

초다시점 영상 합성을 위한 온라인 삼차원 복원 기술

김정호, 김제우, 권인소*
전자부품연구원, *KAIST

요약

본 논문에서는 초다시점 (Super Multi-view) 영상 합성을 위한 영상 기반의 온라인 삼차원 복원 기술들을 소개한다. 복원의 정확성을 높이고자 하는 방법은 크게 두 부류로 나뉜다. 먼저 재투영 오차를 비용 함수(Cost function)으로 정의하고, 이를 Bundle Adjustment로부터 최적화를 수행하는 방법과 카메라의 위치와 삼차원 복원 결과에 대해 확률적인 분포를 정의하고 이를 순차적으로 추정하는 확률적인 필터링(Stochastic filtering)에 기반한 방법이 존재한다. 본 논문에서는 두 방법의 장단점을 분석하고, 이로부터 새로운 확률적 필터링에 기반한 3차원 복원 및 카메라 위치 추정 방법을 제안한다. 이로부터 대공간 환경에 적용하여 성능을 검증한다.

I. 서론

삼차원 복원은 방송 콘텐츠(Contents)제작에서부터 문화재 복원[1]과 지능형 로봇 주행 기술[2] 까지 다양한 응용분야의 핵심 기술이 된다. 그림 1은 삼차원 복원 기술을 기반으로 가상의 시점 영상을 합성하는 예를 보여주고 있다. 이를 위해서 먼저 <그림 1(a)>와 같은 입력 영상들로부터 카메라의 상대적인 움직임을 추정하고, <그림 1(b)>와 같이 장면에 대한 삼차원을 복원한다. 복원된 삼차원 결과를 임의의 시점으로 재투영 (Re-projection)함으로써 <그림 1(c)>와 같이 가상의 시점에 대한 영상을 합성하고, 후처리 (Post-processing) 기술을 이용하여 <그림 1(d)>와 같이 가상 시점 영상을 생성한다. 삼차원 복원 기술은 영상들로부터 영상들 사이의 상대적인 카메라 자세를 계산하고, 이로부터 영상간 대응점들을 삼각측량 (Triangulation)을 적용하여 삼차원의 위치를 계산하게 된다. 하지만 일반적인 삼차원 복원 방법은 대용량의 영상들로부터 오프라인으로 삼차원을 복원하는 기술[3]로 실시간 방송 기술 또는 로봇 주행 기술에 적용하는 것이 어렵다. 또한 다중 카메라

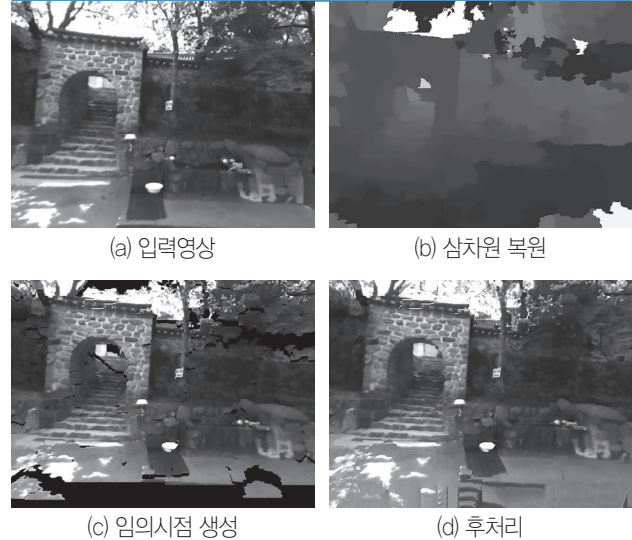


그림 1. 삼차원 복원을 이용한 다시점 영상 합성 기술

라 시스템으로부터 획득한 영상들로 삼차원을 복원하는 문제는 데이터 획득이 용이하지 않기 때문에 최근 단일 카메라 기반의 삼차원 복원 기술들이 많이 연구되고 있다. 단일 카메라로부터 획득한 영상 시퀀스를 이용한 삼차원을 복원 할 경우 시간이 지날수록 오차가 누적되는 경향이 있기 때문에 전체 삼차원 복원 결과의 정확도가 저하된다.

이 문제를 해결하기 위해서 컴퓨터 비전 분야에서는 Bundle Adjustment[4]와 같은 카메라의 위치들과 삼차원 복원 결과를 최적화 하는 기술들이 개발되었고, 이를 실시간에 적용하기 위해서 Local Bundle Adjustment(LBA) 기법들이 제안되었다 [5]. Klein과 Murray는 최적화된 실시간 처리를 위해서 카메라의 위치를 계산하는 부분과 삼차원을 복원하는 부분을 서로 다른 스레드(thread)로 구현함으로써 문제를 해결하였고, 복원 기술로부터 증강현실(Augmented Reality)에 적용하였다 [6]. 로봇틱스 분야에서는 삼차원 복원 문제를 해결하기 위해서 카메라의 위치와 삼차원 복원에 대한 확률적인 모델을 정의하고 이로부터 재귀적으로(Recursively) 확률분포를 추정하는 필터링 방법들이 제안되었다. 예를 들면, Davison은 단일 카메라를 이용하여 카메라의 위치를 삼차원 랜드마크(Landmark)

