

논문 2006-43SC-5-1

차별적인 영상특징들에 적응 가능한 융합구조에 의한 도로상의 물체추적

(Traffic Object Tracking Based on an Adaptive Fusion Framework for Discriminative Attributes)

김 삼 용*, 오 세 영**

(Sam-Yong Kim and Se-Young Oh)

요 약

대부분의 영상을 이용한 물체추적은 적용환경을 단순화하거나 특정한 영상특징만을 적용할 수 있는 제한된 환경에서 잘 동작하기 때문에 이러한 물체추적방법은 지능자동차의 운전자보조시스템이 적용되는 복잡하고 동적인 교통 환경에서 원하는 물체를 추적하기는 어렵다. 이와 같은 물체간의 부분적인 교합이 존재하고 배경과 물체들이 동시에 동적으로 변하는 복잡한 환경에서는 물체의 색상, 외관, 외형 등과 같은 다양한 영상특징들을 적절하게 융합할 수 있는 구조가 요구된다. 본 논문에서는 기존의 파티클 필터를 이용한 적응형 융합구조[1]와 SIFT[2]를 이용한 영상특징 기술자를 강인한 영상특징으로 사용하고 시점, 배경의 동적인 변화에 적응할 수 있도록 학습함으로써 추적의 강건성과 적응성을 향상시킨다. 제안된 알고리즘은 운전자 보조 시스템에서의 차량, 보행자, 자전거와 같은 도로상의 물체추적에 적용하였다.

Abstract

Because most applications of vision-based object tracking demonstrate satisfactory operations only under very constrained environments that have simplifying assumptions or specific visual attributes, these approaches can't track target objects for the highly variable, unstructured, and dynamic environments like a traffic scene. An adaptive fusion framework is essential that takes advantage of the richness of visual information such as color, appearance, shape, and so on, especially at cluttered and dynamically changing scenes with partial occlusion[1]. This paper develops a particle filter based adaptive fusion framework and improves the robustness and adaptation of this framework by adding a new distinctive visual attribute, an image feature descriptor using SIFT (Scale Invariant Feature Transform)[2] and adding an automatic learning scheme of the SIFT feature library according to viewpoint, illumination, and background change. The proposed algorithm is applied to track various traffic objects like vehicles, pedestrians, and bikes in a driver assistance system as an important component of the Intelligent Transportation System.

Keywords : Object tracking, visual attribute, particle filter, SIFT, driver assistance system

I. 서 론

물체추적은 비디오 감시, 사용자 인터페이스, 증강현실, 지능형 공간, 물체기반 비디오 압축, 운전자 보조 등과 같은 다양한 컴퓨터 비전의 응용 분야에 적용되는

기술이나^[3] 물체간의 부분적인 교합, 복잡한 배경, 배경과 물체의 동적인 움직임, 조도변화, 물체의 형태변화, 영상잡음이 존재하는 환경에서는 여전히 어려운 문제이다. 일반적인 추적과정은 크게 물체의 모델링, 예측, 정보비교, 갱신의 반복으로 구성된다. 모델링은 물체의 색상, 외관, 외형, 움직임, 깊이지도 등을 이용해서 물체를 나타내는 방법 등이 있으며 나머지 추적과정은 결정적인 방법과 확률적인 방법으로 분류할 수 있다. 결정적인 방법(deterministic method)은 Meanshift, Camshift

* 학생회원, ** 정회원, 포항공과대학교 전자전기공학부
(Dept. of Electronics and Electrical Engineering,
Pohang University of Science and Technology)
접수일자:2006년4월12일, 수정완료일:2006년8월26일

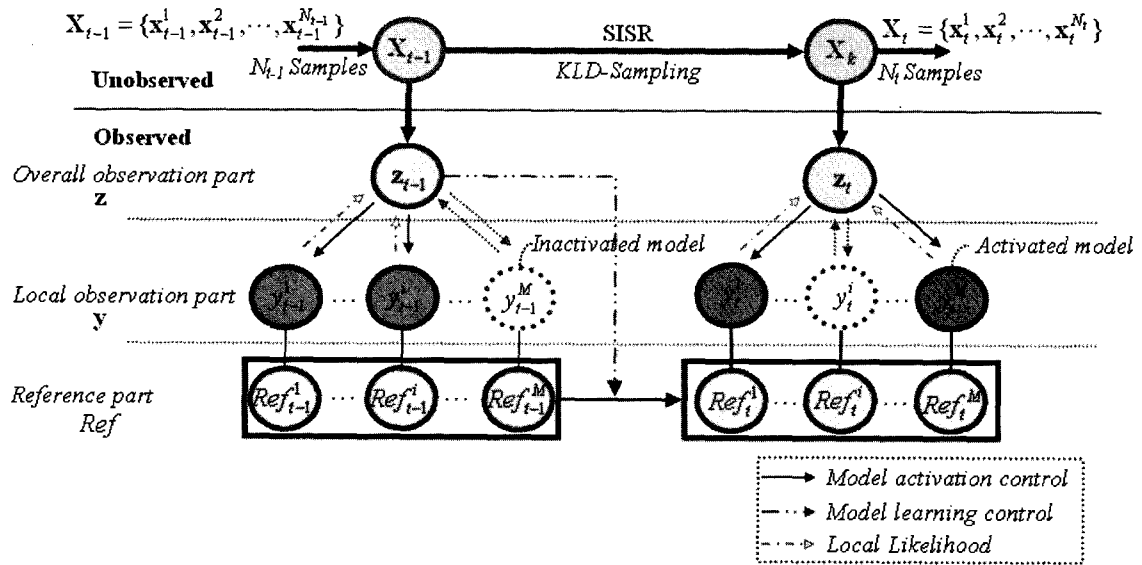


그림 1. 융합 및 적응 구조
Fig. 1. The fusion and adaptation framework.

