

# 웨이블릿을 이용한 계층적 능동형태모델 기반 비디오 추적기술

기 헌 중, Vivek Maik, 정 정 훈, 신 정 호, 백 준 기  
중앙대학교 첨단영상대학원 영상공학과  
시각 및 지능 시스템 연구실

## Hierarchical active shape model-based video object tracking using wavelet transform

Hyunjong Ki, Vivek Maik, Junghoon Jung, 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 : kikingka@wm.cau.ac.kr

### Abstract

This paper proposes a hierarchical approach to active shape model using wavelet transform. The proposed algorithm allows us to use both global shape characteristics and finer details for model deformation. The statistical properties of the wavelet transform of a deformable model are analyzed by principal component analysis and used as priors in the contour's deformation.

### 1. 서 론

지금까지 비디오 추적기술은 움직이는 물체의 위치를 찾는 데 있어 그 성능을 인정받고 있으며, 최근 들어 움직이는 사람의 경우와 같이 추적되는 물체의 형태가 변하는 경우 형태를 동시에 추정하는 영상처리기술이 필요하다. 즉, 정형 모델이 아닌 비정형(non-rigid) 혹은 변형가능(deformable) 모델을 사용해야 한다. 비정형 모델을 이용한 비디오 추적 시스템은 추적 객체에 대한 일련의 정보를 가지고 학습할 수 있기 때문에 객체를 인식할 수 있고, 가림현상(occlusion)을 극복할 수 있기 때문에 강건하게 객체를 추적할 수 있다.

본 논문에서는 비정형 객체를 추출하고 추적하기 위해서 비정형 모델의 일종인 능동 형태 모델(ASM: active shape model)을 적용한다. ASM은 객체 윤곽 정보를 기존의 학습모델들의 통계적인 특성을 바탕으로 해서 좀 더 효과적으로 추정하는 방법이다. 하지만 매 프레임마다 객체

의 윤곽을 모델과 정합시키기 위하여 반복적인 계산이 필요하기 때문에 ASM을 그대로 비디오 보안시스템에 사용하기에는 실시간 구현이 어렵다는 단점을 가지고 있다. 최근 이를 해결하기 위해 계층적인 (hierarchical approach)방법이 소개되었다 [1]. 이는 다해상도 영상 피라미드(pyramid)를 기반으로 하여 낮은 해상도로부터 점차 높은 해상도로 객체의 찾는 방법으로 계산량이나 정확도에 있어서 효과적이다. 본 논문에서는 이것을 발전시켜 웨이블릿 변환기반으로 이용하여, 계층적인 접근은 물론 객체의 윤곽을 찾는 데 있어 좀 더 효과적이고 강건한 방법을 제안하려고 한다.

본 논문은 구성은 다음과 같다. II장에서는 능동형태 모델에 대한 기본 개념을 설명하고, 그것을 이용한 비정형 객체 추적에 대하여 살펴본다. III장에서는 웨이블릿에 대한 간단한 소개를 하고, IV장에서는 본 논문에서 제안하는 웨이블릿 이용한 계층적 ASM 모델에 대하여 설명한다. V장에서는 여러 실험 결과를 비교 정리하고, 마지막으로 VI장에서 향후 보완점과 응용방안을 제안하고 결론을 짓는다.

### 2. 능동형태모델

ASM은 추적하고자 하는 객체윤곽의 학습모델을 사용하기 때문에 불완전하거나 잡음이 많이 들어간 영상에서도 강건하게 객체의 윤곽을 효과적으로 추정할 수 있다 [2].

좀 더 구체적으로 본 논문의 ASM 기반 추적 알고리즘은 다음과 같이 크게 네 가지 단계로 나눌 수 있다: (i) 기준점(Landmark point)의 할당 (ii) PCA(Principal component analysis), (iii) 모델 정합(Model fitting), 그

리고 (iv) 지역구조(Local structure) 모델링.

