

논문 2012-50-3-19

후방 카메라 영상과 3차원 도로지도를 이용한 이동차량의 위치인식

(Localization of A Moving Vehicle using Backward-looking Camera
and 3D Road Map)

최 성 인*, 박 순 용**

(Sung-In Choi and Soon-Yong Park)

요 약

본 논문에서는 실외 도로환경에서 주행하는 차량의 위치를 추정하기 위한 비주얼 오도메트리 기술을 제안한다. 제안하는 방법은 운전자의 이동계획에 따라 차량의 초기위치에서 원거리에 위치한 특정 목적지를 방문한 후 지나온 경로를 따라 다시 초기위치로 정확하게 복귀해야 하는 차량의 위치인식을 위해 사용된다. 위치인식에는 차량 전방의 3차원 정보획득을 위한 스테레오 카메라와 후방의 영상을 획득하는 단일 카메라를 사용한다. 차량이 목적지를 향해 순방향 주행할 때는 전방 스테레오 비주얼 오도메트리(stereo visual odometry)를 이용하여 이동차량의 위치를 추정하고 동시에 도로 및 주변 환경에 대한 3차원 전역지도를 그래프 구조로 생성한다. 차량이 목적지에 도달하여 복귀할 때는 후방의 단일 카메라에서 획득한 2차원 영상과 전역지도를 바탕으로 모노 비주얼 오도메트리(monocular visual odometry)로 위치를 추정한다. 복귀하는 차량의 위치를 정확하게 추정하기 위해서는 효과적인 전역지도의 노드 탐색방법이 요구된다. 후방 카메라의 영상 특징과 전역지도의 각 노드의 영상 특징을 정합하고 지도에 저장된 3차원 좌표를 이용하여 차량의 위치를 추정하였다. 또한 3차원 위치추정에 성공한 이전노드들의 정보를 바탕으로 매 영상 프레임마다 적응적으로 탐색영역을 확장하거나 줄이도록 하였다. 두 개의 서로 다른 경로에 대한 실험을 통하여 제안하는 방법의 성능을 검증하였다.

Abstract

In this paper, we propose a new visual odometry technique by combining a forward-looking stereo camera and a backward-looking monocular camera. The main goal of the proposed technique is to identify the location of a moving vehicle which travels long distance and comes back to the initial position in urban road environments. While the vehicle is moving to the destination, a global 3D map is updated continuously by a stereo visual odometry technique using a graph theorem. Once the vehicle reaches the destination and begins to come back to the initial position, a map-based monocular visual odometry technique is used. To estimate the position of the returning vehicle accurately, 2D features in the backward-looking camera image and the global map are matched. In addition, we utilize the previous matching nodes to limit the search ranges of the next vehicle position in the global map. Through two navigation paths, we analyze the accuracy of the proposed method.

Keywords : Localization, Visual SLAM, Visual odometry

* 학생회원, 경북대학교 전자전기컴퓨터학부

(School of Electrical and Electronic Engineering, Kyungpook National University)

** 정회원, 경북대학교 IT대학 컴퓨터학부

(School of Computer Science and Engineering, Kyungpook National University)

※ 본 논문은 지식경제부 및 정보통신산업진흥원의 IT융합 고급인력과정 지원사업(NIPA-2013-H0401-13-1006)과 지식경제부 산업원천 기술개발사업으로 지원된 연구임(10040927, 운전 안전성 및 편의성 향상을 위한 운전자 시야 중심 차량용 증강현실 정보제공 시스템 기술 개발)

접수일자:2012년9월6일, 수정완료일:2013년2월27일

I. 서 론

이동하는 로봇이나 차량의 위치인식(localization) 기술은 지난 수십 년간 로보틱스(robotics) 분야에서 매우 활발하게 진행된 연구주제로서 유사 시 적적의 환경을 미리 염담하기 위해 사용되는 군사용 수색 로봇이나 달이나 화성 같은 우주 환경에서 사람을 대신하여 탐사 및 일련의 실험을 진행하는 우주탐사로봇을 개발하는데 사용되고 있다^[1~5].

최근에는 위치인식 기술이 민간분야에서도 활용되고 있는 추세이다. 특히 운전자 없이도 자율적으로 목적지 까지 이동하고 다시 시작 지점으로 되돌아오는 무인 이동차량(unmanned ground vehicle, 이하 UGV)에 대한 연구가 활발하게 진행되고 있는 상황이다. 이에 대한 일환으로 미국 국방성인 DARPA에서는 UGV의 주행 기술을 겨루는 “Grand Challenge”라는 국제 대회를 개최하였으며^[6], 국내에서도 현대와 기아사의 협동 주관으로 “무인자율주행자동차 연구경진대회”가 개최되었다. 미국의 구글(Google)사에서는 시각장애인의 운전을 돕는 “Self-Driving Car”를 발표하였으며, 2011년에는 이 “Self-Driving Car”가 미국 네바다 주에서 첫 무인자동차 운전면허를 취득하여 화제를 모은 바 있다^[7].

초창기 무인이동차량의 위치인식 분야에서는 위치 및 자세를 추적하기 위해 바퀴 주행거리계(wheel odometry), GPS(global positioning system) 또는 IMU(inertial measurement unit)와 같은 항법센서가 주로 활용되었다가 최근에는 CCD 카메라 또는 레이저 센서를 이용한 컴퓨터 비전 기반 위치인식 기술이 활발하게 진행되고 있다. 특히, 레이저 센서에 비해 저렴한 비용으로 3차원 거리정보를 획득하는 스테레오 카메라가

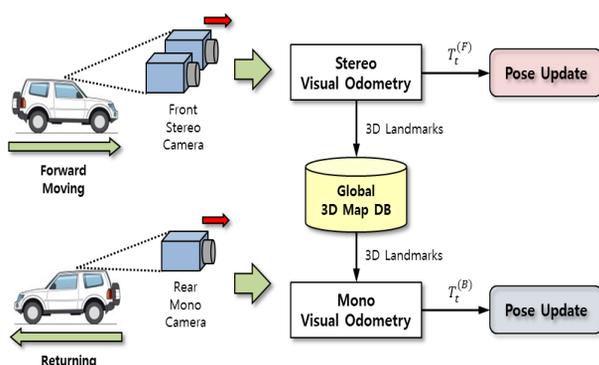


그림 1. 후방 모노카메라를 이용한 이동차량의 위치인식

Fig. 1. Localization of a moving vehicle using a backward-looking monocular camera.

위치인식에 널리 활용되고 있는 추세이다. Howard^[8]와 Nunez^[9] 등은 소형 무인이동차량의 전방에 달린 스테레오 카메라에 특징점 정합 방법을 적용하여 무인차량의 상대모션을 구하는 방법을 제안하였으며, Agrawal^[10]와 Parra^[11], 그리고 Wei^[12] 등은 스테레오 카메라와 GPS 센서를 융합하여 GPS의 위치정보보다 정밀도가 높은 하이브리드 위치인식 기술을 소개하였다. Musleh^[13]는 스테레오 카메라로 획득한 U-V 시차영상 기반으로 지면정보를 제거하여 위치인식의 정확도를 향상시켰다.

본 논문에서는 3차원 컴퓨터비전 기술만을 사용하여 실외 도로환경에 있는 차량의 초기위치로부터 특정 목적지를 방문한 후 지나온 경로를 따라 다시 초기위치로 되돌아오는 차량의 위치를 인식하는 문제를 해결하고자 한다. 제안하는 위치인식 기술은 특히 임무지를 탐색하는 동안 위급상황이 발생했을 때 탐색경로를 따라 신속하게 복귀해야만 하는 군사용 수색로봇의 위치인식에 활용될 수 있다.

3차원 컴퓨터 비전 기반 위치인식 기술은 일반적으로 카메라 영상 프레임 대 프레임 사이의 상대 모션을 계산한 뒤 이를 누적하는 방식으로 현재 이동 중인 차량의 위치를 계산한다. 따라서 시간이 점점 흐름에 따라 추정된 위치정보에는 누적오차가 발생하게 된다. 특히, 동적인 물체가 많이 나타나는 실 주행환경에서 원 거리를 주행하고 또 복귀하고자 할 경우 이러한 누적오차 문제는 더욱 더 심각한 결과를 초래한다.

최근에는 비전 카메라에서 획득한 2차원 영상정보와 3차원 거리정보를 전역지도 작성에 활용한 뒤, 지도 작성시점에 지나쳤던 특정 지역을 재방문하게 되는 경우, 전역지도 정보를 바탕으로 위치정보의 누적오차를 최소화 하는 연구가 보고되고 있다^[14]. 하지만 [14]에 제안된 내용은 단일 카메라만을 사용하기 때문에 차량이 방향을 정반대로 돌려서 복귀하려고 할 경우 카메라의 관측 방향 역시 정반대가 되기 때문에 생성된 전역지도는 누적오차 최소화에 제대로 활용될 수 없다. Zhu등은 [15]와 [16]에서 전·후방으로 두 개의 스테레오 카메라를 장착하여 지도를 생성하고, 또 생성된 지도를 활용하여 위치오차를 최소화시키는 내용을 소개하였다.

