

증강현실을 위한 실시간 마커리스 3차원 객체 추적

Realtime Markerless 3D Object Tracking for Augmented Reality

민재홍*, 이슬람 모하마드 카이들*, 폴 안잔 쿠마*, 백중환*

Jae-Hong Min*, Mohammad Khairul Islam*, Anjan Kumar Paul*, and Joong-Hwan Baek*

요 약

증강현실은 실세계의 정보와 가상의 정보를 연결시키기 위한 매개체가 요구되며, 이러한 매개체를 지속적으로 추적 인식하는 기술을 필요로 한다. 이러한 기술 중에 마커를 이용한 광학 트래킹이 주류를 이루고 있으나, 마커를 부착하는 과정이 불편하고 오래 걸리므로 최근에는 마커리스 트래킹 기법이 활발히 연구되고 있다. 본 논문은 2차원 평면 즉 동일평면상의 특징점들을 트래킹하는 방법이 아닌 3차원 객체에 대한 특징점을 추출하여 실시간으로 트래킹하는 방법을 제안한다. SURF(Speed Up Robust Features)를 이용하여 특징점을 추출하고 이를 POSIT(Pose Object System for Iteration) 알고리즘으로 3차원 객체의 회전과 이동정보를 얻어 실시간으로 객체를 추적한다. 추적 실패시 실시간으로 재추적이 가능하도록 빠른 특징점 추출과 매칭을 통하여 트래킹에 적합한 특징점을 선택하여 객체의 위치와 회전 정보를 얻어 객체를 실시간으로 추적 및 재표현 하였다.

Abstract

AR(Augmented Reality) needs medium between real and virtual world, and tracking and recognition techniques are necessary to track an object continuously. Optical tracking using marker is mainly used, but it takes time and is inconvenient to attach markers onto the target objects. Therefore, many researchers try to develop markerless tracking techniques nowadays. In this paper, we extract features and 3D position from 3D objects and suggest realtime tracking based on these features and positions, which do not use just coplanar features and 2D position. We extract features using SURF, get rotation matrix and translation vector of 3D object using POSIT with these features and track the object in real time. If the extracted features are not enough and it fails to track the object, then new features are extracted and re-matched to recover the tracking. Also, we get rotation matrix and translation vector of 3D object using POSIT and track the object in real time.

Key words : SURF, POSIT, Markerless Object Tracking, Augmented Reality

I. 서 론

가상현실은 게임처럼 공간을 새롭게 창조하지만 증강현실은 실사 현실의 기반 위에 가상의 사물을 결합하여 부가적인 정보들을 제공할 수 있다. 이러한 장점 때문에 가상현실과 달리 현실에 다양한 용

용이 가능하다. 증강현실을 구현하기 위해서는 카메라 및 센서, 컴퓨터 그래픽, 정합(Registration)기술, 디스플레이 기기 등 요소기술이 필요하다. 이러한 요소기술들은 주변 정보 및 사용자의 시선, 자세, 동작 등을 파악하고, 가상정보를 생성하며, 가상과 현실을 일치시켜 사용자에게 전달한다.

* 한국항공대학교 정보통신공학과 (Department of Information and Telecommunication Eng., Korea Aerospace University)

· 제1저자 (First Author) : 민재홍

· 투고일자 : 2010년 3월 29일

· 심사(수경)일자 : 2010년 3월 30일 (수경일자 : 2010년 4월 20일)

