

석사학위논문

키넥트를 이용한
인터랙티브 멀티미디어 작품 제작 연구
(멀티미디어음악작품 <Kinect Illusion>을 중심으로)

지도교수 김 준

동국대학교 영상대학원
멀티미디어학과 컴퓨터음악전공
윤기선

2012

석사학위논문

키넥트를 이용한
인터랙티브 멀티미디어 작품 제작 연구
(멀티미디어음악작품 <Kinect Illusion>을 중심으로)

윤기선

지도교수 김준

이 논문을 석사학위논문으로 제출함.

2012년 1월 일

윤기선의 음악석사학위(컴퓨터음악전공) 논문을 인준함.

2012년 1월 일

위원장: 김정호 (인)

위원: 조형제 (인)

위원: 김준 (인)

동국대학교 영상대학원

목 차

I. 연구의 배경과 목적	1
1. 연구 배경	1
1) 예술에서의 디지털 멀티미디어 기술	1
2) 사례 연구	2
2. 연구 목적	6
II. 기술적 연구	8
1. 키넥트 작동 원리 연구	8
2. 키넥트의 모션 트래킹 연구	9
3. OSC 통신 연구	12
III. 연구 기술의 작품 적용 및 효과	16
1. 작품의 의미와 기술적 접목	16
2. 작품의 구성	17
1) 작품의 전체적인 구성	17
2) 작품의 시스템 구성	18
3) 작품의 무대 구성	19

3. 배우의 움직임에 따른 사운드 효과	21
4. 샤 막을 이용한 입체 영상 효과	24
5. 영상과 퍼포먼스의 구성 및 제작	25
1) section A에서의 기술 적용	25
2) section B에서의 기술 적용	27
3) section C에서의 기술 적용	27
4) section C'에서의 기술 적용	29
5) section D에서의 기술 적용	30
6) section B'에서의 기술 적용	30
 IV. 문제점 및 향후 계획	 32
 참고문헌	 34
 Abstract	 36
 부록 1 : Max/MSP 패치	 38
부록 2 : Quartz Composer 패치	42
부록 3 : 첨부 DVD설명	45

표 목 차

[표-1] MS Kinect SDK와 OpenNI 비교	10
[표-2] 작품의 전체적인 구성	17

그 립 목 차

[그림-1] 사진을 이용한 말의 동작 분석	2
[그림-2] 기계식 모션 캡처 시스템	3
[그림-3] 영화 제작에서의 광학식 모션 캡처 사례	4
[그림-4] 공연에서 모션 트래킹이 활용된 사례	5
[그림-5] Xbox용 키넥트 센서	8
[그림-6] 키넥트에서 쏘는 적외선 픽셀들	9
[그림-7] OpenNI 개요	10
[그림-8] Skeleton API의 신체 인식	11
[그림-9] Synapse app을 통해 트래킹 된 3D depth, Skeleton image	12
[그림-10] OSC를 통한 데이터 흐름도	13
[그림-11] OSC를 통해 받은 트래킹 데이터의 x, y, z값을 추출	14
[그림-12] OSC를 통해 Quartz Composer로 보내지는 트래킹 데이터	14
[그림-13] Quartz Composer에서 받는 트래킹 데이터	15
[그림-14] 시스템 구성도	18
[그림-15] 무대 구성도	20
[그림-16] 손의 위치 값으로 그래놀러 합성 사운드 제어	21
[그림-17] 손의 위치 값으로 주파수변조 합성 사운드 제어	22
[그림-18] 발의 위치 값으로 테이프음악에 적용되는 이펙트 제어	23

[그림-19] 샤 막을 이용해 막 앞/뒤의 배우가 공존하는 효과	24
[그림-20] section A의 퍼포먼스와 영상	25
[그림-21] 오디오 입력을 받아 Particle의 Max Size에 적용	26
[그림-22] [그림-21]의 패치들을 Macro로 묶어 3D Transformation을 적용	26
[그림-23] section B의 퍼포먼스와 영상	27
[그림-24] section C의 퍼포먼스와 영상	28
[그림-25] 가시딩굴 영상 Quartz Composer 패치	28
[그림-26] section C'의 퍼포먼스와 영상	29
[그림-27] section D의 퍼포먼스와 영상	30
[그림-28] section B'의 퍼포먼스와 영상	31

I. 연구의 배경과 목적

1. 연구 배경

1) 예술과 디지털 멀티미디어 기술

현대 사회의 첨단 디지털 기술의 발달은 우리의 문화생활에 많은 변화를 가져왔다. 특히 예술분야에서 그 변화가 두드러지게 나타나고 있다. 기술적으로 표현이 불가능해 창작자들의 머릿속에만 있던 창조적인 표현들이 기술의 발달로 다양한 표현이 가능해졌다. 예술의 범위는 넓어졌고, 장르 또한 모호해지게 되었다. 인간의 상상력을 현실로 이끌어낸 멀티미디어 기술의 발달은 각기 다양하게 표현되는 미디어를 하나의 예술로 통합하여 상호작용을 가능하게 했다. 이러한 멀티미디어 기술과 예술의 조화로 표현되는 작품에서 가장 중요한 것은 서로간의 상호작용이다. 이전의 많은 예술 작품들이 창작자가 만들어놓은 작품을 연주자, 무용수(performer) 또는 감상자가 일방적으로 받아들이고 감상하는 관계였다면, 멀티미디어 기술이 가져온 인터랙티브(interactive) 예술은 작곡가와 연주자, 혹은 창작자와 감상자가 끊임없이 소통하는 예술이다. 이러한 상호작용을 통해 작품의 예술성은 배가된다.

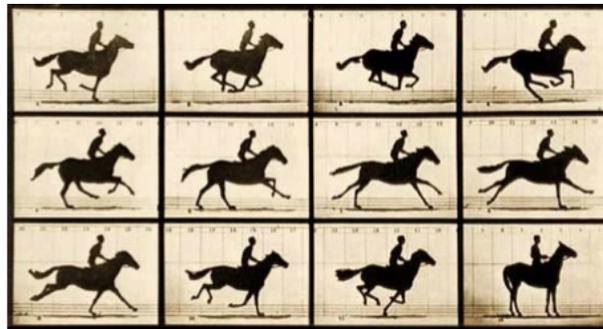
인간은 시각, 청각, 후각, 미각, 촉각의 5가지 감각을 가지고 있다. 흔히들 생각하는 예술 작품은 주로 시각과 청각을 통해 전달된다. 그림 혹은 영상은 시각에 의존한 예술이고, 음악은 청각에 의존한 예술이다. 그러나 현대의 멀티미디어 매체들은 대부분 한 가지의 감각에만 의존하여 감정이나 정보를 전달하지 않는다. 하나의 감각이 다른 영역의 감각을 불러일으키는 것을 공감각이라 하는데, 이러한 공감각적 요소를

답은 현대의 예술 작품에서 창작자가 그 작품에서 표현하고자 하는 의미를 관객이 정확히 파악하지 못해 혼란에 빠지는 경우가 종종 있다. 예술 작품이 예술로서 가치를 인정받기 위해서는 창작자가 표현하고자 하는 의미가 제대로 전달되어야 한다. 때문에 멀티미디어 예술은 각각의 매체들 사이의 상호작용을 적절히 활용하여 작품의 의미를 효율적으로 표현해야 한다. 시각과 청각, 시각과 촉각 등의 다양한 감각으로 창작자의 풍부한 상상력을 현실로 가져올 수 있다.

이러한 멀티미디어 요소들을 결합하여 배우, 무용수와 작품과의 상호작용에 적용하기 위하여 모션 트래킹¹⁾을 사용한다.

2) 사례 연구

모션 트래킹의 일종인 모션 캡처란 3D 공간상에서 피사체의 움직임에 대한 방위를 측정하고, 컴퓨터가 사용할 수 있는 형태의 정보로 기록하여 분석 및 응용하는 기술을 일컫는다.



[그림-1] 사진을 이용한 말의 동작 분석

1) 대상체의 움직임에 대한 데이터를 추출하고 컴퓨터가 사용할 수 있는 형태의 정보로 기록하여 분석 및 응용하는 기술.

[그림-1]과 같이 군사적 목적으로 사진을 이용하여 사람과 동물의 움직임을 연구한 것이 모션 캡처의 기원이라 할 수 있다.

모션 캡처 시스템은 데이터를 추출하는 방식에 따라 광학식, 기계식, 자기식으로 구분된다. 기계식은 [그림-2]와 같이 인체의 각 관절의 움직임을 기계장치 부착을 통해 캡처하는 방식으로, 부착된 기계로 인한 행동의 제약이 없다면 가장 이상적인 경우라 할 수 있다. 그러나 고 중량의 기계 장치를 부착함에 따라 캡처하는 대상이 자연스러운 움직임에 제약을 받게 되고, 이로 인해 동작 데이터의 추출에도 한계가 있다.

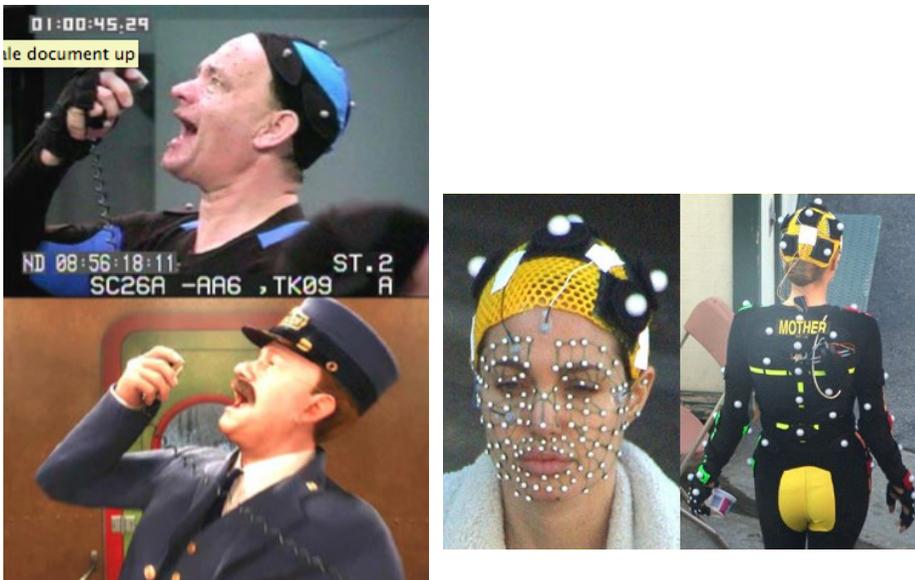


[그림-2] 기계식 모션 캡처 시스템

자기식 모션 캡처 시스템은 대상의 각 관절에 자기장을 발생하는 센서를 부착하고 대상의 움직임에 따른 자기장의 변화를 측정하여 위치 데이터를 추출하는 방식이다. 혹은 그 반대로 자기장의 계측 장치를 대상에 부착하고 자기 발생 장치를 이용하기도 한다. 자기식 모션 캡처 시스템은 광학식에서 발생하는 문제점인 마커(marker)가 가려져 데이터가 손실되는 경우가 없고, 광학식에 비하여 장비의 가격 및 유지비가 저렴하다는 장점이 있다. 하지만 케이블을 사용하기 때문에 대상의 동작에 제약이 있고, 주위의 금속물체 등에 의해 데이터 손실이 발생할 우려가 있다는 단점이 있다.

