creative commons
COMMONS DEED

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

dCollection

석 사 학 위 논 문

목소리의 실시간 사운드 프로세싱을 이용한 인터랙티브 멀티미디어 퍼포먼스 연구

-멀티미디어 작품 <Adsum (I am present)>를 중심으로-

지도교수 김 준

동국대학교 영상대학원 멀티미디어학과

김 부 경

2025

석 사 학 위 논 문

목소리의 실시간 사운드 프로세싱을 이용한 인터랙티브 멀티미디어 퍼포먼스 연구

-멀티미디어 작품 <Adsum (I am present)>를 중심으로-

김 부 경

지도교수 김 준

이 논문을 석사학위 논문으로 제출함

2024년 12월

김부경의 음악석사학위(컴퓨터음악) 논문을 인준함

2025년 1월

위원장 정 진 헌 (인)

위 원 김 정 호 (인)

위 원 김 준 (인)

동국대학교 영상대학원

목 차

Ⅰ. 서 론 ····· 1
1. 연구 배경 및 목적 ····· 1
2. 사례 연구 및 선행 연구 ····· 3

Ⅱ. 기술 연구 ····· 7
1. 목소리의 특성 ····· 7
2. 사운드 시스템 연구 ····· 9
1) 사운드 시스템 ····· 9
2) 사운드 디자인 연구 ····· 10
① 오디오 신호 변환 ····· 10
② FM synthesis ····· 13
③ RM ····· 16
④ granular synthesis ····· 18
⑤ stereo delay ····· 21
⑥ comb-filter ····· 24
⑦ FFT 분석을 이용한 pitch shift ····· 26
⑧ FFT 분석을 이용한 pfft~ xover~ ····· 27
3. 영상 시스템 연구 ····· 29
1) 영상 시스템 ····· 29
2) TouchDesigner를 이용한 영상 제작 ····· 29
① noise ····· 29
② CHOPs ····· 31

③ TOPs····· 33
4. 공연 및 연동 시스템 연구····· 36
1) 공연 시스템····· 36
2) 음악과 영상의 연동 시스템····· 37
① OSC····· 37
② MIDI 매핑····· 40

Ⅲ. 연구 기술의 작품 적용····· 43
1. 작품 소개····· 43
2. 작품 구성····· 45
1) 음악 구성····· 45
2) 영상 구성····· 48
3) 무대 구성····· 49
3. 사운드 및 영상 기술 적용····· 50
1) A Section····· 50
2) B Section····· 52
3) C Section····· 54
4. 연구 기술의 작품 적용 효과····· 56

Ⅳ. 결론····· 58

참고문헌····· 60
ABSTRACT····· 64
부록 : 첨부 DVD····· 67

표 목 차

<표-1> grainflow.live~오브젝트의 파라미터·····20
<표-2> 음악 구성·····45
<표-3> 영상 구성·····48
<표-4> A section에 적용된 사운드 및 영상 기술 적용·····50
<표-5> B section에 적용된 사운드 및 영상 기술 적용·····52
<표-6> C section에 적용된 사운드 및 영상 기술 적용·····54

그 림 목 차

[그림-1] Reggie Watts의 퍼포먼스 영상캡쳐·····3
[그림-2] Beardyman의 퍼포먼스 영상캡쳐·····4
[그림-3] 작품 <상념(想念)>의 영상 장면·····5
[그림-4] 성별에 따른 모음/a/의 주파수 스펙트럼과 포먼트 분포 표··8
[그림-5] 사운드 시스템 설계도·····10
[그림-6] makenote오브젝트를 이용한 MIDI 신호로 변환·····11
[그림-7] FM synthesis의 기본 구조·····13
[그림-8] FM synthesis를 Max로 구현한 패치·····14
[그림-9] poly~ M4L.fm.FMvoice~를 이용한 Max패치·····15
[그림-10] RM과 AM의 변조 진폭 차이·····16
[그림-11] RM을 이용한 트레몰로 음향효과를 구현한 Max 패치·····17
[그림-12] granular synthesis의 과정·····18
[그림-13] grainflow.live~로 granular synthesis를 구현한 Max 패치 19

[그림-14] delay의 기본 원리 · 21
[그림-15] stereo delay 음향효과를 구현한 Max 패치·····22
[그림-16] ping pong delay의 오디오 신호 흐름·····23
[그림-17] /a/ 목소리의 original과 comb-filter를 적용한 파형·····24
[그림-18] comb-fliter음향효과를 Max로 구현한 패치·····25
[그림-19] Max에서 pfft~를 이용한 pitch shift·····27
[그림-20] Max에서 pfft~ xover~의 패치 내부·····28
[그림-21] 그레이디언트 노이즈의 기본 원리·····30
[그림-22] TouchDesigner의 노이즈 타입·····31
[그림-23] Noise CHOP의 사용 예시를 TouchDesigner 패치로 구현·32
[그림-24] Noise CHOP에 따라 변형되는 원의 모양·····33
[그림-25] Noise TOP 패턴의 예시를 나타낸 TouchDesigner 패치···34
[그림-26] Edge TOP과 Feedback TOP을 이용한 예시·····34
[그림-27] Feedback TOP과 Over TOP으로 이미지가 겹치는 과정·· 35
[그림-28] 공연 시스템 설계도····· 36
[그림-29] OSC의 사용 예시를 Max로 구현한 패치····· 38
[그림-30] Ableton Live와 TouchDesigner의 OSC 통신····· 39
[그림-31] Arturia MiniLab MK2····· 41
[그림-32] TouchDesigner에서 MIDI 매핑을 하는 과정·····42
[그림-33] 작품 <Adsum (I am present)>의 초연 장면····· 43
[그림-34] ‘Quis vídeor?‘ 멜로디 악보····· 46
[그림-35] ‘Contínuo hīc adsum’ 멜로디 악보·····46
[그림-36] 작품 <Adsum (I am present)> 무대 구성····· 49

Ⅰ. 서 론

1. 연구 배경 및 목적

컴퓨터와 스마트폰과 같은 멀티미디어 기기들은 어느새 우리 삶에 깊숙하게 침투되어 있다. 과학 기술 발전의 고도화로 이미 세계 인구의 3분의 2가[1] 넘는 사람들이 인터넷을 이용하고 있는 지금, 경제 성장과 정보화 시대라는 말을 넘어 기술적 진보를 경험하며 자라나는 이른바 '알파 세대'[2]가 등장하게 되었다. 이러한 디지털 기술의 발전은 여러 분야에 영향을 미치며 다양한 변화 양상을 나타나게 했다. 특히 예술 분야는 그 변화의 다양성을 수용하여 디지털화된 작품으로 표현하기 시작했다. 인터랙티브 아트(interactive art)[3], 미디어 아트(media art)[4]는 보고 듣는 감상에서 그치는 것이 아니라, 시각에 입각한 공간성의 중요도와 공감각적 동시 활용의 확장으로 다중 감각을 지각하기 위한 기술 활용도가 높아졌다. 높은 미디어 활용도로 인해 새로운 문화 예술 형태들이 계속 생성되며 시대적 흐름의 변화와 더불어 매체 예술의 변화를 이어가고 있다.

이러한 매체 예술의 변화에 따른 대응력은 창작 형태의 변화와 전시 공간의 변화, 그리고 소비 형태의 변화까지 영향을 미친다. 시작점과 끝점이 명확한 형태의 영상을 시청하거나, 화이트 큐브(white cube)[5]의 공간에 설치된 미술품을 수동적으로 관람하던 기존 체계에서 벗어나 창작자는 관람자의 감각을 자극하여 참여를 유도하고, 관람자는 스스로 예술품과 소통하는 매개체

1) Internet usage worldwide - Statistics & Facts (statista.com)
2) 2010년 이후 출생한 이들을 지칭한다.
3) 관객을 참여시켜 작품 또는 예술의 목적을 달성시키는 예술의 한 형태이다.
4) 사진, 영화, 텔레비전의 발명으로 신기술을 활용하는 융합예술이다. 최근 새로운 기계와 대중매체의 등장으로 새로운 매체 기술을 사용하는 예술로 확장되고 있다.
5) 흰색 벽과 천장으로 된 전시 공간의 형태로 20세기 초 모더니즘의 등장으로 도입됐다.

가 되며, 예술품은 관람자에 의해 재창작되어 전시된다. 이는 예술 작품을 감상하는 관람자로 머무르는 것이 아니라 작품의 일부분이 되는 '경험'을 소비하는 형태로 변화시켜 문화 예술을 향유하는 근본적인 방식을 변화시키고 있다.

공연 예술 분야 역시 변화의 다양성을 수용하고 디지털 매체와 융합한 '미디어 퍼포먼스(media performance)' 혹은 '인터랙티브 퍼포먼스(interactive performance)'라고 불리며 공연 예술의 한 형식으로 주목받고 있다.[6] 제한된 무대 공간에서 디지털 매체의 적극적인 활용으로 인해 관객에게 강력한 시청각 효과를 주어 감각을 자극하고 흥미를 유도한다. 또한 기존에 구현하지 못했던 내용 표현의 한계를 극복한 확장된 표현 방법으로 관객에게 능동적 해석을 이끌어내게 된다.

본 논문은 실시간 인터랙티브 멀티미디어 퍼포먼스를 실행할 수 있도록 제작한 작품 <Adsum (I am present)>를 다루고 있다. 실시간 연주와 함께 음향효과와 시각효과를 동시에 제어하기 위해 악기의 형태로 연주하는 것보다 인성으로 표현할 수 있는 다양성을 시도하고자 했다. 여러 악기 중에서 사람의 목소리는 음악에 대한 기본적인 개념을 확립하며 퍼포먼스를 이끄는 데 효과적으로 활용될 가능성이 높다는 점에서 목소리를 중점적으로 사용할 수 있는 사운드 프로세싱 기술을 연구했다. 목소리의 특성과 한계에 연구 기술을 적용하는 방법과 효과를 언급하여 인터랙티브 멀티미디어 퍼포먼스의 새로운 방향을 제시하고자 한다.

6) 정현이, 김형기 "디지털 기술을 활용한 공연 예술의 특성 - 움직임을 이용한 인터랙티브 퍼포먼스를 중심으로-" (디자인융복합연구(구.인포디자인이슈), 2013) 제12권, 1호, 103쪽

2. 사례 연구 및 선행 연구

컴퓨터를 비롯한 다양한 디지털 멀티미디어 기기의 등장으로 멀티미디어를 사용하는 형태의 예술 작업이 대두되고 있다. 이는 장르와 분야를 가리지 않고 새로운 테크닉을 적극적으로 수용하고 활용하여 기술과 예술의 융합을 확장시키고 있다.

공연 예술 분야에서 멀티미디어의 활용은 다양한 목적을 가지고 활용되고 있다. [그림-1][7]은 Reggie Watts가 홀로 무대에서 자신만의 음향효과를 활용하여 즉흥 연주를 하고 있는 장면이다. 음악가이자 코미디언이자 배우인 Watts는 자신의 목소리만으로 비트를 만들어 노래를 하거나, 소음을 중첩시켜 쌓아 연기를 하거나 스탠드 업 코미디를 선보인다.

[그림-1] Reggie Watts의 퍼포먼스 영상캡쳐

Watts는 무대 연출의 한계를 보완하기 위해 멀티미디어 기술을 활용하는 것이 아니라 자신의 목소리만을 사용하여 즉흥적으로 극을 이끌어 가는 모습을 보여준다. 그의 퍼포먼스는 전달하고자 하는 메시지를 관객들이 분명하게 인지하도록 도와주기도 하고, 무거운 토론 주제에 익살스러운 음향효과를 통

7) https://youtu.be/UXyHf_SpUUI?si=uFAOyL-YoLK6xp7I

해 해학적인 표현으로 희화한다. 즉흥성에 맞게 모든 것을 그가 직접 제어하며 연주하기 때문에 퍼포먼스를 극대화해주는 효과를 가져온다.

[그림-2][8]는 Beardyman이 홀로 무대에서 음향효과와 시각효과를 실시간 컨트롤하며 퍼포먼스를 선보이고 있는 장면이다. Beardyman은 음악가이자 코미디언으로 그의 모든 음악은 대부분 그의 목소리만을 사용하여 소리를 생성하고 합성하는 방식으로 음향효과를 추가하여 표현된다.

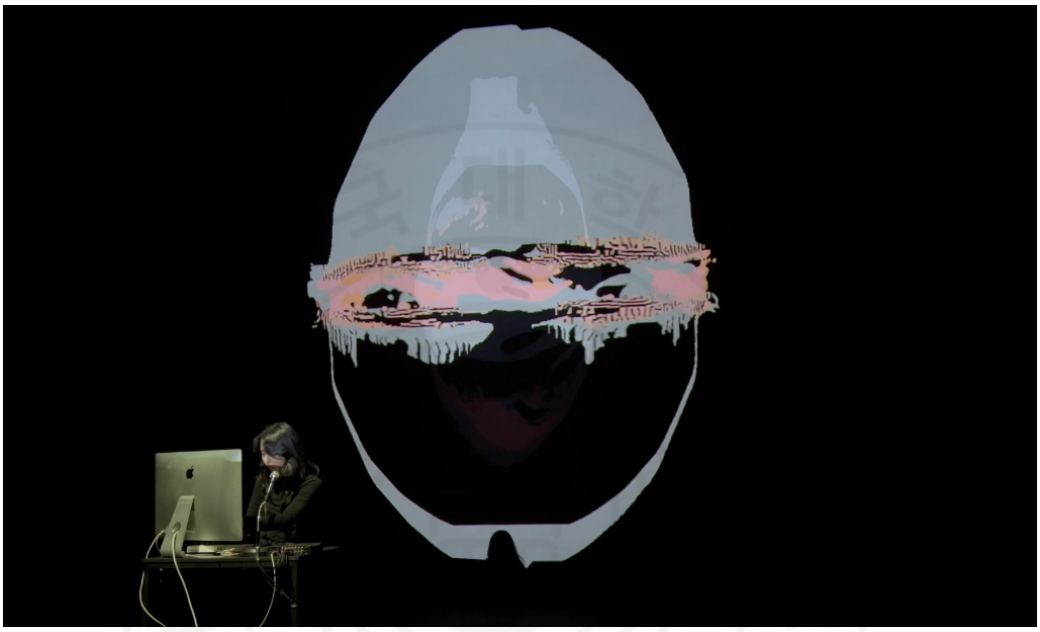
[그림-2] Beardyman의 퍼포먼스 영상캡쳐

Beardyman은 자신이 원하는 쇼를 구현해 내기 위해 스스로 소프트웨어를 개발하고 구성하여 자기만의 라이브 세트를 완성하는 장점을 가지고 있다. 끊임없이 새로운 음악 기술 탐구에 관심을 두고 있으며 청각적 표현과 시각적 표현을 혼합하여 반응을 극대화하기 위해 노력하고 있다.

이와 같이 멀티미디어 기술은 공연 예술 분야에서 적극적으로 활용되고 있을 뿐만 아니라 1인 연주자의 경우 본인의 의도와 목적에 따라 그 효과를 극대화하기 위해 사용된다. 사례 연구를 바탕으로 작품 제작, 연주 그리고 인터랙티브 멀티미디어 퍼포먼스의 가능성을 탐구해 볼 수 있다. 대부분의 소리를 본인의 목소리로 이용한 것과 그 기술을 참고하여 새로운 작품을 창작하고 나아가 퍼포먼스를 진행할 수 있도록 선행 연구를 진행했다.

8) https://youtu.be/3efvGUxPQhQ?si=pcxfvDfAEq8NdwIs

본 논문을 위한 선행 연구로 2023년 11월 목소리를 실시간으로 사운드 프로세싱 한 퍼포먼스 작품 <상념(想念)>을 제작했다. [그림-3][9]은 제작한 작품을 초연하는 영상의 한 장면이다. 목소리를 사용하여 소리를 만들고 이를 무대에서 실시간으로 연주하며 음향효과와 영상효과를 제어하고자 했다.



[그림-3] 작품 <상념(想念)>의 영상 장면

목소리만을 사용해 음악을 제작하는 것은 한계가 존재했다. 목소리는 변화되는 것에 있어 다양한 변화를 야기할 수 없다. 따라서 다양한 음색의 샘플을 사용하여 테이프음악(tape-music)[10]으로 사용했고 무대에서는 목소리로 멜로디 음을 부르며 음향효과와 시각효과를 제어했다.

의도했던 바와 달리 오직 목소리를 사용한 음악적인 구성을 완성하지 못한 것에 부족함이 존재했고 이를 보완할 필요가 있었다. 사람의 목소리를 악기와 같은 소리로 낼 수 있는 방법을 고안했고 소리를 변형하는 것이 아니라

9) https://youtu.be/BHq7I-BK2aw?si=c2ovzlygpzrRwrhX
10) 녹음된 음악을 소스로 활용하여 음악을 연주하거나 작곡하는 유형이다.

신호를 변형하는 것으로 작품 연구를 진행했다. 즉, 마이크를 통해 수음된 사람의 목소리를 컴퓨터로 전달하는 과정에서 오디오 신호 정보를 가상악기로 연주하도록 만드는 것이다.

선행 연구의 결과에 따라 본 논문의 작품 <Adsum (I am present)>는 실시간으로 목소리의 오디오 신호를 MIDI(미디)[11] 신호로 전달하여 연주하는 가상악기 소리와 멜로디 노래로 연주하는 목소리에 음향효과를 더한 음악과 영상의 인터랙티브 퍼포먼스로 제작했다.

11) MIDI는 Musical Instrument Digital Interface의 약자로 신디사이저와 같은 악기들을 컴퓨터와 효과적으로 연결하기 위해 데이터를 처리하고 통신할 수 있는 일종의 규약.

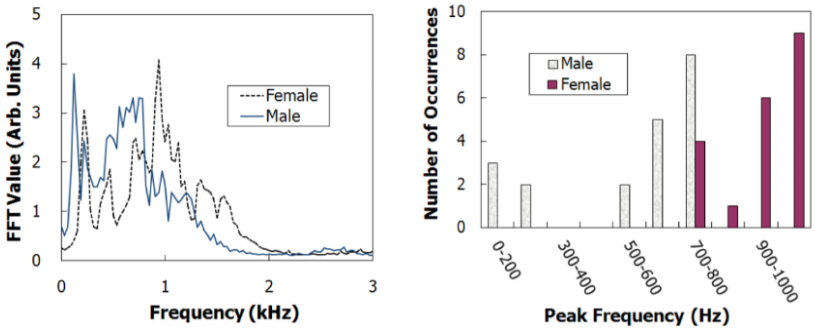
Ⅱ. 기술 연구

1. 목소리의 특성

목소리는 인간의 폐, 성대, 구강과 입술, 그리고 목을 통해 만들어 내는 소리로 목소리를 발현하기 위해서는 호흡(expiration), 발성(phonation), 공명(resound), 조음(articulation)의 네 가지 기본 요소를 필요로 한다. 숨을 들이쉬고 내쉬는 작용으로 시작된 호흡이 성대를 통과해 성대 근육을 서로 부딪히게 하여 만들어 낸 진동으로 소리를 생성하는 것이 발성이다. 이를 우리 몸의 공명 기관에서 소리를 증폭시키고 조음기관을 통해 언어음으로 바꾸어 명확하고 정확한 소리를 내게 된다.[12)]

앞서 언급한 목소리를 발현하기 위한 네 가지 기본 요소는 방법과 구조에 따라 다양한 소리를 만들어낼 수 있는데, 이는 사람의 생김새나 구강구조에 따라 구별점을 갖게 된다. 각기 다른 사람의 목소리는 기본 주파수[13)] 및 포만트 주파수[14)] 등을 통해 비교와 분석이 가능하다.[15)] [그림-4][16)]는 남녀 각 20명으로부터 발성된 모음 아(/a/)의 주파수를 분석하여 성별에 따른 평균값의 주파수 스펙트럼과 포만트 주파수 분포를 표로 나타낸 것이다. 예상 가능한 결과지만, 여성이 남성보다 고음 주파수 영역에 편향되어 있는 것을 알 수 있다. 따라서 목소리는 사람의 성별과 나이 등 신체적 구조가 상이한 개인간의 특성을 반영하여 사람을 구분할 수 있게 된다.

