

밝기 정보와 깊이 정보를 결합한 배경 제거

○

서경민, 이철우

전남대학교 컴퓨터공학과

Background subtraction with combination of intensity and depth information

○

Kyung-Min Seo, Chil-Woo Lee

Dept. of Computer Engineering, Chonnam National University

요약

영상을 전경과 배경으로 분리하는 작업은 영상을 의미 있고 관심의 대상인 전경 영역과 그렇지 않은 배경 영역으로 나눈다는 점에서 매우 유용한 작업이다. 기존의 제안된 방법으로는 intensity 기반, 깊이 기반 그리고 motion 기반 배경 제거 방법 등이 있다. 본 논문에서는 영상내의 intensity 정보와 깊이 정보를 함께 이용하여 영상 내의 배경을 제거하는 방법을 제안한다. 제안하는 방법은 영상 인식과 감시 시스템 등의 전처리로서 활용될 수 있다.

1. 서론

영상이 전경과 배경으로 이루어져 있을 경우 대부분의 의미 있는 정보는 전경 영역에 집중하게 된다. 전경과 배경을 분리한다는 것은 곧 의미 없고 불필요한 배경으로부터 관심의 대상인 전경 영역만을 추출하는 것을 의미한다. 전·배경 분리 과정을 영상 인식, 감시 시스템 등의 전처리로 추가할 경우 정보가 집중된 전경 영역만을 이용하여 처리를 하게 됨으로써 전체 시스템의 성능과 정확성을 향상시킬 수 있게 된다.

Blue-screen과 같은 물리적 제약을 두지 않고, 영상만을 이용하여 전·배경 분리를 하는 방법은 크게 intensity 기반, 깊이 기반, motion 기반의 방법으로 나뉘 볼 수 있다. Intensity 정보를 이용한 배경 제거 방법은 일정한 시간의 영상들로부터 배경 모델을 생성한 후, 새로운 입력과 배경 모델의 intensity를 비교하여 그 유사한 정도가 기준치 이하이면 전경, 이상이면 배경으로 구분하여 배경을 제거하는 방법이다 [1][2]. 이 방법은 실시간 처리가 가능하고 쉽게 구현이 가능하지만, 그림자와 조명 변화 등 주위환경에 민

감한 단점을 지니고 있다. 조명변화에 의한 오류를 줄이기 위해 명암과 색차 신호를 분리하여 처리하는 방법 등이 제안되었으나, 실제 실험 결과 정도의 차이를 보일 뿐 만족할만한 결과를 얻을 수 없었다.

깊이 정보를 이용한 배경 제거 방법은 영상으로부터 깊이 정보를 구한 후 깊이의 정도에 따라 카메라에 가까운 물체는 전경으로 상대적으로 먼 물체는 배경으로 분할하는 방법과[3], 새로운 영상의 깊이 정보와 미리 계산된 배경 모델의 깊이 정보를 비교하여 변화가 있는 부분을 전경으로 보는 방법[4] 등이 있다. 이 방법은 조명 변화와 같은 주위 환경의 변화에 영향을 받지 않지만, 계산량이 많아 처리시간이 오래 걸리고, 일반적으로 2개 이상의 카메라를 사용하여야 하는 단점이 있다.

Motion 정보를 이용한 배경 제거 방법은 배경이 고정된 상황에서 움직이는 물체를 전경으로 분리해내는 방법으로 물체의 움직임이 큰 경우에 오류를 발생하고, 잠음에 민감하며 전경과 배경이 함께 움직일 경우 분할에 실패하게 된다.

본 논문에서는 기존 방법들 중 intensity와 깊이 정보를 이용한 방법을 조합한 보다 견고한 배경제거

방법을 제안한다. 2절에서는 제안하는 방법에 대하여 보다 자세히 설명하고, 3절에서는 제안하는 방법을 실제 영상에 적용한 실험 결과를 보인다. 4절에서는 제안된 방법에서 보완되어야 할 부분과, 차후의 계획에 대하여 정리하며 결론을 맺는다.

