

적분영상 기반 특징 정보 예측을 통한 고속 보행자 검출

Fast Pedestrian Detection Using Estimation of Feature Information Based on Integral Image

김 재 도 *, 한 영 준**

Jae-Do Kim*, Young-Joon Han**

Abstract

This paper enhances the speed of a pedestrian detection using an estimation of feature information based on integral image. Pedestrian model or input image should be resized to the size of various pedestrians. In case that the size of pedestrian model would be changed, pedestrian models with respect to the size of pedestrians should be required. Reducing the size of pedestrian model, however, deteriorates the quality of the model information. Since various features according to the size of pedestrian models should be extracted, repetitive feature extractions spend the most time in overall process of pedestrian detection. In order to enhance the processing time of feature extraction, this paper proposes the fast extraction of pedestrian features based on the estimate of integral image. The efficiency of the proposed method is evaluated by comparative experiments with the Channel Feature and Adaboost training using INRIA person dataset.

요 약

본 논문은 특징 정보 예측을 통한 빠른 보행자 검출 기법을 제안한다. 다양한 크기의 보행자를 검출하기 위해 보행자 모델의 크기나 입력영상의 크기를 변화시킨다. 보행자 모델의 크기를 변화시킬 경우 크기별 모델이 필요하며, 보행자 모델의 크기의 축소시키는 경우 모델 정보를 손상시킨다. 보행자 모델의 다양한 크기별 보행자의 특징을 추출해야 하므로 보행자 특징의 추출은 전체 수행시간 중 가장 많은 시간을 필요로 한다. 따라서 본 논문은 영상 크기에 따라 특징 추출을 반복하지 않고 입력영상에서 얻어진 특징 정보의 예측을 통해 보행자 검출의 특징추출을 수행한다. 제안하는 방법의 효율성을 검증하기 위해 다양한 채널을 가진 ChnFtrs 특징 및 Adaboost 알고리즘을 사용과 학습과 실험을 위한 영상으로 INRIA 보행자 DB를 사용하였다.

Key words : Pedestrian Detection, Adaboost, ChnFtrs Feature, Feature Estimation, Integral Image

* Dept. of Electronic Engineering, Soongsil University(duckjd@ssu.ac.kr, 02-821-2050)

★ Corresponding author(young@ssu.ac.kr, 02-820-0699)

※ Acknowledgment:

“This research was supported by the MSIP(Ministry of Science, ICT & Future Planning), Korea, under the ITRC(Information Technology Research Center) support program (NIPA-2013-H0301-13-2006) supervised by the NIPA(National IT Industry Promotion Agency).”

“This research was supported by Next-Generation Information Computing Development Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT & Future Planning(No. 2012M3C4A7032182).”

Manuscript received Nov. 20, 2013; revised Nov. 27,2013 ; accepted Nov. 27. 2013

1. 서론

경찰청의 최근 10년간 교통사고 발생현황(사망, 부상)의 지표[1]에 따르면 매년 20만 건 이상의 교통사고가 발생하고 이로 인한 부상자는 30만 건 이상, 사망자는 5천 건 이상 발생하고 있다. 보행자 사망자는 전체 교통사고 사망자의 40% 가량으로 높은 비율을 차지하고 있다. 운전자 및 탑승자의 안전을 위한 에어백 의무화, 안전벨트 착용 의무화 등의 안전장치는 크게 활성화된 반면 보행자를 위한 안전장치는 전무한 상황이다. 따라서 보행자 안전에 대한 관심이 점

점 크게 대두된 상황이며, 최근 유럽, 미국을 중심으로 자동차 보행자 안전 시스템의 탑재가 법으로 의무화되고 있다. 유럽에서는 지난 2003년 보행자가 차량과 충돌 시에 받을 충격에 대한 손상을 최소화하기 위해 신체 각 부분에 대한 최대 충격을 규정된 법규를 제정하였다. 2008년 차량 주변 보행자를 검출하여 운전자에게 경고하는 능동 안전 조치 사항을 규정된 법규를 제정하였다. 또한 2014년에는 자동 브레이크 시스템의 유무를 자동차 평가 항목에 추가한다고 발표하였다.

이러한 보행자 보호 시스템은 크게 보행자 감지와 충돌 회피 제어로 나누어진다. 이때 충돌 회피의 기준이 되는 보행자 감지 기술이 핵심이다. 레이더를 이용한 전방 장애물 감지 기술은 그 정확도 면에서는 우월하나 검색시간이 오래 걸려 교통사고 발생 상황에 대처하기 어려운 것이 현실이다. 따라서 발전된 컴퓨터 비전 기술과 비교적 저렴한 카메라 시스템을 이용한 연구들이 많이 연구되고 있다.