와 같이 유사도의 지역 최적해를 반복적으로 찾는 방법이고 확률적인 방법(stochastic method)은 칼만필터(Kalman Filter), 확장형 칼만필터(Extended Kalman Filter: EKF), 파티클 필터(particle filter) 등과 같이 시스템을 모델링한 상태공간을 사용해서 최적상태를 찾는다^[4]. 물체의 특정 모델링을 사용하는 결정적인 방법은 물체간 부분적인 교합, 물체의 갑작스런 움직임, 동적인 배경의 변화가 있는 환경에서는 추적을 실패하기 쉽다. 칼만필터나 확장형 칼만필터는 물체의 상태를 위한 사후확률을 가우시안 형태로 모델링하는 한계가 있으며 일반적인 파티클 필터는 여러 개의 물체를 동시에 추적하는데 처리시간에 대한 부담이 크다^[5].

기존의 방법에서 물체추적의 환경에 대한 강건성을 보완하기 위해 색상, 외관, 외형 등에 의한 여러 가지 영상특징의 다차원적 조합은 동적인 실제 환경을 표현할 수 있는 충분한 정보를 가지고 모호성을 제거할 수 있기 때문에, 이러한 다양한 영상특징들의 가설을 계속 유지하기 위한 확률적인 방법에 기반한 융합구조가 제안되었다^{[11],[6],[7],[8]}. 특히 Soto^[11]는 융합구조에서 특정한 영상특징이 잡음성분을 증가시키기 때문에 적합한 특징점의 조합을 구성하기 위해 적절한 영상특징의 선택과 결합, 영상특징의 기준모델의 갱신 구조를 갖는 적응형 융합 구조를 제안하였다.

본 논문에서는 기존의 적응형 융합구조에 SIFT^[2]에 의한 지역영상기술자를 새로운 뚜렷한 영상특징으로 보충함으로써 보다 안정적으로 물체를 모델링함으로써 추

적성능을 향상시켰으며 이러한 영상특징기술자를 다양한 시점의 물체인식에 적용하기 위해 제안된 시계 클러스터(view-clustering)개념과 부분에 의한 물체표현개념(part-based object representation)^[9]을 이용해서 영상특징기술자를 학습하도록 하였다. 제안된 물체추적방법은 운전자 보조시스템에서 일반적인 복잡한 도로상의 물체 추적을 위해 적용되었다.

II. 본 론

1. 영상특징을 위한 적응 가능한 융합구조

차별적인 영상특징들의 적응 가능한 융합에 기반한 추적의 전체구조는 그림 1과 같다. 상태공간에서 목표물의 가정들을 유지하면서 추적하기 위해 각 영상특징을 표현하는 하위관측부(local observation part) y 와 이 정보들을 융합할 수 있는 상위관측부(overall observation part) z 의 계층적인 관측모델과 이들 사이의 여러 가지 적응신호를 가진 파티클 필터가 사용된다.

가. 영상특징들의 융합구조

베이지안 필터링문제는 비선형이고 가우시안 형태가 아닌 시스템에서 확률밀도함수를 계산할 수 있는 구조를 가지고 있다. 파티클 필터는 가중치를 갖는 일정한 샘플의 집합을 이용해서 이 문제를 해결하기 위한 베이지안의 연속적인 중요 샘플링(sequential importance sampling) 기술이다^{[4],[10]}. 일반적으로 시간 t 에서 모든

가능한 관측 $z_{1:t-1} = \{z_1, \dots, z_{t-1}\}$ 로부터 i 번째 상태 x_t^i 의 신뢰도 $b_t(x_t^i)$ 혹은 사후행렬 $p(x_t^i|z_{1:t})$ 을 추정하는 과정은 예측과 갱신의 과정으로 구성된다. 즉, 상태전이모델 $p(x_t^i|x_{t-1}^i)$ 에 의해 예측되고 베이어스 룰(Bayes rule), 마코브 가정(Markov assumption), 전확률 이론(theorem of total probability)을 이용해서 식 (1)과 같이 갱신된다^[11].

$$\begin{aligned} b(x_t^i) &= p(x_t^i|z_{1:t}) \\ &= \eta_t p(z_t|x_t^i) \sum_{x_{t-1}^i} p(x_{t-1}^i|x_t^i) p(x_{t-1}^i|z_{1:t-1}) \\ &= \eta_t p(z_t|x_t^i) \sum_{x_{t-1}^i} p(x_{t-1}^i|x_t^i) b_t(x_{t-1}^i) \end{aligned} \quad (1)$$

$$\eta_t = 1/p(z_{1:t-1})$$

파티클 필터에서는 사후행렬 $p(x_t^i|z_{1:t})$ 가 가중치 w_t^i 를 갖는 N_t 개의 샘플 $\{x_t^i\}_{i=1, \dots, N_t}$ 로 근사화되며 샘플들의 퇴화를 방지하기 위해 중요가중치에 따라 SISR (sequential importance sampling with re-sampling) 과정을 통해 샘플군을 갱신한다^{[4],[12]}.

그림 1에서같이 하위관측부에서는 각 관측모델 y_t^i 로부터 상위관측부에 해당 상태의 영상특징의 가능성(local likelihood)를 전달하고 상위관측부는 이 특징들을 융합해서 전체 가능성 $p(z_t|x_t^i)$ 을 생성하고 파티클 필터는 이를 이용해 현재 사후확률을 갱신하고 새로운 후보군을 샘플링해서 상태를 추정하기 위한 반복과정을 행한다.

