

HOG 특징과 다중 프레임 연산을 이용한 보행자 탐지

Pedestrian Detection using HOG Feature and Multi-Frame Operation

서 창 진[†] · 지 홍 일^{*}
(Chang-jin Seo · Hong-il Ji)

Abstract - A large number of vision applications rely on matching keypoints across images. Pedestrian detection is under constant pressure to increase both its quality and speed. Such progress allows for new application. A higher speed enables its inclusion into large systems with extensive subsequent processing, and its deployment in computationally constrained scenarios. In this paper, we focus on improving the speed of pedestrian detection using HOG(histogram of oriented gradient) and multi frame operation which is robust to illumination changes in cluttering images. The result of our simulation indicates that the detection rate and speed of the proposed method is much faster than that of conventional HOG and differential images.

Key Words : HOG, Object detection, Object tracking, Background image

1. 서 론

컴퓨터 비전을 이용한 물체 탐지 및 추적 시스템은 일반적으로 3단계로 시스템이 구성되어진다. (1) 물체탐지(object detection) (2) 물체 인식(object recognition) (3) 물체 추적(object tracking) 단계로 이루어지며 강건한 물체 추적 시스템을 구성하기 위해서는 3단계로 구성되어진 각 단계마다 물체 탐지 및 성공률이 전체 시스템의 성능에 영향을 미치게 된다[1].

물체를 탐지하기 위한 단일영상 분석법(single image analysis)은 영상의 색상정보, 에지, 명암정보변화, 히스토그램, 템플릿 등을 이용하여 영상내의 물체를 탐지하는 방법이다. 영상의 색상정보를 이용하는 방법은 영상내의 물체의 특징이 갖는 색상 분포가 일정하여야 한다는 가정이 필요하다. 에지(edge) 및 기울기(gradient) 정보의 분석으로 영상에 존재하는 물체를 탐지하는 방법은 영상 처리과정에서 상당한 오버헤드(overhead)가 발생되어진다[2].

다중 프레임 연산은 두 프레임 이상의 영상에서 깊이정보(depth information)를 추출하여 물체의 탐지에 사용한다. 이 방법의 일반적인 문제점은 이동물체의 움직임이 폐색(occlusion)되는 경우와 프레임 사이의 물체의 움직임이 존재하지 않는 경우 정확한 물체를 탐지하기 어려운 단점이 있다. 하지만 다중 프레임 연산을 이용한 처리가 단일 영상을 이용한 물체 탐지보다 빠른 계산 속도를 가진다는 장점이 있다[3].

이 있다[3].

물체의 특징을 설명할 수 있는 변수가 완전히 구해진다면 물체 탐지와 추적이라는 문제에서 가장 일반적이고 완전한 해답이 된다. 하지만 실제영상 획득과정에서 영상의 미세한 흔들림과 잡음 및 환경의 변화로 인하여 획득된 영상에서 완벽한 변수를 찾아내는 작업은 상당히 어렵고 이러한 정보를 찾기 위한 계산량 또한 매우 방대하다. 물체 탐지 및 추적이라는 동적 영상이해를 수행하는 방법은 다음 두 가지 접근방법에 의하여 분류되어질 수 있다.

- (1) 이론적으로 적합하지만 구현하기 어렵고 많은 계산량을 필요로 하는 작업을 선택하는 방법
- (2) 보다 간단한 모델을 사용하는 대신 영상처리의 계산량을 감소시키는 장점을 선택하는 방법

(1)은 주로 광류(optical flow) 및 영상의 템플릿 및 특징 정보의 계산을 통해 가장 일반적인 이론을 제시하고 증명하였으며, (2)는 물체의 특성을 나타낼 수 있는 적절한 3차원 혹은 2차원의 모델을 만들어 이들 모델의 특징(feature)이 인접영상의 시간에서는 매우 서서히 변화한다는 가정을 사용하고 있다.