12) 배한성, 서혜정, 문선희, 김희선, 조예신 & 박형욱, <성우> (커뮤니케이션북스) 9쪽.
13) 전자기파가 1초 동안에 진동하는 횟수로 사람의 청각기관을 자극하는 주파수 대역을 갖는 파동을 소리라고 한다.
14) 사람의 목소리나 악기 소리에서 특정 주파수 영역이 강조되는 현상.
15) 김봉영, 이은영 & 배명진, “목소리 특성 구분에 대한 사운드 A칼라에 관한 연구” (예술인문사회 융합 멀티미디어 논문지, 2018), 제8호, 13-21쪽.
16) 이창영, “남성과 여성의 음성 특징 비교 및 성별 음성인식에 의한 인식 성능의 향상” (한국전자통신학회 논문지, 2010), 제5권, 6호, 571쪽.



모음/a/에 대한 남성과 여성의 주파수 스펙트럼 평균

모음/a/에 대한 남성과 여성의 포먼트 주파수 분포

[그림-4] 성별에 따른 모음/a/의 주파수 스펙트럼과 포먼트 분포 표

목소리를 이용하여 노래를 부르게 될 때 성별에 따라, 가창자의 특성에 따라 음역대를 나누게 된다. 남성은 테너, 바리톤, 베이스로 구성되어 E2부터 G4까지로 80Hz~400Hz의 음을 부르며 약 8000Hz까지의 배음 주파수[17]를 가지고 있다. 여성은 소프라노, 메조소프라노, 알토로 구성되어 F3부터 C6까지로 200Hz~1000Hz의 음을 부르며 약 17000Hz까지의 배음 주파수를 가지고 있다.[18]

본 논문은 여성의 목소리를 다루는 경우로 여성 음성의 주파수 음역대를 인지하는 것이 중요하다. 사운드 합성 시 중요한 요소인 주파수를 이용할 때 여성의 음정 주파수와 배음 주파수를 이용할 수 있을 뿐만 아니라, 여성이 부를 수 없는 남성의 낮은 음역대를 추가할 수 있기 때문이다. 사람의 목소리는 독창 시 음역대가 좁고, 길이가 짧으며 악기와 다르게 동시 발현이 불가능하다. 또한 동일한 음정을 연주하는 데 있어 발음의 형태에 따라 들리는 것이 상이하다. 해당 한계점들의 보완이 가능한 음향효과와 여성 음역대를 고려하여 사운드 프로세싱을 적용도록 연구했다.

17) 기본음과 기본음의 정수 배가 결합된 주파수.
18) TUNA TORUN, “An Approach On Determining Optimum Acoustic Conditions For Turkish Classical Music”, 2019, p.65

2. 사운드 시스템 연구

1) 사운드 시스템

본 논문에서 다루고 있는 작품 <Adsum (I am present)>는 목소리를 통해 연주되는 사운드를 실시간으로 프로세싱 하여 진행되는 멀티미디어 퍼포먼스이다. 목소리에 적용되는 모든 음향효과는 Max[19]를 활용한 사운드 프로세싱 기술을 기반으로 Ableton Live[20]에서 Max for Live[21]를 통해 구현됐다.

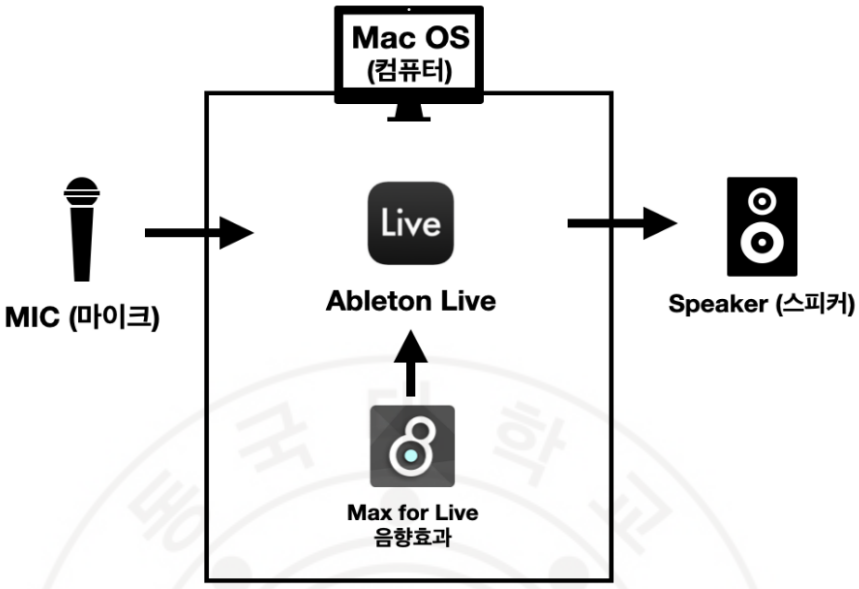
[그림-5]는 본 작품의 실시간 프로세싱을 위한 사운드 시스템 설계도이다. 목소리의 오디오 신호는 마이크를 사용하여 오디오 인터페이스를 통해 컴퓨터로 입력된다. 컴퓨터로 입력받은 목소리의 오디오 신호는 Ableton Live로 전달되어 Max for Live로 제작된 다양한 음향효과와 합쳐진다. Max for Live로 제작된 각 음향효과는 작품의 구성에 따라 Ableton Live의 여러 오디오 트랙으로 분류되어 적용되며, 각 음향효과는 컨트롤러에 매핑되어 사용된다. 목소리의 오리지널 사운드와 음향효과가 적용된 프로세싱 사운드는 최종적으로 스피커를 통해 출력된다.

---

19) Cycling '74에서 개발한 음악과 멀티미디어를 위한 통합 개발 환경 소프트웨어이다. API(Application Programming Interface)를 통해 컴퓨터와 연결하거나 외부 소프트웨어와 연결 혹은 서비스를 제공한다.

20) 독일의 Ableton에서 개발한 음악 제작과 공연을 위한 디지털 오디오 워크 스테이션(Digital Audio Workstation, DAW)이다.

21) Cycling '74의 MAX/MSP를 기반으로 Ableton Live에서 사용자가 자체적으로 음악적 도구(musical tools)를 제작할 수 있는 소프트웨어 플랫폼이다. 사용자가 원하는 대로 악기, MIDI 효과, 오디오 효과를 만들거나 Ableton Live를 제어할 수 있는 기능을 제공한다.



[그림-5] 사운드 시스템 설계도

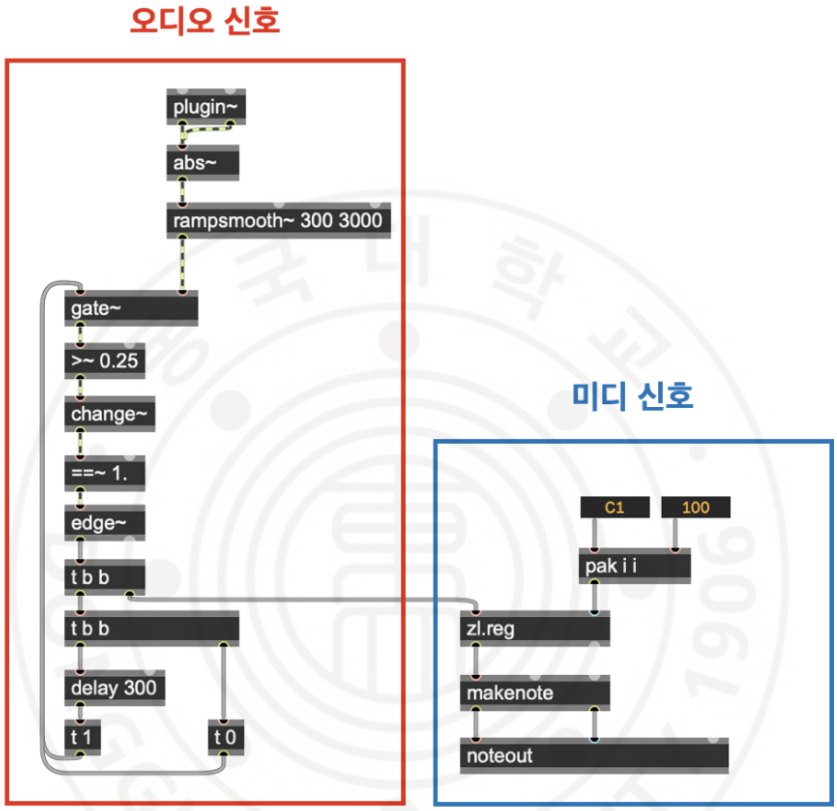
2) 사운드 디자인 연구

① 오디오 신호 변환

마이크를 통해 컴퓨터로 전송되는 목소리는 오디오 신호로 입력된다. 이를 목소리가 아닌 가상의 악기 소리로 출력하고자 할 때 입력된 오디오 신호를 MIDI 신호로 변환하는 과정을 필요로 한다. [그림-6]은 makenote오브젝트를 이용해 오디오 신호를 MIDI 데이터[22]를 담고 있는 MIDI 신호로 변환하는 것을 Max로 구현한 것이다. makenote는 소리를 내고자 하는 MIDI 악기로 전송할 일종의 명령어를 담은 온 오프(on-off) 식의 MIDI 노트를 생성하는 것으로 note number, velocity, duration 정보를 담고 있다. note number는 음정으로 C1은 24에 해당한다. 1씩 늘어날 경우 반음이 높아지고 1씩 줄

22) 소프트웨어 악기와 하드웨어 장치를 연결하기 위한 정보를 담고 있는 디지털 숫자 스트림

어들 경우 반음이 낮아진다. velocity는 음의 세기이고 duration은 음의 지속 시간이다. [그림-6]에 대한 설명은 오디오 신호와 MIDI 신호로 나누어 순차적으로 설명한다.



[그림-6] makenote오브젝트를 이용한 MIDI 신호로 변환

오디오 신호에서 plugin~오브젝트[23]로 입력된 오디오 신호를 절댓값(absolute value)[24]을 받는 abs~오브젝트[25]를 지나 상향 기울기는 300 sample rate[26]로, 하향 기울기는 3000 sample rate으로 설정하는

23) Max for Live의 오디오 입력을 나타낸다.
24) 실수에서 양 또는 음의 부호를 떼어 버린 수
25) absolute의 줄임말로 입력된 신호의 절댓값만을 출력하는 오브젝트이다.
26) 주어진 단위 시간 동안 일정한 간격으로 데이터를 수집하는 것으로 1초 동안 진동한 횟수를 나타내는 주파수로 예를 들면 4Hz는 4sample/sec이다.

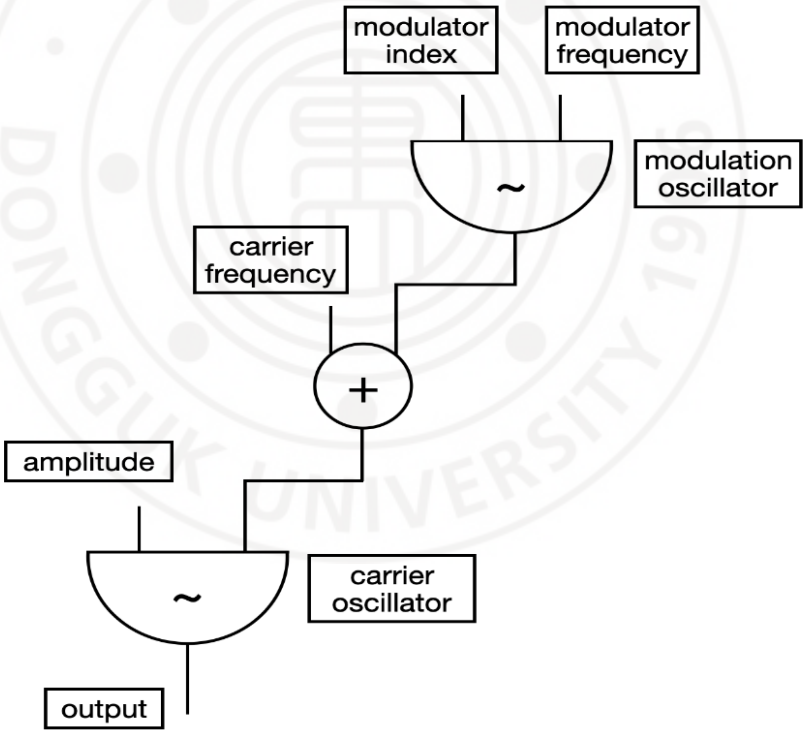
rampsmooth~오브젝트[27]로 입력된다. 오디오 신호의 값이 0.25보다 큰 값을 change~오브젝트[28]로 입력할 경우 -1, 0 또는 1의 값을 출력하게 된다. 만약 1의 값을 출력하여 edge~오브젝트[29]로 입력될 경우 해당 신호를 MIDI 신호로 전달하는 일과 그다음 명령을 수행하도록 지시하는 trigger오브젝트[30](t b b를 뜻함)로 연결된다. 다음 단계로는 명령을 지시하는 새로운 trigger로 연결되어 plugin~으로 입력되던 오디오 신호의 경로를 차단하고(t 0)[31] 300ms(millisecond)[32] 뒤에 다시 통과될 수 있도록(t 1) gate~오브젝트[33]로 연결된다.

오디오 신호로부터 전달받은 입력 신호는 makenote로 전달되기 전 정보를 모아 한 번에 처리되는 zl.reg오브젝트[34]로 연결된다. 오디오 신호의 길이는 duration에 해당하기에 음정과 음의 세기를 필요한 숫자 정보로 설정하여 pak오브젝트[35]로 모아 전달한다. makenote를 통해 전달받은 MIDI 노트 정보는 noteout오브젝트를 통해 소리를 내고자 하는 악기로 출력된다.

27) 입력되는 신호의 샘플을 평활화하는 오브젝트로 변경되는 수신 값에 도달하기 위해 선형(linear)의 기울기를 늘린다.
28) 입력되는 신호 값의 증가, 감소 또는 변화 없음을 모니터링하는 오브젝트로 입력된 값이 이전보다 크면 1을, 같으면 0을, 작으면 -1을 출력한다.
29) 입력되는 신호가 전환되는 것을 감지하는 오브젝트.
30) t는 trigger로 수신된 정보를 값을 설정한 구성 방식에 맞게 출력하는 오브젝트이다. b는 bang의 때리다는 뜻으로 다음 명령을 지시하게 한다.
31) 0은 비활성을 나타내고 1은 활성을 나타낸다.
32) 시간의 단위로 1ms(millisecond)는 1000분의 1초이다.
33) 입력되는 신호의 경로를 정해주는 오브젝트로 오브젝트에 적힌 숫자에 따라 outlet의 개수가 정해진다. 오른쪽 inlet으로 입력된 신호를 왼쪽 inlet의 숫자 정보대로 outlet이 지정되며 0일 경우 모든 outlet이 닫힌다.
34) 목록을 저장하고 출력하는 오브젝트. 오른쪽 inlet에서 받은 목록을 저장하고 bang을 받을 시 왼쪽 outlet으로 출력한다. 왼쪽 inlet에서 받은 목록은 즉시 왼쪽 outlet으로 출력한다.
35) 수신된 항목을 출력할 목록을 결합하는 오브젝트로 inlet에서 인수가 수신될 때마다 전체 목록을 출력한다.

② FM synthesis

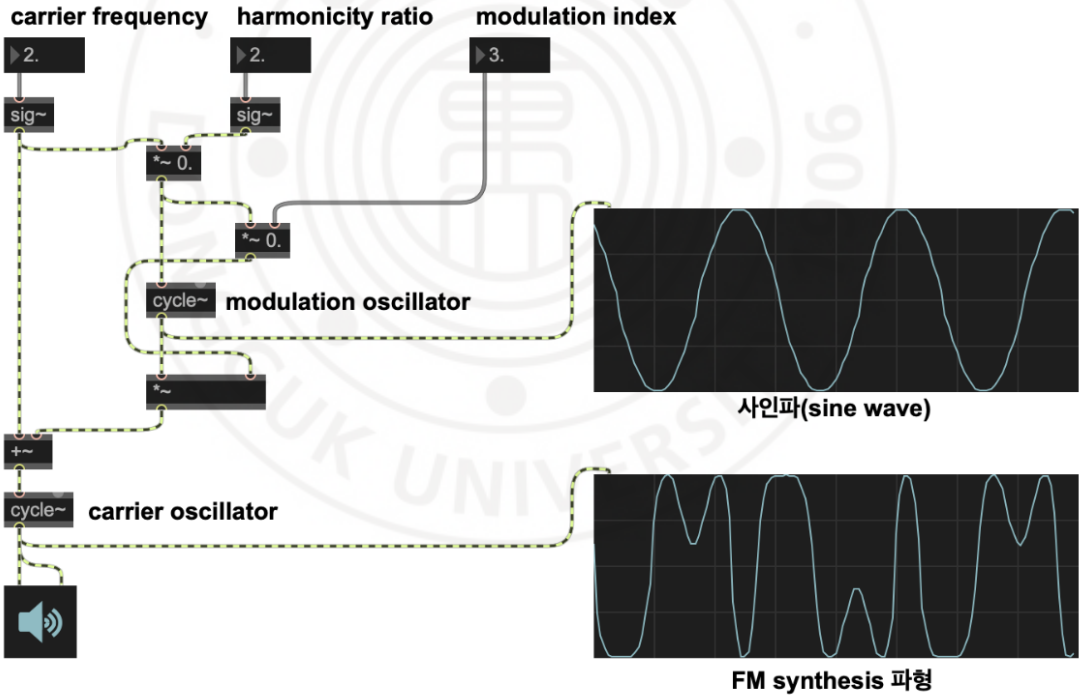
음악에서 modulation은 전조 혹은 리듬의 변화를 뜻하지만 전자음악에서는 변조를 뜻하는 것으로 FM(Frequency Modulation) synthesis는 주파수를 변조하여 사운드를 합성하는 것을 말한다. carrier oscillator의 주파수(carrier frequency)를 변조시키기 위한 modulation oscillator의 주파수(modulator frequency)를 더하는 방식으로 J. M. Chowning[36]이 개발했다. [그림-7]은 FM synthesis의 기본 구조를 그림으로 나타낸 것이다. 다른 합성 방식에 비하여 두 개의 oscillator를 이용하는 간단한 구조로 풍부한 배음을 생성하여 다양한 사운드 합성이 가능한 것이 특징이다.



[그림-7] FM synthesis의 기본 구조

36) 미국의 음악가이자 교수로 스탠퍼드 대학교 재직 당시 '음악과 음향 컴퓨터 연구 센터(CCRMA)'를 설립자이다. FM 합성 알고리즘과 디지털 입체 음향화를 개발했다.