나. 적응방법

같은 물체의 영상특징도 환경에 의해 계속 변하기 때문에 영상특징의 초기가정들로는 안정적인 관측을 할 수 없다. 또한 각 물체의 일정한 파티클에 대한 모든 영상특징의 가능성을 계산하는 것은 시스템의 효율성을 떨어뜨린다. 따라서 관측모델의 적응성을 향상시키고 파티클 필터 구조의 효율성을 증가하기 위한 적응방법이 사용된다. 기본적인 구조는 그림 1과같이 전체관측부가 좋은 영상특징을 나타내는 관측모델을 활성화 혹은 비활성화 시킬 수 있는 신호(model activation control)과 각 관측모델의 기준모델 Ref_t^i 을 학습시키기 위한 신호(model learning control)을 발생시키며 파티클 수를 조절하기 위한 적응방법이 사용된다. Soto는 실제 환경의 모호성과 동력학을 고려하기 위해 파티클의 수렴성을 나타내는 UD(uncertainty

deviation)과 각 영상특징의 우수성을 상태의 확률밀도 함수와 각 관측모델의 가능성과의 유사성에 의해 나타낸 QI(quality of the information)에 의해 전체관측부의 신뢰도를 강화하거나 약화시킬 수 있도록 하였다^[1]. 식 (2)와 같이 물체의 상태 x_t^i 가 위치 $[x, y]$, 크기 $[w, h]$, 속도 $[\Delta x, \Delta y]$, 영상비의 변화율 Δa 를 포함할 때 UD는 상태의 위치, 크기에 대한 분산의 합으로 표현되며 QI는 확률밀도함수간의 유사성을 나타낼 수 있는 KL-divergence로 계산된다.

$$\begin{aligned} UD &= \sigma_x^2 + \sigma_y^2 + \sigma_w^2 + \sigma_h^2 \\ QI_m &= \sum_i f(i) \ln(f(i)/g_m(i)) \end{aligned} \quad (2)$$

σ_k^2 : k 번째 차원에서의 파티클의 분산

$f(i)$: i 번째 상태의 확률밀도함수

$g_m(i)$: i 번째 상태의 m 번째 관측모델의 가능성

(1) 관측모델의 활성화/비활성화 방법

UD가 특정수준 α 보다 클 경우 파티클의 수렴성이 떨어지고 상태의 모호함이 증가한 것이므로 좋은 영상특징을 갖는 관측모델 즉, QI가 큰 관측모델에 활성화 신호를 보낸다. UD가 β ($\alpha > \beta$)보다 작을 경우 파티클이 우수한 가능성을 가지고 수렴이 잘 되는 것이므로 시스템의 효율성을 위해 QI가 낮은 관측모델을 비활성화시킨다. 그 밖의 UD에서는 현재 상태를 유지한다.

(2) 관측모델의 학습

UD가 β 보다 작을 경우 물체추적이 잘되고 있는 상태이므로 시스템의 안정성을 위해서 현재 모델 (MAP_t^i : maximum a posterior)을 이용해서 기존의 기준모델($OldRef_t^i$)를 식 (3)과 같이 학습시킨다.

$$\begin{aligned} NewRef_t^i &= \alpha MAP_t^i + (1 - \alpha) OldRef_t^i \\ \eta &= (1 - \tanh(2(UD - \theta)/\theta))/2 \end{aligned} \quad (3)$$

$NewRef_t^i, OldRef_t^i$: i 번째 관측모델의 갱신된 기준 모델과 기존의 기준모델

η : 가중치

θ : 갱신정도를 조절하기 위한 조정상수

(3) 파티클 수의 적응

고정된 파티클 수를 사용하는 것은 파티클 필터의 수렴성과 계산시간 측면에서 비효율적이기 때문에 실제

사후확률의 오차를 한정할 수 있는 KLD-Sampling^[12]을 이용해서 식 (4)와 같이 파티클 수 N 을 결정한다.

$$N > \frac{k-1}{2\epsilon} \left\{ 1 - \frac{2}{9(k-1)} + \sqrt{\frac{2}{9(k-1)} Z_{1-\delta}} \right\}^3 \quad (4)$$

k : 지원되는 bin의 수

ϵ : 오차의 상위 한계

$Z_{1-\delta}$: $1-\delta$ 의 확신도(confidence level)을 갖는 정규 분포의 변위치(quantile)

다. 기본 관측모델

관측모델을 위해서 기본적인 3가지 영상특징을 사용한다^{[11][6]}.

(1) 색상 모델

색상 모델은 식 (5)와 같이 색상 히스토그램에 대한 기준모델과 가정된 샘플의 색상히스토그램과의 유사도를 측정할 수 있는 Bhattacharyya 거리 D 를 이용하여 모델링한다.

$$D = \left[1 - \sum_{n=1}^N \sqrt{Ref_t^c(n)h_t^c(n)} \right] \quad (5)$$

$$P(h) = \exp(-\gamma_c D^2)$$

N : 색상 bin의 개수

Ref_t^c, h_t^c : 기준 색상모델과 가정된 샘플의 색상모델

γ_c : 조절상수

(2) 외관 모델

외관 모델은 식 (6)과 같이 $N \times M$ 크기로 정규화된 기준외관모델과 가정 샘플에서의 외관모델의 픽셀값간의 유사도와 가우시안 가중치 마스크를 이용해서 모델링한다.

