

이동 카메라 영상에서 움직임 정보와 Support Vector Machine을 이용한 다수 보행자 검출

Multiple Pedestrians Detection using Motion Information and Support Vector Machine from a Moving Camera Image

임 종 석*, 박 효 진*, 김 옥 현**

Jong-Seok Lim*, Hyo-Jin Park*, Wook-Hyun Kim**

요약

본 논문에서는 이동 카메라 영상에서 움직임 정보와 SVM(Support Vector Machine)을 이용하여 다수의 보행자를 검출하는 방법을 제안하였다. 먼저 연속된 영상의 특징점을 이용하여 카메라 자체의 움직임 보상을 한 후 차 영상과 프로젝트 히스토그램을 통해 움직이는 보행자를 검출한다. 차 영상을 이용한 보행자 검출은 간단한 방법이지만 움직임이 없는 보행자는 검출하지 못하는 단점이 있다. 따라서 이러한 단점을 보완하기 위하여 SVM을 이용하여 움직이지 않는 보행자를 검출하였다. SVM은 보행자 검출과 같은 이진 분류 문제에 우수한 성능을 보이는 것으로 알려져 있다. 하지만 영상 내에 보행자가 서로 인접해 있거나 팔과 다리를 과도하게 움직이는 경우 검출하지 못하는 단점이 있다. 그러므로 본 논문에서는 움직임 정보와 SVM을 이용하여 움직임이 없는 보행자와 보행자가 서로 인접해 있거나 과도한 동작을 취하는 경우에도 강건하게 검출할 수 있는 방법을 제안하였다. 본 논문에서 제안된 방법의 성능을 평가하기 위하여 다양한 실세계 영상을 이용하여 수행하였으며, 그 결과 평균 검출률이 94%, FP(False Positive)가 2.8%로 제안된 방법의 우수성을 입증하였다.

Abstract

In this paper, we proposed the method detecting multiple pedestrians using motion information and SVM(Support Vector Machine) from a moving camera image. First, we detect moving pedestrians from both the difference image and the projection histogram which is compensated for the camera ego-motion using corresponding feature sets. The difference image is simple method but it is not detected motionless pedestrians. Thus, to fix up this problem, we detect motionless pedestrians using SVM. The SVM works well particularly in binary classification problem such as pedestrian detection. However, it is not detected in case that the pedestrians are adjacent or they move arms and legs excessively in the image. Therefore, in this paper, we proposed the method detecting motionless and adjacent pedestrians as well as people who take excessive action in the image using motion information and SVM. The experimental results on our various test video sequences demonstrated the high efficiency of our approach as it had shown an average detection ratio of 94% and False Positive of 2.8%.

Keywords : Support Vector Machine, Difference Image, Ego-motion, Projection Histogram

I. 서론

동영상에서 보행자 검출은 지난 수년간에 걸쳐 컴퓨터 비전 연구에서 많은 관심을 끌었으며 현재까지 그것에 대한 연구가 활발히 수행되고 있다[1-3]. 보행자 검출은 보안 또는 감시 시스템, 이동 로봇 시스템, 시각 장애인을 위한 보행 보조 시스템, 자율 주행 자동차 등 실세계 다양한 응용 분야에서 사용될 수 있다. 카메라가 단단한 물체에 고정되어있는 보안 감시 시스템의 경우는 카메라 움직임이 거의 없기 때문에 간단한 방법으로 보행자를 검출할 수 있다. 그러나 이동 로봇 시스템과 같이 카메라가 움직이는

경우는 카메라 자체의 움직임과 보행자의 움직임이 혼합되어 있으므로 움직이는 보행자를 검출하는 것은 간단하지 않다.

카메라 자체의 움직임을 보상하기 위한 여러 가지 방법들이 컴퓨터 비전 연구가들에 의해 제안되었다[4-6]. 일반적으로 가장 많이 사용되는 방법은 연속된 두개의 영상 좌표 시스템 간의 변환 추정을 이용한 전역 움직임 추정 방법이다. 이렇게 추정된 움직임 정보를 이용하여 이전 영상을 보정한 후 차 영상을 수행하여 보행자를 검출한다. 다수의 보행자를 검출하는 기존의 방법에는 상관 기반 방법(correlation based method)[7], 광류 기반 방법(optical flow based method)[8], 성분이나 형태 기반 방법(component or shape based method)[9-11], 움직임 기반 방법(motion based method)[12-13], 차 영상 기반 방법(difference image based method)[14-15], 기계 학습 기반 방법(machine learning based method)[16-18], 부분 검출기 방법(part detector method)[19] 등이 있다.

* 영남대학교 ** 영남대학교(교신저자)