Viola와 Jones[2]는 웨이블릿의 일종인 Haar-like 특징과 같은 단순한 특징들의 조합을 통해 물체를 검출하는 방법을 제안하였다. 부스팅(Boosting) 학습 방법을 사용하여 수많은 특징 중 물체를 식별하는데 유용한 특징들을 점진적이고 자동적으로 찾아내는 방법을 사용하였다.

Dalal와 Triggs[3-4]은 보행자의 특징은 지역적인 그래디언트 분포의 히스토그램(Histogram of Oriented Gradients)에 있다고 보고 밀집된 격자형태의 특징 기술자를 제안하였다. 하나의 특징 채널을 이용하여 보행자를 분류할 경우 특징 추출 단계의 시간이 크게 길리지 않아 수행시간이 빠르나 보행자를 보다 명확히 분류할 수 있는 기준을 만들기 위해 Dollar[5]는 다양한 채널의 영상 특징을 추출하고 이를 종합적으로 이용하여 물체를 검출하는 방법을 제안하였다.

다양한 보행자를 검출하기 위해 입력영상의 부분영역과 모델을 비교하는 작업과 부분영역의 특징들로 조합으로 만들어지는 특징벡터 생성에 있어 반복적인 계산 작업이 수행되게 된다. 이때 소요되는 시간을 줄이기 위해 특징 채널의 적분영상(Integral Image)을 이용하는 방법이 널리 사용되고 있다. 그러나 입력영상의 크기를 조절할 때 마다 특징을 재검출한다.

이를 개선을 위해 Dollar[6]는 자신이 제안한 채널 특징을 사용하면서 보행자 모델과 입력영상을 모두 변화시키며 수행횟수를 줄였고 다양한 크기로 변화된 영상에서 특징을 재검출하지 않고 입력영상의 특징을 기반으로 스케일 값만을 사용하여 특징을 예측함으로써 특징추출 단계를 간소화하였다.

수행시간을 줄이기 위해 또 다른 방법으로 Benenson[7]은 스테레오 비전을 통해 얻은 거리 정보로부터 탐색영역 자체를 축소하는 방법을 제안하였다. 이때 GPU를 사용하여 영상처리가 빠른 계산이 이루어지도록 하였다.

그러나 그 수행속도가 느리거나 스테레오 비전 및 GPU 등과 같은 다른 하드웨어를 사용하여야 함으로써 시스템 구축비용이 매우 큰 문제점을 가지고 있다.

따라서 본 논문은 단일 카메라를 사용하는 동시에 특징추출 단계를 더욱 간소화함으로써 보행자 검출의 수행시간을 단축시키는 방법을 제안한다. 특징영상이 아닌 적분영상을 스케일 값만을 이용하여 예측함으로써 적분영상 생성 단계를 제거하여 더욱 빠른 보행자 검출이 이루어지도록 하였다.

2장에서는 기존 보행자 검출 알고리즘의 세부 단계인 전처리 방법, 특징 추출 방법과 특징 기술자, 보행자 분류를 위한 학습 방법 및 분류기, 보행자 추적 알고리즘에 대해 소개하고, 3장에서 제안하는 적분영상 기반 특징 정보 예측 방법에 대해 상세히 기술한다. 4장에서는 기존 보행자 DB를 사용하여 제안하는 방법의 수행시간 감소 성능 및 검출 성능을 평가한다. 마지막으로 5장에서 결론을 도출하며 끝맺는다.

II. 보행자 검출 알고리즘

보행자 검출은 크게 영상의 화질 개선 및 보행자 후보 검출 또는 관심영역 추출을 수행하는 전처리 단계, 보행자 분류기의 입력인 특징벡터를 얻기 위해 영상 특징을 추출하는 특징 추출 단계, 특징벡터를 입력으로 보행자인지 아닌지 판단하는 보행자 인식 및 분류 단계, 검출된 보행자의 이동 방향 또는 궤적을 검출하는 추적 단계로 이루어진다. 이 중 특징 추출 단계와 보행자 인식 및 분류 단계가 보행자 검출 알고리즘에 있어 가장 큰 부분을 차지한다.

1. 보행자 검출의 특징 추출

일반적인 영상의 특징 정보는 단일 영상에서 얻을 수 있는 색상(RGB, YCbCr, LUV 등), 밝기 값만으로 나타낸 그레이 스케일, 밝기 값의 기울기 방향과 증가율을 나타내는 그래디언트(gradient), 그래디언트의 방향별로 그 수와 양을 히스토그램화 한 그래디언트 히스토그램(gradient histogram), 질감을 나타내는 텍스처(texture) 등이 있고, 다수의 영상에서 얻을 수 있는 영상 특징으로는 프레임 간 변화를 나타내는 차영상(difference image), 시간에 따른 움직임 변화를 나타내는 모션(motion) 등이 있다. 이들 각각을 영상의

특징 채널이라 부른다.

