

# 멀티 센서 퓨전을 이용한 실시간 3D 캐릭터 애니메이션

성만규, 손영우  
 계명대학교 컴퓨터공학부 게임모바일공학전공  
 e-mail:mksung@kmu.ac.kr,

## Realtime 3D Character Animation Through Multi Sensor Fusion

Mankyu Sung, Youngwoo Son  
 Dept of Game & Mobile, Faculty of Computer Engineering, Keimyung University

### 요 약

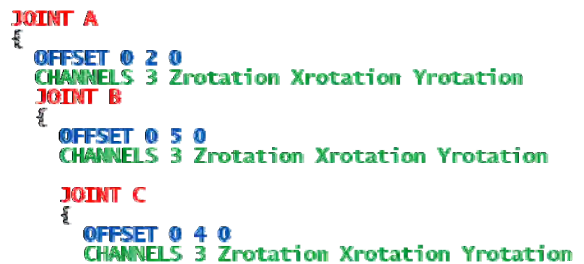
최근 다양한 3차원 뎀스 센서의 등장은, 3차원 캐릭터가 사람의 움직임에 따라 실시간으로 애니메이션 되도록 하였다. 하지만, 센서에 따라, 공간상의 뎀스를 얻는 방식이 다르며, 이 결과 캡처를 가능하는 센싱영역 또한 뎀스의 종류에 따라 많은 차이를 보여 왔다. 본 논문은 두 가지 방식의 멀티의 센서를 결합하여, 동시에 실시간으로 사용함으로써, 하나의 센서만을 사용했을 경우 얻을 수 없는 조인트에 대한 정보를 얻음으로서, 자세한 캐릭터에 대한 스켈레톤을 애니메이션 하는 방법을 제안한다.

### 1. 서론

캐릭터 애니메이션은 게임과 같은 콘텐츠에서 가장 중요한 부분을 차지한다. 게임에서 사용되는 캐릭터 움직임은 키-프레임 방식을 이용하여 제작되었다. 하지만, 이 방식은 미리 정해진 움직임만을 이용할 수 있는 단점이 있다. 뎀스 센서는 공간상의 3차원 위치에 대한 거리정보를 제공하고[1], 이 거리에 대한 정보를 다양한 기계학습 알고리즘에 적용하여 분석하면, 사용자의 3차원 조인트에 대한 위치 및 방향정보를 알 수 있도록 하였다 [2]. 이 정보를 이용하면, 사용자의 움직임을 그대로 따라하도록 3차원 캐릭터의 움직임을 실시간으로 생성 할 수 있었다. 본 논문은 각 뎀스 센서의 장점을 결합한 센서 퓨전 방식을 이용하여 하나의 센서만을 이용했을 경우 얻을 수 없었던 복잡한 움직임을 캡처 할 수 있는 방법을 제안한다.

### 2. 스켈레톤 구조

본 논문에서는 사용자의 움직임을 캡처하여 이를 3차원 캐릭터에 적용하기 위한 계층적인 조인트 구조를 이용한다. 이를 위하여 가장 대표적인 모션캡처 포맷인 BVH를 이용한다. 그림 1은 BVH포맷의 계층구조를 나타낸다. 이 포맷에서 JOINT는 조인트의 이름을 나타내며 OFFSET의 부모 좌표계에서 원점에서 해당 조인트까지의 길이를 나타내는 벡터이며, CHANNELS는 애니메이션을 위해 필요한 정보의 수와 어떠한 정보가 필요한지를 나타낸다.



(그림1) BVH 계층구조

### 3. 센서 퓨전

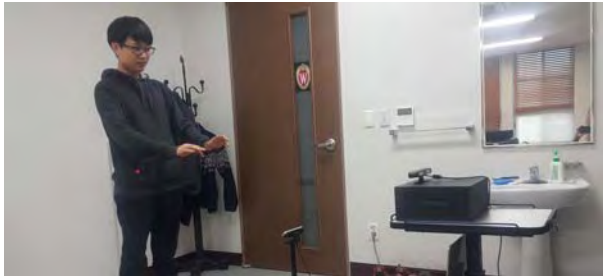
현재 사용되고 있는 뎀스센서는 Intel의 Real sensor, Asus의 Xion등 다양한 종류가 있으며, 센서에 사용되는 캡처 방식에 따라 각각 장단점을 갖고 있다. 표1은 이와 같은 다양한 뎀스센서를 비교 한 것이다.

방식	Time of Flight	스테레오	구조광
Latency	low	medium	medium
Range	Short	Mid	short to mid
Low light	Good	Weak	Good
Accuracy	mm to cm	mm to cm	mm to cm
Speed	Fast	Medium	Medium
Product	Kinect v2	Zed	Kinect v1/

<표1> 뎀스 센서의 종류 [3]

본 연구에서는 두 개의 센서를 동시에 사용하여 각 센서로부터 얻은 정보를 결합하여 BVH 스켈레톤은 실시간으로 애니메이션 하는 방법을 제안한다. 본 연구에서 사용한

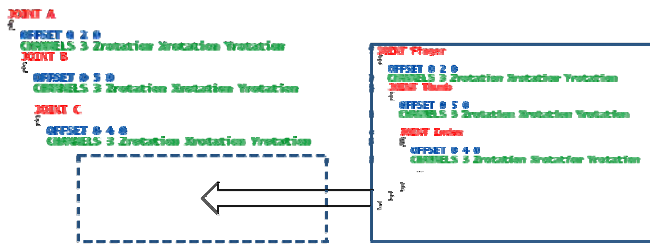
두 가지의 센서는 Asus사의 Xtion pro 센서와 Intel의 Realsensor이다. Xtion pro의 센서의 경우, 중거리에서 전신에 대한 캡처에 적합 하다. 하지만, 사용자의 손가락에 대한 정보를 캡처 하지 못한다. 두 센서는 그림2에서와 같이 근거리와 중거리에 설치하여 동시에 캡처 되도록 하였다. 이 때 구조광 방식의 댄스 센서의 경우 근 거리에 있는 센서가 중거리의 센서를 가리는 현상이 일어나므로, 센서의 위치를 조정하여 공간의 아래 부분과 윗 부분에 센서를 위치시켜 서로간의 간섭현상을 최소화 한다.