투고 일자 : 2011. 7. 30 수정완료일자 : 2011. 10. 28

계재확정일자 : 2011. 11. 1

상관 기반 방법은 한 영상의 국소 부분과 다른 영상의 국소 부분을 분리하여 상관계수법에 의해 이동 물체(특히 보행자)를 검출하는 방법이다. 이 방법은 카메라 움직임이 있어도 이동 물체를 검출할 수 있지만 물체의 외형이 변하는 유동적인 물체를 검출하기 어려운 단점이 있다. 광류 기반 방법은 밝기의 국부적인 시공간 경사와 속도 사이의 관계를 모델링한 광류 방정식(optical flow equation)을 사용하는 방법이다. 그러나 이 방법은 물체 표면의 광강도 값은 일정해야 하며 주위의 인접한 점들은 같은 형태로 움직인다는 속도 평활화 제약(smoothness constraint) 조건을 만족할 경우 좋은 결과를 얻을 수 있지만, 영상 사이의 간격이 넓거나 조명이 변하거나 잡음에 의해 일정한 형태를 유지하지 않을 경우 물체의 형태가 시간이 변함에 따라 변할 경우에는 좋지 못한 결과를 얻을 수 있는 단점이 있다. 성분이나 형태 기반 방법은 이동 물체의 구성 성분이나 외관 특징에 의존하는 방법이다. 구성 성분은 머리, 다리, 왼쪽 팔, 오른쪽 팔을 주로 사용한다. 이 방법은 움직이거나 정지한 물체도 검출할 수 있지만 이동 물체의 크기가 너무 다양하므로 일정한 크기의 물체만을 검출할 수 있는 단점이 있다. 움직임 기반 방법은 영상 내의 움직이는 물체의 동작 패턴을 이용하는 방법이다. 이 방법은 대상 물체의 형태나 크기 변화에 관계없이 검출할 수 있고 연산이 간단한 장점이 있으나 조명이 변하거나 배경의 미세한 움직임에도 이동 물체를 검출할 수 없는 단점이 있다. 차 영상 기반 방법은 영상간의 그레이 레벨 차를 이용하는 방법으로 배경이 변하지 않는 경우 계산이 간단하고 다양한 유형의 이동 물체를 검출할 수 있지만 움직임이 없는 보행자는 검출하지 못하는 단점이 있다. 기계 학습 기반 방법은 다양한 방법으로 추출한 특징들을 훈련을 통해 학습하여 이동 물체를 검출하는 방법이다. 이 방법은 다양한 샘플 데이터를 수집하고 훈련시키는 과정이 필요하고 훈련한 특정 형태의 물체만을 검출할 수 있는 단점이 있다. 부분 검출기 방법은 사람의 몸을 머리-어깨(head-shoulder)부분, 몸통(torso)부분, 다리(legs)부분으로 구분하고 각각의 부분을 훈련하여 종합적으로 판단하여 보행자를 검출하는 방법이다. 이 방법은 각 부분의 특징들이 우수하게 검출되어야 하는 단점이 있다.

실세계 영상에는 움직이는 보행자뿐만 아니라 움직이지 않는 보행자도 존재한다. 움직이는 보행자는 차 영상을 이용하여 간단하게 검출할 수 있지만 움직이지 않는 보행자는 성분이나 형태 기반 방법, 기계 학습 방법 등을 사용하여 검출해야 한다. 대표적인 기계 학습 방법인 SVM(Support Vector Machine)[18]과 Adaboost[17]은 보행자 검출과 같은 이진 분류 문제에 우수한 성능을 보이는 것으로 알려져 있다. 그러나 보행자가 서로 인접해 있거나 팔이나 다리를 과도하게 움직이는 경우에는 기본적인 방식으로 검출하지 못하는 단점이 있다.

따라서 본 논문은 카메라 자체의 움직임 보상을 수행한 차 영상과 프로젝션 히스토그램을 이용하여 움직이는 보행자를 검출하고, SVM을 이용하여 움직임이 없는 보행자를 검출하는 방법을 제안하고자 한다.

본 논문의 구성은 다음과 같다. II장에서는 카메라 자체 움직임 보상을 수행한 차 영상을 이용한 보행자 검출에 대

해 설명한다. III장에서는 SVM을 이용한 보행자 검출에 대해 상세히 기술한다. IV장에서는 제안된 방법에 대해 다양한 영상을 이용하여 실험을 수행하고 그 결과에 대해 고찰한 후, 마지막으로 V장에서 본 논문의 결론을 맺는다.

II. 차 영상을 이용한 보행자 검출

2.1 카메라 자체 움직임 보상

보통 카메라가 고정된 상태에서 보행자를 검출하기 위한 간단하고 빠른 방법은 연속된 두개의 영상 간의 그레이 레벨 차를 사용하는 것이다. 그러나 만약 카메라가 자동차와 같은 모바일 시스템에 탑재되어 있다면 차 영상 방법을 이용하여 보행자를 검출할 수 없다. 왜냐하면 카메라 자체의 움직임으로 인하여 보행자뿐만 아니라 배경까지 추출되기 때문이다. 일반적으로 움직이는 카메라 환경에는 이동 물체의 움직임과 카메라 자체의 움직임과 같은 두개의 독립적인 움직임이 있다. 이들 두개의 움직임은 하나의 영상 내에 혼합되어 있다. 따라서 보행자와 같은 이동 물체를 효과적으로 검출하기 위하여 카메라 자체의 움직임을 찾아내어 제거해야 한다. 이러한 움직임이 제거된 영상을 이용하여 차 영상을 수행하면 간단하게 보행자를 검출할 수 있다.

실세계에서 완벽하게 카메라 자체의 움직임을 찾아내는 것은 잡음 때문에 어렵다. 설정 카메라 자체 움직임 보상이 완벽할지라도 차 영상은 여전히 단일 카메라로 획득한 영상 특징 때문에 보행자 경계 부분에 구조적인 잡음이 포함되어 있다.

카메라 자체 움직임은 영상간의 특징 추적에 의해 평가될 수 있다[4]. 카메라가 움직일 때 연속된 두개의 영상, I_t (현재영상)와 I_{t-1} (이전영상)는 좌표시스템이 서로 다르기 때문에 직접 움직임 보상을 수행할 수 없다. 따라서 I_{t-1} 의 영상 좌표에서부터 I_t 의 영상 좌표로의 변환 과정을 통한 움직임 보상이 필요하다. 변환 과정은 I_t 의 특징 집합과 대응하는 I_{t-1} 의 특징 집합을 이용하여 수행할 수 있다. 본 논문은 특징 집합 선택을 위해 해리스 코너 검출기(Harris corner detector)[20]를 사용한다. 특징 선택 알고리즘은 영상 (I_{t-1})에서 특징 (f_{t-1})을 생성한다. 다음 영상 (I_t)의 특징 (f_t)에 대응하는 특징 집합을 찾기 위해 Shi-Tomasi[21]가 제안한 특징 추적 알고리즘을 적용한다.

