

움직임 관성 센서를 이용한 체험형 댄스 콘텐츠 개발

Development of Experiencing Dance Contents using Motion Inertial Sensors

김종남, Jongnam Kim*, 선덕규, Dukkyu Seon*, 김대웅, Daeong Kim*, 김찬수, Chansu Kim**,
정영기, Youngkee Jung*
호남대학교 *컴퓨터공학과, **게임애니메이션학과

요약 본 연구에서는 관성 센서가 내장된 컨트롤러를 이용하여 정의된 5가지의 동작과 비교하여 가상의 댄스 캐릭터와 함께 즐기는 체험형 댄스 콘텐츠 시스템을 제안한다. 체험자는 화면에 순서대로 출력되는 동작 리스트와 같은 움직임을 취하게 되며 동작의 정확도 및 타이밍에 따라 차등 점수를 부여토록 하였다. 이러한 댄스 콘텐츠 시스템을 위해서는 각 동작에 대해 학습과 실시간으로 사용자의 동작을 구분하는 것이 필요하다. 다수개의 소형 무선 컨트롤러를 손과 발 등 주요부위에 착용하고 논문에 제안된 방법을 활용하게 되면 사용자의 흥미와 몰입감을 더해줄 것으로 기대된다.

핵심어: Accelerometer Sensor, Motion Recognition, Experiencing Contents

1. 서론

최근 개발되고 있는 체험류의 콘텐츠들은 기존 키보드나 마우스 조작에서 탈피하여 별도의 컨트롤러나 영상등을 이용하여 사용자의 직접 체험을 유도하고 있다[1]. 그러나 아직까지 버튼에 의존한 컨트롤러들이 대부분이며 영상을 입력으로 하는 콘텐츠는 영상에서 체함에 필요한 자세인식 등을 높은 수준의 정밀도로 얻기 힘들고 연산 비용 또한 매우 높아 대체로 단순한 게임에 적용되고 있다. 또한 사용자의 전체 모습을 영상으로 얻기 위해서는 카메라와 사용자와의 거리를 확보해야하며 동작을 위해서 보다 넓은 공간이 필요하게 된다.

이러한 문제점들을 해결하기 위해 가속도 센서 및 자이로 센서등을 콘텐츠에 응용하는 방법이 모색되고 있다. 가속도 및 각속도센서는 단위 시간당 속도와 각의 변화를 검출하는 소자로 다축 센서를 통해 외부 장치의 도움 없이 3차원 공간상에서 움직임 측정이 가능하다. 기존까지는 정확성이나 크기가 대중 콘텐츠에 적용하기 부족하였으나 최근 MEMS(Micro Electro Mechanical Systems) 기술의 발달로 소형화 및 저소비전력화 되고 있다. 특히 게임기 및 휴대폰 등 휴대전자기기로 적용되어 간단한 동작 패턴을 인지하여 이를 비트 박스의 음을 내는 휴대전화기나[2] 공간에 드럼을 치듯 움직여 북 소리를 내는 가상 멀티 타악기[3], 컨트롤러를 이동시켜 게임 콘텐츠로의 활용 등 사용자의 움직임을 감지하는 연구가 활발히 진행되고 있다[4]. 동작을 인식

하는 방법에 있어서는 HMM(Hidden Markov Model), 신경망(Neural Network Algorithm) 및 특징 기반 통계적 방법들을 이용하며 이들은 특정 제스처 모델을 미리 가정하여 인식하는 방법을 사용하고 있다[5-7]. 이처럼 모션 인식 기술을 응용하여 다양한 형태로 콘텐츠 개발이 시도되고 있으며 시장 또한 급속히 커져가고 있다.

본 연구에서는 체험형 댄스 콘텐츠 제작을 위해 관성 센서를 이용하여 체험자의 손 동작을 인식하고 정의된 동작과 비교하여 가상의 댄스 캐릭터와 함께 즐길 수 있는 시스템을 구현 하고자 한다. 그동안 체험형 콘텐츠로써 댄스를 활용하는 예는 예전부터 많이 있어왔다. DDR, PUMP, EZ2DJ[8] 등이 예이며 사용자에게 많은 인기를 누렸다. 하지만 위의 제품들은 고가의 전용 장비가 필요하고 큰 공간을 필요로 했기 때문에 특정 장소에서만 접할 수 있었다. 또한 발로 입력을 받는 시스템으로 오랜 기간동안 노출되어 있었기 때문에 빠르게 변화하는 콘텐츠 시장에서 변화가 필요하다.

