일반화된 모델의 변형에 의한 실내 통로공간 추적

Indoor Passage Tracking based Transformed Generic Model

이서진, 남양희
이화여자대학교 디지털미디어학부

Seo-Jin Lee(moon0329@ewhain.net), Yang-Hee Nam(yanghee@ewha.ac.kr)

요약

증강현실에서 3차원적 증강은 입력 비디오나 이미지로부터 3차원 구조 복원 및 추적을 필요로 한다. 이를 위해 제안된 기존의 방법들은 대개 정확한 3차원 모델을 사전에 구축해두고 실시간에 대조하는 방식을 취하였는데, 이 방식은 정확한 측정에 기반한 모델이 있어야 한다는 점과 대상을 일일이 모델링해야 하는 문제점이 있다. 본 논문은 각 대상물별 정밀한 모델 없이 유형별로 일반화된(generic) 모델만을 사용하는 방식을 제안함으로써 중간 이동 등을 허용하는 광범위 이동형 실내 증강현실 응용 가능성을 제시하고자 하였다. 제안한 방법은 일반화된(generic) 모델을 변형(affine transformation)하여 주어진 장면에서 오류 임계치 이내의 정합을 이루는 모델 변형 값인 스케일, 위치, 회전 값을 찾아냄으로써, 그에 따라 정합된 3차원 공간구조에 관해 일관성 있게 증강객체가 배치될 수 있도록 하는 것이다. 이 방법은 정밀 모델링에 드는 시간과 노력 비용을 줄이며, 실험을 통해 크기나 디테일은 다르지만 유사 패턴이 반복되는 통로구조의 복원과 추적에 사용될 수 있음을 보였다.

■ 중심어 : | 증강현실 | 3차원 구조인지 | SLAM |

Abstract

In Augmented Reality, it needs restoration and tracking of a real-time scene structure for the augmented 3D model from input video or images. Most of the previous approaches construct accurate 3D models in advance and try to fit them in real-time. However, it is difficult to measure 3D model accurately and requires long pre-processing time to construct exact 3D model specifically. In this research, we suggest a real-time scene structure analysis method for the wide indoor mobile augmented reality, using only generic models without exact pre-constructed models. Our approach reduces cost and time by removing exact modeling process and demonstrates the method for restoration and tracking of the indoor repetitive scene structure such as corridors and stairways in different scales and details.

■ keyword : | Augmented Reality | 3D Scene Structure Analysis | SLAM |

I. 서론

최근에 새로운 형태의 콘텐츠를 제공 받고자 하는 사

용자의 요구와 함께 증가되고 있는 실세계와 가상 환경을 결합하는 증강현실 시스템의 개발이 활발히 이루어지고 있다. 이 중에서도 가이드 시스템이나 게임 등과

* 본 연구는 서울시의 산학연 협력사업의 전략산업 혁신 클러스터 육성 지원 사업(10581) 수혜에 따른 결과입니다.

접수번호 : #100303-006

접수일자 : 2010년 03월 03일

심사완료일 : 2010년 04월 15일

교신저자 : 남양희, e-mail : yanghee@ewha.ac.kr

같은 증강현실 콘텐츠의 경우, 실제 공간 구조에 맞게 가상 객체들을 정합시켜주어야 한다. 그렇게 하기 위해서는 실세계의 좁은 공간이나 특정한 객체가 아닌, 사용자가 이동하는 환경에 대한 3차원 구조를 알아야 할 필요성이 요구된다. 실세계의 3차원 구조를 알아내는 것은 증강현실 3차원 객체를 실세계 공간의 구조에 맞게 디스플레이 하기 위해 필요하다. 따라서, 실세계 공간의 장소 및 3차원적 정보를 알아내고 연속 영상에서 그 변화를 추적하는 데 많은 연구의 노력이 집중되어 왔다.

하지만 실세계 공간의 구조에 대한 인식 또는 재구축에 대해 연구했던 기존 논문 대부분은 미리 장소를 3차원 적으로 모델링 해 두거나 책상 주변, 방 하나 등의 매우 한정되고 지역적인 영역에만 한하여 적용된다. 따라서 사용자가 여러 공간을 이동하는 모바일 환경에 적용하기 힘든 부분이 있었다. 특히 건물 내를 이동하며 사용하는 증강현실 시스템의 경우 기존 논문들을 그대로 건물 내의 복도와 계단과 같은 이동통로에 적용하여 사용하기에는 제약점이 있었다. 따라서 본 논문에서는 기존 기술로는 해결하기 힘든, 건물 등 실내에서의 이동 환경을 대상으로 한 공간 인식 문제를 해결하고자 한다.

