

증강현실 환경에서의 강건한 카메라 추적을 위한 실시간 입자 필터링 기법

이석한*

요약

본 논문에서는 증강현실 환경에서 보다 강건한 카메라 정보 추정을 위한 입자필터 기반의 카메라 추적 기법에 대해서 설명한다. 실시간 카메라 추적을 위해서는 일반적으로 칼만 필터, 또는 확장 칼만 필터 등이 많이 이용되지만, 카메라의 급격한 흔들림 및 장면의 가려짐 등과 같은 불안정한 조건에서는 정상적인 카메라 추적이 매우 힘들다는 단점이 있다. 본 논문에서 제안하는 입자필터링 기법은 시스템 상태에 대한 추정 표본입자의 가중치를 별도의 가중치 계산과정을 이용하지 않고 가우스 분포를 기반으로 계산하였으며, 카메라 입자를 수렴시키기 위한 별도의 처리과정을 거치지 않고 시스템의 실제 불확실도에 근사화되도록 재표본화된 표본입자 집합을 이용하여 카메라 상태의 추정을 수행하였다. 또한 제안된 방법은 보다 많은 수의 표본 입자를 이용하는 환경에서도 실시간 처리가 가능한 장점이 있다. 실험을 통하여 다양한 환경 하에서 제안된 방법의 효율성과 정확성을 확인하였다.

A Real-time Particle Filtering Framework for Robust Camera Tracking in An AR Environment

Seok-Han Lee*

Abstract

This paper describes a real-time camera tracking framework specifically designed to track a monocular camera in an AR workspace. Typically, the Kalman filter is often employed for the camera tracking. In general, however, tracking performances of conventional methods are seriously affected by unpredictable situations such as ambiguity in feature detection, occlusion of features and rapid camera shake. In this paper, a recursive Bayesian sampling framework which is also known as the particle filter is adopted for the camera pose estimation. In our system, the camera state is estimated on the basis of the Gaussian distribution without employing additional uncertainty model and sample weight computation. In addition, the camera state is directly computed based on new sample particles which are distributed according to the true posterior of system state. In order to verify the proposed system, we conduct several experiments for unstable situations in the desktop AR environments.

Keywords: Camera tracking(카메라추적), augmented reality(증강현실) 3D reconstruction (3차원 재구성)

1. 서론

카메라를 이용하여 취득한 실사 영상에 부가적인 정보를 융합한 다음 이를 사용자에게 제공함으로써 보다 새로운 형태의 컴퓨팅 환경을 구현하고자 하는 증강현실 (AR; Augmented Reality)은 직관적 인터페이스 (perceptual interface), 유비쿼터스 컴퓨팅(ubiquitous computing) 등에 대한 응용 가능성으로 인하여 최근 그 중요성이 강조되고 있다. 일반적인 의미의 증강현실은 사용자에게 제공될 부가적인 정보가 실사 영상에 정합

※ 제일저자(First Author) 이석한
접수일:2010년 12월 11일, 수정일:2010년 12월 29일,
완료일:2010년 12월 30일
* 중앙대학교 첨단영상대학원
weid231@gmail.com
■ 이 논문은 중앙대학교 2009년도 (신진우수연구자/박사후연수과정) 지원 사업에 의한 것임

된 형태로 구현되며, 따라서 사용자는 실제 환경과 가상의 환경이 혼합된 형태의 인터페이스를 통하여 실제 환경과 연계된 상호작용을 수행할 수 있다. 2차원 영상만을 이용하는 시각 기반의 증강현실 구현에 있어서 가장 핵심이 되는 기술은 실시간으로 카메라의 자세 정보를 추정하고 이를 기준 좌표계와 정합시키는 기술이며, 증강현실 구현에 필요한 다른 모든 요소들은 안정적으로 구현된 정합 기법을 토대로 하여 구현된다. 또한 사용자의 위치 및 시선 방향은 한 곳에 고정되어 있지 않고 동적으로 이동하기 때문에 사용자가 자유롭게 움직이는 경우까지 고려된 동적인 환경 하에서의 정합까지 반드시 고려되어야 한다. 본 논문에서는 강건한 카메라 추적 및 좌표계 정합을 위한 입자필터(PF, Particle Filter) 기반의 실시간 카메라 추적 기법에 대해서 설명한다. 3장에서 카메라 추적 시스템을 위한 불확실도(uncertainty) 모델과 이를 기반으로 입자 필터링, 그리고 예측/수정 시스템(prediction/correction system)을 이용한 카메라 추적 모델 및 과정에 대해서 설명한다. 4장에서 설명된 이론적 배경을 바탕으로, 입자 필터링 시스템을 구현하고, 이에 대한 카메라 추적 성능 및 특성을 분석한다.

2. 관련 연구

카메라 추적을 위한 기존의 방법들은 주로 비선형 최적화 등과 같이 막대한 계산량을 필요로 하는 복잡한 과정을 거쳐서 카메라 추적을 수행하기 위한 연구가 많이 진행되었다. 하지만 이러한 방법들은 실시간성이 매우 중요시 되는 증강현실과 같은 분야에 직접 적용시키기가 현실적으로 매우 힘들다[1, 2, 3]. 또한 기존의 일반적인 시각 기반 증강현실 기법들이 갖는 근본적인 문제점은 이들 기법이 갖고 있는 예외적인 상황에 대한 불안정성이다. 즉, 카메라의 급격한 흔들림, 장면의 가려짐 등과 같은 예측 불가능한 현상이 발생했을 경우, 기존의 통상적인 방법들은 공통적으로 큰 오차를 발생시키며 정상적인 카메라 추적에 실패하는 문제가 있다. 또한 이들 기법은 초기화 단계에서 추정된 특징점 정보가 카메라 움직임 등에 의해서 영상에서 사라진 경우에도 카메라의 추적이 불가능하게 되는 문제가 있다. 이와 같은 문제는 AR 시스템 구현 시 빈번하게 발생할 수 있는 문제이며, 따라서 이러한 예외적인 문제 발생 시 빠르게 정상상태로 복구되기 위한 안정화 기법이 필요하다. 이를 위

