

논문 2008-45SP-4-2

# 웨이블릿 기반의 신경망과 불변 모멘트를 이용한 실시간 이동물체 인식 및 추적 방법

(Real-time Moving Object Recognition and Tracking Using The  
Wavelet-based Neural Network and Invariant Moments)

김 종 배\*

(Jong Bae Kim)

## 요 약

본 논문은 실시간 감시 시스템을 위한 웨이블릿(wavelet) 기반의 신경망과 불변 모멘트를 이용한 이동물체 인식과 추적 방법을 제안한다. 제안한 방법의 첫 번째인 움직임 후보영역 검출 단계에서는 연속된 두 프레임간의 차영상 분석 방법을 기반으로 하여 물체의 움직임에 의해 화소값 변화가 발생한 후보영역을 검출한다. 두 번째인 물체 인식 단계에서는 검출된 후보영역에 웨이블릿 신경망(wavelet neural network: WNN) 기반의 인식 방법을 사용하여 추적하고자하는 물체가 포함되어 있는지를 판별한다. 세 번째인 물체 추적 단계에서는 인식된 물체에 웨이블릿 불변 모멘트(invariant moments) 기반의 매칭 방법을 사용하여 인식된 이동 물체를 추적한다. 영상내에서 이동물체를 검출하기 위해 본 논문에서는 이전 영상과 현재 영상간의 화소 밝기 차이에서 적응적 임계값(adaptive thresholding)을 사용하여 주위 환경 변화에 강인한 이동물체 검출이 가능하였다. 또한 물체의 인식과 추적을 위해 웨이블릿 특징값을 사용함으로써, 계산 시간의 감소와 영상의 잡음에 의한 영향을 최소화시킬 수 있을 뿐만 아니라, 물체 인식 정확도가 향상되었다. 제안한 방법을 일반 도로에서 획득한 영상에서 실험한 결과, 자동차 검출율은 92.8%, 프레임당 처리 시간은 0.24초이다. 이것을 통해 제안한 방법은 실시간 지능형 교통 감시 시스템에 유용하게 적용될 수 있음을 알 수 있다.

## Abstract

The present paper propose a real-time moving object recognition and tracking method using the wavelet-based neural network and invariant moments. Candidate moving region detection phase which is the first step of the proposed method detects the candidate regions where a pixel value changes occur due to object movement based on the difference image analysis between continued two image frames. The object recognition phase which is second step of proposed method recognizes the vehicle regions from the detected candidate regions using wavelet neural-network. From object tracking phase which is third step the recognized vehicle regions tracks using matching methods of wavelet invariant moments bases to recognized object. To detect a moving object from image sequence the candidate regions detection phase uses an adaptive thresholding method between previous image and current image as result it was robust surroundings environmental change and moving object detections were possible. And by using wavelet features to recognize and tracking of vehicle, the proposed method decrease calculation time and not only it will be able to minimize the effect in compliance with noise of road image, vehicle recognition accuracy became improved. The result which it experiments from the image which it acquires from the general road image sequence and vehicle detection rate is 92.8%, the computing time per frame is 0.24 seconds. The proposed method can be efficiently apply to a real-time intelligence road traffic surveillance system.

**Keywords :** wavelet transform, neural network, invariant moments

\* 정회원, 서울디지털대학교

(Dept. Computer Eng., Seoul Digital University)

※ 이 논문은 2007년 정부(교육인적자원부)의 재원으로  
한국학술진흥재단의 지원을 받아 수행된 연구임  
(KRF-2007-331-D00715)

접수일자: 2008년2월18일, 수정완료일: 2008년4월23일

## I. 서 론

최근 도로 영상에서 도로 상황 파악, 교통량 분석,  
교통관제 등의 작업을 실시간으로 수행가능한 지능형

도로 교통 자동화 시스템에 대한 요구가 급격히 증가되고 있다. 특히, 도로 교통 감시 위해 도로 영상에서 주행 중인 자동차의 실시간 검출 및 추적은 정확한 교통량 조사 및 관심 차량 추적에 이용될 수 있고, 또한 범죄 예방 및 무인 주차장 시스템 구축과 같은 분야에 실용적으로 적용하기 위해 그 동안 활발히 연구가 진행되어 왔다. 과거의 도로 감시 시스템 연구에서는 도로 영상에서 자동차의 검출 및 추적을 위해서는 입력 영상의 밝기 변화에 따른 화소 정보의 분석이 대표적인 방법이었다<sup>[1~4]</sup>. 더구나, 도로 영상의 획득은 도로 상에 고정된 카메라에 의해 수행됨으로써 복잡한 영상 처리 방법이 필요하지 않은 것이 사실이었다. 하지만, 교통 혼잡 요인으로 차량 간의 거리 감소로 인해 단순히 화소의 밝기 변화에 따른 자동차 검출은 한계에 이르고 있다. 또한, 자동차의 급증으로 인해 도로 대기 환경의 오염이 날로 가중되고 있음으로 획득된 영상의 화질 저하로 인해 실시간 교통 감시에 있어서 많은 어려움이 존재한다.

#### 가. 기존 연구

도로 교통 감시를 위한 기존의 연구에는 인접한 영상 간의 화소 명암차(difference intensity of images)를 이용하여 이동물체를 검출하고, 이를 추적하기 위해 규칙 기반의 추론(rule-based reasoning) 방법을 적용하였다<sup>[5]</sup>. 그리고, Foresti<sup>[6, 7]</sup>는 물체의 모델을 데이터베이스에 저장하여 모델의 특징 정보 매칭을 통해 이동물체를 추적하였으며, Betke과 Davis<sup>[8]</sup>는 이동물체를 템플릿 매칭(template matching)에 의해 인식하고, 인식된 물체의 특징 정보 상관관계를 이용하여 추적하였다. 그러나 이러한 연구들을 실시간 도로 감시에 적용하기에는 다음과 같은 문제점들을 내포하고 있다. 첫 번째는, 차영상(difference image)을 구할 때, 고정된 임계값을 결정하여 이 임계값을 기준으로 이동물체의 움직임에 의해 화소값이 변화된 영역을 검출함으로써, 도로 영상에서 조도 변화와 같은 환경적 요인의 변화에 적용적으로 반응하지 못하는 문제가 발생하여, 이동물체 검출율이 낮아지는 문제가 발생한다. 두 번째는, 영상내에서 이동물체 인식을 위해 그레이(gray) 특징값 기반의 템플릿 매칭 방법, 혹은 윤곽선 정보 기반의 매칭 방법을 사용함으로써 물체의 움직임에 대한 고유한 형태학적인 정보를 사용하지 않음으로 물체 인식률이 저하되는 문제가 발생한다. 그리고 세 번째는, 이동물체 추적을 위해 시간축선상 인접한 프레임들 사이의 밝기 정보의 상관도

를 사용하여 추적함으로써 이동물체들이 서로 겹침이나 크기, 회전, 이동 변환에 의해 오인식되는 문제로 귀결되고, 이것은 연속적인 물체 추적이 불가능해짐으로써 추적 중인 물체를 놓치는 문제가 발생한다.

언급한 이러한 문제점들을 극복하기 위해 여러 가지 많은 연구들이 수행되고 있다. 첫 번째 연구로는 차영상을 계산하기 위해 화소의 밝기 변화를 사용하는 것이 아니라 두 영상간의 모션 정보인 optical flow를 사용하여 움직임이 발생한 부분을 검출하는 방법을 적용하였다<sup>[9]</sup>. 하지만, 실시간 시스템에 적용하기 위해서는 모션 정보에 따른 optical flow 계산에 시간 소모가 많은 단점을 가지고 있다. 두 번째 연구로는 물체에 대한 고유한 정보를 표현할 수 있는 주성분 분석방법(Principal Component Analysis: PCA) 방법을 사용하여 물체를 인식하였다<sup>[10]</sup>. 하지만, 이 방법 역시 320×240 화소 크기 영상에서 주성분을 계산하는데 많은 시간 소모로 인해 실시간 시스템 적용에는 한계가 있다. 세 번째 연구로 이동물체의 추적을 위해 윤곽선에 기반을 둔 active contour 방법<sup>[11]</sup>을 적용하였다. 하지만, 이 방법 역시 확률론적 최적화 방법으로 최종의 적합한 해를 찾는 데 계산 시간의 소모로 인해 실시간 시스템 적용에는 다소 한계를 가지고 있다.

