

## 비전 카메라를 이용한 멀티 터치 입력 장치

### Visual Multi-touch Input Device Using Vision Camera

서 효 동 · 주 영 훈\*

Hyo Dong Seo and Young Hoon Joo

군산대학교 제어로봇공학과

#### 요 약

본 논문에서는 비전 카메라를 이용한 비주얼 멀티 터치 입력 장치를 제안한다. 제안된 입력 장치는 비전 카메라를 이용하므로 기존 스마트 폰에서 사용하는 터치 스크린 입력 장치에 비해 non-touch로 명령을 입력할 수 있다는 장점이 있다. 또한, 제안된 장치는 컴퓨터 계산 시간이 짧아 실시간 구현이 가능하고, 카메라 외에 다른 장치가 필요 없으므로 존재하는 입력장치보다 싸다는 장점이 있다. 이를 위해, 먼저, HSV 컬러 모델과 라벨링 방법을 이용한 영상 처리 알고리즘을 제안한다. 그 다음, 손이 움직이는 영역의 정확성을 개선하기 위해, 기하학적인 특징점, 중심원, 칼만 필터에 기본적인 움직임 인식 알고리즘을 제안한다. 마지막으로, 제안된 장치는 비디오 게임, 스마트 TV, 컴퓨터 등을 원격으로 제어할 수 있음을 실험을 통해 보인다.

**키워드 :** 비주얼 멀티-터치, 레이블링, COG, 손 영역, 칼만 필터

#### Abstract

In this paper, we propose a visual multi-touch air input device using vision cameras. The implemented device provides a barehanded interface which copes with the multi-touch operation. The proposed device is easy to apply to the real-time systems because of its low computational load and is cheaper than the existing methods using glove data or 3-dimensional data because any additional equipment is not required. To do this, first, we propose an image processing algorithm based on the HSV color model and the labeling from obtained images. Also, to improve the accuracy of the recognition of hand gestures, we propose a motion recognition algorithm based on the geometric feature points, the skeleton model, and the Kalman filter. Finally, the experiments show that the proposed device is applicable to remote controllers for video games, smart TVs and any computer applications.

**Key Words :** Visual multi-touch, Labeling, COG, Hand region, Kalman filter

#### 1. 서 론

현대 사회에서 컴퓨터 정보기술의 발전으로 정보기기 사용이 보편화됨에 따라 사용기와 인간 사이에, 자연스러운 상호작용의 중요성이 증가하고 있다[1-4]. 이에 따라 정보기기는 텍스트 위주의 인터페이스에서, 인간의 다양한 제스처 등을 사용하고 사용자 경험을 중요시 하는 MMI(Man-Machine Interface) 방식으로 발전하고 있다[5-8]. 그 중에서도 터치스크린은 사용자 경험을 고려하는 대표적인 장치로서 스마트폰, 태블릿 PC, 게임 등에 많이 이용되고 있다[9,10]. 이러한 방식은 직관적인 도구인 손을 이용하여 직접 정보를 입력하기 때문에 별도의 입력 장치가 필요하지 않고 사용자 훈련 여부와 관계없이 사용할 수 있다는 장점이 있다. 기존의 터치스크린은 구현원리와 동작 방법에 따라 저항막(감압) 방식, 정전용량 방식, 적외선 방식, 초음파 방식 등의 다양한 방식으로 구분된다. 저항막 방식은 모든 것을 이용해

화면을 터치할 수 있다. 그러나 화면 선명도가 떨어지고 충격에 약하다는 단점이 있다. 정전용량 방식은 화면을 살짝 스치듯 만져도 터치 입력이 가능하며 화질 저하도 없다. 하지만 손가락처럼 전자를 유도하는 물질이 아닐 경우 터치 입력이 불가능하고 내구성이 약하다. 이러한 기존 문제점 외에도 대형화 되어가는 터치스크린이나 높은 곳에 위치한 스크린은 손이 닿지 않는 사각지역이 발생하여 터치 입력의 어려움이 있다[11-14]. 이를 극복하기 위해서 비전 기반의 손 동작 인식을 이용한 시스템 개발이 필요하다.