실시간 3차원 구조인지 방법에는 일반적으로 마커를 사용하여 카메라를 통해 이를 인식한 후 마커의 위치와 방향에 맞게 가상 객체를 정합하는 마커 기반 방식과 RFID나 적외선 등과 같은 센서 기반 방법이 있다. 그리고 별도의 센서나 마커 없이 카메라로부터 들어오는 영상만을 분석하여 영상 정합을 위한 정보를 얻어내는 카메라 기반 방법 등이 있으며, 최근에는 카메라를 기반으로 하는 마커프리 증강현실 연구가 주류를 이루고 있는 추세이다[1-3].

카메라를 사용한 기존 논문은 크게 모델기반방법과 특징기반방법으로 나눌 수 있다. 모델 기반 방법은 CAD 모델과 같은 3차원 모델을 이용하여 카메라의 상태를 추적하는 연구이다. 즉, 카메라를 초기화 한 후 2차원 영상에 투영(Projection)된 모델과 영상에서 추출한 특징들과의 에러를 최소화하는 카메라의 위치와 방향을 실시간으로 추적하는 방법이다. 이 방법은 비교적 높은 계산의 효율성과 낮은 복잡성을 보이지만 사전에

대상이 되는 장소에 대한 정확한 사전 모델링이 요구된다. 따라서 건물 내부와 같이 대상 환경이 넓거나 다양한 환경에 적용할 경우 공간을 사전모델링하기 위한 비용과 시간의 부담이 상당히 크다는 단점이 있다.

특징기반방법은 사전 정보 없이 별도의 카메라 초기화 과정을 거친 후, 실시간에 들어오는 영상의 선과 점 등의 특징을 분석하여 실세계 구조나 객체에 대한 3차원 정보를 추정하고 카메라 상태를 예측하는 방법이다. 이 방법은 사전 정보 구축을 위한 노력이나 비용 없이 실시간으로 들어오는 특징들로부터 평면과 같은 3차원 구조를 예측할 수 있다. 하지만 특징의 폐색과 조명에 민감하고 텍스처가 없는 환경에서 비교적 부정확하기 때문에 특징이 별로 없고 조명의 영향을 받는 실내 환경에서는 정확히 3차원 구조를 재구축하기 어려우며 카메라 초기화를 위한 별도의 마커가 필요하다는 제약이 가진다.

따라서 본 논문에서는 이러한 기존논문들의 한계점을 극복한 실내 환경에 적합한 공간인식 방법을 제안하고 특히 사용자의 이동이 대부분 이루어지는 복도와 계단에 이를 적용 하고자 한다.

이 논문에서 제안하는 방법은 일반화 된 3차원 모델을 사용하는 방법으로, 기존 모델기반 방법에 비해 모델링을 위한 비용과 시간의 노력을 줄일 수 있다. 또한 일반화 된 모델을 자동 피팅하여 이동 시 자연스러운 공간인지가 가능하도록 하고 또한 모델과 영상의 텍스처를 함께 카메라 추적에 사용하여 텍스처 특징이 없는 이동 통로에서 정확한 카메라 추적이 가능하도록 한다.

II. 관련연구

2차원 영상에서 3차원 구조를 인지하는 문제는 로봇 비전, 이미지 기반 모델링 분야에서 활발히 연구되고 있는 분야이다. 또한 증강현실 분야에서도 최근 가상 객체와 실세계의 정합을 위하여 센서나 마커를 사용하지 않고 카메라로부터 입력 받은 영상만을 분석하여 실세계의 3차원 구조를 알아내려는 시도가 계속되고 있다.

카메라 기반으로 3차원 구조를 인지했던 기존 연구를 앞에서 언급했던 것처럼 크게 모델 기반과 특징 기반 방법으로 나누어 살펴보겠다.