$$P(h) = \exp\left(-\sum_{i=1}^M \sum_{j=1}^N W_{ij} \frac{(I_{ij} - Ref_{t,ij}^a)^2}{2\sigma_a^2}\right) \quad (6)$$

W_{ij} : 가우시안 가중치 마스크

$I_{ij}, Ref_{t,ij}^a$: 기준모델과 가정샘플간의 (i,j) 에서의 픽셀값

σ_a^2 : 분산

(3) 외형 모델

모델관측의 계산적 효율을 위해서 스네이크 모델과

같이 정교한 외형모델이 아닌 식 (7)과 같이 사각형 모양의 경계의 강도를 측정하는 방법을 사용한다.

$$G(I, x_t) = \frac{1}{LI_{G,\max}} \int_{a \in P(x_t)} |I_G(a) \cdot O_{P(x_t)}| da$$

$$P(h) = \exp(\gamma_s G) \quad (7)$$

$L = \int_{a \in P(x_t)} da$: 경계 주변 $P(x_t)$ 의 길이

$I_{G,\max} = \max_{a \in P(x_t)} |I_G(a)|$: 경계에서의 최대 기울기의 크기

$O_{P(x_t)}(a)$: 위치 a 에서 주변 $P(x_t)$ 에 대한 단위 수 직벡터

γ_s : 조절상수

$G(I, x_t)$ 는 영상 I 에서 상태 x_t 에 의해 생성되는 사각형 경계의 모든 방향에 대한 에지의 강도를 측정한다. 이러한 단순화는 차량과 같이 사각형 형태가 아닌 보행자, 자전거와 같은 물체를 추적을 위해서 좋은 영상특징이 아닐지라도 전체관측부의 가능성을 보조할 수 있다.

2. 영상특징기술자

가. SIFT

영상특징기술자(image feature descriptor)는 시점의 변화, 물체의 크기, 이동 및 회전, 일부 조도변화, 어퍼인이나 3D투영에 변하지 않는 영상내 특징점을 추출하고 기술할 수 있는 방법으로써 물체인식, 영상복구, 비전기반 SLAM^[13]과 같은 다양한 환경에서의 비전문제를 위해 사용된다. 예로써 steerable filter, differential invariants, moment invariants, complex filters, SIFT, BSIFT, PCA-SIFT, G-RIFT, 관심점들간의 상호연관성 등의 다양한 영상특징기술자가 제안되었다^{[9],[14]}. Lowe가 제안한 SIFT는 피라미드 크기 공간에서 가우시안 차 (Difference of Gaussian: DoG)의 최대값들을 핵심점(keypoint)으로 정의하고 방향성분의 히스토그램으로써 기술자를 정의함으로써 물체나 전경의 서로 다른 시점사이의 신뢰성 있는 매칭을 할 수 있는 영상특징기술자로 잘 알려진 방법이다^[2]. 이는 핵심점 추출과 기술자가 배경의 변화에 따라 오차를 갖지만 연속적으로 프레임에서의 무늬(texture)를 갖는 물체들에게는 좋은 반복성(repeatibility)을 갖는다. 또한 상태에 정의되는 관심영역을 설정함으로써 처리시간에 대한 부담을

줄임으로써 전체관측가능성의 융합을 위한 변별력 있는 영상특징으로 사용할 수 있다.

본 논문에서는 SIFT에 의한 관측모델을 영상특징기술자로 명하였으며 특징점 추출, 기술과 매칭을 위해서 (1) 핵심점 위치추정, (2) 핵심점을 위한 기술자 생성, (3) 핵심점 매칭, (4) Hough 변환을 이용한 클러스터링, (5) 최소자승법(least-squares solution)에 의한 검증의 방법을 사용했으며 잘못된 매칭을 줄이기 위해 거리비(distance ratio)^[2]를 사용하였다.

나. 영상특징기술자의 유사도(likelihood)

영상특징기술자는 영상안의 기술자들에 의해 생성되는 $P_{app}(h)$, 위치관계에 의한 $P_{pose}(h)$, 매칭된 특징점들의 수에 의한 $P_{NM}(h)$ 의 조합으로 표현되며 이들은 서로 독립이라 가정되어 식 (8)과 같이 표현된다.

$$P(h) = P_{app}(h) \cdot P_{pose}(h) \cdot P_{NM}(h) \quad (8)$$

연속된 프레임에서 매칭되는 특징점은 우수한 특징점으로 신뢰도 ζ 를 증가시키며 매칭이 되지 않는 경우 신뢰도를 감소시킴으로써 각 특징점에 가중치를 주었다.

$P_{app}(h)$ 는 식 (9)와 같이 입력영상의 영상기술자와 기준모델의 영상 기술자간의 유사성을 표현한다.

$$P_{app}(h) = \exp \left\{ \frac{-\frac{1}{N} \sum_i \zeta_i \|x_i - y_i\|^2}{\sigma_{app}^2} \right\} \quad (9)$$

N : 매칭된 핵심점의 수

x_i, y_i : 입력영상과 모델의 i 번째의 매칭된 핵심점

σ_{app}^2 : 분산

$P_{pose}(h)$ 는 식 (10)과 같이 Hough 변환과 검증에 의해 추정된 물체의 위치, 크기와 기준모델의 위치, 크기사이의 유사성을 나타낸다.