· 게재일자 : 2010년 4월 30일

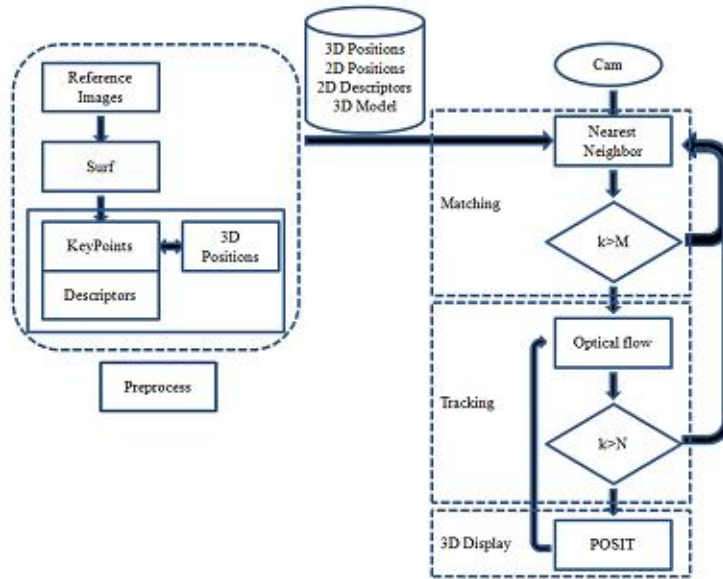


그림 1. 제안된 시스템 블록도
 Fig. 1. Block diagram of the proposed system.

실제 영상에 가상물체를 생성하기 위해 가상물체와 실제 영상 간의 매개체가 요구되며, 이러한 매개체를 지속적으로 추적 인식하여 기준이 되는 좌표계를 설정한다. 추적 및 인식을 위해 기계, 자기, 광학 등이 이용되지만 현재 광학이 가장 높은 정밀도를 보인다. 따라서 마커를 이용한 광학 트래킹이 주류를 이루고 있다. 광학 기반의 트래킹 방법은 조명에 영향을 많이 받게 되므로 이를 극복하기 위해서 IR(infrared) 타깃을 사용하는 방법이 사용되기도 한다[1].

최근에는 번거롭게 마커를 필요로 하지 않고도 트래킹이 가능한 마커리스 트래킹 기법이 연구되고 있다. 이는 인공적으로 프린트한 인쇄물로서의 마커가 아니라 현실세계 속에서의 객체들의 특징점을 추출해 이를 트래킹 한다. 이때 추출되는 특징점이 동일한 평면에 있는 것이 아니라 3차원 공간상의 객체에 대해 대응해야 한다. 대표적인 예로 PTAM이 있으며 작은 AR 공간을 생성하여 추적 및 맵핑을 한다 [2].

일반적으로 객체의 한면 즉 동일 평면에서 특징점을 추출하여 Homography를 이용하여 객체 면의 회전과 이동을 알아낸다. 따라서 객체는 일정한 평면을 가져야 한다. 이러한 문제점을 개선하기 위해 본 논문

은 동일평면상이 아닌 3차원 객체의 특징점을 이용하여 객체의 이동과 회전을 구한다. SIFT(Scale Invariant Feature Transform)를 이용하여 특징점을 추출하면 128의 표현자를 가져 좋은 성능을 보인다. 그러나 사용자가 불편함을 느끼지 않기 위해서는 가상정보와 현실정보를 실시간으로 정확하게 일치시키는 것이 가장 중요하다. 따라서 SIFT의 특징점 추출과정에 연산량이 많아 실시간에 적합하지 않으므로 본 논문에서는 SURF(Speed up Robust Features)을 이용하여 특징점을 추출하였다. 본 논문은 3차원 객체의 특징점을 POSIT(Pose from Orthography and Scaling with Iteration)을 통해 객체의 회전과 이동을 구하며 실시간으로 객체를 표현하고 트래킹하는 방법을 제안하고자 한다[3]-[5]. 그림 1에 제안된 시스템의 개념도를 나타낸다.

2장에서는 3차원 실시간 트래킹 시스템에 대한 개념을 설명하고, 3장에서는 특징점 추출에 사용된 SURF의 구성 요소와 개념을 나타낸다. 4장에서는 추출된 특징점을 실시간으로 추적하는 방법에 사용한 루카스-카나데 알고리즘을 설명한다. 5장은 POSIT의 기본 개념에 대해 설명한다. 6장에서는 트래킹과 게트래킹 등의 실험 결과를 보이며, 7장에서는 결론을 내린다.

