

모바일 매핑 시스템의 GEO 로컬라이제이션

The GEO-Localization of a Mobile Mapping System

전재춘¹⁾

Jaechoon Chon

Abstract

When a mobile mapping system or a robot is equipped with only a GPS (Global Positioning System) and multiple stereo camera system, a transformation from a local camera coordinate system to GPS coordinate system is required to link camera poses and 3D data by V-SLAM (Vision based Simultaneous Localization And Mapping) to GIS data or remove the accumulation error of those camera poses. In order to satisfy the requirements, this paper proposed a novel method that calculates a camera rotation in the GPS coordinate system using the three pairs of camera positions by GPS and V-SLAM, respectively. The propose method is composed of four simple steps; 1) calculate a quaternion for two plane's normal vectors based on each three camera positions to be parallel, 2) transfer the three camera positions by V-SLAM with the calculated quaternion, 3) calculate an additional quaternion for mapping the second or third point among the transferred positions to a camera position by GPS, and 4) determine a final quaternion by multiplying the two quaternions. The final quaternion can directly transfer from a local camera coordinate system to the GPS coordinate system. Additionally, an update of the 3D data of captured objects based on view angles from the object to cameras is proposed. This paper demonstrated the proposed method through a simulation and an experiment.

Keywords : Coordinate transformation, V-SLAM, GPS, Quaternion

초 록

모바일 매핑 시스템 또는 로봇이 GPS (Global Positioning System)와 다중 스테레오 카메라만 탑재 할 경우, V-SLAM(Vision based Simultaneous Localization And Mapping) 에 의한 카메라 외부표정과 3차원 데이터들 GIS데이터와 연계 또는 카메라 외부표정의 누적에러를 제거하기 위해 극부 카메라 좌표계에서GPS (Global Positioning System) 좌표계로 변환이 필요로 한다. 이 요구사항을 만족 시키기 위해, 본 논문은 GPS와 V-SLAM 에 의한 3쌍의 카메라의 위치를 이용하여 GPS좌표계에서 카메라 자세를 계산하는 새로운 방법을 제안하였다. 제안한 방법은 간단한 4단계로 구성되어 있다: 1) 각 3개의 카메라 위치에 기반한 두 평면 법선벡터가 병렬이되도록 하는 사원수 (quaternion)를 계산한다, 2) 계산된 사원수를 통하여 V-SLAM에 의한 3개의 카메라 위치를 변환한다, 3) 변환된 위치에서 두번째 또는 세번째 점이 GPS에 의한 점과 일치하도록 하는 두번째 사원수를 계산한다, 4) 두 사원수의 곱을 통하여 최종 사원수 결정한다. 최종 사원수는 극부 카메라 좌표계에서 GPS좌표계로 변환할 수 있다. 추가적으로, 촬영된 물체 위치에서 카메라를 보는 시야각을 기반으로 물체의3차원좌표 갱신방법을 제안하였다. 본 논문은 제안한 방법을 시뮬레이션과 실험을 통하여 증명하였다.

핵심어 : 좌표변환, V-SLAM, GPS, 사원수

1. 서 론

LBS (Location Based Service)의 일부로서, 모바일 폰 카메라를 통해 사용자의 위치와 자세를 GPS (Global

Positioning System) 좌표계로 변환하여 3차원 지도 정보와 3차원 태그정보를 모바일 카메라 영상에 중첩하는 연구가 활발히 이루어지고 있다. 사용자의 모바일 폰으로 촬영된 동영상에서 하나의 영상 프레임을 3차원 지도정

1) 정회원 · EARTHMININE INC(E-mail: jjc7151@gmail.com)

보가 저장되어 있는 서버로 전송한 후, 서버 내에서 영상처리를 통한 GPS좌표계에서 영상의 로컬라이제이션이 수행되리라 예상된다. 로컬라이제이션 정보를 모바일 폰으로 전송후, 로컬라이제이션된 영상을 기준으로 촬영되고 있는 영상 프레임의 로컬라이제이션은 모바일 폰 내부에서 수행하게 되리라 예상된다. 이 과정에서, 모바일 카메라의 위치와 자세계산을 위해 GPS 좌표계에서 생성된 3차원 지도나 도로상에서 생성된 3차원 파노라마가 필요하다. 인터넷 상에서 제공되는 3차원 파노라마는 구글(google inc.) 스트리트 뷰(maps.google.com)와 어스마인(earthmine inc.)의 wildstylecity.com이 있다(전재춘과 Shibasaki, 2008; Früh 외 2인, 2005; Zhao와 Shibasaki, 2003). 구글 스트리트 뷰는 Früh외 2인이 개발한 레이저 스캐너(Laser range finder) 기반 방법을 채택하고 있으며, 어스마인은 NASA JPL(Jet Propulsion Laboratory California Institute of Technology)에서 개발한 스테레오 광각렌즈(Fish-Eye lens) 카메라 시스템을 채택하고 있으며, 본 저자가 개발한 최적경로 방법기반 seamless 구 파노라마 생성이 적용되어 있다(Chon 외 2인, 2009; Chon 외 3인, 2007). 추가적으로, 사람인식 블러링은 Felzenszwalb가 제안한 훈련된 다중 모델기법이 적용되어 있으며, 인식률(황인, 백인, 흑인)은 84.5%을 결과를 보이고 있다(Felzenszwalb외 3인, 2009). 차량번호 인식은 3D 파노라마 기반 칼라정보를 융합하는 알고리즘이 적용되어 있으며, 93%의 인식률을 보이고 있다. 레이저 스캐너 기반 방법은 모바일 매핑 시스템의 움직임 속도에 의해 3차원 데이터의 밀도가 결정되는 단점이 있으며, 이 시스템의 2차원 모션 모델에 의한 각 레이저 스캐너 데이터간의 얼라이먼트의 문제와 영상 촬영당시의 자세의 불확실성으로 3차원 데이터와 매핑의 문제가 있다. 어스마인의 모바일 매핑 시스템의 경우, 다중 스테레오 카메라 시스템 기반으로 3차원 데이터를 생성하기 때문에 파노라마 화질이 좋고, 3차원 데이터 밀도가 높으며, 3차원 데이터와 매핑문제가 발생하지 않는다. 하지만, 50m 이상의 물체의 정밀도가 떨어지는 단점이 있으며, RTK(Real Time Kinematic) GPS/IMU(Inertial Measuring Unit)에 로컬라이제이션을 의존하고 있기 때문에 GPS 수신 불가능지역과 수신 불량지역에서는 3차원 파노라마 간의 중첩오차가 발생하는 문제가 있다. 본 논문은 건물내와 같이 GPS 수신불능 지역과 수신 불량지역에서 모바일 매핑 시스템의 로컬라이제이션을 하기 위한 것이며, 모바일 매핑 시스템에 탑재된 GPS와 다중

