

비전기반 손 제스처 인식을 통한 로봇 컨트롤¹

김대수, 강행봉

가톨릭대학교 컴퓨터 공학과

e-mail : zootguy81@catholic.ac.kr, hbkang@catholic.ac.kr

Robot Control using Vision based Hand Gesture Recognition

Dae-Soo Kim and Hang-Bong Kang

Dept. of Computer Engineering, Catholic University of Korea

요 약

본 논문에서는 로봇 컨트롤 시스템을 위해 입력 받은 영상부터 몇 가지의 손 제스처를 인식하는 비전기반 손 제스처 인식 방법을 제안한다. 로봇으로부터 입력 받은 이미지는 로봇의 위치, 주변환경, 조명 등 여러 요인에 따라 다양하게 존재한다. 본 논문은 다양한 환경에서 입력되는 영상으로부터 시스템이 로봇 컨트롤을 위해 미리 지정한 몇 가지 제스처를 인식하도록 한다. 먼저 이미지 조명 변화에 강한 손 제스처 인식을 위하여 레티넥스 이미지 정규화를 적용한 후, YCrCb 공간 상에서 입력된 영상에서 손 영역을 검출 후 위치를 추정한다. 인식된 손 영역에서 특징벡터를 추출함으로써 입력 영상내의 존재할 수 있는 손의 크기나 손의 회전각도 등에 상관없이 필요로 하는 제스처를 인식하도록 한다. 제안된 제스처 인식 결과는 로봇컨트롤을 위한 기존의 제스처인식과 비교하여 성능을 측정하였다.

1. 서 론

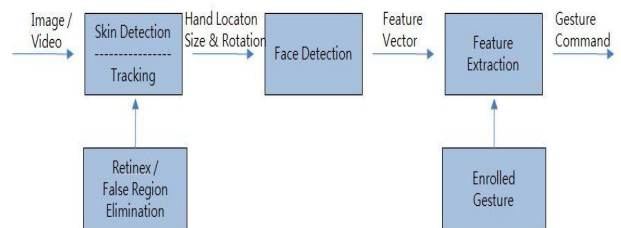
비전기반의 Hand Gesture Recognition 은 게임, 로봇 컨트롤, AR(Augmented Reality) 등과 같이 많은 분야에서 응용이 가능하며 최근에 매우 활발하게 연구되고 있다. 로봇 컨트롤을 위한 Hand Gesture Recognition은 배경이나 조명 등의 영향 뿐만 아니라 손의 크기나 회전각도 및 위치변화에서도 강건해야 한다. 또한 실시간 처리를 위해 효과적이고 빠른 알고리즘이 요구된다. 본 논문에서는 로봇 컨트롤 시스템을 위해 Video로 입력 받은 영상부터 미리 지정된 Hand Gesture 를 인식하는 비전기반의 Hand Gesture Recognition 방법을 제안한다.

손 제스처 인식에 대한 많은 연구가 최근 수년간 활발히 이루어져 왔다. 과거에는 손의 움직임을 측정하기 위해 손가락 끝 혹은 손 영역에 마커를 부착하거나 정적인 이미지로부터 제한된 환경에서의 손 제스처 인식이 이루어져 왔으나, 최근에 이르러서는 비전기반의 손 제스처 인식으로 환경의 변화에서도 둔감한 손 제스처 인식의 연구가 활발해 지고 있다.

손 제스처를 인식하는 가장 널리 알려진 방법은 템플릿 매칭[5]과 PCA[4] 알고리즘이 있다. 하지만 이 두 가지 방법은 손의 크기, 손의 회전각도등에 변경이 있을 경우 손 제스처를 올바르게 인식하기 어렵다. 이러한 손의 크기, 회전 각도등의 문제점을 보완하기 위해 최근 Malima 는 손의 중심점을 계산한 후 특징벡터를 추출하여 손 제스처를 인식하는 방법을 제안하였다[3]. 하지만, 이 방법은 조명이 편중되어 있을 경우 손 영역을 검출하지 못할 뿐 아니라, 손 영역을 제외한 팔목 부분과 같은 피부영역이 많이 노출될 경우는 올바르게 인식되

기 어렵다.

