전경을 포함하는 사각형을 그려줌으로써 배경과 전경을 분리해 주는 알고리즘이다. 첫 번째 단계에서는 사각형 외부를 배경이라 가정하고 포함된 픽셀 값 들을 이용하여 GMM 형태의 배경 모델을 만든다. 또, 사각형 내부를 전경이라 가정하고 포함된 픽셀 값들을 이용하여 GMM 형태의 전경 모델을 만든다. 두 번째 단계에서는 이렇게 구한 배경 모델과 전경 모델을 바탕으로 픽셀 들을 다시 조사하면 확실히 배경인 영역, 확실히 전경인 영역 그리고, 애매한 영역으로 나눌 수 있다. 세 번째 단계에서는 다시 확실히 배경인 영역에 포함된 픽셀 값들 만을 이용하여 보다 정확한 배경 모델을 만들고, 확실히 전경인 영역에 포함된 픽셀 값들만을 이용하여 보다 정확한 전경 모델을 만들 수 있다. 마지막으로 네 번째 단계에서는 세 번째 단계에서 얻은 배경 모델과 전경 모델을 바탕으로 애매한 영역을 배경에 가까운 영역과 전경에 가까운 영역으로 나눌 수 있다.

3 절 의 배경 분리 알고리즘으로부터 얻은 1 차적인 전경 지도를 이용하면 전경을 포함하는 사각형을 쉽게 얻을 수 있고, GrabCut 을 적용하면 최종 전경 지도를 얻을 수 있다. 하지만, 전경을 포함하는 사각형을 얻는 것은 1 차적인 전경 지도를 최대한 활용하는 것이 아니다. 1 차적인 전경 지도로부터 확실히 전경인 영역과 확실히 배경인 영역 그리고 불확실한 영역을 구해낼 수 있다. 이 정보를 바탕으로 GrabCut 의 세 번째 단계부터 적용하면 보다 더 정확한 최종 전경 지도를 얻을 수 있다. 그림 3 의 가운데 위 영상처럼 사각형만 이용하면 가운데 아래 영상처럼 좋은 결과를 얻지 못한다.

5. 실험 결과

실험은 Change Detection Workshop 2012 in CVPR (<http://changedetection.net/>)의 네 가지 데이터 영상을 사용하였다. 첫 번째 영상 highway 는 고속도로 위를 지나가는 차들을 촬영한 영상이고, 두 번째 영상 pedestrians 는 실외 보행자들을 촬영한 영상이다, 세 번째 영상 fall 은 흔들리는 나무를 배경으로 보행자와 차들이 지나다니는 거리를 촬영한 영상이다. 마지막으로 네 번째 영상 office 는 실내에서 사람이 방을 출입하는 것을 촬영한 영상이다.

그림 4 에는 입력 영상과 PBAS 알고리즘을 적용하여 배경을 분리한 1 차적인 전경 지도 그리고 GrabCut 을 적용하여 얻은 최종 전경 지도가 포함되어 있다. 첫 번째 행의 highway 결과를 보면, 1 차적인 전경 지도에서 전경을 정확하게 찾지 못했지만, 최종 전경 지도에서는 전경을 정확하게 찾았음을 확인할 수 있다. 다른 결과에서도 찾고자 하는 전경의 실루엣이 좀 더 정확해 졌음을 확인할 수 있다. 표 1 에서는 1 차적인 전경 지도와 최종 전경 지도를 F-measure 를 통해 비교해 보았다. 배경 분리 알고리즘만을 이용하여 1 차적인 전경 지도를 얻는 것 보다 추가적으로 GrabCut 알고리즘을 적용하여 최종 전경 지도를 얻었을 때, 그 정확도가 향상되었음을 정량적으로 확인할 수 있다.

	PBAS	Proposed
--	------	----------

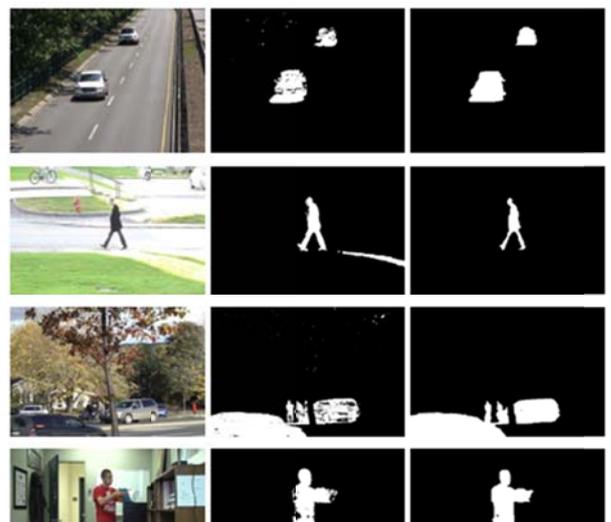


그림 4 입력 영상(왼쪽), PBAS 결과(가운데), GrabCut 결과(오른쪽)

highway	0.734	0.737
pedestrian	0.710	0.827
fall	0.366	0.894
office	0.419	0.433

표 1 실험 영상의 F-measure 비교

Using Flux Tensor with Split Gaussian Models", in proc of IEEE Workshop on Change Detection, 2014

[9] M. De Gregorio and M. Giordano "Change Detection with Weightless Neural Networks", in proc of IEEE Workshop on Change Detection, 2014

6. 결론

본 논문에서는 기존의 배경 분리 알고리즘이 프레임 간 정보만을 이용하는 것에 착안하여 추가적으로 프레임 내 정보를 이용함으로써 정확도를 개선하는 알고리즘을 제안하였다. 프레임 간 정보만을 이용한 기존 알고리즘의 경우 전경 영역을 모두 검출해내지 못하거나 배경 영역을 전경 영역으로 검출하는 상황이 발생하여 정확한 전경 지도를 얻을 수 없다. 하지만, 이를 바탕으로 프레임 내 정보를 추가적으로 이용하면 전경과 배경을 보다 더 정확하게 분리해 낼 수 있다. 기존의 배경 분리 알고리즘을 적용하여 얻은 결과에 추가적으로 GrabCut 을 적용하는 방식이기 때문에 본 논문에서 소개한 PBAS 알고리즘 외에 다른 배경 분리 알고리즘에도 적용할 수 있다.

7. 참고문헌

- [1] M. Hofmann, P.Tiefenbacher, G. Rigoll "Background Segmentation with Feedback: The Pixel-Based Adaptive Segmenter", in proc of IEEE Workshop on Change Detection, 2012
- [2] C. Rother, V. Kolmogorov, and A. Blake "GrabCut: interactive foreground extraction using iterated graph cuts", in ACM SIGGRAPH 2004
- [3] Z. Zivkovic, Improved adaptive Gaussian mixture model for background subtraction, International Conference Pattern Recognition, UK, August, 2004.
- [4] N. Friedman and S. Russell, "Image segmentation in video sequences: A probabilistic approach", In Proceedings Thirteenth Conf. on Uncertainty in Artificial Intelligence, 1997.
- [5] C. Stauffer and W. Grimson, "Adaptive background mixture models for real-time tracking", In Proceedings CVPR, pp. 246-252, 1999.
- [6] E. Hayman and J. Eklundh, "Statistical Background Subtraction for a Mobile Observer", In Proceedings ICCV, 2003.
- [7] A. Elgammal, D. Harwood, and L. Davis. Non-parametric model for background subtraction, in ECCV, 2000.
- [8] R. Wang, F. Bunyak, G. Seetharaman and K. Palaniappan "Static and Moving Object Detection