

## 손동작을 이용한 가상 물체 증강

김일목, 정경부, 최병욱  
 한양대학교 전자컴퓨터통신공학과  
 e-mail : {ilmoek, kbjung, buchoi}@hanyang.ac.kr

### Markerless Augmentation of Virtual Object Using Bare-Hand

Il-Moek Kim, Kyungboo Jung, Byung-Uk Choi  
 Department of Electronics and Computer Engineering Hanyang University

#### 요 약

본 논문은 카메라 영상을 통해 사용자의 손동작을 인식하여 가상의 물체를 증강시키는 인터페이스를 제안한다. 사용자는 영상의 일부를 원형으로 그려주는 특정한 손동작을 취하여 영역을 선택하고 시스템은 이를 인식하여 물체를 증강 시킨다. 손동작을 인식하기 위하여 먼저 손 외곽선을 찾아낸 후, 찾아낸 외곽선의 곡률을 계산하여 손가락의 위치를 알아낸다. 알아낸 손가락의 상대적인 위치와 개수를 이용하여 손동작을 구분한다. 또한 적은 연산량으로도 안정적으로 물체를 증강 시킬 수 있도록 이전 프레임에서 자세 추정에 사용된 특징점들을 이용하여 현재 프레임에서 필요한 인라이어를 찾아 낼 수 있는 방법을 제시한다.

#### 1. 서론

휴대가 간편한 HMD뿐만 아니라 스마트 폰을 이용한 다양한 증강현실 응용이 개발되고 있다. 그와 함께 사용자가 시스템과 상호 작용할 수 있는 많은 방법들이 연구되어지고 있다[1,2]. 사용자와 시스템이 상호 작용할 수 있는 방법은 크게 특별한 인터페이스 장치를 사용하는 방법과 사용자의 음성이나 손동작 등을 사용하는 방법으로 나눌 수 있다. 후자의 경우 번거로운 추가 장치 없이 일상에서 사용하는 행동들을 이용할 수 있어 가상의 물체를 현실과의 괴리 없이 제어할 수 있다. 또한 복잡한 사전 훈련 없이 사용이 가능하기 때문에 사용자가 인터페이스를 이용하는데 편리하다.

손을 인식하여 사용자의 입력을 받는 인터페이스 연구는 여러 가지 형태로 진행되고 있다. [3]은 배경색으로부터 전경색인 손 색상을 분리하고 손가락을 검출 하여 사용자가 현재 비디오 프레임에 메모를 할 수 있는 시스템을 제안한다. 또한 의학에서의 응용으로 손동작을 인식하여 의학 자동화기기에 명령을 내릴 수 있는 시스템도 다양하게 연구되고 있다[4].

본 논문은 위와 같은 응용 연구 분야의 일환으로 사용자의 손(bare-hand) 동작을 인식하여 현실공간의 특정지역에 가상의 물체를 증강시킬 수 있는 인터페이스를 제안한다. 이 인터페이스는 손가락으로 영상의 특정 지역에 사각형을 그려주었을 때 그 위에 물체가 증강 될 수 있도록 한다. 이 인터페이스는 물체 증강을 위한 저작 도구(authoring tool)와 같이 다양한 응용이 가능하다.

또한 물체 증강을 위해 이전 프레임에서 사용되었던 인라이어(inlier) 집합을 이용해 빠르고 안정적인 삼차원 자세추정 방법을 제시한다.

