

강건 예측과 군집화를 결합한 물체의 움직임 감지

장석우[○], 허문행^{*}, 이상훈^{**}

[○]*안양대학교 디지털미디어학과

^{**}한국건설기술연구원 정보진산실

e-mail: {swjang,moonh}@anyang.ac.kr, jason813@kict.re.kr

Object Movement Detection Integrating Robust Estimation and Clustering

Seok-Woo Jang[○], Moon-Haeng Huh^{*}, Sang-Hoon Lee^{**}

[○]*Dept. of Digital Media, Anyang University

^{**}Information Service Division, Korea Institute of Construction Technology

● 요약 ●

본 논문에서는 비디오 데이터로부터 물체의 초기 움직임 영역을 자동으로 검출하는 방법을 소개한다. 제안하는 시스템은 먼저 입력 영상을 받아들인 후 인접된 영상으로부터 일정 크기의 정방향의 블록 단위로 움직임을 나타내는 모션 벡터를 추출한다. 그리고 추출된 모션벡터를 아웃라이어 제거하는 강건 예측 알고리즘에 적용하여 배경에 해당하는 모션벡터와 잡음 및 움직이는 물체에 해당하는 모션벡터를 구분한다. 그런 다음, 군집화 알고리즘을 적용하여 이동하는 물체를 나타내는 모션벡터를 군집화하고, 군집화된 모션벡터에 해당하는 영역의 크기가 일정 수치 값 이상일 때 움직이는 물체가 감지되었다고 판단한다. 본 논문의 실험에서는 제안된 물체의 움직임 감지 방법이 기존의 방법에 비해 성능이 보다 우수함을 보인다.

키워드: 물체 움직임(object movement), 강건 예측(Robust Estimation), 블록 정합(block matching)

I. 서론

최근 들어 움직이는 물체를 효과적으로 감지하기 위한 다양한 연구 방법들이 진행되고 있다. 이 방법들에는 광류(optical flow) 측정을 이용한 방법[1], 차영상을 이용한 방법[2], 모델 생성을 통한 방법[3], 스네이크(snake)와 같은 경계선 기반의 방법[4], 특징 맵(saliency map)을 이용한 방법[5] 등 그 종류가 매우 다양하다.

광류를 이용한 방법은 물체 움직임의 방향을 정확히 측정할 수 있는 반면 조명의 변화에 민감하다. 차영상을 이용한 방법은 고정된 카메라 설정 환경에서 인접된 두 영상에 차연산을 적용하여 물체를 감지하는데, 알고리즘이 간단하면서도 비교적 정확하나 카메라가 고정되어 있어야 한다는 제약사항이 있다. 모델 생성을 이용한 방법은 이동 물체나 배경에 대한 분포 모델을 생성하여 물체를 감지하는 기법으로 물체를 강건하게 추적할 수 있다는 장점이 있지만, 모델을 생성하기 위한 초기 물체의 감지가 어렵다는 단점이 있다. 경계선을 이용한 방법은 에너지 최소화 방법을 사용해 물체의 경계 영역을 정확히 찾을 수 있으나 비교적 빠르게 움직이는 물체를 감지하기 어렵다는 단점을 가지고 있다. 특징 맵을 이용한 방법은 사공간 도메인에서 화소의 밀도 분포를 이용하여 특징 맵을 형성하고, 이를 이용해 이동하는 물체를 검출하는 방법으로, 비교적 강건하게 이동물체를 감지하나 효과적인 맵 통합 방법이 필요하다.

이와 같이 이동하는 물체를 감지하는 많은 기법들이 연구되고 있으나 여전히 해결해야 할 문제들이 남아 있다. 특히, 카메라가 고정되어 있지 않은 상황에서는 차영상을 사용하기 어려우므로 동작 물체를 초기에 정확히 감지하는 작업은 어려운 작업이다. 따라서 본 논문에서는 카메라가 고정되어 있지 않은 상황에서 모션벡터를 기반으로 아웃라이어(outlier) 제거 알고리즘과 군집화(clustering)을 이용해 물체의 초기 움직임을 강건하게 감지하는 방법을 제안한다.

