

색상 정보와 HOG feature를 이용한 보행자 검출 및 추적

한상윤 길태호 황인성 조남익

서울대학교 전기·정보공학부

jaegal88@ispl.snu.ac.kr

Pedestrian Detection Based on the HOG feature and Color Information

Han, Sang-Yoon Kil, Tae-Ho Hwang In-Sung Cho, Nam-Ik

Department of Electrical and Computer Engineering, Seoul National University

School of Electrical Engineering

요약

본 논문에서는 HOG 기반 보행자 검출 및 추적에서, HOG feature의 슬라이딩 윈도우의 수와 피라미드 층 수가 알고리즘의 수행속도와 직접적인 관계가 있다는 것을 확인한다. 그리고 이 결과를 바탕으로 윈도우의 수와 피라미드 층 수를 줄이는 방법을 제안하여 전체적인 보행자 검출 및 추적 속도를 증가시키고자 한다. 구체적으로, 제안하는 알고리즘은 검출 단계에서 색상의 선명도를 이용하여 관심 영역을 프레임 내에 지정함으로써 슬라이딩 윈도우의 수를 줄이고, 부가적으로 피라미드 층 수 또한 줄어들어서 보행자 검출 속도를 향상시킨다. 그리고 추적 단계에서는 보행자로 검출된 윈도우의 색상 정보를 이용하여 검출된 보행자를 빠르고 정확하게 추적하는 하는 방법을 제시한다.

1. 서론

보행자 인식 기술은 영상에서 보행자를 찾는 기술로서, 특히 최근 에 미국과 유럽 등에서는 자동차에 보행자 안전 시스템(AEBS)의 탑재 의무화가 법으로 추진되고 있어서 이에 대한 연구개발이 활발히 이루어지고 있다. 보행자 인식 기술이 자동차에 탑재되기 위해서 가장 중요한 요소는 물론 정확도이며, 실시간 인식을 위해서 알고리즘의 속도 또한 큰 이슈가 되고 있다.

보행자 인식 기술과 관련하여, 다른 검출/인식 알고리즘과 마찬가지로 보행자를 검출하기 위한 feature 추출과 이의 classification에 대한 여러 가지 연구가 진행되고 있다. 기존의 방법들 중에서 현재 가장 널리 쓰이는 HOG feature [1] 및 피라미드 기반 검출/추적기는 그 성능에 있어서 다른 feature를 이용한 검출기에 비해 탁월한 성능을 보이거나 실행 속도가 느리다는 단점을 가지고 있어, 자동차와 관련된 응용에 적용하기에는 어려움이 있다. 따라서 본 논문에서는 영상의 컬러 정보를 이용하여 보행자 검출을 위한 관심 영역의 크기를 줄이면서 성능은 유지할 수 있는 방법을 제시한다. 관심영역의 크기가 줄어들면 검출을 위한 슬라이딩 윈도우의 개수가 줄어들며 따라서 전체적인 검출 속도가 향상되는 결과를 얻을 수 있다. 보행자 추적 과정에서는 검출된 윈도우의 색상 정보에 대한 히스토그램을 사용하여 추적 속도 또한 향상시키고자 한다.

2. 관심영역의 크기와 피라미드 층 수가 수행 속도에 미치는 영향

영상에서의 보행자 검출은 스케일 불변을 위하여 주어진 영상을 그림 1과 같은 피라미드 구조로 만들고, 각 피라미드 이미지마다 HOG feature를 추출하고 그림 2와 같이 슬라이딩 윈도우 기법으로 현재 윈도우 위치 내의 feature가 보행자에 해당하는가를 SVM을 통하여 판단한다. 영상의 크기에 따라 윈도우가 움직이는 양이 달라지므로 각 피라미드 영상마다 계산 시간이 다르며, 윈도우의 크기와 슬라이딩해야 할 범위 등에 따라 계산시간이 달라진다.

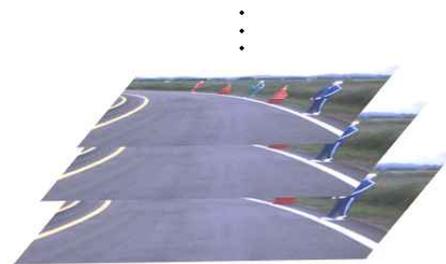


그림 1. 영상 피라미드 구조

본 연구에서 사용한 입력 영상의 크기는 1024×768 로서, 가장 기본 단계 영상의 슬라이딩 윈도우의 크기는 통상 64×128 픽셀크기의 윈도우를 쓰며, 슬라이딩 단위는 4 픽셀 단위로 이루어진다. 따라서 이 영상에 대하여 보행자 검출에 필요한 윈도우의 개수는 60433개이다.

