



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](#)

공학석사 학위논문

터치 슬라이더를 이용한
효과음 검색 인터페이스 연구

Audio Search Interface Using
Touch Sensitive Sliders

2018 년 02 월

서울대학교 융합과학기술대학원

융합과학부 디지털정보융합전공

김 완 수

터치 슬라이더를 이용한 효과음 검색 인터페이스 연구

Audio Search Interface Using
Touch Sensitive Sliders

지도교수 이 교 구
이 논문을 석사 학위논문으로 제출함

2018 년 2 월

서울대학교 융합과학기술대학원
융합과학부 디지털정보융합전공
김 완 수

김완수의 석사 학위논문을 인준함

2018 년 2 월

위 원 장	_____	서 봉 원	(인)
부 위 원 장	_____	이 교 구	(인)
위 원	_____	권 가 진	(인)

국문초록

영화 음향 믹싱 엔지니어는 자신이 변형하려고 하는 효과음을 찾는데 많은 시간을 사용한다. 이는 제한된 시간에 콘텐츠를 만들어야 하는 현재의 영화 제작 구조상 음향 믹싱 엔지니어를 제약한다. 이를 해결하기 위해 음향 믹싱 과정에서 사용하는 음향 콘솔에 트랙별로 음원을 더 잘 시각화해주는 발전이 있었지만 매우 많은 수의 트랙을 사용하는 영화 음향 믹싱 엔지니어를 돕기에는 한계가 있다.

한편, 영화 음향 믹싱 엔지니어를 돕기 위한 연구는 아니지만 목소리, 태핑, 몸 움직임 등 각종 질의(query)를 통한 효과음 검색 연구가 진행되어 왔다. 또한 효과음에서 추출한 소리 크기 변화 곡선, 음 높이 변화 곡선 등과 같은 음향 특성과 사람에게 효과음을 들려주며 몸의 움직임을 사용해서 표현하도록 지시해서 얻은 움직임 데이터를 비교해서 그 연관성을 찾기 위한 소리 투사 연구도 이루어지고 있다.

본 연구는 터치 슬라이더를 이용한 소리 투사 연구를 진행하고 이를 활용해서 영화 음향 믹싱 엔지니어들이 효과음 검색에 사용 가능한 인터페이스를 개발하고 그 사용성을 평가하는 것을 목표로 한다.

주요어 :

터치 슬라이더, 음향 인터페이스, 효과음 검색, 소리 투사, 음향 믹싱 엔지니어, 음향 믹싱 콘솔

학 번 : 2016 - 26030

목 차

제 1 장 서론	1
제 1 절 연구 배경	1
제 2 절 연구 목표	8
제 2 장 선행 연구 및 관련 개념	9
제 1 절 음향 콘솔에서 소리 탐색 연구	9
제 2 절 질의 기반 소리 검색 연구	15
제 3 절 소리 투사 연구	17
제 4 절 터치 슬라이더	19
제 3 장 연구 방법 및 제안 특성	21
제 1 절 연구 문제	21
제 2 절 제안 특성	23
3.2.1 제안 음향 특성	23
3.2.2 제안 움직임 특성	25
제 3 절 소리 투사 실험	26
3.3.1 효과음 데이터	26
3.3.2 소리 투사 실험	29
3.3.3 소리 투사 데이터 처리	31
제 4 절 소리 투사 모델	32
제 4 장 제안 인터페이스	35
제 1 절 인터페이스 개요	35

4.1.1 검색 인터페이스 개요	35
4.1.2 검색 알고리즘 개요	37
제 2 절 인터페이스 평가	39
4.2.1 인터페이스 평가 방법	39
4.2.2 정확도 평가 결과	46
4.2.3 사용성 평가 결과	47
제 5 장 연구 의의 및 논의	56
참고문헌	59
Abstract	62

표 목 차

[표 1] 사전 실험에 사용된 효과음 목록	27, 28
[표 2] 절대값 예측 모델 특성별 유의도	32
[표 3] 1차 미분값 예측 모델 특성별 유의도	32
[표 4] 2차 미분값 예측 모델 특성별 유의도	33
[표 5] 인터뷰 질문	44
[표 6] 평가 대상	45
[표 7] 항목별 인터뷰 결과	53

그 림 목 차

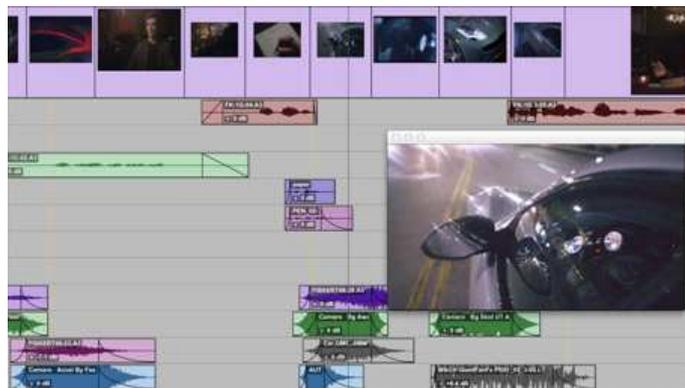
[그림 1] 디지털 오디오 편집 소프트웨어	1
[그림 2] 믹싱 큐시트	4
[그림 3] 현재 음향 믹싱 엔지니어의 작업 환경	4
[그림 4] 음향 믹싱 소프트웨어의 믹싱 화면	5
[그림 5] 채널 줄	5
[그림 6] 채널별 트랙 이름 표시	9
[그림 7] 실시간 파형 표시	9
[그림 8] 파형의 색상으로 주파수 정보 제공	10
[그림 9] 미니맵으로 트랙별 주파수 성분 표현	11

[그림 10] 시간축을 다시 스케일한 파형	12
[그림 11] 태핑을 질의로 사용한 음악 검색	13
[그림 12] 그림을 활용한 효과음 검색 인터페이스	15
[그림 13] 양 팔의 움직임을 사용한 소리 표현	16
[그림 14] 팔 움직임을 사용한 효과음 검색 연구	16
[그림 15] Carmiaux 실험 환경	18
[그림 16] 터치 슬라이더	19
[그림 17] 터치 스크린의 터치 슬라이더	20
[그림 18] BOURNS 사의 PSM Series 터치 슬라이더	30
[그림 19] 소리 투사 실험 환경	30
[그림 20] 효과음 검색 인터페이스 개요 - 1	35
[그림 21] 효과음 검색 인터페이스 개요 - 2	36
[그림 22] 제안 효과음 검색 인터페이스 UI	40
[그림 23] 대조군 효과음 검색 인터페이스 UI	40
[그림 24] 정확도 평가 결과를 나타낸 히스토그램	46
[그림 25] 평균 효과음 검색 시간 그래프	47
[그림 26] 신 내 검색 횟수와 평균 검색 시간 그래프	48
[그림 27] 효과음 분류별 평균 효과음 검색 시간	49

제 1 장 서론

제 1 절 연구 배경

전통적인 영화 음향은 영화가 촬영되는 현장에서 직접 녹음한 소리를 극장에서 그대로 재생하는 것에서 시작되었지만 현대 영화 음향은 각종 효과음 라이브러리들에 미리 녹음되어 있는 소리를 조합하는 방식으로 변화했다. 전통적인 영화 음향 제작 과정에서는 영화의 한 신(scene)에 사용되는 모든 소리는 같은 장소에서 같은 시간에 녹음된 소리로 구성되어 있었다. 이후 아날로그 녹음과 편집 기술의 발달로 제한된 횟수 안에서 아날로그 음향 신호를 복사하고 다른 아날로그 녹음 위에 덮어 쓰는 것이 가능해지면서 촬영 현장에서 녹음된 소리에 다른 시간과 장소에서 녹음된 소리들을 추가할 수 있게 되었다. 현재는 [그림 1]과 같은 디지털 오디오 편집 소프트웨어(digital audio workstation, DAW)와 디지털 효과음 라이브러리들의 발달로 영화 음향에 제한 없는 수의 효과음을 추가해서 편집하는 것이 가능해졌다.



[그림 1] 디지털 오디오 편집 소프트웨어 (AVID protocols)

상업 영화 시장의 세계화로 영화 촬영 현장에서 직접 녹음한 소리는 점점 사용하기 어려워지고, 배우들의 대사를 제외한 영화 음향의 대부분은 현장에서 녹음한 소리가 아닌 효과음 라이브러리에서 녹음된 소리들로 대체되었다. 한 언어로 촬영된 영화가 다른 언어를 쓰는 국가로 수출되는 과정에서 영화의 원본 음향 콘텐츠에 있던 배우들의 대사를 수출 대상국의 목소리 배우들이 다시 녹음하는 더빙(dubbing) 과정을 거치는데, 영화 촬영 현장에서 녹음된 소리들은 배우들의 목소리와 분리할 수 없다. 따라서 다양한 언어로 상영될 것을 전제하고 제작되는 현대 영화의 경우 영화 촬영 현장에서 녹음된 소리들을 사용하지 않고 거의 모든 소리를 따로 녹음된 효과음들로 제작하게 된다.

아주 정적인 영화 장면을 위한 음향 콘텐츠를 만드는 것에도 수많은 효과음들이 필요하며, 정교한 입체 음향 기술의 발달과 함께 점점 더 많은 수의 효과음들을 사용하게 되었다. 사무실에서 업무를 보는 단순한 장면에도 형광등에서 나는 소음, 자판 두드리는 소리, 프린터 동작 소리, 컴퓨터에서 발생하는 소음 등 다양한 소리가 발생한다. 만약 촬영 장면에 움직이는 사람들이 있다면 이들이 각각 움직이며 내는 옷 소리, 발걸음 소리, 의자에서 발생하는 소리 등을 추가해야 한다. 필요한 효과음의 수는 영화 장면이 조금만 복잡해지면 급격하게 늘어나서 로봇들이 전투를 하는 장면과 같은 경우에는 한 신에 수천 개의 효과음을 사용한다. 또한 Dolby atmos, dts X, auro 3D 등 입체 음향 기술들의 경쟁적인 발달과 보급으로 영화 음향 콘텐츠들이 정교해지면서 점점 더 많은 수의 효과음을 사용하는 추세이다.

여러 장소에서 여러 방식으로 녹음된 효과음들이 자연스럽게 영화의 한 장면에서 같이 사용되기 위해서는 이들을 비슷한 장소에서 한 종류의 마이크로 녹음된 것처럼 들리도록 만드는 영화 음향 믹싱(film sound mixing) 과정이 필요하다. 넓은 운동장에서 박수를 치는 소리와 동굴 속에서 박수를 치는 소리가 다른 것처럼 같은 방법에서 발생한 소리라도 장소에 따라 다르게 들린다. 특히 녹음된 소리의 경우 녹음에 사용된 마이크의 종류와 마이크의 녹음 위치 등에 따라 매우 다른 소리가 녹음된다.

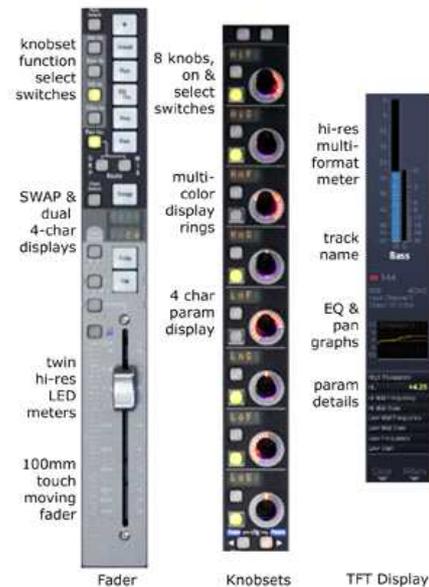
영화 음향 믹싱은 영화의 한 장면에 등장하는 소리들의 크기, 주파수 성분, 동적 변화, 위치, 잔향, 반향 등을 조작해서 영상에 알맞은 자연스럽고 창의적인 콘텐츠를 만드는 과정이다. 음향 믹싱 엔지니어는 음향 편집자가 편집한 결과를 바탕으로 작업을 시작해서 조작성이 필요한 효과음을 찾고 수정하는 작업을 반복적으로 수행한다.

음향 믹싱 엔지니어는 조작성이 필요하다고 생각되는 효과음이 있으면 여러 효과음 트랙들 사이에서 해당 효과음이 어느 트랙에 위치해 있는지 찾아야 한다. 한 영화에서 20개 내외의 효과음 트랙을 사용했던 과거 아날로그 음향 믹싱 과정에서는 영화에 사용된 모든 소리를 시간별로 나타낸 표인 믹싱 큐시트(mixing cuesheet)를 만들어서 사용했었다. [그림 2]의 믹싱 큐시트에는 트랙 이름들이 나타나 있고 각각의 트랙별로 어떤 효과음이 언제 시작해서 언제 끝나는지를 확인할 수 있도록 표현되어 있다.



[그림 4] 음향 믹싱 소프트웨어의 믹싱 화면

[그림 3]의 가운데에 있는 장치는 디지털 음향 믹싱 콘솔(digital mixing console)이다. 디지털 음향 믹싱 콘솔의 한 세로 줄 채널 줄(channel strip)은 디지털 오디오 편집 소프트웨어 상의 한 음향 트랙에 대응하며 이 세로 줄 상에 있는 버튼들과 하단에 위치한 터치 슬라이더를 사용해서 해당 트랙에 있는 음원의 소리를 변화시키는 기능을 수행한다. [그림 5]는 채널 줄 하나를 확대한 그림이다.