II. 아웃라이어 제거

제안된 방법은 먼저 보다 정확하게 모션벡터를 추출하기 위해 확장적 블록 정합을 이용한다[6]. 그런다음, 추출한 모션 벡터를 아웃라이어 제거 알고리즘인 적용적 강건 예측(robust estimation)에 적용하여 카메라의 움직임 정보를 구한다[7]. 제안된 방법에서는 카메라의 동작 인수를 구하기 위해 강건 예측을 사용하고, 강건 예측의 입력으로 모션벡터를 사용한다. 따라서 제안된 방법에서 의미하는 정상 자료는 순수하게 카메라의 동작만을 포함하고 있는 배경 영역의 모션벡터를 의미하고, 오류자료는 물체의 이동이나 잡음으로 형성된 모션벡터를 의미한다.

그러나 기존의 강건 예측은 정상 자료와 오류 자료를 구분하기

어려운 최소화 기법의 초기 단계에서도 임계치를 이용하여 정상 자료와 오류 자료를 분리한다. 따라서 제안된 방법에서는 이런 문제점을 개선한 적응적인 강건 예측을 사용한다. 제안된 방법에서 사용하는 측정 모델은 식 (1)과 같다. 일반적으로, 식 (1)은 어파인 (affine) 모델이라고 불린다.

$$y = y(x, y, \mathbf{a})$$

$$y(x, y, \mathbf{a}) = \begin{bmatrix} u(x, y) \\ v(x, y) \end{bmatrix} = \begin{bmatrix} a_1x + a_2y + a_3 \\ a_4x + a_5y + a_6 \end{bmatrix} \quad (1)$$

식 (1)에서 인수 벡터 \mathbf{a} 는 카메라의 동작 인수로서 기존의 카메라 변환 관계 수식을 이용하여 쉽게 상하각, 좌우각, 경사각, 초점 거리와 같은 조망인자로 변환할 수 있다. 그리고 $u(x, y)$ 와 $v(x, y)$ 는 적응적 강건 예측의 입력으로 사용되는 모션벡터의 수평 성분과 수직 성분을 의미한다. 식 (1)의 측정 모델을 기초로 목적 함수를 정의하면 식 (2)와 같다. 식 (2)에서 w_i 는 i 번째 입력 자료의 가중치를 의미하고, x_i' 와 y_i' 는 i 번째 입력 자료의 수평 성분과 수직 성분을 의미하며, $\hat{x}_i(x_i, y_i, \mathbf{a})$ 와 $\hat{y}_i(x_i, y_i, \mathbf{a})$ 는 측정 모델의 수평 성분과 수직 성분을 의미한다.

$$\chi^2(\mathbf{a}) = \sum_{i=1}^N w_i \left[\frac{y_i' - y(x_i, y_i, \mathbf{a})}{\sigma_i} \right]^2$$

$$= \sum_{i=1}^N w_i \left[\frac{(x_i' - \hat{x}_i(x_i, y_i, \mathbf{a}))^2}{\sigma_{x_i}^2} + \frac{(y_i' - \hat{y}_i(x_i, y_i, \mathbf{a}))^2}{\sigma_{y_i}^2} \right] \quad (2)$$

적응적 강건 예측은 이진 가중치 함수를 사용하는 기존의 강건 예측과는 다르게 연속적인 가중치 함수를 사용한다. 연속적인 가중치 함수로는 시그모이드 함수를 이용한다. 시그모이드 함수는 이진 가중치 함수보다 정상자료와 오류자료의 소속정도를 보다 효과적으로 표현하므로 최소화 기법의 초기 단계에서 범할 수 있는 에러를 최소화할 수 있다. 제안된 방법에서 사용하는 가중치 함수는 식 (3)과 같이 정의한다.