모델 기반 방법이란 가상 객체를 증강할 대상이 되는 실세계의 장소나 객체에 대해 사전에 3차원 모델을 구축해놓는 방법이다. 영상과 매칭한 후 카메라를 초기화하고 모델 정보를 기반으로 카메라 상태를 추적하여 모델의 3차원 구조에 맞게 가상객체를 정합하는 방법을 말한다[4-6]. Heris의 RAPid시스템[5]은 대표적인 실시간 3D 트래커(Tracker)로서 미리 대상에 대한 3차원 모델을 모델링 해두고, 투영된 3차원 모델의 엣지와 이미지에서 강한 대비를 갖는 선들을 이용하여 최적의 카메라 위치를 추정하는 방법이다. 하지만 이 논문의 경우 영상의 크기가 크거나 대상 물체와 카메라의 거리가 멀거나 주변 배경이 복잡한 환경에서는 부정확한 결과를 보이고, 지역적인 물체에는 정확한 결과를 보이지만 공간이 넓은 이동통로에 적용하기 어렵다. 또한 사전에 대상에 대한 정확한 모델링이 요구되어 이동통로를 모델링하기 위해서는 높은 비용과 시간이 요구된다. Andrew[6]은 정확하고 빠른 모델 기반 추적 방법을 위하여 기존의 최적화 방법 대신 로봇 비전 분야에서 쓰이던 Virtual Visual Servoing(VVS)기술을 적용한 방법을 제안하였다. 이 방법 역시 영상의 폐색에 강건한 특징을 보이나 기존의 모델기반방법과 마찬가지로 미리 실세계에 맞는 3차원 모델을 구축해 놓아야만 비교적 정확한 3차원 구조 인지가 가능하기 때문에 이동통로에 적용하기에는 비용과 시간의 부담이 크다.

특징 기반 방법은 사전 정보 없이 별도의 카메라 초기화 과정을 거친 후, 실시간에 들어오는 영상의 선과 점 등의 특징을 분석하여 실세계 구조나 객체에 대한 3차원 정보를 추정하고 카메라 상태를 예측하는 방법이다. Simon[7]은 실시간에 실내 환경에서 텍스처의 선들을 군집화(Clustering) 하는 방법으로 평면을 인지해 실내 환경을 재구축하는 방법을 제안하였다. 이미 알고 있는 3차원 모델로 카메라를 초기화하고 호모그래피(homography)를 구하여 각 평면에 해당되는 영역을 군집화 한다. 이 방법은 비교적 정확하고 효과적으로 3차원 구조를 재구축 해줄 수 있으나 별도의 3D 모델을 사

용하는 카메라 초기화 과정과 벽면을 찾기 위해 영상의 바닥 영역에 해당하는 부분을 사용자가 직접 지정하는 추가적인 과정이 필요하여 사용자가 이동하는 모바일 환경에는 적용하기 힘들다. 또한 텍스처 특징이 적은 이동통로 환경에 적용하기 힘들다. Gee와 Denis[8-9]는 로봇 비전에서 쓰이던 SLAM (Simultaneous Localization and Mapping) 방법을 이용해 실시간으로 영상의 텍스처 특징의 3차원 정보를 읽어오며 특징 점의 군집으로부터 평면을 인지하는 방법을 제안하였다. SLAM 방법은 카메라가 이동할 때 주변 환경을 인지하여 주변 환경에 대한 지도를 구축하고 동시에 지도를 이용하여 카메라의 위치를 추정해 내는 방법이다. 이 방법은 SLAM 방법을 이용함으로써 사전정보 없이 실시간에 들어오는 영상들로부터 3차원 구조를 인지하는 것이 가능하나, 카메라의 초기화를 위해 이미 실험 전에 3차원 위치를 알고 있는 코너 점이 필요하다. 그리고 국소적인 평면 인지는 가능하나 바닥 전체의 넓은 평면 인지는 어려우며 조명과 폐색에 민감하다는 단점이 있다. 따라서 실내조명의 영향을 많이 받고 텍스처 특징이 많지 않은 이동통로에 적용하기에는 한계가 있다. 위 방법보다 개선된 카메라 추적 방법으로 Gee[10]은 SLAM (Simultaneous Localization And Mapping)방법의 카메라 추적 부분에 모델 기반 추적 방법에서 주로 쓰이던 VVS방법을 결합한 새로운 특징기반 추적방법을 제안하였다. 먼저 마커를 이용해 카메라 상태를 초기화 한 후, 실시간으로 들어오는 영상에서 엣지(edge)를 추출하고 3차원 위치를 추정한다. 또한 동시에 3차원 위치가 추정된 엣지에 Visual Servoing 방법을 적용하여 카메라 위치를 추정한다. 이 방법은 기존의 방법에 비해 비교적 정확한 카메라 추적이 가능하나, 카메라 초기화를 위한 인위적인 마커가 필요하며 정확한 추적을 위해서는 영상에 3개 이상의 끝점이 명확한 엣지가 보여야 한다는 제약점을 가진다. 따라서 조명의 영향으로 명확한 끝점을 갖는 선들을 추출하기 어려운 이동통로에 적용하기 힘든 점과 카메라 초기화를 위한 인위적인 마커 또한 사용자가 여러 공간을 이동하는 모바일 환경에서는 큰 제약점으로 작용한다.