[그림 5] 채널 줄 (AVID System 5-MC)

현대 상업 영화의 음향 콘텐츠 제작에는 일반적으로 200개 이상의 트랙이 사용되며 단순한 장면에서도 20개 내외의 효과음이 동시에 재생된다. 이렇게 많은 효과음과 음향 트랙은 음향 믹싱 엔지니어들이 자신이 변화시키고자 하는 효과음이 어느 트랙에 있는지 찾는 과정을 어렵게 만든다. 숙련된 음향 믹싱 엔지니어에게도 시간과 노력이 많이 필요한 과정이며 영화 음향 믹싱을 가르치는 교육 기관에서는 교육 과정에 이를 훈련하는 것이 포함되어 있을 정도로 중요한 문제이다.

음향 믹싱 엔지니어들이 자신이 변화시키고 싶은 효과음을 찾는 기본 과정은 아래와 같다.

1. 찾고자 하는 효과음을 포함하는 5 - 10초 길이의 구간을 설정하고 이를 반복 재생한다.
2. 찾고자 하는 효과음이 있을 것으로 예상되는 트랙 묶음을 음향 콘솔의 중앙으로 불러온다.
3. 음향 콘솔의 트랙별 시각화 정보와 음향 줄 별 음소거, 독주 버튼을 활용해서 찾고자 하는 효과음이 위치한 트랙을 찾는다.

기존 음향 믹싱 엔지니어들은 자신이 찾고자 하는 소리가 어느 트랙에 위치해 있는지 찾는 과정을 돕기 위해 다양한 방법을 사용한다. 가장 기본적인 방법은 음향 스튜디오 내부의 자체적인 규칙을 통해 트랙을 묶음으로 만드는 것이다. 효과음의 주파수 성분에 따라 고음부 효과음 / 중음부 효과음 / 저음부 효과음 등으로 나

누어 이를 별도의 트랙으로 구분하는 방법, 효과음의 길이에 따라 트랙을 구분하는 방법, 효과음의 음향적인 특성에 따라 구분하는 방법, 효과음이 발생하는 물리적인 거리와 그 성격에 따라 배경 소리 / 사람에서 발생하는 소리 / 물건들 사이에서 발생하는 소리 / 특수한 효과음 등으로 구분하는 방법 등 자체적인 규칙을 사용해서 트랙 탐색 과정을 돕는다. 이런 방법을 적극적으로 활용하는 음향 믹싱 엔지니어들의 경우 각각의 트랙 묶음 별로 같은 묶음에 있는 모든 트랙들을 한 번에 음소거 할 수 있는 총괄 페이더(master fader)를 지정해서 사용하기도 한다.

디지털 음향 믹싱 콘솔에도 트랙 탐색 과정을 돕기 위한 여러 기능이 있다. 대부분의 디지털 음향 콘솔들은 채널 줄 8개를 묶어 이를 페이더 뱅크(fader bank)라고 불리는 단위로 만들어 관리한다. 디지털 음향 콘솔에는 이 페이더 뱅크들을 음향 믹싱 엔지니어가 미리 지정해놓은 위치로 자동 정렬하도록 만드는 단축키들이 있어 음향 트랙 탐색 과정을 돕는다. 또한 모든 채널 줄에는 해당 채널 줄에 대응된 음향 트랙만 재생하도록 만드는 버튼인 독주(solo) 버튼과 해당 음향 트랙을 제외하고 재생하도록 만드는 버튼인 음소거(mute) 버튼이 있어 음향 트랙 탐색 과정을 돕는다.

제 2 절 연구 목표

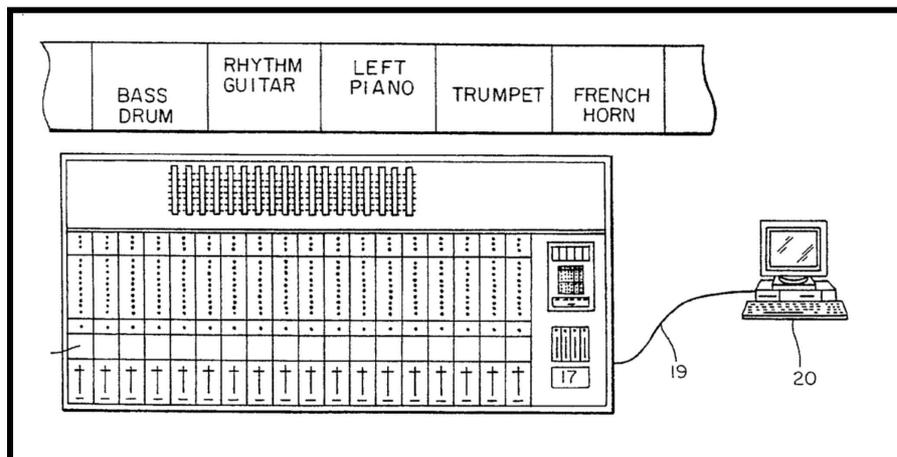
본 연구의 목표는 손가락을 이용한 소리 투사 연구를 통해 효과음에서 추출한 소리 크기 변화 곡선, 음 높이 변화 곡선 등의 음향 특성과 이를 손가락의 움직임으로 표현한 움직임과의 연관성을 찾고 이를 활용해서 영화 음향 믹싱 엔지니어가 찾고 싶은 효과음이 어느 트랙에 있는지 질의(query)를 통해 검색해주는 효과음 검색 인터페이스를 제작하는 것이다.

이를 위한 과정으로 연구는 첫째, 손가락의 1차원 움직임으로 소리를 표현하는 소리 투사 연구를 진행해서 음향 특성과 손가락 움직임 사이의 연관성을 확인하고, 둘째, 효과음 검색 인터페이스를 제작하고, 셋째, 만들어진 효과음 검색 인터페이스를 평가하는 것으로 구성되어 있다.

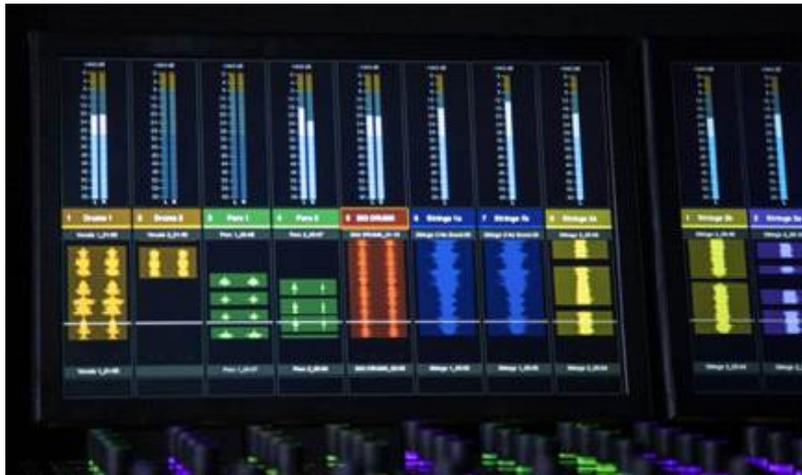
제 2 장 선행 연구 및 관련 개념

제 1 절 음향 콘솔에서 소리 탐색 연구

음향 콘솔에서 자신이 원하는 소리가 어느 음향 트랙에 위치해 있는지 탐색하는 과정을 도와주기 위해 채널 줄 별로 점점 더 많은 정보들을 시각화해서 보여주기 시작했다. 채널 줄 별로 해당 채널 줄에 대응하는 음향 트랙의 이름을 보여주는 발전이 이루어졌다. [그림 6]은 디지털 음향 콘솔에서 트랙 이름을 표시해주는 특허를 보여준다. ([1] US patent 5608807) 최근에는 디스플레이 기술의 발달로 [그림 7]과 같이 실시간으로 각 채널 줄 별로 해당 음향 채널 줄에서 발생하는 소리의 파형을 실시간으로 표시해주는 디지털 음향 콘솔이 개발되었다. ([2] Avid S6)

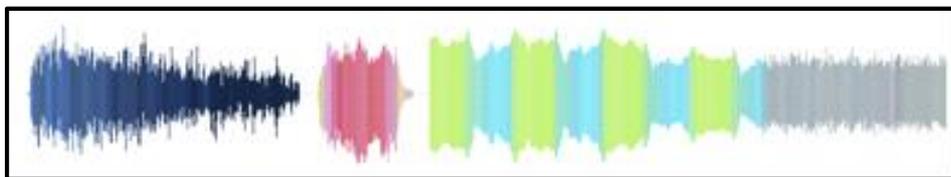


[그림 6] 채널별 트랙 이름 표시



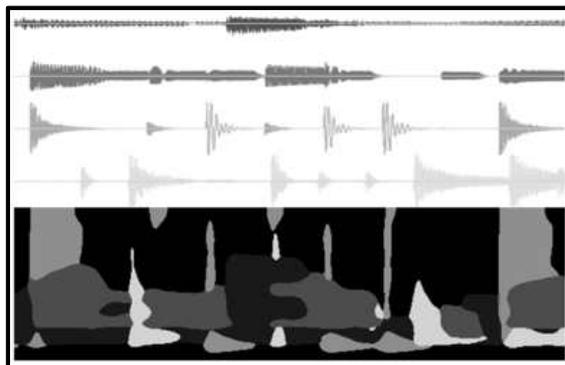
[그림 7] 실시간 파형 표시

디지털 음향 콘솔의 각 음향 채널 줄에서 시각화해서 보여줄 수 있는 디스플레이의 넓이는 제한적인데, 이 제한적인 공간에 더 유용한 정보를 제공하기 위한 연구들이 이루어지고 있다. 이런 연구들은 Rice의 frequency-based coloring of the waveform display to facilitate audio editing and retrieval 연구에서부터 시작되었다. Rice는 디지털 오디오 편집 소프트웨어의 좁은 공간 안에 더 많은 정보를 표시해주기 위해 기존에 단순히 파형만 표시해주던 시각화 방식에서 벗어나 [그림 8]과 같이 파형의 색상을 통해 해당 소음의 주파수 성분을 표시해 주었다.

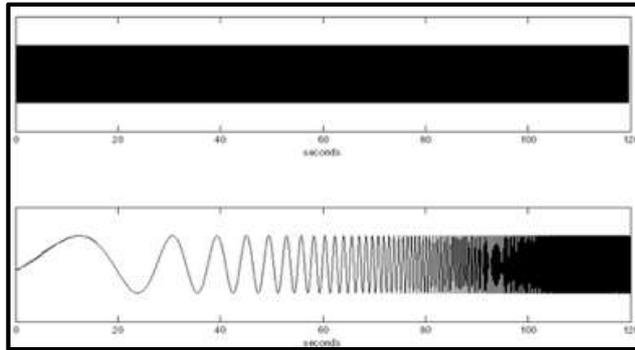


[그림 8] 파형의 색상으로 주파수 정보 제공

이후 비슷한 후속 연구들이 이루어졌다. Gohlke는 track displays in DAW software: Beyond waveform views 연구에서 여러 음향 트랙의 주파수 성분을 상대적으로 분석해서 이를 편리하게 탐색 가능하도록 [그림 9]와 같이 미니맵 형태로 보여주는 연구를 진행했다. ([4] Gohlke, et al. 2010) 또한 Loviscach는 그의 연구 the quintessence of a waveform: focus and context for audio track displays에서 소리 파형의 시간축을 소리의 주파수에 따라 다시 스케일해서 보여주는 연구를 진행했다. ([5] Loviscach, et al. 2011) 이 연구에서 제안한 시각화 방식을 활용해서 낮은 주파수 대역부터 높은 주파수 대역까지 주파수를 점점 높여가며 sine 곡선을 그리는 sine sweep을 표현하면 [그림 10]과 같은 방식으로 소리의 파형을 볼 수 있다. [그림 10]에서 위쪽에 위치한 파형은 기존 방식대로 sine sweep을 표현한 그림이다. 표시되어 있는 시간축의 길이가 너무 길어서 소리의 주파수가 점점 더 높아지고 있다는 정보를 알 수 없다. 하지만 아래에 Loviscach가 제안한 파형은 이 시간축을 다시 스케일해서 표시해서 소리의 주파수가 점점 커지고 있다는 정보를 쉽게 확인 가능하다.



