



파티클 필터를 장착한 가중된 다중 인스턴스 학습을 이용한 전방차량 추적

Forward Vehicle Tracking Based on Weighted Multiple Instance Learning Equipped with Particle Filter

박근호* · 이준환*[†]
Keunho Park and Joonwhoan Lee[†]

*전북대학교 공과대학 컴퓨터공학부

[†]School of Computer Engineering and Science, Chonbuk National University

요약

본 논문에서는 파티클 필터를 장착하고 WMIL(Weighted Multiple Instance Learning)을 이용한 전방차량 추적 알고리즘을 제안하였다. 제안된 알고리즘에서 영상표현은 Haar-like 특징들을 사용하고 차량인식 결과는 추적하고자 하는 전방차량의 위치를 알아내는데 사용된다. 제안된 방식에서 WMIL과 파티클 필터를 결합하기 위해 기존의 외관모델을 이용한 추적에서 탐색영역에서 영상조각의 추적객체 신뢰도 맵을 계산하는 대신에 파티클 필터의 전파, 관측, 추정, 선택 그리고 분류기 훈련 등의 단계를 매 프레임 마다 순차적으로 수행하여 객체의 새로운 위치를 갱신하였다. 제안된 전방차량 추적방식은 실험을 통해 Ada-boost, MIL(Multiple Instance Learning)이나 WMIL 방법을 이용하는 추적에 비해 파티클 필터로 인해 계산량 증가는 불가피하나 추적의 질적인 정확도는 국도, 고속도로, 터널 및 시내도로 등의 실험 동영상에서 추적대상의 위치오차가 평균 4.5화소 정도로 기존의 추적 방법들에 비해 크게 개선되는 것을 확인하였다.

키워드 : 가중된 다중 인스턴스 학습, 파티클 필터, 전방차량, 객체 추적, 비디오 프레임, Haar-like 특징

Abstract

This paper proposes a novel forward vehicle tracking algorithm based on the WMIL(Weighted Multiple Instance Learning) equipped with a particle filter. In the proposed algorithm Haar-like features are used to train a vehicle object detector to be tracked and the location of the object are obtained from the recognition result. In order to combine both the WMIL to construct the vehicle detector and the particle filter, the proposed algorithm updates the object location by executing the propagation, observation, estimation, and selection processes involved in particle filter instead of finding the credence map in the search area for every frame. The proposed algorithm inevitably increases the computation time because of the particle filter, but the tracking accuracy was highly improved compared to Ababoost, MIL(Multiple Instance Learning) and MIL-based ones so that the position error was 4.5 pixels in average for the videos of national high-way, express high-way, tunnel and urban paved road scene.

Key Words : WMIL, Particle Filter, Forward Vehicle, Object Tracking, Video Frames, Haar-like Features

Received: Jan, 26, 2015
Revised : Jun, 19, 2015
Accepted: Jun, 22, 2015
[†] Corresponding author
chlee@chonbuk.ac.kr

1. 서론

오늘날 안전과 편리성에 대한 관심이 높아짐에 따라 사고가 났을 때 운전자를 안전하게 보호해 줄 수 있는 자동차뿐만 아니라 사고를 미연에 방지하며 운전할 수 있는 지능형 자동차에 대한 수요가 증가하고 있다. 이러한 지능형 자동차 안전시스템에 대한 관심이 증가하면서 국내외에서 투자와 개발 또한 증가하는 추세이다. 국외에서는 GM, 포드, 도요타, 폭스바겐, 벤츠, BMW 등 글로벌 자동차업체들은 2020년까지 충돌을 100% 방지하는 지능형 안전자동차를 목표로 개발 중이며, 기술상용화를 위해 IT 업체들과 제휴하고 있다[1]. 한편 지능형 자동차 안전시스템은 초기의 안전벨트, 에어백, 충격흡수 차체 등과 같이 자동차 사고가 발생하였을 때 피해를 최소화하기 위한 수동적인 시스템으로부터, 사고의 회피를 위한 시스템 및 차량 간 거리 측정 등을 통한 무인 자동 운행시스템 등으로 발전해 가고 있다[2].

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

본 논문에서는 이러한 자동차 안전주행을 위한 영상인식 시스템의 한 부분인 전방차량 추적 문제를 다루고 있으며, 다중 인스턴스학습(MIL: Multiple Instance Learning, 이하 MIL이라 표현) 추적방법을 개선한 가중된 다중 인스턴스학습(Weighted Multiple Instance Learning, 이하 WMIL이라 표현)에 파티클 필터(particle filter)를 장착한 효율적인 추적알고리즘을 제안하였다.

제안된 알고리즘에서 추적객체의 영상표현은 Haar-like 특징들을 사용하고 차량인식 결과는 추적하고자 하는 전방차량의 위치를 알아내는데 사용된다. 따라서 외관(appearance)모델은 Haar-like 특징들을 기반으로 매 프레임마다 영상 조각(patch)들의 추적객체의 가능성을 탐지하는 WMIL을 통해 얻은 분류기로 구성된다. 또한 움직임 모델(motion model)로는 전방차량이 대부분 영상의 중앙에 위치하기 때문에 정지된 모델에 비선형 움직임과 비 가우시언 움직임 오차 및 관측 잡음 환경에서도 예측이 가능한 파티클 필터를 적용하였다. 제안된 방식에서 WMIL과 파티클 필터를 결합하기 위해 기존의 외관모델을 이용한 추적에서 객체의 위치를 갱신하기 위해 탐색영역에서 영상조각의 추적객체 신뢰도 맵(credence map)을 계산하는 대신 파티클 필터의 전파, 관측, 추정, 선택 그리고 분류기 훈련 등의 단계를 매 프레임마다 순차적으로 수행하여 객체의 새로운 위치를 갱신하였다.

제안된 전방차량 추적은 실험을 통해 Ada-boost, MIL이나 WMIL 학습방법을 이용하는 추적에 비해 파티클 필터로 인해 계산량 증가는 불가피하나 추적의 질적인 정확도는 국도, 고속도로, 터널 및 시내도로 등의 실험 동영상에서 추적대상의 위치 오차가 평균 4.5화소 정도로 기존의 추적방법들에 비해 크게 개선되는 것을 확인하였다.

본 논문의 2장에서는 배경이론과 제안된 알고리즘을 서술하며 3장에서는 도로 동영상에서의 실험결과를 나열하였으며 4장에서는 결론과 향후과제를 정리하였다.