따라서 본 논문에서 집중하는 실시간 도로 감시 시스템에서 자동차를 인식하기 위해 관심 물체에 관련된 수많은 특징들이 존재하는 것과 이러한 특징들 간의 복잡성이라는 물체 인식에서의 난제에 대한 해결책은, 그들 사이의 연관성을 효율적으로 분석할 수 있어야 하는 것으로 집중된다. 이러한 복잡한 연관성을 적용하기 위해 다양하고 복잡한 신경망 구조를 적용하는 연구가 진행되었다<sup>[12]</sup>. 그러나 많은 특징을 사용할수록 신경망 네트워크 내부의 복잡도가 같이 증가하는 문제가 발생한다. 이러한 문제를 해결하기 위해서 신경망 입력값에 고유한 특징 정보를 유지면서 복잡도를 줄이기 위한 방안이 필요하다.

#### 나. 제안 방법

본 논문에서는 기존 연구들에서 발생한 문제점들을 해결하고 실시간 이동물체의 인식 및 추적을 위해 웨이블릿 기반한 신경망과 불변 모멘트를 이용한 방법을 제안한다. 본 연구에서는 먼저 움직임이 있는 영역 검출을 위해 차영상 분석 방법을 기반으로 하여 연속된 두 프레임의 움직임이 있는 후보영역을 추출한다. 차영상 분석을 통한 후보영역 추출은 입력 영상이 연속적이고,

전 시점에 획득한 영역의 모양이나 위치가 급격히 변하지 않는다는 특징을 이용한 것이다. 그리고 추출된 후보영역에 추적 관심 대상인 물체가 포함되어 있는지를 판별하기 위해 웨이블릿 신경망을 사용한다. 이것은 검출된 후보영역이 가진 특징 정보의 차원을 웨이블릿 변환을 통해 저차원으로 분해하고 저차원의 특징 정보를 신경망 네트워크로 학습시켜 네트워크간의 거리를 정의하여 인식하는 것이다. 이렇게 인식된 이동물체들을 추적하기 위해 웨이블릿 불변 모멘트 기반의 매칭 방법을 사용하여 추적한다. 불변 모멘트는 물체의 이동, 변환, 회전등에 의한 변화와는 무관하게 고유한 정보를 표현할 수 있는 특징을 가지고 있음으로 인해 검출된 물체와 추적중인 물체가 포함된 영상 내에서 존재 가능한 모든 부분 영상들과의 상관 계수 유사도 측정을 통해, 이동물체를 추적한다. 그림 1은 제안한 방법을 나타낸 전체 흐름도이다. 먼저 적응적 임계치 방법을 통해 영상내의 조도 변화에 적응적으로 대응하는 임계값을 사용하여 이동물체에 의해 밝기 변화가 발생된 후보영역을 검출하는 단계와 검출된 후보영역중에서 실제 이동물체가 포함된 영역을 인식하는 단계, 마지막으로 인식된 물체 영역을 추적하는 단계로 이루어진다. 제안한 방법의 입력은 실제 도로 영상이며, 출력은 주행중인 자동차의 수와 도로 상황 등의 정보를 출력한다.

본 논문의 구성은 II장에서 도로 영상에서 움직임이 있는 후보영역을 검출하는 방법을 설명하고, III장에서는 웨이블릿 신경망을 이용한 검출된 후보영역의 물체 인식 방법에 관해 설명하고, IV장에서는 인식된 물체를 추적하기 위해 웨이블릿 불변 모멘트 기반의 매칭 방법에 대해 기술한다. 그리고 실험 결과는 V장에서 보이며, 마지막으로 VI장에서 결론을 맺는다.

## II. 후보영역 검출

고정된 카메라에서 획득한 영상 시퀀스에서 이동물체의 움직임에 의해 화소값이 변화된 후보영역들을 검출하기 위해 연속된 두 프레임간의 차영상 분석 방법을 기반으로 한다. 이 방법은 다른 검색 방법보다 간단하고 빠르며, 특히 도로 영상과 같이 자동차의 움직임에 의해 이전 프레임( $I_{t-1}$ )과 현재 프레임( $I_t$ ) 사이에서 연속적인 변화가 발생하는 경우에 효율적으로 적용되어 질 수 있다<sup>[13]</sup>. 차영상 분석은 식 (1)과 같이 임계값  $\theta$ 에 의해 영상내의 변화가 발생 영역을 구분하게 된다.

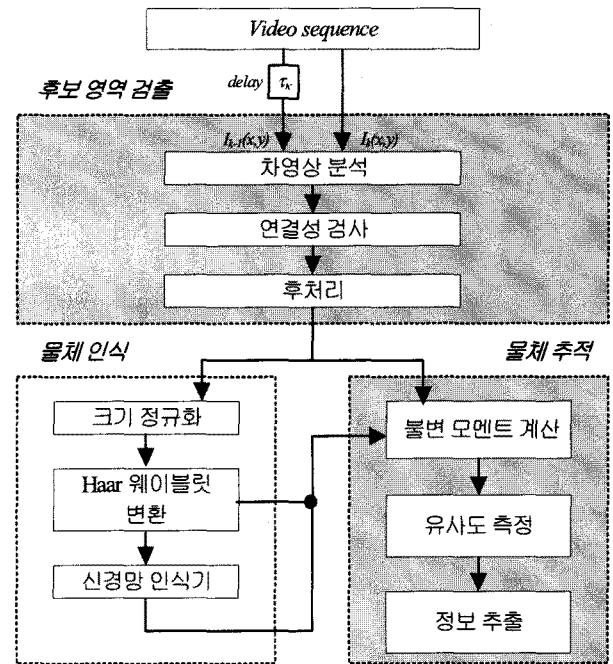


그림 1. 제안한 연구의 흐름도  
Fig. 1. Flowchart of the proposed system.

$$T(x, y) = \begin{cases} 1, & \text{if } |I_t(x, y) - I_{t-1}(x, y)| \geq \theta \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

하지만, 이 방법에서 고정된 임계값을 적용할 경우 간단하면서도 손쉽게 영상내의 변화가 있는 부분을 검출할 수 있지만 실제적으로 이동물체의 움직임에 의해 변화된 영역만을 검출하는데 어려움이 존재한다. 일반적으로 배경에 존재하는 나무와 같은 배경의 물체 움직임이나 혹은 영상 잡음 때문에 발생한 영 미세한 변화까지 검출함으로써 이러한 부분을 배제하기 위해 환경 변화에 적응적으로 대처 가능한 임계값 계산 방법이 필요하다. 그러므로 본 단계에서는 연속된 두 프레임간의 밝기 차이를 고정된 임계값이 아닌 적응적 임계값<sup>[13, 16]</sup>을 이용하여 이동물체의 움직임에 의해 화소값이 변화된 후보 영역을 검출하고, 추출된 후보 영역에 후처리 과정을 통해 배경의 움직임이나 밝기의 변화에 의해 발생한 잡영과 그림자를 제거한다.

### 가. 적응적 임계값

연속된 두 프레임에서 이동물체의 움직임을 검출하기 위한 임계값 계산은 두 프레임에서 얻어진 밝기차를 히스토그램으로 표현한다. 즉, 히스토그램의 가로축은 화소의 밝기 차이의 범위인  $[-255 \sim 255]$ 까지이고, 세로축은 해당 밝기 차이값을 가지는 화소의 확률 분포를 나타낸다. 이러한 히스토그램은 식 (2)와 같이 두 영상

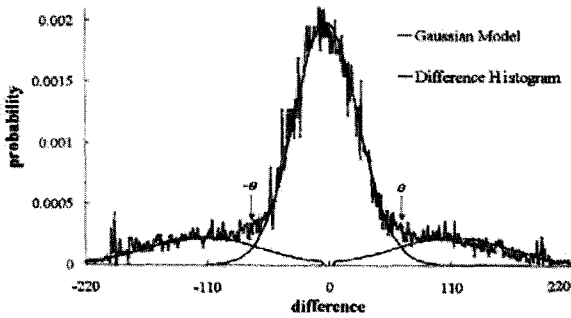


그림 2. 차영상 분석에 따른 히스토그램의 확률적 분포와 가우시안 모델<sup>[13]</sup>  
 Fig. 2. Probability distribution of histogram and Gaussian model according to the difference image analysis.

