DBpia

객체추적을 위한 웨이블릿 기반 계층적 능동형태 모델

준회원 기 현 종*, 신 정 호*, 이 성 원*, 종신회원 백 준 기*

Wavelet transform-based hierarchical active shape model for object tracking

Hyunjong Kim*, Jeongho Shin*, Seong-won Lee*, Joonki Paik* ***Regular Members***

요 약

여기는 본 논문에서는 움직이는 물체 추적을 위한 윤곽선 및 형태 파라미터 추출을 위해 웨이블릿 변환을 이용한 능동형태모델의 계층적인 접근방법에 대해 제안한다. 능동형태 모델의 여러 단계 중 지역구조 모델링은 비정형 객체의 형태를 추출하기 위해 가장 중요한 비중을 차지한다. 제안한 알고리듬은 웨이블릿을 이용하여 계층적인 접근은 물론 지역구조 모델링단계를 웨이블릿 대역 분할을 이용하여 복잡한 환경에서의 객체를 강건하게 추적할 수 있도록 하였다. 또한 비정형객체를 실시간 비디오 추적에 이용하기 위해 웨이블릿을 이용한 계층적 움직임 추정방법을 적용하여 객체의 움직임을 예측, 보정하는 효과적인 방법을 제시하였다. 제안하는 알고리듬은 객체 추적에 대한 성능을 평가하기 위해 다양한 실험영상을 통해 우수함을 확인하였다.

Key Words : hierarchical, active shape model; video tracking; wavelet transform; non-rigid object

ABSTRACT

This paper proposes a hierarchical approach to shape model ASM using wavelet transform. Local structure model fitting in the ASM plays an important role in model-based pose and shape analysis. The proposed algorithm can robustly find good solutions in complex images by using wavelet decomposition. we also proposed effective method that estimates and corrects object's movement by using Wavelet transform-based hierarchical motion estimation scheme for ASM-based, real-time video tracking. The proposed algorithm has been tested for various sequences containing human motion to demonstrate the improved performance of the proposed object tracking.

Ⅰ. 서 론

비디오 영상내에서 객체분할이나 객체추적은 감시 시스템(surveillance system)과 같은 비디오 처리 응용분야뿐만아니라 MPEG-4, MPEG-7등과 같은 비디오 압축 표준에서 기본적이고 필수적인 기술이다. 또한 이들 응용분야 중에는 객체의 대략적인 움직임이나 위치의 검출뿐만 아니라 보다 정확한 객체의 모양과 위치의 추출을 요구하는 분야가 증가하고 있으며, 이러한 요구를 충족시키기 위해 여러 가지 방법들이 제안되었다. 객체추적 알고리듬은 일반적으로 움직임 기반의 접근방법과 모델 기반의 접근방법 그리고 경계선 기반의 접근 방법과 영역 기반의 접근방법으로 크게 나눌 수 있다[1,2]. 객체의 정확한 위치 및 모양을 추출하여 추적하는 방법 중 대표적인 방법으로는 Kass등이 제안한 능동윤곽선 모델(active contour model; ACM)방법이 있지만, 배경의 변화, 객체간의 겹침, 배경과의 가림현

* 중앙대학교 첨단영상대학원 영상공학과 시각및지능시스템 연구실(kikingka@wm.cau.ac.kr)
논문번호 : 2004-07-102, 접수일자 : 2004년 7월 19일
※본 연구는 과학기술부 국가지정연구실(NRL) 사업과 정보통신부 대학 IT연구센터육성(ITRC) 사업지원으로 수행되었습니다.

Copyright (C) 2004 NuriMedia Co., Ltd.
www.dbpia.co.kr

상, 객체 자체의 모양변화 등과 같은 다양한 상황에서 많은 문제점을 않고 있다[3]. 본 논문에서는 이와 같은 문제점을 극복하기 위해 모델기반의 접근방법을 사용한 능동형태모델(active shape model; ASM)을 사용한다. ASM은 객체 윤곽 정보를 기존의 학습모델들의 통계적인 특성을 바탕으로 해서 좀 더 효과적으로 추정하는 방법이다. 하지만, 매 프레임마다 객체의 윤곽을 모델과 정합시키기 위하여 반복적인 계산이 필요하고 연속되는 프레임상에서 객체 윤곽의 이동정보를 예측하기 어려운 단점을 가지고 있기 때문에 실시간 구현을 위한 비디오 시스템보다는 정확성을 요구하면서 처리시간에는 구애를 받지 않는 의료영상처리에 주로 사용되었다[4]. 최근 이를 해결하기 위해 계층적인 (hierachical approach)방법이 소개되었다[5]. 이는 다해상도 영상 피라미드(pyramid)를 기반으로 하여 낮은 해상도로부터 점차 높은 해상도로 객체의 찾는 방법으로 계산량이나 정확도에 있어서 효과적이다. 본 논문에서는 웨이브릿 변환을 기반으로 다해상도 영상을 생성하고, 계층적인 접근은 물론 웨이블릿의 각 부대역(subband) 특성을 효과적으로 이용하여 객체의 윤곽을 찾는 데 있어 좀 더 효과적이고 강건한 방법을 제안하려고 한다. 또한 객체의 움직임 추정을 위한 웨이블릿 기반 계층적 접근을 통해 매 프레임간 객체윤곽의 이동정보를 예측하고 오차를 보정하여 연속되는 프레임상에서 보다 강건하게 실시간 객체 추적을 가능하게 한다[6].

본 논문은 구성은 다음과 같다. II장에서는 능동형태 모델에 대한 기본 개념을 설명하고, 그것을 이용한 ASM에 대하여 살펴본다. III장에서는 본 논문에서 제안하는 웨이블릿 변환을 이용한 계층적인 방법을 설명하고, IV장에서는 웨이블릿 기반 계층적 움직임 추정 알고리듬을 적용한다. V장에서는 다양한 실험영상에 따른 여러 실험 결과을 비교 정리하고, 마지막으로 VI장에서 향후 보완점와 응용방안을 제안하고 결론을 짓는다.

II. 능동형태 모델(ASM)

ASM은 추적하고자하는 객체의 위치나 윤곽을 찾아내고, 나아가서 형태를 분석하는 형태 모델을 의미한다. 이 접근 방법은 (i) 형태나 윤곽선의 통계적인 모델링과 (ii) 새로운 영상을 분석하기 위해 모델 해석의 두 가지로 요약된다. ASM 기반 접근방법은 특히 로봇시각이나, 의료영상, 그리고 비디오 추적등 여러 실용적인 적용분야에 대한 안정적이고 효과적인 해결책을 제시한다. 좀 더 구체적으로 본 논문의 ASM 기반 추적 알고리듬을 다음과 같이 크게 네 가지 단계로 분류하였다: (i) 기준점(landmark point)의 할당, (ii) 주성분분석(principal component analysis; PCA), (iii) 지역구조(local structure) 모델링, (iv) 모델정합(model fitting).