II. 3차원 실시간 트래킹 시스템 개요

3차원 실시간 트래킹 시스템은 전처리과정과 실시간처리과정으로 나뉘지며, 그림 1에서 보면 초기화를 위한 전처리과정에서는 SURF를 이용하여 객체의 특징점들 위치와 표현자를 추출한다. 이 2D 위치와 이미 알고 있는 객체의 3D 위치 정보를 정합한다. 정합된 정보, 표현자(descriptor), 객체의 3D 모델을 카메라에서 들어오는 입력영상과 매칭시켜 트래킹에 필요한 특징점들을 추출한다. 추출된 특징점들이 일정 수 이하가 되면 다시 매칭을 하며, 그 이상이 되면 3D 위치정보를 POSIT에 적용하여 객체의 회전과 이동 정보를 얻어내고 3D 모델을 화면에 표현한다.

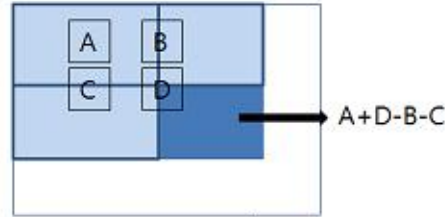


그림 2. 적분영상
Fig. 2. Integral Image.

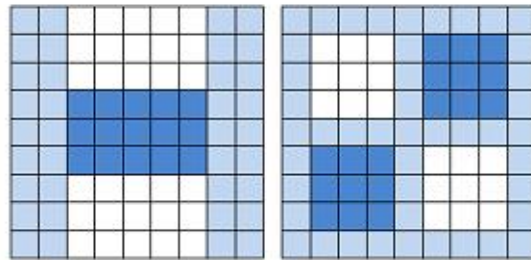


그림 3. 근사화된 가우시안 사각필터
Fig. 3. Approximated Gaussian box filters.

III. SURF(Speed Up Robust Features)

SURF는 특징점 추출 과정과 표현자 생성 과정으로 이루어지며, 특징점 추출 과정은 영상에서 강건한 특징점 위치를 찾는 것이며, 두 번째 과정은 환경 변화에 불변하고 강건한 특징점을 설명하는 표현자를 생성한다. 그림 2에 적분영상의 개념을 보인다.

3-1 특징점 추출

적분영상과 사각필터를 이용한 근사화된 가우시안 2차 미분 9X9 필터를 이용한 헤이시안 검출기로 특징점을 추출한다.

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (1)$$

식 (1)에서 $L_{xx}(x, \sigma)$ 은 점 x에서 x방향으로 영상 I의 2차 가우시안 미분 컨벌루션이고, $L_{xy}(x, \sigma)$ 과 $L_{yy}(x, \sigma)$ 는 xy방향으로의 미분 그리고 y방향으로의 2차 가우시안 미분 컨벌루션을 의미한다. 그림 3이 본 논문에서 사용된 근사화 된 가우시안 사각필터이다.

3-2 표현자(descriptor) 생성

특징벡터를 표현하기 위해 Haar 웨이블릿 특징을 사용하는데 정규화된 부분영상을 이용하여 4개의 특징벡터를 만든다. 그림 4는 웨이블릿 특징 조합을 이용하여 지역적 특징 벡터를 만드는 과정을 보여준다.

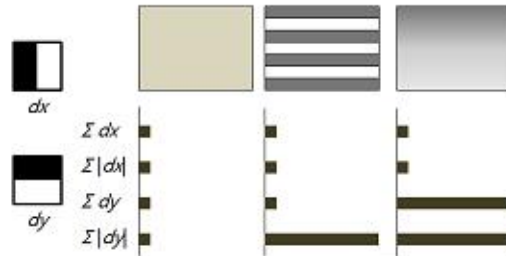


그림 4. Haar 웨이블릿 특징을 가지는 지역적 특징
Fig. 4. Local features of Haar wavelet.

IV. Pyramid Lucas-Kanade Algorithm

루카스-카나데 알고리즘은 다음의 이미지 밝기 제약 방정식을 가진다.

$$\frac{\partial I}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial I}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial I}{\partial t} = 0 \quad (2)$$

$$(\nabla I)^T \mathcal{T}_v + I_t = 0 \quad (3)$$

식 (3)에서 ∇I 는 공간 밝기 변화율이고, I_t 는 시간 밝기 변화율이며, 위 식의 해를 구하기 위해 국부 평탄화를 가정하여 구한다. 그러나 루카스-카나데 알고리즘은 작은 윈도우를 사용하기 때문에 큰 움직임이 있을 경우 움직임을 계산하지 못한다. 따라서 이를 해결하기 위해서 일단 원본 영상에서 영상을 피라미드로 구성하고 상위부터 하위로 추적하여 큰 움직임을 찾아낼 수 있다.