따라서 3축 가속도 센서를 이용한 소형 컨트롤러를 손에 착용하고 여러 가지 동작을 취함으로써 체험자에게 다양한 입력 방식을 제공하고자 한다. 또한 소형 및 저가의 컨트롤러를 사용함으로써 어느 곳에서나 활용이 가능하고 발을 입력으로 하는 시스템에 비해 비교적 자유롭게 동작을 취할 수 있다. 뿐만 아니라 본 콘텐츠를 실시간 입체영상으로 구현함으로써 체험자에게 보다 높은 몰입감을 제공할 수 있다.

2. 제안 시스템 구성

제안된 체험형 댄스 콘텐츠는 사용자가 모션 컨트롤러를 이용하여 음악에 맞춰 순서대로 출력된 동작을 정확한 타이밍에 취함으로써 댄스와 리듬감을 즐길 수 있도록 하였다. 제안된 체험형 댄스 콘텐츠의 시스템 구성도는 <그림 1>과 같다.

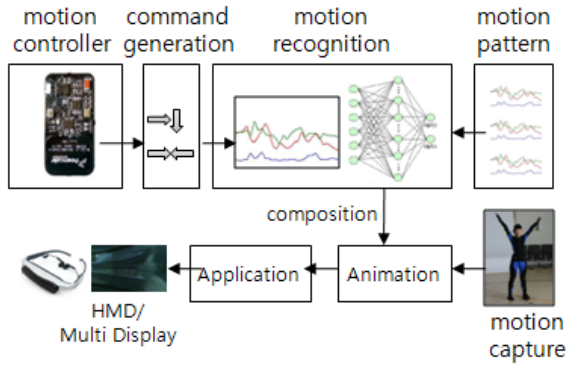


그림 1. 시스템 구성도

제안된 시스템은 크게 다섯 부분으로 구성된다. 첫째, 모션 캡처를 통한 댄스 캐릭터 생성, 둘째, 각 동작의 패턴 생성, 셋째, 체험자의 동작을 통한 가속도 데이터 획득, 넷째, 정의된 패턴과 체험자의 동작 인식, 다섯째, 실시간 입체영상 구현으로 구성된다.

체험자가 바라볼 댄스 캐릭터는 자연스러운 댄스 애니메이션을 얻기 위해 모션 캡처를 활용하였다. 또한 체험자의 동작 인식 결과에 따라 체험자에게 동작의 옳고 그름을 시각적으로 피드백을 주기위해 애니메이션을 임의의 노이즈와 합성하는 방법을 사용하였다. 동작 인식은 3축 가속도 센서가 설치된 모션 컨트롤러를 이용하여 체험자의 손 동작에 대한 가속도 값을 입력받고 이를 정의된 5가지 동작과 비교하게 된다. 이때 제시된 5가지 유형의 동작을 인식하기 위해 학습시킨 신경망을 이용하였으며 3축의 가속도 값의 크기로부터 특징 벡터를 추출하도록 하였다. 그리고 최종적으로 가상환경을 실시간 입체 영상으로 출력하여 체험자에게 보다 높은 흥미와 몰입감을 제공하도록 하였다.

2.1 댄스 캐릭터 제작

모션캡처 기술은 제작기간의 단축과 높은 퀄리티의 사실적인 애니메이션, 제작비용의 절감 등으로 현재 영화, 애니메이션, 게임, 재활의학 등 다양한 산업 분야에 활용되고 있으며 콘텐츠 제작에서는 이미 대중화가 되어있다. 특히 최근에 많은 수의 댄스 콘텐츠가 개발되었는데 모두 모션 캡처를 사용할 정도로 절대적이라 할 수 있다. <그림 2>는 모

션캡처를 활용한 댄스 콘텐츠들의 예를 보여준다.