2.1 기준점(Landmark point) 할당

모델의 윤곽 형태를 정의하기 위해서는 기준점이 우선 할당되어야 한다. 만약 d차원에서 모델의 윤곽 형태를 n개의 기준점으로 표현하기 위해서는 nd개의 요소로 이루어진 벡터를 사용하여 주어진 입력 영상으로부터 객체의 윤곽선을 할당해야 한다. 2차원 영상의 경우에 n개의 기준점을 할당하고자 한다면 기준점의 좌표집합은 $2n$차원의 벡터에 의해서 다음과 같이 표현할 수 있다.

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

기준점의 수는 적용되는 응용분야에 따라 달라질 수 있다. 예를 들면, 사람이나 차량등을 추적하고자 한다면 추적 객체의 위치나 크기등에만 관심이 있기 때문에 적은 수의 기준점으로도 충분히 빠르고 효율적인 추적을 할 수 있다. 반면에, 만약 얼굴인식, 표정인식등을 정확하게 해석하고자 한다면, 보다 많은 수의 기준점들이 필요할 것이다.

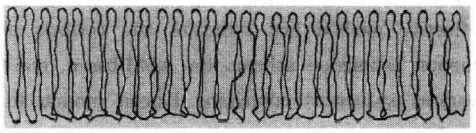


그림 1. 기준점이 할당된 훈련집합

2.2 주성분분석 (principal component analysis)

PCA는 입력 공간내의 변동량이 많은 방향으로 주축(principle axis)을 설정하여 데이터의 중복성을 제거하는 기법이다. ASM의 경우 훈련집합내의 기준점의 정보를 모두 개별적으로 처리하고자 하면 엄청난 계산량이 필요하다. PCA는 훈련집합들의 선형조합으로 이루어진 새로운 변수들을 생성함으로써, 전체 집합내의 변수들의 수를 줄이는 데 그 목적이 있다. 이 과정을 기하학적으로 분석하면 원래 좌표축을 새로운 직교좌표축으로 변환을 시키는 것이고, 한 개의 전체 집합를 대표할 수 있는 최적 라

Copyright (C) 2004 NuriMedia Co., Ltd.
www.dbpia.co.kr

인을 찾는 것이 이 알고리듬의 핵심이다. 전체 데이터를 대표하는 최적라인은 샘플 포인트로부터 라인까지의 수직 거리의 제곱이 최소가 되게 하는 제한조건 하에서 구할 수 있으며, 이렇게 구해진 적은 수의 매개변수를 가지고 우리는 객체의 형태를 효과적으로 모델링할 수 있다.

구체적인 PCA 알고리듬을 살펴보면 다음과 같다.

알고리듬 1: PCA 알고리듬

1. 훈련집합에서 m개의 표본 형태의 평균을 구한다.

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

여기서 x_i는 훈련집합내의 I번째 표본 형태, 혹은 해당 기준점 좌표집합을 의미한다. (i=1,....,m).

2. 훈련집합으로부터 공분산 행렬(covariance matrix) C를 계산한다.

$$C = \frac{1}{m}\sum_{i=1}^{m} (x_i - \bar{x})(x_i - \bar{x})^T. \tag{3}$$

3. 공분산 행렬로부터 고유벡터 ϕ_i,와 고유값 λ_i 을 구한다. 크기가 작은 순서에 따라 구한 고유벡터와 고유값을 정렬한다.

$$\Phi = [\phi_1 | \phi_2 | \cdots | \phi_q], \tag{4}$$

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

4. 주어진 Φ와 $\bar{x}$로부터 각각의 형태는 다음과 같이 표현할 수 있다.

$$x_i \simeq \bar{x} + \Phi b_i \tag{5}$$

여기서

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

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

2.3 지역구조 모델링(Local structure modeling)

실제 입력영상에서 추적하고자하는 객체의 형태를 해석하기 위해서는 초기 기준점들로부터 새로운 기준점들의 집합을 찾아내야 한다. 어떤 기준점을 갱신하고자 하면 그점에 이웃하는 두 기준점들을 연결한 직선의 수직방향을 따라 최적의 윤곽 위치를 탐색한다. 모델의 윤곽을 밝기값의 변화, 즉 에지정보로부터 찾으려고 한다면, 간단히 각 기준점들의 탐색영역내에 가장 강한 에지를 찾으면 되고 그 위치가 새로운 기준점으로 갱신될 것이다. 그러나 실제 객체의 윤곽은 배경에 따라 항상 에지가 가장 강한 부분에 존재하지 않을 수도 있다. 이런 문제를 해결하기 위해서는 훈련집합의 윤곽정보뿐만 아니라 밝기값 정보들을 이용하여 객체의 윤곽을 찾아야 한다. 따라서 본 논문에서 훈련집합의 밝기값 정보를 이용한 Mahalanobis알고리듬을 사용하여 객체의 윤곽을 배경에 강건하고 효과적으로 추정하고자한다. 구체적인 알고리듬은 다음과 같다.

알고리듬 2 : Mahalanobis 알고리듬을 이용한 지역구조 모델링

1. 각각의 기준점에 대하여 양변에 이웃하는 기준점들을 지나는 기울기의 수직방향으로 밝기값의 분포도(기준점을 중심으로 한방향에 대해서 k개의 픽셀)를 정의한다. 밝기값의 분포도로부터 밝기값의 미분분포도를 구하고 분포도내의 모든 값들의 전체합으로 그것을 정규화시킴으로써 미분분포도를 구할 수 있다. 그 결과 각각의 기준점 g_j에 대하여 $2k+1$개의 미분 분포도를 얻는다.
2. 훈련집합에 있는 모든 영상에 대한 각 기준점 g_j의 모든 밝기값 미분분포도를 이용하여 공분산 S_j를 계산한다.
3. 훈련집합에 있는 모든 영상에 대한 각 기준점 g_j들을 평균함으로써 평균 미분 분포도 $\bar{g}_j$을 구한다.
4. 아래와 같은 Mahalanobis거리를 비용함수로 두어 각 기준점에 대하여 적용하고 그것을 최소로 하는 지점을 새로운 기준점으로 할당한다.

$$f(g_{j,m}) = (g_{j,m} - \bar{g}_j)^T S_j^{-1} (g_{j,m} - \bar{g}_j) \tag{7}$$ ■

Copyright (C) 2004 NuriMedia Co., Ltd.
www.dbpia.co.kr

2.4 컬러성분을 이용한 확장된 지역구조 모델링

앞서 2.3장에서 우리는 흑백영상의 밝기값만을 사용하여 Mahalanobis거리를 최소화함으로써 각각의 기준점을 갱신하였다. 하지만, 복잡한 환경에서 강건하게 객체의 윤곽을 추정하기 위해서는 여러 칼라정보를 사용하는 것이 요구된다. 따라서 본 논문에서는 여러 칼라성분을 사용하여 기존의 흑백정보를 이용한 지역구조 모델링을 보다 강건하게 확장하려고 한다. 컬러영상을 RGB, YUV, HSI등 여러가지 채널로 나누어 각 채널에 대해 독립적으로 Mahalanobis거리를 최소화함으로써 지역구조 모델링을 컬러공간으로 확장할 수 있다. 컬러영상은 RGB, YUV, HSI등 여러가지 컬러성분으로 구성되어 있지만 일반적으로 YUV, HSI등에 비교하여 RGB컬러성분을 이용한 지역구조 모델링 결과가 최적의 해를 제시한다[7].

