

웨이블릿 변환을 이용한 움직이는 물체 추적

김 종배*, 이 창우**, 김 항준*
*경북대학교 컴퓨터공학과
**포항 1 대학 전산정보처리과
e-mail : kjbllove@ailab.knu.ac.kr

Detection and Tracking of Moving Objects by Wavelet Transform

Jong-Bae Kim*, Chang-Woo Lee**, Hang-Joon Kim*
*Dept. of Computer Engineering, Kyungpook National University
**Dept. of Computer Information Processing, Pohang College

요 약

본 논문은 도로 상에서 움직이는 물체를 검출하고 웨이블릿 변환을 이용하여 검출된 물체를 추적하는 방법을 제안한다. 제안한 방법은 후보 영역 추출 단계와 물체 판별 단계 그리고 추적 단계로 이루어진다. 첫번째 단계에서는 연속된 두 프레임간의 차영상 분석 방법을 기반으로 움직이는 물체에 의해 발생한 영역과 그 이외의 다른 영역들을 검출한다. 두 번째 단계에서는 검출된 영역에 포함되어 있는 물체를 웨이블릿 변환 계수를 입력하는 신경망을 사용하여 판별한다. 그리고 판별된 물체의 위치 및 크기 정보와 웨이블릿 변환 계수를 이용하여 추적한다. 제안한 방법을 실험한 결과, 물체 추적률은 93%, 프레임당 처리 시간은 0.19ms 이다. 본 논문에서 제안한 방법은 실시간 교통 감시 시스템에 유용하게 적용될 수 있다.

1. 서론

컴퓨터 사용의 증가와 하드웨어 성능의 향상으로 컴퓨터 비전 기술을 실생활에 응용하려는 시도가 여러 분야에서 활발히 진행되고 있다. 특히 비디오 영상에서 물체의 움직임 검출 및 추적은 물체 추적, 침입자 감시 시스템, 교통 감시 시스템등 비디오 기반의 응용프로그램에서 중요한 역할을 한다 [1,2,3]. 영상에서 물체의 움직임을 검출한다는 것은 움직임이 있는 물체와 움직임이 없는 배경을 서로 구별시켜 주는 것이다. 기존의 움직임 검출은 장면 차이(Frame Differencing) [1], 히스토그램의 밝기 값 변화 [2], 색상 변화 [3]등을 이용한 방법들이 있다. 이와 같이 차영상(Difference Image)분석이나 특징 벡터를 이용한 방법은 수행 속도가 빠르다는 장점이 있지만, 고정된 임계값을 적용하기 때문에, 시간에 따라 조명이 변하거나 비디오 영상에 잡음이 있을 때 부적절하다.

그리고 물체를 추적하기 위한 연구에는 3차원 모델에 기반한 방법 [4], 특징값에 기반한 방법 [5]등이 있다. 이러한 방법들은 물체의 움직임 방향이나 속도 뿐만 아니라 다른 정보들까지 추정함으로써 계산 시간이 많이 소모된다.

본 논문에서 움직임이 있는 영역 탐지를 위해 차영상 분석 방법을 기반으로 연속된 두 프레임간의 시간 차이 영상과 현재 이미지와의 배경 영상간의 배경 차이 영상에서 움직임이 있는 후보 영역을 검출한다. 그리고 검출된 후보 영역에 물체를 판별하기 위해 신경망을 사용한다. 신경망의 입력은 검출된 후보 영역의 그레이 레벨의 웨이블릿 변환 계수들을 사용한다. 마지막 단계로 판별된 물체의 위치 및 크기 정보와 웨이블릿 변화 계수를 이용하여 추적한다.

검출된 후보 영역은 도로 영상에서와 같이 조도에 의해 발생한 그림자 영역도 함께 검출된다. 이를 해결하기 위해 검출된 후보 영역에 에지 성분의 투영을

통해 그림자가 포함된 영역을 제거한다. 또한 효율적이고 정확한 물체 판별을 위해 본 논문에서는 웨이블릿 변환을 이용한다.

