

실내외 연동을 위한 markerless 증강현실 구현

김희관*, 조현달*

*(주)제니텀

e-mail:albertk@zenitum.com

A Markerless Augmented Reality Approach for Outdoor/Indoor

Albert Heekwan Kim*, Hyeondal Cho*

*Zenitum, Inc.

요 약

증강현실 기술은 실제 환경에 가상의 물체를 덧씌우는 기술을 말하며, 이는 지리정보의 가시화 같은 작업에 매우 큰 잠재력을 갖고 있다. 하지만 지금까지 연구된 이동형 증강현실 시스템은 사용자의 위치를 파악하기 위해 GPS를 사용하거나 마커를 현장에 붙이는 방식을 사용하였다. 최근 연구들은 마커를 사용하지 않는 방법을 지향하고 있으나 많은 제약을 갖고 있다. 특히 실내의 경우는 GPS정보를 사용할 수 없기 때문에 실내 위치파악을 위해서는 새로운 기술이 필요하다. 최근 무선(RF)기반의 실내 위치 추정 연구가 활발히 수행되고 있지만, 이 또한 다량의 센서와 인식기를 설치해야한다는 제약이 존재한다. 본 연구에서는 한 대의 카메라를 사용하는 SLAM 알고리즘을 이용한 위치 추정기법을 제시하였으며, 추정된 위치를 이용하여 증강현실을 통한 정보 가시화 프로그램을 개발하였다. 이는 실내외 seamless 연동형 u_GIS 시스템의 밑바탕이 될 것이다.

1. 서론

GPS 기술은 이제 대중적인 기술이 되었다. 많은 운전자들이 GPS 기술을 이용한 차량용 항법장치의 도움을 받아 운전을 하고 있으며, 휴대전화의 '친구찾기', '위치추적' 서비스는 인기리에 서비스 되고 있다. 이렇듯 최근 다양한 분야에서 GPS를 이용한 위치기반 응용 서비스의 개발이 이루어지고 연구되고 있다[1]. 이러한 산업계와 학계의 움직임은 향후 위치기반 서비스가 가질 잠재력에서 비롯된 것이라고 볼 수 있다. 하지만 GPS를 이용한 위치기반 시스템에는 한계가 있다. GPS는 인공위성에서 발신하는 마이크로파를 수신할 수 없는 실내와 같은 공간에서는 무용지물이 되어 버리기 때문이다.

이러한 한계에도 불구하고 많은 연구자들은 실내 위치 추정을 위한 연구를 수행하고 있다. 빌딩 내에서 로봇의 위치 기반 서비스를 위해 수행되었던 많은 연구들이 유비쿼터스 컴퓨팅 패러다임의 도입으로 인해 사용자의 위치 추정을 위한 연구로 이어지고 있다. 이러한 연구들에서 GPS가 갖는 기술의 한계를 극복하기 위한 다양한 방법이 모색 되었다. 일부 연구들은 비행기에 사용되는 관성항법 시스템에서 영감을 얻어 관성센서를 사용하여 실내 위치 트래킹 시스템을 개발 하였지만, 대부분의 연구들은 라디오 주파수(RF)의 사용에 초점을 맞춰 개발 되었다.

AT&T에서 개발된 Active-Bat[2]과 MIT에서 개발된 Cricket[3]은 초음파의 도달시간을 사용하였으며,

Active-Badge[4]는 방 이름과 같은 symbolic 정보 전송하는 적외선 태그를 통해 위치 추정을 구현하였다. 하지만 적외선을 이용하는 방식은 적외선의 거리제약 때문에 널리 사용되지 못하였다. PinPoint 3D[5]는 라디오 주파수 신호를 이용한 삼각측량법으로 위치를 추정하게 되며 3미터 정도의 오차를 갖는다. 하지만 이를 위해서는 시스템을 위해 개발된 특별한 장비의 설치를 필요로 한다. SpotON[6]과 UOID[7] 프로젝트는 각각 3D와 2D의 위치 정보를 제공하며 RFID 태그를 사용하였다. Kang et al.은 RFID를 이용하여 자동차의 주차보조 시스템을 개발하였다[8]. SmartFloor[9]는 바닥에 미리 설치된 센서 그리드를 사용하였으며, 정확도는 센서 그리드의 간격에 영향을 받는다.