2. 제안된 추적알고리즘

본 논문에서 제안된 파티클 필터를 장착한 WMIL을 이용하는 추적 알고리즘은 기존의 파티클 필터를 이용한 추적방법과 WMIL을 이용하는 추적방법을 혼합한 방식이다. 즉 Ada-boost를 통한 객체검출기를 구성하는 외관모델을 이용한 추적방법을 발전시킨 MIL 추적, 또 이를 개선한 WMIL 추적방법과 칼만 필터를 개선한 몬테카를로 샘플링 방법의 파티클 필터를 적용하여 이동위치를 예측하는 방법을 적용하고 있다. 그림 1은 본 논문에서 제안하는 추적알고리즘의 발전 과정을 보여주고 있다.

본 절에서는 외관모델을 구성하는 방식의 변화 즉 Ada-boost 학습에 의한 추적방법, MIL에 의한 추적, WMIL에 의한 추적방법 및 파티클 필터를 서술하고 파티클 필터를 장착한 WMIL에 의한 추적방법을 제안한다.

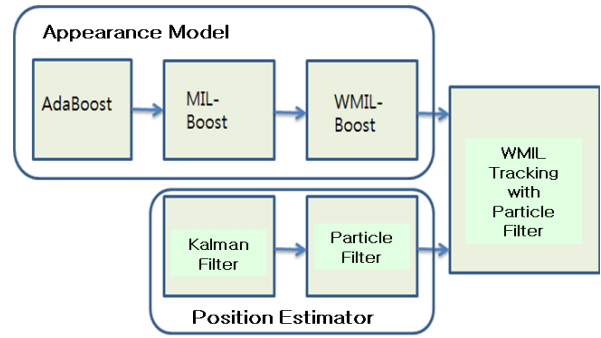


그림 1. 제안된 알고리즘과 관련된 추적알고리즘의 발전
Fig. 1. Development of tracking algorithms related with the proposed one

2.1 외관모델 학습을 이용한 추적

Zhang 등이 제안한 WMIL을 통한 외관모델 구성을 이용하는 추적방법[3]은 일종의 객체검출에 의한 추적(tracking-by-detection)방법으로 그 근원은 Grabner 등이 제안한 온라인 Ada-boost 학습방법을 이용한 객체검출기 구성 및 추적[4]으로 거슬러 올라간다. 또한 이 온라인 Ada-boost 학습방법은 Babenko 등에 의해 MIL방법으로 개선되었으며 [5] WMIL방법으로 발전되었다. 이들 방법의 공통점은 모두 Boosting 학습을 통해 추적객체의 검출기를 구성한다는 것이다. 즉 이들 방법의 핵심은 약 분류기(weak classifier)를 여러 개 조합한 강 분류기(strong classifier)의 형태로 추적할 객체를 탐지하는 검출기(또는 분류기)를 매 프레임마다 구성하고, 이를 이용하여 객체를 검출하고 추적하는데 있다.

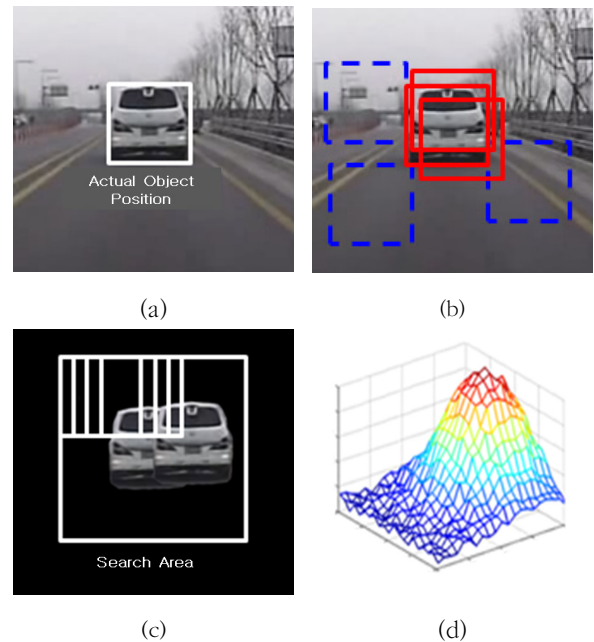


그림 2. 외관모델을 이용한 객체 추적 과정
Fig. 2. Tracking process using appearance model

그림 2는 이러한 Boosting 학습을 통한 외관모델을 이용한 추적과정을 그림으로 보여주고 있다. 그림 2(a)는 추적대상 객체의 바운딩 박스 모델이며, 초기에 객체를 인식하거나 사용자가 강제로 초기화하여 추적대상 객체를 지정한다. 그림 2(b)는 임의의 프레임에서 객체 검출을 위한 학습데이터 영상 조각을 보여주고 있으며 빨강색 박스는 긍정(positive) 샘플(또는 인스턴스), 파랑 색 박스는 부정(negative) 샘플을 표현한다. 이들 샘플들을 이용하여 Boosting 방법으로 학습한 검출기가 구성되면, 다음 프레임에서 움직임 모델을 고려하여 탐색 범위를 제한하며, 그림 2(c)와 같이 탐색 영역내의 영상 조각에 대해서 신뢰도 값을 구해 그림 2(d)와 같은 신뢰도 맵을 구성한다. 이렇게 구성된 신뢰도 맵에서 최대 값을 가지는 부분을 mean-shift 등의 알고리즘으로 찾으며, 바로 이 최대 신뢰도의 위치가 추적 대상객체의 현재 위치로 갱신된다.

여기서 영상 속성으로는 통상 다양한 사각형 패턴을 이용하여 추출하는 Harr-like 특징들을 사용한다[6][7]. 이들 특징들은 한 영상 지점까지의 화소 값을 누적한 적분영상을 이용하면 계산량을 줄이며 빠르게 추출할 수 있다.

한편 이들 방법의 차이점은 학습데이터 샘플의 구성에 따른 학습방법에 있으며, 그림 3은 이들 방법들의 학습데이터를 구성하는 방식을 보여주고 있다. 그림 3에서 Ada-boost 방법에서 빨강색 박스들과 파랑색 박스들은 각각 긍정과 부정 샘플을 의미하며 이들은 온라인 방법으로 강 분류기 객체 검출기를 학습하는데 이용된다. 그러나 만약 연속되는 검출과정에서 추적대상 객체의 위치에 오차를 포함한다면 Ada-boost 방법에서는 오차를 포함하는 부분에서 학습에 필요한 긍정과 부정 샘플을 취득하여 검출기를 학습할 수밖에 없어 오차가 연속적으로 누적될 수 있으며 결국 객체를 놓치게 되는 드리프트 현상이 발생할 수 있다. 이러한 오차의 누적 현상은 온라인 추적과정에서 대상객체의 크기가 변화되거나 또는 가려짐(occlusion)등이 발생할 경우 빈번히 일어난다.