$$\mathbf{W}^k = (w_1^k, w_2^k, \dots, w_j^k, \dots, w_N^k)$$

$$w_j^k = \alpha \cdot w_j^{k-1} + \beta \cdot (1 - \text{sig}^k(x = j; a^k, c^k))$$

$$\text{sig}^k(x; a^k, c^k) = \frac{1}{1 + e^{-a^k(x - c^k)}}$$

where $0 \leq \alpha, \beta \leq 1, \alpha + \beta = 1$

식 (3)에서 \mathbf{W}^k 는 최소화 단계의 k 번째 반복에서 사용하는 가중치 벡터로서 w_j^k 는 0에서 1사이의 범위를 가지는 j 번째 입력 자료의 가중치를 나타낸다. 그리고 x 는 시그모이드 함수의 입력벡터를 나타내고, c^k 는 시그모이드 함수를 수평축을 기준으로 이동시키는 바이어스(bias)를 나타내며, a^k 는 $x=c$ 위치에서 시그모이드 함수의 기울기를 조절하는 역할을 수행한다. 제안된 방법에서는 에러를 오름차순으로 정렬한 누적에러를 사용하므로 일반적인 시그모이드 함수를 사용하지 않고 $x=c$ 위치에서 y 축을 중심으로 회전 이동한 시그모이드 함수를 사용한다.

적응적 강건 예측은 최소화 기법의 반복 과정에서 현재 단계의 가중치 계산 시 이전 단계의 가중치를 반영하도록 정의된다. 또한, 최소화 기법이 반복되면서 에러가 감소함에 따라 시그모이드 함수를 하드리미트(hard limit) 형태로 조율함으로써 정상 자료와 오류 자료를 보다 유연하게 분리한다.

III. 군집화에 의한 물체 추출

모션벡터를 아웃라이어 제거 알고리즘을 통해 입력받아 처리하면 배경 영역의 모션벡터로부터는 카메라의 움직임 정보 파라미터가 추출되며, 이동 물체 영역에 해당하는 모션벡터는 강건 예측의 반복 처리 과정이 진행됨에 따라 아웃라이어로 판단되어 이동물체 모션벡터 부분만 분할되게 된다. 따라서 제안된 방법에서는 이동 물체 영역에 해당된다고 판단된 모션벡터들을 군집화(clustering)하여 잡음을 제거하고 물체 단위로 분할하여 최종적인 이동 물체의 초기 움직임 정보로 활용한다.

일반적으로 샘플을 군집화하는 클러스터링 알고리즘은 여러 종류가 있지만 제안된 방법에서는 사전에 군집화하고자 하는 군집의 개수를 미리 설정하지 않아도 되는 비감독(unsupervised) 클러스터링 알고리즘 중의 하나인 ISODATA 군집화 방법을 사용한다 [8]. 보통 ISODATA 군집화 알고리즘은 분할과 합병을 기반으로 입력 패턴들을 군집화하는 알고리즘으로 개략적으로 다음과 같은 4가지 단계의 과정을 반복적으로 수행한다. 첫째, 초기 파라미터를 결정한다. 둘째, 입력 패턴들을 군집에 할당한다. 셋째, 군집 관련 파라미터를 계산하고 분할과 합병 조건을 확인한다. 넷째, 분할 또는 합병을 수행한다.

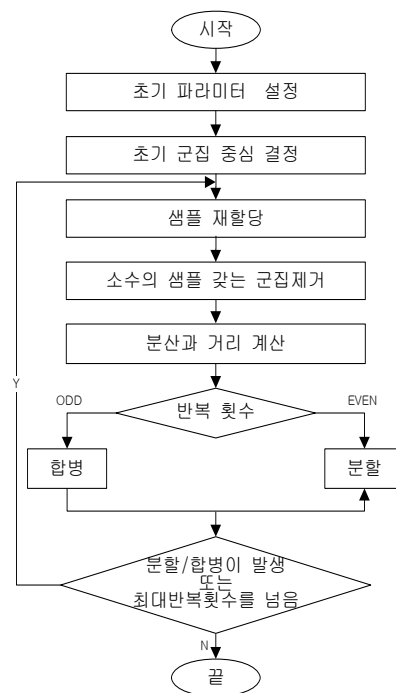


그림 1. ISODATA 군집화
Fig. 1. ISODATA Clustering

위의 과정들은 사용된 파라미터들이 안정된 값을 유지할 때까지 반복된다. 그리고 그림 1은 ISODATA 군집화 알고리즘의 순서도를 나타낸다. ISODATA 알고리즘을 적용했을 때 최종적으로 추출된 군집의 개수는 이동 물체들의 개수를 나타내며, 각 군집이 갖는 평균값은 각 이동 물체의 중심 위치 정보가 된다.