하여 [4, 5] 등과 같은 연구에서는 RANSAC 및 칼만 필터 등을 연동하는 방법 또한 제안되었다. 한편 본 논문과 유사한 입자필터 기반의 카메라 추적 기법으로 [9]의 연구가 있다. 특히 [9]의 연구에서는 카메라 상태에 대한 측정값(measurement) 표본 집합이 정규분포가 아닌 균일분포(uniform distribution)를 갖는다고 가정했으며, 측정값 표본입자의 확률값 추정을 위해서 적합/부적합 샘플을 이용한 가중치 계산 기법(inlier/outlier support weighting)을 별도로 이용하여 측정값을 기반으로 한 개별 표본입자의 확률값을 계산하였다. 이와 같은 방법은 표본입자의 분포가 시스템의 실제 불확실도에 수렴하는 속도가 다소 빨라진다는 장점이 있다. 하지만 이 방법은 확률값 추정을 위한 별도의 가중치 계산 과정을 필요로 하고, 카메라의 움직임에 민감한 결과값을 산출해내기 때문에 최종 결과에 카메라 떨림(jitter) 등과 같은 잡음 및 오차 성분이 존재할 수 있으며, 따라서 사용자에게 제공되는 결과영상이 다소 불안정하게 보이는 단점이 존재할 수 있다. 이를 보완하기 위해서 [9]에서는 별도의 입자 정련 과정(particle annealing) 또한 추가적으로 구현함으로써 안정된 카메라 추적 결과를 산출해내고자 하였다. 하지만 이는 카메라 입자를 가장 높은 확률값을 갖는 표본입자 주변으로 강제로 응집시키는 과정으로, 카메라 입자의 수가 증가하면 연산량이 급격하게 증가하게 되고 최종 응답속도가 현저하게 저하되는 단점이 존재하며 앞서 설명된 상태 벡터의 잡음 및 카메라 떨림 등과 같은 현상을 완전히 제거하지 못하는 문제점 또한 존재할 수 있다. 본 논문에서 제안하는 입자 필터링 방법에서는 시스템 상태의 측정값 분포가 가우스 분포를 갖는다고 가정했으며, 측정값 입자의 가중치 또한 가우스 분포를 기반으로 추정함으로써 별도의 추가적인 가중치 계산 과정을 거치지 않는다. 또한 제안된 방법에서는 기존의 방법과는 달리 입자 정련과 같은 별도의 과정을 수행하지 않고 실제 불확실도에 근사화되도록 재표본화된 카메라 표본 입자를 카메라의 상태 추정에 이용하였다. 이와 같은 경우 표본입자 확률의 수렴속도는 다소 감소하게 되나, 구현 결과에서 확인할 수 있듯이 이로 인하여 최종 결과영상이 받는 영향은 상당히 미미하다. 또한 입자 정련 및 별도의 가중치 계산 같은 추가적인 과정이 필요 없기 때문에 수행시간 측면에서도 상당히 유리하며 따라서 보다 많은 수의 표본입자를 이용한 강건한 추적 수행이 가능하다.

3. 입자필터를 이용한 카메라 정보 추정

3.1 시스템의 개요

본 절에서는 보다 강건한 카메라 정보의 추정을 위한 입자 필터 기반의 카메라 추적 기법에 대해서 설명한다. 입자 필터링은 순차적 몬테카를로 추정법(sequential Monte-Carlo estimation)에 기반한 필터링 기법이며, 칼만 필터, 또는 확장 칼만 필터 등과 같이 시스템 변수 불확실도의 가우스 분포 특성에 대한 가정이 필요 없으며, 또한 시스템의 상태 전이(transition) 및 측정(measure) 과정에서의 비선형 특성이 매우 강한 조건에서도 강건한 필터링 및 추적이 가능하다고 알려져 있다[14]. 입자 필터에서 카메라 상태에 대한 확률값은 3차원 공간과 2차원 영상 사이의 사영관계를 기반으로 계산되며, 이에 사전 확률값은 미리 정의된 확률밀도함수 모델에 의해서 결정된다. 계산된 입자의 확률값을 기반으로 카메라 입자의 재표본화(particle resampling)를 수행하며, 각 입자로부터 카메라 상태에 대한 예측값을 구한 다음 이들을 기반으로 카메라 상태를 추정할 수 있다. 그리고 미리 정의된 카메라 운동 모델에 따라서 각 입자의 카메라 상태 벡터를 조정하고, 다음 주기에서의 카메라 상태를 계산한다.

3.2 카메라의 상태 벡터 및 운동 모델

3차원 공간과 2차원 영상 사이의 사영 변환 관계는 3×4 사영 행렬로 나타낼 수 있다. 하지만 카메라 사영 행렬 원소의 전체 개수는 12개이며, 스케일 변수를 고려하면 11 자유도(DOF, Degree Of Freedom)를 갖는다. 이를 직접 카메라 추적 시스템에 도입할 경우 높은 자유도로 인한 복잡도 및 시스템의 체계화, 해석 등에 많은 어려움이 따를 수 있다. 따라서 다음과 같이 단순화된 형태의 카메라의 상태 변수를 도입한다. 즉, 카메라의 상태를 이동 변위를 나타내는 3×1 벡터 $\mathbf{t} = [t_x \ t_y \ t_z]^T$ 와 회전 변위를 나타내는 4×1 쿼터니온 벡터 $\mathbf{q} = [\mathbf{q}_0 \ \mathbf{q}_x \ \mathbf{q}_y \ \mathbf{q}_z]^T$ 로 구성된 7×1 의 1차원 벡터로 설정한다. 그리고 카메라 영상의 k 번째 프레임에 대한 상태 벡터(state vector)는 다음과 같이 나타낸다.

$$\mathbf{x}_k^c = [\mathbf{t}_k^c, \mathbf{q}_k^c]^T. \quad (1)$$

여기서 윗 첨자 c 는 카메라 좌표계 내에서의 좌

표값을 의미한다. 이때 월드 좌표계 내의 고정된 한 점 \mathbf{p}^w 는 다음의 관계에 의해서 카메라 좌표계 상의 한 점으로 사상된다.

$$\tilde{\mathbf{p}}_k^c = R(\mathbf{q}_k^c)^T (\mathbf{p}^w - \mathbf{t}_k^c) = [\mathbf{R}_k^c \ | \ \mathbf{t}_k^c] \tilde{\mathbf{p}}^w. \quad (2)$$