의 밝기 차이값  $i$ 를  $f(i)$ 에 대입하여 그림 2와 같이 세 개의 가우시안 분포(three modal Gaussian distribution)를 따른다는 가정에서 각 가우시안 분포가 만나는 점을 임계값  $\theta$ 으로 결정한다.

$$f(i) = \frac{\varphi_1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(i-\mu_1)^2}{2\sigma_1^2}} + \frac{\varphi_2}{\sqrt{2\pi}\sigma_2} e^{-\frac{(i-\mu_2)^2}{2\sigma_2^2}} + \frac{\varphi_3}{\sqrt{2\pi}\sigma_3} e^{-\frac{(i-\mu_3)^2}{2\sigma_3^2}} \quad (2)$$

식 (2)의  $\varphi_i, \sigma_i$ , 그리고  $\mu_i$ 는 식 (3)과 같이 계산된다. 여기서 그림 2의 왼쪽에 위치한 가우시안 확률분포의 계산 범위는  $[-255 \sim -\theta]$ , 가운데 위치한 가우시안 확률 분포의 범위는  $[-\theta \sim \theta]$ , 마지막에 위치한 가우시안 확률 분포의 범위는  $[\theta \sim 255]$ 이다.

식 (3)에서는  $m(i)$ 는 차이값  $i$ 를 가지는 히스토그램의 확률 분포 값이다.

$$\begin{aligned} \varphi_1(\theta) &= \sum_{i=-255}^{-\theta-1} m(i) \\ \mu_1(\theta) &= \frac{1}{\varphi_1} \sum_{i=-255}^{-\theta-1} d \times m(i) \\ \sigma_1^2 &= \frac{1}{\varphi_1} \sum_{i=-255}^{-\theta-1} [d - \mu_1(i)]^2 m(i) \end{aligned} \quad (3)$$

그리고 식 (2)의 첫 번째와 세 번째 항은 물체의 움직임 때문에 발생하는 밝기 차이값의 변화들을 나타내며, 두 번째 항은 움직임이 없거나 잡음이 존재하여 생기는 밝기 차이값의 변화들을 나타낸다. 임계값  $\theta$ 는 식 (4)와 같이 적합 기준 함수에 의해 결정된다. 식 (4)는

차영상이 가지는 밝기 차이값이 세 개의 가우시안 분포를 따른다는 가정 하에 밝기 차이값  $i$ 가 가지는 확률 분포 값  $P(i)$ 와 가우시안 분포  $f(i)$ 의 차이가 가장 적은  $i$ 값을 찾아냄으로써 임계값  $\theta$ 가 결정되어진다. 여기서 밝기 차이값  $i$ 의 범위는  $\max(d) \leq i \leq \min(d)$ 이다. 즉, 밝기 차이값의 최하한은  $-255$ 이고, 최상한은  $+255$ 이지만, 일반적인 차이값의 범위는 이보다 더 작을 수 있으므로 최하한과 최상한을 찾아냄으로써 차영상 히스토그램의 확률 분포  $m(d)$ 와 가우시안 분포  $f(i)$ 의 상관관계 계산 범위를 줄일 수 있다.

$$\theta = \text{Min}(|f(i) - m(d)|), \min(d) \leq i \leq \max(d) \quad (4)$$

식 (4)에서 계산된 임계값  $\theta$ 는 연속적인 프레임의 밝기차에 대한 하나의 Change Detection Mask(CDM)을 생성한다. CDM은 움직임 있는 영역과 움직임이 없는 영역으로 구별 시켜주는 이진화된 영상이다. 여기서 적용적 임계값 계산을 모든 입력 이미지를 대상으로 수행함으로써 시간적인 소모가 너무 크게 발생한다. 따라서 제한한 방법에서는 급격한 조도 변화가 발생되지 않는다는 가정을 두고 매 7,680프레임마다 임계값을 구한다. 즉, 입력 영상이 초당 8프레임으로 획득됨으로써 적용적 임계값 계산은 16분에 한 번씩 계산되어진다. 차영상 분석에 의한 얻어진 이진화된 영상내에 산재되어 있는 화소들을 각각의 하나의 동일한 영역으로 할당하기 위해 인접한 화소의 연결성 성분 조사(connected component analysis)하여, 연결되어 있는 화소만 남기고 나머지는 제거한다. 그리고 모폴로지(morphology) 연산을 통해 차영상에 존재하는 잡음 영역들을 제거할 수 있다. 차영상 분석에 의해 얻어진 후보영역은 조도에 의해 발생한 그림자 영역이나 프레임 간의 시간 차이 간격에 의해 도로 배경 영역도 같이 포함될 수 있다<sup>[17]</sup>. 따라서 후처리 과정을 거쳐 정확한 이동물체의 움직임이 있는 후보영역만을 검출하는 것이 필요하다. 이를 위해 후처리 과정에서 입력 영상에서 소벨(Sobel) 에지맵(edge map)  $e(x,y)$ 로부터 검출된  $m \times n$  화소 크기의  $k$ 개의 후보 영역들에 에지맵  $e(x,y)_k$ 를 구하고, 이 에지맵으로부터  $k$ 번째 후보 영역의 투영 벡터인  $v_k, h_k$ 를 얻는다. 수평, 수직 투영 벡터는 식 (5)와 같이 계산한다. 식 (5)에서  $N$ 은 검출된 움직임 후보 영역들의 수를 의미한다.

$$v_k, h_k, k = 1, \dots, N$$

$$v_k = [v_1, v_2, \dots, v_n]_k$$

$$= \left[ \frac{\sum_{i=1}^n e(1, i)_k}{n}, \frac{\sum_{i=1}^n e(2, i)_k}{n}, \dots, \frac{\sum_{i=1}^n e(m, i)_k}{n} \right]_k$$

$$h_k = [h_1, h_2, \dots, h_m]_k$$

$$= \left[ \frac{\sum_{i=1}^m e(i, 1)_k}{m}, \frac{\sum_{i=1}^m e(i, 2)_k}{m}, \dots, \frac{\sum_{i=1}^m e(i, n)_k}{m} \right]_k \quad (5)$$

투영 벡터  $v_k$ 와  $h_k$ 에서로부터 후보 영역 내에 불필요한 영역들을 제거하기 위해 각 투영 벡터에 최대값을 찾고, 찾아진 값을 2로 나누어  $T_v$ 와  $T_h$ 를 구한다. 구하는 계산식은 식 (6)과 같다.

$$T_v = \frac{1}{2} \text{MAX}(v[x] | 1 \leq x \leq m)$$

$$T_h = \frac{1}{2} \text{MAX}(h[x] | 1 \leq x \leq n) \quad (6)$$

수평으로 투영한 벡터  $h$ 에서 위쪽과 아래쪽부터  $T_h$  값을 초과하는 값,  $h(\text{top})$ 와  $h(\text{bottom})$ 를 찾는다. 또한 수직으로 투영한 벡터  $v$ 에서도 마찬가지로 왼쪽과 오른쪽에서  $T_v$ 를 초과하는 값,  $v(\text{left})$ 와  $v(\text{right})$ 를 찾는다. 찾아진 값들은 후보 영역의 크기를 조정하는데 사용된다. 그림 3은 정확한 후보 영역을 검출하는 후처리 과정을 나타낸다.

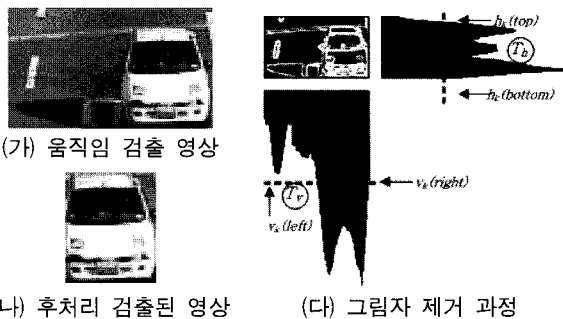


그림 3. 후처리 과정을 통한 그림자 및 배경 영역 제거  
Fig. 3. Remove object shadow and background using the post-processing.