본 논문에서는 스테레오 카메라 한 대와 모노 카메라 한 대를 이용한 차량의 주행 및 복귀 위치인식 기술을 제안한다. 그림 1은 본 논문에서 제안하는 위치인식의 방법론을 보여준다. 차량에는 전방을 바라보는 스테레오 카메라와 차량의 후방을 바라보는 모노 카메라가 갈

이 장착되어 있다.

차량이 목적지를 향해 순방향 주행할 때는 전방 스테레오 카메라를 이용한 위치인식 기술로 이동차량의 위치를 추정한다. 이때, 위치추정 시 사용된 2차원 특징점 및 3차원 특징점은 지도정보로 활용되기 위해 DB에 저장된다. 모노 카메라만을 사용하여 3차원 전역지도를 생성하는 기술은 이미 잘 알려진 바 있다^[26~27]. 하지만 모노 카메라만을 사용하여 3차원 전역지도를 생성하는 경우에는 복원된 지도의 실제 크기(real scale)를 알 수 없다. 따라서 본 논문과 같이 실외환경에서 이동하는 차량의 절대위치를 추정하는 분야에서는 적합하지 못하다. 보정된 스테레오 카메라를 이용하면 영상획득과 동시에 실측 3차원 데이터를 획득할 수 있기 때문에 스테레오 카메라를 순방향 주행시 지도복원에 활용한다.

차량이 목적지에 도착한 후, 반대방향으로 주행하여 출발지로 되돌아올 때에는 차량의 후방을 바라보는 모노 카메라와 DB에 저장된 3차원 지도정보를 활용하여 이동차량의 위치를 추정한다. 차량의 순방향 이동 후에는 3차원 지도정보가 주어지기 때문에 복귀시점에는 연산량이 많은 스테레오비전기반 위치추정 보다는 모노비전기반 위치추정 기법이 적합하다. 모노카메라를 차량의 후방을 바라보도록 설치한 이유는 순방향 주행시점에 획득한 전역지도와의 매칭을 위해서다. DB에 저장된 3차원 지도정보는 순방향 이동시점에 차량의 전방영상을 바탕으로 생성된다. 따라서 복귀 시 작성된 지도를 참고하여 위치를 인식하기 위해서는 차량의 후방을 바라보는 비전센서를 장착하여 DB의 지도정보와 영상시점을 일치시켜야만 한다. 본 논문에서는 스테레오 카메라와 모노카메라를 이용한 위치인식 기술을 각각 스테레오 비주얼 오도메트리(stereo visual odometry)와 모노 비주얼 오도메트리(monocular visual odometry)라고 부른다.

본 논문의 주요 논점은 3차원 전역지도 정보가 주어졌을 때 복귀하는 차량의 위치를 얼마나 정확하게 추정할 수 있는가에 맞춰져 있다. 따라서 전방향 스테레오 센서로 획득하게 되는 차량의 위치정보 보다는 후방향 모노 카메라를 이용한 위치추정 방법론에 주로 초점을 맞추고자 한다. 본 논문의 II장에서는 전방 스테레오 카메라를 이용한 주행차량의 위치인식 및 3차원 전역지도 생성 방법론을 기술하며, III장에서는 생성된 3차원 전역지도와 후방 모노 카메라를 이용하여 복귀중인 차량의 위치를 인식하는 방법에 대해 설명한다. 이어지는 IV장에서는 본 논문에서 제안한 위치인식 기술에 대한

실험 결과를 각각 다루며, 마지막으로 V장에서 본 논문의 결론을 맺는다.

II. 스테레오 비주얼 오도메트리를 이용한 3차원 전역지도 작성

1. 전방향 스테레오 비전 기반 차량의 위치인식

본 논문에서 작성하는 3차원 전역지도는 [25]에 소개된 3차원 위치인식 기술을 기반으로 한다. 주행중인 차량의 자세를 지속적으로 갱신하기 위하여 연속하는 영상 프레임에 대한 상대적인 카메라의 모션을 추정된 뒤 이를 누적시키는 방법을 사용한다. 카메라의 모션을 추정하기 위한 입력 값으로써 $t-1$ 과 t 시간에서 획득한 2.5차원 거리영상을 사용한다. 여기서 2.5차원 거리영상은 2차원 배열에 3차원 거리정보를 저장한 일종의 데이터 구조체를 뜻한다.

2.5차원 거리영상의 3차원 깊이정보는 차량에 설치된 전방향 스테레오 카메라를 사용하여 획득한다. 본 논문에서는 2.5차원 거리영상의 조밀한(dense) 3차원 깊이정보를 획득하기 위해서 ELAS 알고리즘을 사용한다^[17]. ELAS 알고리즘은 고해상도 스테레오 영상으로부터 빠르게 시차영상(disparity image)을 생성하는 스테레오 매칭기술의 한 종류이다. ELAS 알고리즘으로 획득한 시차영상이 주어지면, 3차원 거리정보는 삼각화(triangulation) 기법을 사용하여 계산된다^[18].

2.5차원 거리영상이 주어졌을 때 연속하는 스테레오 영상 프레임에 대한 카메라의 상대모션 $T_{\langle t-1, t \rangle} = [R_{\langle t-1, t \rangle} | t_{\langle t-1, t \rangle}]$ 를 계산하는 과정은 다음과 같다. 먼저 2차원 영상이 입력되면 SURF^[19] 알고리즘을 이용하여 $t-1$ 과 t 시간에서 각각 획득한 영상에 대한 2차원 특징점들을 추출한다. 이후 추출된 2차원 특징점들 사이를 서로 매칭하여 $t-1$ 과 t 프레임 사이의 일치점군을 획득한다. 초기 매칭결과는 잘못된 매칭으로 인한 오류를 포함할 수 있으므로 에피폴라 기하제약과 RANSAC 알고리즘을 적용하여 외치점(outlier)을 제거해야만 한다^[18]. 외치점이 제거된 일치점군이 생성되면, 깊이정보를 가지고 있는 2.5차원 거리영상을 이용하여 $t-1$ 과 t 시점 사이의 3차원-3차원 매칭 관계를 확보할 수 있으며, 이로부터 두 시점 사이의 변환행렬을 계산할 수 있다^[18]. 하지만 스테레오 매칭의 오류로 인해 2.5차원 거리영상의 정보가 부정확한 경우도 있기 때문에, 두 거리데이터 사이의 강제지속 제약과 RANSAC 알고리즘을 접목하여 잘못된 3차원-3차원 매칭결과를 효과적으로 제거

하였다^[8].

강체지속제약에 따라 외치점이 제거된 m 개의 3차원 -3차원 매칭결과 $\{(X_{t-1,i}^c, X_{t,i}^c)\}_{i=1:m}$ 이 주어지면, Umeuama에 의해 제안된 최소 자승해 맞춤(least square fitting) 방법을 사용하여 3차원 상대모션 $T_{\langle t-1,t \rangle}$ 를 계산한다^[20]. 이어서 계산된 변환행렬 $T_{\langle t-1,t \rangle}$ 의 오류를 최소화 하기 위해 Levenberd-Marquardt 알고리즘을 이용하여 재투영오차 값이 최소화 되도록 비선형 최적화 과정을 수행한다^[21]. 모션추정 과정을 통해 $t-1$ 과 t 시간 사이의 변환행렬 $T_{\langle t-1,t \rangle}$ 가 결정되면 t 시간의 카메라좌표계에서 월드좌표계로 이동 하기 위한 변환행렬 $T_{\langle t \rangle}^{cw}$ 는 이전 시간에 계산된 변환행렬을 모두 누적하여 구해진다.

2. 전방향 3차원 전역지도 DB 생성

전역지도를 생성하기 위해 본 논문에서는 그래프 이론에 따라 스테레오 영상 프레임의 모션추정 과정에서 획득한 정보를 노드 단위로 저장하고 관리한다. 연속하는 영상에 대해서 카메라 모션을 추정한 뒤 추정된 모션정보 및 거리정보를 매번 DB화하기 때문에 하나의 스테레오 프레임 당 하나의 노드가 할당 된다. 따라서 본 논문에서 노드의 개념은 하나의 2차원 영상장면(scene)과 영상장면에 대한 3차원 정보로 구성된다고 볼 수 있다. 그림 2는 본 논문에서 설명하는 노드의 개념을 보여주고 있다.