특징 채널 자체가 물체를 분류하는데 특징으로 쓰거나 단일 특징만으로 특정 물체를 기술하기에는 어려움이 많다. 따라서 단일 특징의 지역적인 분포를 조합하거나 다양한 특징 채널을 조합하여 특징 기술자를 생성하여 사용하게 된다.

영상을 사용한 물체 검출 분야에서 고전적이면서도 가장 대표적인 특징으로 Viola와 Jones[2]가 제안한 Haar-like 특징이 있다. 웨이블릿(wavelet)의 일종인 특징으로 그 이웃한 사각영역의 특징 채널 값의 차이를 기반으로 한다. Adaboost와 같은 부스팅 알고리즘을 통해 물체를 식별하는데 유용한 사각 특징을 선정하고 최종적으로 이러한 사각 특징의 조합으로 특징 기술자가 생성된다.

현재 보행자 검출 기술에 있어 가장 기본적인 방법론으로는 Dalal와 Triggs[3-4]가 제안한 HOG(Histogram of Oriented Gradient) 특징이 있다. 물체 영상의 그래디언트를 계산하고 그래디언트의 방향을 기반으로 한 지역적인 그래디언트의 분포를 특징으로 한다. 영역적으로 겹치게 구성하여 이 지역적인 값들을 나열한 특징 기술자를 생성한다.

보행자에 있어 착용한 의복에 따라 색상정보가 다양하고, 보행 중 움직임으로 인해 다양한 외형을 가지게 된다. 따라서 하나의 특징 채널을 사용하여 보행자를 기술하는데 한계가 있음을 인지하고 Dollar[5]는 다양한 특징 채널의 지역적인 특징값들을 후보로 하여 부스팅 알고리즘을 통해 선택된 유용한 특징 특징 채널의 특정 영역의 특징값을 조합한 특징 기술자를 제안하였다. 영상에서 얻을 수 있는 모든 특징 채널이 후보가 될 수 있기 때문에 다양한 환경이 고려된 특징을 선택할 수 있다. 보행자 검출에 있어 HOG 특징보다 그 성능이 뛰어나다고 알려져 있다. 이는 후보 특징으로 HOG 특징도 포함되기 때문이다.

2. 보행자 인식 및 분류

다양한 환경에서 정확하게 보행자를 분류하기 위해서는 다양한 환경에서 획득된 보행자 및 비보행자의 영상이 필요하게 된다. 이러한 데이터를 사용하여 학습을 통해 분류 기준 즉 분류기를 생성하게 된다. 보행자 및 비보행자 학습 영상에서 앞에서 언급한 다양한 방식으로 얻어진 특징 기술자를 특징 공간상에서 보행자와 비보행자로 분류하는 초평면을 찾게 되고 이를 분류기로 사용한다.

이러한 학습을 위해 사용되는 알고리즘은 특징 기술자 생성 방식에 따라 크게 두 가지로 분류할 수 있다. 하나는 물체를 분류하는데 어떤 특징과 특징 조합이 유용한지 정의하기 어려운 경우이고, 다른 방식

은 사용자가 물체를 분류하는데 유용하다고 생각하는 특징 기술자를 정의하는 경우이다.

전자의 경우 사용하게 되는 대표적인 학습 방법은 Adaboost 방식이다. “Adaptive Boosting”의 줄임말로 Yoav Freund 와 Robert Schapire[8]가 제안하였다. 약분류기(특징)들 중에 정확성이 높고 서로 보완적인 약분류기들을 선별하여 이들의 선형 조합으로 강분류기를 구성한다. 가중치를 이용해서 수많은 특징들 중에 작은 수의 좋은 특징을 찾기 위한 매우 효율적인 과정의 학습방법이다. 매 반복마다 이전 분류기에서 제대로 분류가 안 된 학습 데이터의 가중치를 증가시켜 다음 반복에서 가중된 학습 데이터를 잘 분류할 수 있는 특징을 찾게 하는 것이다. Adaboost 학습을 통해 선택된 약분류기의 조합인 강분류기가 최적의 해는 아니나 만족할 만한 성능을 내는 분류기를 자동적으로 찾는 손쉬운 학습 방법이다. 하지만, 학습 시간이 길고 학습 데이터에 의존적인 단점이 있다.

후자의 경우에는 신경망 알고리즘, 유전 알고리즘, 결정트리 등 다양한 학습 방법이 사용되고 있지만 최근 SVM(support vector machine) 알고리즘이 많이 사용되고 있다. SVM 알고리즘은 1970년대 후반에 Vapnik에 의해 제안되었다. 주목받지 못하다가 1990년대 들어 필기 인식[9]등의 분야에 우수한 성능이 입증되어 지금은 많은 분야에서 사용되고 있다[10-11]. 주어진 자료에 대해서 그 자료들을 분리하는 초평면 중에서 자료들과 가장 거리가 먼 초평면을 찾는 방법으로 비선형 분류에서도 커널 트릭을 사용하여 적용할 수 있다.

