

논문 2005-42SP-2-5

밝기순위 특징을 이용한 적외선 정지영상 내 물체검출기법

(Object Detection in a Still FLIR Image using Intensity Ranking Feature)

박 재 희*, 최 학 훈**, 김 성 대**

(Jae-Hee Park, Hak-Hun Choi, and Seong-Dae Kim)

요 약

본 논문에서는 적외선 영상에서 밝기변화를 예측하기 어려운 일정한 크기의 관심 물체를 검출하기 위하여, 밝기순위 특징과 이를 이용한 물체식별기법을 제안한다. 제안하는 밝기순위 특징은 밝기값의 분포가 균일하도록 영상을 정규화하여 나타낸 것으로, 적외선 영상과 같이 검출대상 물체의 밝기분포를 쉽게 예측하기 어려운 경우에 적합한 특징이다. 제안하는 식별기법은 주어진 후보영역이 검출대상 물체의 학습영상들에 대해 밝기순위가 부합하는 정도를 수치화하여 각각의 후보영역을 물체와 비물체로 식별한다. 제안하는 기법을 통하여 별도의 후보영역 선정과정 없이도 일정한 크기의 관심 물체에 대해 화소단위의 검출결과를 획득할 수 있다. 실험에서는 적외선 자동차 영상을 이용하여 밝기순위특징이 적외선 영상 내 물체식별에 적합함을 보이고, 잡음 및 물체의 크기변화, 기울어짐이 존재하는 상황에서의 검출결과를 보인다.

Abstract

In this paper, a new object detection method for FLIR images is proposed. The proposed method consists of intensity ranking feature and a classification algorithm using the feature. The intensity ranking feature is a representation of an image, from which intensity distribution is regularized. Each object candidate region is classified as object or non-object by the proposed classification algorithm which is based on the intensity ranking similarity between the candidate and object training images. Using the proposed algorithm, pixel-wise detection results can be obtained without any additional candidate selection algorithm. In experimental results, it is shown that the proposed ranking feature is appropriate for object detection in a FLIR image and some vehicle detection results in the situation of existing noise, scale variation, and rotation of the objects are presented.

Keywords : Intensity ranking, Object detection, FLIR image

I. 서 론

물체검출(object detection)이란 주어진 영상에서 찾고자 하는 물체의 존재유무를 확인하고, 존재하는 물체의 위치와 크기를 찾는 일련의 과정을 의미한다. 물체검출은 한 장의 입력영상에서 물체를 찾는 정지영상 내 물체검출과 전후 시점의 영상들을 추가적으로 이용하는 동영상 내 물체검출로 나눌 수 있다. 이 때 정지영상 내

물체검출은 자동물체인식, 로봇비전, 의료영상 등 다양한 분야에서 중요한 연구과제로 인식되고 있으며, 많은 연구가 진행 중이다^{[1][2][3]}.

정지영상 내 물체검출은 입력영상 및 물체의 종류, 응용분야에 따라 그 기법에 많은 차이를 보이지만, 일반적으로 그림 1과 같은 정합기반과정(matching-based approach)을 따른다. 이는 입력영상에서 후보영역을 선정하고, 학습영상들과의 유사도(similarity)를 바탕으로 선정된 후보영역이 물체인지 아닌지를 결정(물체식별)하는 과정을 반복하여 물체를 검출하는 기법이다. 이 경우 물체식별을 위한 특징(feature), 유사도 계산방법, 탐색 전략에 따라 그 검출률 및 검출속도가 크게 달라진다.

* 학생회원, ** 정회원, 한국과학기술원 전자전산학과 전기 및 전자공학 전공
(Division of EE, Department of EECS, Korea Advanced Institute of Science and Technology, Daejeon, Korea)

접수일자: 2004년10월3일, 수정완료일: 2004년12월7일

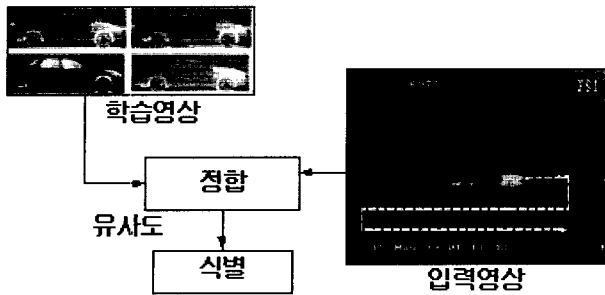


그림 1. 정합기반 물체영역 검출
Fig. 1. Matching-based object detection.

현재까지 얼굴, 자동차, 행인 등 다양한 물체에 대한 많은 정지영상 내 물체검출 기법이 제안되었다. 그러나 이들 기법의 경우, 얼굴검출에서 컬러정보 이용하는 방법 등을 제외하고는 정합기반 검출을 위한 적절한 후보영역 선정 방법을 찾지 못하였으며, 그 결과 검출속도 및 정확도에 있어서 만족할 만한 결과를 얻지 못하고 있다. 또한 일부 기법의 경우, 학습에 적합한 비물체 영상을 획득해야 하는 어려움을 가지고 있다.

본 논문에서 다루고자 하는 전방관측 적외선 영상은 물체로부터 방사되는 적외선의 파장을 밝기값(혹은 컬러값)으로 대응한 영상으로, 물체의 온도에 따라 그 밝기값이 달라진다. 그 결과 같은 물체라 할지라도 영상을 획득할 때의 상태 및 주변 환경에 따라 그 밝기분포가 다양하게 변화하며, 이들 변화는 비선형적이고 예측하기 어렵다^[4]. 이러한 적외선 영상 내에서 관심 물체를 검출하는 문제는 군사적 또는 감시목적으로 그 필요성이 증대되고 있다. 그러나 물체검출을 위한 특징 추출에 관한 연구가 본격적으로 진행되지 않고 있으며, 기존의 CCD (charge-coupled device)영상을 위하여 제안된 기법들에 의존하고 있는 실정이다.

본 논문에서는 적외선 영상과 같이 영상 내 물체의 밝기분포가 다양하게 변화하고, 정확한 윤곽선 획득이 어려운 경우에 적합한 정지영상 내 물체검출 기법을 제안한다. 제안하는 물체검출 기법은 밝기순위 특징(feature)과, 이를 이용하여 고정된 크기의 물체에 대하여 빠르고 정확하게 물체를 식별하는 기법으로 구성된다.

본 논문의 구성은 다음과 같다. 먼저 II장에서는 기존의 물체검출 기법에 대하여 고찰한다. III장에서는 밝기순위 특징을 제안한다. IV장에서는 제안한 특징을 이용한 식별기법을 제안한다. V장에서는 적외선영상 내 자동차 검출과정을 통하여 제안한 기법의 성능을 평가한다.

II. 기존기법 고찰

서론에서 언급한 바와 같이, 정합기반 물체검출과정은 크게 후보영역 선정과정과 선정된 후보영역으로부터 특징을 추출하는 과정, 그리고 선정된 후보영역이 찾과자 하는 물체인지 아닌지를 식별하는 과정(물체 및 비물체 식별과정)으로 구성된다.

