

논문 2004-41SP-5-3

실시간 감시 시스템을 위한 사전 무학습 능동 특징점 모델 기반 객체 추적

(Non-Prior Training Active Feature Model-Based Object Tracking for
Real-Time Surveillance Systems)

김 상 진*, 신 정 호**, 이 성 원**, 백 준 기**

(Sangjin Kim, Jeongho Shin, Seongwon Lee, and Joonki Paik)

요 약

본 논문에서는 사전학습이 필요 없는 능동 특징점 모델(non-prior training active feature model; NPT AFM) 기반에서 광류(optical flow)를 이용한 객체추적 기술을 제안한다. 제안한 알고리즘은 비정형 객체에 대한 분석^[1]에 초점을 두고 있으며, 실시간에서 NPT-AFM을 사용한 강건한 추적을 가능하게 한다. NPT-AFM 알고리즘은 관심 객체의 위치를 파악하는 과정(localization)과 이전 프레임 정보와 현재 프레임 정보를 이용하여, 객체의 위치를 예측(prediction), 보정(correction)하는 과정으로 나눌 수 있다. 위치 파악 과정에서는 움직임 분할(motion segmentation)을 수행한 후 개선된 Shi-Tomasi의 특징점 추적 알고리즘^[2]을 사용 하였다. 예측 및 보정 과정에서는 광류 정보를 사용하여 특징점을 추적하고^[3] 만약, 특징점이 적절히 추적되지 않거나 추적에 실패하면 특징점들의 시간(temporal), 공간(spatial)적 정보를 이용하여 예측, 보정하게 된다. 객체의 형태(shape)대신 특징점을 사용하였으며, 객체를 추적하는 과정에서 특징점들은 능동 특징점 모델(active feature model; AFM)을 위한 학습 집합(training sets)의 요소로 갱신된다. 실험결과, 제안한 NPT-AFM 기반 추적 알고리즘은 실시간에서 비정형 객체를 추적하는데 강건함을 보여준다.

Abstract

In this paper we propose a feature point tracking algorithm using optical flow under non-prior training active feature model (NPT-AFM). The proposed algorithm mainly focuses on analysis non-rigid objects^[1], and provides real-time, robust tracking by NPT-AFM. NPT-AFM algorithm can be divided into two steps: (i) localization of an object-of-interest and (ii) prediction and correction of the object position by utilizing the inter-frame information. The localization step was realized by using a modified Shi-Tomasi's feature tracking algorithm^[2] after motion-based segmentation. In the prediction-correction step, given feature points are continuously tracked by using optical flow method^[3] and if a feature point cannot be properly tracked, temporal and spatial prediction schemes can be employed for that point until it becomes uncovered again. Feature points inside an object are estimated instead of its shape boundary, and are updated an element of the training set for AFM. Experimental results, show that the proposed NPT-AFM-based algorithm can robustly track non-rigid objects in real-time.

Keywords: Active feature model, non-rigid object tracking, feature tracking, optical flow, motion-based segmentation

I. 서 론

연속된 비디오 영상에서 움직임과 형태 분석을 통한

2차원(two-dimensional; 2D) 객체 추적의 문제는 컴퓨터 시각, 비디오 감시, 애니메이션을 위한 움직임 분석, human computer interaction (HCI), 객체 기반 비디오 압축 등 많은 응용 분야에서 그 중요성이 증가하고 있다. 특히, 연속된 프레임에서의 비정형 객체에 대한 추출과 추적^{[4][5][6]}은 비디오 감시 시스템(video surveillance system)의 가장 중요한 기능이 된다.

* 학생회원, ** 정회원, 중앙대학교
(Chung-Ang University)

※ 이 연구는 과학기술부 국가지정연구실 사업과 교육부 BK21 사업 지원으로 수행되었음

접수일자: 2004년7월19일, 수정완료일: 2004년8월5일

논문저자 / 논문명 (Algorithm by)	중요한 특징 (Important Features)	학습 & 가정 (Training & Assumptions)	형태 모델 (Shape Model)	사용한 방법 (Method Used)	목적 (Specific Task)
W4 [7]	그레이 영상 (Grayscale Images)	배경 생성, 템플릿 생성 (Learn background, Learn projection templates)	사람의 실루엣 (Silhouettes of people)	형태 분석 및 추적 (Shape analysis and tracking)	사람 감시 (Surveillance of people)
Wren [8]	한 사람, 고정카메라 (Single person, Fixed camera)	고정 배경 생성 (Learn stationary background)	사람 신체 모델 (Human body model)	블럽 추적 (Blob tracking)	사람 추적 (Human tracking)
Comaniciu [9]	색상 도수분포도 (Color histogram)	초기 영역 표시 (Initial region marked)	형태 모델 사용하지 않음 (타원 영역 필터) (No shape model (elliptical region))	평균 이동을 이용한 도수분포 일치 (Histogram matching using mean-shift)	비정형 객체 (강역) (Track non-rigid objects (region))
Baumberg [10]	B-스플라인, 고른 움직임 (B-splines, uniform motion)	자동으로 공간 형태 생성 (Learn shape Space automatically)	고유 형태 (Eigen shapes)	칼만 필터 (Kalman filter)	사람의 움직임 추적 (Track human motion)
Cootes [11]	학습 집합 (Training sets)	사전에 저장된 형태 정보 이용 (Use prior shape information in database)	사람 모델 (Human model)	주성분 분석, 지역 구조 (PCA, Local structure)	사람 추적 (Track human)
Isard [12]	B-스플라인 사용 (Uses B-splines)	공간 형태 학습, 일정한 속도 (Learn shape space, constant velocity)	공간 형태 알고 있음 (Shape space known)	상태 공간 표본화 (State-space sampling)	형태 공간의 객체 추적 (Track objects of shape space)
Proposed algorithm	특징점, 능동 특징점 모델 (Feature pointer, AFM)	None	특징점 모델 (Feature point model)	광류, 무학습 능동 특징점 (Optical flow, NPT-AFM)	정형 & 비정형 객체 (Rigid & non-rigid objects)

표 1. 추적에 관한 여러 알고리즘과 특징들
Table 1. Some of the object tracking algorithms and their properties.

객체 추출과 추적기술에는 많은 방법들이 있다. 단순하지만 많이 사용하는 방법으로 차분 추적 알고리즘 (difference tracking algorithm)^[7]이 있으며 적응적 배경 갱신 (adaptive background generation)과 배경 차분 (background subtraction)을 그 예로 들 수 있다. 적응적 배경 갱신 알고리즘은 고정 카메라에서 움직이는 객체를 추출하는데 유용하다. 만약 카메라가 고정되어 있다면, 정적인 배경에서의 화소의 밝기 값은 크게 변하지 않기 때문에 고정되어 있거나 변화가 있는 화소 값을 선택적으로 얻을 수 있다. Pfister^[8]는 실시간에서 사용할 수 있는 블럽(blob) 추적으로 배경 영상이 상대적으로 정적이라고 가정한다. 또한 추적된 객체의 색상 도수 분포도(color histogram)를 이용하여 블럽 추적을 수행하는 알고리즘도 있다^[9]. 형태 기반 추적에서는 객체들의 대한 정보를 사전에 학습하여 학습 집합을 얻으며, 객체의 윤곽과 가장 근접한 위치로 학습 집합을 투영 (projection)한다. 이러한 방법으로는 윤곽기반 방법 (contour-based method)^[10]과 능동형태 모델 (active shape model ; ASM)^[11], 상태 공간 표본추출 (state-space sampling) 접근 방법인 응축 알고리즘 (condensation algorithm)^[12] 등이 알려져 있다. 표 1.에 기존의 추적 알고리즘들의 특징이 정리되어 있다.

