

3D영상에 정합되는 스테레오 오디오

Stereo Audio Matched with 3D Video

박성욱 · 정태윤

Sung-Wook Park and Tae-Yun Chung

강릉원주대학교 전자공학과

요 약

본 연구에서는 동일한 내용의 영상을 2D로 감상할 때와 3D로 감상할 때 함께 재생되는 음향이 어떻게 달라져야 하는지를 확인하는 주관적 실험을 수행하고 그 결과를 고찰하였다. 먼저 음향 정보는 음원이 자체적으로 제공하는 정보인 음원의 거리와 방위각(즉 위치) 그리고 음원의 환경 혹은 장면(scene)이 제공하는 정보인 공간감으로 분리가 가능하므로 이에 맞게 동일 내용의 2D/3D 영상이 음원의 위치 선정에 미치는 영향 평가 실험과 동일한 내용의 2D/3D 장면이 음향 공간감에 주는 영향 평가 실험을 수행하였다. 첫 번째 실험 결과 3D 영상을 감상하는 경우 2D 영상을 감상할 때 보다 스크린을 기준으로 음원의 거리와 방위각을 확대하여 인지한다는 결과를 얻을 수 있었다. 이는 2D 영상용 소리보다 거리와 방위각이 큰 3D 영상용 소리를 만들어야 한다는 것을 의미한다. 또한 3D 영상용 소리는 3D 영상뿐만 아니라 2D 영상과도 잘 어울린다는 결과를 얻었다. 두 번째 실험 결과, 3D 영상을 감상하는 경우 2D 영상을 감상할 때 보다 잔향이 더 많은 소리를 선호함을 알 수 있었다. 이는 3D 영상을 감상할 때 공간감이 강화되기 때문으로 해석된다. 본 연구의 결과는 기본적으로 2D 영상용 음향을 제작하던 음향엔지니어가 3D영상용 음향을 제작하는 데 활용할 수 있으며, 2D to 3D 음향을 자동으로 변형하는 연구의 기초가 될 것이다. 더 나아가서 본 연구의 결과를 기반으로 제한된 대역폭에서 2D 와 3D를 동시에 지원하는 방송 시스템을 설계하는데 적용해 본다면, 방송 데이터 규격은 스테레오 영상, 음원의 위치가 강조된 3D 음향과 공간감을 주는 잔향 정보로 구성하는 것이 적절하다고 할 수 있다.

키워드 : 3D 영상, 3D 음향, 3D 방송, 음원의 위치, 공간감, 2D to 3D 변환

Abstract

This paper presents subjective experimental results to understand how audio should be changed when a video clip is watched in 3D than 2D. This paper divided auditory perceptual information into two categories; distance and azimuth that a sound source contributes mostly, and spaciousness that scene or environment contribute mostly. According to the experiment for distance and azimuth, i.e. sound localization, we found that distance and azimuth of sound sources were magnified when heard with 3D than 2D video. This lead us to conclude 3D sound for localization should be designed to have more distance and azimuth than 2D sound. Also we found 3D sound are preferred to be played with not only 3D video clip but also 2D video clip. According to the experiment for spaciousness, we found people prefer sound with more reverberation when they watch 3D video clips than 2D video clips. This can be understood that 3D video provides more spacial information than 2D video. Those subjective experimental results can help audio engineer familiar with 2D audio to create 3D audio, and be fundamental information of future research to make 2D to 3D audio conversion system. Furthermore when designing 3D broadcasting system with limited bandwidth and with 2D TV supportive, we propose to consider transmitting stereoscopic video, audio with enhanced localization, and metadata for TV sets to generate reverberation for spaciousness.