윗 식에서 $R(\mathbf{q})^T$ 는 쿼터니온 벡터 \mathbf{q} 의 3×3 회전 행렬 변환식이며, 윗 첨자 ‘ \sim ’는 좌표점의 동차좌표(homogeneous coordinate) 표현식을 나타낸다. 그리고 윗 첨자 w 는 월드 좌표계 내의 좌표를 의미한다. 3차원 좌표 \mathbf{p}^w 와 영상 내의 \mathbf{p}^c 의 관측값(observation), 즉 \mathbf{p}^w 의 2차원 사영점 \mathbf{p}_k^c 가 주어지면 식 (2)의 관계로부터 카메라의 상태 벡터 \mathbf{x}_k^c 의 추정이 가능하다. 한편 k 번째 프레임에서의 관측값을 \mathbf{y}_k 라고 하고 맨 처음 주기부터 k 번째 프레임까지의 관측값의 집합을 $\mathbf{y}_{1:k}$ 라고 하면, 카메라 정보의 추정 과정은 3차원 좌표의 집합 $Z = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_{k-1}, \mathbf{z}_k\}$ 가 주어졌을 때 관측값의 집합 $\mathbf{y}_{1:k}$ 와 3차원 좌표의 집합 Z 로부터 카메라 상태에 대한 귀납적 조건부 사후 확률(recursive posteriori), 즉 $P(\mathbf{x}_k^c | \mathbf{y}_{1:k}, Z)$ 을 계산하는 과정으로 볼 수 있다. $P(\mathbf{x}_k | \mathbf{y}_{1:k}, Z)$ 는 베이스의 정리(Bayes rule)에 의해서 다음과 같이 재귀적 베이스 필터(recursive Bayes filter)로부터 계산된다.

$$P(\mathbf{x}_k | \mathbf{y}_{1:k}, Z) = \frac{P(\mathbf{y}_k | \mathbf{x}_k, Z)P(\mathbf{x}_k | \mathbf{y}_{1:k-1})}{P(\mathbf{y}_k | \mathbf{y}_{1:k-1})}. \quad (3)$$

한편 k 번째 프레임의 카메라 상태 벡터 \mathbf{x}_k 에 대해서 입자 필터의 N 개의 표본 입자(sample particle)가 다음과 같다고 가정한다.

$$S_k = \{(\mathbf{x}_k^1, w_k^1), (\mathbf{x}_k^2, w_k^2), (\mathbf{x}_k^3, w_k^3), \dots, (\mathbf{x}_k^N, w_k^N)\}. \quad (4)$$

식 (4)에서 윗 첨자는 표본 입자의 순서를 나타낸다. 윗 식에서 n 번째 입자의 가중치(weight) w_k^n 는 3차원 좌표 집합 Z 및 현재 입자의 상태 벡터 \mathbf{x}_k^n 이 주어졌을 때 이에 대한 우도(likelihood)에 비례한다.

$$w_k^n \propto P(\mathbf{y}_k | \mathbf{x}_k^n, Z), \quad \sum_{n=1}^N w_k^n = 1. \quad (5)$$

식 (5)의 관계는 관측값의 분포 특성에 의해서 정의되는 관측값 밀도 함수(observation density)이며, 이를 계산하기 위해서는 입자 샘플 \mathbf{x}_k^n 에 대한 상태 전이 밀도 함수(state transition density) $P(\mathbf{x}_k | \mathbf{x}_{k-1})$ 가 필요하다. 전이 밀도 함수

수 $P(\mathbf{x}_k | \mathbf{x}_{k-1})$ 는 현재 주기에서의 시스템 상태가 다음 주기로 어떻게 변화되는지에 대한 특성 및 확률 분포를 정의해주기 위한 함수이다. 카메라 추적에서의 시스템 상태는 카메라의 위치, 방향 및 회전 성분으로 정의되며, 따라서 $P(\mathbf{x}_k | \mathbf{x}_{k-1})$ 는 결국 각 주기에서 카메라의 움직임이 다음 주기로 어떻게 변화하는지에 대한 카메라 움직임의 통계적인 모델을 정의한다. 본 논문에서는 카메라 움직임 모델을 마르코브 과정(Markovian process)이라고 가정하며, 카메라 상태 전이 함수를 다음과 같이 비선형 함수 $\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{n}_k)$ 이라고 가정한다. 여기서 \mathbf{n}_k 는 상태 벡터 \mathbf{x}_k 와 같은 차원(dimension)의 불확실도(uncertainty)를 나타내는 잡음 벡터(noise vector)이다. 일반적으로 칼만 필터 기반의 추적 시스템에서는 불확실도 모델을 가우스 분포를 갖는 확률 변수(random variable)라고 가정하며, 불확실도의 전과 특성은 가산성(additive)이라고 가정한다. 하지만 본 논문에서는 입자필터를 이용하여 카메라 추적을 수행하며, 따라서 칼만 필터에서와 같은 가우스 분포 및 가산성에 대한 조건이 필요없다. 따라서 본 논문에서는 아래와 같이 정의된 1차 확률보행 모델을 이용하여 카메라 추적을 수행하기로 한다 [12, 13]. 입자 필터를 이용한 카메라 추적에서 상태 전이 밀도 $P(\mathbf{x}_k | \mathbf{x}_{k-1})$ 가 다음과 같이 균일 분포(uniform distribution)를 갖는다고 가정한다.

$$P(\mathbf{x}_k | \mathbf{x}_{k-1}) = U(\mathbf{x}_{k-1} - \mathbf{v}, \mathbf{x}_{k-1} + \mathbf{v}). \quad (6)$$

여기서 \mathbf{v} 는 카메라 운동의 불확실도를 결정하기 위한 변수이며, 이때 카메라의 상태 벡터를 식 (7)과 같이 정의한다.

$$\mathbf{x}_k = \begin{pmatrix} \mathbf{t}_k \\ \mathbf{q}_k \end{pmatrix} = \begin{pmatrix} \mathbf{t}_{k-1} + \mathbf{V}_c \\ \mathbf{q}_{k-1} \otimes \mathbf{Q}(\Omega_c) \end{pmatrix}. \quad (7)$$

카메라 운동의 불확실도가 균일 분포를 갖는다고 가정했으므로, 식 (3.6)의 카메라 위치 및 회전 변위에 대한 불확실도 \mathbf{V}_c 및 Ω_c 는 각각 다음과 같이 균일 분포를 갖는 확률변수로 구성된 3 x 1의 벡터로 주어진다.

$$\mathbf{n} = \begin{pmatrix} \mathbf{V}_c \\ \Omega_c \end{pmatrix} \propto \begin{pmatrix} U^3(\mathbf{t}_{k-1} - \mathbf{v}_t, \mathbf{t}_{k-1} + \mathbf{v}_t) \\ U^3(\boldsymbol{\omega}_{k-1} - \mathbf{v}_r, \boldsymbol{\omega}_{k-1} + \mathbf{v}_r) \end{pmatrix}. \quad (8)$$

여기서 윗첨자는 각 변수의 차원을 나타낸다. 본 논문에서는 상태 전이 밀도가 카메라의 이전 상

태를 평균값으로 갖는 균일 분포라고 가정한다.