1. 후보영역 선정

후보영역 선정과정에서는 적은 계산량으로 영상 내 비물체 영역을 제거하여, 검출대상 후보영역의 수를 줄이는 역할을 수행한다. 선정과정을 통하여 구성된 후보영역 집합은 반드시 물체영역을 포함하여야 하며, 비물체 후보영역은 최대한 적게 포함할수록 좋은 성능을 나타낸다. 이를 위하여 검출 대상의 특성을 잘 파악하여 물체와 가장 비슷한 영역만을 후보영역으로 선정하는 것이 중요하다. 예를 들어, 동영상에서 움직이는 물체를 찾고자 할 경우, 차영상 등을 통한 움직임 정보를 이용하여 후보영역을 선정^[5]할 수 있으며, 컬러영상에서 얼굴을 검출하고자 할 경우, 살색영역 분할과정을 통하여 살색영역을 포함하는 부분만을 후보영역으로 선정^{[6][7]}하기도 한다. 그러나 본 논문에서 다루고 있는 적외선 정지영상 내 자동차 검출의 경우 물체의 컬러정보나 움직임 정보 등을 이용할 수 없다. 이 경우 물체와 같은 크기의 모든 부분영역을 후보영역으로 선정하여야 하며, 그 결과 매우 많은 수의 후보영역에 대하여 물체 및 비물체 식별과정을 반복^[8]하여야 한다. 예를 들어 $320 \times 240 \text{ pixel}^2$ 크기의 영상에서 $70 \times 30 \text{ pixel}^2$ 크기의 물체영역을 검출하고자 할 경우, $(320-70) \times (240-30)$ 개의 식별대상 후보 영역에 대한 식별과정 반복이 필요하다.

2. 물체검출을 위한 특징추출

기존의 정지영상 내 물체검출기법은 사용하는 특징에 따라 밝기정보기반 기법과 형태정보기반 기법으로 분류할 수 있다.

밝기정보기반 기법이란 영상의 밝기 값(intensity)의 분포특성을 이용하는 기법으로 주로 주성분분석법(PCA, principal component analysis) 등의 선형투영기법을 적용하여 사용한다^{[9][10][11]}. PCA는 영상공간의 분포를 가장 잘 표현하는 소수의 기저벡터를 찾아내고, 이 기저벡터들이 생성하는 부공간(sub-space)에 영상을 투영하는 방법이다. 이때 투영오차(projection-error) 또

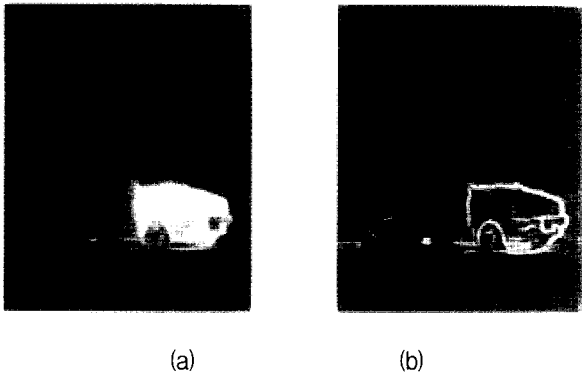


그림 2. 적외선 영상 (a)로부터 획득한 윤곽선 정보(b)
 Fig. 2. Boundary information(b) obtained from a FLIR image (a).

는 부공간 내 학습영상과 표본과의 거리를 이용하여 물체를 식별한다. 이러한 투영기법은 한정된 적은 양의 정보로 물체를 잘 표현할 수 있다는 점에서 효율적인 특징추출기법이라 할 수 있다. 그러나, 기저 벡터의 선형조합으로 영상을 표현하는 방법으로는 적외선 영상과 같이 국부적이고 예측하기 어려운 밝기 변화가 심한 경우 입력영상을 제대로 표현할 수 없다. 또한 많은 경우 후보영역으로 선정된 물체영역과 학습영상 내 물체영역은 서로 정확히 정렬되어 있지 않기 때문에 기저벡터(고유영상)를 이용한 물체 표현은 더욱 어렵다고 할 수 있다.

형태정보기반 기법이란 주어진 후보영상의 형태정보(에지, 모서리 분포)를 이용하여 물체와 비물체를 구별하는 방법이다^{[12][13][14][15]}. 주로 Canny 에지검출기나 Gabor 필터 등을 이용하여 영상 내 물체의 윤곽선을 획득하고, 이를 식별에 이용한다. 이러한 형태정보기반 기법은 조명변화나 환경변화에 따라 영상의 밝기가 심하게 변화하더라도 비교적 변화에 둔감한 특성을 얻을 수 있다. 그러나 적외선 영상의 경우에는 그림 2와 같이 배경과 물체를 정확히 구별할 수 없는 경우가 많고, 많은 경우 물체 내부나 물체 주위에 강한 에지가 존재하기 때문에 신뢰성 있는 윤곽선 획득이 어렵다. 정확하지 않은 윤곽선 정보는 검출실패(false-negative) 또는 오검출(false-alarm)을 야기하며, 그 결과 정확한 물체검출을 위하여 윤곽선 획득에 관한 별도의 연구가 수행되어야 한다.

3. 물체 및 비물체 식별과정

정지영상 내 물체검출은 물체 및 비물체 식별을 위한 모델구성 방법에 따라 학습(learning)기반 기법과 표준형(prototype)기반 기법으로 분류된다.

학습기반 기법은 MLP (multilayer perceptron)나 SVM (support vector machine) 등 정형화 된 구조의 식별기(classifier)를 물체와 비물체로 구분된 예제영상들을 이용하여 훈련시키고 이를 이용하여 물체를 식별하는 기법이다^{[8][11]}. 이러한 학습기반 기법은 주어진 학습영상(training image)에 적합한 판별경계(decision boundary)의 획득을 목표로 하며, 물체의 특징벡터의 통계적 분포를 추정하기 힘든 경우 주로 사용된다. 현재 빠르고 높은 식별률을 가지기 위하여, 식별기 내 제한조건 및 복수의 식별기를 사용하는 기법에 대한 연구들이 진행 중이다. 이러한 학습기반 기법의 적용을 위해서는 충분한 예제영상 획득이 필수적이다. 만약 물체의 예제영상들의 집합이 물체의 변화를 충분히 나타내지 못하거나, 유사물체(clutter) 예제영상이 부족할 경우, 검출률이 떨어지거나 오검출률이 증가하게 된다. 그러나 정확한 판별경계를 얻기 위하여, 유사물체 예제영상들을 어떻게 획득하며, 획득한 예제영상들을 어떻게 선별하여 훈련에 사용할지에 관한 문제는 여전히 난제로 남아있다.

표준형기반 기법은 찾고자 하는 물체를 하나 혹은 소수(小數)의 표준형으로 나타내고 후보영상과 이들 표준형과의 유사도를 이용하여 물체와 비물체를 식별하는 기법이다^[7]. 이 경우 특징벡터들이 분포하는 특징공간 상에서의 유사도의 정의방법, 표준형 패턴의 개수 및 구성방법에 따라 그 성능이 크게 달라진다.

III. 수정된 밝기순위 특징

앞에서 언급한 바와 같이 적외선 정지영상 내 물체검출에서는 적외선 영상에 존재하는 다양한 밝기변화 때문에 안정적인 특징추출이 어려웠다. 본 장에서는 선정된 후보영역이 찾고자 하는 물체인지 아닌지를 확인하기 위하여, 적외선 영상의 다양한 변화에 강인한 수정된 밝기순위 특징에 관하여 제안하고 설명한다.

1. 밝기순위

밝기순위는 각 화소별로 정의되며, 주어진 화소의 영상 내 밝기값 순위를 의미한다. 수학적으로 각 화소에 대한 밝기순위($r(x, y)$)는 영상 내 화소들 중 그 밝기값($i(l, m)$)이 자신의 밝기값($i(x, y)$)보다 같거나 작은 화소들의 비율로 정의되며 식(1)과 같이 계산한다.