그림 2. 모션캡처를 활용한 댄스 콘텐츠

본 연구에 사용된 모션캡처 시스템은 8대의 적외선 카메라가 장착된 디지털 방식의 Eagle 광학식 모션캡처 시스템을 이용하였다. 이 시스템은 사람의 동작을 초당 60 프레임 이상 캡처함으로써 데이터 손실 없이 빠른 시간에 실감 있는 디지털 동작을 획득하기에 적합하다. <그림 3>은 본 연구에 사용된 모션캡처 시스템의 모습이다.



그림 3 Eagle Motion Capture System

모션캡처 과정에서 생성된 디지털 모션 데이터는 33개의 3차원 공간상의 X, Y, Z 좌표를 가지는 점이다. 이 데이터는 Gap과 Spike 필터링 등의 에디팅 과정을 거치고 데이터 변환과정에서 뼈대 구조를 가지는 모션 데이터로 변환을 한 후에 3차원 캐릭터에 적용하여 활용한다[9].

<그림 4>는 모션캡처 데이터를 획득하기 위해 마커를 부착한 액터의 동작을 캐릭터에 적용 시키는 과정을 나타낸다.



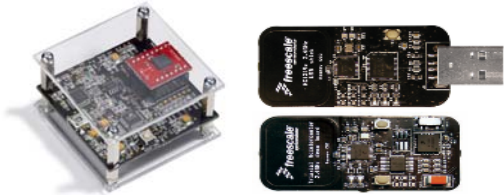
(a) 모션캡처 (b) 모션 데이터 (c) 모션 적용모델

그림 4. 모션 캡처 데이터 적용 과정

2.2 가속도 센서 하드웨어 구성

체험자의 손에 착용될 모션 컨트롤러는 경량 및 소형화,

동작의 간섭이 없는 무선통신, 긴 연속시간을 보장하기 위한 낮은 소비전력, 동작인식을 위한 정밀한 센서 내장, 다수개의 노드구성 가능, 탈부착 기능 등의 조건이 만족되어야 한다. 본 연구에서는 위의 기본 요건들을 대부분 만족시키는 VTI 사의 SCA3000 Demo Kit과 Freescale 사의 RD3157MMA7260Q Evaluation Board를 활용하여 테스트 하였다. 두 제품의 하드웨어 구성은 <그림 5>와 같고 각각의 제품에 내장된 가속도 센서의 자세한 사양은 <표 1> 및 <표 2>와 같다[10-11].



(a) VTI SCA3000 Demo Kit (b) Freescale RD3157MMA7260Q
그림 5. 센서 보드

표 1. 가속도 센서 SCA3000 D01 사양[10]

항목	사양	비고
Axis	x, y, z	3축
Range	±2.0g	g:중력가속도
Sensitivity	1333 count/g	ADC내장
Power voltage	2.35V ~ 3.6V	
Filter	Low-pass	필터 내장

표 2. 가속도 센서 MMA7260Q 사양[11]

항목	사양	비고
Axis	x, y, z	3축
Range	±1.5g/2g/4g/6g	Selectable Sensitivity
Sensitivity	800 mV/g @ 1.5g	ADC내장
Power voltage	2.2V ~ 3.6V	
Filter	Low-pass	필터 내장

2.3 동작 인식

본 장에서는 체험자가 취해야할 5가지 동작에 대해 정의하고 학습된 신경망으로부터 체험자의 제스처를 인식하는 과정과 이를 활용한 인식 성능 알고리즘에 대해 설명한다.

2.3.1 동작 방법

체험자는 <그림 6>과 같이 다섯 동작으로 이루어진 동작 표가 검사 영역에 도달했을때 그와 같은 동작을 취하게 된다. 동작 표는 음악의 박자에 맞춰 자동 생성되며 시작점

에서 검사 영역으로 빠르게 이동하여 순발력과 리듬감을 느낄 수 있도록 하였다. 본 연구에서는 실험을 위해 하나의 컨트롤러를 이용하였으며 왼손을 통해 제스처 인식을 하였다.

