

색, 거리정보를 이용한 강인한 객체추적

이윤형^{*,**}, 정문호^{*}, 박민웅^{**}

한국과학기술연구원 지능로봇연구센터^{*}, 연세 대학교 전기전자공학과^{**}

Robust Tracking using color and depth

Lee, Yoon-Hyung^{*,**}, Jeong, Moon-Ho^{*}, Park, Mignon^{**}.

Korea Institute of Science and Technology Intelligent Robotics Reacher Center^{*},
Yonsei University Electrical and Electronic Engineering^{**}.

Abstract - 이 논문은 비강체 객체에 대한 실시간 추적시 다른 객체에 의한 간섭의 영향을 줄이는 방법을 제시한다. 제시한 방법에서 객체 추적에 대한 알고리즘은 현재 프레임에서 확률적으로 목표의 위치를 찾는 Mean shift 방법에 기초를 두고 있다. 기존의 방법에서는 mean shift의 파라미터로서 색분포만 사용한다. 하나의 파라미터에만 의존하므로 같은 색분포를 갖는 다른 객체가 추적 영역 내에 들어오게 되면 새로운 객체를 기존 객체로 인식하게 되는 문제가 발생한다. 여기서 우리는 강인한 객체추적을 하기 위해 다른 하나의 파라미터로서 거리정보를 이용을 제안한다. 거리정보에 최적화된 흐름 추정(optical flow estimation)방법을 확장 도입한 강인한 에러 기준(robust error norm) 방법을 사용하여 기존의 객체에 대해 더 많은 가중치를 주는 방식으로 mean shift 추적 방법을 기초로 하여 강력하게 추적하는 방법을 제안한다.

1. 서 론

영상 센서의 발전으로 그것의 사용 분야가 점점 확대되고 있다. 그에 따라 그것을 이용한 분야로 서비스 로봇, 보안 시스템, 공장에서의 공정 감시 등을 들 수 있겠다. 그 중 로봇의 지능을 위해 많이 사용되고 있는데 이 분야에서는 대표적으로 객체 추적, 객체 인식, 모션 인식 등에 이용된다. 로봇에서는 어떠한 객체를 추적해야 할 필요성이 부각된다. 일반적으로 객체 추적에서는 컬러의 분포를 이용한 방법[2]을 Mean Shift[1] 방법에 적용하여 확률적으로 다음 위치를 추정한다. 이 논문에서도 객체 추적 알고리즘은 mean shift 기법에 바탕을 두고 있다. 서비스 로봇의 객체 추적에 적용하였을 때 우리는 컬러 정보만을 사용하면 목표 객체를 추적하면서 추적 영역 내에 같은 컬러 정보를 가진 객체가 나타나면 추적에 대한 혼동이 발생함을 발견했다. 이와 같은 문제를 해결하기 위한 방법으로 스테레오 영상 센서로 Depth map을 결정하여 거리 정보를 활용하였다.

거리 정보에 대한 매치 정도를 확인하는 함수로서 영상의 매치 정도에 사용되는 강력한 에러 기준 함수를 사용했다.[9] 섹션 2.1은 컬러 정보를 어떻게 이용하는지에 대해 간략하게 설명하고 있다. 섹션 2.2는 어떻게 거리 정보를 이용하여 각각의 픽셀에 가중치를 주었는지 설명한다. 또한 섹션 2.3에서는 컬러 정보와 거리 정보를 이용하여 기존의 Mean Shift[8]에 거리 정보를 이용하여 실험에 사용된 식을 구현했다. 섹션 2.1과 2.3은 이 논문에서 자세하게 논의하지 않고 간략하게만 보이겠다. 각각의 방법에 대한 자세한 설명은 [2], [1]을 참조하기 바란다.

2. 본 론

2.1 컬러 정보의 이용

컬러 정보의 이용에 대한 표현[2]은 다음과 같은 식을 따른다. 목표 모델: 일반적인 좌표 x 와 y 는 각각 h_x , h_y 으로 정규화 된다는 가정에 의해 목표 모델에서 컬러 u 의 확률은

$$\hat{q}_u = C \sum_{i=1}^n k(\|x_i^* - u\|) \delta(b(x_i^*) - u) \quad (1)$$

후보 모델: 컬러 u 의 확률은

$$\hat{p}_u(y) = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \delta(b(x_i) - u) \quad (2)$$

각각의 픽셀이 갖는 컬러 정보에 대한 가중치 함수[1]는

$$w_i = \sum_{u=1}^m \delta(b(x_i) - u) \sqrt{\frac{q_u}{p_u(y_0)}} \quad (3)$$

위와 같이 정의한다.

