

# 가상 현실 게임 환경에서의 가상 손 제어를 위한 사용자 손 인식 방법\*

김부년, 김종호, 김태영  
서경대학교 컴퓨터공학과

initialkbn@gmail.com, sebhi@naver.com, tykim@skuniv.ac.kr

## A Method of Hand Recognition for Virtual Hand Control of Virtual Reality Game Environment

Boo-Nyon Kim, Jongho Kim, Tae-Young Kim  
Department of Computer Engineering, SeoKyeong University

### 요 약

본 논문에서는 사용자의 손을 인식하여 가상현실 게임 환경에서 가상의 손을 제어할 수 있는 방법을 제안한다. 카메라를 통해 획득한 영상을 통하여 사용자의 손 이동과 가리키는 방향에 대한 정보를 획득하고 이를 이용하여 가상의 손을 게임 화면에 나타낸다. 사용자의 손의 움직임은 가상의 손이 물건을 선택하고 옮기도록 하는 입력 인터페이스로 활용할 수 있다. 제안하는 방법은 비전 기반 손 인식 기법으로 먼저 RGB 컬러영역에서 HSV 컬러영역으로 입력영상을 변환하고 H, S 값에 대한 이중 임계값과 연결 요소 분석을 이용하여 손 영역을 분할한다. 다음으로 분할된 영역에 대하여 0, 1차 모멘트를 적용하고 이를 이용하여 손 영역에 대한 무게 중심점을 구한다. 구해진 무게중심점은 손의 중심에 위치하게 되며, 분할된 손 영역의 픽셀 집합 중 무게중심점으로부터 멀리 떨어진 픽셀들을 손가락의 끝점으로 인식한다. 마지막으로 무게중심점과 손 끝점에 대한 벡터를 통하여 손의 축을 구한다. 인식 안정성과 성능을 높이기 위하여 누적 버퍼를 이용한 떨림 보정과 경계상자를 이용한 처리 영역을 설정하였다. 본 논문의 방법은 기존의 비전 기술을 통한 손 인식 방법들에 비하여 별도의 착용 마커를 두지 않고 실시간으로 처리가 가능하다. 다양한 입력 영상들에 대한 실험 결과는 제안 기법으로 정확하게 손을 분할하고, 안정된 인식 결과를 고속으로 처리할 수 있음을 보여주었다.

### ABSTRACT

In this paper, we propose a control method of virtual hand by the recognition of a user's hand in the virtual reality game environment. We display virtual hand on the game screen after getting the information of the user's hand movement and the direction thru input images by camera. We can utilize the movement of a user's hand as an input interface for virtual hand to select and move the object. As a hand recognition method based on the vision technology, the proposed method transforms input image from RGB color space to HSV color space, then segments the hand area using double threshold of H, S value and connected component analysis. Next, The center of gravity of the hand area can be calculated by 0 and 1 moment implementation of the segmented area. Since the center of gravity is positioned onto the center of the hand, the further apart pixels from the center of the gravity among the pixels in the segmented image can be recognized as fingertips. Finally, the axis of the hand is obtained as the vector of the center of gravity and the fingertips. In order to increase recognition stability and performance the method using a history buffer and a bounding box is also shown. The experiments on various input images show that our hand recognition method provides high level of accuracy and relatively fast stable results.

**Keyword** : Virtual Reality Game, Hand Interface, Hand Recognition Fingertip Extraction

접수일자 : 2010년 02월 10일, 완료일자 : 2010년 03월 16일 \* 교신저자 : 김태영

\* 이 논문은 2009년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(2009-0065960).

\* 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2009년도 문화콘텐츠산업기술지원사업의 연구결과로 수행되었음

