

3차원 공간 데이터 특징을 이용한 포즈 인식 방법

치옥용, 조성재, 엄기현, 조경은
 동국대학교 멀티미디어공학과
 e-mail : cke@dongguk.edu (교신저자)

Posture Recognition Method using 3D Space Data Feature

Yulong Xi, Seungjae Cho, Kyhyun Um, Kyungeun Cho
 Department of Multimedia Engineering, Dongguk University-Seoul,
 26 Pildong 3 Ga, Jung-gu, Seoul 100-715, Republic of Korea

요 약

최근 포즈 인식 기술의 연구는 HCI, 인공지능 등의 분야에서 활발히 진행되고 있다. 하지만 대부분의 연구는 센서와 사용 환경으로부터 많은 영향을 받는다. 본 논문에서는 이러한 종속성을 최소화하여 범용성 있는 포즈 인식 방법을 제안한다. 이 방법을 통해 사람의 다양한 포즈로부터 획득한 특징 데이터를 최적화함으로써 다양한 포즈를 실시간에 인식할 수 있다.

1. 서론

최근 포즈 인식 기술은 HCI(Human-Computer Interaction), 인공지능 등의 다양한 분야에서 연구되고 있다. 이러한 연구들은 주로 RGB-D 카메라에서 획득한 스켈레톤 정보를 사용하여 포즈 인식을 수행하였다 [1][2][3][4]. 그러나 이러한 연구의 경우 사용된 센서의 종류에 따라 스켈레톤 관절의 개수에 차이가 존재한다. 또한 이러한 포즈 인식 방법은 각 센서의 특징에 의존하며 다른 센서를 사용할 경우 바로 적용하기 어려우므로 범용성이 감소한다.

본 연구에서는 이러한 문제를 해결하기 위해 센서의 영향을 최소화하는 데이터 처리 방법을 제안하였다. 이 방법을 사용하여 대량의 포즈 데이터를 실시간에 처리할 수 있다.

2. 관련연구

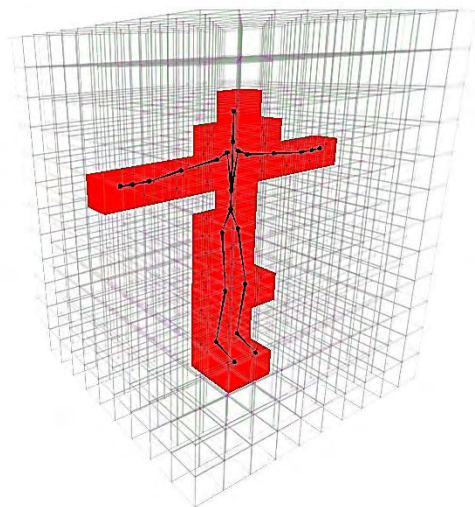
마이크로소프트(Microsoft)사의 키넥트(Kinect)가 등장한 이래로 스켈레톤을 사용한 제스처 인식 및 포즈 인식 연구가 많이 진행되었다 [1][2][3][4]. 이런 연구들 연구마다 사용한 특징의 관절 개수가 차이가 커서 학습된 데이터의 범용성이 떨어진다.

본 논문에서는 사람 모델의 주변 3D 공간 데이터를 이진 배열 특징으로 표현하여 학습을 진행하였다. 이 방법으로 비교적 적은 이진 배열 형식의 데이터로 사람의 모든 포즈를 인식하는 것이 가능해진다. 또한, 공간 데이터를 사용함으로써 사람 모델의 구동 방식에 제한이 없으며 다양한 센서에서 사용할 수 있다.

3. 3D 블록을 이용한 포즈 학습 및 인식

본 연구에서는 사람의 행동 범위를 기준으로 3D 인간 모델의 주변 3D 공간의 크기를 결정한다. 그림 1은 이 3D 공간을 여러 개의 작은 블록으로 나누어

구성한 형태를 나타낸 것이다.