스테레오 카메라만을 이용하여 모바일 매핑 시스템 로컬라이제이션을 위한 것이다.

본 저자가 개발한 카메라 모델(전재춘, 2008; 전재춘과 Sastry, 2008, 2007)에 기반한 V-SLAM(Vision based Simultaneous Localization And Mapping)(Se 외 2인, 2005)을 통하여 계산된 카메라 외부표정과 SIFT(Scale Invariant Feature Transform) 특징점(Lowe, 2004)의 3차원 데이터를 GPS 좌표계로 변환 하기 위해서, 본 논문은 영상 촬영시 동시에 GPS데이터가 수신된 3쌍의 영상을 선택 후, GPS와 V-SLAM에 의한 각 3개의 카메라 위치 기반한 두 평면을 병렬이 되도록 하는 회전과 두평면이 일치 하도록 하는 회전을 통하여 좌표계 변환을 수행 하는 것이다. 각 회전은 사원수(Quaternion)로 표현할 수 있으며(Hamilton, 1844), 두 사원수의 곱에 의한 최종 사원수는 직접적으로 극부 카메라 좌표계에서 GPS좌표계로 변환하는 인자이다. 추가적으로, 본 논문은SIFT 특징점의 공간상 기준으로부터 카메라를 보는 시야각들을 융합하여 SIFT 특징점의 3차원 좌표를 갱신하는 방법을 제안 하였다.

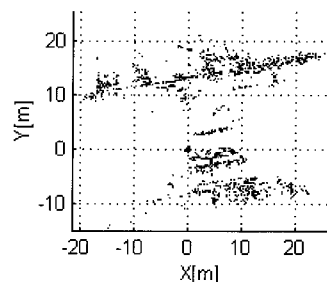
본 논문의 2장은 모바일 매핑 시스템의 로컬라이제이션, 3장은 제안한 좌표변환 방법, 4장은 시뮬레이션, 5장은 실험, 6장은 결론으로써 각각 구성되어 있다.

2. 모바일 매핑 시스템의 로컬라이제이션

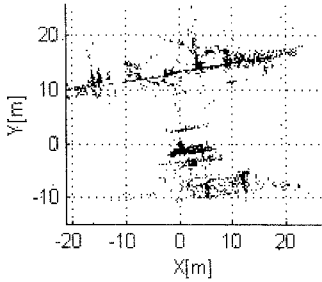
본장은 본 논문에서 사용된 V-SLAM기반 모바일 매핑 시스템의 로컬라이제이션에 대해 소개한다.

2.1 스테레오 카메라 기반 상대표정

첫번째 영상 프레임을 기준 좌표계로 하여 두번째 영상프레임의 상대표정과 두번째 영상 프레임을 기준 좌표계로 하여 첫번째 영상프레임의 상대표정을 본 저자가 개발한 카메라 모델식을 기반한 RANSAC(RANdom



(a) 이전 영상 프레임



(b) 현재 영상프레임

그림 1. SIFT 특징점과 3차원 좌표

SAMPLE Consensus)을 이용하여 최소의 화소에러를 가진 상대표정을 계산하였다 (전재춘, 2008; 전재춘과 Sastry, 2008, 2007). RANSAC 은 카메라 모델식이 필요로 하는 특징점(tie-point, feature correspondence)를 임의로 선택하여 화소 에러가 최소가 되는 쌍을 선택하는 방법이다 (Fischler 와 Bolles, 1981). 이 알고리즘 반복회수 $N = \log(1-p) / \log(1-(1-\epsilon)^s)$ 는 한쌍에 선택되어야 할 특징점들의 수 S , 잘못된 정합된 특징점들의 수의 점유율 ϵ , 요구 성공률 p 로 결정된다. 본 논문에서는 반복회수를 $s=3$, $\epsilon = 50\%$, $p=99\%$ 로 설정 했지만, 영상투영오차가 2화소 이내일 때까지 수행 하였다. 계산과정에서 필요로 하는 특징점의 추출과 정합은 Low가 개발한 SIFT 특징점(tie-point, feature correspondence)을 이용 하였다. SIFT 특징점은 영상 이동, 회전, 스케일과 밝기변화 대하여 불변한 성질을 가지고 있기에, 컴퓨터 비전, 로보비전등 널리 사용되고 있다. 본 논문에서는 SIFT 알고리즘 계산 속도를 비약적으로 향상 시키기 위해 GPU(Graphics Processing Unit)을 사용하고 있다 (Sinha 외 3인, 2006). SIFT 특징점의 3차원 좌표는 공면조건식(coplanar condition)을 이용하여 추정하고 있다. 그림 1은 4쌍의 스테레오 광각렌즈 카메라 시스템으로 촬영된 영상들에서 추출된 SIFT특징점의 3차원 데이터이다.