## 1. 서론

기존의 가상 환경 게임은 키보드나 마우스와 같은 사용자 인터페이스를 통하여 게임 캐릭터가 사용자의 명령을 입력받음으로써 사용자의 행동을 게임에 반영하는 게임이 주를 이루어 게임에 대한 몰입감이 부족하였다. 사용자 인터페이스는 가상 환경 게임의 몰입감을 고조시키기 위한 중요 기술로서 최근에는 실제 사용자의 동작이나 음성 등의 인터페이스를 통하여 게임 환경에 직접적으로 영향을 미치도록 하는 실감형 게임에 대한 관심이 높아져 가고 있다. 이러한 인터페이스에 대한 연구로 사용자와 하나가 되는 인체 친화적인 인터페이스 연구가 활발하게 진행되고 있으며[1] 멀티터치 인터페이스, 센서를 이용한 동작인식에 기반한 실감형 인터페이스, 햅틱 기술을 적용한 촉각 피드백의 인터페이스 등의 인터페이스가 각광을 받고 있다. 하지만 이러한 인터페이스는 멀티터치 스크린, 센서, 햅틱 장비 등 고가의 장비가 필요하기 때문에 많은 어플리케이션에 적용하기가 힘들다는 단점을 가지고 있어 센서 등의 하드웨어를 벗어나 사람의 얼굴이나 신체의 일부를 소프트웨어적으로 인식하여 교육, 게임 등의 다양한 응용 분야에 적용할 수 있도록 하기 위한 사용자 인터페이스 연구들이 시도되고 있다.

고 등[2]은 HSV 컬러 변환과 이중 임계값 기법으로 손 영역을 분할하고, 손 영역의 궤적을 추적하여 동작을 인식하는 인터페이스를 제시하였다. 하지만 이 기법은 손 영역의 분할을 단순화하기 위해 특정 색의 장갑을 착용하여 제약을 두었고, 손의 궤적을 추적하기 위해 영역의 크기 변화를 시작 제스처로 활용하여 카메라 위치를 고정시키기 때문에 동적인 환경에도 적용할 수 없는 단점을 가지고 있다.

최 등[3]은 특정 색 정보를 이용하여 영상에서 손의 대략적 위치를 찾고, 통계학적 색상모델을 사용하여 배경을 제거하여 손 영역을 분할하였다. 그리고 주성분분석법과 신경망회로 기법을 사용하여

손 동작을 인식하였다. 하지만, 손 영역을 분할하기 위해 손목에 특정 색의 밴드를 착용하여야 한다는 제약을 두었다.

Malik 등[4]은 히스토그램 분석을 통한 임계값 기법과 연결 요소 분석 기법으로 영역에서 손가락을 분할하고, 주 축 분석으로 손 끝점에 대한 정보를 추출하는 기법을 제안하였다. 하지만, 이 기법은 모든 손가락의 주축이 평행해야한다는 가정이 전제가 되어서 응용 범위가 제한된다는 단점이 있었다.

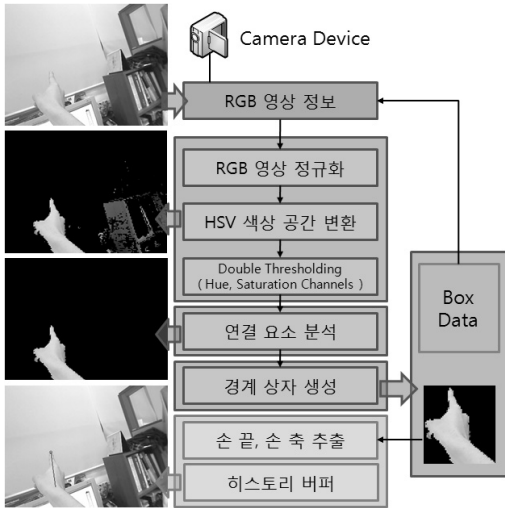
Liang 등[5]은 손 동작을 인식하기 위하여 시뮬레이티드 어닐링(Simulated Annealing) 기법을 이용한 다중 방향 탐색 기법을 제안하였다. 비교적 안정된 인식 결과를 보여주었으나 계산량이 많아 실시간 응용에는 적합하지 않았다.

본 논문에서는 비전 기반의 손 끝점 및 축방향 인식 기법을 제안한다. 먼저 RGB 컬러 영상에서 RGB를 정규화하고, 정규화된 컬러 영상을 HSV(Hue, Saturation, Value) 컬러 영상으로 변환한다. 그리고 조명에 영향을 적게 받는 Hue(이하 H)와 Saturation(이하 S) 채널을 사용하는 이중 임계값과 연결 요소 분석을 이용하여 영상에서의 손 분할 기법을 제안한다. 다음으로 분할된 영역의 무게 중심을 구하고, 이를 이용하여 손 끝점과 축 방향을 추출한다. 또한, 인식의 안정화와 성능 향상을 위하여 손 끝점과 축방향에 대한 히스토리 버퍼를 이용해 이전 영상의 결과 정보를 기억하고, 이를 현재 결과값에 보정함으로써 떨림 등의 오류를 줄였으며 분할된 영상을 모두 포함하는 최소한의 경계 상자(Bounding Box)를 만들어 영상 처리를 가속화 하여 연속되는 영상에서의 인식 속도를 높였다.