전방향 스테레오 비전센서 기반 카메라 위치인식 기술로 $T_{\langle t \rangle}^{cw}$ 가 결정되면 t 시간의 노드 N_t 에는 수식 (1)과 같이 SURF 특징점 $F_{t,k}$ 와 $F_{t,k}$ 에 대응하는 3차원 월드좌표 점 $X_{t,k}^w$ 가 저장된다. 3차원 월드좌표 $X_{t,k}^w$ 는 수식 (2)에 의해 계산된다.

$$N_t = \{[F_{t,k}, X_{t,k}^w]\}_{k=1:n} \quad (1)$$

$$F_{t,k} = [x_{t,k}, \theta_{t,k}, \sigma_{t,k}, d_{t,k}]$$

$$X_{t,k}^w = T_{\langle t \rangle}^{cw} X_{t,k}^c \quad (2)$$

수식 (1)에서 $x_{t,k}$ 와 $\theta_{t,k}$ 와 $\sigma_{t,k}$, 그리고 $d_{t,k}$ 는 각각 t 시간의 k 번째에 해당하는 SURF 특징점 $F_{t,k}$ 의 2차원의 영상좌표, 특징점의 회전각, 특징점의 크기, 그리고 64차원의 특징점 기술자를 뜻하며, $X_{t,k}^c$ 는 $F_{t,k}$ 에 대한 카메라좌표계의 3차원 점을 뜻한다. $X_{t,k}^c$ 는 2차원 영상좌표 $x_{t,k}$ 와 t 시간의 거리영상 I_t^r 에서 수식 (3)과 같이 획득한다.

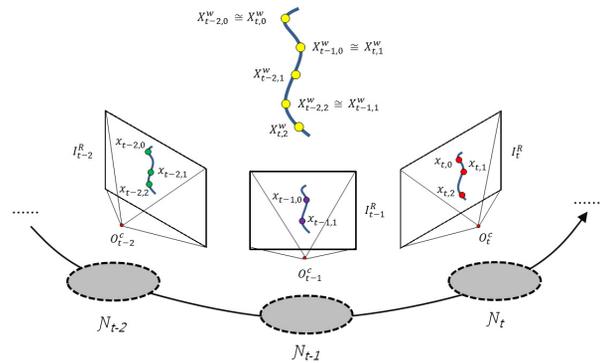


그림 2. 그래프 이론을 사용한 3차원 전역지도 생성
Fig. 2. Global 3D map generation using a graph theorem.

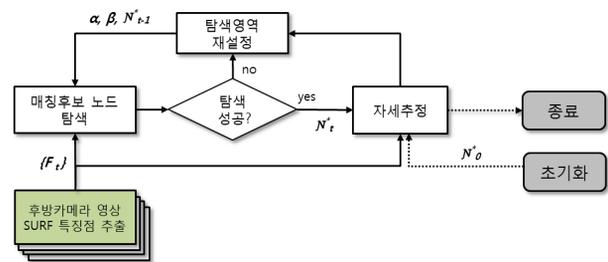


그림 3. 후방 모노 카메라를 이용한 위치추정 과정
Fig. 3. Flow of localization using a backward-looking monocular camera.

$$X_{t,k}^c = I_t^r(x_{t,k}) \quad (3)$$

거리영상 각 노드에 저장되어 있는 SURF 기술자들은 이후 III장에서 소개할 후방 모노카메라를 이용한 전역 위치인식에 활용된다. 따라서 기술자의 고속 탐색이 가능하도록 노드 생성시점에 각 노드별로 FLANN(Fast Approximate Nearest Neighbors) 알고리즘을 사용하여 기술자들에 대한 검색트리를 미리 생성하였다^[22].

III. 후방 모노 카메라를 이용한 3차원 위치인식

본 논문에서 후방 모노 카메라를 이용한 3차원 위치추정 과정은 그림 3과 같이 진행된다. 먼저 $t-1$ 시간의 위치추정에 성공한 노드정보 N_{t-1}^* 과 노드 탐색의 범위를 결정하는 인자 α 와 β 로 t 시간의 자체추정에 사용될 노드 N_t^* 를 결정한다. 노드 N_t^* 탐색시에는 SURF 특징점 매칭 기술 및 에피플라 기하제약을 이용한 외치점 제거 기술이 적용된다. 노드 탐색에 성공하면 노드 N_t^* 에 저장된 3차원 월드점 정보를 이용하여 현재 후방카메라의 자세를 결정한다. 위치추정이 끝나게 되면 $t+1$

에서 사용하게 될 a 와 β 를 수정하게 되고 t 시간의 모션 추정 과정은 마무리하게 된다.

1. 매칭후보 노드 탐색

본 논문에서는 후방 카메라에 획득된 영상과 가장 유사한 2차원 장면을 가진 전역지도상의 한 노드를 카메라의 위치추정을 위한 매칭 후보노드로 정의한다. 전역 지도에 총 m 개의 노드가 저장되어 있으며, $t-1$ 시간에서 카메라 위치추정에 성공한 노드 $N_{i-1}^* = Ni$ ($0 \leq i \leq m$)가 주어졌다고 가정했을 때, t 시간의 매칭후보 노드 N_t^* 를 선택하는 과정을 요약하면 다음과 같다.

먼저 $N_{i+\alpha-\beta} \leq N_{i+\alpha} \leq N_{i+\alpha+\beta}$ ($\alpha, \beta \in \mathbb{Z}$) 노드 구간 각 노드를 모두 순회하면서 후방카메라 영상 I_t 와 특징 유사도를 계산할 노드 N_j ($i+\alpha-\beta \leq j \leq i+\alpha+\beta$)를 선택한다. 여기서 α 는 t 시간의 매칭후보 노드를 미리 예측하는데 사용되는 변수이며, β 는 예측된 t 시간의 매칭후보 노드 $\tilde{N}_t^* = Ni_{i+\alpha}$ 로부터 탐색의 범위를 결정하는 변수이다. α 와 β 를 결정하는 내용은 III-1-라절에서 다루기로 한다.

이어서 선택된 노드 N_j 에 저장된 SURF 특징점들과 후방카메라 영상 I_t 에서 추출한 SURF 특징점들을 서로 매칭한 뒤, 에피폴라 기하제약을 만족하는 2차원 매칭쌍들의 개수를 체크한다. 에피폴라 기하제약을 만족하는 2차원 매칭쌍들의 개수는 이후 t 시간의 카메라 위치추정에 사용될 매칭후보 노드 N_t^* 를 선택하는데 있어서 중요한 척도가 된다. 따라서 효과적인 특징점 매칭 방법과 에피폴라 기하제약 적용 전략이 필요하다. III-1-가절과 III-1-나절에서는 본 논문에서 제안하는 SURF 특징점 매칭방법과 에피폴라 기하제약을 다루도록 한다.

일반적인 경우에는 노드 $N_{i+\alpha-\beta} \leq N_{i+\alpha} \leq N_{i+\alpha+\beta}$ ($\alpha, \beta \in \mathbb{Z}$) 중에서 영상 I_t 와 가장 많은 수의 매칭결과를 보이는 한 노드를 N_t^* 로 결정한다. 하지만 가려짐이 발생한 영상 또는 동적인 물체가 다수 찍힌 영상이 전역지도 생성 시점에 사용될 수 있기 때문에, 이러한 영상에 해당하는 노드는 상대적으로 특징점 매칭 및 에피폴라 기하제약 적용 시 낮은 수의 매칭쌍을 리턴하게 된다. 따라서 단순히 매칭쌍의 수를 체크하여 매칭후보 노드를 결정하는 것보다는 차량의 움직임 방향 및 속도에 따라 확률적인 근거를 추가적인 제약으로 적용하여 매칭후보 노드를 결정하는 것이 좀 더 바람직하다 할 수 있다. 이에 대한 내용은 III-1-다절에서 소개하도록 한다.

가. 특징점 매칭

본 논문에서 특징점의 매칭은 t 시간의 후방카메라 영상 I_t 에서 추출한 n 개의 SURF 특징점 $\{F_i^t\}_{i=1:n}$ 과 어떤 노드 N 에 저장된 m 개의 SURF 특징점 $\{F_j^N\}_{j=1:m}$ 을 서로 비교하여 유사도가 가장 높다고 판단되는 특징점 쌍 집합 $\{(F_i^t, F_j^N)\}_{k=1:n}$ 을 선택하는 과정으로 요약할 수 있다.

