Kinect v2를 이용한 인물 위치 및 행동 인식

*박경무, **김상준, ***이유진, *박구만
*서울과학기술대학교 전자 IT미디어공학과
**서울과학기술대학교 정보통신미디어공학전공
***서울과학기술대학교 미디어IT공학과
mwe226@naver.com

Person Location and Behavior Recognition Using the Kinect v2

*GyeongMoo Park **SangJoon Kim ***YuJin Lee *GooMan Park
*Dept. of Media It Engineering
**Dept. of Information Technology and Media Engineering
***Dept. of Electronics and IT Media Engineering
Seoul National University of Science and Technology

요약

인물의 위치와 행동을 인식하는 것은 여러 분야의 서비스에서 활용할 수 있는 기술이다. 그렇기에 다양한 방식으로 연구되어 왔다. 기존의 방식은 일반 RGB 카메라의 영상에 영상처리 기법과 딥러닝을 사용하여 3차원 공간상의 인물 위치를 인식하는 방식과 라이다와 같이 깊이를 인식 할 수 있는 장치를 활용하여 3차원 공간상 인물의 위치를 인식하는 방식이 있다. 각각의 방식은 RGB 카메라를 이용할 수 있다는 장점, 인식률이 우수하다는 장점을 가지고 있다. 하지만 영상처리 방식은 연산량이 많아 실시간 서비스에 불리하다는 한계점이 있다. 라이다 방식은 기기의 부피가 커 공간제약이 있다는 점과 이동이 불편하다 있다는 한계점이 있다. 본 연구에서는 Kinect와 openFrameworks를 활용하여 공간이 효율적이고 연산량이 적은 방식의 3차원 공간에서 인물 위치 인식과 실시간 이동에 대한 방향 인식을 다룬다.

1. 서론

인물의 위치와 행동을 인식하는 것은 건물의 인원 출입 확인, 잠재적 고객의 행동기반 광고 등 많은 서비스에서 활용 할 수 있다. 그렇기에 활발히 연구가 이루어진 분야이다. 기존 인물 위치 또는 행동 인식 방식은 크게 영상처리, 이미지 기반 딥러닝을 활용하는 방식과 라이다를 이용하는 방식이 있다. 영상처리, 이미지 기반 방식은 상대적으로 연산량이 많아서 실시간 정보 활용이 어렵다는 한계점이 있다.[1] 라이다를 활용하는 깊이 정보 오차는 영상처리 방식에 비해 적은 편이지만 라이다의 부피가 크고 상대적으로 고비용의 장비를 사용하기도 한다. [2]

본 논문에서는 라이다 방식에 비해 상대적으로 부피가 작은 Kinect v2를 사용하고 영상처리 방식에 비해 상대적으로 연산량이 적은 Kinect v2의 RGB sensor, Depth senor 등의 sensor를 활용하여 사용자의 움직임을 인식하는 것에 관해 탐구하여 본다.

2. 관련연구

2.1 Kinect V2

Kinect는 MicroSoft에서 Xbox 출시한 음성, 동작, 인식 카메라이다. Kinect 는 Color Sensor, IR Emitter, IR Depth sensor를 통하여 일반적인 카메라로 촬영한 것과 같은 'RGB 정보', 카메라로부터의 거리의 정보인 'Depth 정보', 검출된 사용자의 골격을 25개로 나누어서 표현해주는 'Skeleton Joints 정보'를 얻는다.[3]

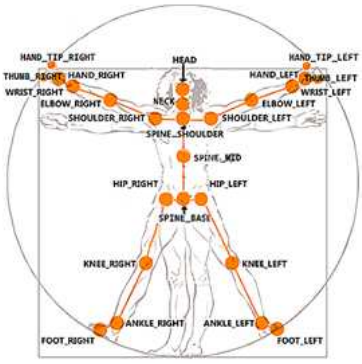

그림 1. Kinect v2가 인식 가능한 25가지 Skeleton Joints의 구성

인물의 행동과 위치를 판단하기 위해 Skeleton Joints 중 SPINE_BASE(골반), SPINE_MID(명치), SPINE_SHOULDER(어깨)의 X, Z 좌표를 활용하여 인물의 이동을 계산하고 계산한 정보를 토대로 이용자의 행동을 파악하고 표시한다.