3.3 표본입자를 이용한 특징점의 사영 및 예측값 취득

식 (7)과 같이 카메라의 상태 벡터가 주어졌을 때, 이에 대한 측정 모델(measurement model)을 2차원 영상으로의 사영점으로부터 구할 수 있다. 3차원 좌표의 집합 $Z = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_{L-1}, \mathbf{z}_L\}$ 가 주어졌을 때, 이에 대한 관측값의 집합 $M = \{\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_{L-1}, \mathbf{m}_L\}$ 은 다음의 관계로부터 취득된다.

$$\tilde{\mathbf{m}}_i = \mathbf{K} \cdot \tilde{\mathbf{R}}(\mathbf{q}_k)^T (\tilde{\mathbf{z}}_i - \tilde{\mathbf{t}}_k) = \mathbf{K}[\mathbf{R}_k | \mathbf{t}_k] \tilde{\mathbf{z}}_i = \mathbf{P}_k \tilde{\mathbf{z}}_i. \quad (9)$$

여기서 첨자 “ \sim ”는 동차좌표계 표현식을 나타내며, \mathbf{K} 는 3 x 3 카메라 교정 행렬이다. 또한 $P(\mathbf{x}_k)$ 는 k 번째 프레임에서의 상태 벡터 \mathbf{x}_k 에 대한 카메라 사영 행렬 변환식을 의미한다. 위의 사영 관계식은 3차원 좌표점 \mathbf{z}_i 가 k번째 프레임의 카메라 상태 벡터 \mathbf{x}_k 에 의해서 영상으로 사영됨을 나타내는 관계식이다. N개의 표본 입자(sample particle)가 주어졌을 때 하나의 3차원 좌표점 \mathbf{z}_i 는 다음과 같이 n 번째 표본 입자에 의해서 2차원 사영점 \mathbf{m}_i^n 으로 사영되며, 이로부터 입자 필터의 예측값(prediction)을 얻을 수 있다.

$$\tilde{\mathbf{m}}_i^n = P(\mathbf{x}_k^n) \tilde{\mathbf{z}}_i = \mathbf{K}[\mathbf{R}_k^n | \mathbf{t}_k^n] \tilde{\mathbf{z}}_i = \mathbf{P}_k^n \tilde{\mathbf{z}}_i. \quad (10)$$

따라서 하나의 3차원 좌표에 대해서 N개의 2차원 사영점 입자 집합을 얻을 수 있다. 본 논문에서는 이들 각각에 대한 관측 밀도 값을 계산한 다음, 이로부터 카메라 상태의 예측값을 구한다. 이들 사영점 입자의 집합은 다음 절에서 설명될 특징점 탐색 영역의 지정에도 이용된다.

3.4 특징점의 검출 및 추적

앞 절에서 설명된 카메라 상태의 측정 과정은 3차원 좌표점 \mathbf{z}_i 및 2차원 사영점 사이의 사영 관계를 기반으로 수행되며, 이를 위해서는 \mathbf{z}_i 의 2차원 사영점의 취득 과정, 즉 특징점 검출(feature detection) 및 매칭(matching) 과정이 필수적이다. 특징점 검출을 위해서는 일반적으로 모서리 검출자(corner detector), SUSAN 모서리 검출자, Shi-Tomasi 모서리 검출자 등이 주로 이용되며, 본 논문에서는 특징점의 검출 및 추적을 위해서 FAST(Features from Accelerated Segment Test) 특징점 검출자[10, 11]과 NCC(Normalized Cross Correlation)를 병용하기로

한다. 즉, 현재 영상에서 FAST 검출자를 이용하여 지정된 영역 내에 위치한 특징점을 검출하고 이에 대한 템플릿을 저장한다. 그리고 순차적으로 입력되는 영상에서 FAST 검출자를 이용하여 새로운 특징점을 검출한 다음, NCC를 이용하여 지정된 특징점 템플릿의 위치를 추정하고 이를 특징점 좌표의 추적 결과로 이용한다. FAST 특징점 검출자는 앞서 언급된 다른 특징점 검출자에 비해서 수행 속도 및 검출된 특징점의 신뢰도 측면에서 보다 효율적인 방법으로 알려져 있다[11]. FAST 및 NCC 기반의 템플릿 매칭을 이용할 경우, 수행 시간 및 계산량의 측면에서 상대적으로 유리한 장점이 있으며, 특징점 검출이 아닌 다른 작업에 더 많은 자원을 할당함으로써 보다 정확하고 효율적인 시스템의 구성이 가능할 수 있다. 한편, 3차원 공간 상에서 각 좌표점 z_i 에 대한 템플릿을 $T(z_i)$ 라고 할 때, NCC를 이용한 템플릿 탐색 과정은 $r_{k,i} = T(z_i) * I_k$ 로 나타낼 수 있다. 여기서 I_k 는 k 번째 영상에서 템플릿을 탐색할 영역이며, $r_{k,i}$ 는 k 번째 영상에서의 i 번째 템플릿 $T(z_i)$ 와 I_k 사이의 NCC 연산 결과, 즉 2차원 교차상관 장(cross correlation field)이다. 또한 연산자 '*'는 템플릿 T 와 기준영상 I 사이의 NCC 연산을 나타낸다. 한편, 본 논문에서는 카메라 추적 과정의 예측값을 이용하여 NCC 연산을 위한 부영상 영역을 결정하는 능동 탐색(active search) 기법을 이용한다. 하나의 3차원 특징점 z_i 에 대해서, 식 (9)로부터 N 개의 표본 입자에 의한 예측값의 집합 $\mathbf{M}_i = \{\mathbf{m}_i^1, \mathbf{m}_i^2, \dots, \mathbf{m}_i^{N-1}, \mathbf{m}_i^N\}$ 을 얻을 수 있으며, 이들에 의해서 영상 내에 사영점의 입자 군(particle cloud)이 형성된다. 식 (10)으로부터 알 수 있듯이 영상 내에서 이들 사영점의 분포는 카메라 상태 벡터의 표본입자 분포에 의해서 결정되며, 따라서 이들 예측값이 분포하는 영역으로부터 카메라 상태가 반영된 부영상 영역을 취득할 수 있다. 이에 대한 예를 그림 1에 보인다. 그림은 중앙의 검은색 사각형의 네 모서리 각각에 대한 템플릿을 생성한 다음 이를 이용하여 NCC 탐색 영역의 취득 및 템플릿 매칭을 수행하는 것을 보여주고 있으며, 녹색 점들로 표시된 부분이 입자 필터에 의해서 생성된 각 특징점에 대한 예측값을 나타낸다. 입자 주변의 타원은 각각의 특징점에 대한 추정값 집합으로부터 계산된 공분산을 의미하며, 그림에서 보이는 바와 같이 타원에 의해서 지정되는 영역이 NCC 탐색을 위한 부영상 영역으로 이용된다.