기존의 형태 기반 알고리즘들은 불완전하거나 잡음이

많이 들어간 영상에서도 강건하게 객체의 외곽을 정합하여 추출할 수 있으며, 객체의 부분 가림 (partial occlusion) 현상 등을 해결할 수 있다는 좋은 성능을 갖고 있지만 사전에 추적하고자 하는 객체마다 훈련 집합을 만들어야 하고, 모델 수립 과정이 반복적으로 이루어지는 등 실시간 환경에서 구현하기 적합하지 못한 단점들을 가지고 있다.

따라서 본 논문에서는 학습 집합을 사전에 확보하는 대신 실시간으로 생성함과 동시에, 소수의 특징점만을 사용하여 객체를 추출하고 추적할 수 있는 무학습 능동 특징점 모델 기반 특징점 추적을 제안한다. 제안한 알고리즘은 움직임 분할을 사용하여 객체를 추출하고 추출된 객체 안의 특징점을 계산한다. 이렇게 얻어진 특징점들은 객체의 강한 에지 또는 경계에 매우 근접하게 되며 선택된 특징점들은 다음 프레임에서 광류 추정에 의해 추적되고 적절하지 않게 추적되거나 추적에 실패한 특징점들은 예측과 보정 과정을 거쳐서 보정된다. 만약 객체의 가림 현상이 있을 경우에는 NPT-AFM이 객체의 보이지 않는 부분의 특징점들을 유지시켜 준다.

제안한 NPT-AFM 알고리즘은 본 논문의 가장 중요한 기여로서, 기존의 형태 기반에서 객체의 가림 현상으로 많은 부분이 보이지 않는 객체를 ASM 원리를 사용하여 추적하고 해석하는 동시에 학습 집합을 실시간으로

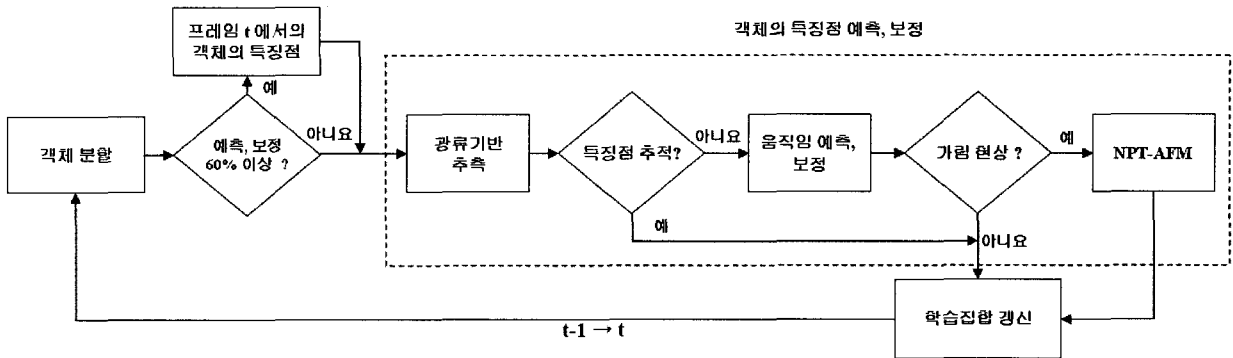


그림 1. 제안한 NPT-AFM 추적 알고리즘의 개요
 Fig. 1. Overview of proposed NPT-AFM algorithm.

로 갱신하기 때문에 실시간 감시 시스템에 응용할 수 있다는 이론적 근거가 된다.

본 논문의 II장에서는 추적 알고리즘 개요를 소개하고 III장에서는 객체 분할, 특징점 추출, 광류 기반 특징점 추적 기술을 소개한다. IV장에서는 NPT-AFM을 이용한 객체 가림 현상 해결을 다루며, V장에서는 실험 결과를 정리한 후 결론과 함께 향후 연구방향을 제시한다.

II. NPT-AFM 추적 알고리즘의 개요

제안한 알고리즘의 전체 개요는 그림 1.에 나타나있다. 여기서 점선으로 된 사각형 영역은 실시간으로 $t-1$ 번째 프레임에서 t 번째 프레임으로 특징점을 추적, 예측, 보정하는 과정을 나타낸다. 객체 분할에서는 움직임 분할과 레이블링(labeling)을 사용하여 영상을 객체의 형태와 객체가 움직이는 방향으로 분할한다. 움직임 분할 후 특징점을 계산하여 광류로 추적한다. 만약, 추적한 특징점이 적절하지 않게 추적되거나 추적에 실패했을 경우 시간, 공간적 예측 방법을 이용하여 특징점을 보정하고 보정한 특징점들에 대하여 목록을 작성하여 특징점 재탐색을 하게 되는데 자세한 알고리즘은 III장의 2절에서 다룬다.

객체를 추적시 실시간으로 학습 집합을 갱신하고 객체의 가림 현상을 판별하여 가림 현상이 생겼을 경우에는 제안한 NPT-AFM을 이용하여 이를 해결한다. 또한 예측, 보정된 특징점이 전체 특징점의 60% 이상일 경우에는 특징점을 다시 계산한다.

본 논문의 실험에서는 객체의 방향을 4방향으로 분할하였고, 추적에 알맞은 특징점을 추출한 후 *Lukas-Kanade* 광류를 사용하여 다음 프레임의 특징점을 예

측하였다. 학습 집합은 70개 단위로 갱신하지만, 만약 50개의 학습집합을 갱신하는 중 객체의 가림 현상이 생기면 50개의 학습 집합을 이용하게 된다.

제안한 특징점 알고리즘의 이점을 요약해 보면 다음과 같다.

- 사전 학습(prior training)된 절차 없이 정형(rigid), 비정형(non-rigid)객체들을 모두 추적할 수 있다. 즉, 학습 집합을 실시간으로 갱신하여 정형, 비정형 객체를 추적한다.
- 움직임의 방향을 추출하고 분할하는 동시에 객체의 특징점까지 추적하기 때문에 객체의 갑작스런 움직임 변화에 강건하다.
- 움직임 분할정보를 이용하여 객체 경계 근처의 특징점과 내부의 특징점을 이용하였기 때문에, 배경이 복잡한 환경에서도 객체의 움직임을 강건하게 추적한다.
- NPT-AFM을 이용하여 객체의 가림 현상을 실시간에서 효과적으로 해결한다.

III. 광류기반 특징점 예측과 보정

이 장에서는 광류기반 특징점 추적에 관한 내용을 설명한다. 전반적인 추적 과정 중에서 특징점이 적절하지 않게 추적되거나 추적에 실패했을 경우, 특징점의 시간, 공간적 정보와 예측된 특징점 주위로 구획(block)을 설정하여 특징점을 재탐색하는 것이 제안된 기술의 핵심이다.