Key Words : 3D Video, 3D Audio, 3D broadcasting, sound localization, spaciousness, 2D to 3D conversion

1. 서 론

음향 및 영상과 관련된 기술은 많은 발전을 하여 일반 사용자들이 가정에서 현실감 있는 음향과 화면을 시청하는 것이 가능하게 되었다. TV의 경우 과거 아날로그 방식의 TV 보다 고화질을 제공하는 디지털 TV로 발전하였고, 최근에는 보다 현장감을 실감나게 전달하기 위하여 공간에 물체가

보이도록 재현하는, 즉 3D 영상을 보여 주는 3D TV가 등장하였다[1].

사실 3D 기술은 일찍부터 존재하였지만 실용화에 번번히 실패해왔다. 실패의 원인으로는 콘텐츠의 부족, anaglyph 기술에 기반한 조악한 화질, 그리고 불완전한 2D to 3D 자동변환기술의 적용을 들 수 있다. 하지만 최근 들어 콘텐츠 개발자[2]의 새로운 사업영역개발의 필요성, 그리고 기술의 발전이 결합하면서 다시 시장에 소개되고 있다.

현재 3D TV는 보급의 초기단계이며, 가정에서는 3D TV 보다는 2D TV를 더 많이 보유하고 있다. 그러므로 아직까지 기존의 지상파 방송국의 경우 3D TV용 콘텐츠를 송출하지 않고 있으며, 3D TV를 보유한 가정에서는 3D 전용

접수일자 : 2011년 2월 19일

완료일자 : 2011년 4월 15일

이 논문은 2010년도 강릉원주대학교 학술연구조성비 지원에 의하여 수행되었음

위성/케이블 방송 혹은 Blu-ray 플레이어를 이용하여 3D 콘텐츠를 시청하여야 한다. 3D 산업의 활성화를 위해서는 3D 콘텐츠가 시장에 풍부하게 공급되는 것이 바람직하는데, 이를 위해서 송출된 하나의 콘텐츠가 3D TV 뿐만 아니라 기존의 2D TV에서도 시청할 수 있도록 하는 기술은 중요하다. 하나의 3D 콘텐츠가 3D TV와 2D TV에서 모두 시청 가능하게 하는, 즉 2D TV에서의 호환성을 가지게 하는 방법으로 3D 영상을 저장하고 전송할 때는, 좌측 눈을 위한 영상과 우측 눈을 위한 영상을 따로 저장 혹은 전송하였다가 2D TV에서는 좌측 혹은 우측 영상 하나만을 재생하고, 3D TV에서는 좌측과 우측 영상을 모두 재생하는 기법을 사용한다[3][4][5]. 하지만 이 연구들은 모두 영상의 호환성은 고려하고 있지만, 음향에 대해서는 고려하고 있지 않다.

본 논문에서는 동일한 내용의 영상이라도 3D로 재생하는 것과 2D로 재생하는 경우 다른 공간 정보를 사용자에게 전달하므로, 이때 각 재생방법에 따라 음향이 어떻게 달라져야 하는지를 실험을 통하여 확인하고자 한다. 그리고 이러한 결과를 이용하여 2D TV와 3D TV를 모두 지원하는 방송은 음향 데이터를 어떤 형식으로 송출하는 것이 좋은지에 대해 의견을 제시하고자 한다.

2. 기존 연구

3D TV, 특히 대화면을 가지는 3D TV가 출현한 것이 얼마 되지 않았기 때문에 3D TV에 적합한 3D audio에 대한 체계적인 연구는 찾아보기 어렵다. Cepric Andre[6]은 여러 사람이 영상을 감상하는 극장에서 3D 음향을 가장 잘 재생할 수 있는 재생 시스템에 대한 연구를 수행하였다. 연구 결과 헤드폰을 이용한 양이(binaural) 재생기술이 가장 적합하다는 결론을 제시하였다. Cepric의 연구는 음장(sound field)를 얼마나 원음에 가깝게 재생하느냐를 중심으로 연구하였고, 3D영상을 위한 음향과 2D영상을 위한 음향이 어떻게 다른지에 대한 고찰은 하지 않았다.