일반적으로 영상에 존재하는 보행자를 탐지하는 방법은 Viola의 Harr 특징기반[4], Dalal의 기울기의 방향성(HOG)을 이용한 방법[5]과 Morkaya의 배경영상을 이용하여 보행자를 탐지하는 기법[6] 등 다양한 연구가 진행되고 있다. 단일 영상을 이용하는 방법은 영상에 존재하는 보행자의 특징 정보를 이용하여 탐지하므로 보행자의 움직임이 없는 경우에도 보행자를 탐지할 수 있는 장점이 있으나 탐지에 많은 연산을 수행하는 단점이 있다. 그리고 다중 프레임 연산을 이용하는 방법은 보행자의 움직임이 없는 경우 보행자를 탐지하기 어려운 단점을 가지고 있으나 배경이미지 생성 또는 간단한 프레임간 연산으로 빠르게 보행자를 탐지할 수 있는

[†] Corresponding Author : Dept. of National Defense Intelligence Engineering, Sangmyung University, Korea
E-mail : cjseo@smu.ac.kr

^{*} Dept. of Automotive Software, YoungDong University, Korea

접수일자 : 2015년 8월 3일

수정일자 : 2015년 8월 16일

최종완료 : 2015년 8월 18일

장점을 가지고 있다. 본 논문에서는 단일영상 분석방법인 HOG 특징과 다중 프레임 연산을 이용하여 배경이미지를 생성하고 생성된 배경을 이용하여 보행자를 탐지하여 단일 영상의 복잡성과 및 다중 프레임 연산이 가지는 단점을 보완하여 보행자의 움직임이 없는 경우에도 보행자를 탐지할 수 있으며 보다 정확성이 높은 보행자 탐지 방법을 제안하고자 한다.

본 논문의 구성은 2장에서 보행자 탐지를 위한 관련연구를 살펴보고 3장에서는 제안하는 방법의 알고리즘을 설명하며 4장에서는 실험결과를 5장에서는 결론 및 향후 연구 과제를 소개한다.

2. 관련 연구

2.1 단일 프레임을 이용한 방법

2.1.1 SIFT(Scale Invariant Feature Transform)

SIFT는 객체 인식, 중복성 검출 및 객체 추적에 사용되며 다양한 응용분야에서 우수한 성능을 보이고 있다. SIFT는 Lowe가 제안한 방법으로 모든 이미지는 고유한 특성을 가진다는 점을 이용하여 영상의 크기, 회전, 잡음에도 불변한 특징을 가지는 키포인트(keypoints)를 생성하여 물체의 특징을 분석한다[7]. SIFT를 이용한 특징 추출과정은 특징점 검출(detection)과정과 특징을 기술(description)과정으로 구분되어지며 특징점 검출과정은 영상에 대한 스케일 공간을 구성하여 가우시안을 이용하여 특징점을 검출하고, 특징점에 대한 보정을 수행한다. 특징점 기술단계에서는 보정된 특징점들에 대한 방향(orientation)을 계산하고, 이를 기반으로 각 특징점을 크기와 회전에 불변하는 기술자(descriptor)로 표현한다. SIFT는 크기에 불변하는 특징을 다양한 스케일의 영상과 각 특징점간의 비교 연산을 수행하므로 우수한 특징점 탐지능력을 가지고 있는 것으로 알려져 있으나 특징점 계산에 많은 계산량을 가지는 단점을 가지고 있다.

2.1.2 HOG(Histogram of Oriented Gradient)

HOG는 Dalal에 의해 제안된 알고리즘으로 단일 영상에서 보행자 검출을 목적으로 제안한 알고리즘으로 슬라이딩 윈도우 기반의 특징 추출 알고리즘이다. 영상에서 검출하고자 하는 검출 대상영역에서 나타나는 색상의 밝기나 방향의 분포 정보를 특징으로 사용하며 영상의 그래디언트의 크기 정보는 색상이 얼마나 급격하게 변화하가에 따라 결정되고 방향정보는 색상이 변화하는 방향에 따라 결정된다[8].

HOG의 특징은 영상분할, 히스토그램생성, 히스토그램 병합의 3단계로 나누어지며 첫 번째, 단계인 영상분할과정은 입력된 영상을 나누는 과정으로 영상에서 추출할 특징을 동일한 크기로 N개의 윈도우로 나누는 작업을 수행한다. 두 번째 단계는 윈도우내의 그래디언트의 크기 및 방향 정보를 이용하여 히스토그램을 생성하는 단계이다. 그래디언트의 크기 및 방향정보는 x축과 y축의 밝기변화를 이용하여 측정되어진다. 세 번째 단계는 이전 단계에서 생성한 N개의 윈

도우에서 각자 생성한 R개의 벡터로 이루어진 히스토그램들을 연결하여 병합하는 과정을 거치게 된다. 병합과정을 거치게 되면 N×R개의 벡터로 이루어진 HOG 특징 벡터가 생성되게 된다.

