

적응적 임계치와 웨이블릿 변환을 이용한 움직이는 물체 검출 및 판별

박혜선^{*} 김중배 이창우 이경미^{*} 김항준^{*}
경북대학교 컴퓨터공학과 인공지능연구소, 구미 1 대학 컴퓨터정보계열^{*}
(hspark^{*}, kjblove, cwlee, kimhj)^{*}@ailab.knu.ac.kr
kmlee@kumi.ac.kr^{*}

Moving Object Detection and Classification using Adaptive thresholding and Wavelet Transform

Hye-Sun Park^{*} Jong-Bae Kim Chang-Woo Lee kyung-Mi^{*} Lee Hang-Joon Kim

Dept. of Computer Engineering, Kyungpook National University

Dept. of Computer Science, Gumi College^{*}

요 약

본 논문에서는 실제 도로 영상에서 움직이는 물체를 검출하고 판별하기 위한 새로운 방법을 제안한다. 제안된 방법은 연속된 영상의 차영상에 적응적 임계값을 적용하여 움직임이 있는 후보 영역을 검출한다. 검출된 후보 영역에 관심의 대상이 되는 물체의 포함 여부를 판별하기 위해 신경망을 사용한다. 신경망의 입력으로 사용되는 특징 벡터들의 차원을 줄이기 위해, 후보 영역의 스케일 공간 웨이블릿 특징 벡터 (scale-space wavelet feature vector)들을 사용한다. 제안된 방법은 비디오 기반의 응용 프로그램에 유용하게 이용될 수 있으며 특히, 시간에 따라 조명이 변하거나 잡음이 포함된 비디오 영상에 대해 좋은 결과를 얻을 수 있다.

1. 서 론

움직이는 물체를 검출하고 판별하는 것은 침입자 감시 시스템, 교통감시 시스템 등 비디오를 기반으로 하는 응용프로그램에서 매우 중요한 역할을 한다. 지금까지 움직이는 물체를 검출하고 인식하기 위한 많은 연구가 이루어지고 있다 [1,2]. 연구되어진 움직임을 검출하는 방법에는 차영상 (difference Image) 분석이나 특징 벡터를 이용한 방법으로, 장면 차이(frame differenceing) [3], 히스토그램의 밝기 값 변화 [4], 색상 변화 [5] 등이 있다. 또 검출되어진 물체를 판별하는 방법으로는 3 차원 모델에 기반한 방법 [6], 특징값에 기반한 방법 [7] 등이 있다. 이 같은 검출 방법은 수행 속도가 빠르다는 장점이 있지만, 고정된 임계값을 적용하기 때문에, 시간에 따라 조명이 변하거나 비디오 영상에 잡음이 있을 때 부적절하며, 판별 방법은 물체의 움직임 방향이나 속도 뿐만 아니라 다른 정보들까지 추정함으로써 계산 시간이 많이 소모된다. 그러므로, 변화하는 환경에 보다 잘 적용할 수 있고 계산되는 시간도 고려한 연구가 요구된다. 곧, 영상의 정보를 보유하면서 특징을 줄이고 동시에 잡음에도 강한 방법이 필요하다.

본 논문에서는 연속된 영상에서 움직이는 물체를 검출하고 판별하기 위한 새로운 방법을 제안한다. 움직임이 있는 후보 검출을 위해 연속된 두 영상간의 차영상에서 적응적 임계치 방법을 사용하여 움직임이 있는 후보 영역을 검출한다. 그리고 검출된 후보 영역의 물체를 판별

하기 위해 검출된 후보 영역에서 그레이 레벨의 웨이블릿 변환 계수들을 입력으로 하는 신경망을 사용하여 실제 도로에서 움직이는 자동차를 판별한다.

2. 움직이는 물체 검출

본 논문에서는 움직임이 있는 후보 영역을 검출하기 위해 연속된 두 프레임간의 차영상 분석 (frame difference analysis) 방법을 사용한다. 차영상 분석 방법은 연속된 두 프레임간의 밝기 차이를 구한 후, 임계값을 사용하여 임계값보다 낮은 밝기 차이를 가진 부분은 움직임이 없는 배경으로 구별하고 임계값보다 큰 밝기 차이를 가진 부분은 움직임이 있는 물체로써 구별한다. 이러한 임계값의 선택은 획득한 영상 안의 잡음 뿐만 아니라 시간에 따라 변하는 조명에 상당히 의존적이다. 따라서 움직임이 있는 후보 영역들을 검출하기 위한 임계값은 유동적으로 선택되어야 한다. 이를 본 논문에서는 적응적 임계치 (adaptive thresholding) 방법을 사용한다. 적응적 임계치 방법은 연속된 두 프레임 간의 밝기 차이를 세 개의 가우시안 분포 (Gaussian Distribution)로 따른다는 가정 하에 [8], 두 프레임 간의 차이값을 나타내는 히스토그램과 세 개의 가우시안 분포가 최적으로 만나는 곳이 임계값으로 결정된다. 그림 1 은 두 프레임 간의 밝기 차이값을 나타낸 히스토그램과 세 개의 가우시안 분포를 나타낸 것이다. 그림 1 에서 임계값은 3 개의 가우시안 모델과 차영상의 히스토그램 사이에서의 적합 기준 함수인, 식 (1)에