이러한 점을 개선한 그림 3의 MIL 추적방법에서는 빨강색 타원은 적어도 하나 이상의 긍정 인스턴스를 포함하는 긍정 백(bag)을 표현하고 있으며 파랑색 점선으로 된 타원은 모조리 부정 인스턴스들을 포함하는 부정 백들을 표현한다. 또한 WMIL방법의 학습에 사용하는 샘플들은 MIL 방법과 유사하나 여기에는 추적대상의 객체와의 거리를 반영한 샘플들의 중요도를 포함하고 이를 학습에 반영한다. 그림에서 노란색 샘플은 가장 중요도가 큰 샘플을 의미한다.

Boosting 학습을 이용하는 강 분류기, 즉 외관모델은

$$H_K(x) = \ln\left(\frac{p(x|y=1)p(y=1)}{p(x|y=0)p(y=0)}\right) \quad (1)$$

로 정의된다. 또한 영상샘플 x 가 긍정일 사후확률 (1)은

$$p(y=1|x) = \sigma(H_K(x)) \quad (2)$$

와 같다. 여기서 $\sigma(z) = 1/(1+e^{-z})$ 의 시그모이드 함수이며

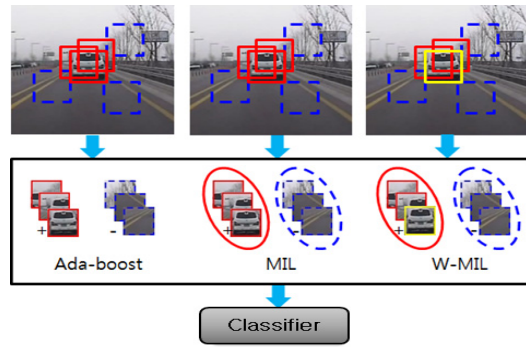


그림 3. Boosting 학습을 이용한 추적에서 학습데이터 샘플
Fig. 3. Training samples(instances) in the tracking algorithms based on boosting

$y \in \{0,1\}$ 은 객체 검출의 부정과 긍정 레이블을 의미한다. 영상샘플 x 는 속성벡터 $F(x) = (f_1(x), f_2(x), \dots, f_K(x))^t$ 로 표현되며, 사전확률 $p(y=0) = p(y=1)$ 이라면 강 분류기는

$$H_K(x) = \ln\left(\frac{p(F(x)|y=1)}{p(F(x)|y=0)}\right) = \sum_{k=1}^K h_k(x) \quad (3)$$

와 같으며, 여기서

$$h_k(x) = \ln\left(\frac{p(f_k(x)|y=1)}{p(f_k(x)|y=0)}\right) \quad (4)$$

는 특징 $f_k(x)$ 를 기반으로 영상 샘플 한 x 의 약 분류기의 분류함수 값에 해당한다. Babenko 등이 제안한 MIL Boosting에서는 y_i 로 레이블된 인스턴스 집합으로 이루어진 i 번째 백을 $X_i = \{x_{i1}, \dots, x_{im}\}$ 라고 할 때 데이터 집합을 $\{(X_1, y_1), \dots, (X_k, y_k)\}$ 으로 정의한다. 따라서 식 (2)와 같이 인스턴스의 확률은 $p_{ij} = \sigma(H_K(x_{ij}))$ 로 계산되며 여기서 강 분류기 $H_K(\cdot)$ 는 식 (3)과 같이 영상의 Haar-like 속성 $f_k(\cdot)$ 과 관련된 약 분류기 $h_k(\cdot)$ 들을 조합하여 구성된다. 식 (4)의 $h_k(\cdot)$ 의 조건부 확률들은 가우시언 모델을 가정하여 $p(f_k(x_{ij})|y_i=1) \sim N(\mu_1, \sigma_1)$ 또는 $p(f_k(x_{ij})|y_i=0) \sim N(\mu_0, \sigma_0)$ 로 가정한다. 여기서 약 분류기를 위한 조건부 확률모델의 파라미터 μ_1 과 σ_1 은

$$\mu_1 \leftarrow \eta \mu_1 + (1-\eta) \frac{1}{N} \sum_{j|y_i=1} f_k(x_{ij}) \quad (5)$$

$$\sigma_1 \leftarrow \eta \sigma_1 + (1-\eta) \sqrt{\frac{1}{N} \sum_{j|y_i=1} (f_k(x_{ij}) - \mu_1)^2} \quad (6)$$

로 갱신된다. 여기서 N 은 긍정 인스턴스의 개수를 의미하며 η 는 학습률에 해당하며 μ_0 와 σ_0 역시 비슷한 방법으로 갱신된다. 또한 i 번째 백의 확률은 인스턴스의 확률을 식 (2)에

의해 계산하고 Noisy-OR 방법으로

$$p_i = 1 - \prod_j (1 - p_{ij}) \quad (7)$$

와 같이 총체화 된다. 식 (7)에서 모든 긍정 인스턴스들은 동일한 중요도를 가지고 총체화됨을 주목할 필요가 있다.

표 1은 추적에 활용되는 MIL알고리즘을 의사코드로 표현한 것이다. MIL 추적에서는 $\Phi = \{h_1, \dots, h_M\}$ 의 약 분류기 풀을 관리하며 그 중에 K 개의 약 분류기를 다음의 평가지표에 따라 탐욕스럽게 약 분류기를 추가해 나간다.

$$h_k = \arg \max_{h \in \Phi} l(H_{k-1} + h) \quad (8)$$

식 (8)에서 $l = \sum_i (y_i \log p_i + (1 - y_i) \log (1 - p_i))$ 는 로그우도(log likelihood) 함수이며 $H_{k-1} = \sum_{m=1}^{k-1} h_m$ 은 $(k-1)$ 개의 약 분류기로 구성된 강 분류기를 의미한다. 표 1에서 N 과 L 은 각각 긍정 인스턴스와 부정 인스턴스의 수를 의미한다.