2.2 다중 프레임을 이용한 방법

2.2.1 Differential Image

차영상을 이용하는 방법은 연속되는 두 프레임 또는 다수의 프레임을 이용하여 각 프레임의 대응점에 대한 차를 구함으로써 물체를 탐지하는 방법이다. 단, 카메라는 고정되어 있어야 하며 조명은 변화는 일정해야 한다. 차영상 $d(i,j)$ 는 이진화된 영상으로써 움직임이 일어난 영역에는 '0' 이 아닌 값을 가진다. 즉, 연속되는 영상의 대응점의 차를 구하면 움직임이 일어나지 않은 경우는 이전 프레임의 정보 값과 현재 프레임의 정보 값이 동일하기 때문에 '0'의 값을 가지게 된다. 차영상은 다음 식(1)에 의해 표현 되어질 수 있다.

$$d(i,j) = 0 \text{ if } |f_1(i,j) - f_2(i,j)| \leq \epsilon \quad (1)$$

f_1, f_2 연속 영상, ϵ 임계치값

차영상에 의한 움직임 분석으로는 물체가 어느 방향으로 움직였는지에 관한 궤적 정보를 얻을 수 없다. 방향에 대한 정보가 필요 할 경우 누적 차영상(cumulative difference image)에 의해 구해 낼 수 있다. 누적 차영상은 작은 물체의 움직임 또는 느린 물체의 움직임에 관계없이 움직임의 방향과 다른 시간과 관련된 움직임 속성에 대한 정보를 가지고 있다.

2.2.2 블록 정합에 의한 방법

블록 정합 방법은 정합 기반 기술로서 탐지하고자 하는 특징 단위로 영상 세그먼트가 선택되어져야 한다[9]. 이 경우 움직임 장 추정은 시간 t상의 참조 영상 중심좌표 x,y를 가지는 세그먼트와 가장 유사한 Δt 상의 영상 세그먼트의 중심 좌표 x', y' 를 찾음으로써 이루어진다. 두 블록간의 유사도 측정은 기본적으로 다음 식(2)와 같다.

$$C(x,y,x',y') = \sum_{m=-\frac{M}{2}}^{\frac{M}{2}} \sum_{l=-\frac{L}{2}}^{\frac{L}{2}} (E_{(x+m,y+l,t-\Delta t)} - E_{(x'+m,y'+l,t)})^2 \quad (2)$$

$E_{(x,y,t)}$: 시간 t 상의 좌표 x, y의 영상 밝기 값

M, L : 영상 세그먼트의 x, y 크기.

블록 정합 기술은 일반적으로 인접 프레임간 특정 물체의 밝기 값이 유사하므로 효율적이긴 하나, 최적의 블록의 크기를 결정하기 어려우며 정합 과정에서 많은 계산이 요구 된다. 이러한 문제점은 특정 흥미점을 추출하여 흥미점들에 대해서만 정합을 함으로서 어느 정도 계산량을 해결할 수 있다.

2.2.3 Optical Flow

생물의 시각 시스템이 이산적 이지만, 광강도(intensity)에 대한 정량화(quantization)는 매우 세밀하기 때문에 거의 연속적인 광강도를 인식하는 것과 같다. 이러한 세밀한 광강도를 가진 두 개의 연속된 영상의 변화를 연속적 흐름으로 나타낼 수 있다. 이러한 연속 정보를 광류(optical flow)라 부른다. 광류는 시간 간격 dt 동안의 움직임에 의해 생긴 변화를 의미하는데 3차원적인 움직임을 2차원 영상에 표현하는 속도장(velocity field)이라 할 수 있다. 순간적인 속도장은 영상면의 모든 점 또는 픽셀에서 2차원의 속도 벡터를 결정하고 뷰 필드(view field)에서 각 점의 움직임은 속도와 방향을 나타낸다. 광류계산은 다음 두 가정에 기반하고 있다[10].

- (1) 물체 표면의 광강도 값은 시간의 변화에 관계없이 일정하다.
- (2) 영상 표면상의 인접해 있는 점들은 같은 형태로 움직인다.