의 위치에 대한 가우시안 (Gaussian) 분포를 추정할 수 있는 Extended Kalman Filter(EKF)를 설계하였다[7]. [8][9]에서는 카메라의 위치와 랜드마크 (Landmark)들의 위치에 대한 분포를 분리해서 추정할 수 있는 새로운 필터링 기법에 대해 제안하였다. 이는 Rao-Blackwellized particle filter (RBPF)에 기반한 방법으로 카메라의 경로에 대한 변수들을 입자(particle)들로 표현하여 추정하고, 나머지 랜드마크들의 위치들은 가우시안 분포를 가정하고 Kalman Filter로부터 계산하였다. Kim et al. [10]은 비선형 관계식으로부터 야기 되는 오차를 해결하기 위해서 Unscented Kalman Filter를 이용한 Unscented FastSLAM 알고리즘을 제안하였다.

〈그림 2〉는 대표적으로 최적화 기법을 이용한 카메라의 위치 추정 방식[5]과 확률적인 필터링 기법을 이용한 카메라의 위치 추정[9] 결과를 보여준다. 정량적인 결과 비교를 위해서 GPS (Global positioning system)로 부터 획득한 좌표들로부터 검증하였다. 〈그림 2(a)〉의 경우 카메라 영상들간

의 대응점을 계산하고 이로부터 상대적인 카메라의 움직임을 추정하는 3-point 알고리즘[11]을 채택하였고, 특정 키프레임 (keyframe)이 지정되었을 경우 카메라의 위치들과 삼차원 랜드마크들의 좌표들을 Bundle Adjustment로부터 최적화하는 방식으로 결과를 획득하였다. 〈그림 2(b)〉의 경우 카메라의 위치와 랜드마크들의 위치에 대한 확률분포들을 Rao-Blackwellized Particle Filter[9]를 이용하여 계산한 결과이다.

〈그림 2〉의 실험결과는 확률적인 필터링에 기반한 방법이 최적화 기반의 방법과 비교하여 정확성이 떨어지는 것을 확인할 수 있다. 그러므로 본 논문에서는 확률적인 필터링 방법에 대한 오차 생성의 근본적인 원인 분석과 더불어 성능을 높이는 방법들을 제안한다.

II. 확률적인 필터링과 한계점

확률적인 필터링을 이용한 삼차원 복원 및 카메라의 자세 추정 기술은 아래와 같은 확률 분포를 추정하는 문제와 동일하다.

$$p(x_t | z_{1:t}, u_{1:t}) \quad (1)$$

여기서 x_t 는 현재 카메라의 위치와 삼차원 랜드마크들의 위치를 포함하는 변수(variable)을 나타낸다. 그리고 $z_{1:t}$ 는 현재까지 획득한 센서의 데이터를 의미하고, $u_{1:t}$ 는 카메라의 이동거리를 예측하게 하는 GPS, IMU 센서 값 또는 등속도 모델 (Constant velocity model)등이 될 수 있다.

Eq. (1)을 순차적으로 계산하기 위해서 베이시안(Bayesian) 형태에 의해 Eq. (2)와 같은 두 개의 식으로 표현이 가능하다. 이는 먼저 현재 카메라의 위치에 대한 예측 분포 (Predictive distribution)을 계산하고, 센서 데이터를 이용하여 예측 분포를 새롭게 갱신하는 형태를 가진다.

Prediction:

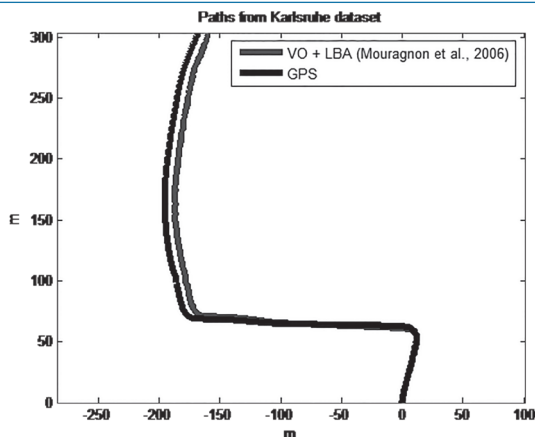
$$p(x_t | z_{1:t-1}, u_{1:t}) = \int p(x_t | x_{t-1}, u_t) p(x_{t-1} | z_{1:t-1}, u_{1:t-1}) dx_{t-1}$$

$$\text{Update: } p(x_t | z_{1:t}, u_{1:t}) = \eta p(z_t | x_t) p(x_t | z_{1:t-1}, u_{1:t}) \quad (2)$$