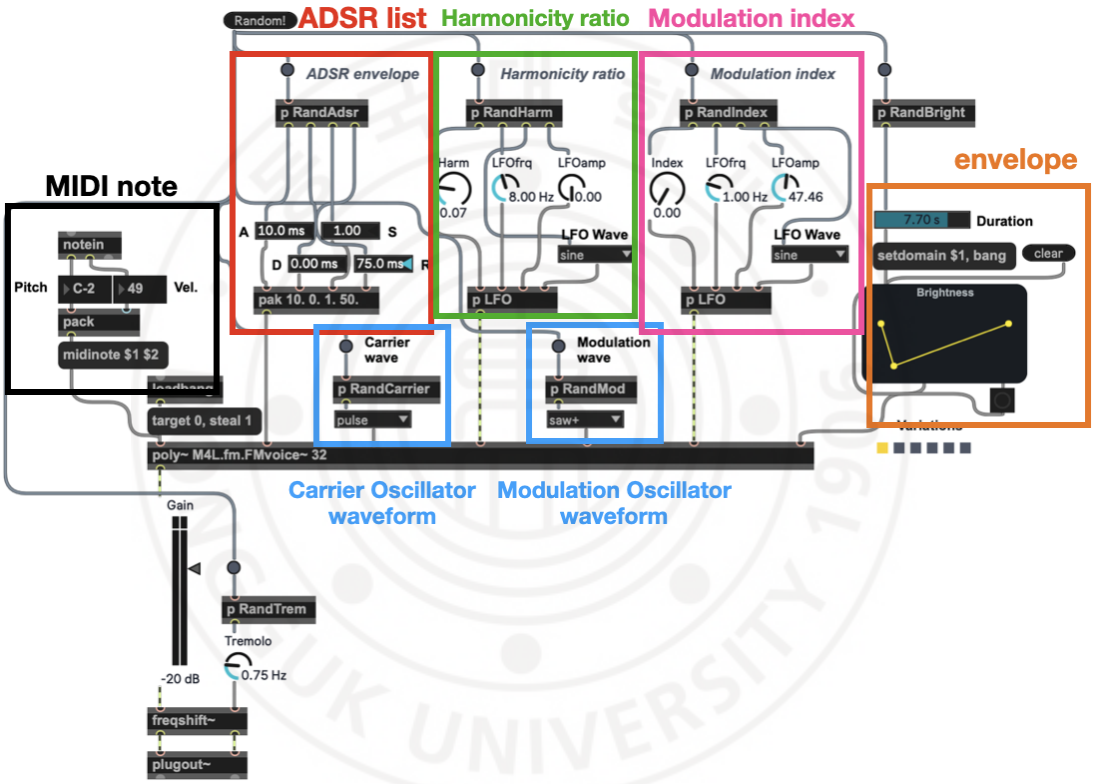
[그림-8]은 Max로 앞서 설명한 FM synthesis의 기본 구조를 구현한 것이다. 기본 사인파(sine wave)의 주파수에 modulator frequency, modulator depth를 조절하면 그림과 같은 새로운 사이드밴드(sideband)[37]를 만들어 기본 사인파보다 많은 주파수를 출력한다. 사이드밴드의 주파수는 carrier frequency와 modulator frequency에 의해 결정되며 두 주파수의 비율 값을 harmonicity ratio라고 한다. 또한 사이드밴드의 강도는 carrier frequency와 modulation depth의 값에 의해 결정되며 강도의 정도를 조절하는 값을 modulation index라고 한다. 값을 조절하는 파라미터의 비율에 따라 음색에 영향을 주며 modulation index 값이 클수록 강도 높은 사이드밴드를 형성하여 음색의 밝기(brightness)에 영향을 준다.



[그림-8] FM synthesis를 Max로 구현한 패치

37) carrier frequency보다 높거나 낮은 주파수 대역.

Max에서 FM synthesis를 구현하는 방법으로 Cycling '74에서 제공하는 다양한 서브패치(subpatch)[38]를 이용할 수 있다. 그중에 Ableton Live에서 구동하기 편리하도록 제공되는 poly~ M4L.fm.FMvoice~를 응용하는 방식을 사용하여 새로운 사운드를 만들어 하나의 악기 소리 제작이 가능하다. [그림-9]는 poly~ M4L.fm.FMvoice~를 이용한 Max패치이다.



[그림-9] poly~ M4L.fm.FMvoice~를 이용한 Max패치

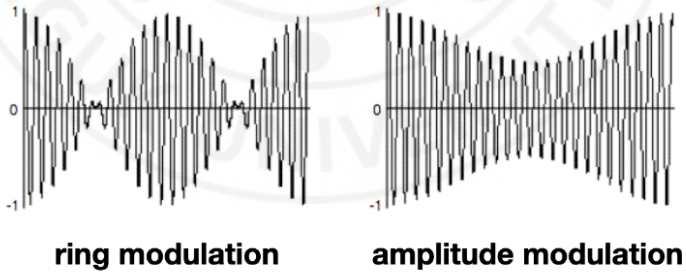
원하는 MIDI note 음정을 지정한 carrier oscillator와 modulation oscillator의 파형(waveform)[39]과 envelope[40]를 설정하고 ADSR[41], Harmonicity ratio,

38) Max에서 패치 작업을 정리하거나 오브젝트들을 그룹화하는 것이다.
39) 시간에 따른 파동 곡선의 모양. sine wave, triangle wave, sawtooth wave, square wave(pulse) 등이 있다.
40) 시간에 따른 소리 변화의 양상.

modulation index의 각 파라미터를 조절하면 음색 변화의 차이를 느낄 수 있어 원하는 사운드를 제작한다. 파형이 triangle wave와 가깝고 index 값이 높을수록 날카로운 음색을 나타내고, 같은 파라미터의 값을 가지고 있더라도 ADSR과 envelope의 차이에 따라 음색이 달라진다. 추가적으로 LFO[42]를 사용하여 기계적인 떨림을 주는 효과를 통해 색다른 음색의 변화를 적용할 수 있다.

③ RM

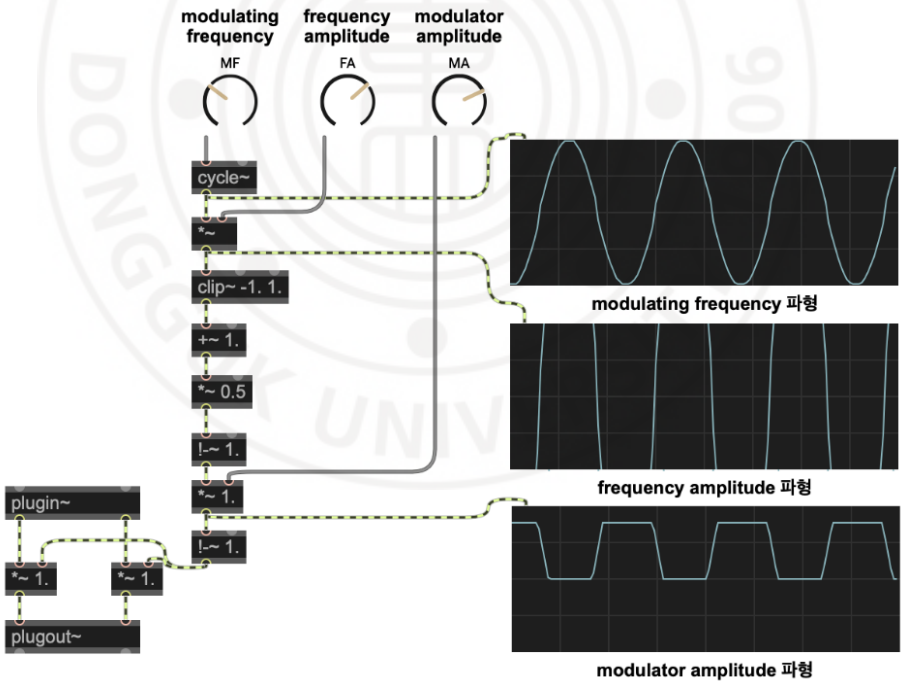
FM synthesis는 주파수를 변조하는 방식이라면, 진폭을 변조하는 방식을 AM(Amplitude Modulation)이라 한다. carrier oscillator에서 발생하는 파형의 진폭에 modulation oscillator의 파형 진폭을 합성하는 것을 말한다. RM(Ring Modulation)은 AM과 마찬가지로 진폭을 변조하는 방식이지만 AM은 단극(unipolar) 변조 값 0에서 1을 사용하는 반면 RM은 양극(bipolar) 변조 값 -1에서 1을 사용하기 때문에 변조 진폭의 차이를 나타낸다. [그림-10][43]은 RM과 AM의 합성에 따른 변조 진폭 차이를 나타낸 파형이다.



[그림-10] RM과 AM의 변조 진폭 차이

41) envelope의 형태를 표현하는 것으로 Attack time, Decay Time, Sustain Level, Release Time을 나타낸다.
42) Low Frequency Oscillator의 약자로 일반적으로 20Hz 미만의 저주파 발진기이다.
43) https://docs.cycling74.com/legacy/max8/tutorials/06_synthesischapter03?q=am (MSP Synthesis Tutorial 3: Using Amplitude Modulation)

[그림-10]과 같이 RM의 경우 변조되는 진폭의 차이가 크기 때문에 비정형적이고 독특한 음색을 나타낸다. 만약 modulation frequency 값이 20Hz 이하일 경우 사운드 합성을 위한 변조를 나타내기보다 트레몰로(tremolo)[44] 효과가 나타나게 된다. [그림-11]은 RM을 이용한 트레몰로 음향효과를 Max로 구현한 것이다. plugin~으로 입력되는 오디오에 변조할 modulator의 주파수를 modulating frequency라 하고, 해당 주파수의 진폭의 정도를 frequency amplitude라고 한다. original sound를 완전히 감쇄하지 않기 위해 연산을 거친 후 modulator amplitude를 이용해 한 번 더 진폭을 조절한다. 주파수와 진폭 값은 트레몰로 효과를 위해 12Hz를 넘지 않게 설정했고 modulator amplitude는 1을 넘지 않게 설정했다. 각 파라미터에 따른 파형 변조는 아래 그림과 같이 나타난다.

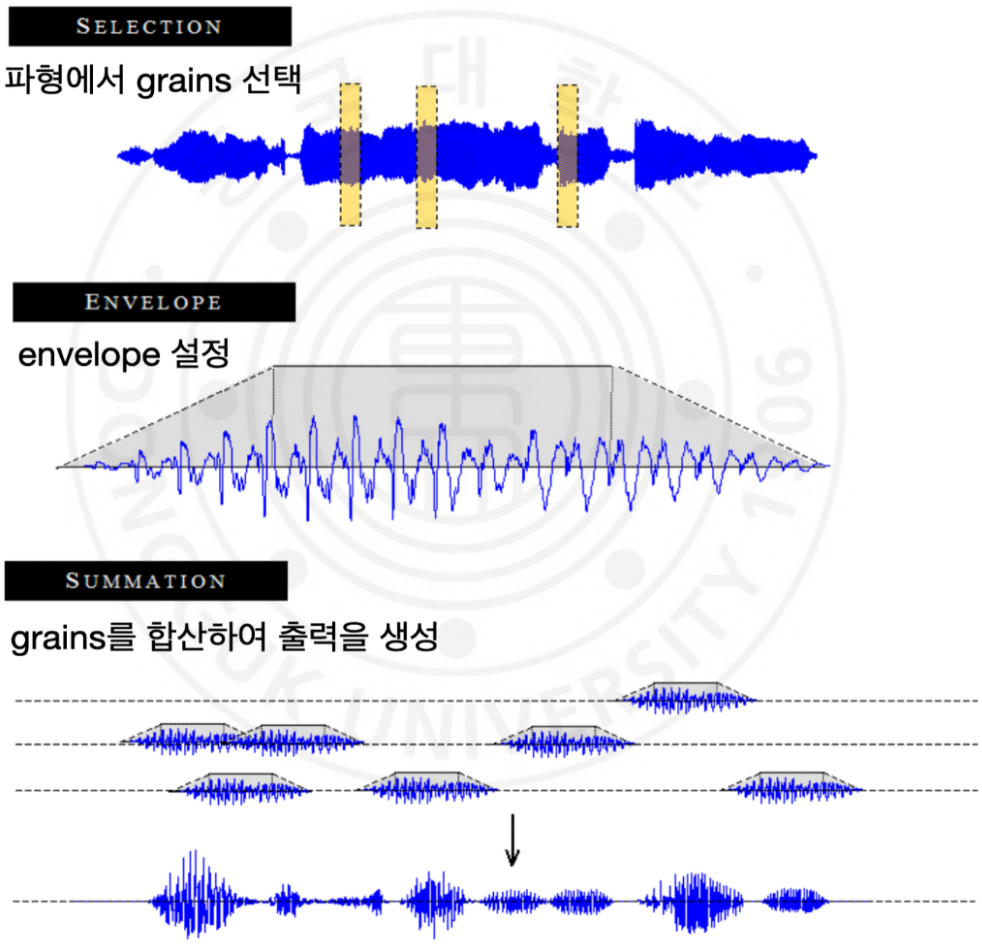


[그림-11] RM을 이용한 트레몰로 음향효과를 구현한 Max 패치

44) 한두 음을 빠르게 반복하거나 음의 볼륨을 조절하여 떨리는 소리를 내는 음악 기법.

④ granular synthesis

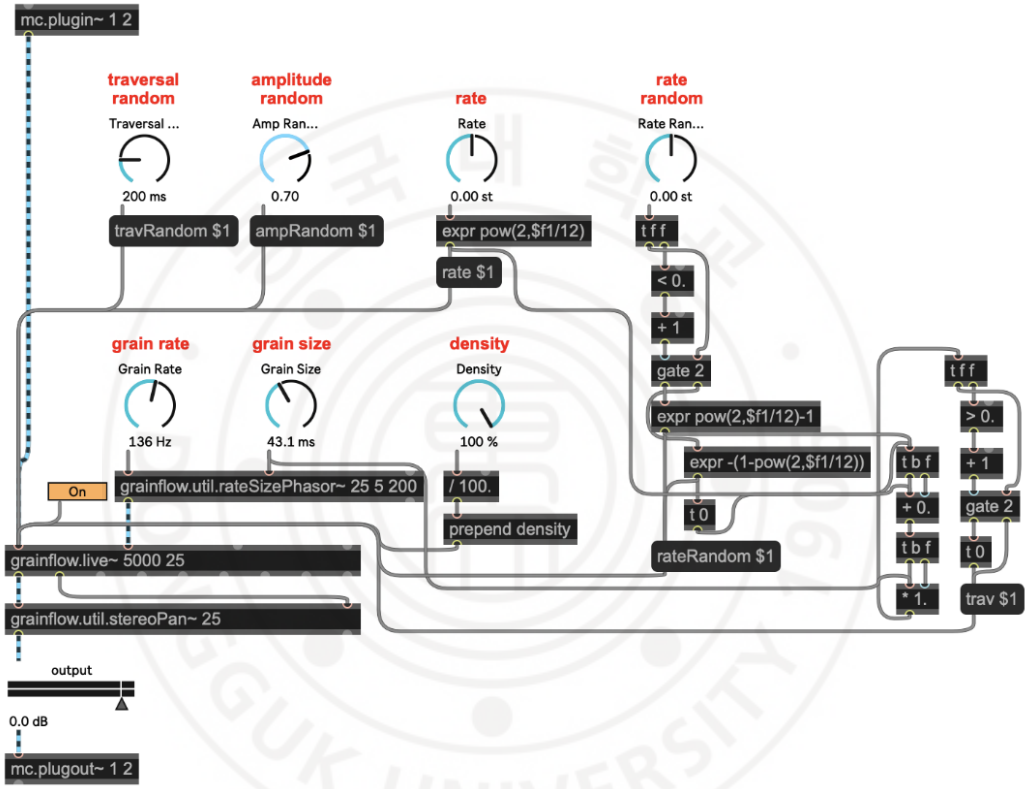
granular synthesis는 오디오를 작은 조각으로 나누어 재합성하는 방식을 의미한다. 나누게 되는 샘플 소스의 길이가 약 1-200ms의 매우 짧은 조각으로 나누는 것이 곡물(grain) 혹은 가느다란 낟알(granule) 같다 하여 붙여진 이름이다. [그림-12][45]는 granular synthesis의 과정을 나타낸 것이다.



[그림-12] granular synthesis의 과정

45) Williamson, John & Murray-Smith, Roderick. "Audio Feedback for Gesture Recognition", (Department of Computing Science, University of Glasgow, 2003) p.7.

Max에서 granular synthesis를 구현하는 경우 munger~오브젝트를 사용하지만, 본 논문의 경우 Ableton Live에 사용하기 적합한 grainflow.live~ 오브젝트를 사용했다. [그림-13]은 grainflow.live~를 응용한 granular synthesis를 구현한 Max 패치이다. 파라미터가 직관적으로 드러나있지는 않지만 지정하여 설정할 수 있다.



[그림-13] grainflow.live~로 granular synthesis를 구현한 Max 패치

grainflow.live~에 입력된 5000은 버퍼 사이즈[46]를 의미하고 25는 grain의 최대 수이다. 사용된 파라미터는 7개로 grain size, grain rate, density, rate, rate random, traversal random, amp random이다. grain size와 traversal random은 ms 단위 값을 적용하며 grain rate의 단위는 Hz이다.

46) 버퍼 사이즈(buffer size)는 컴퓨터로 들어오는 오디오 신호를 처리하는 데 걸리는 시간에 해당하는 샘플 수를 의미한다.

rate와 rate random은 재생 속도를 제어해 음정이 변동되어 semitone[47]단위로 적용한다. density는 확률로 퍼센트(%) 단위로 적용하며 amp random은 음량의 최댓값 1을 넘지 못하기 때문에 0에서 0.95의 값으로 제한하여 적용한다. 7개의 파라미터를 조절하는 방식에 따라 original sound와 전혀 다른 소리를 만들어낼 수 있다. 각 파라미터의 이름에 따른 기능은 <표-1>을 통해 정리했다.

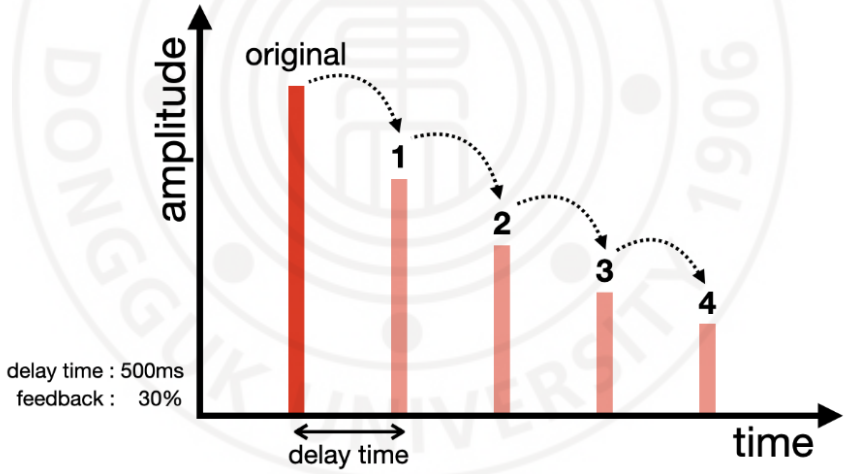
<표-1> grainflow.live~오브젝트의 파라미터

파라미터 이름	파라미터 기능
grain size	샘플의 크기
grain rate	샘플의 속도
density	샘플이 재생될 확률
rate	재생 속도 제어
rate random	재생 속도 변화 값
traversal random	샘플 딜레이 변화 값
amp random	음량 변화 값

47) 47) semitone은 반음을 뜻한다.