결정트리가 각 변수에 따라 수직적으로 분류 기준을 찾지만 신경망은 여러 개의 변수를 가중치를 사용하여 동시에 고려하므로 유연한 분류가 가능하기 때문에 일반적으로 결정트리보다 신경망이 분류율이 좋다. 그리고 SVM이 이러한 신경망보다 분류율이 좋은 것으로 평가되고 있다. 신경망 모델은 분류율이 최대(오류율을 최소)로 하는 것을 목적으로 하지만 SVM은 분류율을 최대로 하되 더 나아가 분류를 구분하는 기준과 데이터의 여백을 최대화 하는 것을 고려하기 때문에 일반적인 성능이 더 우수하다.

III. 적분영상 기반 특징 정보 예측을 통한 보행자 검출 알고리즘

일반적으로 보행자 검출 알고리즘의 세부 단계 중 특징 추출 단계가 가장 많은 수행시간을 소요한다. 다양한 크기의 보행자를 검출하기 위해 다양한 크기의 보행자 모델을 사용하여 수차례 검색하여야 하며 다양한 크기의 입력영상에 대한 특징 추출 단계가 반

복적으로 수행되어야하기 때문이다. 본 논문에서는 채널 특징을 사용한 보행자 검출 알고리즘에서 특징의 최종형태인 적분영상을 사용하여 특징 정보를 예측함으로써 입력영상의 크기 변화, 특징 추출의 반복을 줄인 빠른 특징 추출 기법을 제안한다.

1. 특징 추출 방법

특징 추출 단계는 일반적으로 영상 확대/축소, 특징 추출, 적분영상 생성으로 이루어진다. 영상 확대/축소는 다양한 크기의 보행자를 찾기 위해 보행자 모델 또는 입력영상의 크기를 조절하는 단계이고, 특징 추출은 사용할 여러 특징 채널(컬러, 그레이스케일, 그래디언트 히스토그램 등)의 값을 추출하는 단계이다. 적분영상은 각 채널 특징 별로 생성되며 수많은 반복 작업으로 이루어진 특징 벡터 생성 단계의 수행시간을 단축시키기 위해서 수행된다.

특징 추출 단계의 수행시간은 영상 확대/축소 횟수와 특징의 채널수에 따라 의존한다. 이를 해결하기 위해 Dollar[6]는 영상 확대/축소를 생략하고 최초 입력영상에서 추출된 특징과 확대/축소 스케일 값만을 이용하여 확대/축소된 영상의 특징을 예측하여 수행시간을 줄였다. 그 후 채널 특징별 적분영상을 생성하는 방법을 사용하였다.

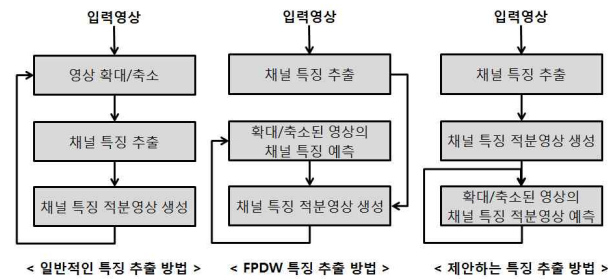


Fig. 1. The comparison of feature extraction method
그림 1. 특징 추출 방법 비교

본 논문에서 제안하는 방법은 최초 입력영상의 채널 특징의 적분영상과 확대/축소 스케일 값을 사용하여 확대/축소된 채널 특징의 적분영상을 예측하는 방법을 제안한다. 제안하는 방법은 영상 확대/축소뿐만 아니라 채널 특징 예측의 반복을 없애 수행시간을 크게 줄일 수 있다. 그림 1은 상기한 특징 추출 방법을 보여준다.

2. 적분영상 예측

적분영상은 쉽게 말해서 현재 픽셀에 이전 픽셀까지의 합이 더해진 영상이다.

적분영상 생성의 일반식은 아래 수식 1과 같다.

$$\Pi(x,y) = \sum_{y=0}^y \sum_{x=0}^x I(x',y') \tag{1}$$

I는 원본영상이고, Π 는 원본영상의 적분영상이다. $\Pi(x,y)$ 는 적분영상의 (x,y)픽셀의 값을 의미하며 이는 원본영상 I의 (0,0)픽셀부터 (x,y)픽셀까지의 모든 픽셀 값의 총합을 말한다.

최근점 이웃 보간법을 적용하여 k배 확대/축소된 영상의 적분영상은 수식 2처럼 나타낼 수 있다.

$$\Pi_R(x,y) = \sum_{y=0}^y \sum_{x=0}^x I_R(x',y') \tag{2}$$

여기서 I_R 은 확대/축소된 영상이고, Π_R 는 확대/축소된 영상의 적분영상이다. 확대/축소된 I_R 에서 적분되는 영역의 크기는 수식 3과 같이 원본영상 I에 대응되는 영역 크기의 k^2 배이다.

$$Area(I_R(x,y)) = k^2 Area(I(x/k,y/k)) \tag{3}$$

여기서 k는 확대/축소 스케일이고 $Area(I(x,y))$ 는 영상의 (0,0)부터 (x,y)까지의 영역크기를 의미한다.