광류에 의한 움직임 분석 기법은 프레임과 프레임사이의 간격이 넓거나, 조명이 변하거나, 노이즈에 의해 일정한 형태를 유지하지 않을 경우, 물체의 형태가 시간이 변화에 따라 변할 경우 등에서는 좋지 못한 결과를 얻을 수 있다.

3. 제안하는 연구방법

최근 연구를 살펴보면 HOG 특징을 기반으로 단일 영상에서 보행자를 탐지하는 연구가 많이 진행 되고 있다. 그러나 단일 프레임을 이용한 방법은 영상에 존재하는 특징 정보를 추출하고 이미 준비되어진 학습 데이터와 비교를 통하여 보행자를 탐지하므로 특징 정보 추출과 학습데이터의 분류에 만족하지 않는 정보가 입력되면 보행자를 탐지하지 못하는 단점을 가지고 있으며 많은 연산을 수행한다는 단점을 가지고 있다. 그리고 다중 프레임 연산을 통하여 보행자를 탐지하는 방법은 다중 프레임에 보행자의 움직임이 없는 경우와 영상의 밝기 정보가 배경과 비슷한 경우 탐지에 실패하는 경우가 발생한다. 따라서 본 논문에서 HOG특징과 다중 프레임을 이용하는 방법의 장점을 채택하여 보행자의 움직임이 없는 경우에도 배경이미지를 생성하여 강건한 보행자 탐지 방법을 제안한다.

제안하는 연구에서는 초기 3 프레임의 영상을 입력받아 HOG 기반 특징을 이용하여 3 프레임에 존재하는 보행자의 위치를 파악하고 각 프레임에 존재하는 보행자의 블록과 배경을 분리하였다. 이렇게 생성된 각각의 블록을 이용하여 보행자와 배경이 겹치지 않는 블록을 초기 블록으로 생성하고 각 프레임에서 겹치는 보행자의 블록은 다음 프레임을 연속해서 입력 받아 배경과 겹치지 않는 영역을 초기 배경에 합성하여 배경이미지를 생성하였다. 본 논문에서는 이렇게 생성된 배경이미지를 이용하여 보행자 탐지에 적용한다. 그러나 초기 생성된 배경이미지를 지속적으로 사용하게 되면 시간의 변화에 따라 현재 입력되는 프레임과 밝기 변화 차이로 프레임 연산에 노이즈가 발생되어 보행자 탐지에 실패

할 수 있다. 이런 경우 본 논문에서는 HOG 기반의 배경 이미지 추출 방법을 다시 적용하여 배경이미지를 재생성시켰다. 그리고 배경이미지와 현재 프레임간 연산에서 영상의 변화가 없는 경우 발생되어지면 배경이미지를 현재 이미지로 변경하여 항상 최적의 배경 이미지를 제공하여 보행자 탐지에 적용하였다. 제안하는 방법으로 배경이미지를 생성하는 과정은 그림 1과 같다.

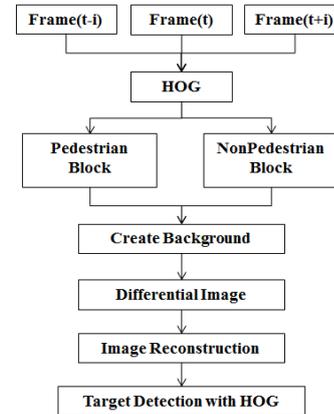


그림 1 제안하는 보행자 탐지 절차
Fig. 1 Proposed method of pedestrian detection

3.1 HOG 기반 보행자 탐지

HOG는 국소영역에 대한 밝기의 크기와 방향을 히스토그램으로 작성된 정보를 특징 정보 벡터로 나타내는 것으로 물체의 형태 정보로 활용된다. 국소영역의 히스토그램을 이용함으로써 밝기의 변화에 영향을 적게 받으며 기하학적인 변화에도 강한 특징을 가지고 있다. HOG 기반의 특징 정보를 추출하기 위하여 입력영상 $I(x,y)$ 에서 x축과 y축의 변화에 대한 그래디언트 정보 θ 와 크기정보 G 을 다음 식(3)을 이용하여 추출한다.