표 1. MIL 추적의 학습 알고리즘
Table 1. Training algorithm for MIL tracking

<p>INPUT: Set of Bags $\{X_i, y_i\}$ containing $X_i = \{x_{i1}, x_{i2}, \dots\}$, $y_i \in \{0, 1\}$</p> <p>1: Update M weak classifiers using $\{x_{ij}, y_i\}$</p> <p>2: for all i, j, initialize $H_{i,j} = 0$</p> <p>3: for $k = 1$ to K do</p> <p>4: set $l^m = 0, m = 1, \dots, M$</p> <p>5: for $m = 1$ to M do</p> <p>6: for $i = 0$ to 1 do</p> <p>7: for $j = 0$ to $N + L - 1$ do</p> <p>8: $p_{ij}^m = \sigma(H_{ij} + h_m(x_{ij}))$</p> <p>9: end for</p> <p>10: $p_i^m = 1 - \prod_j (1 - p_{ij}^m)$</p> <p>11: $l^m = l^m + y_i \log(p_i^m) + (1 - y_i) \log(1 - p_i^m)$</p> <p>12: end for</p> <p>13: end for</p> <p>14: $m^* = \operatorname{argmax}_m l^m$</p> <p>15: $h_k(x_{ij}) \leftarrow h_{m^*}(x_{ij})$</p> <p>16: $H_{ij} = H_{ij} + h_k(x_{ij})$</p> <p>17: end for</p> <p>Output: When $p(y = 1 \cdot) = \sigma(h(\cdot))$, strong classifier $H_K(x_{ij}) = \sum_{k=1}^K h_k(x_{ij})$</p>
--

2.2 WMIL 학습을 이용하는 추적

MIL 추적에서는 긍정 인스턴스들의 중요도를 따지지 않고

즉 모든 긍정 인스턴스의 중요도는 동일하다고 가정하고 강 분류기를 훈련한다. 하지만 긍정 인스턴스의 경우 추적 객체와 거리가 가까운 인스턴스일수록 객체와 비슷한 특징을 가지기 때문에 학습 할 때 그림 3에서와 같이 더 높은 중요도 즉 가중치를 부여하면 학습된 강 분류기는 추적대상 객체를 검지하는데 효율적이다[3]. Zhang 등은 이점에 착안하여 WMIL 학습 알고리즘을 이용한 객체 추적방법을 제안하였다. WMIL 알고리즘에서 영상에서 긍정 인스턴스의 위치를 나타내는 위치함수를 $l(\cdot) \in R^2$ 이라고 하고 c 를 정규화 상수(normalization constant)라고 할 때 인스턴스의 가중치는

$$w_{ij} = \begin{cases} \frac{1}{c} e^{-|l(x_{ij}) - l(x_{10})|} & \text{for } i = 1 \\ 1 & \text{otherwise} \end{cases} \quad (9)$$

과 같이 정의된다. 식 (9)에서 x_{10} 는 현재 프레임의 위치에서 샘플링한 긍정 인스턴스를 의미하고 x_{ij} 는 j 번째 긍정 인스턴스를 의미한다. 또한 $i = 0$ 인 부정 인스턴스 x_{0j} 의 가중치는 1로 고정된다.

WMIL 추적에서 약 분류기는 MIL 추적에서와 같이 가우시안 모델을 활용하나 현재 프레임에서 영상 속성이 모델 파라미터에 기여하는 바는 현재 프레임까지의 속성을 고려한 최대 우도 추정을 통해 평균 μ_1 의 경우는 식 (5)와 같고 표준편차는

$$\sigma_1 \leftarrow \sqrt{\eta \sigma^2 + (1 - \eta) \frac{1}{N} \sum_{j, y_j = 1} (f_k(x_{ij}) - \bar{\mu})^2 + \eta(1 - \eta)(\mu_1 - \bar{\mu})^2} \quad (10)$$

와 같은 방식으로 갱신된다[3]. 여기서

$$\bar{\mu} = \frac{1}{N} \sum_{j, y_j = 1} f_k(x_{ij}) \quad (11)$$

로 현재 프레임에서 긍정 인스턴스의 k 번째 속성의 평균을 의미한다. Zhang 등의 WMIL 추적 학습 알고리즘은 원리적으로 표 1의 라인 10의 Noisy-OR에 의한 백 확률 총체화 부분이

$$p_i^m = 1 - \prod_j (1 - p_{ij}^m w_{ij}) \quad (12)$$

로 바뀌는 것을 제외하고는 표 1의 MIL 추적과 원리적으로 동일하다. 다만 Zhang 등은 표 1의 라인 8과 라인 10의 인스턴스와 백의 확률을 구하는 대신에 인스턴스의 가중치를 고려하고 $(k-1)$ 번째까지의 선택된 약 분류기를 포함하는 강 분류기 H_{k-1} 의 로그 우도의 그래디언트를 사용하여 최대 우도를 가지는 약 분류기를 선택하여 계산량을 줄이려 하고 있다.

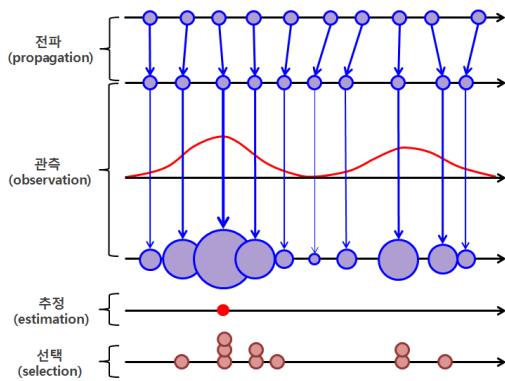


그림 4. 파티클 필터의 진행과정
Fig. 4. Process stages of particle filter

2.3. 파티클 필터를 이용한 추적

파티클 필터[8][9]는 확률적인 문제해결에 있어서 확률 분포함수 대신에 샘플링을 이용하는 몬테카를로 방법의 목적분포로 샘플링을 수행하고 샘플의 중요성을 평가하는 함수에 따라 재 샘플링(importance resampling)을 수행하여 통계적 추론문제를 해결한다. 추적 시스템에서는 적절한 목적분포(proposal distribution)로 랜덤하게 생성된 샘플(파티클)을 가정하고 그 위치에서 추적객체가 존재할 가능성을 중요도로 하여 샘플 수를 늘리거나 줄이는 재 샘플링을 수행하여 시스템의 정보 즉 추적하고자 하는 객체의 위치를 추정하는 방법으로 샘플의 수가 충분하다면 베이저인 최적 추정치에 근접한다. 파티클 필터는 비선형 움직임 모델이나 움직임과 관측 오차에 대한 확률분포(probability distribution)에 대한 제약이 없기 때문에 선형 모델에 가우시안을 가정한 칼만 필터[10]보다 정확하다.