이 논문에서 사용된 컬러 정보는 R, G, B의 전체 데이터를 사용하지 않고 계산의 간략화를 위해 레드를 표현하는 8비트 중 상위 3비트를 사용하였고 그린의 상위 3비트 블루의 2비트를 사용하여 256개의 정보로 축소시켰다. 즉, u 는 8비트를 갖는 정보가 된다.

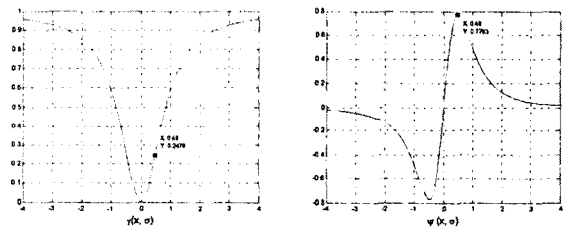
2.2 거리 정보의 이용

기준이 되는 n 개의 데이터 집합을 $\{e_i\}_{i=1, \dots, n}$ 라고 정의 하자. 이 정보에 대해 근사화된 정보 집합을 $\{e_i^*\}_{i=1, \dots, n}$ 라고 표현할 수 있다. 두 집합 사이의 매칭 정도는 정보 n 개 전체에 적용하여 합으로 표현되는 '최소 자승 오차'를 사용한다. 이전 프레임에서의 데이터와 현재 프레임과의 데이터 사이의 오차를 비교하는 척도로 사용된다.

$$E = \sum_{i=1}^n (e_i - e_i^*)^2 \quad (4)$$

이 과정은 이상적으로 정보의 외란이 없는 상태에서는 유용하지만, 실제 근사화된 데이터들은 외부 잡음을 포함하고 있어 잘못된 값을 가질 수 있으므로 적합하지 않다. 영상의 경우는 객체 이외의 배경이 쉬운 예이다. 그래서 각각의 픽셀 위치에서 기준데이터와 근사화된 데이터의 차이에 대해 다른 가중치를 주도록 했다. 가중치 함수로서 최적 흐름 추정 방법을 확장[8, 9] 도입한 '강인한 오차 기준' γ 을 사용한다.

$$E = \sum_{i=1}^n \gamma((e_i - e_i^*), \sigma) \quad (5)$$



<그림 1> 강력한 오차 기준과 그것의 미분

σ 는 스케일 파라미터로서 본 논문에서는 실험을 통하여 결정하였다. 또한,

$$\gamma(x, \sigma) = \frac{x^2}{x^2 + \sigma}, \quad \gamma'(x, \sigma) = \psi(x, \sigma) = \frac{2x\sigma^2}{(x^2 + \sigma)^2} \quad (6)$$

그림에서 보듯이 함수 γ 는 '거절', 다운-가중치의 형태를 가진다. 함수 ψ 는 γ 의 미분으로 결정되는 함수로서 오차 $(e_i - e_i^*)$ 가 주는 영향의 특성을 나타낸다.

σ 는 이상치의 영향이 감소하기 시작하는 위치에 영향을 준다. σ 가 커질수록 γ 함수의 그래프에서 곡선의 급격함의 정도가 완만해진다. 그리고 ψ 함수의 그래프에서는 변곡점의 위치가 영에서 점점 멀어지게 된다. 함수 ψ 에 대한 실험에서 이상치 거절은 γ 의 이차 미분이 0이 되는 시점부터 시작된다. 즉, 강력한 오차 기준 그래프에서 변곡점 위치에 영향을 준다. 이것은 $|e_i - e_i^*| > \sigma/\sqrt{3}$ 가 만족하는 위치부터 이상치로 인정할 수 있음을 의미한다. 우리는 임계값 함수로서

$$m_i = \begin{cases} 1 & |e_i - e_i^*| \leq \sigma/\sqrt{3} \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

를 정의할 수 있다.