수식 3을 이용하여 Π_R 을 다시 표현하면 수식 4와 같이 표현된다.

$$\Pi_R(x,y) = k^2 \sum_{y=0}^{y/k} \sum_{x=0}^{x/k} I(x',y') \tag{4}$$

수식 4의 적분 부분은 수식 1의 정의를 사용하면 수식 5와 같으므로 최종적으로 확대/축소된 영상의 적분영상은 수식 6과 같이 표현된다.

$$\sum_{y=0}^{y/k} \sum_{x=0}^{x/k} I(x',y') = \Pi(x/k,y/k) \tag{5}$$

$$\Pi_R(x,y) = k^2 \Pi(x/k,y/k) \tag{6}$$

따라서 k배로 확대/축소된 영상의 적분영상에서의

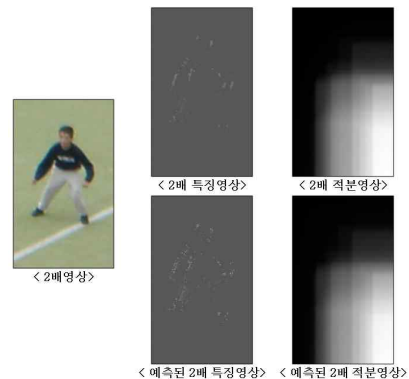


Fig. 2. Original feature and estimated feature
그림 2. 원본 특징과 예측된 특징

한 점의 값은 원본영상의 적분영상에서의 대응점 값에 k^2 배를 하여 쉽게 예측할 수 있다.

그림 2는 입력영상을 2배로 확대한 영상에서 90도~120도 사이의 그래디언트 성분의 특징영상과 이의 적분영상, 입력영상의 원본 크기의 적분영상을 사용하여 예측한 2배 크기의 적분영상 및 이를 미분한 특징영상을 보여주고 준다. 그림에서와 같이 적분영상에서 큰 차이를 보이지 않으며 특징영상 또한 분포에 있어 유사함을 보이는 것을 알 수 있다.

3. 검색 방법

입력영상에서 다양한 크기의 보행자를 검출하기 위해서는 보행자 모델의 크기를 변화시켜 고정된 크기의 입력영상에 적용하거나 하나의 보행자 모델을 다양한 크기로 변화시킨 입력영상에 적용하는 방법, 보행자 모델과 입력영상 모두의 크기를 변화시켜 적용하는 방법으로 크게 나누어 볼 수 있다.

그림 3은 위에서 설명한 대표적인 검색 방법의 개념도이다.

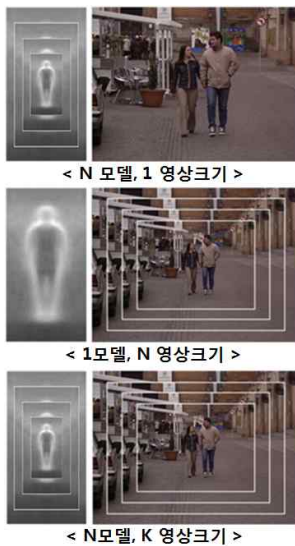


Fig. 3. Searching method of pedestrian detection
그림 3. 보행자 검출의 검색 방법

일반적으로 영상 축소의 경우 그 크기가 작아지면 작아질수록 그림 4처럼 원본의 영상 특성을 쉽게 잃어버리게 된다. 0.5배 축소된 영상의 경우 팔과 몸통 사이의 경계가 불명확하게 변한 것을 볼 수 있다. 이에 비해 2배 확대한 영상에서는 경계가 여전히 명확히 나타남을 볼 수 있다. 특히 보행자 모델은 내부의 부분영역들의 특징을 조합한 벡터이므로 그 영역의

축소는 학습된 모델의 특징을 유지시키지 못해 검출 성능을 크게 저해하는 요인이 된다.



Fig. 4. Effect of image up-sampling and down-sampling
그림 4. 영상의 확대/축소의 영향

이를 방지하기 위해 그림 3처럼 N개의 크기별 모델을 각각 학습시켜 사용할 수도 있다. 그러나 이러한 경우, 다수의 모델을 생성하여야 하므로 학습시간이 크게 증가할 뿐만 아니라 검출 시 다수의 모델을 적용하여야 하기 때문에 검출시간 또한 크게 늘어나게 된다.

또 다른 해결책으로 영상 축소 대신 영상 확대를 사용하는 방법이 있다. 그림 5와 같이 입력영상의 k 배는 보행자 모델의 $1/k$ 배와, 보행자 모델의 k 배는 입력영상의 $1/k$ 배와 동일한 의미를 가진다.