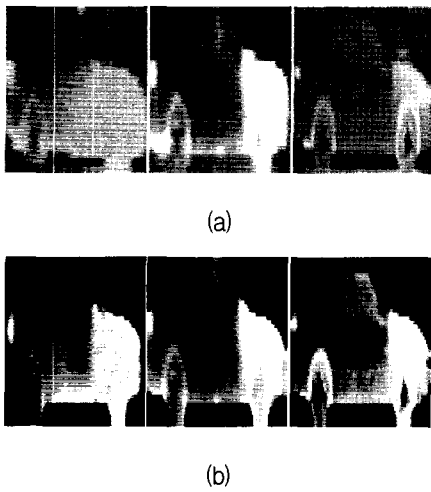


그림 3. 다양한 환경에서 획득한 적외선영상(a)과 대응 순위영상(b)
 Fig. 3. (a) FLIR images obtained in various conditions (b) Intensity ranking image corresponding to (a).

$$r(x, y) = \frac{1}{XY} \sum_{l=0}^{X-1} \sum_{m=0}^{Y-1} u(i(x, y) - i(l, m)) \quad (1)$$

X, Y : 영상크기

$$u(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases} \quad (2)$$

흔히 히스토그램 균등화(histogram equalization)라는 이름으로 입력의 정규화를 위해 널리 사용되고 있는 밝기순위는 적외선 영상과 같이 밝기변화가 심한 영상에서도 그 변화가 작은 특징정보이다. 다양한 환경에서 획득하여 그 밝기변화가 큰 적외선 물체영상들에서 순위영상*을 추출하면 밝기변화가 작아짐을 확인할 수 있다. 적외선 영상의 경우 영상에서 밝기값은 물체의 온도와 관련이 크고, 대부분의 물체에서 열원은 고정된 위치에 있으므로 그림 3에서와 같이 물체영상의 밝기값이 많이 바뀌는 상황이라도, 밝기값의 순위는 열원에 상대적으로 가까운 쪽이 밝은 균일한 특성을 가지게 된다. 또한 밝기가 전체적으로 밝거나 어두워도 이러한 순위영상을 이용하면 입력의 범위가 일정하게 된다.

2. 수정된 밝기순위

밝기순위는 비교적 적외선영상에서의 다양한 밝기변화에 강인한 특징이지만 그림 4의 자동차 후미등(A)과 같이 일부 화소의 밝기가 변할 경우 다른 부분의 밝기

* 밝기순위 값으로 모든 화소의 밝기값을 대체한 영상을 순위영상이라 정의한다.

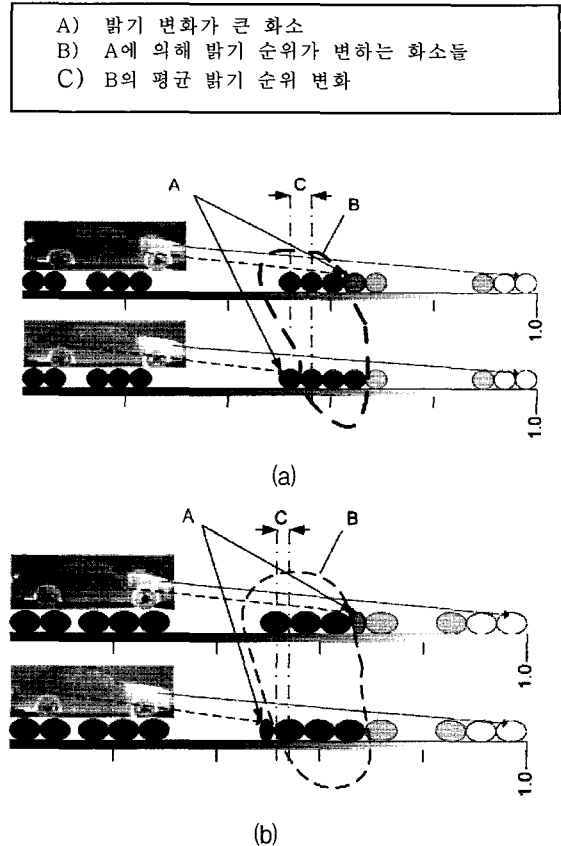


그림 4. 특정화소의 밝기변화에 따른 밝기순위변화 (a) 밝기순위 (b) 수정된 밝기순위
 Fig. 4. Intensity ranking variation caused by intensity change of a pixel (a) Intensity ranking (b) Modified intensity ranking.

순위도 전체적으로 변할 수 있다. 본 논문에서는 이러한 문제점을 해결하기 위하여 학습영상들로부터 각 화소의 위치별로 밝기순위의 항상성을 수치화하고 이를 기반으로 위치별 밝기순위에 신뢰도를 부여한 수정된 밝기순위(modified intensity ranking) 특징을 제안한다.

제안하는 (x, y) 에서의 수정된 밝기순위 $\tilde{r}(x, y)$ 는 자신보다 밝기값이 같거나 작은 화소들의 신뢰도의 합으로 정의된다(식(3)). 이때 $w(x, y)$ 는 (x, y) 에서의 밝기순위의 신뢰도로 그 위치에서 밝기순위가 얼마나 일정한지를 나타낸다.

$$\tilde{r}(x, y) = \sum_{l=0}^{X-1} \sum_{m=0}^{Y-1} w(x, y) u(i(x, y) - i(l, m)) \quad (3)$$

X, Y : 영상크기

$$u(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases} \quad (4)$$

만약 (x, y) 에 위치한 화소의 밝기순위가 비교적 변화가 심할 경우, $w(x, y)$ 에 작은 값을 할당하여 전체

순위에 그 변화가 미치는 영향을 적게 한다. 극단적으로 특정한 임계값을 정하고 밝기순위가 임계값 이상으로 변화할 경우 $w(x, y) = 0$ 으로 정한다면 그 화소는 특징 추출 과정에서 완전히 배제된다.

이러한 신뢰도를 계산하기 위하여 여러 가지 방법을 생각할 수 있지만, 본 논문에서는 그 한 방법으로 학습영상들 내 (x, y) 에 위치한 화소들의 밝기순위의 분산 $\sigma_r^2(x, y)$ 에 기반하여 식(5)과 같이 계산하였다.

$$w(x, y) = \alpha \left(\frac{\exp(-\sigma_r^2(x, y))}{\sigma_{th}^2} \right) \quad (5)$$

α, σ_{th}^2 : 상수

$$\sum_{l=0}^{X-1} \sum_{m=0}^{Y-1} w(l, m) = 1 \quad (6)$$

그림 4는 영상에 따라 밝기순위가 변하는 화소(A)가 존재할 경우, 그에 따라 발생하는 주변화소(B)들의 밝기순위 변동(C)을 나타낸 것이다. 그림 4(a)와 같이 밝기값의 크기를 기준으로 각 화소를 정렬하여 밝기순위를 계산할 경우, 일부 화소의 밝기순위가 변함에 따라 변하는 구간 내 모든 화소의 밝기순위가 변하는 것을 확인할 수 있다. 그러나 그림 4(b)와 같이 수정된 밝기순위에서는 순위변동이 심한 화소의 신뢰도를 작게 부여하여 전체 순위계산에 미치는 영향을 줄임으로써 주변화소의 순위변동 폭을 줄일 수 있다.

