

## 광 흐름과 학습에 의한 영상 내 사람의 검지

도용태<sup>+</sup>

### Human Detection in Images Using Optical Flow and Learning

Yongtae Do<sup>+</sup>

#### Abstract

Human detection is an important aspect in many video-based sensing and monitoring systems. Studies have been actively conducted for the automatic detection of humans in camera images, and various methods have been proposed. However, there are still problems in terms of performance and computational cost. In this paper, we describe a method for efficient human detection in the field of view of a camera, which may be static or moving, through multiple processing steps. A detection line is designated at the position where a human appears first in a sensing area, and only the one-dimensional gray pixel values of the line are monitored. If any noticeable change occurs in the detection line, corner detection and optical flow computation are performed in the vicinity of the detection line to confirm the change. When significant changes are observed in the corner numbers and optical flow vectors, the final determination of human presence in the monitoring area is performed using the Histograms of Oriented Gradients method and a Support Vector Machine. The proposed method requires processing only specific small areas of two consecutive gray images. Furthermore, this method enables operation not only in a static condition with a fixed camera, but also in a dynamic condition such as an operation using a camera attached to a moving vehicle.

**Keywords:** Optical flow, Histograms of Oriented Gradients(HOG), Support Vector Machine(SVM), Video surveillance

#### 1. 서 론

비디오 카메라를 특정 영역의 감시 및 관측(video surveillance and monitoring: VSAM)을 위한 시각 센서로 사용하는 경우가 많이 있다. VSAM에 있어서 가장 중요한 대상은 사람과 차량이다. 전자는 주로 보안의 목적으로, 그리고 후자는 주로 교통 제어의 목적으로 관찰된다. 사람은 크기가 작고, 그 자세와 동작이 유연하며, 의복의 착용으로 컬러가 다채롭게 변화하는 등의 특징으로 인하여 검출과 해석이 차량에 비해 상대적으로 어렵다. 따라서 영상 내 사람의 효과적이고 자동화된 검출을 위한

다양한 방법들이 연구되어 왔다[1].

기존의 카메라 영상 기반 방법들은 우선 사람과 같은 동적인 객체(moving objects)의 영역에 속하는 화소들을 추출한다. 이를 위해서 사용되는 기술로는 차영상(image difference)을 사용하는 방법과 배경차감법(background subtraction)이 대표적이다. 이들 두 기법의 중요한 차이는 전자의 경우 연속된 두 영상 사이의 변화를 살피는 것이고, 후자는 저장되어 있는 배경영상과의 변화를 살피는 것이다. 후자의 시행 결과가 동적 객체의 형태나 특징을 잘 표현한다는 점에서 장점이 있지만, 배경 영상을 환경의 시공간적 변화에 맞추어 갱신하여야 하는 필요성이 따른다. 이 두 기법들은 카메라가 자동차에 장착된 경우와 같이 이동할 때는 모든 화소값들이 변화하기 때문에 적용이 불가능하다. 이에 비해 광 흐름(optical flow)은 임의의 윈도우(window)를 두 개의 연속된 영상에서 찾는 것으로 동적인 촬영 환경에서도 객체의 구분이 가능하다[2]. 이는 인간이 움직이는 중에도 다양한 동적 객체를 감각하는 능력을 가진 것과 비견될 수 있다. 최근 중요성이 증대되고 있는 차량의 자율주행이나 전자식 운전자 보조 (advanced driver-assistance systems: ADAS) 기술에서는 다양한 센서가 사용되고 있는데[3], 특히 이동하는 차량 전방의 사람을 자동으로 검출하는 것은 매우 중요한 주제이다.

대구대학교 전자전기공학부 전자제어공학전공(Division of Electronic Control Engineering, School of Electronic and Electrical Engineering, Daegu University)

201, Daegudae-ro, Gyeongsan-si, Gyeongsangbuk-do, 38453, Korea.

<sup>+</sup>Corresponding author: ytdo@daegu.ac.kr

(Received: Apr. 29, 2020, Revised: May. 25, 2020, Accepted: May. 26, 2020)

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<https://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

만약 카메라를 자동차에 장착하여 자율주행이나 ADAS를 위한 전방 관측의 용도로 사용할 경우, 카메라의 영상 내 화소들은 모두 이동 특성을 가진다. 이 때 광 흐름의 분석이 유용할 수 있으나, 윈도우의 정합(matching)을 위한 탐색 과정으로 인한 계산량이 많고, 큰 질감(texture)이나 특징이 없는 영역에서는 제대로 정합이 이루어지지 않는 문제가 있다.