$$P_{pose}(h) = \exp \left\{ -\frac{\|x_{pose} - y_{pose}\|^2}{\sigma_{pose}^2} \right\} \quad (10)$$

x_{pose}, y_{pose} : 추출된 물체와 매칭된 모델사이의 위치와 크기, $x_{pose} = [xywh]$

σ_{pose}^2 : 분산

$P_{NM}(h)$ 는 매칭된 핵심점의 수의 유사도를 나타낸다.

$$P_{NM}(h) = \exp \left\{ -\frac{(1/N)^2}{\gamma_{nm}} \right\} \quad (11)$$

γ_{NM} : 조정상수

다. 영상특징기술자 기준모델 학습

이전에 소개된 관측모델의 학습방법으로 현재 영상특징기술자의 기술자를 이용해서 기준 모델의 기술자는 갱신된다. 또한 영상특징기술자의 안정성을 위해서 시점과 배경의 변화에 의해 새롭게 나타나거나 없어지는 핵심점을 위한 학습방법이 필요하다. 본 논문에서는 시계 클러스터링과 핵심점 공유를 이용한 물체표현을 이용해서 핵심점을 추가 또는 제거한다^{[9],[15]}.

(1) 핵심점 추가 방법

기존의 기준모델의 핵심점모음(f_{Lib})와는 다른 차별적인 새로운 입력 핵심점(f_{New})가 MAP에 나타날 경우, f_{Lib} 와 f_{New} 사이의 식 (12)와 같이 DM을 이용해서 유사성을 판단한다^[9]. DM은 f_{New} 가 현재 MAP과 기준모델사이의 매칭된 핵심점의 개수($N_{Shared}^{f_{Lib}}$)가 작고 MAP의 중심위치(O_{pose})와 f_{New} 의 중심위치(y_{pose})가 가까우수록 작아짐으로써 f_{New} 의 차별성이 커짐을 나타낸다.

$$DM(f_{New}) = \frac{\exp(-N_{Shared}^{f_{Lib}})}{\exp(-\|O_{pose} - y_{pose}\|^2)} \quad (12)$$

DM이 일정이하로 작아지면 f_{Lib} 에 f_{New} 를 추가한다.

(2) 핵심점 제거 방법

앞서 설명했듯이 기준모델의 핵심점의 신뢰도(ζ)는 반복적인 매칭에 의해서 증가 혹은 감소된다. 낮은 신뢰도를 갖는 핵심점은 시점이나 배경의 변화 혹은 물체

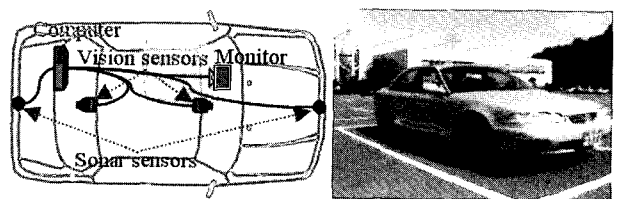


그림 2. 하드웨어 구조 및 테스트베드
Fig. 2. Hardware configuration and the testbed vehicle.

의 형태 변화에 의해 더 이상 나타나지 않는 핵심점으로써 계산시간을 증가시키고, 오탐지율을 높일 수 있으므로 핵심점모음으로부터 제거된다.

III. 실험 결과

지능자동차 (Intelligent Vehicle: IV)의 운전자 보조 시스템 (Driver Assistance Systems: DAS)은 다양한 도로상 물체나 차선을 인식하여 운전자의 인식력을 향상시키고 운전자가 적절히 대응할 수 있도록 적절한 정보를 발생하거나 제어함으로써 안전성과 편의성을 향상시킬 수 있다^[16].

그림 2는 실험을 위해서 기존의 차량인식 및 추적 시스템을 이용한 DAS^[17] 개발에 사용된 IV의 하드웨어구조와 테스트베드이다.

본 논문에서는 물체추적에 문제의 초점을 맞췄기 때문에 물체의 초기상태는 물체가 처음 나타나는 부분을 직접 선택해서 초기화되었다. 사용된 물체의 상태전이 모델 $p(x_t^i|x_{t-1}^i)$ 은 식 (13)과 같다.

$$\begin{aligned} \Delta X_t^i &= \Delta X_{t-1}^i + \kappa N(X_{MAP_t} - X_{t-1}^i) \\ X_t^i &= X_{t-1}^i + \Delta X_t^i \\ \Delta X_t^i &= [\Delta x_t^i \Delta y_t^i \Delta w_t^i \Delta h_t^i]^T \\ X_t^i &= [x_t^i y_t^i w_t^i h_t^i]^T \\ X_{MAP_t} &= [x_t^i y_t^i w_t^i h_t^i] \in MAP_t \\ \kappa &= 4 \times 4 \text{ 대각 상수행렬} \\ N &= 4 \times 4 \text{ 대각 가우시안 잡음행렬} \end{aligned} \tag{13}$$

파티클의 수는 최대 300개로 설정하였으며 평균적으로 211개정도가 사용되었다.

그림 3은 전방의 차량에 대한 각 관측모델의 유사도와 전체 유사도를 보여준다. 외형모델을 제외하고 각

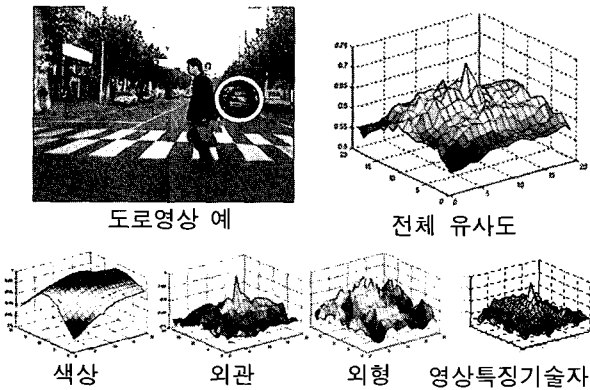


그림 3. 전체유사도와 각 관측모델의 유사도
Fig. 3. Overall likelihood and each local likelihood.