또한 수정된 밝기순위의 항상성을 수치화하여 각 화소별 신뢰도를 재계산하고 새롭게 계산된 신뢰도를 바탕으로 수정된 밝기순위를 다시 계산하는 과정을 반복

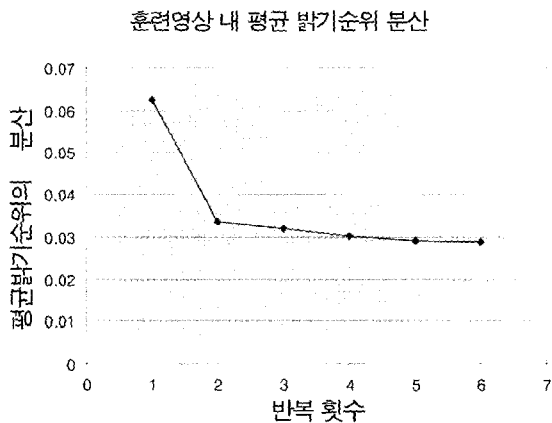


그림 5. 밝기순위 계산과정의 반복에 따른 밝기순위 분산
Fig. 5. Variance of intensity ranking changed by iteration of ranking computation process.

하여 좀 더 안정적인 밝기순위를 획득할 수 있다. 그림 5는 이들 반복과정에 따른 평균 밝기순위 변화량 $(E[\sigma_r^2])$ (식(7))을 측정된 것이다. 그림에서와 같이 수정된 밝기순위 계산과정이 반복됨에 따라 각 화소들의 순위변화를 줄일 수 있음을 확인할 수 있었다. 이때, 약 8회의 반복과정을 거친 후 밝기순위의 분산이 일정 수준으로 수렴하였다.

$$E[\sigma_r^2] = \sum_{l=0}^{X-1} \sum_{m=0}^{Y-1} w(l, m) \sigma_r^2(l, m) \quad (7)$$

IV. 밝기순위 특징을 이용한 물체검출

본 장에서는 제안한 밝기순위 특징의 성질을 이용, 별도의 추가적인 특징추출과정 없이 후보영역의 밝기값을 이용하여 후보영역과 물체와의 유사도를 계산하는 기법을 설명한다.

1. 물체 및 비물체 식별

본 논문에서 검출하고자 한 물체는 그림 3에서 제시한 바와 같이 동일한 위치에서 비슷한 밝기순위를 가진다. 이 경우 밝기순위가 높은 곳의 점들은 순위가 낮은 점들보다 그 밝기값이 커야한다. 이를 이용하여 학습영상으로부터 물체의 밝기순위를 예측하고, 실제 입력이 예측된 밝기순위에 잘 부합하는가를 확인하여 후보영상들 중에서 물체영상을 식별하고자 한다.

우선 같은 물체를 나타내는 크기 정규화된 다수의 학습영상들로부터 각 화소에서의 신뢰도를 얻고, 배경을 제거한 후, 학습영상의 수정된 밝기순위를 획득한다.

그리고 이 밝기순위를 N 개의 등급으로 양자화(quantization)하여, 각 등급(level)이 포함하는 밝기순위 구간을 구한다. 만약 수정된 밝기 순위를 N 개의 등급으로 양자화한 결과, n 번째 등급에 해당하는 밝기순위 구간이 $\tilde{r}_{n, \min} \leq \tilde{r}(x, y) \leq \tilde{r}_{n, \max}$ 와 같다고 할 때, 해당 구간에 속하는 수정된 밝기순위를 가지는 화소들을 모든 n 번째 등급영상벡터 \mathbf{Lv}_n 을 다음과 같이 정의한다.

$$\mathbf{Lv}_n = (lv_n(0,0), lv_n(1,0), \dots, lv_n(X-1, Y-1))^T$$

$$lv_n(x, y) = \begin{cases} w(x, y) & \tilde{r}_{n, \min} \leq \tilde{r}(x, y) \leq \tilde{r}_{n, \max}, (1 \leq n \leq N) \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

본 논문에서는 각 등급의 가중치의 합이 일정하도록

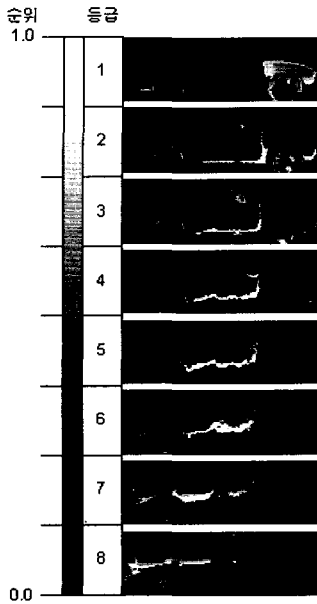


그림 6. 등급영상의 예
Fig. 6. An example of level images.

수정된 밝기순위를 양자화 하였는데 이 경우 n 번째 등급에 해당하는 밝기순위 구간은 $\tilde{r}_{n, \min} = (n-1)/N$, $\tilde{r}_{n, \max} = n/N$ 로 결정되며 각 등급영상벡터 내 화소값의 합은 항상 일정하게 된다. 위의 경우, 각 등급영상벡터와 같은 크기로 정규화된 후보영상의 입력밝기벡터 $\mathbf{I} = (i(0,0), i(1,0), \dots, i(X-1, Y-1))^T$ 와 의 내적만으로 각 등급에 해당되는 후보영상 내 영역에서의 평균 밝기값(m_n)을 구할 수 있다.

$$m_n = \mathbf{L}v_n \cdot \mathbf{I} \quad (9)$$

상대적으로 밝기순위가 높은 점들에서 구한 평균 밝기값 m_a 와 상대적으로 밝기순위가 낮은 점들로부터 구한 평균 밝기값 m_b 가 있을 때, 입력 후보영상 \mathbf{I} 가 물체라면 $m_a > m_b$ 이어야 한다. 만약 밝은 등급영역의 평균 밝기값이 어두운 등급영역의 평균 밝기값보다 더 작다면 물체가 아닐 가능성이 크다. 따라서 본 논문에서는 평균 밝기값의 순서가 바뀐 크기를 후보영상과 모델과의 에러로 정의하고 식(10)과 같이 계산하였다.

$$\varepsilon = \sum_{i=1}^{N-1} \sum_{j=i+1}^N (m_i - m_j) u(m_i - m_j) \quad (10)$$

2. FFT(Fast Fourier Transform)를 이용한 빠른 물체검출

앞 절에서는 등급영상을 이용하여 특징추출과정 없이 입력 후보영역을 물체와 비물체로 식별하는 기법에 대해 설명하였다. 본 절에서는 이러한 식별기법을 전체 영상에 적용하는 방법에 대해 설명한다.

일정한 크기의 물체를 검출하기 위하여 그림 7에서와 같이 왼쪽 위 꼭지점의 좌표가 (x, y) 이고 크기가 X_C, Y_C 인 후보영역을 앞 절에서 제안한 에러함수를 기반으로 물체 또는 비물체로 식별하고자 한다.

이때, 후보영역의 밝기벡터를

$$\mathbf{C}^{xy} = (c^{xy}(0,0), \dots, c^{xy}(X_C-1, Y_C-1))^T \text{라 하고}$$

전체 입력영상벡터를 $\mathbf{I} = (i(0,0), \dots, i(X-1, Y-1))^T$ 라 할 때, $c^{xy}(m, n) = i(m+x, n+y)$ 인 관계가 성립한다.