손 동작 인식은 글러브 데이터를 이용하는 방법과, 비전을 이용하는 방법으로 구분할 수 있다. 글러브 데이터를 사용하는 방법은 손의 각 부위에 굴절계측기를 장착한 장갑의 출력 값을 이용하는 기기 기반의 측정 방식으로 손동작에 대한 정확한 위치를 얻을 수 있다. 하지만 고가의 센서가 필요하고, 사람과 센서간의 연결선이 필요하기 때문에 사용이 불편하고 손동작의 행동 반경이 제한되는 단점이 있다. 비전을 이용하는 방법은 별도의 기계적인 장치를 직접 사용하지 않고, 카메라를 이용한다. 하지만 조도의 변화에 민감하여 피부 뿐만 아니라 피부와 비슷한 색도 피부로 인식하여 잡음이 심하기 다양하고 복잡한 환경에서 손 형상만을 추출하여 인식하기 어렵다. 고로, 다양한 환경에서의 손 영역을 추출하기 위한 알고리즘이 필요하며 실시간에서의 응용을 위한 빠른 검출 속도를 필요로 한다.

접수일자 : 2011년 11월 19일

완료일자 : 2011년 12월 13일

+ 교신저자

본 논문은 2011년도 추계학술대회 우수논문으로 선정되어 본 논문지에 게재됨.

본 논문에서는 손 동작 인식에 대한 기존 연구의 문제점을 개선하여, 비전을 이용한 비주얼 멀티-터치 방법 및 무게 중심을 이용한 손 영역 특징점 추출 방법을 제안한다. HSV 컬러 모델을 이용하여 피부 영역을 추출하고 손 후보 영역 검출을 위해 레이블링 기법을 이용한다. COG(Center Of Gravity)를 이용하여 손의 중심을 찾고 원을 그린다. 원을 확대 시켜 가면서 레이블링 된 손 영역의 손 끝 이진신호를 탐색하여 손 끝 특징점을 추출한다. 손의 빠른 움직임에서는 특징점들의 원활한 추적이 불가능한 경우가 많기 때문에 칼만 필터를 이용하여 오차를 보정하고 추적을 수행한다. 마지막으로, 제안된 장치는 비디오 게임, 스마트 TV, 컴퓨터등을 원격으로 제어할 수 있음을 실험을 통해 보인다.

## 2. 손 영역 특징점 시스템 개요

본 논문에서는 별도의 하드웨어를 손에 부착하지 않는 손 영역 특징점 추출 기법을 제안한다. 제안하는 특징점 추출 시스템은 그림 1과 같다.

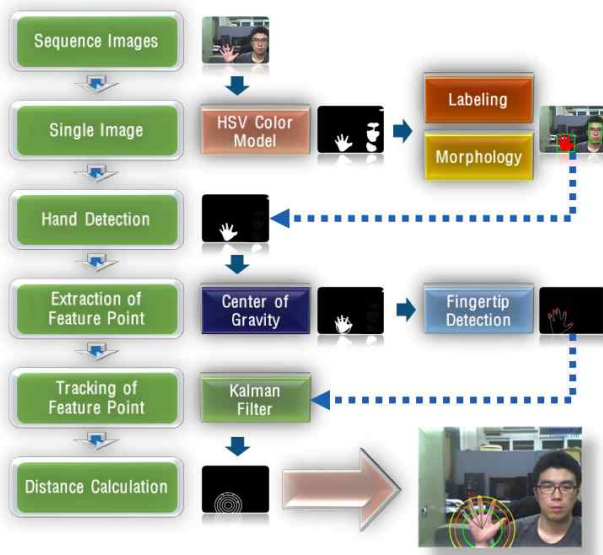


그림 1. 제안된 비주얼 멀티 터치 시스템  
Fig. 1. The proposed visual multi-touch system

웹 카메라로부터 입력받은 RGB 색상정보는, 조도의 변화에 강인한 HSV 컬러 모델로 변환하여 피부 영역을 추출한다. 손 영역을 추출하기 위해서 레이블링 기법으로 객체를 생성 한 후, 일정 크기 이 객체는 필터링을 거쳐 최종적인 손 영역을 추출한다. 손 영역 이외의 잡음들을 제거하기 위해 모폴로지 필터링을 수행한다. 손 영역의 특징점은 무게 중심법을 이용하여 원을 그리고 레이블링 된 손 영역까지 원을 확장 및 탐색하여 손 끝의 이진신호를 이용하여 손 끝 특징점을 찾는다. 찾아진 특징점들은 칼만 필터를 이용하여 오차를 보정하고 추적하게 된다. 손 끝 특징점과 손 중심점의 거리를 계산하여 터치 유무를 감별한다.