서 e 가 최소화가 되는 밝기차, d 로 결정한다.

$$e = \frac{1}{n} \sum_{d=-255}^{255} [P(d) - m(d)]^2 \quad (1)$$

$m(d)$ 는 밝기차의 확률이고 $P(d)$ 는 세 개의 가우시안 분포로 이루어진 혼합확률밀도이다.

$$P(d) \approx \begin{cases} p_1(d) = \frac{w_1}{\sqrt{2\pi\sigma_1}} \exp\left[-\frac{(d-\mu_1)^2}{2\sigma_1^2}\right] & -255 \leq d < -\theta \\ p_2(d) = \frac{w_2}{\sqrt{2\pi\sigma_2}} \exp\left[-\frac{(d-\mu_2)^2}{2\sigma_2^2}\right] & -\theta \leq d < \theta \\ p_3(d) = \frac{w_3}{\sqrt{2\pi\sigma_3}} \exp\left[-\frac{(d-\mu_3)^2}{2\sigma_3^2}\right] & \theta \leq d < 255 \end{cases} \quad (2)$$

w_i, μ_i, σ_i 는 각각 모집단의 비율, 평균, 그리고 분산을 나타낸다. $P(d)$ 는 식 (2)와 같이 세 개의 가우시안 분포값으로 근사된다.

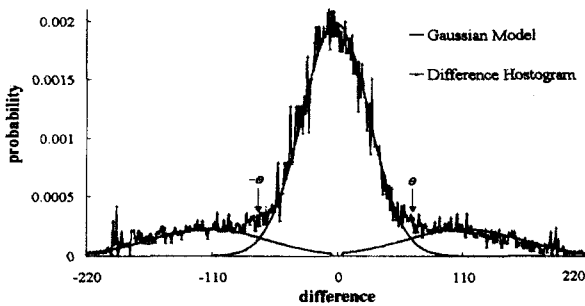


그림 1. 3 개의 가우시안 모델과 차영상의 히스토그램

식 (1)에 의해 계산된 적응적 임계값 d 를 통해 움직임이 있는 영역과 움직임이 없는 영역으로 구별시켜주는 이진화 영상인 CDM (Change Detection Mask)을 얻을 수 있으며, 식 (3)과 같이 정의 된다.

$$CDM(x,y) = \begin{cases} 1, & \text{if } |I(x,y) - I_{t-1}(x,y)| > d_t \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

식 (3)에서, $I_t(x, y)$ 와 $I_{t-1}(x, y)$ 는 t 시간에서의 현재 프레임과 이전 프레임의 픽셀의 밝기 값이며, d_t 는 t 시간에서의 적응적 임계값이다. 그림 2는 적응적 임계값을 사용하여 CDM 영상을 생성하는 과정을 보여 준다. (가)와 (나)는 연속된 입력 영상이고, (다)는 입력 영상의 차영상이다. (다)는 움직임에 의해 변화된 영역 외에, 많은 잡음을 포함하고 있음을 볼 수 있다. 이는 시간에 따른 조명 변화와 영상 내에서 생겨나는 잡음에 의해 발생한 것이다. 이러한 문제점을 해결하기 위해 적응적 임계값을 차영상에 적용한 후, 후처리 과정으로 연결성 성분 조사와 모폴로지 연산을 사용하여 CDM을 생성한다.

