

비정형 객체추적을 위한 계층적 영상과 Kalman Filter 기반 능동형태모델

강진영, 기현종, 신정호, 백준기
중앙대학교 첨단영상대학원 영상공학과
시각 및 지능 시스템 연구실

Hierarchical image and Kalman filter-based active shape model for non-rigid object tracking

Jinyoung Kang, Hyunjong Ki, Jeongho Shin, and Joonki Paik

Image Processing and Intelligent Systems Laboratory

Department of Image Engineering

Graduate School of Advanced Imaging Science, Multimedia, and Film, Chung-Ang University

E-mail : zippykang@wm.cau.ac.kr

Abstract

In this paper, we present a hierarchical approach of an enhanced active shape model for video tracking. Kalman filter is used. To estimate a dynamic shape in video object tracking. The experimental results show that the proposed hierarchical active shape model using Kalman filter is efficient.

I. 서론

고정 형태를 갖지 않는 객체를 추출하고 추적하는 문제는 의료영상 분야나 보안시스템 등에서 날로 그 중요성을 더해가고 있다. 특히, 비디오 영상에서 움직이는 객체를 인식하고 추적하기 위한 연구들이 많이 진행되고 있으며, HCI(human computer interaction), 영상 합성, 애니메이션을 위한 모션캡처, 보안시스템 등에 널리 응용되고 있다.

기존의 모델 기반 컴퓨터 비전 기법은 정형(rigid) 객체를 대상으로 하여 주로 연구 되었으나, 움직이는 사람과 같이 추적하는 물체의 형태가 변하는 경우의 추적기술의 개발이 필요하다. 즉 정형모델이 아닌 비정형(non-rigid) 혹은 가변모델(deformable model)을

사용해야 한다. 이러한 비정형 모델을 이용한 객체추출 및 추적 시스템은 객체를 강건하게 추적할 수 있으며 가림현상(occlusion) 또한 극복할 수 있다.

본 논문에서는 비정형 객체를 추출하고 추적하기 위해서 비정형 모델의 일종인 능동형태모델(ASM: active shape model)을 적용한다. ASM은 객체 외형의 정보를 학습시키고 이 정보를 바탕으로 객체의 평균 형태에 변형을 가하여 새로운 이미지 내에서 객체를 찾는 방법이다.

ASM을 그대로 비디오 보안시스템에 사용하기에는 실시간 구현이 어렵다는 단점을 가지고 있다. 매 프레임마다 객체의 윤곽을 모델과 정합시키기 위하여 반복적인 계산이 필요하기 때문이다. 따라서 본 논문에서는 ASM기반 객체추적 기법의 뛰어난 성능을 유지하면서, 동시에 실시간 구현이 가능한 시스템을 구현하기 위해 계산량을 줄이고, 연속적인 영상에서 다음영상의 객체의 위치를 Kalman filter를 이용해서 추정하고자 한다.

II. 능동형태모델

ASM은 찾고자 하는 객체 외곽 형태의 선형정보(prior)를 이용한다. 따라서 불완전하거나

잡음이 많이 들어간 영상에서도 강건하게 객체의 외곽을 정합하여 추출할 수 있다. 이 모델은 학습 집합(training set)을 필요로 하는데, 사용자가 직접 훈련 영상으로부터 객체의 외곽에 기준점들을 표시하여 여러 가지 형태 모델을 구축해야 한다.

2.1 기준점 할당

객체의 윤곽선을 잘 표현하기 위해서 주어진 입력영상으로부터 적당한 기준점(Landmark point)을 할당한다. 이 작업은 주로 수동으로 이루어진다.

2차원 영상에서 우리는 n 개의 기준점을 $2n$ 차원의 벡터에 의해서 다음과 같이 표현할 수 있다.

$$x = [x_1, \dots, x_n, y_1, \dots, y_n]^T \quad (1)$$

본 논문에서 적용될 실험에 사용될 기준점의 수는 수동적으로 42($n=42$)로 했다. 기준점의 역할은 모델의 윤곽 형태를 조절하는 것이다.

2.2 Principal Component Analysis(PCA)

n 개의 기준점의 집합은 객체의 형태를 표현한다. 학습집합(training set)에서 각각의 형태는 2차원 공간에서 존재하지만, 우리는 PCA 기술을 기반으로 하여 좀 더 적은 수의 매개변수를 가지고 객체의 형태를 효과적으로 모델링할 수 있다. 구체적인 PCA 알고리즘을 살펴보면 다음과 같다.

1. 훈련집합에서 m 개의 표본 형태의 평균을 구한다.

$$\bar{x} = \frac{1}{m} \sum_{i=1}^m x_i \quad (2)$$

여기서 m 은 각각의 형태에 따른 훈련집합을 의미한다 ($i=1, \dots, m$).