2.2 상대표정 누적

각 영상 프레임에서 추정된 카메라 상대표정 (이동 t , 회전 R)을 이용하여 외부표정 요소누적은 다음과 같이 표현된다.

$$\begin{aligned}
 t_{13} &= R_{12}t_{23} + t_{12} \\
 t_{14} &= R_{23}t_{34} + t_{13} \\
 &\vdots \\
 t_{1n} &= R_{n-2n-1}t_{n-1n} + t_{1n-1} \\
 \therefore t_{1n} &= \sum_{i=3}^n R_{i-2i-1}t_{i-1i} + t_{12}
 \end{aligned} \tag{1}$$

여기서, $(t_{i,i+1}, R_{i,i+1})$ 는 i 번째 영상 프레임 좌표계에서 $i+1$ 번째 영상 프레임의 외부표정 요소를 의미한다. 각 영상 프레임의 회전 행렬간의 곱을 통하여 현재 자세를 표현하는 회전행렬을 계산할 수 있다.

$$R_{1n} = \prod_{i=2}^n R_{i-1}$$

이동 t 와 회전 R 을 하나의 행렬로 표현하면 4×4 행렬이 된다. $m = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix}$ 이용하여 식 (1)을 다음과 같이 간략하게 표현 할 수 있다.

$$m_{1n} = \prod_{i=2}^n m_{i-1i}$$

2.3 사원수를 이용한 상대표정 누적

본 논문은 Hamilton이 개발한 사원수 기반 상대표정을 누적하였다 (Hamilton, 1844). 일반적으로 사용되는 Euler 좌표계에서 회전행렬은 짐발락(Gimbal Lock)문제와 공간상에서 물체회전 표현이 불편한 단점있다. 사원수 $q(w, x, y, z)$ 는공간상에서 물체회전 표현이 간단하며 짐발락 문제가 발생하지 않고, 계산량이 적기 때문에 컴퓨터 그래픽스, 로봇틱스, 위성 자세 제어 등 널리 사용되고 있다.

$$q = w + xi + yj + zk = w + (x, y, z) = \cos\left(\frac{\theta}{2}\right) + \vec{u} \sin\left(\frac{\theta}{2}\right) \tag{4}$$

여기서, \vec{u} 는 회전축 벡터, θ 는 회전각도이다. 각 영상 프레임으로 계산된 회전행렬 R 은 다음식을 통하여 사원수로 바꿀 수 있다.

$$q_w = \frac{1}{4s}, \quad q_x = (R_{32} - R_{23})s, \quad q_y = (R_{13} - R_{31})s, \quad q_z = (R_{21} - R_{12})s \tag{5}$$

여기서, $s = \frac{1}{2\sqrt{1+R_{11}+R_{22}+R_{33}}}$ 이며, $1+R_{11}+R_{22}+R_{33} > 0$ 일 경우이다. 다음식은 이와 반대로 사원수를 회전 행렬 변환하는 것이다.

$$R = I + 2QQ + 2SqQ \tag{6}$$

여기서, $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $Q = \begin{bmatrix} 0 & -q_x & q_y \\ -q_x & 0 & -q_z \\ -q_y & q_z & 0 \end{bmatrix}$, $Sq = qw$ 이다. 식(1)

을 사원수로 표현하면 다음과 같다.

$$t_{1n} = \sum_{i=2}^n q_{i-2i-1} t_{i-2} q_{i-2i-1}^{-1} + t_{12} \tag{7}$$

여기서, q' 는 공액 (conjugate number) 사원수로써 $q^{-1} = W(x, y, z)$ 이며, $t' = qq^{-1} = \left\{ \cos\left(\frac{\theta}{2}\right) + \vec{u} \sin\left(\frac{\theta}{2}\right) \right\} \left\{ \cos\left(\frac{\theta}{2}\right) - \vec{u} \sin\left(\frac{\theta}{2}\right) \right\} = (\vec{u} - \vec{u}(\vec{u} \cdot \vec{t})) \cos\left(\frac{\theta}{2}\right) + (\vec{u} \times \vec{t}) \sin\left(\frac{\theta}{2}\right) + \vec{u}(\vec{u} \cdot \vec{t})$ 이다. 식(2)를 사원수로 표현하면 다음과 같다.

$$q_{1n} = \prod_{i=2}^n q_{i-1i} = q_{12} q_{23} \cdots q_{n-2n-1} q_{n-1n} \tag{8}$$

여기서, 사원수 곱 $q = q_1 q_2$ 는 $qw = qw_1 w_2 - x_1 x_2 - y_1 y_2 - z_1 z_2$, $qx = w_1 x_2 + x_1 w_2 + y_1 z_2 - z_1 y_2$, $qy = w_1 y_2 - x_1 z_2 + y_1 w_2 + z_1 x_2$, $qz = w_1 z_2 + x_1 y_2 - y_1 x_2 + z_1 w_2$ 이다.