피라미드의 층 수가 올라갈 때마다 피라미드의 크기가 1.05배씩 작아 지도록 했으므로 필요한 슬라이딩 윈도우의 개수는 $(1/1.05)^2$ 배씩 작아 지며, 총 필요한 윈도우의 개수는 405036개가 된다. 이와 같이 초기에 관심영역이 설정되지 않고 전체 영역에 대하여 윈도우를 적용해야 한다면, 필요한 윈도우의 개수가 상당하며 이 때문에 한 프레임 당 연산속도가 약 0.7 sec에서 0.9 sec 정도로 실시간 처리가 어려운 속도가 나온다. 따라서, 특정한 기준에 따라 이미지 전부가 아닌 이미지 내의 특정한 영역에 대해서만 이미지 피라미드를 쌓아서 그에 따른 feature를 계산하는 것이 계산 속도의 향상에 크게 도움이 될 것이다. 이를 위하여 본 연구에서는 보행자는 일반적으로 도로나 주변 배경과 색상으로도 구분되는 가정을 바탕으로, 색상이 주변과 대비되는 곳을 중심으로 윈도우를 제한함으로써 전체적인 검출 및 추적속도를 향상시키는 방법을 제안한다. 즉, 도로는 대부분 어두운 색이며, 보행자의 옷이 어두운 색상이어도 도로나 주변 환경과는 어느 정도 구분이 되므로 컬러를 바탕으로 관심 영역을 제한할 수 있다. 만약 보행자의 색상이 도로나 주변과 거의 구분이 안될 정도이거나 주변 조명이 어두워서 색상이 구분이 안될 정도의 환경이라면 HOG 피쳐 추출도 어렵기 때문에 영상 전체에 대한 검출도 실패할 수 있다. 구체적으로, 본 연구에서는 입력 이미지를 RGB 영상에서 HSV 영상으로 변환하고 S값과 V값이 주변과 대비되는 부분, 즉 컬러가 대비되는 영역에 대해서만 윈도우를 적용하도록 했다.



그림 2. 슬라이딩 윈도우

3. 실험 결과

실험에 사용한 영상은 KATRI 모의 주행로에서 보행자를 차량 내의 카메라를 이용하여 촬영한 것이다. 비교적 주변 지대가 평탄한 곳으로 실험에 앞서 가정하였던, 보행자가 주변 환경에 대하여 구분되는 색상을 지니고 있다는 조건을 만족하는 곳이다. 1024×768 픽셀 크기의 영상을 이용하였으며, 영상 피라미드는 최대 10개의 층을 가지도록 하였다. 슬라이딩 윈도우는 64×128 픽셀 사이즈를 사용하였으며, SVM을 위한 트레이닝 세트로는 KATRI 시험로 내의 영상 및 서울대학교 정밀연구소 인근 영상을 사용하였다.

그림 3은 프레임 내에서 색상값이 많은 관심 영역을 구하는 알고리즘으로서, 첫 단계에서는 원 영상의 RGB 값을 HSV 값으로 변환시켜 S 값과 V 값이 각각 120과 40 이상인 영역으로 제한한 바이너리 맵을 만든다. S나 V값이 낮더라도 주변과 대비되는 경우 몇 단계의 판단조건을 추가하여 주변 색상과 다른 영역을 찾을 수 있다. 다음 단계에서는, 위 과정의 조건을 만족하는 영역에 1, 만족하지 못하는 영역에 0을 할당하였으며 이를 모폴로지 연산을 통하여 영역 내의 비는 부분을 채

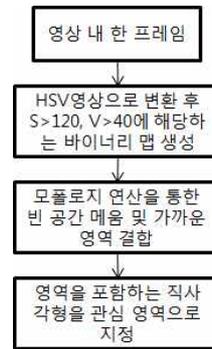


그림 3. 관심영역 설정 알고리즘

우거나 가까운 거리에 위치하는 바이너리 영역들을 서로 묶어주어 하나의 영역으로 만든다. 이 영역을 포함하는 직사각형 영역을 관심 영역으로 두어, 해당 영역의 이미지 피라미드를 생성하여 검출에 이용한다.

그림 4는 그림 3의 알고리즘을 실행 영상의 한 프레임에 적용한 결과로, (a)의 원본 영상에서 (b), (c) 단계를 거쳐 (d)의 관심 영역을 도출해낸 것이다. 총 5 부분의 관심 영역을 도출해내었고, 이에 따라 필요한 슬라이딩 윈도우의 개수는 약 1/40 가량으로 줄어들었다. 이는 슬라이딩하는 공간이 줄어들어 한 피라미드 내에서의 연산량 감소도 있지만, 관심 영역의 크기가 줄어들었기 때문에 피라미드 층 수가 감소하는 효과도 있었기 때문이다. SVM으로 보행자를 검출하는 연산을 해야하는 윈도우의 개수가 줄어들었기 때문에 초당 프레임 수는 약 30 프레임 ~ 40 프레임 정도로 증가하였다.