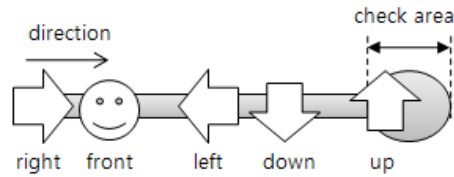


그림 6. 동작 표 및 인식구간

두 가지가 콘텐츠에 반영되는데 체험자가 주어진 동작 표와 동일한 동작을 취했는지와 정확히 검사 영역에 도달했을때 동작을 취했는지를 판별하게 된다. 그리고 실제 동작을 취한 시간보다 가속도 데이터의 전송이 느리기 때문에 약간의 지연 시간을 고려하여 검사하게 된다. 또한 노이즈나 작은 움직임에도 동작으로 오인할 수 있기 때문에 일정 크기 이상의 가속도 발생에만 인지할 수 있도록 하였다.

2.3.2 동작 정의

본 연구에서 구분을 목표로 하는 5개의 동작은 <그림 7>과 같다. 모든 동작은 몸 가운데에서 시작하며 상, 하, 좌, 우, 앞으로의 움직임을 구분해야 한다. 사용자마다 움직임의 범위와 행동 패턴이 너무나도 다양하기 때문에 본 연구에서는 사용자가 움직이는 데에 불편함이 없는 범위 안에서 아래와 같은 제약 조건을 두었다.

- (제약 조건 1) 동작은 되도록 기본자세에서 시작 하여 기본자세로 끝나야 한다.
- (제약 조건 2) 동작은 되도록 직선 이동을 해야 한다.
- (제약 조건 3) 1회의 동작은 정해진 시간 내에 움직여야 한다.

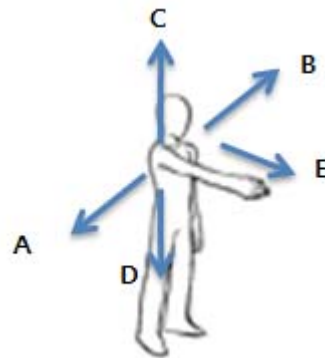








그림 7. 5가지 방향의 체험자 움직임

기본자세를 포함한 각 여섯 동작은 <표 3>과 같다.

표 3. 체험자 동작 정의

동작	방향	동작	방향	동작	방향
	기본 자세		좌측		우측
	상향		하향		정면

2.3.3 패턴 정의

컨트롤러에서 얻어지는 원 데이터는 미세한 움직임에도 가속도가 발생하여 위치추적이나 동작인식에 문제가 발생할 수 있으므로 이전 프레임과의 평균화 필터링을 통해 노이즈를 제거하였다. 필터링으로 데이터의 사실성과 정밀도가 떨어지지만 단방향 운동에 대해 동작을 인식 하기 때문에 오히려 큰 움직임만을 얻을 수 있는 이점이 있다. <그림 8>은 필터링 적용 전과 필터링 적용 후의 가속도 신호의 그래프를 나타낸다.

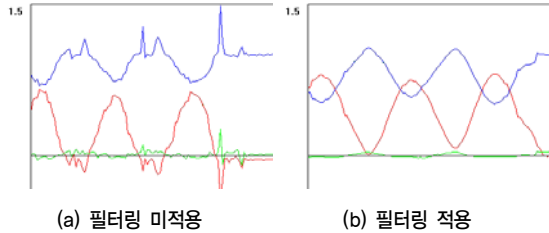


그림 8. 필터링

<그림 9>는 각각의 동작에 따른 가속도 신호 유형을 그래프로 나타낸 것이며 <그림 10>은 체험자가 동작을 취하는 모습이다.

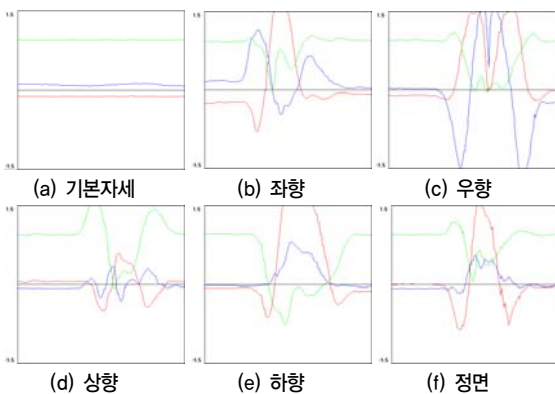


그림 9. 각 동작별 가속도 그래프



그림 10. 동작 패턴

2.3.4 동작 인식

본 연구에서는 가속도 센서의 신호를 10ms 마다 획득하도록 하였으며 +2g와 -2g 사이의 값을 255레벨로 샘플링 하였다. 획득한 각 축의 가속도 신호로부터 연속된 세 신호 획득 구간, 즉 $t-1$, t , $t+1$ 구간에서 가속도 변화가 임계치보다 큰 경우 모션이 발생한 경우로 간주하며 (1)과 (2)를 이용하여 3축 각각의 가속도 신호 크기를 구한다.

