

배경 분리를 이용한 보행자 검출 개선

*이상훈 **조남익

서울대학교

*tkd1088@ispl.snu.ac.kr

Improvement of Pedestrian Detection using Background Subtraction

*Lee, Sang-Hoon **Cho, Nam-Ik

Seoul National University

요약

최근 영상 내에서 보행자를 검출하는 기술이 발전하면서 보행자 검출 기술이 다양한 분야에서 응용되고 있다. 영상 내에서 보행자들을 검출함으로써 보행자의 통행량이나 이동경로를 분석할 수 있고, 위험 지역이나 보안 지역에 진입하려는 보행자에게 경고를 줄 수도 있다. CCTV와 같이 고정된 카메라를 이용하여 촬영된 영상의 경우 배경 분리 기술을 적용할 수 있는데, 배경 분리 기술을 통해 영상 내에서 움직이는 물체의 영역을 검출해 낼 수 있다. 본 논문에서는 영상의 배경 분리 결과를 이용하여 보행자 검출의 정확도를 높이고자 한다. 영상 내에서 보행자를 검출 했을 때, 보행자 외에 다른 영역이 보행자로 검출되는 상황이 발생할 수 있다. 이로 인해 보행자 검출의 정확도가 낮아진다. 하지만 배경 분리 결과를 이용하여 전경 부분에서만 보행자가 검출되도록 하고 배경 부분에서는 보행자가 검출되지 않도록 한다면, 보행자가 아닌 영역이 보행자로 검출되는 현상을 막을 수 있다. 실제 HDA Person Dataset에서 실험을 해본 결과, 정량적인 성능 향상을 확인 할 수 있었다.

1. 서론

보행자 검출 기술은 그림 1처럼 영상 내에서 서 있거나 걷고 있는 사람의 영역을 검출하는 기술이다. 이를 이용하면 다양한 분야에 응용할 수 있다. 자동차의 전방에 달린 카메라에 적용한다면, 운전 중 전방에 있는 보행자에 대한 경고를 할 수도 있다. 또, 건물 밖이나 안에 설치된 감시카메라에 적용한다면, 촬영 중인 지역에 대해 보행자에 대한 모니터링이 가능하다. 또, 위험지역이나 보안지역에 침입하려는 보행자에게 경고를 줄 수도 있다.

영상으로부터 보행자를 검출하는 기본적인 방법은 다음과 같다. sliding window 방식을 이용하여 전체 영상 내에서 부분 영역들을 추출하고, 각각의 영역에서 특징벡터를 계산한다. 그 후, 계산된 특징벡터들을 보행자에 해당하는 것과 보행자에 해당하지 않는 것으로 분류한다. 이 때, 특징벡터가 보행자에 해당하는 것으로 분류되면 해당 지역을 보행자 영역으로 검출하게 되는 것이다.

초기의 보행자 검출 기술은 [1]의 HOG 특징벡터를 추출하고, SVM을 이용하여 추출된 특징벡터들을 분류하는 방식으로 수행되었다. HOG란 Histogram of Oriented Gradients의 약자로, 일정 영역 내에서 여러 가지 방향에 대한 Gradient를 구하고, 그 분포를 특징벡터로 이용하는 것이다. [2]에서는 YUV채널의 밝기 값, 여러 가지 방향에 대한 gradient의 크기 등 여러 가지 채널 정보를 포함한 총 10차원의 특징벡터를 이용하였다. 최근에는 Deep-learning의 발전과 함께 Convolutional Neural Network를 이용하는 방법들이 많이 등장하고 있다. 깊은 네트워크를 학습함으로써, 더 좋은 특징벡터를 추출하고 더 정확한 분류가 가능해지고 있다. [3]은 특징벡터들을 분류할 때, 보행

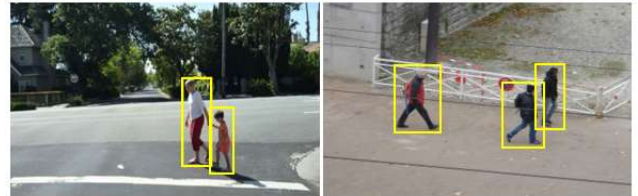


그림 1 보행자 검출 기술

자 여부만 분류하는 것이 아니라 보행자의 자세나 배경의 종류 등 좀 더 세부적인 분류를 통해 보행자 검출 성능을 높였다. [4]는 보행자를 한 번에 검출하는 것이 아니라, 보행자의 다양한 부분을 검출한 후 결과들을 조합하여 최종적으로 검출하는 방법을 이용하였다.