광학식 모션 캡처 시스템은 대상의 움직임을 얻어내기 위한 적절한 부위에 반사 마커(marker, 카메라에서 발산하는 적외선을 반사)를 부착하고 여러 대의 카메라가 2차원 이미지로 피사체를 촬영한 후, 그 이미지를 다시 3차원 위치 데이터로 계산하여 데이터를 추출하는 방식이다. 이 방식은 현재 활용되고 있는 모션 캡처 시스템 중에서 가장 진보된 형태이다. 광학식 모션 캡처는 고속으로 촬영할 수 있어 유실되는 데이터가 거의 없고, 대상의 움직임을 제약하는 요소가 없어 자유로운 동작 표현이 가능하여 매우 섬세한 동작을 추출해 낼 수 있다.

다음의 [그림-3]은 실제 영화 제작에서 사용된 광학식 모션 캡처의 사례이다. 왼쪽 그림은 애니메이션 “The Polar Express”에서 톰 행크스의 퍼포먼스 캡처 장면과 최종 영상이다. 오른쪽 그림은 영화 “Beowulf”에서 안젤리나 졸리의 모션 캡처 장면이다.



[그림-3] 영화 제작에서의 광학식 모션 캡처 사례

최근에는 영화나 애니메이션뿐만 아니라 공연 예술에서도 모션 트래킹을 사용한 사례를 자주 찾아볼 수 있다.



[그림-4] 공연에서 모션 트래킹이 활용된 사례

[그림-4]의 왼쪽 그림은 무용수가 들고 움직이는 손전등(라이트) 불빛의 움직임을 트래킹하여 공연에서 활용하는 사례이다. 오른쪽 그림도 같은 원리로, 무용수가 들고 있는 막대의 끝에 부착되어있는 적외선 LED의 움직임을 트래킹 하여 공연 영상에 적용시킨 사례이다. 이 두 사례 모두 광학식 모션 캡처 시스템의 반사체 대신 직접 발광체를 이용한 시스템이라 할 수 있다.

2. 연구 목적

현대 멀티미디어 작품들은 배우나 무용수의 움직임(퍼포먼스) 및 예술적 표현이 작품과 상호작용해 음악과 영상 예술로 표현되는 방법을 많이 이용하고 있다. 이런 방법은 실시간으로 이루어지며 매번 다른 결과물이 추출되는 특징이 있다. 이를 인터랙티브 아트라고 부른다.

한편, 보급되는 컴퓨터와 카메라의 처리 속도가 빨라지게 되면서 멀티미디어 작가들은 배우, 무용수의 움직임을 트래킹하기 위해서 여러 가지 트래킹 방법들을 사용하기 시작했다. 지정된 색깔만을 트래킹하는 컬러 트래킹(color tracking)²⁾을 주로 사용하기 시작하여 최근에는 적외선 LED를 트래킹³⁾하거나, 센서 또는 마커 등을 부착하여 퍼포머의 움직임을 트래킹하여 왔다. 하지만 컬러 트래킹이나 적외선 LED, 광학식 모션 트래킹을 하기 위해서는 트래킹을 위한 색깔에 맞는 의상을 착용하거나 적외선 LED를 위한 장치, 센서, 마커 등을 부착해야 하는 등 움직임이나 표현에 있어 많은 제약이 따른다. 때문에 트래킹 데이터가 정확하지 않거나 트래킹 데이터의 종류가 LED, 센서, 마커 등의 수량에 따라 한정될 수밖에 없다.

본 연구에서는 이러한 단점들을 보완할 수 있는 키넥트(Kinect)⁴⁾를 사용하여 트래킹을 시도하였다. 키넥트와 컴퓨터만의 간단한 시스템 구성으로 퍼포머가 아무 장치 없이 자유롭게 공연할 수 있기 때문이다. 또한 움직임에 따라 동시에 트래킹 되는 관절 데이터의 수가 많고,

2) 디지털 비디오카메라(DV cam), 웹 캠(web cam) 등 컴퓨터와 연결된 외부영상장치로부터 입력된 영상의 색상을 분석하고 지정된 색상의 위치를 추적하는 기술.

3) 적외선 필터를 장착한 후, 카메라로 촬영된 영상이 컴퓨터에 입력되어 적외선 LED의 위치 값만을 추적하는 기술.

4) 컨트롤러 없이 이용자의 신체를 이용하여 게임과 엔터테인먼트를 경험할 수 있는 MS Xbox360과 연결해서 사용하는 주변기기.

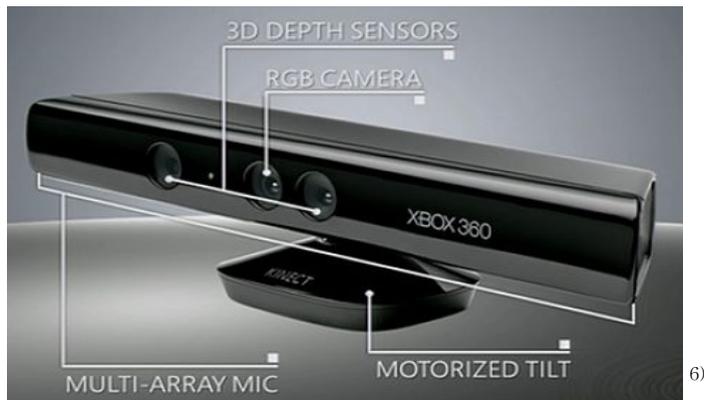
depth image 및 RGB image⁵⁾도 함께 사용할 수 있기 때문이다.

5) Red, Green, Blue 세 가지 빛을 기본으로 혼합하여 색을 만드는 영상. 일반적인 카메라 영상.

II. 기술적 연구

1. 키넥트 작동 원리 연구

배우의 움직임과 동작을 인식하기 위해서는 카메라를 이용한 트래킹이 필요하다. 이를 모션 트래킹이라고 하는데 일반적으로 특정 색깔, LED, 센서, 마커 등을 이용한다. 하지만 이러한 데이터 추출은 여러 가지 환경에 따라 발생하는 변수가 많고, 추출할 수 있는 데이터의 종류도 한계가 있기 때문에 많은 부족함이 있다. 본 연구에서는 이러한 불편함과 한계를 벗어나기 위해 키넥트를 이용한다.

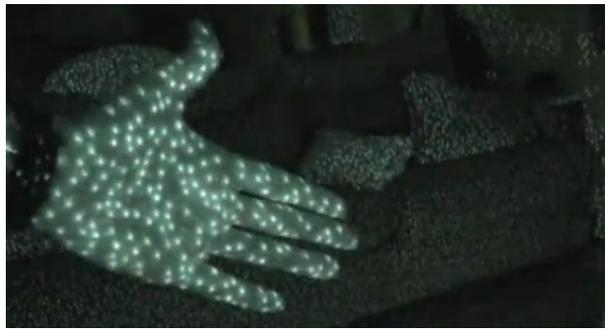


[그림-5] Xbox용 키넥트 센서

키넥트는 [그림-5]에서 보는 바와 같이 기본적으로 3개의 렌즈로 구성되어있다. 왼쪽과 오른쪽 끝에 있는 렌즈가 3D depth sensors이다. [그림-6]과 같이 왼쪽의 적외선 프로젝트에서 적외선을 픽셀 단위의

6) <http://www.xbox.com/kinect>

무수히 많은 점들로 쏘아주면 오른쪽의 적외선 카메라가 피사체에 반사 되어 돌아오는 것을 인식한다. 이것이 깊이 인식 카메라, 적외선 카메라이다. 가운데 렌즈는 RGB 카메라로 일반 카메라와 같은 RGB 영상을 인식한다. 이와 같은 키넥트는 Xbox용의 게임 컨트롤러뿐 아니라, USB 인터페이스를 가지고 있는 PC와의 연결을 통한 활용이 가능하다.



[그림-6] 키넥트에서 쏘는 적외선 픽셀들

2. 키넥트의 모션 트래킹 연구

게임 콘솔 Xbox용인 키넥트를 컴퓨터에서 사용하기 위해서는 두 가지 개발 환경이 있다. 마이크로소프트(MS, microsoft)에서 배포하는 키넥트용 SDK(software development kit)와 키넥트의 비공식 SDK라고 할 수 있는 오픈소스(open source)⁷⁾ OpenNI⁸⁾가 있다. 다음의 [표-1]은 마이크로소프트의 키넥트용 SDK와 오픈소스인 OpenNI를 비교한 내용이다.