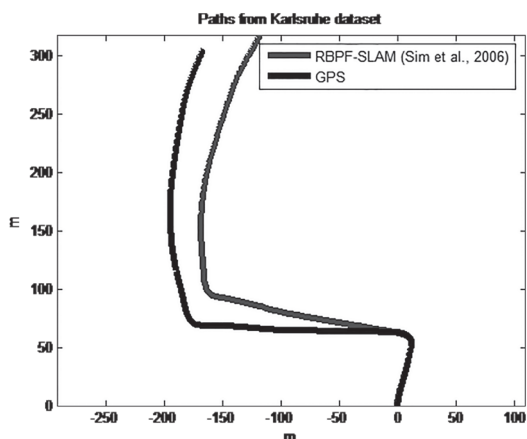
Eq. (2)에 포함된 두 확률 분포, $p(x_t | x_{t-1}, u_t)$, $p(z_t | x_t)$ 는 Eq. (3)과 같이 카메라의 모션 모델 (motion model)과 카메라의 관측 모델(measurement model)로부터 정의된다.

$$\begin{aligned} x_t &= f(x_{t-1}, u_t) + w_t \\ z_t &= h(x_t) + v_t \end{aligned} \quad (3)$$

여기서 w_t 와 v_t 는 모션 모델과 관측 모델의 불확실성을 나타내는 변수(variable)이다. 불확실성에 대한 확률분포를 표현



(a) Local Bundle Adjustment(LBA) 기반의 카메라 위치 추정 방식 결과



(b) Rao-Blackwellized Particle Filter (RBPF)기반의 카메라 위치 추정 방식 결과

그림 2. 최적화 기법과 확률적인 필터링 기법을 이용한 카메라 위치 추정 결과 비교

하기 위해서 일반적으로 가우시안(Gaussian) 분포를 가정하는 Kalman Filter[7]기반의 방법들과 입자들로 임의의 분포를 표현하는 입자 필터 기반의 방법[9]으로 대표될 수 있다. 하지만 가우시안 분포를 가정함으로써 필연적으로 오차를 발생시킨다. 입자로 표현할 경우 모든 경우에 대한 확률이 아닌 샘플링(Sampling)된 변수에 대한 확률만을 정의할 수 있기 때문에 입자들을 어떻게 생성하느냐에 따라서 그 성능을 좌우하게 된다.

확률적인 필터링의 경우 Eq.(2)에서와 같이 먼저 예측(Prediction) 단계에서는 현재 상태 변수 x_t 에 대한 확률분포를 예측하게 되고, 현재 센서 데이터 z_t 로부터 예측된 확률 분포를 갱신하게 된다. 이때 예측단계에서는 현재 상태 변수에 대한 불확실성이 증가하게 되고 갱신 단계는 증가된 불확실성을 감소시키게 된다. 하지만 이때 증가된 양만큼 항상 감소하지 않기 때문에 시간이 지날수록 불확실성은 증가되어 나중에는 필터가 발산(Divergence)하는 경우가 발생한다. 이는 일반적으로 다양한 문제로 인하여 갱신이 제대로 이루어지지 않을 경우 주로 발생을 한다. 이를 필터의 일관성(Consistency)문제라 한다.

III. 제안된 확률적 필터링 기술

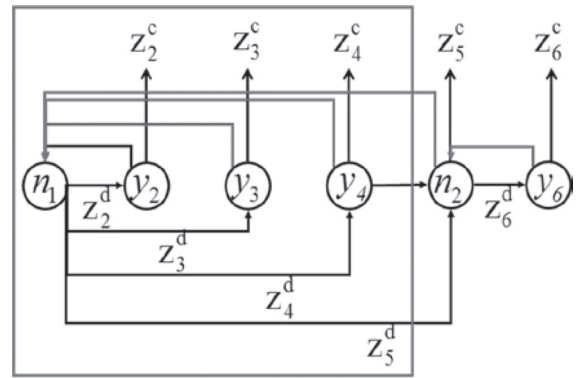
위에서 언급했던 확률적 필터링을 설계할 때 필요한 모션 모델을 정의하기 위해서 우리는 카메라 기반의 움직임 추정(Motion Estimation) 방법[12]과 RANSAC[13] 방법을 결합한 입자 생성 방법론을 제안한다. 이는 카메라의 위치 추정을 위한 3-point 알고리즘[11]에서 임의의 3점의 삼차원 좌표와 그에 해

당하는 영상의 특징점(feature)을 선택하고, 이로부터 하나의 상대적인 위치관계를 계산할 수 있다. 이 계산된 위치관계를 하나의 입자로 정의하고, 가중치는 대응점 중 계산된 카메라 위치에 대해 삼차원 좌표를 재투영 하였을 경우 영상 좌표와의 거리가 작다고 판단되는 대응점의 수로 정의한다. 그러므로 Eq. (4)와 같이 모션 모델에 대한 확률 분포를 정의할 수 있다.

$$p(x_t | x_{t-1}, r_t^{[i]}) = \frac{N_{inlier}(x_{t-1}, r_t^{[i]})}{N_T} \delta(x_t - (x_{t-1} \oplus r_t^{[i]})) \quad (4)$$

여기서 $r_t^{[i]}$ 는 RANSAC에서 획득한 하나의 카메라 움직임 추정 결과값으로 가중치는 추정된 모션 값이 가지는 인라이어(Inlier) 수 $N_{inlier}(x_{t-1}, r_t^{[i]})$ 에 비례한다. 여기서 N_T 는 전체 대응점 수를 나타낸다.