본 논문은 레티넥스 이미지 정규화를 통하여 이미지 조명 변화를 보정하여 입력된 영상의 YCrCb 공간상의 스킨영역을 보다 정확하게 추출한다. 또한, 기존연구에서 문제가 되었던 피부의 많은 노출에도 정확한 손 제스처 인식을 위하여 스킨 영역에 세션화 알고리즘을 적용하여 새로운 손 중심점(COG) 계산 알고리즘을 제안한다. 제안된 기법은 트레이닝과 같은 사전 작업 및 데이터가 필요 없는 최소한의 계산으로 효율적인 손 제스처 인식을 수행하며, 실험을 통하여 기존의 방법보다 좋은 결과를 얻음을 보인다.



[그림 1] 손 제스처 인식 시스템

2. 제스처 인식

2.1 레티넥스 이미지 정규화를 통한 스킨 영역 추출

¹ 본 연구는 문화관광부 및 한국문화콘텐츠진흥원의 지역문화산업연구센터(CRC)지원사업의 연구결과로 수행되었음.

레티넥스는 입력된 값과 입력된 값의 이웃되는 값들의 로그 함수의 차이를 통해 너무 밝아진 이미지 또는 너무 어두워진 이미지를 보정하기 위하여 사용된다. SSR(Single-Scale Retinex)는 식(1)과 같이 구할 수 있다. 그러나 SSR에서의 로그 그래프는 어두운 영역은 보정할 수 있지만, 피부가 빛에 반사되어 너무 밝아진 부분은 보정할 수 없다. 이를 보완하기 위해서 로그 그래프를 $x = y$ 에 대칭 시킨 대칭 로그 그래프와 결합하여 식 (2)를 얻어낸다.

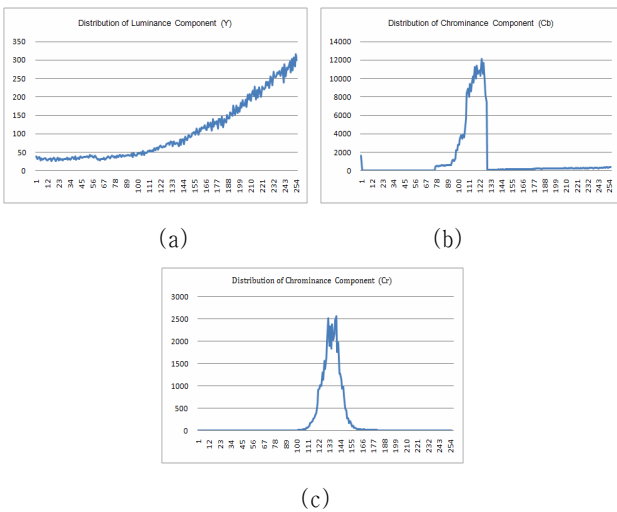
식(2)를 통하여 그래프 처리를 한 후 이미지 결합을 위하여 식(3)과 같이 적용하면 새로운 영상결과를 얻는다[2].

$$R(x, y) = \log(I(x, y)) - \log(F \otimes I(x, y)) \quad (1)$$

$$\hat{R}(x, y) = L(I(x, y)) - L(F \otimes I(x, y)) \quad (2)$$

$$R_{new}(x, y) = \alpha \hat{R}(x, y) + (1 - \alpha)L(I(x, y)) \quad (3)$$

여기서 $I(x, y)$ 는 입력 이미지이고 F 는 가우시안 필터이다. 이때 L 은 대칭되는 로그 그래프를 결합한 그래프이다. 이렇게 처리된 이미지는 밝기에 대하여 보정된 결과를 얻을 수 있다.



[그림 2] 스킨 영역에 대한 YCrCb 분포 그래프, a) b) c) 각각 Y, Cr, Cb 에 대한 스킨 영역 분포 그래프

보정된 이미지를 통하여 피부를 인식하기 위해 RGB를 비롯하여 YCrCb, YIQ, HSV 등의 공간상 영역을 고려할 수 있다. 그러나 실험을 통하여 YCrCb 공간상의 영역에서 가장 높은 스킨 인식률을 보였다. 따라서, 본 논문에서는 YCrCb 영역을 고려하여 스킨 영역 추출하였으며, [그림2]은 200 개의 샘플 데이터의 실험을 통하여 스킨 영역에 대한 YCrCb의 분포도 그래프를 보여주고 있다. 실험에서는 식(4)와 같은 임계구간을 도출해 낼 수 있었으며, 레티넥스 처리된 이미지의 스킨영역을 추출할 수 있었다. 스킨 영역으로 추출된 이미지는 이진화 이미지로 처리한다.