연속된 두 영상간의 특징추적이 완료되면 카메라 자체의 움직임은 변환 모델을 사용하여 추정될 수 있다. 본 논문은 카메라 자체 움직임 추정을 위하여 여러 가지 변환 모델 중에서 양선형 모델(bilinear model)을 사용한다. 양선형 모델은 연속된 영상간의 간격 길이에 상관없이 대부분의 카메라 자체 움직임을 추정할 수 있는 비선형 변환 모델이다. 본 논문의 실험에서 사용된 양선형 모델은 다음과 같다.

$$\begin{bmatrix} f_x^t \\ f_y^t \end{bmatrix} = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 & a_7 \end{bmatrix} \begin{bmatrix} f_x^{t-1} & f_y^{t-1} & 1 & f_x^{t-1}f_y^{t-1} \end{bmatrix}^t \quad (1)$$

여기서 (f_x^{t-1}, f_y^{t-1})은 이전영상에서 생성한 특징들이고, (f_x^t, f_y^t)은 그 특징들을 현재영상에서 추적한 특징들이다. 이것을 통하여

변환 모델의 파라미터를 계산한다. 한편, 변환 모델(T_i)이 주어지면 최소 자승 최적화(least square optimization)에 대한 비용 함수(cost function)는 다음과 같이 정의된다.

$$J = \frac{1}{2} \sum_{i=1}^N (f_i^t - T_{i-1}^t (f_i^{t-1}))^2 \quad (2)$$

여기서 N 은 특징의 수이다. 카메라 자체 움직임 보상을 위한 모델 파라미터는 비용(cost)을 최소화함으로써 추정된다. 그러나 변환 모델이 계산될 때 일부 특징은 이동 물체인 보행자와 관련 있는 특징들이기 때문에 부정확한 변환을 일으킬 수 있다. 따라서 이들 특징들은 다음 조건을 만족한다면 특징 집합에서 제거되어야 한다.

$$|f_i^t - T_{0i-1}^t (f_i^{t-1})| \geq \epsilon \quad (3)$$

여기서 ϵ 은 미리 정의된 임계값이고 T_0 는 전체 특징 집합을 사용하는 초기 추정이다. 그러면 최종 추정 T 가 계산된다.

카메라 자체 움직임 보상을 위하여 영상 I_{t-1} 은 영상 I_t 와 비교하기 전에 변환 모델을 사용하여 변환된다. 즉, 영상 I_{t-1} 에서 해리스 코너 검출기를 이용하여 추출한 특징들을 영상 I_t 에서 추적한 후, 이 특징들을 이용하여 초기 변환 모델을 추정한다. 그리고 식(3)을 이용하여 최적의 변환 모델 추정에 방해가 되는 특징들을 제거하여 최종 변환 모델을 계산한다. 따라서 각 픽셀 (x, y) 에 대하여 연속된 두 영상간의 차 영상은 다음과 같이 최적의 변환 모델을 사용하여 변환한 후 카메라 자체 움직임 보상된 영상을 사용하여 계산한다.

$$I_{diff}(x, y) = |(I_{t-1}(T_{t-1}^{-1}(x, y)) - I_t(x, y))| \quad (4)$$

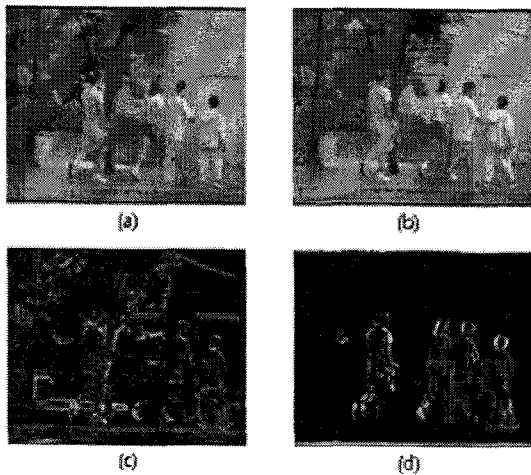


그림 1. 차 영상 결과.

Fig. 1. Results of difference image.

그림 1은 카메라 자체 움직임 보상 없이 수행한 차 영상과 카메라 자체 움직임 보상을 적용한 차 영상을 비교한 것이다. 그림 1(a)는 이전 영상이고, 1(b)는 현재 영상, 1(c)는 카메라 자체 움직임 보상 없이 수행한 차 영상, 1(d)는 카메라 자체 움직임 보상을 적용한 차 영상이다.

2.2 보행자 검출을 위한 전처리 과정

이전 단계에서 생성된 차 영상은 보행자 검출을 효과적으로 수행하기 위하여 이진 영상으로 변환된다. 이진 영상은 임계값에 크게 영향을 받는다. 만약 임계값이 너무 작으면, 영상 내에 불필요한 잡음이 생성된다. 이러한 잡음은 보행자 검출 시 심각한 문제를 발생시킨다. 만약 임계값이 너무 크면, 영상 내에 보행자의 손이나 발과 같은 이동 물체의 작은 일부분만 남겨둔 채 아무것도 남아있지 않을 수 있다. 이와 같은 경우 보행자 검출은 실패하게 된다. 따라서 적절한 임계값의 결정은 매우 중요하다. 본 논문은 많은 실험을 통하여 얻은 값(50)을 사용한다.

