

3D Skeleton Model을 이용한 제스처 인식

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

Gesture Recognition Using a 3D Skeleton Model

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

요 약

본 논문에서는 3D Skeleton Model로부터 획득된 관절 정보를 이용하여 제스처를 인식할 수 있는 방법을 제안한다. 사람의 신체 크기나 비율은 다르더라도 구조는 같다는 사실을 바탕으로, 관절과 관절이 이루는 각도를 이용해 제스처를 인식한다. 몇 가지 제스처를 선정한 뒤, 실험을 통해 제안한 방법의 인식률을 측정해 보았다. 또한 동적 제스처 인식을 위한 기초를 다지기 위해 이동 방향과 이동 거리, 이동 위치를 측정하는 실험을 해 보았다.

1. 서론

기술이 발전해 나감에 따라 전자 기기들의 크기와 무게는 줄어들어 휴대가 용이해 졌으며, 조작 방법은 사용자가 어렵지 않게 쓸 수 있도록 간편하게 발전하였다. 하지만 휴대성과 불구하고 특정 기기를 필요로 하는 것은 사용자에게 불편함으로 인식되고 있다. 이러한 사용자들의 요구에 맞춰 발전한 기술 중 하나로 제스처 인식(Gesture Recognition)을 꼽을 수 있다. HCI(Human-Computer Interface)에 속하는 이 기술은 별다른 입력 장치 없이 카메라로 촬영한 영상에서 사람의 자세나 행동을 인식하여 동작하는 기술을 말한다[1].

기존의 방법들을 살펴보면 사람의 신체 중 제스처를 판단할 기준을 가진 부분을 찾아내 특징 추출 및 제스처 패턴 모델링을 통해 제스처 인식을 했었다. 그러나 기술이 발전함에 따라 깊이 카메라를 통해 사람의 관절 정보를 추적해 사람의 골격을 특정할 수 있는 방법이 나오면서 신체 부분을 특정하고 자세를 추측하는 어려움을 덜 수 있게 되었다[2][3].

본 논문에서는 이렇게 얻어진 골격 정보를 토대로 제스처 인식을 할 수 있는 방법을 제안한다.

2. 본론

본 논문에서는 제스처 인식을 실험하기 위해 키넥트(Kinect)가 제공하는 NUI API를 사용해 3D 스켈레톤 모델(Skeleton Model)을 얻었다. 키넥트의 스켈레톤 모델로부터 얻을 수 있는 관절정보는 총 20가지이다.

각 관절 정보는 x, y, z의 좌표 정보를 담고 있다. 본 논문에서는 이 정보를 토대로 사용자가 어떠한 제스처를 취하고 있는지를 파악할 것이다.

3. 제안하는 제스처 인식 방법

신체의 크기나 비율은 사람마다 차이를 가지고 있지만 관절의 연결 순서는 모든 사람에 있어 동일하다. 인접한 두 개의 관절을 선택해서 그 관절들 사이의 방향 벡터를 찾는다면 모든 사람들을 대상으로 자세를 판단할 수 있는 기준을 얻을 수 있을 것이다. 본 논문에서는 양 팔의 관절 간의 방향 벡터를 가지고 판단 할 수 있는 제스처들을 대상으로 정적 제스처에 대한 실험을 해 보았다.

이동 제적을 통해 동적 제스처를 구하는 이전 단계로 오른쪽 손목 관절의 이동 정보를 받아 이동 방향과 이동 거리를 구하는 실험을 해 보았다. 그리고 같은 방향으로 같은 거리를 움직였을 때, 스켈레톤 모델의 정보를 기준으로 이동 위치를 확인 할 수 있는지를 함께 실험해 보았다.

<표 1> xy좌표에 대한 방향 구분

1	337.5° ~ 22.5°	
2	22.5° ~ 67.5°	
3	67.5° ~ 112.5°	
4	112.5° ~ 157.5°	
5	157.5° ~ 202.5°	
6	202.5° ~ 247.5°	
7	247.5° ~ 292.5°	
8	292.5° ~ 337.5°	

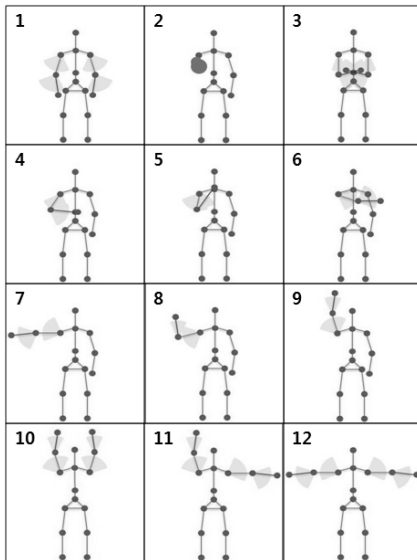
4. 제스처 인식 제약 사항

스켈레톤 모델은 카메라를 통해 받은 영상을 분석해서 작성된 것이므로 좌표 값에 오차가 생기는 것을 완벽하게 방지할 수 없다. 따라서 본 논문에서는 관절 사이의 방향 벡터를 측정하는데 x, y 좌표평면을 표1의 8방향으로 나누어 대략적인 방향으로 사용했다.

