



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

석 사 학 위 논 문

카메라 트래킹을 이용한 멀티미디어 작품 제작 연구

- 멀티미디어음악 <할머니의 방>을 중심으로 -

지도교수 김 준

동국대학교 영상대학원 멀티미디어학과

허 남 영

2 0 1 5

석사학위논문

카메라 트래킹을 이용한 멀티미디어 작품 제작 연구

- 멀티미디어음악 <할머니의 방>을 중심으로 -

허남영

지도교수 김준

이 논문을 석사학위 논문으로 제출함

2015년 6월

허남영의 음악석사(컴퓨터음악) 학위 논문을 인준함

2015년 7월

위원장: 정진헌



위원: 김준



위원: 박상훈



동국대학교 영상대학원

목 차

I. 연구 배경과 목적	1
1. 연구 배경.....	1
2. 연구 목적.....	3
II. 기술적 연구	5
1. 객체 인식 및 추적 방법 선행 연구.....	5
2. 작품에 적용한 기술 구현.....	8
1) 차영상을 활용한 트래킹 환경 구현.....	8
2) 광류를 활용한 트래킹 환경 구현.....	11
III. 연구 기술의 작품 적용	14
1. 작품의 내용	14
2. 작품의 구성	14
1) 작품의 구조	14
2) 작품의 시스템과 무대 구성	16
3. Section별 작품의 구성 및 효과	18
1) Section A	18
2) Section B	19

① 음악 알고리즘	19
② 영상 알고리즘	21
3) Section C	23
① 음악 알고리즘	23
② 영상 알고리즘	24
4) Section C'	25
① 음악 알고리즘	25
② 영상 알고리즘	25
5) Section D	26
6) Section B'	27
IV. 결론 및 문제점	28
참고문헌	30
Abstract	33
부록 1 :	35

표 목 차

[표-1] 객체 추적 방법 비교.....	4
[표-2] 작품 구조.....	15

그 립 목 차

[그림-1] 공연에서 객체추적이 활용된 사례	2
[그림-2] 그레이 영상 변환	5
[그림-3] 프레임 비교 및 차영상	6
[그림-4] 차영상의 이진화	7
[그림-5] 광류의 흐름	8
[그림-6] 차영상 구현 과정	9
[그림-7] 차영상을 활용한 트래킹 환경 구현 패치	10
[그림-8] 광류 구현 과정	11
[그림-9] 광류를 활용한 트래킹 환경 구현 패치	12
[그림-10] 시스템 구성도	16
[그림-11] 무대 구성도	17
[그림-12] 작품에 쓰이는 이미지들	18
[그림-13] 퍼포머의 등장.....	19
[그림-14] 사물 소리에서 음악소리로의 변화 과정.....	20
[그림-15] 가상악기를 활용하기 위한 과정.....	21

[그림-16] Section B의 영상 출력 알고리즘.....	22
[그림-17] Section B에서 출력되는 영상.....	22
[그림-18] Section C의 음악 알고리즘.....	23
[그림-19] Section C의 영상 알고리즘.....	24
[그림-20] Section C의 영상과 무대.....	25
[그림-21] jit.gl.mesh오브젝트의 구성.....	26
[그림-22] jit.gl.mesh오브젝트 반영 영상.....	26
[그림-23] Section D의 영상과 무대.....	27

I. 연구 배경과 목적

1. 연구 배경

예술은 인간의 사상이나 감정을 표현하는 것으로 유사(有史) 이래 미술·음악·문학·무용 등 다양한 형태로 존재해왔다. 이러한 예술은 기술의 발달로 인해 인간의 상상에만 머물던 생각을 조금 더 현실적으로 표현할 수 있는 폭이 넓어졌다. 기술의 발달로 예술의 새로운 분야인 미디어아트(media art)가 탄생하였다. 미디어아트는 현대 커뮤니케이션(communication)의 주요 수단인 대중매체를 활용한 예술로 비디오아트(video art)를 시작으로 컴퓨터아트(computer art)·멀티미디어아트(multimedia art)·디지털아트(digital art) 등으로 발전해 왔다. 이러한 여러 미디어아트를 모두 포함하여 뉴미디어아트(new media art)로 명명하기도 한다.

뉴미디어아트에서 강조되는 것은 관객과의 소통이다. 창작자의 작품을 관객에게 단편적으로 전달하는 것이 아닌 창작자와 작품, 작품 내 구성요소, 또는 관객 간의 상호작용을 통해 예술의 이해도를 높이는 것이다. 관객과의 소통이 원활하기 위해서는 인간의 감각에 직관적으로 다가가는 것이 유효하다. 무대 위에서 행해지는 작품에 있어 관객이 퍼포머(performer)·물체·빛 등의 동작·움직임을 직관적으로 받아들인다면 창작자가 의도한 의미를 쉽게 이해할 것이다.

최근 객체추적 기술과 동작인식 기술은 창작자의 의도를 관객에게 직관적으로 전달하기 위한 방법으로 주목받고 있다. 객체추적 기술과 동작인식 기술은 게임·교육·과학·의료 등 여러 분야에 활용되고 있는

기술로 예술분야에서도 적극적으로 활용하고 있다. 객체의 움직임을 인식하기 위한 방법으로는 객체의 색상을 추적하는 컬러 트래킹(color tracking)¹⁾, 눈에 보이지 않는 적외선을 이용한 적외선 트래킹²⁾, 자이로센서나 인체감지센서 등을 사용하는 센서 트래킹³⁾ 등이 있다.

[그림-1]의 왼쪽 그림은 2011년 윤기선의 <Kinect Illusion>이라는 작품으로 모션 센서(motion sensor)인 Kinect⁴⁾를 사용하였다. Kinect 앞에 위치한 퍼포머의 움직임을 감지해 파티클(particle)⁵⁾ 영상을 생성한 장면이다. Kinect를 활용한 공연 기술은 컴퓨터와 Kinect라는 간단한 구성으로 퍼포머의 관절 움직임까지 인식할 수 있는 효율적 기술이지만 모션 센서의 범위가 제한적이어서 퍼포머가 무대 공간을 충분히 활용할 수 없다는 단점이 있다. 또한 여러 명의 퍼포머를 동시에 인식하는데 용이하지 않아 등장하는 퍼포머의 수도 한정적이다.