#### 2. 본론

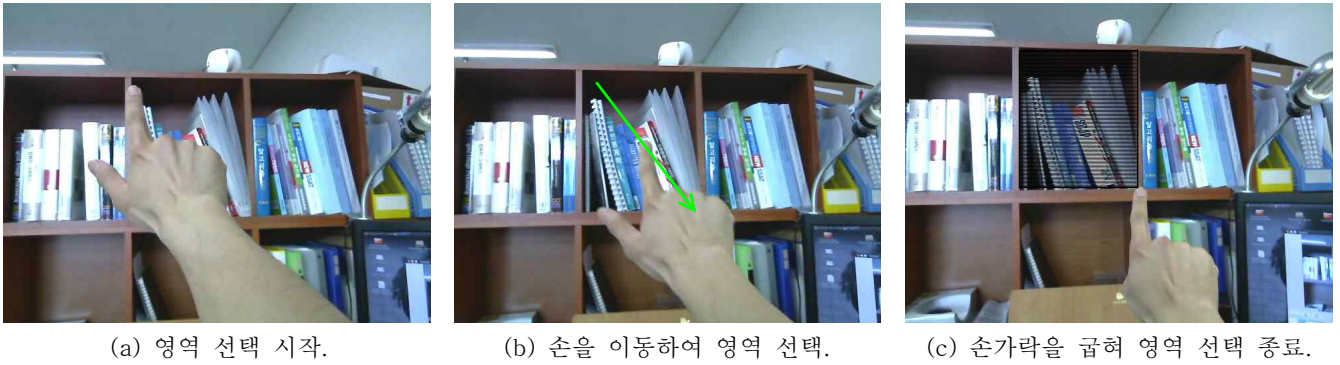
##### 2.1. 사용자 지정 영역 검출

손은 주어진 색상 정보를 기준으로 각 픽셀의 유사성 판단을 수행하여 이진화 한다. 먼저 다양한 환경에서 인터페이스를 사용할 수 있도록 간단한 손동작을 통해 색상을 추출 할 수 있는 방법을 이용하여 손 색상을 지정한다[5]. 이진화 된 영상에서 노이즈 제거, 외곽선 검출과 같은 후처리과정을 거친 후 검출된 손의 동작을 인식하기 위하여 손가락 검출을 수행한다. 검출된 손에서 손가락을 추출하기 위하여 세 점 사이의 곡률을 계산하는 식 (1)을 이용한다.

$$K_l(P) = \frac{1}{2} \left[ \frac{1 + \overrightarrow{P_1P} \cdot \overrightarrow{PP_2}}{\|\overrightarrow{P_1P}\| \|\overrightarrow{PP_2}\|} \right] \quad (1)$$

여기서  $K_l$ 는 점  $P$ 에서의 곡률을 계산하는 함수이고,  $P_1$ 과  $P_2$ 는 각각 점  $P$ 의 이전 점, 이후 점이다. 손가락은 계산된 곡률값이 높은 지역에서 검출 된다[6].

손가락 검출 결과를 이용하여 손동작을 인식한다. 본 논문에서 사용될 손동작은 가상 물체가 증강될 영역지정 시작 동작과 종료 동작 두 가지로 제한한다. 그림 1과 그림 2는 각각 시작과 종료 동작을 보여준다. 사용자가 그림 1의 동작과 같이 두 개의 손가락을 화면에 보이게 되면 영역 선택이 시작된다. 이후 손을 우측 하단으로 움직여 사각형을 만들고 영역 지정이 다 되었을 때, 그림 2와 같이 하나의 손가락만을 보여 종료되었음을 알릴 수 있다. 그림 2의 회색 사각형은 물체가 증강될 위치를 나타낸다. 물체 증강은 회색 사각형을 중심으로 상대적



(a) 영역 선택 시작.

(b) 손을 이동하여 영역 선택.

(c) 손가락을 굽혀 영역 선택 종료.

그림 1. 손가락을 이용하여 물체 증강 영역을 선택. 선택할 사각형 영역의 시작 부분에서 두 개의 손가락을 펼쳐 보여주며 손을 이동하여 사각형의 우측 하단 지점으로 이동. 이후 손가락을 굽혀 영역 선택을 종료.

인 위치에서 이루어진다. 사용자의 영역 지정이 완료된 이후에는 자세추정과 물체 증강 처리 과정으로 넘어간다.