본 논문은 다음과 같은 순서로 구성되어 있다. 2장에서는 본 논문에서 제안한 손 끝점 및 축 방향 인식 기법에 대해서 설명한다. 3장에서는 실험 결과를 기술하고, 4장에서 결론을 맺는다.

## 2. 손 끝점 및 축 방향 인식 기법

[그림 1]은 본 논문에서 제안하는 기법의 전체적인 과정을 보여준다.



[그림 1] 제안하는 기법 전체적인 과정

### 2.1 HSV 컬러 영역 변환 및 이중 임계값 기법

카메라에서 얻어지는 RGB 영상 정보는 카메라의 기중에 따라서 조금씩 차이를 보여주고 있기 때문에 RGB 영상 정보를 정규화 해주는 작업이 필요하다. 그리고 RGB 컬러 영역은 조명의 영향에 따라 RGB 영상 정보 값이 크게 변하기 때문에 조명의 영향을 덜 받는 HSV 컬러 영역으로의 색상 변환이 필요하다. 본 논문에서는 HSV 컬러 영역에서도 비교적 조명의 영향을 적게 받는 Hue(이하 H)와 Saturation(이하 S) 성분의 이중 임계값을 이용하여 카메라 이미지에서 피부색 영역을 추출한다. 먼저, 카메라에서 들어오는 비정규화된 RGB 영상 정보를 (식 1)을 적용하여 정규화하고, 정규화된 RGB 영상 정보를 이용하여 (식 2)를 적용하여 HSV 컬러 영역으로 변환한다.

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B} \quad (\text{식 1})$$

$$b = \frac{B}{R+G+B}$$

$$H = \cos^{-1} \left( \frac{\frac{1}{2}[(r-g) + (r-b)]}{\sqrt{(r-g)^2 + (r-b)(g-b)}} \right) \quad (\text{식 2})$$

$$S = 1 - \frac{3}{r+g+b} [\min(r,g,b)]$$

$$V = \frac{1}{3}(r+g+b)$$

[그림 2]와 같이 다양한 샘플 영상을 만들고, 배경을 임의적으로 제거하여 피부색에 대한 영상 정보만을 남긴다. 피부 영역만 포함된 샘플 영상들의 각 픽셀에 대해서 (식 2)를 적용하여 HSV 컬러 영역으로 변환하고, H와 S 성분 영역의 빈도수를 누적시켜 평균 픽셀 수를 계산하여 H와 S에 대한 평균 히스토그램을 만들어 임계값을 결정하였다. [그림 2]와 같은 100여개의 샘플 영상을 활용하여 H는 200~250, S는 0.24~0.59의 값에서 피부색이 분포하는 것을 알 수 있었다.

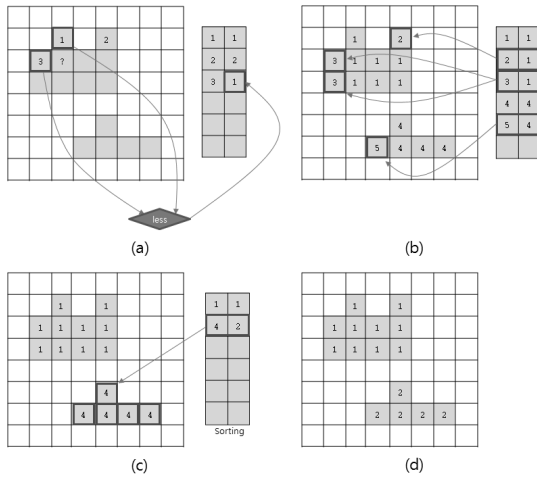


[그림 2] 손 영역에 대한 다양한 샘플 영상

### 2.2 연결 요소 분석을 통한 손 영역 분할

지금까지 제시한 H와 S의 이중 임계값을 이용한 기법은 영상에서 피부색을 추출할 수 있지만 손 영역뿐만 아니라 주변의 잡음까지 함께 추출되어 정확한 손 영역 분할을 할 수 없다. 따라서 본 논문에서는 RGB 정규화와 HSV 영상 변환, H와 S성분의 이중 임계값 적용 후 연결 요소를 분석하

