

패턴인식 기술을 이용한 인터랙티브 설치작품 제작연구

하승연, 김 준

동국대학교 영상대학원 멀티미디어학과

A Study on the Interactive Installation Art Work Using Pattern Recognition System

Seung-Yun Ha, Jun Kim

Department of Multimedia Dongguk Univ. Graduate School of Digital Image & Contents

요 약

과학 기술의 발전에 의해 음악, 미술, 공학의 경계가 무너지고 서로 접목되면서 종합적인 장르를 만들어 내고 있다. 20세기 중반 이후 나타나기 시작하던 설치미술 역시 이러한 발전에 의한 장르 융합이 이루어지면서 기존의 작가의 의도를 감상하는 차원을 넘어 감상자에게 역동적으로 반응하고 살아 움직이는 작품으로 발전하고 있다. 본 논문은 최근 미디어아트 분야에서 활발하게 나타나고 있는 CV(Computer Vision)를 이용하여 소리 합성과 음악 믹싱을 위한 설치 작품 제작을 목표로 하였다.

설치작품을 제작하고자 한다.

I. 서 론

1.1 작품의 의도

현대 예술에서 공학은 빼놓을 수 없는 분야이다. 예술작품을 위한 공학기술의 이용이나 공학의 미적표현을 위한 예술작품 등은 최근 그 경계가 모호해지면서 새로운 장르로 결합되었다. 이러한 장르 중 미디어아트 분야는 대중적, 예술적, 상업적 성과를 가능하게 하여 국내외에서 많은 연구가 이루어지고 있다. 최근 CV 기술을 이용한 연구와 작품들은 컴퓨터와 인간의 커뮤니케이션의 많은 가능성을 보여주고 있다.

설치 미술은 주어진 공간을 활용하여 다양한 감각의 요소를 통해 작가의 메시지를 전달한다. 하지만 흔히 말하는 설치 미술에 있어서 감각의 자극은 작품과 감상자의 커뮤니케이션이 이루어진다고 하기 힘들다. 작가가 설치해 놓은 매체만을 보고 듣고 느끼는 것일 뿐 감각과 작가의 메시지가 상호작용을 하고 있지는 않다.

인터랙티브 작품은 예술의 소통문제를 컴퓨터의 대화형 조작 환경으로 보완하고 있다. 따라서 본 논문에서는 CV 기술을 이용하여 사용자에게 작품의 메시지에 대한 흥미를 유발하고 보다 더 쉽게 접근할 수 있는 인터랙티브한

1.2 사례 조사

실시간 사운드 처리를 위한 TUI(Tangible User Interface)¹⁾는 오래전부터 연구되어 왔다. 대표적인 인터페이스로는 MIT Media Lab의 AudioPad(그림 1)와 스페인의 Pompeu Fabra University의 reactable(그림 2)이 있다. 두 인터페이스 모두 테이블 위에 블록을 위치시키고 이동하는 방법으로 컨트롤을 하고 있으며 블록의 위치와 회전 값을 인식한다는 점이 동일하다. 화려하고 직관적인 디자인으로 세밀한 부분까지 제어가 가능하지만 설치를 목적으로 한 인터페이스가 아니라서 사운드 처리에 대한 사전지식이 없으면 사용이 불편하다는 단점이 있다. 본 논문의 작품에서는 전문가를 위한 인터페이스보다는 누구나 쉽게 조작하여 소리를 만들 수 있도록 하는 것에 중점을 두고 진행하였다.