식별을 위한 등급영상벡터 $\mathbf{L}v_n$ 은 식 (8)과 같이 $lv_n(x, y) (0 \leq x \leq X_C-1, 0 \leq y \leq Y_C-1)$ 를 통하여 정의된다. 그리고 후보영역 밝기벡터 \mathbf{C}^{xy} 와 검출대상 물체와의 에러 ε^{xy} 는 식 (10)을 이용하여 다음과 같이 계산한다.

$$m_k^{xy} = \mathbf{L}v_k \cdot \mathbf{C}^{xy} \quad (1 \leq k \leq N) \quad (11)$$

$$\varepsilon^{xy} = \sum_{i=1}^{N-1} \sum_{j=i+1}^N (m_i^{xy} - m_j^{xy}) u(m_i^{xy} - m_j^{xy}) \quad (12)$$

이때

$$lv_k^R(m, n) = \begin{cases} lv_k(X_C-1-m, Y_C-1-n) & 0 \leq m \leq X_C-1 \text{ and } 0 \leq n \leq Y_C-1 \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

로 정의된 함수 $lv_k^R(m, n)$ 을 이용하여 식 (11)을 전개하면

$$\begin{aligned} m_k^{xy} &= \mathbf{L}v_k \cdot \mathbf{C}^{xy} \\ &= \sum_{n=0}^{Y_C-1} \sum_{m=0}^{X_C-1} lv_k(m, n) \times c^{xy}(m, n) \\ &= \sum_{n=0}^{Y_C-1} \sum_{m=0}^{X_C-1} lv_k^R(X_C-1-m, Y_C-1-n) \times c^{xy}(m, n) \\ &= \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} lv_k^R(X_C-1+x-(m+x), Y_C-1+y-(n+y)) \times i(m+x, n+y) \end{aligned} \quad (14)$$

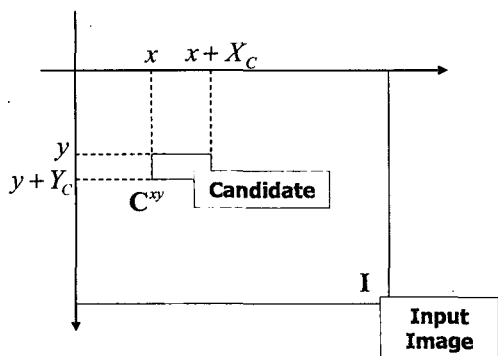


그림 7. (x,y)에 위치한 후보영역
Fig. 7. the candidate at (x,y).

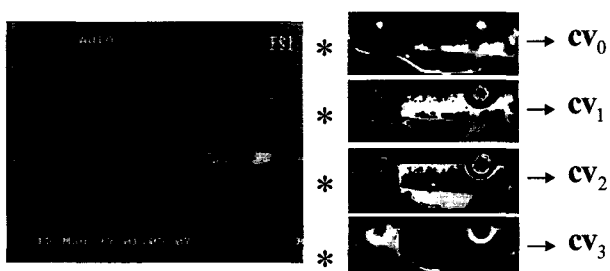


그림 8. 등급영역별 평균영상 계산
Fig. 8. Computation of mean images corresponding to each level region.

와 같은 형태가 되고, $m + x$ 와 $n + y$ 를 각각 a 와 b 로 치환하고 두 함수 $lv_k^R(x, y)$ 과 $i(x, y)$ 의 이차원 컨볼루션(convolution) 결과를 $cv_k(x, y)$ 라 하면, 다음과 같이 $m_k^{xy} = cv_k(x + X_C - 1, y + Y_C - 1)$ 인 관계를 확인할 수 있다. 그림 8과 그림 9는 후보영역의 각 등급 해당 영역 평균 밝기 m_k^{xy} 를 구하기 위한 컨볼루션 연산 과정과 그 결과인 $cv_k(x, y)$ 를 영상으로 구성하여 도시한 것으로 각 컨볼루션 영상은 일종의 지역통과여과기를 통과한 입력영상과 같은 형태를 나타낸다.

$$\begin{aligned}
 m_k^{xy} &= \sum_{a=-\infty}^{\infty} \sum_{b=-\infty}^{\infty} lv_k^R(X_C - 1 + x - a, Y_C - 1 + y - b) \times i(a, b) \\
 &= lv_k^R(m, n) \Big|_{m=x+X_C-1, n=y+Y_C-1} \\
 &= cv_k(x + X_C - 1, y + Y_C - 1) \tag{15}
 \end{aligned}$$

(*: convolution)

식 (15)의 결과를 식 (12)에 대입하면 다음과 같다.

$$\varepsilon^{xy} = \sum_{i=1}^{N-1} \sum_{j=i+1}^N r \left(cv_i(x + X_C - 1, y + Y_C - 1) - cv_j(x + X_C - 1, y + Y_C - 1) \right) \tag{16}$$

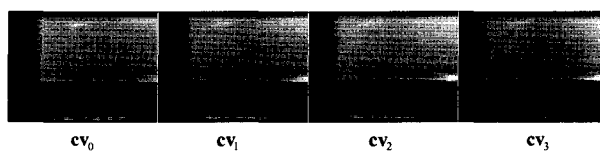


그림 9. 등급영역별 평균 밝기 영상
Fig. 9. Mean images corresponding to each level.

$$r(t) = \begin{cases} t, & t \geq 0 \\ 0, & t < 0 \end{cases} \tag{17}$$

즉, 왼쪽 위 꼭지점의 좌표가 (x, y) 이고 크기가 X_C, Y_C 인 후보영역의 검출대상 물체와의 에러가 꼭지점의 위치 (x, y) 를 입력으로 하는 함수형태로 나타나는 것이다.

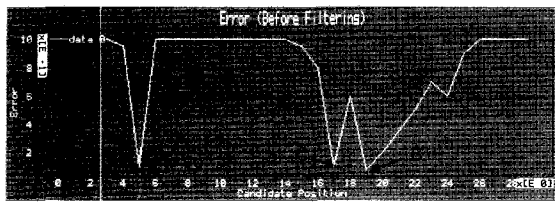
이때, 이차원 함수 $cv_k(x, y) = lv_k^R(x, y) * i(x, y)$ 는 $lv_k^R(x, y)$ 와 $i(x, y)$ 에 충분한 0 삽입(zero-padding)을 가정할 때, 이산푸리에변환(DFT) $\mathcal{J}(\cdot)$ 및 역변환 $\mathcal{J}^{-1}(\cdot)$ 을 이용하여 다음과 같이 계산할 수 있으며 고속푸리에변환(FFT) 알고리즘을 이용하여 빠른 시간 내에 함수 $cv_k(x, y)$ 를 획득할 수 있다.

$$\begin{aligned}
 cv_k(x, y) &= lv_k^R(x, y) * i(x, y) \\
 &= \mathcal{J}^{-1} \left\{ \mathcal{J}(lv_k^R(x, y)) \times \mathcal{J}(i(x, y)) \right\} \tag{18}
 \end{aligned}$$

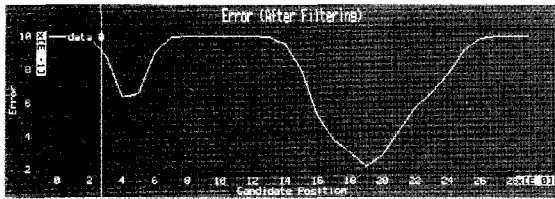
이러한 방법을 통하여 물체를 검출할 경우 다음과 같은 장점이 있다.

첫째, 고정된 크기의 물체에 대하여 적은 연산량으로 화소단위의 정확도를 가지는 물체검출이 가능하다. 기존의 정합기반 물체검출에서는 각 후보영역에 대하여 개별적인 후보영역 식별과정을 수행하여야 하고, 이 경우 서로 이웃하는 후보영역의 식별과정 사이에는 연산의 중복성이 존재한다. 이를 식 (18)과 같이 FFT를 이용하여 후보영역 선정과정과 식별과정을 통합함으로써 연산량을 크게 줄일 수 있다.