[그림 9] 미니맵으로 트랙별 주파수 성분 표현



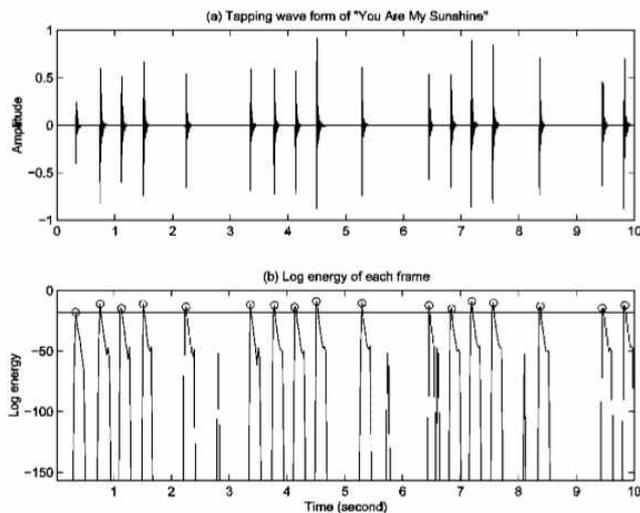
[그림 10] 시간축을 다시 스케일한 파형

그러나 이렇게 시각화를 통해 음향 믹싱 엔지니어의 효과음 탐색 과정을 보조하는 방법은 사용하는 트랙 수가 제한되어 있는 음악 믹싱 엔지니어들에게는 효과적인 발전 방향이지만 200개 이상의 트랙을 사용하는 영화 음향 믹싱 엔지니어를 돕기에는 제한적이다. 영화 음향 믹싱 콘솔의 경우 일반적으로 24개 또는 32개의 음향 채널 줄을 사용하기 때문에 시각화된 정보를 한눈에 보는 것조차 불가능하기 때문이다.

제 2 절 질의 기반 소리 검색 연구

질의 기반 소리 검색 연구는 예시를 통한 질의(query by example)의 일환으로 비언어적인 질의를 통해 소리를 검색하기 위한 연구이다. 기존 질의 기반 소리 검색 연구들에서는 그림, 손가락을 활용한 태핑, 박수, 목소리, 몸동작 등을 활용해서 소리를 검색하는데 활용한다. 이런 연구들은 소리를 언어적으로 표현하는 것의 한계를 극복하는 목표를 가지고 있다.

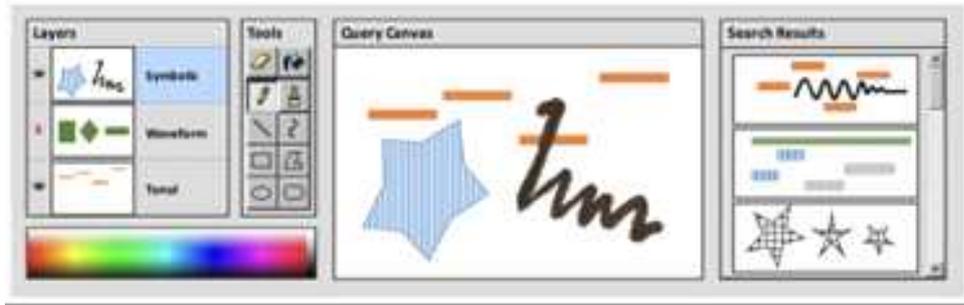
Jang은 소리 검색을 위해서 박수를 활용하는 연구를 제안했고 Hanna는 태핑을 사용하는 연구를 제안했다. [그림 11]은 손가락을 위, 아래로 움직이는 태핑을 사용해서 소리를 검색하는 연구에서 피험자에게 음악을 들려주며 이를 손가락으로 표현한 결과를 보여준다. 박수와 손가락 태핑은 음악이나 박자감이 있는 효과음을 표현하기에는 매우 유리하다. 또한 다른 질의 방식들에 비해 개인별 차이가 적다고 알려져 있다. 하지만 박자감이 없는 효과음은 표현하지 못한다는 문제점이 있다.



[그림 11] 태핑을 질의로 사용한 음악 검색

Pauws는 그의 연구 "CubyHum: a fully operational" query by humming" system에서 허밍을 질의로 사용하는 소리 검색 연구를 진행했다. ([8] Pauws et al. 2002) 이런 허밍을 질의로 활용하는 연구들은 앞서 박자감을 활용한 연구들과 비슷하게 소리의 높낮이 변화나 음악의 멜로디와 같은 소리의 특성을 활용해서 소리를 검색한다. 따라서 이런 특성이 없는 소리들의 경우 표현하는데 한계가 있다.

Keen은 Searching for Audio by Sketching Mental Images of Sound: A Brave New Idea for Audio Retrieval in Creative Music Production에서 현업에서 음악 제작을 하는 음향 전문가들의 인터뷰를 통해 기존 텍스트 기반 소리 검색 방식은 소리에 대한 심리적 그림(mental image)을 잘 반영하지 못한다는 점을 발견했다. 또한 그는 이런 텍스트 기반 소리 검색을 대체하기 위해 그림을 질의로 사용해서 효과음 데이터베이스를 검색하는 인터페이스를 제안했다. ([9] Keen, et al. 2016) 이 연구에서는 찾고자 하는 소리의 심리적 그림을 [그림 12]와 같이 상징적 층(symbolic layer), 파형 층(waveform layer), 톤 층(tonal layer)으로 나누고 각각의 층들을 그려서 소리를 검색하는 인터페이스를 제시했다. 이는 큰 데이터베이스에서 원하는 효과음을 검색하기 위한 연구로 한 질의 그림에 많은 정보를 요구하기 때문에 사용자가 질의를 그리는데 많은 시간이 걸린다. 이는 제한된 규모의 데이터베이스에서 찾고자 하는 소리를 빠르게 검색하는 본 연구의 목표와 맞지 않는다.



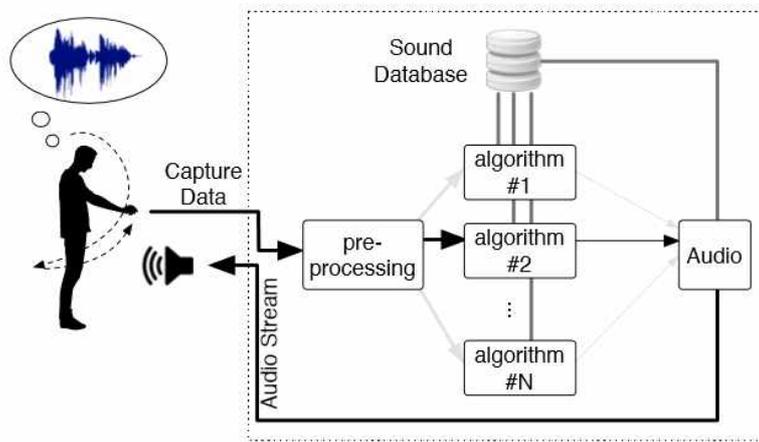
[그림 12] 그림을 활용한 효과음 검색 인터페이스

목소리를 질의로 사용해서 데이터베이스 내 효과음을 검색하는 연구들도 진행되고 있다. ([10] Cartwright, et al. 2013) ([11] Zhang, et al. 2016) 질의 목소리와 데이터베이스 내 효과음의 소리 크기, 음의 높이, 선명도 등을 DTW (Dynamic Time Warping) 를 통해 비교해서 가장 유사한 효과음을 찾는 효과음 검색 인터페이스가 제안되기도 했다. ([10] Cartwright, et al. 2013) 또한 Zhang은 질의 목소리와 데이터베이스 내 효과음에서 미리 학습된 SAE (Stacked Auto-Encoder)를 통해 소리 특성을 추출하고 이를 Kullback-Leibler divergence와 DTW를 사용해서 비교하는 소리 검색 인터페이스를 제안했다. ([11] Zhang, et al. 2016)

Nymoan과 Caramiaux는 몸의 움직임을 사용해서 소리를 검색하는 인터페이스를 제안했다. ([12] Nymoan, et al. 2012) ([13] Caramiaux, et al. 2011) 하지만 목소리나 몸의 움직임을 사용해서 효과음을 검색하는 방법은 쉽게 피로해지기 때문에 반복으로 효과음을 검색해야 하는 음향 믹싱 엔지니어가 활용하기에는 한계가 있다. 또한 몸의 움직임을 질의로 사용해서 소리를 검색하는 인터페이스의 경우 질의의 시작과 끝을 정확하게 특정하지 못한다는 단점이 있다.



[그림 13] 양 팔의 움직임을 사용한 소리 표현



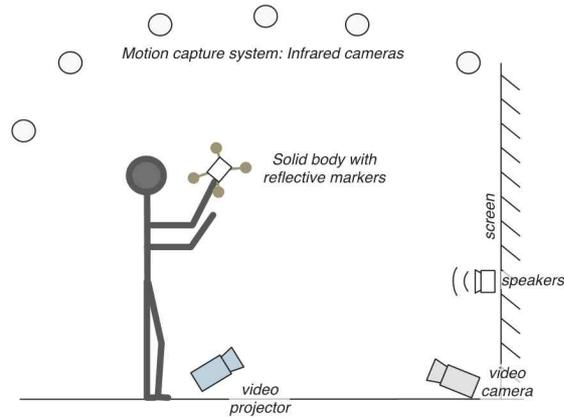
[그림 14] 팔 움직임을 사용한 효과음 검색 연구

제 3 절 소리 투사 연구

Godøy가 처음 정의한 소리 투사(sound tracing) 연구는 사람의 움직임과 소리의 연관성을 분석하는 연구이다. ([14] Godøy, et al. 2006) 소리 투사는 사람의 물리적인 움직임 데이터를 사용하는 경우가 대부분이지만 사람의 목소리를 통한 소리 모사에 적용해서 넓은 범위의 연구로 보기도 한다. 사람의 움직임과 소리라는 두 시계열(time-series) 데이터를 비교해서 사람의 움직임이 소리의 어떤 특성을 표현하는지 찾기 위해 이루어진다. 또한 궁극적으로 이 연구 결과들을 바탕으로 새로운 음향 디자인 도구, 새로운 악기, 새로운 음향 검색 방법 등을 개발하는데 활용하려고 하는 목표를 가지고 있는 연구들이다.

Carmiaux는 효과음을 일반적인 환경에서 자주 접하는 효과음(causal sound)과 이들을 변형한 효과음(noncausal sound)로 나누고 이를 몸 움직임을 통해 표현하는 실험을 하고 피실험자 대상 인터뷰를 진행했다. ([15] Carmiaux, et al. 2014) [그림 15]는 Carmiaux의 실험 환경을 보여준다. Carmiaux는 모션 캡처 시스템(motion capture system)을 활용해서 사람이 반사 마커가 있는 장치를 들고 소리를 표현하도록 지시했다. 이 실험에서는 사람들이 일반적인 환경에서 접하지 못하는 소리를 표현하는 경우 음향 특성 곡선을 표현하지만 일반적인 환경에서 접하는 효과음을 표현하는 움직임은 그 소리가 나도록 하는 몸 움직임을 모사하는 것을 발견했다. 예를 들어 일상에서 쉽게 접하는 새가 날개를 움직이는 소리를 들려주며 이를 표현하라고 요구한 경우에는 새가 날개를 움직이는 모양으로 팔을 움직이는 반면 일상에서 쉽게 접하지 못하는 특수 효과음을 들려주면 효과음의 음향적 특성을 표현한다는

것을 발견했다.



[그림 15] Carmiaux 실험 환경

Nymoен은 효과음을 들려주며 두 팔을 자유롭게 사용해서 표현하도록 하는 실험을 통해 팔의 높이와 효과음의 주파수 중심(spectral centroid)의 연관성과 도입 부분이 강한 성질이 있는 충돌적인(impulsive) 효과음들과 그렇지 않은 효과음들에서 팔 움직임의 가속도 차이가 있음을 보여주었다. ([12] Nymoен, at el. 2012)

Kozak는 다양한 리듬으로 조작한 소리의 리듬에 맞춰 몸을 움직이라고 지시한 실험을 통해 소리와 몸의 움직임에는 복잡한 관계가 있고 이는 소리의 구조와 감정과 같은 요소들에 영향을 받는다는 것을 보여주었다. ([16] Kozak, et al. 2011)

이런 소리 투사 관련 연구들의 목적은 소리와 몸 움직임의 연관성을 통해 효과음 데이터베이스를 검색하는 것과 몸과 목소리 등의 표현을 통해 음향 디자이너가 원하는 소리를 합성 ([17] Rocchesso, et al. 2015) 하는 것이다.

제 4 절 터치 슬라이더



[그림 16] 터치 슬라이더

과거 아날로그 음향 콘솔의 터치 슬라이더는 소리의 크기를 조절하기 위한 장치로 사용되어 음향 페이더(audio fader)로 불려왔다. 그러나 [그림 16]과 같은 디지털 음향 콘솔의 터치 슬라이더(touch sensitive motorized linear encoder)는 단순히 소리의 크기를 조절하는 장치에서 벗어나 잔향의 정도, 소리의 위치 등 다양한 음향 특성을 조절하는데 사용되는 다기능 디지털 입출력장치로 발전했다. 터치 슬라이더는 축전식 센서를 이용해 손가락 접촉 여부를 탐지하고 선형 가변저항을 이용해 슬라이더의 위치를 감지하고 모터를 이용해 슬라이더의 위치를 제어한다.