[그림-1] 공연에서 객체추적이 활용된 사례

- 1) 객체의 색상·밝기 등을 분석하고 설정된 색상의 객체를 추적하는 기술
- 2) 적외선을 발산하는 객체를 적외선 카메라로 촬영하여 위치를 추적하는 기술
- 3) 객체에 센서를 부착하여 객체의 움직임에 따른 센서 값의 변화를 반영하는 기술
- 4) Microsoft에서 개발한 별도의 컨트롤러 없이 사람의 움직임을 감지해 화면에 반영하는 Xbox의 주변기기
- 5) 3차원 컴퓨터 그래픽에서 많은 양의 미세한 입자를 중력이나 바람의 영향을 고려해서 움직이게 하는 시스템

[그림-1]의 오른쪽 그림은 2014년 11월 중국 상해아트페스티벌에서 공연된 창작 뮤지컬 <투란도트>의 한 장면이다. 해당 장면은 RTLS(Real Time Location System)⁶⁾를 사용한 것으로 실시간으로 무대 위 퍼포머의 위치를 추적해 무대 바닥과 배우의 의상에 3D 그래픽을 투영한 것이다. RTLS는 조명과 무대 환경에 영향을 받지 않고 여러 퍼포머들의 위치에 따라 자유로운 표현이 가능한 기술이지만 구성 기술이 까다롭고 공연장의 구조도 고려해야하며 상당히 고가의 비용이 든다는 단점이 있다.

이러한 기존의 방법들은 퍼포머의 움직임 제한에 따른 무대 활용 폭이 좁거나 고비용과 기술적 부담으로 창작자의 의도를 충분히 표현하기 힘들다. 따라서 표현의 제약을 줄이는 공연기술이 요구되고 있다.

2. 연구 목적

본 연구는 객체추적 기술을 활용하여 표현의 제약을 최소화하고 저비용으로 손쉽게 관객이 공감할 수 있는 멀티미디어 공연을 수행하는 시스템을 갖추는 것이 목적이다. 따라서 컴퓨터 비전(computer vision) 기술인 차영상(difference image)과 광류(optical flow)를 적용해 표현하는 공연 시스템을 구축하고자 한다. 차영상과 광류를 활용한 기술은 컴퓨터와 카메라만으로 객체추적 환경을 구축할 수 있기 때문에 복잡한 장비와 비용을 줄일 수 있고 무대의 활용 폭을 넓힐 수 있다.

[표-1]은 공연에서 사용되는 다양한 객체 추적 방법을 비교한 것이다.

6) 근거리 및 실내와 같은 제한된 공간에서의 위치확인 및 위치추적 서비스

구분	컬러 트래킹	적외선 트래킹	센서 트래킹	Kinect	RTLS	카메라 트래킹
퍼포머의 트래킹 도구 부착 여부	○	○	○	×	○	×
무대 활용 폭	넓음	넓음	넓음	좁음	넓음	넓음
비용	낮음	낮음	낮음	낮음	높음	낮음

[표-1] 객체 추적 방법 비교

II. 기술적 연구

1. 객체 인식 및 추적 방법 선행 연구

객체를 인식하고 추적하는 방법 중에 대표적으로 활용되는 것이 객체가 움직이는 영역을 검출하여 그것을 추적하는 방법이다. 객체의 움직임 영역을 활용한 방법 중 가장 많이 쓰이는 방법은 영상 간의 차영상을 이용한 방법과 광류를 활용한 방법이 있다.

차영상을 이용한 방법은 영상의 인접한 프레임(frame)을 비교하여 변화한 부분의 픽셀(pixel) 차이를 추출하고 추출된 영상을 시간에 따라 추적하는 방법이다. 차영상은 카메라로부터 입력되는 컬러 영상을 그레이(gray) 영상으로 변환 후 프레임 사이의 차이를 구한다. 그 다음 처리 속도를 빠르게 하고 데이터의 활용을 용이하게 하기 위해 임계값을 설정하여 색상을 흑과 백으로 표현하는 이진화(binartization)를 진행한다. 차영상의 과정은 다음과 같다.



(a)

(b)

[그림-2] 그레이 영상 변환

카메라로 입력되는 [그림-2]의 (a)와 같은 컬러 영상을 (b)처럼 그레이 영상으로 변환 한다. 그레이 처리한 영상을 이전 영상인 [그림-3]의 (a)와 현재 영상인 (b)의 픽셀 값을 비교하여 감산(減算)하면 (c)와 같이 변화가 없는 부분은 표시되지 않으며 변화가 일어난 픽셀만 도출된다.



(a) Frame n



(b) Frame n+1



(c) 차영상

[그림-3] 프레임 비교 및 차영상

[그림-3]의 (c)는 영상의 밝고 어두운 차이를 나타내는 그레이 스케일 (gray scale)⁷⁾ 값을 구한 것이다. 여기서 도출된 값은 가장 어두운 값

7) 하양에서 검정 사이의 회색의 점진적인 단계 범위로 명도 차의 척도

인 0으로부터 가장 밝은 값인 255까지로 표현 된다. [그림-4]와 같이 대상 피사체의 움직임 영역을 명확하게 하고 컴퓨터가 다루는 데이터의 양을 줄이기 위해 특정 임계값을 설정해 픽셀의 명암 값을 흑색인 0값과 백색인 255값으로 이진화 한다. 여기서 특정 임계값은 조도와 피사체의 움직임 속도에 따라 다르게 적용된다.



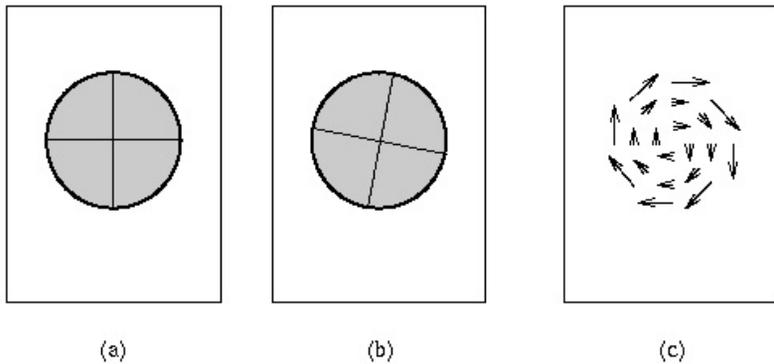
[그림-4] 차영상의 이진화

광류는 “연속하는 두 프레임 간 차이를 통해 움직임을 추정하는 방법 중 하나로서 명도를 서서히 변화시켜 영상에서 발생하는 움직임의 방향과 속도를 벡터적으로 나타내는 방식이다.”⁸⁾ 차영상 과정 중 그레이 영상으로 변환하는 과정까지는 동일하며 이후 픽셀별로 벡터(vector)⁹⁾ 값을 구한다. 이 때 모든 픽셀의 이동 벡터를 얻으려면 상당한 연산을 요구하므로 피사체의 표면·모서리 등의 특징 점을 선택해 계산한다. 특징 점의 개수를 지정하고 계산하는 방식은 Lucas - Kanade 방식,

8) 이해정 외, “Edge Detection과 Lucas-Kanade Optical Flow 방식에 기반한 디지털 영상 안정화 기법” 『로봇학회 논문지』, 제5권 제2호 (2010), 87쪽.