1. 객체 분할

추적하고자하는 객체를 영상의 배경으로부터 추출하

기 위하여 광류를 사용한 움직임 분할과 레이블링을 이용하였다. 광류는 극히 짧은 시간 동안에 영상에서의 특정한 점의 밝기(intensity) 값이 변화하지 않는다는 가정 하에 객체의 움직임을 계산하는 방법이다. 만약 연속 영상에서 공간, 시간적 밝기 분포를 $s_c(x, y, t)$ 라고 한다면, 광류는 식(1)과 같이 정의할 수 있다.

$$\frac{ds_c(x, y, t)}{dt} = 0. \quad (1)$$

식(1)을 연쇄법칙(chain rule)을 적용하면

$$\frac{\partial s_c(x, t)}{\partial x} v_1(x, t) + \frac{\partial s_c(x, t)}{\partial y} v_2(x, t) + \frac{\partial s_c(x, t)}{\partial t} = 0. \quad (2)$$

와 같이 표현되고, 이로부터 광류는 다음과 같이 정의할 수 있다.

$$\xi_{oj}(v(x, t)) = \langle \nabla S_c(x, t), v(x, t) \rangle + \frac{\partial S_c(x, t)}{\partial t} = 0. \quad (3)$$

하지만 식(2)는 미분을 근사적으로 계산하기 위해서 주변 화소들과의 차분을 이용하기 때문에 광류의 결과는 잡음에 민감한 특성을 갖게 된다. 잡음의 영향을 줄이기 위해 널리 사용하는 기술로 *Horn-Schunk* 과 *Lucas-Kanade* 알고리즘 등이 있으며 *Lucas-Kanade* 는 부분적인 블록들에서 움직임의 밝기 변화가 일정하다는 가정 하에 움직임 모델을 제안하였고, 가정에 대한 오류(E)는 다음과 같이 정량화할 수 있다.

$$E = \sum_{x \in Neighbor} (\xi_{oj})^2. \quad (4)$$

식(4)의 오류를 최소화하기 위하여 움직임 벡터를 계산할 때, 미세한 움직임을 제거하여 움직임 영역을 분할하였고 움직임 영역을 레이블링 하였다. 객체 분할 알고리즘은 그림 2.에서 볼 수 있다.

Lucas-Kanade 알고리즘은 광류를 계산할 때, 반복적 계산을 없앴기 때문에 실시간으로 구현가능하다는 장점이 있다. 그림 3에서 보듯이, 연속된 프레임에서 움직임을 추출하여 영상 좌표를 기준으로 4방향으로 객체의 움직임을 분할하고 형태학적 연산(morphological operation)을 적용하여 영역의 작은 틈(hole)을 제거하였다. 하지만 낮은 공간 그래디언트(low spatial gradient)의 영향으로 잡음이 함께 추출되기 때문에 객체와 잡음을 분간하기 위해서 다음과 같은 척도(criterion)를 사용하였다.

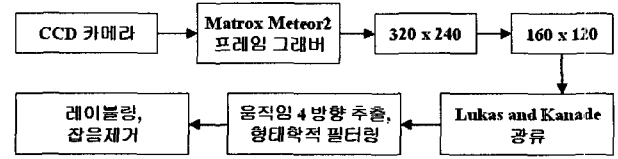


그림 2. 객체 분할
Fig. 2. Object segmentation.

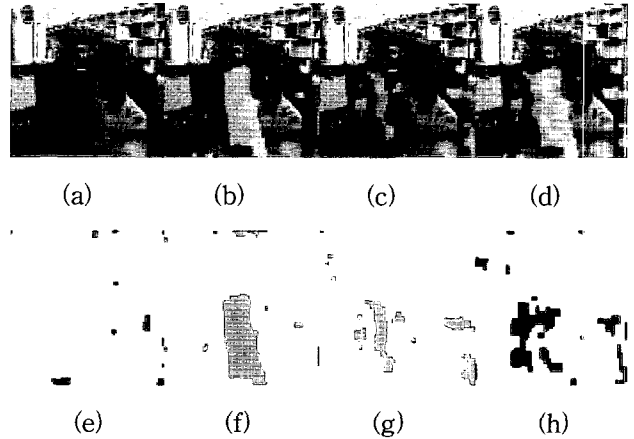


그림 3. 4방향 움직임 추출: (a) 원영상 (b) 수평 움직임 (c) 수직 움직임 (d) 수평, 수직 움직임 (e) 0° 움직임 (f) 180° 움직임 (g) 90° 움직임 (h) 270° 움직임
Fig. 3. 4 Direction motion segmentation: (a) original image (b) horizontal motion (c) vertical motion (d) horizontal and vertical motion (e) 0° motion (f) 180° motion (g) 90° motion (h) 270° motion.



그림 4. 잡음 제거
Fig. 4. Noise removal.

$$C_i = \frac{\sum_{(x, y) \in A} (|I_t(x, y) - I_{t-1}(x, y)|)}{\sum_{(x, y) \in A} 1} < t_n, \quad (5)$$

여기서 A 는 i 번째 분할영역을 나타내고, 식 (5)는 두 프레임 I_t 과 I_{t-1} 의 정규화된(normalized) 차분을 나타낸다. 우리는 실험에서 문턱치(threshold) 값 t_n 을 0.55로 사용하였으며 결과는 그림 4.에서 볼 수 있다.

2. 특징점 추출과 예측

객체 분할 후 각 영역 안에서 Shi-Tomasi의 특징점 추출 알고리즘^[1]을 사용하여 관심 객체 안의 특징점을 추출하고 소수의 특징점을 광류로 예측하였다. 알고리즘 수행 순서는 다음과 같다.

알고리즘 1: 광류기반 특징점 추출 및 예측.

- (단계1.) 영상에 가우시안 저역통과필터(gaussian low-pass filter)를 적용한다.
- (단계2.) 미분 연산을 수행한다. d_x 는 x방향으로의 1차 미분을 의미하며, d_{xx} 는 x방향으로 2차 미분을 의미한다.
- (단계3.) 모든 화소에서 다음과 같은 변형(deformation) 행렬을 계산한다.

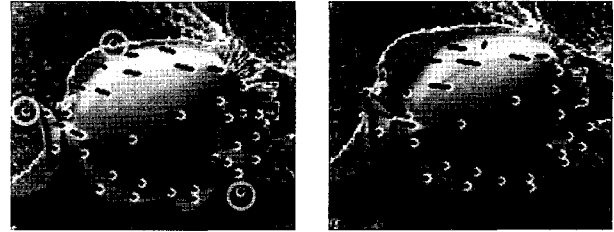
$$D = \begin{matrix} \sum d_{xx} & \sum d_{xy} \\ \sum d_{yx} & \sum d_{yy} \end{matrix} \quad (6)$$

- (단계4.) 변형 행렬의 충분히 작은 고유값(eigenvalue)을 계산하고 비극대 억제(non-maxima suppression)를 수행한다. (여기서 지역 최대값(local maxima)은 5근방 (5x5 neighborhood)에 존재한다고 가정한다.)
- (단계5.) 임계값 보다 작은 고유값을 버린다.
- (단계6.) 두 특징점들 사이의 거리를 측정하여 임계 거리에 만족하지 못하면 버린다. 본 논문에서는 임계값을 0.05, 임계거리를 10으로 사용하였다.
- (단계7.) 특징점 추출 후 광류를 사용하여 다음 프레임에서 특징점을 예측한다.

하지만 객체 분할과정에서 움직임 분할이 실제 객체보다 조금 크게 추출되기 때문에 영상의 복잡한 정도에 따라서 간혹 객체 외부에서 특징점이 추출될 수 있다. 이는 관심영역에 대한 특징점이 아니기 때문에 식(7)과 같이 프레임 사이의 특징점의 이동 거리를 사용하여 제거하였으며 결과는 그림 5.에서 볼 수 있다.

$$d = \sum_{t=1}^{N,M} \sqrt{(x_i^{t+1} - x_i^t)^2 + (y_i^{t+1} - y_i^t)^2} < t_n \quad (7)$$

여기서 t 는 프레임 수, i 는 특징점의 개수를 나타낸다. 본 실험에서는 5 프레임 내에서 특징점의 이동거리



(a) (b)

그림 5. 객체 밖의 특징점 제거: (a) 2nd 영상 (b) 7th 영상
Fig. 5. Feature point removal out of object: (a) 2nd frame (b) 7th frame.

를 계산하였으며 t_n 을 1.3으로 사용하였다.

3. 특징점 보정과 예측

연속적인 프레임에서 특징점을 추적할 경우 적절히 추적되지 않거나 추적에 실패할 경우가 있으며 그 이유는 다음과 같이 두 가지로 요약할 수 있다.

