

증강현실 응용을 위한 손 끝점 추출과 손 동작 인식 기법

이정진[†], 김종호^{**}, 김태영^{***}

요 약

본 논문에서는 증강현실 응용을 위한 손 끝점 추출과 손 동작 인식 기법을 제안한다. 먼저 RGB 컬러 영역에서 HSV 컬러 영역으로 영상을 변환하고, H, S 값에 대한 이중 임계값과 영역 확장 기법, 연결 요소 분석을 이용한 손을 분할한다. 다음으로 가상 키보드와 마우스의 입력을 위해서 모폴로지 연산과 감산을 통하여 검지와 엄지 손가락 끝점을 추출한다. 마지막으로 마우스 버튼 클릭의 역할을 하기 위하여 손바닥의 무게중심점에 대한 엄지와 검지 손가락 끝점의 각도를 계산하여 엄지와 검지 손가락의 떨어지고, 붙는 동작을 인식한다. 다양한 입력 영상들에 대한 실험 결과는 제안 기법이 고속으로 정확하게 손을 분할하고, 손 끝점을 추출하여 손 동작을 인식할 수 있음을 보여주었다. 제안 기법은 증강 현실 응용을 위한 정보 입력 인터페이스로 사용될 수 있다.

Fingertip Extraction and Hand Motion Recognition Method for Augmented Reality Applications

Jeongjin Lee[†], Jong Ho Kim^{**}, Tae-Young Kim^{***}

ABSTRACT

In this paper, we propose fingertip extraction and hand motion recognition method for augmented reality applications. First, an input image is transformed into HSV color space from RGB color space. A hand area is segmented using double thresholding of H, S value, region growing, and connected component analysis. Next, the end points of the index finger and thumb are extracted using morphology operation and subtraction for a virtual keyboard and mouse interface. Finally, the angle between the end points of the index finger and thumb with respect to the center of mass point of the palm is calculated to detect the touch between the index finger and thumb for implementing the click of a mouse button. Experimental results on various input images showed that our method segments the hand, fingertips, and recognizes the movements of the hand fast and accurately. Proposed methods can be used the input interface for augmented reality applications.

Key words: Fingertip Extraction(손 끝점 추출), Hand Motion Recognition(손 동작 인식), Hand Interface(손 인터페이스), Augmented Reality Application(증강 현실 응용)

※ 교신저자(Corresponding Author) : 김태영, 주소 : 서울특별시 성북구 정릉동(136-704), 전화 : 02)913-1362, FAX : 02)919-0345, E-mail : tykim@skuniv.ac.kr

접수일 : 2009년 7월 7일, 수정일 : 2009년 9월 30일

완료일 : 2009년 11월 18일

[†] 정회원, 가톨릭대학교 디지털미디어학부 전임강사 (E-mail : imagination@catholic.ac.kr)

^{**} 준회원, 서경대학교 컴퓨터공학과

(E-mail : sebhi@naver.com)

^{***} 정회원, 서경대학교 컴퓨터공학과 부교수

※ 본 연구는 2009년도 가톨릭대학교 교비연구비의 지원으로 이루어졌음. 또한, 본 연구는 2009학년도 서경대학교 교내연구비 지원에 의하여 이루어졌음. 또한, 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2009년도 문화콘텐츠 산업기술지원사업의 연구결과로 수행되었음.

1. 서론

정보기술의 고도화로 인하여 언제 어디서나 컴퓨터를 사용할 수 있는 인간 중심의 컴퓨팅 환경인 유비쿼터스 컴퓨팅(Ubiquitous Computing) 환경으로 발전함에 따라서 정보 단말기는 휴대성과 착용성을 고려해 소형화 되고 있다. 이러한 유비쿼터스 정보 단말기를 위한 증강현실용 디스플레이(Display) 장치로는 투시형 HMD(Head Mounted Display)가 널리 사용되고 있다. 이러한 HMD와 같은 착용형 디스플레이 장치의 사용 편리성을 높이기 위하여 3차원 공간 상에서 정보를 입력할 수 있는 인터페이스 기술에 대한 요구가 증대되고 있다. 증강현실 응용을 위한 인터페이스 기법으로 PDA 기반의 펜과 터치스크린 기반의 인터페이스[1], 작은 마커를 부착한 펜을 이용하는 인터페이스[2], 마커 기반 입력 인터페이스[3] 등이 제안되었다. 하지만, 이러한 기법들은 입력을 위하여 별도의 장치가 필요하다는 단점이 있었다. 또한, Kim 등은 실제 키보드와 마우스 없이 허공에서 손과 손가락의 움직임으로 문자 입력 기능을 수행하기 위해 손가락과 손등에 착용하는 센서를 이용한 착용형 키보드를 제안하였다[4]. 하지만, 이 기법 역시 센서를 착용해야 한다는 단점이 있었다. 따라서 유비쿼터스 환경에서 키보드와 마우스와 같은 별도의 입력 장치 없이도 3차원 증강현실 공간에서 정보를 입력하기 위하여 손 영역을 분할하고, 손 동작을 인식하여 인터페이스로 사용하는 기법에 대한 필요성이 증가되고 있다.