파티클 필터의 수행은 그림 4와 같이 전파(propagation), 관측(observation), 추정(estimation), 그리고 선택(selection) 네 단계로 나뉜다. 파티클 필터에서 시간 t 에서의 확률분포는 관측으로부터 얻어진 가중치를 가지는 N 개의 파티클 $P = \{(x_t^{(n)}, w_t^{(n)}) | n = 1, \dots, N\}$ 로 근사화 한다. 여기서 $x_t^{(n)}$ 은 시간 t 에서 파티클의 상태를 나타내는 벡터이며, $w_t^{(n)}$ 은 시간 t 에서 파티클의 가중치이다. 전파 단계는 시간 t 에서의 파티클들을 움직임 모델과 움직임 오차를 가정하여 시간 $t+1$ 에서의 파티클들로 전파하는 단계이다. 본 논문에서는 추적객체가 정지되어 있다고 가정하며 목적분포는 평균이 0이며 분산이 1화소 거리의 정규 분포로 설정하였다. 관측 단계는 파티클과 추적 대상객체의 유사도를 구하여 이에 따른 가중치를 부여하는 단계이다. 각 파티클에 대한 가중치는

$$w_t^{(n)} = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{d_t^2}{2\sigma^2}} \quad (13)$$

의 가우시안 형태의 함수를 사용하여 계산된다. 여기서 σ 는 가우시안 함수의 표준편차이며, d_t 는 파티클과 대상객체

의 바타차라(Bhattacharyya) 거리로

$$d_t = \sqrt{1 - \rho(p, q)} \quad (14)$$

로 계산된다. 식 (14)에서 $\rho(p, q)$ 는 바타차라 계수[11]로 파티클과 대상객체의 특징함수인 $p(k)$ 와 $q(k)$ 로부터

$$\rho(p, q) = \sum_k \sqrt{p(k)q(k)} \quad (15)$$

와 같이 구하며, 식 (13)의 가중치는

$$w_t^{(n)} = \frac{w_t^{(n)}}{\sum_{n=1}^N w_t^{(n)}} \quad (16)$$

과 같이 정규화(normalization) 과정을 거친다. 추정 단계는 전파 단계와 관측 단계에서 구한 파티클의 상태와 가중치를

$$x_t = \sum_{n=1}^N w_t^{(n)} x_t^{(n)} \quad (17)$$

과 같이 가중치를 이용하여 선형 결합하고 추적하고자 하는 객체의 위치를 구하는 단계로, 여기서 x_t 는 시간 t 에서 추정된 객체의 위치이다. 선택 단계에서는 다음번 전파, 관측, 그리고 추정의 일련에 과정에서 사용할 새로운 파티클들을 재 샘플링(re-sampling) 해주는 단계이다. 관측 단계에서 구한 가중치에 비례하여 새로운 파티클들을 무작위로 선택하기 때문에 가중치가 큰 이전 파티클들과 같은 위치에서 새로운 파티클들이 다른 위치보다 많이 생성된다.

2.4 제안된 파티클 필터를 장착한 WMIL-Boost 추적

본 논문에서는 WMIL과 파티클 필터를 결합한 객체추적 알고리즘을 제안한다. 제안된 추적방법에서 WMIL은 추적대상의 온라인 외관모델을 구성하기 위한 분류기 학습에 사용된다. 즉 현재 추적대상 객체의 위치에서 얻은 긍정 인스턴스들과 식 (9)의 거리에 따른 가중치 포함하는 긍정 백과 각각의 부정 인스턴스에 해당하는 부정 백들을 구성하여 WMIL 방법으로 약 분류기를 갱신하고 선택하는데 활용된다. 반면에 제안된 방법에서 파티클 필터는 현재 프레임에서 선택된 약 분류기들을 이용하여 최초 프레임의 추적대상을 포함하는 영상 조각과 각 파티클 위치에서의 영상 조각들을 비교하여 식 (13)의 가중치를 부여하는데 사용된다.

본 논문에서 제안하는 객체추적 알고리즘은 표 2와 같다. 제안된 알고리즘은 블랙박스에 연결된 카메라 등을 통해 촬영된 전방차량의 동영상과 초기 차량이 위치를 입력받아 매 프레임 전파, 관측, 추정, 선택단계를 거치며 객체의 위치 x_t 를 출력한다.

표 2. 제안된 WMIL과 파티클 필터를 결합한 객체추적 알고리즘

Table 2. Proposed object tracking algorithm using WMIL and particle filter

<p>INPUT: Input image sequence from $t=1$ to T $\{I_t\}_{t=1}^T$, and object position x_1 at $t=1$</p> <ol style="list-style-type: none"> 1: Set of positive instances centered at x_1 in I_1(positive bag) S_1^+ and set of negative instances(negative bag) S_1^-. 2: WMIL training for strong classifier $H_{K,1}$ by Haar-like features of S_1^+ and S_1^- 3: Initialize particle set $P_1 = \{x_1^{(n)} n=1,\dots,N\}$ around x_1 by uniform random numbers 4: Select object s^{target} at x_1 in I_1 5: for $t=2$ to T do 6: Propagation: Update particle set from $\{x_{t-1}^{(n)}\}_{n=1}^N$ to $\{x_t^{(n)}\}_{n=1}^N$ 7: Observation: Calculate weights $\{w_t^{(n)}\}_{n=1}^N$ of particles by Har-like feaures of s^{target} and $\{x_t^{(n)}\}_{n=1}^N$ 8: Estimation: Estimate object position x_t using linear combination of $\{x_t^{(n)}\}_{n=1}^N$ and $\{w_t^{(n)}\}_{n=1}^N$ 9: Get the set of positive(negative) instances $S_t^+(S_t^-)$ around x_t in I_t 10: WMIL training for strong classifier $H_{K,t}$ by Haar-like features of S_t^+ and S_t^- 11: Selection: Get new particle set P_t using $\{x_t^{(n)}\}_{n=1}^N$ and $\{w_t^{(n)}\}_{n=1}^N$ 12: end for <p>OUTPUT: Set of postions of object at t $\{x_t\}_{t=1}^T$</p>

전과단계는 기존의 파티클 집합 $\{x_t^{(n)}\}_{n=1}^N$ 에 움직임 모델과 목적 확률분포에 해당하는 움직임의 불확실성을 결합하여 갱신하는 단계이다. 본 연구의 추적 대상인 전방차량의 경우 대부분 영상의 중앙에 머물러 있기 때문에 움직이지 않는다고 가정하였으며, 움직임의 불확실성은 좌우, 상하로 $N(0,1)$ 의 정규분포로 외란이 작용하여 다음 프레임에서의 파티클이라 가정한다.