⑤ stereo delay

delay는 오디오 신호를 지연시켜 original sound와 출력하는 방식으로 에코(echo)[48]와 같은 소리의 반복을 만들어내는 음향효과이다. 이는 delay의 가장 큰 특징이며, 시간의 경과에 따라 반복되는 소리의 볼륨이 계단식으로 감소한다 하여 시간 기반(time-based)[49] 음향효과로 분류된다. [그림-14]는 delay의 기본 원리를 도식화한 것으로 피드백(feedback)은 반복되는 소리의 양을 나타내고 delay time은 반복되는 소리의 간격을 나타낸다. delay time이 500ms, 피드백은 30%로 0.5초마다 첫 번째 반복된 소리는 original보다 30% 감소하고, 두 번째 반복된 소리는 첫 번째 반복된 소리보다 30% 감소하는 등의 형태로 소리가 반복된다는 것이다.



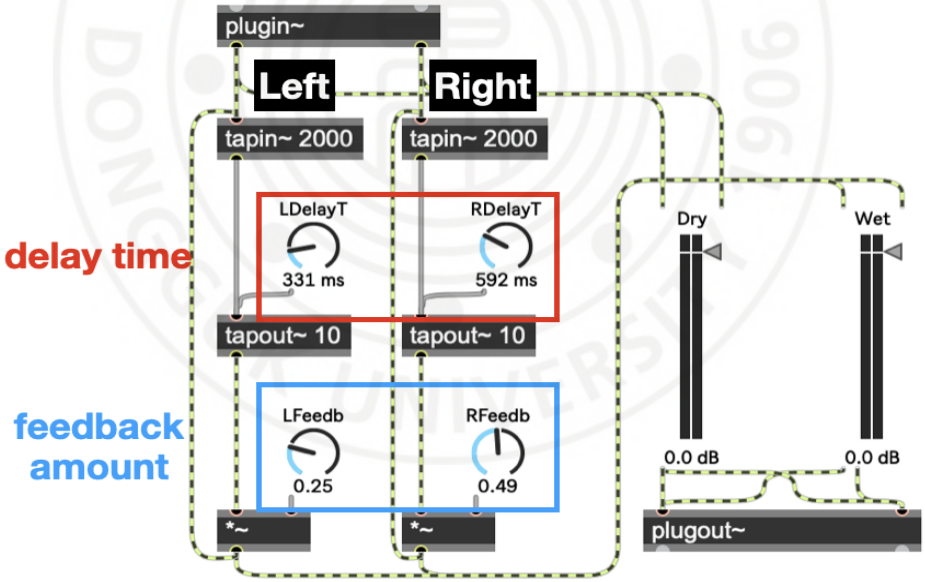
[그림-14] delay의 기본 원리

delay는 오픈 릴식의(reel-to-reel)[50] 테이프 녹음 시스템을 사용하던 때 테이프 루프(tape loop)[51]로 인해 처음 기술이 발전하게 됐다. 이후 디지털 프

48) 울림이나 반향을 뜻하는 것으로 소리가 벽과 같은 표면에 반사된 후 들리게 되는 것이다.
49) 오디오 신호의 타이밍을 다루어 소리의 음색, 공간감, 깊이감을 만들어내는 효과이다.
50) 정식 명칭은 'reel to reel tape recorder'로 릴(reel)에 감긴 마그네틱테이프에 강한 자장을 걸어 소리를 기록, 재생하는 장치이다.

로세싱이 보급되면서, 아날로그가 디지털로 변환되는 컨버터의 버퍼에 기록된 디지털 오디오 신호를 사용자의 설정에 따라 재생하는 방식으로 개발됐다. delay 효과를 적용한 오디오 신호인 wet과 적용되지 않은 오디오 신호인 dry를 혼합하여 출력이 가능하도록 한다.

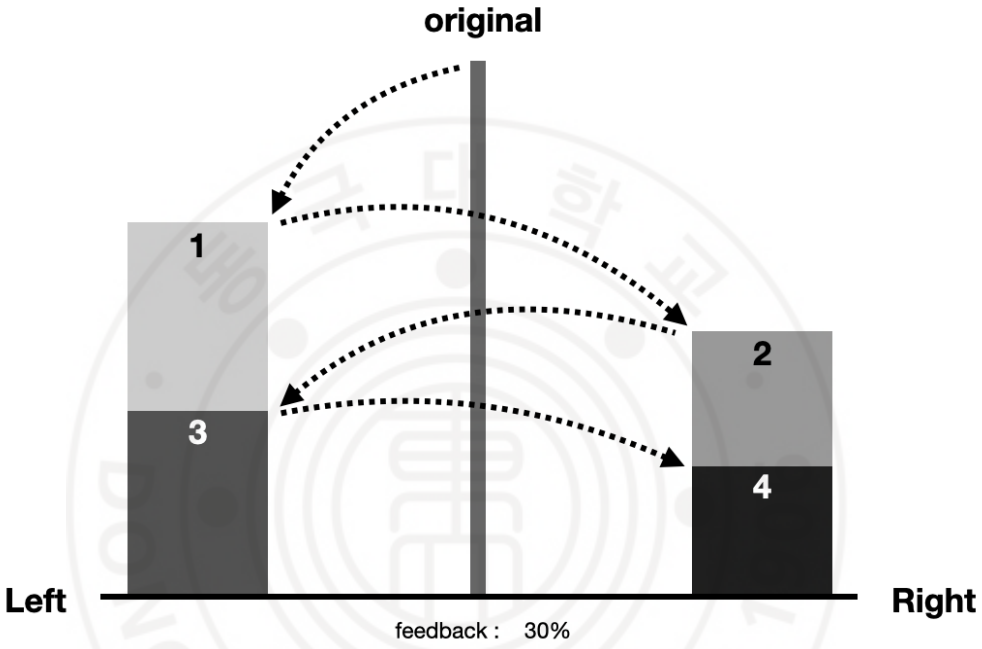
해당 원리를 바탕으로 Max에서 delay 효과를 구현할 수 있는 것이 tapin~오브젝트와 tapout~오브젝트이다. tapin~은 입력된 오디오 신호를 지연시키기 위해 저장하는 공간이며, tapout~은 저장된 오디오 신호를 내보내어 delay time을 조절하는 역할을 한다. tapout~에서 나온 신호를 다시 tapin~에 입력시키는 것으로 피드백을 발생시키며, 왼쪽과 오른쪽의 delay time을 다르게 설정하면 stereo delay 음향효과를 생성하게 된다. [그림-15]는 stereo delay 음향효과를 Max로 구현한 것이다.



[그림-15] stereo delay 음향효과를 구현한 Max 패치

51) 녹음된 마그네틱테이프를 자르고 붙이는 방식으로 끝을 이어 원 혹은 루프를 만들어 반복 재생하는 기법.

stereo delay에서 왼쪽과 오른쪽의 두 채널로 소리를 출력하는 것이 아니라 소리가 반복될 때마다 왼쪽 혹은 오른쪽으로 신호를 번갈아 가며 처리하는 것을 ping pong delay라고 한다. [그림-16]은 ping pong delay의 오디오 신호 흐름을 도식화한 것이다.

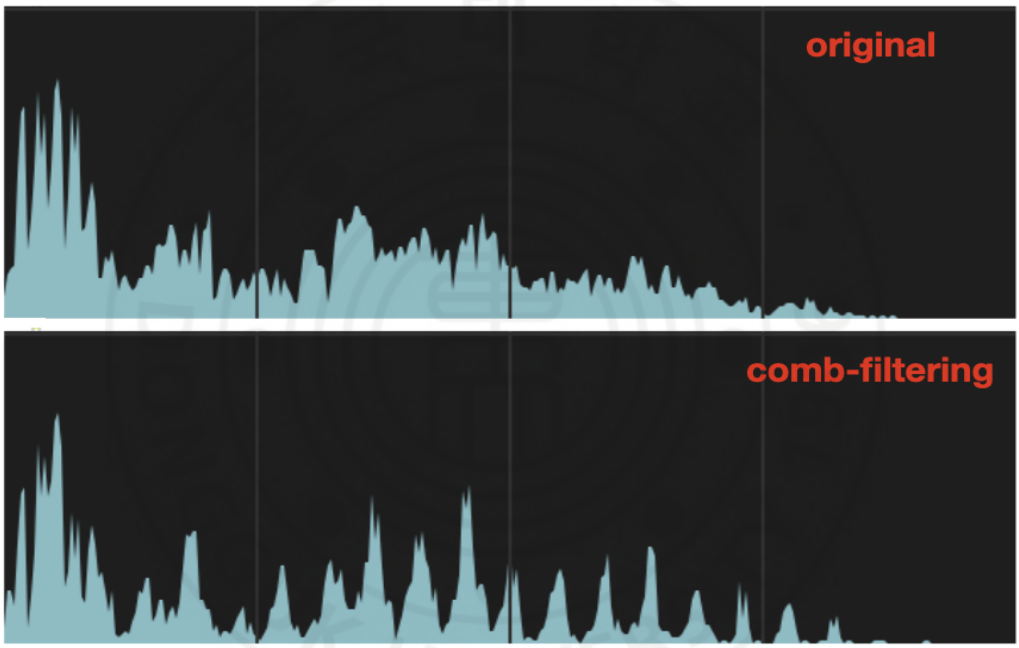


[그림-16] ping pong delay의 오디오 신호 흐름

피드백이 30%인 ping pong delay는 첫 번째 반복된 소리를 왼쪽으로 original의 30% 감소하고, 두 번째 반복된 소리는 오른쪽으로 첫 번째 반복된 소리보다 30% 감소하는 등의 형태로 반복하여 처리한다. 탁구공이 탁구대의 양쪽을 튀어 다니듯이 ping pong deay 또한 역동성을 느끼게 하여 소리가 이동하는 듯한 착각을 불러일으킨다. 이를 통해 소리의 너비와 깊이감을 부여하여 공간감을 생성하는 데 도움이 된다.

⑥ comb-filter

comb-filter는 직접음과 반사음이 서로 시간적인 차이를 두고 섞이면서 특정 주파수를 증감하는 현상을 의미한다. [그림-17]은 모음/a/를 부르는 목소리의 original 파형과 comb-filter를 적용한 파형의 모습이다. 아래 그림과 같이 주파수 스펙트럼의 모양이 빗살 모양과 유사한 형태로 나타난다 하여 comb-filter라고 불린다.



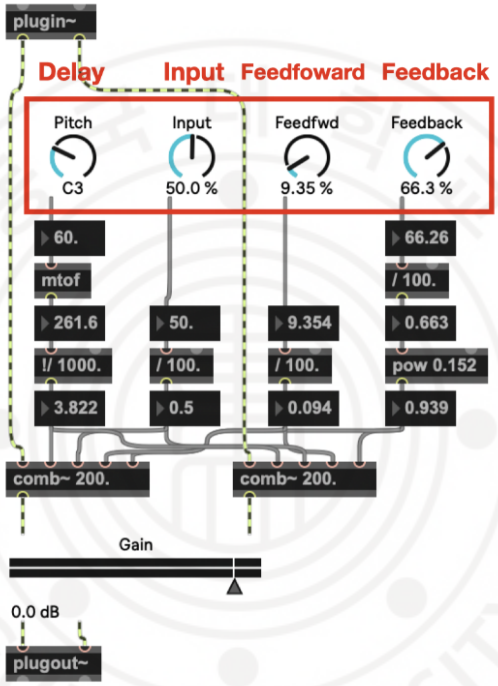
[그림-17] /a/ 목소리의 original과 comb-filter를 적용한 파형

Max에서 comb~오브젝트를 이용해 comb-filter 음향효과를 구현할 수 있으며, 해당 오브젝트는 아래의 공식을 따른다.[52] 신호가 입력되는 시점을 t로 두고 입력되는 오디오 신호를 출력 오디오 신호와 혼합하여 comb-filter 음향효과를 출력하는 y_t를 계산하는 공식이다.

52) https://docs.cycling74.com/legacy/max8/tutorials/15_delaychapter06 (MSP Delay Tutorial 6: Comb Filter)

$$y_t = ax_t + bx_{(t-d)} + cy_{(t-d)}$$

a는 입력 신호, b는 이전에 받은 입력 신호, c는 출력 신호, d는 a, b, c의 delay 시간을 지정하는 양이다. comb~는 입력 신호의 값(ax_t)을 Input, d를 Delay, $bx_{(t-d)}$를 Feedfoward, $cy_{(t-d)}$를 Feedback으로 하여 조절한다.



[그림-18] comb-filter음향효과를 Max로 구현한 패치

[그림-18]은 comb-filter 음향효과를 Max로 구현한 패치이다. Delay는 ms 단위 값으로 지연 시간을 조절하는 것으로 1ms만큼 미세하게 조절한다. 예를 들어 지연 시간이 2ms(1/500초)인 경우 500Hz 간격으로 강조가 발생하여 해당 주파수 사이에서 감쇄가 발생한다. 이와 같이 지연 시간에 따라 적용되는 comb-filter의 주파수는 반비례하기 때문에 MIDI number[53]로 입력한

53) 0에서 127 사이의 정수로 MIDI에서 음정을 나타낸다.

음정의 주파수를 변환하면 입력되는 오디오의 음정에 맞는 정확한 지연 시간의 계산이 가능하다. Feedback 값은 최댓값이 1이나, 1에 가까워질수록 강조나 감쇄가 극단적으로 나타나며 1보다 큰 값일 경우 강한 공진[54]으로 comb-filter가 실행되지 않는다.

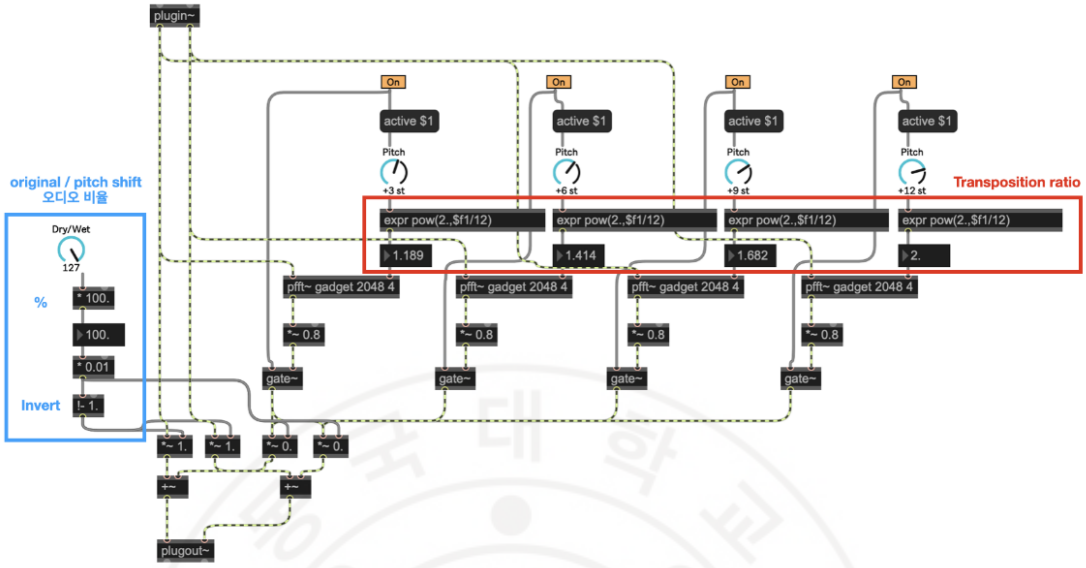
comb-filter는 보통 마이크와 스피커의 일정하지 못한 거리 차이나, 두 개 이상의 마이크의 거리 차이로 발생하는 음색의 왜곡으로 음향적인 측면에서는 음향 품질을 저하시키는 것으로 인식될 수 있지만 사운드 프로세싱에서는 소리를 넓히거나, 미세한 조절로 변화하는 차갑고 날카로운 음색을 의도적으로 사용할 수 있다.

⑦ FFT 분석을 이용한 pitch shift

푸리에 변환(Fourier transform)은 time domain[55]을 frequency domain[56]으로 변환하는 수학적 기법을 의미하는 것으로 시간에 대한 함수를 주파수 성분으로 분해하는 과정을 말한다. 푸리에 변환 방식을 빠르게 연산하는 것을 고속 푸리에 변환인 FFT(Fast Fourier Transform)[57]라고 한다. FFT는 다양한 분야에서 활용되고 있는데 음악에서는 주파수 성분을 분석하고 음 간의 미세한 차이를 구분하는 데에 유용하게 사용된다. Max에서 `fft~`오브젝트를 이용해 FFT 분석이 가능하지만 신호를 수정할 때마다 아티팩트(artifact)[58]가 생성된다. `fft~`의 단점을 해결하여 사용이 편리하고 효율적인 `pfft~`오브젝트를 이용해 프로세싱 패치를 연산 및 실행한다.

[그림-19]는 Max에서 `pfft~`를 이용한 pitch shift 음향효과 패치로 pitch shift는 음색의 왜곡 없이 음의 높이를 변화시키는 음향효과를 말한다.

54) 진동하는 진폭이 급격하게 늘어나거나 늘어난 현상.
55) 신호나 데이터가 시간에 따라 어떻게 변하는지를 나타내는 방법.
56) 신호에 포함된 주파수 성분인 진폭과 위상에 의해 다루어지는 변환영역.
57) 1822년 프랑스의 푸리에(Jean Baptiste Joseph Fourier)가 제시한 분석 이론.
58) 인공물 또는 오류로 해석되며 원치 않는 데이터 변경이나 선명도 손실을 나타낸다.



[그림-19] Max에서 **pfft~**를 이용한 pitch shift

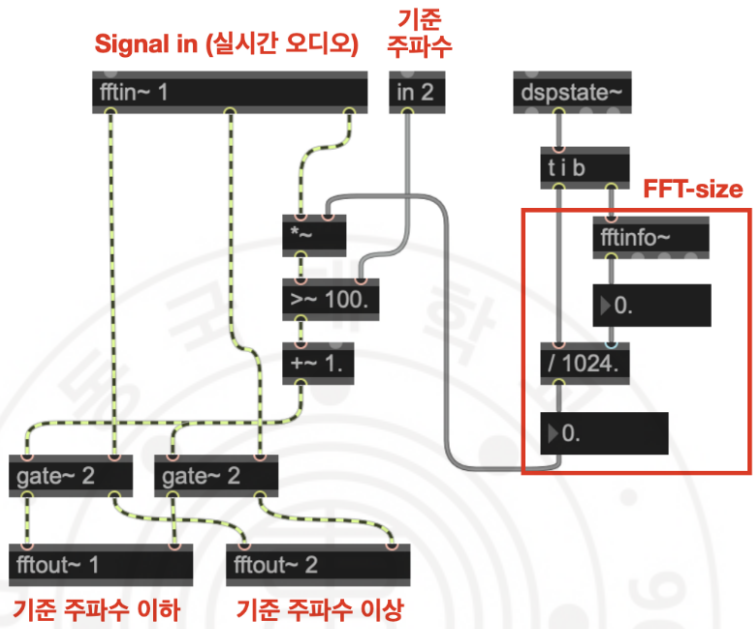
음의 높낮이는 semitone 단위로 조절하지만, 옥타브(octave) 단위를 사용하는 pfft~ gadget오브젝트를 위해 수학식을 계산하는 expr오브젝트[59]를 거쳐 Transposition ratio(비율)을 입력한다. original과 pitch shift 오디오 비율은 0에서 127의 값을 0에서 1값으로 변환하여 입력된 값으로 조절된다. 각 파라미터에 원하는 음정의 높낮이를 설정한 뒤 오디오를 입력하면 화음 관계가 형성되는 harmonizer 효과를 내게 된다.