터치 슬라이더는 음향 믹싱 엔지니어들에게 가장 친숙한 입력 장치로 모든 음향 믹싱 콘솔에 기본적으로 장착되어 있다. 영화 음향 믹싱 엔지니어를 교육하는 교육 기관에서는 눈을 감고 터치 슬라이더를 조작하는 훈련을 할 정도로 터치 슬라이더 사용의 능숙도는 음향 믹싱 엔지니어의 가장 기본적인 기술 중 하나이다. 음향 믹싱 엔지니어는 음향 믹싱 중 대부분의 시간에 터치 슬라이

드 위에 손을 올려놓고 작업을 한다. 최근에는 iPad 등 터치 스크린을 활용한 기기의 보급으로 물리적인 터치 슬라이더를 대체하는 가상 터치 슬라이더도 사용되고 있다. [그림 17]에는 터치 스크린을 사용하는 기기의 디지털 오디오 편집 소프트웨어에 터치 슬라이더가 나타나 있다.



[그림 17] 터치 스크린의 터치 슬라이더

제 3 장 연구 방법 및 제안 특성

제 1 절 연구 문제

본 연구에서는 디지털 음향 콘솔의 터치 슬라이더를 활용해서 영화 음향 믹싱 엔지니어가 찾고 싶은 효과음이 어느 트랙에 있는지 질의를 통해 검색해주는 효과음 검색 인터페이스를 제작하고 그 사용성을 확인하고자 한다. 이 연구를 통해 음향 믹싱 엔지니어가 쉽고 빠르게 자신이 원하는 효과음이 있는 위치를 찾을 수 있도록 도와주는 연구 목표를 가지고 있다. 이를 위해 두 가지 연구 문제를 제시한다.

연구 문제 1

터치 슬라이더를 이용한 소리 표현과 음향 특성 사이의 관계는?

먼저 터치 슬라이더를 이용한 소리 표현과 음향 특성과의 관계를 알아보하고자 한다. 이 연구에서는 사람이 효과음을 터치 슬라이더를 사용해서 소리를 표현하는 경우에 효과음의 어떤 특성들을 표현하는지 확인하고 이 특성들이 기존 몸 움직임 질의를 사용한 효과음 표현과 어떤 차이가 있는지 살펴볼 것이다. 이 연구에서는 기존 관련 연구인 음향 투사 연구들에서 사용한 연구 방법들을 활용할 것이다. 또한 이 연구의 결과로 터치 슬라이더를 사용해서 영화 음향 믹스 내에서 효과음 검색이 가능한지 기술적인 타당성을 확인하려 한다.

연구 문제 2

터치 슬라이더를 이용한 효과음 검색 인터페이스의 사용성은?

터치 슬라이더를 이용한 효과음 검색 인터페이스가 실제로 음향 믹싱 엔지니어들의 효과음 검색 과정을 도와주는지 평가한다. 이 평가는 현업에서 일하고 있는 음향 믹싱 엔지니어들을 대상으로 수행하고 이를 통해 제안하는 인터페이스의 활용 가능성을 검증하고자 한다. 평가는 제안하는 인터페이스가 기존 인터페이스에 비해 효과음 검색 시간을 얼마나 단축시키는지를 평가하는 정량적 평가와 인터뷰를 통한 정성적 평가로 이루어진다.

제 2 절 제안 특성

3.2.1 제안 음향 특성

소리를 나타내기 위해 사용하는 음향 특성으로는 음압 변화 곡선, 음색 변화 곡선 등 다양한 특성들이 있다. 본 연구에서는 이들 중 기존 음향 투사 연구들에서 몸의 움직임과 관계가 있다고 알려진 음향 특성들과 음향 인지에 주로 활용되어온 음향 특성들인 음압 변화 곡선, 소리 크기 변화 곡선, 주파수 중심 변화 곡선, 원점 건넌 밀도를 선택했다. 본 연구에서는 이 음향 특성들과 터치 슬라이더의 움직임 특성을 비교해서 이들 중 유의한 특성들이 무엇인지 확인할 것이다.

1. 음압 변화 곡선(intensity curve)

소리의 음압 시간에 따라 어떻게 변화하는지를 나타내는 곡선이다. 음압은 소리가 갖는 물리적인 압력의 크기이다. 소리를 나타내기 위한 음향 특성들 중 가장 기본적인 특성들 중 하나이다. 본 연구에서는 음압의 절대값이 아닌 로그 스케일된 데시벨(dB) 단위의 음압을 사용한다.

2. 소리 크기 변화 곡선(loudness curve)

소리의 크기가 시간에 따라 어떻게 변화하는지를 나타내는 곡선으로 사람이 인지하는 소리의 크기의 변화를 표현한다. 전술된 음압 변화 곡선과는 달리 주파수와 음압에 따른 인지된 소리 크기를 나타낸다. 본 연구에서는 ISO 226:2013의 소리 크기 일치 곡선

(equal loudness contour)을 사용해 phon 단위로 나타낸 소리 크기를 사용한다.

3. 주파수 중심 변화 곡선(spectral centroid curve)

주파수 중심 변화 곡선은 시간-주파수 특성에서 시간별 주파수 중심이 어떻게 변화하는지를 나타낸다. 특정 시각의 주파수 중심은 아래의 공식으로 표현된다.

$$SC = \frac{\sum_{k=0}^{N/2} f_k |X(k)|^2}{\sum_{k=0}^{N/2} |X(k)|^2}$$

(f_k : k 번째 DFT(Discrete Fourier Transform) 필터의 중심 주파수, $X(k)$: k 번째 DFT 필터의 값)

4. 원점 건넌 밀도(zero crossing rate)

원점 건넌 밀도는 소리의 파형이 원점을 건너는 밀도를 시간에 따라 나타낸 음향 특성으로 소리의 크기와 주파수 등 다양한 소리의 성질에 의해 영향을 받는다.

기존 음향 투사 관련 연구들에서 발견한 결과에 따라 본 연구에서는 전술한 음향 특성들에 대해 각각 그 절대값, 미분값, 2차 미분값 특성을 추출해서 이를 음향 특성으로 활용할 것이다.

3.2.2 제안 움직임 특성

본 연구에서는 터치 슬라이더를 통해 수집된 1차원 시계열 데이터를 다양한 방법으로 변형해서 사용한다. 실험에 사용할 10cm의 터치 슬라이더를 통해 수집된 기본 값은 0 ~ 100mm 단위의 시간에 따른 위치 데이터이다. 본 연구에서는 시간에 따른 위치 데이터를 1차 미분해서 얻은 움직임의 속도와 2차 미분해서 얻은 가속도를 구해서 이를 추가적인 움직임 특성으로 활용한다. 이는 기존 음향 투사 연구들 제안했던 방법들 중 하나이다.

제 3 절 소리 투사 실험

터치 슬라이더를 사용한 1차원 소리 투사 실험을 진행했다. 이 실험의 목표는 사람들이 터치 슬라이더를 사용해서 소리를 표현하는 경우 어떤 소리 특성을 어떻게 표현하는지를 알아보기 위한 것이다. 소리 투사 실험은 선행 연구들[12, 14, 15]에서 진행한 소리 투사 실험과 유사한 방법으로 이루어졌다. 본 연구에서 기존 선행 연구들의 소리 투사 실험과 다르게 설계된 부분은 효과음이 시작되는 순간을 예측 불가능 하도록 설계한 부분이다.

3.3.1 효과음 데이터

실험은 인공적으로 합성된 효과음 20개, 녹음된 효과음 15개, 사람의 움직임에서 발생하는 효과음인 폴리 효과음 15개로 총 50개의 효과음을 사용했다. 사용된 효과음은 0.5 ~ 4초 길이의 효과음들이며 주어진 효과음들은 5초 안에서 불규칙한 순간에 시작해서 사용자가 효과음의 시작 시각을 예측할 수 없는 실험 환경을 만들었다. 이런 예측 불가능함은 기존 몸 움직임 관련 소리 표현 실험들 대부분이 효과음이 시작하게 전에 일정한 간격으로 청각 자극을 주어 효과음이 시작하는 순간을 알려주는 것과 대조적이며 음향 믹싱 엔지니어의 효과음 검색 환경에 맞춘 실험 환경이다. 또한 소리 투사 연구들에서 사용한 효과음들은 대부분 소리의 크기나 음의 높이와 같은 특성의 변화가 일정한데 비해 본 실험에 사용한 효과음들은 그렇지 않다. 실험에 사용된 효과음은 Vancouver Film School의 student sound library에서 제공하는 효과음들에서 선정했다. 사용된 효과음은

아래 [표 1]과 같다.

Original File Name	Type
Plasma_Cannon_Sci-Fi_Fire_Energy_Boom	합성음
Doepfer_Power_Down	합성음
Beeps_Filtered_Ui	합성음
Boost_Hoover_Quick	합성음
Creating_Magic	합성음
Electric_Blast_Zap_Electricity	합성음
Electric_Hit_Impact_Buzz_Zap	합성음
Energy_Charge	합성음
Gun_Percussive_Echo_Sci-Fi	합성음
Interface_HUD_Sci-Fi	합성음
Magic_Orb_03	합성음
Mechanical_Servo_Cinematic_Slow_Down	합성음
Mega_Punch	합성음
Spaceship_Alien_Engine_Scream	합성음
Spaceship_Engine_Synth_Airy_Fast_Dopper_ST	합성음
Spell_Magic_Electric_Bright	합성음
Stinger_Buildup_Bell_Dissolve	합성음
Synth_Engine_Slow_Down	합성음
Synth_Error_Warning	합성음
Synth_Gun_Shot_Hollow	합성음
Air_Tank_Valve_Hiss_Shut_Off	녹음 효과음
Birds_Exotic_Chirping_Various	녹음 효과음
Cow_Moo_Dairy_Various	녹음 효과음
Horse_Trot_Gravel	녹음 효과음

Plane_Prop_By_Buzz_Mid_ST	녹음 효과음
Sea_Lion_Breath	녹음 효과음
Spray_Once	녹음 효과음
Chainsaw_Various	녹음 효과음
Clock_Tick_Slow_01	녹음 효과음
Door_Heavy_Large_Metal_Close	녹음 효과음
Door_Sliding_Metal_Rattle	녹음 효과음
Alarm_Beeps_Fire_Suit_ST	녹음 효과음
Bike_Crash_Impact_Pavement	녹음 효과음
Camera_Auto_Zooms	녹음 효과음
Machine_Grinder_Large_On_Off	녹음 효과음
Balloon_Squeaks_2	폴리
Beer_Pouring	폴리
Bone_Skull_Old_Break_Snap	폴리
Book_Paper_Page_Turn	폴리
Broom_Sweep_Floor	폴리
Cement_Drag_Bric_Stone	폴리
Cloth_Pull_Smooth_Short	폴리
Cork_Pop_Wine_02	폴리
Drag	폴리
Switches_9	폴리
Hammer_Small_Wood_1	폴리
Dirt_Impact_Ground_Heavy_01	폴리
Umbrella_Open	폴리
Bite_Potato_Chips_Various	폴리
Knock_Door_Wood_Triple	폴리

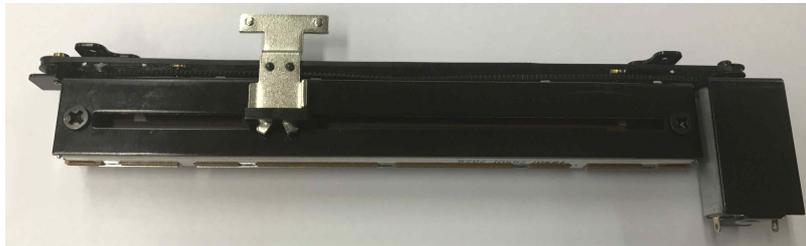
[표 1] 소리 투사 실험에 사용된 효과음 목록

3.3.2 소리 투사 실험

소리 투사 실험은 영화 음향 믹싱 경험이 없는 20대, 30대로 구성된 남녀 대학원생을 대상으로 진행되었다. 피험자는 3.3.1의 [표 1]에 제시된 효과음을 들으며 이를 터치 슬라이더를 통해서 실시간으로 표현하도록 지시되었다. 세부적인 실험 과정은 다음과 같다.

1. 피험자는 터치 슬라이더 근처에 손을 올리고 모니터 헤드폰을 착용한다.
2. 피험자는 모니터 헤드폰을 통해서 들려오는 소리를 터치 슬라이더를 사용해서 실시간으로 표현하도록 지시받는다.
3. 피험자가 실험 환경에 익숙해지도록 돕기 위한 예시 효과음 5개를 순서대로 들으며 터치 슬라이더를 사용해서 소리를 표현하는 연습을 한다. 한 효과음에 대한 표현이 끝난 다음 피험자가 터치 슬라이더에서 손가락을 떼면 터치 슬라이더는 다음 효과음이 재생되기 전에 자동으로 가운데 지점(50mm)으로 이동해서 새로운 입력을 받을 준비를 한다.
4. 피험자는 [표 1]에 나타난 효과음을 무작위 순서로 들으며 이를 터치 슬라이더를 사용해서 표현하도록 지시받는다. 3번 과정과 같이 매 효과음에 대한 표현이 끝나면 터치 슬라이더의 위치가 초기화된다.

실험에는 [그림 18]에서 볼 수 있는 디지털 음향 콘솔의 터치 슬라이더로 많이 사용되는 BOURNS사의 PSM Series 터치 슬라이더를 사용했고 Arduino를 DAQ (Data AcQuisition) 장비로 활용해서 움직임 데이터와 터치 데이터를 수집했다. 데이터는 약 30Hz의 샘플링 속도로 수집되었다. 효과음 재생에는 스튜디오 모니터 헤드폰으로 자주 사용되는 ATH M50 헤드폰을 사용했다.