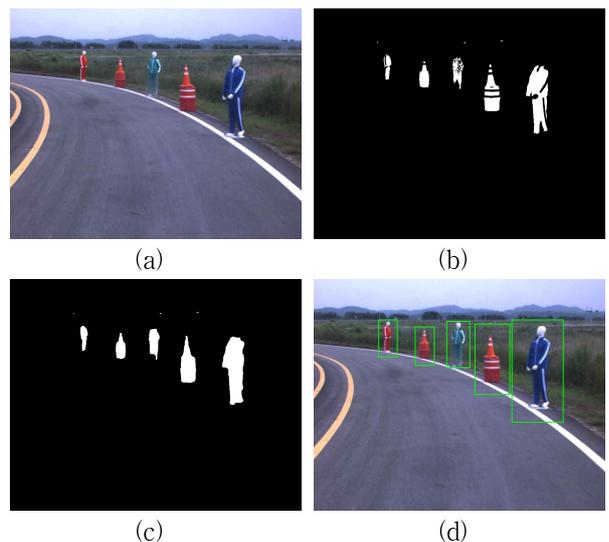


그림 4. (a) 원본 영상 (b) HSV의 S와 V값을 이용한 바이너리 맵 영상 (c) 모폴로지 영상 (d) 관심 영역

표1은 본 논문의 색상의 선명도를 이용한 HOG feature 기반 알고리즘과 HOG cascades, Haar cascades들의 알고리즘을 여러 실험 영상들에 대하여 비교한 것이다. 첫 번째 행의 hsv color based ROI hog*가 본 논문의 FPS를 의미하며, 차래대로 각각 일반적인 HOG[1], cascade를 이용한 Haar feature의 이용[2], HOG cascade를 이용한 방법[3]이 있다. 다른 영상과는 달리 seq4에서는 본 논문의 hsv color

based ROI HOG*의 FPS 성능이 떨어지는 것을 확인할 수 있는데, 이는 그림 5에서 보는 바와 같이 seq4의 경우에는 도로 경계의 연석과 같이 색상이 선명한 부분이 많은 프레임이 많았기 때문이다. 즉, 더 많은 영역에 대하여 윈도우를 적용해 봐야 하므로 수행속도가 떨어지게 된다.

표 1. 알고리즘별 FPS(초당 프레임 수)

평균 FPS(평균 초당 프레임 수)	Seq 1	Seq 2	Seq 3	Seq 4
hsv color based ROI HOG*	34.8	31.5	35.1	20.4
HOG[1]	1.7	1.3	1.2	1.5
Haar cascades[2]	8.9	10.7	8.8	9.5
HOG cascades[3]	5.6	5.7	5.6	5.8

추적단계에서는 검출된 보행자의 HSV 색상 중 H 값의 분포 히스토그램을 이용하여 다음 프레임에서 비슷한 H 값 분포를 가진 영역을 추적한다. H값은 0에서 180사이의 값을 5 단위로 분할하여 36개의 빈을 사용한다. 이를 이용하여 camshift 알고리즘[4]을 적용하면 빠른 초당 프레임 수를 얻을 수 있다.



그림 5. Seq4의 한 프레임

4. 결론

본 논문에서는 HSV 색상 정보를 이용한 관심 영역의 축소를 이용하여 영상 내에서의 보행자 검출 및 추적에서의 초당 프레임 수 이득에 대하여 알아보았다. 실험 결과, 영상에서의 보행자 검출 및 추적 속도의 초당 프레임 수가 상승함을 확인할 수 있었다. 다만, 보행자의 색상값이 도로나 주변환경보다 높다는 초기 가정에 위배되는 영상의 경우에는 윈도우를 적용해야 할 범위가 넓어져서 수행속도가 다소 떨어지는 경우가 생긴다. 이 경우 컬러 대비도를 사용한 saliency detection 방법과 유사하게, 주변 색상과 대비되는 부분을 찾도록 알고리즘을 보다 정교화 할 필요가 있다.

감사의 글

이 논문은 2014년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No.2009-0083495).

참고문헌

[1] Dalal, Navneet, and Bill Triggs. "Histograms of oriented

gradients for human detection." Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. Vol. 1. IEEE, 2005.

[2] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. Vol. 1. IEEE, 2001.

[3] Zhu, Qiang, et al. "Fast human detection using a cascade of histograms of oriented gradients." Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on. Vol. 2. IEEE, 2006.

[4] Allen, John G., Richard YD Xu, and Jesse S. Jin. "Object tracking using camshift algorithm and multiple quantized feature spaces." Proceedings of the Pan-Sydney area workshop on Visual information processing. Australian Computer Society, Inc., 2004.

[5] Dollár, Piotr, et al. "Integral Channel Features." BMVC. Vol. 2. No. 3. 2009.