손 동작을 인식하기 위한 기존 연구들은 다음과 같다. Malik 등은 히스토그램 분석을 통한 임계값 기법과 연결 요소 분석 기법으로 손가락을 분할하고, 주축(principal axis) 분석으로 손 끝점을 추출하는 기법을 제안하였다[5]. 하지만, 이 기법은 모든 손가락의 주축이 평행하다는 가정이 전제가 되어서 응용 범위가 제한된다는 단점이 있었다. Liang 등은 손 동작을 인식하기 위하여 시뮬레이티드 어닐링(Simulated Annealing) 기법을 이용한 다중 방향 탐색 기법을 제안하였다. 비교적 강인한 인식 결과를 보여주었으나 계산량이 많아 실시간 응용에는 적합하지 않았다[6]. Sato 등은 손가락으로 물체를 치는 간단한 손 동작 인식을 위한 기법을 제안하였다[7,8]. 이를 위하여 먼저 단색의 배경으로부터 살색의 손 영역을 검출

하고, 템플릿 매칭 기법으로 손 끝을 인식하였다. 이 기법은 손이 위치한 배경이 살색과 구분이 되는 단색이어야 한다는 단점이 있었다. 김 등은 RGB 컬러 모델을 YCbCr로 변환하고, 경계 검출 기법, 임계값과 영역 성장법을 이용한 손 영역 분할 기법을 제안하였다[9]. 이 기법은 배경과 손 영역이 명확히 구분되는 영상에만 적용이 가능하다는 문제점이 있었다.

본 논문에서는 증강현실 응용을 위한 손 끝점 추출과 손 동작 인식 기법을 제안한다. 먼저 RGB 컬러 영역에서 HSV 컬러 영역으로 영상을 변환하고, H, S 값에 대한 이중 임계값과 영역 확장 기법, 연결 요소 분석을 이용한 손 분할 기법을 제안한다. 다음으로 가상 키보드와 마우스의 입력을 위해서 모폴로지 연산과 감산을 통하여 검지 손가락, 엄지 손가락 끝점을 추출하여 키보드 입력의 역할을 하는 기법을 제안한다. 마지막으로 손바닥의 무게중심점에 대한 엄지와 검지 손가락 끝점의 각도를 계산하여 엄지와 검지 손가락의 떨어지고, 붙는 동작을 인식하여 마우스 버튼의 역할을 하는 기법을 제안한다.

본 논문은 다음과 같은 순서로 구성되어 있다. 2장에서는 본 논문에서 제안한 손 끝점 추출과 손 동작 인식 기법에 대하여 설명한다. 3장에서는 실험 결과를 기술하고, 4장에서는 결론을 맺는다.

2. 손 끝점 추출과 손 동작 인식 기법

손 끝점 추출과 손의 동작을 인식하기 위해서는 먼저 영상으로부터 손 영역을 분할한 뒤 손가락의 끝점들을 추출하고 이들의 위치 관계를 통하여 동작을 인식해야 한다. 본 논문에서는 이중 임계값과 영역 확장 기법을 이용한 손 분할 기법과 형태학적 필터링과 감산 기법을 이용한 손 끝점 추출 기법을 제

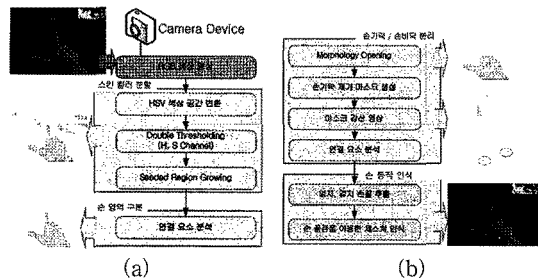


그림 1. 본 논문에서 제안한 기법. (a) 손 영역 분할 과정 (b) 손 동작 인식 과정

안한다. 또한 추출된 손 끝점들의 위치 관계를 이용하여 손이 가리키는 지점과 이를 선택하는 기법을 제시한다. 그림 1은 본 논문에서 제안한 기법의 전체적인 과정을 보여준다.

