

윈도우 프로그램 제어를 위한 실시간 손 형상 인식

위승정, 김종민, 양환석, 이웅기*

조선대학교 대학원 수학·전산통계학부

mrjjoung@hanmail.net, badhack@image.chosun.ac.kr,

lee@image.chosun.ac.kr

Real Time Hand Shape Recognition for Window Program Control

Seung-Jung Wi, Jong-Min Kim, Hwan-Seok Yang, Woong-Ki Lee*

Division Mathematics and Computer Science and Statistic

Graduate School, Chosun. Univ.

* Division Mathematics and Computer Science and Statistic College
of Natural Sciences, Chosun. Univ.

요 약

본 연구는 손의 형상을 복잡한 배경환경에서 손 영역을 안정적으로 검출, 인식하여 윈도우 플레이어의 기능을 제어하는 시스템을 제안하였다. 손의 형상이 매우 복잡하기 때문에 2차원 형상의 불변량에 해당하는 에지의 방향성 히스토그램을 이용하여 인식을 행한다. 이 방법은 복잡한 배경에서 피부색을 지닌 손 영역이 정확히 추출되며 손 형상을 인식하는데 있어서 수행속도가 빠르고 조명변화에 덜 민감하기 때문에 실시간 손 형상 인식에 적합하다. 본 논문에서 제안한 방법을 윈도우 플레이어 제어에 적용한 결과 안정적으로 제어 할 수 있었다.

1. 서론

손은 인간에게 있어서 가장 중요한 도구의 하나로 그 동작과 모양이 의미하는 바를 인식하는 것은 정보 시스템과 인간 사이의 자연스러운 인터페이스를 실현할 수 있다는 점에서 오랜 기간동안 많은 사람들의 연구테마로 다루어져 왔다. 그러나 복잡한 환경으로부터 사람의 손 형상만을 정확히 추출하고 인식하여 의미 있는 입력정보로 사용하는 것은 어려운 일이다. 기본적으로 정확한 손 영역 분할을 위한 알고리즘이 필수이며 손 형상의 분석과 인식을 위한 빠른 처리 속도 또한 고려해야할 요소이다. 기존의 연구들은 크게 손 형상을 분석하기 위해 Data Glove를 사용하는 방법[1]과 3차원 모델을 사용하거나 스테레오 비전을 이용하여 추출된 손의 3차원 기하학적인 정보를 분석하는 방법[2], 손의 2차원 패

턴 정보를 분석하는 방법[3]으로 분류할 수 있다. Data glove를 사용하는 방법은 정밀도가 낮고 전체 시스템과 접속하기 위해 반드시 연결선이 필요하다는 점에서 자연스러운 인터페이스 구축에는 많은 문제점을 안고 있다. 3차원 기하 정보를 이용하는 방법은 스테레오 영상을 이용한 3차원 정보추출과 복잡한 손의 3차원 운동 모델을 구축하기 위한 많은 계산 량이 필요하므로 고성능의 하드웨어가 필요하다. 단점이 있다. 2차원 정보를 이용하는 방법들은 사람의 피부색의 컬러분포를 사전 학습과정에서 계산하여 입력 영상 내 해당 피부 컬러 분포에 해당되는 영역만을 추출하는 방법을 적용하였으나 대부분 단일 배경이거나 조명 변화가 거의 없다고 가정하였

다.

본 논문에서는 손의 형상 변화로 인한 손 영역의 국부적인 조명변화와 복잡한 배경으로부터 손 영역을 추출한 후 화소간의 휘도 경도를 구해 에지 방향성 히스토그램을 만들고 이 히스토그램을 평활화 하여 얻어진 특징 벡터를 형상 구분의 척도로 사용하였다. 본 논문에서 이 알고리즘을 이용하여 실시간 손 형상을 윈도우 플레이어 제어를 위한 명령으로 사용하였다. 우리가 구현한 시스템은 수행 속도가 빠르고 복잡한 배경에서 안정적으로 손 형상을 추출하여 인식한다. 또한 PC와 USB 카메라로 구현되므로, 일반 사용자들도 쉽게 사용할 수 있는 저 비용 시스템을 구축하였다.

2. 전체 시스템 개요

본 연구에서는 피부색의 컬러 분포를 이용한 손 영역 추출과 에지 방향성 히스토그램 알고리즘을 이용하여 손 형상을 인식하여 윈도우 플레이어 제어에 적용하여 구현했다. 제안된 알고리즘의 전체 구성도를 그림 1에 나타내었다.