## 3. 손 영역 추출

### 3.1 피부 영역 검출을 위한 HSV 색상 모델

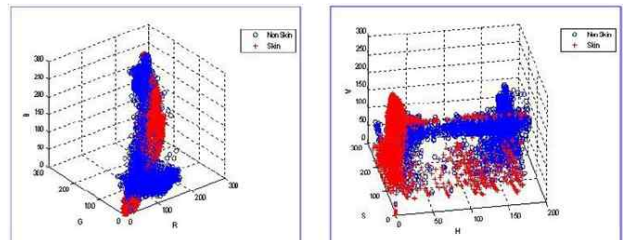
손 후보 영역을 추출하기 위한 첫 번째 작업은 피부 영역을 추출하는 것이다. 전체 컬러 공간에서의 피부는 매우 적은 영역만을 차지하고 있다. 피부색만의 분포를 적절한 모델로 근사화 하여 배경색과 피부색의 구분이 가능하다. 웹 카메라로부터 입력받은 RGB 색상정보는 다음 식을 통해 조도 변화에 강인한 HSV 컬러 모델로 변환하고 피부 영역을 검출한다.

$$H = \cos^{-1} \left[ \frac{\frac{1}{2}[(R-G)-(R-B)]}{\sqrt{(R-G)^2+(R-B)(G-B)}} \right] \quad (1)$$

$$S = 1 - \frac{3}{R+G+B} [\min(R,G,B)]$$

$$V = \frac{1}{3}(R+G+B)$$

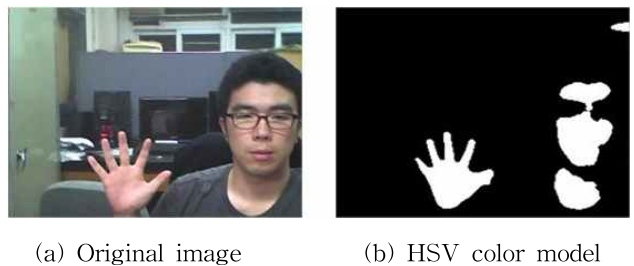
H는 색상의 분포, S는 순수한 색상에 하얀 빛이 포함된 정도인 채도, V는 빛의 밝기인 명도를 나타낸다. V의 값을 제거하거나 조절해 준다면 조명에 강인한 색상 모델을 만들 수 있다. 그림 2는 RGB 영역과 HSV 영역에서의, 피부 및 배경의 색상 분류를 나타낸다.



(a) RGB Color model (b) HSV color model

그림 2. 피부와 배경의 색상 분류  
Fig. 2. Color classification of skin and background

색상 표현의 한계가 있는 RGB 영역에서는 피부와 배경 분리가 어려웠지만 HSV 영역에서는 원활한 분리가 가능하다. 그림 3은 HSV 색상 모델을 이용한 피부 검출 영상이다.



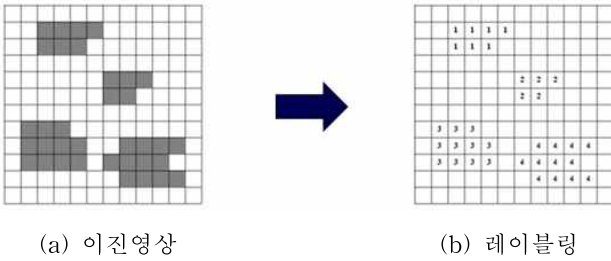
(a) Original image (b) HSV color model

그림 3. 피부 영역 추출  
Fig. 3. Extraction of skin region

### 3.2 손 영역 추출을 위한 레이블링

손 후보 영역에서 최종적인 손 영역을 추출하기 위해서 레이블링을 이용한다. 레이블링 기법은 이진영상에서 수행되는 영역 구분 방법이다. 그림 4의 (a)에서는 이진영상 내에서 이웃 화소로 뭉쳐있는 영역을 검출하고 (b)에서 모든 픽셀에 고유한 번호를 붙여 하나의 객체라고

규정한다. 한 화소가 다른 화소들에 연결되어 있는지를 알기 위하여 8-이웃화소를 통해 손 후보 영역을 레이블링 한다.