9) 크기와 방향을 동시에 나타내는 물리량

Horn - Schunck 방식, Buxton - Buxton 방식 등 여러 가지가 존재한다. 사물이 [그림-5]의 (a)에서 (b)로 변화한다면 특징 점의 움직임 방향과 크기는 (c)처럼 화살표의 방향과 크기로 표현할 수 있다.



[그림-5] 광류의 흐름

10)

2. 작품에 적용한 기술 구현

1) 차영상을 활용한 트래킹 환경 구현

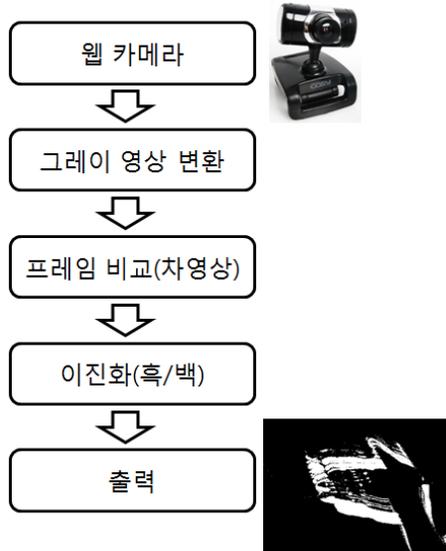
객체의 움직임을 수치화하여 컴퓨터에서 활용할 수 있도록 입력도구로 웹(web) 카메라를 이용했다. 컴퓨터에서 입력된 데이터를 처리하는 도구는 Max11)를 사용했다.

전체적인 차영상 구현 과정은 [그림-6]과 같다. 웹 카메라로 입력된 영상을 그레이 영상으로 변환하고 프레임을 비교하여 차영상을 구현

10) <http://user.engineering.uiowa.edu/~dip/lecture/Motion2.html>

11) Cycling74에서 개발한 음악, 사운드, 비디오, 인터랙티브(interactive) 미디어 응용 프로그램 저작 도구

다. 차영상을 흑과 백으로 이진화하고 이를 출력한다.



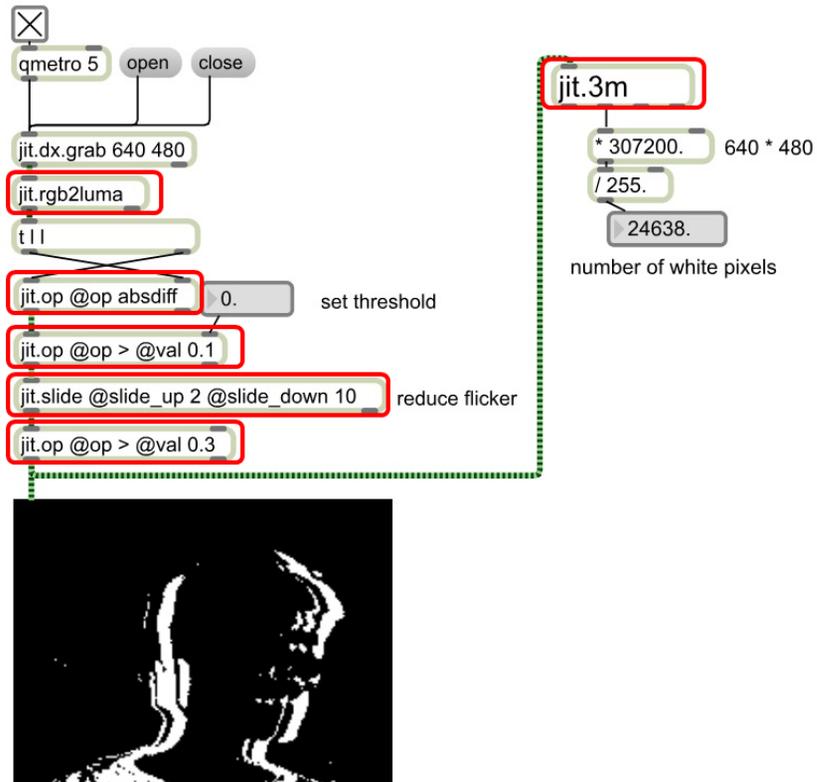
[그림-6] 차영상 구현 과정

웹 카메라가 컴퓨터에 제공하는 값 중 명암 값을 구하기 위해 영상을 그레이로 변환하는 과정은 [그림-7]과 같이 Max의 `jit.rgb2luma` 오브젝트를 이용하였다.

그레이 변환 영상의 인접 프레임을 실시간으로 비교하여 두 프레임간의 차이 값을 얻기 위해 2개의 값 또는 하나의 단위로 된 연산을 수행하는 `jit.op` 오브젝트를 활용하였다. 인접 프레임에서 비교된 차이 값은 차이의 크기 값 즉 절댓값만이 필요하고 음수와 양수 정보는 필요하지 않기 때문에 `jit.op`의 연산자는 절댓값을 결과로 보여주는 `absdiff`를 사용하였다.

인접 프레임 차이의 절댓값은 가장 어두운 값인 0으로부터 가장 밝은 값인 255까지 생성된다. 그런데 인접 프레임 차이의 절대 값은 주변

환경, 특히 빛 또는 조명에 영향을 받는다. 빛 또는 조명의 변화는 영상정보의 변화를 야기해 이 또한 특정 움직임으로 간주되기 때문에 jit.op의 val연산자를 사용하여 임계값을 조정하였다. 본 연구에서는 임계값을 0.1로 설정하여 값이 0.1 이상인 데이터만 사용하도록 하였다. 임계값이 0.1 이상이면 픽셀이 백색으로 표시되어 이진화가 완료된다.