[그림 18] BOURNS의 PSM Series 터치 슬라이더



[그림 19] 소리 투사 실험 환경

3.3.3 소리 투사 데이터 처리

소리 투사 실험 결과 수집된 터치 데이터와 움직임 데이터, 음향 특성 데이터는 세 단계의 전처리 과정을 통해 소리 투사 데이터로 만들어졌다. 첫 단계는 수집된 전체 데이터에서 피험자가 손가락을 터치 슬라이더에 댄 순간부터 효과음에 대한 표현을 끝내고 손가락을 떼 순간까지의 데이터만 잘라내는 전처리 과정이다. 다음으로는 음원에서 추출한 음향 특성 데이터와 피험자의 터치 데이터의 시작 지점을 맞춰 사용자의 반응 속도를 보정하는 단계를 거쳤다. 마지막으로 얻어진 선형 보간 (linear interpolation)을 통해 20Hz 일정한 샘플링 속도로 변환했다. 소리 투사 실험 결과 얻어진 데이터는 10 명이 50 개의 효과음을 20Hz 단위로 샘플링한 결과로 총 8,562 개의 데이터로 구성되어 있다.

제 4 절 소리 투사 모델

3장 3절에서 수집한 소리 투사 데이터를 활용해 터치 슬라이더를 활용한 소리 투사 결과를 활용해서 유의한 특성을 추출하고 이를 활용해 소리 투사 모델을 만들었다. 3장 2절에서 제안한 음향 특성과 움직임 특성들을 활용한 다인자 선형 모델(multivariate linear model)을 만들고 유의한 특성을 찾아냈다. 다인자 선형 모델은 음압 변화 곡선, 소리 크기 변화 곡선, 주파수 중심 변화 곡선, 원점 건넌 밀도 터치 슬라이더의 움직임을 얼마나 잘 예측하는지를 확인하도록 3가지 모델로 나누어 만들었다.

모델 1 : 터치 슬라이더의 절대값(위치) 예측 모델

특성	β	p-value
음압 변화 곡선 - 절대값	0.82	0.12
소리 크기 변화 곡선 - 절대값	0.71	0.07
주파수 중심 변화 곡선 - 절대값	-0.06	0.09
원점 건넌 밀도 - 절대값	-0.02	0.42

[표 2] 절대값 예측 모델 특성별 유의도

모델 2 : 터치 슬라이더의 1차 미분값(속도) 예측 모델

특성	β	p-value
음압 변화 곡선 - 1차 미분	0.43	< 0.001
소리 크기 변화 곡선 - 1차 미분	0.35	< 0.001
주파수 중심 변화 곡선 - 1차 미분	0.02	< 0.01
원점 건넌 밀도 - 1차 미분	0.17	0.36

[표 3] 1차 미분값 예측 모델 특성별 유의도

모델 3 : 터치 슬라이더의 2차 미분값(가속도) 예측 모델

특성	β	p-value
음압 변화 곡선 - 2차 미분	0.61	< 0.01
소리 크기 변화 곡선 - 2차 미분	0.33	0.04
주파수 중심 변화 곡선 - 2차 미분	0.18	0.08
원점 건넌 밀도 - 2차 미분	-0.05	0.39

[표 4] 2차 미분값 예측 모델 특성별 유의도

실험 결과는 선행 연구 결과와 차이가 있었다. 손의 움직임을 사용한 선행 연구의 경우 손의 절대적인 높이가 음압 변화 곡선이나 주파수 중심 변화 곡선과 연관이 있다는 결과가 있었지만 터치 슬라이더를 사용한 효과음 표현 실험에서는 슬라이더의 절대적인 위치는 정보를 포함하고 있지 않았다. 하지만 위치를 미분한 속도에서는 효과음 검색에 사용될 수 있는 정보가 포함되어 있었다.

이런 결과는 터치 슬라이더라는 입력 장치의 특성으로 때문으로 보인다. 터치 슬라이더에서 손가락이 위치한 정확한 위치는 음향 특성들과 비교했을 때 유의한 결과가 나오지 않은 것에 비해 그 변화량과 음향 특성들의 변화량 사이에서는 유의한 결과가 나왔다는 것은 터치 슬라이더를 사용해서 소리를 표현할 때 한 효과음 안에서는 그 음향 특성의 변화를 표현하지만 터치 슬라이더의 각 위치별로 특정 소리 크기나 특정 주파수의 소리가 나타나지는 않는다는 것을 의미한다.

음향 특성들 중 원점 건넌 밀도와 터치 슬라이더의 움직임 사이에 유의미한 연관성이 발견되지 않은 점도 기존 선행 연구들과 차이가 있는 부분이다. 선행 연구 결과들에서는 주파수 중심 변화

곡선에 비해서 원점 건넌 밀도가 소리 투사에 더 유의한 관계가 있음을 보여주는 연구들이 있었지만 이는 터치 슬라이더에서는 적용되지 않았다.

소리 투사 실험 결과에 따라 효과음 검색 인터페이스에 사용할 소리 투사 모델의 특성으로 1차 미분한 음압 변화 곡선과 1차 미분한 소리 크기 변화 곡선, 1차 미분한 주파수 중심 변화 곡선을 선정했다. 터치 슬라이더 움직임의 2차 미분 결과와 음압 변화 곡선의 2차 미분 결과 사이에도 유의한 결과가 발견되었지만 같은 기본 특성을 활용한 음압 변화 곡선의 1차 미분 결과와 터치 슬라이더 움직임의 1차 미분 결과가 더 좋은 결과를 보여주어 이는 활용하지 않았다. 효과음 검색 인터페이스에는 선정된 특성들을 활용한 다인자 선형 모델을 사용했다.

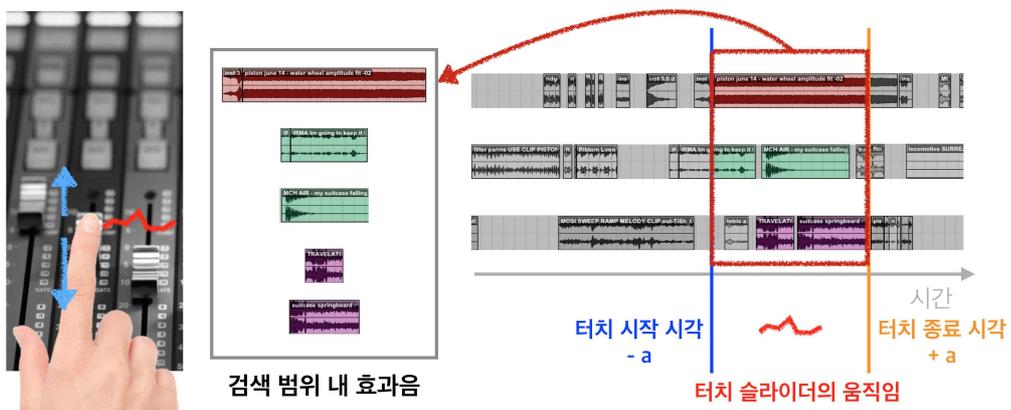
제 4 장 제안 인터페이스

제 1 절 인터페이스 개요

4.1.1 검색 인터페이스 개요

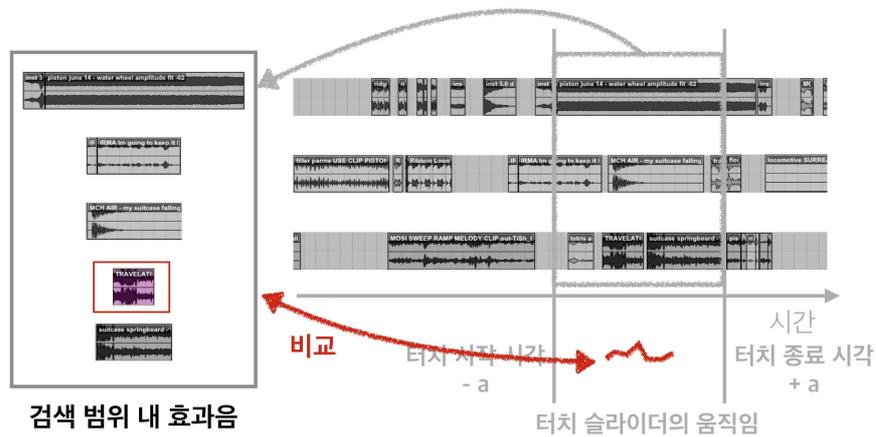
터치 슬라이더를 이용한 효과음 검색 인터페이스는 아래의 과정을 통해 작동한다.

- (1) 영화 음향 믹싱 엔지니어가 음향 믹스를 재생하는 도중 자신이 변화시키고 싶은 소리가 생기면 터치 슬라이더를 통해 효과음을 표현한다. 검색 인터페이스는 사용자가 슬라이더를 터치한 시각과 터치를 종료한 시각에 여유를 두고 효과음 검색 범위를 설정하고 이렇게 설정된 효과음 검색 범위 사이에서 현재 효과음이 재생되고 있는 모든 트랙에 있는 모든 효과음이 효과음 검색 대상이 된다.



[그림 20] 효과음 검색 인터페이스 개요 - 1

- (2) 효과음 검색 인터페이스는 효과음 검색 대상과 터치 슬라이더의 움직임, 터치 슬라이더의 터치 시간을 바탕으로 음향 믹싱 엔지니어의 의도를 파악해서 해당 효과음을 검색해 준다. 만약 검색 결과값이 유사한 효과음이 둘 이상일 경우 이를 음향 디자이너가 효과음을 검색하는데 사용한 터치 슬라이더 주변의 터치 슬라이더로 불러와서 이후 효과음 탐색 과정을 돕는다.



[그림 21] 효과음 검색 인터페이스 개요 - 2

4.1.2 검색 알고리즘 개요

제안하는 검색 알고리즘은 터치 슬라이더를 통한 효과음 질의가 포함하고 있는 움직임 정보와 터치 정보를 조합에서 사용자가 표현했다고 판단하는 효과음을 검색한다.

(1) 슬라이더 움직임

효과음 특성의 시간에 따른 변화와 슬라이더의 움직임을 비교하기 위한 알고리즘은 3장 4절의 결과를 활용해 터치 슬라이더의 속도를 예측하도록 만들었다. 터치 슬라이더의 속도를 예측하는 선형 모델 $f(P_n, L_n, SC_n)$ 에서, 터치 슬라이더의 움직임과 효과음 특성 사이의 거리 알고리즘은 다음과 같다.

$$d = \frac{1}{b-a} \sum_{n=a}^b |V_n - f(P_n, L_n, SC_n)|$$

a : 효과음과 슬라이더 움직임 중 늦게 시작하는 값의 index

b : 효과음과 슬라이더 움직임 중 먼저 끝나는 값의 index

V_n : n 번째 슬라이더 속도

P_n : 효과음의 n 번째 음압 1차 미분값

L_n : 효과음의 n 번째 크기 1차 미분값

SC_n : 효과음의 n 번째 주파수 중심 1차 미분값

(2) 터치 시간

터치 슬라이더는 몸 움직임의 질의로 사용한 기존 연구들과 달리 사용자가 슬라이더 위에 손을 올린 시각과 사용자가 슬라이더 위에서 손을 떼 시각을 정확하게 알 수 있고 이를 통해 사용자가 효과음을 들으며 표현한 동작의 시작과 끝을 정확하게 알 수 있다. 이를 검색 대상 효과음의 길이와 비교해서 그 차이를 사용해 효과음 검색에 활용한다. 본 논문에서는 사용자가 터치 슬라이더에 손을 얹은 시각 이전 1초와 사용자가 터치 슬라이더에서 손을 떼 시각 이후 1초 사이에 재생되고 있는 모든 음원을 검색 대상 효과음으로 지정했다.

제 2 절 인터페이스 평가

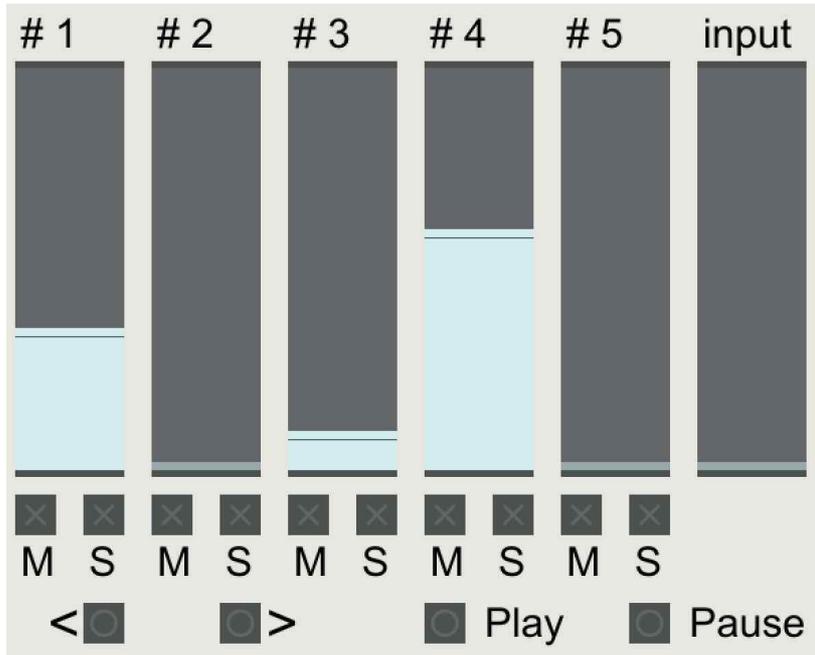
4.2.1 인터페이스 평가 방법

제안 인터페이스의 평가는 4절에서 제안한 소리 투사 모델을 사용한 효과음 검색 알고리즘의 검색 성능을 평가하기 위한 검색 정확도 평가와 이를 실제 음향 믹싱 엔지니어들이 활용하는 경우 효과음 검색 과정을 도울 수 있는지를 평가하기 위한 사용성 평가로 나누어 진행했다.