IV. 실험 결과

입력 영상으로는 카메라가 고정되지 않은 환경에서 촬영한 다양한 종류의 동영상을 활용하였다. 그림 2는 인접된 영상으로부터 블록 정합 알고리즘을 이용하여 움직임 정보를 나타내는 모션벡터를 추출한 예이다. 그림 2의 모션벡터 추출 결과에서 확인할 수 있듯이 다양한 환경에서 촬영한 영상에서 부분적으로 약간의 오류가 발생하고 있으나 비교적 정확하게 모션 벡터를 구했음을 시각적으로 확인할 수 있다.

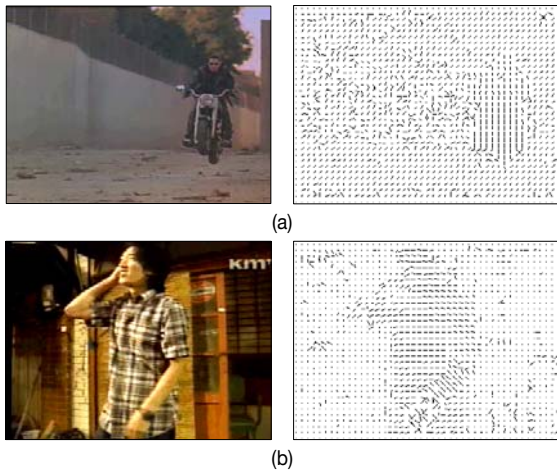


그림 2. 모션벡터 추출
Fig. 2. Motion Vector Extraction

그림 3은 모션벡터를 아웃라이어 제거 알고리즘의 입력으로 받아들이어 아웃라이어라고 판단한 잡음과 이동 물체 영역에 해당하는 모션벡터 부분만을 제거한 결과이다. 그림 3의 좌측의 결과는 이동 물체 영역과 잡음에 해당하는 모션벡터가 제거되었으나, 입력영상의 화질이 전체적으로 낮아 잡음 부분이 상당수 제거되지 않고 남아 있다. 그림 3의 우측의 결과는 영상 화질이 낮아서 잡음이 많이 발생 하였으나 비교적 정확하게 잡음에 해당하는 모션벡터를 제거하였다.

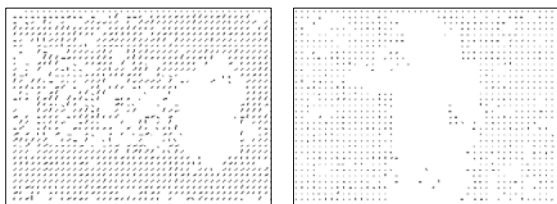


그림 3 잡음 제거
Fig. 3 Outlier 제거

그림 4는 그림 3에서 아웃라이어라고 판단된 모션벡터들을 ISODATA 군집화하여 영역이 매우 작은 것들과 영역의 밀집도(compactness)가 너무 낮은 것들은 제거하고, 나머지를 이동 물체 영역이라고 판단한 결과이다. 그림 4의 결과는 이동 물체 영역 내에 아웃라이어라고 판단하여 제거된 몇몇의 블록 영역을 제외하고는 비교적 정확하게 이동 물체의 움직임을 감지할 수 있음을 확인하였다.



그림 4. 물체 군집화
Fig. 4. Object Clustering

그리고 제안된 방법의 성능을 비교 평가하기 위해서 기존의 방법 중 차영상을 이용한 방법과 일반적인 아웃라이어 제거 방법, 그리고 제안된 아웃라이어 제거와 클러스터링을 이용한 방법을 비교 실험하였다. 물체의 움직임이 포함된 다양한 비디오 데이터에 각각의 알고리즘들을 적용하였으며, 성능 평가는 각 알고리즘의 초기 물체 감지 기능의 성공과 실패를 백분율로 측정하였다.