2. 밝기 정보와 깊이 정보의 결합

서론에서 언급하였듯이, 배경 제거에 관한 기존 방법들은 각각 보완될 수 있는 장단점을 가지고 있다. 본 논문에서는 intensity 정보를 이용한 방법으로 하되, 보다 정확한 결과를 얻기 위하여 깊이 정보를 추가로 결합하는 방법을 제안한다. 제안된 배경 제거 시스템의 전체 구성은 그림 1과 같다.

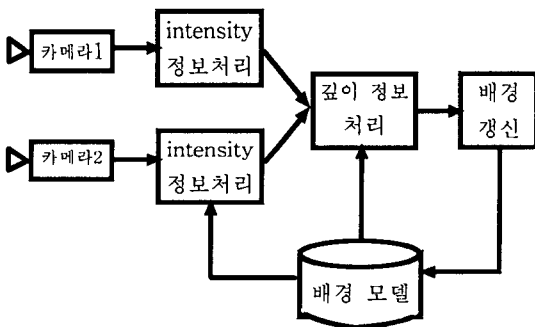


그림 1. 제안한 배경 제거 알고리즘의 전체 구성

입력 영상의 intensity 값과 미리 생성된 배경 모델의 intensity 값이 유사하지 않으면 입력 영상에 전경, 그림자 혹은 노이즈 등이 포함된 것으로 볼 수 있다. Intensity 처리 부분에서는 입력 영상과 배경 모델의 intensity 특성을 비교해서 유사하지 않은 부분들을 추출하고 이 영역을 전경의 후보 영역으로 다음 단계에 넘겨주게 된다. 다음 단계의 깊이 정보 처리 단계는 전 단계에서 구해진 후보 영역 내의 깊이를 계산하여 전경 영역만을 추출하는 단계이다.

깊이 정보를 복원하는 과정에 전 단계인 intensity 처리 단계에서 계산된 정보를 재사용 할 경우 중복되는 처리 시간을 줄일 수 있다. 배경 갱신의 단계는 전 단계에서 구해진 전·배경의 결과를 배경 모델에 반영하여 보다 정확한 결과를 얻을 수 있도록 해주는 부분이다.

2.1 밝기 정보 기반 배경 제거

입력 영상으로부터 전경 후보 영역을 생성하는 단계로서, 입력 영상 I_{new} 와 배경 모델 I_{back} 의 유사한 정도를 $S(I_{new}, I_{back})$ 라고 하면,

$$\text{if } (S(I_{new}, I_{back}) < \text{threshold}) \quad I_{new} = \text{전경 후보 영역,} \\ \text{otherwise} \quad I_{new} = \text{배경 영역이다.}$$

여기서 $S(I_{new}, I_{back})$ 은 단순히 대응되는 pixel값의 차로부터 구할 수도 있고, 보다 더 안정적인 결과를 얻기 위하여 block단위로 그 통계적 특성을 비교하여 구할 수도 있다. 이때 구하여진 값 $S(I_{new}, I_{back})$ 은 입력 영상과 배경 모델이 유사할수록 커지게 된다.

본 논문에서는 두 영상간의 유사한 정도를 계산하기 위하여 block간의 상관값(correlation)을 이용하고, 상관값을 계산하기 위해 필요한 block내의 평균값과 표준편차를 저장해두어 다음 단계인 깊이 정보 처리 단계에서 대응점을 찾기 위한 상관값 계산에 재사용되도록 하였다. 현재 처리 중인 입력 영상과 배경 모델의 block을 b_1, b_2 라 하면, block b_1 과 b_2 간의 상관값은 다음 식에 의하여 계산된다.

$$\rho_{b_1, b_2} = \frac{\sigma_{b_1, b_2}}{\sigma_{b_1} \sigma_{b_2}} = E \left[\left(\frac{b_1 - \mu_{b_1}}{\sigma_{b_1}} \right) \left(\frac{b_2 - \mu_{b_2}}{\sigma_{b_2}} \right) \right]$$

여기서,

$$\mu_{b(u,v)} = \frac{1}{(2n+1)(2m+1)} \sum_{i=-n}^n \sum_{j=-m}^m I(u+i, v+j)$$

은 점 (u, v) 를 중심으로 하는 크기 $(2n \times 2m)$ 인 block의 평균치이고, block 내의 분산 $\sigma_{b(u,v)}$ 는 다음 식에 의해 계산된다.