여 가장 큰 영역을 손 영역이라고 가정하였다. 먼저, HSV 영상 변환과 H, S성분의 이중 임계값 적용 후 생성된 이진 영상을 스캔하여 주변 픽셀과의 관계를 고려하여 번호를 할당하고 할당되는 번호의 증가 테이블을 작성한다. 그리고 작성된 증가 테이블을 참조하여 각 픽셀의 고유 번호를 할당한다. 마지막으로는 각 연결 요소의 픽셀 수를 파악하여 큰 영역부터 작은 영역까지 오름차순으로 번호를 재 할당함으로써 연결 요소 성분에서 가장 큰 영역을 찾아 낼 수 있다. [그림 3]은 연결 요소 분석을 통한 영역 분할의 과정을 나타낸다.



[그림 3] 연결 요소 분석을 통한 영역 분할의 과정

- (a) 근접 픽셀의 레이블이 충돌이 될 때, 서로 다른 레이블 중 작은 숫자의 레이블을 할당
- (b) (a)를 모든 픽셀에 수행한 뒤, 증가 테이블의 정리를 통해, 레이블 1과 4만 남게되고, 이를 픽셀에 적용
- (c) 분리된 레이블의 픽셀 수를 기준으로 큰 영역에서 작은 영역으로 오름차순 정렬하여, 레이블을 다시 할당
- (d) 연결 요소 분석을 통해 생성된 결과

### 2.3 손 끝점과 손 축 방향 추출

본 논문에서는 경계 상자에서 손 영역의 무게 중심은 (식 3)을 통하여 구할 수 있다. 손 영역의 판단 함수 I를 사용하여 경계 상자 내에서 손 영역에 대한 좌표(x,y)를 판단하여, 무게 중심점 (Cx, Cy)를 구한다. 무게 중심점을 구한 뒤에 무게 중심점으로부터 y값이 작은 영역들에 대하여 가장 먼 거리에 있는 점을 손 끝점으로 추출하고, 손 끝점과 무게 중심점으로 생성되는 벡터를 축으로 추출하였다. 이렇게 추출된 손 끝점은 혼합 현실에 응용되어 포인팅과 손 오브젝트의 위치 등에 사용될 수 있고, 손 축은 손 오브젝트의 회전 정보로 활용되어 사용자의 손의 방향에 따라 정합되는 손 오브젝트가 동기화되어 움직이므로 실감나는 표현이 가능하다.

$$\mu_{x\_order, y\_order} = \sum_{x,y} (I(x,y) \cdot x^{x\_order} \cdot y^{y\_order}) \quad (식 3)$$

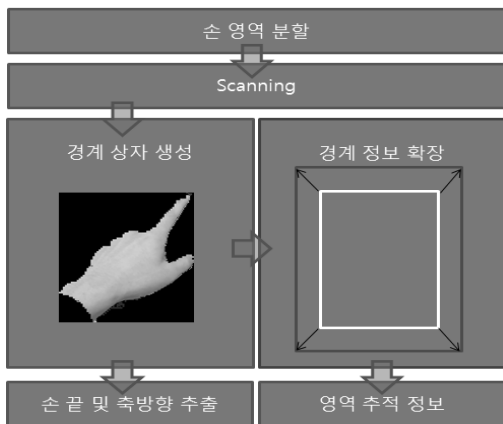
$$C_x = \frac{\mu_{10}}{\mu_{00}} \quad C_y = \frac{\mu_{01}}{\mu_{00}}$$

### 2.4 히스토리 버퍼를 사용하여 떨림 현상 보정

본 논문에서 제안하는 컬러를 이용한 영역 분할은 배경 영상의 복잡도, 조명, 환경 등에 매우 민감한 반응을 보이고 있다. 제안하는 기법이 비교적 정확한 영상 분할을 하여도 모든 영상 프레임에서 예외적인 부분이 발생할 수 있기 때문에 이러한 문제를 해결해야한다. 본 논문에서는 카메라의 움직임의 최적화를 위하여 각 프레임의 손 끝점과 축 방향에 대한 히스토리 버퍼를 사용하여 떨림 현상을 제거하는 방법을 제안한다. 각 프레임의 손 끝점과 축방향에 대한 평균 필터 및 중간 필터의 개념과 유사한 보정 방법을 통해 떨림 현상을 제거한다[6].