### III. 웨이블릿 신경망 기반의 물체 인식

물체의 움직임에 의해 검출된 후보영역들 중에서 관심 있는 물체가 포함되어 있는 후보영역만을 판별하기 위해 후보영역에 포함된 특징 정보를 기반으로 하여 환

경 변화에 덜 민감하고, 실시간 처리가 가능한 물체 인식 방법이 필요하다. 제안한 방법의 적용 분야는 실시간 처리 시스템으로써,  $m \times n$  크기를 가지는 영역들은 인식하기 위해 후보영역 내에 포함된 특징 정보를 최대한 유지하면서 특징 정보의 차원을 줄일 수 있는 방법이 요구된다. 따라서 본 단계에서는 물체 인식을 위해 영상의 고유한 정보를 유지하면서 다해상도(multiresolution) 표현이 가능한 웨이블릿 변환 방법을 사용한다. 본 단계의 물체 인식 과정은 해당 후보영역으로부터 웨이블릿 변환을 통해 다해상도 영상으로 표현한 후, 다해상도로 표현된 영상들 중에서 최소한의 정보를 유지할 수 있는 해당 레벨의 다해상도 영상내의 웨이블릿 계수를 추출한다. 추출된 웨이블릿 계수는 많은 특징을 가진 정보로써 이들 사이를 각 특징별로 선형으로 분리가 가능하지 않는 문제를 가지고 있다. 이를 해결하기 위해 여러 개의 퍼셉트론(perceptron)을 여러 층으로 연결하여 비선형 분리가 가능한 신경망을 사용한다.

#### 가. 웨이블릿 변환

웨이블릿 변환은 최근 들어 신호 처리나 패턴 인식 분야에서 가장 많이 쓰이고 있는 다해상도 변환 방법중의 하나이다<sup>[14-15,17]</sup>. 이 변환 방법은 공간 영역과 주파수 영역에서 그 지역적 특성을 잘 나타낸다는 특성으로 인해 어떤 정보를 얻고자할 때 많이 쓰이고 있다. 웨이블릿 변환은 공간과 주파수의 두 영역에서 영상을 표시할 수 있기 때문에 영상에 에너지등과 같은 공간적 특성과 저주파수 영역에 에너지가 밀집되어 있는 영상의 주파수적 특성을 보다 효율적으로 나타낼 수 있는 특징이 있다. 게다가, 도로 영상과 같이 환경오염에 의해 노이즈 포함된 영상의 경우 노이즈의 영향을 최소화할 수 있는 장점 역시 가지고 있다. 웨이블릿 변환에 의한 영상의 특징 추출 방법은 영상을 주파수 대역별로 분할하고, 분할된 부 밴드들을 이용하여 한 영상을 대표하는 특징 벡터를 얻는다. 분할된 저대역(lowpass) 부밴드들은 원 영상이 가지는 에너지의 대부분을 소유하고 있으며, 그리고 고대역(highpass) 부밴드(sub-band)는 집중도가 떨어지지만 인간 시각에 민감한 영상의 에지 정보를 보유하고 있다<sup>[14, 16]</sup>.

따라서 본 단계에서 자동차 영상이 가진 특징 정보의 복잡도를 줄이기 웨이블릿 변환을 통한 얻어진 웨이블릿 계수를 자동차 인식을 위한 특징값으로 사용한다. 본 연구에서 사용한 것은 Haar 웨이블릿 변환을 이용하여 영상을 분해함으로써, 각 부밴드 영상의 해상도를

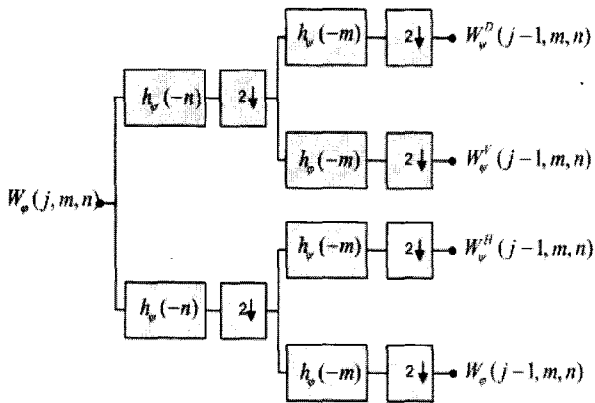


그림 4. 웨이블릿 변환 과정  
Fig. 4. Process of the wavelet transform.

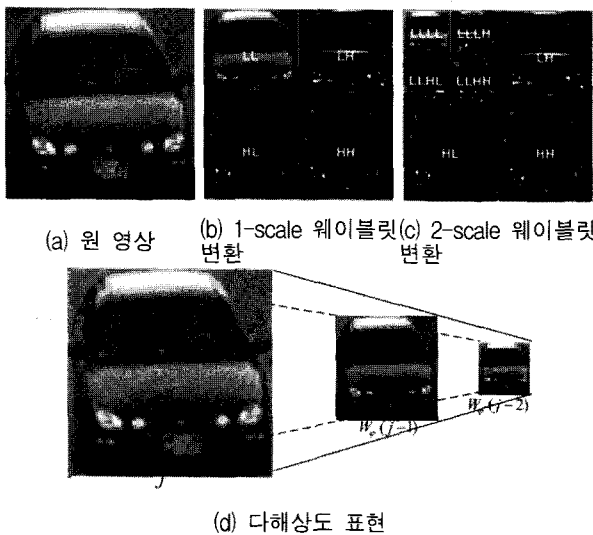


그림 5. 영상의 웨이블릿 변환과 다해상도 표현  
Fig. 5. Image representation of Wavelet transform and multiresolution.

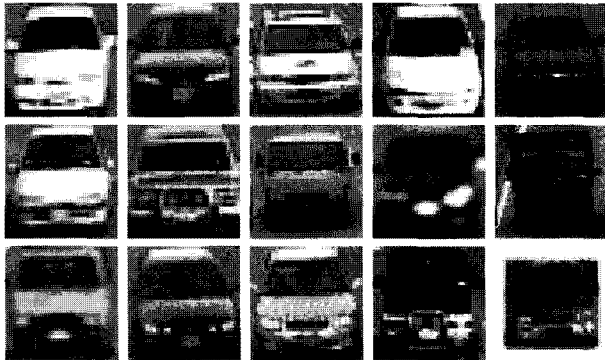
줄일 수 있고, 또한 그 결과 계산의 복잡성을 낮은 저해상도의 영상에서 작업함으로써 영상의 고유한 성질을 내포함과 동시에 작업 시간의 소모를 줄일 수 있는 이점을 가지고 있다. 그림 4는 웨이블릿 변환 과정을 그림으로 도식화한 것이다.  $m \times n$  화소 크기의 원영상  $f(x,y)$ 를 입력으로 하고 이를  $\varphi(X)$ 로 정의되는 mother wavelet을 변이시키고 확대/축소시킴으로써 얻어지는 함수들의 집합을 웨이블릿이라고 하며, 원영상  $W_\varphi(j,m,n)$ 를 수평, 수직 다운 샘플링(down-sampling) 함수  $h_\varphi(-n)$ 와  $h_\psi(-n)$ 에 의해 4개의 서브 밴드로 분리된 웨이블릿 변환 영상,  $W_\varphi(j-1,m,n)$ ,  $W_\psi^H(j-1,m,n)$ ,  $W_\psi^V(j-1,m,n)$ , 그리고  $W_\psi^D(j-1,m,n)$ 으로 분리된다.

그림 5는 입력 영상을 1, 2-level 웨이블릿 변환에 의

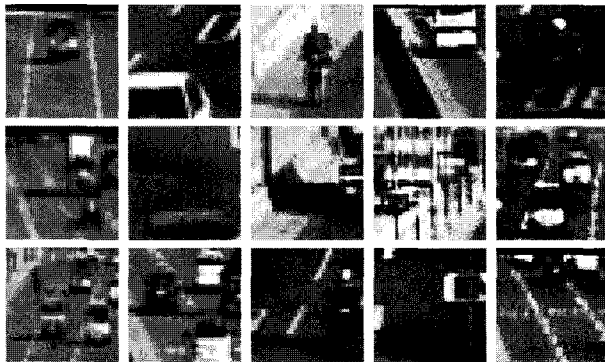
해 생성된 서브 밴드 대역들의 영상들과 다해상도로 표현한 그림이다. 웨이블릿 변환을 2차원적으로 적용할 때, 영상의 크기가  $m \times n$ 이라면 최종적으로 네 개의  $m/2 \times n/2$  크기의 분할된 영상(LL, LH, HL, HH)이 분석 결과로 나온다. 여기서 LL은 저주파성분, LH는 영상의 수평, HL은 수직, 그리고 HH는 대각선 고주파성분을 담고 있다. 그림 5와 같이  $32 \times 32$  화소 크기의 자동차 영상에 2-level 웨이블릿 변환을 적용한 그림으로써 웨이블릿 변환을 통한 각 서브 밴드 부분에는 자동차의 특징적인 에지 정보를 담고 있는 것을 알 수 있다. 특히 웨이블릿 서브 밴드 부분에 수평 성분은 자동차의 루프와 아래쪽, 그릴의 윗부분, 범퍼, 그리고 대각 성분은 자동차 몸체의 코너 부분을 나타낸다.