$$0 < Y < 255, 100 < Cr < 120, 140 < Cb < 170 \quad (4)$$

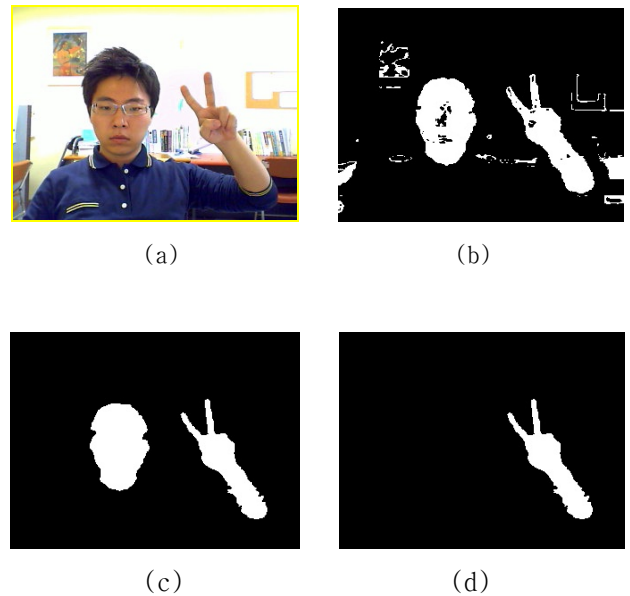
이러한 방법은 기존의 많은 연산을 필요로 하였던 트레이닝 연산이나 매칭을 위한 데이터가 필요 없으므로 효율적인 계산 속도를 가지기 때문에 로봇에 적용할 수 있다.

2.2 False-Region Elimination

2.1을 거쳐 검출된 영상의 영역들은 스킨 영역과 비슷한 색상을 가진 다른 영역이 스킨 영역으로 판별 될수 있기 때문에 이러한 영역은 스킨 영역에서 제외시킬 필요가 있다. 이를 위해 스킨 영역으로 판별된 영역 중 전체 스킨 영역의 20%가 되지 못하는 영역은 스킨 영역이 아닌 것으로 간주한다. 또한, 스킨으로 인식 되더라도 스킨영역 내에 noise가 생길 수 있다. 이를 제거하기 위하여 스킨 영역 안에 스킨이 아니라고 인식된 작은 부분[그림3](b)은 스킨으로 간주한다.

마지막으로 조명등의 의해 손 영역의 에지 부분에서 매끄럽지 못하고 일그러질 수 있으므로 이를 위해 이미지를 Gaussian Smoothing 처리를 거치도록 한다.

[그림3]은 이러한 과정을 통하여 수행한 결과를 보여준다.



[그림 3] 입력이미지에 대한 피부영역과 손 영역 추출

a) 입력 영상, b) 스킨 영역 추출, c) False Region Elimination, d) 얼굴 영역 제거

2.3 Haar-Like Feature

2.2 에서 얻어진 스킨 영역에서 손이 아닌 얼굴 영역을 제거한다. 이를 위해서 Adaboost 방법을 사용한다. Adaboost는 환경 요소에 강인하면서 로봇에 적용될 수 있는 빠른 연산이 가능한 얼굴 검출 방법이다[1].

Adaboost는 여러 개의 약한 분류기를 이용하여 강인한 분류기를 생성하는 기법으로 Integral Image를 이용하여 빠르게 분류할 수 있다. 이미지 (x) 가 주어졌을 때, 강인한 분류기 $H_M(x)$ 는 M개의 약한 분류기 h_m 의 선형 조합으로 식(5)과 같이 생성할 수 있다.

$$H_M(x) = \text{sign}\left(\sum_{m=1}^M h_m(x)\right) \quad (5)$$

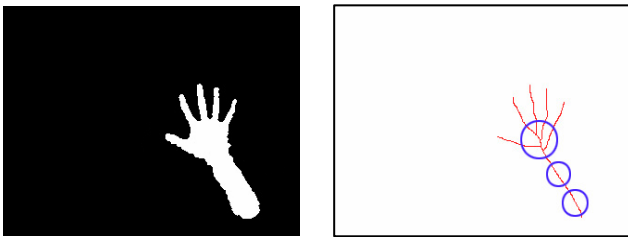
검출된 얼굴영역은 손을 인식하기 위하여 제거 되어진다.

2.4 특징 추출 및 인식

2.3 방법을 통한 입력 영상[그림 3](d)는 세선화 과정을 거쳐 손 영역에 대한 중심점(COG : Center of Gravity)를 구하게 되며, COG를 중심으로 얻어진 특징벡터를 기준으로 손에 대한 제스처를 인식하게 된다.