$$G(x,y) = \sqrt{\Delta I_x^2 + \Delta I_y^2} \tag{3}$$

$$\theta(x,y) = \tan^{-1} \frac{\Delta I_y}{\Delta I_x}$$

$$\Delta I_x = I(x+1,y) - I(x-1,y) \quad \Delta I_y = I(x,y+1) - I(x,y-1)$$

입력영상에서 계산되어진 그래디언트의 방향과 크기정보를 이용하여 8×8 픽셀 크기를 하나의 셀(cell)로 나타내고, 이 셀 내부의 정보를 이용하여 밝기 변화에 대한 히스토그램을 작성한다. 여기서 작성된 방향성분은 0~360도를 0~+180도 기준으로 작성해서 나타내었으며, 매 20도씩 9개의 방향에 대한 히스토그램을 작성하였다. 검출된 크기와 그래디언트 히스토그램을 셀 단위로 영상을 구분하여 값을 누적하는 것은 영상에서 예지성분이 분포하는 영역을 구별하기 때문이다. 각 셀에서 작성된 밝기에 대한 그래디언트 히스토그램을 2×2블록으로 정규화를 실시한다. 이는 부분적인 조명의 차이, 각종 영상에서 발생하는 노이즈의 영향을 줄여 준다.

입력영상이 64×128픽셀인 경우 50%의 오버랩을 적용하여

수평으로 7개, 수직으로 15의 블록이 생성되어지며 전체 105개의 블록에 대한 정규화가 이루어지고 정규화 과정은 L2 norm을 이용하였다. 본 논문에서는 하나의 셀이 가지는 bin의 수는 9개이기 때문에 3,780 차원의 특징벡터를 얻을 수 있다. 이렇게 작성된 특징벡터를 선형 SVM을 이용하여 보행자 탐지에 적용하였다. 본 논문에서는 보행자 탐지에 대표적으로 사용되는 INRIA Person Dataset을 이용하였으며 표 1은 INRIA Person Dataset에 대한 설명이다.

표 1 INRIA Person Dataset

Table 1 INRIA Person Dataset

Data Name	INRIA Person Dataset	
Site	http://pascal.inrialpes.fr/data/human/	
Train Data	person image	2,416
	non-person image	1,218
Test Data	person image	1,132
	non-person image	453

3.2 HOG 특징과 다중프레임을 이용한 배경 생성

영상의 다양한 분석으로 움직이는 표적의 위치, 크기, 형태 등의 특징 정보를 추출할 수 있다. 본 논문에서는 HOG 특징과 다중프레임 연산을 기반으로 보행자를 탐지한다. HOG 특징과 다중프레임 연산을 사용하는 이유는 다른 추정 기법보다 간단하고 배경이 프레임과 프레임 사이에서 심하게 변하고 복잡한(clutter)경우에도 효율적으로 적용되어 질 수 있기 때문이다.

영상에서 보행자를 탐지하기 위한 방법으로 배경이미지를 사용하는 방법은 연산의 구조가 타 방법에 비하여 매우 간단하므로 빠른 속도로 보행자를 탐지할 수 있다. 그러나 도로상에서 움직이는 보행자의 움직임이 지속적으로 발생하는 경우 일반적인 다중 프레임 연산으로 배경이미지를 추출하기 어려운 단점이 있다. 그래서 본 논문에서는 HOG 특징을 이용하여 먼저 초기 보행자의 위치를 탐지하고 탐지되어진 보행자 위치 영역과 배경 영역을 블록으로 정하고 다중 프레임을 이용하여 배경이미지를 추출하는 방법을 제안한다.

환경이 미세한 변화에 대응하기 위하여 영상에 노이즈 제거 필터를 적용함으로써 표적이 아닌 배경의 움직임이나 밝기의 변화에 의해 발생한 노이즈를 제거하였다. 연속된 프레임을 이용하여 표적을 추출해내는 기본적인 방법은 식 (4)와 같다.

$$a_{ij}(t) = f_{ij}(t) - f_{ij}(t-i), \quad b_{ij} = f_{ij}(t+i) - f_{ij}(t) \quad (4)$$

$$T_{ij}(t) = a_j(t) * b_{ij}(t)$$

하나의 배경이미지로 차영상 분석에 적용하여 물체를 탐지하는 방법은 배경의 밝기 변화가 심한 환경에서는 적용하기가 어렵다. 그래서 본 논문에서는 현재 탐지하고 있는 보행자 탐지에 3회 이상 실패를 하게 되면 HOG 특징과 다중프레임 연산으로 배경이미지를 재생성 하도록 하였으며 보행자의 탐지가 없는 경우 현재 프레임을 배경이미지로 변경하였다.