관측단계에서는 최초 프레임에서의 추적대상 객체 영상조각 s^{target} 과 파티클 위치에서의 영상조각 $r(x_t^{(n)})$ 을 WMIL로 선택된 약 분류기 관점에서 식 (15)의 거리를 구한다. 즉 표 1에서 강 분류기 $H_k(\cdot)$ 대신에 이를 구성하는 선택된 $h_k(\cdot)$ 를 기반으로 $p(k) = h_k(s^{target})$ 과 $q(k) = h_k(r(x_t^{(n)}))$ 를 구하여 이들 값들의 식 (15)의 거리를 구하였다. 선택되는 약 분류기의 개수는 식 (15)에서 바타차라 거리를 구하는데

사용되는 벡터들의 차원과 같다. 본 논문에서 제안된 방법은 Song 등이 제안한 MIL 학습과 파티클 필터를 결합한 방법 [12]과 차이가 있다. 즉 그들 방법에서는 파티클 위치에서 단순히 강 분류기 $H_k(\cdot)$ 로 파티클의 가중치를 얻어서 추적대상 객체의 변화에 적응하도록 하였다. 하지만 이들 방법은 객체가 가려짐 등에 의해 사라졌다가 나타나는 경우 추적대상 객체의 외관 모델을 잃어버릴 수 있다. 따라서 본 논문에서는 초기 추적대상 객체와 파티클 위치의 영상객체를 온라인 학습으로 얻은 우수한 약 분류기를 관점에서 비교함으로써 가려짐 등에 의한 문제를 해결하려 시도하였다.

제안한 방법의 추정단계에서는 전과단계와 관측단계에서 구한 파티클들의 (x,y) 좌표 $\{x_t^{(n)}\}_{n=1}^N$ 는 식 (17)과 같이 선형결합하여 객체의 위치 x_{t+1} 을 추정하였다. 선택단계에서는 다음번 전과, 관측, 그리고 추정의 일련에 과정에서 사용할 새로운 파티클 집합 $P = \{x_t^{(n)}|n=1,\dots,N\}$ 을 가중치를 바탕으로 다시 샘플링(re-sampling)하는 단계이다. 관측 단계에서 구한 가중치에 비례하는 확률로 새로운 파티클들을 무작위로 선택하기 때문에 가중치가 큰 이전 파티클들과 같은 위치에서 새로운 파티클들이 다른 위치에 비해 많이 생성된다. 제안된 방법의 실험에서 파티클의 개수는 500개로 한정하였다.

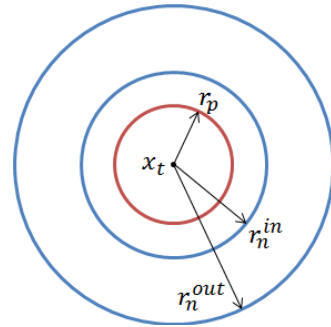


그림 5. 긍정 인스턴스와 부정 인스턴스의 채취 영역
 Fig. 5. Regions for sampling the positive and negative instances

제안된 추적방법의 실험에서 학습을 위한 긍정 인스턴스와 부정 인스턴스의 샘플 추출은 그림 5와 같은 방법으로 진행된다. 표 2의 단계 2와 9의 긍정 인스턴스들은 그림 5의 x_t 를 중심으로 $r_p = 8$ 이내의 영역의 모든 화소를 중심으로 바운딩 박스 크기의 영상조각을 채취하였으며 부정 인스턴스 들은 $r_n^{in} \leq r \leq r_n^{out}$ 범위의 무작위로 추출된 65개의 화소를 중심으로 하는 영상조각들이다. 본 논문에서 긍정 백은 하나이며, 부정 백은 부정 인스턴스 수만큼을 가정하였다. 또한 WMIL 학습에서 Haar-like 속성 즉 약 분류기는 150개를 추출하였으며 이 중에 매 프레임 15개의 약 분류기가 선택되어 식 (13)의 바타차라 거리를 계산하는데 이용되었다. 식 (13)에서 $\sigma^2 = 1$ 로 하였다.

3. 제안된 알고리즘을 이용한 전방차량 추적 실험

본 논문에서는 전방차량 동영상에 대해 Ada-boost, MIL, WMIL 그리고 제안된 추적 알고리즘을 적용하여 성능을 관찰하였다. 실험은 3.3GHz의 CPU와 8GB의 RAM의 PC의 환경에서 수행하였다. 그림 6은 이들 알고리즘을 이용한 추적 실험의 사용자 인터페이스를 보여주고 있다.

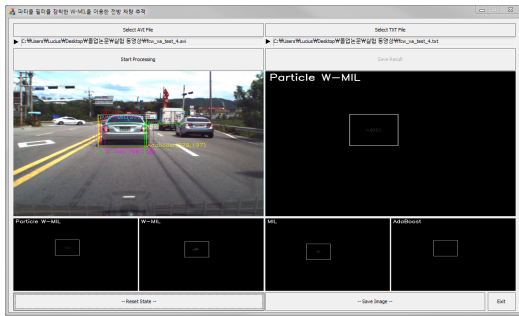
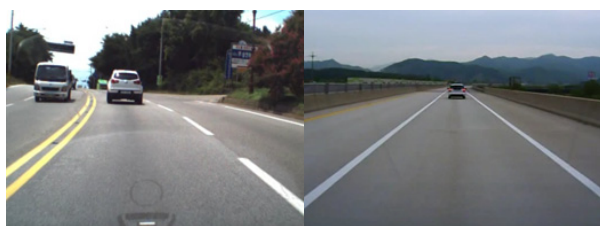


그림 6. 제안된 알고리즘 실험 인터페이스

Fig. 6. Interface for the experiment of proposed algorithm

3.1 실험 동영상 및 성능측정 방법

실험에 사용된 동영상은 각각 다른 환경에서 촬영된 블랙박스 영상이다. 그림 7(a)는 국도주행에서 촬영한 영상이다. 국도에서 차량은 빠른 속도로 이동하고 도로가 구불구불하며 반대편 차선의 차량통행이 많기 때문에 알고리즘의 복잡한 환경에서 객체를 추적하는 능력을 평가할 수 있다. 그림 7(b)는 고속도로에서의 주행을 촬영한 영상이다. 고속도로에서는 차간 거리를 길게 유지해야 하므로 알고리즘의 작은 객체를 추적하는 능력을 평가할 수 있다. 그림 7(c)는 시내 도로에



