

실시간 Slit-scanning과 Spectral Delay를 활용한 인터랙티브 설치작품 연구

최 홍 찬, 김 준

동국대학교 영상대학원 멀티미디어학과 컴퓨터음악전공

A Study on an Interactive Installation Art Using Real-time Slit-scanning with Spectral Delay

Hongchan Choi, Jun Kim

Computer Music Lab., Department of Multimedia,
Graduate School of Image and Contents
Dongguk University

요 약

특수한 사진촬영방법 중 하나인 슬릿-스캐닝(slit-scanning) 기법을 컴퓨터를 통해 실시간으로 구현하고, OpenGL에 기반을 둔 3D 그래픽 및 스펙트럴 딜레이(Spectral Delay)를 사용한 오디오-비주얼(Audio-visual) 작품을 제작한다. 이 작품은 카메라로 입력된 관객 자신의 영상에 시간적 왜곡을 부여하는 효과를 중심으로 한다. 또한 관객이 내는 소리의 음량과 음고에 따라 영상의 변화를 일으키고, 작품으로 입력된 소리를 조작하여 새로운 형태로 바꾸어 출력함으로써 시청각적인 재미를 느낄 수 있도록 한다. 본 논문에서는 오디오-비주얼 인터랙티브 설치작품의 제작 과정에서 요구된 기술적 문제들을 해결한 방법들을 제시한다.

I. 서 론

인류의 역사를 통틀어 상호작용성(interactivity)을 갖춘 가장 오래된 미디어를 꼽으라면 단연코 거울일 것이다. 거울은 자신의 모습을 비추어 볼 수 있으며, 자신의 행동에 따라 거울 속에 비추어진 상의 모습도 항상 함께 움직인다. 이러한 원시적 상호작용성을 갖춘 거울 만큼 오래된 것은 미디어는 없을 것이다.

본 논문의 주제가 되는 작품 <Mirror II>는 이 거울의 원시적 상호작용성에 기반을 둔다. 동시에 그것을 배반함으로써 예술적인 재미를 얻고자 하였다. 자신의 모습을 볼 수 있다는 점에 있어서는 같으나, 그 모습에 현재와 과거가 뒤섞여 있다는 것을 느끼게 되면 관객은 시각적 혼동을 일으키며, 그것에 대해 본능적으로 신기하게 느낄

수밖에 없다. 가장 친숙한 미디어가 자신의 예측을 배반한다는 것만으로 예술적 재미의 소재가 되는 것이다.

이 작품의 중심 소재는 슬릿-스캐닝(slit-scanning)이라는 시각적 효과이다. 슬릿-스캐닝이란 빛이 통과할 수 있는 작은 틈과 움직일 수 있는 슬라이드를 가진 특수한 사진기를 이용하여, 슬라이드에 필름을 부착한 후 한 방향으로 천천히 움직이며 작은 틈으로 들어오는 빛에 필름을 노출시키는 촬영기법이다. [1] 본 작품에서는 이 슬릿-스캐닝 기법이 실시간으로 구현되어있다.

또한, 입력된 음성의 음고 및 음량에 따라 영상을 변화시키고, 입력된 음성을 다수의 주파수 대역으로 분리하여 딜레이 효과를 부여한 뒤 재결합하여 출력하는 스펙트럴 딜레이(Spectral Delay)를 사용 하였다.

영상 부분은 Jitter, 음향 부분은 MSP를 이용하여 프로 그래밍 되었는데, 이와 같이 본 작품은 <Cycling '74>의 멀티미디어 프로그래밍 환경인 Max/MSP/Jitter에 기술적 기반을 두고 있다.

그림 1은 본 논문에서 제시한 기술적 요소들을 설치

작품형태로 구성하였을 때의 구성도이다.