3.5 입자 필터의 예측값에 대한 확률 밀도

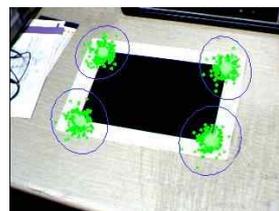
식 (3)으로부터 카메라 상태에 대한 사후확률 (posterior)을 얻기 위해서는 조건부 확률 $P(\mathbf{y}_k | \mathbf{x}_k, Z)$ 를 계산해야 한다. 이를 위해서 특징점 추적과정으로부터 취득된 특징점의 추정값 $\mathbf{y}_{k,i}$ 및 식 (9)의 관계로부터 계산된 k 번째 프레임의 예측값 집합 Z 를 이용한다. 먼저 k 번째 프레임에서 카메라 상태의 예측값이 추정값과 일치할 확률이 가우스 분포를 갖는다고 할 때, 이들 사이의 관계식을 다음과 같이 나타낼 수 있다.

$$\mathbf{y}_k = h(\mathbf{x}_k) + \mathbf{v}_k, \mathbf{v}_k \sim N(0, C). \tag{11}$$

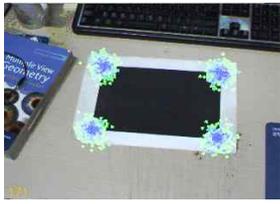
여기서 C 는 예측값 집합으로부터 계산된 공분산 행렬을 나타내며, $h(\mathbf{x})$ 는 카메라 입자 \mathbf{x} 에 의한 특징점 사영 과정을 의미한다. k 번째 프레임에서 i 번째 특징점에 대한 n 번째 입자를 $\mathbf{x}_{k,i}^n$ 라고 할 때, 이에 대한 예측값과 추정값 $\mathbf{y}_{k,i}$ 가 일치하는 경우에 대한 상대적인 가중치 q_k^n 를 다음과 같이 계산할 수 있다.

$$q_k^n = P(\mathbf{y}_k | \mathbf{x}_{k,i}^n, Z) = P(\mathbf{v}_k = \mathbf{y}_k - h(\mathbf{x}_{k,i}^n)) \\ \propto \frac{1}{2\pi \cdot |C|^{1/2}} \cdot \exp\left(-\sum_{i=1}^L (\mathbf{m}_i^n - \mathbf{y}_{k,i}^n)^T C^{-1} (\mathbf{m}_i^n - \mathbf{y}_{k,i}^n) / 2\right). \tag{12}$$

식 (12)에서 추정값 $\mathbf{y}_{k,i}$ 는 앞서 계산된 2차원 교차상관에서 최대값을 갖는 좌표로부터 취득할 수 있다. 따라서 식 (12)에 의해서 각 카메라 입자에 의한 예측값과 NCC를 이용한 추정값 사이의 차분에 의해서 결정되는 가중치 분포를 얻을 수 있다. 이후 카메라 입자들은 그림 2에 보이는 바와 같이 높은 확률값을 갖는 카메라 입자를 중심으로 재표본화된다. 본 논문에서는 전체 입자 평균 $\bar{\mathbf{x}}_k$ 을 상태 추정값으로 이용한다.



(그림 1) 예측값 입자의 공분산을 이용한 능동 탐색 영역 취득



(그림 2) 카메라 입자 재표본화 수행 이전 및 이후의 예측값 입자 분포



(그림 3) 실험에 사용된 CCD 카메라 및 3차원 좌표 추적기

4. 실험 과정 및 결과

실험은 2.4GHz의 인텔® Core2-Quad CPU 및 2GB RAM을 갖는 컴퓨터를 이용하여 진행되며, 영상 취득에는 320 x 240 크기의 RGB 영상을 30fps로 출력할 수 있는 IEEE1394 카메라를 이용한다. 또한 그림 3과 같이 추적 결과 확인을 위해서 카메라에 부착된 3차원 좌표 추적기로부터 취득된 카메라 정보를 입자 필터링 결과와 비교한다. 카메라 교정 및 내부 파라미터의 취득은 Zhang의 방법[7]을 이용하여 수행하였다.

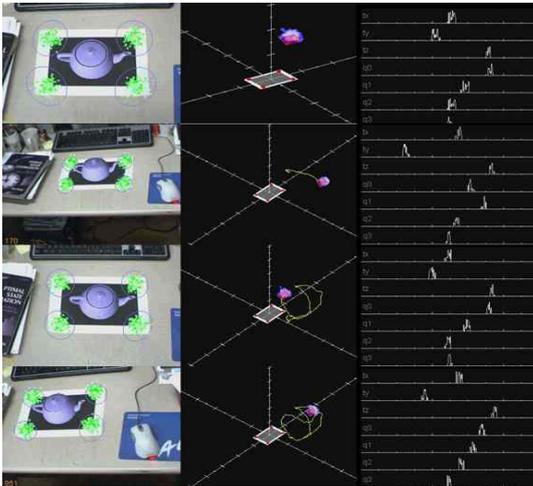
4.1 안정된 조건에서의 카메라 추적

그림 4에 입자 필터를 이용한 카메라 추적 결과를 보인다. 카메라 변수의 초기화는 그림에서와 같은 사각형 패턴의 네 모서리 좌표 및 템플릿을 이용하여 수행되며, 카메라 변수 및 입자 필터의 초기화가 완료되면 매 프레임마다 앞서 설명된 입자 필터 과정이 순차적으로 수행된다. 추적 결과에 대한 직관적인 확인을 위해서 OpenGL을 이용하여 생성한 가상의 그래픽 모델을 카메라 정보를 이용하여 영상으로 사영시킨 다음, 이를 실사 영상에 합성하여 출력하였다. 또한 그림의 녹색 점들은 영상으로 사영된 예측값 입자를 나타내며, 입자의 색상이 밝을수록 높은 값의 확률을 갖는 입자를 의미한다. 한편, 가운데 열의 영상은 3차원 공간에서의 기준 좌표계 시점에 대한 추적 과정을 보이며, 3차원 공간