본 논문에서는 카메라 영상의 관측 영역 내에서 사람을 검출하기 위한 단계별 방법을 제시한다. 이 때 카메라는 고정되어 있거나, 혹은 움직이는 상황을 모두 감안하였고, 이 두 가지 경우에 모두 유용한 광 흐름 기술을 활용한다. 광 흐름의 계산량을 줄일 수 있도록 일차원의 검출선 (detection line)을 영상 내 적절한 위치에 지정하고, 이 검출선에서 주목할만한 변화가 감지될 때만 주변의 제한된 영역에서 광 흐름을 찾는다. 또한 광 흐름 특징만으로는 사람의 유무를 확정하기 어렵기 때문에 Histograms of Oriented Gradients(HOG) 특징[4]을 기반으로 학습한 Support Vector Machine(SVM)을 사용하여 최종적으로 사람의 관측 영역 내 존재를 판정한다. 제안하는 방법은 간단하지만 효율적이고 체계적인 판단이 가능하다.

본 논문의 구성은 다음과 같다. 먼저 2장에서 여러 가지 관련된 기법들을 간단하게 소개하고, 이어 3장에서는 제안하는 기법의 세부 내용을 서술한다. 4장에서는 실험의 결과를 보이며, 마지막으로 5장에서 결론을 제시한다.

## 2. 카메라 영상에서의 사람 검출 기법

본 논문에서 제안하는 방법은 단계별 처리에 의한다. 그림 1은 본 논문의 방법을 요약하여 보여준다. 즉, 영상 내 관심 영역(region of interest: ROI)에서 사람의 출현을 초기에 검출하기 적절한 위치에 검출선(detection line)을 설정하고, 카메라의 연속 출력 영상에서 그 검출선의 화소들을 관찰한다. 이 경우 영상의 기본적인 처리 비용은 매우 낮다. 일단 검출선에서 어떤 주목할만한 변화가 검출되면, 검출선 주변에 설정된 제한된 영역 내에서 두드러지는 점들을 추출하고, 그 특징점들의 광 흐름을 두 연속 영상으로부터 계산한다. 광 흐름 계산의 결과가 어

떤 동적 객체의 존재를 시사하면, HOG 특성을 검출하고, SVM에 의하여 사람인지를 최종 판정한다. 2.1절과 2.2절에서는 이와 같은 단계별 처리 과정 중에 활용되는 기존의 주요 기법을 간략히 서술한다.

본 논문의 방법은 세 가지 측면에서 기여하는 바가 있다. 먼저, 체계적이고 효과적인 단계별 처리 구조를 제안하였다. 둘째로, 최초의 검출은 일차원적인 화소열의 처리에 의하기 때문에 매우 간단하다. 셋째로, 카메라가 고정되어 있는 경우나 카메라가 움직이는 경우 모두에서 적용이 가능하다.

### 2.1 특징점 추출과 광 흐름의 계산

VSAM이나 ADAS를 위해 사용된 카메라 영상에서 차량이나 사람과 같은 동적 객체를 검출하는 것은 그 사용 목적을 고려할 때 중요하다. 또한 동적 객체를 검출하는 것은 추적이나 해석과 같은 보다 고차원적인 영상 처리 및 컴퓨터 시각 작업의 기초가 된다. 예를 들어 사람의 행위 영상을 분석함으로써 해당 상황을 자동으로 분류하는 작업[5]은 인공지능적 기능을 갖춘 시각 센서 시스템의 개발에 유용할 수 있다.

영상에서 움직임을 검출하기 위한 기존의 기법들은 크게 네 가지로 분류될 수 있다[6]. 즉, 상관(correlation) 기반 방법, 에너지 기반 방법, 파라미터 모델 기반 방법, 그리고 차영상 기반 방법이다. 광 흐름은 차영상에 기반한 방법들 중의 하나로서, 영상의 연속된 프레임에서 빛의 흐름 정보를 활용해 움직이는 대상을 파악한다. 이 때 두 가지의 가정을 사용하는 데, 그 첫 번째는 연속된 두 영상에서 객체가 이동하더라도 해당 화소의 값은 변하지 않는다는 것이다. 동영상을 획득할 때 짧은 시간 동안에 다수의 영상 프레임을 획득하게 되면 연속 영상들 사이의 환경 변화는 실제로 크지 않기 때문에 이 가정은 대체로 유용하다. 두 번째로, 사용된 두 영상 프레임에서 동적 객체의 움직임은 작다는 가정이다. 이 가정은 특히 사람처럼 상대적으로 움직임이 느린 대상의 경우 연속 영상 간 짧은 시간 간격을 고려할 때 유효하다.