### 2.1 기준점(Landmark point) 할당

기준점을 적용하기 적당한 주어진 입력영상으로부터 객체의 윤곽선을 할당한다. 기준점의 역할은 모델의 윤곽 형태를 조절하는 것이다. 2차원 영상에서 우리는 n 개의 기준점을 2n차원의 벡터에 의해서 다음과 같이 표현할 수 있다.

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

### 2.2 PCA

PCA 기술을 기반으로 하여 좀 더 적은 수의 매개변수를 가지고 객체의 형태를 효과적으로 모델링할 수 있다. 훈련집합에서 m개의 형태를 가지고 있다고 가정하자.

구체적인 PCA 알고리즘을 살펴보면 다음과 같다.

#### PCA algorithm

1. 훈련집합에서 m개의 표본 형태의 평균을 구한다.

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

여기서  $x_i$ 는 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  $\phi_j$ (단,  $j = 1, \dots, q$ )는 S에서 q개의 가장 큰 고유값(eigen value)에 해당하는 고유벡터(eigen vector)를 의미한다.

4. 주어진  $\Phi$ 와 x로부터 각각의 형태는 다음과 같이 나타낼 수 있다.

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

여기서 벡터  $b_i$ 는

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

과 같이 정의되는 변형 가능한 모델 파라미터의 집합을 나타낸다. PCA 알고리즘의 3번째 단계에서 q값은 데이터의 전체적인 분산값의 98%에 해당하는 모델형태의 개수를 사용한다. ■

### 2.3 모델정합(model fitting)

모델 좌표프레임의 형태인 x와 영상 좌표프레임의 형태인 y를 정합하기 위해서 다음과 같은 에러 함수를 최소화함으로써 최적의 모델위치와 형태를 찾을 수 있다.

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

여기서 W는 대각행렬이며, 그 대각성분은 각각 기준점에 대한 가중치를 나타낸다. 그리고 M은 회전성분  $\theta$ , 전이성분 t, 스케일링 성분 s로 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 파라미터의 집합( $\theta, t, s$ )이 결정되면, 모델좌표 프레임으로 투영된 y값은 다음과 같이 주어진다.

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

결국 모델의 파라미터는 다음과 같이 갱신된다.

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

윤곽을 따라 위 단계들의 결과로써 최적의 기준점의 변위를 얻을 수 있다.

### 2.4 지역구조 모델링 (Local structure modeling)

통계적이고 변형가능한 모델은 PCA와 모델 정합과정을 거쳐 기준점을 할당함으로써 표현할 수 있다. ASM을 사용하여 입력영상의 형태를 해석하기 위해서는 영상의 모델을 최적으로 정합시키는 기준점의 집합을 찾아내야 한다.

우선, j번째 기준점을 가로지르는 K만큼의 길이를 갖는 지역적인 윤곽선의 정규화된 미분값을  $g_j$ , 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 다음 식과 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 m개의 표본만큼 이동된  $g_j$ 을 나타낸다. 실제적으로 본 논문에서는 좀 더 근접한 윤곽을 찾기 위해서 계층적인 기술을 사용한다.

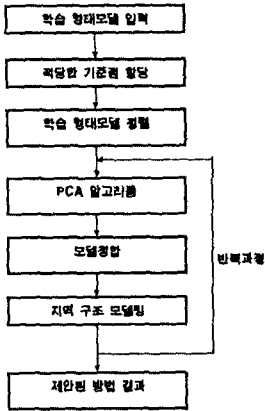


그림1. 전체적인 ASM 알고리즘 흐름도

### 3. 웨이블릿 변환

웨이블릿 변환은 웨이블릿 기저를 사용하여 신호를 분해하는 일종의 서브밴드 필터링이라 할 수 있다. 웨이블릿 변환은 입력신호를 여러 계층으로 분해할 수 있는데, 레벨이 증가할수록 주파수에 대한 해상도는 좋아지는 반면에, 시간에 대해서는 해상도가 떨어진다. 또한 이 때 사용하는 웨이블릿 기저함수에 따라 분해능력이 달라지는 데, 본 논문에서는 Daubechies-4기저를 이용하여 대역분할하였다.