2. 훈련집합으로부터 공분산 행렬(covariance matrix)을 구한다.

$$S = \frac{1}{m} \sum_{i=1}^m (x_i - \bar{x})(x_i - \bar{x})^T \quad (3)$$

3. 다음과 같은 행렬을 만든다.

$$\Phi = [\phi_1 | \phi_2 | \dots | \phi_q] \quad (4)$$

여기서 ϕ_j (단, $j=1, 2, \dots, q$)는 S 에서 q 개의 가장 큰 고유값(eigen value)에 해당하는 고유벡터(eigen vector)를 의미한다. 이 때 q 는 데이터의 전체적인 분산 값에 의

98%에 해당하는 모델형태의 개수를 사용한다.

4. 주어진 Φ 와 \bar{x} 로 부터 각각의 형태는 다음과 같이 간략화 시킬 수 있다.

$$x_i \approx \bar{x} + \Phi b_i \quad (5)$$

여기서

$$b_i = \Phi^T (x_i - \bar{x}) \quad (6)$$

벡터 b 는 변형 가능한 모델 파라미터의 집합을 정의한다. b 의 요소를 변하시키면서 형태를 변화시킬 수 있다.

2.3 모델정합(Model fitting)

모델 좌표프레임 x 와 영상 좌표프레임 y 를 정합하기 위해서 다음과 같은 에러 함수를 최소화함으로써 이루어질 수 있다.

$$E = (y - Mx)^T W (y - Mx) \quad (7)$$

여기서 W 는 각각 기준점에 대한 가중치를 나타낸다. 그리고 M 은 회전성분 θ , 전이성분 t , 스케일링 성분 s 로 이루어진 기하학적인 변환을 나타내는 행렬이다. 하나의 점을 $[x_0, y_0]^T$ 으로 표현한다면, 기하학적인 변환행렬은 다음과 같이 정의 될 수 있다.

$$M \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = s \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (8)$$

기하학적인 파라미터가 결정되면, 모델좌표 프레임으로 투영된 값은 다음과 같이 주어진다.

$$x_p = M^{-1} y \quad (9)$$

결국 모델의 파라미터는 다음과 같이 갱신된다.

$$b = \Phi^T (x_p - \bar{x}) \quad (10)$$

최적으로 갱신된 기준점의 조합은 영상좌표 프레임 y 에 새로운 형태를 만들어낸다. 새로운 형태는 식 (7)을 사용하여 가장 근접한 형태를 찾는 데 이용될 수 있다. M 을 통해서 최적의 위치를 도출한 뒤, 이 새로운 형태는 주어진 학습집합에 주요 요소에 포함된 Φ 로 투영될 수 있다. 이 과정은 모델파라미터 b 를 갱신한다. 그 결과로 주요 요소에 해당되는 비슷한 편차만이 모델 파라미터에 영향을 줄 수 있다. 모델 파라미터를 계산한 후에 x 에 의해 주어지는 새로운

형태는 식 (5)에 의해 구할 수 있고, 적당한 반복을 거쳐, 최종적인 형태가 x 로 얻어진다.

2.4 지역구조 모델링

ASM을 사용하여 입력영상의 형태를 해석하기 위해서는 영상과 모델을 최적으로 정합시키는 기준점의 집합을 찾아야 한다.

우선 j 번째 기준점을 가로지르는 K 만큼의 길이를 갖는 지역적인 윤곽선(profile)의 정규화된 미분값을 $g_j(j=0, \dots, n)$, 해당하는 평균과 분산을 각각 \bar{g}_j 와 S_j 로 둔다. 가장 근접한 윤곽을 구하기 위해서 모델의 표본과 평균사이의 Mahalanobis 거리를 최소화함으로써 얻을 수 있다. Mahalanobis 거리는 다음과 같이 나타낼 수 있다.

$$f(g_{j,m}) = (g_{j,m} - \bar{g}_j)^T S_j^{-1} (g_{j,m} - \bar{g}_j). \quad (11)$$

여기서 $g_{j,m}$ 은 해당되는 경계의 수직방향에 존재하는 한 m 개의 표본만큼 이동된 g_j 를 나타낸다. 실제적으로 본 논문에서는 좀 더 근접한 윤곽을 찾기 위해서 계층적인 기술을 사용한다.

III. 계층적 접근과 Kalman filter

ASM의 계층적인 접근은 여러 가지 형태와 크기의 객체를 다루는 데 있어서 필수적이다 [3]. 계층적 모델은 비 계층적 모델의 방법에 비해 계산량이나 성능면에 있어 상당히 효과적이고 우월하다. Baumberg는 Kalman filter가 형태를 정합하는데 있어서의 성능 향상을 위해 계층적 스네이크 모델에 어떻게 적용될 수 있는지 연구했다[1]. 그는 계층적으로 표본화의 정도를 작은 값에서 큰 값으로 변화시켜가면서 기준점의 수를 변화시켰다. 본 논문에서는 스네이크 알고리즘 대신에 ASM방법을 사용하였다. Cootes는 자신의 ASM을 계층적으로 확장하였다 [2]. 그가 제안한 방식은 일반적인 계층적 방식과 마찬가지로 저해상도에서 해를 구하고 고해상도로 옮겨간 다음 그 해를 수정한다.