광 흐름 계산법에는 여러 가지가 있지만, Lucas와 Kanade가 제안한 방법이 널리 사용되고 있다[7]. 기존의 광 흐름 방법은 작은 윈도우를 사용하기 때문에 이 윈도우 보다 큰 움직임이 발생하였을 경우 움직임을 계산하지 못하는 단점이 있었는데, Lucas-Kanade 방법은 이러한 단점을 피라미드 기법을 사용해 개선하였다. 광 흐름 속도 벡터는 국부 최소자승법(local least squares)으로 계산하는데, 영상 내 화소값이 주변부와 두드러지게 다른 특징점을 검출하여 사용한다. 이때 특징점은 Harris와 Stephens가 제안한 코너 검출기(corner detector)[8]가 자주 사용된다. 이른바 Harris 코너 검출기는 기존 Moravec의 코너 검출기를 개선한 것으로 다음의 절차에 의한다[2]: (i) 컬러 영상을 명암(gray) 영상으로 변환, (ii) 공간적 미분 계산, (iii) 구조 tensor 셋업, (iv) Harris 응답 계산, (v) 비최대치 억제.

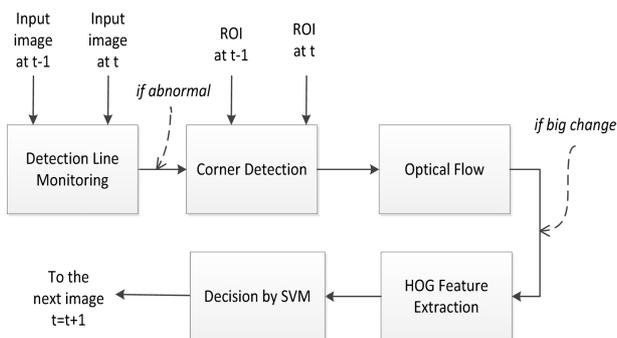


Fig. 1. Block diagram of the proposed method

## 2.2 HOG 특징의 추출과 SVM 학습

Dalal과 Triggs는 영상 내 사람을 검출하기 위하여 기계학습에 기반한 방법을 제안한 바 있다[4]. 사람 영상 데이터 베이스(DB)를 이용하여 HOG 특징을 추출하고, 이를 SVM의 학습에 이용하였다. 학습된 SVM은 이후 카메라로부터 입력되는 새로운 영상에서 사람을 판정하는데 사용하게 된다. 이 방법은 정확도가 높아 널리 사용되고 있다.

HOG 특징 추출은 다수의 영상처리 단계를 통하여 얻어지는데, 셀(cell), 블록(block), 그리고 윈도우(window)의 단위로 처리한다. 먼저 셀은 8×8 화소 배열로 구성되며, 화소값 기울기에 대해서 방향과 크기를 계산한다. 계산된 방향은 0~ $\pi$  각도의 범위에서  $\pi/9$  각도의 간격으로 각각의 저장소에 할당하여 히스토그램을 얻는데, 계산된 크기를 고려하여 가중치를 준다. 블록은 셀보다 상위 단위이며 2×2의 셀 배열로 구성된다. 따라서 하나의 블록에서 총 36개의 히스토그램을 얻을 수 있으며, 이들 히스토그램은 정규화한다. 이 보다 상위 단위인 윈도우는 128×64 크기의 화소 배열로 구성된다. 직립해 있는 사람의 키와 폭은 영상 내에서 이 윈도우의 크기와 근사하다고 가정한다.

윈도우 내에서 블록을 중첩하며 이동하면서 방향별 히스토그램 값을 계산한다. 만약 블록 간에 50% 중첩을 가정하면, 총 15×7=105개의 블록에 대해 계산이 이루어지므로, 최종적으로는 3780개의 값을 가지는 형태 서술자 벡터(shape descriptor vector)가 얻어진다. 블록 중첩의 비율을 낮추면 처리 시간을 줄일 수 있으나 대신 검출의 정확도는 낮아질 수 있다. HOG는 한 장의 영상만을 사용하기 때문에 VSAM 장치뿐 아니라 스마트 자동차의 안전 주행을 위한 센서장치에도 활용될 수 있다.

한편 SVM은 Vapnik과 Lerner가 일찍이 제안한 특징기반 개체 분류를 위한 기계학습 기법이다[9]. 개체들을 두 종류로 구분하는데, 두 그룹 사이의 최대 간격에서 초평면(hyper-plane)을 찾는다. HOG 기법의 경우, 3780개 특징을 기반으로 SVM에 의하여 사람과 사람이 아닌 영상을 구분한다.

HOG와 SVM을 이용하는 방법은 검출의 정확도가 높다고 알려져 있다. 그러나 HOG 특징 벡터의 전 영상에서의 추출을 위한 계산량이 많아서 실시간 활용에 제약이 되고 있다. 이에 대한 해결책으로 Xu 등은 영상 내 에지(edge)의 양에 기반하여 선택적으로 HOG와 SVM을 적용하는 방법을 제안한 바 있으며 [10], Do와 Ban은 이를 더욱 정교화하여 검출의 정확도를 개선한 바 있다[11].