변환된 이진 영상에 존재하는 불필요한 잡음을 제거하기 위하여 본 논문은 고립점 검출을 이용한다. 독립적으로 존재하는 하나의 픽셀은 제거되고, 가로방향이나 세로방향으로 연속된 두개의 픽셀만이 존재하는 경우 마찬가지로 제거된다. 이것을 수행하기 위해 3x3이나 3x4, 4x3 마스크를 회전하여 검출한 후 제거한다.

그림 2는 고립점 검출을 사용하여 잡음을 제거한 영상이다. 2(a)는 그림 1(d)에 임계값 50을 사용하여 변환된 이진 영상이고, 2(b)는 (a)에 고립점 검출을 이용하여 잡음이 제거된 영상이다.

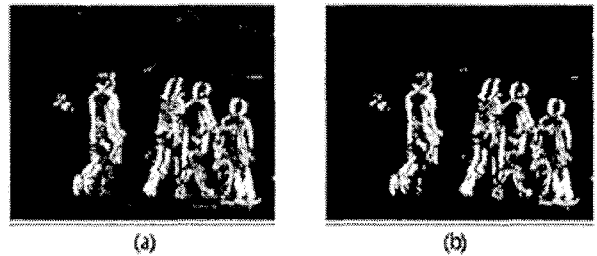


그림 2. 고립점 잡음 제거.

Fig. 2. A noise removal using an isolated point detection.

2.3 프로젝션을 이용한 보행자 검출

전처리 과정을 통하여 잡음이 제거되면 보행자를 검출하기 위하여 프로젝션을 수행한다. 프로젝션은 수평 방향과 수직 방향으로 수행하면서 0보다 큰 픽셀 값의 개수를 카운트하여 히스토그램을 만든다. 다음 식 (5)와 식 (6)은 수평 방향과 수직 방향에 대해 프로젝션 히스토그램을 생성하는 식이다.

$$H_x = \sum_{y=0}^{n-1} I_{nd}(x, y) \quad (5)$$

$$V_y = \sum_{x=0}^{m-1} I_{nd}(x, y) \quad (6)$$

식 (5)와 식 (6)에서 $I_{nd}(x, y)$ 는 잡음이 제거된 이진 영상, x 는 0에서 $m-1$, y 는 0에서 $n-1$ 까지이고, m 은 영상의 세로 크기, n 은 영상의 가로 크기이다.

이렇게 생성된 프로젝션 히스토그램은 인간의 체형 정보를 이용하여 보행자를 검출하는데 사용하고, 또한 SVM을 이용하여 보행자를 검출할 때 스케일에 대한 사전 정보로 이용할 수 있으므로 보행자 검출 시간을 단축시킬 수 있다. 한편, 보행자는 자동차나 동물 등과 같은 이동 물체와는 외관 형태가 다르므로 체형 정

보를 이용하여 간단하게 검출할 수 있다. 그러나 이 방법은 여러 명의 보행자가 서로 인접해 있는 경우 보행자를 분리 검출하지 못하는 단점이 있지만 보행자가 존재하는 영역을 검출할 수 있고 독립된 보행자는 빠르게 검출할 수 있는 장점이 있다.

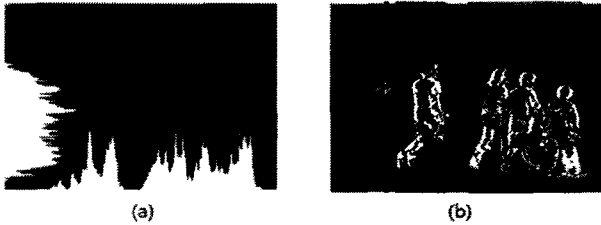


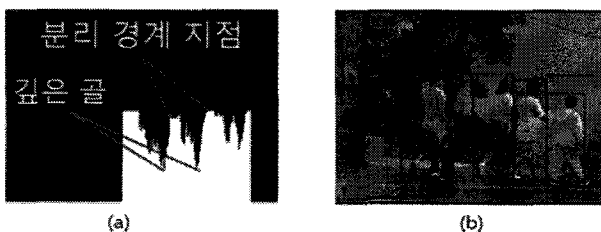
그림 3. 프로젝션 히스토그램과 보행자 검출 결과.

Fig. 3. The projection histogram and detected pedestrians.

그림 3은 프로젝션 히스토그램과 이것을 이용하여 보행자를 검출한 결과를 보여주고 있다. 그림 3(a)에서 가로축은 수직 방향 프로젝션 히스토그램이고, 세로축은 수평 방향 프로젝션 히스토그램을 나타낸다. 그림 3(b)에서 독립된 한 명의 보행자는 검출되었지만 나머지 세 명의 보행자는 서로 인접해 있기 때문에 검출되지 못했음을 알 수 있다. 이와 같이 여러 명의 보행자가 인접해 있는 경우 프로젝션 히스토그램을 이용하면 큰 바운딩 박스가 생성된다. 따라서 큰 바운딩 박스가 생성되면 여러 명의 보행자가 그 바운딩 박스 내에 존재한다는 것을 의미하므로, 이 영역에 대해 추가적인 보행자 검출 과정을 수행해야 한다.

여러 명의 보행자가 서로 인접해 있는 경우 보행자를 검출하기 위해 본 논문에서는 세로 방향 프로젝션을 수행한다. 바운딩 박스의 최상단 지점에서부터 세로 방향으로 0보다 큰 픽셀 값의 존재 여부를 확인한다. 그리고 바운딩 박스의 최상단 지점에서부터 처음으로 나타나는 픽셀 위치까지의 거리를 히스토그램으로 나타낸다. 이때 세로 크기의 1/3에 해당하는 상단부분에 대해서만 프로젝션을 수행한다. 그 이유는 대부분 바운딩 박스의 상단부분에 보행자의 머리가 위치하고 있고, 또한 보행자의 머리를 구분할 수 있으면 보행자를 검출한 것과 같기 때문이다. 이렇게 세로 방향 프로젝션을 수행하여 만든 히스토그램에서 깊은 골 다음에 나타나는 꼭대기(peak) 부분이 두 보행자를 분리하는 경계 지점이 된다. 이렇게 구한 분리 경계 지점을 이용하여 새로운 바운딩 박스를 만들고 바운딩 박스 내에 존재하는 공백을 제거하여 최종적인 보행자를 검출함으로써 서로 인접한 보행자도 검출할 수 있다.