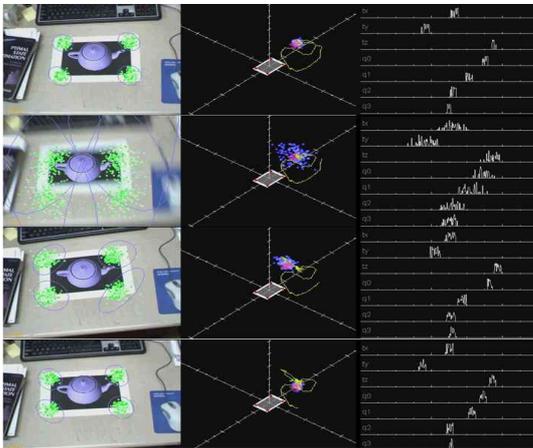
상에서의 카메라의 위치 및 특징점, 그리고 카메라의 궤적을 나타내고 있다. 푸른색 점으로 표시된 부분은 각각의 입자를 나타내며, 입자의 색상은 각 입자가 갖는 확률값을 의미한다. 즉, 그림에서 확률값이 높을수록 백색에 가까운 색상을 가지며, 확률값이 작을수록 푸른색을 갖는다. 가장 우측 열의 그림은 각 카메라 상태 변수에 대한 가중치 히스토그램을 보이며, 이로부터 각 프레임에서 입자의 확률값 분포를 확인할 수 있다. 그림으로부터 비교적 큰 시점 변화에도 큰 오차 없이 카메라 추적을 수행하는 것을 알 수 있으며, 일반적인 카메라 추적 기법에 비해서 우수한 추적 성능을 보인다는 것을 확인할 수 있다. 특히 순간적으로 발생하는 카메라의 급격한 움직임으로 인한 작은 화면 떨림 및 흐려짐(motion blur) 등과 같은 문제에 발생시에도 안정된 추적을 수행하는 것을 확인할 수 있다.

4.2 급격한 카메라 흔들림에 대한 추적

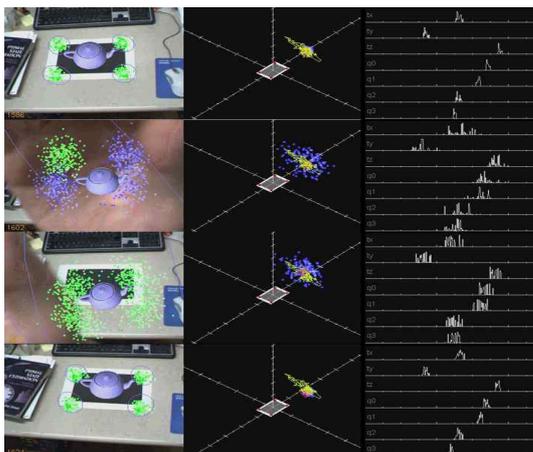
그림 5에 카메라의 급격한 흔들림에 대한 추적 결과를 보인다. 그림으로부터 매우 불안정한 상태의 카메라 운동이 발생한 경우에도 강건한 카메라 추적이 가능함을 확인할 수 있으며, 정상 상태로 복구하는데 매우 짧은 시간이 소요됨을 확인할 수 있다. 그림 8은 각 주기에서 각 카메라 변수에 대한 확률 분포를 보여준다. 그림에서 x축은 입자의 인덱스, y축은 프레임 수를 의미하며, 그래프의 가로축과 평행한 방향의 곡선이 입자의 확률값 분포를 나타낸다. 카메라의 운동이 정상 상태일 때는 앞서 설명된 바와 같이 카메라 입자의 확률값 분포가 입자 재표본화 과정에서 최대 확률값을 중심으로 가우스 분포를 갖도록 수렴하며, 따라서 그래프에서 카메라 운동이 정상 상태인 구간, 즉 1 ~ 1100 번째 프레임 등과 같은 구간에서는 입자의 분포가 하나의 입자를 중심으로 수렴하게 된다. 하지만 그림 5와 같이 카메라의 급격한 흔들림 등의 예외 상황 발생 시에는 영상의 흐려짐 현상 때문에 입자의 가중치 계산에 막대한 오차가 발생한다. 이러한 경우, 카메라 입자의 분포에 대한 공분산이 급격히 증가하며, 입자 필터의 상태 또한 불안정하게 된다. 그림 8의 그래프에서 1100 ~ 1600 번째 프레임 사이의 추적 구간이 이와 같은 예외상황 구간이며, 추정 오차로 인하여 입자 분포가 넓은 구간으로 분산된다. 한편, 정상상태로 돌아온 경우, 카메라 상태의 사전 확률값 추정이 정상적으로 수행되며, 따라서 입자 분포가 빠르게 정상상태로 복구되는 것을 확인할 수 있다.



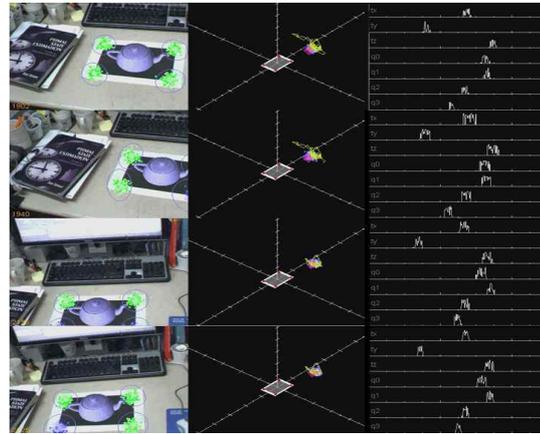
(그림 4) 안정된 환경에서의 카메라 추적 과정



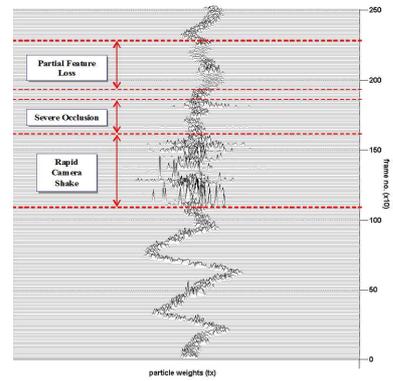
(그림 5) 급격한 흔들림에 대한 카메라 추적



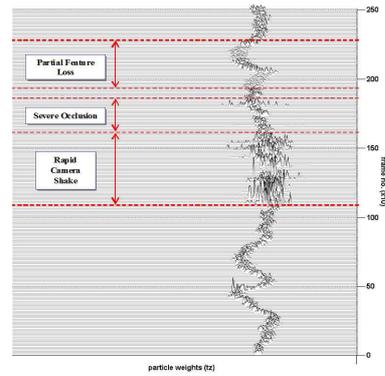
(그림 6) 장면 가려짐에 대한 카메라 추적 성능



(그림 7) 특징점 유실에 대한 카메라 추적 성능



(a)



(b)

(그림 8) 카메라의 x, z 축 이동 변위에 대한 표본입자의 확률값 분포. (a) tx, (b) tz.

