

깊이 카메라 기반 손동작 공간터치 시스템 구현

안양근, 정광모
전자부품연구원
e-mail:ykahn@keti.re.kr

Depth Camera Based Hand Gesture Spatial Touch System Implementation

Yang-Keun Ahn, Kwnag-Mo Jung
Korea Electronics Technology Institute

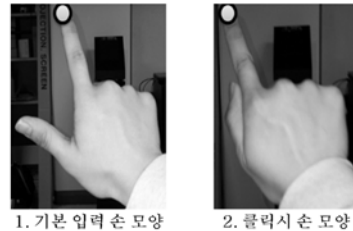
요 약

본 논문에서는 Depth 카메라를 이용하여 검지 끝을 인식하고 공간 터치 손 제스처를 인식하는 방법에 대해 제안한다. 제안된 방법은 손의 형태학적으로 엄지 끝을 추정하고 보정하는 방법을 제안하고, 추정된 검지와 엄지 끝의 위치를 이용해 마우스 이동, 클릭을 구현하여 문자 입력 시스템에 적용하였다. 제안된 방법을 실험하기 위하여 실제 디스플레이와 Depth 카메라를 하드웨어적으로 구성하고 마우스 기반 콘텐츠를 제작하여 이용하였다.

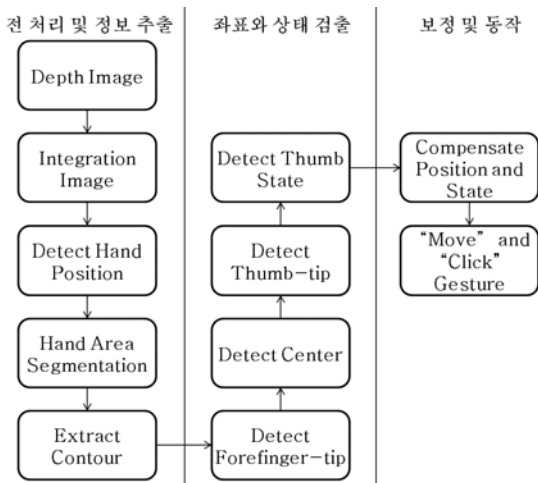
1. 서론

최근 모바일 기기, 스마트 TV 등의 보급이 활발해지면서 기타 입력 장치 없이 공간에서 동작할 수 있는 터치 시스템을 개발하는 연구가 활발히 진행되고 있다. 기존에 2D 카메라, 적외선 카메라 등을 이용한 손동작 인식 연구가 계속 진행되어왔지만, 손 영역 추출, 환경에 대한 강건함, 인식의 실제 편의성, 인식의 정확도등이 문제가 되어왔다.

본 시스템에서의 기본적인 입력은 오른손을 통해 이루어진다. 사용자는 오른손을 이용하여 마우스 좌표를 이동시키거나 클릭을 조작한다. 기본 손 입력 모양은 주먹을 쥔 상태에서 엄지와 중지를 편 'V'형태이다.



(그림 2) 손 입력 모양 정의



(그림 1) 시스템 순서도

본 논문이 제안하는 시스템은 손 영역 검출, 검지 끝 검출, 손 중심점 검출, 엄지 끝 검출 엄지의 상태 검출, 좌표와 상태 값의 보정의 순서로 이루어진다.

2. 공간 터치 손 모양 및 위치 정의

사용자 손 입력 시 검지의 끝부분이 곧 마우스의 좌표가 된다. 기본 입력 모양에서 사용자가 엄지를 접어 검지 혹은 중지에 붙이면 마우스의 좌측 버튼이 눌러진다. 사용자가 엄지를 접은 상태에서 엄지를 떼면 마우스의 좌측 버튼이 떨어진다.

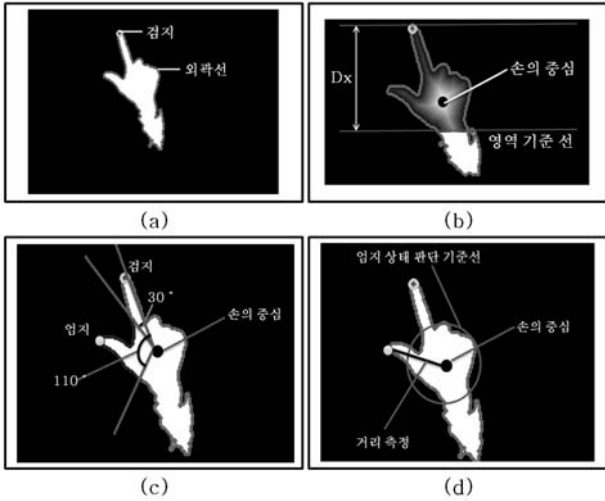
사용자가 입력을 위해 손을 내민다면 Depth 카메라의 영상에서 손은 가장 가까운 영역이 된다. 영상 내에서 가장 깊이 값이 작은 영역을 찾고 해당 깊이 값에서 5cm까지의 영역을 손 영역이라 가정하고 손 영역을 구분시킨다.