## 2.5 경계 상자를 통한 영상 처리 가속화 기법

입력 영상에서 손 영역이 차지하는 비율은 통상적으로 불과 20~30%에 지나지 않는다. 그렇기 때문에 다음 영상의 영상 처리에서 위와 같은 영상 처리 과정을 수행한다면 매우 큰 낭비가 된다. 따라서 본 논문에서는 불필요한 영상처리 과정을 줄이기 위하여 경계 상자<sup>1)</sup>를 생성하여 영상처리를 가속화 하였다. 경계 상자를 통한 영상처리 가속화 기법은 [그림 4]와 같다. 먼저, 분할된 손 영역을 스캔하여 손 영역의 경계 정보<sup>2)</sup>를 얻고, 경계 상자를 생성한다. 경계 정보에 대해서는 프레임속도에 적절한 오프셋을 주어 확장한 뒤, 다음 프레임의 영역 추적 정보로 활용한다. 다음 손 끝점 및 축방향 기법에서는 여기서 생성된 경계 상자를 활용한다.



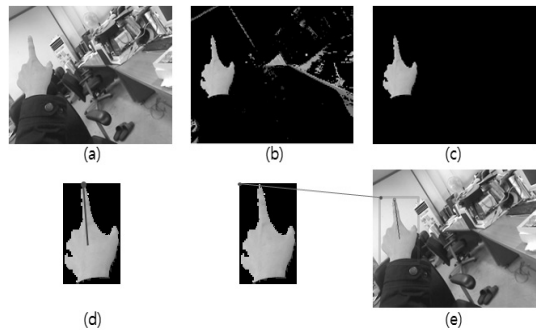
[그림 4] 경계 상자를 통한 영상처리 가속화 과정

## 3. 실험 결과

실험은 AMD Phenom(tm) II X4 920 2.80GHz 프로세서에 4GB 주 메모리를 갖는 시스템에서 수행되었다. 그래픽스 하드웨어는 ATI Radeon HD 4850을 사용하였고, Visual Studio 2005 환경에서

개발되었다. 사용된 카메라의 입력 영상의 크기는 320x240이었다. 평균적으로 입력 영상에 대한 제안 기법의 처리 속도를 측정한 결과 약 30fps의 성능을 보여주었다.

[그림 5]는 본 논문에서 제안하는 기법의 수행 과정을 단계별로 보여주었다. (a)는 카메라에서 받은 입력 영상이고, (b)는 RGB 색상 영상을 HSV로 변환하고 이중 임계값을 적용하여 피부색을 검출한 영상이다. (c)는 연결 요소 분석으로 손 영역을 남긴 영상이고, (d)는 (c)의 결과에서 손의 영역을 포함하는 최소한의 경계 상자이고, (e)는 손 끝점과 축방향을 검출한 최종 결과 영상이다. 실험 결과 본 논문에서 제안한 기법은 매우 빠른 연산 속도와 정확한 인식 결과를 확인하였다.



[그림 5] 제안하는 기법의 수행 과정

이 등[7]의 논문에서 제안하는 손 끝점 추출과 손 동작 인식 기법은 정확한 손 영역을 추출하였고, 실시간 손 동작 인식이 가능하였다. 하지만 손 끝점에 대한 인식과 동작을 이용한 입력 인터페이스를 위한 방법으로 실제 사용자의 손이 어떠한 모양으로 객체를 가리고 있는지에 대한 정보는 알 수 없다. 이에 반하여 본 논문에서 제안하는 방법은 손 끝점 추출과 함께 손이 가리고 있는 방향을 인식함으로써 실제 사용자의 손 모양을 게임에 증강시킬 수 있다. 본 논문에서 제안하는 방법의 성능을 비교하기 위하여 손 끝점에 대한 인식