2.2 openFrameworks

openFrameworks는 여러 가지 라이브러리를 통합하여 이용의 편의성을 높인 프레임워크를 지원하는 open source C++ toolkit이다. 윈도우, 맥OSX, 리눅스, iOS, 안드로이드 5개의 OS 와 4종류의 IDE를 지원하는 Cross-platform toolkit이기도 하다. 또한 다양한 Addon을 사용하여 필요한 라이브러리를 편리하게 추가할 수 있다. [4]

본 논문에서는 Kinect관련 라이브러리와 OpenCV, GUI, 등의 addon을 활용하여 Kinect의 RGB 정보, Depth 정보, Skeleton joints 정보를 다루고 영상에 도형을 표시한다.

3. 본론

본 연구는 인물의 이동 방향과 위치를 인식하는 것을 목표로 한다. 구체적으로 인물의 이동 방향은 하면, Kinect Camera 정면을 기준 하여 "N(북쪽)" 방향이라고 정의한다. 그리고 시계방향으로 45°씩 회전한 방향을 각각 "NE", "E", "SE", "S", "SW", "W", "NW"로 정의한다. 또한 사용자의 좌표가 일정 수준 이상으로 움직이지 않는 상태인 "정지" 까지 총 9개의 사용자 이동 방향을 인식한다.

또한 인물의 위치를 Kinect 로부터 1.8m 떨어진 지점을 중앙으로 하여 60cm 반경의 원 안에서 인식한다. 인식한 이동 방향과 위치를 openFrameworks의 GUI를 통하여 방향 표시 기호와 위치 표시기호를 이용하여 표시 한다.

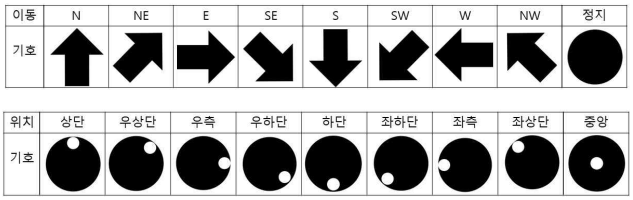

그림 2. 인물의 이동 방향 표시 기호(상)와 인물의 위치 표시 기호(하) (Kinect 전면 기준)

인물의 이동 방향 표시 기호는 시간에 따른 인물의 이동 방향을 표시하고 위치 표시 기호는 0.1mesc 마다 갱신하여 인물의 위치를 큰 원안의 작은 원으로 표시한다.

사용자의 이동을 인식하기 위해 적절한 Skeleton Joint를 선정한다. Skeleton Joint는 신체의 중앙 지점에 있는 SPINE_BASE, SPINE_MID, SPINE_SHOULDER의 좌표의 평균값을 사용하였다. 위와 같은 좌표선정으로 오른발을 사용한 이동과 왼발을 사용한 이동 사이의 좌표값의 차이를 줄일 수 있었다.

Kinect로 부터 사용자의 3차원 공간정보를 X, Y, Z 좌표 형태로 수신한다. (Kinect 전면 기준 좌우 움직임은 X, 앞뒤 움직임은 Z 좌표의 변화로 나타난다.) 그 후 시간간격을 두어 X, Z를 수신하여 좌표의 시간에 따른 변화를 구한다. 그 변화량을 식(1,2)을 이용하여 크기와 방향을 구하여 사용자의 이동을 크기와 각도를 가진 벡터값으로 변환한다.

$$\sqrt{(x_2-x_1)^2+(z_2-z_1)^2} \quad (1)$$

$$\arctan\left(\frac{z_2-z_1}{x_2-x_1}\right) \quad (2)$$

여기서 x_1은 x좌표 이고 x_2는 x_1보다 0.1msec뒤의 x좌표이며 z_1은 z 좌표 z_2는 z_1보다 0.1mesc뒤의 z좌표이다.

움직임의 크기가 일정 이상이 되면 움직임이 있는 것으로 인식하고 각도를 통해 이동 방향을 정한다. 움직임의 크기가 특정 값 미만이면 움직임이 없는 것으로 인식한다 . 움직임의 크기는 30CM 이상의 움직임부터 움직임이 있는 것으로 설정하였다.