후방카메라 영상에서 추출한 한 개의 질의(query) 특징점 $F_Q \in \{F_i^t\}$ 에 대한 매칭결과 $F_R \in \{F_j^N\}$ 을 노드 N 에 있는 m 개의 특징점 집합 $\{F_j^N\}_{j=1:m}$ 과 비교하여 찾는다 가정하자. 특징점 F_Q 의 기술자를 $d_Q = [q_1, q_2, \dots, q_{64}]$, k 번째 특징점 F_{R^k} 의 기술자를 $d_{R^k} = [r_1^k, r_2^k, \dots, r_{64}^k]$ 라고 하였을 때, 특징점 F_Q 의 매칭결과 F_R 은 수식 (4)와 같이 최근접 이웃법을 사용하여 구해진다.

$$\text{dist}^E(d_Q, d_{R^k}) = \sqrt{\sum_{i=1}^{64} (q_i - r_i^k)^2},$$

$$k^* = \underset{k}{\operatorname{argmin}} \{ \text{dist}^E(d_Q, d_{R^k}) \}_{k=1:N}, \quad (4)$$

$$f: F_Q \rightarrow F_R = F_{R^{k^*}}$$

실외 주행상황에서는 가로수 잎이나 건물의 벽면 또는 도로지면과 같이 패턴이 반복되는 텍스처 부분에서도 특징점이 추출되며 이러한 영역에 대한 특징점의 기술자는 유일성 (uniqueness) 또는 구별성 (distinctiveness)이 나쁘다고 할 수 있다. 따라서 거리가 가장 가까운 기술자 $d_{R^{(1)}}$ 과 두 번째로 가까운 기술자 $d_{R^{(2)}}$ 를 구한 뒤 이 둘의 유사도를 수식 (5)와 같이 비교하여 모호하다고 판단될 경우에는 매칭에서 제외하도록 한다. 본 연구에서 λ 는 다양한 실험 결과를 토대로 0.6으로 고정 설정하였다.

$$\|d_Q - d_{R^{(1)}}\| / \|d_Q - d_{R^{(2)}}\| < \lambda \quad (0 < \lambda \leq 1.0) \quad (5)$$

비록 수식 (5)를 사용하여 SURF 기술자의 유일성이 어느 정도 보장되는 특징점에 대해서만 매칭을 결정하였다 하더라도 기술자의 유일성이 매칭 기준으로써 완전하게 반영되는 것은 아니다. 특징점 매칭이 뷰 시점 차가 크지 않는 연속 영상 시퀀스 사이에서만 일어난다는 것을 가정하면, 특징점 추출 단계에서 획득한 각 특징점의 회전 및 크기 정보를 추가적인 매칭 제약으로 활용할 수 있다. 질의 특징점 F_Q 와 기술자 유사도 비교로 획득한 매칭점 F_R 의 크기 및 회전각 쌍을 각각 (θ_Q, θ_R) , (σ_Q, σ_R) , $(0 \leq \theta_Q, \theta_R \leq 2\pi, 0 < \sigma_Q, \sigma_R)$ 이라고 하였을

때, 특징점 F_Q 와 F_R 이 최종 매칭이 되기 위해서는 다음 수식 (6)과 수식 (7)의 조건을 모두 만족해야 한다.

$$|\theta_Q - \theta_R| < \gamma \quad (6)$$

$$\xi_{QR} = \xi(\sigma_Q, \sigma_R) = \begin{cases} \sigma_Q / \sigma_R & \sigma_Q < \sigma_R \\ \sigma_R / \sigma_Q & otherwise \end{cases} \quad (7)$$

$$\xi_{min} \leq \xi_{QR} \leq 1.0$$

다양한 실험 결과를 토대로 γ 와 ξ_{min} 는 각각 $\pi/18$ (약 10°)와 0.9로 설정하였다.

나. 에피폴라 기하제약을 이용한 2차원 장면 판별
3차원 공간상의 한 장면 (scene)에 대해 각각 다른 위치에서 동시에 바라보며 획득한 두 영상이 있다고 가정하면, 공간상의 한 3차원 점 $X (X \in \mathbb{R}^3)$ 와 두 영상에 맺힌 각각의 투영점 $x (x \in \mathbb{R}^2)$ 와 $x' (x' \in \mathbb{R}^2)$ 는 에피폴라 기하 관계에 놓여있다고 말한다^[18]. 기본행렬 (fundamental matrix)은 에피폴라 기하의 대수적인 표현을 말한다. 에피폴라 기하에 따라 2차원 매칭점 x 와 x' 의 관계는 동차좌표계(homogeneous coordinate)로 표현하였을 때 수식 (8)과 같이 나타낼 수 있으며 이때의 행렬 $M_F (M_F \in \mathbb{R}^{3 \times 3})$ 를 기본행렬이라고 부른다.

$$x'^T M_F x = 0$$

$$\begin{bmatrix} u' & v' & 1 \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = 0 \quad (8)$$

수식 (8)에서의 $M_F x$ 는 에피폴라 선을 의미하며, 수식의 정의에 따라 x' 는 에피폴라 선 위의 한 점이여야만 하므로 기본행렬 M_F 는 에피폴라 기하 제약을 의미한다고 볼 수 있다.

두 영상 사이의 모든 2차원 매칭쌍들은 원칙적으로 에피폴라 제약에 따라 수식 (8)을 만족해야만 한다. 다시 말해, 두 영상 사이의 에피폴라 기하관계를 설명하는 기본행렬 M_F 가 주어졌다고 가정했을 때, 후방카메라 영상 I_t 의 SURF 특징점들과 노드 N 에 저장된 SURF 특징점들을 매칭하여 획득한 2차원 매칭쌍 집합 $\{(x_i^I, x_j^N)k\}_{k=1:n}$ 은 모두 수식 (8)을 만족해야만 한다. 여기서 수식 (8)을 만족하는 2차원 매칭쌍을 내치점(inlier correspondence)이라고 부르며, 이렇게 구해진 내치점의 개수는 이후 매칭후보 노드를 선택하는데 있어서 중요한 척도가 된다.

기본행렬 M_F 를 구하기 위해서는 일반적인 경우 8점

알고리즘(8-point algorithm)을 사용한다^[23]. 하지만 $\{(x_i^I, x_j^N)k\}_{k=1:n}$ 에는 매칭실패로 인한 오류결과도 같이 포함될 수 있기 때문에, 본 논문에서는 8점 알고리즘과 RANSAC 알고리즘을 함께 접목하여 신뢰할 수 있는 기본행렬 M_F 를 계산함과 동시에 오류결과가 제거된 내치점을 선별하도록 하였다.

다. 가우시안 함수를 이용한 모션 필터링

후방카메라 영상 I_t 와 전역지도 내의 어떤 한 노드 N_k 가 주어졌을 때, (가)절과 (나)절에 소개된 내용을 바탕으로 에피폴라 기하제약을 만족하는 SURF 매칭쌍의 개수 c_k 를 구하는 함수 $\Psi(N_k, I_t)$ 를 수식 (9)와 같이 정의한다.

$$c_k = \Psi(N_k, I_t) \quad k \in [i + \alpha - \beta, i + \alpha + \beta] \quad (9)$$

수식 (9)에 따라 영상 I_t 와 매칭후보 노드 구간 $[N_{i+\alpha-\beta}, N_{i+\alpha+\beta}]$ 에서 에피폴라 기하제약을 만족하는 SURF 매칭쌍의 개수 집합 $\{c_k\}_{k=i+\alpha-\beta:i+\alpha+\beta}$ ($\alpha, \beta \in \mathbb{Z}$)가 결정되면, 구해진 매칭쌍의 개수를 참고하여 카메라의 3차원 자세를 추정하기 위한 매칭후보 노드 N_t^* 를 최종적으로 선택해야 한다. 매칭후보 노드 N_t^* 는 다음과 같이 결정된다.

$$n = \operatorname{argmax}_{k \in [i + \alpha - \beta, i + \alpha + \beta]} (\{c_k\})$$

$$N_t^* = N_n \quad (10)$$

수식 (10)은 영상 I_t 와 에피폴라 기하제약을 만족하는 가장 많은 매칭쌍을 가진 어느 한 노드 N_t 를 매칭후보 노드로 선택하는 것을 의미한다. 이를 바꿔 설명하면 후방카메라 영상 I_t 와 2차원 장면의 특징이 가장 유사한 어느 한 노드 N_n 를 선택하겠다는 뜻으로 해석할 수 있다.

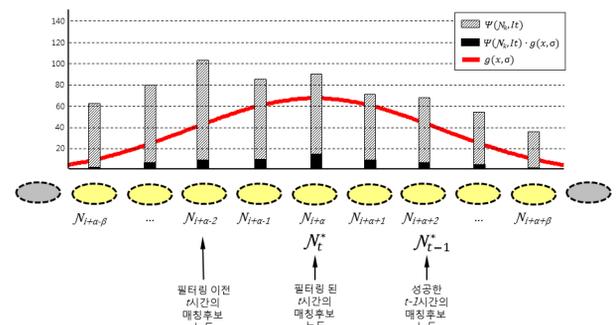


그림 4. 가우시안 함수를 이용한 필터링 결과 ($\alpha=2, \beta=4$)

Fig. 4. Motion filtering using Gaussian distribution function.