$$|A_t(x,y,z) - A_{t-1}(x,y,z)| > T \quad (1)$$

$$|A_{t+1}(x,y,z) - A_{t-1}(x,y,z)| > T \quad (2)$$

여기서, A , A_{t-1} , A_{t+1} 은 $t-1$, t , $t+1$ 에 대한 각 신호의 샘플링 프레임을 나타내며, T 는 임계치를 나타낸다.

동작이 발생한 시점으로부터 최대 1.2초 구간을 동작 구간으로 정하고 이 전 상태를 인지하기 위해 동작 발생 시점에서 5 프레임을 동작구간에 추가하였다. 그리고 15 프레임 동안 일정 값 이상의 가속도 신호의 변화가 발생하지 않을 경우 데이터 큐(Queue)로의 저장을 멈추게 된다. 또한 체험자가 취해야할 동작 표가 검사영역에 도달하지 않은 상태에서 동작이 이뤄질 경우에는 인지를 하지 않도록 하였다.

임계값에 의한 모션 정보를 이용하여 검출된 동작 구간은 <그림 11>에서와 같다.

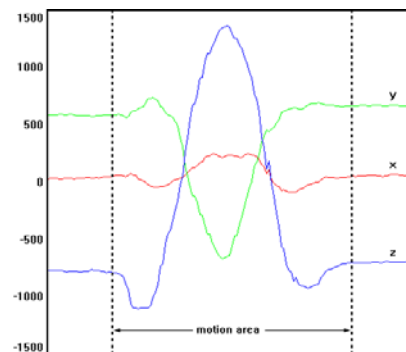


그림 11. 동작 구간

기본자세를 기준으로 하여 센서의 연속 프레임으로부터 추출된 가속도 신호 x , y , z 의 크기의 변위를 구하고 이를 중심으로 제스처 인식을 위한 특징 벡터로 추출하여 사용하였다.

다섯 가지 유형의 동작을 인식하기 위해 한 개의 은닉층을 가지는 MLP(Multi-Layer Perceptron)를 이용하였으며 개선된 backpropagation 알고리즘[6]으로 구현하였다. 사용된 알고리즘에서는 P개의 입력 패턴과 N개의 출력 뉴런 사이에서 목적값과 실제 출력값 사이의 오차 누적값을 계산하고 전체 오차값이 미리 정의된 임계치보다 작은 값을 가질 때까지 가중치를 계속 갱신하도록 하는 방법을 사용하였다 [5].

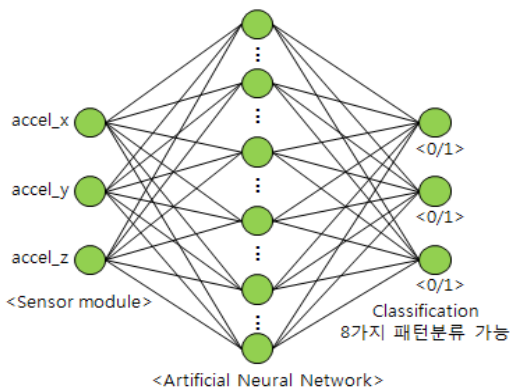


그림 12. 신경망 알고리즘에 의해 동작 패턴이 분류되는 과정

〈그림 12〉는 신경망 알고리즘에 의해 동작 패턴이 분류되는 과정이다. 센서를 통해 가속도 데이터가 측정되면 3축의 데이터는 샘플링 과정을 거쳐 사전에 학습 데이터에 의해 학습이 이루어진 신경망의 입력으로 들어간다. 신경망 은닉층의 개수는 10~20개 정도로 변화하며 실험한 결과 12개 이후부터는 성능향상이 없었기 때문에 12개로 고정하였다 [12].