$$\sigma_{b(u,v)} = \frac{\sum_{i=-n}^n \sum_{j=-m}^m \{ I(u+m, v+n) - \mu_{b(u,v)} \}^2}{(2n+1)(2m+1)}$$

계산된 상관값은 -1에서 1까지의 범위 내에 있으며, 1에 가까울수록 block 간의 유사도가 높은 부분이다. 배경 모델과 유사도가 높은 입력 부분은 배경 영역으로 판단하여 제거되고, 상관값이 임계값 이하로 낮은 부분은 전경의 후보 영역으로 판단하여 다음 단계의 입력으로 쓰이게 된다.

2.2 깊이 정보 기반 배경 제거

전 단계인 intensity 기반 처리 단계에서 입력 영상과 배경 모델의 유사한 정도를 계산하여 전경의 후보 영역을 구하였다. 깊이 정보 처리 단계는 입력 영상으로부터 깊이 정보를 구하고, 구하여진 깊이 정보를 이용하여 전 단계에서 구한 전경의 후보 영역 중 전경만을 다시 분리하는 단계이다. 영상의 깊이 정보를 복원하기 위한 여러 방법들 중 본 논문에서는 2대의 카메라를 사용하는 일반적인 스테레오 방법을 이용하였다. 스테레오 방법은 2장 이상의 영상이 주어질 경우 각 영상의 대응점간의 편차(disparity)로부터 물체의 깊이(depth) 정보를 복원하는 방법으로 다음과 같이 나타낼 수 있다.

$$Z = f \frac{B}{d}$$

Z = 깊이 정보
 f = focal length
 B = 카메라간의 거리
 d = 대응점간의 시차

스테레오 방법으로 영상의 깊이 정보를 구하기 위해서는 먼저 각 영상간의 대응점을 찾아야 한다. 영상간의 대응점을 찾기 위하여 본 논문에서는 block간의 유사도를 평가하였다. 대응점을 찾기 위한 유사도의 평가는 전 단계와 같이 상관값을 이용하고, 이때 전 단계에서 구하여진 block의 μ 와 σ 값을 재사용하여 계산의 중복을 없앴다. 이렇게 구하여진 상관값 중 최고 상관값을 갖는 위치가 대응점을 나타낸다고 볼 수 있다. 대응점 탐색은 입력 영상의 전 영역에 대하여 이루어지는 것이 아니고, 전 단계인 intensity 기반 처리 단계에서 구하여진 입력 영상들의 후보 영역 내에서만 이루어지게 된다.

전 단계에서 계산된 정보를 재사용하고, 전경의 후보영역 내에서만 대응점을 찾기 때문에, 두 방법을 각각 실행한 경우에 비해 처리시간이 단축된다. 또한, 일반적인 stereo 시스템의 경우, 처리해야 할 정보의 크기는 (영상의 크기 x 탐색 영역)이지만, 본 방법에서 필요한 정보의 크기는 (후보 영역 x 가변 탐색 영역)이 된다. 여기서 가변 탐색 영역이란, 후보 영역의 크기가 미리 주어진 탐색 영역보다 작을 경우, 대응점을 찾기 위한 탐색 영역의 크기가 후보 영역의 크기

로 줄어드는 경우이다.

구하여진 입력 영상의 깊이 정보와 배경 모델의 깊이 정보를 비교하여 값의 변화가 없거나, 임계값보다 작은 경우는 노이즈나 조명의 변화 등으로 인한 intensity의 변화로 볼 수 있으므로, 배경 영역으로 평가되고, 해당 영역의 깊이 정보에 충분히 큰 변화가 있었다면 전경으로 평가된다.

2.3 배경 모델의 갱신

영상이 전경과 배경으로 분리되면, 분리된 배경 영역을 배경 모델에 반영하여, 주위 환경의 변화에 대응하도록 하였다. 본 논문에서는 전경이 배경보다 가까이 있다고 가정하고, 배경의 깊이보다 입력 영상의 깊이가 크다면-대상이 카메라로부터 더 멀리 위치한다면- 새로운 배경으로 판단하여, 배경 모델의 깊이 정보를 새로운 값으로 대체하도록 하였다.