본 논문에서는 보다 강건한 결과를 위해 각각의 컬러채널을 독립적으로 사용하는 것이 아니라 컬러채널간의 상관성을 이용한 세 가지 조합방법을 이용하였다.

첫 번째로는 각 RGB성분에 대해 Mahalanobis거리를 가장 최소화하는 지점으로 아래 식과 같이 갱신하는 방법이다.

$$\min\{\arg\min_m f_R(g_{j,m}), \arg\min_k f_G(g_{j,k}), \arg\min_l f_B(g_{j,l})\}, \quad (8)$$

f_R, f_G, f_B는 식 (7)에 의해 정의된다.

그러나 이 방법으로 RGB성분중의 한 성분이 잘못된 위치에서 해를 가질 경우, 그 해가 최종적인 해가 되기 때문에 에러가 발생할 수 있다.

두 번째로 각 RGB성분에 대해 Mahalanobis거리를 최소화하는 지점들의 평균값으로 아래 식과 같이 갱신하는 방법이다.

$$\frac{1}{3}(\arg\min_m f_R(g_{j,m}) + \arg\min_k f_G(g_{j,k}) + \arg\min_l f_B(g_{j,l})), \quad (9)$$

이 방법의 경우 평균값은 비용함수의 최적의 해에 해당되는 값이 아니기 때문에 RGB성분중의 한 성분에 따라 에러를 누적할 수 있는 경우가 있다.

이런 문제점을 극복하기 위해 마지막 방법으로 각 RGB성분에 대해 Mahalanobis거리를 최소화하는 지점들의 중간값으로 아래 식과 같이 갱신하는 방법이다. 이 방법으로는 한 성분이 잘못된 위치에서 해를 가지는 경우에도 에러를 최소화할 수 있다.

$$median(\arg\min_m f_R(g_{j,m}) + \arg\min_k f_G(g_{j,k}) + \arg\min_l f_B(g_{j,l})). \quad (10)$$

2.5 모델 정합(model fitting)

훈련 집합의 한 샘플에서 유도된 모델의 형태를 실제 비디오 영상의 객체와 정합하기 위해서는 최적의 위치와 형태를 갖는 모델 파라미터를 구해야 한다. 이것은 모델 좌표프레임의 형태인 x와 영상 좌표프레임의 형태인 y를 정합하기위해서는 다음과 같은 에러 함수를 최소화하는 방식을 사용할 수 있다.

$$E = (y - Mx)^T W (y - Mx), \quad (11)$$

여기서 W는 대각행렬이며, 그 대각성분은 각각 기준점에 대한 가중치를 나타낸다. 그리고 행렬 M은 회전성분 θ, 전이성분 t, 스케일링 성분 s로 이루어진 기하학적인 변환을 나타내는 행렬이다. 가중치 요소는 객체의 윤곽을 따라 선택된 이전의 기준점과 갱신된 기준점 사이 값의 변위에 관련이 있다. 변위가 크면 클수록 행렬에 해당되는 가중요소들은 작아지고, 반면에 변위가 작아질수록 가중요소들은 커진다. 하나의 점을 $[x_0, y_0]^T$으로 표현한다면, 기하학적인 변환행렬은 다음과 같이 정의 될 수 있다.

$$M\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = s\begin{bmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{bmatrix}\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}. \quad (12)$$

기하학적인 파라미터의 집합(θ, t, s)이 결정되면, 모델좌표 프레임으로 투영된 y값은 다음과 같이 주어지고,

$$x_p = M^{-1} y, \quad (13)$$

최종적으로 모델의 파라미터는 다음과 같이 갱신된다.

$$b = \Phi^T (x_p - \bar{x}). \quad (14)$$

윤곽을 따라 위 단계들의 결과로써 최적의 기준점의 변위를 얻을 수 있다. 최적으로 갱신된 기준점의 조합은 영상좌표 프레임 y에 새로운 형태를 만들어 낸다. 새로운 형태는 식 (11)을 사용하여 가장 근접한 형태를 찾는 데 이용될 수 있다. M을 통해서 최적의 위치를 도출한 뒤, 이 새로운 형태는 주어진

Copyright (C) 2004 NuriMedia Co., Ltd.
www.dbpia.co.kr

훈련 집합에 주요 요소에 포함된 Φ로 투영될 수 있다. 이 과정은 모델 파라미터 b를 갱신한다. 그 결과로 주요 요소에 해당되는 비슷한 한 편차만이 모델 파라미터에 영향을 줄 수 있다. 모델 파라미터를 계산한 후에 x에 의해 주어지는 새로운 형태는 식 (5)에 의해 구할 수 있고, 이 새로운 형태는 식 (11)에 의해 다음 반복과정을 위해 사용된다. 적당한 반복을 거쳐, 최종적인 형태가 x로 얻어진다.

Ⅲ. 웨이블릿 변환을 이용한 계층적인 접근

이 장에서는 ASM 기반 추적에 사용하기 위한 웨이블릿 이론을 설명하고, 웨이블릿 변환을 사용하는 두가지 기본적인 근거를 제시한다. 이를 근거로 웨이블릿 기반 계층적 모델 정합 알고리듬을 제안하고자 한다.

3.1 웨이브릿 이론

최근 여러 분야에서 응용되고 있는 웨이블릿 변환 방법은 공간 및 주파수 영역에서 분리되어 얻어진 부대역들을 제공함으로써 체계적이고 효율적인 영상해석 및 신호분할을 가능하게 한다. 변환에 사용되는 웨이블릿(wavelet)이라 불리는 기저 함수는 이동과 확장을 통해서 공간 및 주파수 영역에 따라 분리된 성분을 추출하는 기능을 갖는다. 웨이블릿 변환을 2차원 영상에 1회 적용하였을 때 4개의 성분으로 나누어 진다. 이 4개의 성분들은 시각에 민감하고 정보의 대부분을 가지고 있는 저주파 성분과 수직성분, 수평성분, 대각성분으로 구성된다.

2차원 영상의 웨이블릿 변환은 그림. 2의 대역분할 부분에서 나타난 것과 같이 필터 뱅크 구조를 구성하여 수평 방향과 수직 방향으로 1차원 이산 웨이블릿 변환을 반복적으로 수행해서 얻어진다.

그림. 2에서 표시된 LL, LH, HL, HH은 분해된 대역을 나타낸다. 첫 번째 영문자는 영상을 수평 방향으로 필터링하는 방법을, 두 번째 영문자는 수직 방향으로 필터링하는 방법을 의미한다. 이 때, L과 H는 각각 저주파 대역 통과 필터(h(n))와 고주파 대역 통과 필터(g(n))를 이용하여 필터링함을 의미한다. 분해 과정에서는 먼저 입력 영상을 수평 방향에 대해서 각각 저주파 대역 통과 필터링과 고주파 대역 통과 필터링한 후 저주파 대역과 고주파 대역으로 분해 한다. 그리고, 분해된 각각의 저주파 대역과 고주파 대역에 대하여 다시 수직 방향으로 저주파 대역 통과 필터링과 고주파 대역 통과 필터링한 후에 분해 대역 LL, LH, HL, HH을 얻는다.