이러한 시스템들은 위치 추정을 위해 특별한 기반시설을 설치해야 하는 단점이 있다. 그렇기 때문에 기존에 설치되어 있는 기반시설을 이용해 위치를 추정하려는 시도가 생겼다. 텔레커뮤니케이션 망과 휴대전화망, WiFi망 들은 비교적 잘 정비되어 있다. 휴대전화망은 도심에서도 적절한 정확도를 제공할 수 있을 만큼 많이 설치되어있다. 휴대전화망을 사용하는 방법은 신호의 세기를 이용하여 현재 포함되어있는 셀을 구별할 수 있으며, 그렇기 때문에 정확도는 셀의 밀도와 하나의 셀이 커버하는 영역의 크기에 영향을 받는다. 물론 서로 다른 셀에서 나오는 각각의 신호의 세기를 이용해 삼각측량법을 적용할 수도 있다[10]. 이러한 방식은 wireless LAN에도 적용될 수 있다.

최근에는 빌딩 안에 wireless LAN 망이 잘 정비되어 있으므로, 실내 위치 추정을 위해 WLAN을 사용하려는 시도가 많이 생겨났다. Microsoft Research Group에서 개발된 RADAR[11]는 건물 내에서 WLAN신호의 세기를 측정하여 미리 작성된 데이터 테이블과 비교하여 위치를 추정한다. Battiti et al. 은 신경망을 적용한 WLAN 핑거프린팅 방법을 제안하였다[12].

앞서 보았듯이 RF기반 실내 위치 추정 시스템은 많은 연구자들에게 연구되어 지고 있다. 하지만 이러한 방식에도 몇 가지 제약점이 존재한다. 첫째로, 위치의 정확도가 좀 부족하다는 것이다. 방법마다 차이가 있겠지만 앞서 살펴본 RF기반의 기술들은 5m에서 10m의 오차를 갖는다. 둘째로, RF기반의 기술은 위치 추정이 이뤄지는 공간에 설치된 RF 센서 또는 인식기의 개수에 의해 정확도가 정해지기 때문에 위치의 정확도를 높이기 위해서는 더 많은 하드웨어를 필요로 한다. 셋째로, RF기반의 기술은 센서의 2차원 위치를 제공하기 때문에 센서를 착용한 사용자의 자세를 추정하기 어렵다.

이러한 문제점들을 해결하기 위해서 이 논문에서는 카메라 한 대만을 센서로 사용하는 MonoSLAM 알고리즘을 적용한 새로운 위치 추정 기법에 대해 설명하겠다. 또한 이 알고리즘을 사용해 증강현실을 구현하고 이를 이용한 실내 위치기반 정보 가시화 시스템인 ARVisualizer에 대해 설명하도록 하겠다.

2. MonoSLAM

SLAM(Simultaneous Localization and Mapping)은 본래 로봇공학 연구그룹에서 시작된 방법으로, 로봇의 현재 위치 추정함과 동시에 주변 환경을 맵으로 만들어 나가는 방법을 말한다. 일반적으로 로봇공학에서는 로봇의 주행거리계와 초음파 또는 레이저를 사용하는 거리센서 값을 일련의 베이지안 추정기법을 통해 로봇의 현재 위치를 추정하고 주위 환경의 맵을 생성한다[13]. 연구자들은 차차 카메라를 이용해 SLAM을 수행하고자 하는 연구를 시도하였다. 초기에는 스테레오 카메라를 로봇에 장착하여 SLAM을 수행하도록 하였다. 그 후, Davison et al. 에 의해 카메라 한 대만을 사용하는 SLAM방법이 제안되었고, Davison은 연구를 점차 진행시켜 MonoSLAM 알고리즘을 개발하였다[14,15,16].

MonoSLAM은 자유롭게 움직이는 wide-angle 카메라를 유일한 센서로 사용하여 실시간으로 SLAM을 수행하도록 하였다. MonoSLAM은 30fps로 동작하며, 카메라의 자세를 추정함과 동시에 맵을 생성한다. 이때 맵은 카메라로 획득된 영상에서 추출한 작은 이미지패치를 랜드마크로 갖는다. MonoSLAM은 카메라 한 대만을 사용해 비교적 정확한 위치를 추정할 수 있는 점에서 굉장히 효율적인 알고리즘이지만 실내공간에서 동작하도록 개발 되었고, 카메라의 움직임이 부드러워야 한다는 제약조건을 갖는다.

MonoSLAM은 확장 칼만 필터를 사용하여 카메라의 위치와 맵을 유지하고 재설정해 나간다.