- 특징점 주변에 있는 객체 영역에 자기 가림 (self-occlusion)현상 또는 다른 개체에 의한 가림 (occlusion by another object) 현상이 생겼을 때
- 특징점이 추적 객체의 경계에 존재하고 배경화소가 변할 때

특징점을 제대로 추적하지 못하거나 실패할 경우, 추적 객체의 이전 정보들과 예측된 특징점들 사이의 위치 정보를 이용하여 특징점을 보정한다. 제안한 특징점 예측과 보정 알고리즘은 다음과 같다.

알고리즘 2: 특징점 예측과 보정.

- 1) 시간적 예측: t 번째 프레임의 특징점(v_i^t)이 $t+1$ 번째 프레임에서 추적에 실패했다고 가정 한다면 식(8)을 이용하여 특징점을 예측한다.

$$\hat{v}_i^{t+1} = v_i^t + \frac{1}{K} \sum_{k=0}^{K-1} m_i^{t-k}, \quad (8)$$

$$m_i^t = v_i^t - v_i^{t-1},$$

여기서 i 는 특징점의 개수를 나타내며 K 는 평균 움직임 벡터를 계산하기 위한 프레임 수를 나타낸다.

- 2) 공간적 예측: $t+1$ 번째 프레임에서 추적에 실패한 특징점을 예측하는 다른 방법은 t 번째 프레임에서

$t+1$ 번째 프레임으로 추적에 성공한 특징점들의 정보를 이용하는 것이다. 특징점 사이의 벡터정보와 벡터들 사이의 각 등이 좋은 예라고 할 수 있다.

시간, 공간적 예측 방법은 특징점을 예측하기 위하여 조합될 수 있으며, 본 실험에서는 시간적 예측에서 K 를 7로 설정하였고, 공간적 예측에서는 추적에 성공한 특징점들의 평균 벡터를 사용하였다. 정형 객체에는 공간적 예측의 가중치를 높였고 비정형 객체에는 시간적 가중치를 높였다.

3) **예측된 특징점의 재탐색:** 시간, 공간적 방법으로 예측, 보정된 특징점의 목록을 작성하고 예측, 보정된 특징점 주위에 구획을 설정하여 다음 프레임에서 구획안의 특징점을 재탐색하여 특징점이 존재 할 경우 갱신한다.

만약, 예측된 특징점이 전체 특징점의 60% 이상일 경우에는 추적 객체에 대한 특징점을 다시 추출한다.

IV. 무학습 능동 특징점

일반적으로, 2차원영상에서 비정형 객체에 대한 모델 알고리즘들은 객체를 표현하기 위하여 추적 객체의 경계에 대한 정보를 많이 이용한다. 비정형 객체에 대한 모델 방법 중의 하나인 능동 형태 모델은 추적하고자 하는 객체에 대하여 사전에 학습 집합을 만들었다는 전제 하에서, 그림 6. 같이 불안정한 영상에서의 객체의 윤곽 정합, 객체에 대한 적은 정보를 이용한 객체의 부분 가림 현상 해결 등 많은 장점을 가지고 있다^[13].

하지만, 사전에 추적하고자 하는 모든 객체에 대하여 학습 집합을 만드는 작업은 복잡하고 어려운 일이며 모델 수립 과정이 반복적으로 이루어진다는 점에서 실시간 환경에 적합하지 못한 단점들을 가지고 있다.

본 논문에서는 학습 집합을 사전에 확보하는 대신 실시간으로 학습 집합을 생성함과 동시에, 소수의 특징점만을 이용하여 객체를 추출하고 추적할 수 있는 무학습 능동 특징점 모델을 제안한다. NPT-AFM 알고리즘은 (i) 특징점을 이용한 기준점 할당(landmark point), (ii) PCA(principal component analysis), (iii) 특징점 모델 정합(model fitting), 그리고 (iv) 특징점 모델 재구성과 객체의 가림 현상 해결 순서로 나눌 수 있다.

1. 특징점을 이용한 기준점 할당

일반적으로 비정형 객체에 대한 모델을 사용하는 알고리즘들은 객체의 경계에 적당한 기준점을 수동으로 할당

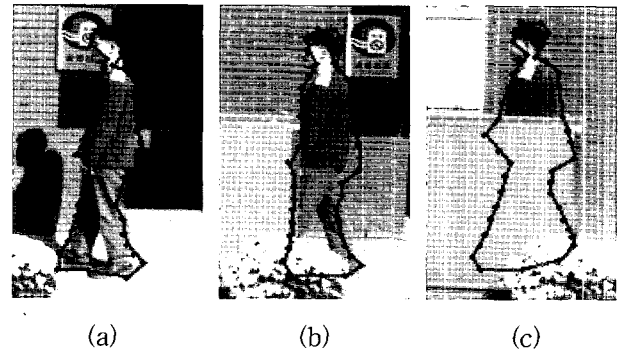


그림 6. 부분 가림 현상에서의 ASM을 사용한 모델 정합: (a) 45th 영상 (b) 65th 영상 (c) 79th 영상

Fig. 6. Model fitting for partial occlusion using ASM: (a) 45th frame (b) 65th frame (c) 79th frame.

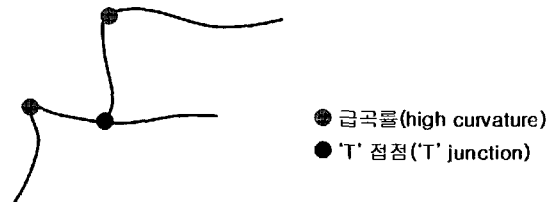


그림 7. 좋은 기준점 선택

Fig. 7. Choice good landmark point.

한다. 좋은 기준점이란 선택된 점들의 위치가 프레임 사이에 크게 변하지 않으며 그림 7.과 같이 급곡률(high curvature)과 'T' 접점('T' junction)에 위치한다. 특징점도 이와 같은 조건을 만족하므로 좋은 기준점이라 할 수 있으며 본 연구에서는 객체의 윤곽에 대한 특징점 뿐만 아니라 객체 내부의 특징점도 기준점으로 사용하였다. 객체의 외부의 특징점들은 이 장의 3절에 설명할 특징점 모델 정합에서 배경과 객체의 경계를 나타내는 중요한 요소로 사용된다. 객체 내부의 특징점들을 기준점으로 사용한 중요한 이유는 구획 맞추기(block matching) 혹은 광류 등 움직임 정보를 이용한 추적에서는 객체의 경계에 대한 특징점들은 내부의 특징점들에 비하여 객체를 추적하는 과정에서 실패할 확률이 높기 때문이다.

n 개의 기준점들은 객체의 특징을 나타내고 이들의 집합을 학습 집합이라 한다. 본 논문에서는 학습 집합을 실시간으로 생성함과 동시에 Procrustes Analysis^[14]을 사용하여 영상좌표에 정렬(alignment)하였다. 학습 집합은 70개 단위로 지속해서 갱신하였으며 이는 객체의 자기 가림 혹은 객체들의 가림 현상이 발생할 경우, 이전의 가장 가까운 객체에 대한 정보들이 특징점 모델 정합을 적용했을 경우, 객체와 가장 가까운 특징점 모델을 생성할 수 있기 때문이다.

2. PCA

n 개의 특징점을 $2n$ 차원 벡터를 사용하면 다음과 같이 표현할 수 있다.

$$x = [x_1, \dots, x_n, y_1, \dots, y_n]^T. \tag{9}$$

학습 집합에서 각각의 특징점들은 2차원 공간에 존재 하지만, 계산의 효율성을 위해 PCA를 사용하여 모집단(parameter)을 줄여서 특징점을 모델링 하였다^[15].