웨이블릿 변환은 최근 들어 신호 처리나 패턴 인식 분야에서 가장 많이 쓰이고 있는 변환 방법중의 하나이다 [6, 7]. 이 변환 방법은 공간 영역과 주파수 영역에서 그 지역적 특성을 잘 나타낸다는 특성으로 인해 영상에서 특징 정보를 얻고자 할 때 많이 쓰이고 있다. 웨이블릿 변환은 공간과 주파수의 두 영역에서 영상을 표시할 수 있기 때문에 에너지등과 같은 공간적 특성과 저주파수 영역에 에너지가 밀집되어 있는 영상의 주파수적 특성을 보다 효율적으로 나타낼 수 있는 특징이 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) 부밴드는 집중도가 떨어지지만 인간 시각에 민감한 영상의 에지 정보를 보유하고 있다. 그림 1은 제안한 움직임 물체 검출 및 추적 방법을 나타낸 흐름도이다.

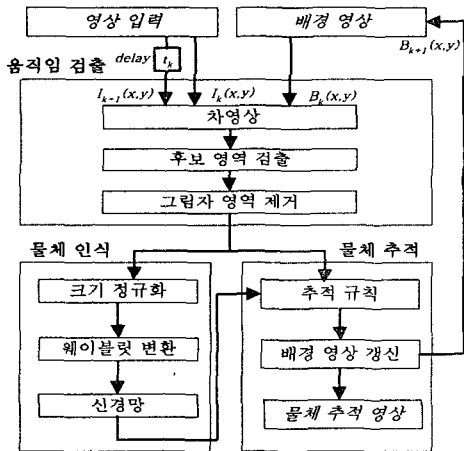


그림 1. 물체 추적 흐름도

2. 움직임 검출

본 논문에서는 움직임이 있는 후보 영역 검출을 위해 연속된 프레임간의 차영상 분석에 의한 방법을 기반으로 하고 있다. 본 장에서는 연속된 두 프레임과 배경 이미지와 현재 프레임의 밝기 차이를 이용하여 움직임이 있는 후보 영역을 검출하고, 그림자 제거 과정을 통해 배경의 움직임이나 밝기의 변화에 의해 발생한 잡영과 그림자를 제거한다.

2.1 차영상 생성

현재 프레임과 이전 프레임간의 차영상(D_i)에서 변화된 픽셀 값이 임계값을 초과하는 부분을 1, 나머지는 0으로 할당된 이진화 영상(T_i)을 생성한다. 현재

프레임과 배경 이미지와의 차영상(D_b)도 같은 방법으로 이진화 영상(T_b)을 생성한다. D_i 영상 내에서는 움직이는 물체뿐만 아니라 정지해 있는 물체 영역도 검출될 수 있다. 그림 2는 후보 영역 검출 과정을 나타낸 그림이다. 그림 2 (a), (b), (c)는 각각 이전 프레임(I_{k-1})과 현재 프레임(I_k), 배경 이미지(B_k)이다. 이전 프레임과 현재 프레임의 밝기 차이 영상 $D_i : D_i = I_{k-1}(x,y) - I_k(x,y)$ 이고, 배경 이미지와 현재 프레임의 밝기 차이 영상 $D_b : D_b = I_k(x,y) - B_k(x,y)$ 이다. (d)와 (e)는 식 (1)과 같이 구해진 시간 차이 영상(T_i), 배경 차이 영상(T_b)이다. 임계값 t 는 적응적 임계값(adaptive thresholding) 알고리즘을 이용하여 구한다. [8].

$$T_i(x,y) = \begin{cases} 1, & \text{if } |D_i(x,y)| > t \\ 0, & \text{otherwise} \end{cases}, T_b(x,y) = \begin{cases} 1, & \text{if } |D_b(x,y)| > t \\ 0, & \text{otherwise} \end{cases} \quad \text{식 (1)}$$

2.2 후보 영역 검출

차영상 분석에 의한 얻어진 이진화된 두 영상에서 1로 할당 되어져 있는 공통된 영역은 물체의 움직임에 의해 발생된 영역이고 (그림 2 (f): $M_i = T_i \cap T_b$), 나머지 영역(그림 2 (g): $U_i = M_i - T_i$)은 정지해 있는 물체이거나 환경 변화에 의해 발생된 영역이다.

움직임이 검출된 영역을 하나의 동일한 영역으로 할당하기 위해 인접한 픽셀의 연결성 성분을 조사하여, 일정한 수 이상 연결되어 있는 픽셀만 남기고 나머지는 잡음 영상으로 처리하여 제거한다. 그림 2 (h)는 차영상 분석에 의한 움직임 있는 후보 영역 (Minimum Bounding Region, MBR) 검출된 영상이다.