⑧ FFT 분석을 이용한 pfft~ xover~

FFT 분석으로 원하는 특정 주파수를 기준으로 저음역과 고음역으로 분리시켜 음향효과 적용이 가능해진다. 분리된 주파수에 적용된 각 음향효과들은 서로 간의 간섭을 줄여 해당 기능을 명확하게 활용할 수 있는 장점이 있다.

59) C언어를 기반으로 수학 계산을 수행하는 오브젝트.

[그림-20]은 Max에서 특정 음역대를 분리할 수 있도록 만든 pfft~ xover~의 패치 내부이다.



[그림-20] Max에서 pfft~ xover~의 패치 내부

분리하고자 하는 음역대의 기준이 되는 주파수를 지정하면 >~오브젝트[60]에서 기준보다 높은 경우 1을 출력하고 그렇지 않은 경우 0을 출력한다. 이를 +~오브젝트[61]에서 1을 더하여 gate~오브젝트로 전달하면 기준보다 높은 경우 2가 되어 fftout~2오브젝트[62]로, 기준보다 낮은 경우 fftout~1오브젝트로 출력하여 고음역과 저음역의 분리가 가능해진다.

60) 입력되는 신호의 크기를 비교하는 오브젝트.
61) 입력되는 신호와 오브젝트에 적힌 숫자를 더해주는 오브젝트.
62) pfft~오브젝트에 로드되는 패치의 출력을 제공하는 것으로 지정되는 숫자에 따라 outlet의 순서가 정해진다.

3. 영상 시스템 연구

1) 영상 시스템

목소리를 이용한 실시간 인터랙티브 멀티미디어 작품을 만들기 위해 목소리와 프로세싱 된 사운드에 각각 반응하는 영상을 제작했다. 사운드 시스템에서 사용하는 프로그램인 Ableton Live와 연동하기에 효과적이고 데이터 정보를 호환하기 적합한 TouchDesigner[63]를 선택하여 영상을 제작하는 프로그램으로 사용했다. Ableton Live에서 TouchDesigner로 전달하는 오디오 정보는 OSC[64]를 통해 전송되며 전하고자 하는 오디오 정보는 선택할 수 있다. TouchDesigner로 전송된 정보와 연동된 영상은 프로젝터를 통해 출력된다.

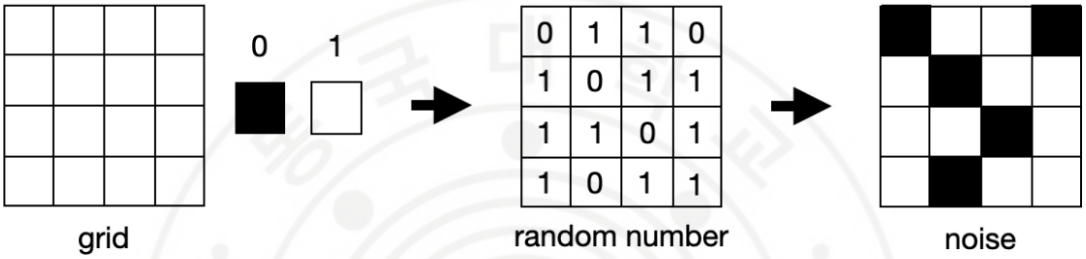
2) TouchDesigner를 이용한 영상 제작

① noise

20세기 이전에 noise(노이즈)는 방해가 되는 소리만을 의미했다. 그러나 1900년에서 1920년대 사이에 축음기, 전화기, 라디오와 같은 음향 재생 기술의 부상으로 노이즈를 측정하고 표현의 도구로 다루기 시작했다. 동시에 물리학자들의 브라운 운동[65] 이론이 개발되고 수학자들의 이론적 접근으로 측정 이론이 공식화되어 통신 시스템과 정보 전송에 적용하게 되면서 노이즈는 모든 종류의 신호, 정보의 오류와 편차의 의미로 개념적 변화를 겪게 됐다.

63) Derivative가 개발한 실시간 인터랙티브 콘텐츠를 위한 노드 기반의 프로그래밍 언어이다.
64) 음악적인 용도로 Adrian Freed와 Matt Wright가 CNMAT에서 개발한 것으로 사운드 데이터 전송을 위한 네트워크 통신 규약이다.
65) 영국의 로버트 브라운(Robert Brown)이 발견한 것으로 액체 혹은 기체 안에 떠서 움직이는 작은 입자의 불규칙한 운동을 말한다.

컴퓨터 그래픽스(computer graphics)[66]와 같이 컴퓨터로 이미지나 영상을 나타내는 과정에서 노이즈는 1980년대 텍스처 매핑(texture mapping)[67]으로 인해 개발됐다. 물체의 시각적 속성을 변조하기 위해 사용되던 RNG[68]는 표현의 한계가 있어, 수학적 기반으로 그리드에 걸쳐 그레이디언트(gradient)[69]를 생성하고 보간하는 노이즈 함수를 개발하기 시작했다. [그림-21]은 그레이디언트 노이즈의 기본 원리를 나타낸 그림이다.



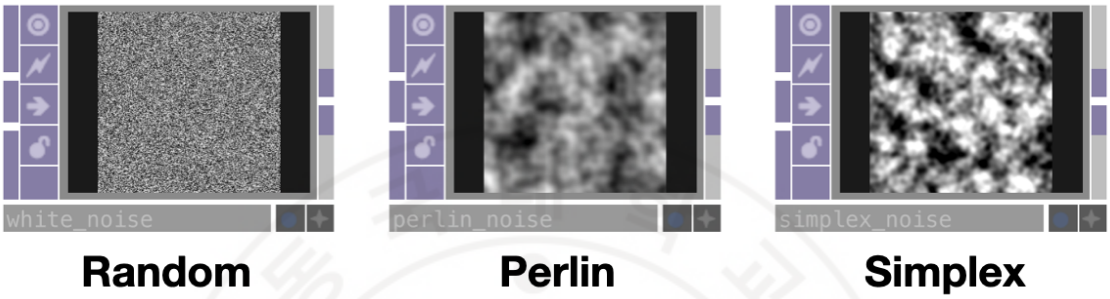
[그림-21] 그레이디언트 노이즈 기본 원리

그레이디언트 노이즈에서 가장 대표적인 것은 Ken Perlin[70]이 처음 개발한 Perlin noise와 이를 개선한 Simplex noise이다. 완전한 무작위 난수를 사용하는 방식이 아닌 의사 난수(pseudo random)[71]를 사용하여 전체적인 변화폭을 줄여 텍스처를 자연스럽게 생성하는 것이 Perlin noise이고, 고차원에서 Perlin noise의 한계를 개선한 것이 Simplex noise이다.

TouchDesigner에서 이러한 텍스처와 이미지를 조작하는 오퍼레이터 중 Noise TOP오퍼레이터[72]를 통해 노이즈 함수들과 그에 따른 이미지를 확인

66) 컴퓨터에 의한 영상처리로 CG로 약칭되는 경우가 많다.
67) 3D 물체의 표면에 사실감을 부여해 주기 위한 기술이나 과정을 말한다.
68) Random Number Generator의 약자로 난수를 발생하는 장치이다.
69) 기울기, 변화도, 미적분의 벡터 등을 나타낸다.
70) 뉴욕대 컴퓨터 공학과 교수로 Perlin noise, Simplex noise뿐만 아니라 실시간 인터랙티브 캐릭터 애니메이션과 컴퓨터 유저 인터페이스를 개발하거나 참여한 인물이다.
71) 지정된 초깃값을 이용하여 난수를 생성하는 것으로 주로 알고리즘으로 사용한다.
72) 노이즈 패턴을 생성하고 결합하는 데 사용하는 오퍼레이터이다.

한다. 앞서 설명한 Perlin과 Simplex뿐만 아니라, Random, Sparse, Hermite 등의 노이즈 함수를 사용할 수 있도록 타입별로 분류되어 있다. [그림-22]는 TouchDesigner의 노이즈 타입 중 Random, Perlin 그리고 Simplex를 타나낸 것이다.



[그림-22] TouchDesigner의 노이즈 타입

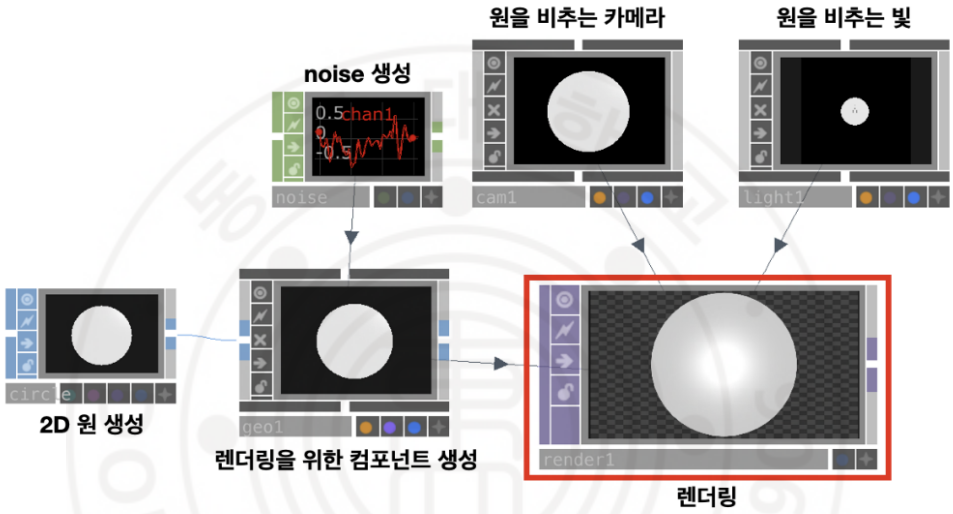
Random은 white noise 기반으로 동일한 강도를 갖는 완전한 무작위 노이즈 패턴을 생성하고 있다. 반면에 Perlin은 Random보다는 자연스러운 패턴을 생성하고 있으며, Simplex는 Perlin보다 더 부드러운 패턴을 생성하고 있다. 난수를 생성하여 입력하는 seed 혹은 주기와 진폭을 조절하면 패턴의 모양이 변화한다. 해당 노이즈 패턴을 그대로 사용하거나 이미지와 결합하여 사용함으로 영상이나 이미지를 제작한다.

② CHOPs

TouchDesigner에는 6개의 오퍼레이터 패밀리(operator family)[73]를 제공하고 있는데 CHOPs는 그중 하나의 패밀리 오퍼레이터이다. 오디오와 움직임을 생성하고 숫자 정보를 처리하거나 연산할 수 있으며, MIDI 데이터 혹은

73) 체계 구조와 경로를 형성하는 1개의 컴포넌트 COMPs(Components)와 5개의 TOPs(Texture Operators), CHOPs(Channel Operators), SOPs(Surface Operators), DATs(Data Operators), MATs(Material Operators)가 있다.

각종 기기 및 프로토콜 간의 연동으로 데이터를 처리할 수 있는 기능을 제공한다. 음악과 영상을 연동하고 제어하기에 필수적인 기능을 담고 있어 CHOPs는 멀티미디어 작품을 제작하는 데에 다양하게 사용이 가능하다. [그림-23]은 CHOPs 중 Noise CHOP을 사용한 예시를 TouchDesigner 패치로 구현한 것이다.



[그림-23] Noise CHOP의 사용 예시를 TouchDesigner 패치로 구현

Noise CHOP은 -1에서 +1 범위 값에서 불규칙한 파동을 생성하는 오퍼레이터로 모양이 다른 여러 개의 커브를 주기, 진폭 등을 조절하여 생성할 수 있다. 먼저 2D 원을 생성하여 원을 렌더링 할 수 있도록 Camera COMP오퍼레이터[74]와 Light COMP오퍼레이터[75] 그리고 Geometry COMP오퍼레이터[76]를 생성한 후 Render TOP오퍼레이터[77]로 렌더링 시켜준다. Noise CHOP으로 하나의 불규칙한 파동을 임의로 생성하고 Noise CHOP 파라미터의

74) 3D 오브젝트의 바라보는 시점을 Render TOP에 제공하는 오퍼레이터.
75) 3D 오브젝트에 빛을 투사하여 색상, 밝기, 분위기를 제어하는 오퍼레이터.
76) Render TOP에 렌더링 할 3D 오브젝트를 나타내는 오퍼레이터.
77) TouchDesigner에서 모든 3D 장면을 렌더링 하는 데 사용되는 오퍼레이터.

translate z 좌표에 frame rate[78]만큼 움직이도록 'absTime.frame'[79]이라는 명령어를 입력한다. Noise CHOP의 정보를 Geometry COMP의 파라미터 translate z 좌표에 입력하면 Noise CHOP이 생성하는 파동의 모양과 움직임대로 원의 크기가 변하는 것처럼 나타난다. [그림-24]는 Noise CHOP에 따라 변형되는 원의 모양을 나타낸 것이다.

[그림-24] Noise CHOP에 따라 변형되는 원의 모양

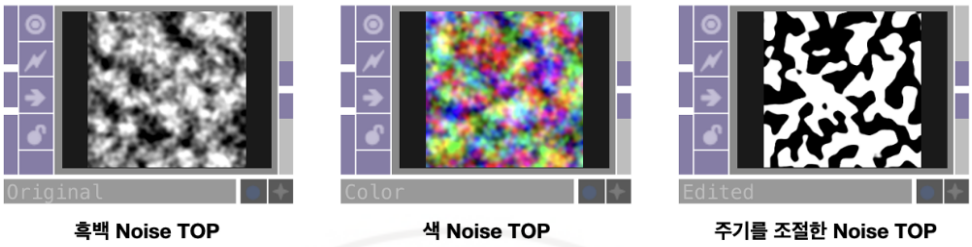
본 작품의 경우 Geometry COMP의 파라미터 translate x, y, z 좌표에 입력하기 위한 서로 다른 주기와 진폭을 가진 Noise CHOP을 생성하여 목소리의 저음과 전체 음량 값 그리고 stereo delay 음향효과가 적용된 목소리와 연동했다. 목소리의 크기에 따라 원의 크기가 달라지고 stereo delay 음향효과가 적용될 경우 원의 움직임이 빨라져 여러 개의 원이 분산되는 것처럼 나타내어 청각 정보를 시각화할 수 있도록 표현했다.

③ TOPs

TOPs는 실시간 GPU 기반으로 2D 이미지를 생성하고 결합하거나 수정하는 기능을 제공하는 오퍼레이터이다. 파일이나 네트워크에서 이미지와 동영상을 읽고 쓰기가 가능하여 거의 모든 이미지 작업에 사용할 수 있다. CHOPs와 마찬가지로 TOPs에도 존재하는 Noise TOP으로 다양한 노이즈

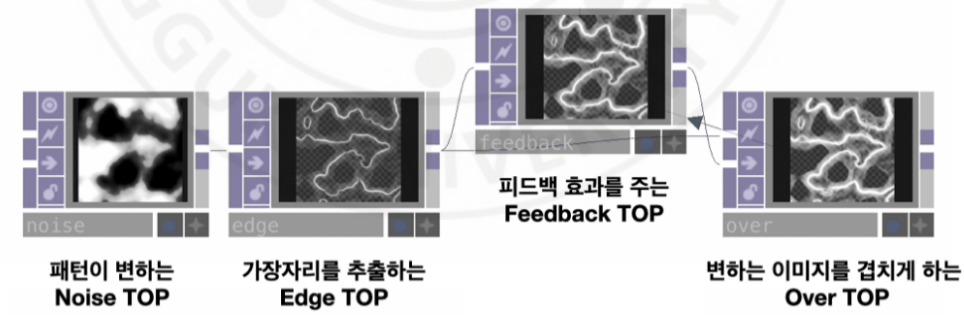
78) 초당 진행되는 프레임의 수로 숫자가 높을수록 화면에 자연스러운 움직임을 나타낸다.
79) 'Absolute Time.frame'으로 프레임 수를 절대적인 시간으로 두고 프로세스를 0부터 카운트한다는 명령어이다.

패턴을 생성하고 결합하는 데 사용할 수 있다. [그림-25]는 Noise TOP으로 생성할 수 있는 패턴의 예시를 TouchDesigner 패치로 나타낸 것이다.



[그림-25] Noise TOP 패턴의 예시를 나타낸 TouchDesigner 패치

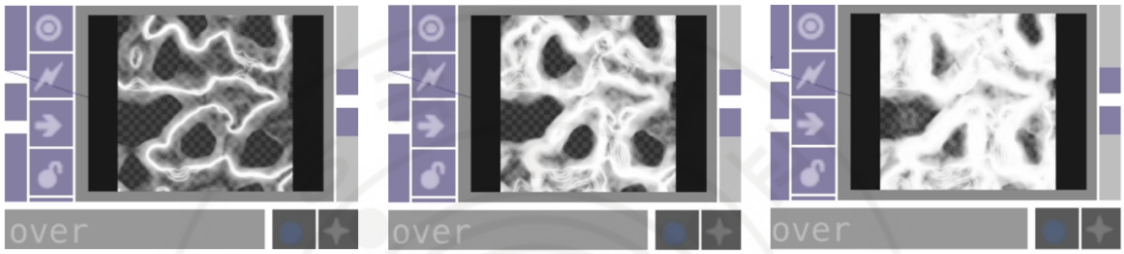
Noise TOP이 생성하는 패턴 이미지를 반복적으로 나타내기 위해 생성된 패턴 이미지를 순차적으로 겹쳐지게 표현하도록 한다. 이미지를 생성하고 사라지는 것을 음악에 맞춰 효과적으로 표현하기 위해 변화하는 이미지의 잔상이 남도록 Feedback TOP오퍼레이터[80]를 사용하고 효과를 분명하게 시각화하기 위해 패턴 이미지의 가장자리를 추출하는 Edge TOP오퍼레이터[81]를 사용한다. [그림-26]은 Edge TOP과 Feedback TOP을 이용한 TouchDesigner 패치로 피드백 효과의 예시를 나타낸 것이다.



[그림-26] Edge TOP과 Feedback TOP을 이용한 예시

80) 출력 이미지를 재입력하여 피드백 효과를 만드는 데 사용하는 오퍼레이터.
81) 이미지의 가장자리를 강조하여 나타내는 오퍼레이터로 서로 인접해 있는 픽셀의 값을 비교해 임곗값이 큰 것이 출력값이 더 높다.

Noise TOP의 파라미터 translate z 좌표에 ‘absTime.frame’ 명령어를 입력한 뒤 속도를 낮추기 위해 0.001을 곱하여 패턴 이미지에 변화를 준다. 변화하는 패턴의 가장자리를 Edge TOP로 추출하고 이를 Feedback TOP을 거친 것과 거치지 않은 Edge TOP를 합하게 하는 Over TOP오퍼레이터[82]와 연결한다. [그림-27]은 앞선 설명에 따라 추출한 가장자리의 패턴 이미지가 점점 겹치는 과정을 나타낸 것이다.