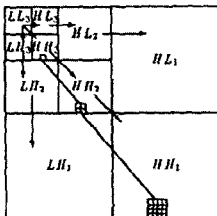


그림 2. 웨이블릿을 이용한 영상 신호 분할

### 4. 제안한 방법

앞서 서론에서 언급했듯이 ASM을 비디오 추적에 적용할 경우 매 프레임마다 모델 정합을 수행하면 계산량이 많기 때문에 효과적인 계산량을 위해 계층적인 방법을 사용할 수 있다 [3]. 따라서 본 논문에서는 웨이블릿 변환을 이용하여 ASM을 계층적으로 확장하였다. 기존의 계층적인 방법을 적용할 경우, 지역구조 모델링을 수행하는 데 있어 각 기준점에 수직방향에 따라 모든 방향에지를 고려한 그레디언트(gradient)영역에서 찾기 때문에 주위 배경에 강한 에

지영향을 받아 잘못된 윤곽을 찾는 경우가 발생할 수 있다. 하지만 본 논문에서 제안하는 방식은 일반적인 계층적 방식과 마찬가지로 저해상도에서 윤곽을 추정하지만, 웨이블릿 특성을 이용하여 각 기준점의 수직 방향에 따라 다른 그레디언트영역에서 물체를 윤곽을 찾으려고 한다. 구체적으로 살펴보면, 그림 2에서 보듯이 n 레벨상에서 기준점 수직방향에 수평인 경우 수평방향 에지인 HLn대역을 이용하고, 수직인 경우에는 LHn대역을, 대각인 경우에는 HHn대역을 이용하여 지역구조 모델링을 하는 데 있어 효과적인 결과를 가져올 수 있다.

제안하는 전체적인 알고리즘은 다음과 같다.

#### 제안 알고리즘 :

1. 초기 시작 레벨의 정의한다 ( $L$ =시작레벨)
2.  $L$ 레벨 상에서의 이미지에서 모델정합
3. 각각의 기준점에 대한 수직방향의 각도에 따른 주어진 웨이블릿 서브밴드에서 기준점 위치를 갱신
4. PCA알고리즘을 통해 모든 기준점 갱신
5. 레벨을 낮춘다. ( $L \rightarrow L-1$ )
6.  $L \geq 0$  일 때 까지 단계 2)~단계 4)를 반복한다. ■

### 5. 실험 결과

본 논문에서 제안된 알고리즘의 성능을 평가하기 위해 여러 가지 실험을 수행하였다. 실험에 사용된 PCA알고리즘을 위한 학습모델의 수는 50개가 사용되었다. 각각의 모델형태는 그림3에 나타나있다.

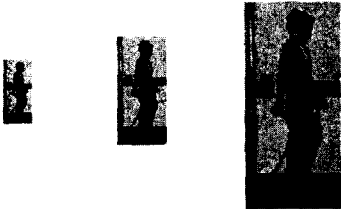


그림 3. 50개의 학습모델

첫 번째 실험에서는 계층적인 방법과 비계층적인 방법을 적용한 결과를 비교 정리하였다. 그림4는 각 레벨에 따른 이미지 사이즈를 보여준 것이고, 그림5는 초기 형태와 비계층적인 방법과 비계층적인 방법이 적용된 결과를 나타내었다. 계층적인 방법이 보다 효과적인 윤곽을 추출하고 있음을 확인할 수 있다.

두 번째 실험에서는 본 논문에서 제안하는 웨이블릿을 이용한 방법의 결과를 보여주고 있다. 그림6의 (a)는 2레벨에서의 웨이블릿 변환을 보여주고 있으며, 그림 6의 (b)는 본 논문에서 제안하는 웨이블릿의 각 서브밴드를 이용하여 찾은 결과이다. 기존의 계층적인 방법만 이용한 것보다 주위환경에 보다 더 강한 결과를 얻고 있음을 확인할 수 있다.