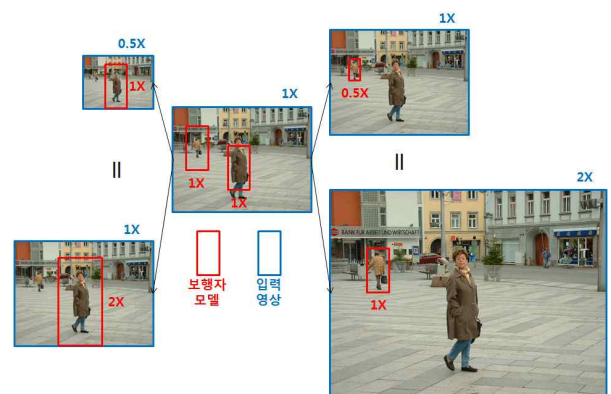


Fig. 5. Resizing relation between pedestrian model and input image
그림 5. 보행자 모델과 입력 영상의 확대/축소 상호 관계

따라서 본 논문에서는 보행자 모델보다 큰 보행자를 검출하기 위해서는 보행자 모델을 확대시키고, 작은 보행자를 검출하기 위해서는 입력영상을 확대시키

는 방법을 사용하여 축소 시 발생할 수 있는 정보 손실을 최소화하는 방법을 사용하였다.

IV. 실험 결과 및 분석

제안하는 방법은 최초 입력영상 크기의 특징 적분 영상을 스케일 파라미터만을 사용하여 다른 크기의 특징 적분영상을 예측하므로 입력영상의 크기를 변화시키고 얻은 특징 영상과의 비교를 통해 예측 성능을 평가하고, 아울러 영상 크기 변화와 반복적인 특징 추출을 수행하는 방법과의 비교를 통해 수행시간 감소 성능을 평가한다. 아울러 최종적인 검출율 비교를 통해 제안하는 방법과 기존 방법과의 성능 차이를 평가한다.

실험 및 성능 평가를 위해 INRIA 보행자 DB[12]를 사용하였다. 적분 영상의 예측 오차를 측정하기 위해 학습용 보행자 영상(보행자 영역만 잘려진) 1,208개를 사용하였고, 보행자 검출기 학습을 위해 INRIA 학습용 보행자 영상과 그 좌우 반전 영상으로 구성된 2,416개의 영상과 랜덤 생성한 24,360개의 비보행자 영상을 사용하였다. 보행자가 포함된 INRIA 실험 영상 288개를 사용하여 특징 추출 단계의 수행시간과 검출 성능을 비교하였다.

실험에 사용된 PC의 성능은 다음과 같다. 프로세서는 Intel(R) core(TM) i7-2600K 3.4GHz, 16GB RAM, 그래픽카드 NVIDIA Geforce GTX 560Ti, 저장장치로 삼성 SSD 840 series (MZ-7TD500)를 사용하였다.

1. 적분영상의 예측 오차

INRIA 훈련용 보행자 영상의 크기는 96×160 픽셀이고 영상 높이와 영상 내 보행자 높이 비율은 1 : 0.6 가량이다. 예측 오차는 2배 확대한 영상의 그래디언트 특징영상과 예측한 2배 크기의 그래디언트 특징영상을 비교하여 측정되었다. 예측한 2배 크기의 그래디언트 특징영상은 크기 96×160 픽셀의 원본영상의 특징 적분영상을 사용하여 2배 크기인 192×320 픽셀 적분영상을 예측한 후 이를 미분하여 얻었다. 그래디언트 방향은 0°~180°사이의 값을 가지며 30°씩 분할하여 총 6개의 특징영상을 구성하였다. 한 영상의 평균 오차를 기반으로 최소오차, 최대오차 및 평균오차를 측정하였다. 표 1과 같이 최소 1에서 최대 10 가량의 값 차이를 보이고 평균적으로 3.5 가량의 오차를 보였다. 픽셀 당 최대 255의 값 차이를 가질 수 있는 것에 비해 매우 작은 오차를 가지는 것으로 나타났

다.

Table 1. Error of integral image estimation
표 1. 적분영상 예측 오차

각도 \ 오차	최소오차	최대오차	평균오차
0°~30°	0.8	9.4	3.4
30°~60°	0.6	6.5	2.3
60°~90°	1.2	11.5	4.9
90°~120°	0.9	10.7	3.9
120°~150°	0.5	5.5	2.3
150°~180°	1.4	11.1	4.3

그림 6은 그래디언트의 각도에 따른 원본 특징영상 및 예측된 특징영상을 보여준다. 예측된 특징영상의 경우 원본 특징영상에 비해 그 값이 더 크게 나타나는 것을 볼 수 있다. 그러나 특징벡터 구성 시 부분영역 특징값을 바로 쓰는 것이 아니라 검출 윈도우 영역의 전체 특징값과 비교하여 그 비율을 사용하기 때문에 보행자 검출에 있어 큰 영향을 주지 않는다.