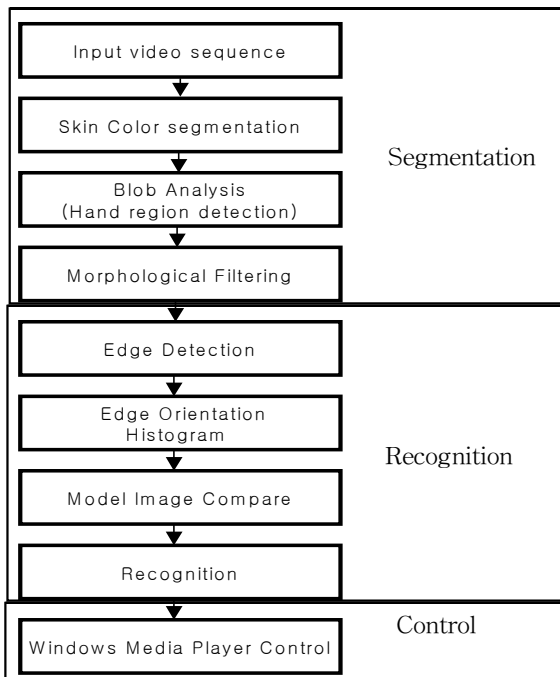


그림 1 시스템 전체 구성도

입력 영상이 들어오면 각 단계에서 제안된 알고리즘의 연산이 수행되며, 최종적으로 손 영역이 추출된 영상을 얻게된다. 추출된 영상을 모델 영상과 비교하여 손 형상에 대한 어휘를 획득한 후 윈도우 플레이어를 제어한다. 각 단계에서의 결과 영상은 순차적

으로 다음 입력으로 사용되게 된다.

3.1 Skin Color filtering

칼라 정보만을 이용하여 손 영역을 검출할 경우에는 살색 계통의 물체가 배경영역에 존재하여 함께 검출될 수 있다. 따라서 보다 정확하고 빠른 검출을 위하여 본 논문에서는 다음과 같은 방법을 제안했다. 이 방법은 간단하고 빠르게 RGB공간 안에서 살색을 추출 할 수 있다. RGB 데이터의 영상이 입력되면 살색은 RED 값이 다른 GREEN이나 BLUE보다 크다고 가정했다. 식1과 같이 피부색에 해당하는 부분은 그 픽셀을 1로 세팅하여 주고 이 이외에 해당하는 픽셀의 위치에서는 살색이 아닌 영역으로 판정하여 0의 값으로 세팅을 하게 된다.

$$I_2(x,y) = \sum_{x,y} \begin{cases} 1 & (\text{RED} \geq \text{BLUE}) \&\& (\text{RED} \geq \text{GREEN}) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

3.2 손 영역 추출

매우 정확한 칼라 필터라도 영상에서 손만을 추출하기는 쉽지 않다. 그림 2에서 보는 바와 같이 얼굴과 손에서 손만을 추출하기 위해서는 본 논문에서는 효율적인 방법을 사용했다.

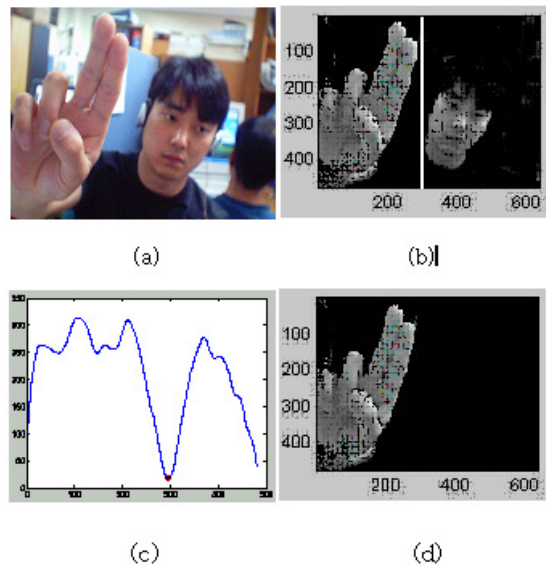


그림 2 (a) 640*480 입력영상, (b) 칼라 필터 적용 (c) 손과 얼굴 분포와 분할 지점 (d) 손 영역 추출

우리의 방법은 칼라 필터가 적용되고 남아 있는 픽셀들의 대부분이 얼굴이나 손에 속해 있기 때문에 횡 좌표 X축으로 살색 픽셀들의 수를 계산한다. 두 영역에서 손만을 분리한다(labeling). 보통 손이 머리

보다 카메라에 더 크게 보임에 따라, 필터 된 픽셀들의 가장 높은 계수를 가진 쪽을 손으로 귀착시킨다. 이 알고리즘은 다이내믹 상황에 매우 좋은 결과를 산출했다.