관측모델이 비교적 전체 유사도의 최대값, MAP의 위치로부터 가우시안형태와 유사한 형태를 가지고 있지만 각기 다른 영상특징의 관측모델에 의해 약간씩 다른 형태의 유사도를 발생하여 전체 유사도를 강인하게 모델링하게 된다.

실험에서는 안정적인 변별력을 갖는 색상 모델, 외관 모델, 영상특징기술자를 초기의 관측모델로 설정했다. 그림 4에 제안된 추적방법을 테스트하기 위한 연속적인 도로영상 예를 보였다. 실험 영상들에는 차량 2대(V#1, V#2), 보행자 2명(P#1, P#2), 자전거 1대(B#1)를 포함하고 있으며 차차량과 물체의 움직임으로 인한 시점과 외관, 형태의 연속적인 변화가 나타나며 물체들 사이의

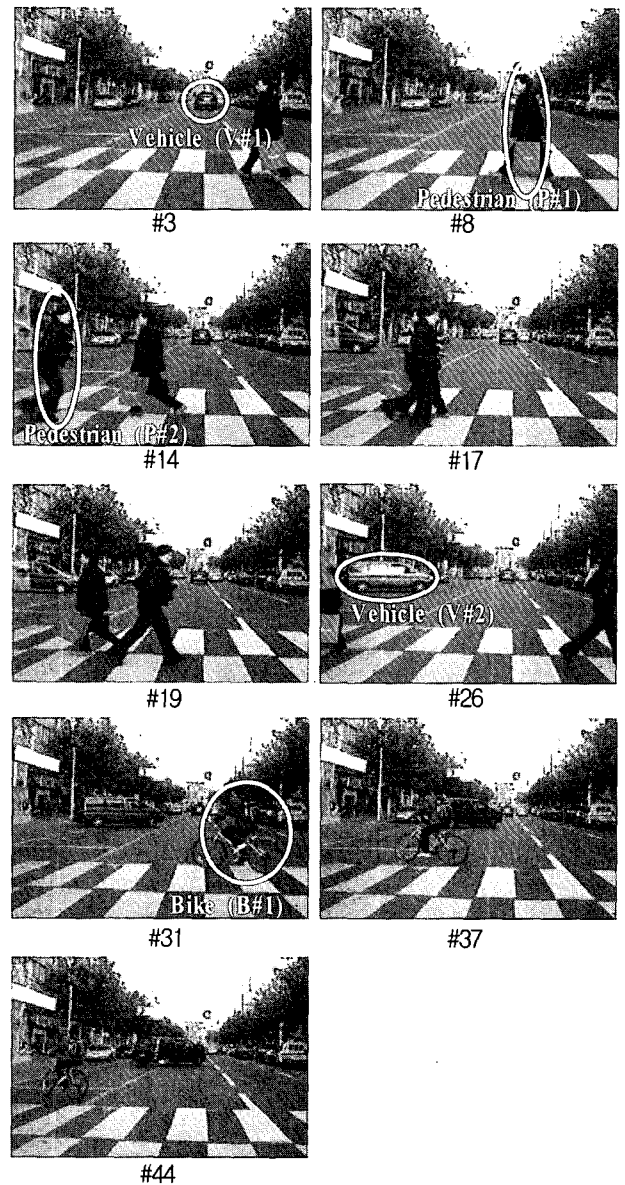


그림 4. 테스트 도로영상
Fig. 4. Test traffic scenes.

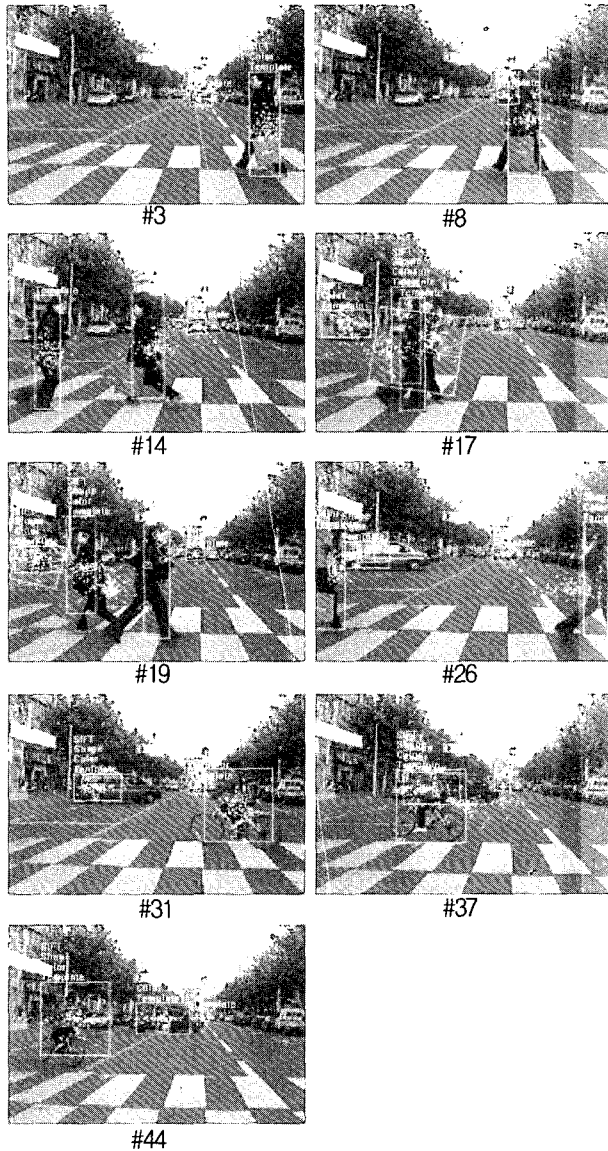


그림 5. 추적결과
Fig. 5. Tracking results.