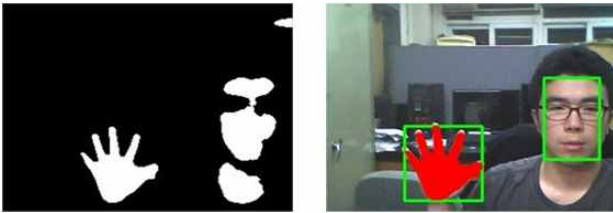
(a) 이진영상 (b) 레이블링

그림 4. 레이블링 과정  
Fig. 4. Labeling process

레이블링 된 영상에서 손 영역 크기 이하의 레이블은 식 (2)에 의해 잡음으로 판단하여 제거한다.

$$f(x) = \begin{cases} 255, & \text{labelsize} \geq \text{threshold} \\ 0, & \text{other} \end{cases} \quad (2)$$

그림 5는 레이블링을 통한 손 영역 추출 영상을 나타낸다.



(a) HSV color model (b) Hand labeling

그림 5. 손 영역 추출  
Fig. 5. Extraction of hand region

### 3.3 영상의 잡음 제거를 위한 모폴로지 기법

조명에 의해 손실된 손영역을 보정하고 미세하게 영상 곳곳에 분포한 작은 노이즈를 제거하기 위해 모폴로지(morphology) 기법을 이용한다. 모폴로지 알고리즘에는 배경과 객체 중심을 둔다면 배경 확장과 객체를 축소시키는 침식연산, 배경 축소와 객체 확장 시키는 팽창 연산, 미세한 잡음을 제거하는 열림 연산, 객체 내의 빈 공간을 메워주는 닫힘 연산으로 구성되어 있다.

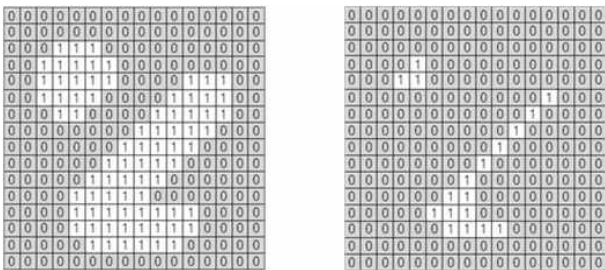


그림 6. 침식 연산  
Fig. 6. Erosion operation

그림 6은 침식 연산을 나타낸다. 객체의 크기는 줄어들고

배경은 확대된다. 영상의 객체와 배경 사이에 잡음이 있을 때 제거하거나, 전체 영상에서 아주 작은 물체를 제거할 때 이용된다. 영상에서의 돌출부는 감소시키고, 내부 돌출부는 증가시켜서 서로 닿은 물체를 분리 하기도 한다.

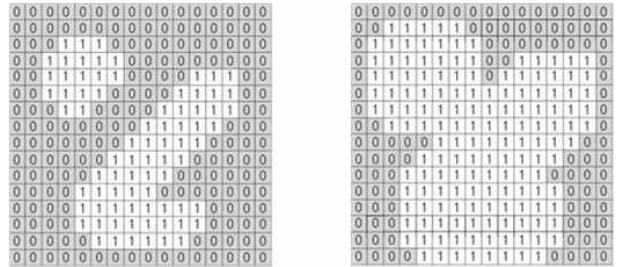


그림 7. 팽창 연산  
Fig. 7. Dilation operation

그림 7은 팽창 연산을 나타낸다. 객체 내부의 돌출부는 감소하고 외부의 돌출부는 증가시켜서 물체의 크기를 확장하고 배경을 축소하는 기법이다. 객체 내부에 발생한 구멍과 같은 공간을 채우거나 짧게 끊어진 영역을 연결 할 때 이용된다.