2.4 Vision based Simultaneous Localization And Mapping (V-SLAM)

광속블록조정법(Bundle Block Adjustment)은 영상프레임 간의 소수의 타이포인트(tie-point, feature correspondence), 극소수의 지상기준점을 이용하여 모든 타이포인트의 3차원 위치와 모든 영상프레임의 외부표정을 한꺼번에 행렬을 통하여 추정하는 방법이다. 지상기준점만 주어지지 않고, 비데오 영상처럼 많은 영상프레임과 타이포인트가 존재 할 경우, 현존하는 개인용 컴퓨터에 적용이 용이하지 않다. 로봇 비전분야에서는 이경우를 효율적으로 다루기 위해 SLAM방법을 널리 사용하고 있다. 두 영상 프레임간의 관계만을 고려한 식(7)이 직접적으로 모든 영상프레임의 외부표정을 결정하게 되면, 영상프레임의 외부표정 누적오차가 다소 빠르게 증가 될 것이다. 그러나, 동일 물체(Natural landmark or tie-point)가

여러 영상 프레임에서 촬영된다면, 가장 근접한 거리에서 측정된 좌표가 오차를 적게 포함하게 되고, 여러 영상 프레임에서 이 물체를 보는 시각이 다르게 되므로 측정오차를 줄일 수 있게 된다. 그러므로 본 논문은 이러한 문제를 해결하기 위하여 V-SLAM (Vision based Simultaneous Localization And Mapping)을 사용 하였다. V-SLAM 이란 갱신된 물체의 3차원 좌표기반으로 외부표정을 다시 계산하는 방법이다 (Milford 와 Gordon, 2008; Se 외 2인, 2005; Thrun et al, 2000). V-SLAM에 관련된 연구들은 기본적으로 로봇이나 모바일 매핑 시스템의 모션을 위한 엔코더등 추가적인 센서를 이용하고 있으며 영상 프레임간의 촬영 단위가 수 센티미터 이다. 센서와 영상 데이터 융합을 위해 EKF(Extended Kalman Filter)을 이용하여 상대표정 갱신할 수 있다 (Civera, 2009). 본 논문에서는 추가적인 센서를 이용하지 않고 2m, 5m 또는 10m 단위로 영상이 촬영되기 때문에, EKF의 동역학적 특성을 고려한 상태 방정식이 의미가 없기 때문에 EKF기반 외부표정 요소 갱신부분은 사용하지 않는다. 사후확률(posterior probability) 방법을 기반한 SLAM에서 3차원 데이터 갱신은 입자필터(particle filter)와 미리 설정된 격자 지도를 이용하고 있다(Thrun 외 3인, 2000). 촬영간격이 먼 거리의 경우, 하나의 근접한 물체가 여러 영상 프레임에서 촬영되지 않고 두 영상 프레임 또는 세개의 영상 프레임만 촬영되기 때문에 사후확률방법 적용이 용이하지 않다. 본 논문은 물체 위치를 중심으로 스테레오 카메라 시스템을 바라보는 시각 θ 로 융합하여 물체의 3차원 좌표를 추정하는 방법을 제안 하였다.

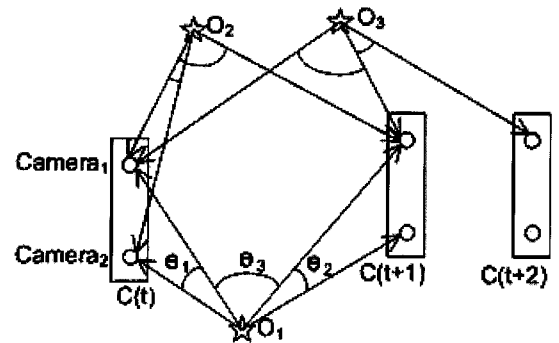


그림 2. 3. 시각융합방법을 통한 물체의 3차원 좌표 갱신

그림 2에서 물체 O₁는 첫번째 두번째 스테레오 카메라

3. 좌표변환

시스템의 위치 $C(t)$ 와 $C(t+1)$ 에서 3차원 좌표 P_1 와 P_2 가 측정된 경우이다. 이 경우, $IC(t+1)-C(t)$ 인 기선(baseline)을 이용하여 물체 O_1 의 3차원 좌표 P_3 를 다시 계산할 수 있으며, $P_{=1,2,3}$ 기반으로 최종 P 를 결정할 수 있다. 다음 식을 통하여 이 3개의 측정 좌표를 융합하였다.

$$P = \frac{\theta_1 P_1 + \theta_2 P_2 + \theta_3 P_3}{\theta_1 + \theta_2 + \theta_3} \quad (9)$$

여기서, 각 $P_{=1,2,3}$ 는 극부 $C(t)$ 좌표계로 변환된 물체의 3차원 좌표이다. 시각 θ 크면 클 수록, 스테레오 영상에서 시각차이(disparity)가 크거나, 스테레오 카메라 시스템의 기선이 크다는 것을 의미한다. 또한 물체와 시스템 간의 거리가 멀어 질수록 시각 θ 가 작아진다. 그림 2에서 물체 O_2 의 경우는 한 스테레오 카메라 시스템에서 측정된 것과 $C(t)$ 와 $C(t+1)$ 에서 측정된 것을 융합하는 경우이고, O_3 는 스테레오 카메라 시스템에서 측정되지 못하고 $C(t), C(t+1), C(t+2)$ 에서만 측정된 것을 의미한다.