## 3. 제안하는 방법

본 논문에서는 카메라의 연속 영상 내에서 사람을 검출하기 위한 효율적인 체계를 서술한다. 제안하는 방법은 광 흐름과 HOG 특징에 기반하기 때문에 고정된 카메라와 이동하는 카메

라 모두에 적용 가능하다. 본 절에서는 이들 두 센싱 환경에서의 처리 방법을 각각 서술한다.

### 3.1 고정된 카메라를 사용한 검출

VSAM의 용도로 특정 지역을 향하여 고정된 카메라를 가정한다. 대체로 관측 지역에는 사람들이 다니는 구역(예를 들어 보로나 통로)이 있으며, 이 구역을 영상 내 감시 영역으로 특정한다. 그리고 그 영역에 사람이 출현할 때 초기에 검출할 수 있도록 영역의 적절한 위치에 검출선(detection line)을 지정한다. 이 경우 카메라 영상의 처리는 그 검출선 상의 화소들에 대해서만 실시하기 때문에 기본적인 계산 비용은 매우 낮다. 만약 사람 등의 동적 객체가 출현하면 검출선 화소값들이 크게 변화하게 되며, 적절한 판정식을 사용하면 이를 검출할 수 있다.

검출선을 지나는 사람을 효과적으로 검출하기 위한 판정식으로 우리는 식 (1)을 사용하였다. 식 (1)은 현재 시간  $t$ 의 명암 영상(gray image)  $I_t$ 에 설정된 검출선에 존재하는 화소들의 밝기값을 합하여  $S_t$ 를 계산한다. 이 방법은 매우 간단하면서도 검출선 상에서 일어나는 다양한 변화를 잡음에 강건하게 잘 검출할 수 있다.

$$S_t = \sum_{(x,y) \in d} I_t(x,y) \quad (1)$$

여기서  $(x,y)$  좌표의 화소들은 검출선  $d$ 상에 있다.

검출선 상의 화소합은 다음 식 (2)에 의하여 현저한 변화가 있을 때 Flag\_Detection을 set한다.

$$\begin{aligned} \text{Flag\_Detection} &= 1 \text{ if } |S_t - D_t| > \delta, \\ &= 0 \text{ otherwise.} \end{aligned} \quad (2)$$

여기서  $D_t$ 는 현재 영상에서 비교를 위한 기준값이고,  $\delta$ 는 임의의 임계치(threshold)이다. 기준값은 식 (3)과 같이 연속 영상에서 주변 상황의 변화에 따라 적응적으로 갱신할 수 있다.

$$D_t = \alpha S_t + (1 - \alpha) D_{t-1} \quad (3)$$

여기서  $\alpha$ 는 가중치로 [0,1]의 범위를 가진다.  $\alpha$ 가 1에 가까울수록 최근의 관찰값을 활용하는 특성을 가지며, 0에 가까워질수록 기존의 기준치를 유지하는 특성을 가지게 된다.

만약 Flag\_Detection이 set되었다면, 이는 검출선 화소들에 주목할 무엇인가에 의하여 큰 변화가 발생하였음을 의미한다. 따라서 주변에서 광 흐름을 계산하여 이동 객체의 존재 가능성을 검증한다. 먼저 검출선 인근의 관심 영역(region of interest: ROI)에서 Harris 코너 검출기를 기반으로 특징점들을 찾고, 그 특징점들에 대해서 광 흐름을 계산한다. 단조로운 배경에서는 특징점들이 거의 검출되지 않으며, 질감이 있는 배경이라 하더라도 지정된 영역에서의 특징점들의 분포는 미리 알고 있기 때문에 사람 등의 이동 객체의 진입에 의하여 특징점 분포의 변화가 발생하면 이를 쉽게 감지할 수 있다.

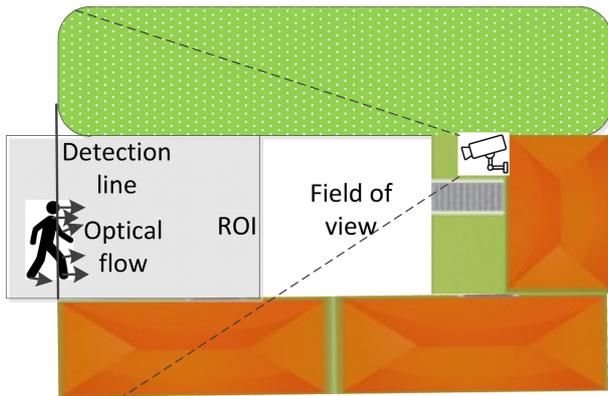


Fig. 2. Human detection using a stationary camera.