Klein[11]은 듀얼코어 컴퓨터에서 병렬 스레드 처리

로 tracking과 mapping을 구분하여 처리함으로써 기존의 SLAM시스템보다 빠르고 많은 특징 점들을 제어할 수 있는 시스템을 제안하였다. 이 논문은 기존의 방법보다 높은 정확성과 처리 속도를 보이지만 트래킹을 통해 찾아낸 특징들이 나타내는 것이 무엇을 의미하는지 알 수 없다. Robert[12]는 작은 평면이지만 가능하며 Gerhard[13]은 특정 건물의 정확한 모습을 나타내는 이미지 맵이 필요하여 일반화된 이동통로 전체에 적용하기에는 문제가 있다.

앞에서 살펴본 것과 같이 기존의 논문들을 이동통로에 적용하여 3차원을 재구축하기에는 여러 가지 제약점이 있었다. 따라서 본 연구는 이러한 기존논문의 한계점을 극복하기 위해 3차원 구조를 재구축하는 문제에 대해 다음과 같이 해결하고자 한다.

첫째, 정확한 모델 구축에 소모되는 노력 및 시간을 줄인다. 이동통로와 같은 넓은 공간에 대하여 기존 연구와 같이 사전에 정확한 정보를 구축하기 위해서는 상당한 노력과 시간이 소요된다. 그래서 본 연구에서는 실제 환경에 맞게 정확히 모델링 되지는 않았지만, 실제 환경에 맞게 변형될 수 있는 일반화 된 모델을 사용하여 사전 정보 구축의 어려움을 줄이고자 한다.

둘째, 크기가 다른 이동통로에 적용 가능하게 한다. 본 연구에서는 건물의 복도와 계단과 같은 이동 공간의 공통된 특성을 일반화 한 3차원 모델을 정의하고 일반화 된 모델을 장소에 맞게 변형하여, 비슷한 구조를 가지는 건물에서도 별도의 모델링 과정 없이 공간구조인지가 가능하게 한다.

셋째, 카메라 초기화를 위한 사전 정보를 최소화 한다. 기존 논문에서는 미리 알고 있는 모델을 영상에 인위적으로 피팅(Fitting)하거나 마커(Marker)를 사용하여 카메라 초기화(camera calibration)를 해주어야 했다. 본 연구에서는 일반화 된 모델을 피팅 하는 과정에서 모델을 바라보고 있는 카메라의 초기 상태를 자동으로 구함으로써 카메라 초기화를 위한 사전 정보와 사용자 입력을 최소화 하였다.

III. 실시간 공간구조 인지 알고리즘

본 논문은 단일 카메라로부터 실시간으로 들어오는 영상을 분석하여 텍스처가 분명하지 않은 넓은 구조의 3차원 정보를 추정해 내는 것을 목표로 한다. 본 논문에서는 두 가지의 접근방법으로 이동통로 내에서의 3차원 구조 인지 문제를 해결하였다.

첫 번째로 일반화 된 3차원 모델을 정의하고, 이를 실제 환경에 맞게 자동으로 피팅시켜 준다. 본 논문에서는 특정한 건물에 제한되지 않는 일반화 된 계단, 복도 모델을 구축하고, 특정 건물에 맞게 이를 자동으로 피팅하는 방법을 사용한다. 일반화 된 장소모델을 사용함으로써 정확한 장소모델을 구축하기 위한 시간 및 노력을 줄였으며 기존논문에서 필요했던 인위적인 카메라 초기화 과정을 자동 피팅 과정으로 대체하였다.

두 번째로 이동통로 특징에 맞는 카메라 추적방법을 제안한다. 일반화 된 3차원 모델을 피팅한 후 모델의 특징 코너 점과 선 정보를 카메라 추적에 사용하고 카메라 추적과 동시에 추적에 사용될 영상의 텍스처 특징점을 추출하는 SLAM(Simultaneous Localization And Mapping) 방법을 적용한다. 이때 본 논문에서는 특징만을 사용하는 SLAM 방법의 부정확성을 보완하기 위해 특징기반으로 카메라 상태를 추정한 후 모델의 특징을 사용하여 이를 보정해 주는 방법을 사용한다.