본 논문에서는 ASM 단계중 객체의 윤곽을 추정하는 지역구조모델링에서 계층적인 접근과 효과적인 에지맵을 위해 웨이블릿을 사용하고자 한다.

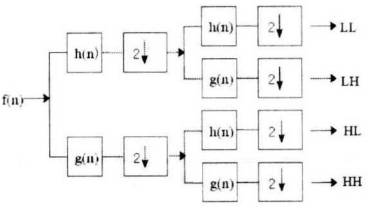


그림 2. 2차원 신호상의 웨이브릿 변환을 위한 대역 분할 과정

3.2 대역간 에지 상관성을 이용한 웨이브릿 기반 계층적인 모델정합

능동형태 모델을 이용한 비디오 추적 시스템은 입력영상에서 추적하고자하는 객체의 형태나 크기가 변화하는 경우에도 강건하게 객체의 위치와 윤곽정보를 해석해야 한다. 그러나 초기윤곽이 실제 입력영상과 상이할 경우 만족스럽지 못한 결과를 보인다. 훈련집합으로부터 최적의 모델정합을 얻기 위해서 우리는 계층적인 접근, 효과적인 에지맵이라는 두가지 개념을 제안하였다. 첫번째로 입력영상의 계층적인 접근은 지역구조 모델링에 있어 기준점 갱신을 위한 탐색영역 길이를 줄일 수 있도록 한다. 탐색영역이 길면 초기윤곽으로부터 멀리 떨어진 입력영상에서 객체윤곽의 추정이 가능한 반면, 계산량은 증가하고, 오히려 주위 배경에 영향으로 에러를 가중시킬 수 있다. 그러므로 입력영상의 계층적인 접근은 작은 탐색영역길이를 가지고도 객체윤곽을 추정하는 데 있어 성능과 효율성면에서 모두 우수하기 때문에 실제 비디오 추적 시스템에 적용할 수 있다[5]. 구체적으로 살펴보면, 우선 작은 탐색영역길이로 저해상도 레벨에서 어느 정도 윤곽을 추정한 후, 동일한 탐색영역길이를 가지고 점점 고해상도 레벨로 이동하면서 객체윤곽의 정확도를 향상시키는 것이다.

두 번째로 기존의 계층적인 방법의 경우 객체의 윤곽을 추정하기 위해서 각 기준점의 에지 방향에 관계없이 훈련집합으로부터 미분 밝기값의 분포를 고려하여 최적의 위치를 갱신하였다. 그러나 각 기준점에서 탐색구간의 방향성을 나타내는 에지맵에서 각 기준점을 갱신한다면 보다 효과적인 결과를 얻

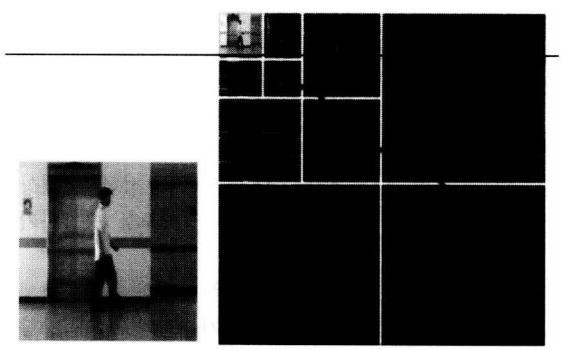
Copyright (C) 2004 NuriMedia Co., Ltd.
www.dbpia.co.kr

을 수 있다.

이 두 가지 개념을 모두 수용하기 위한 도구로서 웨이블릿 변환을 이용하여 계층적인 접근과 효과적인 에지맵을 구현하였다. 웨이블릿 변환은 입력영상을 계층적으로 분할하여 방향에 따라 수평, 수직, 대각방향의 고주파성분을 얻을 수있다. 결국 우리는 계층적인 접근은 물론 방향에 따른 고주파성분들을 기준점의 탐색구간 방향에 따라 효과적인 에지맵으로 적용할 수 있다.

알고리듬3 : 제안하는 웨이블릿 기반 계층적인 모델정합

1. 가장 낮은 해상도로 $n = n_{\max}$과 같이 초기 레벨을 정의한다.
2. 각 기준점에 대해서 양변에 이웃하는 기준점들을 지나는 기울기의 수직방향으로 탐색구간방향을 구한다.
3. n레벨에서 탐색구간방향에 따라 웨이블릿의 각 서브밴드를 사용하여 새로운 기준점을 계산한다. 각 기준점에 대해서 탐색구간방향이 수평, 수직, 대각방향일 경우, 그림 3같이 각각 HL_n,LH_n, HH_n의 웨이블릿 서브밴드영역에서 구한다.
4. 새로운 기준점을 모델에 정합하기 위해 모델의 위치나 형태파라미터를 갱신한다.
6. 만약 $L > 0$이면, $L \leftarrow L-1$ 로 레벨을 낮춘다.
5. 각 레벨에 따라 정해진 반복횟수만큼 단계1에서 단계6까지를 반복한다. ■



(a) (b)

그림 3. (a) 원영상, (b)제안하는 웨이브릿 서브밴드의 모델정합 단계

Ⅳ. 웨이블릿을 이용한 계층적 움직임 추정을 통한 객체 이동

ASM은 연속되는 프레임상에서 적용할 경우 각 프레임간의 움직임이 적기 때문에 어느 정도까지는 객체윤곽을 추적하지만, 점점 기준점들의 에러가 누적되면서 밝기값 변화가 큰 배경들에 고정되어 추적에 실패하는 경우가 발생한다. 실시간 비디오시스템에서 ASM을 적용하기 위해서 객체 윤곽의 이동정보를 효과적으로 예측하기 위한 방법이 추가로 제시되어야 한다. 따라서 본 논문에서 객체의 움직임을 추정하여 매 프레임마다 객체의 움직임 방향과 크기를 구한 다음, 현재 프레임의 ASM의 기준점 집합들을 움직임에 따라 다음 프레임으로 이동시켜 연속되는 프레임에 적용하고자 한다. 움직임을 추정하기 위해 앞서 사용한 웨이블릿을 통해 계층적인 접근으로 보다 정확하고 계산량에 효율적인 방법을 사용한다.