검출선에서의 판정이 특징점 검출과 광 흐름 계산을 유발하고, 이로부터 사람의 존재가 추정될 때는 ROI 내 광 흐름 벡터들이 밀집한 부분에서 HOG 특징을 추출한다. 그리고 HOG 특징은 SVM을 사용하여 분류한다. 그림 2는 고정된 카메라에 의한 사람 검출 환경을 개괄적으로 보여 준다.

### 3.2 움직이는 카메라에 의한 검출

본 논문에서 제안하는 방법은 자동차에 부착된 카메라와 같이 동적인 센싱 환경에도 적용이 가능하다. 우선 카메라가 부착된 차량의 자기 주행 차선(ego-lane)을 검출한다. 우리는 간편함과 검출의 강건함의 측면에서 유리한 [12]의 방법을 사용하였다. 즉, 차량의 전방에 설정된 다수의 스캔 라인을 따라 횡방향 검색하면서 임계치를 적용하여 두드러진 밝기 변화를 검출하고, 여기에 허프 변환(Hough Transform)을 적용하여 차량 좌우의 직선을 찾는다.

찾아진 주행 차선을 좌우 경계로 하여 카메라 전방의 일정한 거리에 횡방향으로 검출선을 지정한다. 만약 검출선에서 주목할 만한 변화가 감지되면, 사람이 전방에 있을 가능성을 가정하고 앞의 식 (2)에서와 같은 방식으로 Flag\_Detection을 set한다. 다만, 정적 카메라 촬영 환경과는 달리 차량에 부착된 카메라는 진동이나 주행 차선의 부정확한 검출 등으로 인하여 검출선의 폭이 지속적으로 변화하게 된다. 따라서 정적인 환경에서 사용한 식 (1)을 동적 카메라 촬영 환경에 적용하면 사람이 없더라도 움직임 중에 값의 변화가 자주 발생한다. 그러므로 카메라가 차량에 부착된 경우에는 식 (1) 대신에 아래의 식 (4)와 같이 탐색선 상에 존재하는 화소값의 분산(variance)  $V$ 를 계산하였다. 즉, 일반적인 도로 표면은 대체로 비슷한 화소값의 분포를 가지지만, 사람이 있으면 이로 인하여 주목할만한 화소값의 변화가 발생한다.

$$V_i = (1/n) \sum_{(x,y) \in d} (I_i(x,y) - \overline{I_i(x,y)})^2 \quad (4)$$

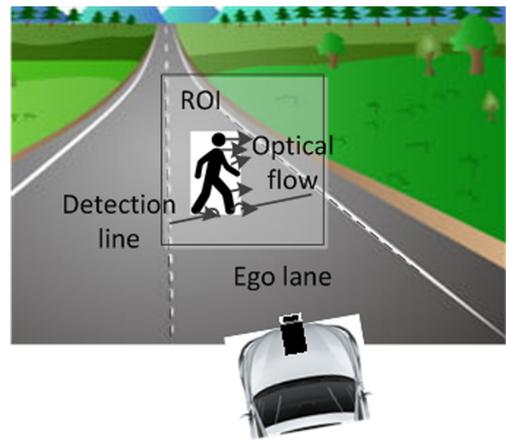


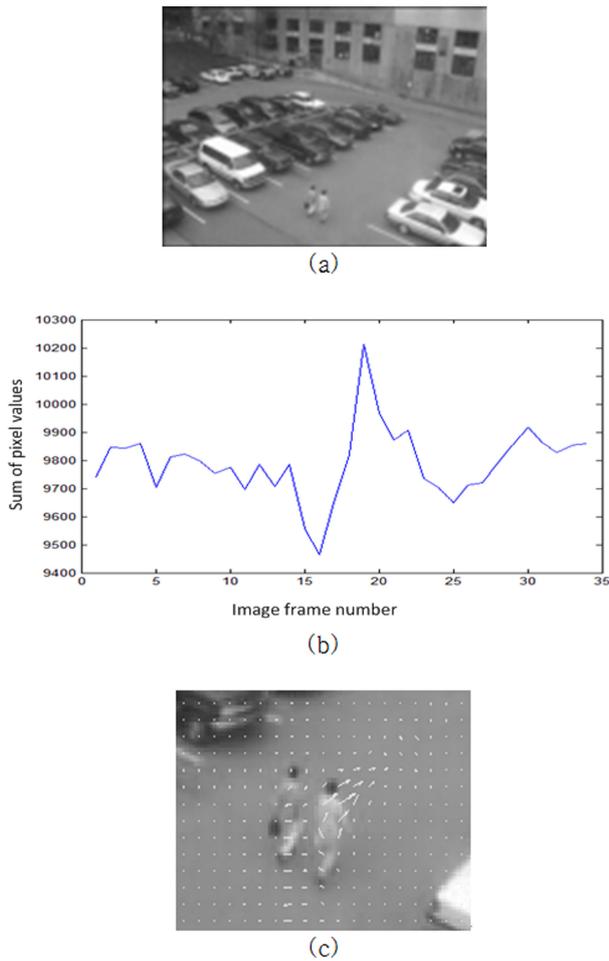
Fig. 3. Human detection using a camera attached to a moving vehicle.