카메라의 자세와 위치, \hat{x}_p 는

$$\hat{x}_p = \begin{pmatrix} r \\ q \end{pmatrix} = (x \ y \ z \ q_0 \ q_x \ q_y \ q_z)^T$$

위 와 같이 나타내어지며, r은 카메라의 3차원 공간상의 위치를 나타내고, q는 카메라의 자세를 쿼터니언을 사용해 나타내고 있다.

카메라의 상태를 나타내는 벡터 \hat{x}_v 는 \hat{x}_p 와 함께 카메라의 속도와 각속도 정보를 포함하고 있다. 맵을 구성하는 랜드마크인 특징점은 3차원 위치 벡터로 나타내어진다.

$$\hat{y}_i = (x \ y \ z)^T$$

전체 상태를 나타내는 상태벡터는 아래와 같이 다변량 가우시안 형태로 나타내어진다.

$$\hat{x} = \begin{pmatrix} \hat{x}_v \\ \hat{y}_1 \\ \hat{y}_2 \\ \vdots \end{pmatrix}, P = \begin{bmatrix} P_{xx} & P_{y_1x} & P_{y_2x} & \dots \\ P_{y_1x} & P_{y_1y_1} & P_{y_1y_2} & \dots \\ P_{y_2x} & P_{y_2y_1} & P_{y_2y_2} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

시스템은 초기 위치를 알고 있는 특징점을 통해 카메라의 위치를 초기화 한다. 그 이후에 들어오는 영상 프레임 부터는 Davison이 "Active Search" 라고 명명한 방법을 통해 각 특징점에 대한 3σ 검색영역을 결정한다. 각 특징점들은 11×11 pixels의 평면 영상으로 이루어져 있으며, 각 특징점들은 Active Search 영역 안에서 Correlation을 검사를 통해 실제 위치를 측정되게 된다. 각 특징점들의 예측된 위치와 실제위치간의 차이는 확장 칼만 필터에 제공되어 카메라의 상태와 전체 맵을 갱신하는데 사용된다.

이 논문에서는 핸드 헬드 컴퓨터에 맞춰 MonoSLAM을 수정하였고, MonoSLAM에서 얻어진 위치정보를 이용하여 이미지 스트림에 가상의 가구를 덧입히는 증강현실 기술을 구현하고자 하였다.

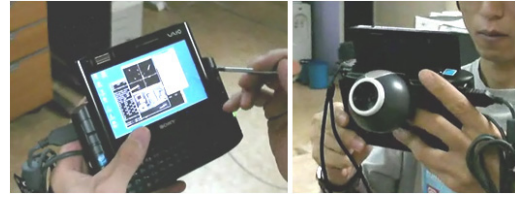
3. ARVisualizer

3.1 하드웨어 구성

앞서 설명한 MonoSLAM 연구에서, Davison et al 은 wide-angle IEEE 1394 카메라를 사용하였다. 1394 카메라는 데이터 전송율이 꽤 안정적이며, 이들이 사용한 wide-angle 카메라는 넓은 시야각으로 인해서 작업공간에 분포된 특징점들이 더 자주 중복적으로 관측될 수 있다. 이러한 이점에도 불구하고 IEEE 1394 카메라를 핸드헬드 컴퓨터에 적용 하는 것은 약간의 문제가 있다. 대부분의 IEEE 1394 카메라들은 전원을 호스트 컴퓨터에서 끌어오는 6핀 규격을 사용하는데, 시중에 판매되는 대부분의 모



<그림 1> Davison의 MonoSLAM에 사용된 IEEE1394 카메라(왼쪽)와 4pin과 6pin 커넥터(오른쪽)



<그림 2> ARVisualizer에 사용된 UMPC와 USB카메라

바일 컴퓨터는 4핀을 사용하는 IEEE 1394a 규격만을 제공하기 때문이다[17].

이러한 제약으로 인해서 우리는 Davison의 MonoSLAM을 USB 카메라로 동작시킬 수 있도록 알고리즘을 수정해야 했다. 핸드 헬드 컴퓨터로는 1.2Ghz 로 동작하는 Intel Core2 Solo U2200를 사용한 SONY VGN-UX58LN UMPC (Ultra Mobile PC)를 사용하였다. MonoSLAM을 성능이 제한된 UMPC상에서 동작시키기 위해 부분적으로 알고리즘을 수정하였고 최적화 하였다.