Setzu[7]는 HD TV 화면에서 인물이 중앙에 보일 때, 실험 참가자에게 여러 방위각에서 발생한 소리를 들려주고 소리와 화면이 불일치하는 방위각이 얼마인지를 알아보는 실험을 수행하였다. 그 결과 전문가의 경우 약 10도, 일반인의 경우 약 20도 정도 차이가 나는 소리를 들려주어도 크게 불편함을 느끼지 못했다는 결과를 제시하였다. Setzu의 연구는 2D TV에 대한 연구로서 3D TV에 대한 연구는 진행하지 않았다.

Werner[8][9]는 실물크기의 인물이 있는 2D 화면에서 물체의 거리, 방향을 잘 표현할 수 있는 오디오 재생 시스템이 어떤 것인지를 평가하였다. 또한 영상을 녹화할 때 사용한 시점(viewpoint) 위치가 아닌 다른 위치에서 화면을 보고 스테레오 스피커로 재생되는 소리를 듣는다면, 시점 위치에서 화면을 보면서 듣는 소리와 다른 결과를 보임을 실험으로 확인하였다. 이 실험 결과를 확장하여 시점이 바뀌는 다른 재생환경, 즉 2D 영상을 시청하다가 3D 영상을 시청하면, 소리도 이를 반영하여 변경되어야 한다는 의미로 해석할 수 있다.

3. 실험 및 결과

본 연구에서는 먼저 2D영상과 3D영상으로 인지된 시각적 공간정보가 음향으로 인지된 청각적 공간정보에 얼마나

영향을 미치는지 실험을 통하여 알아보려 하였다. 만일 사용자가 2D영상을 보면서 청취하는 음향과 3D영상을 보면서 청취하는 음향이 서로 달라야 한다고 응답한다면 이를 이용하여 2D화면을 3D화면으로 시청할 때 어떤 음향 정보를 얼마만큼 보정하면 보다 적절한 음향 청취가 가능하게 될지 알 수 있을 것이다.

공간 정보를 표현하는 음향의 인지속성(perceptual attribute)은 문헌에 따라 여러 가지 분류되고 또한 다른 용어로 소개되고 있는데[10][11] 본 연구에서는 음원(sound object or sound source)이 가지는 속성과 장면(sound scene or environment)이 가지는 속성으로 분리하여 고찰하였다 [9]. 이 중에서도 음원 속성과 관련하여, 기존 연구[6][7][11]를 참조하여, 수평 위치와 거리를 주요 속성으로 선정하여 관찰하였고, 장면 속성과 관련하여 공간감(spaciousness)을 주요 속성으로 선정하여 관찰하였다. 이와 같이 음향의 인지속성을 분리하여 관찰한 것은 2D 영상을 3D 영상으로 표현하면 물체의 위치감, 거리감 뿐만 아니라 깊이감 및 공간감이 강화된다는 것을 고려한 것이다.

실험은 그림 1과 같이 일반적인 가정에서의 TV 시청 환경과 유사하도록 대형화면을 중앙에 두고 스테레오 스피커로 소리를 재생하는 환경을 조성하여 수행하였다.

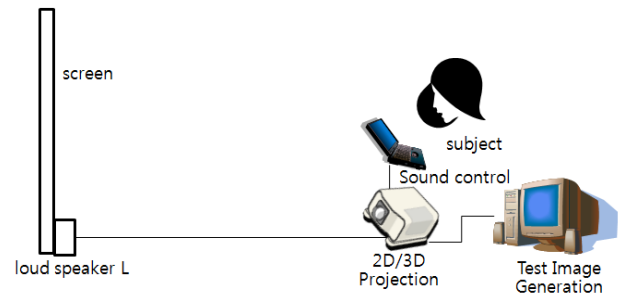


그림 1. 실험 환경

Fig. 1. Experimental Environment

3.1 영상의 차이가 음원의 위치 선정에 미치는 영향에 대한 실험

2D영상과 3D영상을 시청할 때 음원 자체가 제공하는 속성인 수평 위치와 거리가 얼마나 다르게 인지되는지를 확인하기 위한 실험을 수행하였다.