여기서  $\overline{I_i(x,y)}$ 는 검출선 상에 있는  $n$ 개 화소들 밝기값의 평균이다. 자동차에 장착된 카메라 전방의 검출선을 이용할 때, 사람의 경우 발이나 다리가 우선 검출되므로, 광 흐름은 그 검출선의 위쪽을 중심으로 계산한다. 대체로 사람이 없는 경우 이 영역은 도로 표면에 해당하므로 Harris 코너 검출기에 의해 얻어지는 특징점들의 숫자가 많지 않으며, 광 흐름 벡터도 카메라의 이동에 의한 것으로 예측이 가능하다. 따라서 특징점들의 분포와 광 흐름 벡터의 변화는 사람이 존재하는 가능성이 크다는 것을 의미한다. 이후 절차로 HOG 특징 벡터를 추출하고, 이를 기반으로 SVM을 이용하여 최종 판정함은 정적인 카메라의 경우와 동일하다. 그림 3은 이상과 같은 방법을 적용하는 환경의 예를 보여 준다.

## 4. 결과와 예시

본 논문에서 제안하는 방법을 검증하기 위하여 비디오 영상을 직접, 혹은 인터넷 검색을 통하여 획득하고, 이를 실험에 활용하였다. 총 16개의 영상을 사용하였는데, 그 중 8개는 고정된 카메라를 사용하는 정적인 촬영 환경, 그리고 나머지 8개는 자동차에 부착된 카메라에 의한 동적인 촬영 환경이다. 각 환경에서 4개 영상은 사람이 존재하지 않는 경우이고, 나머지 4개는 사람이 존재하는 경우이다.

그림 4는 정적인 센싱 환경에서 검출선에서의 변화에 의하여 광 흐름을 계산한 결과 예를 보여 준다. 사용된 영상은 CMU의 Robotics Institute에서 VSAM[13]의 목적으로 영상 간 시간 간격이 1초 이상 느린 비정기적 속도로 획득한 것으로, 240×320 크기의 ppm 형식이다. 그림 4(a)와 같이 명암 영상으로 변환한 입력에 식(1)을 사용하여 검출선 화소들의 밝기값의 합을 관찰한 결과 그림 (b)와 같은 변화 그래프를 얻을 수 있었다. 카메라가 보행자의 후면에 설치되어 있어, 화소값들은 사람의 검은색 머리가 검출선에 다다를 때 낮아졌다가 사람의 몸통이 지날 때는 밝은 색 의상 때문에 급격히 높아졌다. 이 때 광 흐름을

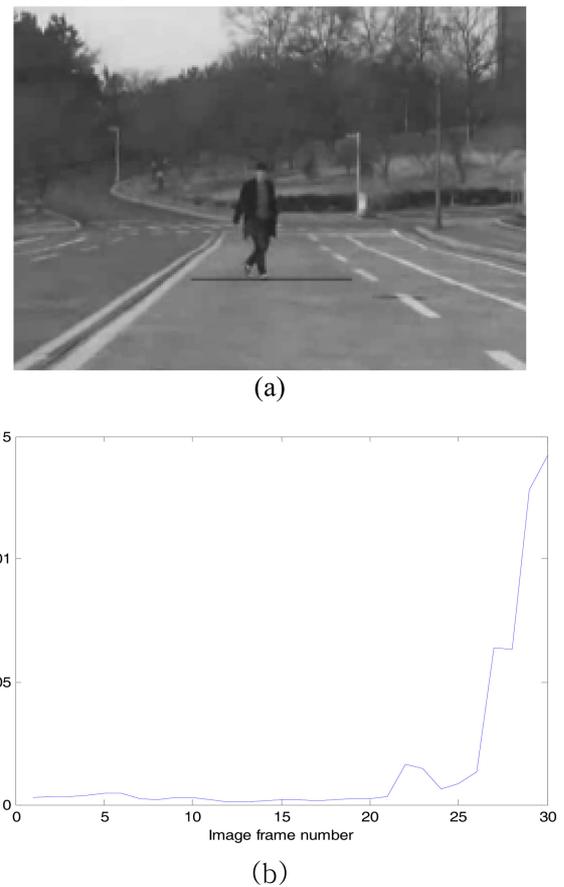


**Fig. 4.** Optical flow of a human image captured by a fixed camera: (a) Input image, (b) Detection line values defined by Eq. (1), (c) Optical flow vectors.