[그림-7] 차영상을 활용한 트래킹 환경 구현 패치

또한 영상의 깜빡임 현상을 줄이고 이동하는 객체의 형상을 예술적으로 표현하기 위해 jit.slide 오브젝트를 사용하였는데 이는 영상의

흑과 백의 변화를 부드럽게 해주는 역할을 한다. `jit.slide` 오브젝트를 거쳐 나온 데이터는 다시 한 번 `jit.op`의 `val` 연산자를 통과시켜 임계값 이하의 노이즈를 제거한다.

본 연구에서는 이진화가 완료된 영상에서 백색으로 표시되는 픽셀의 수를 작품에 활용하였는데 백색으로 표시되는 픽셀 값을 다음과 같이 구하였다. `jit.3m` 오브젝트를 사용하여 이진화 영상의 평균값을 도출하고 본 연구에서 차영상 트래킹 환경을 구현하기 위해 설정된 화면 크기인 640×480 값을 곱한 후 다시 명암의 최댓값인 255로 나누어 순수한 백색 픽셀 값을 구하였다. 백색 픽셀 값은 객체의 움직임 값을 의미한다.

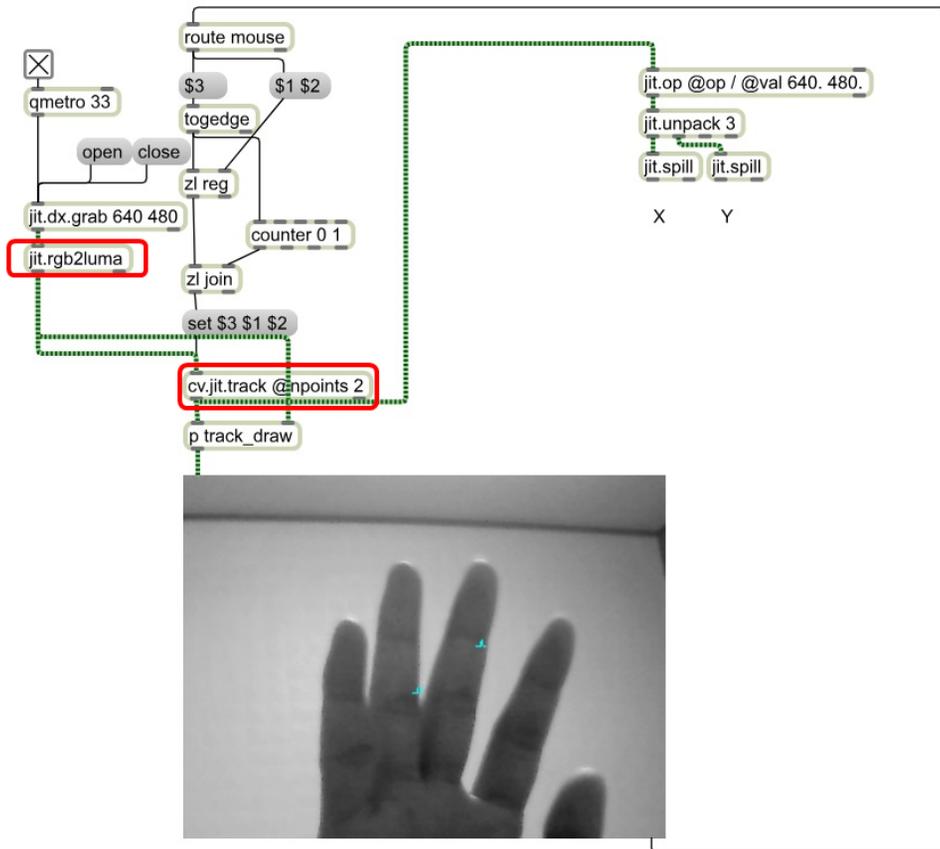
2) 광류를 활용한 트래킹 환경 구현

광류를 활용한 트래킹 역시 앞서 객체의 움직임을 찾아내기 위해 이용한 웹 카메라와 컴퓨터를 동일하게 사용한다. 또한 그레이 영상 변환 과정도 같은 순서로 진행한다. [그림-8]은 광류의 구현 과정을 나타낸 것이다.



[그림-8] 광류 구현 과정

특징점을 추출하는 방식은 Lucas - Kanade 방식을 선택하였다. “Lucas - Kanade 알고리즘은 밝기 항상성(brightness constancy), 시간 지속성(temporal persistence), 공간 일관성(spatial coherence)의 3가지 가정을 기반으로 특징역역을 추적한다.”¹²⁾ [그림-9]는 광류를 활용한 트래킹 환경 구현을 위한 Max 패치다.



[그림-9] 광류를 활용한 트래킹 환경 구현 패치

12) 김성민 외, “초음파 영상에서의 Optical Flow 추적 성능 향상을 위한 전처리 알고리즘 개발 연구” 「전자공학회 논문지」, 제47권 SC편 제5호 (2010), 26쪽.

특징점을 추출하고 객체를 트래킹하기 위해 Jean-Marc Pelletier¹³⁾가 Lucas - Kanade 방식을 기반으로 개발한 Max 오브젝트인 cv.jit.track 오브젝트를 사용하였다. cv.jit.track는 컴퓨터 비전(computer vision) 기술을 이용한 오브젝트이다. 컴퓨터 비전이란 인간 눈의 기능과 동일한 형태를 컴퓨터에 행하게 하는 기술로써 기본적인 목적은 한 장 또는 여러 장의 영상으로부터 3차원 세계의 특성, 즉 물체의 형상이나 위치 같은 기하학적 특성(geometric property)과 물체의 속도나 운동 같은 동적 특성(dynamic property)을 계산하는 것이다.

본 연구에서는 광류를 활용한 트래킹의 대상 객체를 2개로 하였다. 따라서 cv.jit.track의 속성 중 트래킹 대상 객체 수를 설정할 수 있는 npoints를 사용하였고 설정은 2로 하였다. 트래킹 하고자 하는 화면상의 객체를 마우스로 클릭하면 대상 객체를 실시간으로 트래킹 하게 된다.

13) International Academy of Media Arts & Sciences (IAMAS) 소속 당시 개발

III. 연구 기술의 작품 적용

1. 작품의 내용

치매 환자인 할머니는 마치 어린 아이와 같아 모든 것이 신기하고 서툴다. 늘 사고를 치고 보살핌이 필요한 할머니지만 가끔 좋아하는 음악에 눈망울이 선명해 진다. 어릴 적부터 음악을 사랑하여 음악을 직업으로 하고 싶었지만 주변 상황이 여의치 못해 꿈을 이루지 못한 것이다. 하지만 평생 음악을 사랑한 무의식을 치매라는 병도 지우지는 못했다. 할머니의 주변에 존재하는 익숙했지만 낯선 사물들이 이상한 소리를 내는 듯 하더니 이내 음악으로 변한다. 이와 같이 치매를 앓고 있는 할머니가 느끼고 행동하는 것들을 작품으로 표현하고자 하였다.