3.2 Camera Tracking

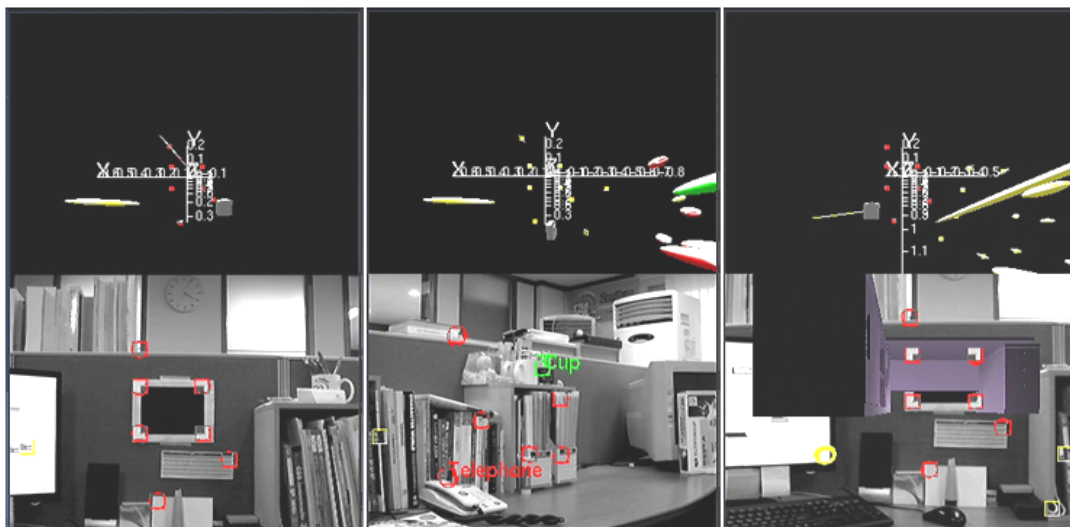
Davison의 연구에서처럼, 우리의 시스템도 현재의 카메라 위치와 자세, 운동변수를 추정한다. 또한 마찬가지로 이미지패치로 구성된 맵을 생성하며 관리한다.

시스템은 4개의 이미 알고 있는 특징점을 통해 초기의 위치와 자세를 정하게 된다. 이때 사용되는 특징점은 Shi와 Tomasi가 제안한 방법을 통해 얻은 이미지패치를 사용한다[18]. 이 특징점을 맵의 랜드마크로 사용하는 EKF 기반 SLAM 과정에서 특징점의 측정은 저장된 이미지 패치와 현재 영상에서의 Active Search를 통해 예측된 특징점 간의 NCC(Normalized Cross Correlation)을 통해 측정된다. 만약 현재의 영상에서 맵에 저장된 랜드마크들이 관

측되지 않거나 부족하게 되면 Corner detector를 통해 새로운 특징점을 찾고 3차원 공간상의 위치를 추정하여 맵에 등록하게 된다. 이렇게 등록된 3차원 정보를 갖는 랜드마크들을 반복적으로 관측함으로써 EKF update를 통해 현재 카메라의 위치와 자세를 추정하게 된다.

시점에 따른 이미지패치의 변화를 고려해서 특징점 관측의 정확도를 향상시키기 위해 Molton의 Locally planar patch feature에 관한 연구를 적용해 보았다. 이 연구에서 각 특징점에 해당하는 이미지패치는 평면으로 가정하였고, 초기 등록시 카메라의 시야각과 이미지패치의 법선벡터를 일치시켰다. 그럼으로써 이후 특징점을 관측할 때, 특징점의 시점에 따른 왜곡을 고려할 수 있도록 하였다. 이 연구에 의하면 미리 왜곡하여 등록한 이미지패치들은 그 패치의 법선벡터가 실제와 차이가 있더라도 카메라의 위치와 자세의 추정의 정확도를 크게 향상 시켰다고 한다[19].

하지만 우리는 실험을 통해 수직벡터를 EKF를 통해 계속 갱신해 나가는 것이 그것을 통해 얻을 수 있는 이점에 비해 너무 많은 부하를 준다는 것을 발견하였다. 그렇기 때문에 우리는 초기에 설정한 수직벡터를 갱신하지 않기로 하였다. 또한 성능 향상을 위해 NCC를 Search window의 모든 위치에서 수행하는 대신, 발견된 Fast corner들에서만 수행하도록 하였다.



<그림 3> Implementation Results: (좌에서 우로) (1) 3차원 맵의 초기화, (2) 생성된 맵을 통해 구현한 3차원 주석. (3) 3차원 영상에 3D Collada 모델을 덧씌운 결과.