그림 4는 세로 방향 프로젝션 히스토그램과 그것을 이용하여 보행자를 검출한 결과 영상을 보여주고 있다. 그림 4(a)는 세로 방향 프로젝션 히스토그램이고, 그림 4(b)는 최종적으로 검출한 보행자에 대한 바운딩 박스를 현재 영상에 나타낸 결과 영상이다.



(a)

(b)

그림 4. 세로 방향 히스토그램과 보행자 검출 영상.

Fig. 4. The vertical histogram and pedestrians detected image.

III. SVM을 이용한 보행자 검출

3.1 Support Vector Machines

SVM은 학습 자료로 주어지는 두개의 클래스를 구분하는 초평면(hyperplane)을 만드는 선형 분류기이다. 두 클래스에 속하는 선형 분리 가능한 데이터에 대한 이진 분류 문제에서 데이터들을 분류하기 위해 직선으로 표현되는 선형 분류 함수는 무수히 존재할 수 있다. 이와 같이 수없이 많이 존재하는 분류 함수 중에서 분류 함수인 일차 직선을 데이터 점과 만날 때까지 확장한 폭을 선형 분류기의 마진(margin)이라고 정의하면 마진이 최대가 되는 마진의 중심을 이등분하는 초평면이 하나 나타난다. 이러한 최적 분류 초평면을 OSH(Optimal Separating Hyperplane)이라 한다. 그림 5는 선형 공간에서 두 클래스 간에 생성된 최적 분류 초평면을 보여주고 있다. SVM은 학습을 통해 이러한 최적 분류 초평면을 찾는 과정이다.

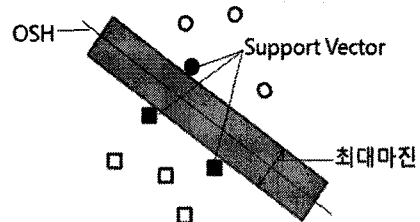


그림 5. 선형 공간에서의 최적 분류 초평면.

Fig. 5. An optimal separating hyperplane in a linear space.

만약 입력 공간이 선형 분리 가능하다고 가정하면 최적 분류 초평면은 고차원 공간에서 식 (7)로 표현된다.

$$ux + b = 0 \tag{7}$$

여기서 x 는 학습 자료 벡터이고, w 는 초평면의 단위 법선 벡터이다. N 개의 벡터 x_i 의 클래스 색인이 $y_i = \pm 1$ 일때, 학습 자료는 이 두 값의 쌍들의 집합 $\{x_i, y_i\}$ 로 주어진다. y_i 는 그 부호에 따라 각 클래스를 구분한다. 모든 학습 자료가 정확하게 분류되기 위해서는 식 (8)의 제약 조건을 만족해야 한다.

$$y_i(ux_i + b) - 1 \geq 0, \forall i = 1, 2, \dots, N \tag{8}$$

w 와 b 만 주어진다면 최적 분류 초평면을 알 수 있으므로 모든 학습 자료가 정확히 어느 클래스에 속하는지 판별 할 수 있고, 마진의 폭도 계산할 수 있다.

주어진 학습 자료로부터 최적 분류 초평면에 대한 최적의 w 와 b 를 찾아내는 것이 중요하다. 모든 학습 자료에 적합한 최대 마진을 가지는 최적의 w 와 b 를 w 와 b 의 공간에서 찾아내기 위해 여러 가지 최적화 기법 중에서 라그랑지안 최적화(Lagrangian optimization) 기법을 이용하여 구할 수 있다. 그러면 이제 라그랑

지만 함수 $Q(\alpha)$ 를 최대화시키는 최적화 문제가 되어 라그랑제 승수 $\{\alpha_i\}_{i=1}^N$ 는 이차 계획법(Quadratic programming)을 이용하여 계산한다.

$$\text{Maximize: } Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (9)$$

$$\text{Subject to: } \sum_{i=1}^N \alpha_i y_i = 0, \alpha_i \geq 0 \quad (10)$$

위의 식 (9)는 입력 데이터의 내적항 $\{x_i^T x_j\}$ 과 출력 데이터 y 에 의해 결정되는 함수이므로 $Q(\alpha)$ 를 최대화하는 α_i^* 를 구한 후 아래의 식 (11)를 사용하여 최적 가중치 벡터 w^* 를 구한다.

$$w^* = \sum_{i=1}^N \alpha_i^* y_i x_i \quad (11)$$

그리고 아래의 식 (12)를 사용하여 최적 바이어스 b^* 를 구한다.

$$b^* = 1 - w^* x_i, y_i = 1 \quad (12)$$

이때 support vector는 $\alpha_i^* \neq 0$ 이고, $\alpha_i^* > 0$ 인 x_i 이다.

최종적으로 임의의 입력 벡터 x 가 주어지면 판별 함수 $f(x)$ 에 의해 분류된다.

$$f(x) = \text{sgn}(\sum_{i=1}^N \alpha_i^* y_i x_i^T x_i + b^*) \quad (13)$$

이 방법은 비선형인 경우 커널 함수를 이용하여 적용할 수 있다. 커널 함수는 비선형 입력 공간을 선형 특징 공간으로 사상시키는 함수이다. 따라서 식 (13)에서 내적은 커널 함수로 대체될 수 있다. 결과적으로 거리 측정은 입력 공간이 선형이든 비선형이든 상관없이 유도될 수 있다.