### 2.2. 카메라 자세추정 (Camera Pose Estimation)

먼저 선택된 영역에서 특징점들을 추출하여 DB에 저장한다. 저장된 특징점들은 원 중심 부분을 기준으로 한 상대적인 위치정보와 함께 현재 프레임에서 인라이어로 선택되었는지를 나타내는 정보를 갖게 된다. 저장된 특징점 위치 정보와 현재 프레임에서 검출된 특징점들을 이용하여 삼차원 자세 추정을 수행한다[7].

자세 추정을 위해서는 영상 내에 특징이 두드러지게 나타나 밝기 변화나 크기의 변화가 있더라도 유지 될 수 있는 점이 필요하다. 이와 관련되어 최근 많은 연구가 진행되고 있으며 본 논문에서는 빠른 연산시간을 갖으면서도 높은 정합률을 갖는 SURF 특징점을 이용하였다 [8]. 하지만 특징점 정합(matching)이 정확히 되지 않는 경우가 있기 때문에 모든 정합점을 자세추정에 이용할 수 없다. 만약 모든 정합점을 이용하여 자세 추정을 할 경우 자세 추정이 올바르게 되지 않거나 증강 결과가 심하게 흔들리게 된다. 추정 결과를 보정하기 위하여 RANSAC 알고리즘을 사용한다[9]. 먼저 자세추정에 필요한 최소 개수의 정합점들을 임의로 선택하고 자세 추정을 수행한다. 추정결과를 이용하여 나머지 정합점들의 재투영 오차 (Reprojection Error)값들을 계산한다. 모든 정합점들중에 오차값이 충분히 작은 정합 집합인 인라이어 집합을 이용하여 다시 자세를 추정하고 재투영 오차의 총합을 계산한다. 위의 과정을 일정 횟수동안 반복해서 수행하면서 인라이어 집합들 중 인라이어의 개수가 원래 정합점들 개수의 특정 비율 이상이 되지 않은 경우를 제외하고 재투영 오차의 총합이 가장 작은 집합을 선택한다. RANSAC을 사용하게 되면 반복 횟수에 따라 거의 올바른 위치에 물체를 증강시킬 수 있다. 증강 오류는 반복 횟수를 증가시키면 줄일 수 있다. 하지만 반복 횟수를 증가시키게 되면, 많은 연산 시간이 들게 되어

실시간으로 물체를 증강시키기 힘들게 된다. 이러한 문제를 해결하기 위하여 이전 프레임의 인라이어를 이용하여 자세 추정을 빠르게 할 수 있도록 한다. 그림 2는 인라이어 추적을 사용해 안정적인 자세 추정을 하기 위한 단계를 간단히 표현하였다.

자세 추정의 첫 단계로 카메라 영상의 특징점을 추출하여 DB의 특징점과 정합 시킨 후 RANSAC을 이용하여 인라이어를 찾아낸후 DB에 저장한다.

다음 반복 과정 부터는 RANSAC를 이용하지 않고 이전 프레임에서의 인라이어 집합의 모든 정합점을 현재 프레임에서의 특징점에서 추적하여 자세를 추정한다. 이전 프레임에서의 인라이어들과 현재 프레임의 특징점들을 정합시키는 방법으로 인라이어 집합의 정합점들을 추적할 수 있다. 그림 3의 직선은 두 프레임 사이의 정합된 결과를 보여주고 있다. 특징점들 중 정합되지 않는 점들을 아웃라이어로 분류하고 DB의 인라이어 목록에서 제거한다. 그림 3의 좌측 영상에 직선으로 연결되지 않은 점들이 정합에 실패한 점이다. 추적된 정합점들은 이

