

# 움직이는 물체검출을 위한 영상 좌표계 변환에 관한 연구

김용진<sup>0</sup> 이일병

연세대학교 컴퓨터과학과

{yjkim98<sup>0</sup>, yblee}@csai.yonsei.ac.kr

## The study of the Image Geometric Transforms for Moving Object Detection

Yongjin Kim<sup>0</sup>, Yillbyung Lee

Dept. of Computer Science, Yonsei University

### 요 약

배경이 움직이는 카메라에서 영상을 획득하여 특징점을 추출하고 특징점을 이용해 영상 좌표계 변환 파라미터를 추정한다. 추정된 파라미터를 이용하여 영상내의 움직이는 물체를 검출하기 위해 카메라의 Ego-motion을 보정하는 영상 좌표계 변환 방법을 소개하고, Ego-motion 보정을 통해 연속된 두 영상에서 움직이는 물체를 검출하는 실험을 수행한 내용의 논문이다.

### 1. 서 론

물체추적은 컴퓨터 비전에서 중요한 부분을 차지한다. 물체추적을 위한 연구도 다양하게 진행되고 있으며, 많은 연구 성과가 있다. 물체추적은 주요 산업, 의료, 군사 분야에 널리 활용되고 있고, 특히 최근에는 Robotics에서 중요한 연구 분야이다. 본 논문에서는 카메라와 같은 영상입력장치를 통해 입력되는 영상의 흐름(sequence)에서 이동 물체를 검출하고, 추적하기 위해 카메라에서 입력된 영상을 보정하기 위한 방법을 제안하고자 한다. 카메라에서 입력된 영상이 고정된 배경이고 움직이는 물체를 검출하고자 할 때에는 직관적이고 가장 단순한 방법으로 시간  $t$ 와 시간  $t+1$ 에서 추출된 영상 프레임의 차를 이용하면 움직이는 물체가 쉽게 검출이 가능하다. 영상이 획득된 환경, 특히 영상에서 빛의 밀도(density)에 의한 영향을 많이 받을 수도 있으나, 이러한 고정된 배경 영상에서 움직이는 물체를 추적하는 기법으로 optical flow method나 block match method 등의 방법[1]을 기반으로 하는 연구들이 있다. 배경이 고정되지 않은 환경, 즉 카메라의 상하, 좌우 이동이 있을 경우 단순하게 영상에서 시차가 있는 두 프레임의 차를 이용하여 이동 물체를 검출하는 것은 거의 불가능하다. 배경이 고정되지 않고 움직이는 환경에서 이동물체를 검출하고 추적하기 위해 카메라의 움직임 파라미터를 이용하여 움직이는 에지(edge)를 추출하여 움직이는 물체를 추적하는 연구[2]와 같은 선행 연구가 있었으나, 이 경우 카메라의 파라미터를 선형하여 알고 있는 카메라의 모션을 통해 영상을 변형(transformation)하여 카메라의 움직임을 보상하는 방법을 사용하였으나, 이와 같은 방법은 카메라의 정확한 제원을 알 수 없는 경우 사용하기 좋은 방법이 아니다.

그리고 두 영상 프레임 간의 라인(line)을 특징으로 하고 특징 매칭을 통해 모션 또는 회전을 예측하는 관련 연구도 있다. 라인 특징을 사용할 경우에는 직선으로 분명히 구분되는 물체의 경우 모션 추정이 가능하지만, 사람과 같이 굴곡이 있는 물체는 추적하기 어려운 점이 있다. 그리고 색상과 움직임을 이용하여 배경과 물체를 분리하여 추적하는 연구[3]등이 있다. 이러한 연구의 대부분은 사람을 추적하기 위해 사용되고 있고, 움직이는 물체가 어떠한 색상 정보를 갖고 있는지 알 수 없으므로 불특정의 움직이는 물체에 대해 적용하기 위해 별도의 처리 과정이 필요하다. 그 외에도 비교적 적은 움직임을 갖는 배경 영상의 학습을 통해 배경으로 인지 시키고 영상에서 움직임이 큰 물체가 나타났을 때, 학습된 배경 이외의 물체를 추출하여 추적하는 연구도 있다[4]. 본 논문에서는 직관적이고 단순한 방법인 시차가 있는 두 영상 프레임의 차를 이용하도록 한다. 두 영상 프레임의 차를 직접적으로 연산하면 앞서 설명한 것과 같이 배경과 물체를 분리할 수 없다. 그래서 카메라에서 획득된 영상의 보정, 즉 ego-motion compensation 과정을 통해 배경 영상의 움직임을 보정하고 보정된 영상과 시차가 있는 프레임의 차를 계산하여 움직이는 물체를 검출할 수 있다.