2.1 색상 변환 및 이중 임계값 기법

카메라 장치로부터 얻어지는 RGB 영상 정보는 같은 물체라도 조명에 따라 영상 정보 값이 크게 바뀌기 때문에 조명의 영향을 덜 받는 색상 변환이 필요하다. 본 논문에서는 카메라로부터 입력된 RGB 영상 정보를 HSV(Hue, Saturation, Value) 영상 정보로 변환하여 V에 비하여 조명의 영향을 덜 받는 H, S 성분을 이용하여 손을 포함하는 살색 영역을 추출한다. 먼저 RGB 영상 정보를 식 (1)과 같이 정규화하고 정규화 된 각 성분 r, g, b를 식 (2)를 통하여 HSV 영상 정보로 변환한다[10].

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

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

$$S = 1 - \frac{3}{r+g+b} [\min(r, g, b)], V = \frac{1}{3}(r+g+b) \quad (2)$$

손 영역에 해당하는 H, S값의 임계값을 구하기 위해 그림 2와 같이 다양한 경우의 피부색 및 조명에 대하여 샘플(Sample) 영상들을 만들고, 그림 3과 같이 손 영역이 아닌 부분을 수동으로 제거하였다. 손 영역만 포함된 샘플 영상들의 각 픽셀(Pixel)들에 대해서 H, S값의 빈도수를 각각 누적시킨 후 평균 픽셀

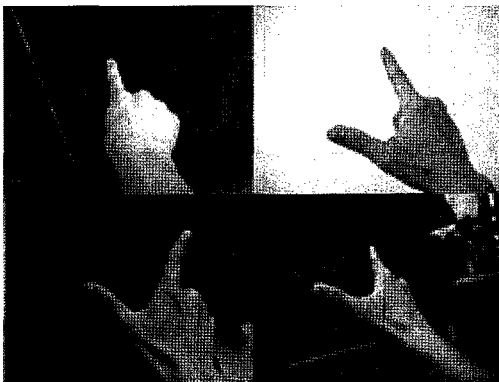


그림 2. 손 영상들의 예

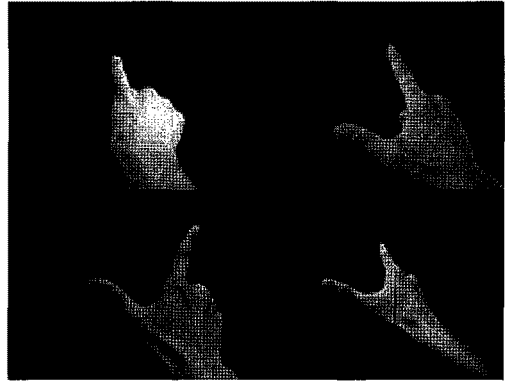


그림 3. 손 영역만 포함된 영상

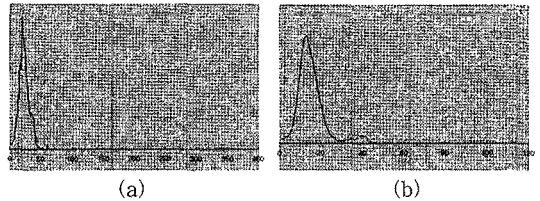


그림 4. 평균 히스토그램 분포 (a) H값 (b) S값

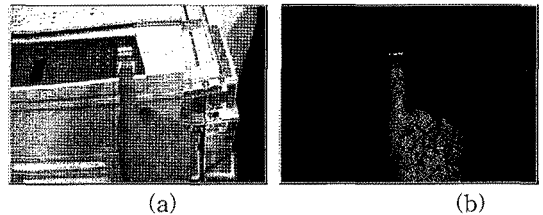


그림 5. 이중 임계값을 적용한 분할 결과. (a) 샘플 영상 (b) 샘플 영상에 대한 이중 임계값 적용 결과

수를 계산하여 H값과 S값에 대한 평균 히스토그램을 만들었다(그림 4). H값은 0-41의 값에서 손 영역이 많이 분포하였고, S값은 7-33의 값에서 손 영역이 많이 분포됨을 알 수 있었다. 히스토그램을 통하여 구한 H, S에 대한 이중 임계값을 입력 영상에 적용한 결과는 그림 5와 같다.