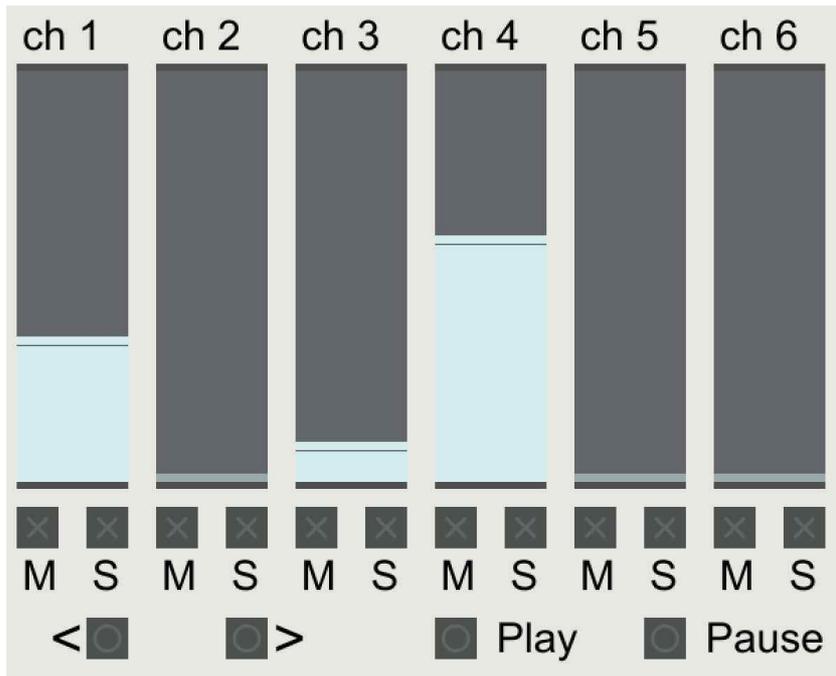
1. 평가를 위한 시제품 인터페이스

본 연구에서 제안하는 효과음 검색 인터페이스를 평가하기 위한 시제품 인터페이스는 음향 믹싱 콘솔에서 음향 탐색을 위해 제공하는 기본적인 기능들을 가지고 있도록 구현되었다. 시제품 인터페이스는 여러 트랙의 음향 신호를 사용자의 입력에 맞춰 실시간으로 재생하며 각 채널별 소리 크기 역시 실시간으로 표현해 주어야 한다. 또한 사용자의 터치 입력이 발생한 경우 해당 구간에서 재생되고 있는 음원들의 소리 특성을 분석하는 연산 시간이 소요된다. 본 연구에서는 이런 연산 시간을 최대한 단축하기 위해 실험을 진행하기 전에 미리 트랙별로 음향 특성들을 추출해 놓은 상태에서 실험을 진행했다.

평가를 위한 시제품 인터페이스는 Microsoft Surface 태블릿 PC와 인터랙티브 음향 미디어 저작 도구인 Max MSP를 사용해서 구현되었다. ([18] Max MSP)



[그림 22] 제안 효과음 검색 인터페이스 UI



[그림 23] 대조군 효과음 검색 인터페이스 UI

[그림 22]은 본 연구에서 제안하는 인터페이스의 UI로 음향 신을 재생하다가 찾고자 하는 음원이 재생되면 왼쪽의 input 터치 슬라이더를 사용해서 움직임을 입력하도록 구성되어 있다. 사용자가 input 터치 슬라이더를 통해 찾고자 하는 소리를 표현하면 효과음 검색 알고리즘이 사용자가 찾는 것으로 예상되는 효과음을 5 순위까지 나열한다. [그림 23]는 본 연구에서 제안하는 인터페이스와 비교하기 위한 대조군 인터페이스의 UI로 대부분의 콘솔에서 사용하는 채널 기반의 음향 인터페이스이다.

본 연구에서 제안하는 효과음 검색 인터페이스의 경우 표시되는 전체 음향 채널 6 개 중 한 개를 사용자의 입력을 받기 위한 채널로 사용하기 때문에 대조군 인터페이스에 비해 실제로 표시되는 음향 채널의 수가 하나 적다. 두 인터페이스를 공평하게 비교하기 위해서는 실제 음향 믹싱 상황에서의 제약 조건인 면적을 동일하게 유지하는 것이 적당하기 때문이다.

두 인터페이스 모드 채널 묶음 사이를 이동하기 위한 좌 우 버튼과 소리를 재생하고 정지하기 위한 버튼이 있다. 버튼이 눌린 트랙의 소리를 제외한 다른 모든 소리를 끄는 독주 (Solo) 버튼과 버튼이 눌린 트랙의 소리를 재생하지 않도록 만드는 음소거 (Mute) 버튼도 공통적으로 제공된다.

2. 평가 방법

평가에는 10 초 길이의 음향 신(auditory scene) 10 개가 활용되었다. 각 음향 신은 총 30 개의 음향 트랙으로 구성되어 있고 평균적으로 17.2 개의 개별 효과음들이 음향 트랙에 분산되어 있다.

각 음향 신에는 6 개의 목표 음원이 있다. 음향 믹싱 엔지니어는 이 6 개의 목표 음원들 중 3 개는 본 연구에서 제안한 새로운 효과음 검색 인터페이스를 사용해서, 나머지 3 개는 대조군 효과음 검색 인터페이스를 사용해서 효과음을 검색하도록 지시받는다. 음향 믹싱 엔지니어는 이 과정을 10 개 음향 신에 대해서 반복해서 수행하게 된다. 세부적인 평가 방법은 다음과 같다.

1. 피험자에게 10 초 길이의 음향 신들 중 무작위 순서로 골라진 한 개를 2 회 반복해서 들려준다.
2. 피험자에게 들려준 음향 신에는 총 6 개의 목표 음원이 있는데, 이 6 개의 목표 음원들 중 하나를 들려주고 해당 음원이 위치한 트랙을 찾으라고 지시한다. 피험자는 본 연구에서 제안한 효과음 검색 인터페이스와 대조군 효과음 검색 인터페이스 중 임의로 선택된 하나를 사용해서 효과음을 검색하게 된다. 만약 첫 목표 음원을 찾기 위해서 대조군 효과음 검색 인터페이스가 사용되었다면 다음 목표 음원은 제안된 효과음 인터페이스를 사용해서 찾게 된다. 이렇게 번갈아가며 6 개의 목표 음원이 위치한 트랙을 찾는다. 결과적으로 한 음향 신의 목표 음원 6 개 중 3 개는 제안된 인터페이스를 사용해서, 나머지 3 개는 대조군 인터페이스를 사용해서 찾게 된다. 실험 과정 중에는 두 인터페이스에 모두 제공되는 독주 버튼과 음소거 버튼을 자유롭게 사용할 수 있다.
3. 피험자는 실험 과정 2를 총 10 개의 음향 신에 대해 반복적으로 수행한다.

6개의 음원들 중 어떤 음원을 제안하는 효과음 검색 시스템을 통해 검색하고 어떤 음원을 대조군 효과음 검색 인터페이스로 검색하도록 하는지와 어떤 음향 신을 먼저 실험하는지는 피험자별로 무작위로 결정했다.

3. 평가 방법

(1) 정확도 평가

제안하는 효과음 검색 알고리즘의 검색 정확도를 평가하기 위한 방법으로는 일반적인 정보 검색 알고리즘의 평가 방식을 사용했다. 평가자가 입력 (input) 터치 슬라이더를 통해 입력한 움직임으로 효과음 검색 알고리즘이 사용자가 표현한 것으로 예상한 효과음들을 순위별로 나열한 결과에서 정답 효과음의 등수를 기준으로 평가했다.

(2) 사용성 평가

본 연구의 주된 목표는 터치 슬라이더를 활용한 효과음 검색 인터페이스를 통해 현업에서 일하는 영화음향 믹싱 엔지니어를 돕는 데 있다. 따라서 효과음 검색 인터페이스를 평가하기 위해서는 단순한 검색 성능 평가에 이어 실제 목표 사용자를 대상으로 인터페이스의 사용성 검증 평가가 필요하다. 사용성 검증 평가는 정량적 평가와 정성적 평가로 나누어 진행했다.

사용성의 정량적 평가 기준은 효과음의 트랙 위치를 찾는 데 걸리는 평균 시간으로 했다. 이를 위해서 시제품 인터페이스의 사

용 과정에서 발생하는 로그를 기록하였다. 사용성의 정성적 평가는 사용자를 대상으로 한 실험 후 인터뷰를 통해 평가했다. 실험 후 인터뷰를 통해서 제안하는 인터페이스의 유용성(항목 1, 항목 2)과 직관성(항목 3)을 확인하고자 하였다. 각 질문에는 자유로운 형식으로 답변하도록 하였으며 피험자의 답변에 따라 추가 질문을 통해 각 답변이 의미하는 바를 명확하게 알 수 있도록 했다.

항목	질문
1	제안된 효과음 검색 인터페이스가 디지털 음향 믹싱 콘솔에 장착되어 나온다면, 얼마나 자주 사용하겠는가?
2	제안된 효과음 검색 인터페이스가 디지털 음향 믹싱 콘솔에 장착되어 나온다면, 얼마나 유용하게 사용하겠는가?
3	제안된 효과음 검색 인터페이스가 효과음을 얼마나 직관적으로 검색하여 준다고 느꼈는가?

[표 5] 인터뷰 질문

3. 평가 대상

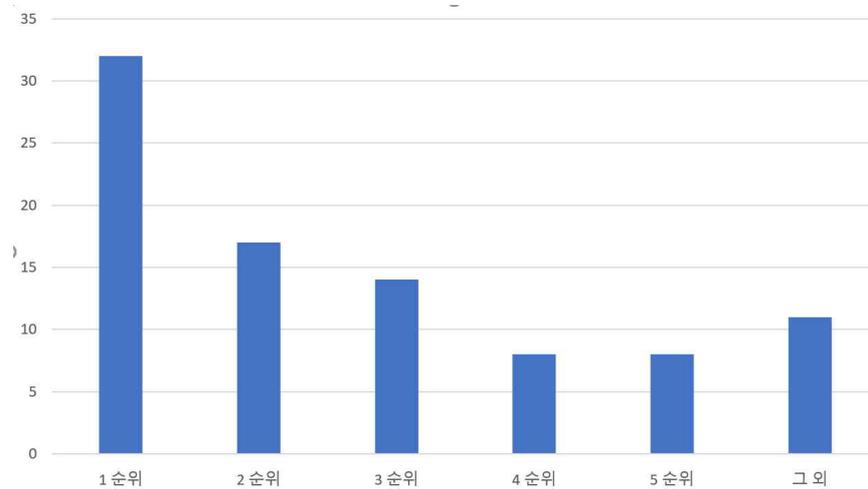
현업에서 일하고 있는 3명의 음향 믹싱 엔지니어를 대상으로 평가를 진행했다.

피험자 #	비고
1	방송국 소속 드라마 음향 믹싱 엔지니어 음향 관련업 종사 경력 15년 음향 믹싱 경력 7년
2	방송국 소속 드라마 음향 믹싱 엔지니어 음향 관련업 종사 경력 8년 음향 믹싱 경력 4년
3	광고 제작 스튜디오 소속 음향 믹싱 엔지니어 음향 관련업 종사 경력 5년 음향 믹싱 경력 5년

[표 6] 평가 대상

4.2.2 정확도 평가 결과

본 연구에서 제안한 효과음 검색 알고리즘의 정확도 평가 결과는 다음과 같다. 10 개의 신에서 3 개의 효과음을 찾는 과제를 3 명의 피험자가 시도하여 총 90 번의 사용자 질의에 대한 평가 결과이다. 전체 질의들 중 32 번은 원하는 소리를 1순위로 찾았다. 또한 전체 90번 중 70%인 63번은 원하는 소리가 3순위 안에서 찾았다.