일반적인 경우 매칭후보 노드 N_t^* 는 수식 (9)로 정확하게 구할 수 있다. 하지만 차량 주행 중에는 전방향 스테레오 카메라로 획득한 영상에 다수의 동적인 물체가 포함될 수 있다. 이와 비슷하게 차량 복귀 중에도 후방 카메라에는 다수의 동적인 물체가 포함될 수 있다. 따라서 전역지도에 저장된 SURF 기술자와 후방카메라 영상에서 획득된 SURF 기술자 사이에서 매칭 불일치 현상이 자주 발생할 수 있다. 이러한 노드에 대해서는 상대적으로 작은 수의 매칭쌍 개수를 반환하기 때문에 예측된 t 시간의 매칭후보 노드 \tilde{N}_t^* 와 측정된 N_t^* 는 실제로 많은 차이를 보이게 된다. 따라서 예측된 매칭후보 노드 \tilde{N}_t^* 가 최종 매칭후보 노드 N_t^* 를 선택하는데 기여할 수 있도록 필터링 과정을 거치는 것이 중요하다.

전역지도의 각 노드는 스테레오 카메라의 동작 속도에 맞춰 일정한 시간 간격으로 저장된다. 그리고 예측된 매칭후보 노드 $\tilde{N}_t^* = Ni+a$ 를 중심으로 $Ni+a-\beta \leq Ni+a \leq Ni+a+\beta$ ($\alpha, \beta \in \mathbb{Z}$) 노드 구간을 탐색하며 최적의 N_t^* 를 찾기 때문에, 노드 \tilde{N}_t^* 와 가까운 노드에서 실제 매칭후보 노드 N_t^* 가 선택될 확률이 높다. 따라서 예측 매칭후보 노드의 인덱스 값 $i+a$ 를 중심으로 하는 1차원 가우시안 함수 $G(x, \sigma)$ 를 가중치로 적용하여 수식(12)와 같이 에피폴라 기하제약을 만족하는 SURF 매칭쌍의 개수 ck 를 필터링 하도록 한다.

$$G(x, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}} \quad (11)$$

$$\hat{c}_k = \Psi(N_k, I_t) \cdot G(x, \sigma) \quad (12)$$

$$k \in [i + \alpha - \beta, i + \alpha + \beta], x \in [-\beta, \beta]$$

본 논문에서는 β 값의 크기에 따라 노드의 탐색영역이 결정되기 때문에 가우시안 필터의 크기도 이에 맞게 자동적으로 조절되어야만 한다. 따라서 β 값의 변화에 맞춰 가우시안 분포 함수가 항상 2σ 가 유지되도록 수식 (13)과 같이 자동적으로 σ 값을 결정한다.

$$\sigma = \beta/2 \quad (13)$$

그림 4는 가우시안 필터링 이전 및 이후의 결과를 보여준다. 필터링 이전의 경우 예측된 t 시간의 매칭후보 노드 \tilde{N}_t^* 와 결정된 매칭후보 노드는 많은 차이를 보이는 반면, 필터링 적용 이후에는 예측한 매칭후보 노드

\tilde{N}_t^* 와 최적으로 매칭된 노드 N_t^* 가 일치하는 것을 알 수 있다.

2. 3차원 위치추정

매칭후보 노드 탐색으로 t 시간의 매칭후보 노드 N_t^* 가 선택되고 나면, 탐색 과정 중에 획득한 후방카메라 영상 I_t 와 노드 N_t^* 사이의 에피폴라 기하제약을 만족하는 2차원-2차원 매칭쌍 $\{(x_k^I, x_k^N)\}_{k=1:n}$ 정보를 참조하여 2차원-3차원 매칭쌍 $\{(x_k^I, X_k^{wN})\}_{k=1:n}$ 을 획득한다. 여기서 X_k^{wN} 는 노드 N_t^* 에 저장되어 있는 k 번째의 3차원 월드점을 말한다. 총 6개 이상의 2차원-3차원 매칭쌍이 주어지다면, 직접선형변환 (Direct Linear Transform, 이하 DLT) 방법을 사용하여 수식 (14)의 에너지함수가 최소가 되도록 3×4 투영 행렬 P 를 계산한다^[23].

$$\min_P \sum_i^n = (x_i^I - P X_i^{wN})^2 \quad (14)$$

DLT 방법을 사용하여 생성된 투영행렬 $P=K[R|t]$ 는 내부 파라미터를 포함한 상삼각행렬(upper triangular matrix) K ($K \in \mathbb{R}^{3 \times 3}$)와 회전정보를 가진 직교행렬(orthogonal matrix) R ($R \in \mathbb{R}^{3 \times 3}$), 그리고 3×1 이동벡터 t 의 조합으로 이루어져 있다. 따라서 투영행렬 P 를 QR 분해(QR-decomposition)하여 회전행렬 R 과 이동벡터 t 를 구하면 전역지도상에서 현재 후방카메라의 위치를 획득할 수 있다^[23].

3. 매칭후보 노드의 탐색 구간 설정

매칭후보 노드 탐색 시 탐색의 구간을 설정하는데 사용되는 변수로써 a 와 β 가 있다고 설명하였다. a 는 $t-1$ 시간에 카메라 위치추정을 성공한 노드 N_{t-1}^* 의 인덱스 정보와 결합되어 t 시간의 관심 탐색지역을 결정하며, β 는 결정된 관심지역으로부터 탐색의 범위를 설정하는 변수이다. a 와 β 값에 따라 노드의 탐색 구간이 달라지기 때문에 정확한 매칭후보 노드 선택에 영향을 줄 수 있다. 따라서 적절한 a 와 β 값을 선택하는 것이 중요하다.

먼저 t 시간의 a 값 a_t 는 수식 (15)과 같이 구해진다. 수식 (15)의 $idx(\cdot)$ 는 이전 시간에 카메라 위치추정에 성공한 노드들의 인덱스를 반환하는 함수이다.

$$\alpha_t = \left\lfloor \frac{1}{n} \sum_{i=1}^n (i dx(N_{t-i}^*) - i dx(N_{t-i-1}^*)) \right\rfloor \quad (15)$$

수식 (15)에 따라 t 시간의 α_t 는 이전 시간에 매칭을 성공한 노드들의 인덱스 간격을 평균한 값이 된다. 전방 스테레오 카메라와 후방 모노 카메라가 동기화 되어 있으며, 전역지도의 각 노드는 스테레오 카메라의 영상 획득 속도에 맞추어 일정한 거리 간격을 유지하며 저장 되기 때문에 후방 카메라를 이용한 위치인식 중에도 차량의 이동속도나 방향이 급변하지 않는 이상 정확도가 크게 벗어나지 않는 예측을 하게 된다.

다음, t 시간의 노드 탐색 범위를 결정하는 변수 β_t 는 수식 (16)과 같이 결정된다.

$$\beta_{\min} \leq \beta_t \leq \beta_{\max}$$

$$\beta_t = \begin{cases} \beta_{t-1} - 1 & t-1\text{시간에 카메라 자세 추정에 성공하는 경우} \\ \beta_{t-1} + 1 & \text{이외의 경우} \end{cases} \quad (16)$$

수식 (16)에서 β_t 는 $t-1$ 시간의 카메라 자세추정 성공 여부에 따라 β_{t-1} 에서 1이 감소하거나 또는 증가된 값을 가지게 된다. 이는 카메라 위치추정에 성공하면 탐색의 범위를 점차 줄이고, 위치추정에 실패한 경우에는 탐색의 범위를 점차 늘리겠다는 뜻이다. 탐색의 범위를 무한정 줄여들거나 늘어나는 현상을 방지하기 위하여 β_t 값이 가질 수 있는 최소치 β_{\min} 와 최대치 β_{\max} 를 설정하였다. 본 논문에서는 다수의 실험 결과를 토대로 β_{\min} 과 β_{\max} 를 각각 3과 15로 설정하였다.

IV. 실험

1. 실험 시스템

본 논문에서는 그림 5와 같이 차량의 바깥쪽 지붕에 설치할 수 있는 다중카메라 시스템을 구축하였다. 제작된 시스템의 고정 리그에는 총 3대의 카메라가 장착되어 있다. 전방을 바라보는 2대의 카메라는 스테레오 카메라를 구성하기 위한 것으로, 약 25cm의 베이스라인 거리를 가지며 향후에 가변적으로 조절할 수 있도록 설계되어 있다. 스테레오 카메라 정중앙에는 180° 반대 방향을 바라보도록 1대의 모노 카메라가 배치되어있다.

모든 카메라의 CCD 센서는 PointGrey사의 2/3인치 CCD를 사용하는 Grasshopper 모델로 구성되었다. Grasshopper 모델은 Firewire 1394b 인터페이스를 사용하며, 모든 센서는 1394b 프로토콜과 제작사로부터 제공된 라이브러리에 의해 동기화 되어 있다. 전후방



(a)



(b)

그림 5. (a) 전방 스테레오 및 후방 모노카메라의 구성
(b) 차량외부에 장착된 모습

Fig. 5. (a) forward & backward multi-camera system
(b) mounted example

CCD 센서에는 모두 C-mount형 8mm f1.4 렌즈가 장착되어 있다.

