

피부색 검출 및 특징점 추적을 통한 원거리 손 모션 제스처 인식

윤종현, 김성영
국립금오공과대학교 컴퓨터공학과
e-mail : yjh1911@kumoh.ac.kr

Hand Motion Gesture Recognition at A Distance with Skin-color Detection and Feature Points Tracking

Jong-Hyun Yun, Sung-Young Kim
Dept of Computer Engineering, Kumoh National Institute of Technology

요 약

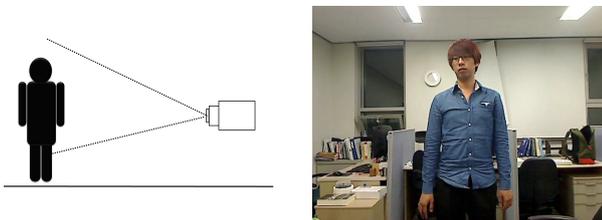
본 논문에서는 손 모션에 대하여 피부색 검출을 기반으로 전역적인 모션을 추적하고 모션 벡터를 생성하여 제스처를 인식하는 방법을 제안한다. 추적을 위하여 Shi-Tomasi 특징점 검출 방법과 Lucas-Kanade 옵티컬 플로우 추정 방법을 사용한다. 손 모션을 추적하는 경우 손의 모양이 다양하게 변화하므로 초기에 검출된 특징점을 계속적으로 추적하는 일반적인 방법으로는 손의 모션을 제대로 추적할 수 없다. 이에 본 논문에서는 프레임마다 새로운 특징점을 검출한 후 옵티컬 플로우를 추정하고 이상치(outlier)를 제거하여 손 모양의 변화에도 추적을 통한 모션 벡터 생성이 가능하도록 한다. 모션 벡터들로 인공 신경망을 사용한 판별 과정을 수행하여 최종적으로 손 모션 제스처에 대한 인식이 가능하도록 한다.

1. 서론

최근 컴퓨터와 사람간의 자연스러운 인터페이스에 대한 연구가 각 분야에서 활발하며, 그 중에서도 마커 또는 센서, 모션 장갑 등의 특수한 장치가 필요치 않은 카메라를 기반으로 하는 손의 모양 및 모션을 사용한 비전 기반 인터페이스가 각광받고 있다.

본 논문에서 제안하는 방법은 원거리에서 손의 모션에 대한 제스처 인식을 통하여 전자칠판 및 스마트 TV 원거리 원격 제어 응용 등의 기반이 된다.

2. 시스템 환경 및 구조



(그림 1) 카메라와 사용자의 위치 및 입력 영상

제안하는 시스템의 환경은 (그림 1)과 같이 카메라가 사람 키 보다 약간 낮은 높이에 설치되어 있고, 사용자가 원거리에서 카메라 방향을 바라보고 있어 카메라 입력영상에 사람의 양손과 얼굴 및 상반신이 나타나는 상황이다. 전체적인 시스템 구조는 아래 (그림 2)처럼 입력 영상에서 피부색을 검출하고, 피부색 영역에 대한 특징점을 추적하

여 모션 벡터를 생성한 후 최종적으로 손 모션에 대한 제스처를 인식한다.



(그림 2) 시스템 구조

3. 모션 벡터 생성 및 제스처 인식

3.1 피부색 모델링을 통한 피부색 검출



(그림 3) 일반적인 피부색 검출 결과

HSV 컬러공간에서 일반적인 피부색의 범위를 이용한 픽셀기반 피부색 검출[1] 방법은 (그림 3)과 같이 배경에서 피부가 아닌 픽셀들이 같이 검출되게 된다. 그러므로 실제 피부색만을 검출하기 위하여 사용자에게 초기에 (그림 4)와 같이 손 영상을 등록받아서 일반적인 피부색 검출을 통해 검출되는 픽셀들에 대해 H-S 2차원 히스토그램으로 피부색을 모델링한다.



(그림 4) 손 영상 및 H-S 2차원 히스토그램

생성된 피부색 모델로부터 입력 영상 픽셀들을 임계값 Θ 를 기준으로 식 (1)을 만족할 때 피부색으로 분류한다.

$$p(HS | skin) > \Theta \quad (1)$$

피부색으로 분류된 영상에서 잡음을 제거하기 위하여 (그림 5)와 같이 외곽선을 검출하고 일정 길이 외곽선 이하의 객체들을 제거한다.