본 논문에서는 Baumberg의 방법과 Cootes의 방법 모두에 기반을 두고 계층적 ASM 알고리즘을

제안하였다.

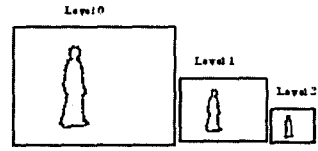


그림 1. 계층적 능동형태모델

또한 다음 상태를 예측할 수 있는 Kalman filter를 이용해서 연속되는 영상에서 이전 영상의 정보를 가지고 다음영상의 정보를 예측하여 보다 빨리 연속적인 영상에서 물체를 추적하는 것에 한걸음 더 다가가 고자 한다.

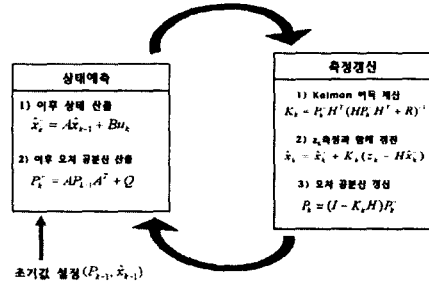


그림 2. Kalman filter 동작의 상태도

IV. 실험 결과

본 절에서는 제안된 알고리즘의 성능을 실험을 통하여 확인한다. 그림 3은 초기윤곽을, 그림 4는 원영상(Level0), 원영상의 1/2(Level1), 원영상의 1/4(Level2)로 모델정합을 한 결과를 보여주고 있다.



그림 3. 초기 윤곽

그림 5는 계산량을 줄이기 위해서 기준점의 개수를 달리해서 계층적으로 모델을 정합한 결과이다. 그림 6은 이전의 연속된 두 영상(k-2, k-1번째)에서 중심점을 찾고 이 점을 Kalman filter를 이용하여 다음(k번째)점을 추정 한 후 그 점을 중심으로 k번째 영상의 윤곽을 찾은 것이다. 이러한 결과 또한 계층적으로 수행하게 되면 계산량을 줄이게 된다. 그림 7은 가림현상에서

있어서도 능동형태 모델의 강건함을 보여주는 실험 결과이다.



그림 4. 정합결과 (Level0, Level1, Level2)



그림 5. 기준점의 개수에 따른 계층적 정합결과



그림 6. Kalman filter이용 다음 영상 윤곽 추정



그림 7. 가림현상에서 모델정합

실험결과에서 볼 수 있듯이 능동형태모델은 가림현상에 강건하며, 계층적 모델을 사용할 경우 계산량의 줄임과 동시에 좀 더 정확한 물체의 윤곽을 찾을 수 있었다. 또한 Kalman filter를 이용하면 연속적인 영상에서 다음 영상에서의 물체를 추적할 수 있기 때문에 실시간 객체추적기술에 쓰일 수 있다.

V. 결론

본 논문에서는 비정형 물체를 강건하게 추적하기 위한 계층적 ASM 기반 비디오 추적 시스템을 제안하였다.

제안된 시스템은 기존의 ASM 기반 추적 시스템을 실시간으로 구현하기 위하여 계층적으로 구현하였으며, 각 계층에서의 계산량을 줄이기 위하여 기준점의 개수를 적응적으로 조절하였다. 또한 Kalman filter를 이용하여 다음 영상의 위치를 알 수 있으므로 해서 비정형 객체를 실시간으로 추적할 수 있으며 가림현상을 극복할 수 있음을 확인하였다.

본 시스템의 성능을 더욱 향상시키기 위하여 앞으로 컬러정보를 이용할 수 있도록 확장할 것이며 사람뿐만 아니라 여러 가지 다양한 객체를 추적할 수 있도록 시스템을 개선할 것이다. 또한, 변형가능한 모델의 시간적 상관 관계를 모델 예측에 이용하여 더욱 강건하고 빠른 시스템을 제안할 것이다.

감사의 글

본 연구는 교육부의 BK21 사업과, 과학기술부의 국가 지정연구실 사업 지원으로 이루어 졌습니다.

참고문헌

- [1] A. Baumberg, "Hierarchical shape fitting using an iterated linear filter," *Image and Vision Computing*, vol. 16, pp. 329-335.
- [2] T. F. Cootes, C. J. Taylor, and A. Lanitis, "Active shape models: evaluation of a multi-resolution method for improving image search," *Proc. British Machine Vision Conference*, pp. 327-336, 1994.
- [3] A. Koschan, S. Kang, J. Paik, B. Abidi, and M. Abidi, "Color active shape models for tracking non-rigid objects," *Pattern Recognition Letters*, vol. 24, no. 11, pp. 1751-1765, July 2003.