### 2. 본 문

위에서 언급한 것과 같이 영상의 보정을 위해서 영상에서 특징점을 추출하고 특징점 매칭을 이용하여 영상의 보정을 위한 변형(transformation) 과정을 수행하기 위한 파라미터를 산출한다. 영상의 변형은 선형적인 변형과 비선형적 변형이 있다. 영상 획득 당시에 가까이 있는 배경 물체와 멀리 있는 배경 물체는 서로 비선형적으로 변형되므로 영상에 적절한 변형 알고리즘을 적용하여야 한다. 본문에서는 Lucas-Kanade method[5]를 이용하

“본 연구는 과기부 뇌신경정보학사업으로부터 부분적인 지원을 받아 수행되었음.”

여 물체의 특징점을 추출하고, 추출된 특징점을 이용하여 ego-motion을 추정하기 위해 bilinear 변형 모델(transformation model)과 최적화를 위한 비용함수를 이용하여 optimization model의 파라미터를 산출하여 카메라 ego-motion 보상을 수행한다.

2.1 특징점 추출

특징점 추출을 위해 사용된 Lucas-Kanade method는 KLT corner detector라고 알려져 있고, Kanade-Lucas-Tomasi(KLT) 추적 알고리즘[6,7]에서 사용되었다. 영상에서 KLT corner detector에 의해 선택된 특징점은 엄밀히 말해 corner가 아니라 영상에서 작은 탐색 윈도우(small search window)를 이용하여 윈도우 내에서 두 직교 방향(x, y)에 대해 그래디언트(gradient)가 높은 값을 나타내는 점(point)로 나타낼 수 있다[8].

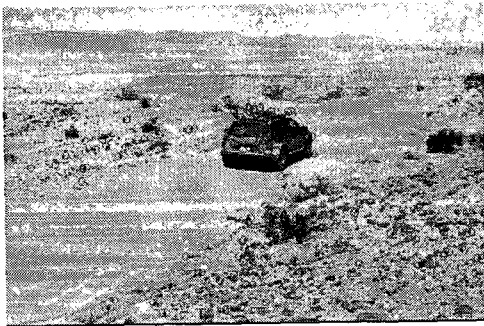


그림 1 Feature Point Extraction

움직이는 물체를 추적하기 위해 배경이 움직이는 영상에서 시차가 있는 두 영상 프레임을 직접 비교하지 않은 것은 두 영상 프레임의 view point가 다르기 때문이다. 스테레오 비전 시스템에서 획득된 영상과 같이 영상 시퀀스에서 인접한 시간의 두 영상 프레임은 다른 view point를 갖지만 매우 비슷하여 두 영상 프레임 간에는 큰 변화가 없다고 가정할 수 있다. 영상에서 특징점은 3개의 변수를 가진 함수로 표현할 수 있다. 즉, 위치와 시간에 함수로 표현되고 위에 언급한 두 영상 프레임에서 특징점은 아래 (식 1)과 같이 나타낼 수 있다.

$$I(x, y, t) = I(x - \xi, y - \eta, t + \tau) \quad ; \text{(식 1)}$$

특징점을 추출하기 위해 size가 N x N인 탐색 윈도우에서의 coefficient matrix를 정의하여야 한다. 각각의 특징점  $x = (x, y)$ 이고, covariance matrix는 아래 (식 2)와 같이 나타낼 수 있다.

$$C = \sum_{\substack{x \\ y}} \begin{bmatrix} g_x^2 & g_x g_y \\ g_x g_y & g_y^2 \end{bmatrix} \quad ; \text{(식 2)}$$

$$\text{where } \nabla I(x) = \begin{bmatrix} \text{round} & \text{round} \\ \text{round} & \text{round} \end{bmatrix} = [g_x, g_y]$$

(W : size N x N with x as center)

(식 2)에서 covariance matrix를 계산하여, covariance matrix의 고유값  $(\lambda_1, \lambda_2)$ 을 구하고, 탐색 윈도우 내에서  $\min(\lambda_1, \lambda_2) > \theta$ 인 특징점을 추출한다. 여기서  $\theta$ 는 미리 정의된 threshold 이다.