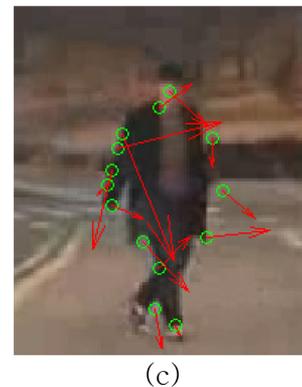
계산하면 사람들 주변에서만 두드러진 광 흐름 벡터가 그림 (c)와 같이 얻어졌다. 다만 결과를 확실히 볼 수 있도록 (c)에서는 사람 부근 영역만 확대하여 보였다. 광 흐름의 계산에는 Lucas-Kanade의 방법 [7]을 사용하였으며, 탐색 윈도우의 크기는 9×9로 하였고, 실험 프로그램은 Matlab을 사용하여 코딩하였다.

카메라를 차량에 부착하고 획득한 영상에 대한 예는 그림 5에 보인 바와 같다. 초당 15 프레임의 속도로 획득한 1080×1920 크기의 컬러 영상에서 480×640 크기의 카메라 전방 영역을 png 형식으로 변환하여 사용하였다. 그림 (a)와 같이 해당 영역 내에 검출선을 설정하였고, 식 (4)를 사용하여 검출선 화소들의 밝기값 분산을 관찰하였을 때, 그림 (b)와 같이 사람이 검출선에 도달한 시점부터 분산이 급격하게 증가하기 시작하였다. 그림 (c)는 검출선 인근 ROI 내에서 검출된 코너점들에서 계산된 광 흐름 벡터들 중 임계치 이상의 것들을 보여 준다.

실험 영상에 본 논문에서 제안한 방법을 적용한 결과는 표 1에서 보였다. 고정된 카메라에 의한 영상 촬영의 경우, 모든 실험 영상에서 정확하게 사람이 있는 경우와 없는 경우를 구분하



**Fig. 5.** Human detection by a moving camera: (a) Input image, (b) Detection line values defined by Eq. (4), (c) Significant optical flow vectors computed at corner points within ROI.

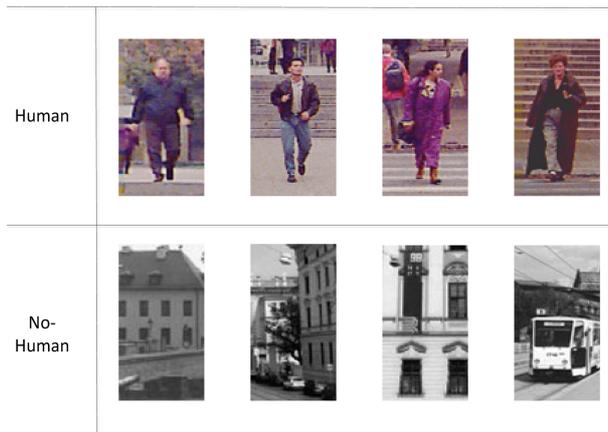


여 검출하였다. 이는 환경의 변화가 정적이기 때문에 사람의 출현으로 인한 변화가 쉽게 감지될 수 있는 이유로 인한 것이다. 한편 차량에 장착된 카메라로 전방의 사람을 검출한 경우에도 사람을 잘 검출하였으나, 사람이 존재하지 않음에도 사람이 있는 것으로 오검출한 경우가 한 번 발생하였다. 전체 실험 결과를 인식 성공률로 수치화하면 15/16, 즉 94%이었다.

사람은 의복의 착용이나 제스처 및 카메라와의 각도에 따라 매우 다양한 형태로 나타나기 때문에 실험을 위해 직접 촬영

**Table 1.** Results using detection line and optical flow for 4 images/set

Camera condition	Image sets	Correct detections / Total number
Stationary	Human	4/4
	No_Human	4/4
Moving	Human	4/4
	No_Human	3/4
Total	All images	15/16

**Fig. 6.** Example images misclassified when using the method of [11], but successfully classified by HOG and SVM.**Table 2.** Recognition results for 30 images/set

Method	Image set	Correct detections
Do and Ban [11]	Human	26
	No_Human	22
HOG+SVM	Human	30
	No_Human	30

한 제한된 수의 영상만으로는 그 성능을 판단하기가 어렵다. 따라서 제안하는 방법의 마지막 단계인 HOG와 SVM에 의한 판정 성능 확인을 위하여 MIT의 Center for Biological and Computational Learning에서 제공하는 Pedestrian Dataset[14]을 사용하였다. 해당 DB의 영상들은 128×64 화소 크기의 ppm 형식으로써, 영상의 크기를 조정할 필요 없이 실험에 바로 사용할 수 있었다. MIT DB에서 선택한 사람 영상 30장과 인터넷 검색을 통하여 획득한 사람이 없는 영상 30장을 실험에 사용하였다. 사용된 영상들은 상당한 복잡성과 다양성을 지니도록 하여 난이도를 높였다. 실험의 결과 HOG와 SVM에 의한 방법은 모든 경우에서 사람의 존재를 정확하게 판정하였다. 얻어진 결과를 이전 연구의 결과인 에지 기반의 방법 [11]과 비교하였을 때 결과는 표 2에 제시한 바와 같았다. 그림 6은 [11]의 방법을 적용하였을 때 오판정된 영상들의 예이다.