PCA 알고리즘을 간략하게 소개하면 다음과 같다. 여기서, m 개의 학습 집합을 생성하였다고 가정하고 x_i 는 각각의 특징점 모델에 따른 학습 집합을 나타낸다. ($i = 1, \dots, m$)

1) 특징점 모델 집합에서 m 개의 표본에 대한 평균을 구한다.

$$\bar{x} = \frac{1}{m} \sum_{i=1}^m x_i \tag{10}$$

2) 학습 집합으로부터 공분산 행렬(covariance matrix)을 계산한다.

$$S = \frac{1}{m} \sum_{i=1}^m (x_i - \bar{x})(x_i - \bar{x})^T. \tag{11}$$

3) 공분산 행렬로부터 다음과 같은 값을 얻을 수 있다.

$$\Phi = [\phi_1 | \phi_2 | \dots | \phi_q], \tag{12}$$

여기서 ϕ_j (단, $j = 1, \dots, q$)는 S 에서 q 개의 가장 큰 고유값에 해당하는 고유벡터를 나타낸다.

4) 주어진 Φ 와 \bar{x} 로부터 각각의 훈련 집합은 다음과 같이 간략화 시킬 수 있다.

$$x_i \simeq \bar{x} + \Phi b_i \tag{13}$$

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

벡터 b 는 변형 가능한 특징점 모델의 모집단을 나타내며 b 의 요소를 조절하여 특징점 모델을 변화시킬 수 있다. PCA 알고리즘의 3번째 단계에서 q 값은 데이터의 전체적인 분산 값에 98%에 해당하는 특징점의 개수를 나타낸다.

3. 특징점 모델 정합

학습 집합에 의한 특징점 모델을 실제 비디오 영상의

객체와 정합하여 최적의 위치와 특징점 모델을 갖는 모집단을 얻어야한다. 이것은 모델 좌표 프레임의 특징점 모델인 x 와 영상 좌표프레임의 특징점 모델인 y 를 정합하기위하여 아래와 같은 에리 함수를 최소화함으로써 이루어질 수 있다.

$$E = (y - Mx)^T W (y - Mx). \tag{14}$$

여기서 W 는 대각행렬이며, 그 대각성분은 각각 기준점에 대한 가중치를 나타내고 식(15)에서 알 수 있듯이, M 은 회전성분 θ , 전이성분 t , 스케일링 성분 s 로 이루어진 기하학적인 변환을 나타내는 행렬이고 가중치 요소는 객체의 특징점을 따라 선택된 이전의 기준점과 갱신된 기준점 사이 값의 변위와 관련이 있다. 변위가 크면 클수록 행렬에 해당되는 가중요소들은 작아지고, 반면에 변위가 작아질수록 가중요소들은 증가한다. 하나의 특징점을 $[x_0, y_0]^T$ 으로 표현한다면, 기하학적인 변환행렬은 다음과 같이 정의 될 수 있다.

$$M \begin{matrix} x_0 \\ y_0 \end{matrix} = s \begin{matrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{matrix} \begin{matrix} x_0 \\ y_0 \end{matrix} + \begin{matrix} t_x \\ t_y \end{matrix}. \tag{15}$$

기하학적인 모집단의 집합(θ, t, s)이 결정되면, 특징점 모델 좌표 프레임으로 투영된 y 값은 다음과 같이 나타낼 수 있고,

$$x_p = M^{-1}y, \tag{16}$$

결국 특징점 모델의 모집단은 다음과 같이 갱신된다.

$$b = \Phi^T (x_p - \bar{x}). \tag{17}$$

갱신된 기준점의 조합은 영상좌표 프레임 y 에 새로운 특징점 모델을 만들며 식 (14)을 사용하여 가장 근접한 특징점 모델을 찾는데 이용될 수 있다. M 을 통해서 최적의 위치를 도출한 후, 새로운 특징점 모델은 주어진 학습 집합에 주요 요소에 포함된 Φ 로 투영되고, 이러한 과정은 모집단 b 를 갱신시킨다. 그 결과, 주요 요소에 해당되는 비슷한 한 편차만이 특징점 모델의 모집단에 영향을 줄 수 있다. 특징점 모델의 모집단을 계산한 후에 x 에 의해 주어지는 새로운 특징점 모델은 식 (13)에 의해 얻을 수 있다. 아래의 그림은 51개의 학습 집합을 가지고 제안한 시간적 예측방법으로 모델 정합을 했을 때의 결과이다.

4. 특징점 모델 재구성과 객체의 가림 현상 해결 기준점 할당, PCA, 모델 정합을 이용하여 객체에 대한

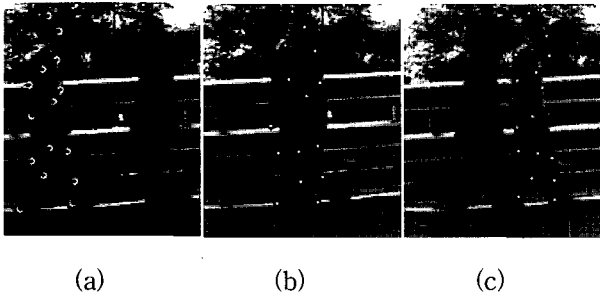


그림 8. 무학습 능동 특징점 모델을 사용한 모델 정합: (a) 광류기반 특징점 추적(40rd 영상), (b) 모델 정합(74th 영상) (c) 모델 정합(92nd 영상)
 Fig. 8. Model fitting in NPT-AFM: (a) optical-based feature tracking(40rd frame), (b) model fitting (74th frame), (c) model fitting(92nd frame).

적은 특징점 정보로 객체에 특징점을 모델링 하고 객체의 부분 가림 현상을 해결 할 수 있다. 하지만, 그림 6.의 ASM 정합 과정과 그림 8. 제안한 NPT-AFM의 정합 과정에서 알 수 있듯이, 모델 정합후의 특징점 모델은 학습 집합으로부터 만들어진 특징점 모델이기 때문에, 가림 현상을 해결한 후의 객체와 정확하게 일치하지 않을 가능성이 높다. 이를 해결하기 위해서 ASM과 같이 객체의 형태 정보를 이용하는 알고리즘에서는 사전에 학습 집합 생성할 때, 기준점에 해당하는 점들에서 물체와 수직이 되는 방향으로 각각의 밝기 값들을 저장하고 추적 시 밝기 값들을 비교하여 객체의 윤곽을 찾는지만, 모델 수렴 과정에서 반복적 계산이 이루어지기 때문에 실시간 처리에서는 적합하지 않다.

본 논문에서는 실시간 처리를 위하여 모델 정합 후, 객체의 외부의 특징점을 객체의 윤곽과 가장 가까운 내부로 옮기기 위하여 외부의 특징점을 특징점 모델의 평균점 방향으로 이동하면서 특징점 주위의 3근방을 탐색하여 3근방이 모두 움직임 분할 영역 안에 있는 점을 찾아 모델을 재구성하였다. 그리고 외부의 특징점이 특징점 모델의 60% 이상이면 객체에 대한 특징점을 다시 계산하였다. 아래의 그림 9은 객체의 모델 정합 후 특징점 재구성을 보여준다.

모든 추적 알고리즘에서 객체의 가림 현상은 중요하게 다루는 요소 중 하나이다. 제안한 NPT-AFM에 대한 이론적인 설명을 전제로 객체의 가림 현상을 해결하는 방법을 설명한다.

알고리즘 3: NPT-AFM을 이용한 객체의 가림 현상 해결.