1) 경계 상자 : 경계 정보로 생성되는 이미지

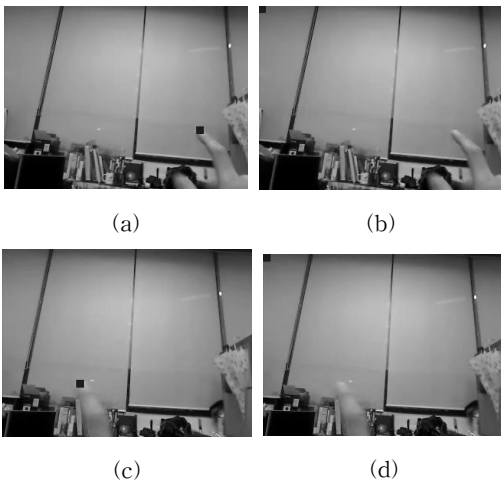
2) 경계 정보 : 경계 좌표 값

률과 인식 속도를 비교한 결과는 [표 1]과 같다. 속도는 카메라를 통하여 획득한 동영상에 대하여 손 끝점 인식 후 손 끝점을 나타내는데 소요된 fps(frame per second)이며, 인식률은 전체 동영상의 프레임 수에 대한 인식 성공률이다. 동영상은 약 700프레임으로 구성된 손 영상을 사용하였다.

[표 1] 손 끝점 인식 성능 비교 결과

비 교	속 도	인식률
제안하는 방법	38~43 fps	91.4%
[6]의 방법	31~41 fps	82.8%

인식률에 영향을 주었던 영상들은 [그림 6]과 같이 손의 일부만 포함된 영상으로 기존 방법에서는 손 끝점이 인식되지 않았지만 제안하는 방법은 손 끝점을 인식할 수 있었다.

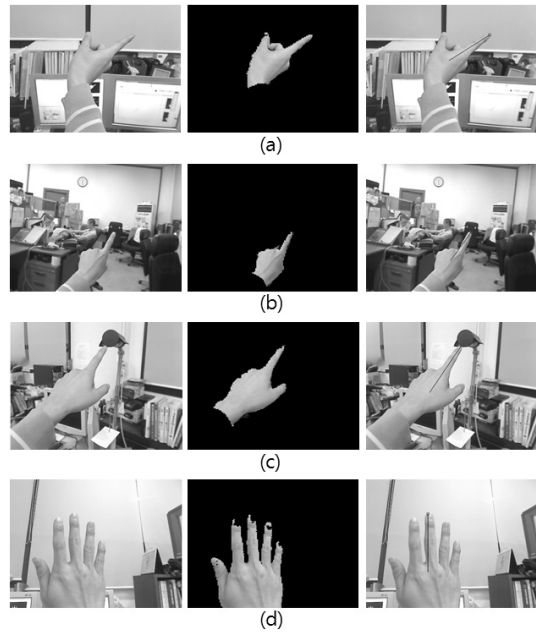


[그림 6] 손 끝점 인식을 비교 (a), (c) 제안하는 방법의 인식 결과 (b), (d)[6]의 방법의 인식 결과

또한 기존 방법에 비하여 인식에 소요되는 최저/최고 인식 속도의 차이가 심하지 않아 비교적 안정적인 속도로 인식이 가능함을 알 수 있다.

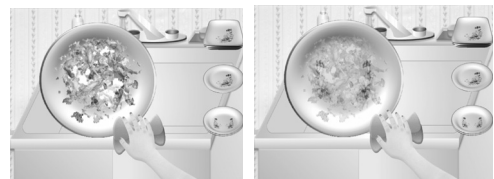
[그림 7]은 다양한 배경과 다양한 손 모양에 대하여 손 끝점 인식뿐만 아니라 축 방향을 추출한

결과를 보여준다. 제안하는 방법은 [그림 7](d)에서 보는 바와 같이 하나의 손가락이 객체를 가리키고 있을 시의 끝점 인식뿐만 아니라 2개 이상의 손가락이 객체를 가리킬 때에도 끝점 인식이 가능하다.