〈그림 3〉은 제안된 방법으로 입자들을 생성한 결과를 보여준다. 사용자가 미리 정해둔 파라미터(parameter)가 필요 없이 자동으로 영상으로부터 움직임에 대한 확률분포를 적응적(Adaptive)으로 계산할 수 있다.



n : key frame \longrightarrow forward filtering
 y : non key frame \dashrightarrow backward marginalization

그림 4. 제안된 키프레임 기반의 카메라 경로 추정 방법의 상태 변이도

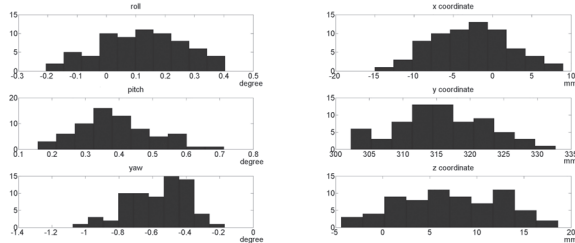
필터의 일관성을 향상시키기 위해서 불확실성을 증가시키는 예측단계를 줄이고, 불확실성을 감소시키는 갱신단계의 횟수를 증가시키는 것이 필요하다. 이를 위해서 일반적으로 키프레임(keyframe)기반의 카메라 움직임 추정 및 삼차원 복원 기술들이 소개되었다[5][6]. 그러므로 본 논문에서는 〈그림 4〉와 같이 키프레임 기반의 확률적인 필터링 방법을 제안한다. 이를 위해서 Eq. (5)와 같은 키프레임의 카메라 경로 $n_{1:k}$ 와 삼차원 복원을 위한 랜드마크의 위치들 M 에 대한 확률 분포를 계산한다.

$$p(n_{1:k}, M | z_{1:t}, u_{1:t}) = p(n_{1:k} | z_{1:t}, u_{1:t}) p(M | n_{1:k}, z_{1:t}, u_{1:t}) \quad (5)$$

Rao-Blackwellized particle filtering 기법을 채용함으로써, $n_{1:k}$ 와 M 에 대한 동시 확률 분포를 추정하는 문제를 카메라의 경로를 추정하는 문제와 삼차원을 복원하는 문제로 나누어 독



(a) 입력 영상



(b) 이동 변이

(c) 이동 각도

그림 3. RANSAC 기반의 샘플 생성 방법론

립적으로 해결할 수 있기 때문에 문제의 복잡도를 줄일 수 있다. Eq. (5)에서 확률분포 $p(n_{1:k} | z_{1:t}, u_{1:t})$ 는 현재 프레임이 키프레임으로 지정이 되었을 경우에만 생성되는 변수로 키프레임에서만 예측을 수행하기 때문에 예측의 횟수를 줄일 수 있다. 갱신 횟수를 높이기 위해서 현재 프레임이 키프레임이 아닐 경우에도 영상으로부터 키프레임의 위치들을 추정할 수 있는 방법이 필요하다. 이를 위해서 이전 키프레임에서 현재 프레임 사이의 상대적인 위치 r_t 를 잠재 변수 (Latent variable)로 정의하고, 키프레임 카메라 위치를 현재의 관측치 z_t 로부터 Eq. (6)과 같이 새롭게 추정한다.

$$p(n_{1:k} | z_{1:t}, u_{1:t}) = \int_{u_t} p(n_{1:k}, r_t | z_{1:t}, u_{1:t}) du_t \quad (6)$$

Eq. (6)에서 동시 분포 (joint distribution)은 Bayes' theorem에 의해 아래와 같이 정의 될 수 있다.

$$\frac{p(n_{1:k}, r_t | z_{1:t}, u_{1:t})}{p(z_t | n_{1:k}, r_t) p(r_t | u_t) p(n_{1:k} | z_{1:t-1}, u_{1:t-1})} \quad (7)$$

여기서 $p(z_t | n_{1:k}, r_t)$ 는 가능성 분포 (likelihood distribution)를 나타내고, $p(r_t | u_t)$ 는 <그림 3>과 같이 RANSAC으로부터 획득한 입자들로부터 분포를 표현할 수 있다.

가능성 분포는 현재 카메라의 위치에서 랜드마크들을 투영하여 현재 관측치와 얼마나 유사한 위치에 있는가를 검사함으로써 Eq. (8)과 같이 계산이 가능하다.

$$\begin{aligned} p(z_t | n_{1:k}, r_t) &= \int_M p(z_t, M | n_{1:k}, r_t) dM \\ &= \int_M p(z_t | M, n_{1:k}, r_t) p(M | n_{1:k}, r_t) dM \\ &= \frac{1}{L} \sum_{l=1}^L d(z_t^l, m_l, n_{1:k}, r_t) \end{aligned} \quad (8)$$

여기서

$$d(z_t^l, m_l, n_{1:k}, r_t) = \begin{cases} 1 & \|z_t^l - z(m_l, (n_{1:k} \oplus r_t))\| < \sigma_l \\ 0 & \text{otherwise} \end{cases}$$

$(n_{1:k} \oplus r_t)$ 는 이전 키프레임의 위치와 현재 프레임의 상대적인 위치의 조합으로 계산한 현재 카메라의 위치를 나타낸다. $z(\cdot)$ 는 재투영 함수를 나타내고, σ_l 는 Eq. (10)에서 계산된 랜드마크들의 불확실성을 나타낸다.