4.1 기존의 움직임 추정방법

비디오 프레임에서 움직임 추정기법에는 예측오차를 최소가 되도록 반복하는 PRA(pel recursive algorithm)와 탐색 구간내 블록의 왜곡이 최소인 위치를 찾아 움직임 벡터를 결정하는 BMA(block matching algorithm)가 있다. 비디오 추적 시스템에서는 PRA보다는 예측 오차 성능은 떨어지지만 하드웨어 구성이 용이한 BMA방법이 주로 사용되고 있다. BMA방법 중 전역 탐색방식은 추정 블록당 소요되는 계산량이 지나치게 많으며, 대상 영상의 움직임에 대해 적절한 탐색 영역과 블록 크기를 결정하기 어렵다. 전역탐색방법의 문제점중 하나인 계산량 문제를 해결하기 위해 three step search, locarithmic search등의 방법들이 제안되었으나, 전역 탐색에 비해 추정오차가 크고 추정벡터의 신뢰성이 현저히 떨어진다. 탐색영역과 블록크기 결정의 어려움을 극복하기 위해 계층적 블록정합(hirearchical block matching)기법이 제안되었지만, 마찬가지로 계층간의 움직임 추정에 따라 국부 최소를 나타내는 변위값을 가지게 된다. 이러한 문제점을 극복하기 위해 본 논문에서는 웨이블릿 변환을 이용하여 각 부대역에 있는 블록간의 에너지값을 비교하여 국부 최소로 빠질 가능성이 있는 블록에 대해 해상도를 증가시켜 보다 움직임 추정을 하려고 한다.

Copyright (C) 2004 NuriMedia Co., Ltd.
www.dbpia.co.kr

4.2 웨이블릿 기반 계층적 움직임 추정 알고리듬

본 논문에서는 기존의 움직임 기법 중 고속 알고리듬의 움직임 추정에서 흔히 발생할 수 있는 국부 최소확률을 줄이고, 이동벡터값의 정확도를 높이기 위해 웨이블릿을 이용한다. 일반적으로 계층적인 접근을 위해 여러번의 저역 통과 필터링을 하면서 영상이 몽롱화(blur)되고, 탐색범위를 벗어나는 큰 움직임을 갖는 블록일 경우 국부 최소로 빠질 확률이 비교적 크다. 이러한 국부 최소의 문제를 해결하기 위해 본 논문에서는 웨이블릿 변환된 각 블록의 부대역 신호 특성에 따라 적응적으로 움직임 추정에 참여하는 계층 수를 결정하여 보다 정확한 이동벡터를 추정하도록 하였다. 먼저 계층1의 저주파 신호에 대해서 블록 단위로 메모리 내에 저장된 이전 프레임의 계층 신호와 현재 프레임의 계층 신호간의 차를 구하고, 각 계층간 저주파 대역 에너지 값들을 비교함으로써 계층적 움직임 추정시 고주파 성분의 손실로 움직임이 잘못 찾아질 가능성이 있는 블록을 판단한다. 이 때, 차신호를 구하는 이유는 움직임이 없는 블록을 배제함으로써 계산량을 줄이는 데 있다. 식 (15)은 계층간의 에너지 관계를 나타낸다.

$$E_{k,L} = E_{k+1,L} + E_{k+1,H}, \tag{15}$$

여기서 E_k는 계층 k의 에너지이고, L은 저주파 대역을 H는 고주파 대역을 나타낸다. 어떤 계층의 저주파 대역 에너지와 고주파 대역 에너지의 비교는 그 계층과 그보다 해상도가 한 단계 낮은 계층의 저주파 대역 에너지를 비교함으로써 이루어진다. 구체적으로 보면, 계층1과 계층2의 저주파 대역 에너지의 비, $(4\times E_{2,L})/E_{1,L}$을 구한다. 그 값이 임계값보다 작으면 계층1의 저주파 대역의 블록은 고주파 성분을 많이 잃은 블록으로 움직임 추정에 적합하지 않다고 판단하여 해상도가 높은 계층1에서 움직임 추정을 한다. 그렇지 않은 블록은 계층을 한 단계 높여서 계층2와 계층3의 저주파 대역 에너지의 비, $(4\times E_{3,L})/E_{2,L}$을 구한다. 그 값이 임계값보다 작으면 계층2와 계층1을 이용하여 계층적으로 움직임 추정을 하고, 그렇지 않은 블록은 계층을 한 단계 올라간다. 이러한 판단은 계층 k와 계층 $(k-1)$의 저주파 대역 에너지 비교까지 계속된다.

각 경우 계층1에서 구해지는 최종 이동벡터 $MV_1(x,y)$는

$$MV_1(x,y) = 2^{n-1}\cdot MV_n(x,y) + \sum_{m-2}^{n} 2^{n-m}\cdot \triangle MV_{n-m+1}, \quad n \geq 2, \tag{16}$$

와 같이 나타낼 수 있다.

본 알고리듬의 흐름도는 그림4에 나타나있다.

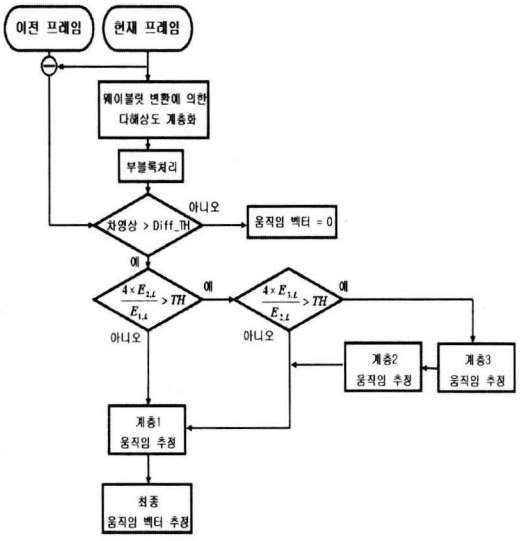


그림 4. 웨이블릿 기반 계층적 움직임 추정 흐름도

4.3 움직임 추정을 통한 객체 이동

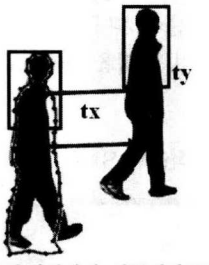


그림 5. 매 프레임간 추적영역내 이동벡터들의 평균에 따라 객체윤곽의 이동정보를 보정하는 과정

사람은 대표적인 비정형객체이기 때문에 팔이나 다리부분은 프레임간 움직임이 유동적이기 때문에 객체전체에 대한 이동정보를 예측하는 데 어려움이 따른다. 그러므로 비교적 움직임이 가장 작은 머리부터 가슴부분까지의 움직임 영역을 추적영역으로 설정하여 매 프레임간 추적영역내의 추정된 이동벡터들의 평균을 객체의 이동정보로 계산한다. 그 결과, 매 프레임당 객체의 이동방향과 크기만큼 기

Copyright (C) 2004 NuriMedia Co., Ltd.
www.dbpia.co.kr

(a) (b) (c) (d) (e) (f)

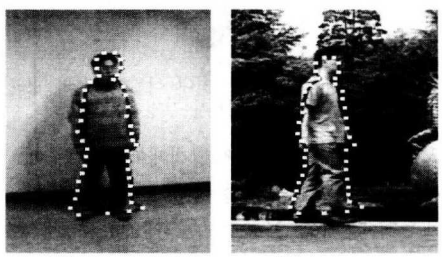
그림 6. 실험에 사용된 여러 영상: (a) 실내_1, b)실내_2, (c)실외_1, (d) 실외_2, (e) 실외_3, (f) 실외_4

준점들의 집합을 이동시키면 연속적인 프레임내에서 보다 안정적인 객체윤곽을 추정할 수 있다. 그러나, 이 경우 이동정보 예측에 의한 보정에도 불구하고 프레임간 객체 윤곽의 에러가 누적되어 일정한 개수이상의 프레임에 도달하면 추적에 완전히 실패하는 경우가 발생한다. 따라서 주기적으로 추적영역을 중심으로 다시 초기윤곽을 잡고, 상기의 수행과정을 반복해야한다.