그림 1. 이동 통로 재구축 시스템 흐름도

[그림 1]은 두 가지의 접근방법이 본 연구의 알고리즘에 적용된 과정을 보여준다. 먼저 실시간으로 카메라를 통해 들어오는 영상에 미리 구축되어 있는 일반화 된 모델을 자동으로 피팅시킨다. 이때, 모델의 크기 값

그림 2. 전체 시스템 다이어그램

이 실제 환경에 맞게 변형되며 모델을 바라보고 있는 카메라의 초기 위치와 방향이 추정된다(Camera Calibration). 초기 모델 피팅 과정이 끝나면, 카메라의 움직임에 따라 모델을 연속적으로 피팅시키기 위하여 카메라의 위치와 방향을 실시간에 추적한다.

이때, 피팅 된 모델의 코너 점이 카메라 추적에 사용되며, 카메라 추적과 동시에 실시간에 카메라 영상으로부터 텍스처 코너 점이 추출되어 카메라 추적에 사용된다. 카메라 추적 단계에서는 이동통로 특성에 맞는 카메라 추적을 위해 특징기반으로 카메라 상태를 추정한 후 모델을 기반으로 이를 보정해 준다.

1. 일반화된 3차원 모델 정의

본 논문에서 다루고 있는 건물 내에서의 이동의 경우, 대부분 복도와 계단으로 이루어진다. 따라서 본 논문에서는 공간구조 인지를 복도와 계단으로 한정하였다. 건물의 복도와 계단을 살펴보면 대부분 일반적이며 공통적인 구조를 가진다. 복도는 대부분 직육면체의 구조를 가지고 있으며, 계단은 같은 크기의 층계가 경사를 이

루며 주기적으로 반복되는 구조를 가진다. 따라서 본 연구에서는 이와 같은 이동통로의 특성에 기반을 두어 특정 건물에 한정되지 않으며 크기가 정해지지 않은 일반화된 3차원 복도, 계단 모델을 그림 3과 같이 정의하였다.

그림 3. 일반화된 3차원 모델

이러한 3차원 모델의 크기를 장소에 맞게 피팅 파라미터를 자동으로 추출함으로써 공간구조 인지의 상황 적응력을 높이고자 하였다.

2. 일반화된 3차원 모델 피팅

본 논문에서는 투영된 일반 모델과 실제 카메라로 입

력 받은 이미지에서 실제 대상 공간에 맞는 모델 파라미터와 카메라 초기 파라미터를 구한다. 일반화된 모델은 모델의 특정한 한 면을 바라보는 방향으로 투영되며 모델 피팅 과정을 통해 구하고자 하는 파라미터 P 는 식 (1)과 같이 9차원 벡터로 이루어진다.

$$P = \{sx, sy, sz, cx, cy, cz, roll, yaw, pitch\} \quad (1)$$

파라미터 P 는 모델의 크기 파라미터(sx, sy, sz)와 카메라의 초기상태 위치(cx, cy, cz), 회전값($roll, yaw, pitch$)로 이루어지며, 2차원 이미지 평면에 투영된 모델의 선과 카메라로부터 들어오는 비디오 이미지에서 추출한 공간 구조의 선들의 차이를 최소화 하는 값으로 구해진다.

[그림 2]의 모델 피팅 과정은 먼저, 비디오 이미지에서 공간구조를 이루는 주요 선들만을 추출해내는 이미지 전처리 과정을 거친다. 그리고 전처리 과정을 거쳐 추출된 선과 일반화 된 3차원 모델을 투영시킨 모델 선과의 차이를 가장 최소화 하는 카메라의 초기 상태와 모델 크기를 구하는 파라미터 최적화 단계를 거친다.

그림 4. 모델 피팅 결과

먼저 일반화 된 3차원 모델을 카메라로부터 입력된 영상에 피팅하기 위하여 노이즈 제거, 엣지 추출, 선 추출 등의 이미지 처리 방법을 사용하여 영상에서 공간구조에 해당되는 기본적인 선들만을 추출하게 된다. 그리고 안정성이 높은 비선형 최적화 알고리즘인 Levenberg-Marquardt(LM)방법을 사용하여 파라미터를 추정한다[14]. 본 논문에서는 Moving Edge알고리즘[15]을 사용하여 전처리 과정에서 정의된 제어점 주변의 선들을 찾아 LM의 에러를 계산하였다. LM 방법을 수행하는 과정에서 반복한 실험 결과를 통해 얻은 오류 임계치(error threshold)보다 에러율이 낮으면 모델 피팅 과정을 끝내고 [그림 1]의 카메라 추적과정으로 넘어간다. [그림 4]는 파라미터 P 가 자동으로 구해진 후 일반화된 모델이 실제 영상에 피팅된 모습이다.