(a) National Highway

(b) Expressway



(c) Urban Road

(d) Tunnel

그림 7. 실험에서 사용된 동영상

Fig. 7. Video images used for experiment

서의 주행을 촬영한 영상이다. 시내 도로는 차간 거리가 비교적 가깝기 때문에 알고리즘의 큰 객체를 추적하는 능력을 평가하기 할 수 있으며 작은 객체 움직임의 변화가 예상되었다. 그림 7(d)는 터널내부에서의 주행을 촬영한 영상이다. 터널내부는 어둡기 때문에 전방 차량이 다른 환경에 비해 잘 보이지 않는다. 따라서 알고리즘이 조명환경이 변화된 어두운 환경에서 객체를 추적하는 능력을 평가할 수 있다. 실험에 사용된 동영상들은 초당 30프레임의 재생속도이며 재생시간은 국도 32초, 고속도로 51초, 시냇길 31초, 마지막으로 터널 26초이다. 각 영상프레임의 해상도는 640 × 480이다.

본 연구의 성능평가는 추적 알고리즘의 속도와 정확성 측면에서 측정하고 비교하였다. 추적알고리즘의 속도는 초당 처리 프레임의 속도(fps:frame per second)를 측정하여 비교하였다. 정확도는 사람이 직접 동영상 시스의 매 프레임을 눈으로 보고 전방 차량의 위치를 표시한 실측 정보(ground truth)의 바운딩 박스의 중앙값과 알고리즘에 의해 추정된 위치의 바운딩 박스의 중앙값의 유클리디언 거리(euclidean distance)를 측정하였다.

3.2 실험 및 검토

본 논문의 추적 알고리즘의 속도는 실험 동영상 각각에 대하여 10번의 추적을 수행하고, 이들의 평균 fps를 계산하였다. 측정 결과 표 3과 같이 Ada-boost, MIL, WMIL, 그리고 제안된 알고리즘으로 넘어갈수록 알고리즘에 새로운 개념들이 추가되어 점점 속도가 느려지는 것을 볼 수 있다.

표 3. 추적 알고리즘의 속도 비교(fps)

Table 3. Comparison of algorithms in terms of processing speed(fps)

	Adaboost	MIL	WMIL	Proposed
N. highway	61.41	28.30	25.40	22.90
Ex. way	60.30	28.97	25.01	20.71
U. road	65.73	28.40	24.39	22.71
Tunnel	62.07	26.47	24.86	19.18
Average	62.38	28.03	24.92	21.37

실험 동영상 각각에 대하여 추적 시작으로부터 200 프레임 동안 실제 객체 바운딩 박스의 중심과 추적된 바운딩 박스의 중심의 유클리디언 거리를 측정된 값을 그래프로 나타낸 것은 그림 8과 같다. 실험에 사용된 알고리즘들은 전부 확률기반 추적 알고리즘이다. 따라서 같은 영상으로 실험을 실시하여도 추적결과가 다를 수 있다. 따라서 다양한 결과를 보여주기 위해 각각의 도로환경에서 세 번씩 실험을 실시하였다.

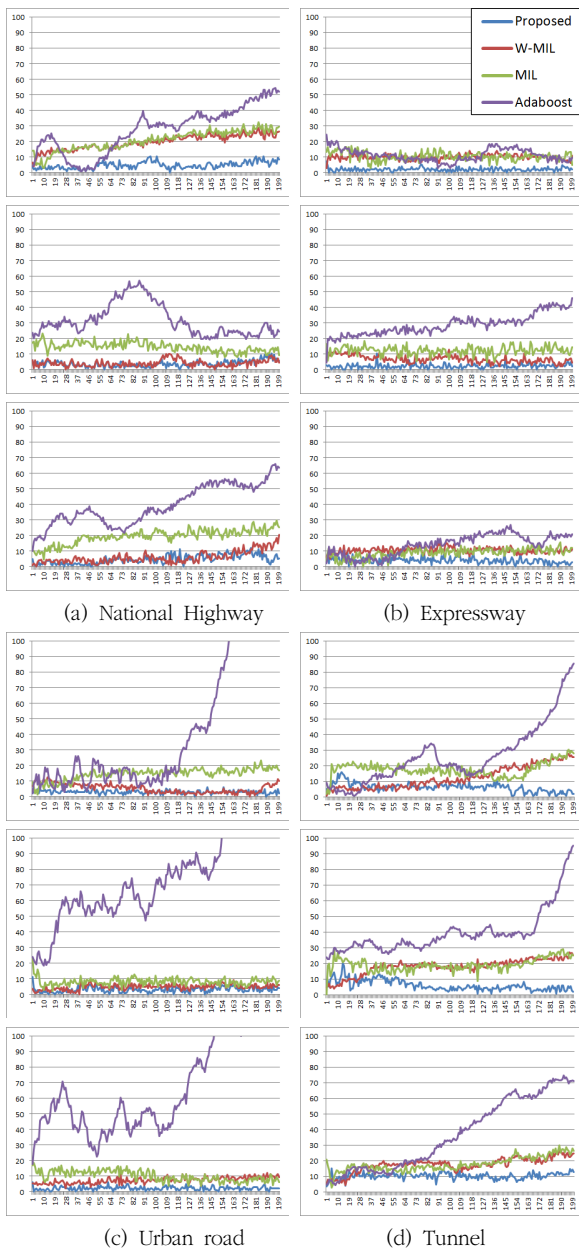


그림 8. 실험영상의 추적오차
Fig. 8. Tracking error in the experiment videos

국도의 경우 도로의 곡률 정도가 심하고 전방 차량의 이동이 심하여, 기존의 추적알고리즘의 추적결과가 정확하지 않은 경우도 있다. 하지만 제안된 알고리즘의 결과는 이들에 비해 양호한 것을 볼 수 있다. 고속도로의 경우 관찰환경과 전방 차량의 상대가 안정적 이므로 모든 알고리즘이 정확한 추적을 하는 것을 볼 수 있다. 시내도로의 경우 신호등의 영향으로 영상에서 전방 차량의 크기변화 및 가려짐 등이 심하다. 따라서 백의 개념을 사용하지 않는 Ada-boost 알고리즘의 경우 드리프트가 심하게 발생하나, 학습 샘플이 백의 개념인 다른 MIL 알고리즘의 경우 안정적으로 추적을 하는 것을 볼 수 있다. 터널 동영상의 경우 어두운 환경의 영향으로 입

력영상의 잡음이 심하고, 영상의 강도(intensity)가 전체적으로 낮게져 Haar-like 특징 추출이 다른 환경에 비해 용이하지 못하다. 따라서 기존의 학습 알고리즘들의 결과는 다른 환경에 비해 다소 떨어지는 편이지만, 예측필터로서 파티클 필터를 결합한 제안된 알고리즘은 안정적인 추적을 해냄을 볼 수 있다.