기존의 다중프레임 연산을 이용한 배경이미지 추출방법은 프레임간 대응 픽셀 연산을 수행하므로 영상에 존재하는 보행자의 위치가 서로 겹치거나 카메라가 미세한 흔들림으로 생성된 배경이미지에 보행자의 잔상이 남는 고스트(ghost) 현상 및 노이즈가 발생할 가능성이 있다. 본 논문에서는 이러한 고스트 현상과 노이즈를 방지하기 위하여 HOG 특징을 이용하여 영상의 각 프레임에 존재하는 보행자의 위치 정보를 이용하고 보행자 블록과 배경 영역을 분리하였으며 미세한 환경의 변화에 대응하기 위하여 배경 영역을 16×16의 블록으로 분할하여 각 프레임의 배경 블록의 평균값을 구한 뒤 블록 단위로 배경이미지 생성에 적용하였다. HOG 특징과 다중 프레임 연산을 이용한 배경이미지 추출방법은 다음 그림 2와 같다.

```

while( CheckObjectBlock )
{
    if( CheckObjectBlock != PedestrianBlock_PreviousFrame)
        Replace Background_Block Frame(t+i) to BackgroundImage;
    else
        CheckObjectBolck = CheckPedestrianBlock(Frame(t+i));
    if( AllBlockOK ) break;
}
    
```

그림 2 HOG 기반 배경이미지 추출방법

Fig. 2 Background Image extraction using HOG feature

기존의 다중 프레임을 이용하여 배경을 추출하는 방법에서는 보행자의 보행동작으로 인하여 고스트 현상이 발생할 가능성이 있으나 제안하는 블록단위 배경이미지 추출 방법을 사용하였을 경우는 깨끗한 배경이미지를 추출할 수 있다. 본 논문에서는 보행자 탐지를 이용하여 배경이미지를 생성하기 때문에 환경의 변화에 강한 배경이미지를 생성할 수 있다. 제안하는 방법으로 생성된 배경이미지는 그림 3과 같다.

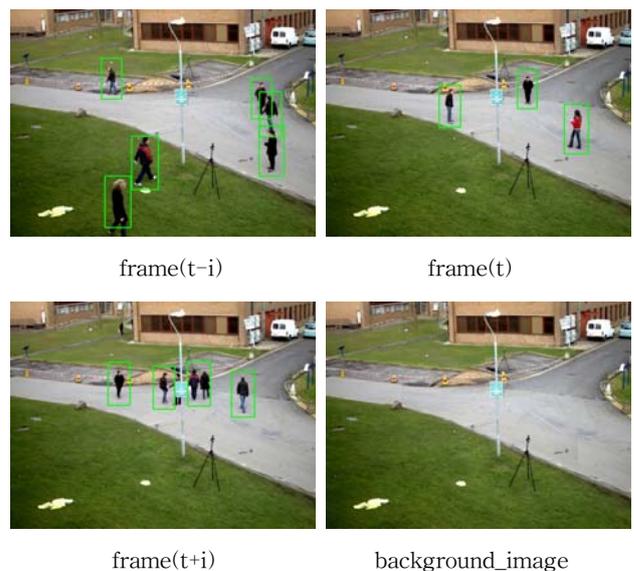


그림 3 생성된 배경 이미지

Fig. 3 Created background image

3.3 보행자 탐지와 블록 영역 확장

도로에서 움직이는 보행자 밝기 정보와 배경의 밝기 정보 차이가 작은 경우 완벽한 형태의 보행자를 추출하기 어려운 경우가 발생되어진다. 이런 경우 한명의 보행자가 다수의 보행자로 판단되는 오인식이 발생할 가능성이 있으며 연산 결과 정보의 추출량이 작은 경우 보행자가 노이즈로 분석되어 탐지에 실패하는 경우가 발생한다.