[그림 7] 다양한 배경과 손 모양에 대한 손 끝점 및 축 방향 인식 결과

다음 [그림 8]은 본 논문에서 제안한 방법을 이용하여 인식된 사용자의 손을 증강시킨 응용 프로그램의 결과로 추출된 사용자의 손 끝점을 이용하여 가상의 손을 이동하고, 축 방향을 통하여 손이 가리키는 방향을 프로그램에 적용하도록 한 프로그램 결과이다.





[그림 8] 손을 이용한 접시 닦기 프로그램

## 4. 결 론

본 논문에서는 비전 기반의 손 끝점 및 축방향 인식 기법을 제안하였다. 제안하는 방법은 조명에 영향을 적게 받는 Hue와 Saturation을 이용한 이중 임계값과 연결 요소 분석을 이용하여 손 영역을 분할하고 분할된 영역에 대한 모멘트 값과 특징점들을 이용하여 손의 끝점과 축 방향을 인식한다. 인식 안정성과 처리 속도를 높이기 위하여 누적 버퍼를 이용한 떨림 보정과 경계상자를 이용한 처리 영역을 설정하여 기존 방법 보다 인식률 및 성능을 향상 시켰다. 또한 하나 이상의 손가락이 객체를 가리킬 시에도 손 끝점 인식이 가능하다. 실험 결과 다양한 입력 영상들에 대하여 제안 기법이 기존 방법 보다 안정된 인식 결과를 고속으로 처리할 수 있음을 보여주었다. 또한 본 논문에서 제안한 방법을 통하여 다양한 배경과 손 모양에서도 손 끝점을 이용하는 포인팅과 축방향을 오브젝트와 동기화하여 사용자의 손을 응용 프로그램에 증강시킴으로서 실감나는 표현을 할 수 있었다. 제안 기법은 몰입감과 직관적 입력 방식이 필요한 혼합현실 기반 게임의 사용자 인터페이스로의 활용이 가능하다.

## 참고문헌

- [1] 이안재, “제품 가치를 높이는 인터페이스 기술”, SERI경제포커스, 2008. 10. 27.
- [2] 고민삼, 이광희, 김창우, 안준호, 김인중, “비전 기반 제스처 인식을 이용한 사용자 인터페이스

구현”, 한국컴퓨터종합학술대회 논문집, pp.507-511, 2008.

- [3] 최준영, 서병국, 박종일, “손 동작을 이용한 인간친화형 증강현실 인터페이스”, 한양대학교, 2009. 1.
- [4] S. Malik, C. McDonald, G. Roth, “Hand tracking for interactive pattern-based augmented reality”, Proceedings of International Symposium on Mixed and Augmented Reality, pp.117-126, 2002.
- [5] W. Liang, Y. Jia, F. Sun, B. Ning, T. Liu, X. Wu, “Visual Hand Tracking Using MDSA Method”, IMACS Multiconference on Computational Engineering in Systems Applications, pp.255-259, 2006.
- [6] 윤종현, 이범종, 박종승, “히스토리 버퍼를 사용하여 떨림 현상을 줄이는 마커 추적”, 한국정보과학회, pp.448-452, 2006. 10.
- [7] 이정진, 김중호, 김태영, “증강현실 응용을 위한 손 끝점 추출과 손 동작 인식 기법”, 한국멀티미디어학회, 2010. 02.



김 부 년 (Boo-Nyon Kim)

2009년-현재 서경대학교 컴퓨터공학과 학사 과정

관심분야 : 증강현실, 실시간 렌더링, 게임

---



김 종 호 (Jongho Kim)

2008년 2월 서경대학교 컴퓨터공학과 학사

2010년 2월 서경대학교 전자컴퓨터공학과 석사

2010년 3월-현재 서경대학교 전자컴퓨터공학과 박사과정

관심분야 : 증강현실, 실시간 렌더링, 모바일3D

---



김 태 영 (Tae-Young Kim)

1991년 2월 이화여자대학교 전자계산학과 학사

1993년 2월 이화여자대학교 전자계산학과 석사

1993년 3월-2002년 2월 한국통신 멀티미디어연구소  
선임연구원

2001년 8월 서울대학교 컴퓨터공학부 박사

2002년 3월-현재 서경대학교 컴퓨터 공학과 조교수

관심분야 : 실시간 렌더링, 모바일3D, 증강현실, 불륨  
그래픽스

---