### 4. 무게 중심을 이용한 손 영역 특징점 추출

본 장에서는 추출된 손 영역으로부터 거리 계산을 통한 터치 동작을 생성하기 위해 손의 중심점과 다섯 손가락의 끝점을 추출한다.

#### 4.1 손 영역 무게 중심

손바닥 영역은 상대적으로 크고 원과 비슷한 모양을 가지며, 손가락은 상대적으로 작고 긴 사각형 모양의 형태학적인 특징을 가지고 있다. 손의 형태와 손바닥의 중심을 기술하는데 모멘트 이론을 적용 할 수 있다. 영상  $I(x,y)$ 의 모멘트는 식 3과 같이 주어진다.  $I(x,y)$ 는 영상의  $x,y$  위치에서 픽셀의 값을 나타낸다.

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q I(x,y) dx dy \quad p,q=0,1,2,\dots \quad (3)$$

손 영역에서의 무게 중심은 식 4를 통해 구할 수 있고 이는 픽셀들의 합으로 나타낼 수 있다.

$$m_{00} = \sum_x \sum_y I(x,y) \quad (4)$$

식 5를 이용하여 손의 1차 모멘트를 구할 수 있다.

$$m_{01} = \sum_x \sum_y x I(x,y), \quad m_{10} = \sum_x \sum_y y I(x,y) \quad (5)$$

1차 모멘트를 이용하여 손의 중심점을 구할 수 있다. 중심점을 구하는 방법은 식 6과 같다.

$$x_c = \frac{m_{10}}{m_{00}}, \quad y_c = \frac{m_{01}}{m_{00}} \quad (6)$$

그림 8은 레이블링된 손 영역을 바탕으로 손 영역 무게 중심의 결과 영상을 나타낸다.

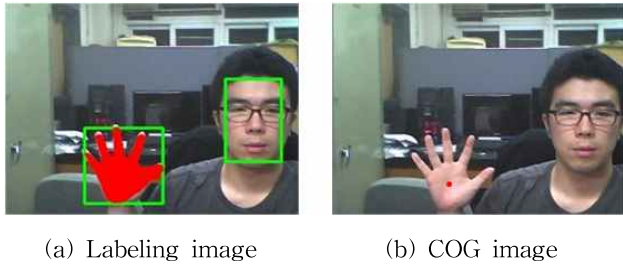


그림 8. 손 영역의 무게 중심  
Fig. 8. COG of hand region

#### 4.2 이진신호를 이용한 손 끝 특징점 추출

손의 무게 중심에 그려진 원은 손 과 맞닿는 부분의 이진신호를 나타내기 위해 생성한다. 원이 손의 레이블링 된 영역까지 확장하면서 손 끝의 이진신호를 탐색한다. 그림 9는 손 영역의 이진신호를 나타낸다.

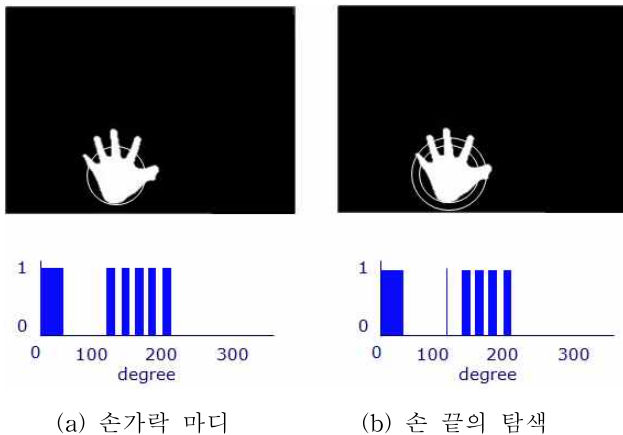


그림 9. 손 영역의 이진신호  
Fig. 9. Binary signal of hand region

그림 9-(a)는 손가락 마디의 이진신호를 나타낸다. 손 끝의 이진신호는 마디보다 각이 좁기 때문에 그림 9-(b)와 같이 나타나게 된다.