(그림 1) 사람 모델의 3D 공간 데이터

이러한 3D 공간을 구성하는 블록들은 1차원 배열 형태로 3D 인간 모델과의 교차 여부를 저장한다. 각 블록은 3D 인간 모델과 교차할 경우 교차 여부 배열의 해당 블록의 인덱스 값을 1로 설정하며 교차가 발생하지 않은 블록에 해당하는 인덱스 값은 0으로 설정한다. 이런 3D 블록의 집합은 하나의 이진 배열로 정리할 수 있으며 이를 통해 다음과 같이 사람의 포즈 한 개를 표현할 수 있다.

$$P=\{0,0,0,1,0,1\dots,0,0,0\}.$$

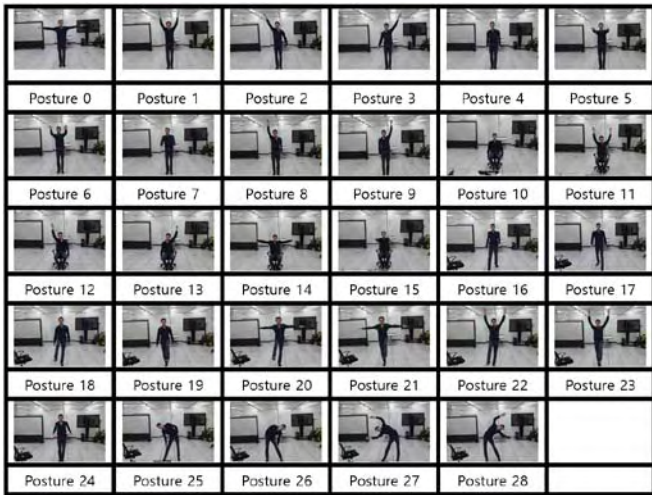
사람이 특정 포즈를 여러 번 취할 경우 동일한 포즈라도 매번 차이가 발생한다. 또한 센서를 통해 수집한 데이터에도 오차가 존재한다. 이러한 오차를 고

려하여 본 연구에서는 하나의 포즈를 여러 번 학습한 후 SVM(Support Vector Machine)으로 분류하였다. 대량의 포즈 데이터를 분류하기 위해서는 Multi-Class Classification 과 One-Vs-Rest 방식을 적용하였다.

4. 인식 실험

실험은 Intel(R) Core(TM) i5-4690 CPU, 8GB RAM, GeForce GTX 760 의 PC 환경에서 진행되었다. 본 포즈 학습 및 인식 시스템은 Unity3D 엔진을 사용하여 구현하였다.

본 연구에서는 그림 2 와 같이 전신, 상체, 앉은 상태 등의 3 가지 종류의 총 29 개의 포즈를 학습하였다. 각 포즈는 약 150~200 회 학습 데이터가 누적되었다. 본 시스템을 통해 포즈를 학습한 후 포즈 인식 시 평균 94.03%의 인식률을 달성하였다. 또한 포즈 인식의 속도는 평균 0.009ms 을 달성하였다. 본 실험을 통하여 다양한 포즈를 실시간에 정확하게 인식할 수 있음을 확인하였다.



(그림 2) 본 시스템을 통해 학습된 29 개의 포즈

5. 결론

본 연구에서는 사람 모델의 행동 범위를 포함하는 3D 공간 데이터를 특징으로 추출하여 포즈 학습 및 인식을 수행하는 방법을 제안하였다. 본 시스템은 실험을 통해 실시간에 다양한 포즈를 인식 가능성이 확인되었다. 본 기법은 특정 센서에 종속되지 않는 사람 모델을 구동하여 포즈 학습 및 인식을 수행함으로써 범용성을 확보하였다. 또한 최적화된 특징 데이터를 사용함으로써 실시간에 대량의 포즈를 인식할 수 있다. 추후 연구에서는 라이브러리 형식으로 발전시킴으로써 다양한 응용 분야에 적용하고자 한다.

감사의 글

본 연구는 미래창조과학부 및 정보통신기술진흥센터의 대학 ICT 연구센터육성 지원사업의 연구결과로 수행되었음 (IITP-2015-H8501-15-1014).

참고문헌

- [1] Orasa Patsadu, Chakarida Nukoolkit, Bunthit Watanapa, "Human gesture recognition using Kinect camera", JCSSE (2012) 28-32
- [2] Thi-Lan Le, Minh-Quoc Nguyen Thi-Thanh-Mai Nguyen, "Human posture recognition using human skeleton provided by Kinect", ComManTel (2013) 340-355
- [3] Junhong Wu, Jun Cheng, Wei Feng, "3D Dynamic Gesture Recognition Based on Improved HMMs with Entropy", International Conference on Information and Automation (2014) 213-218
- [4] Chang-Beom Park, Seong-Whan Lee, "Real-time 3D pointing gesture recognition for mobile robots with cascade HMM and particle filter", Image and Vision Computing (2011) 51-63