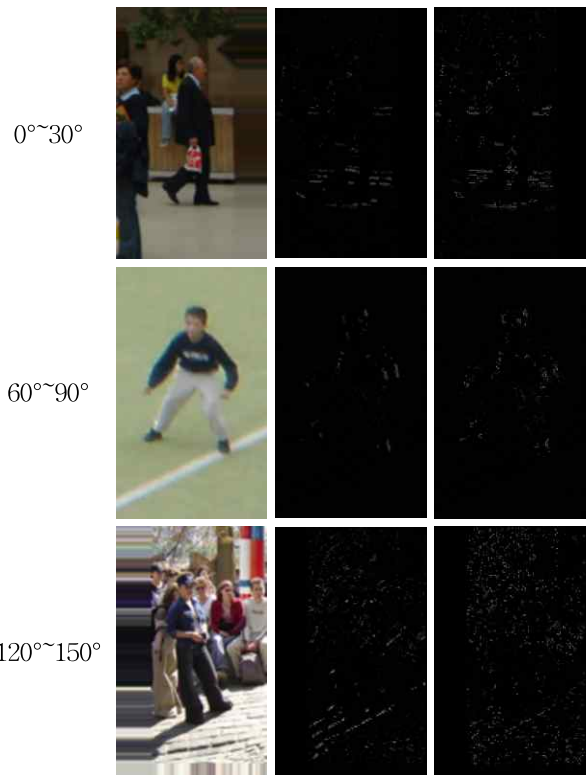


Fig. 6. Original feature and estimated feature images
그림 6. 원본 특징영상 및 예측된 특징영상

2. 검출 성능 비교

Dollar[5]가 제안한 보행자 검출 알고리즘을 사용하였다. 보행자 모델 크기는 96×160 픽셀, 특징 채널은 총 10개를 사용하였다. 검출 성능은 보행자 영역 크기가 80픽셀 이상인 보행자에 대해 실험하였다. INRIA 보행자가 포함된 실험 영상 288개 중 116개 영상이 보행자 영역(여백포함)의 크기가 80픽셀 이상인 보행자를 포함하고 있었고, 총 보행자 수는 288개이다. 표 2는 실험을 통해 얻어진 결과를 보여준다. 반복적인 영상 확대/축소와 특징 추출이 이루어지는 일반적인 추출방법과 반복적인 영상 확대/축소 및 특징 추출을 수행하지 않고 특징의 적분영상을 사용하여 예측하는 방법은 대동소이한 결과를 보였다.

Table 2. 검출 성능 비교

표 2. Detection performance comparison

성능 방법	검출율	장당 오검출 수
일반적인 추출방법	87.5%	190개
제안하는 추출방법	88.2%	193개

그림 7은 보행자 검출 결과 영상이다. 동일한 영역에 나타나는 사각영역을 그룹핑 하지 않은 결과이며, 최초 학습 시 보행자 내부 및 부근의 오검출은 적용되지 않아 많은 오검출을 보인다. 본 논문은 보행자 검출 특징 및 학습 기법의 개선보다는 보행자 검출 시간을 줄이는데 중점을 두어 보행자 검출 성능을 높이기 위한 하드트레이닝(hard training)을 수행하지 않았다. 결과 영상에서도 표 2의 결과와 마찬가지로 제안하는 방법의 결과가 큰 차이가 없음을 보여준다.

3. 수행시간 비교

보행자 모델의 확대는 일반적인 방법과 제안하는 방법에 동일하게 적용되는 부분이므로 제외하고 입력 영상 확대 측면에 대해서만 수행시간을 비교하였다. 입력영상의 크기 확대는 확대된 입력영상의 높이 크기와 보행자 모델의 높이 크기의 비율이 원본 입력영상의 크기 기준으로 보행자 모델의 높이 크기가 80 픽셀 이상의 비율을 가질 때까지만 입력영상을 확대하였다. 따라서 원본 영상에서 높이 48 픽셀 이상인 보행자를 검출할 수 있다. 단, 여백이 포함된 영역이 원본영상을 벗어나지 않아야 한다.



Fig. 7. The Result image of pedestrian detection
그림 7. 보행자 검출 결과 영상

Table 3. 특징 추출 단계 수행시간 비교
표 3. Processing time comparison of feature extraction

방법 시간(ms)	일반적인 방법	제안하는 방법
이미지 확대 평균수행시간	49.53	0
특징추출 평균수행시간	158.6	0
적분영상생성 평균수행시간	21.53	51.8
총 수행시간	229.66	51.8

표 3에서와 같이 제안하는 방법은 일반적인 특징 추출 방법에 비해 특징 추출 단계에서 영상의 확대 및 반복적인 특징 추출 단계를 수행하지 않아 수행시간이 4배 이상 빠른 것으로 나타났다. 또한 표에서 알 수 있듯이 반복적인 특징 추출이 가장 많은 시간이 소요됨을 알 수 있다.