실험은 2D영상과 3D영상에 대하여 13명의 참가자가 7 가지 위치에 대하여 2회씩 랜덤한 순서(총364회=2(2D/3D영상)*13명*7위치*2회)로 실시하였다. 실험 방법은 실험 참가자에게 먼저 2D영상으로 인물 한명이 등장하는 화면을 랜덤하게 14회(=7위치*2회)보여 주고, 참가자가 직접 스테레오 스피커에서 나는 소리를 그 물체와 동일한 위치에 설정하도록 요구하였다. 2D 영상을 이용한 실험을 완료한 이후 3D 영상을 이용하여 동일한 실험을 수행하였다. 소리는 5초 분량의 한 문장을 남성 아나운서가 읽는 것이며, 소리의 위치와 거리를 조정하기 위하여 단순히 음압만 조정하였다.

그림 2는 본 실험에서 사용된 영상들 중에 하나를 보여 준다. 실험에 사용된 영상은 한 사람이 의자에 앉아 있는 영상으로서 그림 3 과 같이 모두 7 가지 위치에서 3D로 촬영하였다. 2D 영상은 3D영상에서 좌측 영상을 사용하였다.

각 위치의 좌표값은 표 2에 표시하였다. 이 위치들은 차이를 식별할 수 있는 정도[12]인 방위각 약 10도씩의 차이가 있도록 설정하였고, 거리는 3D영상으로 보았을 때 1번과

2번은 화면에서 돌출하도록, 3-4번은 화면에 위치하고, 5-7 번은 화면 뒤에 위치하도록 선정하였다.



그림 2. 3D 영상을 만들기 위한 좌안영상(좌측)과 우안영상(우측) (6번 위치에 해당)

Fig. 2. Image for left eye (left) and image for right eye (right) for 3D which is for #6 position

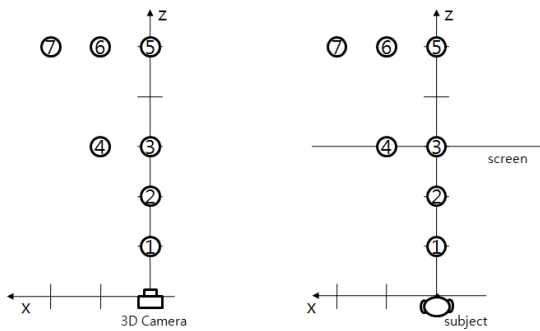


그림 3. 7 가지 인물 위치 (좌측: 촬영, 우측: 재생)
Fig. 3. 7 positions of a person (left: positions in recording, right: positions in reproduction)

표 1 실험 화면에서 인물의 위치
Table 1 Location of visual objects

번호	x좌표(m)	z좌표(m)	거리(m)	방위각(도)	비고
1	0	1	1	0	화면 전방
2	0	2	2	0	화면 전방
3	0	3	3	0	화면
4	-1	3	3.16	18.43	방위각 약20도
5	0	5	5	0	화면 후방
6	-1	5	5.10	11.31	방위각 약10도
7	-2	5	5.39	21.80	방위각 약20도

실험의 신뢰도를 향상하기 위하여 제시된 화면이외의 시각 정보를 배제하였다. 이를 위하여 참가자는 소리를 위치시킬 때 소리를 발생시키는 컴퓨터의 화살표 버튼을 보지 않고 눌러서 소리의 위치를 옮길 수 있는 프로그램을 작성하여 사용하였다. 또한 매 실험 수행 직전에 기준이 되는 화면과 그 화면에서의 소리를 먼저 들려주어 이를 기억하고 위치를 선정할 화면과 소리를 들려주어 비교할 수 있도록 하였다. 본 실험에서는 3번 위치를 기준으로 사용하였다.