겹치는 상태가 계속 발생한다.

그림 5는 물체추적 결과를 보여준다. 물체의 외접 사각형 주위의 점들은 각 물체의 파티클의 위치를 나타내고 상단의 주석은 전체관측 가능성을 위해 사용된 영상 특징들이다. 'Color'는 색상모델, 'Template'은 외형모델, 'Shape'은 외관모델, 'SIFT'는 영상특징기술자를 의미하며 노란색 사각형은 MAP안에 있는 SIFT의 핵심점들을 나타낸다. 물체추적시 약간의 위치의 오차가 있지만, 각 관측의 기준모델을 학습하고 적당한 영상특징을 선택하거나 제거하여 융합함으로써 복잡한 환경에서도 추적이 실패하지 않음을 확인할 수 있다.

그림 6은 각 물체추적시 나타나는 UD 지수를 나타낸다. 적응 방법에 의해 UD는 좋은 QI를 갖는 관측모델

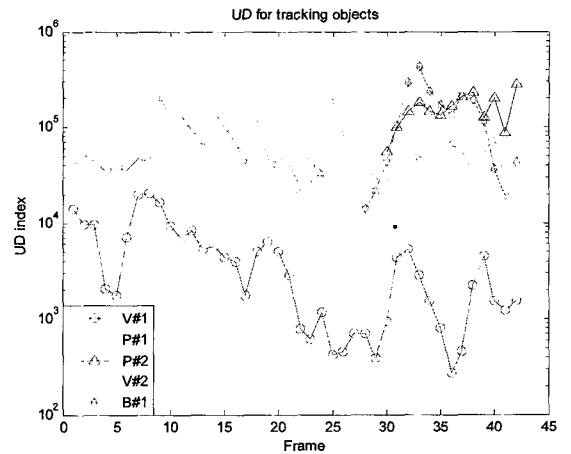


그림 6. 각 물체를 위한 UD 지수
Fig. 6. UD indices for each object.

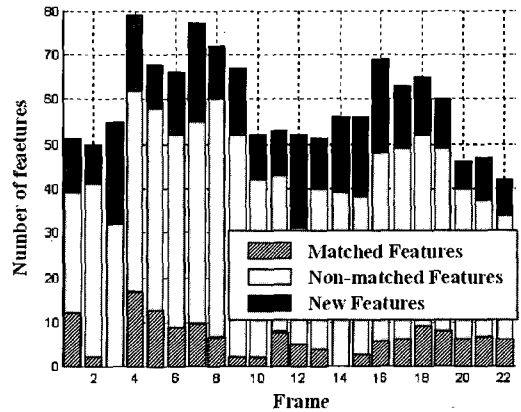


그림 7. 영상특징기술자의 학습
Fig. 7. Learning of the image feature descriptors.

을 활성화 시킨다. 예를 들어, V#1과 P#1의 경우에, 7번 영상과 9번 영상사이에서 두 물체가 겹침으로써 UD가 일정 수준을 벗어난다. UD를 낮추기 위해 V#1을 위한 전체 관측가능성은 영상특징기술자를 비활성화 시키고 P#1을 위한 전체 관측가능성은 외형 모델을 활성화 시켰다.

그림 7에서는 V#2를 위한 영상특징기술자의 변화를 보였다. 시점과 배경이 변함에 따라 핵심점의 추가 및 제거 방법에 의해 핵심점모음이 갱신됨을 확인할 수 있다. 시점과 배경의 변화에 상대적으로 매칭되는 핵심점에 비해 새롭게 추가되는 핵심점이 많음을 확인할 수 있다.

추적과정에서 활성화된 영상특징들의 빈도수를 이용해서 각 영상특징들의 유용성을 측정할 수 있고 특정 영상특징을 융합에서 제외한 후 실시한 물체추적의 결과를 통해 해당 영상특징의 중요도를 측정할 수 있다. 표 1은 이를 이용한 각 영상특징의 유용성과 중요도를

표 1. 각 영상특징의 유용성과 중요성

Table 1. Usefulness and importance for each visual attributes.

	색상	외관	외형	영상특징기술자
유용성	72%	53%	53%	88%
중요도	33%	11%	24%	35%

※ 유용성=(해당 영상특징이 활성화된 영상수/전체 영상수)×100(%)

중요도=(해당 영상특징을 제외하여 추적에 실패한 영상수/전체영상수)×100(%)

보여준다. 영상특징기술자가 사용빈도수가 비교적 높음으로써 변별성이 좋고 유용한 영상특징임을 알 수 있으며 색상모델과 영상특징기술자가 없을 경우 가장 낮은 추적성능을 보임으로써 중요한 영상특징으로 사용됨을 확인할 수 있다. 하지만 독보적인 영상특징은 존재하지 않았으며 적응 가능한 융합과정을 적용했을 경우 실패율이 3%인 경우에 비교했을 추적성능이 떨어짐으로써 낮은 중요도를 갖는 영상특징이라도 적절한 융합이 중요함을 보여준다.