2.2 영역 확장 기법을 이용한 손 영역 분할

이중 임계값을 이용한 손 영역 추출 기법은 영상에서 살색 영역을 추출할 수 있지만, 그림 5의 결과에서 보는 것과 같이 정확한 손 영역 분할이 어렵다. 따라서 본 논문에서는 HSV 영상 정보 변환과 이중 임계값 적용 후 영역 확장 기법을 추가로 적용하였

다. 먼저 이중 임계값 기법을 통하여 추출된 영상의 픽셀들을 시드(Seed)로 지정하여 큐(Queue)에 삽입하고, 큐에 삽입된 픽셀을 기준으로 4방향 인접 픽셀들의 H, V 값의 차가 일정한 값 이내일 때 해당되는 픽셀을 큐에 삽입한다(그림 6). 큐에 삽입된 모든 시드가 없어질 때까지 이와 같은 과정을 반복하여 영역을 확장한다. 영역 확장을 위한 인접 픽셀들 간 차이

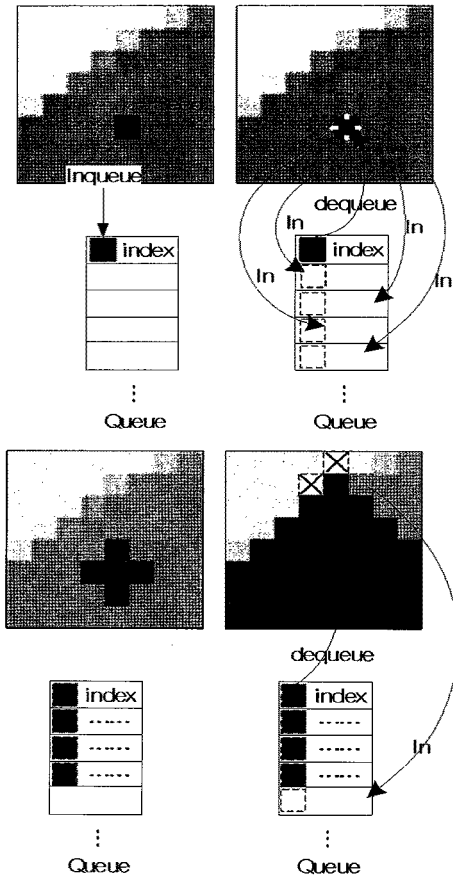


그림 6. 영역 확장 기법 과정

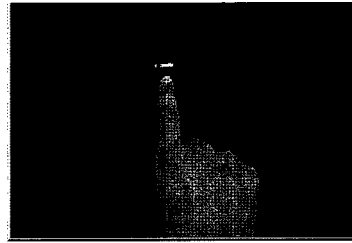


그림 7. 그림 5(a)의 샘플 영상에 대한 이중 임계값 적용 후 영역 확장 기법 적용 결과 영상

의 범위는 H는 ± 10 , V는 ± 4 로 지정하였다. 그림 7은 그림 5(a)의 샘플 영상에 대해 이중 임계값을 적용 후 영역 확장 기법을 적용한 결과 영상이다.

입력 영상에는 손 이외에도 살색에 가까운 영역이 있을 수 있다. 본 논문에서는 입력 영상에서 손 영역이 살색에 가까운 영역들 중 가장 크기가 큰 영역으로 가정한다. 따라서 영역 확장법을 적용한 영상에 연결 요소 분석[11] 기법을 적용하여 가장 큰 연결 요소를 추출하면 손 영역을 그림 8과 같이 정확하게 분할할 수 있다. 또한, 첫 프레임에서 분할된 손 영역을 정보를 기반으로 경계 상자(bounding box)를 적당한 오프셋을 주어 생성한 후 다음 프레임에서는 경계 상자 내의 영상에 대해서만 앞의 영상 처리 과정을 수행하여 연산 속도를 더욱 가속할 수 있다.