정적이던 할머니는 음악에 심취해 방안 이곳저곳을 돌아다니며 음악 소리로 변한 주변 음과 어울리는 독립적 소리를 간병인과 함께 발생하게 된다.

2. 작품의 구성

1) 작품의 구조

본 작품의 구조는 [표-1]과 같이 A-B-C-C'-D-B'의 여섯 부분으로 나뉜다. 사운드 및 음악은 D를 제외한 나머지 전 부분이 알고리즘 음

악(algorithm music)으로써 실시간으로 생성된다. 영상 역시 D를 제외한 나머지 전 부분이 알고리즘으로 실시간 생성된다.

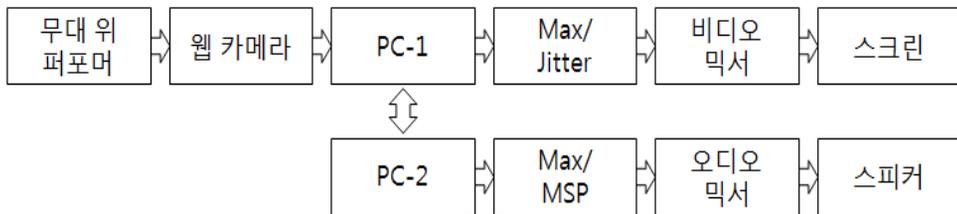
section A와 B에서는 할머니가 혼자 등장하여 음악으로 변해가는 사물의 소리를 느끼며 점차 동화되어 가고 section C에서는 간병인이 등장하여 할머니와 함께 동일한 사운드를 생성한다. section C'에서는 할머니와 간병인이 서로 다른 사운드를 생성하며 전개되고 section D에서는 강렬한 테이프음악에 맞춰 두 퍼포머의 움직임이 극대화 된다. section B'에서는 할머니가 안정을 되찾으며 작품이 마무리 된다.

구성	시간	내용	사운드	영상	카메라
section A	0:00~ 0:32	할머니의 방에 존재하는 사물들의 소리가 각각 등장	-사물 소리	각 사물 이미지	
section B	0:32~ 2:02	등장한 사물들의 소리가 점차 음악으로 변형	-그래놀러 사운드	각 사물 이미지의 혼재	
section C	2:02~ 3:32	할머니와 간병인이 움직임이며 연속적 사운드를 생성	-그래놀러 사운드 -프로세싱 사운드	차영상 (차영상 트래킹)	On
section C'	3:32~ 5:02	할머니와 간병인이 서로 다른 두 개의 사운드를 생성	-그래놀러 사운드 -프로세싱 사운드	색상 파티클 (광류 트래킹)	
section D	5:02~ 6:02	할머니가 간병인의 혼란이 가중	-테이프음악 -주파수변조 사운드	사물 이미지 파괴	Off
section B'	6:02~ 7:18	간병인의 노력으로 안정의 도모 및 퇴장	-그래놀러 사운드	각 사물 이미지 혼재	

[표-2] 작품 구조

2) 작품의 시스템과 무대 구성

<할머니의 방>은 전체적으로 한 개의 웹 카메라와 두 대의 컴퓨터로 구성된다. 두 대의 컴퓨터 중 한 대는 음악작업에 사용되며 나머지 한 대는 영상작업에 사용된다. 본 작품의 전체적인 시스템 구성은 [그림-10]과 같다.



[그림-10] 시스템 구성도

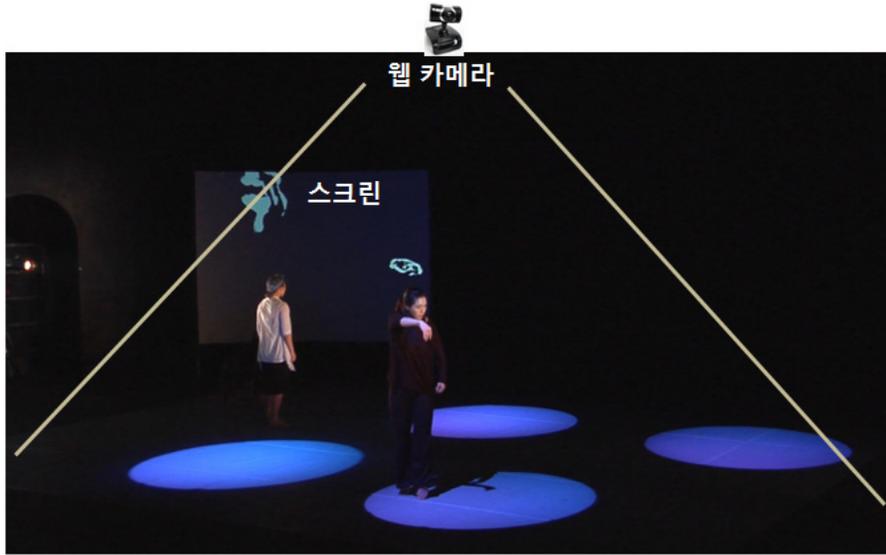
추적 대상이 되는 퍼포머의 중첩을 피하기 위해 무대 위 천장에 수직으로 웹 카메라를 설치하여 퍼포머의 움직임에 부감(high angle)¹⁴⁾으로 얻고 이 영상 데이터를 PC-1로 입력한다. PC-1로 입력한 영상 데이터는 Max를 통해 [그림-7]과 [그림-9]와 같이 가공이 되어 PC-1과 PC-2에서 각각 영상과 오디오작업에 사용된다. 이때 PC-1과 PC-2 사이의 데이터 전송은 Wi-Fi¹⁵⁾를 이용한 UDP(user datagram protocol)¹⁶⁾ 무선통신을 활용하였다.

14) 카메라 앵글의 일종으로 카메라를 높은 위치에 놓고 피사체를 내려다보는 촬영 각도를 말함

15) 고성능 무선통신을 가능하게 하는 무선랜 기술

16) 인터넷에서 정보를 주고받을 때, 서로 주고받는 형식이 아닌 한쪽에서 일방적으로 보내는 방식의 통신 프로토콜

[그림-11]과 같이 웹 카메라가 무대 위에 설치되어 무대 전체를 인식할 수 있고 퍼포머 간에 중첩도 피할 수 있어 무대의 활용 폭이 넓으며 트래킹도 원활히 수행할 수 있다.