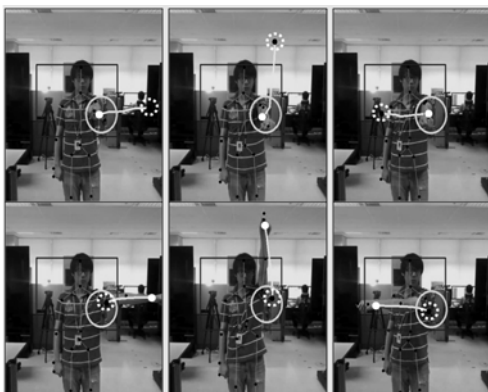
방향 벡터를 구할 때 사용자가 정면을 바라보고 서있는 상태만을 생각하고 실험을 진행하였다. 또한 스켈레톤 모델을 추출할 때 각 관절 점이 서로 겹칠 경우 오작동을 할 가능성이 있으므로 손목, 팔꿈치, 어깨 부분을 전부 볼 수 있도록 사용자의 키보다 조금 더 높은 높이에서 촬영하도록 키넥트를 배치하였다.

5. 실험

실험을 위해 몇 가지 자세를 선정한 후, 그에 대한 스켈레톤 모델에서 인접한 관절끼리 이루고 있는 방향 벡터들을 얻어냈다. 그리고 미리 선정해둔 자세들을 취하고 있는 사용자를 키넥트로 촬영하여 스켈레톤 모델을 추출한 후, 양 팔의 관절들이 인접한 관절들과 이루고 있는 방향 벡터들을 구한다. 이렇게 구해진 방향 벡터들을 미리 선정해둔 제스처들의 방향 벡터 정보와 비교해서 현재 사용자가 취하고 있는 정적 제스처가 무엇인지를 찾아낸다. 본 논문에서는 총 12가지의 자세를 선정하였으며, 해당 자세를 반복해 봄으로써 인식률과 오차율에 대한 데이터를 수집했다. 이 때 제시된 제스처 모델은 거울 모드의 자세이다.



(그림 1) 정적 제스처 샘플



(그림 2) 동적 제스처 실험

오른쪽 손을 앞으로 쪽 내민 상태로 상하 좌우로 움직여 이동 방향과 이동 거리를 측정하였다. 또한 좌우로 움직일 때 오른쪽 어깨를 중심으로 원을 그려 좌측에서 움직였는지, 우측에서 움직였는지 동적제스처를 구분하였다.

6 실험 결과

실험 결과 3번 자세를 제외한 모든 정적 제스처에 대해 100%의 인식률을 얻을 수 있었다. 96%의 인식률을 얻은 3번 자세는 실험을 시작할 때 우려했던 신체가 서로 겹쳐 스켈레톤 모델을 제대로 구해내지 못한 경우이다.

동적 제스처의 경우 신체 안쪽으로 손을 이동하는 경우 스켈레톤 모델이 일그러지는 현상이 정적 제스처보다 심하게 일어나서 이동 방향과 이동 거리를 찾는 데 계속해서 인식 오류가 일어났다. 해결 방법으로는 각 관절이 서로 영향을 크게 미치지 않는 범위에서 동작을 하는 동적 제스처를 설정하거나 정적 제스처와 마찬가지로 신체가 겹치더라도 정확한 스켈레톤 모델을 추출할 수 있는 방법이 필요할 것이다.

7. 결론

본 논문에서는 신체의 크기나 비율과는 관계없이 관절들이 이루고 있는 각도들을 이용해 제스처를 구분하는 방법을 제안하였다. 이를 실험하기 위해 키넥트로 부터 3D 스켈레톤 모델을 추출하여 인식률을 측정해 보았으며, 실험 결과를 통해 높은 인식률을 얻을 수 있었다.

그리고 동적 제스처 인식을 위한 준비 단계로서 몇 가지 간단한 테스트를 통해 동적 제스처를 인식할 때 고려해야 할 점을 찾아보았다.

참고문헌

[1] 김문환 외 2인 “유비쿼터스 로봇과 휴먼 인터랙션을 위한 제스처 추출” 제어로봇시스템학회 논문지, 제 11권 제12호, 2005년 12월, pp. 1062-1067
 [2] 조선영 외 3인 “키넥트 센서 데이터를 이용한 손 제스처 인식” 방송공학회논문지, 제 17권 제3호, 2012년 5월, pp. 447-458
 [3] 조선영 외 3인 “키넥트 센서 기반 슈팅 게임을 위한 팔 제스처 인식” 정보과학회논문지, 소프트웨어 및 응용 제 39권 제 10호, 2012년 10월, pp. 796-805