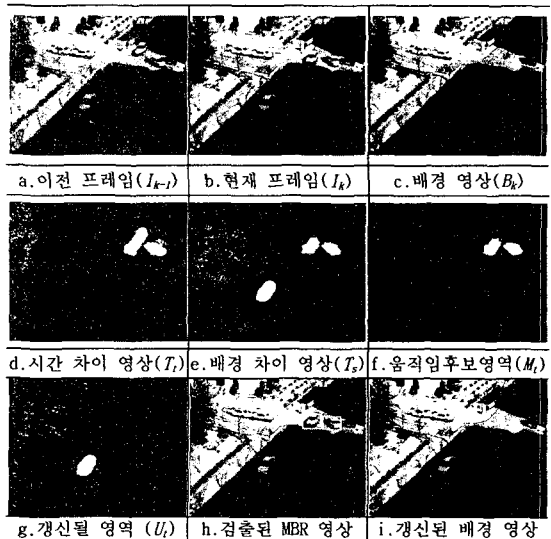


그림 2. 움직임 후보 영역 검출

차영상 분석에 의해 얻어진 사각형 후보 영역(MBR)은 조도에 의해 발생한 그림자 영역이나 프레임 간의 간격에 의해 배경 영역도 같이 얻어 질 수 있다. 따라서, 그림자 영역 제거 과정을 거쳐 정확한 움직임이

있는 후보 영역만을 검출한다 [9]. 검출된 $m \times n$ 픽셀 크기의 후보 영역에 에지맵 $E_k(x, y)$ 을 구하고, 이 에지맵에서 투영 벡터인 v_k, h_k 를 얻는다. 그림 3은 정확한 후보 영역을 검출하는 그림자 제거 과정을 나타낸다. 수평, 수직 투영 벡터를 구하는 식(2), (3)과 같다.

$$v_k = (v_1, \dots, v_m) = \sum_{i=1}^m \{E_k(x_i, y_1), \dots, E_k(x_i, y_n)\} \quad (2)$$

$$h_k = (h_1, \dots, h_n) = \sum_{j=1}^n \{E_k(x_1, y_j), \dots, E_k(x_m, y_j)\} \quad (3)$$

투영 벡터 v_m 와 h_n 에서 최대값을 찾고, 그 값을 2로 나누어 T_v 와 T_w 를 구한다. 구하는 방법은 식(4)와 같다.

$$T_v = \frac{\text{Max}(v_i | 1 < i < m)}{2}, T_w = \frac{\text{Max}(h_j | 1 < j < n)}{2} \quad \text{식(4)}$$

수평으로 투영한 벡터 v_m 에서 위쪽과 아래쪽부터 T_v 값을 초과하는 값, $T_v(\text{top})$ 와 $T_v(\text{bottom})$ 를 찾는다. 또한 수직으로 투영한 벡터 h_n 에서도 마찬가지로 왼쪽과 오른쪽에서 T_h 을 초과하는 값, $T_h(\text{left})$ 와 $T_h(\text{right})$ 를 찾는다. 찾아진 값들은 검출된 후보 영역에서 그림자 부분을 제거하기 위해 사용된다.