세 번째 실험은 고정점을 개수를 변화시켜면서 적용한 결과이다. 그림7에서 보듯이 적은 고정점으로도 충분히 물체의 윤곽을 추정하고 있음을 확인할 수 있다. 마지막 실험은 물체가 배경에 가려진 경우 얼마나 강건하게 물체를 추적하는 지 확인하는 실험이다. 그림8에서 보면, 약 50% 정도 가림현상이 있었지만, 물체의 전반적인 윤곽을 찾고 있음을 확인할 수 있다.



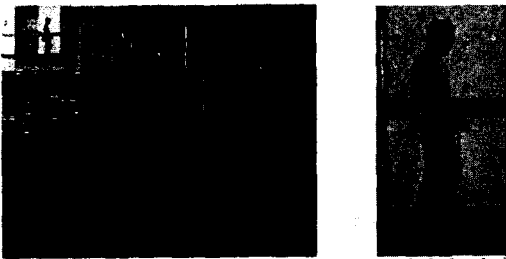
(a) 레벨 2 (b) 레벨 1 (c) 레벨 0

그림 4. 계층적인 방법에 사용한 세가지 다른 해상도



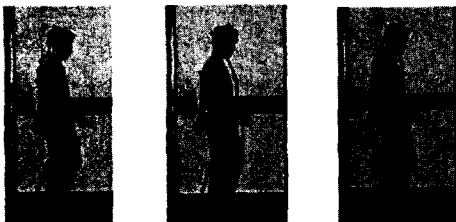
(a) (b) (c)

그림 5. (a) 초기 고정점 위치 (b) 계층적인 방법의 결과 (c) 비계층적인 방법의 결과



(a) 2레벨 웨이블릿 변환 (b) 제안한 방법

그림 6. 웨이블릿 변환 결과 및 제안한 방법의 결과



(a) 42개 (b) 21개 (c) 12개

그림 7. 고정점 개수에 따른 결과



(a) 가림현상 전 (b) 가림현상 후

그림 8. 가림현상에 따른 결과

## 6. 결론

본 논문에서는 비정형 물체를 효율적으로 추적하기 위해 웨이블릿을 사용하여 계층적인 ASM 비디오 추적 알고리즘을 제안하였다. 제안한 알고리즘은 ASM 기반 추적 시스템을 실시간으로 구현하기 위하여 계층적인 웨이블릿을 이용하였으며, 보다 효과적인 모델정합을 위해서 웨이블릿의 여러 서브밴드를 효율적으로 이용하였다. 앞으로 본 시스템의 성능을 더욱 향상시키기 위하여 컬러 웨이블릿을 이용하여 주위환경에 보다 강건하게 적용할 수 있도록 확장할 것이며, 사람뿐만 아니라 여러 가지 다양한 객체를 추적할 수 있도록 시스템을 개선할 것이다. 제안한 알고리즘은 공항, 지하철 등 사람들로 붐비는 공공장소나 도로와 같이 객체의 움직임이 빠른 환경에서도 적합한 보안시스템으로 사용될 수 있으며, 영상 합성, 애니메이션 등 여러 분야에 저비용 시스템을 제공함으로써 폭넓게 응용될 수 있다.

## 감사의 글

본 연구는 교육부의 BK21 사업과, 과학 기술부의 국가 지정연구실 사업 지원으로 이루어졌습니다.

## 참고문헌

- [1] S. Kang, A. Koschan, H. Zhang, J. Paik, and M. Abidi, "Hierarchical approach to enhanced active shape model for color video tracking," *Proc. Int. Conf. Image Processing*, September 2002.
- [2] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models - their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38-59, 1995.
- [3] T. F. Cootes, C. J. Taylor, and A. Lanitis, "Active shape models: evaluation of a multi-resolution shape method for improving image search," *Proc. British Machine Vision Conference*, pp. 327-336, 1994.