3. 영상특징 기반 카메라 추적 및 특징 추출

카메라의 움직임에 따라 모델을 연속적으로 피팅 하기 위하여 카메라(또는 사용자)의 위치 및 방향을 추적한다. 본 연구에서는 앞에서 언급했던 것처럼 SLAM을 기반으로 한 영상특징기반 카메라 추적방법을 제안한다.

SLAM은 이동로봇의 위치추적에 쓰이던 방법으로 이동로봇의 위치인식(Localization)과 벽, 통로등과 같은 환경에 대한 지도 작성(Mapping)을 동시에 수행하는 작업을 말한다. 이는 본 연구와 같은 카메라 기반 시스템에서 카메라의 위치를 추적하고 동시에 주변의 특징들의 정보를 저장하는 것을 의미한다. 그리고 본 논문에서는 SLAM과정에서 주변의 특징들에 대한 정보를 저장하고 있는 지도를 특징맵이라고 정의한다 [16-19].

특징들의 정보를 저장하는 특징맵 구축(Mapping)하는 과정은 다음과 같다. 본 논문에서 특징맵은 모델 피팅 과정을 거쳐 구해진 3차원 모델의 코너점과 영상으로부터 추출된 텍스처 특징점으로 이루어진다. 먼저 초기 피팅 과정을 거쳐 구해진 모델로부터 4개의 코너 점의 3차원 위치가 맵에 저장된다, 실시간 영상으로부터 2차원 텍스처 특징점을 추출하면, 카메라의 움직임과 특징점의 위치변화를 기반으로 텍스처 특징점의 3차원

위치가 추정되어 특징 맵에 저장된다. 이때 텍스처 특징점은 FAST코너 추출방법[20]을 이용하여 영상에서 추출되며 Inverse Parameterization방법[17]을 통해 3차원 위치가 추정된다. 추출된 코너 점을 저장하고 매칭하는 문제는 카메라의 움직임에 따라 특징을 추적해 나가면서 특징의 3차원 위치와 카메라의 위치와 방향을 예측해야 하기 때문에 비교적 정확한 방법이 요구된다. 따라서 본 연구에서는 스케일과 방향에 강건한 SIFT(Scale Inverse Feature Transform Description) [21]기술자를 이용하여 코너 점을 저장한다.

위치 인식과정(Localization)은 크게 특징 맵에 저장된 특징점들의 위치와 확장칼만필터(Extended Kalman Filter) 방법을 사용하여 카메라 상태를 추적하는 1차 카메라 추적방법과 모델의 선 정보와 [그림 5]와 같이 LM방법을 이용하여 2차로 카메라를 보정해주는 단계로 이루어진다. 칼만필터는 관측 방정식을 이용하여 이동로봇의 상태변수를 추정하는데 사용되는 대표적인 최적 추정 알고리즘으로 SLAM을 구현하는데 가장 많이 쓰이는 방법이며 확장칼만필터는 비선형식을 선형화하여 칼만필터를 적용한 방법이다.

그림 5. LM방법으로 보정 전(위), 보정 후(아래)

확장칼만필터는 전 프레임에서의 카메라 상태를 기반으로 현재의 카메라 상태를 예측하는 단계(Predict Step)와 현재프레임에서 들어온 특징과 예측된 카메라로부터 투영된 특징지도 안의 특징들의 위치 차이를 이용한다. 카메라 상태와 특징 상태를 보정(Update Step)

해주는 단계를 재귀적으로 실행해 주며 연속적으로 카메라 상태와 특징 상태의 정확성을 높여준다.

특징점 외에도 모델 엣지는 추적에 있어 중요한 정보인데, 본 논문에서는 확장칼만필터만을 사용한 카메라 추적결과를 모델엣지를 기반으로 LM방법을 사용하여 카메라 값을 보정해준다. 이는 이동통로 환경에서 보다 정확한 카메라 추적이 가능하도록 하였다. [그림 6]은 에러 변화량을 나타내는 그래프이다. 그래프에서 알 수 있듯이 LM 방법으로 보정한 결과 에러율이 더 낮다.