(그림 5) 피부색 검출결과, 잡음영상, 잡음제거 결과

3.2 특징점 추적을 통한 모션 벡터 생성

손의 모션 제스처를 취할 시 손은 움직이고 얼굴과 나머지 손은 비교적 적게 움직이므로 (그림 6)처럼 영상에서 피부색이 검출되는 부분 전체에 대해서 전역적인 모션을 검출한다. 추적을 위해서 Shi-Tomasi 특징점 검출 방법 [2]과 영상 피라미드를 사용한 Lucas-Kanade 옵티컬 플로우 추정 방법 [3]을 사용한다.

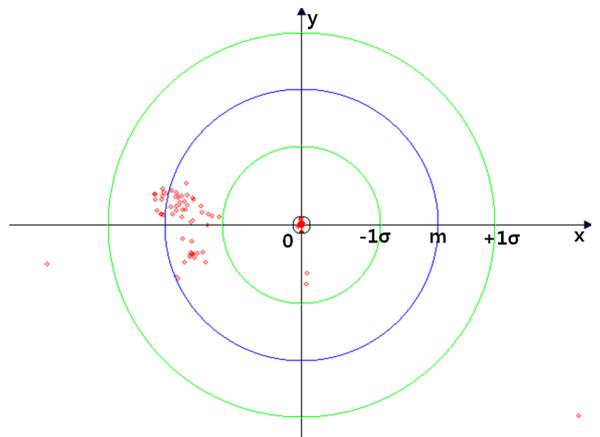


(그림 6) 피부색 영상에 대한 전역적인 모션 추적



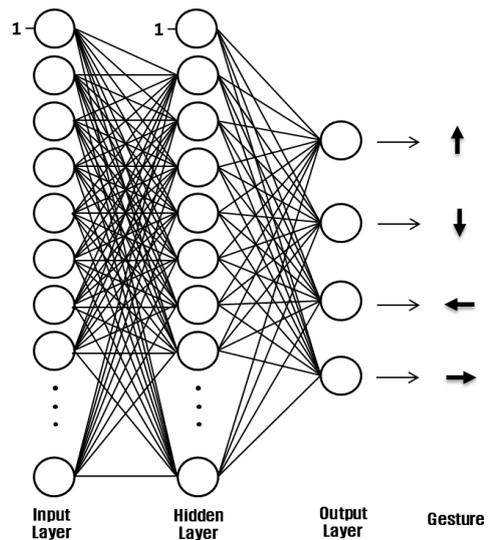
(그림 7) 손 모션시 다양한 손의 형태

일반적인 최소 옵티컬 플로우 추정시 초기 추출된 특징점을 계속 사용하는 방법은 위 (그림 7)과 같이 자연스러운 손 모션 시 손의 모양이 다양하게 변화하기 때문에 특징점들도 다르게 나타나게 되므로 사용할 수 없다. 이에 본 논문에서는 프레임마다 새로운 특징을 추출하고 다음 프레임에서 이를 추적하여 각 특징점들에 대한 변위를 얻는다. 이 때 정확한 모션벡터 생성을 위하여 (그림 8)과 같이 나타나는 이상치(outlier)들을 변위들의 평균값에서 $-1\sigma \sim +1\sigma$ (σ : 표준편차) 영역의 값들만을 사용하여 제거한다. 선택된 영역의 변위들의 평균값으로 모션벡터의 크기를 결정하고, 분산되어 있는 변위 값들을 가장 잘 표현하는 방향을 결정하기 위하여 주성분분석(Principal Component Analysis)[4]을 이용하여 주방향을 결정해 각 프레임에 대한 모션벡터를 생성하게 된다.