배경 분리 기술은 그림 2와 같이 고정된 카메라에 의해 촬영된 동영상에 대해, 배경에 해당하는 영역과 전경에 해당하는 영역을 구분하여 배경을 제외한 전경 영역만을 검출해 내는 기술이다. 배경 분리 기술은 다양한 동영상 분석 기술로 많이 이용된다. 동영상에 등장하는 물체 영역만 얻을 수 있기 때문에, 물체 인식이나 사람의 동작 분석에 사용할 수 있다. 또, 변하지 않는 배경 영역과 그렇지 않은 전경 영역을 따로 압축하는 동영상 압축 기술에도 사용될 수 있다.

대부분의 배경 분리 기법은 크게 두 가지 과정으로 이루어져 있다. 첫 번째는 모델링 과정인데, 과거에서부터 축적된 영상들을 이용하여 배경 모델을 만든다. 아주 간단하게는 영상들의 평균 영상을 배경 모델로 사용할 수도 있고, 다양한 확률 모델을 이용하여 배경 모델을 만들 수도 있다. 다음은 결정 과정인데, 새로 들어온 영상을 배경 모델과



그림 2 배경 분리 기술

비교하여 배경에 해당하는 영역과 전경에 해당하는 영역을 구분하는 과정이다. 배경을 모델링할 때 어떤 방법을 사용했는지에 따라 다양한 비교 방법이 사용된다. 이 두 과정은 픽셀 단위로 독립적으로 수행되며, 모든 픽셀에 대해 수행하고 나면 그림 2의 오른쪽처럼 최종적인 전경 지도를 얻을 수 있다.

배경 분리 기법은 배경 모델을 만들 때 모델링 방법에 따라 크게 두 가지 방법으로 나뉜다. 먼저, 파라미터를 이용하는 방법은 평균과 분산 등의 파라미터를 이용해 모델을 나타내는 방법으로 Gaussian mixture model(GMM)을 이용하는 것[5]이 대표적이다. 두 번째 방법은 파라미터를 이용하지 않고 모델을 만드는 방법으로서 누적된 픽셀 값들의 집합을 모델로 이용한다. [6]에서는 정해진 개수의 픽셀 값들의 집합을 모델로 저장하고 있다가 kernel을 이용하여 새 픽셀값과 모델 사이의 거리를 계산하여 배경 여부를 결정한다. 최근에는 [7]에서 처럼 GMM을 이용하여 배경 모델을 만들되 Flux Tensor를 이용해 영상간의 움직임 정보를 추가로 이용하기도 한다.

본 논문에서는 영상의 배경 분리 결과를 참고하여 보행자 검출의 정확도를 높이고자 한다. 영상 내에서 보행자를 검출 했을 때, 보행자 외에 다른 영역이 보행자로 검출 되는 상황이 발생할 수 있다. 이로 인해 보행자 검출의 정확도가 낮아지는 것이다. 이 때, 배경 분리 결과를 이용하여 전경에 해당하는 영역에서만 보행자가 검출되도록 하고 배경에 해당하는 부분에서는 보행자가 검출되지 않도록 한다면, 보행자가 아닌 영역이 보행자로 오검출 되는 현상을 막을 수 있다. 실제로 HDA Person Dataset의 9개 영상에 대해 실험을 해본 결과, recall을 유지한 채 precision이 증가하는 것을 확인할 수 있었다.

2. 본론

아래 그림 3은 한 입력 영상에 대해 보행자를 검출 해 본 것이다. 총 두 명의 보행자가 복도를 걸어가는 영상인데, 보행자 검출 결과 두 명의 보행자와 함께 오른쪽 벽에 설치된 소화기가 검출되는 것을 확인할 수 있다. 즉, 두 개의 올바른 검출과 함께 하나의 잘못된 검출이 발생한 것이다.