스킨 이라고 인식 되어진 영역에 대해서 손 부분은 팔 영역에 비하여 그 너비가 넓기 때문에 더 큰 원을 그릴 수 있다. 따라서, 세선화를 거친 입력 이미지는 선을 따라 원을 그리게 되고, 스킨 영역 안에 그릴 수 있는 가장 큰 원은 손 영역에

위치하게 되므로 가장 큰 원의 중심이 손에 대한 중심점(COG)이 된다.[그림 4]



[그림 4] 손 영역에 대한 COG 추출

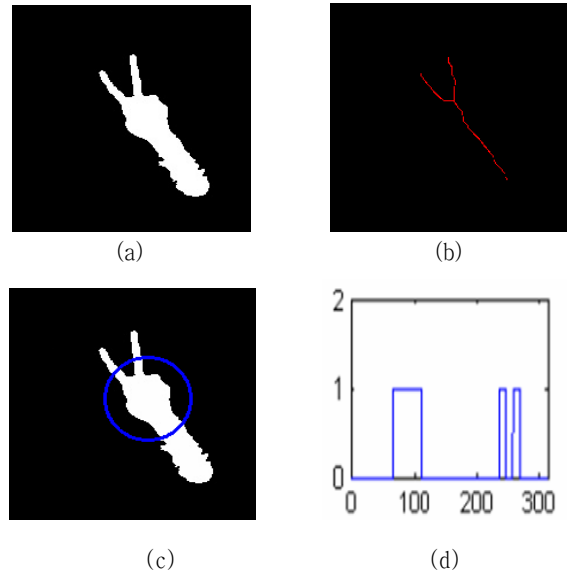
영역의 중심점(COG)으로 가장 큰 원의 반지름을 확장하여 모든 손가락을 지나치는 원을 그린 후, 원 둘레를 따라 시계방향으로 특징 벡터를 얻는다.

특징벡터는 시스템에 등록된 제스처의 특징벡터와 식(6)와 같이 유사도를 구한다. 이때 특징벡터는 원형으로부터 얻어지기 때문에 벡터를 회전하여 최소의 거리를 얻는다.

$$\text{dist}(x, s) = \min \sum_{i=1}^l \sqrt{(x_i - s_{ri})^2} \quad (6)$$

이때, r은 벡터의 회전각이다.

얻어진 특징벡터는 이진화된 이미지의 피부영역을 지나감에 있어서 손목을 포함하는 손가락수 + 1 만큼의 영역을 지나게 된다[3]. [그림5]은 손가락 두 개의 경우에 대하여 세선화와 특징벡터를 추출하는 과정을 보여주고 있다.

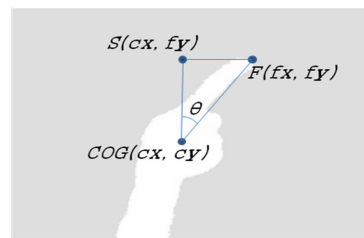


[그림 5] 두 손가락에 대한 세선화과정과 특징벡터 추출, a) 손 영역 추출, b) 손 영역에 대한 세선화, c) 손 영역의 특징벡터 추출, d) 특징 벡터 그래프

2.5 손가락 방향에 대한 인식

2.4 를 통하여 얻어진 특징벡터를 바탕으로 로봇의 상, 하, 좌, 우 컨트롤에 적용할 수 있다. 특징 벡터를 통하여 손 제스처가 손가락 하나로 인식될 경우 [그림 6]과 같이 손가락의 끝점 $F(fx, fy)$ 와 COG의 좌표값 $C(cx, cy)$ 을 계산하여 COG에 수직으로 그어진 선분 위의 임의의 점 $S(cx, fy)$ 를 얻을 수 있다. 여기서 얻어진 점 $S(cx, fy)$ 는 점 $F(fx, fy)$ 와 COG가 수직으로 만나는 점이다.

여기서 얻어진 $S(cx, fy)$ 는 손가락의 방향성을 계산하는데 사용되며, 식(7)을 사용하여 COG의 수직으로 그어진 선분을 중심으로 손가락이 이루는 θ 를 계산할 수 있다.



[그림 6] 손가락 방향에 대한 인식

$$\sin \theta = \frac{SF}{FC} \quad (7)$$

그러나, 사인 그래프는 90°를 기준으로 대칭성을 띄게 되므로 다른 각도에 대해서 같은 값을 가질 수 있다. 예를 들어