## 5. 결 론

VSAM의 용도로 사용되는 CCTV 카메라와 자동차의 자율주행 및 ADAS를 위한 시각 센서용 카메라를 대상으로 영상 내에서 사람을 검출하기 위한 방법을 서술하였다. 제안한 방법은 먼저 영상 내에 지정된 검출선 상의 화소 특성을 연속 영상에서 관찰하고, 특이한 사항이 발생하는 경우에 광 흐름을 통하여 확인한 후, HOG와 SVM에 의하여 최종 감별하는 단계별 접근법을 사용하였다. 각 단계에서 다음 단계를 호출하기 위해서는 임계치의 설정이 필요하며, 연속 영상에서 임계치는 쉽게 결정하고 적응적으로 갱신 가능하다. 제안한 방법은 단지 두 장의 연속적으로 획득한 명암 영상만으로 실행될 수 있다는 점에서 현실적인 장점을 가지고 있으며, 실제 영상을 사용한 실험의 결과 정확하게 사람을 검출할 수 있었다. 이는 실험에서 비교한 이전 기법 [11]이 HOG 기반 기법의 계산량 문제를 극복하기 위하여 전단계의 처리를 실시했을 때 검출의 정확도가 낮아지는 문제가 발생한 것과 비교할 때 개선된 결과이다. 다만 실험에서 사용된 영상이 제한적이므로, 보다 다양한 영상을 여러 가지 환경에서 획득하여 시험함으로써 제안된 방법의 장단점과 제약 사항을 구체화하는 작업이 필요하며, 이는 향후 연구 주제로 남아있다.

## 감사의 글

이 연구는 2017학년도 대구대학교 학술연구비지원으로 수행되었음(과제번호: 20170426).

## REFERENCES

- [1] D. T. Nguyen, W. Li, and P. O. Ogunbona, "Human detection from images and videos: A survey", *Pattern Recognit.*, Vol. 51, No. 3, pp. 148-175, 2016.
- [2] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*, 3rd Ed., Thomson, 2008.
- [3] J. K. Baruah, R. Bera, and S. Dhar, "Ranking of sensors for ADAS—an MCDM-based approach", in *Advances in Communication, Devices and Networking*, R. Bera, S. K. Sarkar, and S. Chakraborty, Eds., Springer, Singapore, 2018.
- [4] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection", *Proc. Comput. Vis. Pattern Recognit.*, Vol. 1, pp. 886-893, San Diego, USA, 2005.
- [5] E. J. Alreshidi and M. Bilal, "Characterizing human behaviours using statistical motion descriptor", *Signal Image Process.*, Vol. 10, No. 1, pp. 15-25, 2019.
- [6] K. Souhila and A. Karim, "Optical flow based robot obstacle avoidance", *Int. J. Adv. Robot. Syst.*, Vol. 4, No. 1, pp. 13-16, 2007.
- [7] J. Bouguet, "Pyramidal implementation of the Lucas Kanade feature tracker description of the algorithm", *Tech-*

- nical Report OpenCV Document Intel Microprocessor Research Labs*, 1999.
- [8] C. Harris and M. Stephens, "A combined corner and edge detector", *Proc. 4<sup>th</sup> Alvey Vis. Conf.*, Vol. 15, pp. 147-151, 1988.
- [9] V. Vapnik and A. Lerner, "Pattern recognition using generalized portrait method", *Autom. Remote Control*, Vol. 24, No. 6, pp. 774-780, 1963.
- [10] G. Xu, X. Wu, L. Liu, and Z. Wu, "Real-time pedestrian detection based on edge factor and Histogram of Oriented Gradient", *Proc. IEEE Conf. on Inf. Autom.*, pp. 384-389, Shenzhen, China, 2011.
- [11] Y. Do and J. Ban, "Edge-based method for human detection in an image," *J. Sens. Sci. Technol.*, Vol. 25, No. 4, pp. 285-290, 2016.
- [12] U.-D. Kim and Y. Do, "Vision sensing for the ego-lane detection of a vehicle," *J. Sens. Sci. Technol.*, Vol. 27, No. 2, pp. 1-5, 2018.
- [13] R. Collins, A. Lipton and T. Kanade, "Introduction to the special section on video surveillance", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 22, No. 8, pp. 745-746, 2000.
- [14] <http://cbcl.mit.edu/software-atasets/PedestrianData.html> (retrieved on Apr. 29, 2020).