3.2 영상의 차이가 음원의 위치 선정에 미치는 영향에 대한 실험 결과

그림 4는 2D 혹은 3D 영상이 주어졌을 때 이를 보고, 영상 내 인물의 위치와 동일하다고 생각되는 장소에 음원을 위치시킨 결과의 평균을 표시한 것이다. 3D 영상을 보고 7

가지 위치에 대하여 응답한 결과에는 원 표식을 겹쳐 2D 영상에 대한 결과와 구분하였다. 이 결과를 보면 3D 영상을 보고 응답한 결과 위치는 2D 영상을 보고 응답한 결과 위치에 비하여 기준 위치 3번보다 거리가 멀거나 방위각이 큰 4, 5, 6, 7 번에서는 거리와 방위각이 더욱 크다. 3번 위치가 전후 방향으로 0 인, 스크린에 맺히는 영상임을 감안한다면 3D 로 영상을 감상하면 2D로 감상할 때보다 소리도 더 멀리 위치시켜야 한다는 것을 알 수 있다.

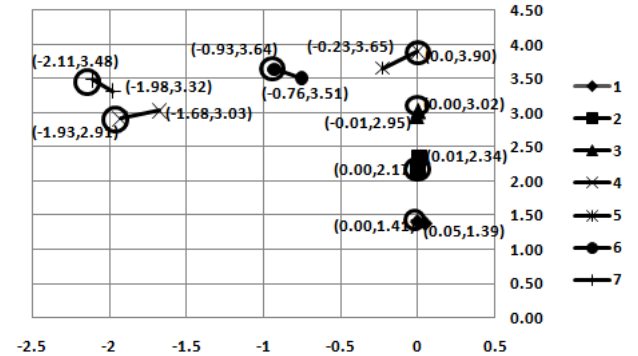


그림 4. 3D 영상을 보고 응답한 음원의 평균위치(원 표시)와 2D 영상에 응답한 결과 비교

Fig. 4. Comparison of average responses of sound positions when heard with 3D video(marked with empty circles) and 2D video

기준 위치 3번 보다 거리가 작은 1, 2번의 경우는 거리가 좀 더 짧거나 거의 동일하다. 거리가 거의 동일한 1번 위치의 경우는 실험에서 일반 가정과 유사한 음향 재생환경으로 사용한 스테레오 시스템이 음압만을 제어해서는 스피커보다 앞으로 음상을 위치시키지 못하기 때문에 이런 결과를 얻은 것으로 생각할 수 있다. 하지만 2번 위치에서의 반응을 고려한다면 3D 영상을 감상할 때 2D 영상 보다 상대적으로 앞으로 소리를 위치시키는 것이 필요함을 알 수 있다.

그림 5는 수평 방향(x축)에 대하여 실험 참가자들이 2D 혹은 3D 영상이 주어졌을 때 이를 보고, 영상 내 인물의 위치와 동일하다고 생각되는 장소에 음원을 위치시킨 결과의 분산을 나타낸 것이다. 분산들을 보면, 전반적으로 3D 영상을 보았을 때의 값이 2D 영상을 보았을 때의 값보다 작아서 실험 참가자들이 소리의 위치를 보다 일관성 있게 선정하였음을 알 수 있다. 즉 3D 영상이 2D 영상보다 다수의 실험 참가자에게 공간 정보를 안정적으로 제공하므로 3D 영상과 함께 재생되는 음향의 수평 위치 정보를 결정할 때 보다 적절히 결정하여야 할 것이다. 상대적으로 2D 영상과 함께 재생되는 음향의 경우 위치 정보가 부적절하게 선정되더라도 시청자가 영상과의 이질감 등을 잘 느끼지 못할 가능성이 많다고 볼 수 있다.