1) 실제로 물건을 만지고 느끼고 잡고 옮기는 등의 행위를 통해 디지털 정보를 조작하는 인터페이스 기술

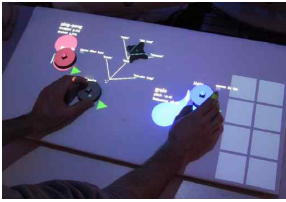


그림 1 : AudioPad



그림 2 : reactable

II. 본 문

2.1 구조

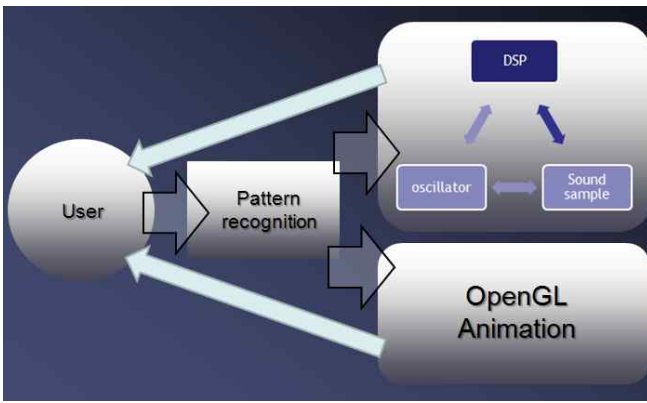


그림 3 : 기술적 구조

이 작품은 감상자가 소리를 듣고 디스플레이 되는 화면을 보고 인터페이스를 만지는 여러 감각과의 상호작용을 통해 간단한 조작만으로 소리의 추가, 합성, 변조 등을 가능하게 하여 음악 또는 소리에 대한 흥미와 이해를 유발시키는 것이 목적이다.



그림 4 : 패턴의 예

사용자가 특정 패턴(그림 4)이 새겨진 블록을 카메라가 설치된 테이블 위에 위치시키거나 이동, 회전하게 되면 패턴의 종류와 블록의 위치에 해당하는 소리와 영상이 사용자에게 전달되게 된다. 사용자가 조작한 패턴의 인식 정보는 Max/MSP²⁾로 전송된다. Max/MSP가 전송받은 신호 중 필요한 정보를 분류하고 목적에 맞는 데이터로

2) 음악, 소리, 멀티미디어 등을 그래픽 환경에서 실시간으로 제어할 수 있는 오브젝트 기반의 컴퓨터 프로그램.

변환하여 Jitter³⁾와 MSP로 보내고 영상과 소리의 처리가 이루어지게 된다.

2.2 패턴 인식

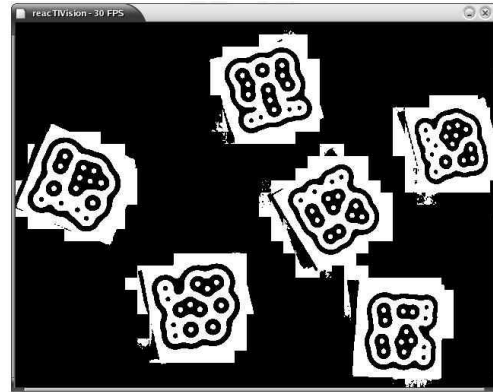


그림 5 : reactTVision

패턴은 reactTVision⁴⁾(그림 5)에서 인식하며 여러 가지 패턴은 각각 하나의 독립적으로 object로 Max/MSP에서 정보로 처리한다. 패턴이 새겨진 블록의 움직임이 카메라를 통해 reactTVision으로 보내지면 그 화상에서 패턴을 찾아내고 패턴의 정보 리스트를 Max/MSP로 전송한다. 이렇게 전송되는 패턴의 리스트는 최대 초당 30번 갱신되어 처리되며 새로운 정보(블록의 상태에 변화가 없는 경우)가 발견되지 않거나 부적합한 정보 리스트를 전송받으면 Max/MSP에서 소리와 영상을 위한 분석기로 보내지 않게 된다. 또한 reactTVision과 카메라 사이에서 생기는 인식 오류로 인해 패턴의 종류나 위치 유무에 잘못된 정보가 들어오는 것을 방지하기 위해 허용오차범위 이내의 리스트가 2/15초 이상 유지될 때만 처리하도록 하였다. 이 과정을 리스트필터링이라 하고 필터링중 리스트와 리스트 사이에 큰 편차가 생기는 경우 근사계산하여 처리한다.