전방향 스테레오 카메라는 Camera Calibration Toolbox^[24]를 사용하여 보정된 상태이며, 추출된 보정 파라미터로 매 시간 획득하는 스테레오 영상 프레임을 교정(rectification)하였다. 또한, 교정으로 발생하는 영상 에일리어싱(aliasing) 현상을 막기 위해, 교정된 스테레오 영상 프레임에 저역 필터링(low-pass filtering)을 적용하였다.

실험 데이터를 획득하기 위해 차량에는 휴대용 랩탑이 설치되었다. 차량에 설치된 랩탑은 640x480 해상도의 전후방 컬러 영상을 프레임 손실 없이 실시간으로 저장한다. 획득된 데이터를 이용하여 모든 실험은 오프라인에서 i7급 CPU에 4GB 메모리를 장착한 데스크탑

으로 진행하였다. 본 논문에서 제안한 모든 알고리즘은 C++기반으로 작성하였다.

2. 차량 주행환경 (경로)

본 논문에서 제안하는 알고리즘의 성능을 검증하기 위해 경북대학교 대구캠퍼스 내에서 실험이 진행되었으며, 그림 6에 보이는 것과 같이 크게 두 가지의 주행경로에 대해서 실시되었다. 첫 번째 경로의 실험에서는 경북대 후문을 시작지점(1a)으로 하여 총 다섯 개의 중간지점(1b~1f)를 방문한 후 경북대 본관 앞 화단(1g)까지 이동하였다. 반대로 차량의 복귀 시에는 본관 앞 화단(1g)으로부터 각 중간지점(1f~1b)를 역으로 따라가면서 경북대 후문(1a)으로 이동하였다.

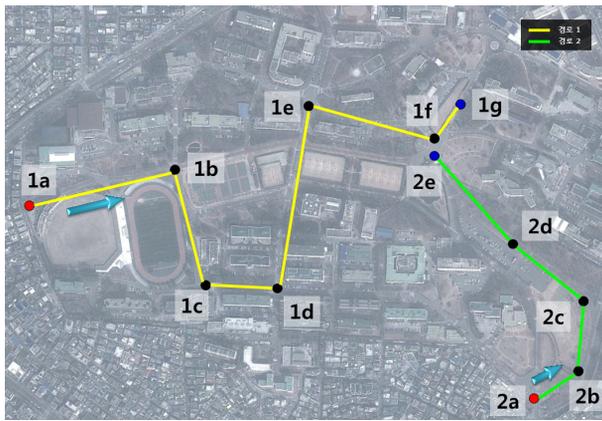


그림 6. 두 가지 경로에 대한 실험차량의 이동계획
Fig. 6. Two experimental paths

두 번째 경로에 대한 실험에서는 경북대 정문을 시작지점(2a)으로 하여 총 세 개의 중간지점(2b~2d)를 방문한 후 경북대 분수대 앞(2e)까지 이동하였다. 첫 번째 실험과 마찬가지로, 복귀 실험은 경북대 분수대 앞(2e)에서 정문(2a)까지 이동하였다. 첫 번째 실험과 두 번째 실험의 차량 왕복 이동거리는 각각 약 2,745m와 1,212m에 이른다.

실험 진행방법으로 우선, 전방향 스테레오 카메라를 이용하여 순방향으로 진행되는 주행차량에 대한 위치인식의 정확도를 평가한다. 스테레오 카메라로 획득한 위치인식의 정확도를 판단하기 위해 스테레오 영상로깅 시점에 함께 획득한 GPS 정보를 비교 및 분석에 활용하였다. 실험에 사용된 GPS 모듈은 UIGOODS사의



그림 8. 두 가지 실험경로에 대한 GPS의 궤적
Fig. 8. GPS trajectory of two experimental paths



그림 7. 경로1 실험에서 MonoVO와 매칭된 노드의 영상 예
(a) 프레임 20 (b) 프레임 1509 (c) 프레임 1827 (d) 프레임 286 (매칭실패)
Fig. 7. Scene of matched nodes (a) frame 20 (b) frame 1509 (c) frame 1827 (d) frame 286 (failed).

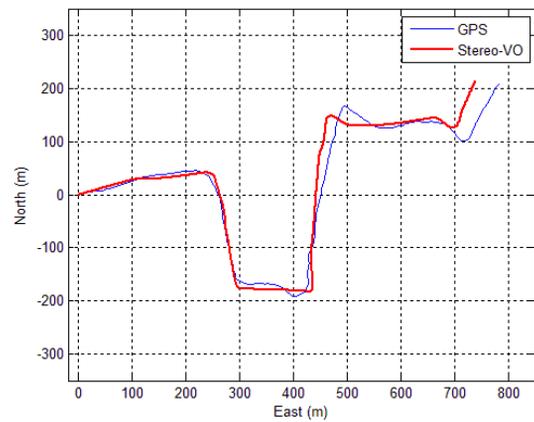
UIGGUB02-R001 모델이다. UIGGUB02-R001는 개방된 지역에서 약 3m급의 정확성을 가지며, 위치정보는 최대 4Hz의 속도로 갱신된다. GPS의 원(raw) 데이터는 WGS84 위경도 좌표계로 출력되기 때문에 이를 직교좌표계인 TM 좌표계로 변환하였다. GPS의 정확도를 검증하기 위해 매회 수신된 NMEA 문자열의 GST 정보를 활용하였으며, 실험동안 GPS의 위치오차가 허용범위 이내에서 꾸준히 유지되는 것을 확인하였다. 그림 8은 실험동안 획득한 GPS의 궤적을 구글어스 지도에 오버레이한 결과를 보여준다. 그림 8에서 알 수 있듯이 획득된 GPS 정보는 계획된 이동경로를 따라 정확하게 로깅된 것을 확인할 수 있다.

전방향 스테레오 카메라를 이용하여 전역지도가 생성되고 나면, 각 경로의 도착지점에서 역방향으로 차량을 주행시키면서 후방에 장착된 모노카메라와 생성된 전역지도 정보를 이용하여 위치인식 실험을 진행하였다.

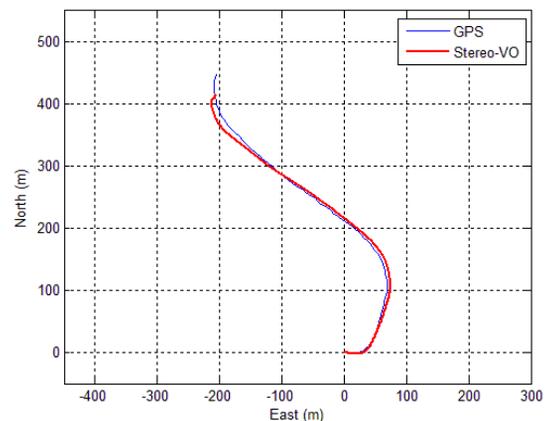
3. 전방향 스테레오 카메라의 위치추정 결과 분석

그림 9-(a)와 9-(b)는 본 논문에서 제안한 스테레오 비주얼 오도메트리 (Stereo Visual Odometry, 이하 StereoVO)로 획득한 차량 위치인식 결과와 GPS로 획득한 위치정보를 비교해 놓은 그래프이며, 각각 '경로1'과 2의 결과에 해당한다. '경로1'에 대한 StereoVO 실험에서는 총 1,810장의 스테레오 영상 프레임이 사용되었으며, '경로2'에서는 총 810장의 영상 프레임이 사용되었다. 그림 9-(a)와 9-(b)의 결과에서 알 수 있듯이 장거리 주행에 대한 실험이 진행됐음에도 불구하고 전체적인 GPS와 StereoVO의 궤적이 상당히 유사한 것을 알 수 있다.

'경로1' 실험의 경우 1a에서 1d지점까지는 GPS와 StereoVO 모두 비슷한 위치 결과를 보이고 있으나, StereoVO의 경우 1e와 1e 중반지점부터 누적오차가 발생하여 점점 GPS의 위치와 큰 차이를 보인다. 실제 1d에서 1e로 구간은 경사가 심한 내리막 및 오르막 노면이 반복하는 지점이다. 해당 구간에서 스테레오 카메라에는 주로 도로 및 하늘 장면이 크게 획득되기 때문에 모션추정에 매우 불리하다. 1e에서 1g구간의 경우 전역 위치인식 관점에서 봤을 때에는 StereoVO와 GPS의 결과가 큰 차이를 보이지만, 1e와 1g구간만 놓고 봤을 때에는, StereoVO와 GPS의 궤적이 매우 유사한 것을 알 수 있다. 따라서 1e에서 1e구간 사이의 오차만 향후에 극복되면 좀 더 나은 결과를 얻을 것으로 예상된다.



(a)



(b)

그림 9. GPS 위치정보와 StereoVO로 추정된 위치인식 결과 분석 (a) '경로1' (b) '경로2'

Fig. 9. Trajectory comparison between GPS and StereoVO (a) 'path1' (b) 'path2'.