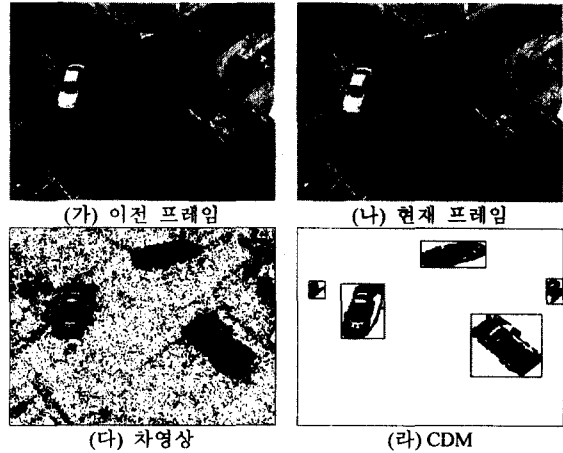


그림 2. 실제 도로 영상에서 움직이는 물체 검출 결과

3. 움직이는 물체의 판별

추출된 CDM 영상에서 물체를 판별하기 위해 신경망을 사용한다. 검출 단계에서 추출되어진 CDM에 해당하는 각 영역의 크기를 64×64 픽셀 크기로 정규화 한다. 크기를 정규화한 영상은 2-level haar 웨이블릿 변환을 통해 생성된 각 서브밴드에서 스케일 공간 웨이블릿 특징을 신경망의 입력으로 사용한다. 이러한 특징 값으로 2-level haar 웨이블릿 변환을 통한 저주파, 수평, 수직, 대각 영역의 4×4 윈도우에 대응하는 특징들의 평균과 2, 3 차 중심모멘트 계수들을 사용한다. 이 특징값들은 식 (4), (5), (6) 같이 계산한다. N 은 윈도우 크기이고, S 는 웨이블릿 변환을 통한 서브 밴드의 윈도우 영역이다.

$$M = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} S(i, j) \quad (4)$$

$$\mu_2 = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (S(i, j) - M)^2 \quad (5)$$

$$\mu_3 = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (S(i, j) - M)^3 \quad (6)$$

신경망의 학습에 사용된 알고리즘은 오류 역 전파 학습 (standard error back-propagation) 알고리즘이며, 입력노드 192, 히든 노드 20, 출력 노드 1 개로 구성된다. 제안한 방법에서 자동차를 판별하기 위해, 신경망의 학습으로는 523 개의 자동차가 포함된 영상과 740 개의 자동차가 포함되지 않은 영상들을 사용하였다. 물체를 판별하기 위해 사용되어지는 스케일 공간 웨이블릿 특징을 이용한 신경망의 구조는 그림 3과 같이 나타난다

4. 실험 결과

실험에 사용된 영상은 다양한 움직임 있는 물체들과 많은 조도 변화, 영상내의 잡음을 포함하고 있다. 실험된 영상은 고정된 초점에서 초당 5 프레임으로 획득된 320×240 크기의 실제 도로의 영상들이다.

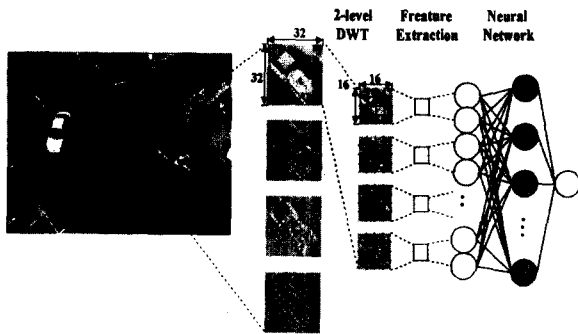


그림 3. 스케일 공간 웨이블릿 특징을 이용한 신경망의 구조

실험은 펜티엄 IV 1.4Ghz PC 에서 Matlab 5.3 으로 하였다. 한 프레임 당 평균 처리 시간은 0.19 초이다. 제안한 방법은 물체를 검출하고 판별하는 시간을 줄이는데 매우 효과적이다. 제안한 방법의 효율성을 알아 보기 위해, 다른 방법과 비교해 보았다. 첫 번째 비교한 방법은, Papageorgiou [1]가 개발한 것으로, 이 방법은 물체를 판별하기 위해 웨이블릿 특징 벡터들을 이용하여 SVM (support vector machine)을 사용한다. 두 번째 비교한 방법은, 물체를 판별하기 위해 목표 형상 파라미터들의 분산값 [2]을 사용한다. 표 1 은 실제 도로 영상 안에서, 제안된 방법과 두 가지 다른 방법의 판별 성공률을 비교해 본 결과이다. 차영상에 따른 잡음 비율값 (SNR)은 가우시안 잡음을 나타낸다.