4.3 장면의 가려짐 및 특징점 유실에 대한 카메라 추적성능

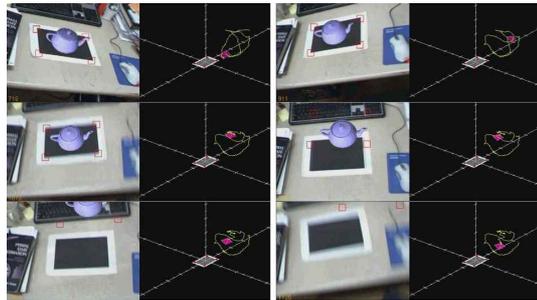
카메라 영상이 가려지는 경우, 영상으로부터의

정보 취득이 불가능해지며, 따라서 입자 가중치 계산시 막대한 오차가 발생한다. 그림 6에 이와 같은 장면의 가려짐 현상이 발생한 경우에 대한 추적 성능을 보여주고 있다. 장면이 가려진 경우에도 시스템이 안정상태로 돌아오면 입자의 분포가 빠르게 정상상태로 복구되는 것을 확인할 수 있으며, 이때의 카메라 입자의 분포를 그래프의 1600 ~ 1850번째 구간에서 보인다. 한편, 카메라의 움직임, 사용자의 동작 등에 의해서 일부 특징점이 유실되는 현상이 발생할 경우 통상적인 카메라 추적 기법은 추적 오차 때문에 추적에 실패하는 문제가 있다. 하지만 그림 7에서 보이는 바와 같이 입자 필터를 이용한 카메라 추적에서는 제한된 수준이기는 하지만 카메라 추적이 수행됨을 확인할 수 있다. 그림 9, 10에 UKF (unscented Kalman filter)와 입자필터의 카메라 추적 성능을 보인다. 그림 9의 UKF를 이용한 카메라 추적에서 급격한 흔들림이 발생하는 경우 앞의 결과와는 달리 추적에 실패하는 것을 알 수 있다.

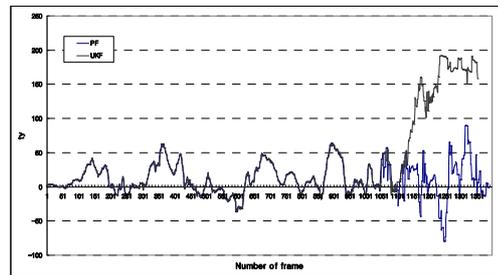
4.4 카메라 추적 오차 및 입자 수에 따른 카메라 추적 성능

입자 필터는 필터링에 사용되는 입자의 수에 의해서 성능이 크게 좌우되는 특성이 있다. 즉, 최대한 많은 수의 입자를 이용할수록 보다 정확한 결과를 얻을 수 있으나, 입자의 수가 늘어날수록 계산량 또한 증가한다는 문제가 있다. 그림 11, 12에 입자의 수와 수행 성능 사이의 상관관계를 보인다. 그림 12는 입자필터로부터 계산된 카메라 추적 결과와 그림 3에서 보인 3차원 좌표 추적기를 통하여 얻어진 카메라 변위 측정값 사이의 오차를 보여준다. 예상된 바와 같이 입자의 수가 증가할수록 알고리즘의 수행 시간 또한 증가하며, 이에 반하여 추정값의 정확도는 향상되는 것을 확인할 수 있다. 또한 그래프로부터 약 500개 정도의 입자가 이용될 때 수행 시간과 정확도 측면에서 안정된 카메라 추적의 수행이 가능하다는 것을 알 수 있다. 특히 그림 11은 입자의 수에 대한 기존의 방법 및 제안된 방법의 수행시간을 비교하여 보여준다. 그림에서 표본입자의 수가 약 700개 이상이 되면 기존의 방법은 초당 프레임 수가 약 16~17fps 정도로 하락하며 실시간 처리에 어려움이 발생하기 시작한다는 것을 알 수 있으며, 또한 입자의 수가 1000개를 초과하는 경우 정상적인 실시간 카메라 추적이 불가능하다는 것 또한 확인할 수 있다. 하지만 제안된 방법의 경우 1200개의 입자에 대한 프레

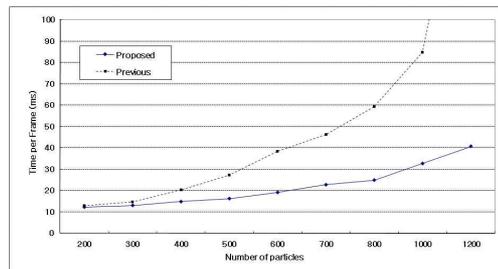
임당 소요시간이 약 40ms이며, 따라서 실시간 처리에 큰 어려움이 없음을 확인할 수 있다. 그림 13에 기존의 방법과 제안된 방법의 카메라 정보 추정 결과 그래프를 보인다. 기존의 방법의 경우 제안된 방법에 비해서 다소 빠른 응답특성 및 민감한 추적결과를 보임을 알 수 있으나 시스템 상태의 추정 결과에 카메라 떨림(jitter) 및 잡음이 존재함을 확인할 수 있다. 반면에 제안된 방법의 경우 잡음 및 떨림 현상이 기존의 방법에 비해서 개선된 것을 알 수 있으며, 따라서 보다 안정된 결과영상을 출력함을 확인할 수 있다.