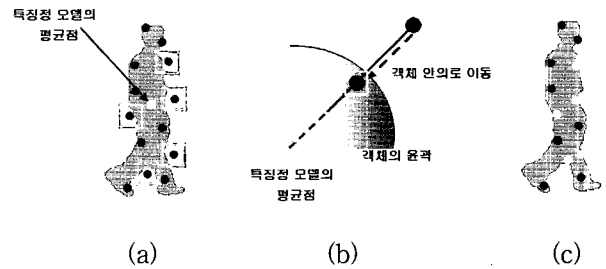


그림 9. 특징점 모델의 재구성:(a)특징점 모델 정합 (b)특징점 재구성 방법 (c)특징점 모델 재구성 결과
 Fig. 9. Reconstruction of feature model: (a) feature model fitting (b) method feature reconstruction (c) result of feature model reconstruction.

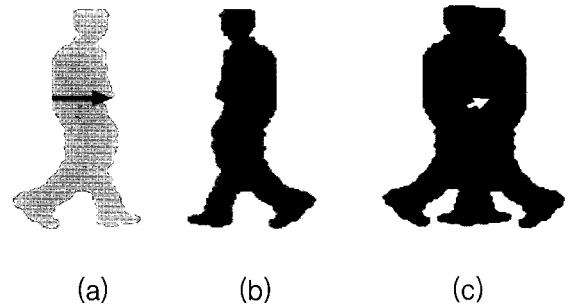


그림 10. 객체의 가림 현상 판별 : (a) 객체 1의 레이블과 방향 (b) 객체 2의 레이블 영역과 방향 (c) 객체의 가림
 Fig. 10. Occlusion distinction of object (a) label and direction of object 1(b) label and direction of object 2 (c) occlusion of object.

(단계1.) 추적 객체에 대한 가림 현상의 존재여부를 판별한다. 본 실험에서는 추적 객체의 레이블링 영역이 1.6배 이상일 경우, 레이블링된 영역 안에서 움직임 방향과 그 크기의 평균을 계산하여 객체의 가림 현상을 판별하였다. (그림 10.에서 볼 수 있듯이, 레이블링된 영역의 크기, 객체의 움직임에 대한 갑작스런 변화를 조사하여 객체 가림 현상의 존재 여부를 판별하였다.)

(단계2.) 학습 집합을 이용하여 특징점 모델 정합을 한다. 실시간으로 70개의 학습 집합을 갱신하는데 50개의 학습 집합을 갱신하는 중 객체의 가림 현상이 생겼을 경우에는 50개의 학습 집합을 이용하여 모델 정합을 하였다. (모델 정합과 제안한 예측과 보정과정을 조합하여 다른 객체에 의한 전체 가림 현상도 추적 객체의 방향이 갑자기 변하지 않는다는 가정 하에 해결할 수 있다.)

(단계3.) 객체의 가림 현상 해결 후 특징점 모델을 재구성을 한다. 가림 현상 해결 후의 시점은 가림 현

상을 판별하기 전의 레이블(L) 영역과 방향(D) 정보를 이용하였다. (가림 현상 전과 방향 벡터의 방향과 크기가 비슷하고 레이블 영역이 $0.8 L < L < 1.2L$ 만족할 때 특징점 모델을 재구성한다.)

V. 실험결과

우리는 실내, 실외에서 해상도(resolution) 320 x 240 연속 비디오 영상을 취득하여 정형, 비정형 객체에 대한 실험을 하였다. 본 논문의 실험영상들에서 노란 바탕에 검은(밝은) 색상을 가지고 있는 원들은 성공적으로 추적되고 있는 특징점을 나타내고, 파란 바탕에 흰(어두운) 색상을 가지고 있는 원들은 적절히 추적되지 않거나 추적에 실패한 특징점에 대하여 특징점 예측과 보정, 혹은 모델 정합이 적용된 특징점을 나타낸다.

1. 정형 객체 추적

정형 객체에 대한 실험으로 실내에서 *Pelco spectra pan-tilt-zoom(PTZ)* 카메라를 사용하여 로봇 영상을 획득하였으며 *Pelco spectra PTZ*는 카메라의 특성상 pan, tilt, zoom에 대한 상대좌표 값을 얻어 올 수 있기 때문에 PTZ 카메라가 움직이지 않을 때의 프레임을 추출하여 제안한 알고리즘을 적용하였다. 아래의 그림 11은 정형 객체에 대한 추적을 보여준다.

2. 비정형 객체(사람) 추적

비정형 객체에 대한 실험 영상은 실내, 실외에서 *Sony SCC DC-393* 칼라 비디오 카메라를 사용하여 획득하였으며 렌즈는 초점(focal length)변화가 2.9 mm에서 7.5 mm 사이의 자동 조리개(auto-iris) 렌즈를 사용하였다. 특징점에 목록(index)을 작성하여 적절히 특징점을 추적하지 못하거나 추적에 실패한 특징점들은 제안한 특징점 예측과 보정에서 목록을 보고 특징점을 갱신하게 되며, 특징점 예측과 보정에서는 추적에 성공한 특징점들이 사용된다. 그림 13.에서 (c)와 (d) 프레임 사이에 예측과 보정된 특징점의 개수가 전체 특징점의 60%가 넘기 때문에 (c)의 13개의 특징점 모델이 9개의 특징점 모델로 변한 것을 알 수 있다.

3. NPT-AFM을 이용한 객체의 가림 현상 해결

제안한 NPT-AFM 기반 추적 알고리즘은 크게 객체

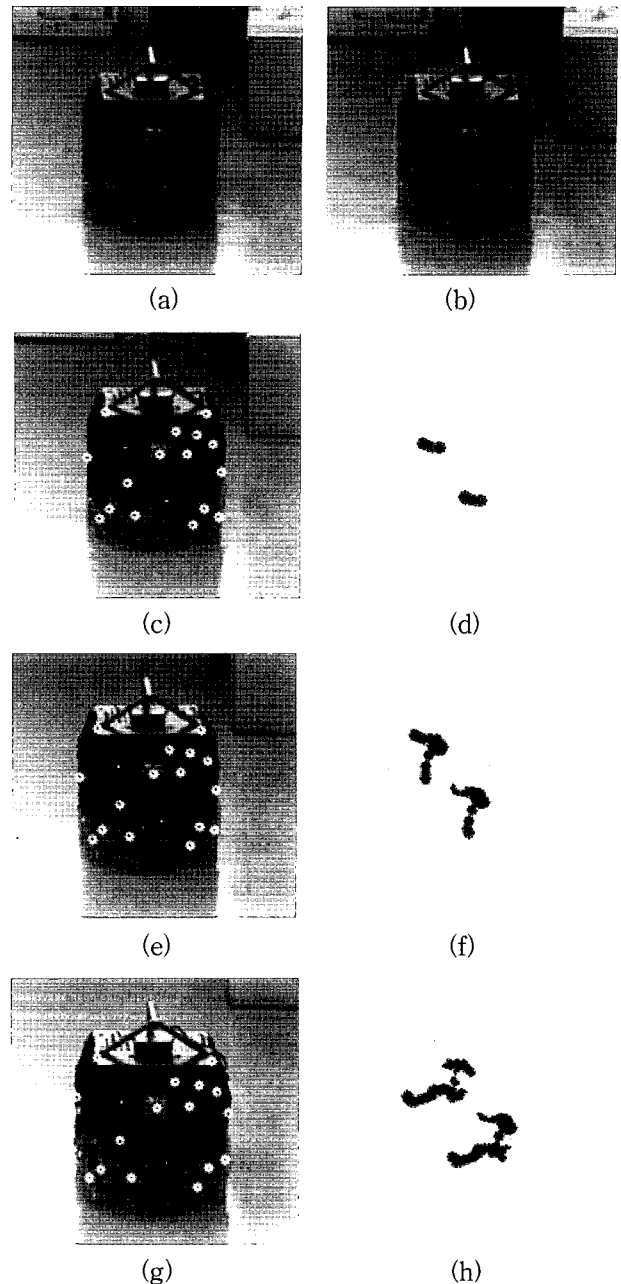


그림 11. 정형 객체의 특징점 추적과 추적 특징점의 자취: (a) 원 영상 (b) 움직임 분할 (c) 3rd 영상 (d) 3rd 영상에서의 자취 (e) 10th 영상 (f) 10th 영상에서의 자취 (g) 34th 영상 (h) 34th 영상에서의 자취

Fig. 11. Feature tracking of rigid object and trace of tracked feature point: (a) original image (b) motion segmentation (c) 3rd frame (d) trace in 3rd frame (e) 10th frame (f) trace in 10th frame (g) 34th frame (h) trace in 34th frame

의 가림 현상 판별, 특징점 모델 정합, 특징점 모델 재구성 과정을 거쳐 객체의 가림 현상을 해결하며 그림 14.