2.3 3D Animation

Max는 많은 데이터를 가진 리스트 중 Visualization에서 사용되는 정보만을 분류하고 OpenGL⁵⁾ Animation을 위한 3D 모델의 종류, 위치, 회전값 등의 목적에 맞는 데이터로 변환하여 Jitter로 보내진다. 3D Animation은 640x480 pixel의 공간을 가지고 있으며 테이블에 놓여진 블록의 위치와 Animation 처리되는 영상에서의 3D 모델

3) Max/MSP 내에서 영상을 실시간으로 제어 할 수 있는 오브젝트 기반의 컴퓨터 프로그램.

4) Pompeu Fabra University에서 개발한 오픈소스 패턴 인식 프로그램

5) Open Graphics Library. 2차원 및 3차원 그래픽을 위한 컴퓨터 산업계의 표준 API

의 위치는 동일하다.

Animation에 사용된 3D 모델은 사운드샘플이나 이펙터의 종류에 따라 다른 형태를 지니며 블록의 회전으로 소리의 크기, 음의 높이, 이펙터의 파라미터가 변하게 되면 그 양의 많고 적음에 따라 3D 모델의 크기나 움직임의 변화가 생긴다.

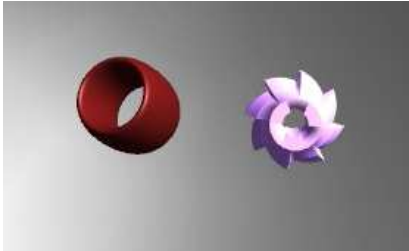


그림 6 : 3D model의 예

2.4 Audio Processing

MSP로 보내진 리스트의 정보에 따라 소리를 재생, 합성하거나 변조되는데 패턴의 종류마다 다른 기능을 가지고 있기 때문에 여러 가지 조합이 가능하다. 예를 들어 두 개의 서로 다른 파형 oscillator 정보를 가진 블록을 조합할 때 각각의 블록은 독립적으로 자신이 가진 음색을 출력하지만 두 블록의 위치에 따라 (블록이 서로 만나거나 거리가 가까워지면) 한 개의 oscillator가 다른 oscillator의 amp에 연결되어 Amplitude modulation이 생긴다. 이는 oscillator뿐만 아니라 sound sample일 경우에도 해당되며 블록의 종류나 위치의 조합에 의해 DSP가 결정된다. 블록 회전은 각각의 기능에 대한 파라미터를 변화시키는데 시계반대방향으로 회전하면 해당 파라미터의 양이 증가하게 되고 시계방향으로 회전하면 양이 감소한다. 회전값은 블록이 테이블에 놓이는 시점이 0이며 1/15초 이내에 10도 이상 회전하게 되면 패턴인식 오류방지를 위해 필터링 처리한다. 또한 reactIVision에서 보내오는 회전값은 증가나 감소에 한계가 없기 때문에 파라미터가 가진 범위 내에서 처리하기 위해 1/15초 마다 수치를 비교하여 이전 프레임 값과 다음 프레임 값의 차이를 0에서 360사이에서 증가하거나 감소시킨다.

2.5 구현

서로 다른 기능을 가진 블록 들의 3D model은 패턴의 정보와는 별개로 단일 또는 링크(블록 들의 조합)된 사운드 시그널의 변화, Volume이나 Frequency에 의해 animation 처리되어 3D model들의 움직임은 사운드 컨트롤과 인터페이스 컨트롤의 상태를 동시에 디스플레이 해준다. 패턴을 교환하거나 추가할 때 사운드 샘플의 looping 시점이 동일하도록 1~2개의 looping point를 설정하여 여러 개의 사운드 샘플들이 동시에 재생되더라도 하나의 음악으로 인식되게 하였다.

블록을 테이블에 위치시키면 해당 블록이 가진 사운드 샘플이 재생된다. 샘플은 블록이 테이블에서 벗어날 때까

지 반복되며 이펙터 블록을 샘플 블록 옆에 위치시키면 두 블록의 색이 같아지면서 서로 연결되었다는 것을 표시한다. 이때 이펙터 블록을 회전시키면 사운드 샘플에 이펙터가 적용되며 그 효과는 두 블록 사이의 거리가 멀어지거나 블록이 테이블을 벗어날 때까지 지속된다. 블록의 테이블상의 좌, 우 위치는 Pan을 조절한다. 샘플 블록을 추가할 때마다 새로운 사운드 샘플이 재생되고 여러 개의 블록이 테이블 위에 위치하게 되면 하나의 완성된 음악이 만들어진다.