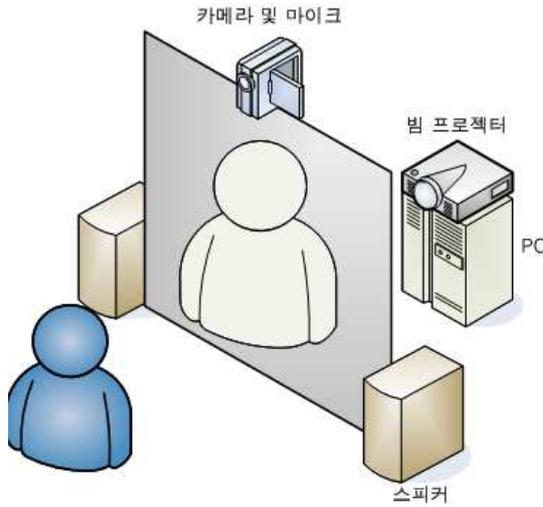


그림 1. 설치작품 구성도

Fig. 1. A Block diagram of the installation artwork

II. 실시간 Slit-scanning 기법

2.1 슬릿-스캐닝(Slit-scanning)

본 작품에서 중요한 부분을 차지하는 시각적 혼동의 근원은 슬릿-스캐닝이라는 전통적 사진 촬영기법이다. 일반적으로 사진(still image)이라는 미디어는 시간적 개념이 없으나, 사진 촬영 과정에 시간의 흐름을 강제적으로 적용함으로써 현실과는 다르게 왜곡된 현상을 얻게 된다.

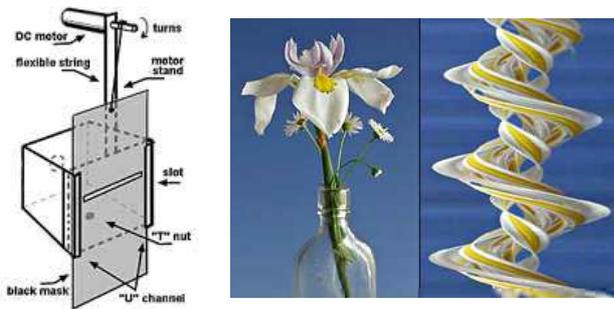


그림 2. 슬릿-스캐닝 장치 및 촬영 결과 [1][2]

Fig. 2. The slit-scanning device and result pictures

이 기법을 사용하여 작품을 제작하기 시작한 시기는 확실치 않으나 1960년대 정도로 추정해 볼 수 있다. [2] 또한, 이 기법을 사용하여 영상들 중 가장 널리 알려진 것은 ILM(Industrial Light and Magic)에 의해 제작된 <Star Trek: The Next Generation>이라는 TV 시리즈물에서 우주선 공간을 뛰어 넘는 특수효과 장면이다.[3]

본 작품에서 사용된 기법은 컴퓨터를 이용한 실시간 영상 처리를 기반으로 구현한 것으로, 전통적인 슬릿-스캐닝을 응용하여 다른 방식으로 영상 효과를 보여준다.

2.2 Jitter를 이용한 실시간 슬릿-스캐닝의 설계

본 작품에서 영상의 입출력을 위해 사용되는 가장 중요한 수단은 카메라와 빔 프로젝터이다. 앞서 소개한 바 있는 Jitter는 Max를 기반으로 하여 실시간 영상처리를 위한 다양한 기능을 제공하며, Max는 일반적인 PC에서 사용할 수 있기에 카메라 및 빔 프로젝터와 쉽게 연결할 수 있다.

본 작품의 핵심이라 할 수 있는 슬릿-스캐닝 기법을 Jitter를 사용하여 구현하는 것이 매우 중요한 작업이었으며, 효율적인 가변 스캔라인(variable scan-line)의 구현을 위해서 일반적인 지터 오브젝트를 사용한 GUI 프로그래밍을 최대한 배제하고, Java Script를 사용하여 프로그래밍 하였다.

그림 3은 본 작품의 영상처리 과정을 나타낸 구성도이다.