그림 3 입력 영상(왼쪽)과 보행자 검출 결과(오른쪽)

이 세 개의 검출 결과를 두 개의 올바른 검출 결과와 하나의 잘못된 검출 결과로 구분할 수 있는 방법으로서 본 연구에서는 배경 분리

기술을 사용한다. 복도를 걸어가고 있는 두 명의 보행자는 움직이는 물체이므로 배경 분리 결과 전경으로 나타날 것이고, 오른쪽의 소화기는 배경으로 나타날 것이기 때문이다. 실제로 배경 분리를 적용시켜 본 결과, 그림 4의 왼쪽과 같이 전체 영상에서 두 보행자 영역, 열려있다 닫힌 문 영역을 제외한 배경이 제거 되었다. 문제가 되었던 소화기 영역이 제거되었으니, 배경이 분리된 영상에 대해 보행자 검출을 적용해 보았다. 예상대로 소화기 영역이 보행자로 잘못 검출되는 것을 방지할 수 있었다. 그런데, 원래 잘 검출할 수 있었던 보행자 한명을 검출하지 못하게 되었다. 즉, 배경 분리 결과가 완벽하지 않아서 보행자의 실루엣을 제대로 보존하지 못하는 경우, 원본 영상에서 잘 검출하던 보행자를 검출하지 못하게 되는 것을 확인할 수 있다.

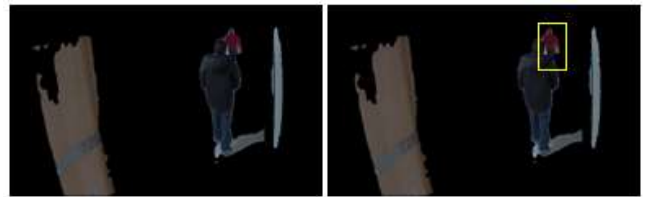


그림 4 배경 분리 결과(왼쪽)와 보행자 검출 결과(오른쪽)

즉, 원본 영상에서 배경을 제거한 후 보행자 검출을 수행하는 것이 역효과를 가져다 줄 수 있다는 것이다. 이 역효과를 방지하기 위해서 그림 5와 같이 새로운 방법을 제안하고자 한다. 먼저, 입력 영상에 대해 보행자 검출 결과와 배경 분리 결과를 각각 얻는다. 그 후, 배경 분리 결과를 참고하여 보행자로 검출된 영역 각각에 대해 영역 내부의 전경과 배경의 비율을 계산한다. 보행자로 검출된 영역 내부의 전경 비율이 일정 수준 이상인 경우 검출 결과를 유지하고, 일정 수준 이하인 경우 검출 결과를 취소하도록 한다. 그림 5에서 보행자 검출 결과 총 세 개의 검출이 나타났는데, 두 개의 올바른 검출의 경우 검출 영역의

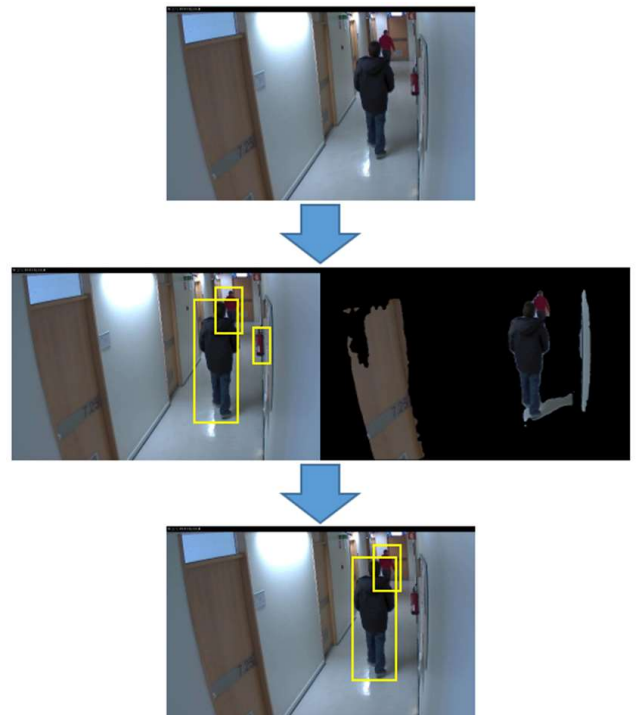


그림 5 제안하는 방법의 전체 구조(입력 영상(위), 보행자 검출 및 배경 분리 결과(가운데), 최종 보행자 검출 결과(아래))

반 이상(각각 71%, 64%)이 전경에 해당하고, 하나의 잘못된 검출은 검출 영역이 전부 배경에 해당한다. 따라서 최종 검출결과로는 두 개의 검출 결과만 유지하게 되는 것이다.