2.2 영상좌표계 변환

카메라를 통해 입력된 비디오 sequence의 연속된 프레임은 카메라의 움직임과 물체의 움직임으로 구성될 수 있으며, 이러한 카메라의 움직임을 ego-motion이라 하고, 카메라 ego-motion은 동영상의 이미지 프레임에서 특징점을 추출 및 추적하여 예측할 수 있다[8,9]. 고정된 카메라가 아닌 비고정 카메라의 경우, 연속된 두 이미지 프레임 간에는 서로 다른 좌표계를 갖는다. ego-motion compensation는 서로 다른 좌표계를 갖는 두 영상을 직접적으로 비교할 수 있도록 프레임 영상의 변형 과정을 통해 서로 다른 좌표계를 하나의 좌표계로 일치시키는 것이다. 즉, 시간 t-1에서의 이미지 프레임  $I^{t-1}$ 과 시간 t에서의 이미지 프레임  $I^t$ 가 있을 때,  $I^{t-1}$ 의 좌표계를 변형과정을 통해  $I^t$ 와 같게 하여 두 영상의 차를 직접 비교함으로써 배경이 고정된 카메라에서 연속된 두 영상 프레임에서 영상차를 통해 움직이는 물체를 찾아내는 것과 같은 효과를 나타낼 수 있다.

본 논문에서는 이미지 좌표계의 변형을 위한 방법으로 projective/bilinear model을 적용하여 좌표계 변환을 수행한다. projective/bilinear model은 대부분의 카메라 움직임, 즉 변환(translation), 줌(zooming), 수직 및 수평 이동(tilting, panning), 회전(rotation) 등의 선형적인 변형과 비선형적인 변형을 나타낼 수 있다[10]. projective model은 8개의 파라미터( $m_i, i=1,2,\dots,8$ )를 이용하여, 일반적으로 아래와 같은 식으로 표현된다.

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} m_1 & m_2 & m_3 \\ m_4 & m_5 & m_6 \\ m_7 & m_8 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad ; \text{(식 3)}$$

where  $(x, y)$  : the original coordinates  
 $(x', y')$  : the transformed coordinates

위 (식 3)에서 transformed coordinates  $(x', y')$ 는 다음과 같이 식에 의해 구해진다.

$$x' = \frac{m_1x + m_2y + m_3}{m_7x + m_8y + 1} = \frac{u}{w}$$

$$y' = \frac{m_4x + m_5y + m_6}{m_7x + m_8y + 1} = \frac{v}{w}$$

연속된 영상 프레임  $I^{t-1}$ 와  $I^t$ 에서 서로 연관된 특징점의 집합을  $f^{t-1} = [f_x^{t-1}, f_y^{t-1}]^T$ 와  $f^t = [f_x^t, f_y^t]^T$ 라고 하고, bilinear model로 projective model을 근사화하면 다음의 (식 4)와 같이 나타낼 수 있다.

$$f'_x = q_1 f_x^{t-1} f_y^{t-1} + q_2 f_x^{t-1} + q_3 f_y^{t-1} + q_4$$

$$f'_y = q_5 f_x^{t-1} f_y^{t-1} + q_6 f_x^{t-1} + q_7 f_y^{t-1} + q_8 \quad ; \quad (\text{식 } 4)$$

위 (식 4)에서 8개의 파라미터( $q_i, i=1,2,\dots,8$ )은 카메라의 ego-motion을 나타낸다. 아래 (식 5)와 같이 cost function을 이용하여 최소제곱 최적화로 전체 영상에 적용하기 위한 transformation model의 파라미터를 추정할 수 있다.

$$J = \frac{1}{2} \sum_{i=1}^N (f_i^{t-1} - T_{i-1}^t(f_i^{t-1}))^2 \quad ; \quad (\text{식 } 5)$$

where N : number of features,  
T : transformation model

### 3. 실험 결과 및 향후 과제

위 본문에서는 언급한 것과 같이 배경이 고정되지 않은 환경에서 영상차를 이용하여 이동하는 물체를 검출하기 위해서는 움직이는 물체 이외의 움직임, 즉 카메라의 ego-motion을 보정하여야 한다. ego-motion 보정한 경우와 보정하지 않은 경우에 대하여 각각 시간 t-1에서의 이미지 프레임  $I^{t-1}$ 과 시간 t에서의 이미지 프레임  $I^t$ 의 영상차를 이용하여 검출 가능 여부를 실험하였다.