3D 영상에서는, 평균적으로 보았을 때, 인물의 방위각이 기준으로부터 커질 수 록 (1, 2, 3, 5 → 6 → 4 → 7) 분산이 커지는 경향을 보이는데, 이는 실험에서 제공된 기준 소리의 위치인 3번 위치와 멀어질 수 록 응답의 일관성이 떨어진 결과로 볼 수 있다.

그림 6은 전후 방향(z축)에 대하여 실험참가자들이 2D 혹은 3D 영상이 주어졌을 때 이를 보고, 영상 내 인물의 위치와 동일하다고 생각되는 장소에 음원을 위치시키고 그 값의 평균과 분산을 나타낸 것이다. 전후 방향의 분산을 살펴

보면 x축의 경우와 동일하게 3D 영상을 보았을 때 2D 영상을 보고 전후 위치를 선정했을 때 보다 작은 분산을 가지므로 3D 영상을 보았을 때 보다 일관성 있는 응답을 한다고 볼 수 있었다. 그러므로 3D영상과 함께 제공되는 음향의 전후 정보는 2D 영상과 함께 제공되는 음향 보다 적절하게 선정되어야 할 것이다.

2D 영상과 3D 영상을 통틀어 인물의 위치가 기준 위치인 3번에서 멀어 질수록 분산이 커지는 것을 볼 수 있다. 1번, 2번의 경우 4-7번 경우 보다 분산이 큰 데, 이는 스테레오 재생시스템이 음원을 스피커 앞으로 재생하는 능력이 좋지 않기 때문으로 해석할 수 있다.

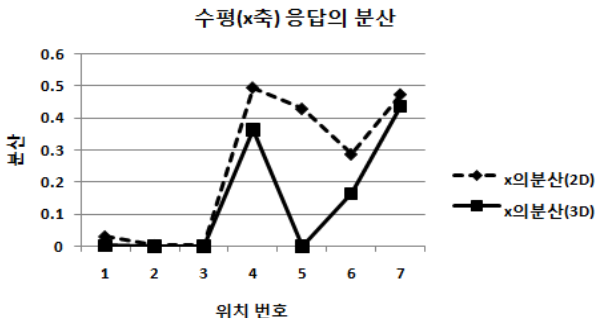


그림 5. 2D/3D영상에서 x축에 대한 음원 위치선정실험 결과

Fig. 5. Experimental result of sound localization in x axis under 2D video and 3D video

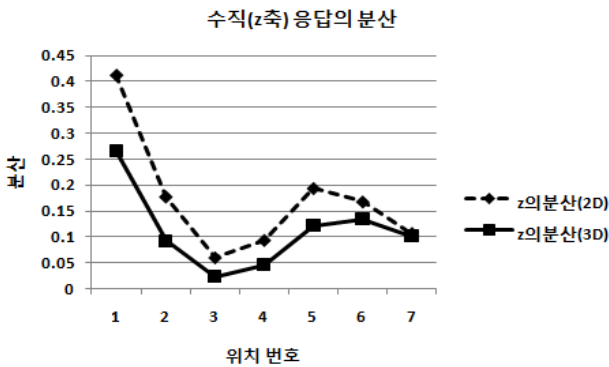


그림 6. 2D/3D영상에서 z축에 대한 음원 위치선정실험 결과

Fig. 6. Experimental result of sound localization in z axis under 2D video and 3D video

3.3 영상의 차이가 장면과 관련된 음향 공간감에 미치는 영향에 대한 실험

2D영상과 3D영상을 시청할 때 장면(scene)과 관련된 속성인 공간감이 얼마나 다르게 인지되는지를 확인하기 위한 실험을 수행하였다. 공간감은 반향(reflection)과 잔향(reverberation)에 의해서 결정되는데[13] 본 연구에서는 잔향을 이용하여 공간감을 표현하였다.