#### 4.3 칼만필터를 이용한 특징점 추적

전 단계에서 찾은 손 영역에 대한 특징점은 특히 조명 변화에 따라 평균 10 프레임의 카메라 프레임 수 별로 빈번한 위치 변화가 이루어진다. 여기에 손 움직임의 변화가 동시에 이루어지기 때문에 정확한 손 영역 특징점을 추적하기가 힘들다. 따라서 손 영역 특징점의 위치를 추정하는 알고리즘이 필요하다. 따라서, 본 연구에서는 매 프레임의 측정값을 토대로 예상 지점을 확률적 모델로서 측정할 수 있는 칼만 필터를 이용하였다.

### 5. 실험 및 결과 고찰

본 연구에 사용된 하드웨어 시스템은 IBM 호환 컴퓨터 (CPU : Intel Q6600, RAM : 4GB)이며 1440\*900의 해상도를 갖는 모니터 2대가 사용되었다. 입력 영상은 640\*480의 해상도를 가지는 24bit RGB 영상이고 웹 카메라로 실시간

처리하였다. 실험은 실내에서 진행하였으며 진행과정은 그림 10과 같다.

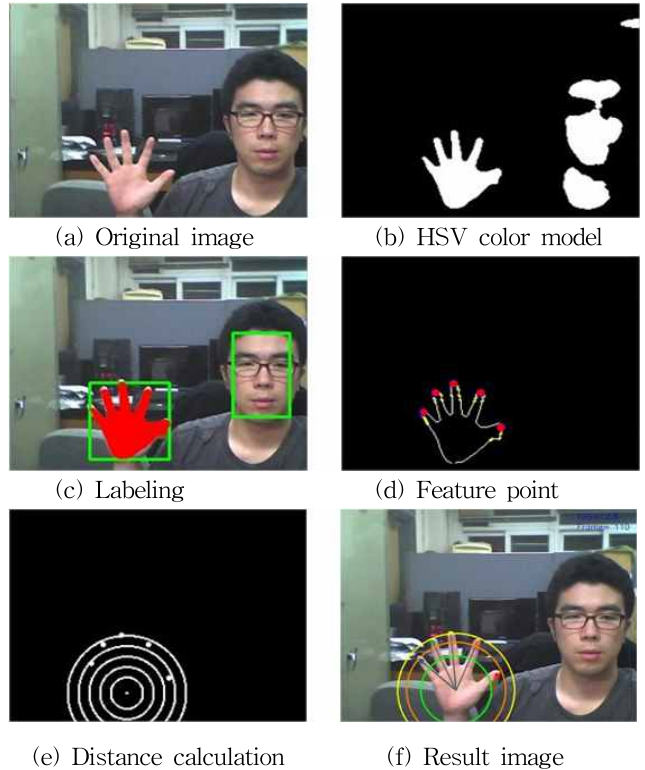


그림 10. 손 영역 특징점 추출 과정  
Fig. 10. Extraction process of feature points of hand region

그림 10에서 (a)는 웹카메라로부터 받은 24bit RGB 영상이고 (b)는 HSV 색상 모델로 변환 후 피부 영역만을 추출한 영상이다. (c)는 레이블링을 통해 손 영역을 추출한 영상이며 (d)는 손 끝 특징점을 추출한 영상이다. (e)는 손 끝 특징점과 중심점의 연결을 통해 거리 계산을 하는 영상이고 (f)는 최종 결과 영상을 나타낸다.

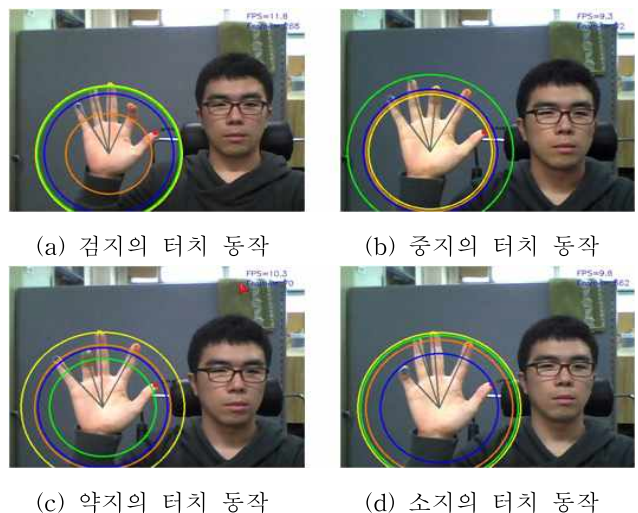


그림 11. 손가락 끝의 클릭 동작 인식 실험  
Fig. 11. Recognition experiments of click operation of fingertips

그림 11은 손가락 각각의 터치 동작을 나타낸다. 검지, 중지, 약지 및 소지는 각각의 터치 명령을 수행할 수 있으며 손가락 끝은 이진신호 판별을 위한 원이 그려져 있다.