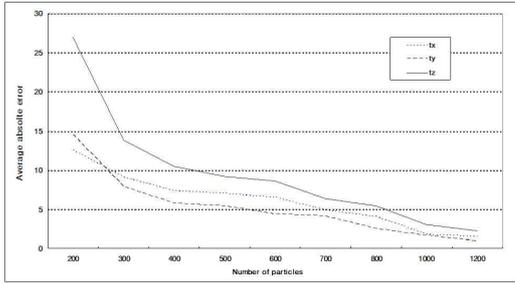
(그림 9) UKF를 이용한 카메라 추적 결과



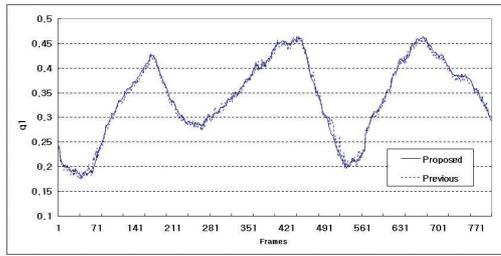
(그림 10) y축 이동 변위에 대한 UKF와 입자필터의 추적 성능 비교 (t)



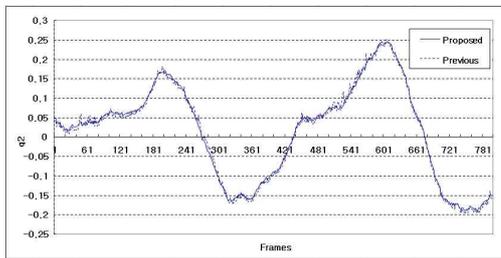
(그림 11) 입자 수에 대한 프레임당 소요 시간



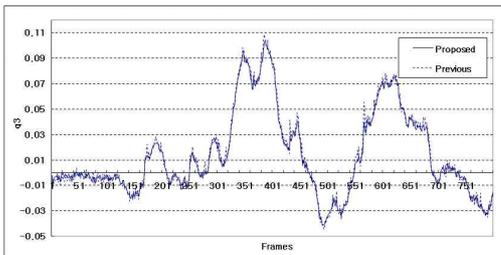
(그림 12) 카메라 입자 수에 따른 위치변위 오차



(a)



(b)



(c)

(그림 13) 제안된 방법 및 기존의 방법에 의한 카메라 정보 추정 결과. (a) q_1 , (b) q_2 , (c) q_3 .

5. 결론

본 논문에서는 보다 강건한 카메라 정보의 추정을 위한 입자 필터 기반의 카메라 추적 기법을 제안했다. 제안된 카메라 추적 기법은 카메라

상태에 대한 임의의 수의 표본 입자를 생성한 다음, 각 표본 입자에서 카메라 상태에 대한 확률값을 계산한다. 카메라 상태에 대한 확률값은 3차원 공간과 2차원 영상 사이의 사영관계를 기반으로 계산되며, 이에 대한 사전 확률값은 사전에 정의된 확률 밀도 함수 모델에 의해서 결정되었다. 계산된 입자의 확률값을 이용하여 카메라 상태 벡터를 조정하고 다음 주기에서의 카메라 상태를 추정하였다. 제안된 카메라 추적 기법은 시스템의 상태 전이(transition) 및 측정(measure) 과정에서의 비선형 특성이 매우 강한 조건에서도 강건한 필터링 및 추적이 가능함을 실험을 통해서 확인하였다. 특히 기존의 방법에 비해서 수행 시간 및 카메라 떨림/잡음 억제 측면에서 개선된 결과를 얻을 수 있었으며 보다 안정된 결과 영상을 출력함을 실험 및 구현 결과를 통하여 확인하였다.

참고 문헌

- [1] D. Koller, G. Klinker, E. Rose, D. Breen, R. Whitaker, and M. Tuceryan, "Realtime vision-based camera tracking for augmented reality applications," ACM Symposium on Virtual Reality Software and Technology, pp.87-94, 1997.
- [2] T. Drummond and R. Cipolla, "Real-time visual tracking of complex structures," IEEE Transactions on Pattern Analysis and Machine Intelligence Vol. 24, No.7, pp.932-946, 2002.
- [3] I. Gordon and D.G. Lowe, "Scene modelling, recognition and tracking with invariant image features," International Symposium on Mixed and Augmented Reality, 2004.
- [4] A Chiuso, P Favaro, H Jin, and S Soatto, "Structure from motion causally integrated over time," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.24, No.4, pp.523-535, 2002.
- [5] D. Nister, "Preemptive RANSAC for live structure and motion estimation," Machine Vision and Applications, Vol.16, No.5, pp.321-329, 2005.
- [6] R. Azuma, "A Survey of Augmented Reality," Teleoperators and Virtual Environments, pp. 355-385, Vol. 6, No.4, Aug. 1997.
- [7] Z. Zhang, "A Flexible New Technique for Camera Calibration and the Application," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1330-1334, Vol. 22, No. 11, Nov. 2000.

[8] A. Davison, I. Reid, N. Morton, and O. Stasse, "Mono SLAM: Real-Time Single Camera SLAM," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1052-1067, Vol. 29, No. 6, Jun. 2007.

[9] M. Pupilli and A. Calway, "Real-time camera tracking using a particle filter," In Proceedings of the British Machine Vision Conference. (BMVC'05), pp. 43-57, Oxford, UK, Sep. 2005.

[10] E. Rosten, and T. Drummond, "Fusing points and lines for high performance tracking," In Proceedings of IEEE International Conference on Computer Vision (ICCV'05), pp. 1508-1511, Beijing, China, Oct. 2005.

[11] E. Rosten, and T. Drummond, "Machine learning for high-speed corner detection," In Proceedings of European Conference on Computer Vision (ECCV'06), pp. 430-443, Graz, Austria, May. 2006.

[12] S. Soatto and P. Perona, "Reducing Structure From Motion: A General Framework for Dynamic Vision Part 1: Modeling," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 933-942, Vol. 20, No. 9, Sep. 1998.

[13] S. Soatto and P. Perona, "Reducing Structure From Motion: A General Framework for Dynamic Vision Part 2: Implementation and Experimental Assessment," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 943-960, Vol. 20, No. 9, Sep. 1998.

[14] D. Simon, "Optimal State Estimation," John Wiley & Sons, Hoboken, NJ, 2006.

[15] R. Hartley, and A. Zisserman, "Multiple-View Geometry in Computer Vision," Cambridge University Press, Cambridge, 2003

이 석 한



1999년: 중앙대학교 전자공학과 졸업 (공학사)

2001년: 중앙대학교 첨단영상대학원 영상공학과 (공학석사)

2009년: 중앙대학교 첨단영상대학원 영상공학과 (공학박사)

2001년~2004년: (주)LG전자 주임연구원

2009년~2010년: 중앙대학교 첨단영상대학원 CT연구소 전임연구원 (post-Doc.)

2010년~현재: 중앙대학교 첨단영상대학원 연구전담교수

관심분야: 3D computer vision, real-time camera tracking, image processing