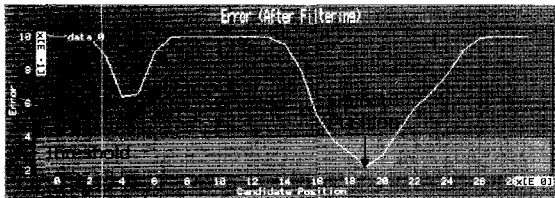
둘째, 후보영역에 대한 물체와의 에러값이 하나의 이차원 신호 형태로 나타남으로써, 손쉽게 정확한 물체의 위치 추정이 가능하다. 단순히 에러값의 임계화 과정만으로 후보영역을 식별할 경우 잡음에 의한 오검출 및 하나의 물체 주위에 존재하는 다수의 물체로 식별된 후보영역 발생에 대한 대처가 어렵다. 그러나 에러값이 하나의 탐색창의 위치에 따른 이차원 신호로 나타날 경



(a)



(b)



(c)

그림 10. 에러함수의 필터링 및 임계화를 통한 물체 위치 추정
 Fig. 10. Filtering of error function and estimation of object position by thresholding.

우 주위의 에러분포를 참고하여 이러한 오류를 줄일 수 있다.

본 논문에서는 저역통과여파기(low pass filter)를 거친 에러신호 $\epsilon(x, y) = \epsilon^{xy}$ 에 대하여 1차 임계화하여 에러가 임계값 이하인 영역집합을 찾고, 연속성을 가지는 후보영역 집합 내에서 국부최소(local minimum) 위치를 찾음으로써 물체의 영상 내 위치를 결정하였다. 예를 들어 탐색창의 위치에 따른 후보영역의 물체와의 에러 함수를 획득한 결과 그림 10(a)와 같이 잡음에 의한 오검출 후보(A)가 존재하고, 같은 물체에 대하여 임계치 이하의 에러값을 갖는 위치가 다수 존재(B, C)할 경우, 가우시안(Gaussian) 여파기를 이용하여 그림 10(b)와 같은 신호를 획득할 수 있고, 이를 임계화하여 그림 10(c)와 같이 에러가 최소인 점을 물체영역의 위치로 결정한다.

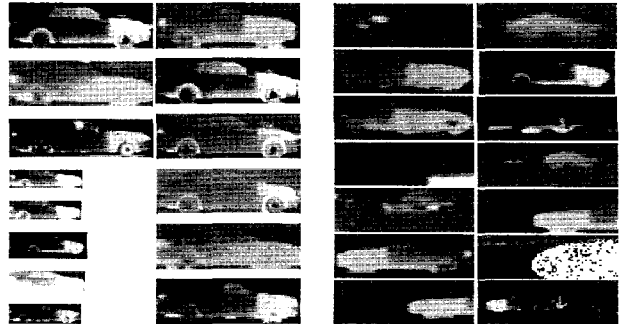
V. 실험 결과 및 분석

1. 밝기순위 특징

그림 11의 학습영상과 그림 12의 실험영상을 이용하여 특징벡터공간 상에서 학습영상 내 모델과 실험영상



그림 11. 학습영상
 Fig. 11. Training images.



(a) (b)

그림 12. 물체(a) 및 비물체(b) 실험영상
 Fig. 12. Test images : objects(a) non-objects(b).

과의 거리(error)를 측정함으로써 특징벡터의 성능을 비교하였다.

이때, 학습영상집합은 특정차종, 특정자세의 자동차를 검출하기 위하여 획득한 데이터베이스를 기반으로 하며, 고정된 자동차를 밤과 낮, 가열 전후 등 다양한 상황에서 적외선 카메라로 촬영한 영상이다. 실험영상 중 물체영상은 여러 자동차가 동시에 움직이는 상황에서 촬영한 적외선 영상 중 해당차종, 해당자세의 자동차가 등장하는 부분을 추출하여 획득한 것이다. 그리고 비물체 실험영상은 같은 자세의 다른 차종의 자동차, 같은 차종의 다른 자세의 자동차 등의 등장 부분과, 무작위로 추출한 부분 적외선 영상을 혼합하여 구성하였다.

각 실험영상에 대한 거리는 학습영상들과의 거리 중 최소거리(minimum distance)로 정하였다. 이 때, 비교하고자 하는 두 영상의 거리는 특징 벡터의 종류에 따라 다르게 정의하였다. 영상 밝기 값, PCA 는 각각의 영상을 특징 벡터로 변환한 후 특징 벡터간의 유클리드 거리(Euclidian distance)로 계산하였다. 밝기 순위 특징의 경우는 순위 영상들의 차이 제곱을 신뢰도에 따라 가중치 합(weighted squared sum of distance)으로 구하였다.

그림 13은 영상의 밝기값, 첫 번째부터 일곱 번째까지의 PCA 계수(PCA(0~6)), 네 번째부터 일곱 번째까지의 PCA 계수(PCA(3~6)) 및 수정된 밝기순위 특징을 이용하여 측정한 실험영상과 4장의 학습영상들과의 거리 중 최소거리를 실험영상이 찾고자 하는 물체인 경우(target)와 아닌 경우(non-target)로 구별하여 도시한

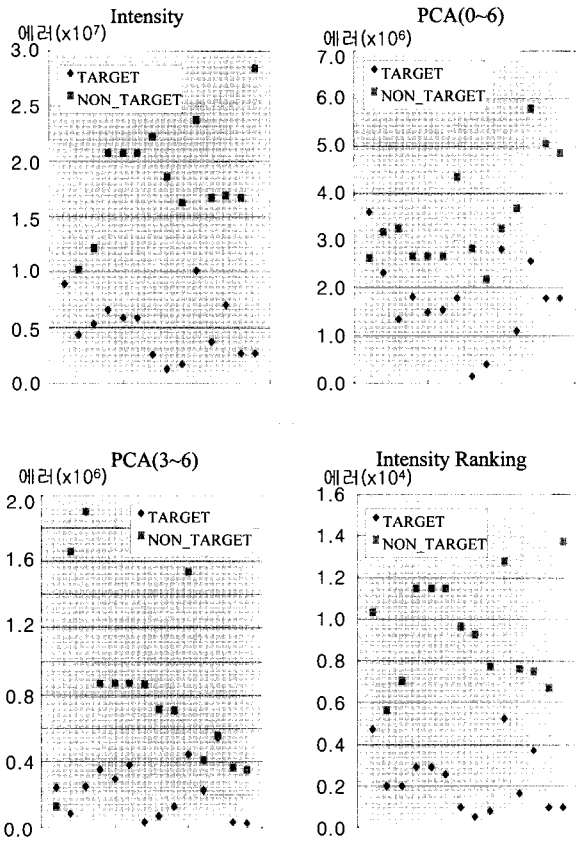


그림 13. 특징벡터 성능비교 (학습영상 4개)
 Fig. 13. Performance comparison of feature extraction methods (four training images).

것이다. 공정한 실험을 위하여 히스토그램 균등화 기법을 이용하여 각 영상을 정규화한 후 특징을 추출하였다. 실험결과, 밝기값 및 PCA 계수를 특징으로 이용하여 특징 공간을 구성할 경우, 일부 물체 실험영상에서 추출된 특징벡터가 비물체 실험영상에서 추출한 특징벡터보다 학습 특징벡터에 더욱 가까운 거리 역전 현상이 심각하게 발생하는 것을 확인할 있었다. 그러나 제안하는 수정된 밝기순위 특징을 이용한 경우 물체 실험영상은 학습영상과의 거리가 가깝고 비물체 실험영상은 먼 위치에 존재하므로 특징공간 상에서의 거리를 이용하여 물체와 비물체를 구분할 수 있음을 확인할 수 있었다.