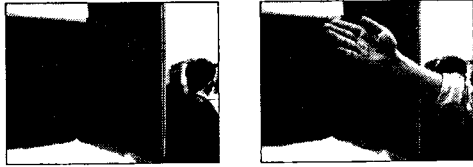
3. 실험결과

3.1 실험 환경

제안된 방법은 Pentium II PC와 Visual C++ 6.0을 사용하여 구현되었다. 영상 내에서 처리되는 block의 크기는 (5x5)로 하고, 대응점을 찾기 위한 최대 탐색 영역의 크기는 60으로 하였다. 기본적인 설정과 임계값은 반복되는 실험 과정에서 경험적으로 구해졌다.

3.2 실험 결과

그림 2와 그림 3은 실제 영상을 이용하여 배경 제거를 수행한 결과이다. 그림 2는 입력 영상 내에 전경과 전경의 그림자가 함께 있는 경우이다. 이 경우, 명암과 색차 성분을 분리하는 기존 방법만으로는 그림자 영역을 완전하게 제거할 수 없음을 볼 수 있다. 본 논문에서 제안된 방법을 이용할 경우, 그림자 영역의 깊이 정보는 변하지 않으므로 배경으로 분류가 되어, 기존 방법보다 정확한 배경 제거의 결과를 얻을 수 있었다. 그림 3은 전경의 주위에 유리와 같이 매끄러운 면이 있어 전경을 반사하는 경우이다. 이와 같은 경우 빛의 반사에 의해 기존 방법으로는 전경만을 추출하기 힘들다. 그러나 본 논문에서 제안하는 방법을 사용함으로써 깊이 정보에 변화가 없는 영역들을 제거하여 전경 영역만을 추출할 수 있었다.



(a) 배경 영상

(b) 입력 영상



(c) 기존 방법

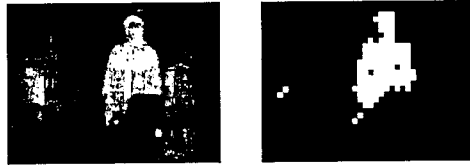
(d) 제안된 방법

그림 2. 그림자가 포함된 경우.



(a) 배경 영상

(b) 입력 영상



(c) 기존 방법

(d) 제안된 방법

그림 3. 전경이 반사되는 경우.

4. 결론

제안된 방법은 기존 방법 중 intensity 정보를 이용한 방법과 깊이 정보를 이용한 방법을 조합함으로써 전·배경 분할의 성능을 개선할 수 있었다. 각 단계별 처리에 필요한 계산들을 효율적으로 공유함으로써, 전체적인 처리 시간을 두 가지 방법을 각각 사용하였을 때에 비하여 줄일 수 있었다. 그러나 처리를 block 단위로 수행하여 전경 영역의 정밀한 윤곽선을 얻을 수 없었다. 차후에는 보다 정밀한 결과를 얻기

위한 후처리 작업의 추가와 각종 임계값의 적절한 선택 문제에 대하여 보완할 예정이다.

[참고문헌]

- [1] Richard J. and M. Ibrahim Sezan, "Video Background Replacement without A Blue Screen", ICIP99. 28A03.
- [2] Ahmed Elgammal, David Harwood and Larry Davis, "Non-parametric Model for Background Subtraction", ECCV 2000, p751~767.
- [3] Christopher Eveland, Kurt Konolige and Robert C. Bolles, "Background Modeling for Segmentation of Video-Rate Stereo Sequences", CVPR 1998, p266~271.
- [4] Yuri Ivanov, Aaron Bobick and John Liu, "Fast Lighting Independent Background Subtraction", IJCV No2. June 2000. p199~207.
- [5] S. Peleg and H. Rom, "Motion Based Segmentation", CVPR 1990, p109~113.
- [6] Cassandra Swain and Tsuhan Chen, "Defocus-based Image Segmentation", ICASSP-95, volume4, p2403~2406.
- [7] Kurt Konolige, "Small Vision Systems: Hardware and Implementation", Proc. ISPR, Hayama, 1997.