나. 인식단계

본 단계에서는 이동물체의 움직임에 의해 검출된 후보 영역들 중에서 자동차가 포함된 영역들을 인식하기 위해 Haar 웨이블릿 변환을 통해 얻은 웨이블릿 계수를 입력으로 하는 신경망을 이용한다.  $m \times n$  화소 크기의 후보 영역내의 모든 화소값을 이용하는 것이 아니라, 2-level 웨이블릿 변환된 각 서브밴드 부분의 웨이블릿 계수를 사용한다. 이전 단계에서 설명한 웨이블릿의 특성을 이용하여 추출된 영역에서 웨이블릿 변환을 통한 계수들을 신경망의 입력으로 사용함으로써, 입력 패턴의 차원을 감소시키고, 동시에 신뢰성 있는 물체 인식이 가능하도록 한다. 본 연구에서는 다층 퍼셉트론(MLP: Multi Layer Perceptron)을 사용하고, 이 MLP는 비선형 활성화함수인 시그모이드(sigmoid) 함수를 가지고 있고, 은닉 층으로 불리는 중간층을 가지며, 각 층의 뉴런들은 인접한 층의 다른 모든 뉴런들과 상호 연결되어 있다는 특징으로 인해 다양한 특징 정보를 지닌다. 그리고 MLP의 학습은 오류 역전파 방식(error-back propagation)에 의해 수행된다. 학습에 사용한 영상은  $32 \times 32$  화소 크기의 그레이(gray) 색상의 자동차의 정면 영상이며, 자동차가 포함된 영상 3400개, 그렇지 못한 영상 2600개를 사용하였다.  $32 \times 32$  화소 크기 영상을 2-level 웨이블릿 변환을 적용하여  $16 \times 16$  화소 크기의 웨이블릿 변환된 영상(LLLL, LLHL, LLLH, LLHH 서브 밴드)에서 수직, 수평 대각선 성분의 계수값 40개씩을 추출하여 신경망의 학습을 위한 입력으로 사용하였다. 그림 6은 신경망 학습에 사용된 예제 영상이다. 영상에서 자동차의 크기는 영상 획득 장치와 자동차 사이의 거리에 의해 다르게 나타난다. 따라서 검출된



(가) 자동차 영상



(나) 비자동차 영상

그림 6. 신경망 학습에 사용한 예제 영상  
Fig. 6. Example images for training NN.

후보 영역들을 신경망에 입력하기 위해 정보 손실을 최소화하면서 크기 정규화를 적용해야 한다. 그러므로 검출된 후보영역을  $32 \times 32$  화소 크기로 정규화한 후, 2-level 웨이블릿 변환을 통해 각 서브 밴드의 계수들을 신경망에 입력하여 검출된 후보영역에 자동차가 포함된 영상인지를 판별한다. 그림 7은 자동차 인식을 위해 웨이블릿 변환 계수들을 입력으로 하는 신경망의 구조이다. 신경망의 입력 노드 수는 입력 특징값의 개수에 의해 결정되고, Hidden 노드의 수는 신경망 학습에 의해 결정된다.

신경망 학습은 예제 샘플들의 인식 결과와 실제 결과의 차이인 에러값이 최소가 되는 방향으로 진행된다. 본 단계에서 사용한 신경망 알고리즘의 순서는 다음과 같다. 먼저 출력층이  $k$ 개, 은닉층이  $j$ 개, 입력층이  $i$ 개로 구성되어 있을 때, 네트워크의 연결 강도를 결정하는 연결 강도  $w$ 는 랜덤하게  $-0.5$ 부터  $+0.5$  할당하고 오프셋(offset) 역시 랜덤하게  $-0.05$ 부터  $+0.05$ 까지 할당한다. 학습을 위해  $32 \times 32$  영상을 2-level 웨이블릿 변환에 의해 변환된  $8 \times 8$  화소 크기의 4개의 서브 밴드 영역,  $W_{\varphi}(j-2, m, n)$ ,  $W_{\psi}^H(j-2, m, n)$ ,  $W_{\psi}^V(j-2, m, n)$ ,

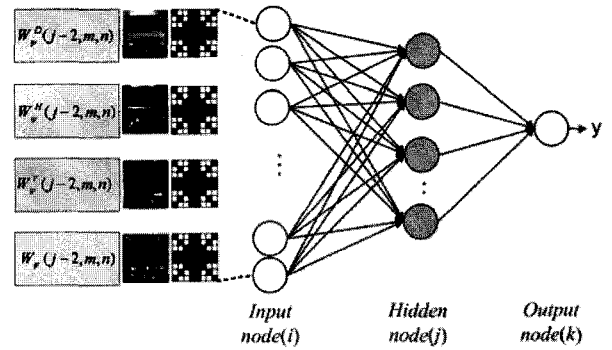


그림 7. 신경망 구조도

Fig. 7. The structure of neural network.

그리고  $W_{\psi}^D(j-2, m, n)$ 에 수직, 수평, 그리고 대각선 성분들을 입력한다. 그리고 은닉층  $j$ 번째의 유니트에 대한 입력의 합은  $N_{oj} = \sum W_{ji} O_{oi}$ 이고, 노드  $j$ 번째의 출력은  $O_{oj} = f(N_{oj})$ 이다. 여기서  $f$ 는 시그모이드 활성화 함수(activation function)이므로  $O_{oi} = 1/(1 + \exp(-O_{oj}))$ 가 된다. 또한 출력층  $k$ 번째의 유니트에 대한 입력의 합은  $N_{ok} = \sum W_{kj} O_{oj}$ 이고,  $k$ 번째 유니트의 출력은  $O_{ok} = f(N_{ok})$ 가 된다. 이러한 학습 과정을 통해 출력된 출력층의 값과 실제 학습 패턴의 원하는 값과는 차이가 있기 때문에 이 차이인 에러를 최소화하는 연결 강도  $w$ 값을 반복 학습을 통해 수정한다.

#### IV. 웨이블릿 불변 모멘트 기반의 물체 추적

본 연구에서는 실시간으로 처리해야 함으로 간단하면서도 효율적인 웨이블릿 불변 모멘트를 이용하여 이동물체를 추적한다. 본 단계에서는 이전 프레임에서 검출된 영역과 현재 영상에서 검출된 영역과의 유사도 계산을 통해 물체 추적을 수행하는데, 이를 위해 사용한 특징 정보는 영상해석에서 유용한 특징 정보중에 하나인 불변 모멘트를 이용한다. 불변 모멘트 특징 정보는 주요 객체의 크기 및 회전, 이동에 불변한 특성으로 인해 물체를 구분하는데 좋은 정보이며, 패턴 인식 분야에 물체를 기술하는데 있어 우수한 특성을 가지는 것으로 알려져 있다<sup>[15]</sup>. 2차원 연속함수  $f(x, y)$ 를 갖는 웨이블릿 변환된 영상의 웨이블릿 계수에서, 함수  $f(x, y)$ 에 의해  $(p+q)$ 차수의 모멘트  $m_{pq}$ 는

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \quad p, q = 0, 1, \dots \quad (7)$$

,이고 중앙모멘트(central moment)  $\mu_{pq}$ 는

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \hat{x})^p (y - \hat{y})^q f(x, y) dx, dy \quad (8)$$

$$\hat{x} = \frac{m_{10}}{m_{00}}, \hat{y} = \frac{m_{01}}{m_{00}}$$

이다.  $\eta_{pq}$ 로 정의하는 정규화된 중앙모멘트는

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^\gamma} \quad \gamma = \frac{p+q}{2} + 1, (p+q) = 2, 3, \dots \quad (9)$$

이다. 위의 2, 3차 모멘트로부터 해당 영역의 크기, 회전 등에 불변한 특징을 가지는 7개의 모멘트 텀을 식 (10) 과 같이 계산한다.

$$\begin{aligned} m_1 &= \eta_{20} + \eta_{02} \\ m_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ m_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ m_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ m_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ &\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ m_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ &\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ m_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ &\quad + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (10)$$

식 (10)의 불변 모멘트 값을 가지고 유사도 계산은 이전 프레임에서 검출된 영역들과의 현재 프레임에서 검출된 물체 영역의 특징 벡터를  $m1 \sim m7$  대해 유클리디안 거리(euclidean distance)를 사용하여 거리가 최소인 현재 프레임의 영역을 이전 프레임의 영역에 대응시킴으로써 추적을 수행한다.