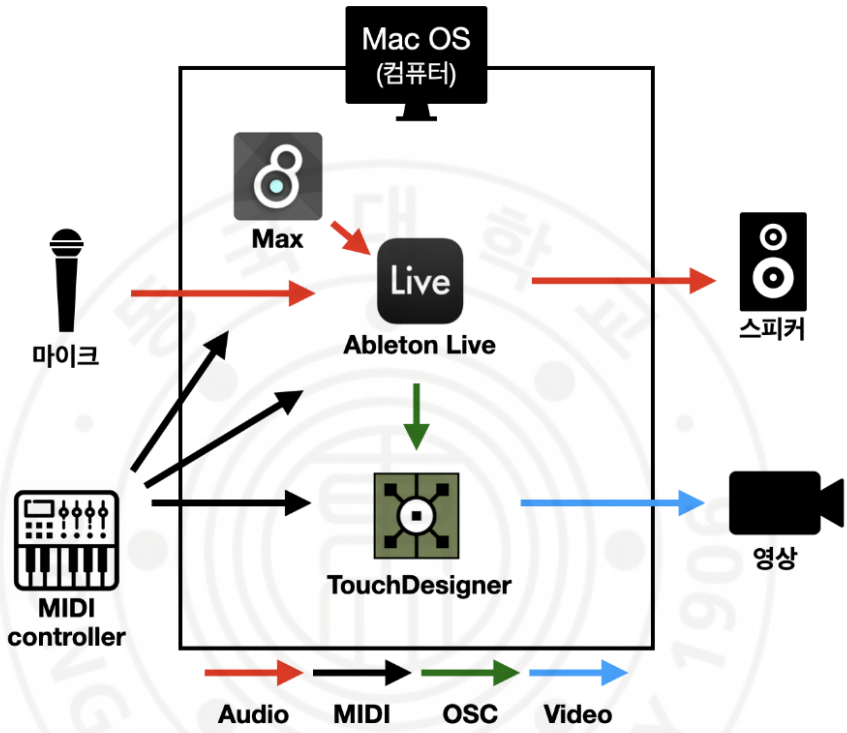
[그림-27] Feedback TOP과 Over TOP으로 이미지가 겹치는 과정

본 작품의 경우 Noise TOP의 파라미터 translate x, z 좌표에 리듬감 있는 목소리로 연주되는 악기의 음량 값과 연동하고 Level TOP오퍼레이터[83]를 연결한 뒤 목소리의 음량 값과 연동했다. 악기의 리듬감에 따라 패턴 이미지가 변화하고 목소리 크기에 따라 밝기와 선명도가 조절되어 혼란스럽고 요동치는 분위기와 분명해지고 흐려지는 존재감을 시각화했다.

82) input1을 input2 위에 겹치도록 배치하는 오퍼레이터.
83) 이미지의 대비, 밝기, 감마, 블랙 레벨, 색상 범위, 불투명도 등을 조정하는 오퍼레이터.

4. 공연 및 연동 시스템 연구

1) 공연 시스템



[그림-28] 공연 시스템 설계도

[그림-28]은 무대에서 사용되는 사운드와 영상의 총 시스템 설계도이다. 목소리에 사용한 마이크로 전달한 오디오 신호는 오디오 인터페이스를 통해 컴퓨터로 입력된다. 이는 Ableton Live로 입력되어 오디오 트랙마다 적용된 Max for Live를 이용해 사운드 프로세싱을 거치게 된다. 각 음향효과는 MIDI 매핑으로 연결된 MIDI 컨트롤러로 제어된다. 영상과 연동할 오디오 정보들은 OSC를 통해 Touch Designer에 실시간으로 정보가 전달된다. 만들어진 영상은 전달받은 오디오 데이터 정보와 실시간 제어로 프로젝터를 통해 출력된다.

2) 음악과 영상의 연동 시스템

① OSC

OSC는 UDP(User Datagram Protocol) 통신을 기반으로 한 네트워크 프로토콜[84]로 다양한 디바이스 간의 데이터 교환을 가능하게 한다. Open Sound Control[85]의 약자인 OSC는 주로 실시간 사운드 처리와 다양한 프로그램 간의 연동에 용이하며 높은 정확성과 낮은 지연 시간을 가지는 것이 특징이다. 비슷한 전송 계층 프로토콜로는 TCP/IP(Transmission Control Protocol /Internet Protocol)가 존재하지만 TCP/IP 방식의 경우 보안 검사 과정에서 지연 시간이 발생하기 때문에 다양한 프로그램이 연동되는 멀티미디어 작품에 활용하기 적합하지 않다.

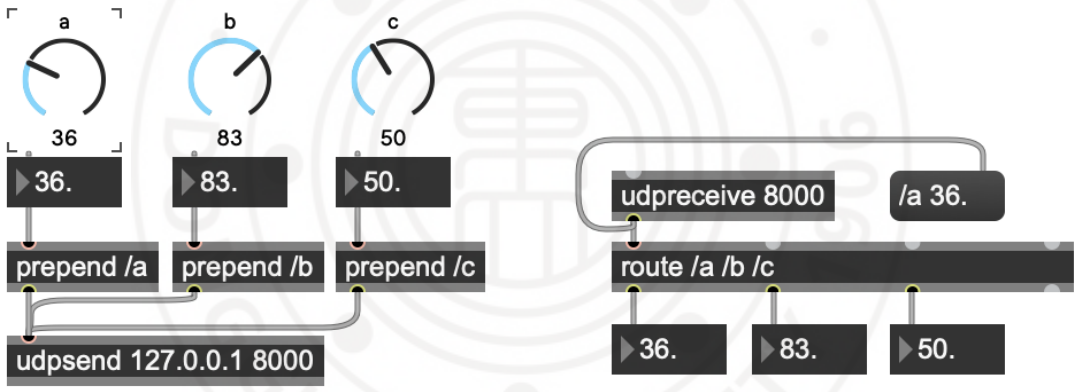
UDP는 IP(Internet Protocol) 종류 중 하나로 IP 주소(IP address)[86]로 통신하며 0에서 255까지의 네 자리 숫자 포맷의 형식을 취하고 있다. 모든 컴퓨터는 고유한 IP 주소를 형성하고 있어 다른 컴퓨터와 데이터 송수신을 원할 경우 해당 컴퓨터의 IP 주소를 통해 연결이 가능하다. 하나의 컴퓨터 내에서 데이터 송수신을 원할 경우 IANA(Internet Assigned Numbers Authority)[87]에서 지정한 '127.0.0.1'[88]을 통해 연결이 가능하다. 이는 컴퓨터가 IP 주소와 관계없이 자신을 식별하는 데 사용된다.

UDP는 TCP/IP보다 더 작은 패킷(packet)[89]으로 전송하기 때문에 지연을 줄여 속도가 빠르고, 일부 패킷이 누락되어도 전송이 중단되지 않아 효율적이다. 반면에 데이터를 온전하게 전송하고 데이터 도달을 확인하여 실패 시

84) 통신 장비 사이에서 데이터 통신을 원활하게 하기 위한 통신 규약.
85) 음악적인 용도로 Adrian Freed와 Matt Wright가 CNMAT에서 개발했다.
86) 인터넷에 연결된 모든 장치에 할당된 고유 식별 번호.
87) 인터넷의 원활한 운영을 위해 IP 주소, 최상위 도메인 등을 관리하는 곳이다.
88) 특수 목적의 IP 주소로 로컬 호스트(localhost) 또는 루프백(loopback) 주소라고 말한다.
89) 데이터를 보내기 위한 경로를 배정할 때 효율성을 높이기 위해 데이터를 나누는 조각.

재전송을 시도하는 TCP/IP와 달리 UDP는 전송 여부와 손실 여부를 알 수 없어 신뢰성이 낮다. 하지만 지속적으로 변하는 데이터를 빠르게 전송해야 하는 온라인 게임이나 실시간 스트리밍의 경우 UDP를 사용하는 것이 적합하다.

실시간으로 오디오 데이터를 통신하고 프로세싱 하는 경우에도 UDP를 사용하는 것이 적합하기 때문에 이를 기반으로 한 OSC는 음악, 비주얼, 멀티미디어 커뮤니케이션 등에서 사용된다. OSC를 사용하기 위해서는 IP 주소, port number(포트 번호), address pattern(주소 패턴), arguments(인수)를 필요로 한다. [그림-29]는 앞선 설명의 OSC의 송수신을 위해 필요한 4가지 요소를 바탕으로 OSC의 사용 예시를 Max로 구현한 것이다.



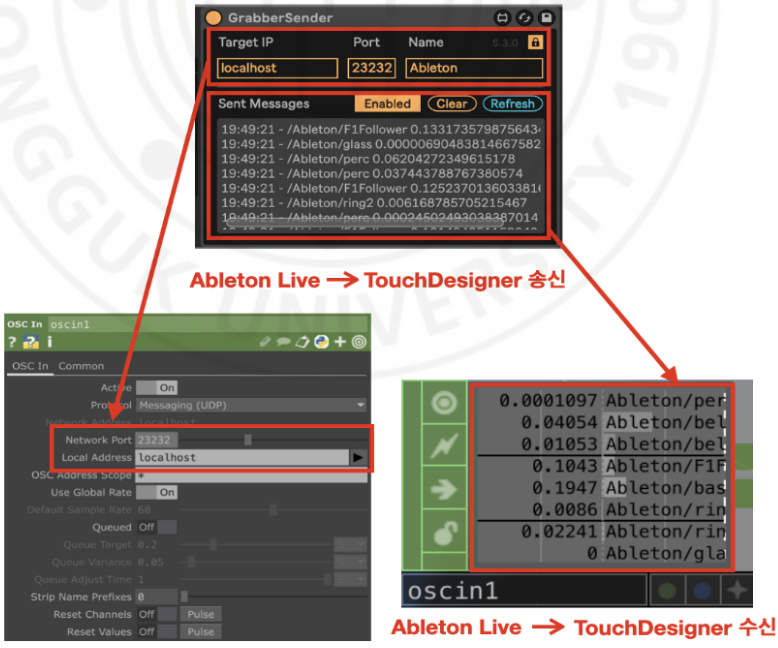
[그림-29] OSC의 사용 예시를 Max로 구현한 패치

Max에서 udpsend오브젝트[90]와 udpreceive오브젝트[91]를 이용해 OSC를 사용한다. [그림-29]는 동일 컴퓨터 내에서 데이터를 송수신하는 경우로 IP 주소는 127.0.0.1으로 지정하고 포트 번호는 8000으로 설정한다. 포트 번호는 데이터 처리의 통로를 식별하는 것으로 IP 주소와 함께 사용한다. 16비트의 길이를 갖게 되어 2^{16}인 65536개 즉, 0에서 65535의 범위 내의 숫자로 지정하

90) 데이터를 송신하는 오브젝트.
91) 데이터를 수신하는 오브젝트.

도록 되어 있다. 주소 패턴은 데이터의 종류를 구분 지어 보내도록 도와주며 '/'와 함께 붙여 사용한다. [그림-29]과 같이 각 a, b, c 파라미터를 실시간으로 조절하며 데이터를 전송할 때, 그 값을 /a, /b, /c에 해당하는 경로로 수신하여 인수로 나타내고 있다.

본 논문에서 인터랙티브 멀티미디어 퍼포먼스를 위해 사용된 Max, Ableton Live, TouchDesigner 또한 서로 다른 프로그램으로 연동과 실시간 제어에 용이한 OSC를 사용했다. Max와 Ableton Live의 경우 별도 OSC 연결 없이 Max for Live로 연동되지만, Ableton Live의 사운드 정보를 TouchDesigner로 전송하기 위해 OSC를 사용했다. Ableton Live의 오디오 트랙마다 적용되는 Max for Live의 음향효과가 다르기 때문에 각 음향효과에 맞춰 변화하는 영상효과를 다르게 나타내기 위해 전달하는 사운드 정보를 개별화할 필요가 있다. 따라서 각 오디오 트랙마다 해당 오디오 트랙의 사운드 정보만을 전송하는 OSC를 설정하여 TouchDesigner로 전송했다.



[그림-30] Ableton Live와 TouchDesigner의 OSC 통신

[그림-30]은 Ableton Live와 TouchDesigner의 OSC 연동 과정을 나타낸 것이다. Ableton Live에서 전송하고자 하는 오디오 정보를 선택하고 해당 오디오 트랙에 GrabberSender[92]플러그인(plug-in)[93]을 사용하여 송신하는 주소와 포트 번호를 설정한다. TouchDesigner에서 OSC를 수신하기 위해 OSC In CHOP오퍼레이터[94]를 사용한다. 송신 단계에서 설정한 주소와 포트 번호를 파라미터에 입력하면 GrabberSender에서 실시간으로 송신하는 메시지 정보와 OSC In CHOP에서 수신하는 메시지 정보가 일치하는 것을 나타낸다.

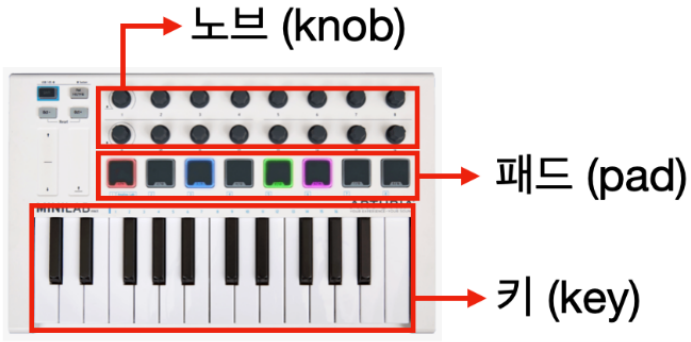
② MIDI 매핑

인터랙티브 멀티미디어 퍼포먼스를 진행하고자 할 때, 음악과 영상을 제어하는 파라미터를 효과적으로 조절할 수 있도록 MIDI 매핑(MIDI Mapping)을 사용한다. MIDI 매핑은 컴퓨터 소프트웨어 프로그램의 음향효과를 버튼이나 키보드 같은 하드웨어 컨트롤러와 연결시켜주는 것을 말한다.

MIDI 매핑을 하려면 먼저 사용하는 소프트웨어 프로그램과 MIDI 컨트롤러가 필요하다. 본 연구작품에서 사용된 프로그램 중 매핑을 필요로 하는 프로그램은 Ableton Live와 TouchDesigner이고, MIDI 컨트롤러는 Arturia의 MiniLab MK2이다. 연구한 음향효과마다 적용하고자 하는 파라미터 값을 유연하게 조절할 수 있는 노브(knob), 패드(pad), 키(key)를 모두 갖추고 있는 컨트롤러로 MiniLab MK2를 사용했다. 본 논문을 위해 사용된 컨트롤러의 모습은 아래 [그림-31][95]과 같다.

---

92) Showsync에서 제공하는 음악적 도구 'LiveGrabber'중 하나로 네트워크 상에서 데이터 정보를 OSC 전송할 때 사용한다.
93) 컴퓨터에서 사용하는 기존 프로그램에서 특정 기능을 추가하는 소프트웨어 구성 요소.
94) TouchDesigner로 수신되는 OSC 메세지를 받는 데 사용되는 오퍼레이터.
95) https://www.arturia.com/minilab-mkii/resources



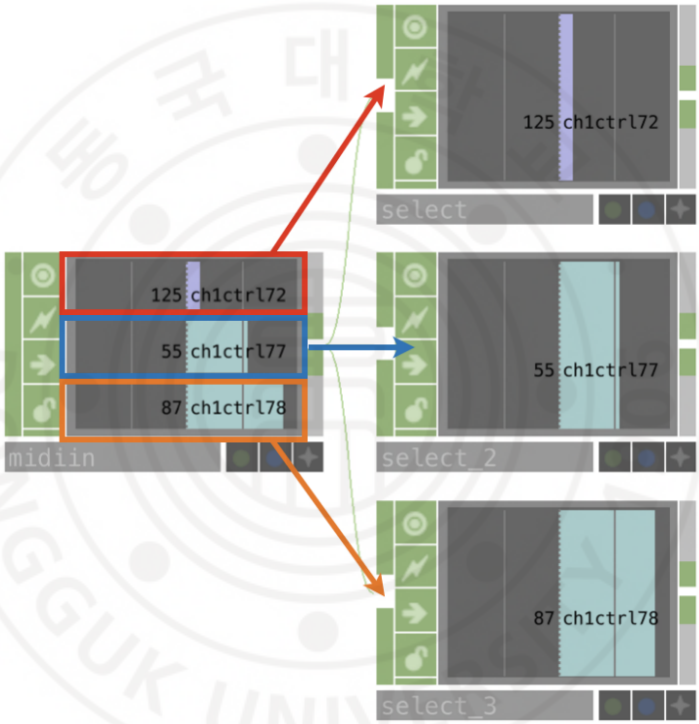
[그림-31] Arturia MiniLab MK2

Ableton Live의 매핑 방법은 간단하다. 컨트롤러를 연결한 뒤, MIDI 맵 모드를 활성화하여 적용하고자 하는 음향효과를 클릭하고 컨트롤러를 움직이면 된다. 예를 들어 pitch의 매개변수 값을 -127에서 127까지 자유롭게 제어하고 싶을 경우, MIDI 맵 모드를 활성화하고 pitch 파라미터를 클릭한 상태에서 컨트롤러의 1번 노브를 움직이면 매핑이 적용된다. 노브, 패드, 키를 선택한 기준은 적용하려는 기능의 형태에 따라 분류되었는데, 숫자 값을 리니어(linear) 하게 주려고 할 경우 노브를 사용했고, 음향효과를 선택적으로 적용하기 위한 on, off의 경우와 입력하고자 하는 파라미터의 값을 숫자로 입력하지 않고 랜덤 값을 호출해서 명령하는 경우 패드를 사용했다. 키의 경우 패드와 마찬가지로 버튼적 기능을 수행할 수 있지만, 사용하고자 하는 음정을 직관적으로 명령하기 위해 사용했다.

TouchDesigner의 매핑 방법 역시 간단하다. MIDI In CHOP오퍼레이터[96]를 사용해 MIDI 컨트롤러 내에서 사용하고자 하는 노브를 움직이면 MIDI In CHOP에서 해당 노브의 채널 번호가 나타나며 오퍼레이터에 활성화된 것을 나타낸다. 활성화시킨 각 컨트롤러는 Select CHOP오퍼레이터[97]를 이용해 선별한 뒤 TouchDesigner에서 조절하고자 하는 파라미터와 연결하여 사용한

96) TouchDesigner와 연결된 MIDI 장치의 MIDI 정보를 수신하고 수행하는 오퍼레이터.
97) CHOPs에서 채널을 선택하거나 지정할 때 사용하는 오퍼레이터.

다. [그림-32]는 TouchDesigner에서 MIDI 매핑을 하는 과정을 나타낸 것이다. MIDI 데이터[98]는 0에서 127의 숫자 값만 사용하기 때문에 필요에 따라서 Select CHOP에 Math CHOP오퍼레이터[99]를 이용해 파라미터가 요구하는 숫자 범위를 설정한 뒤 연결하여 사용한다. 대부분의 파라미터는 OSC를 통해 음악과 연동되어 있어 TouchDesigner에서의 MIDI 매핑은 영상을 전환하거나 제어하기 위해 사용했다.