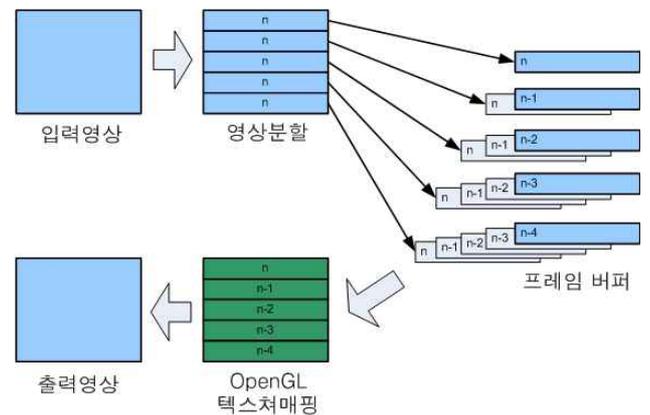


그림 3. 슬릿-스캐닝 효과를 위한 영상처리과정

Fig. 3. Video processing for slit-scanning effect

입력되는 영상 프레임(frame)을 2개 이상의 영역으로 분할한 후, 분할된 영상 프레임들(n번째 프레임)을 각각 다른 크기를 가진 버퍼에 저장한다. 서로 다른 크기를 가진 버퍼들에 저장되었던 프레임들을 순차적으로 읽어 들여 결합시키면, 가장 최근에 촬영된 프레임(n번째 프레임)부터 가장 긴 버퍼에 저장되었던 오래전에 촬영된 프레임(n-4번째 프레임)까지 한 화면에 동시에 나타나게 된다.

이렇게 분할되었던 영상을 그대로 결합하지 않고 OpenGL¹⁾을 사용하여 생성한 평면 모형의 개체에 텍스처 매핑(texture mapping)²⁾을 통해 투영(project)한다. 투영된 영상은 OpenGL의 개체에 대한 다양한 조작(크기, 회전각, 위치, 시점 변환 등)을 통해 다채로운 변화가 가능하게 된다.

1) 2D와 3D를 정의한 컴퓨터 산업 표준 응용 프로그램 인터페이스(API)

2) 컴퓨터로 생성된 그래픽이나 3차원 개체에 세부 묘사 및 질감, 색 등을 부여하는 방법

III. 상호작용을 위한 음향 분석

3.1 음향적 조건 및 MSP를 이용한 음향 분석

본 작품은 공개된 장소에 설치된다는 점으로부터 여러 가지 제약 조건을 지니게 된다. 특히, 공개된 장소의 음향 환경은 의미 있는 소리와 무의미한 잡음이 뒤섞여 있는 상태로, 작품의 올바른 상호작용을 위해 의미 있는 소리만을 취사선택할 수 있어야 한다.

MSP는 Max를 기반으로 하여 실시간 음향처리를 위한 다양한 기능을 제공한다. 하지만 작품의 구현에 있어서 필요한 기능인 음고의 감지, 음량의 최저 및 최고점 제한, 어택(attack)의 검출 등은 자체적으로 제공되지 않기 때문에 외부 오브젝트(external object)인 <fiddle~> 을 사용하였다. 이를 사용하여 입력되는 소리를 다양한 각도로 분석하여 여러 가지 결과 값을 얻을 수 있다.

또한 이 오브젝트에 특정 메시지를 입력하여 조건을 설정하면, 처리할 음량의 범위, 추출할 음고의 개수, 어택의 검출 등 얻고자 하는 분석결과를 원하는 형태로 제한할 수 있다. [4]

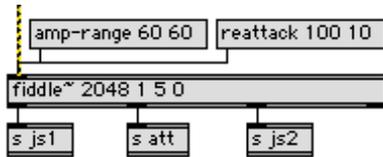


그림 4. 외부오브젝트 <fiddle~>의 사용
Fig. 4. Using a External Object <fiddle~>

본 작품에서는 설치 환경의 주변 잡음을 고려하여 음고를 인식할 대상을 60dB 이상으로 제한하였다. 지나치게 작은 음량 즉, 주변 환경 잡음에 시스템이 일일이 반응할 경우 관객과의 상호작용이 무의미해지기 쉽기 때문이다. 또한, 어택의 재검출 시간 범위는 100ms로 제한하였는데, 이는 어택이 검출된 후 100ms내에 또 다른 어택이 검출된다면 두 개의 어택을 동일한 음악적 사건으로 간주하겠다는 의미이다.

3.2 분석 정보와 OpenGL 영상 효과의 관계 성립

2.2절에서 언급한 바와 같이, 화면의 상하로 평면들이 나란히 배열되어 있고 이들은 입력되는 소리의 음고에 따라 움직이게 된다. 음고가 낮은 음에는 아랫부분에 있는 평면들이 반응하며, 높은 음에는 윗부분의 것들이 반응한다. (그림 5.)