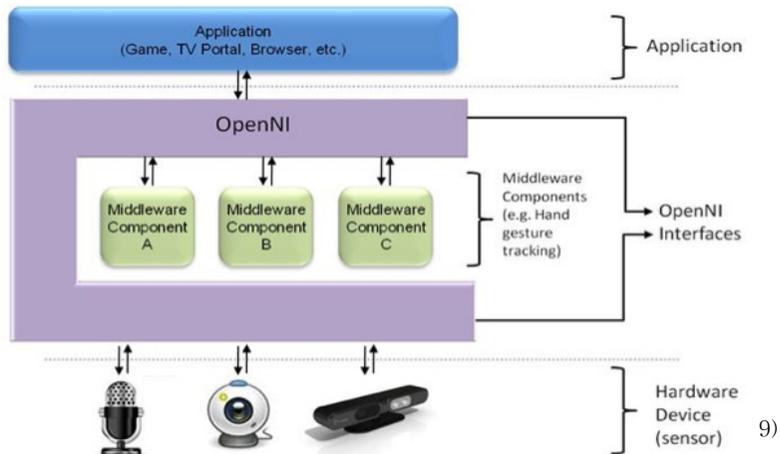
7) 무상으로 공개된 소스코드 또는 소프트웨어

8) 키넥트를 개발하고 있는 PrimeSense를 중심으로 개발하고 있는 API(application programming interface)의 집합

[표-1] MS Kinect SDK와 OpenNI 비교

	MS Kinect SDK	OpenNI
OS	Only Windows 7	Windows XP, Vista, 7 Mac OS X, Linux
Programming Language	C++, C#	C, C++, C#, Java
License	비상용으로만 사용가능	상용으로 사용가능
Tracking Range	0.8 ~ 4m	0.5 ~ 9m

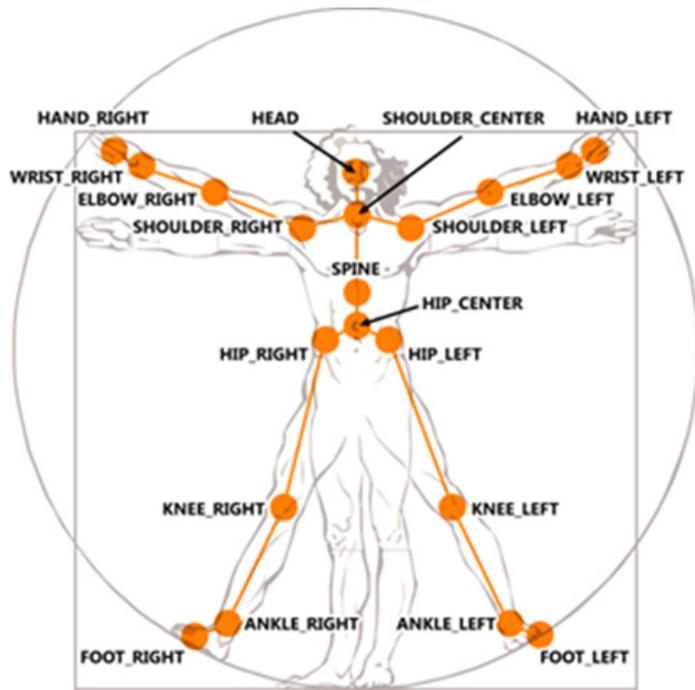
[표-1]에서와 같이 OpenNI는 마이크로소프트의 키넥트용 SDK에 비해 호환되는 OS가 다양하고 트래킹 범위도 더 넓다. 또한 마이크로소프트 키넥트 SDK는 상용으로 사용이 불가능하지만, OpenNI는 상용으로도 사용이 가능하다. 때문에 본 연구에서는 MacBook Pro를 사용하여 조금 더 자유로운 환경에서 프로그래밍하기 위해 OpenNI를 사용하였다.



[그림-7] OpenNI 개요

[그림-7]은 응용프로그램들과 하드웨어를 연결해주는 OpenNI의 전체적인 개요이다.

OpenNI 라이브러리에서 제공되는 API 중 Skeleton API를 이용해 관절을 트래킹하였다. 이 API는 키넥트 앞의 사람을 최대 두 명까지 뼈대 형태로 인식한다. 단순히 형태로만 인식하는 것이 아니라 신체의 각 부분을 정확하게 인식하는 것이다. [그림-8]은 Skeleton API에서 인식하는 신체의 각 부분이다.



10)

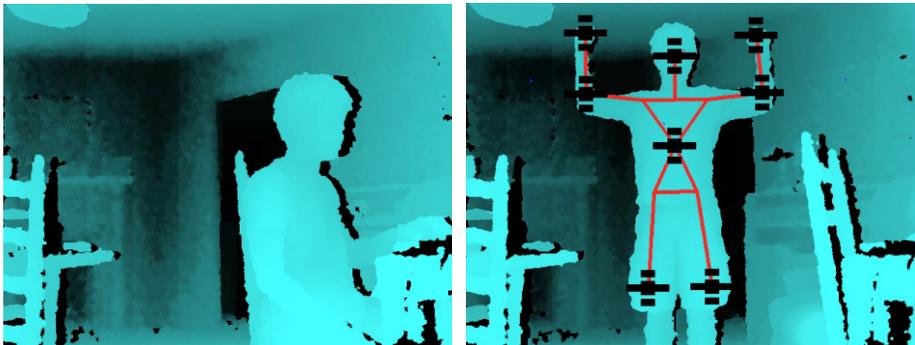
[그림-8] Skeleton API의 신체 인식

키넥트가 배우의 움직임을 촬영하면 OpenNI 라이브러리로 프로그래

9) OpenNI Userguide

10) Kinect SDK Programming Guide

밍 된 Synapse app이 키넥트의 관절 트래킹 데이터와 3D depth image를 추출한다. 키넥트가 트래킹 할 수 있는 데이터는 [그림-8]과 같이 종류가 더 많지만, Synapse app에서는 손, 팔꿈치, 머리, 몸통, 발, 무릎의 관절만을 트래킹 한다. 트래킹 된 관절 값은 OSC¹¹⁾를 통해 Max/MSP¹²⁾로 보내지며, 각 관절 별로 데이터가 분류된다. [그림-9]는 Synapse app을 통해 촬영 된 3D depth image 및 각 관절과 몸통을 트래킹 한 모습이다. 각각의 데이터들을 사용해 실험해 본 결과 손, 팔꿈치, 머리, 몸통의 데이터가 가장 효과적인 것으로 나타났다. 이러한 이유에서 이 네 가지 데이터를 작품에 적용하였다.



[그림-9] Synapse app을 통해 트래킹 된
3D depth image, Skeleton image

13)

3. OSC (Open Sound Control) 통신 연구

배우의 움직임을 추적하여 얻어진 데이터는 소리와 영상을 만들기 위

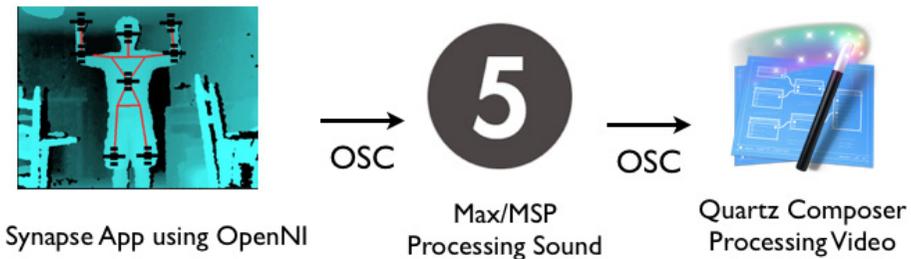
11) 음악과 관련된 다양한 데이터를 프로토콜을 이용해 송수신 하는 방식.

12) Cycling74에서 제작한 MIDI, 사운드 컨트롤 응용프로그램.

13) <http://synapsekinect.tumblr.com>

해 각 프로그램으로 보내진다. 소리는 Max/MSP를 이용하고, 영상은 Quartz Composer¹⁴⁾를 이용하여 제작한다. Synapse app에서 트래킹 된 데이터를 Max/MSP와 Quartz Composer로 전송해야 하는데 이를 위해 OSC 전송방식을 사용하여 데이터를 송수신 한다.

OSC 전송방식은 유무선상에 네트워크가 설치되어 있고 IP 주소가 정확할 경우, 하드웨어와 소프트웨어 사이의 데이터들을 빠른 속도로 송수신 할 수 있다. Max/MSP와 Quartz Composer는 모두 OSC 방식으로 데이터를 송수신 할 수 있는 기능이 있으며, 하나의 컴퓨터 안에서 로컬 방식으로 데이터를 주고받아 소리와 영상에 적용할 수 있다.¹⁵⁾ [그림-10]은 OSC를 통한 전체 시스템의 데이터 흐름도이다.



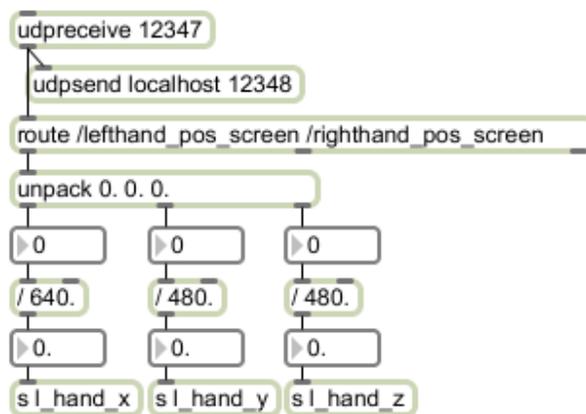
[그림-10] OSC를 통한 데이터 흐름도

[그림-11]과 [그림-12]는 Synapse app의 데이터가 OSC를 통해 Quartz Composer로 전송되는 Max/MSP 패치이다. Synapse app의 데이터를 받기 위해 `udpreceive` 오브젝트를 사용하였으며, Quartz Composer로 데이터를 보내기 위해 `udpsend` 오브젝트를 사용하였다. 모든 관절 값의 트래킹 데이터가 동시에 전송되기 때문에 `route` 오브젝

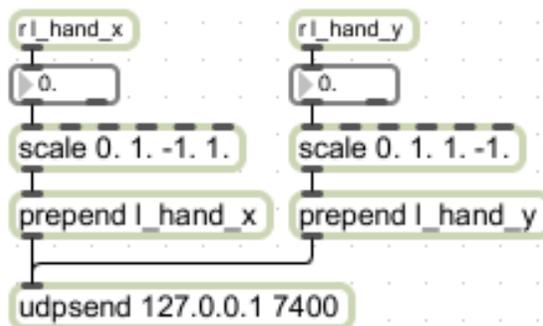
14) Apple에서 만든 Visual Programming 환경.

15) 김영민, “인터랙티브 멀티미디어 콘텐츠를 위한 적외선 LED 트래킹 시스템 연구” 한국 공학·예술학회, 논문지 2009.

트를 이용하여 각 관절 데이터를 분류해준다. 그리고 각각의 관절 값은 x, y, z축의 세 값이 일괄 처리되기 때문에 unpack 오브젝트로 각각의 값들을 분류해준다. 트래킹 된 값들은 Max/MSP와 Quartz Composer에서 주로 사용하는 0~1의 소수(float number)가 아닌 Synapse app의 640×480 픽셀 값으로 들어온다. 때문에 x, y, z축의 각 픽셀 값으로 나눠주어 0~1의 소수로 만들어준다.