(그림 8) 특징점들의 x,y 방향 변위값

3.3 인공 신경망 기반 제스처 인식



(그림 9) 다층 퍼셉트론의 구조

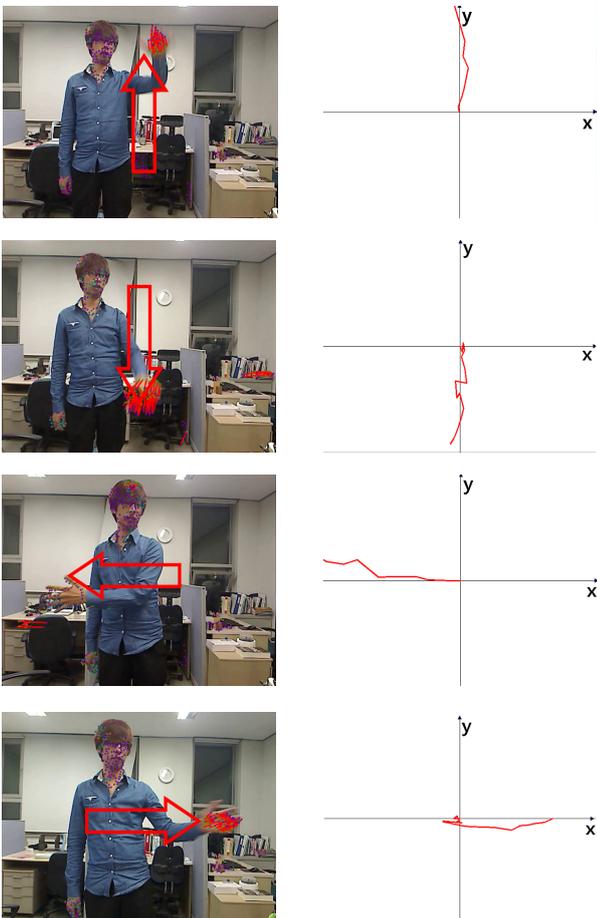
상, 하, 좌, 우 제스처에 대하여 다층 퍼셉트론은 (그림 9)와 같이 입력 뉴런 11개, 은닉 뉴런 11개, 각 제스처에

대한 출력 뉴런 4개로 구성했으며 활성화 함수로 시그모이드 함수(sigmoid function)를 사용하고 오류 역전파 학습 알고리즘(error back propagation learning algorithm)을 사용하여 학습하였다. 출력 층의 결과를 토대로 최종 손 모션에 대한 제스처를 판별한다.

4. 구현 및 실험결과

Intel Core 2 Duo E7500 2.94Ghz, 4GB RAM의 PC환경에서 Logitech Webcam Pro 9000 으로부터 640x480 해상도의 영상을 입력받아 사용하고, C++언어와 Visual C++ Compiler를 사용하여 구현하였다.

실험결과 아래 (그림 10)과 같이 제스처의 손 모션을 추적하여 모션 벡터를 생성한 후, 인공 신경망을 사용하여 손 모션에 대한 제스처를 인식할 수 있었으며, 피부색 검출 및 모션 벡터 생성 과정에서 초당 15프레임의 속도로 실시간 구현이 가능하였다.



(그림 10) 손 모션 추적 결과 생성된 모션 벡터들

5. 결론

본 논문에서는 피부색 검출 및 특징점 추적을 통한 원거리에서 손의 모션에 대한 제스처 인식 방법을 제안하였다. 그러나 사용한 피부색 기반 방법은 색상변화에 민감하고 아주 유사한 색상을 가지는 배경이 있을 경우 구분이

불가능한 단점을 가지고 있으며, 손 모션 추적 시 전역적인 모션을 추적하기 때문에 얼굴이 움직이는 경우 및 양손을 사용한 모션 등의 다양한 제스처에 대해서는 인식할 수 없다. 이러한 문제점의 개선을 위해서는 지속적인 연구가 필요하다.

Acknowledgment

본 논문은 중소기업청에서 지원하는 2012년도 산학연 공동기술개발사업(No. C0025247)의 연구수행으로 인한 결과물임을 밝힙니다.

참고문헌

- [1] Oliveira, V.A., Conci, A. "Skin Detection using HSV color space" In: Sibgrapi, Nr. 12. 2009.
- [2] Shi J., Tomasi C. "Good features to track" In Proc. IEEE Conf. on Computer Vision and Pattern Recognition(CVPR), pp. 594-600, June 1994.
- [3] Lucas B.D., Kanade T. "An iterative image registration technique with an application to stereo vision" In Proc. Image Understanding Workshop, pp. 121-130, 1981.
- [4] Jolliffe I.T. "Principal Component Analysis" Springer-Verlag, New York, 1986.