### V. 실험 결과

제안한 방법을 실험하기 위해 일반 도로와 고속도로, 그리고 건물의 옥상에서 촬영한 도로 영상등 서로 다른 시간대에 획득한 5,300여대의 자동차가 포함된 영상으로 실험하였다. 영상 획득을 위해 Sony사의 DCR-PC3 디지털 카메라를 사용하였고, 팬티엄4 IBM호환 PC에서 윈도우 XP 환경(CPU: 3.2GHz, RAM:2GM)의 VC++언어로 구현하였다. 입력 영상은 초당 8 프레임이고, 320×240 화소 크기의 칼라 영상이다. 영상내에 자동차의 크기는 약 23×19에서 72×67 화소 크기이다. 그림 8은 제안한 방법을 적용한 시스템의 실험 결과가 표시된

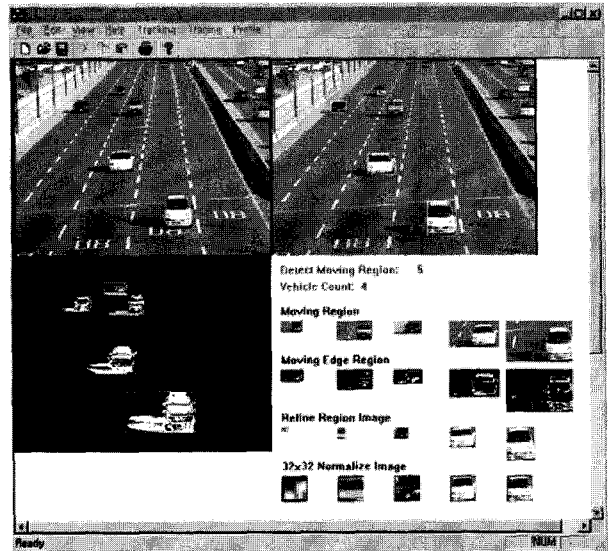


그림 8. 제안한 방법을 적용한 시스템의 인터페이스  
Fig. 8. Systeme interface of the proposed system.

인터페이스 그림이다. 인터페이스 그림에서 위쪽은  $t-1$ 와  $t$ 시간에 획득한 영상이고, 왼쪽아래는 두 영상에 적응적 임계치 방법을 통해 움직임 후보 영역을 검출한 결과 영상이다. 오른쪽 하단은 검출된 후보 영역들에 후처리 과정과 신경망 입력을 위해 크기를 정규화한 결과 영상들이다. 신경망 학습을 위해 네트워크의 상태를 결정하는 연결강도 웨이트  $w$ 는 -0.5와 0.5 범위내에서 랜덤하게 부여하고, 오프셋(offset)은 -0.05와 0.05 범위로 랜덤하게 초기화하였다. 신경망 학습에 사용한 학습 데이터는 500개(자동차:200개, 비자동차 영상:300개)를 사용하였고, 입력 노드의 수는 160개, hidden 노드 수는 실험을 통해 최적의 인식률을 출력할 수 있는 320개를 사용하고, 출력 노드 수는 1개로 구성하였다. 그리고 신경망의 학습을 위해 학습에 사용한 학습 횟수는 신경망 출력값과 학습 값의 에러차이가 더 이상 감소하지 않는 횟수인 10,000번으로 설정하였다.

제안한 연구에 구성된 각 단계별 수행 시간은 표 1과 같다. 먼저 후보 영역 검출 단계에서 사용한 적응적 임계값은 매 프레임마다 계산하지 않으므로 미리 계산되어 있다는 가정하여 수행시간을 계산하였다. 제안한 방법의 전체 수행에 있어서 인식 단계에서 가장 많은 시간이 소모되었음을 알 수 있다. 이것은 입력 영상의 웨이블릿 변환과 함께 특징 정보 추출로 인해 시간 소모가 발생하였다. 따라서 제안한 방법은 초당 8프레임을 입력을 받지만 처리하는 데는 약 초당 5프레임 처리가 가능함을 알 수 있다.

실험을 통해 영상 내에서 움직이는 물체가 사람이나



표 1. 제안한 연구의 단계별 수행 시간

Table 1. Processing time of each step.

단계	평균 수행 시간 (sec/frame)
후보 영역 검출	0.08
물체 인식	0.13
물체 추적	0.03
전체 수행 시간	0.24

표 2. 자동차 인식 실험 결과

Table 2. Results of vehicle recognition.

실험 데이터	성공	실패	인식률(%)
Video #1 (자동차:496, 비자동차:32)	487 (465/22)	41	92.23
Video #2 (자동차:1636, 비자동차:43)	1586 (1555/31)	93	94.47
Video #3 (자동차:899, 비자동차:145)	779 (653/126)	265	74.61
Video #4 (자동차:183, 비자동차:121)	247 (136/111)	57	81.25
Video #5 (자동차:2103, 비자동차:61)	1987 (1945/42)	177	91.82
평균			86.87

혹은 이륜자동차가 포함되어 있을 때 움직임 후보 영역 검출 단계에서 이 영역들도 같이 검출되었으나, 신경망 입력 통한 후보 영역 인식 단계에서는 자동차가 포함된 영역만을 검출하였다. 표 2는 제안한 방법을 적용한 자동차 인식 결과이다.

표 2에서 실험 비디오 영상 3, 4번째에 오인식률이 높게 나온 것은 해당 비디오 영상내에서 인도에 보행자가 포함되어 있음으로 여러 명의 보행자가 동행할 경우 자동차로 오인식되는 문제가 발생되었다. 이러한 문제는 기존 연구[17]에서 제시한 바와 같이 도로 상의 차선을 검출하고 해당 차선 상에서만 움직임을 검출함으로써 비디오 영상 3과 4번의 오인식 문제를 해결할 수 있을 것이다. 따라서 인도가 포함된 실험 결과를 제외하면 평균 인식률은 약 92.8%의 자동차 인식률을 보인다.

가. 후보영역 검출 결과

입력 영상에서 후보영역 검출 단계를 평가하기 위해 적응적 임계치 방법과 고정 임계치 방법을 사용한 영역 검출 결과를 비교 평가 하였다. 이때 고정 임계값은 100으로 설정하고, 실험을 위해 이동물체 영역을 수동으로 검출한 다음에 실험을 통해 검출된 영역과의 위치 유사성 비교 분석을 실시하였다. 그림 9에서 실험데이

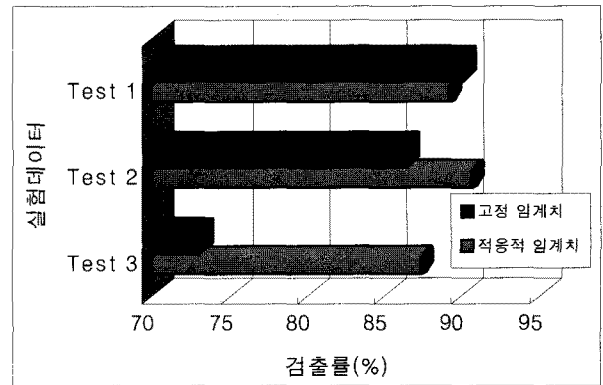


그림 9. 후보영역 검출 결과 비교

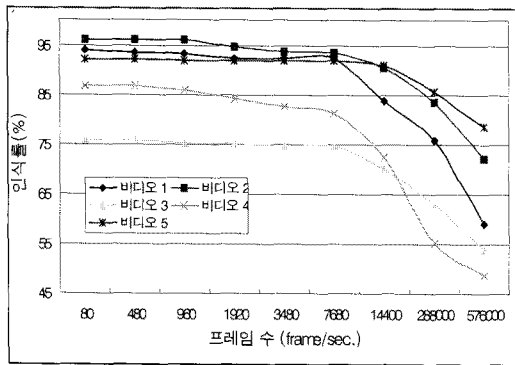
Fig. 9. Compare with results of candidate detection.