루카스-카나데 알고리즘은 밝기 항상성, 시간 지속성, 공간 일관성의 3가지 가정을 기초로 한다. 밝기 항상성은 추적할 영역의 내부는 시간이 지나도 그 값이 일정함을 나타내며, 시간지속성은 프레임간의 움직임이 작아서 그 변화가 매우 적다는 것이다. 또한 공간 일관성은 공간적으로 인접한 점들은 동일한 객체일 가능성이 많다는 것을 나타낸다[6].

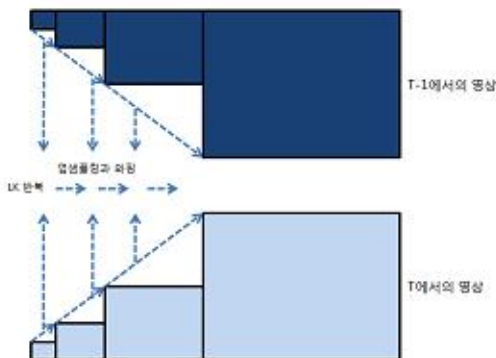


그림 5. 피라미드 옵티칼 플로우 추정
Fig. 5. Estimation of pyramid optical flow.

그림 5에서 보면 루카스-카나데 옵티칼 플로우와 엡샘플링과 외평을 반복하여 움직임을 추정한다. 피라미드 루카스-카나데 옵티칼 플로우는 최상위 계층부터 옵티컬 플로우를 계산하여 시간 지속성과 공간일관성의 가정들로부터 나오는 작고 일관된 움직임이 야기하는 문제점들을 해결한다. 이전 계층의 움직임 정보는 다음 계층의 움직임 추정시 시작점으

로 사용한다. 이러한 연산을 최하위 계층까지 반복 수행하여 빠르고 큰 움직임에 강건하게 추적할 수 있다.

V. POSIT(Pose from Orthography and Scaling with Iteration)

그림 6은 카메라에 따른 객체의 회전 및 이동과 객체의 2차원 평면 투영을 나타낸다. 객체의 회전과 이동정보 즉 객체의 포즈에 따라 2차원 평면에 투영되는 객체의 윤곽 및 특징점들의 위치가 다름을 알 수 있다. 본 논문에서는 3차원 객체의 포즈를 알아내는 방법으로 Posit을 사용하였으며 Pose는 위치값 3개와 회전값 3개로 표현할 수 있다 [6].

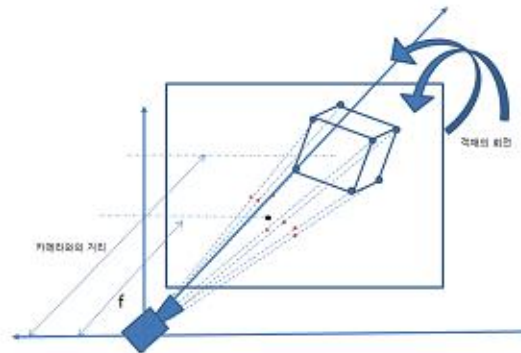


그림 6. POSIT 알고리즘의 회전과 이동의 추정에 대한 객체와 카메라의 상관도
Fig. 6. Relationship between object and camera for estimating rotation and translation in the POSIT algorithm.

객체의 포즈를 알기 위해서는 입력영상에서 최소한 4개의 점들과 그에 대응하는 3D 점들을 알고 있어야 한다. POS(Pose from Orthography and Scaling) 방법은 객체의 점들이 카메라로부터 충분히 멀리 있어서 실제 점들 사이의 깊이를 무시할 수 있다는 것을 가정하며 상수로 사용한다. 따라서 구해진 회전행렬과 이동벡터는 내부 파라미터 행렬로부터 구할 수 있다. 또한, 크기 변화는 오로지 객체와 카메라 사이의 거리에 따라 이루어지며, 객체의 3D pose에 대한 폐쇄형 해(close-d-form solution)가 존재하게 된다. 이렇게 해서 얻어진 새로운 점들을 POS의 입력점들로

다시 반복하여 수행하면 그 점들이 수렴한다.