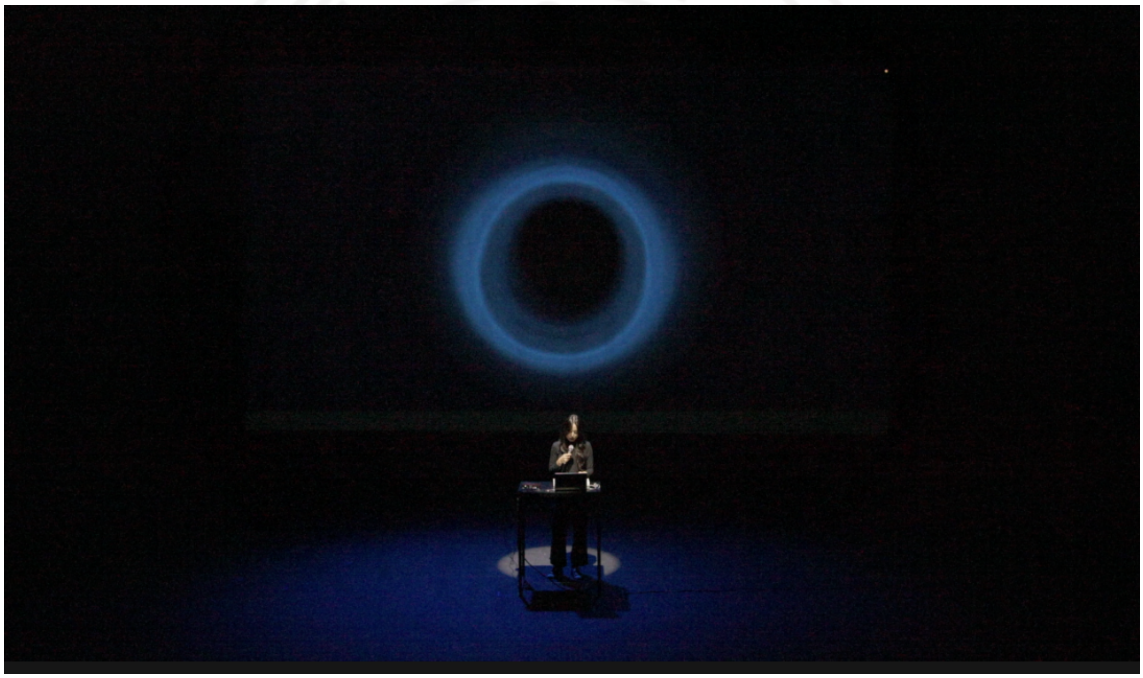
[그림-32] TouchDesigner에서 MIDI 매핑을 하는 과정

98) 음악 작업을 수행하는 데 있어 필요한 MIDI 정보를 전달하는 디지털 숫자 스트림.
99) CHOPs에서 산술 연산을 수행하며 스케일을 조정할 수 있는 오퍼레이터.

Ⅲ. 연구 기술의 작품 적용

1. 작품 소개

연구된 목소리의 실시간 사운드 프로세싱을 이용한 인터랙티브 멀티미디어 작품 <Adsum (I am present)>은 2024년 11월 9일 동국대학교 이해랑예술극장에서 진행된 'SEEING SOUND, LISTENING IMAGE(보는 소리, 듣는 영상) 2024'에서 초연되었다.



[그림-33] 작품 <Adsum (I am present)>의 초연 장면

Adsum은 라틴어로 '존재하다'라는 뜻과 함께 '여기에 있다'는 상태를 나타내는 단어이다. 작품 <Adsum (I am present)>는 보이지 않는 존재의 불명료한 개념에 대한 이해와 존재의 의미에 관한 물음을 담고 있다. 나무, 꽃, 새와 같이 '존재하는 것'은 우리 눈에 보이지만, '있다'라고 말하나 눈에 보이

지도, 귀에 들리지도, 손으로 만져지지 않는, 이를테면 생각, 믿음, 사랑과 같은 '존재'의 '존재 의미'에 대한 물음을 이야기하는 것이다.

작품에서 반복하여 부르는 두 문장의 노랫말도 라틴어이다. 하나의 문장에 여러 의미를 내포하고 있어 부를 때마다 의미는 다르게 해석될 수 있다. 먼저 'Quis vídeor?'는 어느, 누구, 무엇, 누군가 등의 quis와 보이다, 인지하다, 눈에 띄다, ~처럼 보이다의 vídeor가 더해진 것으로 '어느 것이 보이는가', '무엇처럼 보이는가', '어떤 것으로 보이는가' 등을 뜻한다. 이는 존재 의미에 대해 개인 스스로가 고찰한다는 것을 나타낼 뿐만 아니라 존재의 입장에서 타자에게 존재 의미에 대해 질문하는 것을 나타낸다. 'Contínuo hīc adsum'는 지속적으로, 연속적으로, 끊임없이, 계속해서의 contínuo와 여기의 hīc 그리고 존재하다, 현존하다의 adsum이 더해진 것으로 '나는 계속 여기 있다', '지속적으로 있다', '끊임없이 존재한다' 등으로 존재의 입장에서 자신의 존재 상태를 대변하는 것을 나타낸다.

'내가 보이는가'라는 질문으로 말하고 있듯 눈에 보이지 않는 불명료한 개념에 대한 추상적인 이미지를 영상으로 표현하고, 계속해서 변화하는 생각을 묘사하기 위해 여러 패턴의 영상을 음악과 함께 전환시켰다. 음악은 목소리로 쌓여가는 악기 소리와 멜로디 형식으로 연주되는 질문과 대답이 음향효과와 겹쳐진다. 점점 조금씩 고조화되는 음악은 분명한 듯 분명하지 않은 각자의 존재에 대한 개념이 개인에게 인식되고 확립되어가는 과정을 나타낸다.

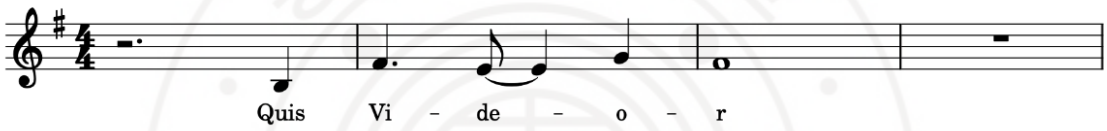
2. 작품 구성

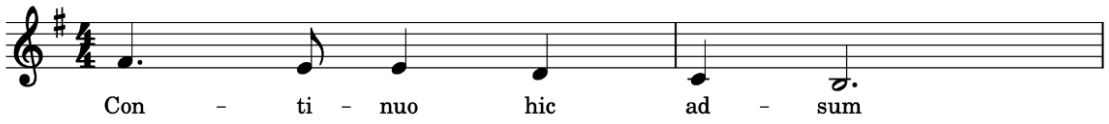
1) 음악 구성

본 작품은 존재가 인식되어 혼란을 겪다 점차 존재의 의미가 확립되어가는 과정을 표현하기 위해 A-B-C의 구성으로 나누었다. E단조의 음악으로 저음부를 강조하고 공간감을 주는 음향효과들을 통해 전체적으로 어둡고 신비스러운 분위기를 조성했다. <표-2>는 음악 구성에 따라 목소리의 활용법과 그에 따른 악기 그리고 음향효과를 표로 정리하여 나타냈다.

<표-2> 음악 구성

section	목소리	악기	음향효과
A	숨소리(breathing) 속삭임(whisper)	pad bass synth piano	audio to MIDI FM synth comb-filter stereo delay
B	진성(modal voice)	pad bass synth pad bell pluck	audio to MIDI FM synth comb-filter ping pong delay pitch shift granular synth ring modulation
C	흉성(chest voice)	pad bass synth pad bell pluck synth piano synth bass	granular synth comb-filter stereo delay pitch shift ring modulation

A section에서는 사람이 일반적으로 소리를 내는 방식으로 표현한 것이 아닌, 입으로 호흡을 뱉는 소리를 마이크로 입력하여 받은 오디오 신호를 통해 만들어진 FM synth 악기 소리들이 쌓이게 된다. 하나씩 쌓여가는 pad, bass, synth piano 등 과 해당 악기들에 더해지는 음향효과들로 존재가 처음 인식되어 점차 확장되어가는 것을 표현했다. 목소리로 'Quis vídeor?'를 작게 속삭인다. 속삭이던 목소리는 조금씩 멜로디의 형태로 연주되고 `pfft~` `xover~`로 높은 음역대의 소리들에 적용된 음향효과가 덧붙여지면서 B section으로 넘어간다. [그림-34]와 [그림-35]는 'Quis vídeor?'와 'Contínuo hīc adsum'의 멜로디를 그린 악보이다.

[그림-34] 'Quis vídeor?' 멜로디 악보

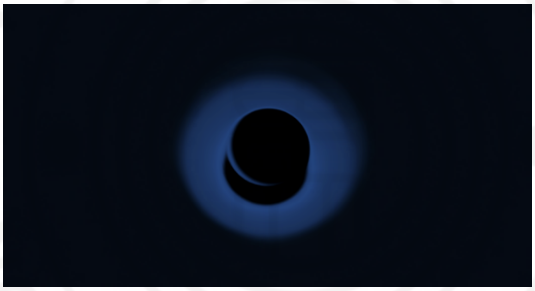
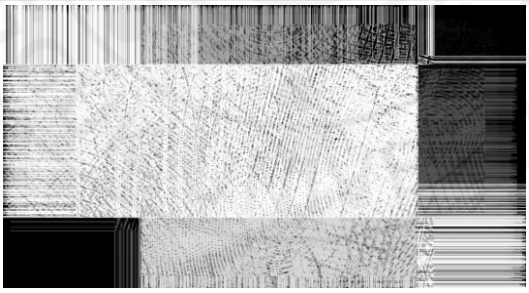
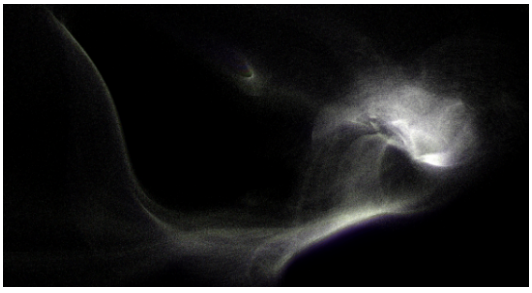
단순한 선율로 연주되는 'Quis vídeor?' 멜로디에 음향효과를 하나씩 더해가며 적용했다. 하나의 문장에 여러 의미를 내포하고 있다는 것과 하나의 존재에 여러 의미를 가지고 있다는 것을 나타내고자 했다. comb-filter, ring modulation, ping pong delay, pitch shift를 반복적으로 사용하여 멜로디 음이 교차하는 방식으로 혼란을 겪는 것을 표현했다. 질문하던 멜로디에 대답하듯 'Contínuo hīc adsum'의 멜로디가 연주되며 C section으로 넘어간다.

[그림-35] 'Contínuo hīc adsum' 멜로디 악보

C section은 B section에 비해 멜로디 음의 연주 빈도가 높아지며 original sound보다 음향효과가 적용된 오디오 비율도 높아진다. 저음역대를 강하게 표현하는 흉성을 사용함과 동시에 멜로디의 저음부만 추출해 FM synthesis 방식으로 더 깊은 저음 악기 소리를 추가함으로 긴장감을 더했다. ‘Contínuo hīc adsum’의 멜로디에 granular synthesis를 통해 소리를 변형시켜 delay시간을 랜덤하게 하고 pitch shift를 이용해 음정과 화음 수를 늘려 목소리가 많아지는 것처럼 느끼게 했다. 음악의 마지막 부분에 도달할수록 혼란을 표현하기 위해 반복적으로 사용하던 음향효과의 소리를 조금씩 소멸시키며 여전히 불명료하기는 하나 점차 존재의 의미를 확립해가는 것을 표현했다.

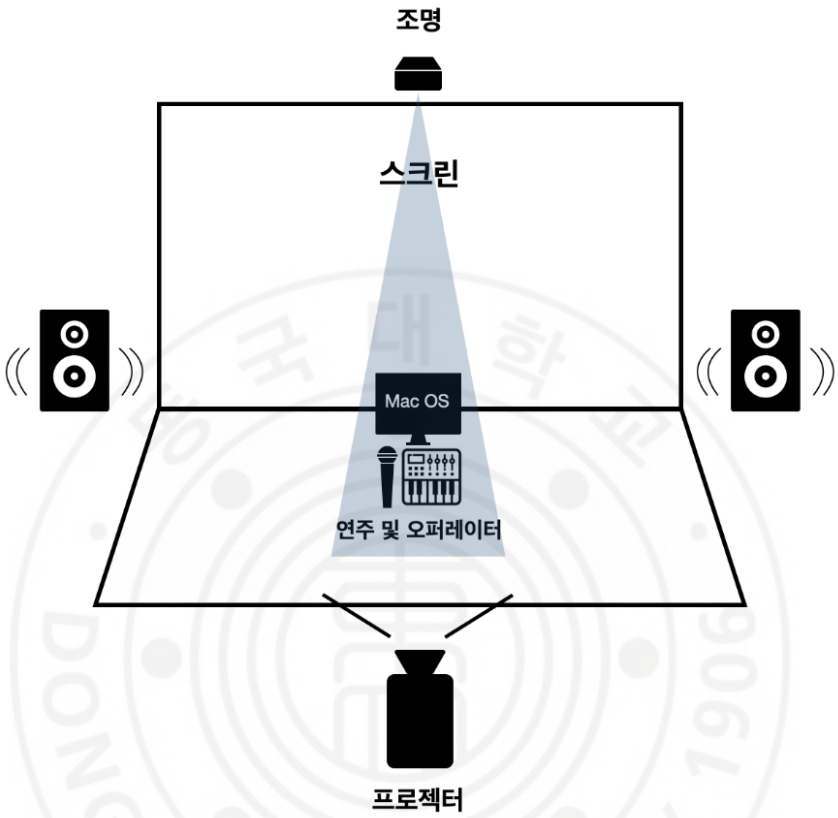
2) 영상 구성

영상은 음악에 맞추어 A-B-C의 구성으로 3가지 패턴을 사용했다. 사용된 패턴들은 음악의 전체적인 분위기와 각 구성에 맞는 의미를 시각적으로 표현한다. 단조 음악의 minor 한 느낌과 어둡고 신비한 분위기를 흑백과 차가운 파랑색으로 표현하고 저음부의 울림을 나타내기 위해 분산되는 영상효과를 음악과 연동하여 사용했다. <표-3>는 영상 구성에 따라 진행되는 이미지와 의미를 정리하여 나타낸 것이다.

<표-3> 영상 구성

section	이미지	의미
A		존재 생성과 인식
B		존재 의미에 대한 혼란
C		존재 의미 확립 그리고 소멸

3) 무대 구성



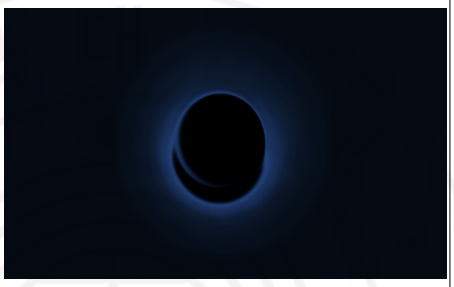
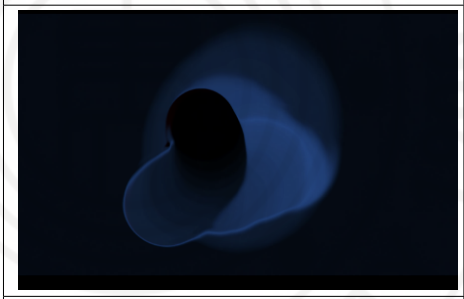
[그림-36] 작품 <Adsum (I am present)> 무대 구성

[그림-36]은 작품 <Adsum (I am present)>의 무대 구성이다. 목소리 연주와 시스템 제어를 직접 혼자 해야 하기 때문에 마이크, 컴퓨터 그리고 MIDI 컨트롤러는 함께 배치한다. 목소리로 연주되는 모든 소리와 사운드 프로세싱을 거친 소리는 공연장의 스피커로 출력되고 음악과 연동되는 영상은 무대를 비추는 프로젝터로 송출된다. 오퍼레이터이자 연주자는 스크린을 등지고 무대의 중앙에 배치하여 송출되는 영상과 퍼포먼스를 함께 볼 수 있도록 했다.

3. 사운드 및 영상 기술 적용

1) A Section

<표-4> A section에 적용된 사운드 및 영상 기술 적용

section	사운드 프로세싱	영상 변화	사운드와 연동
A	audio to MIDI FM synthesis comb-filter stereo delay		pad + bass
			synth piano
			멜로디 목소리

음악은 입으로 호흡을 뱉는 소리를 내어 입력된 오디오 신호를 변환(audio to MIDI) 하여 FM synthesis로 만든 pad와 bass 악기 소리를 만들며 시작한다. Ableton Live에서 최종으로 출력되는 음량 값과 TouchDesigner의 최

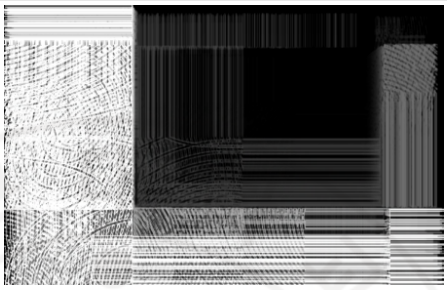
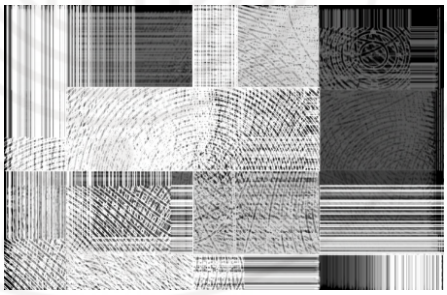
종으로 출력되는 영상의 불투명도(opacity) 값을 OSC로 연동하여 음악이 시작되면 불투명도 값이 1이 되어 영상이 보이고 음악이 끝나면 불투명도 값이 0이 되어 검은 화면이 되도록 했다.

A section의 영상은 pad에 맞추어 생성된 이미지로 인식하게 된 어떠한 존재를 추상적으로 나타내고, bass 리듬에 맞추어 두둥실 떠 있는 모양으로 생동감 있는 모습으로 표현했다. 또다시 오디오 신호 변환으로 연주되는 synth piano에 comb-filter와 stereo delay를 적용하여 날카로운 음색의 소리를 분산시켰다. 오른쪽, 왼쪽으로 움직이는 소리와 MIDI 컨트롤러와 매핑된 comb filter의 주파수 조절을 시각적으로 나타냄과 동시에 존재의 의미가 생성되는 과정을 표현하고자 TouchDesigner에서 이미지의 translate x, y 좌표와 음향효과가 적용된 소리를 연동하여 방향의 움직임을 추가했다.

악기가 쌓이고 ‘Quis vídeor?’의 멜로디가 목소리로 연주되면 점진적으로 추가하여 적용되는 음향효과들에 따라 하나의 원의 모습을 띄던 이미지가 다수로 쌓이며 형태가 조금씩 붕괴되어 나타난다. TouchDesigner에서 이미지의 scale x, y좌표와 목소리의 음량 값을 연동하여 값이 커질수록 이미지의 크기와 형태의 변형도 커지도록 했다. 이는 어떠한 존재가 인식되면 주입되는 생각으로 끊임없이 그 개념과 의미를 생성하고 확장해나가듯 생성된 이미지가 정체하지 않고 입력되는 목소리로 인해 변형되어가는 과정을 나타냈다.