3. 특징 추출

입력된 영상을 이진화 한 후 Labeling 알고리즘을 이용하여 여러 군집으로 나타내고 크기가 임계치 보다 작은 군집을 제거한다. 최종 추출된 군집들을 영상에서 손끝 추출을 수행하기 위해 외곽선을 추출한다. 본 논문에서는 외

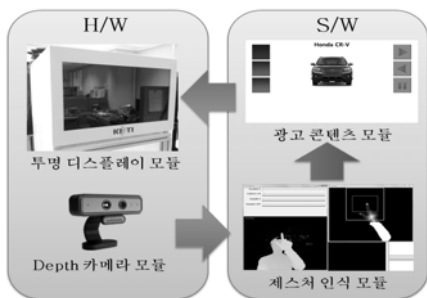
곽선 추출 기법은 공개 라이브러리 OpenCV에서 제공하는 외곽선 추출 함수를 사용하였다[1].

그림 3(a)에서 검지의 끝부분은 손 영역에서 가장 높은 위치에 있다. 즉 손 영역의 외곽선을 이루는 점 중 가장 높이가 있는 점이 검지의 끝점이 된다.



(그림 3) 특징 추출 (a) 손 외곽선과 검지의 끝 (b) 손 중심점 검출 (c) 엄지 끝점 검출 (d) 엄지의 상태 검출

그림 3(b)에서 손의 중심점은 손 영역에서 가장 안쪽에 위치한다. 거리변환 영상에서 가장 큰 값을 갖는 점이 손의 중심점이다. 이때 손목부분에서 가장 큰 값이 존재할 수 있음을 고려하여 검지 끝에서 아래로 특정 영역까지만 검출의 대상으로 한다. 그림 3(c)에서 손 영역의 외곽선을 이루는 모든 점들 중 해당 점과 손의 중심, 검지의 끝이 이루는 각이 30°~110° 사이인 모든 점은 엄지 끝의 후보 점이다. 엄지의 끝과 손의 중심, 검지의 끝이 이루는 작은 벡터의 내적을 이용하여 계산한다. 그림 3(d)에서 엄지가 펴져있거나 접혀있는 상태를 검출하기 위해 엄지의 끝점과 손 중심점의 거리 값을 이용한다. 검출한 엄지 끝점과 손 중심점의 거리를 계산하여, 이 값이 특정 임계값 이하이면 엄지를 접은 상태로 간주한다. 실제 손을 이동할 때 좌표의 떨림은 보완하기 위해 Low-Pass Filter[2]를 사용하였다.

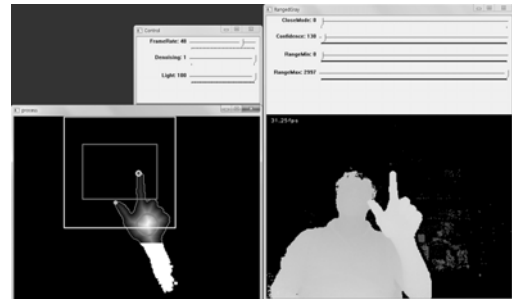


(그림 4) 시스템 구성도

4. 시스템 구성

본 논문이 제안하는 공간터치 기술성을 평가하기 위해 실제 마우스 기반 콘텐츠를 구현하였다. 시스템은 사용자의 손 영상을 깊이 카메라가 입력받고 입력된 데이터를 인식 모듈에서 분석하여 좌표 및 클릭 상태를 구한 뒤, 인식 결과를 콘텐츠에 적용한다. 그림 4는 간략한 시스템 구성 및 흐름을 보여준다.

그림 5는 인식모듈의 실행 화면이다. 파라미터를 조절하기 위한 파라미터 화면과 입력 영상을 출력하는 메인 화면, 손끝 검출 및 특징 영역을 출력하기 위한 손끝 검출 화면들로 구성되어 있다.



(그림 5) 인식 모듈

5. 실험결과

개발환경은 운영체제 Window7 OS 기반, 개발 툴 Visual Studio 2010, Win API Console c++ 환경에서 개발되었고, 하드웨어의 구성은 SoftKinetic사의 DS325 깊이 카메라, LG 투명 디스플레이, 데스크탑 PC Intel i5 CPU, 2.80GB에서 개발되었다. 인식에 사용되는 각 파라미터는 손의 최소 거리 40cm, 손의 최대거리 80cm, 손의 최소 넓이 1000, 이전 엄지 상태에 대한 가중치 0.75, 이전 좌표에 대한 최소 가중치 0.75, 이전 좌표에 대한 최대 가중치 0.98, 좌표의 최소 변화 2000, 좌표의 최소 변화 9000로 각각 사용하였다. 인식 알고리즘 부분만 고려했을 때 결과적으로 60Frame-rate의 속도를 보였으며 좌표의 떨림이 없고 반응속도 또한 뛰어났다.

6. 결론

본 논문에서는 공간 터치 기반 제스처 인식을 위해 형태학적인 엄지의 검출 방법과 좌표에 대한 보정 방법을 제안하였고, 실험을 위해 마우스 기반 콘텐츠를 구현하였다. 결과적으로 알고리즘 수행속도가 빠르며, 오차가 적고 반응속도가 빠른 시스템을 구현할 수 있었다.

참고문헌

[1] Intel Corporation. Open Source Computer Vision Library reference manual. December 2000
 [2] http://en.wikipedia.org/wiki/Low-pass_filter