그림 14는 그림 11의 학습영상 집합의 평균영상(mean-image)을 획득하고, 밝기값, PCA 계수, 밝기순위 특징 및 등급영상을 이용한 검출기법을 이용하여 측정된 평균영상과 실험영상과의 거리를 물체, 비물체로 구별하여 도시한 것이다. 실험결과 제안한 밝기순위 특징을 이용하여 특징을 추출한 경우, 한 장의 평균영상만을 이용할 경우에도 특징공간에서의 거리를 이용하여 물체와 비물체를 구분할 수 있음을 확인할 수 있었다.

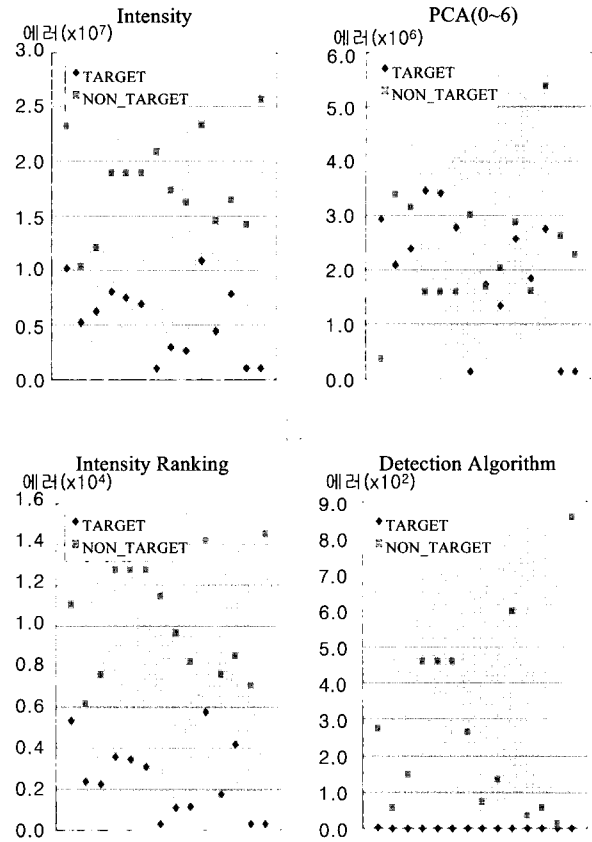


그림 14. 특징벡터 성능비교 (1개 모델)
 Fig. 14. Performance comparison of feature extraction methods (using one mean image).

이는 각 영상의 밝기순위와 평균영상의 밝기순위가 크게 다르지 않기 때문이며, 이를 통해 밝기순위 특징이 적외선 영상에서 존재하는 국부적 밝기변화에 강인한 특징임을 알 수 있다. 또한 평균영상으로부터 제안하는 방법을 이용하여 등급영상을 생성하고, 생성된 등급영상을 이용하여 식(10)의 에러값(거리척도)을 측정된 경우(Detection Algorithm), 물체실험영상과 평균영상과의 오차가 가장 작고 비물체 실험영상은 물체실험영상보다 항상 그 에러가 크게 측정됨을 확인할 수 있었다. 이는 제안한 거리척도가 적외선 영상의 특징을 잘 반영하였기 때문으로 생각된다.

2. 밝기순위 특징을 이용한 물체검출 기법

제안하는 밝기순위 특징을 이용하여 실외환경에서 획득한 2가지 차종의 자동차가 포함된 적외선 영상들로부터 영상 내 자동차 영역을 검출하였다.

검출대상 자동차의 영상 내 크기는 대략 $70 \times 20 \text{ pixel}^2$ 정도로 고정되어 있고, 그림 11과 같이 각 차종별로 서로 다른 시간(밤, 낮) 및 상황(미가열, 가열)에서

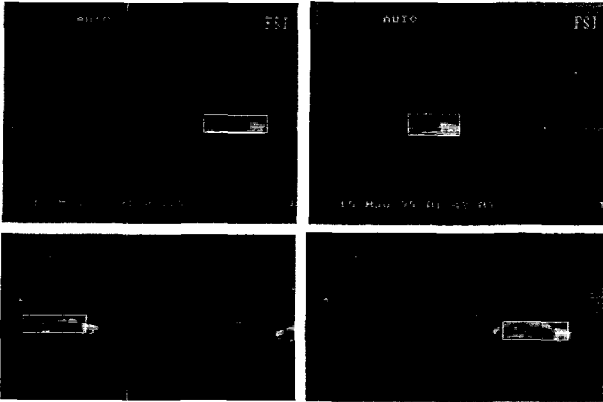


그림 15. 적외선 영상 내 자동차 검출결과
Fig. 15. Results of vehicle detection in a FLIR image.

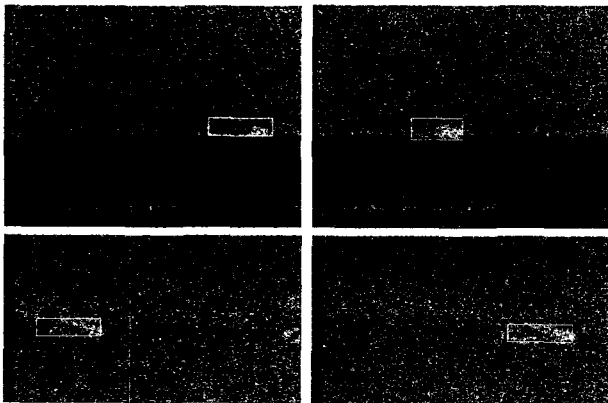


그림 16. 강한 잡음이 첨가된 상황에서의 검출결과
Fig. 16. Detection results in the situation of noise existence.

획득한 IV장의 학습영상을 이용하여 각 화소별 신뢰도 및 평균영상을 획득하였다. 이들 신뢰도를 기반으로 평균영상 내 밝기순위를 양자화하여 그림 6과 같은 등급영상을 구성하였으며 이를 이용하여 물체를 검출하였다. 이때 $320 \times 240 \text{ pixel}^2$ 크기의 입력 영상 및 각 등급영상은 $512 \times 256 \text{ pixel}^2$ 크기로 0 삽입을 수행하고 IV.2절에서 제시한 방법으로 물체의 위치를 검출하였다.

검출하고자 하는 자동차와 같은 차종의 자동차 영상으로부터 획득한 등급영상을 사용하여 검출과정을 수행한 결과는 그림 15의 흰 사각형과 같다. 그림에서와 같이 배경이 균일할 경우 매우 정확한 검출결과를 획득할 수 있었으며, 배경이 복잡하거나 검출 대상 자동차가 다른 차종의 자동차와 겹쳐 있을 때에도 비교적 안정적인 검출결과를 얻을 수 있었다.

그림 16은 그림 15의 실험영상에 대하여 크기 80의 표준편차를 가지는 가우시안 잡음을 첨가한 후 검출과정을 수행한 결과이다. 그림에서와 같이 영상 내 잡음 성분이 매우 강하여 육안으로 물체의 위치 확인이 어려

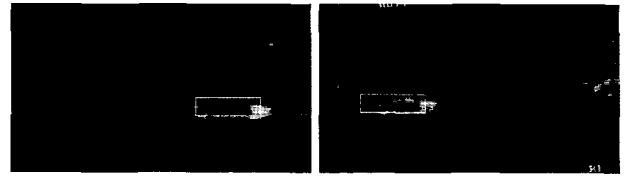


그림 17. 크기가 변한 경우 및 기울어진 물체에 대한 검출결과
Fig. 17. Results of detecting enlarged or rotated objects.

운 경우에도 비교적 정확한 검출결과를 획득할 수 있었다. 이는 제안한 검출기법은 후보 영역 내 밝기의 부분 평균값을 바탕으로 물체와의 유사도를 측정하기 때문이다.