[그림-11] 무대 구성도

3. Section별 작품의 구성 및 효과

1) Section A

Section A에서는 치매에 걸린 할머니의 고독함과 그 고독함으로 느껴지는 방안의 일상적 소음들을 표현하였다. 사운드는 시계, 휠체어, 유리잔, 고물 TV 등 4개의 사물 소리를 8초 간격으로 페이드인·페이드아웃(fade in·fade out) 처리하여 플레이한다.

시계, 휠체어, 유리잔 사운드는 샘플링(sampling)¹⁷⁾하여 groove~오브젝트로 플레이하였고, 고물 TV 사운드는 noise~오브젝트를 사용하여 화이트 노이즈(white noise)로 표현하였다. 최초 시계 사운드가 플레이 되고 플레이가 끝나면 자동으로 휠체어 사운드가 플레이 되며 이어서 유리잔과 고물 TV 사운드가 자동으로 플레이 된다.

영상은 [그림-12]와 같은 이미지들이 사운드의 플레이 시간과 순서대로 페이드인·페이드아웃하며 화면에 보인다. 영상 역시 순차적으로 자동 플레이 되는 것이다.



[그림-12] 작품에 쓰이는 이미지들

17) 기존의 음원이나 아날로그 형식으로 존재하는 자연음을 녹음하여 활용하는 기법

[그림-13]은 퍼포머의 등장 장면이다. 퍼포머를 스크린과 빔 프로젝터 사이에 위치시켜 그림자를 이용한 영상을 연출한 것이다.



[그림-13] 퍼포머의 등장

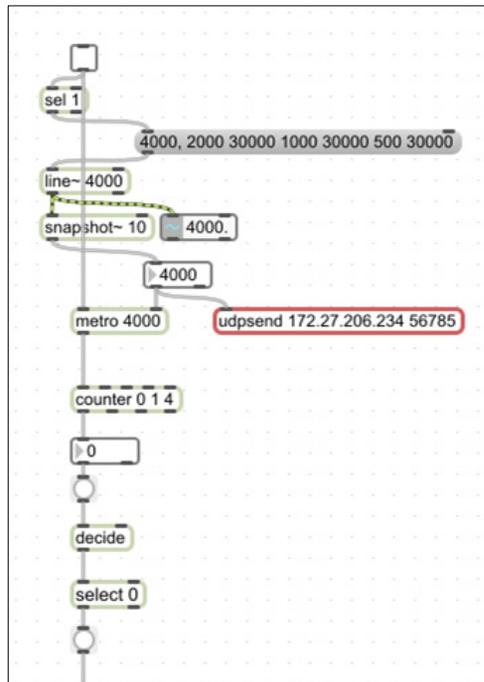
2) Section B

Section B에서는 할머니가 일상적 사물의 소리들을 점차적으로 음악으로 느끼게 되며 그것에 심취해가는 과정을 그렸다.

① 음악 알고리즘

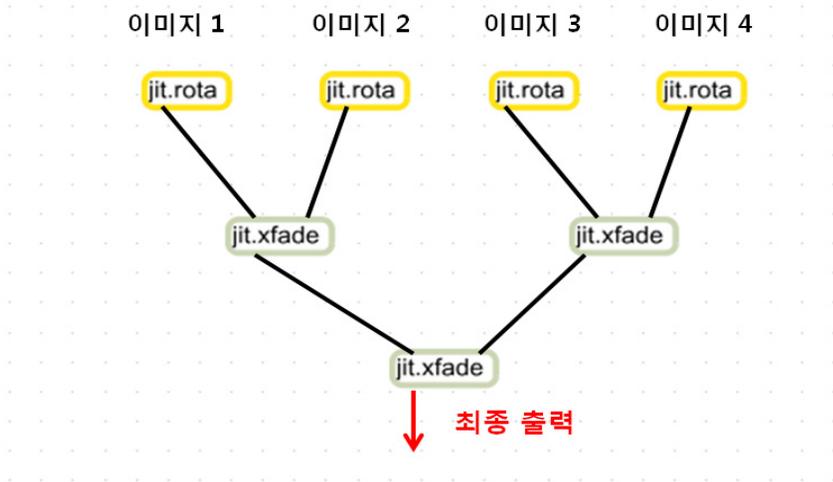
음악의 속도가 점점 빨라지고 음량이 점차 커지는 효과를 구현하기 위해 [그림-14]와 같이 line~오브젝트를 사용하였다. 다중 값을 연속적으로 표현하기 위해서도 line~를 사용하였으며 숫자 값의 사용을

위하여 snapshot~오브젝트를 활용하여 시그널(signal) 값을 숫자 값으로 변환하였다. 한편 기계적인 딱딱한 변화를 피하고 예측할 수 없는 신선함을 불어넣기 위해 decide오브젝트를 사용하였다. decide는 0과 1 값 중 랜덤(random)하게 선택하는 오브젝트다. 이와 같이 생성된 신호들은 Section A의 서브패치(subpatch)로 보내져 Section A에서 활용했던 사운드들을 다시 활용하게 된다.



[그림-14] 사물 소리에서 음악소리로의 변화 과정

또한 음악적 색깔을 더하기 위한 목적으로 [그림-14]와 유사한 과정으로 생성된 신호를 [그림-15]와 같이 가상악기(virtual studio technology instrument)에 연결하여 음악을 생성하였다.



[그림-16] Section B의 영상 출력 알고리즘

jit.rota 오브젝트는 각 사물 이미지의 크기와 회전을 변화시키기 위해 사용됐고 jit.xfade 오브젝트는 이미지들이 자연스럽게 섞이도록 하기 위해 쓰였다.

[그림-17]은 실제 공연에 적용된 사물 이미지 혼재 영상이다.