'경로2' 실험에서는 주행구간이 상대적으로 '경로1'보다 짧았지만, 장소가 나무가 깊이 우거진 가로수길이었기 때문에 카메라렌즈의 광량부족 현상에 따른 모션블러, 특징점 매칭 불일치 등의 어려움이 있었다. 그렇지만 그림 9-(b)의 결과와 같이, StereoVO로 추정된 차량의 위치정보는 GPS의 위치정보와 비교했을 때에도 상당히 안정적인 것을 알 수 있다.

'경로1' 실험에서는 구간 1c에서 1d사이 그리고 1e에서 1f사이의 GPS 오차가 심하게 나타난 것이 관찰된다. 그럼에도 불구하고 StereoVO는 GPS와 최소 오차거리 약 2.3m에 평균 17.79m의 거리오차를 보였다. GPS와 StereoVO는 각각 1,338.5m과 1,287.2m의 주행거리를 기록하였다. '경로2' 실험에서는 StereoVO와 GPS의 최소 오차거리 약 1.12m에 평균 9.5m의 거리오차를 보였다. GPS와 StereoVO는 각각 609.2m과 597.8m의 주행거리를 기록하였다. GPS의 결과를 정답(ground truth)이라

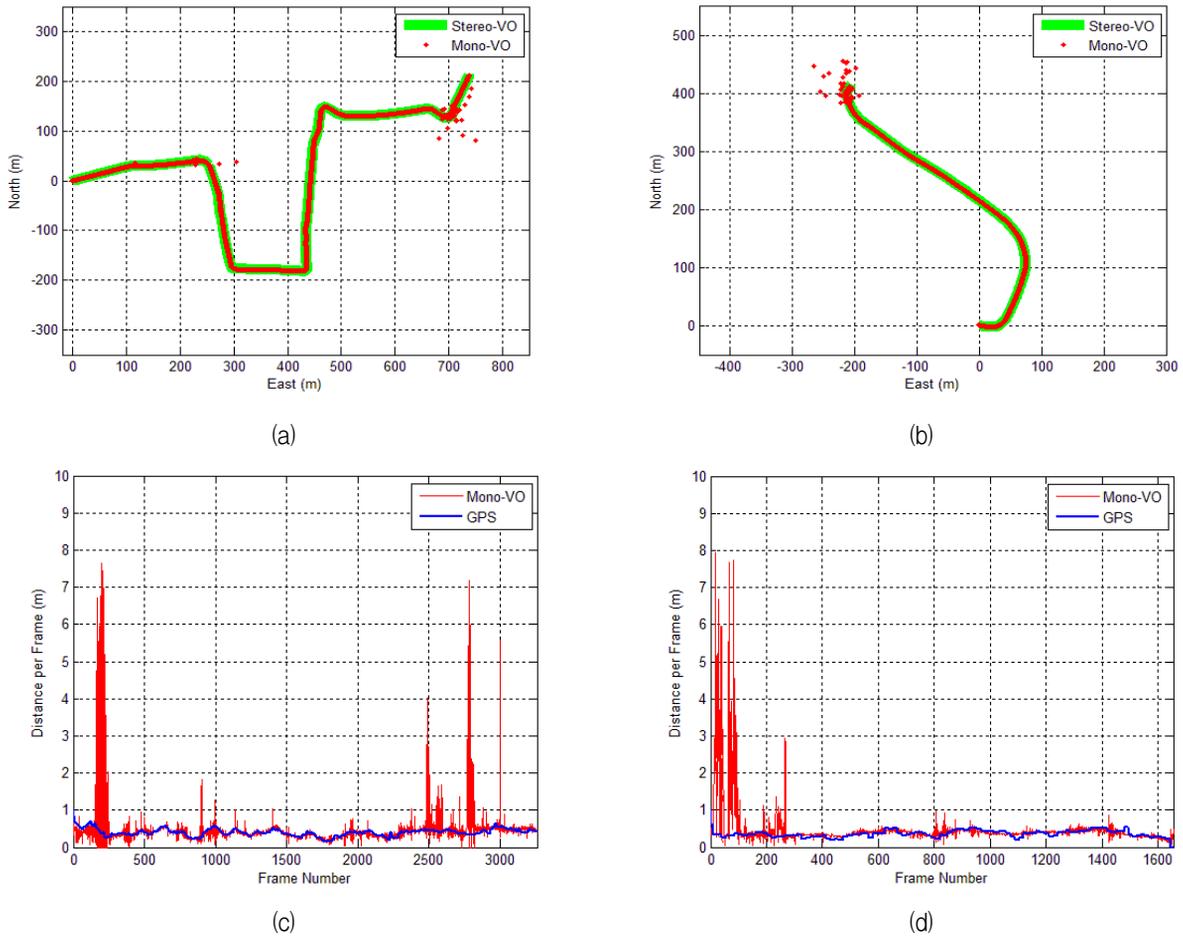


그림 10. MonoVO를 이용한 차량의 위치인식 결과 분석

(a) '경로1'의 이동궤적 (b) '경로2'의 이동궤적 (c) '경로1'의 프레임 당 이동거리 (d) '경로2'의 프레임 당 이동거리

Fig. 10. Localization result of MonoVO (a) estimated trajectory on 'path1' (b) estimated trajectory on 'path2'

(c) moving distance per frame on 'path1' (d) moving distance per frame on 'path2'.

가정했을 때, '경로1'과 '경로2'에 대한 StereoVO의 주행 거리 오차와 오차율을 표 1과 같다. 프레임 당 평균 처리속도는 표 3과 같이 '경로1'의 경우 0.933초, '경로2'의 경우 0.989초가 소요되었다.

4. 후방향 모노 카메라의 위치추정 결과 분석

다음 실험은 초기위치에서 목적지로 주행하며 생성한 전역지도가 주어졌을 때, 본 논문에서 제안한 모노 비주얼 오도메트리(Mono Visual Odometry, 이하 MonoVO) 알고리즘을 적용하여 목적지로부터 초기위치로 다시 복귀하는 차량의 위치를 추정하고 이에 대한 정확도를 분석하였다. 본 실험은 StereoVO 실험으로 미리 작성된 전방향 3차원 전역지도를 활용하였으며, 총 두 가지의 복귀 주행경로에 대해 진행되었다.

실험을 진행하기에 앞서 '경로 1'과 '경로 2'의 경우

모두 StereoVO 알고리즘이 종료된 위치에서부터 복귀 주행 실험이 진행된다고 가정한다. 따라서 차량의 3차원 초기위치를 계산하는데 필요한 0번째 시간의 매칭 후보 노드 N_0 는 전역지도에 저장된 마지막 노드로 각각 대체하였다. 이후 1번째 영상 프레임부터 마지막 프레임까지 매칭후보 노드 탐색 및 위치추정, 그리고 탐색 범위 재설정 모두 자동으로 진행된다. 1번째 노드 탐색에 필요한 α_1 와 β_1 값은 각각 3과 5로 설정하였다. 복귀 실험에 사용된 영상 프레임은 '경로 1'의 경우 총 3,260장이었으며, '경로 2'는 총 1,622장의 영상이 실험에 사용되었다.

그림 10-(a)와 10-(b)는 본 논문에서 제안한 전방향 StereoVO 및 후방향 MonoVO로 추정한 차량의 이동궤적을 동시에 보여준다. 녹색의 선으로 표시된 궤적은 전방향 StereoVO로 획득한 차량의 이동궤적이며, 붉은

표 1. StereoVO의 위치인식 오차 분석
Table 1. Accuracy analysis on StereoVO.

분류	총 프레임 수	주행거리	오차	오차율
경로1	1,810	1,234.1m	51.3m	4.1%
경로2	810	609.2m	11.4m	1.8%

표 2. MonoVO의 위치인식 오차 분석
Table 2. Accuracy analysis on MonoVO.

분류	총 프레임 수	실패 프레임 수	성공율
경로1	3,260	52	98.4%
경로2	1,622	46	97.2%

표 3. 프레임 당 평균처리속도 (sec)
Table 3. Average computation time per frame (sec).

분류	StereoVO	MonoVO
경로1	0.933	0.729
경로2	0.989	0.715

색 점으로 표시된 궤적은 후방향 MonoVO로 획득한 복귀주행 차량의 이동경로를 나타낸다. ‘경로1’과 ‘경로2’의 실험 모두 비슷한 실험장소에서 많은 위치오차가 발생하였다. 오차가 발생한 위치는 넓은 공터가 나타나는 지역으로써, MonoVO가 전역지도상의 잘못된 3차원 정보로 자세를 추정하려다 실패한 경우라고 할 수 있다. 초기 몇 장의 영상 프레임만을 제외하고는 마지막 지점까지 매우 안정적인 위치추정 결과를 보여주었다. ‘경로1’에서는 총 52장의 실패 프레임이 발생하였으며, ‘경로2’에서는 총 46장의 실패 프레임이 발생하였다. 따라서 위치추정 성공율은 각각 98.4%와 97.2%를 기록하였다. 그림 7은 MonoVO로 모션추정에 성공한 경우와 실패했을 경우에 대한 매칭후보 노드의 영상을 보여준다.