(그림2) 댄스 센서 퓨전 세팅

#### 4. BVH포맷 설계

본 연구에서는 댄스센서가 제공해 주는 25개의 조인트와 손가락 20개의 조인트를 결합하여 총 45개의 조인트를 갖는 BVH를 설계하였다.



(그림3) BVH 설계

#### 5. 실시간 캐릭터 애니메이션

모션캡처 데이터는  $n$ 개의 프레임으로 이루어져 있다.

$$F_i = \{p_i^0, q_i^0, q_i^1, q_i^2, \dots, q_i^m\} \quad (1)$$

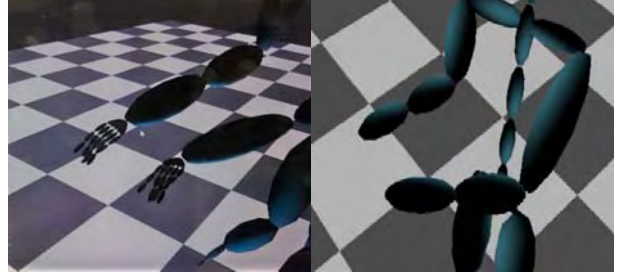
$p_i^0$ 는  $i$ 프레임에서의 루트 조인트의 3차원 위치,  $q_i^0$ 는 루트의 오리엔테이션,  $q_i^j$ 는  $j$ 번째 로컬 오리엔테이션을 의미한다. 댄스센서로부터 받은 조인트  $i$ 의 글로벌 오리엔테이션 데이터를  $I^i$ 라고 가정하고 BVH상의 조인트  $i$ 의 부모 조인트  $j$ 의 글로벌 오리엔테이션을  $q^j$ 라고 하자.  $q^j$ 는 루트로 부터 조인트  $j$ 까지의 모든 조상 조인트의 로컬 오리엔테이션을 곱하여 구한다. 실시간으로 연동되는 조인트  $i$ 의 로컬 오리엔테이션은  $I^i$ 와 조인트  $i$ 의 부모조인트인 조인트  $j$ 의 글로벌 오리엔테이션의 차이를 통해 구할 수 있다.

$$q^i = I^{-i} \cdot q^j \quad (2)$$

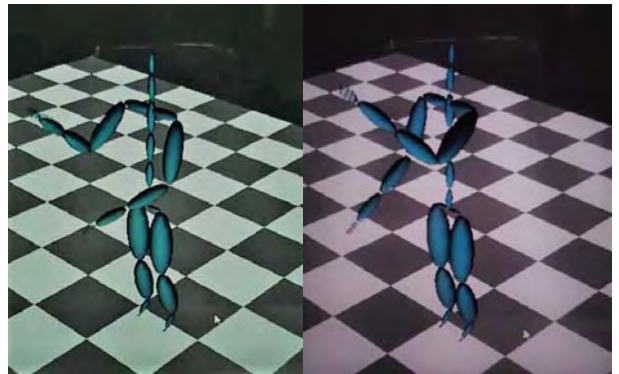
#### 5. 실험결과

본 연구를 위해 그림2와 같은 근거리, 중거리 센서를 위치시킨 후에 동시에 캡처하였다. 렌더링은 OpenGL를 이용

하였으며, 빠른 렌더링을 위하여 GPU Shader를 사용하였다. 개발환경은 Window7에 Visual Studio 2015를 이용하였다. 그림4에서 쉽게 알 수 있듯이 센서 퓨전을 이용하면 팔과 함께 손가락의 움직임은 애니메이션 할 수 있다. 하지만 하나의 센서(중거리 센서)만을 이용할 경우, 손가락 애니메이션은 얻을 수 없다.



(그림4) 센서 퓨전을 이용한 캐릭터 애니메이션 비교  
그림 5는 센서 퓨전을 이용하여 동시에 전신 캡처와 동시에 손가락 캡처를 수행한 스크린 샷 이다.



(그림5) 전신+손가락 동시 애니메이션

#### 6. 결론

본 연구는 두 개의 서로 다른 댄스센서를 동시에 이용하여 세밀한 3차원 애니메이션을 생성하는 기법을 제안하였다. 본 연구에서 가장 어려운 부분은 근거리 센서와 중거리 센서의 물리적 위치를 서로간의 간섭을 최소화하도록 위치시키는 것이다. 이를 위해 추후 다양한 실험을 통해 최적의 위치를 파악 할 예정이다.

#### Acknowledgment

이 논문은 2016년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(2015R1D1A1A01059066)

#### 참고문헌

[1] Kinect Hack, [https://www.wired.com/2011/06/mf\\_kinect/all/1](https://www.wired.com/2011/06/mf_kinect/all/1)  
 [2] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp and M. Finocchiro, Real-Time Human Pose Recognition in Parts from Single Depth Images, In Proc. CVPR., 2010  
 [3] Comparison of 3D imaging technologies, <http://www.ti.com/lscds/ti/sensing-products/optical-sensors/3d-time-of-flight-sensors-technology-comparison.page>  
 [4] BVH format, <https://research.cs.wisc.edu/graphics/Courses/cs-838-1999/Jeff/BVH.html>