2.3 손 끝점 추출 및 손 동작 인식

분할된 손 영역에서 검지 손가락과 엄지 손가락의 끝점들은 다음과 같은 과정을 통하여 추출할 수 있다(그림 9).

본 논문에서는 7회의 연속된 침식 모폴로지 필터링과 확장 모폴로지 필터링으로 구성된 오프닝(Opening) 연산[10]을 적용하여 손 영역에서의 손가락 영역을 분리하였다. 손가락이 제거된 영상에서 무

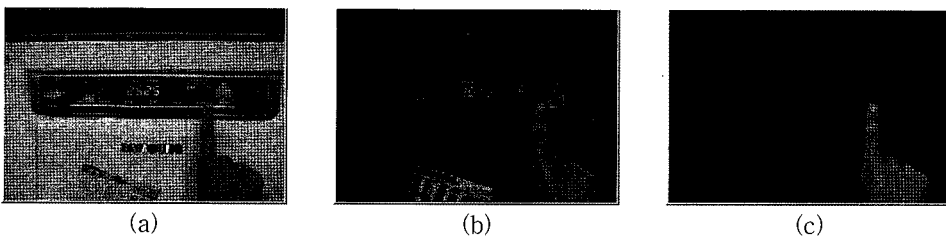


그림 8. 연결 요소 분석을 이용한 손 영역 분할. (a) 입력 영상 (b) 입력 영상에 대한 이중 임계값 기법과 영역 확장 기법 적용 결과 영상 (c) 연결 요소 분석을 통한 최종 손 분할 영상중 임계값을 적용한 분할 결과

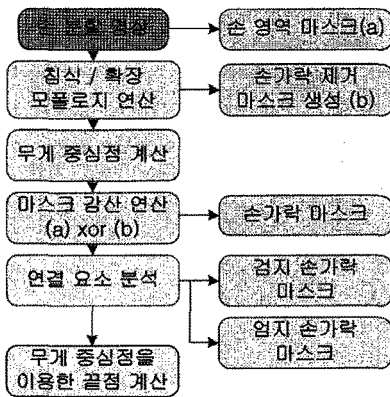


그림 9. 손 끝점 추출 과정 개요

계 중심점을 계산하고 손 영상과 손가락 제거 영상의 감산 연산을 통하여 손가락 영상을 생성할 수 있다. 손가락 영상에서는 검지 손가락, 일부 엄지 손가락 등이 포함되는데 연결 요소 분석을 통하여 가장 큰 연결 요소로 분류되는 검지 손가락 영역을 분리할 수 있다. 검지 손가락 영역과 무게 중심점의 위치 관계를 이용하여 검지 손가락 영역을 제외한 나머지 영역에서 엄지 손가락 영역을 추출한다. 검지 손가락 영역과 엄지 손가락 영역에 대해서 무게 중심점으로부터 가장 먼 거리에 있는 점을 각 손가락들의 끝점으로 추출한다(그림 10).

손 영역에서 추출된 검지 손가락 끝점(x_1, y_1), 엄지 손가락 끝점(x_2, y_2), 손바닥의 무게중심점(x_c, y_c)에 대해서 그림 11과 같은 방식으로 엄지 손가락과 검지 손가락이 무게중심점에 대하여 이루는 각도($\beta - \alpha$)를 구하고 이를 이용하여 두 손가락이 크게

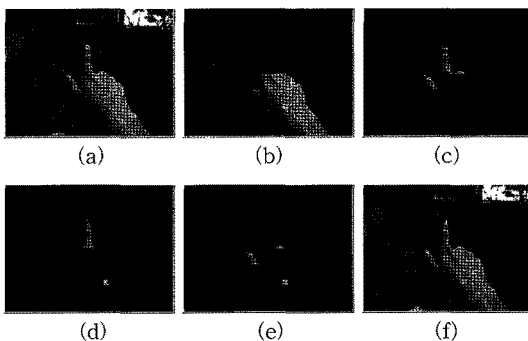


그림 10. 손 끝점 추출 과정. (a) 입력 영상 (b) 손가락 제거 영상 (c) 손가락 영상 (d) 손바닥의 무게중심점(녹색)과 검지 손가락 끝점(붉은색) 추출 (e) 엄지 손가락 끝점(파란색) 추출 (f) 최종 결과