Eq. (5)에서 삼차원 랜드마크에 대한 확률 분포는 카메라의 경로가 주어졌을 경우 각 랜드마크마다 서로 독립이라는 사실이 성립하므로 전체 랜드마크들의 삼차원을 추정하는 문제를 Eq. (9)과 같이 각 랜드마크마다 독립적으로 위치를 추정하는 문제로 바꿀 수 있다.

$$p(M | n_{1:k}, z_{1:t}, u_{1:t}) = \prod_{l=1}^L p(m_l | n_{1:k}, z_{1:t}, u_{1:t}) \quad (9)$$

<그림 5>와 같이 주어진 카메라의 위치와 영상 사이의 대응점들을 이용하여 다시 삼각측량(Triangulation)을 수행하고 위치에 대한 분포, $p(m_l | n_{1:k}, z_{1:t}, u_{1:t})$ 를 갱신한다. 그리고 각 랜드마크 마다의 불확실성을 계산하기 위해서 마지막 키프레임으로 재투영하고 관측치 사이의 거리를 아래와 같이 계산한다.

$$e_l = \|z_N^l - z(\hat{m}_l, n_N)\| \quad (10)$$

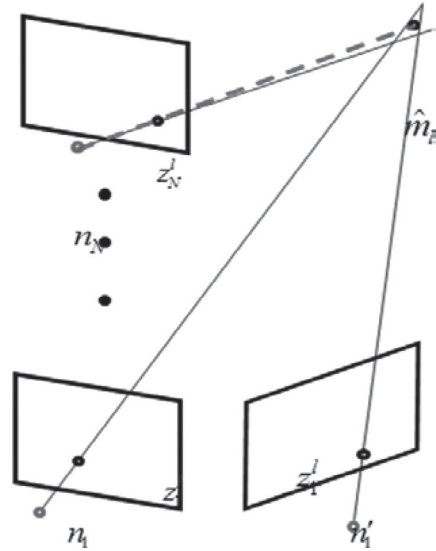


그림 5. 삼차원 복원을 위한 랜드마크 추정 방법

IV. 실험 결과

삼차원 복원 및 카메라 위치 추정에 대한 실험을 위해서 <그림 6>과 같이 실외환경에서 Hand-Held 스테레오 카메라로부터 총 10667 영상을 약 400m 거리에서 획득하였다.



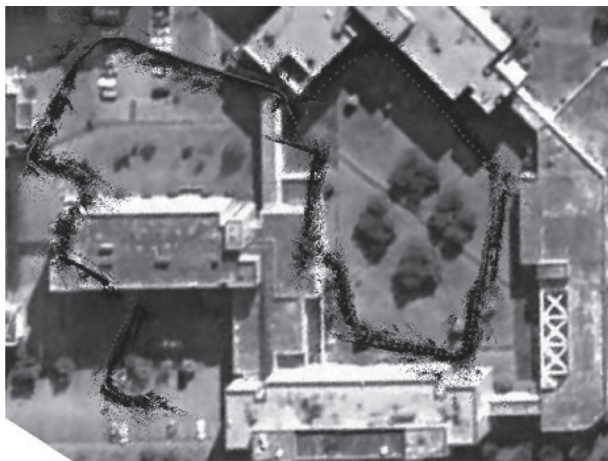
그림 6. 실외환경에서 획득한 테이스 영상

이중에 325개의 영상을 키프레임으로 지정하였고, 110076개의 랜드마크의 위치를 계산하였다. 제안된 알고리즘은 C++로 구현되었고, 2.4GHz CPU에서 약 10 프레임을 처리할 수 있다. <그림 7(a)>는 카메라를 이용한 움직임 추정 방법[12]에 기반한 삼차원 복원 결과를 보여준다 <그림 7(b)>와 (c)는 제안된 필터

링 방법에 기반한 카메라 위치 추정 및 삼차원 복원 결과를 보여준다. 본 결과에서 보여주듯 제안된 확률적인 필터링 방법은 <그림 7(a)>에서 보여준 것처럼 시간이 지나면서 오차가 누적되는 현상을 방지하여 정확한 삼차원을 복원할 수 있음을 보여준다. 제안된 방법의 성능을 검증하기 위해서 카메라의 위치 추정 결과를 Local Bundle Adjustment(LBA) 기반의 방법[9]과 비교하였다.