모바일 매핑 시스템이 GPS 수신 불량지역과 건물내와 같이 GPS수신 불량지역을 통과할때, 스테레오 카메라 시스템에 의한 카메라 외부표정의 에러누적 제거가 필요하며, 추출된 3차원 데이터를 GIS데이터와 연계하기 위해 GPS좌표계 변환이 필요하다. 본 논문은 GPS와 V-SLAM에 의한 3쌍의 카메라 위치점을 이용한 좌표변환 방법을 제안 하였다. 이 3쌍에서 영상간의 거리는 근접거리가 될 수도 있고, 원거리일 수도 있다. 기존 영상은 모바일 매핑 시스템이 교차로와 같이 일시 정지한 경우이며, 수신지역에서 수신불량지역이나 수신 불량지역으로 진입전과 빠져 나온 이후에 각각 3쌍의 영상이 선택되어야 한다. 2쌍의 영상간의 외부표정을 GPS 좌표계로 매핑할 경우, 모바일 매핑 시스템의 Pitch와 Yaw를 결정할 수 있으면, 1쌍을 추가 할 경우 나머지 Roll을 결정할 수 있다.

그림 3은 제안한 방법의 진행 과정을 나타낸 것이다.

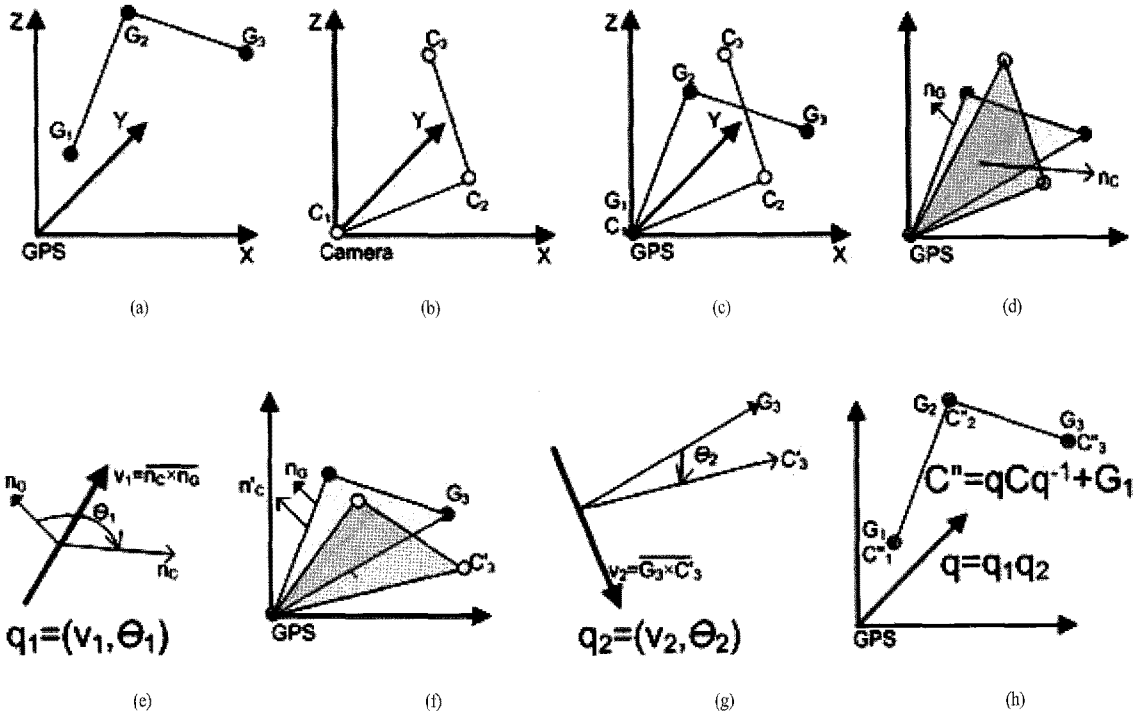


그림 3. VO(visual odometry)와 GPS 위치의 3쌍을 이용한 좌표변환; (a)GPS에 의한 카메라 위치, (b) VO에 의한 카메라 위치, (c) GPS좌표계에서 3쌍의 카메라 위치, (d) Roll 모션을 위한 두 법선벡터, (e) 두 병렬 법선벡터, (f) 두 병렬 평면, (g) VO에 의한 위치를 GPS에 의한 위치상에 매핑, (h) GPS 좌표계로 변환된 최종 카메라 자세.

그림 3(d)와 같이 일차적으로 GPS와 V-SLAM에 의한 각 3 점을 기반으로 평면 방정식을 만든다. 그림 3(e) 각 평면의 법선벡터의 일치화사원수를 이용한다. 사원수 q_1 는 두 법선벡터 nG 와 nC 의 내적식 $\cos(nG \cdot nC)^{-1}$ 으로 계산된 회전 각 θ_1 와 두 벡터의 외적 $nG \times nC$ 으로 계산된 회전축으로 결정된다. 이 과정은 GPS좌표계에서 모바일 매핑의 Roll 회전을 결정하는 것이다. 그림 3(g)는 모바일 매핑 시스템의 나머지 회전인 Pitch와 Yaw를 결정하는 사원수 q_2 를 결정하는 것을 나타낸 것이다. 사원수 q_1 에 의해 변환된 카메라 위치 C' 에서 벡터 C'_1, C'_2 또는 C'_1, C'_3 을 벡터 G_1, G_2 또는 G_1, G_3 와 일치화하는 조건으로 q_2 를 계산한다. q_2 계산은 q_1 계산 과정과 동일하다. 카메라 좌표계에서 GPS좌표계로 변환하는 최종 사원수 $q = q_1, q_2$ 는 계산된 두 사원수의 곱을 통해 계산하며, 변환식은 다음과 같다.