III. 결 과

본 논문의 작품을 통하여 단순하게 테이블에 블록을 올려놓고 이동하고 회전시키는 것만으로 짧은 시간에 누구나 쉽게 음악을 만들 수 있는 컨트롤러의 구현에 성공하였으나 조작의 단순함으로 인해 오랜 시간 접하게 되면 지루해지는 문제가 있다. 그 때문에 사운드 샘플과 이펙터, 3D 모델의 종류와 양을 늘리는 방법을 시도해보았지만 컴퓨터의 과부하로 인해 3D Animation의 FPS (Frames Per Second)가 급격히 감소하거나 시스템이 멈춰버리는 현상이 생겼다. 이러한 문제는 그림 7과 같이 reactIVision과 Max/MSP를 서로 다른 컴퓨터시스템에서 구동하고 TCP/IP를 이용하여 이들 간의 정보를 전송시켜 많은 부분에서 생기는 시스템의 한계를 극복할 수 있었다.

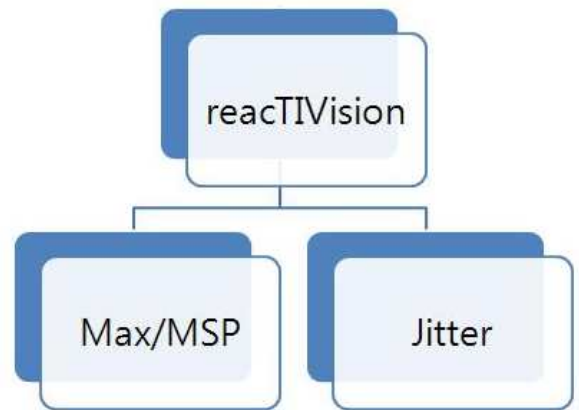


그림 7 : TCP/IP를 이용한 구조

IV. 향후 연구 방향

본 논문의 작품에서 사용된 인터페이스의 3D Animation은 단순히 블록의 위치와 종류만을 표시하는 것에 불과하지만 Animation의 기법이나 효과 등을 추가하여 3D Animation만으로도 사용자에게 흥미를 줄 수 있는 인터페이스로 발전시키고자 한다. 또한 카메라의 각도, 주위 환경의 밝기, 패턴의 크기에 의해 생기는 오류

등을 여러 가지 실험을 통하여 개선해 나갈 것이다.

기술적인 부분과 미적구현에 있어서 부족한 부분을 개선하여 누구나 단순한 조작으로도 시각, 청각, 촉각을 동시에 만족시키고 경험 할 수 있는 설치 작품으로 발전하기를 기대해 본다.

VI. REFERENCES

Patten, J., Recht, B., Ishii, H., "Audiopad: A Tag-based Interface for Musical Performance", in Proceedings of Conference on New Interface for Musical Expression (NIME '02)

http://tangible.media.mit.edu/content/papers/pdf/audiopad_NIME02.pdf

Kaltenbrunner, M. & Bovermann, T. & Bencina, R. & Costanza, E. "TUIO - A Protocol for Table Based Tangible User Interfaces" Proceedings of the 6th International Workshop on Gesture in Human-Computer Interaction and Simulation (GW 2005)

<http://mtg.upf.edu/reactable/pdfs/GW2005-KaltenBoverBencinaConstanza.pdf>

MIT Media Lab - Tangible Media Group.

<http://tangible.media.mit.edu/>

Tangible and Embedded Interaction 2008,

<http://www.tei-conf.org/>

Ishii, H., "Tangible Bits: User Interface Design towards Seamless Integration of Digital and Physical Worlds" (in Japanese). IPSJ Magazine, Vol. 43, No. 3, March 2002, pp. 222-229.

<http://sigchi.org/chi97/proceedings/paper/hi.htm>

TangibleTable, <http://www.tangibletable.de/>