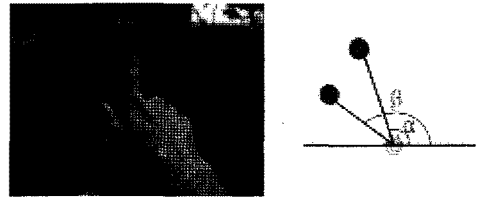


그림 11. 두 손가락 사이의 각도 측정

벌어졌을 때와 그렇지 않았을 때의 손 상태를 구분할 수 있다.

$$\tan(\alpha - \beta) = \frac{\tan\alpha - \tan\beta}{1 + \tan\alpha \tan\beta}, \tag{3}$$

$$\tan\alpha = \frac{y_1 - y_c}{x_1 - x_c}, \tan\beta = \frac{y_2 - y_c}{x_2 - x_c}$$

이렇게 추출된 검지 손가락의 위치와 검지 손가락과 엄지 손가락이 이루는 각도는 증강현실 응용에서 물체의 포인팅(Pointing), 선택, 드래그(Drag) 등의 기능에 사용될 수 있다. 먼저 물체의 포인팅은 검지 손가락 끝점에 대한 좌표를 이용할 수 있다. 다음으로 엄지와 검지 손가락이 벌어져 있을 경우와 붙어 있을 경우를 마우스의 왼쪽 버튼을 누르는 동작으로 대응시키면, 버튼의 클릭(Click), 드래그 등에 적용이 가능하다.

3. 실험 결과

실험은 Intel Core2Quad 6600 2.4GHz 에 2GB 주 메모리를 갖는 시스템에서 수행되었다. 그래픽스 하드웨어는 NVIDIA GeForce 8800을 사용하였고, Visual Studio 2005 환경에서 개발되었다. 사용된 카메라의 입력 영상의 크기는 320×240이었다. 다양한 입력 영상들에 대하여 제안 기법의 처리 속도를 측정 한 결과 약 20fps(frame per second)의 성능을 보여 주었다.

그림 12는 제안 기법의 수행 과정을 단계별로 보여주었다. 그림 13은 다양한 입력 영상들에 대하여 무게중심점(녹색), 검지 손가락의 끝점(붉은색), 엄지 손가락의 끝점(파란색)을 추출한 결과를 보여주었다. 엄지와 검지 손가락이 최대한 벌려진 상태(그림 13(a))로부터 바로 옆으로 붙은 상태까지(그림 13(b))를 보여주었다. 실험 결과 제안 기법으로 정확한 손 동작을 인식할 수 있음을 확인하였다.

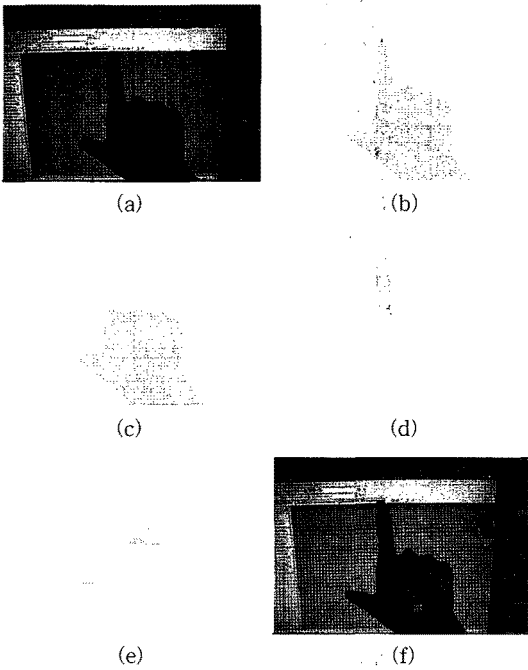


그림 12. 제안 기법의 수행 과정. (a) 입력 영상 (b) HSV 변환 후 이중 임계값 적용 영상 (c) 침식 및 확장 연산 적용 영상 (d) (b) - (c) 감산 영상 (e) 엄지 끝점 추출 결과 영상 (f) 최종 추출 결과 영상