그림 6. LM 방법을 이용한 카메라 위치 보정 결과 에러 그래프

IV. 구현 및 실험 결과

본 논문의 시스템은 윈도우즈 XP Professional 2.40GHz, 2.00GB RAM, NVIDIA GeForce 7600GT 비디오 가속기가 달린 PC에서 구현되었다. 카메라는 초당 30프레임 처리 속도의 CCD센서 카메라를 사용하였다. Visual C++ 6.0 환경에서 영상처리 라이브러리 OpenCV를 사용하여 구현되었으며 실험 영상의 전체 해상도는 320X240이다.

[그림 7]은 모델 피팅 과정으로부터 구해진 모델의 상태와 카메라 초기 상태를 기반으로 실시간에 카메라 추적을 한 결과 이미지이다. 실험 결과 이동통로에서 비교적 정확한 카메라 추적이 가능 하였으며 [그림 8]과 같이 프레임의 변화에 따라 영상과 3D 모델에 대한 에러율이 줄어드는 것을 확인할 수 있었다. [그림 8]의

피팅1은 [그림 7]의 복도의 피팅 결과를 나타내며 피팅 2는 [그림 8]의 계단에서 실험한 결과를 나타낸다. 계단의 경우는 사용자가 미리 계단의 수를 지정해야한다.

그림 7. 모델 추적 결과

또한 [표 1]에서 보이는 결과처럼 실시간 추적이 가능하다는 것을 1프레임 당 모델 추적의 각 단계의 처리 시간을 통해 알 수 있다. 또한 [표 2]에 기존 연구와의 실행 결과를 비교하였다.

그림 8. 모델 피팅 과정의 어려움

표 1. 모델 추적 처리 시간

처리 과정	1프레임 당 처리 시간 (약 7 frame/sec)
전처리	약 0.005 sec
EKF를 이용한 추적	약 0.08 sec
LM 방법을 이용한 결과 보정	약 0.0016 sec
총	약 0.0766 sec

V. 결론 및 향후 연구 과제

본 논문에서는 카메라로 입력된 영상에서 실내 건물의 이동통로의 3차원 구조 재구축 방법을 제안하였다. 이를 위해, 단일 카메라를 이용하고 실시간에 가능하며, 특정 건물에 한정되지 않는 알고리즘을 제안하였다.

실험결과, 본 논문에서 제안한 공간구조 인지 알고리즘은 특정건물에 정확히 모델링이 되지 않은 일반적인 3차원 모델만으로도 대상 건물에 맞는 공간구조 인식이 가능한 결과를 보였으며 비교적 실시간성을 유지하였다. 우리는 향후, 크기 외에도 모델의 일반적인 특성을 가지되, 특정 건물에 따라 달라질 수 있는 상황변수에 대한 추가적인 연구를 수행할 예정이다. 또한 현재 PC기반에서 구현된 시스템을 모바일 플랫폼에 적용하여야 하며, 모바일 증강현실 애플리케이션의 개발이 뒤따라야 할 것이다.

표 2. 기존 연구와 제안 연구의 실행 결과 비교

	모델기반 방법	특징기반 방법			제안 방법
	Virtual Visual Servoing[6]	Walls Detection [7]	Visual SLAM [8-9]	Model based SLAM [10]	
대상 환경	사전에 모델링된 장소	텍스처 특징이 있는 벽	텍스처 특징이 있는 영역	텍스처 특징이 있는 영역	기본 육면체형 복도와 계단
카메라 초기화	사용자의 인위적 모델 피팅	3차원 모델 삽입	3차원 정보를 알고 있는 마커 삽입	3차원 정보를 알고 있는 마커 삽입	모델 피팅에 의한 자동화
상황 적응력	모델링 된 대상 객체, 장소외 적용 못함	텍스처 특징이 있는 영역만 가능	텍스처 특징이 있는 영역만 가능	텍스처 특징이 있는 영역만 가능	기본 육면체형 건물에 적용 가능
사용자 추가 입력	없음	바닥부분 영역 지정	없음	없음	없음

참 고 문 헌

- [1] G. Simon, A. W. Fitzgibbon, and A. Zisserman, "Markerless Tracking using Planar Structures in the Scene," International Symposium on