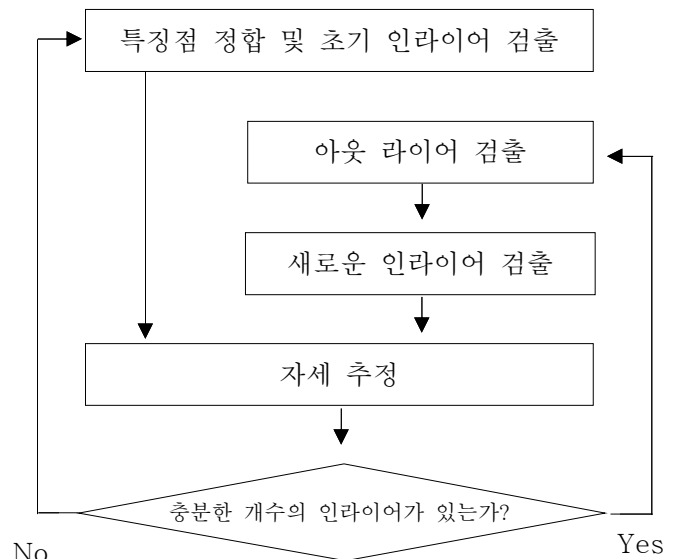


그림 2. 삼차원 자세 추정.

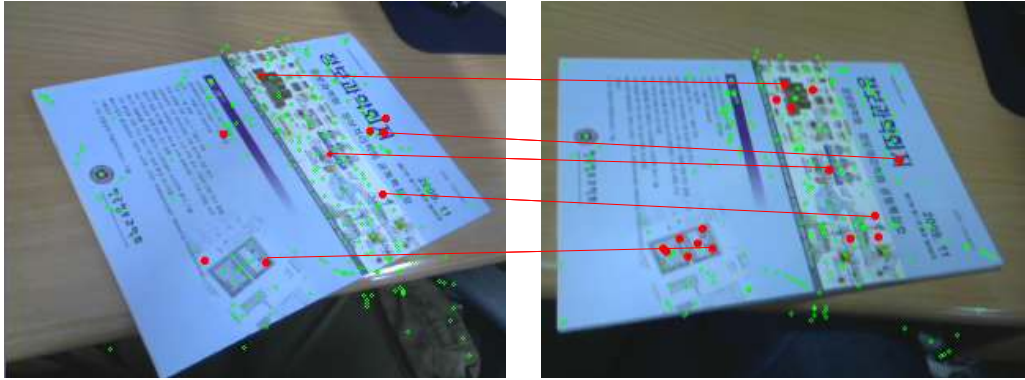


그림 3. 이전 프레임의 인라이어와 현재 프레임의 특징점 사이의 대응점을 이용하여 계산된 호모그래피를 사용하여 새로운 인라이어 검출. (좌) 이전 프레임. (우) 현재 프레임.

전 프레임에서 자세추정에 사용되었던 정합점들의 부분 집합이기 때문에 현재 프레임에서도 올바른 자세 추정을 할 수 있는 데이터가 된다.

현재 프레임에서 쉽게 찾아낼 수 있는 모든 정합점들을 찾아내서 인라이어로 포함시키면 최대한 최초 데이터를 계산할 시간을 단축시킬 수 있다. DB에 저장된 특징점 정보와 현재 프레임 사이의 호모그래피를 계산하여 자세추정을 원하는 영상 모든 특징점을 변환한다. 위치가 변경된 특징점들을 현재 프레임의 같은 위치에서 발견하게 되면 이를 새로운 인라이어로 사용한다. 현재 프레임의 해당 위치에 해당 특징점의 존재 여부를 판단하기 위해서 해쉬 테이블(hash table)을 이용하였다. 그 결과 모든 특징점들을 정합해볼 필요 없기 때문에 짧은 연산시간에도 인라이어 집합을 재구성 할 수 있었다. 그림 3은 호모그래피를 이용하여 인라이어 집합을 재구성한 결과이다. 우측 영상의 좌측 영상으로 연결되지 않은 점들이 새로 찾아낸 인라이어들이다. 과정이 반복 되면서 불필요한 인라이어들이 제거되고 필요한 점들만이 남게 된다.