터 첫 번째의 경우는 적응적 임계치 방법보다 고정 임계치 값을 사용한 영역 검출 결과가 더 좋음을 보였다. 이유는 전역적으로 영상의 밝기 변화가 발생했기 때문에, 확률적 통계에 기반으로 한 적응적 임계치 방법은 계산의 소모로 인해 정확한 임계값 추정이 어려움을 나타내었다. 하지만 일반적으로 고정된 카메라로부터 획득된 영상의 변화는 전역적인 변화보다는 국소적인 변화가 대부분이기 때문에 적응적 임계치 방법이 고정 임계치 방법보다 효율적임을 알 수 있다.

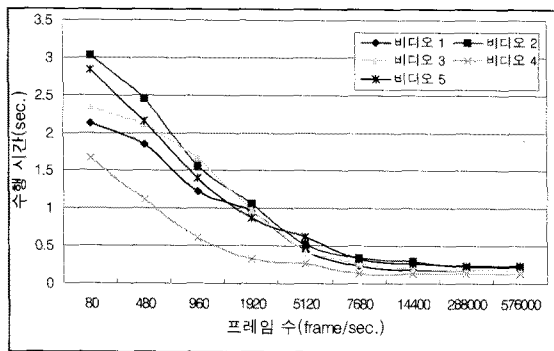
실험 결과를 통해 적응적 임계치 방법을 후보 영역 검출에 사용함으로써 고정적 임계치보다 검출율이 보다 뛰어남을 알 수 있다. 다만 환경의 변화가 적은 곳에서는 적응적 임계값의 계산에 시간 소모로 인해 불필요할 것이다. 이러한 적응적 임계치 계산을 번 몇 프레임마다 적용하는 것이 좋은 지에 결정하기 위해 실험 비디오 영상에서 적응적 임계값 계산 주기에 따른 자동차 검출율과 수행 시간 비교를 그림 10에서 나타낸다. 그림에서 인식률의 변화는 적은 프레임 당 임계치 값을 계산함으로써 인식률이 향상됨을 알 수 있다. 하지만, 자른 임계치 계산으로 인해 수행 시간이 많이 소모된다. 따라서, 인식률과 수행 시간의 타협 프레임 수는 약 7,680 프레임 경우, 인식률에서는 이 기점을 정적으로 인식률이 급격히 감소되었으며, 수행 시간 역시 이 기점 이후에는 수행 시간의 감소가 더딘 점으로 인해 적응적 임계값 계산은 매 7,680 프레임, 즉 16분마다 한 번씩 수행하였다.

나. 물체 인식 결과

후보 영역 검출 단계에 의해 검출된 후보 영역들중에서 자동차가 포함된 영역을 인식하기 위해 본 연구에서는 웨이블릿 계수를 입력하는 하는 신경망을 사용하여



(가) 인식률 변화



(나) 수행 시간 변화

그림 10. 입력 프레임당 적응적 임계치 계산에 따른 인식률과 수행 시간 비교

Fig. 10. Recognition rates and performance time according to the adaptive thresholding value per frame.

자동차가 포함되어 있는지 인식하였다. 이 방법의 장점은 노이즈에 강인함과 계산 시간의 단축으로 인해 실시간 처리 시스템에 적용이 가능한 점이다. 먼저 노이즈의 강인함을 평가하기 위해 가우시안(Gaussian) 노이즈가 포함시킨 후보 영역에서 밝기 값과 주성분분석(PCA)을 통해 고유벡터(eigenvector) 영상, 그리고 웨이블릿 계수 추출하여 신경망의 입력으로 사용하여 자동차 인식 결과와 함께 수행 시간을 비교하였다. PCA 정보는 다차원 변수들을 축소, 요약하는 차원의 단순화와 더불어 일반적으로 서로 상관되어 있어 변수 상호간의 복잡한 구조를 효율적으로 분석하는데 사용됨으로써 웨이블릿 계수들과의 성능 비교에 적용하였다. 실험을 위해  $m \times n$  화소 크기의 후보 영역을  $16 \times 16$  화소 크기로 정규화 시킨 수직, 수평, 대각성분을 추출하여 160개의 밝기값 정보와 PCA 고유벡터 영상 정보, 그리고 웨이블릿 계수를 추출하였다. 표 3은 가우시안 노이즈를 랜덤하게 영상내에 포함시킨 후 여러 특징 정보를 기반으로 한 물체 인식 결과를 비교한 것이다. 실험한 결과 웨

표 3. 특정 정보에 노이즈 강인함 실험 결과

Table 3. Robustness of noise according to features.

Test data	Noise(%) Features	0	5	10	15	20
		Video #1	Intensity	82.77	80.30	72.92
	PCA	85.42	85.23	78.03	75.19	68.93
	Wavelet	92.23	91.10	83.33	79.92	74.05
Video #2	Intensity	89.76	88.03	79.10	70.46	64.86
	PCA	91.19	90.00	87.37	82.73	78.14
	Wavelet	94.47	93.63	88.92	86.54	82.25
Video #3	Intensity	72.51	69.16	65.71	60.06	53.26
	PCA	71.93	70.31	68.20	63.60	58.72
	Wavelet	74.61	72.61	69.06	66.86	59.48

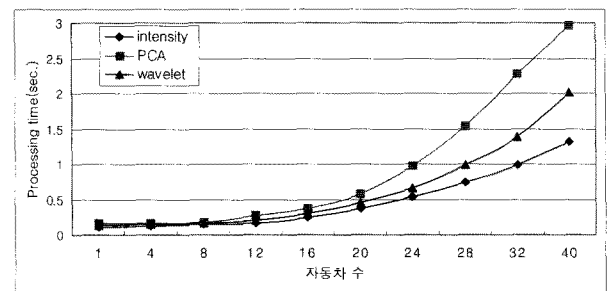


그림 11. 영상내 포함된 자동차 수에 따른 특징 정보별 수행 시간 비교

Fig. 11. Flowchart of the proposed system.

이블릿 계수를 인식을 위한 특징 정보로 사용함으로써 밝기값 정보와 PCA 고유벡터 영상 정보보다 물체 인식에 있어서 노이즈에 강인함을 볼 수 있다.

그림 11은 실험에서 사용한 특징 정보에 따른 수행 시간을 비교하였다. 실험 결과 영상내 포함된 주행중인 자동차의 수가 많아질수록 수행시간이 기하급수적으로 증가함을 알 수 있다. 이것은 이동물체의 수가 많아짐에 따라 후보 영역이 많이 검출되고 이에 따른 특징 추출과 인식에 많은 수행 시간이 필요한 것을 의미한다. 실험 결과 그림 11에서 밝기값 특징 정보를 사용한 인식의 수행 시간을 다른 정보들 보다 정보의 가공 단계가 필요치 않으므로 비교적 빠른 수행 시간을 보였지만, 표 3에서와 같이 인식 결과가 저조함을 알 수 있다. 따라서 웨이블릿 정보를 이용하여 물체 인식에 사용한 결과 수행 시간 및 인식 결과가 비교적 좋음을 알 수 있다.

다. 물체 추적 결과

물체 인식 단계에서 인식된 영역들을 추적하기 위해

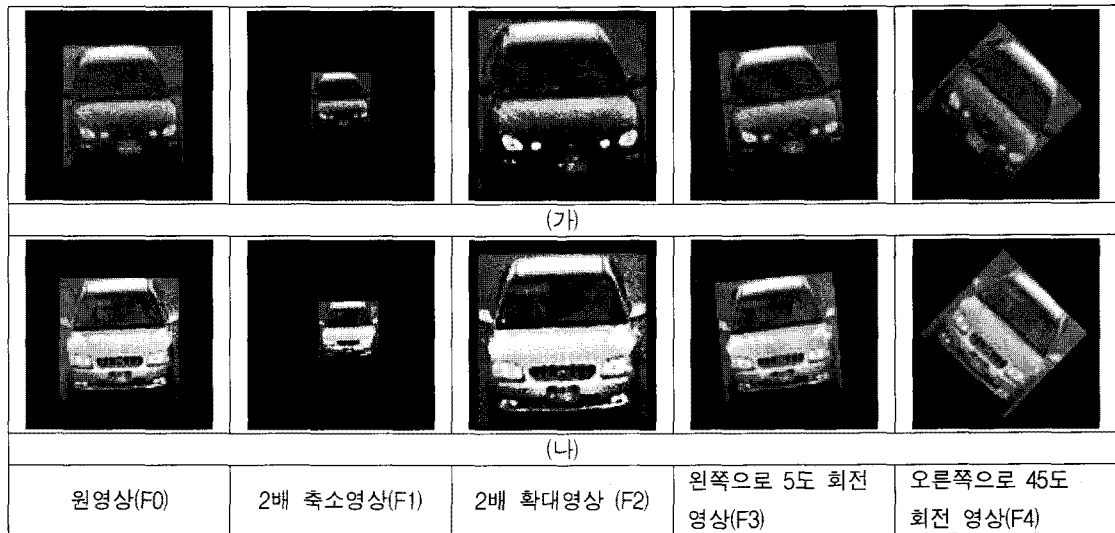


그림 12. 불변 모멘트의 강인함 실험을 위한 변환 영상  
 Fig. 12. Transformed image for experiment of invariant moments.