$$C'' = pCp^{-1} + G1 \tag{10}$$

4. 시뮬레이션 결과

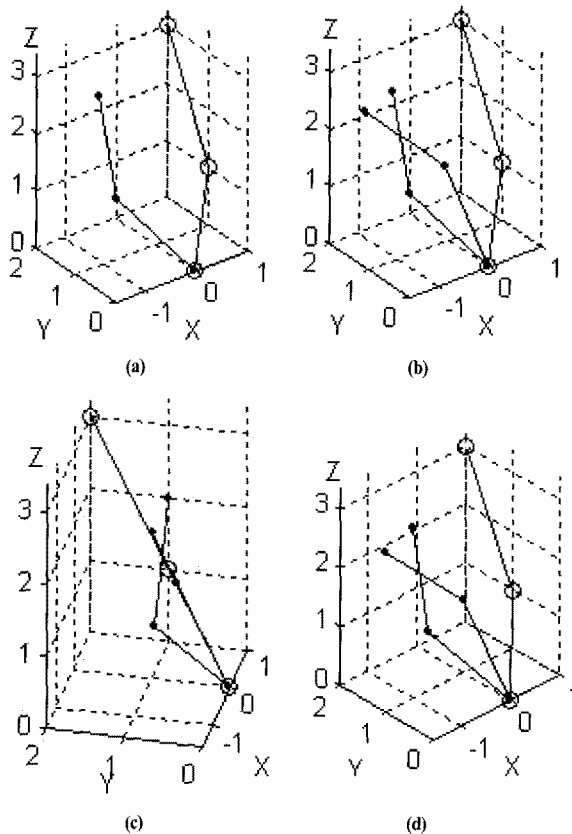
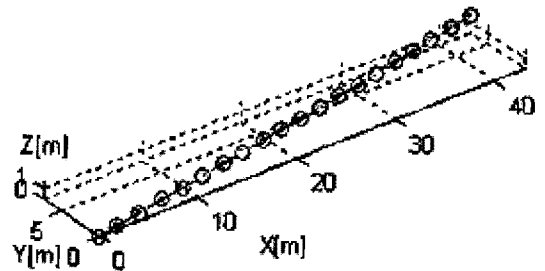


그림 4. 제안한 방법의 시뮬레이션 결과

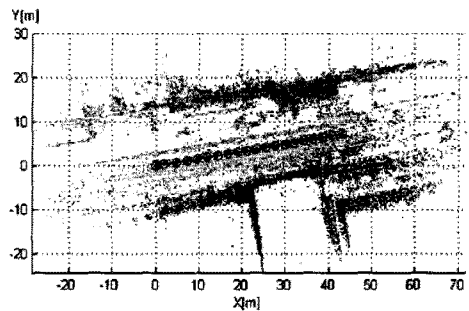
본 논문에서는 3차원 좌표와 회전은 각각 미터[m]와 각도[Deg] 단위로 사용한다. 식(10)을 증명 하기 위해, GPS좌표계에서 3점의 좌표 $G[0\ 0\ 0; 1\ 1\ 1; 1\ 2\ 3]$ 와, 이 좌표를 $[10\ -5\ 100]$ 만큼 회전한 카메라 좌표계에서 좌표 $C[0\ -0.9543\ -1.5497; 0\ 0.7408\ 0.4461; 0\ 1.2412\ 3.3763]$ 를 사용 하였다. 그림 4(a)은 각 3쌍의 좌표를 나타내고 있으며, 두 평면의 법선벡터의 일치화를 위한 사원수 $q_1[0.7114\ -0.3213\ 0.1435\ 0.6084]$ 에 의한 카메라 좌표는 그림 4(b)에 추가적으로 표시되어 있다. 그림 4(c)는 그림 4(b)를 다른 시점에서 본 것으로서 GPS좌표와 변환된 카메라 좌표가 병렬인 것을 확인할 수 있다. 그림 4(d)는 추가적으로 계산된 $q_2[0.8952\ 0.1819\ -0.3639\ 0.1819]$ 와 q_1 를 곱한 $q[0.6368\ 0.0893\ 0.0388\ 0.7648]$ 를 이용한 최종 카메라 좌표를 추가적으로 도시한 결과이다.

5. 실험 결과

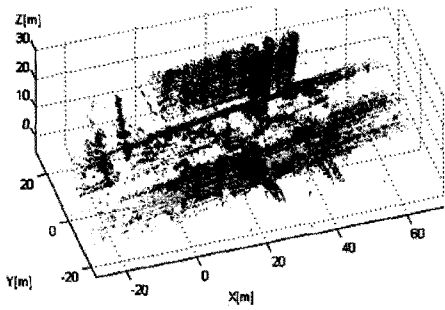
본 논문은 실험을 통하여 제안한 방법을 증명 하기 위해, 스테레오 카메라 캘리브레이션(전재춘 2007, 2008)과 카메라와 IMU캘리브레이션(전재춘, Shibasaki, 2008)을 위한 2m간격으로 촬영된 20개의 영상 프레임을 사용했다. 카메라 상대표정 계산시, 계산 오차를 줄이기 위해