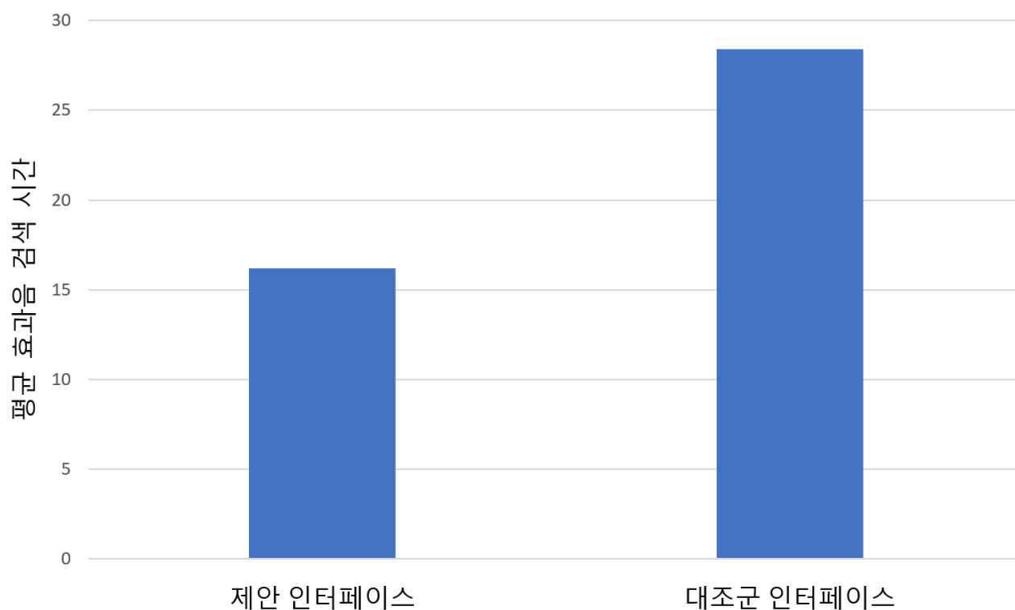
[그림 24] 정확도 평가 결과를 나타낸 히스토그램

실험 결과 5순위 안에서 찾지 못한 소음들에는 시간에 따라 큰 변화가 없는 배경 효과음들과 매우 짧은 시간에 비슷한 소리가 많이 연사되는 총소리와 같은 효과음들이 있었다.

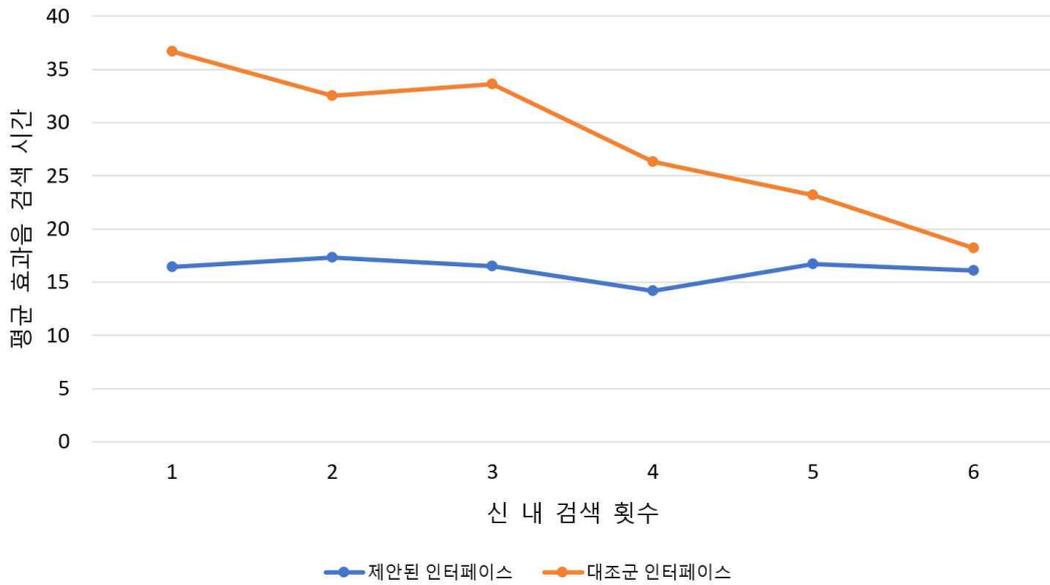
4.2.3 사용성 평가 결과

1. 정량적 평가

사용성에 대한 정량적 평가 결과 본 연구에서 제안하는 인터페이스를 사용한 경우 주어진 효과음을 찾는데 평균 16.2초(표준편차 5.3)가 소요되었고 대조군 인터페이스를 사용한 경우 평균 28.4초(표준편차 12.7)가 소요되었다. 본 연구에서 제안한 인터페이스를 활용하는 경우 효과음 검색 시간을 유의미하게 줄일 수 있다는 것을 확인했다. 두 평균값의 차이를 보기 위해 검증한 T-test (+ Bonferroni correction, 양측 검정) 결과 유의미한 차이($P < 0.001$)를 확인했다.



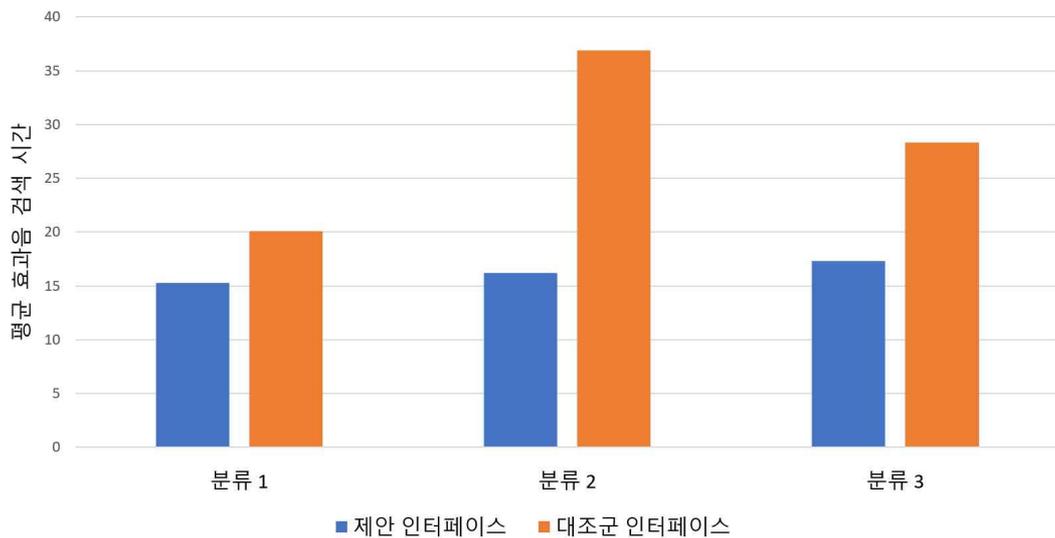
[그림 25] 평균 효과음 검색 시간 그래프



[그림 26] 신 내 검색 횟수와 평균 검색 시간 그래프

피험자들은 한 신에서 총 6 개의 목표 음원을 검색하게 되는데 그 과정에서 두 인터페이스들 간에 명확한 차이가 발생했다. 대조군 인터페이스의 경우 한 신 내에서 효과음 검색을 반복함에 따라 효과음 검색에 필요한 시간이 감소하는데 비해 새로 제안된 인터페이스의 경우 한 신 내에서 효과음 검색을 반복함에 따라 효과음 검색에 필요한 시간이 크게 변하지 않다는 것이 발견되었다. 이는 기존 인터페이스의 경우 트랙의 상대적인 위치가 항상 고정되어 있기 때문에 한 신 안에서 효과음을 반복해서 검색하면서 해당 신에 있는 효과음들의 위치를 암기하기가 용이한데 비해 제안하는 인터페이스의 알고리즘이 판단한 순위에 따라 효과음의 위치가 계속 변해서 암기가 어렵기 때문으로 생각된다. 하지만 10 초 길이의 음향 신을 12 번 이상 반복 재생한 다음 찾는 신 내 6 번째 검색하는 검색 대상 음원의 경우에도 제안하는 인터페이스에 비해 검색 시간이 더 많이 소요되었다.

인터페이스를 사용해서 효과음을 검색하는데 필요한 시간은 효과음의 종류에 따라서도 차이를 보여주었다. 평가 대상 효과음 60개(신 당 6개 X 10 신)를 효과음의 길이와 음향적 특성에 따라 분류해서 각각의 분류별로 효과음을 검색하는데 사용한 시간을 비교하면 [그림 27]과 같다.



[그림 27] 효과음 분류별 평균 효과음 검색 시간

분류 1 : 충격성 효과음

분류 2 : 2초 이하의 비충격성 효과음

분류 3 : 2초 이상의 비충격성 효과음

대조군 인터페이스를 사용해서 효과음을 검색하는 경우 검색 대상 효과음의 길이와 충격성/비충격성 효과음인지 여부에 따라 효과음 검색 시간에 차이가 있었다. 대조군 인터페이스에서는 소리의 크기가 실시간으로 시각화된 정보를 바탕으로 효과음을 검색하

기 때문에 (분류 1)의 충격성 소음의 경우 시각화된 소리 크기 정보가 순간적으로 커졌다가 줄어드는 것을 보고 다른 소리들에 비해 쉽게 찾는다는 것을 알 수 있었다. 대조군 인터페이스를 통해 효과음을 찾는데 가장 오래 걸리는 효과음은 (분류 2)의 2초 이하 비충격성 효과음이다. 비충격성 효과음은 시각화되는 정보를 보고 찾고자 하는 효과음인지 여부를 바로 판단하기 어려운데 효과음 재생 시간까지 짧은 경우 평균 효과음 검색 시간이 많이 늘어났다. 본 연구에서 제안한 효과음 검색 인터페이스를 활용한 경우 효과음의 종류가 변하더라도 검색 시간에 큰 차이가 없었다.

2. 정성적 평가

제안 인터페이스의 정성적 평가는 사용성 평가를 시행한 이후 피험자들을 대상으로 인터뷰를 진행하는 방법으로 이루어졌다. 본 질문 항목에 대한 인터뷰 코딩 기준은 아래와 같다.

인터뷰 코딩 기준(Coding Manual) - 항목 1

제안된 효과음 검색 인터페이스가 디지털 음향 믹싱 콘솔에 장착되어 나온다면, 얼마나 자주 사용하겠는가?

1) 대부분의 효과음 검색에 사용

- 대부분의 효과음 검색에 사용하겠다고 이야기한 경우
- 효과음이 어디 위치해 있는지 암기하고 있는 경우를 제외하면 효과음 검색에 사용하겠다고 이야기한 경우

2) 효과음 검색에 자주 사용

- 효과음 검색에 자주 사용하겠다고 이야기한 경우
- 사용 빈도를 묻는 추가 질문에서 반 정도의 상황에서 사용하겠다고 이야기한 경우

3) 효과음 검색에 가끔 사용

- 효과음 검색에 가끔 사용하겠다고 이야기한 경우
- 사용 빈도를 묻는 추가 질문에서 4번 중 한번 정도의 상황에서 사용하겠다고 이야기한 경우

4) 거의 사용 안함

- 효과음 검색에 사용하지 않겠다고 이야기한 경우

인터뷰 코딩 기준 - 항목 2

제안된 효과음 검색 인터페이스가 디지털 음향 믹싱 콘솔에 장착되어 나온다면, 얼마나 유용하게 사용하겠는가?

1) 매우 유용함

- 매우/아주 유용할 것이라고 이야기한 경우
- 언제쯤 상용화 가능한지, 지금 장비에 바로 적용 가능한지를 물어보는 등 직접적인 사용 의사 표시

2) 유용함

- 유용할 것이라고 이야기한 경우
- 간접적인 사용 의사 표시

3) 별로 유용하지 않음

- 별로 유용하지 않을 것이라고 이야기한 경우
- 기능이 있으면 사용해 보겠지만 특별한 관심은 없다고 이야기하는 경우

4) 전혀 유용하지 않음

- 전혀 유용하지 않을 것이라고 이야기한 경우

인터뷰 코딩 기준 - 항목 3

제안된 효과음 검색 인터페이스가 효과음을 얼마나 직관적으로 검색하여 준다고 느꼈는가?

1) 직관적이다

- 본 질문에 직관적이라고 답변하고, 이후 추가 질문에서도 이 직관성이 인터페이스를 처음 접한 다음부터 시간이 지난 이후까지 직관적이라고 답변한 경우

2) 처음에는 직관적이거나 점점 직관성이 떨어졌다

- 본 질문에 직관적이지 않다고 답변하고, 이후 추가 질문에서 인터페이스를 처음 접했을 때는 직관적이라고 생각했으나 인터페이스를 사용하면서 점점 직관적이지 않다는 느낌이나 판단하기 어려웠다고 답변한 경우

3) 직관적이지 않다

- 본 질문에 직관적이지 않다고 답변하고, 이후 추가 질문에서도 인터페이스를 처음 접한 다음부터 시간이 지난 이후까지 직관적이지 않다고 답변한 경우

실험 후 인터뷰 결과는 아래의 [표 7]과 같다.

	항목 1	항목 2	항목 3
피험자 1	가끔 사용	매우 유용함	처음에는 직관적, 점점 직관성 떨어짐
피험자 2	자주 사용	매우 유용함	처음에는 직관적, 점점 직관성 떨어짐
피험자 3	가끔 사용	매우 유용함	직관적

[표 7] 항목별 인터뷰 결과

실험 후 인터뷰 결과 본 연구에서 제안하는 인터페이스의 유용성과 개선해야 할 점을 파악할 수 있었다. 인터뷰는 만약 이런 음향 검색 인터페이스가 현재의 음향 믹싱 콘솔에 존재한다면 이를 얼마나 유용하게 사용할지, 얼마나 자주 사용할지를 묻는 질문과 제안한 인터페이스가 얼마나 직관적인지를 묻는 기본 질문을 토대로 진행했다.