자극에 대한 평면들의 반응은 회전의 형태로 나타나며, 입력되는 음량에 따라 평면의 회전 속도가 달라진다. 또한, 평면에 적용된 회전은 스크립트 프로그래밍을 통해 가속도와 관성을 고려해 구현된 것으로, 평면에 대한 자극이 중단되더라도 평면의 회전은 마치 바람개비처럼 서서히 멈추게 된다. 이로 인해 동작 또는 멈춤의 단순한 움직임보다 더욱 다양한 모습을 보여주게 된다.

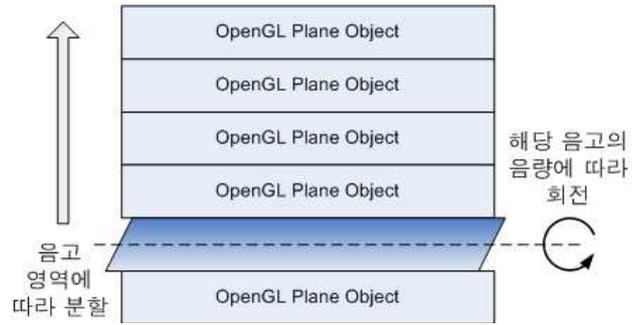


그림 5. 음고와 음량, 3D 그래픽의 관계 설정
Fig. 5. Mapping parameters; Pitch, Amplitude and 3D Graphics

IV. 스펙트럴 딜레이(Spectral Delay)

4.1 주파수 분할 및 재결합을 이용한 음향 처리

일반적으로 딜레이 효과는 입력되는 신호의 주파수 영역 전체에 동일하게 적용된다. 그러나 본 작품에서 사용된 딜레이 효과는 입력 신호를 FFT 분석을 통하여 512 개의 주파수 단편(bin)으로 분할 한 뒤, 각각의 단편에서 서로 다른 지연시간과 피드백을 가진 딜레이 효과를 적용시킨다. 그리고 분할되었던 신호를 다시 결합하여 최종 출력을 만들어 낸다.

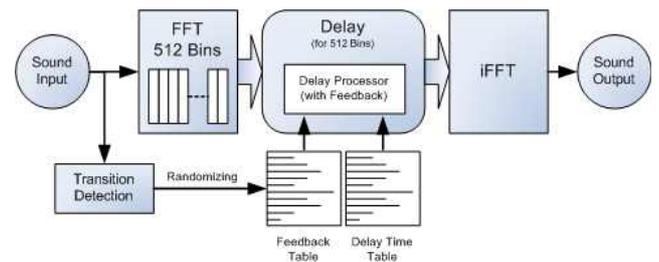


그림 6. 스펙트럴 딜레이의 신호 흐름도
Fig. 6. A Signal Flow Diagram for Spectral Delay

이 과정은 MSP가 제공하는 <pFFT~> 오브젝트 및 다양한 푸리에 주파수 변환(FFT) 관련 오브젝트들의 조합으로 가능하며, 디지털 신호 처리에 대한 깊은 지식이 없는 음악인들도 쉽게 프로그래밍이 가능 하다.

V. 결과 및 시연

인터랙티브 멀티미디어 작품을 제작할 때 가장 중요한 것은 다양한 변수들 간의 관계 설정(mapping)이다. 이렇게 성립된 관계 아래 발생하는 작품과 청자 간의 상호작용의 질에 따라 작품의 예술적 수준이 결정된다는 것을 보여준 다수의 사례가 그것을 증명한다. [7]



그림 7. <Mirror II> 작품의 실제 시연
Fig. 7. A Demonstration of <Mirror II>

작품의 기저에 깔린 원시적 상호작용성의 응용이라는 측면에 있어서는 분명히 성공했다고 할 수 있으나, 실시간 슬릿-스캐닝이 주는 시각적 재미에 많은 부분 의존하고 있음을 부인할 수는 없다.

또한, 작품의 전반에서 발견되는 변수 간의 일대일 대응형식의 관계 설정은 보다 다차원, 고차원적인 것으로 발전되어야 하며, 사용자와 시스템 간의 상호작용뿐만 아니라, 시스템 내 미디어 간의 상호작용 역시 고려되어야 할 것이다.