찾아낸 인라이어 집합을 이용하여 삼차원 자세를 추정하게 되면 한 번의 추정 과정만으로도 안정적인 결과를 얻을 수 있다. 자세 추정이 완료 되면 가상의 물체를 증강 시킨다.

### 3. 실험 및 결과

손을 이용하여 가상 물체를 증강 시키는 세 가지 실험을 진행하였다. 먼저 두 가지의 다른 환경에서 손을 이용하여 특정 지역을 선택하고 물체를 증강시켜 보았다. 또한 자세 추정 속도를 비교하기 위한 실험을 진행하였다.

#### 3.1. 손을 이용한 가상 물체 증강 영역 설정

그림 4는 사용자가 손을 이용하여 책을 증강 영역으로 지정하고 그 위에 가상 물체가 증강되는 것을 보여준다.

#### 3.2. 자세추정 비교 실험

실험을 통하여 자세 추정의 안정화 정도와 속도를 비교하였다. 결과 비교를 위한 영상은 처음부터 자세 추정을 시작하여 천천히 증강 영역이 보이지 않을 때까지 움직이도록 하였다. 그림 6은 제안하는 방법을 이용하여 가상 물체를 증강시켰을 때와 RANSAC을 이용하여 모

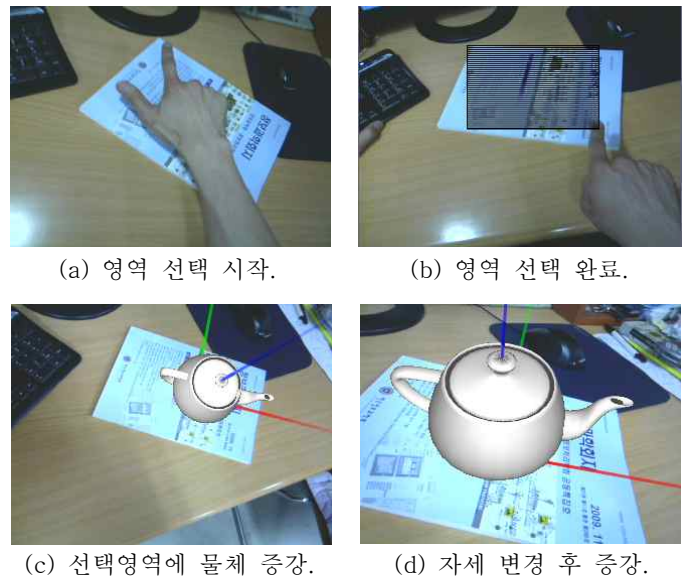


그림 4. 손을 이용하여 책을 물체 증강 영역으로 지정하고 그 위에 가상 물체 증강.



(a) 책장에 물체 증강 (b) 다른 자세로 물체 증강

그림 5. 책장에 가상 물체를 증강.

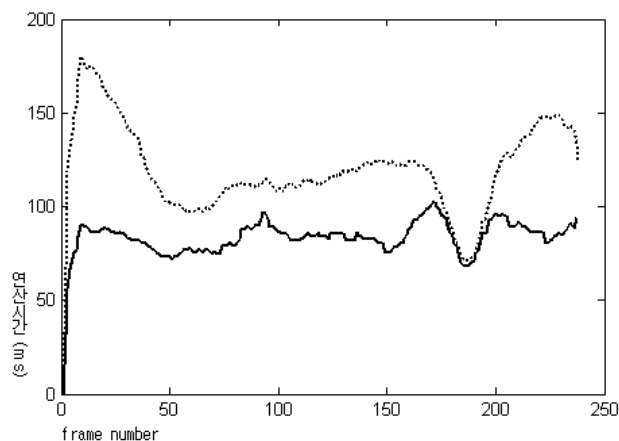


그림 6. (점선) 모든 프레임에서 자세추정을 하였을 때, (실선) 제안하는 방법을 이용하여 자세 추정 횟수를 줄인 결과.