3. 시스템 구현 및 실험

3.1 콘텐츠 적용

구현된 댄스 콘텐츠 시스템의 하드웨어 구성은 〈그림 13〉과 같다.

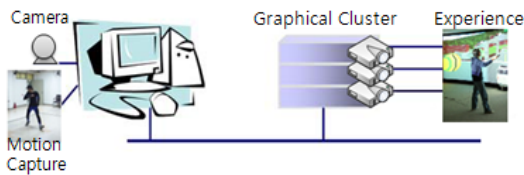


그림 13. 시스템 환경

보다 사실적인 콘텐츠 체험을 위해 멀티 디스플레이 시스템에 화면을 출력하게 되는데 콘텐츠에 적용된 다수의 뷰 카메라를 그래픽 클러스터를 통해 각각의 채널로 영상을 분리하여 출력하게 된다.

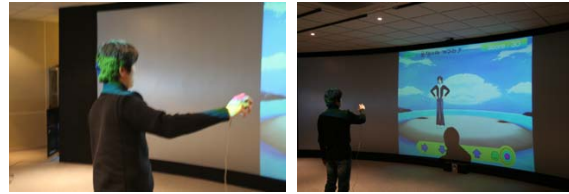
제작된 체험자 댄스 콘텐츠는 실시간 입체영상으로 보여지게 되는데 빠른 시간에 입체영상을 적용시키기 위하여 VR 엔진인 Virtools를 사용하였다[13]. 스크립트 프로그래밍 방식인 Virtools는 기능별로 구현된 빌딩 블록의 조합으로 콘텐츠를 쉽고 빠르게 작성할 수 있으며 리소스를 비주얼하게 관리할 수 있다.

〈그림 14〉는 Virtools를 이용하여 실시간 입체영상을 제작하기 위한 과정이다.



그림 14. 제작 과정

〈그림 15〉는 본 연구에서 제작한 컨트롤러를 활용하여 멀티 디스플레이 시스템 환경에서 콘텐츠에 적용, 시연 모습이다.



(a) 시연중인 체험자 모습 (b) 게임 화면
그림 15. 콘텐츠 시연 모습

3.2 실험 및 결과

제스처 인식 성능을 평가하기 위하여 그림 5에서 나타난 다섯 가지 종류의 제스처에 대해 인식성능을 평가 하였다. 단일 동작의 제스처 데이터를 학습시킨 다음 연속 동작의 제스처 프레임으로부터 한 동작의 제스처를 분리한 다음 학습시킨 신경망을 이용하여 인식 성능을 평가하도록 하였다. 제스처 학습 데이터베이스는 5명의 사람으로부터 5가지 자세를 통해 3축의 신호 크기를 추출하여 구현하였으며 이 데이터는 제스처 인식기를 학습하는데 사용되었다.

학습 결과에 대한 평가를 하기 위해 학습 실험자의 단일 동작에 대해 동작 인식 테스트를 하였다. 실험자는 다섯 동작에 대해 기본자세를 취한 후 일정 시간 뒤 무작위로 출력된 동작 순서에 맞게 각각 30번씩 테스트 하도록 하였다. 〈표 4〉는 인식 결과에 대한 Confusion Matrix를 나타낸 것이다.

표 4. 다섯 동작의 인식 결과에 대한 Confusion Matrix

인식 대상 \ 인식	인식 대상					인식대상 동작 수	인식 동작 수
	C[0]	C[1]	C[2]	C[3]	C[4]		
좌향 C[0]	27	1	2	-	-	30	27
우향 C[1]	-	26	-	-	4	30	26
상향 C[2]	3	2	24	-	1	30	24
하향 C[3]	-	-	-	29	1	30	29
정면 C[4]	-	1	3	-	26	30	26
인식률 (%)	90	87	80	97	87	150	132

위 실험은 학습 대상자가 인식 실험을 하였고 다음 동작을 위해 기본자세를 취할 시간이 충분하였기 때문에 비교적 높은 인식률을 보였다. 따라서 학습에 참여하지 않은 피 실험자 5명에 대해 콘텐츠 진행 상황과 같은 환경에서 시뮬레이션 해보았다. 동작 표는 5분간 1초에서 3초 사이를 두고 랜덤하게 출력되게 하였다. <표 5>는 피 실험자 5명의 시뮬레이션에 의한 동작인식 결과이다.