그림 17은 영상 내 물체의 크기가 학습영상 내 물체와 다르거나 물체가 기울어져 있는 경우에 대한 실험결과이다. 그림에서와 같이 입력영상 내 물체의 크기 및 기울어진 정도가 변화하더라도, 학습영상과 크게 다르지 않을 경우에는 비교적 정확한 물체의 위치를 찾아낼 수 있었다.

실험결과 $320 \times 240 \text{ pixel}^2$ 크기의 입력영상 한 장에 존재하는 대략 $70 \times 20 \text{ pixel}^2$ 크기의 후보영역 55,000개의 식별과정을 수행하고, 후처리과정을 통하여 물체영역을 결정짓는데 1.5초 정도의 시간이 소요되었다.

VI. 결 론

본 논문에서는 적외선 영상에 적합한 정지영상 내 물체검출 기법을 제안하였다.

일반적으로 적외선 영상 내 물체는 영상의 획득환경과 물체에 따라 그 밝기가 매우 다양하게 바뀌고, 정확한 경계를 획득하기 어렵기 때문에 기존의 CCD 영상을 위하여 제안된 방법으로는 만족할 만한 검출결과를 얻기 어렵다. 또한 정지영상 내에서 관심 물체를 찾고자 할 경우 물체의 대략적인 위치에 대한 정보를 얻기 힘들기 때문에 검출과정을 수행하는데 많은 시간이 필요하다고 알려져 있다.

본 논문에서 제안한 기법은 이러한 문제점들을 해결하고자 하였으며 다음과 같은 특징을 가진다.

첫째, 제안한 밝기순위 특징은 적외선영상 내 예측하기 어렵고, 국부적인 밝기변화에 강인한 특징으로 기존의 밝기기반 특징에 비하여 적외선 영상에서 관심물체를 비물체와 구별하는 능력이 우수하고 잡음에 강인하며, 형태기반 특징에 비하여 안정적으로 추출이 가능하다.

둘째, 제안한 물체검출기법을 이용하여 적은 연산량으로 화소단위의 정확도를 가지는 물체식별 및 후처리 과정을 통한 보다 정확한 물체위치 선정이 가능하다. 이는 기존의 정합기반 물체검출 기법의 후보영역 선정 과정과 물체식별과정이 통합된 형태로서 이웃하는 후보영역들에 대한 개별적인 식별과정 사이에 존재하는 계산의 중복성을 효과적으로 제거한 결과이다. 실험결과, 320×240 pixel² 크기의 입력영상에서 물체영역을 결정 짓는데 약 1.5초의 시간이 소요되었다. 이는 특별한 후보영역선정 기법을 적용할 수 없는 상황에서 기존의 검출기법이 모든 가능한 후보영역에 대한 식별과정을 수행하는데 5초 이상 걸리는데 비하여 매우 빠른 계산속도이다.

제안한 기법을 이용하여, 실제 물체검출 과정에 적용, 실험해 본 결과 그 밝기변화가 다양하고 잡음 성분이 매우 강한 경우에도 정확한 검출결과를 획득할 수 있었다. 그러나 훈련영상과 검출대상 물체의 크기 차이가 많이 나거나, 영상 내 물체가 심하게 기울어진 경우, 물체의 진행방향이 달라진 경우에는 정확한 물체 위치 검출에 실패함을 관찰할 수 있었다. 이들 문제에 대한 해결책으로 다양한 크기, 회전, 자세에 대한 학습영상집합을 획득하고, 각 경우에 대하여 개별적인 검출과정을 수행하는 방법을 생각할 수 있다. 그러나 보다 근본적인 문제해결을 위해서는 이러한 변화에 불변한 특징에 관한 연구가 수행되어야 할 것이다.

참 고 문 헌

- [1] D. M. Weber and D. P. Casasent, "Quadratic Gabor filters for object detection," *IEEE Trans. on Image Processing*, vol. 10, no. 2, pp. 218-230, 2001.
- [2] M. A. Khabou and P. D. Gader, "Automatic target detection using entropy optimized shared-weight neural networks," *IEEE Trans. on Neural Networks*, vol. 11, no. 1, pp. 186-193, 2000.
- [3] R. P. Broussard, S. K. Rogers, M. E. Oxley, and G. L. Tarr, "Physiologically motivated image fusion for object detection using a pulse coupled neural networks," *IEEE Trans. on Neural Networks*, vol. 10, no. 3, pp. 554-563, 1999.
- [4] 신호철, 최해철, 이진성, 조주현, 김성대, "자동 목표물 인식 시스템을 위한 클러스터 기반 투영기법과 혼합 전문가 구조," 대한전자공학회 논문지, 제 40권, SP편, 제3호, pp. 19-32, 2003.
- [5] M. Irani and P. Anandan, "A unified approach to moving object detection in 2D and 3D scenes," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 6, pp. 577-589, 1998.
- [6] R. L. Hsu, M. Abdel-Mottaleb and A. K. Jain, "Face detection in color images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 696-706, May. 2002.
- [7] C. Garcia and G. Tziritas, "Face detection using quantized skin color regions merging and wavelet packet analysis," *IEEE Trans. Multimedia*, vol. 1, no. 3, pp. 264-277, Sep. 1999.
- [8] H. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp. 23-38, Jan. 1998.
- [9] H. I. Hahn and R. N. Strickland, "Wavelet transform methods for object detection and recovery," *IEEE Trans. on Image Processing*, vol. 6, issue 5, pp. 724-735, 1997.
- [10] J. Zhou, J. Wu, and X. Zhang, "Vehicle detection in static road images with PCA-and-wavelet-based classifier," *2001 IEEE Intelligent Transportation Systems*, pp. 740-744, 2001.
- [11] M. Oren, E. Osuna, C. Papageorgiou, T. Poggio, and P. Sinha, "Pedestrian detection using wavelet templates," *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 193-199, Jun. 1997.
- [12] J. E. Dayhoff, A. K. Jain, and B. Kamgar-Parsi, "Aircraft detection: a case study in using human similarity measure," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, issue. 12, pp. 1404-1414, 2001.
- [13] D. M. Gavrilu and V. Philomin, "Real-time object detection for "smart" vehicles," *Proc. 7th IEEE Int'l Conf. Computer Vision.*, vol. 1, pp. 87-93, 1999.
- [14] 신기선, "방향성 기반의 거리 지도를 이용한 형태 정합 알고리즘," 한국과학기술원 석사학위논문, 2001.
- [15] D. P. Huttenlocher and C. F. Olson, "Automatic target recognition by matching oriented edge pixels," *IEEE Trans. on Image Processing*, vol. 6, issue. 1, pp. 103-113, Jan 1997.

— 저 자 소 개 —



박 재 희(학생회원)
 2001년 연세대학교 전자공학전공
 학사 졸업.
 2003년 한국과학기술원 전자전산
 학과 석사 졸업.
 2003년~현재 한국과학기술원
 전자전산학과 박사과정
 재학 중.

<주관심분야 : 영상처리, 패턴인식, 컴퓨터비전>



최 학 훈(정회원)
 2001년 포항공과대학교
 전자전기공학과 학사졸업.
 2003년 한국과학기술원
 전자전산학과 석사졸업.
 2003년~현재 삼성전자 영상
 디스플레이사업부 Core
 Tech 연구원

<주관심분야 : 영상처리, 컴퓨터비전>



김 성 대(정회원)
 1977년 서울대학교 전자공학과
 학사 졸업.
 1979년 한국과학기술원 전기 및
 전자공학과 석사 졸업.
 1983년 프랑스 INPT ENSEEIHT
 박사 졸업

1984년~현재 한국과학기술원 전자전산학과 전기
 및전자공학전공 교수

<주관심분야 : 영상처리, 영상통신, 컴퓨터비전,
 VLSI 구현 등>