(a) RTK('o')와 V-SLAM('·')에 의한 카메라 위치



(b) RTK(적색점, 'o')과 V-SLAM(초록색, '·')에 의한 3차원 점군과 카메라 위치

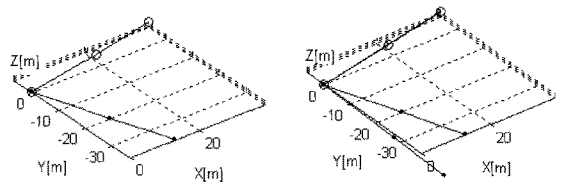


(c) 다른 시점에서의 3차원 점군과 카메라 위치

그림 5. V-SLAM에 의한 카메라 위치와 3차원 점군(point cloud)

40m이상의 거리인 SIFT특징점은 제외를 하였다. 하지만, 식9에 의한 3차원 좌표 갱신을 수행 했다. 그림 5는 V-SLAM에 의해 계산된 카메라 외부표정에서 첫번째 영상프레임의 자세를 RTK GPS/IMU로 계산된 것을 이용한 결과로써, 마지막 영상 프레임의 누적 위치오차는 0.1733m 이며, 각 좌표축에서 누적 위치오차는[0.1310 0.0650 0.0930] 이었다.

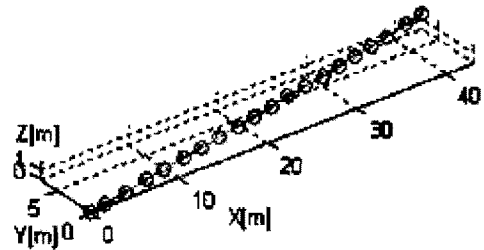
그림 6(a)은 GPS좌표계에서 첫번째 영상 프레임의 자세가 주어지지 않았을때, SIFT 특징점의 3차원 좌표들이 다. 그림 6(b)는 제안한 방법을 적용하기 위해, 임의로 선택된10번째와 18번째의 영상프레임의 외부표정이며, 그림 6(c)는 제안한 방법을 적용할때, 최종 외부표정을 추가적으로 도시 한 결과이다. RTK GPS/IMU와 V-SLAM에 의해 계산된 카메라 위치점들중 선택된 3쌍의 카메라 위치점은 각각 [0 0 0; 21.1180 3.4470 0.6920 ; 38.3590 6.3880 1.4120], [0 0 0; 6.9176 -20.1396 -0.0948; 12.7723 -36.6191 0.0069]이다. RTK GPS/IMU



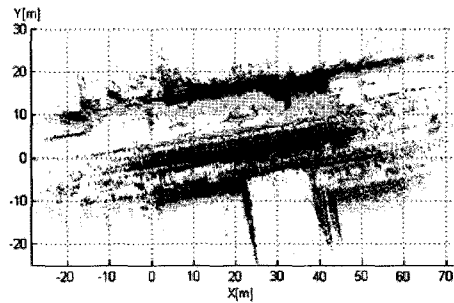
(b) 선택된 3쌍

(c) 제안한 방법에 의한 결과

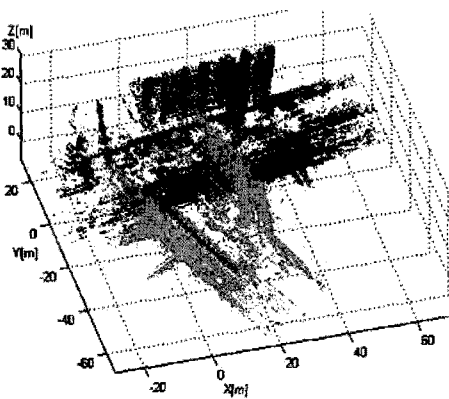
그림 6. 초기 데이터와 선택된 3쌍



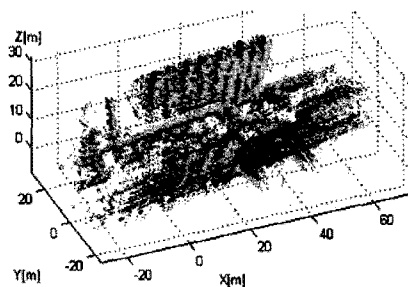
(a) RTK 'o'와 제안한 방법 '·'에 의한 카메라 위치



(b) RTK(적색점, 'o') 과 제안한 방법 (초록색, '·')에 의한 카메라 위치와 3차원 점군과 카메라 위치



(a) 중첩 3차원 점군



(c) 다른 시점에서 3차원 점군

그림 7. RTK와 제안한 방법에 의한 카메라 위치와 3차원 점군.

표 1. RTK와 제안한 방법에 의한 회전행렬 비교

RTK GPS/IMU에 의한 회전 행렬	제안한 방법에 의한 회전 행렬
[0.1714 0.9852 -0.0008; -0.9845 0.1712 -0.0388; -0.0381 0.0074 0.9992]	[0.1656 -0.9860 -0.0175; 0.9788 0.1665 -0.1195; 0.1208 0.0026 0.9927]

에 의한 영상 첫번째 프레임의 위치와 자세가 [고도: 56.249, UTM east: 564256.547, UTM north: 4191653.880, roll: 0.426, pitch: 2.185, yaw: 80.124]로써 기준좌표 [0 0 0 0.426 2.185 -80.124]로 설정하였다. 계산된 사원수는 $q_1[0.7624 -0.0401 0.0454 -0.6443]=q_2[0.8826 -0.2298 -0.2779 -0.3015]$ $q_3[0.8637 -0.0529 0.3879 -0.3173]$ 였다.