V. 실험 결과

여러 실험영상들은 SONY 3-CCD DCR-TRV900 비디오 카메라를 사용하여 320X240의 크기로 취득되었다. 5.1절에서는 흑백영상에서 계층적인 방법, 비계층적인 방법, 웨이블릿에 기반한 계층적인 방법을 사용한 실험결과를 비교하고 5.2절에서는 여러 가지 칼라 공간에서 실험된 ASM 결과를 보인다. 5.3절에서는 기준점의 개수와 탐색구간의 길이의 변화에 따른 여러 결과를 분석하고, 마지막으로 5.4절에서는 실시간 비디오추적을 위해 웨이블릿 기반 계층적 움직임 추정을 통한 ASM 추적결과를 보여준다.

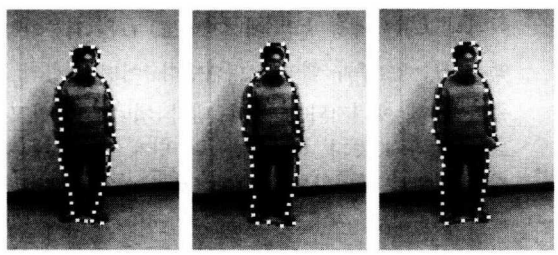
5.1 계층적인 ASM



(a) (b)

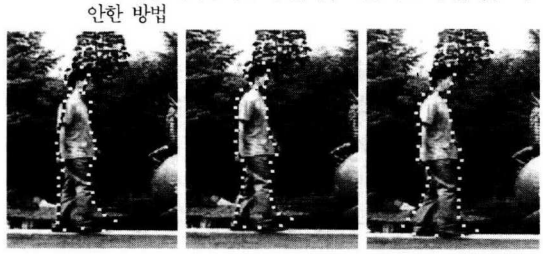
그림 7. 초기 기준점들을 할당한 실험영상: (a) 실내_1, (b) 실외_1

그림 6에서는 실험에 사용된 여러 테스트 영상을 나열하였다. 본 실험에서는 PCA를 위한 훈련집합으로 50개의 형태가 사용되었고, 각각의 기준점을 갱신하는 데 사용된 탐색영역길이는 7로 동일하게 설정하였다. 그림 7은 모든 훈련집합의 평균 기준점들을 초기 기준점들로 할당한 영상을 보여주고 있다. 지역구조 모델링을 수행하는 데 있어 탐색영역길이는 Mahalanobis 거리를 최소화하기 위한 탐색구간을 결정한다. 탐색영역길이가 길고, 탐색구간이 넓을수록 계산량의 증가나 복잡도는 증가하지만 성능향상은 반드시 탐색영역의 길이에 정비례하지는 않는다. 객체의 크기에 따른 적당한 탐색영역길이와 각 기준점에 방향에 따른 효과적인 밝기값 변화 맵를 사용하는 것이 효과적으로 모델 정합을 가능하게 한다. 제안한 방법의 성능을 기존 방식들과 비교하기 위하여 본 논문에서는 계층적인방법, 비계층적인 방법, 웨이브릿 기반 계층적인 방법을 사용한 결과를 각각 그림 8, 9에 나타내었다.



(a) (b) (c)

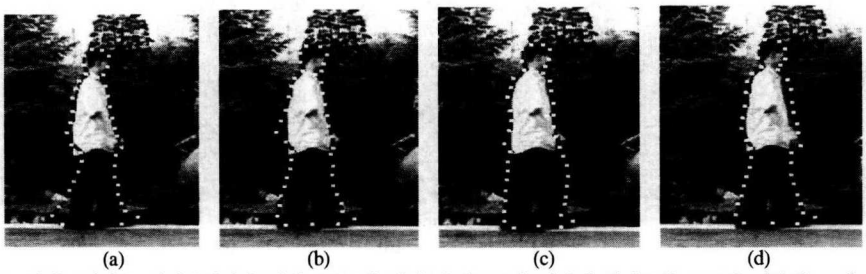
그림 8. 실내영상에 대한 3가지 다른 방법에 따른 모델 정합 결과: (a) 비계층적인 방법, (b) 계층적인 방법, (c) 제안한 방법



(a) (b) (c)

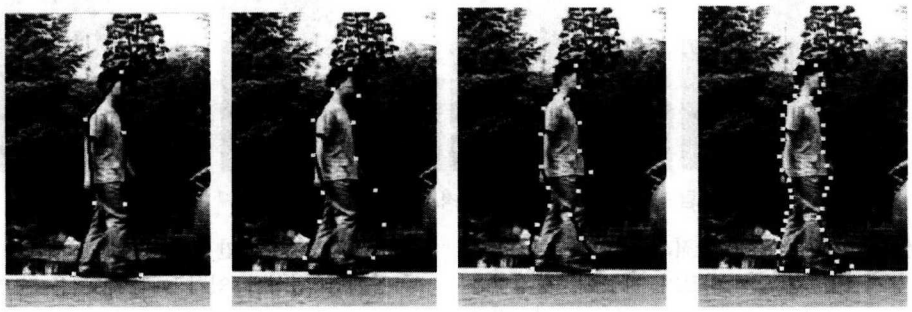
그림 9 실외영상에 대한 3가지 다른 방법에 따른 모델 정합 결과: (a) 비계층적인 방법, (b)계층적인 방법, (c) 제안한 방법

Copyright (C) 2004 NuriMedia Co., Ltd.
www.dbpia.co.kr



(a) (b) (c) (d)

그림 10. 실외_4영상을 웨이블릿기반 칼라 ASM에 따라 흑백모드과 세가지 칼라조합 모드를 사용한 모델정합 결과: (a) 기존의 흑백영상을 이용한 모드, (b) 최소값모드, (b) 중간값모드, (c) 평균값모드



(a) (b) (c) (d)

그림 11. 기준점의 개수에 따른 실외_1 영상의 모델정합 결과: (a) 7, (b) 14, (c) 21 ,(d) 42

기존의 계층적인 방법과 제안한 웨이블릿 기반의 계층적인 방법에서, 0레벨, 1레벨, 2레벨은 차례대로 원영상의 해상도, 1/2영상의 해상도, 1/4영상의 해상도를 의미한다. 본 논문에서는 2레벨, 1레벨, 0레벨에서 각각 반복횟수를 5,5,1로 수행하였다. 제안하는 방법에서는 Daubechies(7,9) 웨이블릿 필터를 사용하였다. 비계층적인 방법의 경우는 10번의 반복을 수행하였다.