V. 실험 및 결과

본 실험에서는 Intel Core i5 CPU 750 2.67 GHz 프로세서와 램 4 GB를 가진 시스템을 이용하였으며, 3차원 객체로는 높이 26 cm 가로 20 cm 세로 18 cm의 모형집을 사용하였다. 정면 좌우 측면의 영상들을 참조영상으로 삼아 SURF를 통하여 특징점들을 추출한다. 이로부터 표현자와 이에 대응하는 3차원 위치정보를 초기 정보로 생성한다. 피라미드 유틸칼 플로우를 통해 추적된 객체의 특징점들을 POSIT을 통해 회전과 이동정보를 얻는다. 이를 객체의 외곽 윤곽에 대해 회전과 이동을 적용하여 와이어프레임으로 표현한다.

표 1에서는 추적링 처리 시간을 보인다. 초기화에 51.66 ms의 시간이 소요되며, 이는 참조영상들의 특징점들로부터 적절한 점을 선택하기 위한 처리시간이며, 추적링은 평균 6.22 ms이 걸렸다. POSIT을 포함한 전체 추적링 시간은 58.22 ms 이다.

표 1. 추적링 처리 시간
Table 1. Processing time for tracking

Initializing	51.66 ms
Tracking	6.22 ms
POSIT	0.24 ms
Extracting Features & Selecting Features	0.10 ms
Total	58.22 ms

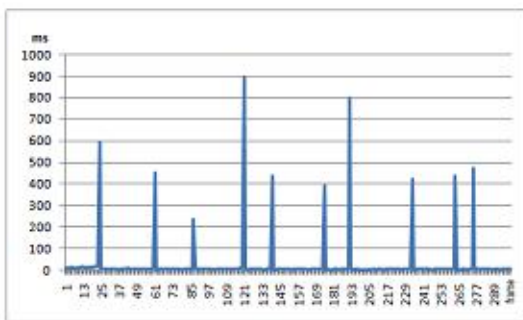


그림 7. 프레임별 추적링 시간
Fig. 7. Tracking times per each frames

그림 7은 프레임별 추적링 시간을 나타낸다. 재매칭 실험을 위해 빈번히 카메라를 객체로부터 벗어나게 하였으며, 재매칭에 평균 465 ms의 시간이 필요하였다. 추적링은 재매칭 시간에 비해 극히 작은 시간이 소요되었다. 추적링 실행시 참조영상과 다시 매칭을 통해 추적링에 적합한 특징점들을 찾을 때까지의 시간을 보여주며, 매칭되는 특징점의 수가 5개 이상일 경우에 추적을 시작하였다. 추적링에 필요한 적절한 특징점들의 개수들을 설정함에 따라 느려지고 빨라질 수 있으나 정확한 매칭에 중점을 두었으며, 매칭되는 각각의 특징들 상대적인 위치변위가 특정 범위를 벗어나지 않게 지정하여 정확도를 높였다.