4. 손 형상 인식

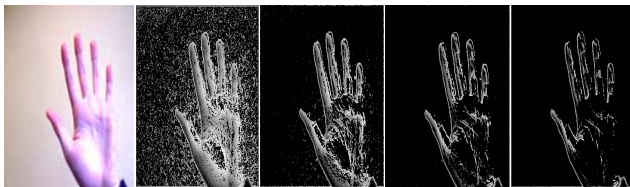
영상이 지닌 밝기의 산술적 차이만을 비교하여 손과 같이 복잡한 물체의 형상을 인식한다는 것은 매우 어려운 일이다. 본 논문에서는 실시간으로 손 형상을 인식하기 위해 우리는 간단하고 빠른 알고리즘(에지 방향성 히스토그램)을 선택했다.

4.1 에지 추출

에지 방향성 히스토그램을 생성하기 위해서는 먼저 에지를 추출해야 하는데 에지를 추출하는 방법은 식 2를 이용하여 에지를 구한다. 영상으로부터 추출한 에지를 가지고 에지의 방향성을 이용하여 에지 방향성 히스토그램을 생성하면 이것을 영상의 특징정보로 사용할 수 있다.

$$magnitude = \sqrt{s_x^2(i, j) + s_y^2(i, j)} \quad (2)$$

위에 식에서 생성한 에지 값은 모든 픽셀에 대하여 생성된다. 따라서 실제로 에지가 아닌 부분에서도 에지 값을 갖게 되므로 특정 픽셀이 에지인지 아닌지를 판별하는 과정이 필요하다. 특정 픽셀이 에지인지 아닌지를 판별하기 위해서 임계치에 따라 에지로 판정된 픽셀만을 대상으로 에지 방향성 히스토그램을 생성한다. 각 임계치별 에지영상은 그림 3과 같다.



정규화 영상 T=1 T=2 T=3 T=4
그림 4 임계값(T)에 따른 에지영상

4.2 에지 방향성 히스토그램 생성

임계치로 얻어진 에지영상을 16 단계의 방향으로 에지 방향성 히스토그램을 생성한다.(식 3)

$$direction = \tan^{-1} \frac{s_y(i, j)}{s_x(i, j)} * \frac{36}{\pi} \quad (3)$$

$$\theta(x, y) \arctan[I(x, y) - I(x-1, y), I(x, y) - I(x, y-1)]$$

식 3으로부터 얻어진 픽셀들을 36저장소에 매핑하고 픽셀들의 수로부터 얻어진 방향성 히스토그램을 모델영상과 비교하여 실시간으로 손 형상을 인식하게 된다. 그림 5는 방향성 히스토그램을 생성한 모습이다.

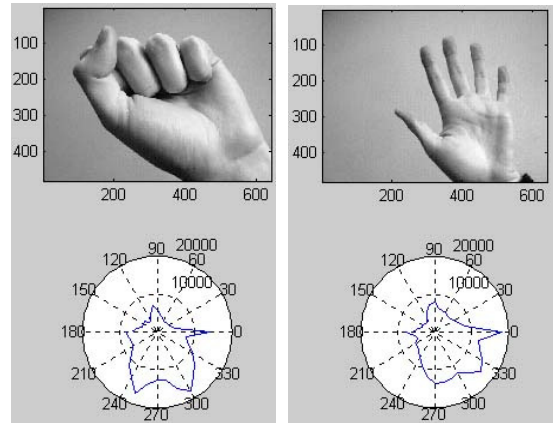


그림 5 방향성 히스토그램

5. 실험 및 결론

실험은 범용 USB 카메라(초당30프레임)를 장착한 펜티엄 4 1.2MHz 메모리 256MB를 가진 PC에서 수행하였으며, 프로그램은 Visual C++사용하여 구현하였다. 윈도우 플레이어를 제어하기 위해서 사용한 영상은 6가지의 의미를 부여한 손 형상을 사용하였다. 각 영상의 크기는 640*480사이즈를 실시간을 제어한다. 그림 6과 같이 실시간으로 손 형상을 인식 후 그 결과를 윈도우 미디어 플레이어 핸들을 얻어서 윈도우 플레이어 제어 메시지로 사용하였다. 표 1에서는 실험에 사용된 영상과 인식 결과를 수치로 나타내었다.