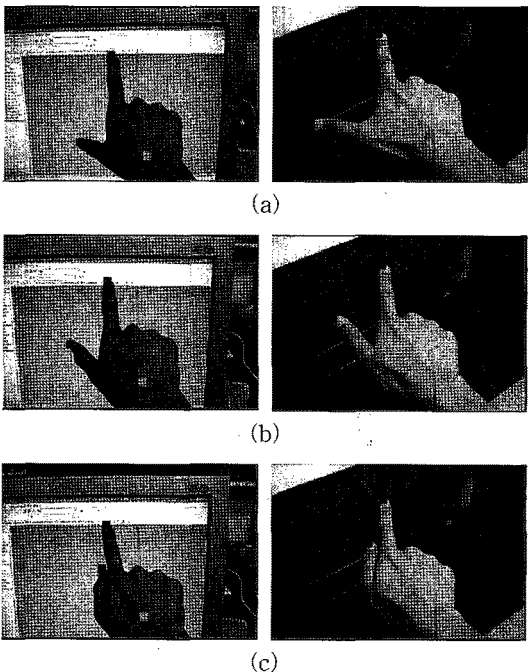


그림 13. 손끝 동작 인식 과정. (a) 손끝이 벌어진 상태 (b) 중간 상태 (c) 붙은 상태

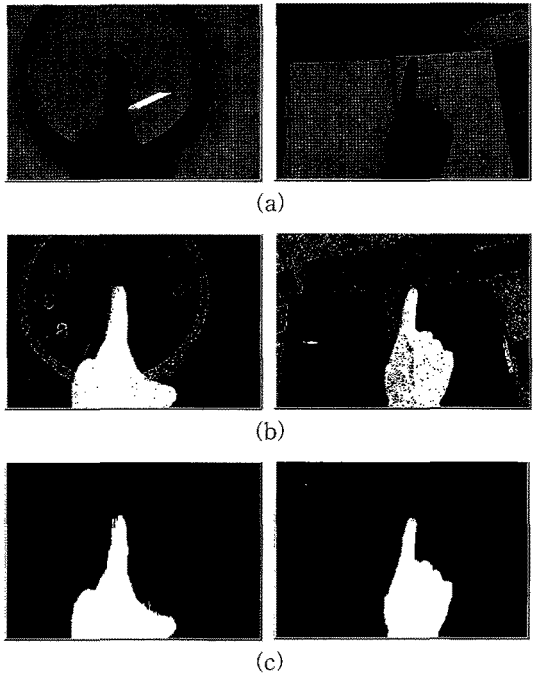


그림 14. 손 분할 결과 비교. (a) 원 영상 (b) 기존 기법[12]의 분할 결과 (c) 제안 기법의 분할 결과

하 등[12]의 논문에서 제안한 손 분할 기법과 본 논문에서 제안한 기법의 손 분할 결과를 그림 14에 비교하였다. 그림 14(b)의 기존 기법의 손 분할 결과에 비하여 그림 14(c)의 손 분할 결과는 정확한 결과를 보여주었다. 따라서 정확한 손 분할은 사용자 손 끝 동작 인식을 가능하게 한다. 또한, 제안 기법은 김 등[9]의 논문에서 제안한 손 분할 기법에 비하여 분할 결과의 정확성은 비슷하였지만, 정량적으로 평균 약 3배 정도의 속도 향상을 보여주었다. 결과적으로 제안 기법은 기존 기법들[9,12]에 비하여 고속으로 정확한 분할이 가능하였다.

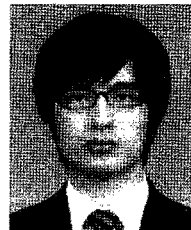
4. 결 론

본 논문에서는 증강현실 응용을 위한 손 끝점 추출과 손 동작 인식 기법을 제안하였다. 먼저 RGB 컬러 영역에서 HSV 컬러 영역으로 영상을 변환하고, H, S 값에 대한 이중 임계값과 영역 확장 기법, 연결 요소 분석을 이용한 손 분할 기법을 제안하였다. 다음으로 모폴로지 연산과 감산을 통하여 검지 손가락, 엄지 손가락 끝점을 추출하여 키보드 입력의

역할을 하는 기법을 제안하였다. 마지막으로 손바닥의 무게중심점에 대한 엄지와 검지 손가락 끝점의 각도를 계산하여 엄지와 검지 손가락의 떨어지고, 붙는 동작을 인식하여 마우스 버튼의 역할을 하는 기법을 제안하였다. 다양한 입력 영상들에 대한 실험 결과 제안 기법으로 빠르고, 정확하게 손을 분할하고, 손 끝점을 추출하여 손 동작을 인식할 수 있음을 확인하였다. 제안 기법은 증강 현실 응용을 위한 정보 입력 인터페이스로 사용될 수 있다.