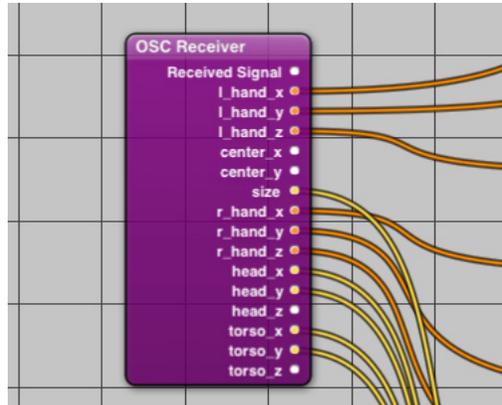


[그림-11] OSC를 통해 받은 트래킹 데이터의 x, y, z값을 추출



[그림-12] OSC를 통해 Quartz Composer로 보내지는 트래킹 데이터

[그림-13]은 Quartz Composer가 OSC Receiver 오브젝트로 각각의 관절 데이터를 받는 패치이다. OSC Receiver에서 받은 각 관절의 이름은 [그림-12]의 prepend 오브젝트를 이용하여 지정한다.



[그림-13] Quartz Composer에서 받는
트래킹 데이터

Ⅲ. 연구 기술의 작품 적용 및 효과

1. 작품의 의미와 기술적 접목

연구 기술이 적용된 작품은 멀티미디어 음악극 “A tale of June”의 일곱 번째 파트인 <Kinect Illusion>이다. 주인공 June은 선천적으로 듣지 못하고 말하지 못하는 청각장애인이다. 그렇기 때문에 소리가 아닌 수화나 다른 어떠한 수단을 사용하여 자신의 의사를 전달하고 감정을 표현해왔다. June이 새롭게 자신의 몸동작으로 의사를 전달하고 감정을 표현하는 새로운 방법을 알아가는 과정을 내용으로 담고 있다.

Illusion의 세계에 들어오면서 내면에 있던 자신의 다른 자아를 만나게 된다. 소극적이고 조용한 June과는 달리 역동적이고 활발한 자아가 몸을 이용해 감정을 표현하는 것을 본다. 그리고 자신의 동작에 의해 영상이 바뀌는 신기한 경험을 하게 되고, 자아와 June 자신 사이에 생기는 그 Illusion을 공감하게 된다. 또한 자신은 알지 못한 채 동작에 맞추어 소리가 발생되어 같이 전달이 되며, 또 다른 감정이 표현되는 세계를 경험하게 된다.

June의 앞에 불꽃같은 형상이 생기게 되고, June의 손이 움직임에 따라 그림이 그려진다. June은 선을 그리며 새로운 표현 방법을 알게 된 자신의 심리를 표현한다. 마지막으로 몸 전체의 움직임을 통해 소리와 영상을 만들어 내며, 더 크고 격정적인 감정 표현을 배우게 된다.

2. 작품의 구성

1) 작품의 전체적인 구성

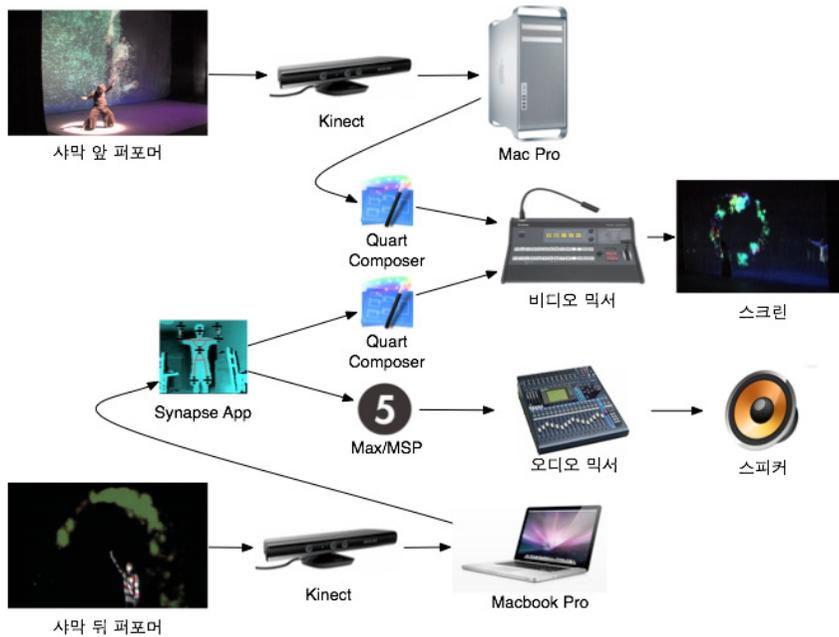
[표-2] 작품의 전체적인 구성

구성	section A	section B	section C
시간	0:00 ~ 2:23	2:23 ~ 3:47	3:47 ~ 3:58
내용	Illusion 세계 상자 속의 자아가 깨어 나와 드러남.	June이 불꽃을 발견. 불꽃을 이용해 자신의 감정을 표현함.	가시덩굴로 변한 그림을 잡으려 하지만 계속 사라짐.
사운드	테이프음악	테이프음악 그래놀러 사운드	테이프음악 주파수변조 사운드
영상	바람	불꽃	가시덩굴
Kinect	-	June 트래킹	June 트래킹
구성	section C'	section D	section B'
시간	3:58 ~ 5:13	5:13 ~ 6:03	6:03 ~ 6:48
내용	자아가 본격적으로 활발하게 꿈틀댐.	자아와 함께 June도 깨어나 활발히 감정을 표현.	Illusion의 세계가 끝나고 다시 현실로 돌아옴.
사운드	테이프음악 그래놀러 사운드	테이프음악 그래놀러 사운드	테이프음악
영상	RGB, 파티클	depth	불꽃
Kinect	자아 트래킹	June 트래킹	June 트래킹

본 작품의 전체적인 구성은 [표-2]와 같이 A - B - C - C' - D - B'로 이루어지며, 총 6개의 부분으로 나뉜다. June 안의 자아만 깨어나 감정을 표현하다가 이내 June도 자아를 깨닫고 함께 몸을 이용해 감정 표현을 하게 된다. section A, B, C까지는 주로 손만을 이용해 정적인 감정을 표현했다면, section C'부터는 몸 전체의 움직임을 통해 역동적인 감정 표현을 한다. 영상과 사운드 역시 June의 감정 표현을 중심으로 함께 발전하여 점차 역동적인 영상과 강렬한 음색으로 표현된다.

2) 작품의 시스템 구성

본 작품의 전체적인 시스템 구성은 [그림-14]와 같다.



[그림-14] 시스템 구성도

샤 막¹⁶⁾ 앞과 뒤의 배우 앞에 설치된 2대의 키넥트는 USB 포트를 통해 2대의 컴퓨터에 각각 연결된다. 샤 막 앞의 배우를 인식하는 키넥트는 Mac Pro에 연결되고, Quartz Composer에서 RGB 영상을 사용하여 영상을 만들어낸다. 샤 막 뒤의 배우를 인식하는 키넥트는 MacBook Pro에 연결되고, Synapse app을 통해 배우의 관절을 인식하여 트래킹한다. 트래킹 된 데이터들은 OSC를 통해 Max/MSP로 보내져 사운드를 생성한다. 그리고 Max/MSP에서 필요한 데이터 처리를 거쳐 다시 Quartz Composer로 보내져 트래킹 데이터에 따라 영상이 변하게 된다. Mac Pro와 MacBook Pro에서 생성된 영상들은 비디오 믹서를 통해 전환된다.

3) 작품의 무대 구성

무대를 구성할 때 주안점을 둔 것 중 하나는 샤 막 뒤의 배우가 키넥트의 트래킹 범위 안에서 충분히 움직이도록 하는 것이다. 키넥트가 원활히 트래킹 하기 위해서는 너무 가깝지 않은 적당한 거리, 즉, 제한된 활동 범위가 필요하다.¹⁷⁾ 그래서 최적의 키넥트 트래킹 범위를 위해 많은 시행착오를 거쳤다. 처음에는 배우가 키넥트의 정면에 위치해 움직이도록 했으나 샤 막 뒤의 공간이 좁아 트래킹이 원활하지 못했다. 때문에 키넥트를 샤 막 뒤 가운데에 두고, 배우가 대각선에서 움직이도록 구성을 하여 트래킹 범위를 해결했다.

다른 한 가지는 키넥트가 적외선을 사용하여 트래킹 하기 때문에 조

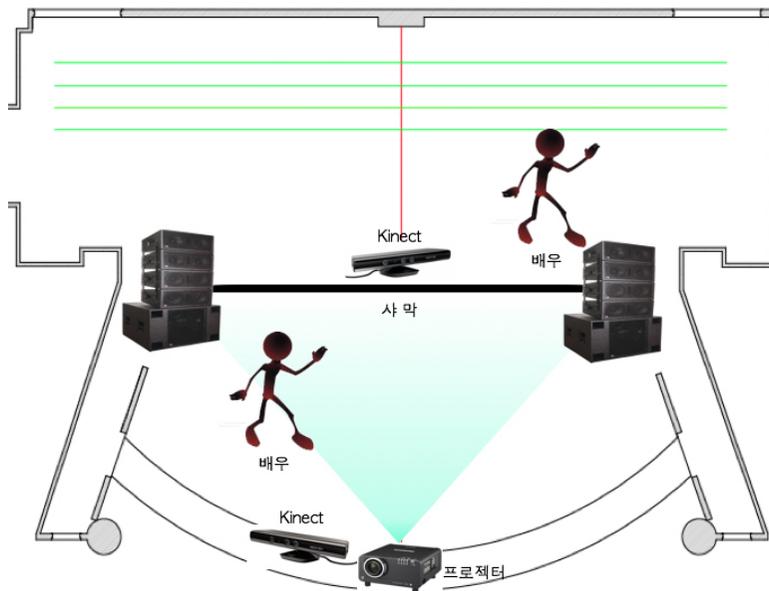
16) 연극에서 쓰이는 막의 종류이며 일반 스크린과는 다른 망사 소재로 되어있다. 영상의 상이 멧히기도 하지만 투과하는 성질을 가지고 있어 입체효과를 나타낼 수 있다.

17) Xbox용 키넥트의 설명서에는 사용자가 한 명일 경우에 1.8m 정도 떨어진 거리에서 사용하게 되어있다.