그림 12는 비주얼 멀티-터치를 위한 응용프로그램에서의 실험 결과이다. 윈도우7과의 연동을 통해 실험을 진행하였다. 각각의 손가락에는 클릭, 뒤로 가기, 스크롤 다운 및 스크롤 업의 터치 동작이 부여되었다.

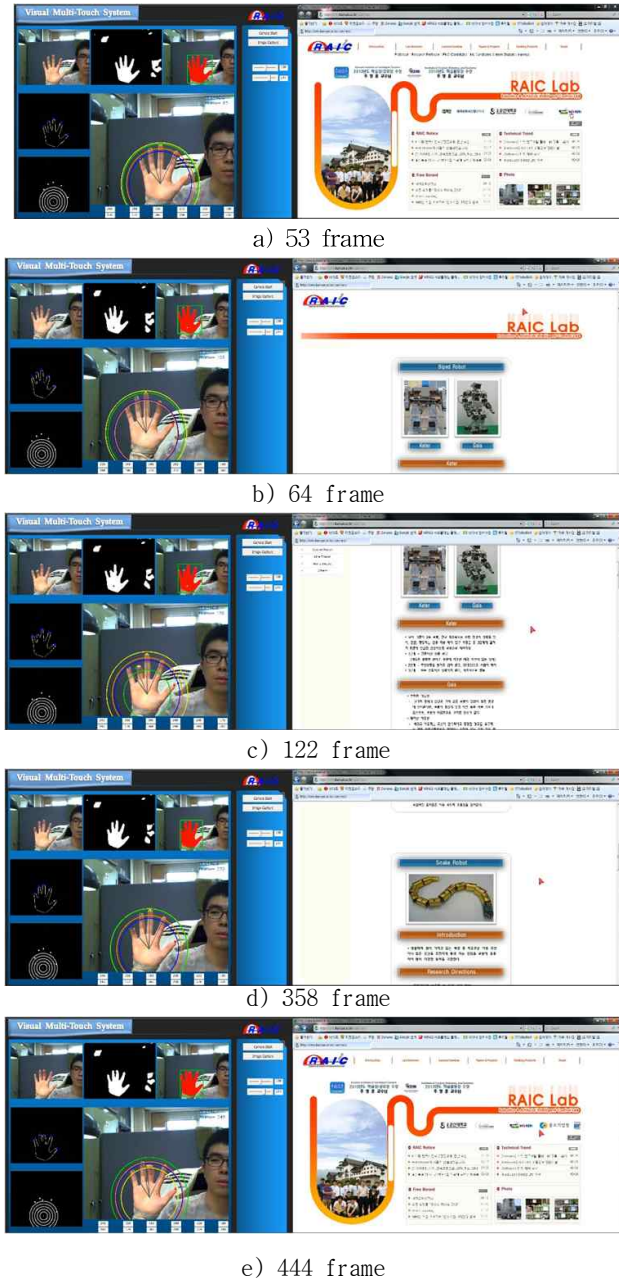


그림 12. 비주얼 멀티-터치 응용프로그램  
Fig. 12. Visual multi-touch application

## 6. 결론

본 논문에서는 무게 중심을 이용한 손 영역 특징점 추출 기법을 제안하였다. 제안된 기법은 영상을 통한 비주얼 멀티-터치 입력 방식에 이용된다. 웹 카메라로부터

입력받은 입력 영상을 조도 변화에 강인한 HSV 색상 모델로 변환하여 피부색 영역을 추출하였다. 레이블링을 통해 최종 손 영역을 추출하고 임계값 이하의 레이블은 제거하였다. 레이블링 과정에서 생긴 잡음은 모폴로지 기법을 이용하여 제거하였다. 전처리 과정을 통해 얻어진 손 영역을 바탕으로 무게 중심을 구하고 원을 그려 손 끝의 이진신호를 판별하고 특징점을 추출하였다. 손 끝 특징점과 손 무게 중심을 연결하여 거리 계산을 통한 터치 동작을 구현하였다. 손가락을 구부릴 때 임계값 이하가 되면 터치 동작으로 인식하게 된다.