그림 7(a)은 사원수 q 에 의해 변환된 카메라 위치를 나타내고 있다. 하지만, 그림 7(b)와 7(c)와 같이 SIFT 특징점의 3차원 좌표가 오차를 보이고 있다. 표 1은 계산된 사원수 q 를 통해 계산된 회전 행렬과 RTK GPS/IMU에 의한 회전 행렬을 비교한 결과이다. 두 행렬이 다소 유사하지만, 다른 이유는 모바일 매핑 시스템의 이동이 직선 운동할때 취득된 20개의 영상 프레임 이기 때문이다. 이 경우 모바일 매핑의 roll회전 계산 오차가 발생한다. 이 오차를 줄이기 위해서는 직선 운동이 아닌 좌우 회전 운동이 필요하다.

6. 결 론

본 논문은 RTK GPS와 V-SLAM을 통하여 계산된 3쌍의 카메라 외부표정을 이용하여 GPS 좌표계에서 카메라 자세를 계산하는 방법을 제안 했다. 제안한 방법은 각 3쌍을 기반한 두 평면의 법선벡터가 병렬이 되도록 하는 사원수와 3쌍의 카메라 위치가 공간상에서 일치되도록 하는 사원수를 곱하여 최종 사원수를 계산 하는 것이다. 우리는 최종 사원수를 통하여 직접적으로 극부 카메라 좌표계에서 GPS좌표계로 변환할 수 있다. 추가적으로, 다중 기선기선기반을 통한 촬영된 물체의 3차원 좌표 갱신방법을 제안하였다.

제안한 방법은 고가의 IMU를 수입하기 어려운 우리나라 현실에 기여 할 것으로 예상되며, GPS수신이 불량한 도심지 내부와 건물 내부에서 3차원 데이터 취득 후 Geo-로컬라이제이션에 도움을 줄 것으로 예상된다.

참고문헌

전재춘 (2008), 벡터내적 기반 카메라 자세 추정, 대한원격탐사학회지, 대한원격탐사학회, 24(6), pp. 641-644.

전재춘, Sastry, S. (2008), 벡터내적기반 카메라 외부파라메타, 응용: 절대표정, 제어로봇시스템학회지, 제어로봇 시스템학회, 14(1), pp. 70-74.

전재춘, Sastry, S. (2007), 벡터내적 기반 카메라 외부 파라메타응용: 카메라와 레이저스캐너간의 캘리브레이션, 대한원격탐사학회지, 대한원격탐사학회, 23(6), pp. 553-558.

전재춘, Shibasaki, R. (2008), IMU/GPS와 카메라 좌표계 간의 물리적 관계를 위한 수학적 모델, 한국측량학회지, 한국측량학회, 제 26권, 제 6호, pp. 611-616.

Chon, J, Fuse, T., Shimizu, E, and Shibasaki, R. (2007), 3D Image Mosaicking Using Multiple Projection Planes, *IEEE Tr. on SMC-B*, 37(3), pp. 771-783.

Chon, J Kim, H, and Lin, C., (2009), Seam-line determination for image mosaicking: A technique minimizing the maximum local mismatch and the global cost. *ISPRS Journal of Photogrammetry and Remote Sensing*, doi:10.1016/j.isprsjprs.2009.09.001

Civera, Javier, Oscar G. Grasa, Andrew J. Davison and J. M. M. Montiel (2009), 1-Point RANSAC for EKF-Based Structure from Motion, *IROS 2009*, CD-ROM.

Felzenszwalb, P. F., Girshick, R.B., McAllester, D., Ramanan, D., (2009), Object Detection with Discriminatively Trained Part Based Models, To appear in *IEEE Trans. on PAMI*.

Fischler, M.A., Bolles, R.C. (1981), RANSAC random sampling consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Communications of ACM*, 26, pp. 381-395.

Früh ,C., Jain,S., and Zakhor, A. (2005), Data Processing Algorithms for Generating Textured 3D Building Facade Meshes from Laser Scans and Camera Images,

- International Journal of Computer Vision*, 61(2), pp. 159-184.
- Hamilton, William Rowan (1844), On quaternions, or on a new system of imaginaries in algebra, *Philosophical Magazine*, 25(3), pp. 489-495.
- Lowe, David G. (2004), Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, 60(2), pp. 91-110.
- Milford, Michael J., Gordon F (2008), Mapping a Suburb With a Single Camera Using a Biologically Inspired SLAM System, *IEEE Tr. on Robotics*, 24(5), pp.1038-1053.
- Se, S. Lowe, D.G. Little, J. J. (2005), Vision-based global localization and mapping for mobile robots, *IEEE Tr. on Robotics*, 21(3), pp. 364-375.
- Sinha, Sudipta N, Jan-Michael Frahm, Marc Pollefeys and Yakup Genc (2006), GPU-Based Video Feature Tracking and Matching, *EDGE 2006*, workshop on Edge Computing Using New Commodity Architectures, Chapel Hill, May 2006.
- Thrun, S., Fox, D., Burgard, W., and Dellaert, F. (2001), Robust Monte-Carlo localization for mobile robots, *Artificial Intelligence*, 128(1), pp. 99-141.
- Zhao, H., Shibasaki, R. (2003), A vehicle-borne urban 3-D acquisition system using single-row laser range scanners, *IEEE Tr. on SMC-B*, 33(4), pp. 658-666.

(접수일 2009. 9. 16, 심사일 2009. 10. 12, 심사완료일 2009. 10. 14)