본 실험에서는 각 실험결과를 정량적으로 분석하기 위해 수동적으로 할당한 최적의 기준점들과 각 방법에 따른 결과 기준점들의 차이들의 평균제곱오차를 사용하여 에러값을 측정하였다. 표1은 제안한 방법이 기존의 방법들과 비교하여 모델정합에 있어 에러 수치상으로도 보다 향상된 결과를 보여준다. 제안한 방법을 동일한 탐색영역길이를 가지고 실험했을 때 보다 향상된 모델정합 결과를 보여주고 있다. 이런 흑백영상의 결과를 토대로 우리는 칼라영상에 대해 보다 효과적인 모델정합을 위해 다음 절에서는 제안한 방법을 여러 다른 칼라 공간에 따라 확장하려고 한다.

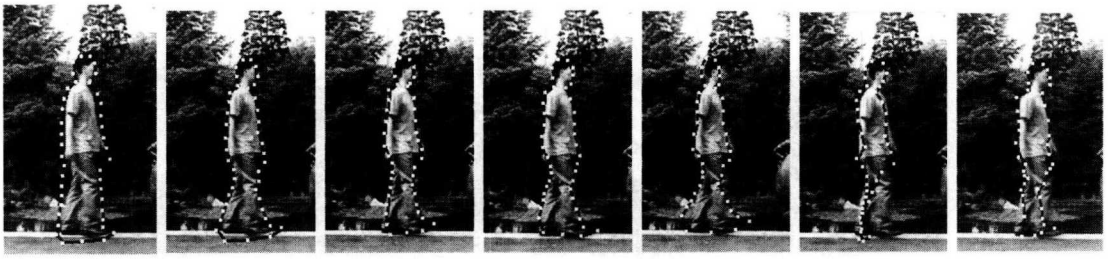
표 1. 여러 실내, 실외영상에 따른 기존과 제안하는 방법을 사용한 결과의 평균제곱오차

테스트 영상	비계층적인 방법	계층적인 방법	제안한 방법
실내_1	319.54	274.25	245.45
실내_2	227.35	193.12	185.24
실외_1	323.14	294.84	229.41
실외_2	294.48	254.03	255.12
실외_3	348.54	314.14	291.25

5.2 확장된 칼라기반 ASM

복잡한 환경에서 객체의 윤곽을 보다 강건하고 정확하게 추정하기위해서 본 논문에서 흑백성분이외에 RGB성분의 특성을 따른 여러 분석을 통하여 모델정합을 수행하였다. 실험에 사용된 방법은 기존 흑백성분만을 사용한 방법을 포함하여 절대값, 최소모드, 중간값 모드, 평균값 모드 등 네 가지 방법을 사용하여 모델정합의 성능을 표2에 비교하였다. 표2에 분석된 네 가지 다른 모드의 평균제곱오차를 볼 때, 중간값 모드를 사용한 방법이 여러 다른 방법과 비교해 보다 나은 성능을 가져온다. 평균값모드 또한 비교적 만족할만한 성능을 가져오는 것을 확인

Copyright (C) 2004 NuriMedia Co., Ltd.
www.dbpia.co.kr



(a) (b) (c) (d) (e) (f) (g)

그림 12. 여러 탐색영역길이에 따른 실외_1 영상의 모델정합 결과: (a) 3, (b) 5, (c) 7, (d) 9, (e) 11, (f) 13, (g)15

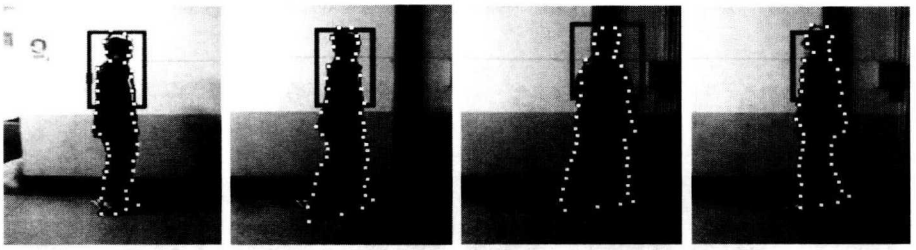


그림 13. 웨이블릿 기반 계층적 움직임 추정을 통한 객체 추적

할 수 있다. RGB성분 일반적으로 비슷한 동적인 범위와 통계적인 분포를 따르기 때문에, 중간값을 기반으로 한 방법이 효과적으로 에러를 최소화할 수 있다. 그림 10은 기존 흑백영상만 사용한 방법과 최소값, 중간값, 평균값모드를 적용한 칼라기반 ASM결과를 보여주고 있다.

표 2. 확장된 칼라 ASM을 기존의 흑백정보를 사용한 결과와 세가지 모드에 다른 결과에 따른 에러의 평균제곱오차

테스트 영상	기존 흑백모드	최소값 모드	중간값 모드	평균값 모드
실내_1	245.45	225.14	214.84	294.24
실외_1	229.41	245.14	201.49	254.11
실외_2	255.12	246.41	221.14	231.11
실외_4	214.14	204.14	189.14	200.14

5.3 탐색영역길이와 기준점의 갯수에 따른 분석

앞서 언급했던 것처럼 탐색영역길이와 기준점의 개수는 모델정합에 있어 중요한 영향을 끼친다. 우선 기준점의 개수에 따른 분석을 위해 여러 기준점의 개수에 따라 제안한 방법을 실험해 보았다. 기준점의 개수가 많을수록 모델정합결과는 좀 더 정확하지만 계산량이나 수행시간이 증가하는 단점을 가지고 있다. 따라서 기준점의 개수는 응용분야에 따라 달라져야 한다. 예를 들면, 사람이나 차량등을 추적하고자한다면 우리는 추적 객체의 위치나 크기 등에만 관심이 있기 때문에 적은 수의 기준점으로도 충분히 빠르고 효율적인 추적을 할 수 있다. 반면에, 만약 우리가 얼굴인식, 표정 인식등을 수행하고자 한다면, 보다 많은 수의 기준점의 수가 사용될 필요가 있을 것이다. 그림 11은 각 기준점의 갯수에 따른 모델정합의 결과를 보여주고 있다. 사람 추적의 경우, 본 실험결과에 따르면 42개의 기준점 개수가 모든 면에서 최적의 모델정합결과를 가져오고 있다. 또한 탐색영역길이의 경우, 탐색영역길이가 너무 길거나 짧은 경우는 좋은 결과를 가져오지 못한다. 탐색영역길이기 너무 짧을 경우, 초기 윤곽에서 실제 객체의 윤곽을 포함하지 않을 수도 있고, 반면에 너무 긴 경우는 부적절한 윤곽을 찾을 확률이 커질 수 있다. 많은 실험을 통해 7~9 정도의 탐색영역길이가 최적의 모델정합결과를 가져오는 것을 알 수 있었다. 그림 12은 여러 탐색영역길이랑 따른 모델정합결과를 보여주고 있다.