든 프레임에서 자세를 추정했을 때 연산 시간의 차이를 그래프로 보여준다. 약 5프레임에서 최초 자세 추정이 완료되며, 이후 본 논문의 방법을 이용하면 추가적인 자세 추정이 필요 없으므로 연산 시간이 80ms를 유지하게 된다. 하지만 RANSAC만을 사용하게 되면 모든 프레임에서 새로이 자세 추정을 해야 하기 때문에 그래프에 나타난것과 같이 많은 연산 시간이 필요 했다. 이후 영상의 상태에 따라 인라이어를 잃게 되어 100프레임에서 다시 자세 추정을 하게 되고, 약 170프레임에서 증강시킬 영역이 영상에서 완전히 사라졌다. 이후 다시 자세 추정을 하게 되고 인라이어들을 추적하게 된다.

이 때 의 차이가 두드러지게 나타나는 것을 확인 하였다. 본 실험에서 SURF 특징점 추출과 정합을 하는데 걸린 시간은 약 50ms이다.

#### 4. 결론

본 논문은 손동작을 이용하여 영상의 특정 영역을 지정하는 방법을 제안하였고 지정된 영역의 자세추정을 수행하여 가상 물체를 증강시켰다. 사용자와의 상호작용에 의해 가상의 물체를 증강시킬 수 있고 언제 어디서나 사용자가 원하는 영역에 가상 물체를 증강시킬 수 있게 되었다. 또한 적은 연산량으로도 안정적인 실시간 자세 추정을 하기 위하여 이전 프레임의 인라이어를 이용하여 현재 프레임의 인라이어를 찾아내는 방법을 제시하였으며 실험을 통해 개선된 성능을 입증하였다.

#### 5. 참고문헌

[1] Mark Billinghurst and Hirokazu Kato, "Collaborative Augmented Reality," Communications Of The ACM, volume 45, no.7, pp. 64-70, July 2002.

- [2] Gun A. Lee, Gerard J.Kim, Mark Billinghurst. "Immersive authoring: What You eXperience Is What You Get (WYXIWYG)," Communications of the ACM, Volume 48, Issue 7, pp. 76-81, July 2005.
- [3] Taejin Ha and Woontack Woo, "Bare Hand Interface for Interaction in the Video See-Through HMD Based Wearable AR Environment," Entertainment Computing - ICEC 2006, Volume 4161, pp. 354-357, 2006.
- [4] J. Wachs, H. Stern, Y. Edan, M. Gillam, C. Feied, M. Smith and J. Handler., "A Real-Time Hand Gesture System based on Evolutionary Search," Applications of Soft Computing, Volume 36, pp. 153-162, 2006.
- [5] 김일목, 정경부, 정승도, 최병욱. "손 영역 검출을 기반으로 한 다양한 환경에서 손 색상 추출", 컴퓨터 소사이어티(멀티미디어2), 전자공학회 2009 하계학술대회, 2009.
- [6] Antonis A. Argyros and Manolis I.A. Lourakis, "Vision-Based Interpretation of Hand Gestures for Remote Control of a Computer Mouse," Computer Vision in Human-Computer Interaction, pp. 40-51, 2006.
- [7] Z. Zhang. "A flexible new technique for camera calibration," IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11): 1330-1334, November 2000.
- [8] Herbert Bay, Andres Ess, Tinne Tuytelaars, Luc Van Gool, "Speeded-Up Robust Feature", Computer Vision and Image Understanding, " volume 110, Issue 3, pp. 346-359, 2008.
- [9] Martin A. Fischler and Robert C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," Readings in computer vision: issues, problems, principles, and paradigms, pp. 726-740, 1987.