실험은 2D영상과 3D영상으로 준비된 4 가지 화면에 대하여 각각 10명의 참가자가 화면에 가장 적절한 음향을 선정하는 방식으로 수행되었다(총80회=2(2D/3D영상)*4가지 영상*10명). 실험 참가자에게 들려준 음향은 서로 다른 공간감을 표현하기 위해 잔향의 크기가 조정된 4 가지 소리로

서 남성 아나운서가 14초 분량의 두 문장을 읽은 것이다. 표 3는 실험에 사용된 소리의 잔향 특성을 나타낸다. 잔향에 의한 공간감은 직접음(direct sound)과의 상대적 크기에 의해서 결정된다. 실험에 사용된 직접음은 하나이므로 고정되었다고 한다면, 실내(enclosed space)의 경우, 잔향의 크기가 커면 넓은 공간 그리고 멀리서 들리는 소리의 느낌이 나며, 잔향의 크기가 작으면 좁은 공간 그리고 가까이서 나는 소리의 느낌이 난다[12]. 실험에서 사용한 소리는 원음에 추가로 잔향을 부가한 것으로써 열린 공간의 경우에 해당하는 잔향, 평균적인 실내의 잔향, 평균 실내보다 조금 많은 잔향, 콘서트홀 정도의 잔향을 가지도록 설계하였다.

표 2 실험에 사용된 소리의 잔향 특성
Table 2. Reverberation added to 4 sounds

ID	잔향크기	비고
1	-	잔향 없음. 열린 공간 (open space)과 유사한 환경, 가까이서 들리는 느낌
2	-32.04dB	원음 크기의 2.5%. average room [?], 작은 공간, 가까이서 들리는 느낌
3	-26.02dB	원음 크기의 5%, 중간 공간
4	-18.06dB	원음 크기의 12.5%, reverberant room [7], 넓은 공간, 멀리서 들리는 느낌



그림 7. 2D/3D영상이 장면과 관련된 음향 공간감에 미치는 영향을 알기 위해 사용한 화면 (좌측상단부터 해변, 캠퍼스, 축구장, 댄스실)

Fig. 7. Video clips used to understand impact on spaciousness when watched together in 2D and 3D mode (beach, campus, soccer stadium, and dance room from top-left to bottom-right)

시각정보가 음향의 깊이감과 공간감에 미치는 영향을 확인하기 위하여 공간의 크기가 서로 다른 4가지 영상을 준비하였다. 그림 7는 본 실험에서 사용된 영상들을 보여 준다. 실험에 사용된 영상은 크고 열린 공간을 나타내는 해변 장면, 작고 열린 공간을 나타내는 캠퍼스, 크고 닫힌 공간을 나타내는 축구장 그리고 작고 닫힌 공간을 나타내는 댄스실로 구성하였다.

3.4 영상의 차이가 장면과 관련된 음향 공간감에 미치는 영향에 대한 실험 결과

각 장면 별로 2D 영상으로 보았을 때와 3D 영상으로 보았을 때, 그 화면과 가장 어울린다고 느껴지는 소리를 선택하였고, 그 결과를 그림 8에 나타내었다. 그림 8의 각 그

래프는 하나의 장면에 대한 응답으로, 특정한 소리를 몇 명의 실험 참가자가 선호했는지를 2D 영상과 3D 영상으로 구분하여 표시하였다.

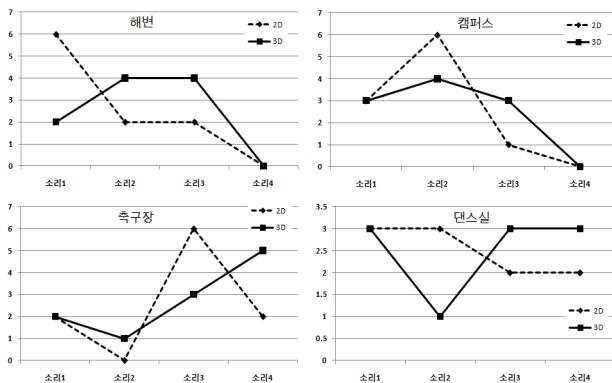


그림 8. 영상의 차이가 장면과 관련된 음향 공간감에 미치는 영향에 대한 실험 결과(좌측상단부터 해변, 캠퍼스, 축구장, 댄스실의 결과임)