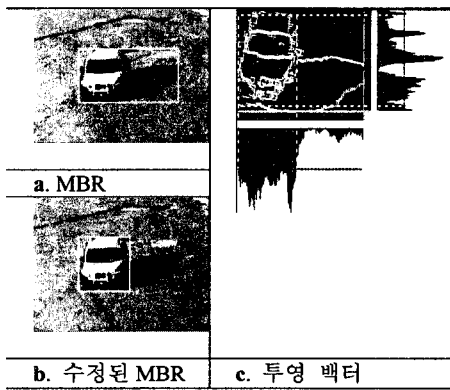


그림 3. MBR의 수정

3. 신경망을 이용한 물체 판별

검출된 자동차 후보 영역에 물체가 포함되어 있는지를 판별하기 위해 신경망을 사용한다. 신경망의 입력으로는 2-level haar 웨이블릿 변환을 통한 저주파, 수평, 수직, 대각 영역의 계수들을 사용한다. 그레이 레벨의 자동차 영상에서 multi-level 웨이블릿 변환을 통한 계수들을 신경망의 입력으로 사용함으로써, 입력 패턴의 차원을 감소시키고, 동시에 신뢰성 있는 물체 판별이 가능하다. 제안한 방법은 일반 도로에서 여러 시간대에 연속적으로 움직이는 자동차 영상으로 실험하였다. 신경망의 학습에 사용된 알고리즘 오류 역전파 학습 알고리즘이며, 학습에 사용된 영상은 64 x 64, 32 x 32 픽셀 크기의 자동차의 정면 그레이 영상이며,

자동차가 포함된 영상 230 개, 그렇지 못한 영상 400 개를 사용하였다. 영상에서 자동차의 크기는 영상 획득 장치와 자동차 사이의 거리에 의해 다르게 나타난다. 따라서 검출된 후보 영역들을 신경망에 입력하기 위해 정보 손실을 최소화하며, 정규화를 적용해야 한다. 그러므로 후보 영역을 크기 비율에 따라 32x32, 64x64 픽셀 크기로 정규화한다.

32x32 픽셀 크기로 정규화한 영상은 1-level 웨이블릿 변환하고, 64x64 픽셀 크기로 정규화한 영상은 2-level 웨이블릿 변환을 통해 각 서브 밴드의 계수들을 신경망에 입력하여 검출된 후보 영역에 자동차가 포함된 영상인지를 판별한다. 그림 4는 자동차 판별을 위해 웨이블릿 변환 계수들을 입력하는 신경망의 구조이다. 본 논문에서 사용된 신경망의 입력노드 1024, 히든노드 350, 출력 노드 2개로 구성된다.

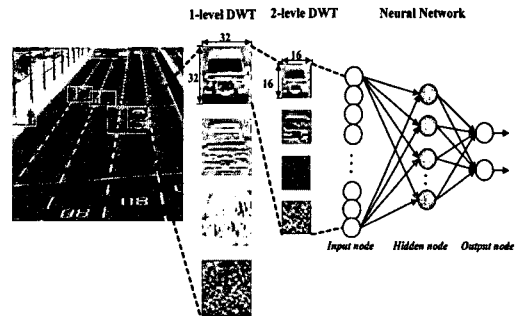


그림 4. Multi-level 웨이블릿 변환

4. 물체추적

물체의 추적은 웨이블릿 변환에서 얻어진 특징값들과 판별된 물체의 중심 좌표, 크기 정보를 이용한다. 웨이블릿 변환으로 얻어진 특징값으로는 아래 식(5), (6), (7)과 같이 판별된 물체의 평균(M)과 이차(U_2), 삼차(U_3) 모멘트이다. $S(i,j)$ 는 웨이블릿 변환된 영상 (decomposition image)이다.

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

$$\mu_2(I) = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (S(i, j) - M(I))^2 \quad (6)$$

$$\mu_3(I) = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (S(i, j) - M(I))^3 \quad (7)$$

판별된 물체의 추적 규칙은 아래와 같다.

- 단계 1 : 현 프레임에 판별된 물체와 이전 프레임에서 인식된 물체 사이의 중심좌표의 유사도를 비교한다.
- 단계 2 : 현 프레임에서 판별된 물체의 크기 정보와 이전 프레임에서 판별된 물체의 크기 유사도를 비교한다.
- 단계 3 : 현 프레임에서 특정 임계치를 만족하는 물체가 둘 이상이면, 웨이블릿 특징 값으로 얻어진 모

멘트들을 비교하여 결정한다.

추적 단계가 끝나면 배경 영상을 갱신한다. 갱신 방법은 그림 2 (g) 영상에서처럼 움직임이 있는 영역을 제외한 나머지 영역에 해당하는 부분을 이용하여 새로운 배경 영상(B_{k+1})을 생성한다.