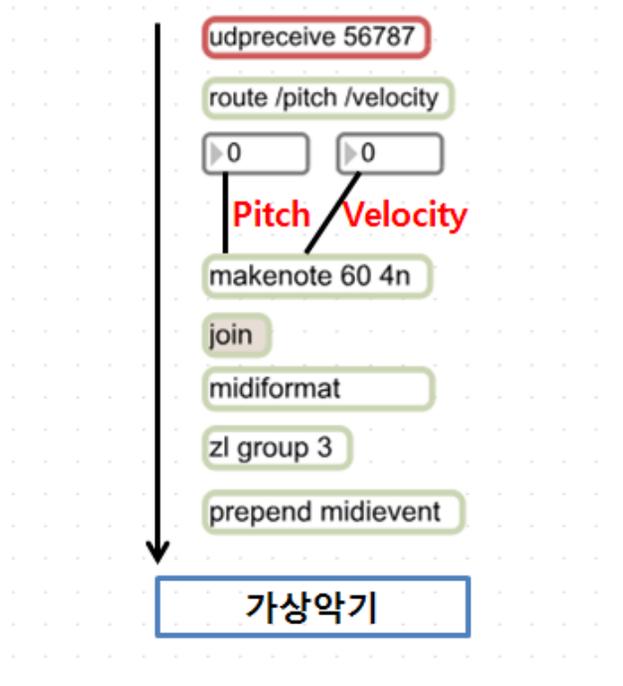
[그림-17] Section B에서 출력되는 영상

3) Section C

Section C는 간병인이 등장하여 할머니와 교감을 나누는 부분이다.

① 음악 알고리즘

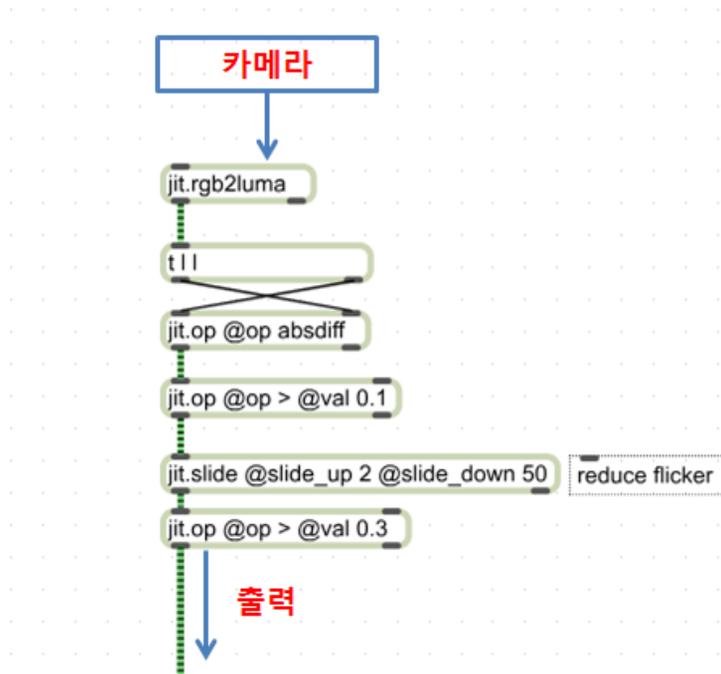
멜로디(melody)는 영상처리를 담당하는 PC-1에서 생성된 백색 픽셀의 값을 PC-2에서 받아 [그림-18]과 같이 피치(pitch)와 벨로시티(velocity)로 활용하고 이를 MIDI(musical instrument digital interface) 데이터로 변환하여 가상악기에서 생성하였다.



[그림-18] Section C의 음악 알고리즘

② 영상 알고리즘

Section C의 영상은 차영상 기술을 활용하는 것으로 무대 위에서 움직이는 할머니와 간병인의 움직임을 카메라로 촬영하고 각 프레임을 비교하여 움직임이 발생하는 픽셀을 백색으로 표시하여 영상 출력한다. [그림-19]는 Section C의 영상 Max 패치다.



[그림-19] Section C의 영상 알고리즘

[그림-20]은 실제 공연에서 무대 위 퍼포머의 움직임을 실시간으로 스크린에 형상화하는 장면이다.



[그림-20] Section C의 영상과 무대

4) Section C'

Section C'는 무대의 조명과 스크린의 영상이 바뀌며 할머니와 간병인이 서로 다른 멜로디를 생성한다.

① 음악 알고리즘

Section C'의 오디오 작업에 사용되는 데이터는 PC-1에서 광류를 활용한 할머니와 간병인의 객체별 추적 데이터를 PC-2에서 받은 것이다. 나머지 멜로디 생성과정은 Section C와 동일하며 할머니는 피아노, 간병인은 전자기타를 연주한다.

② 영상 알고리즘

Section C'의 영상은 Section C의 영상에 기하학적 모양과 색을 표현하는 jit.gl.mesh오브젝트를 사용하여 변화를 주었다. jit.gl.mesh오브젝트의 입력 값의 종류는 [그림-21]과 같고 실제 입력 값은 퍼포머의 움직임에 따라 실시간으로 바뀐다.

```
jit.gl.mesh @draw_mode points @color 1 1 1 1 @antialias 1 @scale 2 2 2 @position -1 -1 1 @transform_reset 2
```

[그림-21] jit.gl.mesh오브젝트의 구성

jit.gl.mesh오브젝트를 사용한 결과 [그림-22]와 같이 퍼포머의 움직임에 따라 스크린에 추상적 영상이 생성되는 것을 볼 수 있다.



[그림-22] jit.gl.mesh오브젝트 반영 영상

5) Section D

Section D는 할머니의 감정이 고조되는 부분이다. 음악은 큰 북소리가 강조된 데이프음악이 흐르고 퍼포머가 무대의 사방 가장자리에 위치할 때 주파수변조 사운드가 생성된다. 영상은 [그림-23]과 같이 Section B를 그대로 사용하지만 데이프음악의 음량이 커지고 빠르기가 증가하면서 좀 더 복잡하고 밝은 색으로 표현된다.



[그림-23] Section D의 영상과 무대

6) Section B'

Section B'는 작품의 결말 부분으로 할머니가 점차 안정을 되찾아 일상으로 돌아가는 모습을 표현했다. Section B의 구성과 같지만 알고리즘을 반대로 작동시켜 점차 잦아드는 분위기를 연출하고자 했다.

IV. 결론 및 문제점

본 연구는 광류 트래킹을 이용하여 간단한 장비로 관객이 직관적으로 느낄 수 있는 멀티미디어 공연을 만들 수 있도록 하는 것이다. 작품 <할머니의 방>은 퍼포머가 아무런 장치를 하지 않고 무대의 모든 공간을 활용해 인터랙티브 공연을 할 수 있는 시스템을 구축하는데 목적을 가지고 있다. 연구를 통해 얻은 성과는 다음과 같다.

첫째, 인터랙티브 공연을 위해 별도의 장비 없이 카메라만으로 퍼포머를 추적하여 데이터를 추출했다. 장비의 제약이 줄어들어 공연에 활용되는 공간을 넓게 활용하는데 효과적이었다.