일반적으로 많이 사용하는 커널 함수는 다음과 같다.

1. $K(x, x') = (x^T x')$
2. $K(x, x') = (x^T x' + 1)^p$
3. $K(x, x') = \exp(-\frac{\|x - x'\|^2}{2\sigma^2})$
4. $K(x, x') = \tanh(\beta_0 x^T x' + \beta_1)$

본 논문에서는 3번 커널을 사용하여 SVM을 구현하여 보행자 검출을 수행하며, $\sigma^2 = 1$ 또는 $\sigma^2 = 4$ 로 설정하여 실험한다.

3.2 SVM을 이용한 보행자 검출 시스템

본 논문에서 구현한 보행자 검출 시스템은 그림 6과 같이 훈련 시스템과 검출시스템으로 구성된다. 다양한 크기와 다양한 형태의 보행자 영상과 보행자가 없는 영상을 훈련 모델에서 훈련시킨다.

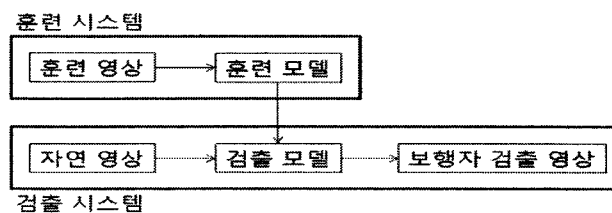


그림 6. 보행자 검출 시스템 구성도.

Fig. 6. Block diagram of pedestrian detection system.

이렇게 훈련이 완료된 모델은 실세계 영상으로부터 보행자를 검출하는 검출 모델로 사용된다. 이러한 훈련 모델과 검출 모델에서 보행자를 분류하는데 사용되는 핵심 엔진은 SVM 알고리즘이다. SVM 알고리즘은 훈련에 의해 여러 영상의 통계적 특성들을 포착할 수 있는 도구로 잘 알려져 있다. 그리고 훈련에 사용하는 영상의 특징들은 HOG(Histograms of Oriented Gradient)[22] 서술자를 이용하여 추출한다. HOG는 국소영역에 대한 밝기의 분포 방향을 히스토그램으로 만들어 이를 특징 벡터로 나타낸 것으로써 물체의 형상 특징을 나타내는데 사용된다. 이렇게 추출된 특징은 SVM을 이용하여 보행자를 훈련하고 검출하는데 사용된다.

훈련 영상의 크기는 25×30에서부터 126×151까지 1.5배의 스케일로 크기를 다양화 하고, 보행자가 있는 영상은 앞모습, 뒷모습, 옆모습 등 다양한 형태의 보행자 영상을 사용한다. 그리고 보행자가 없는 훈련 영상은 다양한 배경에서 임의로 획득한 영상을 사용한다. 본 논문에서는 총 5442개(보행자가 있는 영상 2740개, 보행자가 없는 영상 2802개)의 훈련 영상을 사용하여 훈련한다.



그림 7. 훈련 예제 영상들.

Fig. 7. Examples of training image.

그림 7은 훈련에 사용된 보행자가 있는 영상과 보행자가 없는 영상의 예를 보여주고 있다. 그림 7(a)는 보행자가 있는 예제 영상이고, 그림 7(b)는 보행자가 없는 예제 영상이다.

훈련이 완료되면 실세계 영상을 입력받아 영상의 전 영역에 걸쳐 여러 크기의 스케일을 이용하여 보행자 검출을 수행한다. 이때 프로젝션을 통해 검출된 보행자의 크기와 보행자 영역을 이용하면 보행자 검출 시간을 많이 단축시킬 수 있다.

IV. 실험 및 고찰

본 논문에서 사용된 실험 영상은 실세계의 다양한 시내 영상을 이용하였다. 비디오 카메라를 손으로 들고 천천히 걸으면서 촬영한 동영상을 320×240 컬러 영상으로 프레임 단위로 캡처하였다. 이 영상을 펜티엄 IV 3GHz 프로세서의 윈도우 XP 운영

체제에서 Visual C++ 6.0을 이용하여 구현하였다.

실험은 먼저 프로젝션을 이용하여 보행자를 검출한 후, 검출된 보행자들이 존재하는 영역과 보행자의 크기 정보를 사용하여 SVM을 이용한 보행자 검출 결과를 결합하는 방식으로 진행하였다. 실험에 사용한 영상은 4가지 종류의 서로 다른 배경을 가진 비디오 영상을 가지고 실험하였다.

제안한 방법의 성능을 평가하기 위하여 보행자를 보행자로 정확하게 검출한 보행자 수를 전체 보행자 수로 나눈 검출률 DR(Detection Rate), 잘못 검출한 보행자 수를 전체 보행자 수로 나눈 MR(Miss Rate)과 보행자가 아닌데도 불구하고 보행자로 잘못 검출한 FP(False Positive)를 계산하였다. 표 1은 4개의 비디오 영상을 이용하여 실험한 성능 평가 결과를 나타낸 것이다.

표 1. 성능 평가 결과.

Table 1. Performance Evaluation.

	V1	V2	V3	V4	평균
총 보행자 수	1213	1325	1187	1458	1296
MR	4%	7%	5%	10%	6.5
DR	96%	93%	95%	90%	94%
FP	0%	4%	0%	7%	2.8%

표 1에 나타난 것과 같이 평균 94%의 검출률을 보이고 있으며 FP도 평균 2.8%의 결과를 나타내고 있다. V2와 V4의 결과가 다소 좋지 않은 이유는 보행자들이 너무 가까이 인접해 있어 분리 검출하지 못했기 때문이다.

