

증강현실에서 가려짐 표현을 위한 실시간 영상 트래킹 및 조밀 복원 알고리즘

강현 김재현 구분기

한국전자통신연구원

hkang@etri.re.kr

Development of real-time camera tracking and dense reconstruction algorithm
for representing occlusion effect in augmented reality contents

Kang, Hyun Kim, Jae-Heon Koo, Bonki

Electronics and Telecommunications Research Institute

요약

증강현실 콘텐츠에서 가려짐 표현은 사실감을 위한 필수적이다. 가려짐이란 가상 물체의 일부가 실제 사물에 의해서 가려져서 표현되는 것을 의미한다. 이는 카메라의 트래킹 정보와 함께 현재 관찰하고 있는 실제 세계의 3차원 스캔 정보를 같이 획득해야 한다. 카메라 트래킹이란 현재 카메라가 세상의 어디에 위치해있는지, 어떤 자세로 바라보고 있는지에 대한 정보를 실시간으로 획득하는 기술이다. 3차원 스캐닝이란 실제 물체를 이루는 모든 점들의 3차원 위치 관계를 파악하는 과정이다. 가상물체의 3차원 위치와 자세가 실제물체의 위치/자세와의 관계를 통해서 가상물체의 그려야할 부분과 그리지 말아야할 부분을 판단할 수 있다.

본 논문에서는 트래킹과 스캐닝을 동시에 처리하여 가려짐 표현이 가능한 증강현실용 트래킹 기술을 제안한다. 카메라 트래킹이 실내외에서도 제약없이 동작하기 위해서 스테레오 카메라를 활용하였다. 트래킹은 카메라 프레임간 특징점들의 상호 관계를 파악하는 방법에 의해서 구하였다. 스테레오 카메라 이미지를 통해서 매 프레임마다 실제세계의 3차원 깊이정보를 파악하게 되고, 이를 앞서구한 카메라 위치자세를 통해서 3차원 깊이 데이터를 병합하는 과정으로 스캐닝기술을 구현하였다.

1. 서론

증강현실 기술은 현실 세계위에 가상 물체를 합성하여 현실의 정보 외에도 컴퓨터로부터 획득한 부가정보를 도시해주는 기술이다. 아즈마는 증강현실 시스템이 가지는 특징을 1) 현실과 가상을 결합할 것(Combine real and virtual), 2) 실시간으로 상호작용 할 것(interactive in real time), 3) 3차원 공간상에 있을 것(Registered in 3-D)으로 정의하였다[1]. 비어있는 방을 촬영하고 있는 카메라 시스템에서 가상의 의자를 놓기 위해서는 카메라로 찍은 영상을 분석하여 카메라와 방사이의 관계를 알아내야 한다. 또한 카메라가 움직여도 가상의 의자가 실제로 배치되어있는 것처럼 보여야 하므로 카메라의 움직임 정보를 알아내야 한다. 이는 카메라 트래킹 기술로 해결할 수 있다.

가상의 의자 앞에 실제의 책상이 있을 때 책상과 의자의 관계에 따라 의자의 어떤 부분은 가려져야 한다. 그렇지 않으면 현실감 없는 영상을 제공하게 된다. 따라서 현재 세계에 대한 3차원 정보를 획득하여야 하는데 이는 3차원 스캐닝 기술로 해결할 수 있다.

카메라 트래킹의 기존 사례 연구는 마커기반 방법이 있다. 이는 영상에서 분명한 특징을 나타내는 마커를 활용하여 트래킹을 수행하는데, 구현의 용이성과 비교적 적은 연산량 만으로 트래킹을 수행할 수 있어 많은 응용시스템에 적용되어 왔다. 대표적인 기술로는 AR toolkit [2]과 Vicon [3]등의 제품이 있다. 그러나 이 기술은 공간의 제

약이 있다는 단점을 가지고 있다.

임의 공간에 대한 카메라 트래킹 기술은 마커없이 영상에 나타나 는 자연스러운 특징을 사용한다. 대표적인 기술인 PTAM은 영상내에서 특징점을 구하고 영상간에 연관성을 구하여 이들을 3차원 맵으로 구축한다. 이러한 연관성 분석을 통한 트래킹 정보 획득과 3차원 맵을 만드는 과정을 동시에 하는 프레임워크를 만들었다[4]. 이는 3차원 맵을 구성하는 작업과 카메라 트래킹 작업을 병렬적으로 수행하여 실시간성을 확보하는 특징을 가지고 있다. 그러나 맵이 영상내 특징점을 기반으로만 만들기 때문에 대략적인 윤곽만이 나오며 스캐닝으로 얻는 데이터에 비해 현저히 적은 특징을 가지고 있다.