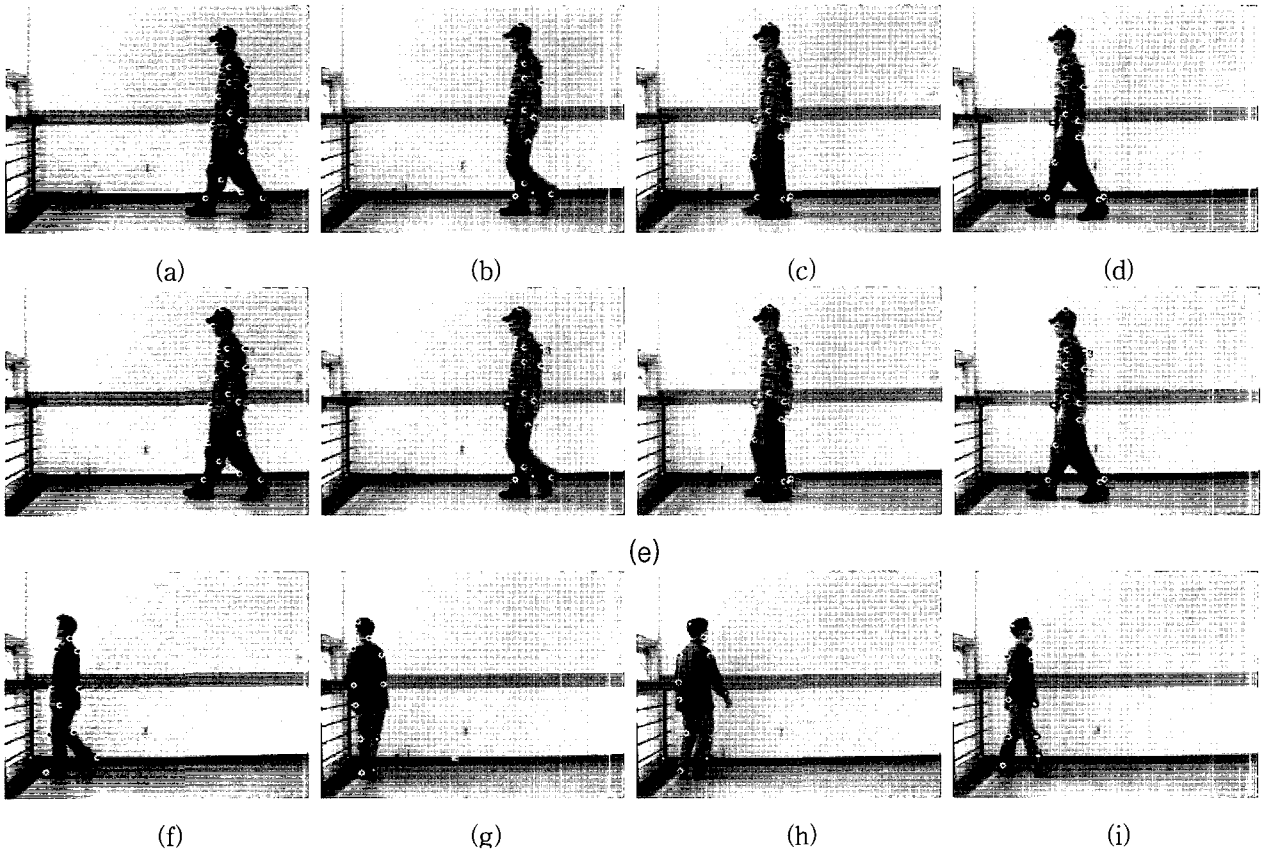


그림 12. 실내에서 비정형 객체의 특징점 추적과 특징점 목록: (a) 5th 영상 (b) 13rd 영상 (c) 32nd 영상 (d) 39th 영상 (e) 5th, 13rd, 32nd, 39th 영상에 대한 각각의 특징점 목록 (f) 27th 영상 (g) 58th 영상 (h) 75th 영상 (i) 113rd 영상
 Fig. 12. Feature tracking of non-rigid object and feature index in indoor: (a) 5th frame (b) 13rd frame (c) 32nd frame (d) 39th frame (e) Feature index in 5th, 13rd, 32nd, 39th frame, respectively (f) 27th frame (g) 58th frame (h) 75th frame (i) 113rd frame.

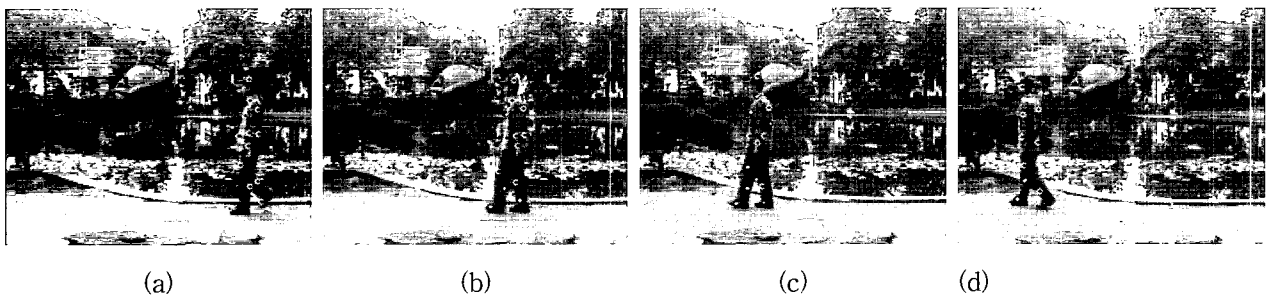


그림 13. 실외에서 비정형 객체의 특징점 추적: (a) 12nd 영상 (b) 46th 영상 (c) 98th 영상 (d) 133rd 영상
 Fig. 13. Feature tracking of non-rigid object in outdoor: (a) 12nd frame (b) 46th frame (c) 98th frame (d) 133rd frame.

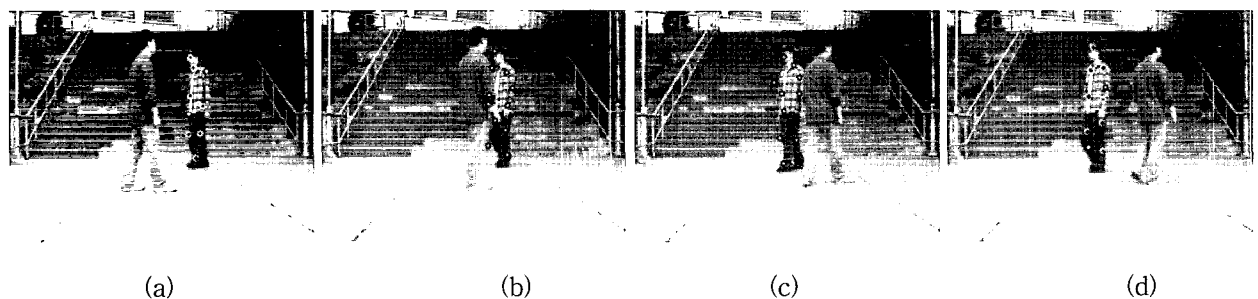


그림 14. NPT-AFM을 이용한 객체의 가림 현상 해결: (a) 106th 영상 (b) 125th 영상 (c) 165th 영상 (d) 177th 영상
 Fig. 14. Occlusion handling using NPT-AFM: (a) 106th frame (b) 125th frame (c) 165th frame (d) 177th frame.