본 논문에서 제안한 컬러 정보를 이용한 실시간 손 형상 인식을 이용한 윈도우 플레이어 제어는 수행속도가 매우 빠르며, 조명 변화와 복잡한 배경에 무관하게 안정적으로 손 영역을 검출 및 인식할 수 있어서 원활한 윈도우 플레이어 제어가 가능하였다. 그러나 단지 영상의 컬러 정보만을 이용하기 때문에 배경에 살색 영역이 많이 포함되거나 얼굴과 손이 겹쳐서 나타날 경우 손을 잘 찾지 못하는 경우가 발생하였고, 손의 방향이 많이 틀어진 경우 그 손 형상을 다른 손 형상으로 인식하는 오류를 범하기도 하였다. 향후 실시간 손 형상을 보다 정확하고 안정

적인 인식을 위해서는 손 방향의 변화를 해결할 수 있는 연구가 계속 수행되어야 한다.

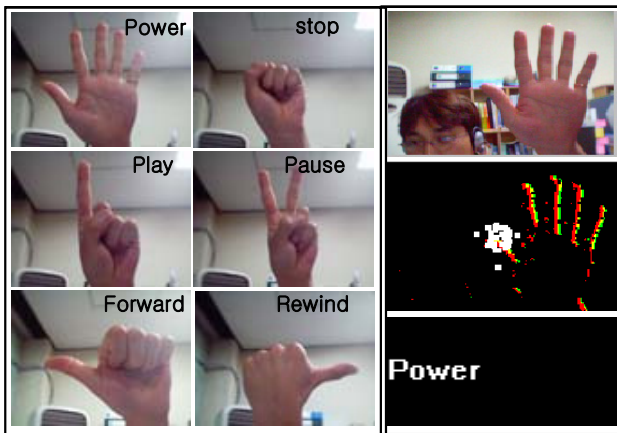


그림 6 (a) 윈도우 플레이어제어를 위한 입력 영상

(b) 실시간 손 형상 인식 결과

표 1. 실험 결과

영상의 종류	영상의 개수	성공률(%)	실패율(%)
단순한 영상	12	96.2	0.38
복잡한 영상	12	88.5	12.5

6. 참고문헌

- [1] J. Ohya and Y. Kitamura, etc, "Real-Time Reproduction of 3D Human Images in Virtual Space Teleconferencing" in Proc. of '93 IEEE Virtual Reality Annual Int. Symp. pp.408-414, 1993.
- [2] J. Segen and S. Kumar, "Shadow Gestures: 3D Hand Pose Estimation Using a Single Camera," CVPR99, vol. 1, pp. 479-485, Fort Collins, Colorado, June, 23-25, 1999.
- [3] K. Imagawa, S. Lu and S. Igi, "Color-Based Hand Tracking System for Sign Language Recognition," Proceedings of International Conference on Automatic Face and Gesture Recognition, Killington, 1996.
- [4] S. Ahmad, "A Usable Real-Time 3D Hand Tracker," Conference Record of the Asilomar Conference on Signals, System and Computers, pp. 1257-1261, 1994.
- [5] V.Pavlovic, R. Sharma, and T.Huang, "Visual

Interpretation of Hand Gestures for Human-Computer Interaction: A Review", IEEE Trans. on PAMI, Vol.19, No.7, pp.677-695, 1997.

[6] William T. Freeman and Michael Roth, "Orientation histograms for hand gesture recognition," Intl. Workshop on Automatic Face and Gesture Recognition, IEEE Computer Society, pp. 296-301, June 1995, Zurich, Switzerland

[7] Tony Heap and Ferdinando Samaria." Real-time hand tracking and gesture recognition using Smart Snakes," Proc. Interface to Human and Virtual Worlds, June 1995, Montpellier, France.

[8] D Chi, King N. Ngan "Face Segmentation Using Skin-Color Reference map in Videophone Applications", IEEE Transactions on Circuits and Systems for Vidoe Technology, vol.9, June 1999

[9] 이칠우, "복수 모델영상의 상위도 패턴을 이용한 손형상 인식", 멀티미디어 학회 논문지 제2권, 4호 pp. 400-408, December 1999