그림 8은 본 논문에서 제안한 방법으로 보행자를 검출한 결과를 나타낸 영상이다.

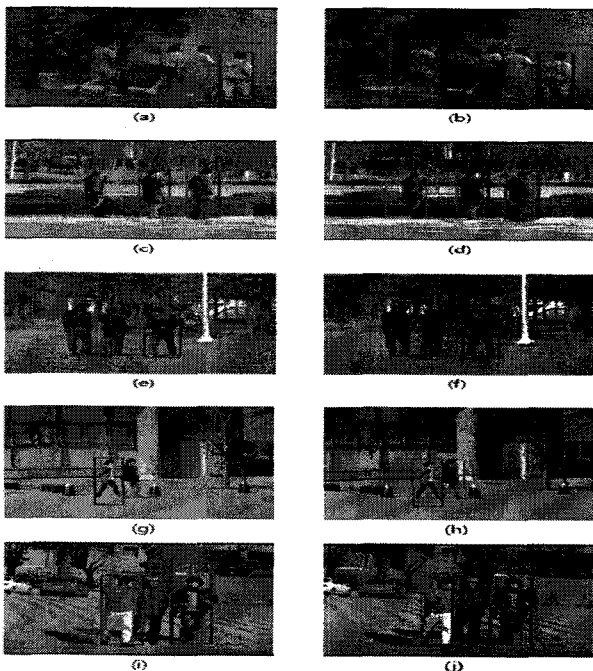


그림 8. 보행자 검출 결과.

Fig. 8. The result of pedestrian detection.

그림 8의 왼쪽 열은 SVM을 이용하여 보행자를 검출한 결과

를 나타낸 것이고, 오른쪽 열은 SVM을 이용하여 보행자를 검출한 결과와 움직임 정보와 프로젝션을 이용하여 보행자를 검출한 결과를 하나로 결합한 것이다. 그림 8에서 빨간색 바운딩 박스는 SVM을 이용하여 검출한 보행자를 표시한 것이고, 녹색 바운딩 박스는 움직임 정보와 프로젝션을 이용하여 검출한 보행자를 표시한 것이다. 그림(i)와 (j)에서 중간에 있는 보행자는 움직임이 거의 없으므로 프로젝션 방법으로 검출이 불가능하지만 SVM 방법으로 검출이 가능함을 알 수 있다. 그림에서 볼 수 있는 바와 같이 SVM을 이용한 보행자 검출이 움직임 정보와 프로젝션을 이용한 보행자 검출 보다 좋지 않은 결과를 보임을 알 수 있다. SVM을 이용한 방식은 보행자가 서로 인접한 경우나 보행자가 팔이나 다리를 과도하게 움직이는 경우에는 검출하지 못하는 단점이 있으나 보행자 검출 시간이 빠르고 움직임이 없는 보행자도 검출할 수 있는 장점이 있다. 움직임 정보와 프로젝션을 이용한 방식은 움직임이 없는 보행자는 검출할 수 없으나 보행자가 팔이나 다리를 과도하게 움직이는 경우와 서로 인접한 보행자도 검출할 수 있는 장점이 있다. 따라서 이 두 가지 방법을 결합함으로써 서로의 단점을 보완할 수 있고 보행자 검출을 더욱 우수하게 수행할 수 있다.

그림 9는 4가지 비디오 영상에 대해 Adaboost와 SVM 알고리즘을 이용한 검출률과 본 논문에서 제안한 알고리즘을 이용한 검출률을 비교한 것이다. 그림에 나타난 것과 같이 본 논문에서 제안한 방법이 다른 방법보다 더 우수함을 알 수 있다.

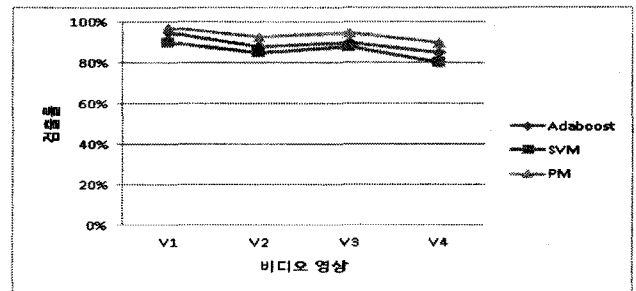


그림 9. 검출률 비교.

Fig. 9. Comparison of detection rate.

V. 결론

본 논문에서는 이동 카메라 영상에서 움직임 정보와 SVM을 이용하여 다수의 보행자를 검출하는 알고리즘을 제안하였다. 움직이는 단일 카메라를 사용하는 경우 이동하는 보행자의 움직임과 보행자 검출을 위해 사용된 카메라 자체의 움직임으로 인해 보행자를 검출하는데 많은 어려움이 있다. 따라서 제안한 알고리즘은 카메라 자체 움직임 보상을 수행하여 차 영상을 통해 독립된 보행자를 검출하고 프로젝션 히스토그램을 이용하여 서로 인접한 보행자를 검출하였다. 그리고 움직임이 없는 보행자를 검출하기 위해 SVM 알고리즘을 이용하였다. 따라서 본 논문은 보행자가 서로 인접해 있거나 보행자가 팔이나 다리를 과도하게 움직이는 경우에도 검출하였고, 움직임이 없는 보행자 또한 검출하였다.

실제 영상을 이용하여 실험한 결과 제안한 알고리즘은 4가지 종류의 서로 다른 배경을 가진 비디오 영상 586개의

프레임에서 평균 94%의 검출률과 평균 2.8%의 FP(False Positive)를 나타내어 그 우수성을 확인할 수 있었다.