본 논문에서는 완벽한 보행자를 탐지하기 위하여 프레임 연산이 끝난 이후 탐지되어진 정보의 무게중심을 구한 뒤 3×3픽셀을 하나의 블록으로 선정하고 이웃하고 8방향 영역 확장(region growing)기법을 이용하여 표적의 영역을 확장하였다. 즉, 이웃하는 블록과 비교하면서 보행자 영역을 확장해 나간다. 영역확장 방법은 식 (5)와 같으며 m, n은 필터 크기이며, BITCOUNT는 유효한 픽셀의 임계치로 이 값에 의해서 표적의 표면인지 아닌지를 결정한다.

$$Fimage(x,y) = \begin{cases} 1 & \text{if } (\sum_{i=1}^m \sum_{j=1}^n I(x+i,y+j)) \geq BITCOUNT \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

영역 확장시 블록 단위로 비교함으로써 얻어질 수 있는 장점은 노이즈에 민감하게 반응하지 않는다는 것과 처리속도를 증가시킬 수 있다는 것이다. 영역확장 기법은 그림 4와 같다.

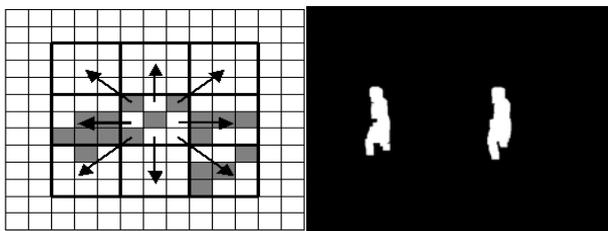


그림 4 8-방향 영역확장 기법
Fig. 4 8-direction region growing

4. 실험결과

제안하는 알고리즘의 성능을 평가하기 위하여 교내 캠퍼스 도로에서 움직이는 보행자를 640×480 크기로 촬영(영상1, 영상2)하여 실험에 사용하였다. 실험에 사용한 시스템은 인텔 i5-3470 CPU, Radeon HD 7800 그래픽카드, Visual Studio 2012와 OpenCV 2.4.11를 사용하였으며 입력영상은 초당 30프레임이다. 단일 프레임(HOG)를 이용한 실험에서는 초당 5프레임 정도의 처리속도를 보였으며 다중프레임(차연산)을 이용하는 방법에서는 초당 30프레임 처리속도를 보였다. 제안하는 연구 방법은 초기 배경이미지 생성에 HOG의 특징을 사용하므로 초당 5프레임 정도의 처리속도를 가졌으나 배경이미지 생성 이후 초당 30프레임을 처리하였다.

실험에 의한 결과 단일 프레임(HOG)에서는 영상에 존재하는 특징 정보만을 이용하여 보행자를 탐지하므로 학습자 탐지에 사용되는 보행자 DB의 능력에 따라 보행자 탐지 성

능 차이를 보이는 것으로 알 수 있었다. 본 논문에서는 보행자 탐지에 대표적으로 사용되는 INRIA person dataset을 이용하여 실험 영상에 적용하였다. 그리고 다중 프레임(차연산)을 이용하는 방법은 영상에 보행자의 움직임이 없는 경우 보행자 탐지에 실패하는 것을 확인할 수 있었다. 제안하는 연구 방법은 프레임에 존재하는 보행자의 정보와 배경 이미지를 사용하므로 단일영상(HOG)과 다중 프레임(차연산)을 적용하는 기법에 비하여 우수한 보행자 탐지 성능을 가지는 것을 확인할 수 있었다. 실험 결과는 표 2와 같다.

표 2 실험에 의한 결과
Table 2 Result of simulation

	영상1			영상2		
	성공	실패	탐지율	성공	실패	탐지율
단일프레임 (HOG)	30	7	81.08%	46	8	85.18%
다중프레임 (차연산)	31	6	83.78%	47	7	87.03%
제안하는 방법	34	3	91.89%	51	3	94.44%



그림 5 보행자 탐지 결과
Fig. 5 Result of pedestrian detection

영상의 외곽 영역에 생성된 보행자는 신규로 진입하는 보행자로 판단되어질 수 있으나 영상의 중간 지점에서 다른 보행자와 충돌 및 교차 진행에 의해서 둘 이상의 보행자가 하나의 보행자로 판단되어질 수 있다. 그리고 영상의 진입 반대 방향으로 소멸하는 보행자는 영상의 외부범위로 벗어난 보행자로 탐지에서 제외하였다. 실험결과 다음과 같은 문제점이 발생되었다.