V. 문제점 및 개선방안

<Mirror II>는 MacBook, MacBook Pro³⁾를 이용하여 제작되었으며, 제작 과정에 사용된 수음장치 및 영상입력장치 역시 MacBook 제품군에 탑재된 기본 장치들을 사용하였다. 고성능 워크스테이션이 아닌 작업 환경에서도 수급할 수 있을 만한 성능을 보여주어야, 공개된 장소에 저렴한 구성의 시스템으로 설치가 가능하기 때문이다. 또한, 설치작품은 그 특성상 매우 많은 변수가 존재하고, 그들을 완전히 예측한다는 것은 불가능하다. 그렇기 때문에, 범용성과 안정성을 위해 예술적, 기술적으로 포기해야 하는 부분도 있기 마련이다.

5.1 문제점 및 개선방안

첫째로, 가장 큰 기술적 문제점은 저가의 보급형 카메라에서 발견되었다. 웹-캠(web-cam)이라 불리는 소형 카메라들은 자동 백색 균형(auto white balance) 기능을 내장하고 있다. 따라서 피사체의 색상과 거리, 주변의 밝기 등에 따라 영상의 전체적인 색조가 지속적으로 변화하는 문제를 낳는다.

MacBook에 탑재된 카메라에는 자동 백색 균형 기능을 비활성화 할 수 있는 방법이 없으므로, 백색 균형 및 초

점을 수동으로 설정할 수 있는 카메라를 별도로 장착함으로써 이 문제를 해결할 수 있다.

둘째로, 마이크와 스피커가 한 장소에 설치되어야 하는 설치 작품의 특성상 피드백 문제가 발생할 수 있다. 하지만, 지향성이 강한 마이크를 사용하여 수음하고, 스피커의 위치와 방향을 적당히 조절함으로써 피드백의 양을 억제할 수 있다.

셋째로, 작품의 영상은 입력되는 소리의 음고에 반응하는데, 남녀 간의 음성의 음역대가 상이함으로 인해 영상의 반응에 제한이 발생한다. 영상이 반응하는 음고의 범위는 남성 테너(tenor)의 일반적 음역대인 A2(110Hz)~A4(440Hz) 정도로 설정되어 있다. 이러한 이유로 소프라노의 음역대(C4~C6)를 가진 사람이 가장 낮은 목소리로 발생할 경우, 중간 정도에 위치한 평면이 회전하게 된다.

이 문제의 해결을 위해 남성 및 여성의 음성을 분석하여 얻어낸 평균 음고 주파수(average pitch frequency)를 통해 성별을 인식할 수 있는 기능을 추가해야 한다. 뿐만 아니라 아동의 음성도 성인의 것과는 다른 음역대를 갖고 있으므로 이에 대한 대처가 필요할 것으로 보인다. [6]

넷째로, Max의 GUI를 사용한 프로그래밍보다 스크립트 프로그래밍을 통해 시스템을 제작하는 것이 그 효율 면에서 여러모로 우수하지만, 실시간 슬릿-스캐닝은 여전히 큰 부하를 가진 작업이다. 스캔라인의 수가 늘어날수록 시스템의 부하가 증가하여 영상의 초당 프레임 수(frame per second)가 감소하게 되며 영상미를 떨어뜨리는 결과를 낳는다.

이 문제는 Max/MSP/Jitter 기반으로 프로그래밍된 알고리즘을 바탕으로 C언어 등으로 이식하여 독립적인 소프트웨어를 개발함으로써 다소 해결할 수 있을 것이다.

VI. 참고문헌

- [1] A. Davidhazy, *Slit-scan Photography*.
<http://www.rit.edu/~andpph/text-slit-scan.html>
- [2] G. Levin, *An Informal Catalogue of Slit-Scan Video Artworks*.
http://www.flong.com/writings/lists/list_slit_scan.html,
2005-2007
- [3] Wikipedia, *Slit-scan Photography*.
<http://en.wikipedia.org/wiki/Slit-scan>
- [4] M. S. Puckette, T. Apel, D. D. Zicarelli, "Real-time audio analysis tools for Pd and MSP." Proceedings, ICMC 1998.
- [6] W. H. Abdulla, N. K. Kasabov, "Improving speech recognition performance through gender separation." 2001.
- [7] S. Wilson, *Information Arts: Intersections of Arts, Science, and Technology*. MIT Press, 2002.

3) Intel Core Duo CPU를 장착한 매킨토시 노트북 제품군으로, 2006년 Apple Computer Inc.에 의해 출시