명에서 발생하는 적외선을 제한해야 했다. 본 공연에 앞서 시험 공연을 했던 소극장에서는 조명과 무대와 거리가가 가까워 적외선의 간섭이 매우 심해 트래킹에 문제가 발생하였다. 하지만 본 작품이 공연된 이해랑 예술극장은 조명과 무대와 거리가 멀고, 공간이 넓어 조명의 색상을 블루 계열과 화이트 계열로 조절함으로써 적외선의 간섭을 피할 수 있었다. 공연장에서 키넥트를 사용하여 공연을 한다면, 키넥트의 제한된 트래킹 범위 안에서 원활한 배우의 움직임을 만드는 것, 그리고 조명에서 나오는 적외선과의 간섭 문제를 해결하는 것이 무대 구성의 관건이라 할 수 있다.

반면 샤막 앞 배우의 움직임은 키넥트의 관절 트래킹을 이용하는 것이 아니라 키넥트의 RGB 영상만을 사용하기 때문에 샤막 뒤의 배우보다 활동 범위의 제한 없이 자유롭게 넓게 움직이도록 구성했다.

다음의 [그림-15]은 전체적인 무대 구성도이다.

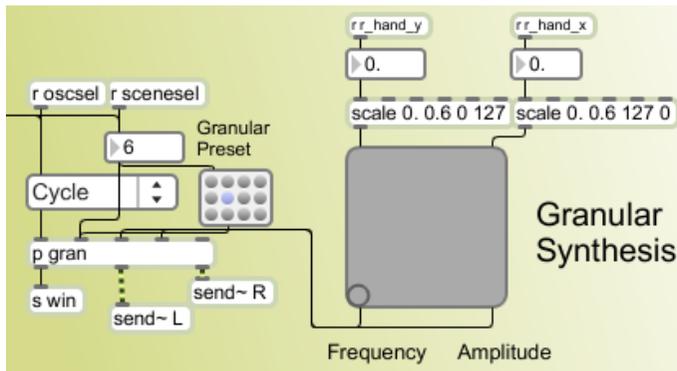


[그림-15] 무대 구성도

3. 배우의 움직임에 따른 사운드 효과

실시간으로 사운드의 생성과 변화를 위해 배우의 움직임에 따라 소리가 변화하도록 구성하였다.

먼저 양 손의 위치 값을 이용해 사운드를 제어 한다. 손 위치의 x축 값은 그래놀러 합성(granular synthesis)¹⁸⁾과 주파수변조 합성(FM synthesis)¹⁹⁾의 주파수(frequency)를 실시간으로 제어 한다.



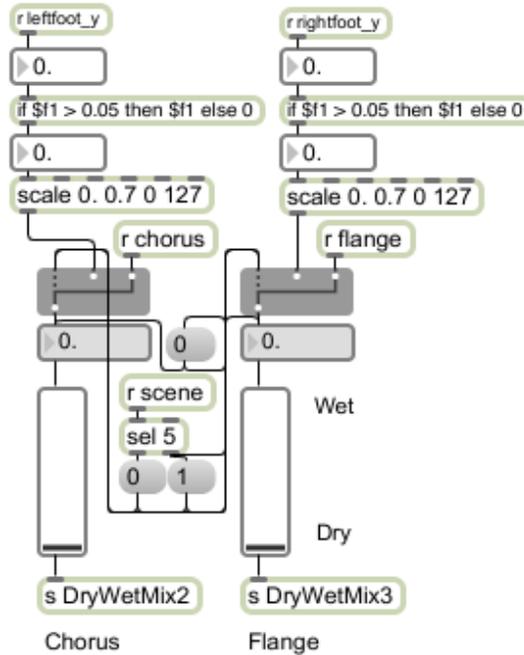
[그림-16] 손의 위치 값으로
그래놀러 합성 사운드 제어

그리고 [그림-16], [그림-17]과 같이 손 위치의 y축 값은 그래놀러 합성의 음량 값(amplitude)과 주파수변조 합성의 modulation index를 실시간으로 제어 한다. 손의 위치 값은 픽셀 값으로 트래킹되어 Max/MSP로 들어오지만 데이터를 수신하는 패치에서 0~1의 소수로

18) 낱알같이 짧은 소리들을 모아서 하나의 이어진 다른 소리를 만드는 음향 합성.

19) 가산 합성법, 감산 합성법 등과 함께 신시사이저를 통한 합성법 중의 하나. 피변조파를 사인(sine)파로 하고, 변조파도 그와 같은 영역을 가진 오디오 신호로 하면, 출력되는 변조파는 고차 배음을 많이 포함한 파형이 됨. 이 원리를 이용하여 새로운 형태의 파형을 얻어 음을 만들어내는 방법.

(flanger)²¹⁾ 효과의 원 신호(wet)/변화된 신호(dry) 비율을 제어한다.



[그림-18] 발의 위치 값으로 테이프음악에 적용되는 이펙트(chorus, flange) 제어

마지막으로, 양 손, 양 팔꿈치로 이어지는 곡선의 형태로 주파수 변조 합성의 음량 엔벨로프(amp envelope)를 변화시켰다. 실시간으로 음량 엔벨로프를 변화시켜 사운드의 형태가 달라지도록 했으나 두드러지는 효과가 적어 공연 시에는 사용되지 않았다.

- 20) 하나의 음에 다른 음을 가미하여 우아하고 중후하게 만든 것으로 보통 전자 피아노, 전자 기타와 같은 악기나 신디사이저와 함께 사용된다. 사운드에 깊이와 부피를 더해준다.
- 21) 딜레이를 응용한 이펙터의 하나. 입력된 음을 약간 지연시켜(1ms-10ms정도로) 원 음에 더하는 방법. 지연 시간을 실시간 변화시키면 그에 따라 필터의 특성도 변화하여 특정 주파수의 정점이 움직이게 된다. 그 결과 특정 주파수 성분이 강조되어 독특한 플랜징 사운드가 생성된다.

4. 샤 막을 이용한 입체 영상 효과

[그림-19]는 샤 막에 맺히는 영상을 샤 막 앞, 뒤의 배우가 같이 바라 보며 함께 공존하는 입체 효과를 표현하기 위해 샤 막을 사용한 것을 보여준다.



[그림-19] 샤 막을 이용해 막 앞/뒤의 배우가 공존하는 효과

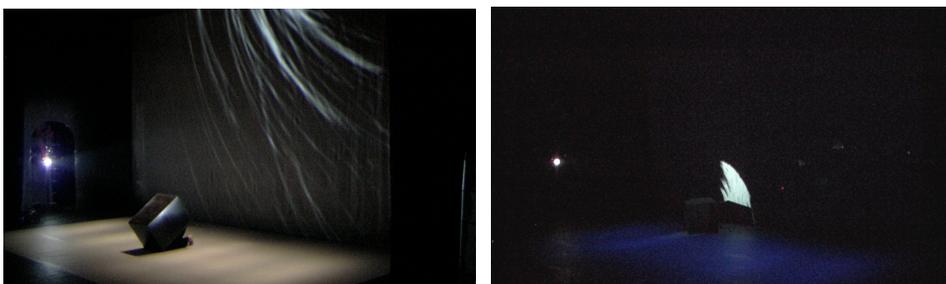
샤 막 앞의 배우는 샤 막 뒤 June의 또 다른 내면의 자아를 표현하기 때문에 샤 막에 맺히는 영상을 통해 June이 자아와 교감하는 효과를 나타냈다.

5. 영상과 퍼포먼스의 구성 및 제작

작품에 사용된 영상은 모두 Quartz Composer로 제작하였다. 배우의 움직임은 Synapse app에서 트래킹되어 Max/MSP로 보내진다. 그 중 영상을 제어하는데 효과적인 데이터(손, 머리, 몸통)만을 선택해 약간의 데이터 스케일링을 한다. 스케일링 된 데이터들은 다시 OSC를 통하여 Quartz Composer로 전송된다. 파티클(particle)²²⁾을 이용한 바람 영상, 불꽃 영상, 가시뎅굴 영상, RGB 영상과 파티클을 이용한 영상, depth image를 이용한 영상의 총 다섯 종류의 영상이 사용되었다.

1) section A에서의 기술 적용

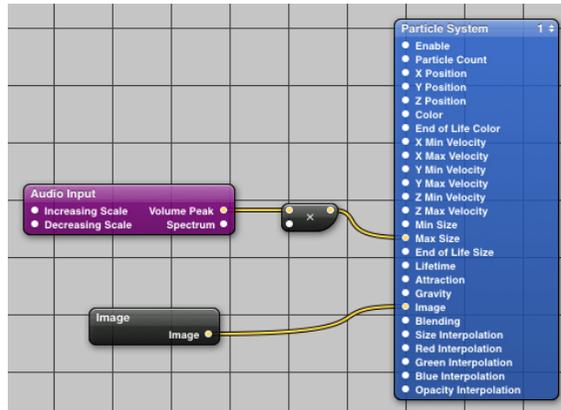
section A는 상자 안에 감춰있던 자아가 서서히 드러나는 단계이다. 음악과 영상이 모두 울리는 소리, 바람 소리를 형상화하여 감정을 전개했다. 테이프음악의 음량 값은 바람 영상의 두께를 실시간으로 변화시킨다.



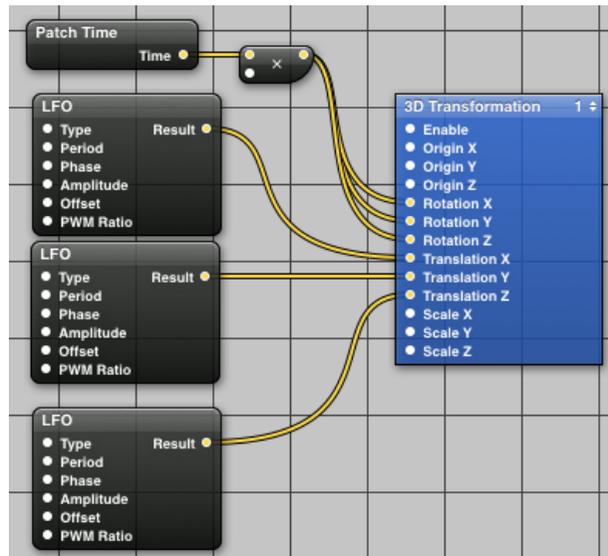
[그림-20] section A의 퍼포먼스와 영상

22) 3차원 컴퓨터 그래픽에서 많은 양의 미세한 입자를 중력이나 바람의 영향을 고려해서 움직이게 하는 시스템, Quartz Composer의 Particle System 패치.