표 4는 그림 8의 그래프의 데이터를 전체 프레임에 대해 평균을 계산한 결과이다. 계산 결과 제안된 알고리즘은 프레임 당 평균 4.5 화소의 오차로 가장 정확한 결과를 제공하는 것을 알 수 있다.

표 4. 추적 알고리즘의 평균오차(유클리디안 화소거리)
Table 4. Average tracking error for tracking algorithm(Euclidian pixel distance)

	Adaboost	MIL	W-MIL	Proposed
N. highway	33.0	18.3	10.0	4.5
Ex. way	18.1	10.7	9.1	2.8
U. road	65.2	11.0	5.9	2.8
Tunnel	34.5	18.5	16.2	7.7
Average	37.7	14.6	10.3	4.5

4. 결론

본 논문에서는 자동차 안전주행을 위한 영상인식 시스템의 한 부분인 전방차량 추적 문제를 다루고 있으며, 파티클 필터를 장착하고 WMIL을 이용한 효율적인 추적알고리즘을 제안하였다. 제안된 알고리즘에서 영상표현은 Haar-like 특징들을 사용하고 차량인식 결과는 추적하고자 하는 전방차량의 위치를 알아내는데 사용된다. 외관모델은 Haar-like 특징들을 기반으로 매 프레임 마다 WMIL을 통해 영상 조각의 추적객체의 가능성을 탐지하는 분류기로 구성된다. 또한 움직임 모델로는 전방차량이 대부분 영상의 중앙에 위치하기 때문에 정지된 모델에 비선형 움직임과 비 가우시언 움직임 오차 및 관측 잡음환경에서도 예측이 가능한 파티클 필터를 적용하였다. 제안된 방식에서 WMIL과 파티클 필터를 결합하기 위해 기존의 외관모델을 이용한 추적에서 객체의 위치를 갱신하기 위해 탐색영역에서 영상조각의 추적객체 신뢰도 맵(credence map)을 계산하는 대신 파티클 필터의 전파, 관측, 추정, 선택 그리고 분류기 훈련 등의 단계를 매 프레임 마다 순차적으로 수행하여 객체의 새로운 위치를 갱신하였다.

제안된 알고리즘을 이용한 전방차량 추적은 실험을 통해 Ada-boost, MIL이나 WMIL 학습방법을 이용하는 추적에 비해 파티클 필터로 인해 계산량 증가는 불가피하나 추적의 정확도는 국도, 고속도로, 터널 및 시내도로 등의 실험 동영상에서 추적대상의 위치오차가 평균 4.5화소 정도로 기존의 추적방법들에 비해 크게 개선되는 것을 확인하였다.

특히 제안된 시스템은 파티클 필터의 가중치 계산을 위해

초기 추적대상 객체를 보존하고 이용하기 때문에 매 프레임 추적대상 객체를 변경하는 방식에 비해 객체의 가려짐 등에 강인할 것으로 예상된다.

References

[1] H. S. Chun, *Development Trends of Intelligent Vehicle Safety Technology*, Electronics and Telecommunications Research Institute, 2012.

[2] S. H. Yoon, and others, *Trends in Image Recognition SoC for Smart Vechicles*, Electronics and Telecommunications Research Institute, 2012.

[3] B. Babenko, M. H. Yang, and S. Belongie, "Visual Tracking with Online Multiple Instance Learning," *IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL, pp. 983-990, 2009.

[4] P. Viola, M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 511-518, 2001.

[5] G. J. Edwards, C. J. Taylor, and T. F. Cootes, "Interpreting Face Images using Active Appearance Models," *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, Nara, pp. 300-305, 1998.

[6] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, Issue. 6, pp. 681-685, 2001.

[7] H. Grabner, M. Grabner, and H. Bischof, "Real-time Tracking via On-line Boosting," *In Proceedings British Machine Vision Conference (BMVC)*, Vol. 1, pp. 47-56, 2006.

[8] K. Zhang, and H. Song, "Real-time visual tracing via online weighted multiple instance learning," *Pattern Recognition*, Vol. 46, Issue. 1, pp. 397-411, 2013.

[9] P. Viola, J. C. Platt, and C. Zhang, "Multiple Instance Boosting for Object Detection," *Advances in Neural Information Processing Systems*, pp. 1417-1426, 2007.

[10] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Journal of Fluids Engineering*, Vol. 82, Issue. 1, pp. 35-45, 1960.

[11] M. Isard, and A. Blake, "Condensation-Conditional Density Propagation of Visual Tracking," *International*

Journal of Computer Vision, Vol. 29, Issue. 1, pp. 5-28, 1998.

[12] F. Gustafsson, "Particle Filter Theory and Practice with Positioning Applications," *Aerospace and Electronic Systems Magazine, IEEE*, Vol. 25, Issue 7, pp. 53-82, 2010.

[13] E. S. Chio, C. H. Lee, "Feature Extraction Based on The Bhattacharyya Distance," *Pattern Recognition*, Vol. 36, Issue 8, pp. 1703-1709, 2003.

[14] Y. Song, Q. Li, "Visual Tracking Based on Multiple Instance Learning Particle Filter," *Proceedings of the 2011 IEEE International Conference on Mechatronics and Automation*, pp 1063-1067, August 7-10, Beijing, China.

저 자 소 개



박근호(Keunho Park)

2013년 : 전북대학교 컴퓨터공학부 공학사
2013년 ~ 현재 : 전북대학교 컴퓨터공학부 석사과정

관심분야 : 영상처리, 컴퓨터비전, 인공지능
Phone : +82-10-2446-1670
E-mail : khpark0224@gmail.com



이준환(Joonwhoan Lee)

1980년 : 한양대학교 전자공학과 공학사
1982년 : 한국과학기술원 전자공학과 공학석사
1990년 : 미국 미주리대학 전기 및 컴퓨터공학과 공학박사
1990년~현재 : 전북대학교 컴퓨터공학부 교수

관심분야 : 영상처리, 컴퓨터 비전, 인공지능
Phone : +82-63-270-2406
E-mail : chlee@chonbuk.ac.kr