한편 깊이정보를 활용한 카메라 트래킹 기술로, 카메라 한 대로 연속된 프레임간을 스테레오 가정하여 매 프레임간의 스테레오 맵스영상을 사용하여 실제 물체의 기하정보를 수 밀리미터내로 스캔이 가능하면서도 카메라 트래킹에 필요한 정보를 획득하였다[5]. 그러나 연산량이 너무 많고 조명 변화에 취약한 단점을 가진다.

본 논문에서는 임의공간 영상 합성용 트래킹 기술을 제안한다. 스테레오 카메라를 사용하여 실내외에서도 트래킹이 가능하며 조명에 안정적인 효과를 가지며, 맵스카메라 수준의 조밀맵을 생성할수 있다. 이를 통해서 가상 물체와 실제 물체의 정합시에 가려짐 효과가 제시할 수 있다.

2. 제안 시스템

본 연구에서는 가상물체의 실제 물체에 의한 가려짐 표현하기 위해서 두가지 모듈로 구성된 방법을 제안한다. 하나는 카메라 위치/자세 트래킹 모듈이며, 다른 하나는 조밀복원 모듈이다. 카메라 트래킹 모듈은 실시간으로 입력된 영상을 분석하여 카메라의 3차원 위치와 자세를 알아내는 기능을 갖는다. 조밀복원 모듈은 영상을 분석하여 깊이 데이터를 생성하고 이들을 카메라 트래킹 모듈의 위치/자세를 고려하여 깊이 데이터를 융합한다.

카메라 트래킹 모듈은 영상 프레임 안에서 특별한 정보를 가진 점들을 먼저 찾는다. 특징점은 사전정보가 있는 여부에 따라 마커 혹은 내추럴 마커 기법으로 나눌 수 있으며, 제안된 시스템에서는 안정적인 성능을 위해서 마커기법을 도입하였다. 단위 기술로 내추럴 마커 기법도 사용하였으나 본 논문의 범위에 속하지 않는다. 마커는 전반사 원형 스티커를 사용하였다. 카메라 근처에 있는 LED 빛에 의해서 마커만 도드라지게 영상에서 나타난다. 이진화 알고리즘과 연결요소 분석법등을 통하여 마커의 여부와 위치를 알 수 있다.

제안된 시스템은 스테레오 카메라를 사용하므로 이 마커는 좌우 영상에 동시에 나타날 수 있다. 같은 마커의 경우 에피폴라 라인 상에 있어야 하며 이를 통해 좌우영상에서 연관성을 가진 마커 상의 쌍을 구할 수 있다. 매 프레임 마다 수행하면 마커 상의 쌍의 연관리스트를 수집할 수 있다. 좌우 카메라에서 나타난 마커 상의 쌍집합을 통해서 RANSAC을 통해 랜덤하게 샘플을 수집한다. 수집된 샘플로 현재 카메라의 위치/자세를 Levenberg-Marquardt방법을 통해 구한다[6].

조밀 복원 모듈은 깊이 데이터를 생성하는 스테레오 영상처리부와 깊이데이터를 저장하는 복셀처리부로 나뉜다. 스테레오 맵스 맵을 구하는 방법은 Semi Global Matching 기법을 사용하였다[7]. 이 맵스 맵을 시간적으로 늘리며 공간적으로 확장하기 위해서는 복셀 해쉬 기법을 사용하였다 [8]. 매 프레임마다 계산된 카메라의 위치/자세와 깊이 데이터를 가지고 월드좌표계의 깊이 데이터들을 만들어 낸다. 이를 복셀 해쉬 기법을 통해서 데이터를 수집을 한다.