<그림 8>은 제안된 방법과 LBA의 결과를 비교해서 보여준다. LBA의 경우 Levenberg-Marquardt (LM)을 반복하여 최적화를 하는데 <그림 8(a)>의 경우 15번의 LM을 수행한 결과로 제

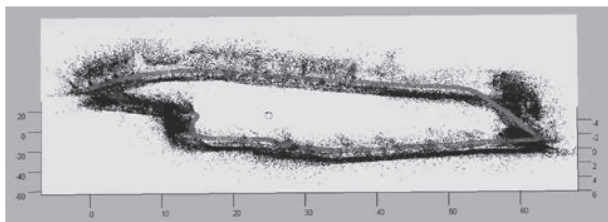
안된 방법과의 처리속도가 유사하다. <그림 8(b)>의 경우 40번의 LM을 수행한 결과로써 이 경우는 제안된 필터링 방법과 유사한 성능을 보이는 반면 많은 계산량이 요구된다. 즉 LM의 반복횟수가 증가할수록 LBA 결과는 제안된 방법으로 획득한 결과와 유사하다는 것을 확인할 수 있다. Bundle Adjustment 기반의 방법과 다르게 제안된 필터링 방법은 현재 센서 데이터만을 사용하여 정확성과 더불어 계산의 효율성을 보장할 수 있다. 뿐만 아니라 입자 필터링 (Particle filtering)기반의 방법의 경우 Open Multi-Processing(OpenMP)로 프로그래밍으로 구현이 용이하다.



(a) Visual Odometry기반의 카메라 위치 추정 및 삼차원 복원 결과

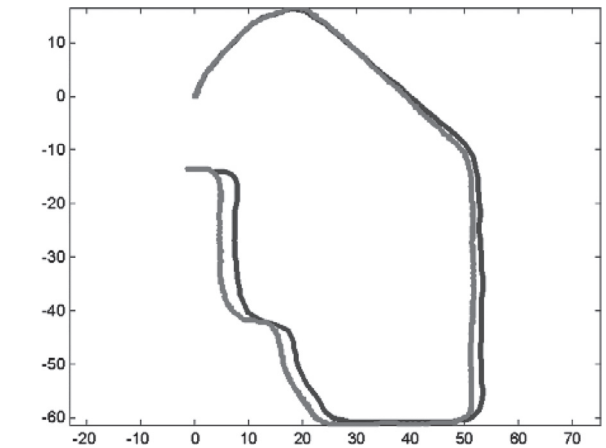


(b) 제안된 확률적 필터링을 이용한 카메라 위치 추정 및 삼차원 복원 결과

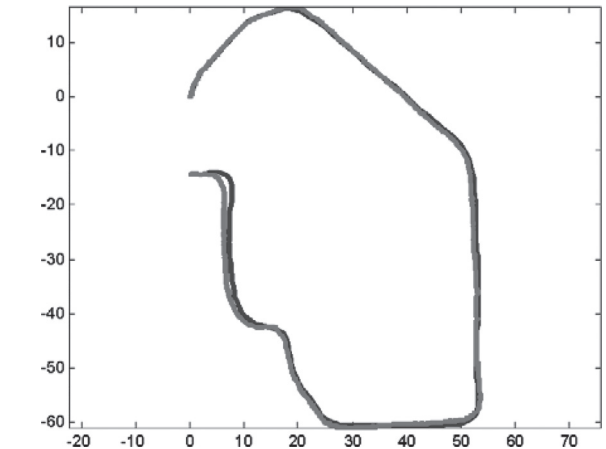


(c) 제안된 방법론의 삼차원 시점 뷰

그림 7. 카메라 위치 추정 및 삼차원 복원 결과



(a) 15번의 LM으로 최적화를 수행 결과



(b) 40번의 LM으로 최적화를 수행한 결과

그림 8. 다른 횟수의 LM기반의 LBA와 제안된 필터링 방법의 성능 비교
(빨강: LBA기반의 카메라 위치 추정, 파랑: 제안된 방법으로 획득한 카메라의 위치 추정)

<표 1>은 제안된 방법을 구현하였을 때 처리 속도를 보여주고 있다. 영상의 화소수가 320x240일 때 카메라의 위치를 추정하는 방법의 경우 27ms 정도가 소요되고, 확률적 필터링 방법은 76ms 가 소요된다.

표 1. 제안된 방법의 평균 처리 시간

기능	처리속도 (ms)
카메라 위치 추정	27
확률적 필터링	76
전체	103

그러므로 제안된 방법론은 초당 약 10프레임의 영상을 처리할 수 있다.

〈그림 9〉는 〈그림 2〉에서 보여준 결과에서 제안된 필터링 방법을 사용했을 때 얻어진 결과를 추가하여 비교한다. 제안된 방법은 기존의 확률적 필터링 방법에 비해 우수한 결과를 보이며, 최적화 기법인 Local Bundle Adjustment 기반의 방법 대비 우수한 결과를 보임을 확인할 수 있다. 기존의 state-of-the-art 필터링 방법인 RBPF 방식 보다 우수한 결과를 보이는 이유는 필터링의 근본적인 문제인 확률 모델링과 필터의 일관성을 유지하도록 하였기 때문이다.