IV. 결 론

적응 가능한 융합구조는 시스템의 모호성을 모델링하고 적응성을 추가함으로써 추적시스템의 장점을 더한다. 본 논문에서는 변별력이 뛰어난 영상특징인 영상특징기술자를 기존의 적응 가능한 융합구조에 보완함으로써 물체추적의 성능을 향상시켰다. 영상특징기술자는 시점, 조도, 배경의 변화에서도 좋은 변별력을 가지며 자동으로 학습이 가능하기 때문에 전체가능성이 상태를 잘 표현할 수 있도록 할 수 있도록 유효한 정보를 전달할 수 있다.

논문에서 사용된 영상특징이외에 스테레오 카메라를 이용한 깊이지도, 효율적인 외형모델, SIFT보다 우수한 Harris-Laplacian이나 G-RIF 등의 영상특징기술자 등의 보다 우수한 관측모델들의 융합을 통해서 강인하면서 효율적인 추적성능을 낼 수 있을 것이다. 또한 현재 추적시스템의 기본구조와 학습방법을 이용해서 도로상의 물체를 인식할 수 있는 운전자 보조시스템에 대한 연구의 인식을 향상에 기여할 것으로 기대한다.

참 고 문 헌

- [1] A. Soto, "A probabilistic approach for the adaptive integration of multiple visual cues using an agent framework," Doctoral dissertation, Tech. Report CMU - RI - TR - 02 - 30, Oct. 2002.
- [2] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91 - 110, Nov. 2004.
- [3] D. Comaniciu, V. Ramesh, P. Meer, "Kernel-Based Object Tracking," *IEEE T. Pattern Anal.*, vol. 25, issue 5, pp. 564 - 577, May 2003.
- [4] C. Yang, R. Duraiswami, and L. Davis, "Fast Multiple Object Tracking via a Hierarchical Particle Filter," in *Proc. 10th IEEE Int. Conf. Computer Vision*, vol. 1, pp. 212 - 219, Oct. 2005.
- [5] K. Deguchi, O. Kawanaka, and T. Okatani, "Object tracking by the mean-shift of regional color distribution combined with the particle-filter algorithm," in *Proc. 17th Int. Conf. Pattern Recognition*, vol. 3, pp. 506 - 509, Aug. 2004.
- [6] T. Xiong and C. Debrunner, "Stochastic car tracking with line- and color-based features," *IEEE T. Intell. Transp.*, vol. 5, no. 4, pp. 324 - 328, 2004.
- [7] F. Amell and L. Petersson, "Generic Fusion of Visual Cues Applied to Real-World Object Segmentation," in *Proc. 2005 IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, pp. 2954 - 2959, 2005.
- [8] J. Triesch, C.V.D. Malsburg, "Self-Organized Integration of Adaptive Visual Cues for Face Tracking," in *Proc. 4th IEEE Int. Conf. on Automatic Face and Gesture Recognition*, pp. 102 - 107, 2000.
- [9] S. H. Kim and I. S. Kwon, "Scalable Representation and Learning for 3D Object Recognition using Shared Feature-based View Clustering," in *Proc. Asian Conf. on Computer Vision 2006*, pp. 561 - 570, Jan. 2006.
- [10] B. H. Han, Y. Zhu, D. Comaniciu, and L. Davis, "Kernel-based Bayesian filtering for object tracking," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition 2005*, vol. 1, pp. 227 - 234, June 2005.
- [11] S. Thrun, "Probabilistic algorithms in robotics," *AI Magazine* vol. 22, no. 1, pp. 93 - 109, 2000.
- [12] D. Fox, "KLD-Sampling: Adaptive particle filter," *Advances in Neural Information Processing Systems*, vol. 14, 2001.
- [13] V.K. Mansinghka, "Towards Visual SLAM in Dynamic Environments," MIT 16.412J Final

Project Writeup, 2004.

- [14] Y. Ke, R. Sukthankar, "PCA-SIFT: A More Distinctive Representation for Local Image Descriptor," in Proc. Computer Vision and Pattern Recognition, vol. 1, pp. 511 - 517, 2004.
- [15] D. G. Lowe, "Local Feature View Clustering for 3D Object Recognition," in Proc. 2001 IEEE Computer Society Conf Computer Vision and Pattern Recognition, vol. 1, pp. 1-682 - 1-688, 2001.
- [16] R. Bishop, Intelligent Vehicle Technology and Trends. Artech House, 2005, pp. 25 - 37.
- [17] S. Y. Kim, S. Y. Oh, et al. "Front and rear vehicle detection and tracking in the day and night times using vision and sonar sensor fusion," in Proc. 2005 IEEE/RSJ International Conf. Intelligent Robots and Systems, pp. 2173 - 2178, 2005.

저 자 소 개



김 삼 용(학생회원)
 2000년 성균관대학교
 전자공학과 학사.
 2003년 포항공과대학교 전자전기
 공학과 석사.
 2006년 현재 포항공과대학교 전자
 전기공학과 박사과정.

<주관심분야 : 지능자동차, 신경망 및 진화연산,
 컴퓨터 비전>



오 세 영(정회원)
 1974년 서울대학교
 전자공학과 학사.
 1978년 Case Western Reserve
 Univ. 전자공학과 석사.
 1981년 Case Western Reserve
 Univ. 전자공학과 박사.

1981년~1984년 University of Illinois 조교수.
 1985년~1988년 University of Florida 조교수.
 현재 포항공과대학교 전자전기공학과 정교수.
 <주관심분야 : 신경망, 진화연산, 이동로봇, 지능
 자동차, 지능시스템>