3. 실험 결과 및 분석

제안된 시스템은 스테레오 카메라, 영상처리용 PC로 구성되어 있으며 마커기반의 트래킹 모듈을 사용하였기 때문에 마커가 붙은 피사체 및 스테레오 카메라에 LED 전구를 부착하였다. 카메라 트래킹 모듈에 비해 연산량이 많은 조밀복원 모듈은 CUDA로 구현하였다. 스테레오 카메라는 글로벌 서터의 1280x960해상도와 45HZ의 동기화된 좌우영상을 출력하는 oCamS-1CGN-U제품[9]을 사용하였다. 조밀복원의 결과는 Raycasting기법을 CUDA로 구현하여 영상화 하였다.

표 1. 제안된 시스템 주요 성능치

마커 트래킹 처리속도	스테레오영상 처리속도	복셀 해쉬 처리속도	전체 시스템 성능
35 ms 이상	11.7ms 이상	1.2ms 이상	최대 45 fps

주된 결과는 그림 1에 제시하였으며, 시스템의 주요성능치는 표1에 표시하였다. 카메라 영상획득, 마커트래킹, 조밀복원이 각각 독립적으로 수행하며 이들 간의 데이터 흐름이 있는 구조로 구현되었으며

최종 전체 성능은 45 frame per seconds를 이루었다(CPU: i7 6700K, GPU:nVidia 1080TI 11GB, RAM:16GB, OS: Windows 10)

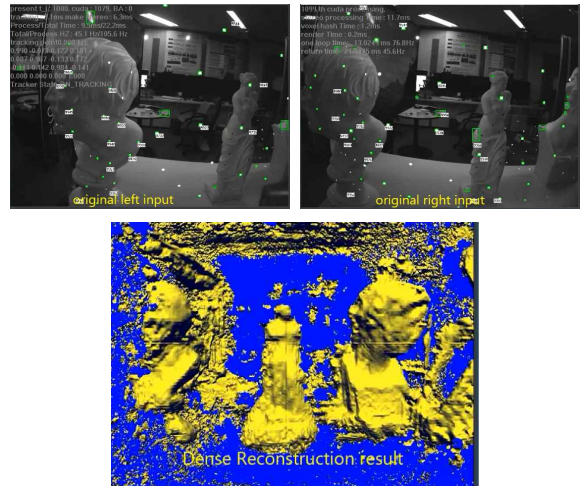


그림 1. 제안된 시스템 결과물 영상(입력 좌우, 조밀복원출력)

4. 결론 및 향후 연구방향

현재 출시된 증강현실 시스템은 카메라의 위치/자세의 트래킹에만 집중되고 있는데 이는 소형기기의 연산 처리능력이 문제일 수 있다. 조밀복원을 통해 실제 세상에 대한 정보를 수집하게 되면 더 현실감있는 증강현실 시스템이 나타날 것이다. 본 연구가 포함된 프로젝트는 이를 목표로 필요한 기술을 축척해 나가고 있다. 조밀복원에서는 특징점 기반의 데이터와 조밀 복원된 데이터간의 상호관계를 이용한 연구를 계획하고 있다. 또한 스테레오 영상처리를 다중영상 영상처리로 확장하여 깊이 데이터의 품질을 높여려 하고 있다.

감사의 글

이 논문은 2018년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (과제번호 : 2017-0-01849, 과제명 : 실내의 임의공간 실시간 영상 합성을 위한 핵심 원천기술 및 개발툴킷 개발 (Development of Core Technology for Real-Time Image Composition in Unstructured In-outdoor Environment))

5. 참고문헌

- [1] R.T. Azuma, "A Survey of Augmented Reality," In Presence: Teleoperator and Virtual Environments 6(4), 1997
- [2] H. Kato et al., "Marker Tracking and HMD Calibration for a video-based Augmented Reality Conferencing System," IWAR 1999
- [3] <https://www.vicon.com/>
- [4] G. Klein et al., "Parallel Tracking and Mapping for Small AR Workspaces," ISMAR 2007
- [5] R. A. Newcombe, "DTAM: Dense tracking and mapping in real-time," ICCV 2011
- [6] 김재현 외, "스테레오 비주얼 SLAM을 이용한 핸드헬드 스캐너 구현," 한국로봇종합학술대회 2015
- [7] H. Hirschmuller, "Accurate and Efficient Stereo Processing by Semi Global Matching and Mutual Information," CVPR 2005
- [8] M. Nießner, "Real-time 3d reconstruction at scale using voxel hashing," TOG, 32(6):169, 2013.
- [9] <http://withrobot.com/camera/ocams-1cgn-u/>