이제 실험에 사용된 목표 모델과 후보 모델을 정의한다.

목표 모델: 최초 목표 객체 모델에 대한 거리 정보에는 다른 객체는 포함되지 않는다고 가정한다. 모델의 n 개의 픽셀 위치 집합

을 $\{x_i^*\}_{i=1, \dots, n}$ 이라고 하자. 함수 $b(x_i)$ 와 $d(x_i)$ 는 각각 픽셀 위치 x_i 에서 컬러 정보를 갖고, 거리 정보를 갖는다고 정의한다. 특정 컬러 u 를 갖는 거리 정보를 다음과 같이 나타낼 수 있다.

$$e_i^* = d(x_i^*) \sum_{u=a}^b \delta[b(x_i^*) - u]_{i=1, \dots, n} \quad (8)$$

또한 정보에 대한 평균으로써

$$\bar{e}^* = \frac{1}{n} \sum_{i=1}^n d(x_i^*) \sum_{u=a}^b \delta[b(x_i^*) - u] \quad (9)$$

를 정의 한다.

후보 모델 : 현재 프레임에서 n_h 개의 특정 컬러 u 를 갖는 위치 x_i 에 대한 거리 정보는 아래와 같이 표현한다.

$$e_i = d(x_i) \sum_{u=a}^b \delta[b(x_i) - u]_{i=1, \dots, n_h} \quad (10)$$

목표 모델과 후보모델과의 차이를 강력한 오차 기준함수 $\gamma(|e_i - e_i^*|, \sigma)$ 에 적용한다. 이 방법을 사용함으로써, 새로운 객체가 다른 거리정보를 갖고 후보 모델의 정보에 포함되었다면, γ 에 의해 이상치에 해당하는 정보를 분류할 수 있다. 본 논문에서는 기존 데이터에 대해서 e_i^* 를 사용하지 않고 \bar{e}^* 를 사용하였다. 이유는 객체의 스케일에 따라 이전 프레임에서의 거리정보의 개수와 현재 프레임에서의 개수가 다를 수 있기 때문이다. 기존 객체의 이동으로 인해 $|e_i^* - e_i| = e_i^*$ or e_i 의 결과를 가질 때 잘못된 연산을 수행할 수 있기 때문이다.

거리정보는 컬러값에 분포와는 달리 일정하다고 가정할 수 없다. 정보가 일정하지 않고 지속적으로 변화하는 값이므로 현재 프레임에서 임계값 함수 m_i 를 사용하여 새로운 목표 모델을 만들어야 한다.

$$e_i^* = e_i m_i, \quad \bar{e}^* = \frac{1}{n_h} \sum_{i=1}^{n_h} e_i m_i \quad (11)$$

n_h 는 기존 현재 프레임에서 n_h 개의 데이터가 있지만 임계값 함수에 의해 개수가 바뀌게 되므로 새롭게 표기하였다.

2.3 Mean Shift와의 결합

Mean-Shift의 이전 위치 \hat{y}_0 에서 다음 위치 \hat{y}_1 를 추정하는 방법[8]에 각각의 픽셀의 거리 정보를 이용한 식은

$$\hat{y}_1 = \frac{\sum_{i=1}^{n_h} x_i \gamma_i w_i g \left(\left\| \frac{\hat{y}_0 - x_i}{h} \right\|^2 \right)}{\sum_{i=1}^{n_h} \gamma_i w_i g \left(\left\| \frac{\hat{y}_0 - x_i}{h} \right\|^2 \right)} \quad (12)$$

다음과 같이 표현한다. 컬러 정보에 대한 가중치 함수 w_i 는 (3)을 통해 구할 수 있다. 또한 커널 g 에 대한 설명은 [2]를 참조하기 바란다.

3. 결 론

이 논문의 실험을 위해 Bumblebee 스테레오 영상센서를 통해 입력되는 320 X 240 영상과 거리 정보를 사용하였다.

<그림 2>의 왼쪽 영상은 기존에 사용된 Mean Shift 방법을 사용하였고, 오른쪽 영상은 거리 정보를 함께 이용한 방법이다. 특정 컬러 u 는 피부색에 대한 정보를 사용하였다.