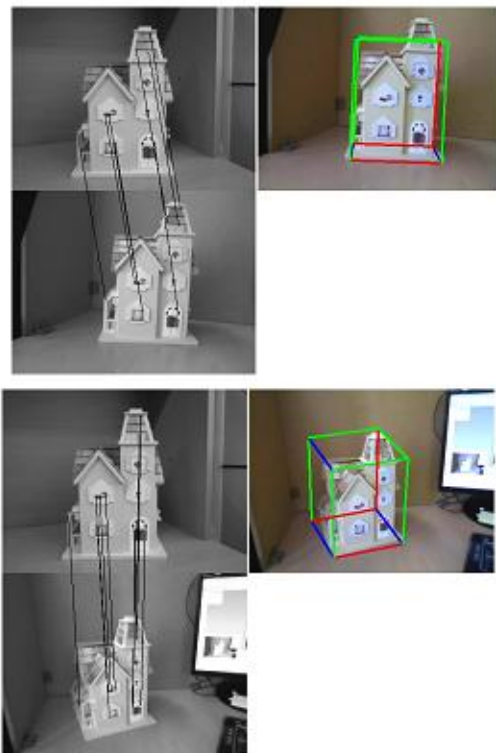


그림 8. 결과 영상들
Fig. 8. Result images

그림 8의 왼쪽 그림은 참조영상과 현재 영상간의 특징점 매칭을 보여주며, 오른쪽 그림은 회전과 이동 정보를 이용하여 와이어 프레임으로 표현하였으며, 붉은 점은 매칭된 특징점들을 나타낸다. 또한 객체의 회전과 이동에 대해서 매칭된 특징점들이 정확하게 추적되고, 그 결과를 와이어 프레임으로 실제 객체와 일치함을 볼 수 있다.

VI. 결 론

본 논문에서는 마커리스 증강현실 시스템을 위한 물체 추적 알고리즘을 제안하였다. SURF를 이용하여 특징점을 추출하고 이를 POSIT 알고리즘으로 3차원 객체의 회전과 이동정보를 얻고 이를 실시간으로 추적하였다. 추적 실패시 실시간으로 재추적이 가능하도록 빠른 특징점 추출과 매칭을 통하여 트래킹에 적합한 특징점을 선택하여 객체의 위치와 회전 정보를 얻을 수 있었고 실험을 통하여 실시간 트래킹이 가능함을 확인할 수 있었다.

감사의 글

본 연구는 경기도지역협력연구센터 (GRRRC) 프로그램에 의해 한국항공대학교 차세대방송미디어기술 연구센터의 지원으로 수행되었으며, 일부 KAGERIC의 지원에 의함.

참 고 문 헌

[1] 김용준, 이수용, 이준석, 노경희, "혼합현실 기반 이터닝 기술 동향", 전자통신동향분석 제24권 제1호, Feb. 2009

[2] G. Klein and D. Murray, "Parallel Tracking and Mapping for Small AR Workspaces", 6th IBBS and ACM International Symposium on In Mixed and Augmented Reality, pp. 225-234, Jun. 2008.

[3] H. Bay, A. Ess, T. Tuytelaars, and L. Gool, "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346-359, Sept. 2008.

[4] D. Dementhon and L. Davis, "Model-based object pose in 25 lines of code", International Journal of Computer Vision, Vol. 15 No. 1-2, pp. 123-141, Jun. 1995.

[5] G. Parin and A. Knoll, "Fully Automatic Real-Time 3D Object Tracking using Active Contour and Appearance Models", Journal of Multimedia, Vol. 1, No. 7, Nov/Dec. 2006.

[6] B. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," In International Joint Conference on

Artificial Intelligence, pp. 674-679, 1981.

민 재 홍 (閔 載 泓)



1995년 2월 : 한국항공대학교 항공통신 정보공학과(공학사)
 1998년 2월 : 한국항공대학교 항공통신 정보공학과(공학석사)
 2010년 4월 : 한국항공대학교 항공통신 정보공학과 박사과정
 관심분야 : 영상처리, Augmented Reality, Stereo Vision

Mohammad Khairul Islam



2000년 7월 : Shahjalal University of Science & Technology, Bangladesh (BS)
 2007년 8월 : 한국항공대학교 정보통신 공학과 (공학석사)
 2007년 9월 ~ 현재 : 한국항공대학교 정보통신공학과 박사과정
 관심분야 : 멀티미디어, 영상처리,

컴퓨터비전

Anjan Kumar Paul



2000년 1월 : Khulna University, Bangladesh (BE)
 2006년 7월 : Indian Institute of Science, India (M.Tech)
 2006년 9월 ~ 현재 : 한국항공대학교 정보통신공학과 박사과정
 관심분야 : Augmented Reality, 멀티미디어,

컴퓨터비전

백 중 환 (白 重 煥)



1981년 2월 : 한국항공대학교 항공통신 정보공학과(공학사)
 1987년 7월 : (미)오클라호마주립대학교 전기 및 컴퓨터공학과(공학석사)
 1991년 7월 : (미)오클라호마주립대학교 전기 및 컴퓨터공학과(공학박사)
 1992년 3월 ~ 현재 : 한국항공대학교

항공전자 및 정보통신 공학부 교수
 관심분야 : 영상처리, 패턴인식, 멀티미디어