배경 분리 결과를 이용하여 보행자 검출 결과가 개선되었는지 정량적으로 확인해보기 위해서 HDA Person Dataset에 대해 실험을 해 보았다. 본 실험에서는 총 13개의 영상 중에서 보행자 검출에 적절하지 않은 2개의 영상과 배경 분리에 적절하지 않은 2개의 영상을 제외하고 9개의 영상을 이용하였다. 두 가지 실험을 진행하였는데, 첫 번째 실험은 영상에 대해 보행자 검출만을 수행하는 것이고, 두 번째 실험은 영상에 대해 보행자 검출을 수행 한 후 배경 분리 결과를 참고하여 최종 검출 결과를 얻는 것이다. 배경 분리 알고리즘으로는 PBAS[6]를 사용하였고, 보행자 검출 알고리즘으로는 LDCF[8]를 사용하였다.

총 9개의 영상에 대해 실험해본 결과는 아래 표 1과 같다. 단순히 보행자 검출만을 수행했을 때에 비해, 보행자 검출 후 배경 분리 결과를 참고하였을 때 recall은 1.3%p 감소하면서 precision이 16.5%p 증가하게 되는 것을 확인 할 수 있었다. 즉 잘 검출한 보행자를 놓치는 역효과를 최소화 하면서, 오검출을 많이 감소시킬 수 있었다. 그림 6과 같이 실제 실험 영상으로 부터도 그 효과를 확인 할 수 있다.

영상	보행자 검출		배경 분리 + 보행자 검출	
	precision	recall	precision	recall
cam17	0.215	0.525	0.217	0.525
cam19	0.273	0.818	0.553	0.818
cam40	0.850	0.512	0.853	0.485
cam50	0.314	0.885	0.755	0.886
cam53	0.304	0.515	0.319	0.515
cam56	0.131	0.414	0.132	0.414
cam57	0.198	0.607	0.198	0.607
cam58	0.140	0.817	0.141	0.818
cam59	0.112	0.348	0.120	0.349
평균	0.342	0.641	0.507	0.628

표 1 배경 분리 결과를 이용한 보행자 검출 실험 결과

3. 결론

본 논문에서는 영상의 배경 분리 결과를 참고하여 보행자 검출의 정확도를 높이고자 하였다. 영상 내에서 보행자를 검출 했을 때, 보행자 외에 다른 영역이 보행자로 검출 되는 상황이 발생할 수 있는데, 배경 분리를 이용하여 이러한 상황을 방지하고자 하였다. 간단하게 배경 분리와 보행자 검출을 순차적으로 수행할 수도 있지만, 그 경우 역효과가 발생하는 것을 알 수 있었다. 보행자 검출 수행 후, 배경 분리 결과를 참고하여 검출 결과들을 다시 분류함으로써 오검출을 거를 수 있게 되었다. 실제 실험 결과 recall이 크게 감소하지 않으면서 precision이 증가하는 것을 확인할 수 있었다. 배경 분리 결과를 이용하면 보행자 검출을 수행할 영역을 축소할 수 있을 것인데, 이를 이용하여 보행자 검출 시간을 단축하는 것을 후속 연구로 삼을 수 있을 것으로 보인다.

감사의 글

이 논문은 경찰청과 치안과학기술연구개발사업단의 지원을 받아 수행된 치안과학기술연구개발사업임 (PA-C000001).



그림 6 배경 분리 결과를 이용한 보행자 검출 실험 결과(입력 영상(왼쪽), 보행자 검출 결과(가운데), 배경 분리 결과를 이용한 보행자 검출 결과(오른쪽))

참고문헌

- [1] Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). Vol. 1. IEEE, 2005.
- [2] Dollár, Piotr, et al. "Fast feature pyramids for object detection." IEEE Transactions on Pattern Analysis and Machine Intelligence 36.8 (2014): 1532-1545.
- [3] Tian, Yonglong, et al. "Pedestrian detection aided by deep learning semantic tasks." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015.
- [4] Girshick, Ross, et al. "Deformable part models are convolutional neural networks." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015.
- [5] N. Friedman and S. Russell, "Image segmentation in video sequences: A probabilistic approach", In Proceedings Thirteenth Conf. on Uncertainty in Artificial Intelligence, 1997.
- [6] Hofmann, Martin, Philipp Tiefenbacher, and Gerhard Rigoll. "Background segmentation with feedback: The pixel-based adaptive segmenter." Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on. IEEE, 2012.
- [7] R. Wang, F. Bunyak, G. Seetharaman and K. Palaniappan "Static and Moving Object Detection Using Flux Tensor with Split Gaussian Models", in proc of IEEE Workshop on Change Detection, 2014
- [8] Nam, Woonhyun, Piotr Dollár, and Joon Hee Han. "Local decorrelation for improved pedestrian detection." Advances in Neural Information Processing Systems. 2014.