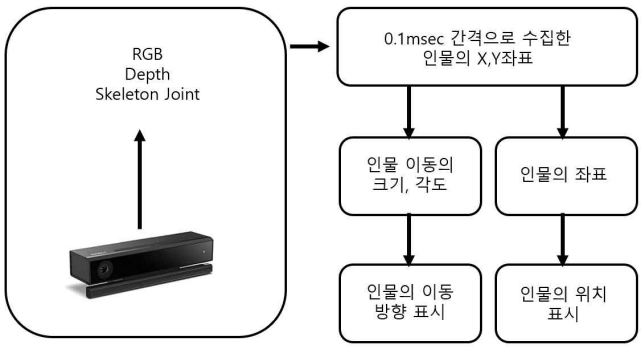

그림 3. 인물의 이동 방향과 인물의 위치를 인식하는 과정

4. 실험

인물의 행동과 위치 인식률을 확인하기 위해 실험을 구성한다. cmat 라이브러리의 random 함수를 이용하여 8가지 방향 중 무작위 방향을 지시한다. 실험 대상은 무작위로 지시된 방향으로 이동한다.

이동의 일관성을 위해 먼저 Kinect 로부터 약 1.8m 떨어진 지점을 중심점으로 설정한다. 그 중심점을 기준으로 하여 약 50cm(한걸음) 떨어진 8가지 방향에 표식을 해둔다. 실험 대상은 무작위 지시 방향에 따라 표식이 있는 지점으로 이동하고 프로그램이 이 이동을 잘 인식하는지 무작위 방향과 방향 인식 결과를 비교한다. 무작위 지시 방향과 인식 방향이 다르다면 인식 실패, 같다면 인식 성공으로 정의한다. 이를 50회 5번 반복하여 인식률을 계산하였다.

그림 4. 설계된 실험을 수행하는 모습 (고정된 Kinect, 실내 환경에서 진행)

5. 결론

Kinect를 활용하여 Depth 정보가 포함된 인물의 Skeleton joints좌표를 얻을 수 있었고 openFrameworks Addon을 통해 좌표값을 수식에 대입하여 인물의 위치를 확인하고 인물의 이동 방향을 확인 할 수 있었다. 기본의 영상 처리를 통하여 인물의 위치를 확인하는 방식에 비하여 연산이 줄어들어 실시간 연산을 통한 빠른 정보 제공에 유리하며 라이다 방식에 비하여 장비의 부피가 작고 비용이 저렴한 Kinect를 사용하여 장소 조건에 덜 구애받는 인물 위치 인식을 구현할 수 있었다.

표.1. 설계된 실험에 따른 실험 진행 결과

실험 번호	시험 횟수	인식 횟수	인식률
1	50	48	96
2	50	46	92
3	50	47	94
4	50	48	96
5	50	49	98
평균	50	47.6	95.2

방향을 판단하는 방식은 0.1msec 단위로 X, Z 좌표를 비교하여 움직임의 여부를 판단한 뒤 각도를 반영하는 것이다. 그렇기 때문에 0.1msec 안에 일정 값보다 작은 수치로 움직인다면 움직임이 없는 것으로 인식하여 인식에 실패하게 된다. 하지만 대부분의 움직임에 대해서는 약 95%의 인식률을 보여준다.

향후 연구는 Kinect를 기준으로 인물이 위치하는 지점을 동적으로 조정하여 인물의 위치를 능동적으로 설정하는 것을 목표로 한다.

감사의 글

이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2017-0-00217, 투명도와 레이어 가변형 실감 사이니지 기술 연구)

참고문헌

[1] 노요한, 김민정, 이도훈 “사람 행동 인식에서 반복 감소를 위한 저수준 사람 행동 변화 감지 방법” 멀티미디어학회논문지, pp. 432-442, 2019.

[2] 이지상, 홍승환, 박일석, 손홍규 “지상라이다 취득 점군자료와 영상을 이용한 실내 환경에서의 인물 감지.” 한국측량학회 학술대회자료집, pp. 244-249, 2017.

[3] 이새봄, 정일홍 “키넥트를 사용한 NUI 설계 및 구현” 한국디지털콘텐츠학회 논문지, pp. 473-480, 2014.

[4] Internet Openframeworks 구조 https://openframeworks.cc/about/