에서 볼 수 있듯이, 다양한 변화를 가지고 있는 비정형 객체의 가림 현상에서도 강건하다. 또한 다른 객체에 의한 전체 가림 현상이 생겼을 경우에도 추적 객체의 방향이 갑자기 변하지 않는다는 가정 하에 해결할 수 있다. 그림 14의 (b)는 객체의 가림 현상 판별 후의 모델 정합을 보여주며 (d)는 가림 현상 후의 모델 재구성 결과를 보여주고 있다.

VI. 결 론

본 논문에서는 NPT-AFM을 이용하여 정형, 비정형 객체에 대한 새로운 추적 알고리즘을 설명하였다. 본 논문에서 제안한 추적 알고리즘을 간략히 설명하면, 추적 객체의 움직임 영역을 분할하고 분할된 영역 안에서 특징점을 계산하여 특징점을 광류로 추적한다. 적절하게 추적되지 않거나 추적에 실패한 점에 대해서는 특징점들의 공간, 시간적 정보를 이용하여 예측하고 보정하게 되며 예측, 보정한 점에 대하여 목록을 만들어 특징점을 재탐색 한다. 또한 객체를 추적하는 과정에서 실시간으로 학습 집합을 갱신하여 객체의 가림 현상시 특징점 모델을 만들어 이를 해결한다. 실험 결과 NPT-AFM은 정형, 비정형 객체에 대하여 강건한 추적 결과를 보여주고 있으며 학습 집합을 실시간으로 갱신한다는 점에서 실시간 감시 시스템에서 응용될 수 있을 뿐만 아니라 다른 새로운 알고리즘에 응용될 수 있다는 중요한 기여를 한다.

우리는 NPT-AFM 알고리즘에서 객체 윤곽 특징점과 객체 내부의 특징점을 분리하여 따로 계산할 수 있는 알고리즘을 구상하고 있다. 객체에 대한 윤곽 특징점과 내부 특징점을 분리할 경우 추적 객체의 윤곽정보를 이용하여 객체를 분리하고 인식할 수 있기 때문에

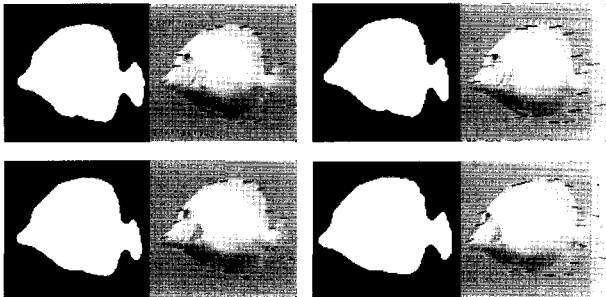


그림 15. 색상 분할을 이용한 객체의 윤곽 특징점 추출 및 추적

Fig. 15. Boundary feature point of object detection and tracking using color segmentation.

실시간 처리에서는 물론, 정확성을 요구하는 영상 후처리(video-post processing), 객체기반 압축(object-based coding) 등에서도 제안한 NPT-AFM 알고리즘이 응용될 수 있다. 단순한 실험으로 추적 객체에 대하여 색상 분할을 적용하여 객체 윤곽에서의 특징점을 추출하여 실험을 해 보았으며 그림 15에서 볼 수 있다.

참 고 문 헌

- [1] C. Eroglu, M. Tekalp, and B. Sankur, "Non-Rigid Object Tracking Using Performance Evaluation Measures as Feedback," *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, pp. 323-330, 2001.
- [2] J. Shi and C. Tomasi, "Good features to track," in *Proc. IEEE int. Conf. Computer Vision and Pattern Recognition*, pp. 593-600, 1994.
- [3] D. Bruce, Lucas and Takeo Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," In *Proc. DARPA image Understanding Work-shop*, pp 121-130, 1981.
- [4] S. McKenna, Y. Raja, and S. Gong, "Tracking Contour Objects Using Adaptive Mixture Models," *Image and Vision Computing*, pp 225-231, 1999.
- [5] R. Plankers and P. Fua, "Tracking and Modeling People in Video Sequences," *Computer Vision and Image Understanding*, pp 285-302, 2001.
- [6] D. Comaniciu, V. Ramesh and, P. Meer, "Kernel-based object tracking," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 25, No. 5, pp. 564 -577, May 2003.
- [7] I. Haritaoglu, D. Harwood, and L.S. Davis, "W-4: Real-time surveillance of people and their activities," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, pp. 809-830, Aug 2000.
- [8] C. R. Wren, A. Azerbayejani, T. Darrell, and A. P. Pentland, "Pfinder: Real-time tracking of the human body," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 780-785, July 1997.
- [9] D. Comaniciu, V. Ramesh and P. Meer, "Real-time tracking of non-rigid objects using mean shift," in *Proc. IEEE int. Conf. Computer Vision and Pattern Recognition*, vol. 2, Hilton Head Island, SC, pp. 142-149, June 2000.
- [10] A. M. Baumberg, "Learning deformable models for tracking human motion," Ph.D. dissertation, School of Comput. Studies, Univ. of Leeds, U.K.,

Oct 1995.

[11] 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.

[12] M. Isard and A. Blake, "Condensation conditional density propagation for visual tracking," Int. J. Comput. Vis, vol. 29, no. 1, pp. 5-28, 1998.

[13] 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.

[14] C. Goodall. Procrustes methods in the statistical analysis of shape. Journal of the Royal Statistical Society B, 53(2): 285-339,1991.

[15] 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.

저 자 소 개



김 상 진(학생회원)
 2003년 강남대학교 전자공학과
 학사졸업.
 2004년 현재 중앙대학교 첨단영
 상대학원 영상공학과
 석사과정.

<주관심분야: 비디오 처리, 비디
 오 감시 시스템, 영상 후 처리, 객체 기반 압축, 컴
 퓨터 비전>



이 성 원(정회원)
 1988년 서울대학교 제어계측
 공학과 학사졸업.
 1990년 서울대학교 제어계측
 공학과 석사졸업
 2003년 University of Southern
 California 전기공학과
 박사졸업

2004년 현재 중앙대학교 첨단영상대학원
 영상공학과 연구교수.
 <주관심분야: 미디어프로세서 및 SOC설계, 멀티
 미디어 신호처리, Power-Aware Computing>



신 정 호(정회원)
 1994년 중앙대학교 전자공학과
 학사졸업.
 1998년 중앙대학교 전자공학과
 석사졸업.
 2001년 중앙대학교
 첨단영상대학원
 영상공학과 박사졸업

2004년 현재 중앙대학교 첨단영상대학원
 영상공학과 연구교수
 <주관심분야: 영상처리, 영상복원, 해상도개선, 영
 상융합, 컴퓨터 비전>



백 준 기(정회원)
 1984년 서울대학교 제어계측
 공학과 학사졸업.
 1987년 노스웨스턴대학교 전기 및
 컴퓨터 공학과 석사졸업.
 1990년 노스웨스턴대학교 전기 및
 컴퓨터 공학과 박사졸업.

2004년 현재 중앙대학교 첨단영상대학원
 영상공학과 교수
 <주관심분야: 영상복원, 신호처리, 반도체>