- (1) 둘 이상의 보행자가 교차진행시 하나의 보행자로 병

합되어 탐지된 이후 분리되는 경우.

(2) 둘 이상의 보행자가 교차진행시 하나의 보행자로 병합되어 탐지된 이후 분리되지 않는 경우

(1)의 경우는 보행자의 과거 이동 궤적을 고려하여 계속 탐지할 수 있으나 (2)의 경우는 한명의 보행자는 탐지에 실패하게 되는 경우이다. 그림 5는 제안하는 보행자 탐지 방법으로 다양한 보행자를 탐지한 결과이다.

5. 결론 및 향후연구과제

본 논문은 HOG 기반 특징 추출과 다중 프레임 연산을 이용한 배경이미지 생성으로 보행자 탐지 방법을 제안하였다. 실험결과 단일 영상만을 사용하는 방법과 다중프레임 연산만 수행하는 방법보다 제안하는 방법이 우수함을 알 수 있다. 제안하는 연구 방법은 고정된 카메라와 움직이는 객체 SCMO(stationary camera and moving object)에 관한 연구이다. 향후 연구과제는 MCMO(moving camera and moving object)에 관련한 보행자 탐지 및 추적에 관한 연구가 필요할 것으로 보인다.

본 논문의 적용분야는 교통량 측정, 무인감시 시스템, 지능형 신호등제어 시스템에 적용되어 질 수 있다.

References

[1] Dollar P., Wojek C., Schiele B., Perona P., "Pedestrian detection: A benchmark," Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, pp. 304-311, June 2009.

[2] Enzweiler M., Gavrilu D.M., "Monocular Pedestrian Detection: Survey and Experiments," Pattern Analysis and Machine Intelligence, IEEE Transactions on Volume 31 Issue 12, pp. 2179-2195, Dec. 2009.

[3] Rodrigo Benenson, Mohamed Omran, Jan Hosang, Bernt Schiele, "Ten Years of Pedestrian Detection, What Have We Learned?," Computer Vision - ECCV 2014 Workshops LNCS Volume 8926, pp. 613-627, 2015.

[4] Yao Wentao, Deng Zhidong, "A Robust Pedestrian Detection Approach based on Shapelet Feature and Haar Detector Ensembles," Tsinghua Science and Technology Volume 17 Issue 1, pp. 40-50, Feb. 2012.

[5] Zhiqian Chen, Kai Chen, Chen J., "Vehicle and Pedestrian Detection Using Support Vector Machine and Histogram of Oriented Gradients Features," Computer Sciences and Applications (CSA), 2013 International Conference, pp. 365-368, Dec. 2013.

[6] Ö. Morkaya and S. Korukoğlu, "Pedestrian and Vehicle Tracking with Adaptive Background Subtraction and Adaptive Object Matching By Using Simple Object Features," Journal of Computing, vol. 3, no. 12, pp. 71-75, Dec. 2011.

[7] Lowe, D. G., "Distinctive Image Features from Scale-Invariant Keypoints", IJCV, pp. 91-110, 2004.

[8] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," In Proc. of CVPR 2005, pp. 886-893, June 2005.

[9] Shan Zhu, Kai-Kuang Ma, "A New Diamond Search Algorithm for Fast Block-Matching Motion Estimation," Image Processing, IEEE Transactions on Volume 9 Issue 2, pp. 287-290, Feb. 2000.

[10] S. Ince, J. Konrad, "Occlusion-Aware Optical Flow Estimation," IEEE Transactions on Image Processing Volume 17 Issue 8, pp. 1443-1451, August 2008.

저 자 소 개



서 창 진 (徐 彰 辰)

1999년 부산대 대학원 멀티미디어 석사과정 졸업 2003년 동대학원 박사과정 졸업(공학박사), 현재 상명대 국방정보공학과 조교수

E-mail : cjseo@smu.ac.kr



지 홍 일 (池 弘 壹)

2002년 충북대 대학원 컴퓨터공학과 석사과정 졸업 2007년 동대학원 박사과정 졸업(공학박사), 현재 영동대 자동차소프트웨어학과 조교수

E-mail : jih161@yd.ac.kr