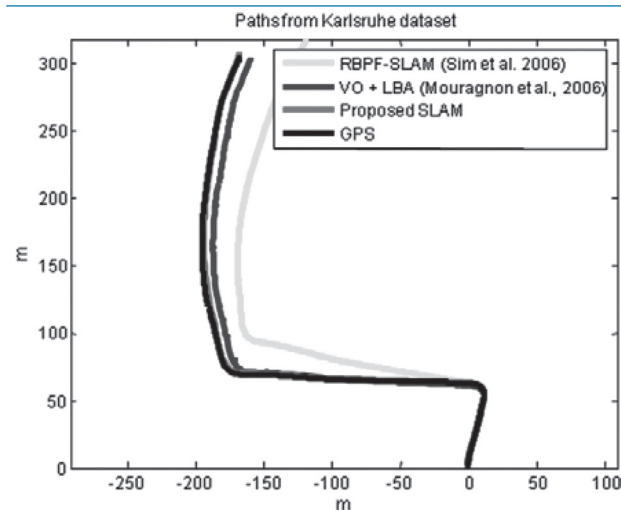


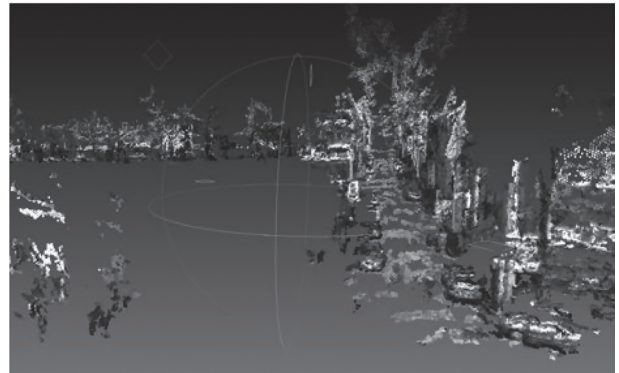
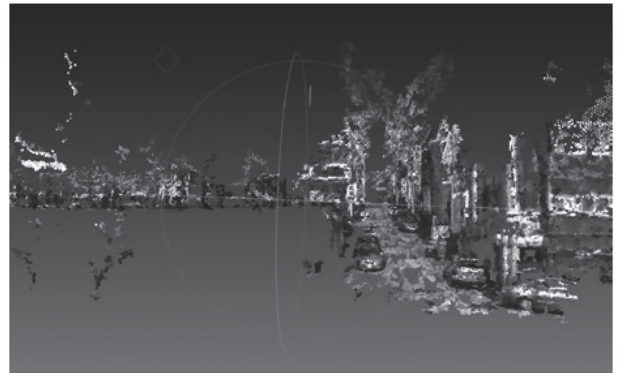
그림 9. Karlsruhe 데이터셋을 이용한 성능 비교

획득한 모든 데이터를 이용하여 오프라인에서 카메라의 경로를 추정하고 삼차원을 복원하는 전역 (Global) Bundle Adjustment 방법론[3]은 오차의 누적을 줄이는 가장 좋은 방법론이다. 하지만 이 방법은 온라인으로 처리를 해야 하는 로봇 주행 기술 및 실시간 방송용 콘텐츠 제작 기술등과 많은 응용 분야에 적용이 불가능하다. 하지만 제안된 방법은 우수한 성능을 보일 뿐만 아니라 온라인으로 처리를 할 수 있는 장점이 있기 때문에 위에서 언급했던 다양한 응용분야에 적용이 가능하다. 그리고 제안된 필터링 방법의 경우 기존의 확률적 필터링의 근본적인 문제를 해결할 수 있음을 보였다.

〈그림 10〉은 대전 어은동 부근의 실외 환경에서 획득한 영상으로 삼차원 복원한 결과를 보여준다. 카메라가 이동한 거리는



(a) 입력 영상들



(b) 삼차원 복원 결과

그림 10. 어은동 데이터셋을 이용한 삼차원 복원

대략 400m 정도 이고 이때 획득한 영상 4500장으로부터 카메라의 위치를 추정하고 주변환경에 대한 삼차원을 복원한 결과를 보여주고 있다. 조밀한(dense) 삼차원 복원을 위해서 카메라의 위치와 영상을 이용하여 CMVS[14]를 통하여 렌더링을 수행하였다.

V. 결론

본 논문에서는 초다시점 영상 생성 및 합성을 위한 온라인 삼차원 복원 기술에 대해 소개하였다. 기존의 방법은 영상으로부터 카메라의 점진적 (Incremental) 위치 변화를 추정하는 방법과 이 때 발생하는 오차를 줄이기 위해서 최적화 기법인 Bundle Adjustment를 채용하였다. 하지만 대공간 환경에서는 오차가 발생하면 누적이 되기 때문에 최적화 기반의 기술들