표 5. 시뮬레이션 인식률

피실험자	총 동작 수	인식률(%)				
		좌향	우향	상향	하향	정면
1	167	93	86	89	94	89
2	145	85	69	91	91	83
3	187	89	73	89	88	85
4	180	87	88	80	93	87
5	153	91	89	84	93	84
평균	166	89.00	81.00	86.60	91.80	85.60

왼손에 컨트롤러를 쥐고 우측으로 이동하게 되면 동작이 부자연스럽기 때문에 동작을 크게 하지 않을 경우 인식률이 낮게 나온 결과가 발생하였다. 이는 양손 착용 형태의 모델로 확장하여 오른손을 이용하게 되면 해결 될 것이다.

4. 결론

본 연구에서는 체험자가 컨트롤러를 착용하고 음악에 맞춰 흐르는 동작 순서를 취함으로써 기존의 단순 버튼 입력을 탈피하고 직관적이며 사용자의 조작감과 몰입감을 높일 수 있도록 하였다. 체험자의 동작 인식은 가속도 센서로 이루어진 센서 모듈과 측정된 데이터를 이용하여 동작 패턴을 분류해 주는 신경망 알고리즘으로 구성되었다. 실험 결과 비

교적 높은 인식률이 측정되었으나 대부분 단순한 동작으로 이루어져 있어 빠르고 복잡한 패턴이 있는 상황에서 올바른 동작 인식은 쉽지 않을 것이다. 따라서 다양한 체험형 콘텐츠로의 적용을 위해 차후 가속도 센서를 활용하는 방법과 멀티 컨트롤러로의 확장, 콘텐츠 체험에 앞서 체험자의 동작을 높은 가중치로 학습에 적용시켜 인식률을 높이는 방법에 대해 연구하고자 한다. 본 시스템을 통하여 다양한 체험형 콘텐츠로 적용시 사용자의 흥미와 몰입감을 더해줄 것으로 기대된다.

참고문헌

- [1] Johanna H ysnieni, Perttu H m l inen, Laura Turkki, "Wizard of Oz Prototyping of Computer Vision Based Action Games for Children", IDC 2004, College Park, Maryland, USA, June 1-3, 2004.
- [2] E.-S. Choi, W.-C. Bang, S.-J. Cho, J. Yang, D.-Y. Kim, and S.-R. Kim, "Beatbox Music Phone: Gesture-based Interactive Mobile Phone using a Tri-axis Accelerometer", In Proceedings of IEEE International Conference on Industrial Technology 2005 (ICIT2005), Hong Kong, December 14-16, 2005, pp. 97~102.
- [3] 최은석, 손준일, 방원철, 김연배, "관성 센서를 이용한 위치기반 가상 멀티 타악기", HCI2006 논문집 1권, 2006, pp.379~385.
- [4] P. Keir, J. Elgoyhen, M. Naef, J. Payne, M. Horner, P. Anderson, "Gesture-recognition with Non-referenced Tracking", IEEE Symposium on 3D User Interfaces 2006 (3DUI'06), Alexandria, Virginia, USA, March 25-26 2006.
- [5] 김계경, 김혜진, 조수현, 이재연, "로봇 동작 제어를 위한 제스처 인식", IPIU, 2004.
- [6] M. Brand, N. Oliver, and A. Pentland, "Coupled Hidden Markov Models for Complex Action Recognition," IEEE Proc. Computer Vision and Pattern Recognition(CVPR), 1997, pp.994-999.
- [7] B. Kosko, "Neural Networks and Fuzzy System," Prentice-Hall, Englewood Cliffs, NJ, 1992.
- [8] 정승우, "차세대 게임 인터페이스", 한국특허정보원, 2004.12.
- [9] 김남호, "모션캡처 시스템을 이용한 모바일 3D 댄스 콘텐츠 제작 연구", 한국콘텐츠학회논문지, 2006.
- [10] SCA3000D01, <http://www.vti.fi/>.
- [11] MMA7260Q, <http://www.freescale.com/>.
- [12] 임성빈, "관성센서 모듈과 신경망 알고리즘을 이용한 동작인식 시스템의 개발", 2006.
- [13] Virtools, <http://www.virttools.com/>.