향후 본 연구를 기반으로 실시간 보행자 검출을 수행하고 추적하는 연구가 필요하다.

참고 문헌

- [1] L. Zhao and C. E. Thorpe, "Stereo- and neural network-based pedestrian detection," *IEEE Trans. Intelligent Trans. Systems*, Vol. 1, No. 3, pp. 148-154, 2000.
- [2] T. Darrell, G. Gordon, M. Harville, and J. Woodfill, "Integrated person tracking using stereo, color, and pattern detection," *Proc. of the Conference on Computer Vision and Pattern Recognition*, pp. 601-609, 1998.
- [3] B. Leibe, E. Seemann, and B. Schiele, "Pedestrian detection in crowded scenes," *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 878-885, 2005.
- [4] A. Censi, A. Fusiello, and V. Roberto, "Image stabilization by features tracking," *Proc. of the International Conference on Image Analysis and Processing*, pp. 665-667, 1999.
- [5] S. Srinivasan and R. Chellappa, "Image stabilization and mosaicking using the overlapped basis optical flow field," *Proc. of IEEE International Conference on Image Processing*, pp. 420-425, 1997.
- [6] M. Irani, B. Rousso, and S. Peleg, "Recovery of ego-motion using image stabilization," *Proc. of the IEEE Computer Vision and Pattern Recognition*, pp. 454-460, 1994.
- [7] H. Inoue, T. Tachikawa, and M. Inaba, "Robot Vision System with a Correlation Chip for Real-time Tracking, Optical flow and Depth Map Generation," *Proc. of the IEEE International Conference on Robotics and Automation*, pp. 1621-1626, 1992.
- [8] S. Yamamoto, Y. Mae, and Y. Shirai, "Real-time Multiple Object Tracking based on Optical Flows," *Proc. of the Robotics and Automation*, Vol. 3, pp. 2328-2333, 1995.
- [9] A. Broggi, M. Bertozzi, and A. Fascioli, "Shape-based Pedestrian Detection," *Proc. of the IEEE Intelligent Vehicles Symposium 2000*, pp. 215-220, 2000.
- [10] A. Mohan, C. Papageorgiou, and T. Poggio, "Example-based object detection in images by components," *IEEE Trans. Pattern Anal. Mach. Intell.* Vol. 23, No. 4, pp. 156-177, 2001.
- [11] P. Viola, M. J. Jones, and D. Snow, "Detecting pedestrians using patterns of motion and appearance," *Proc. of IEEE International Conference on Computer Vision*, pp. 734-741, 2003.
- [12] H. Mori, N. M. Charkari, and T. Matsushita, "On-Line Vehicle and Pedestrian Detection based on Sign Pattern," *IEEE Trans. on Industrial Electronics*, Vol. 41, No. 4, pp. 384-391, 1994.
- [13] S. A. Niyogi and E. H. Adelson, "Analyzing and Recognizing Walking Figures in xyt," *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 469-474, 1994.
- [14] M. Yachida, M. Asada, and S. Tsuji, "Automatic Analysis of Moving Image," *IEEE Trans. Pattern Anal. Mach. Intell.* Vol. 3, No. 1, pp. 12-20, 1981.
- [15] J. S. Lim and W. H. Kim, "Detection and tracking multiple pedestrians from a moving camera," *Proc. of the 1st International Symposium on Visual Computing*, pp. 527-534, 2005.
- [16] C. Papageorgiou and T. Poggio, "A trainable system for object detection," *Int. Journal of Computer Vision*, Vol. 38, No. 1, pp. 15-33, 2000.
- [17] S. Paisitkriangkrai, C. Shen, and J. Zhang, "Fast Pedestrian Detection Using a Cascade of Boosted Covariance Features," *IEEE Trans. Circuit and System for Video Technology*, Vol. 18, No. 8, pp. 1140-1151, 2008.
- [18] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Cambridge University Press, 2000.
- [19] B. Wu and R. Nevatia, "Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet based Part Detectors," *Int. Journal of Computer Vision*, Vol 75, No. 2, pp. 247-266, 2007.
- [20] C. Harris and M. J. Stephens, "A combined corner and edge detector," *Proc. of the 4th Alvey Vision Conference*, pp. 147-152, 1998.
- [21] J. Shi and C. Tomasi, "Good features to track," *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 593-600, 1994.
- [22] N. Dalal and B. Triggs "Histograms of Oriented Gradients for Human Detection," *Int. Conference on Computer Vision and Pattern Recognition*, pp. 886-893, 2005.



임 종 석(Jong-Seok Lim)

正會員

1991년 2월 계명대 물리학과(이학사)
1996년 8월 대가대 전산통계학과(이학석사)
2004년 2월 영남대 컴퓨터공학과(공학박사)
2005년 3월 ~ 2009년 8월 영남대학교 객원교수
2009년 9월 ~ 현재 영남대학교 시간강사

※주관심분야 : 영상처리, 패턴인식, 컴퓨터 비전



박 효 진(Hyo-Jin Park)

2008년 2월 영남대 컴퓨터공학과(공학사)
2010년 2월 영남대 컴퓨터공학과(공학석사)
2010년 3월 ~ 현재 영남대 LED-IT융합산업센
터연구원

※주관심분야 : 영상처리, 패턴인식, 문서영상



김 옥 현(Wook-Hyun Kim)

正會員

1981년 2월 경북대학교 전자공학과(공학사)
1983년 2월 경북대학교 전자공학과(공학석사)
1983년 2월 일본 쓰쿠바 대학 공학연구과(공학박사)
1983년 3월 ~ 1993년 12월 전자통신연구원

1994년 3월 ~ 현재 영남대학교 컴퓨터공학과 교수

※주관심분야 : 영상처리, 패턴인식, 시각정보처리