## 참 고 문 헌

- [1] R. Cowie, "Emotion Recognition In Human-Computer Interaction," Signal Processing Magazine, IEEE, Vol. 18 pp. 32-80, January 2001.
- [2] T. Huang, and V. Pavloic, "Hand Gesture Modeling, Analysis, and Synthesis," International Workshop on Automatic Face and Gesture-recognition, pp. 73-79, June 1995.
- [3] F. Karray, M. Alemzadeh, J. Saleh, and M. Arab, "Human-Computer Interaction: Overview on State of the Art," Int. Journal on Smart Sensing and Intelligent Systems, Vol. 1, No. 1, pp. 137-159, March 2008.
- [4] Francis K. H. Quek, "Toward a Vision-Based hand Gesture Interface," Conference on Virtual Reality Software and Technology, pp. 23-26, August. 1994.
- [5] J. Han, "Low-Cost Multi-Touch Sensing through Frustrated Total Internal Reflection," In Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology, pp. 115-118, 2005.
- [6] S. Oviatt, "Advances in Robust Multimodal Interface Design," IEEE Computer Graphics and Applications, Vol.23 pp. 62-68, October 2003.
- [7] C. Spence, "Cross-modal Attention and Multisensory Integration: Implications for Multimodal Interface Design," Int. Conference on Multi-modal Interfaces, Vol. 5, November 2003.
- [8] M. Kolsch, R. Bane, T. Hollerer, and M. Turk, "Multimodal Interaction with a Wearable Augmented Reality System," IEEE Computer Graphics and Applications, Vol. 26, No. 3 pp. 62-71, June 2006.
- [9] M. Kim, and J. Yoon, "Implementation of Mobile Game Interface through an Operating Interface Analysis of Touch-screen Devices," Journal of Society of Design Science, No. 81, pp. 231-244, 2009.
- [10] J. Rantala, R. Raisamo, J. Lylykangas, V. Surakka, J. Raisamo, K. Salminen, T. Pakkanen, and A. Hippula, "Methods for Presenting Braille Characters on a Mobile Device with a Touchscreen and Tactile Feedback," IEEE Trans. on Haptics, Vol. 2, No. 1, pp. 28-39, 2009.
- [11] F. Wang, H. Deng, K. Ki, and Q. Ting, "A Study on Image Splicing Algorithm for Large Screen

Multi-touch Technique," Machine Vision and Human-machine Interface, pp. 526-529, April 2010.

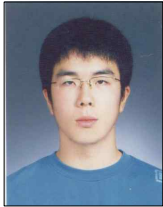
[12] V. I Pavlovic, R. Sharma, and T.S. Huang, "Visual Interpretation of Hand Gestures for Human-computer Interaction: a Review," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 677-695, July 1997.

[13] M. Ishikawa and H. Matsumura, "Recognition of a Hand-gesture Based on Self-organization Using a Data Glove," ICONIP '99, Vol. 2, pp. 739-745, 1999.

[14] S. S. Fels and G. E. Hinton, "Glove-talk: A Neural Network Interface between a Data Glove and a Speech Synthesizer," IEEE Trans. Neural Networks, Vol. 4, No. 1, pp. 2-8, January 1993.

---

저 자 소 개



**서효동(徐孝東)**  
 2010년 군산대 전자정보공학부 졸업.  
 2010년~현재, 동 대학원 전자정보공학부 석사과정.  
 관심분야는 지능형 로봇, 인간-로봇간 상호작용, 지능제어

Phone : 063-469-4706  
 E-mail : [kinelian@kunsan.ac.kr](mailto:kinelian@kunsan.ac.kr)



**주영훈(周永焄)**  
 제 21권 제 5호 참조