평가 대상자 3명 모두 해당 본 연구에서 제안하는 인터페이스가 실제 음향 믹싱 콘솔에 적용되는 경우 효과음 검색에 도움이 될지를 묻는 질문에 매우 유용할 것이라는 답변을 했다. 해당 효과음 검색 인터페이스의 활용 빈도를 묻는 질문에는 피험자 중 한 명은 자주 사용하겠다는 답변을 했고 다른 두 명은 가끔 사용할 것이라고 답변했다.

관련 답변에 대한 추가 인터뷰에서 그 이유를 알 수 있었다. 효과음이 어느 트랙에 위치했는지를 검색 / 탐색하는 많은 경우에는 해당 효과음이 어디쯤에 위치하고 있는지를 어느 정도 예상하고

있는 상태에서 찾기 시작하기 때문에 이를 찾는데 큰 어려움이 없지만 가끔 어느 트랙에 있는지 예상하지 못하는 소리를 찾아야 하는 경우가 있다고 답변했다. 이렇게 쉽게 예상하지 못하는 경우에는 믹싱 과정을 멈추고 해당 소리가 나는 트랙을 찾는데 상당한 시간을 소비하는 경우에 유용하게 사용할 것이라는 답변이 있었다.

영화 음향 믹싱 단계는 처음 편집된 결과물을 접한 다음 각 음향 신 별로 기본적인 믹싱 작업을 하는 단계와 여러 신들을 연속해서 재생하면서 문제가 있는 효과음들을 찾아 작업하는 단계, 전체 영화를 재생하면서 문제가 있는 효과음들을 찾아 작업하는 단계 등으로 나뉘는데 처음 단계에서는 한 신을 계속 반복해서 작업하기 때문에 각 효과음들이 어디에 위치해 있는지 어느 정도 암기가 가능해서 제안하는 인터페이스가 활용되지 않을 것이라고 답변했다. 하지만 나머지 두 단계에서는 효과음의 위치를 암기하는 것이 불가능하기 때문에 제안된 인터페이스가 자주 활용될 것이라고 답변했다.

본 연구에서 제안한 인터페이스의 직관성을 묻는 질문에서 피험자 #3 은 직관적이었다고 답변했지만 피험자 #1 과 피험자 #2 는 인터페이스를 처음 접했을 때는 직관적이었으나 점점 직관적이지 못하는 답변과 판단하기 어려웠다는 답변을 했다. 이후 후속 질문들을 통해 그 이유를 질의한 결과 처음에는 직관적이라고 느꼈지만 시간이 지날수록 소리의 어떤 특성을 표현해야 하는지를 파악하지 못해서 답답한 느낌을 가지게 되었다고 답변했다.

터치 슬라이더를 사용해서 소리를 표현하라고 하는 모호한 지시 대신 소리의 크기를 표현하라거나 소리의 피치를 표현하라고 지시 했으면 좋겠다는 답변도 있었다. 이런 답변은 여러 음향 특성들을 활용해서 사람의 소리 표현을 잘 나타내도록 하는 기존 소리 투사 연구들이 인터페이스를 제작하는데 직접적인 도움을 주지 않을 수 있다는 것을 의미한다.

제 5 장 연구 의의 및 논의

본 연구에서 제안하는 터치 슬라이더를 이용한 효과음 검색 인터페이스는 영화 음향 믹싱 엔지니어가 자신이 변화시키고자 하는 효과음의 위치를 찾는 시간을 줄여줄 것이다. 이는 제한된 시간 안에서 음향 콘텐츠를 제작해야 하는 현재의 영화 음향 믹싱 산업에서 만들어지는 콘텐츠의 질을 높이는데 기여할 것으로 예상된다. 또한 영화 음향 믹싱을 위해 사용하는 음향 콘솔에 요구되는 채널 수를 줄여 영화 음향 믹싱을 공부하는 학생들의 진입 장벽을 낮춰줄 것이다.

본 연구는 음향 믹싱 엔지니어를 돕기 위한 기존 연구들이 음향 시각화를 통해 효과음 탐색을 보조하는 방향으로 발전하고 있는 것에 비해 음원을 질의를 통해 직접 검색하도록 제안한 첫 연구이다. 또한 음원 투사 관련 연구 결과를 처음으로 실제 인터페이스에 적용해서 평가했다.

본 연구는 터치 슬라이더를 사용하는 제한된 형태의 소리 투사 연구에서 음향 특성으로 음압 변화 곡선, 소리 크기 변화 곡선, 주파수 중심 변화 곡선, 원점 건넌 밀도 변화 곡선을 제안하고 움직임 특성으로는 위치, 위치의 1차 미분 결과, 위치의 2차 미분 결과를 제안했다. 터치 슬라이더를 활용한 소리 투사 실험을 진행해서 터치 슬라이더를 사용한 소리 투사는 기존의 자유로운 소리 투사와 다르게 터치 슬라이더 상의 절대적인 위치와 음향 특성에 연관이 없고 그 변화량인 시간에 대한 1차 미분 결과가 유의미한 연관이 있음을 찾아냈다. 이런 제한된 소리 투사에 관한 연구는 현재

발전하고 있는 휴대형 태블릿 등 모바일 환경의 음향 저작 도구에서 소리를 검색하거나 합성하는데 응용될 수 있을 것이다.

소리 투사 실험 결과를 바탕으로 터치 슬라이더를 활용한 효과음 검색 인터페이스를 제안하고 시제품을 구현해서 사용자 평가를 진행했다. 사용자 평가 결과 본 연구에서 제안한 효과음 검색 인터페이스가 같은 수의 채널을 사용한 대조군 효과음 검색 인터페이스에 비해 효과음 검색 시간이 평균 28.4 초에서 16.2 초로 유의미하게 줄어드는 것을 확인했다. 이런 검색 시간 감소는 2 초 이하의 비충격성 효과음에서 가장 큰 것으로 나타났다. 또한 인터뷰를 통한 사용자 평가에서는 본 연구에서 제시한 효과음 검색 인터페이스가 상용화될 경우 효과음이 어느 트랙에 위치해 있는지 암기하지 못한 믹싱의 중, 후반 단계에서 매우 유용하게 사용될 수 있을 것이라는 평가를 받았다.

본 연구에서는 인터페이스의 학습 효과에 대해 고려하지 못했다. 그 결과 인터페이스의 정성적 평가에서 알 수 있는 것처럼 사용자가 특정 소리를 검색하기 위해 터치 슬라이더를 어떻게 움직여야 하는지에 대한 의문이 들도록 만들었다. 소리 크기와 주파수 중심 변화 등 여러 특성이 융합된 알고리즘을 바탕으로 소리를 검색하도록 하는 본 연구에서 제안하는 알고리즘은 사용자의 첫 시도에서 원하는 소리를 검색하도록 하는 것에는 도움이 되었지만 이후 정확도가 늘지 않았다. 즉, 소리 특성과 움직임 특성의 관계를 가장 잘 설명하는 모델을 활용하는 것이 음향 검색 인터페이스 등에 활용되었을 경우 사용자에게 가장 직관적이지 않을 수 있다는 것을 알 수 있었다. 앞으로 음향 투사 관련 연구가 실제 인터페이스 형태로 활용되는 경우에 참고해야 할 것이다. 후속 연구에

서는 소리의 크기 또는 주파수 성분을 표현하라고 직접적으로 지시하고 이를 활용하는 효과음 검색 인터페이스를 제작하여 본 연구에서 제안한 효과음 검색 인터페이스와 사용성을 비교하는 실험을 진행할 계획이다.

후속 연구를 통해 기존 효과음 검색 인터페이스와 본 연구에서 제안한 효과음 검색 인터페이스를 같이 활용하는 복합적인 효과음 검색 인터페이스를 제작하면 본 연구의 연구 결과를 개선할 수 있을 것으로 보인다. 본 연구에서 제안한 효과음 인터페이스를 단독으로 사용하는 경우 한 음향 신을 반복적으로 믹싱하는 초기 믹싱 단계에서 사용자가 음향 신 내 효과음들의 트랙 위치를 전혀 암기하지 못해서 한 신을 반복하는 과정에서 효과음 검색에 필요한 시간이 줄어들지 않는 현상이 발견되었다. 이런 문제점을 해결하기 위해서 채널 줄의 일부는 기존 효과음 검색 인터페이스의 페이더뱅크 방식을 유지하고 나머지 부분만 새로운 효과음 검색 인터페이스를 사용하기 위한 채널 줄로 제공하면 음향 믹싱 엔지니어들의 효과음 검색 과정을 더 효과적으로 도와줄 수 있을 것으로 생각된다.

본 연구 자체는 영화 음향 믹싱 엔지니어가 처한 환경과 음향 믹싱에 사용하는 효과음들을 대상으로 진행되어 직접 활용 범위는 넓지 않지만 응용 범위는 넓다. 앞으로 후속 연구를 통해 더 폭넓은 사용자의 요구에 맞는 효과음 검색 시스템으로 발전될 수 있을 것으로 예상된다.

참 고 문 헌

- [1] Theodore M B. Audio Mixer Sound Instrument I.D. Panel. 1997.
- [2] The AVID webpage. www.avid.com. Accessed 26 Mar. 2017.
- [3] Rice, Stephen V. "Frequency-based coloring of the waveform display to facilitate audio editing and retrieval." Audio Engineering Society Convention 119. Audio Engineering Society, 2005.
- [4] Gohlke, Kristian, et al. "Track displays in DAW software: Beyond waveform views." Audio Engineering Society Convention 128. Audio Engineering Society, 2010.
- [5] Loviscach, J*rn. "The Quintessence of a Waveform: Focus and Context for Audio Track Displays." Audio Engineering Society Convention 130. Audio Engineering Society, 2011.
- [6] Jang, Jyh-Shing Roger, Hong-Ru Lee, and Chia-Hui Yeh. "Query by tapping: A new paradigm for content-based music retrieval from acoustic input." Pacific-Rim Conference on Multimedia. Springer Berlin Heidelberg, 2001.
- [7] Hanna, Pierre, and Matthias Robine. "Query by tapping system based on alignment algorithm." Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on. IEEE, 2009.

- [8] Pauws, Steffen. "CubyHum: a fully operational" query by humming" system." ISMIR. 2002.
- [9] Knees, Peter, and Kristina Andersen. "Searching for Audio by Sketching Mental Images of Sound: A Brave New Idea for Audio Retrieval in Creative Music Production." Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval. ACM, 2016.
- [10] Cartwright, Mark, and Bryan Pardo. "SynthAssist: Querying an Audio Synthesizer by Vocal Imitation." NIME. 2014.
- [11] Zhang, Yichi, and Zhiyao Duan. "Imisound: An unsupervised system for sound query by vocal imitation." Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on. IEEE, 2016.
- [12] Nymoen, Kristian, et al. "A statistical approach to analyzing sound tracings." Speech, sound and music processing: Embracing research in India. Springer Berlin Heidelberg, 2012. 120-145.
- [13] Caramiaux, Baptiste, Frédéric Bevilacqua, and Norbert Schnell. "Sound selection by gestures." New Interfaces for Musical Expression (NIME). 2011.
- [14] Godøy, Rolf Inge, Egil Haga, and Alexander Refsum Jensenius. "Exploring music-related gestures by sound-tracing: A preliminary study." (2006): 27-33.

- [15] Caramiaux, Baptiste, et al. "The role of sound source perception in gestural sound description." *ACM Transactions on Applied Perception (TAP)* 11.1 (2014): 1.
- [16] Kozak, Mariusz, Kristian Nymoen, and Rolf Inge Godøy. "Effects of spectral features of sound on gesture type and timing." *International Gesture Workshop*. Springer Berlin Heidelberg, 2011.
- [17] Rocchesso, Davide, et al. "Sketching sound with voice and gesture." *interactions* 22.1 (2015): 38-41.
- [18] Cycling 74 webpage. www.cycling74.com. Accessed 26 Mar. 2017.

Abstract

Audio Search Interface Using Touch Sensitive Sliders

Wansoo Kim

Digital Contents and Information Studies

The Graduate School

Seoul National University

Film sound mixing engineers spend a lot of time looking for the sound they want to transform. This limits the current film production sound mixing engineers who have to create content in a limited time. There has been a development to better visualize sound sources for each track in the audio console used in the sound mixing process, but there is a limit to helping movie sound mixing engineers who uses a very large number of tracks.

On the other hand, research on sound effects through various queries such as voices, tapping, and body movements has been carried out although it is not a study for helping a movie sound mixing engineer.

The purpose of this study is to develop a sound projection study using a touch slider and to develop a novel interface that movie sound mixing engineers can use to search sound effects and to evaluate their usability.

keywords : Touch Slider, Audio Interface, Sound Effect Search, Audio Mixing Engineer, Audio Mixing Console

Student Number : 2016 - 26030