참 고 문 헌

- [1] D. Wagner, T. Pintaric, and D. Schmalstieg, "The invisible train: a collaborative handheld augmented reality demonstrator," *Proceedings of ACM SIGGRAPH 2004 Emerging Technologies*, pp. 12, 2004.
- [2] D. Bandyopadhyay, R. Raskar, and H. Fuchs, "Dynamic shader lamps : painting on movable objects," *Proceedings of IEEE and ACM International Symposium on Augmented Reality*, pp. 207-216, 2001.
- [3] 한갑중, 황재인, 최승문, 김정현, "증강현실 기반의 3차원 도자기 모델링 시스템," *한국 HCI학회 논문지*, 2권, 2호, pp. 19-26, 2007.
- [4] Y. S. Kim, B. S. Soh, and S. G. Lee, "A new wearable input device: SCURRY," *IEEE Transactions on Industrial Electronics*, Vol. 52, No.6, pp. 1490-1499, 2005
- [5] S. Malik, C. McDonald, and G. Roth, "Hand tracking for interactive pattern-based augmented reality," *Proceedings of International Symposium on Mixed and Augmented Reality*, pp. 117-126, 2002.
- [6] W. Liang, Y. Jia, F. Sun, B. Ning, T. Liu, and X. Wu, "Visual Hand Tracking Using MDSA Method," *IMACS Multiconference on Computational Engineering in Systems Applications*, pp. 255-259, 2006.
- [7] T. Sato, K. Fukuchi, and H. Koike, "OHAJIKI Interface: Flicking Gesture Recognition with a High-Speed Camera," *Lecture Notes in Computer Science*, Vol.4161, pp. 205-210, 2006.
- [8] T. Sato, K. Fukuchi, and H. Koike, "Camera-based Flicking Gesture Recognition and Game Applications," *Adjunct Proceedings of the 19th annual ACM Symposium on User Interface Software and Technology*, 2006.
- [9] 김성진, 김태영, 임철수, "발달장애인을 위한 혼합현실 기반 상황훈련 시스템," *한국컴퓨터그래픽스학회 논문지*, Vol.14, No.2, pp. 1-8, 2008.
- [10] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital image processing using matlab*, Prentice-Hall, 2004.
- [11] F. Chang, C. J. Chen, and C. J. Lu, "A linear-time component-labeling algorithm using contour tracing technique," *Computer Vision and Image Understanding*, Vol.93, No.2, pp. 206-220, 2004.
- [12] 하태진, 우운택, "Video see-through HMD 기반 증강현실을 위한 손 인터페이스," *한국 HCI학회*, pp. 169-174, 2006.



이 정 진

2000년 2월 서울대학교 기계항공
학부 학사
2002년 2월 서울대학교 컴퓨터공
학부 석사
2005년 3월 New York Institute of
Technology 경영학 석사
2008년 8월 서울대학교 컴퓨터공
학부 박사

2007년 10월~2009년 2월 울산대학교 의과대학 영상의
학과 연구교수
2008년 1월~현재 (주) 클리니컬 이미징 솔루션 기술이사
2009년 3월~현재 가톨릭대학교 디지털미디어학부 전
임강사
관심분야 : 컴퓨터 애니메이션, 영상 정합, 영상 분할, 가
상 내시경, 소셜 컴퓨팅



김 종 호

2008년 2월 서경대학교 컴퓨터
공학부(학사)
2008년~현재 서경대학교 컴퓨터
공학부(석사)
관심분야 : 실시간 렌더링, 모바일
그래픽스, 가상 현실



김 태 영

1991년 2월 이화여자대학교 전자
계산학과 학사
1993년 2월 이화여자대학교 전자
계산학과 석사
1993년 3월~2002년 2월 한국통
신 멀티미디어연구소 선
임연구원
2001년 8월 서울대학교 컴퓨터공학부 박사
2002년 3월~현재 서경대학교 컴퓨터 공학과 조교수
관심분야 : 실시간 렌더링, 모바일3D, 증강현실, 블롭그
래픽스