[그림-21]과 [그림-22]가 실시간 두께 변화를 나타내는 패치이다.



[그림-21] 오디오 입력을 받아 Particle의 Max Size에 적용



[그림-22] [그림-21]의 패치들을 Macro로 묶어 3D Transformation을 적용

배우가 상자 안에서 퍼포먼스를 하는 초반에는 상자에 영상을 축소,

집중시켜 상자에 무언가가 있다는 기대감을 가지도록 했다.

2) section B에서의 기술 적용

section B는 주인공인 June이 자신의 손의 움직임을 따라다니는 불꽃 형상을 발견하게 되고, 그 불꽃을 이용해 그림을 그리며 자신의 감정을 표현해 나가는 단계이다. 샤 막 뒤에 있는 June의 오른손 움직임을 트래킹하여 불꽃 영상의 위치 값과 그래놀러 합성 사운드의 주파수와 음량을 실시간으로 제어하였다. 또한 샤 막을 사이에 두고 샤 막 뒤의 June과 샤 막 앞의 자아가 그려지는 불꽃을 함께 바라보고 공유하며 퍼포먼스를 한다.



[그림-23] section B의 퍼포먼스와 영상

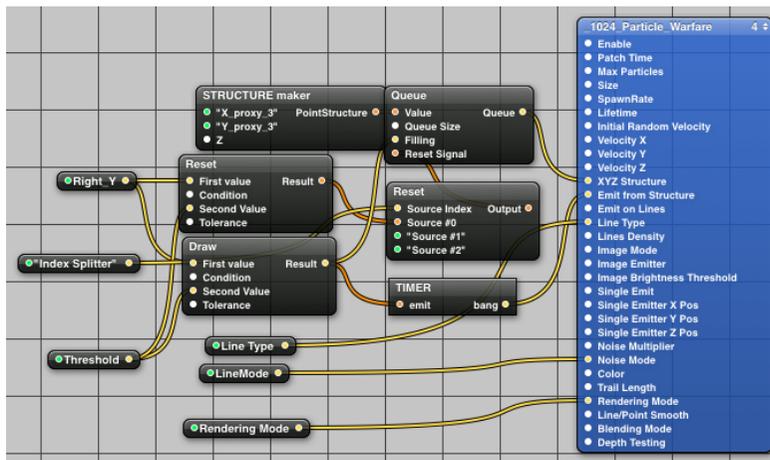
3) section C에서의 기술 적용

section C는 손으로 그려지는 영상이 불꽃에서 가시딩굴 형상으로 바뀌는 단계이다. 짧은 시간의 파트이지만 강렬하게 그려지고 June이 잡

으려 하면 사라지기를 반복하는 효과를 주었다. section B와 마찬가지로 June의 오른손의 움직임을 트래킹하여 영상을 그려냈다. 한 가지 다른 점이 있다면, section C의 가시덩굴 영상은 실제로 잡을 수 없는 아쉬움을 표현하기 위해 스톱(threshold)²³⁾를 정해 June이 손을 허리 아래로 내리게 되면 영상이 사라지는 효과를 더했다.



[그림-24] section C의 퍼포먼스와 영상



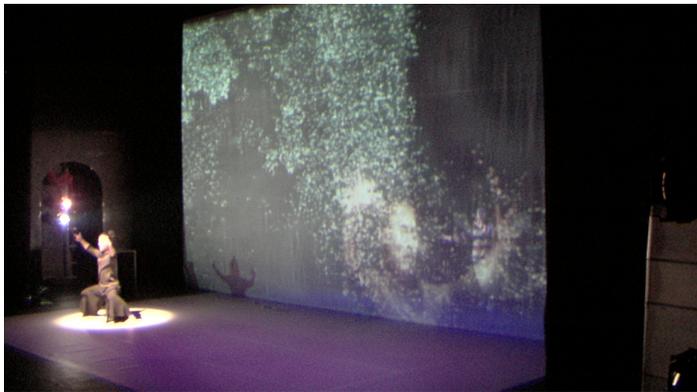
[그림-25] 가시덩굴 영상 Quartz Composer 패치

23) 한계점. 한계점의 기준 값을 정해 기준 값보다 높거나 낮은 값이 입력될 때 작동 시키거나 정지시킨다.

[그림-25]는 `_1024_Particle_Warfare`패치²⁴⁾를 사용하여 트래킹된 손의 위치 값으로 가시뎅굴 영상을 그려내고 스레숄드에 의해 영상을 사라지게 하는데 사용된 Quartz Composer 패치이다.

4) section C'에서의 기술 적용

section C'는 June의 자아가 본격적으로 깨어나 숨겨져 있던 감정을 적극적으로 표현하는 단계이다. 자아의 역동적인 움직임을 키넥트의 RGB 카메라로 촬영 해 자아의 몸에서 나오는 파티클 영상을 만들어낸다.



[그림-26] section C'의 퍼포먼스와 영상

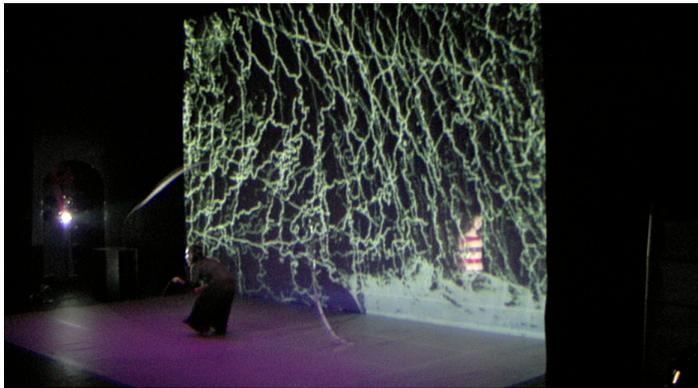
파티클 영상은 OpenCL²⁵⁾를 활용하여 Quartz Composer로 제작하였다. section C'에서는 파티클 영상과 테이프음악으로만 진행된다.

24) 오디오 비주얼 설치작품을 주로 하는 '1024'라는 회사에서 만든 Quartz Composer용 플러그 인이다.

25) Open Computing Language, 개방형 범용 병렬 컴퓨팅 프레임워크, 3D 그래픽스 및 컴퓨터 오디오에 대한 산업계의 개방형 표준이다.

5) section D에서의 기술 적용

section C'가 자아의 역동적인 감정 표현이었다면, section D는 자아의 감정표현을 이어받아 June의 역동적인 감정 표현을 이끌어내는 단계이다. section C까지 June이 손의 움직임만을 감정 표현의 수단으로 사용했다면, section D에서는 몸 전체의 움직임과 넓은 활동 반경으로 감정을 표현한다. section D는 키넥트의 depth 영상을 사용하였는데 depth 영상에 감마(밝기) 필터를 거치고, 테두리 검출 필터를 거친 영상을 사용하였다. 또한 테두리가 일정 시간동안 딜레이(delay) 되어 남아있는 피드백 영상을 사용하여 단순한 움직임이 아니라 잔상이 남는 움직임으로 자연스럽게 물 흐르는 느낌을 주도록 하였다.



[그림-27] section D의 퍼포먼스와 영상

6) section B'에서의 기술 적용

section A, B, C, C', D가 Illusion의 세계에서 자아와 공존하며 현실에서는 표현할 수 없는 감정을 표현하였다면, section B'는 자아와 헤

어지며 다시 현실 세계로 돌아오게 되는 아쉬움과 허무함을 표현하였다. 적용된 영상은 section B의 불꽃 영상이 다시 사용되었지만, 그려지는 불꽃이 아닌 자아와 June이 머리 위에 떠 있는 불꽃을 마주 바라보도록 했다. Illusion의 세계에 미련을 두고 끝나는 효과를 의도했다.



[그림-28] section B'의 퍼포먼스와 영상

IV. 문제점 및 향후 계획

본 작품 <Kinect Illusion>은 키넥트의 모션 트래킹과 RGB, depth 카메라 기능을 이용하여 배우의 움직임을 사운드와 영상 요소에 접목시킨 멀티미디어 음악작품이다. 배우와 사운드, 영상의 결합을 통해 창작자의 예술적 표현을 더욱 확장시키기 위한 다양한 기술과 예술적 기법이 도입되었다. 특히 감상자는 동작에 따라 변화하는 소리, 그려지는 영상의 변화를 실시간 감각적으로 느낄 수 있으며, 다양한 동작과 시도를 통해 많은 결과를 얻을 수 있다. 또한 앞으로 본 연구의 더욱 완성적인 결과물을 위해 필요한 연구의 요소들이 있다.

첫째, 동작과 소리와 상호작용에서 얻어지는 상관관계이다. 동작에 의해 만들어지는 소리가 음악적으로 표현될 수 있으며, 이 때 소리는 동작과 상관관계를 가지며 만들어져야 한다. 그렇기 때문에 동작에 적합한 소리를 만들어 내기 위한 연구가 더 필요하다.

둘째, 동작과 영상간의 상관관계에도 더 적합한 결과물을 얻기 위한 연구가 필요하다. 관객이 그려지는 영상과 배우의 동작과의 연관성을 느끼도록 영상을 제작해야 한다.

셋째, 키넥트는 적외선을 이용하여 트래킹 하기 때문에 무대의 조명에서 발생하는 적외선과의 간섭을 해결해야 한다. 이 문제를 해결하기 위해서는 조명과의 거리를 멀게 하고, 조명의 밝기를 어둡게 하거나, 조명기에 필터를 넣는 등 공연에 제한이 생기지 않는 적정 수준으로 적외선을 줄이는 방법을 연구해야 한다.

마지막으로는 키넥트의 트래킹 범위와 속도에 대한 연구이다. 키넥트의 기술 사양은 본 작품 <Kinect Illusion>이 공연된 이해랑 예술극장의 무대 넓이에 적합하다. 하지만 조명이나 그밖에 변수들 때문에 완벽

한 트래킹을 위해서는 배우가 움직이는 범위를 제한하거나, 특정 영역에서만 트래킹이 되도록 하는 방법을 연구해야 한다. 또한 배우의 움직임이 빠를 경우 3D depth image와 RGB image에는 문제가 없었지만 관절 트래킹 데이터의 정확성이 떨어지는 문제가 있었다. 연속적으로 이어지는 데이터로 트래킹 되어야 하지만, 배우의 움직임이 빠르면 데이터가 순간적으로 손실되거나 틀어지는 현상이 발생했다. 그렇기 때문에 배우의 움직임이 빨라지더라도 데이터가 최대한 정확하고 연속적으로 트래킹 되도록 하는 방법을 연구해야 한다.