둘째, 객체 추적을 통한 실시간 인터랙티브 공연 시스템을 구축하였다. 객체 추적을 통해 얻어진 데이터로 음악과 영상을 실시간으로 제어하여 관객으로 하여금 퍼포머의 움직임과 음악, 영상이 유기적으로 결합돼 있다는 것을 직관적으로 느끼게 하는데 효과적이었다.

하지만 본 연구에서 몇 가지 문제점이 있었고 향후 연구를 통해 보완해 나갈 계획이다. 문제점은 다음과 같다.

첫째, 객체의 움직임보다 조명의 변화에 더 민감하게 반응한다는 것이다. 픽셀 간의 차이로 움직임을 검출하기 때문에 조명 변화가 생기면 다른 영상 값을 가지게 되므로 정확한 검출이 어려웠다.

둘째, 카메라 설치 장소에 제약이 있다. 하나 이상의 다중 객체를 추적할 때 객체들이 서로 겹치면 추적에 오류가 생기기 때문에 카메라를 무대의 지붕에 설치해야 했다. 무대 지붕이 낮으면 카메라가 인식하는 범위도 줄어들어 그만큼 활용할 수 있는 공간이 줄어들게 되었다.

본 작품에서는 카메라만을 사용해 공연에 활용되는 공간을 넓히고 관객이 무대의 퍼포머 동작에 의한 음악과 영상제어를 실감할 수 있도록 기획되었다. 연구를 통해 드러난 문제점들을 보완해 나간다면 더 많은 인터랙티브 멀티미디어 작품 창작자들이 간편하면서도 효과적인 공연을 할 수 있을 것으로 기대된다.

Keyword (검색어): 멀티미디어음악(multimedia music), 인터랙티브 아트(interactive art), Max/MSP, Jitter, 광류(optical flow), 차영상(difference image)

E-mail: meyousick@naver.com

참고문헌

1. 참고논문

- 강동구 외, “객체 추적 카메라 제어를 위한 고속의 움직임 검출 및 추적 알고리즘” 「방송공학회 논문지」, 제7권 제2호 (2002)
- 김성민 외, “초음파 영상에서의 Optical Flow 추적 성능 향상을 위한 전처리 알고리즘 개발 연구” 「전자공학회 논문지」, 제47권 SC 편 제5호 (2010)
- 김태희, “옵티컬 플로우와 마스킹에 의한 실시간 인터랙티브 비디오 개발” 「한국콘텐츠학회논문지」, 제11권 제6호 (2011)
- 오명관 외, “차영상에서의 히스토그램을 이용한 적응적 임계값 결정” 「한국콘텐츠학회 논문지」, 제4권 제3호 (2004)
- 오상협, “차영상을 이용한 잡음 제거 및 얼굴 영역 검출 하드웨어 구현” 「경북대학교 대학원 전자전기컴퓨터학부 회로 및 임베디드 시스템공학전공」, (2012)
- 윤기선 “키넥트를 이용한 인터랙티브 멀티미디어 작품 제작연구” 「동국대학교 영상대학원 멀티미디어학과」, (2012)

- 이동빈 “실시간 동작 추적을 이용한 멀티미디어음악 제작 연구” 「동국대학교 영상대학원 멀티미디어학과」, (2010)
- 이혜정 외, “Edge Detection과 Lucas-Kanade Optical Flow 방식에 기반한 디지털 영상 안정화 기법” 「로봇학회 논문지」, 제5권 제2호 (2010)
- 전유진 “컬러 트래킹을 이용한 멀티미디어음악 창작 연구” 「동국대학교 영상대학원 멀티미디어학과」, (2005)
- 정하영, “광류기반 다중 이동 객체의 실시간 인식 및 추적 시스템에 관한 연구” 「관동대학교 대학원 전자통신공학과」, (2010)
- 조성련, “얼굴인식을 위한 누적 차영상 기반의 얼굴영역 검출에 관한 연구” 「연세대학교 대학원 전기전자공학과」, (2001)
- 조혜리, “광류를 이용한 3차원 운동 추정 기법에 대한 연구” 「서울대학교 대학원 전기공학부」, (2001)
- 최준환 “무용 동작인식을 이용한 인터랙티브 멀티미디어작품 제작 연구” 「동국대학교 영상대학원 멀티미디어학과」, (2014)

2. 기타

- B. D. Lucas, T. Kanade, “An Iterative Image Registration Technique with an Application to stereo Vision”, Proceedings of the 1981 DARPA Imaging Understanding Workshop. (1981), pp. 121-130.
- 주홍식 외, “차영상과 컬러정보를 이용한 목표물 계수 알고리즘” 「한국정보기술응용학회 2004년도 춘계학술대회 디지털 컨버전스 (Digital Convergence)와 경영혁신」, (2004)

3. 인터넷

- Max/MSP
<http://www.cycling74.com>
- wikipedia
http://en.wikipedia.org/wiki/Optical_flow
http://en.wikipedia.org/wiki/Lucas%E2%80%93Kanade_method
- optical flow
<http://user.engineering.uiowa.edu/~dip/lecture/Motion2.html>
- 사례연구
<http://vinyl-i.com/Work/#!vlab/0077>

Abstract

A Study on multimedia music performance by Camera Tracking (Focusing on the multimedia music <Grandmother's Room>)

Namyoung Hur

<Grandmother's Room> is a multimedia music performance using a camera tracking and real-time visual image. This piece is controlled by performer that using a camera tracking by optical flow and difference image at the same time. The goal of this study is to find the way of artistic expression with sound and image by using camera.

This study suggests concept of music from listening music to watching music, and you provide a new multimedia works by camera tracking. The sound of <Grandmother's Room> is analyzed by computer and converted to data, and it used to change parameters of sound and visual image.

Using a Max/MSP programs to analyze the values of camera tracking, and this programs used the real-time sound as a data to apply to the image. A data which is an analysis of optical flow and difference image controlled the image.

Jitter is used for image production that controlled by camera tracking data. performer is harmony with sound effects and image which controlled by camera.

Researching this system, to analyze the optical flow and difference image from camera in real time is define a mutual relation between sound and image. Through such technology, the camera tracking was able to control the music and image. In addition, The difference image and the optical flow technologies were found to be effective when the audience intuitively can feel the systematic combination of movement of performer, music and image.

부록 1 : 첨부 DVD

1. 할머니의 방.mp4 : 공연실황
(2013년 11월 15일, 이해랑 예술극장)
2. Max Patch : <할머니의 방> 공연을 위한 Max 패치