MonoVO로 추정된 차량의 위치결과는 전역지도의 정확도에 영향을 받는다. 따라서 전역지도의 정확도가 GPS 급에 미치지 않는 이상 GPS의 절대 위치정보와 MonoVO로 추정된 위치정보를 서로 비교하는 것은 의미가 없다. 본 실험에서는 GPS와 MonoVO의 상대적인 이동거리를 비교하는 것으로 정확도를 비교하도록 한다. 그림 10-(c)와 10-(d)는 MonoVO와 GPS의 프레임당 이동거리를 프레임의 흐름에 따라 제시한 결과이다. 위치추정에 실패한 부분에서는 상대적으로 프레임당 이동거리가 높게 나타났다. 그럼에도 불구하고 전체적으로 GPS의 프레임당 이동거리와 상당히 유사하게 획득된 것을 알 수 있다. MonoVO의 프레임당 평균 처리속도는 ‘경로 1’과 ‘경로 2’에서 각각 0.728초와 0.715초를 기록하였다.

V. 결 론

본 논문에서는 전방향 스테레오 카메라와 후방향 모노 카메라를 조합하여 목적지를 방문하고 이동경로를 따라 복귀하는 차량의 위치인식에 대한 기술을 소개하였다. 목적지를 향해 순방향으로 이동하는 차량의 위치를 추정하고 복귀시점에 활용될 3차원 전역지도를 생성하기 위하여 전방향 스테레오 카메라로 비주얼 오도메트리를 구성하였다. 스테레오 비주얼 오도메트리로 획득된 2차원 특징점 및 3차원 월드점은 그래프 구조에 따라 노드단위로 DB에 저장하였다. 차량이 목적지에 도달하여 시작위치로 복귀하는 경우 차량의 후방을 바라보는 모노 카메라의 비주얼 오도메트리가 작동한다. 복귀주행 차량의 위치를 정확하게 계산하기 위해서는 전역지도의 효과적인 노드 탐색방법이 요구된다. 본 논문에서는 3차원 모션추정에 성공한 이전노드들의 정보를 바탕으로 매 프레임마다 적응적으로 탐색영역을 확장하거나 줄이도록 하였다. 또한, 매칭이 일어날 후보노드를 미리 예측하고 필터링을 한 번 더 거침으로써 3차원 모션추정의 성공률을 높이도록 하였다. 실제 주행환경에서 획득한 두 개의 장거리 이동경로에 대하여 실험을 진행하였으며, 두 실험 경로 모두 97% 이상의 매칭 성공률을 보여 정확한 위치 추정이 가능함을 알 수 있었다.

참 고 문 헌

- [1] R. Madhavan and E. Messina, "Iterative registration of 3d lader data for autonomous navigation," in Proceedings of the The IEEE Symposium on Intelligent Vehicles, pp. 186-191, 2003.
- [2] R. Madhavan, T. Hong and E. Messina, "Temporal range registration for unmanned ground and aerial vehicles," Journal of Intelligent and Robotics Systems, vol. 44, no. 1, pp. 47-69, 2005.
- [3] N. Vandapel and M. Hebert, "3D Rover localization in airborne lader data," The International Symposium on Experimental Robotics, July 2002.
- [4] N. Vandapel, R. Raghavendra, R. Donamukkala, and M. Hebert, "Unmanned ground vehicle navigation using aerial lader data," The International Journal of Robotics Research, vol. 25, no. 1, pp. 31-51, 2006.
- [5] J. J. Biesiadecki and M. Maimone, "The Mars Exploration Rover Surface Mobility Flight Software: Driving Ambition," in Proceedings of the IEEE conference on Aerospace, 2006.

- [6] E. Jones, B. Fulkerson, E. Frazzoli, D. Kumar, R. Walters, J. Radford, R. Mason, "Autonomous Off-Road Driving in the DARPA Grand Challenge," The IEEE Symposium on Position Location, And Navigation 2006.
- [7] "Google driverless car", Google, http://en.wikipedia.org/wiki/Google_driverless_car
- [8] A. Howard, "Real-time stereo visual odometry for autonomous ground vehicles," in Proceedings of the IEEE conference on Intelligent Robots and Systems, pp. 22-26, 2008.
- [9] P. Nunez, R. Vazquez-Mart, A. Bandera, and F. Sandoval, "Combined Constraint Matching Algorithm for Stereo Visual Odometry based on Local Interest Points," in Proceedings of the IEEE conference on Intelligent Robots and Systems, pp. 3095-3100, 2008.
- [10] M. Agrawal and K. Konolige, "Real-time Localization in Outdoor Environments using Stereo Vision and Inexpensive GPS," in Proceedings of the IEEE International Conference on Pattern Recognition, pp. 1063-1068, 2006.
- [11] I. Parra, M. A. Sotelo, D. F. Llorca, C. Fernandez, A. Llamazares, N. Hernandez and I. Garcia, "Visual odometry and map fusion for GPS navigation assistance," The IEEE Symposium on Industrial Electronics, pp. 832-837, 2011.
- [12] L. Wei, C. Cappelle, Y. Ruichek and F. Zann, "GPS and Stereovision-Based Visual Odometry: Application to Urban Scene Mapping and Intelligent Vehicle Localization," The International Journal of Vehicular Technology, vol. 2011, special page 1-13, 2011.
- [13] B. Musleh, D. Martin, A. Escalera and J. M. Armingol, "Visual ego motion estimation in urban environments based on U-V disparity," The IEEE Symposium on Intelligent Vehicles, pp. 444-449, 2012.
- [14] G. L. Mariottini and S. I. Roumeliotis, "Active Vision-Based Robot Localization and Navigation in a Visual Memory," The IEEE International Conference on Robotics and Automation, 2011.
- [15] Z. Zhu, T. Oskiper, S. Samarasekera, H. Sawhney and R. Kumar, "Ten-fold improvement in visual odometry using landmark matching," in Proceedings of the IEEE International Conference on Computer Vision, 2007.
- [16] Z. Zhu, T. Oskiper, S. Samarasekera, R. Kumar and H. Sawhney, "Real-time Global Localization with a Pre-Built Visual Landmark Database," in Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 2008.
- [17] A. Geiger, M. Roser and R. Urtasun, "Efficient large-scale stereo matching," in Proceedings of the Asian Conference on Computer Vision, 2010.
- [18] R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision," 2nd Ed, Cambridge University Press, 2004.
- [19] H. Bay, A. Ess, T. Tuytelaars, L. V. Gool, "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346-359, 2008.
- [20] S. Umeyama, "Least-Square Estimation of Transformation Parameters Between Two Patterns," In IEEE Transaction on Pattern Analysis and Machine Intelligence, 4(13): 376-380, April, 1991.
- [21] K. Levenberg, "A method for the solution of certain non-linear problems in least square," The Quarterly of Applied Mathematics, pp. 164-168, 1944.
- [22] M. Muja and D. G. Lowe, "Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration", in International Conference on Computer Vision Theory and Applications, 2009.
- [23] E. Trucco and A. Verri, "Introductory Techniques for 3-D Computer Vision," pp. 132-134, 1998.
- [24] "Camera Calibration Toolbox for Matlab," http://www.vision.caltech.edu/bouguetj/calib_doc/
- [25] S. I. Choi, L. Zhang, and S. Y. Park, "Stereo Vision based Motion Adjustment of 2D Laser Scan Matching," in Proceedings of the International Conference on Image and Vision Computing New Zealand, 2011.
- [26] D. Nister, O. Naroditsky, J. Bergen, "Visual Odometry," in Proceedings of the International Conference on Computer Vision and Pattern Recognition, pp. 652-659, 2004.
- [27] A. J. Davison, I. Reid, N. Molton and O. Stasse, "MonoSLAM: Real-time single camera SLAM," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 29, No. 6, pp 1052-1067, 2007.

— 저 자 소 개 —



최 성 인(학생회원)
2007년 2월 금오공과대학교 컴퓨터공학부 졸업 (공학사)
2009년 8월 경북대학교 전자전기컴퓨터학부 졸업 (공학석사)

2009년 9월~현재 경북대학교 전자전기컴퓨터학부 박사과정
<주관심분야 : 3차원 정합 및 복원, 로봇위치인식>



박 순 용(정회원)-교신저자
1991년 2월 경북대학교 전자공학과 졸업 (공학사)
1993년 2월 경북대학교 전자공학과 졸업 (공학석사)
2003년 5월 미국 뉴욕주립대학교 스토니브룩 졸업 (공학박사)

1993년~1999년 한국원자력연구소
로봇개발연구실 선임연구원
2004년~2005년 한국전자통신연구원
디지털콘텐츠연구단 선임연구원
2005년~현재 경북대학교 컴퓨터학부 부교수
<주관심분야 : 3차원 컴퓨터비전 및 로봇비전>