Keyword (검색어) : 컴퓨터음악(computer music), 멀티미디어음악(multimedia music), 키넥트(Kinect), 모션 트래킹(motion tracking), Max/MSP, 퀴츠 컴포저(Quartz Composer), OSC(Open Sound Control), 샤 막(sha screen), 그레놀러 합성(granular synthesis), 주파수 변조 합성(frequency modulation synthesis)

E-mail : yksilj@gmail.com

참고문헌

1. 단행본

- 정성환, 「오픈소스 OpenCV를 이용한 컴퓨터 비전 실무 프로그래밍」 (홍릉과학출판사, 2007)
- 조재원 「멀티미디어와 인터랙티브 아트」 (한국학술정보, 2001)
- Roads, Curtis “The Computer Music Tutorial”
(The MIT Press, 1996)
- Rush, Michael “New Media in Art” (Thames & Hudson, 2005)

2. 참고논문

- Alberto Menache, “Understanding Motion Capture for Computer Animation and Video Games”, (2000)
- 김민경, “비보잉 동작에 의한 오디오-비주얼 작품 제작 연구”
「동국대학교 영상대학원 멀티미디어학과」, (2010)
- 김영민, “적외선LED의 트래킹을 이용한 멀티미디어 음악작품 연구”
「동국대학교 영상대학원 멀티미디어학과」, (2009)

- 김종현, “뇌파와 모션 디텍션을 이용한 멀티미디어 음악제작 연구” 「동국대학교 영상대학원 멀티미디어학과」, (2010)
- 박상범, “동작 인식과 Max, MSP, Jitter를 이용한 멀티미디어음악 시스템 연구” 「동국대학교 영상대학원 멀티미디어학과」, (2009)
- 윤민철, “멀티미디어 기술을 이용한 인터랙티브 미디어 제작 연구” 「동국대학교 영상대학원 멀티미디어학과」, (2010)
- 이범렬 외, “모션캡처 기술 동향” 「ETRI, 전자통신동향분석」, 제 22권 제 4호 (2007)

3. 인터넷

- Max/MSP Forum
<http://www.cycling74.com/forums/index.php>
- Kinect
<http://www.xbox.com/kinect>
<http://openkinect.co.kr>
<http://www.kinecthacks.nl>
<http://www.kinecthacks.net>
- Quartz Composer
<http://http://kineme.net/>
<http://1024d.wordpress.com/>

Abstract

Research on interactive multimedia productions with Kinect

(Focus on multimedia music - 'Kinect Illusion')

Yoon, Ki Seon

The development of advanced digital technology in modern society, brought many changes in the cultural life, especially in the field of art. As technology develops, artistic expression which was only inside the artist's head became possible. Multimedia arts should convey its meaning to the audience efficiently. By using variable senses like vision, auditory and tactile, artists can bring it to the real world what they think.

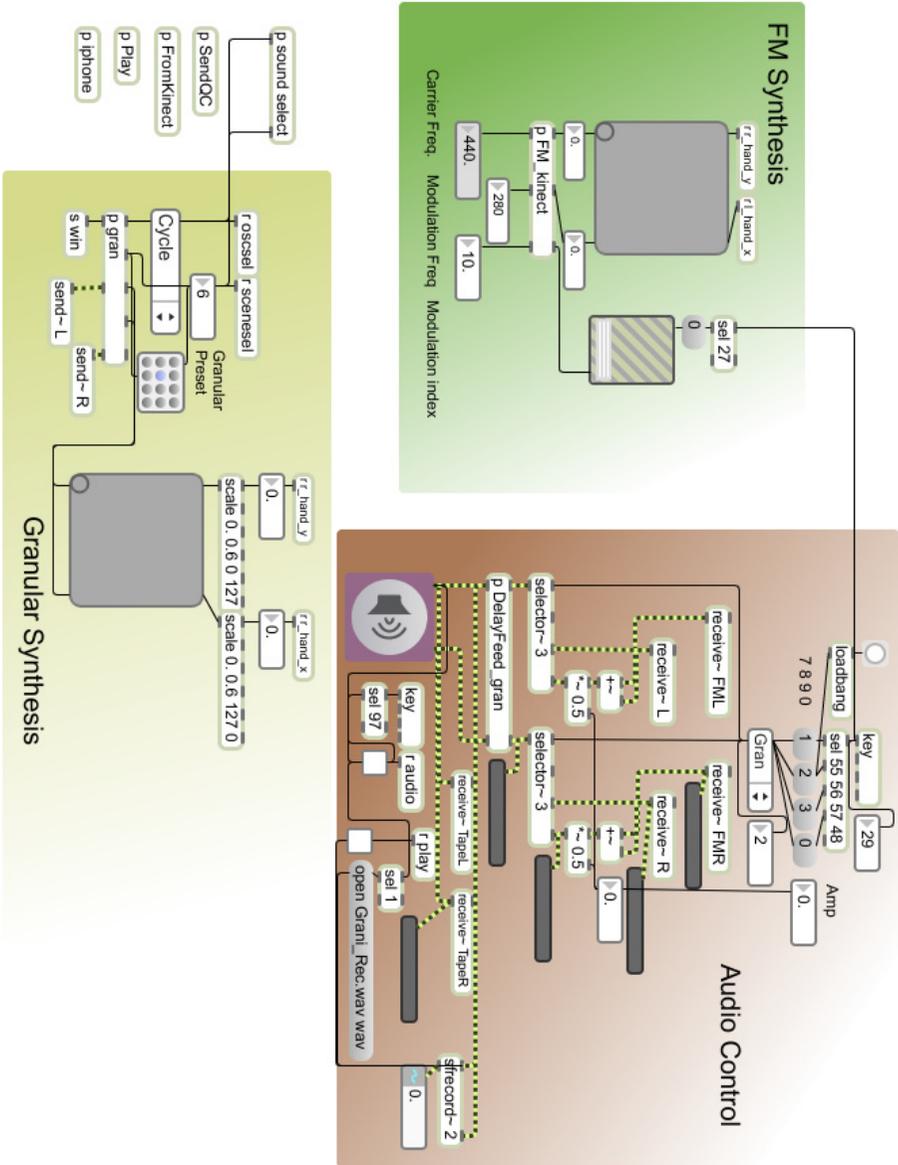
Contemporary multimedia pieces involve interactive music and visual performance through actor or dancer. For color tracking or infrared LED, optical motion tracking it needs costume that has specific color, device or sensor to be tracked and it constrains the movement or expression of the performer. Furthermore, tracking data can not be enough accurate because of limited number of sensors and devices for tracking.

In this study, tracking has been done using Kinect(without a controller

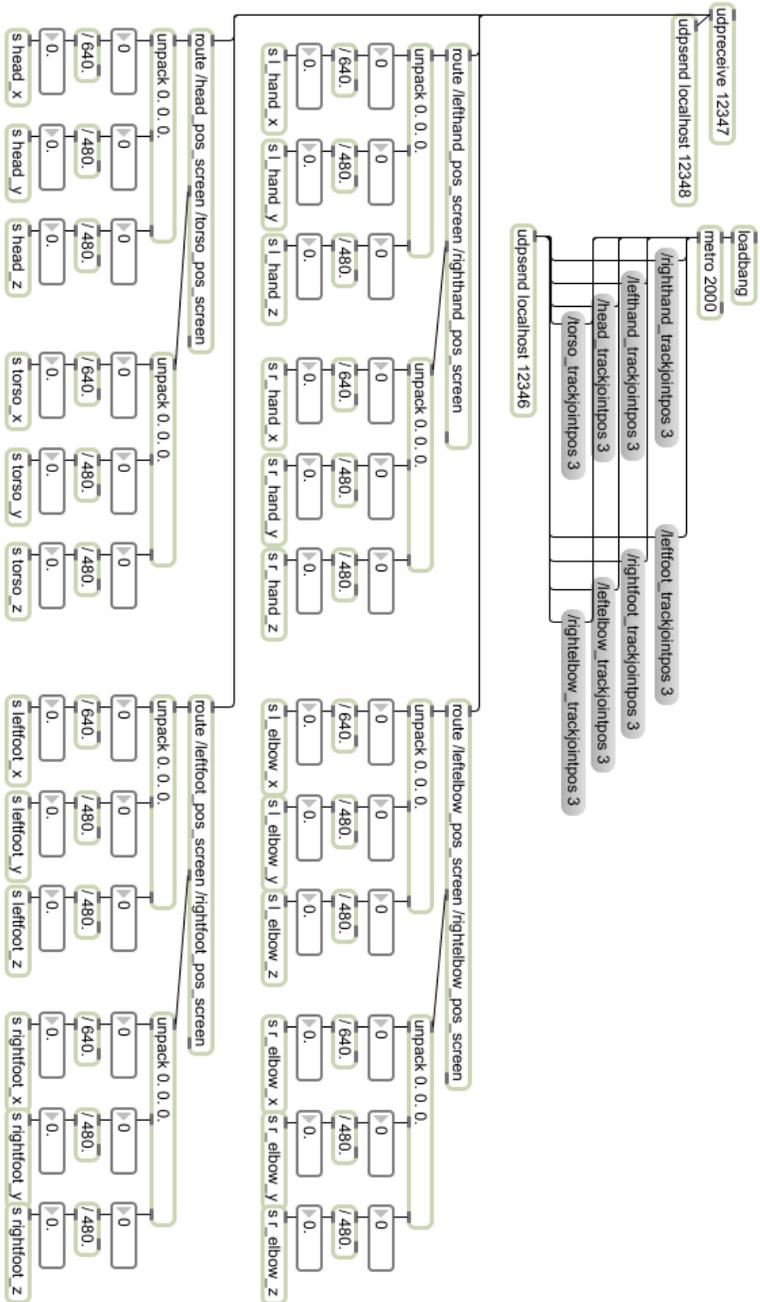
by using the performer's body to experience games and entertainment in connection with using MS Xbox360 peripherals) to overcome disadvantages. By using Kinect and computer, performer can be free from any devices or sensors, also can be tracked accurately by using skeleton data, depth image and RGB image.