4. 결론

우리는 성능의 제약이 있는 하드웨어 상에서 MonoSLAM기반의 증강현실 시스템을 구현하였다. 이를 위해 MonoSLAM알고리즘에 USB카메라를 사용할 수 있도록 몇몇 클래스를 개발하였고, Davison의 알고리즘을 UMPC에 맞춰 최적화 시켰다. 또한 collada/KML 기반의 3차원 영상을 실시간영상에 덧씌움으로써 모바일 증강현실 시스템의 실용성을 증명하였다. 또한 이 시스템은 Intel Core2 Solo 단일 코어를 사용하여 개발되었지만, 만약 멀티코어를 사용하여 맵 생성과 카메라 위치 추적을 서로 다른 코어에 스레드로 동작시킬 수 있다면 시스템의 성능이 크게 개선 될 수 있음을 발견하였다[20].

감사의 글

본 연구는 건설교통부 첨단도시기술개발 사업 - 지능형국토정보기술혁신 사업과제의 연구비지원 (07국토정보C05)에 의해 수행되었습니다.

참고문헌

- [1] 이성호, 민경욱, 김재철, 김주완, 박종현, "위치기반서비스 기술 동향", ETRI 전자통신동향분석 제20권 제3호, 2005, pp. 33-42.
- [2] A. Ward, A. Jones, and A. Hopper, "A new location technique for the active office," IEEE Personal Communications,4(5), 1997, PP.42 - 47.
- [3] N. B. Priyantha et al., "The Cricket Location-Support System," In Proc. 6th Ann. Int'l Conf. Mobile Computing and Networking (MOBICOM 2000), 2000, pp.32-43.
- [4] R. Want et al., "The Active Badge Location System," ACM Trans. Information Systems, 1992, pp.91-102.
- [5] Jay Werb and Colin Lanzl, "Designing a positioning system for finding things and people indoors," IEEE Spectrum, 1998, pp.71 - 78.
- [6] J. Hightower, et al., "SpotON: An indoor 3D location sensing technology based on RF signal strength," Technical Report UW CSE 00-02-02, Department of Computer Science and Engineering, University of Washington, Seattle, WA, Feb. 2000.
- [7] Y. Jang, et al., "Management Plan of Urban Object IDentification through Status- Analysis of Existing Object Management Code," The Journal of GIS Association of Korea, Vol. 16, No. 1, 2008, pp. 51-64.
- [8] K. Kang, J. Kim, "A System of Guiding Path for Parking Lots based on RFID to Consider Real-Time Constraints," The Journal of GIS Association of Korea, Vol. 16, No. 1, 2008, pp. 65-77.
- [9] Orr, Robert, et al., "The Smart Floor: a Mechanism for Natural User Identification and Tracking," Human Factors in Computing Systems, 2000.
- [10] "Location Technologies for GSM, GPRS and UMTS Networks," SnapTrack. 2003.
- [11] Bahl, P. and Padmanabhan, V.N..(2000),"RADAR: An in-building RF-based user location and tracking system," In IEEE Infocom 2000, Tel Aviv, Israel.
- [12] Roberto Battiti, et al., "Location-aware computing: a neural network model for determining location in wireless lans," Technical Report DIT-5, Universit'a di Trento, Dipartimento di Informatica e Telecomunicazioni, 2002.
- [13] S. Thrun, W.Burgard, D. Fox,"Probabilistic Robotics," MIT Press, 2005.
- [14] Davison, A. J., Murray, D. W., "Simultaneous localization and mapbuilding using active vision," IEEE Trans. Pattern Anal. Mach. Intell., IEEE Computer Society, Washington, DC, USA, v.24, n.7, 2002, p.865 -80. ISSN 0162-8828.
- [15] A. Davison, "Real-time simultaneous localization and mapping with a single camera," In Proc. International Conference on Computer Vision, Nice, 2003.
- [16] Davison, A. J., Molton, N. D., "MonoSLAM: Real-time single camera slam," IEEE Trans. Pattern Anal. Mach. Intell., IEEE Computer Society, 2007.
- [17] IEEE 1394 Interface, Wikipedia: <http://en.wikipedia.org/wiki/FireWire>.
- [18] J. Shi and C. Tomasi, "Good Features to Track," Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1994, pp.593-600.
- [19] N.D. Molton, A.J. Davison, and I.D. Reid, "Locally Planar Patch Features for Real-Time Structure from Motion," Proc. 15th British Machine Vision Conf., 2004.
- [20] Klein G., Murray D., "Parallel Tracking and Mapping for Small AR Workspaces," In Proc. International Symposium on Mixed and Augmented Reality, 2007.