- Augmented Reality, pp.120-128, 2000.
- [2] L. Libuda and K. Kraiss, "Identification of Natural Landmarks for Vision Based Navigation," Proceedings of IEEE International Conference on Robotics and Automation, 2000.
- [3] R. Gerhard and T. W. Drummond, "Going Out: Robust Model-based Tracking for Outdoor Augmented Reality," IEEE International Symposium on Mixed and Augmented Reality, 2006.
- [4] I. C. Andrew, M. Eric, and C. Francois, "Robust Model-based Tracking for Robot Vision," IEEE International Conference on Intelligent Robots and Systems, 2004 .
- [5] C. Harris, *Tracking with Rigid Objects*, MIT Press, 1992.
- [6] I. C. Andrew and P. Muriel, "Real-time Markerless Tracking for Augmented Reality: The Virtual Visual Servoing Framework," IEEE Transactions on Visualization and Computer Graphics, Vol.12, issue.4, pp.615-628, 2006.
- [7] G. Simon, "Automatic Online Walls Detection for immediate Use in AR Tasks," IEEE International Symposium on Mixed and Augmented Reality, 2006.
- [8] A. P. Gee, C. Denis, M. Walterio, and C. Andrew, "Discovering Planes and Collapsing the State Space in Visual SLAM," British Machine Vision Conference, 2007.
- [9] C. Denis, G. Andrew, C. Andrew, and M. C. Walterio, "Ninja on a Plane: Automatic Discovery on Physical Planes for Augmented Reality using Visual SLAM," IEEE International Symposium on Mixed and Augmented Reality, 2007.
- [10] A. P. Gee and M. C. Walterio, "Real-Time Model-Based SLAM Using Line Segments," International Symposium on Visual Computing, 2006.
- [11] G. Klein and D. Murray, "Parallel Tracking and Mapping for Small AR Workspaces," In ISMAR'07, Nara, Japan, 2007(11).
- [12] R. O. Castle and D. W. Murray, "Object Recognition and Localization while Tracking and Mapping," Proceedings of the 8th IEEE/ACM International Symposium on Mixed and Augmented Reality, 2009.
- [13] G. Schall, H. Grabner, M. Grabner, P. Wohlhart, D. Schmlstieg, and H. Bischof, "3D Tracking in Unknown Environments Using On-Line Keypoint Learning for Mobile Augmented Reality," In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp.285-288, 2008.
- [14] H. P. William, P. F. Brian, A. T. Saul, and T. V. William, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, 1992.
- [15] P. Bouthemy, "A Maximum Likelihood Framework for Determining Moving Edges," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.11, No.5, pp.499-511, 1989.
- [16] N. Karlsson, E. D. Bernardo, J. Ostrowski, L. Goncalves, P. Pirjanian, and M. E. Munich, "The vSLAM Algorithm for Robust Localization and Mapping," Proceedings of IEEE International Conference on Robotics and Automation, 2005.
- [17] J. M. M. Montiel, J. Civera, and J. D. Andrew, "Unified Inverse Depth Parameterization for Monocular SLAM," Proceedings of Robotics: Science and Systems, 2006.
- [18] J. D. Andrew, D. R. Ian, D. M. Nicholas, and S. Olivier, "MonoSLAM: Real-Time Single Camera SLAM," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.29,

issue.6, pp.1052-1067, 2007.

[19] <http://www.doc.ic.ac.uk/~ajd/index.html>

[20] R. Edward and T. Drummond, "Machine Learning for High-Speed Corner Detection," In European Conference on Computer Vision, Vol.1, pp.430-443, 2006.

[21] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, Vol.60, pp.91-110, 2004.

저 자 소 개

이 서 진(Seo-Jin Lee)

준회원

- 2003년 2월 : 한국기술교육대학교 디자인공학과(공학사)
- 2007년 9월 ~ 현재 : 이화여자대학교 디지털미디어학부 미디어공학 석사과정

<관심분야> : HCI, 영상처리, Computer Graphics

남 양 희(Yang-Hee Nam)

정회원

- 1989년 : 이화여자대학교 전자계산학과 졸업(이학사)
- 1991년 : KAIST 전산학과 졸업(공학석사)
- 1997년 : KAIST 전산학과 졸업(공학박사)

- 1998년 ~ 1999년 : 스위스 로잔공대 박사후연구원
- 1999년 ~ 2000년 : 일본 ATR 연구소 초청연구원
- 2000년 ~ 2002년 : 세종대학교 컴퓨터공학과 전임강사
- 2002년 ~ 2006년 : 이화여자대학교 디지털미디어학부 조교수
- 2006년 ~ 현재 : 이화여자대학교 디지털미디어학부 부교수

<관심분야> : 가상/증강현실, 영상처리, HCI