부록-1 (Max/MSP 패치)

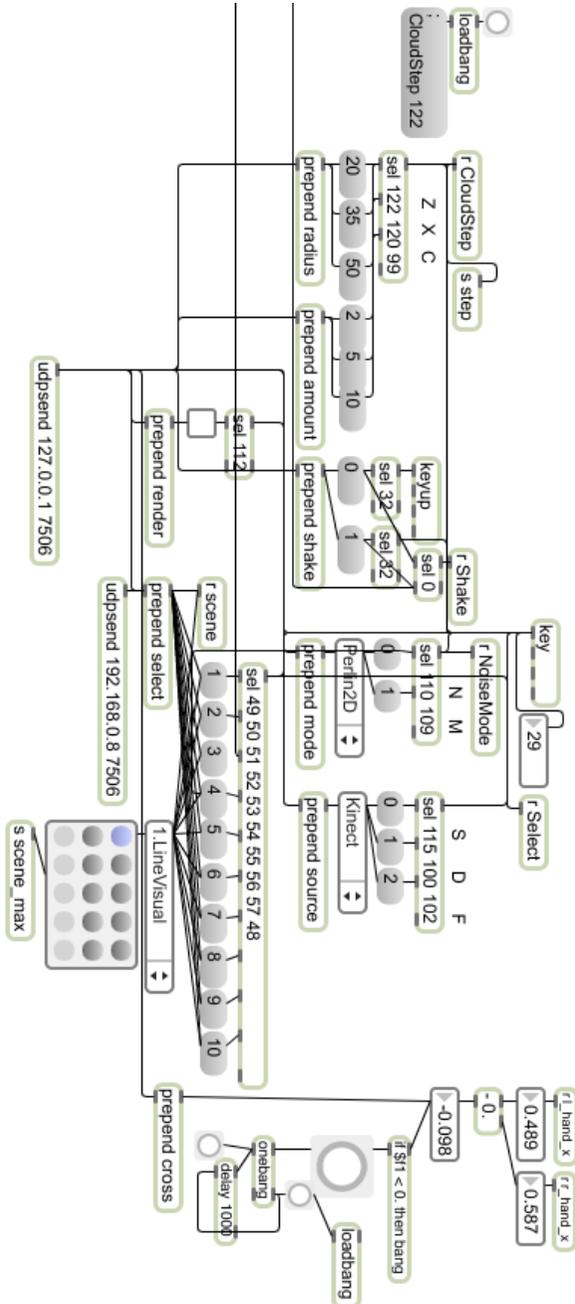
1. 합성 사운드와 전체적인 오디오 컨트롤 패치



2. 트래킹 데이터를 수신하여 각각의 데이터로 추출하는 패키지

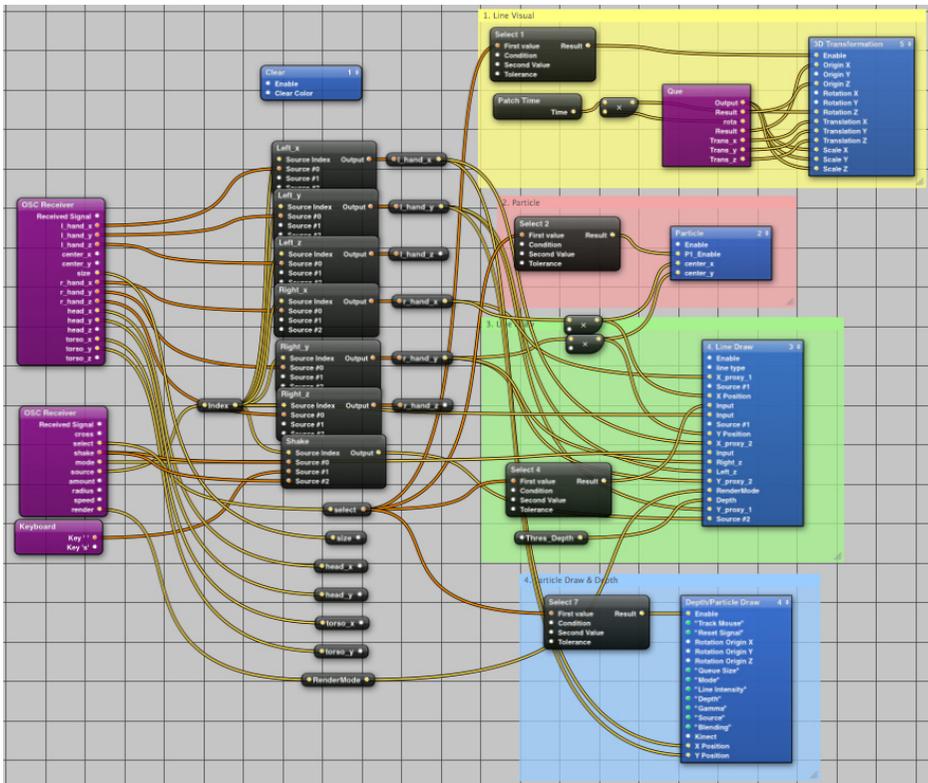


4. section 선택과 각각의 파라미터 선택할 수 있는 컨트롤 패치

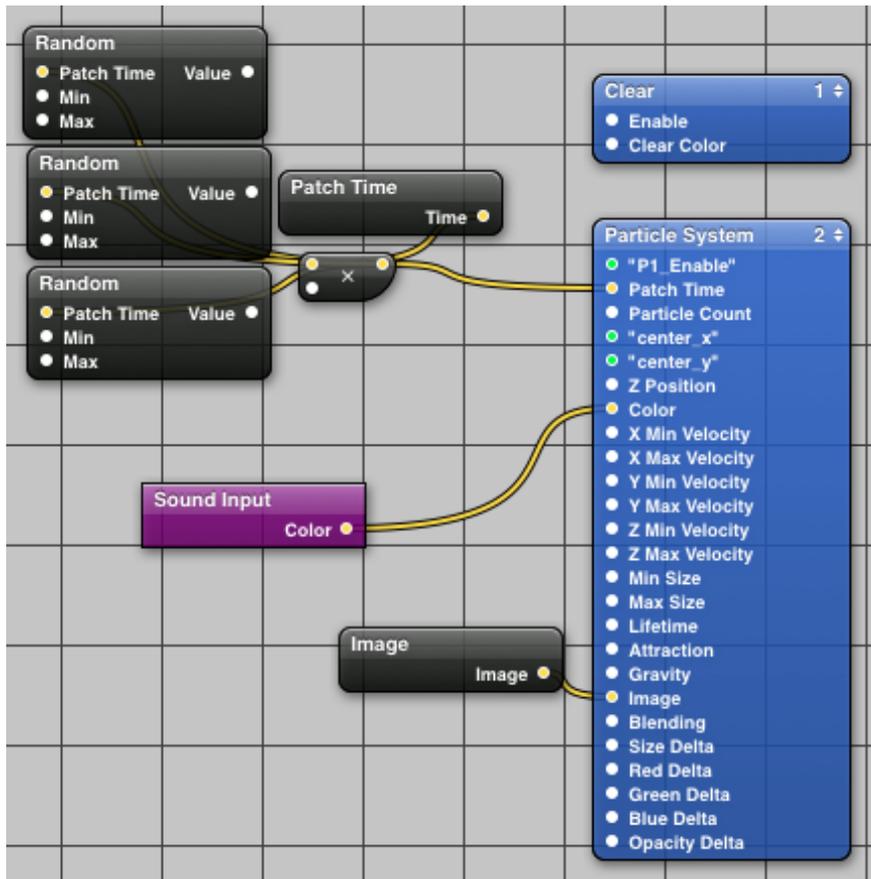


부록-2 (Quartz Composer 패치)

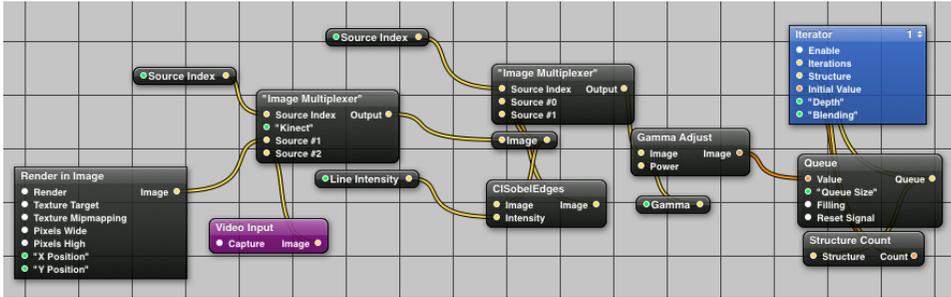
1. 트래킹 데이터를 수신 및 각 section 별 영상 패치



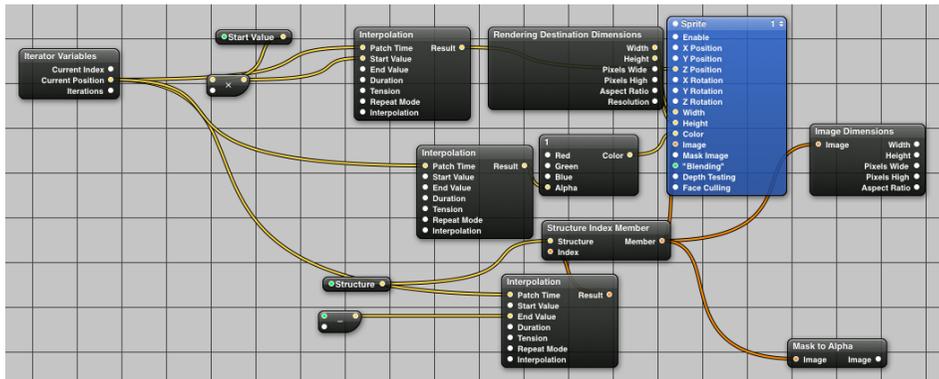
2. 파티클 영상 패치



3. 키넥트의 depth image를 이용한 영상 패치



4. Iterator 패치의 루트 매크로 패치



부록-3 (첨부 DVD설명)

1. Kinect Illusion : 2011년 11월 18, 19일 이해랑 예술 극장
<Kinect Illusion>의 공연실황
2. Kinect Illusion : Max/MSP 패치
3. Kinect Illusion : Quartz Composer 패치
4. Kinect Illusion.wav : 테이프음악