V. 결 론

보행자 검출에 있어 다양한 크기의 보행자를 검출하기 위해 보행자 모델 및 입력영상의 크기 확대/축소가 필요하다. 그러나 영상의 축소는 영상 정보의 손실을 가져오므로 보행자 모델의 확대, 입력영상의 확대를 통해 영상의 축소와 동일한 효과를 내면서 영상의 정보 손실이 없게 보행자를 검색하여야 한다. 입력영상의 확대 시 검출에 사용되는 채널특징을 다시 추출하는 반복적인 작업이 필요하고 이로 인해 수행시간이 크게 증가한다. 본 논문은 이러한 문제를 해결하고자 보행자 검출에 있어 반복 시행 및 처리로 인해 가장 많은 수행시간이 수행되는 특징 추출 단계의 계산 복잡도를 단순화하는 방법을 제안하였다. 빠른 계산을 위해 특징을 적분영상화 하는데, 이 최초 특징 적분영상을 기반으로 반복적인 특징 추출 없이 확대된 영상의 특징 적분영상을 예측함으로써 특징 추출 단계의 계산시간을 크게 감소시켰다. 다양한 실험을 통해 예측된 적분영상 오차는 매우 미미함을 보였고, 그로 인해 최종적인 보행자 검출 성능의 하락을 발생하지 않았다. 이는 더 좋은 성능을 내기 위해 더 다양하고 많은 채널의 특징을 사용하는 방법에 있어 본 논문에서는 수행시간을 크게 향상 시키는 방법을 제시하였다. 하드웨어의 성능 향상과 발전을 통해 더 빠른 보행자 검출이 이루어질 수 있으나 알고리즘

의 최적화는 그 비용을 최소화하면서 동일한 성능을 낼 수 있는 이점을 가진다.

References

- [1] http://www.index.go.kr/egams/stts/jsp/potal/stts/PO_STTS_IdxMain.jsp?idx_cd=1614
- [2] P. A. Violar, and M. J. Jones, "Robust real-time face detection", International Journal of Computer Vision, Vol. 57, No. 2, pp. 137-154, May 2004
- [3] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection", In IEEE Conference on Computer Vision and Pattern Recognition, pp. 886-893, 20-25 June, 2005
- [4] N. Dalal, B. Triggs, and C. Schmid, "Human Detection Using Oriented Histograms of Flow and Appearance", In Proceedings of the 9th European Conference on Computer Vision, Volume Part II, pp.428-441, Feb. 2005
- [5] P. Doll'ar, Z. Tu, P. Perona, and S. Belongie, "Integral channel features", In British Machine Vision Conference, 7-10 September, 2009.
- [6] P. Doll'ar, S. Belongie, and P. Perona, "The fastest pedestrian detector in the west", in British Machine Vision Conference, 21 Aug.-3 Sep., 2010
- [7] R. Benenson, M. Mathias, R. Timofte, and L. Van Gool, "Pedestrian detection at 100 frames per second", In IEEE Conference on Computer Vision and Pattern Recognition, pp. 2903-2910, 16-21 June, 2012
- [8] Y. Freund, and R. E. Schapire, "A decision theoretic generalization of on-line learning and an application to boosting", In Computational Learning Theory, Springer Berlin Heidelberg, pp. 23-37, 1995
- [9] Corinna Cortes, and Vladimir Vapnik, "Support vector networks", Machine Learning, Vol. 20, No. 3, pp. 273-297, 1995
- [10] Christopher J. C. Burges, "A tutorial on Support Vector machines for Pattern Recognition", Data Mining and Knowledge Discovery, Vol. 2, No. 2, pp. 121-167, June 1998
- [11] Alex J. Smola, and Bernhard Schölkopf, "A tutorial on support vector regression", Statistics and Computing, Vol. 14, No. 3, pp. 199-222, Aug. 2004
- [12] <http://pascal.inrialpes.fr/data/human/>

BIOGRAPHY

Jae-Do Kim (Student Member)

He received the B. S., and M. S. degrees in electronics engineering from Soongsil University, Seoul in 2007 and 2009. He is doing a Ph. D. program in Soongsil University. His research interest includes applications of computer vision on autonomous vehicles and HCI.

Young-Joon Han (Member)

He received the M. S., and Ph. D. degrees in electronics engineering from Soongsil University, Seoul in 1998 and 2003, respectively. And became an assistant professor at School of Electronicengineering in Soongsil UniversityE in 2003. He is

currently an associate professor. His research interest includes applications of computer vision on autonomous vehicles, robots, and HCI.