첫 번째 줄 그림에서 객체에 대한 초기화를 해주었다. 앞서 가정 했듯이 거리 정보에 대한 최초 평균의 연산을 위해 시작할 때 객체는 하나로 하였다. 두 번째 줄 영상에서는 기존 추적중인 영역에 같은 컬러 분포를 갖는 객체(손)가 등장 하면서 각각의 추적이 어떻게 반응하는지를 설명해 준다. 왼쪽의 영상에서는 영역의 중심이 아무런 변화를 갖지 않는 것을 확인할 수 있고, 오른쪽 영상에서는 새로운 객체의 등장에 따라 영역의 중심이 변화되는 것을 확인할 수 있다. 즉, 같은 특정 컬러 u 를 갖지만 다른 거리 정보를 사용함으로써 두 객체간의 혼동이 없다.

<그림 3>은 거리정보를 이용하여 사용되는 가중치를 보여준다. 다른 거리정보(손)가 (7)에 의해 이상치로 인지되면서 제거되는 현상을 확인할 수 있다.

<그림 2>에서 왼쪽과 오른쪽 영상의 추적기의 영역 스케일에 변화가 없다. 일반적인 Mean Shift에서는 Bhattacharyya 계수를 이용하여 영역을 크게 하여 계수를 계산하고, 작게 하여 계수를 계산하여 그 둘을 비교하여 영역의 스케일을 적용하는 적응 스케일[1] 방법으로 영역의 스케일을 변화시킨다. 하지만, 본 논문에서는 거리 정보를 사용하면서 스케일 변화에 대처하는 방법은 아직 찾아내지 못하였다. 따라서 이 실험에서는 기존 Mean Shift 방법에서도 스케일의 변화를 하지 않고 단지 객체를 추적함에 따른

영역의 중심의 변화를 비교하는 것에 중점을 두었다. 스케일의 변화가 아닌 같은 컬러 분포를 갖는 두 객체의 혼동에 강인하게 반응하는데 그 목적이 있다.

앞으로 영역에 대한 스케일 변화에 대처할 수 있는 방법을 찾아낸다면 보다 강력한 방법을 만들 수 있을 것으로 기대한다.



<그림 2> 실험 영상



<그림 3> 강력한 예러 기준 가중치

[참 고 문 헌]

- [1] D. Comaniciu, V. Ramesh and P. Meer, Real-Time Tracking of Non-Rigid Objects using Mean Shift, Computer Vision and Pattern Recognition (2000) 142-149.
- [2] K. Nummiaro, E. Koller-Meier, L. Van Gool, "Color Features for Tracking Non-Rigid Objects", ACTA Automatica Sinica, 2003.
- [3] S. Konishi, A.L. Yuille, J. Coughlan, S.C. Zhu, "Fundamental Bounds on Edge Detection: An Information Theoretic Evaluation of Dierent Edge Cues," IEEE Conf. on Comp. Vis. and Pat. Rec., Fort Collins, 573(579), 1999.
- [4] S.J. McKenna, Y. Raja, S. Gong, "Tracking Colour Objects using Adaptive Mixture Models," Image and Vision Computing, 17:223(229), 1999.
- [5] D.Comaniciu, P.Meer, "Mean Shift Analysis and Applications", IEEE Int'l Cinf. Comp.Vis., Kerkyra, Greece, 1197-1203, 1999.
- [6] C. Wren, A. Azarbayejani, T. Darrell, A. Pentland, "Pnder: Real-Time Tracking of the Human Body," IEEE Trans. Pattern Analysis Machine Intell., 19:780(785), 1997.
- [7] S.Birchfield, "Elliptical head using intensity gradients and color histograms", in Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Santa Barbara, CA, 1998, pp.232-237.
- [8] M. Black and P. Anandan. The robust estimation of multiple motions: Affine and piecewise smooth flow fields. Computer Vision and Image Understanding, in press. Also Tech. Report P93-00104, Xerox PARC, Dec.1993.
- [9] M. J. Black and P. Anandan. A framework for the robust estimation of optical flow. In ICCV-93, pp. 231-236, Berlin,