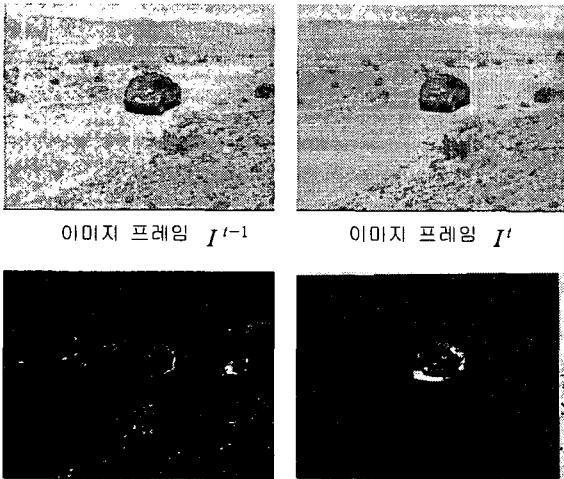


그림 2 직접비교(좌), ego-motion 보정후 비교(우)

위 그림 2의 실험영상에서 알 수 있듯이 배경의 움직임으로 물체와 배경을 구분하기 힘든 좌측 영상에 비하여 ego-motion을 보정한 우측 영상의 경우 움직이는 물체가 두드러지게 나타난다. 단순히 ego-motion을 보정하여 배경의 움직임이 모두 제거하였다고 할 수는 없다.

본 논문에서 KLT 알고리즘을 이용하여 연속된 영상에서 특징점을 추출하고 연속한 두 영상의 특징점을 매칭하여 영상좌표의 변환 파라미터를 추정하고 ego-motion을 보정하는 방법을 사용하였다. 영상의 좌표계

변환에 의한 움직임 검출 방법은 영상 환경에 따라 성능차가 나타나고 많은 한계점을 갖고 있다. 물체가 영상의 앞/뒤 방향으로 움직이거나 배경의 움직임과 비슷한 속도로 물체가 움직일 경우 좋지 않은 영상보정 결과가 나왔다. 특징점 매칭을 이용하여 연속된 프레임에서 특징점의 이동벡터 산출하고, 이를 이용하여 배경으로 추정되는 좌표와 물체로 추정되는 좌표를 분리하여 변환 파라미터 추정에 활용하면 성능이 향상될 것이다. 향후과제로 실용적인 ego-motion 보정 기법을 연구하여 좌표계 변환 모델과 함께 적용하여 배경이 움직이는 영상에서도 영상차를 이용하여 움직이는 물체를 검출하고, 이동물체를 지속적으로 추적할 수 있도록 위치예측(estimation)을 위한 particle 필터 또는 칼만 필터를 적용한 물체추적 시스템을 연구한다.

### 참고문헌

- [1] A. Murat Teklap, "Digital Video Processing", Prentice Hall, 1995.
- [2] Don Murray, Anup Basu, "Motion tracking with an active camera.", IEEE Transactions on Pattern Analysis and Machine Intelligence, 16(5):449-459, 1994.
- [3] Di Zhong, Shih-Fu Chang, "Long-term moving object segmentation and tracking using spatio-temporal consistency", Proc. Of IEEE Image Processing, vol.2, pp.57-60, 2001
- [4] Shu-Ching Chen, Mei-Ling Shyu, Peeta, S. Chengcui Zhang, "Adaptive Background Learning for Vehicle Detection and Spatio-Temporal Tracking" IEEE Transactions Intelligent Transportation Systems, 4(3):154-167, 2003.
- [5] BD Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision (IJCAI)," Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI '81):674-679, 1981.
- [6] J. Shi, C. Tomasi, "Good Features to Track", IEEE Conference on Computer Vision and Pattern Recognition, 1994.
- [7] C. Tomasi, T. Kanade, "Detection and tracking of point features.", Carnegie Mellon University Technical Report CMU-CS-91-132, 1991.
- [8] A. Censi, A. Fusiello, V. Roberto, "Image Stabilization by feature tracking", the 10th International Conference on Image Analysis and Processing:665-667, 1999.
- [9] G. Foresti C. Michelon, "A robust feature tracker for active surveillance of outdoor scenes", Electronic Letters on Computer Vision and Image Analysis 1(1):21-34, 2003
- [10] Yiwei Wang, Doherty, J.F., Van Dyck, R.E., "Moving object tracking in video", Applied Imagery Pattern Recognition Workshop, 2000. Proceedings. 29th:95-101, 2000