5.4 실시간 비디오 추적을 위해 웨이블릿 기반 계층적 움직임 추정을 통한 객체의 이동정보 보정

그림 13은 웨이블릿 기반 계층적 움직임 추정하여 객체의 이동정보 보정에 의한 실시간 비디오 추적결과를 보여준다. 머리와 가슴부위를 추적영역으로 설정하고, 매 프레임간 추적영역의 움직임 정보를 이용하여 객체윤곽의 이동정보를 보정하였다. 움

Copyright (C) 2004 NuriMedia Co., Ltd.
www.dbpia.co.kr

그림 14. 가림현상이 존재하는 영상에서의 객체추적 결과

직임 추정을 하는 데 있어 웨이블릿 부대역의 저주파 대역 에너지를 비교하는 임계값 TH는 실험적으로 0.7를 사용하였다. 그림 14는 ASM의 장점인 가림현상 극복에 대한 실험결과이다. 사전 훈련집합으로부터 객체윤곽이 보정되기 때문에 어느정도 가림현상이 존재하더라도 모델정합에 의해 객체윤곽을 추정할 수 있다.

V. 결 론

본 논문에서는 비정형 물체를 효율적으로 추적하기 위해 웨이블릿을 사용하여 계층적인 ASM 비디오 추적 알고리듬을 제안하였다. 제안한 알고리듬은 ASM 추적 시스템을 계층적인 웨이블릿의 여러 부대역을 효율적으로 이용하여 보다 효과적인 모델정합으로 비디오 추적을 구현하였다. 또한 컬러정보를 이용하여 주위환경에 보다 강건하게 적용할 수 있도록 확장하였고, 프레임간 웨이블릿 기반 계층적 움직임 보정을 통해 계산량을 줄여 실시간 비디오 추적을 가능하게 했다. 그러나, 비디오 영상내 여러 사람을 동시에 추적할 경우 추적하는 객체들사이의 겹침현상에 대한 문제점과 컬러기반 ASM을 사용하였기 때문에 빛의 조건에 따른 알고리듬의 성능저하를 해결하기위한 과제를 안고 있다. 이러한 문제를 극복한다면, 제안한 알고리듬은 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.

참 고 문 헌

[1] L. Vincent, P. Soille, "Watersheds in Digital Spaces : An Efficient Algorithm Based on Immersion Simulations," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 1022-1034, Jun. 1991.

[2] D.K. Park, H. S. Yoon, C. S. Won, "Fast Object Tracking in Digital Video," *IEEE Trans. Consumer Electronics*, vol. 46, no. 3, pp. 785-790, 2000.

[3] M. Kass, A. Witkin, D. Terzopoulos, "Snakes: Active Contour Models," *International Journal of Computer Vision*, vol. 1, pp. 321-332, 1998.

[4] T. F. Cootes, A. Hill, C. J. Taylor C.J., and J. Haslam, "The use of active shape models for locating structures in medical images," *Information Processing in Medical imaging*, pp.33-47, 1993.

[5] 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.

Copyright (C) 2004 NuriMedia Co., Ltd.
www.dbpia.co.kr

[6] 장준필, 김진태, 장태규, 최종수, "Wavelet 변환을 이용한 계층적 움직임 추정", *제5회 신호처리 합동학술대회 논문집*, pp. 406-410, 1992년 9월.

[7] 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.

[8] 기현종, Vivek Maik, 정정훈, 신정호, 백준기, "웨이블릿을 이용한 능동형태모델 기반 비디오 추적기술," *대한전자공학회 추계학술대회*, pp. 161-164, 2003년 11월.

[9] R. Plankers, P. Fua, "Tracking and modeling people in video sequences," *Comp. Vision and Image Understanding*, vol. 81, pp. 285-302, 2001.

[10] T. J. Cootes, C. J. Taylor, D. H. Cooper, and J. Gragam, "Training models of shape form sets of examples," *In British Machine Vision Conference*, pp.9-18, September 1992.

[11] O. Rioul and M. Vetteri, "Wavelets and signal processing," *IEEE Signal Processing Mag.*, vol. 8, no. 4, pp. 14-38, Oct. 1991.

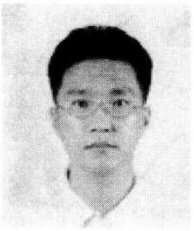
[12] S. Tanimoto and T. Pavlidis, "A hierarchical data structure for picture processing," *Comput. Graphics Image Process.* vol. 4, pp. 104-119, 1996.

[13] A. Baumberg, "Hierarchical shape fitting using an iterated linear filter," *Image and Vision Computing*, vol. 16, pp. 329-335, 1996.

[14] S. M. Smith, "Reviews of Optical Flow, Motion Segmentation, Edge Finding and Corner Finding," *Techinical Report, Dept. of Clinical Neurology*, Oxford University, 1997.

[15] A. Murat Tekalp, "Digital Video Processing," Prentice hall signal processing series, 1995.

기 현 종(Hyunjong Ki) 준회원

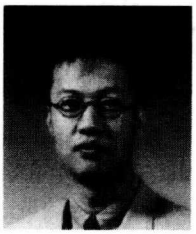


2003년 2월 : 중앙대학교 전기전자공학부 (공학사)

2003년 3월~현재 : 중앙대학교 영상공학과 (공학석사)

<관심분야> 영상복원, 웨이블릿, 객체추적

신 정 호(Jeongho Shin) 정회원



1994년 2월 : 중앙대학교 전자공학과 (공학사)

1998년 2월 : 중앙대학교 대학원 전자공학과 (공학석사)

2001년 2월 : 중앙대학교 첨단영상대학원 영상공학과(공학박사)

2003년 4월~현재 : 중앙대학교 첨단영상대학원 영상공학과 연구교수

<관심분야> 영상처리, 영상복원, 해상도개선, 영상융합, 컴퓨터 비전

이 성 원(Seongwon Lee) 정회원

1988년 2월 : 서울대학교 제어계측공학과 (공학사)

1990년 2월 : 서울대학교 대학원 제어계측공학과 (공학석사)

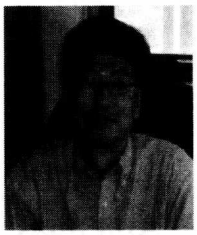
2003년 2월 : University of Sourthern California 전기공학과 (공학박사)

2004년 3월~현재 : 중앙대학교 첨단영상대학원 영상공학과 연구교수

<관심분야> 임베디드시스템 및 SOC설계, 멀티미디어 신호처리, Power-Aware Computing

Copyright (C) 2004 NuriMedia Co., Ltd.
www.dbpia.co.kr

백 준 기(Joonki Paik) 정회원



1984년 2월 : 서울대학교 제어계측공학과 (공학사)
1987년 2월 : 노스웨스턴대학교 전기 및 컴퓨터공학과(공학석사)
1990년 2월 : 노스웨스턴대학교 전기 및 컴퓨터공학과 (공학박사)
1993년 3월~현재 : 중앙대학교 첨단영상대학원 영상공학과 교수
<관심분야> 영상복원, 신호처리, 반도체

Copyright (C) 2004 NuriMedia Co., Ltd.
www.dbpia.co.kr