5. 실험 결과

제안한 방법을 실험하기 위해 도로상에서 획득한 128 대 자동차가 포함된 동영상으로 실험하였다. 영상 입력을 위해 Sony 사의 DCR-PC3 디지털 카메라를 사용하였고, 팬티엄II IBM 호환 PC 에서 윈도우 환경의 VC++ 언어로 구현하였다. 입력 영상은 초당 8 프레임이고, 320x240 픽셀 크기의 칼라 영상이다. 실험을 통해 영상 내에서 움직이는 물체가 사람이나 혹은 이를 자동차가 포함되어 있을 때 움직임 후보 영역 검출 단계에서 이 영역들도 같이 검출되었으나, 신경망 입력 통한 후보 영역 판별 단계에서는 자동차가 포함된 영역만을 검출하였다. 표 1 은 가우시안 노이즈를 입력 영상에 추가한 후 실험한 결과 명암(Intensity) 피쳐(feature)보다 웨이블릿 피쳐가 노이즈에 강한 자동차 판별률을 나타내었다. 표 2 는 제안한 방법으로 움직이는 자동차를 추적한 추적률을 나타낸다. 그림 5 는 제안된 방법으로 실제 도로 영상에서 자동차를 추적한 실험 결과이다.

표 1. 웨이블릿 피쳐와 명암 피쳐를 이용한 자동차 판별률

판별률(%)	노이즈(%)				
	0	5	10	15	20
웨이블릿 피쳐	98	96	94	91	84
명암 피쳐	99	96	93	87	79

표 2. 물체 추적률

총 자동차의 수	추적된 자동차의 수	추적률 (%)
128	119	93

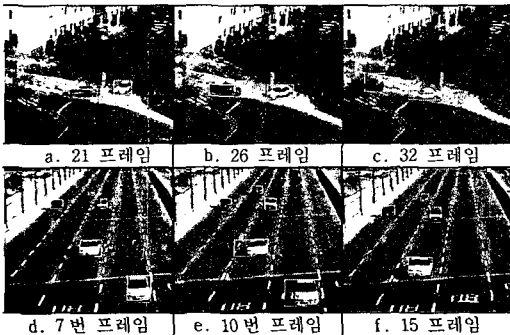


그림 5. 자동차 영역 추적결과

6. 결론

본 논문은 웨이블릿을 이용한 판별 및 추적 방법을 제안하였다. 연속적으로 입력되는 영상에서 움직임이 있는 영역 검출을 위해 차영상 분석 방법을 사용하고, 검출된 후보 영상에 자동차가 포함되어 있는지 판별하기 위해 웨이블릿 변환 계수들을 입력으로 하는 신경망을 이용하였다. 또한, 판별된 자동차는 중심 좌표와 크기 정보, 그리고 웨이블릿 변환에서 얻어진 모멘트를 이용하여 추적하였다. 일반 도로상에서 여러 시간대에 연속적으로 움직이는 자동차 영상으로 실험해 본 결과, 자동차 추적률은 93%이고, 프레임당 처리 시간 0.19ms 미만의 결과를 얻었다. 또한 판별된 자동차의 추적에 사용된 정보가 실제 도로 상황에서 자동차를 추적하기 위해 충분하다는 것을 보였다.

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 교통 감시 시스템 구현에 관해 연구하고자 한다.

참고문헌

- [1] 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.
- [2] 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.
- [3] 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.
- [4] D.Koller, K.Daniilidis, T. Thorhallson, and H. Nagel, "Model-based object tracking in traffic scenes," In European conf. Computer Vision, pp. 437-452, Italy, May 1992.
- [5] Badenas, J.; Pla, F. "Segmentation based on region-tracking in image sequences for traffic monitoring", Pattern Recognition, Proceedings. Fourteenth International Conference, vol 2, pp. 999-1001, 1998
- [6] T.S.Yun, "Support Vector Machine-based Solder Joint Inspection", Ph.D. thesis, KyungPook National University, 2000
- [7] Y.J.Jeon, J.I.Kim, "A Study on the Feature Extraction using the Wavelet Transform in Satellite Remote Sensing Image", Proceeding of The 1th KISPS Summer Conference, vol. 1, no. 1 pp. 237-240, 2000.
- [8] N. Habibi, A. Moini, N. Burgess, " Automatic thresholding for change detection in Digital video", in Proc. SPIE Visual Communications and Image Processing, Vol. 4067, pp. 133-142, 2000.
- [9] M. Betke, E. Haritaoglu, L. S. Davis, "Real-time multiple vehicle detection and tracking from a moving vehicle", Machine Vision and Applications, 12 , pp. 69-83, 2000