2) B Section

<표-5> B section에 적용된 사운드 및 영상 기술 적용

section	사운드 프로세싱	영상 변화	사운드와 연동
B	audio to MIDI FM synthesis comb-filter stereo delay pitch shift granular synthesis ring modulation		bell pluck
			synth pad
			멜로디 목소리

TouchDesigner와 매핑된 MIDI 컨트롤러를 통해 노이즈 패턴으로 이루어진 B section의 영상이 전환된다. TouchDesigner에서 Noise TOP의 translate x, z 좌표와 synth piano를 연동했다. 짧고 빈도가 많아 리듬감을 형성하는 bell pluck[100]에 맞추어 노이즈 패턴이 빠르게 변화하여 인식된 존

100) 단어의 뜻은 (털을) 뽑다는 의미로 줄을 뜯는 듯이 통통 튀는 소리를 뜻한다.

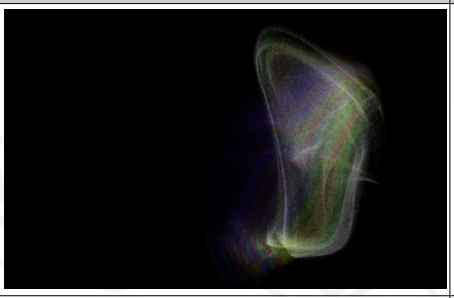
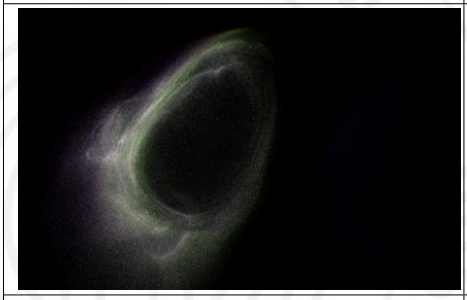
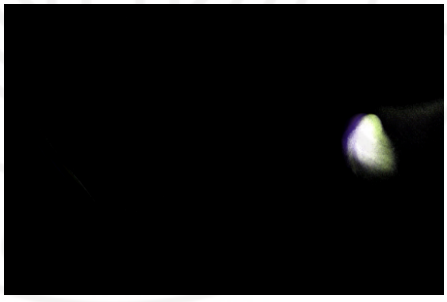
재의 의미가 가중되어 가고 있음을 표현했다. 목소리로 연주되는 멜로디에 오디오 신호 변환을 거친 후 FM synthesis를 이용해 신비스러운 느낌의 synth pad를 추가했다. Ableton Live의 아르페지에이터(arpeggiator)[101]를 적용하여 1Hz의 사인파 LFO를 synced rate[102]에 매핑해 128분 음표부터 4분 음표까지의 범위로 재생하도록 설정했다. 이를 TouchDesigner에서 노이즈 패턴을 픽셀 화하는 quantize step 수와 연동하여 재생되는 음표 범위에 따라 픽셀 수가 랜덤하게 변화하도록 했다. 빠르게 변화하는 노이즈 패턴과 픽셀의 수를 통해 존재에 대한 물음으로 인해 끊임없이 생성해 내는 그 의미에 대한 생각으로 혼란스러움을 표현했다.

멜로디를 연주하는 목소리는 comb-filter, ping pong delay, pitch shift에 granular synthesis가 추가되어 작은 단위로 쪼개어진 목소리를 랜덤하게 배치한다. 분산되는 음향효과에 ring modulation으로 트레몰로 효과를 더해 혼란스러움을 중첩시킨다. 목소리 멜로디의 음량 값은 노이즈 패턴의 밝기와 연동되어 값이 커지면 노이즈 패턴이 선명하게 채워지고 값이 작아지면 사라지듯 줄어든다. 멜로디로 인해 연주되는 존재의 물음에 그 의미가 선명해지고 이내 사라지기를 반복하는 것을 통해 불명료한 존재 의미에 대한 개념으로 인해 겪는 혼란을 표현했다.

101) 코드의 음을 한 번에 연주하지 않고 차례로 연주하는 도구.
102) 재생되는 속도를 설정하는 것으로 ms 단위나 비트 단위로 설정할 수 있다.

3) C Section

<표-6> C section에 적용된 사운드 및 영상 기술 적용

section	사운드 프로세싱	영상 변화	사운드와 연동
C	granular synthesis comb-filter stereo delay pitch shift ring modulation		granular synthesis
			멜로디 목소리
			전체 음악

MIDI 컨트롤러를 통해 10만 개의 점(point)들로 이루어진 C section으로 영상이 전환된다. 분위기가 고조된 음악의 목소리에 granular synthesis를 적용하여 글리치(glitch)[103]한 느낌을 주고, synth piano에 반복적으로 적용하여 일종의 리듬감을 조성해 긴장감을 더했다.

103) 예기치 못한 오류로 생기는 기기적 소리, 소음 또는 이를 음악적 도구로 사용하는 장르.

TouchDesigner에서 불규칙한 숫자 값을 생성하는 Noise CHOP의 채널을 만들고 채널을 9개로 설정하여 9가지 숫자 값을 끊임없이 만들게 한다. 각 채널의 숫자 정보는 필요에 따라 translate x, y, z 축과 rotate x, y, z 축 그리고 scale x, y, z 축에 연동하여 사용한다. synth piano에 적용된 granular synthesis를 연동하여 순간 숫자 정보를 받는 트리거(trigger)[104]역할로 사용해 rotate의 3가지 축과 scale의 3가지 축에 연동했다. synth piano에 적용된 granular synthesis의 음향효과 소리가 날 때마다 점들로 이루어진 모양과 크기가 변화되고 음향효과가 적용된 멜로디 목소리는 translate의 3가지 축과 연동되어 점들로 이루어진 형태의 위치가 변화된다.

모든 악기 연주가 끝을 향해 가고 음향효과가 적용된 멜로디 목소리가 서서히 소멸되면 변화하던 점들의 위치가 모여들게 됨으로 가중된 혼란으로 급변하던 의미가 점차 확립되는 것으로 나타냈다. 음악이 모두 끝이 나면 최종 출력 음량 값과 연동되어 있던 불투명도가 0이 되어 영상도 끝이 난다.

104) 총의 방아쇠를 뜻하는 사격 용어로 어느 특정한 동작에 반응해 다른 필요한 동작을 발생하게 하는 것이다.

4. 연구 기술 적용의 예술적 효과

본 논문은 실시간 사운드 프로세싱 기술을 활용하여 목소리가 가진 예술적 가능성을 탐구하고 이를 인터랙티브 멀티미디어 퍼포먼스에 적용함으로써 새로운 표현 방식을 제시하는 데 목표를 두었다. 목소리는 인간이 가진 가장 본질적이고 일상적인 소리의 매체이지만, 동시에 그 표현의 폭과 가능성은 여전히 무궁무진하다. 본 논문은 이러한 목소리의 잠재력을 기술적 도구와 결합하여 확장하고, 목소리의 다양한 발현 방식이 예술적 감각으로 전환될 수 있음을 보여주고자 했다.

작품의 구성은 목소리의 발현 방식을 중심으로 A(호흡), B(말), C(멜로디)의 구성으로 나누며 '존재 의미'에 대한 서사를 내포하고 있다.

A는 '존재의 생성'을 나타낸다. 숨소리와 속삭임 같은 비일상적 발성법을 활용하여 일반적인 목소리의 표현 방식에서 벗어난 미묘하고 섬세한 음향적 변화를 탐구했다. 숨소리와 속삭임은 단순하지만 내재된 긴장감과 정적을 불러일으킨다. 이를 comb-filter를 사용하여 날카로운 음색을 극대화하고, 블루 컬러 계열의 단조로운 시각적 요소와 결합하여 존재가 감지되고 인식되는 과정을 표현했다.

B는 '존재 의미의 혼란'을 표현했다. 진성을 활용하여 사람의 목소리가 가진 복잡성과 다층적 특성을 강조하며, 이를 시각적으로는 혼란스럽고 정교하게 얽힌 노이즈 패턴으로 형상화했다. granular synthesis와 ping pong delay로 하나의 소리가 여러 음으로 파생되고 이동함으로 단순히 들리는 소리뿐만 아니라 목소리가 가진 서사적 속성과 역동성을 느끼도록 했다. 이를 통해 목소리가 단순한 의사소통 도구를 넘어 감정과 서사의 매개체임을 시사하고자 했다.

C는 '존재 의미의 확립'을 표현했다. 목소리를 통해 감정을 전달하는 강력한 역할을 강조했다. pitch shift와 RM을 통해 음정과 음량의 변화는 감정의

고조와 해소를 나타내며, 이를 시각적으로는 컬러의 변화로 나타냈다. 목소리와 감정의 상관관계를 극대화하여 감정의 흐름에 따라 시청각적 요소가 상호작용하도록 했다.

목소리와 기술이 결합한 새로운 예술 형식을 탐구하여 목소리를 예술적 매체로 재해석하고, 실시간 사운드 프로세싱을 통해 단순한 소리의 변화가 아닌 하나의 예술적 표현의 도구로 새로운 경험과 메시지를 전달하고자 했다. 결과적으로 인터랙티브 멀티미디어 퍼포먼스의 발전 가능성과 예술적, 기술적 발전의 교차점을 끌어올리게 된다. 앞으로 다양한 예술 분야에서 활용될 가능성을 지니며 기존 음악과 퍼포먼스 형식의 한계를 넘어설 수 있는 잠재력을 제시한다.

Ⅳ. 결론

본 연구는 목소리를 이용해 실시간으로 프로세싱 한 소리와 상호작용하는 영상을 제작하여 직접 퍼포먼스 가능한 멀티미디어 작품 제작에 관한 연구이다. 사람의 목소리로 악기 소리를 만들고, 하나의 목소리로 다양한 음색의 소리를 표현하여 주제와 분위기에 맞는 영상과 연동할 뿐만 아니라 연주자가 음향효과와 영상효과를 제어할 수 있도록 시스템을 설계했다.

마이크를 통해 수음되는 목소리에 Ableton Live와 Max for Live를 사용하여 제작된 audio to MIDI, FFT 분석, FM synthesis, RM, granular synthesis, comb-filter, stereo delay, pitch shift 음향효과를 음악 구성과 특성에 맞게 적용했다. 음향효과를 적용함으로 사람이 낼 수 없는 목소리의 음색을 구현하고, 목소리 특성의 한계에 벗어나 다양한 형태의 퍼포먼스에 적용할 수 있는 가능성을 발견했다.

음악과 연동되는 영상은 TouchDesigner를 사용하여 제작했다. 목소리를 통해 연주되는 여러 소리와 음향효과가 적용된 소리들을 TouchDesigner와 OSC 통신으로 연동했다. 원활한 데이터 전송을 통해 MIDI 컨트롤러의 실시간 제어에 맞추어 곧바로 반응하는 소리와 영상으로 사운드 시각화를 느끼게 했다.

본 연구는 인터랙티브 멀티미디어 퍼포먼스 작품으로 실시간으로 진행되는 만큼 예상치 못한 변수가 존재할 가능성이 있다. 연구를 통해 발생한 문제점은 다음과 같다. 음악에 연주되는 모든 소리를 목소리로만 이용하기 위해 마이크로 전달받은 오디오 신호를 통해 작품이 전개된다. 오디오 신호를 숫자 값으로 변환하여 TouchDesigner에서 연산한 후 위치(translate), 회전(rotate), 크기(scale)의 x, y, z 축에 적용한다. 하지만 공연장의 공간 음향 특성과 마이크의 수음 위치에 따라 기존에 연산한 숫자 값보다 큰 폭의 차를 내는 경우 영상의 이미지의 위치나 상태가 의도했던 바와 크게 달라진다.

이러한 문제를 최소화하기 위해 사전에 공간 음향 특성을 분석하는 작업이 필수적이다. 공연 공간의 음향 특성을 미리 측정하고 마이크의 배치와 설정을 최적화함으로써 변수의 영향을 줄일 수 있다. 또한, 오디오 신호 값의 변화 폭을 제한하거나, 다양한 공간에서 사용할 때도 일관적으로 볼륨 데이터를 전달받을 수 있는 기술을 개발하는 방안이 필요하다.

그럼에도 본 작품을 연구하며 목소리 연주의 장점을 부각함과 동시에 기존에 생각한 한계적 특성을 극복하였기에 목소리로 파생된 여러 소리들을 직접 연주하고 실시간 제어가 가능했다. 특히, 목소리라는 인간적인 요소가 기술과 결합되어 관객과의 감성적 교감을 더욱 풍부하게 만들었으며, 몰입감을 높이는데 기여했다. 향후에도 기술과 예술이 상호 보완적으로 진화하며 다양한 창작의 형태와 독창적인 표현 방식이 시도되고 확장되길 기대한다.

Keyword(검색어)
컴퓨터음악(computer music), 소리시각화(sound visualization), TouchDesigner, 실시간 사운드 프로세싱(real-time sound processing), Max, 목소리(voice), 인터랙티브 멀티미디어 음악(interactive multimedia music), Ableton Live

E-mail: bukyngmusic@dgu.ac.kr

참 고 문 헌

1. 단행본

- 루트비히 비트겐슈타인, 이영철 옮김「문화와 가치」(책세상, 2020)

- 발터 벤야민, 반성완 편역「발터 벤야민의 문예이론」(민음사, 2005)

- 배한성, 서혜정, 문선희, 김희선, 조예신, 박형욱, 「성우」(커뮤니케이션북스, 2013)

- 이관규, 「be sound be visionary TouchDesigner Guide」(be sound be visionary, 2023)

- 이선일, “하이데거『존재와 시간』”, 「철학사상」, 별책 제2권 제12호 (2003)

- Gallagher, Mitch. 「The Music Tech Dictionary: A Glossary of Audio-Related Terms and Technologies」, (Course Technology, 2009)

- Holmes, Thom. 「Electronic and Experimental Music, Second Edition」, (Routledge, 2002)

- Manning, Peter. 「Electronic and Computer Music, Revised and Expanded Edition」, (Oxford University Press, 2004)

- Manzo, V. J. 「Max/MSP/Jitter for Music: A Practical Guide to Developing Interactive Music Systems for Education and More」, (Oxford University Press, 2011)

- White, G;eem D & Louie, Gary J.「The Audio Dictionary: Third Edition, Revised and Expanded」, (University of Washington Press, 2005)

- Vaughan, Tay.「Multimedia: Making It Work (Sixth Edition)」, (McGraw-Hill, 2003)

- Yeang, Chen-Pang.「Transforming Noise: A History of Its Science and Technology from Disturbing Sounds to Informational Errors, 1900-1955.」(Oxford University Press, 2024)

2. 참고논문

- 김봉영 & 이은영 & 배명진, 「목소리 특성 구분에 대한 사운드 A칼라에 관한 연구」, (예술인문사회 융합 멀티미디어 논문지, 제2권 제8호, 2018)

- 김진주, 「목소리와 기타의 실시간 사운드 프로세싱을 이용한 인터랙티브 멀티미디어 퍼포먼스 제작 연구 (멀티미디어 작품 <Do not fear>를 중심으로)」(동국대학교 영상대학원 멀티미디어학과, 2022)

- 이관규, 「피아노 연주와 무용수의 움직임을 이용한 인터랙티브 멀티미디어 작품 제작 연구」(동국대학교 영상대학원 멀티미디어학과, 2021)

- 이창영, 「남성과 여성의 음성 특징 비교 및 성별 음성인식에 의한 인식 성능의 향상」, (한국전자통신학회 논문지, 제5권 제6호, 2010)

- 오이링, 「베이스기타의 실시간 사운드 프로세싱을 이용한 인터랙티브 멀티미디어 작품 제작 연구 -멀티미디어음악 작품 <Dive>를 중심으로-」 (동국대학교 영상대학원 멀티미디어학과, 2024)

- 정현이 & 김형기. 「디지털 기술을 활용한 공연 예술의 특성 -움직임을 이용한 인터랙티브 퍼포먼스를 중심으로-」, (디자인융복합연구(구.인포디자인이슈), 제12권 제1호, 2013)

- 조환희, 「베이스 트롬본과 피아노의 실시간 사운드 프로세싱을 이용한 멀티미디어작품 제작 연구 (멀티미디어음악 작품 <Yes, I am.>을 중심으로」(동국대학교 영상대학원 멀티미디어학과, 2019)

- Torun, Tuna. 「An Approach on Determining Optimum Acoustic Conditions for Turkish Classical Music」(The Graduate School of Natural and Applied Sciences of Middle East Technical University, 2019)

- Williamson, John & Murray-Smith, 「Roderick. Audio Feedback for Gesture Recognition」 (Department of Computing Science, University of Glasgow, 2003)

3. 웹사이트

Ableton Live: https://www.ableton.com/

CCRMA: https://ccrma.stanford.edu/

Max: https://cycling74.com/

TouchDesigner: https://derivative.ca/

ABSTRACT

A Study on Interactive Multimedia Performance using Real-time Sound Processing of Voice

Kim, Bu Kyoung

Department of Multimedia
Graduate School of Digital Image and Contents
Dongguk University

This study focuses on exploring the artistic potential of the human voice using real-time sound processing technology and applying it to interactive multimedia performance to propose new modes of expression. Considering that the voice is the most intrinsic sound medium that humans possess and that its range of expression is infinite, this thesis aims to demonstrate that the various ways in which the voice can be

expressed can be expanded in an artistic sense by combining it with technological tools.

The composition of the work is divided into three sections, A (breath), B (speech) and C (melody), which focus on the expression of the voice. Each section contains a narrative that explores 'being' and its meaning. The first section, A, uses breath and whispering to explore the creation of being. Technically, a comb-filter is used to maximise the sharp timbre of these sounds, combined with monotonous visuals in a blue colour scheme to represent the process of being perceived. Section B uses real and granular synthesis and ping-pong delay to emphasise the complexity and multi-layered nature of the voice, which is then translated into chaotic visual noise patterns. Section C emphasises the role of the voice in conveying emotion, using pitch shifts and RM to express the rise and fall of emotion, and colour changes to maximise audio-visual interaction.

This study focused on the design and implementation of a physical piece to realise a new art form combining voice and technology. Using Ableton Live and Max for Live, the voice was recorded through a microphone and various sound effects such as FFT analysis, audio to MIDI conversion and granular synthesis were applied to create timbres that would not be possible with a conventional voice. This proves that the voice can evolve from a tool for communication to an artistic medium. This sound transformation process was combined with video production using TouchDesigner, and the voice and video were designed to interact in real time via OSC communication. A control system using MIDI controllers allowed the simultaneous manipulation of sound and image to maximise immersion.

This paper presents new artistic possibilities for combining real-time sound processing technology with the voice to push the boundaries of traditional performance forms. The human and emotional medium of the voice plays an important role in strengthening the connection with the audience and enhancing the sense of immersion by combining it with technology. This research has demonstrated the potential of interactive multimedia performance and the expansion of creative methods that combine technology and art.

In particular, this study has provided new tools and ideas for creators by implementing a real-time performance system centred on the voice, and has the potential to be practically applied in various artistic fields. Real-time processing technology has shown that it can expand the scope of artistic experience, and this approach will enable more original and innovative creative forms as technology advances in the future.

부록 : 첨부 DVD

Adsum (I am present) : 2024년 11월 9일 이해랑 예술극장 공연 영상

Adsum_Max : 작품에 사용된 Max 패치 폴더