표 1. 판별률과 그 평균시간

SNR [db]	판별률(%)			평균계산시간(s)
	0	5	10	
제안한 방법	96.5	93	84	0.19
Lipton's	96.2	93	79.8	0.45
Papageorgiou's	95	89.4	83.6	1.27

표 1 을 비교 분석해 본 결과, 잡음이 많아질수록 판별률이 낮아 지는 것은 세 가지 모두 같으나, 제안된 방법에서의 계산된 시간의 평균과 판별률은 대략 다른 방법보다 빠르다는 것을 알 수 있었다. 제안된 방법으로 실제 도로 상의 움직이는 자동차를 검출하고 판별해 보았다. 그림 4 는 제안한 방법의 실험 결과를 보여 준다. 실험 결과 움직이는 자동차를 잘 검출하고 판별하였다. 그러나 움직이는 자동차가 아닌 움직임이 있는 영역도 자동차와 특징이 유사할 경우, 잘못 판별되어지기도 하였다.

5. 결론

본 논문에서는 연속된 영상에서 움직이는 물체를 검출하고 판별하기 위한 새로운 방법을 제안한다. 실험에 사용된 실제 도로 상에서 움직임이 있는 영역을 검출하고, 검출된 영역에 관심의 대상이 되는 물체가 포함되어 있는지를 판별하였다. 움직임이 있는 영역을 검출하기 위해 적응적 임계값을 사용하여 연속적으로 입력되는 영상

에서 움직이는 물체들을 검출하였다. 또한, 검출된 후보 영역에 자동차가 포함되어 있는지 판별하기 위해, 움직이는 영역의 스케일 공간 웨이블릿 변환 계수들을 사용하였다. 자동차 판별 단계에서 웨이블릿 특징을 사용한 것은 매우 효과적일 뿐 아니라, 계산 시간도 감소하였다. 제안된 방법은 자동 비디오 감시 시스템에 유용하게 적용될 수 있다. 향후 보다 효과적인 특징추출을 통하여, 계산량을 줄이고 동시에 판별률을 높이는 연구를 하고자 한다.

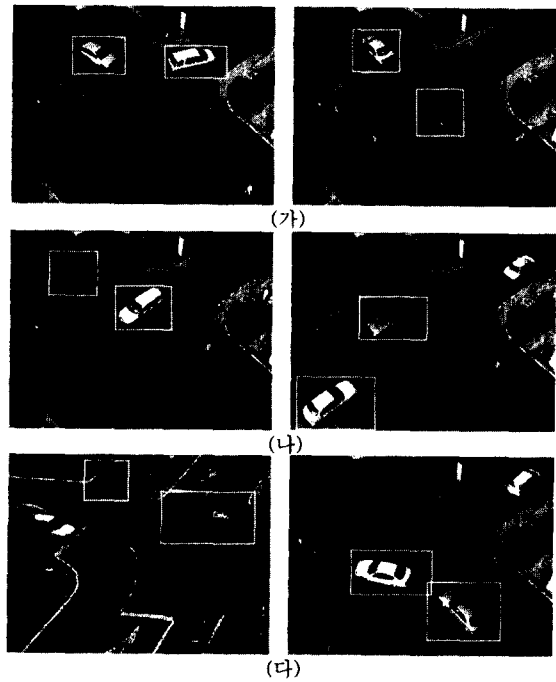


그림 4. 자동차 영역 판별 결과

6. 참고문헌

1. Papageorgiou, C. P., and Poggio, T.: 'A Trainable Object Detection System: Car Detection in Static Images', *MIT Technique Report*, 1999.
2. Lipton, A. J., Fujiyoshi, H., Patil, R. S.: 'Moving Target Classification and Tracking from Real-time Video', *IEEE Image Understanding Workshop*, pp. 129-136, 1998.
3. M.Kilger "A Shadow Handler in a Video-based Real-time Traffic Monitoring System", *IEEE Workshop on Applications of Computer Vision, Proceedings*, pp. 11-18, 1992.
4. 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.
5. K.I.Kim., K.K.Kim., S.H.Park., K.Jung., H.J.Kim., "VEGA VISION: A Vision System For Recognizing Vehicle License Plates", *ISCE*, vol.2, pp. 176-181, 1999.
6. D.Koller, K.Daniilidis, T.Thorhallson, and H.Nagel, *conf. Computer Vision*, pp. 437-452, Italy, May 1992.
7. Badenas, J ; Pla, F. "Segmentation based on region-t racking in image sequences for traffic monitoring", *Pattern Recognition, Proceedings. Fourteenth International Conference*, vol 2, pp. 999-1001, 1998.
8. N.Habili, A.Moini, N.Burgess, "Automatic Shareholding for change Detection in Digital Video", in *Proc. SPIE Visual Communications and Image Processing*, vol 4067, pp.133-142, 2000.