은 근본적인 해결책일 수 없다. 이에 반해 카메라의 위치와 주변 3차원 지도에 대한 확률 분포를 추정하는 확률적 필터링 방법은 카메라의 움직임 모델과 센서의 관측치 모델을 어떻게 정의하느냐에 따라 그 성능이 좌우되고, 확률 분포의 불확실성(Uncertainty)이 점차 증가하여 필터가 발산하는 필터의 일관성 문제가 발생한다. 이 문제를 해결하기 위해서 제안된 확률적 필터링 방법은 영상으로부터 상대적인 움직임 분포를 추정함으로써 카메라의 움직임 모델을 영상을 이용하여 적응적으로 정의할 수 있기 때문에 정확한 확률 모델링이 가능하다. 센서의 데이터 없이 현재 상태를 예측함으로써 불확실성이 증가하는 원인이 되고, 센서 데이터를 이용하여 예측된 분포를 갱신함으로써 불확실성이 감소하게 된다. 제안된 키프레임(keyframe) 기반의 방법론은 예측단계를 획기적으로 감소시키고, 갱신단계는 기존의 순차적 방법과 같은 횟수를 유지하기 때문에 기존 방법 대비 필터의 일관성이 향상되는 장점을 가지고 있다. 제안된 방법론의 정확성은 대공간에서 위치 추정의 오차율을 계산함으로써 측정하였고, 현재 state-of-the-art의 최적화 기반 방법론인 지역 Bundle Adjustment보다 우수한 성능을 보이며, 기존 확률적 필터링 보다 우수한 정확성을 보인다.

Acknowledgement

본 연구는 미래창조과학부 ‘범부처 Giga KOREA 사업’의 일환으로 수행하였음. [GK13D0200, 실시간 인터랙션을 제공하는 초다시점 단말 기술 개발]

참고 문헌

- [1] A. Banno, T. Masuda, T. Oishi and K. Ikeuchi, Flying Laser Range Sensor for Large-Scale Site-Modeling and Its Applications in Bayon Digital Archival Project, *International Journal of Computer Vision*, vol. 78, pp. 207–222.
- [2] J. Kim and I. S. Kweon, Vision-based navigation with pose recovery under visual occlusion and kidnapping, *IEEE International Conference on Robotics and Automation*, 2010.
- [3] Noah Snavely, Steven M. Seitz, Richard Szeliski, Photo tourism: Exploring photo collections in 3D, *ACM Transactions on Graphics (SIGGRAPH Proceedings)*, vol. 25, pp. 835–846, 2006
- [4] B. Triggs, P. McLauchlan, R. Hartley and A. Fitzgibbon, Bundle adjustment – a modern synthesis, *LNCS (Vision Algorithms: Theory and Practice)*, 2000.
- [5] E. Mouragnon, M. Lhuillier, M. Dhome, F. Dekeyser, and P. Sayd, Real Time Localization and 3D Reconstruction, *IEEE Conference on Computer Vision and Pattern Recognition*, pp.363 – 370, 2006.
- [6] G. Klein and D. Murray, Parallel Tracking and Mapping for Small AR Workspaces, *International Symposium on Mixed and Augmented Reality*, 2007.
- [7] A. J. Davison, I. D. Reid, N. Molton, O. Stasse, MonoSLAM: Real-Time Single Camera SLAM, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, 1052–1067, 2007.
- [8] M. Montemerlo, W. Whittaker and S. Thrun, FastSLAM: A factored solution to the simultaneous localization and mapping problem, In *Proceedings of the AAAI National Conference on Artificial Intelligence*, pp. 593–598, 2002.
- [9] R. Sim, P. Elinas, M. Griffin and J. J. Little, Vision-based SLAM using the Rao-Blackwellized Particle Filter, *IJCAI Workshop Reasoning with Uncertainty in Robotics*, 2005.
- [10] C. Kim, R. Sakthivel and W. K. Chung, Unscented FastSLAM: A Robust and Efficient Solution to the SLAM Problem, *IEEE Transactions on Robotics*, vol. 24, no. 4, pp. 808–820, 2008.
- [11] R. Haralick, C. Lee, K. Ottenberg and M. Nolle, Review and Analysis of Solutions of the Three Point Perspective Pose Estimation Problem, *International Journal of Computer Vision*, vol. 13, no. 3, pp. 331–356, 1994.
- [12] D. Nister, O. Naroditsky and J. Bergen, Visual Odometry, *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 652–659, 2004.
- [13] M. Fischler and R. Bolles, Random Sample Consensus: a Paradigm for Model Fitting with Application to Image Analysis and Automated Cartography, *Communications ACM*, 1981.
- [14] Yasutaka Furukawa and Jean Ponce Accurate, Dense, and Robust Multi-View Stereopsis *IEEE*

Transactions on Pattern Analysis and Machine Intelligence, Vol. 32, Issue 8, Pages 1362-1376, August 2010.

약 력



김 정 호

2004년 경북대학교 공학사
2006년 한국과학기술원 공학석사
2013년 한국과학기술원 공학박사
2012년~현재 전자부품연구원 멀티미디어 IP 연구 센터 선임연구원
관심분야: 로봇비전, SLAM, 물체 추적



김 제 우

1997년 서울시립대학교 공학사
1999년 서울시립대학교 공학석사
1999년~현재 전자부품연구원 멀티미디어IP연구센터 책임연구원
관심분야: 멀티미디어 신호처리 및 차세대 방송 기술



권 인 소

1981년 서울대학교 공학사
1983년 서울대학교 공학석사
1990년 Carnegie Mellon University 공학박사
1991년~현재 한국과학기술원 교수
관심분야: 컴퓨터 비전, 3D 비전, 카메라 보정, 물체 인식 및 분류, SLAM