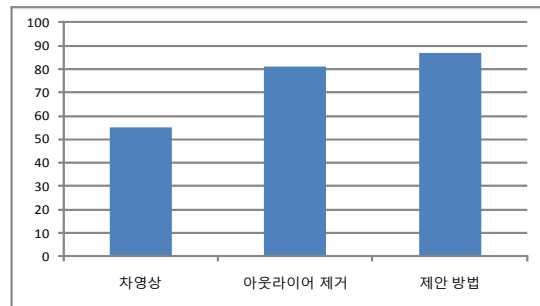


그림 5. 성능 비교
Fig. 5. Performance Comparison

그림 5는 비교 결과를 그래프로 보여준다. 차영상을 이용한 방법은 카메라가 고정되어 있는 환경에서 정확성 있게 동작하므로 카메라의 움직임이 거의 없는 환경을 제외하고는 이동 물체를 감지하지 못했다. 기존의 아웃라이어 제거 알고리즘은 어느 정도 정확성 있게 이동물체를 검출하였으나, 군집화 알고리즘을 결합한 제안된 방법보다는 성능이 좋지 못했다. 제안된 방법은 어떤 영상을 사용하느냐에 따라 성능에 다소 차이가 있을 수 있으나, 다양한 비디오 영상에 대해 실험한 결과 비교적 안정적으로 물체의 초기 이동을 감지하는 것을 확인하였다.

IV. 결 론

본 논문에서 제안한 물체의 움직임 감지 방법은 대부분의 이동 물체 검출 기법이 가지고 있는 가장 어려운 문제 중의 하나인 카메라가 고정되지 않은 환경에서 촬영된 다양한 영상에서도 물체의 초기 움직임을 효과적으로 감지할 수 있음을 보여주었다는데 그 의미가 있다. 또한, 본 논문에서 검출된 물체의 초기 이동 정보는 이후의 처리 단계, 예를 들어, 물체의 모델 생성 및 추적, 카메라의 동작을 제어하는 서보 장치의 초기화 설정 등에 매우 유용하게 사용될 수 있을 것으로 기대된다.

향후 연구로는 제안된 방법으로 검출된 초기 이동 물체를 모델로 생성하여 물체를 강건하게 추적하는 방법에 대한 연구를 수행할 예정이다.

참고문헌

- [1] S.P.N. Singh, P.J. Csonka, K.J. Waldron, "Optical Flow Aided Motion Estimation for Legged Locomotion," International Conference on Intelligent Robots and Systems, pp.1738~1743, 2006.
- [2] Dae-Sik Jang, Seok-Woo Jang, Hyung-II Choi, "2D Human Body Tracking with Structural Kalman Filter," Pattern Recognition, Vol. 35, No. 10, pp. 2041~2049, October 2002.
- [3] T. Parag, A. Elgammal, A. Mittal, "A Framework for Feature Selection for Background Subtraction," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1916~1923, 2006.
- [4] M. Yokoyama, T. Poggio, "A Contour-based Moving Object Detection and Tracking," International Workshop on Visual Surveillance and Performance Evaluation, pp. 271~276, 2005.
- [5] Shan Li, M.-C. Lee, "An Efficient Spatio-temporal Attention Model and Its Application to Shot Matching," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 17, No. 10, pp. 1383~1387, 2007.
- [6] Seok-Woo Jang, Hyung-II Choi, "A Strategy of Matching Blocks at Multi-Levels," International Journal of Intelligent Systems, Vol. 17, No. 10, pp. 965-975, 2002.
- [7] Seok-Woo Jang, Gye-Young Kim, Hyung-II Choi, "Estimating Motion Parameters Using a Flexible Weight Function," IEICE Transactions on Information and Systems, Vol.E89-D, No.10, pp. 2661~2669, 2006.
- [8] H. Jin, Y. Feng, "Towards an Automatic Road Lane Marks Extraction based on ISODATA Segmentation and Shadow Detection from Large-scale Aerial Images, 24th FIG International Congress, 2010.