표 4. 그림 12의 7개 불변 모멘트 값  
 Table 4. Invariant moments of Fig. 12.

모멘트		$m_1$	$m_2$	$m_3$	$m_4$	$m_5$	$m_6$	$m_7$
그림 12(가)	F0	6.6701	20.4295	26.8532	26.2848	53.4558	36.6806	53.0279
	F1	6.6623	20.3711	26.8825	26.2612	53.4531	36.6138	53.0038
	F2	6.6681	20.4271	26.8511	26.2686	53.4323	36.6612	53.0059
	F3	6.6636	20.4339	26.9025	26.2375	53.6298	36.6418	52.9148
	F4	6.6634	20.4252	26.8448	26.2551	53.3996	36.6461	52.9866
그림 12(나)	F0	6.8703	19.9073	28.6234	29.4273	58.7908	39.6708	58.8078
	F1	6.8685	19.9117	28.6467	29.4210	58.8080	39.6732	58.7950
	F2	6.8673	19.9429	28.6050	29.3472	58.6005	39.5711	58.7503
	F3	6.8647	19.8761	28.4718	29.4889	58.8255	39.6635	58.8065
	F4	6.8636	19.8774	28.6102	29.4016	58.7651	39.6232	58.7434

본 연구에서는 웨이블릿 계수들의 불변 모멘트 값 매칭을 통해 연속된 이미지상에 존재하는 이동물체들을 추적한다. 제안한 방법의 효율성을 검증하기 위해 검출된 자동차 예제 영상 2개를 크기, 회전, 이동 변환을 통해 그림 12와 같이 생성하였다. 표 4는 그림 12와 같이 생성된 영상들과 원영상과의 7개의 모멘트 값을 계산한 것이다. 실험 결과에서 알 수 있듯이 원영상의 크기 및 회전 등의 변환에 의해서도 같은 영상일 경우 유사한 모멘트 값을 가지는 것을 알 수 있다. 따라서 물체 추적을 위해 웨이블릿 불변 모멘트 기반의 방법을 사용함으로써 이동물체의 크기, 회전, 그리고 이동변환에도 올바르게 추적할 수 있음을 알 수 있다.

### VI. 결 론

본 논문은 도로 교통 감시 시스템을 위한 실시간 이

동물체 검출, 인식 및 추적 방법을 제안하였다. 연속적으로 입력되는 영상에서 움직임이 있는 영역 검출을 위해 적응적 임계치를 사용한 차영상 분석 방법을 사용하고, 검출된 후보 영상에 자동차가 포함되어 있는지 판별하기 위해 웨이블릿 변환 계수들을 입력으로 하는 신경망을 이용하였다. 그리고 판별된 자동차 영역들을 추적하기 위해 웨이블릿 계수에 기반으로 한 모멘트 값을 이용하였다. 일반 도로상에서 여러 시간대에 연속적으로 움직이는 자동차 영상으로 실험해 본 결과, 자동차 검출율은 92.4%이고, 프레임당 처리 시간 0.24초미만의 결과를 얻었다. 본 논문에서 제안한 방법은 잡음에 의해 저하된 영상에 상대적으로 좋은 성능을 보였고, 실시간 처리를 가능하게 함으로써 교통 감시 시스템에 유용하게 적용될 수 있다. 향후 연구로는 전체적인 도로 교통 감시 시스템을 구현에 관해 연구하고자 한다.

## 참고 문헌

- [1] J. Zhou, D. Gao, D. Zhang, "Moving Vehicle Detection for Automatic Traffic Monitoring", IEEE Trans. on Vehicular Technology, Vol. 56, no. 1, pp. 51-59, 2007.
- [2] T. Alexandropoulos, S. Boutas, V. Loumos, E. Kayafas, "Real-time change detection for surveillance in public transportation", IEEE Conference on Advanced Video and Signal Based Surveillance, pp. 58-63, 2005.
- [3] J. C. McCall, M. M. Trivedi, "Video-based lane estimation and tracking for driver assistance: survey, system, and evaluation", IEEE Trans. on Intelligent Transportation Systems, Vol. 7, no. 1, pp. 20-37, 2006.
- [4] R. Cucchiara, C. Grana, M. Piccardi, A. Prati, "Detecting moving objects, ghosts, and shadows in video streams", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 25, no. 10, pp. 1337-1342, 2003.
- [5] R. Cucchiara, M. Piccardi, M. P. Mello, "Image analysis and rule-based reasoning for a traffic monitoring system", IEEE Trans. on Intelligent Transportation Systems, Vol. 1, no. 2, pp. 119-130, 2000.
- [6] G. L. Foresti, "Object recognition and tracking for remote video surveillance", IEEE Trans. on Circuits and Systems for Video Technology, Vol. 9, no. 7, pp. 1045-1062, 1999.
- [7] G. L. Foresti, V. Murino, C. Regazzoni, "Vehicle recognition and tracking from road image sequences", IEEE Trans. on Vehicular Technology, Vol 48, no. 1, pp. 301-318, 1999.
- [8] M. Betke, E. Haritaoglu, L. S. Davis, "Highway scene analysis in hard real-time", IEEE Conference on Intelligent Transportation System, pp. 812-817, 1997.
- [9] A. Talukder, S. Goldberg, L. Matthies, A. Ansar, "Real-time detection of moving objects in a dynamic scene from moving robotic vehicles", IEEE/RSJ International Conference on Intelligent Robots and Systems, Vol. 2, pp. 1308-1313, 2003.
- [10] W. Junwen, Z. Xuegong, "A PCA classifier and its application in vehicle detection", IEEE International Joint Conference on Neural Networks, Vol. 1, pp. 600-604, 2001.
- [11] P. R. Liu, M. Q.-H. Meng, P. X. Liu, "Moving object segmentation and detection for monocular robot based on active contour model", Electronics Letters, Vol 41, no. 24, pp. 1320-1322, 2005.
- [12] J. B. Kim, K. K. Kim, H. J. Kim, "Learning-Based Approach For License Plate Recognition", Proceeding of The 1th KISPS Summer Conference, Vol. 1, no. 1, pp. 273-276, 2000.
- [13] J. B. Kim, H. J. Kim, "Efficient Region-Based Motion Segmentation for a Video Monitoring System", Pattern Recognition Letter, Vol. 24, no. 1, pp. 113-128, 2003.
- [14] J. B. Kim, H. J. Kim, "Multiresolution-Based Watersheds for Efficient Image Segmentation", Pattern Recognition Letter, Vol. 24, no. 1, pp. 473-488, 2003.
- [15] R. C. Gonzalez, Digital Image Processing, Prentice Hall, 2004.
- [16] C. E. Bae, J. B. Kim and H. J. Kim, "Moving Object Segmentation Using Adaptive Thresholding and K-Means Clustering", the KISS. Fall Workshop on CVPR, pp. 23-24, 2001.
- [17] J. B. Kim, C. W. Lee, K. M. Lee, T. S. Yun, H. J. Kim, "Wavelet-based vehicle tracking for Automatic Traffic Surveillance", IEEE TENCON'01, Vol. 1, pp. 313-316, 2001.

## — 저 자 소 개 —



김 종 배(정회원)

2000년 부산대학교

컴퓨터공학과 공학사

2002년 경북대학교

컴퓨터공학과 공학석사

2004년 경북대학교

컴퓨터공학과 공학박사

2006년~현재 서울디지털대학교 컴퓨터공학부  
조교수

&lt;주관심분야 : 인공지능, 영상처리, 패턴인식&gt;