Fig. 8. Experimental results regarding spaciousness (beach, campus, soccer stadium, dance room from top-left to bottom-right)

열려 있으면서 넓은 공간인 해변 영상의 경우, 비교적 잔향의 크기가 작은 소리에 대한 선호가 높았다. 특히 2D 영상의 경우에는 열린 공간의 느낌을 많이 주는 소리1에 대한 선호가 많았다. 3D 영상의 경우 소리2와 3에 대한 선호도가 높았는데 실험 후 인터뷰에 따르면 2D 영상은 평면적이어서 잔향이 느껴지지 않는 소리가 어울렸지만 3D 영상은 입체적이어서 잔향이 조금 있는 소리가 더 어울린다는 의견을 보였다. 열려 있으면서 좁은 공간인 캠퍼스 영상의 경우, 어느 정도 잔향이 있는 소리 2에 대한 선호가 높았다. 2D 영상의 경우 소리2에 대한 선호가 뚜렷했지만, 3D 영상의 경우 소리1, 2, 3에 대해 고른 선호를 보였다. 이는 열리고 작은 공간이라는 상충되는 복합적인 정보가 전달되어 실험 참가자마다 중점적으로 선택한 정보가 달라지면서 나타난 현상으로 보인다.

크고 닫힌 공간인 축구장의 경우 2D 영상에서는 소리 3을 3D 영상에서는 소리 4를 선호하였다. 이는 실험 참가자들이 기본적으로 장면이 주는 내용에 영향을 받아 2D 영상과 3D 영상을 막론하고 잔향이 있는 소리를 선호하지만 3D 영상의 경우 강화된 입체감으로 보다 잔향이 큰 소리 4를 선호한다고 해석할 수 있다. 작고 닫힌 공간인 댄스실의 경우 2D 영상에서는 소리1과 2에 대한 선호도가 높았으며, 3D 영상에서는 선호도가 소리1, 3, 4로 복합적으로 나타났다. 실험 후 인터뷰 결과에 따르면 3D영상을 소리1을 선호한 부류는 인물이 커서 가까운 느낌이 나는 잔향이 없는 소리를 선호했다고 응답한 반면, 소리3, 4를 선호한 부류는 닫힌 공간이어서 잔향이 많은 소리를 선호했다는 응답을 하였다. 2D영상에서 보다 3D 영상에서 소리3, 4를 선호한 경우가 많다는 사실을 기초하여 추론해 볼 때 3D 영상의 경우 강화된 공간 정보가 실험 응답자들에게 전달되었다고 생각할 수 있다.

4. 고 찰

영상의 차이가 음원의 위치 선정에 미치는 영향에 대한

실험 결과를 종합해 보면 실험 참가자들은 3D 영상을 보았을 때 2D 영상을 보았을 때보다 음원의 위치에 대하여 일관된 반응을 보였으며, 음원의 위치 선정에 있어서 3D 영상을 보았을 때 2D 영상을 보았을 때보다 거리는 스크린에서 멀수록 더 멀어지고 방위각은 기준 위치에서 멀수록 더 커지는 결과를 얻을 수 있었다. 이는 2D 영상용 소리보다 거리와 방위각이 큰 3D 영상용 소리를 만들어야 한다는 것을 알 수 있다.

하나의 데이터로 2D와 3D를 모두 지원하는 3DTV 방송 환경에서 단지 하나의 음향만을 전송할 수 있다면 어떤 음향을 선정하는 것이 더 많은 시청자에게 만족감을 주는지를 알아보기 위한 추가 실험을 수행하였다. 음원의 위치를 선정하는 실험을 통하여 얻은 평균값들을 10명을 대상으로 2D영상과 3D영상과 함께 보여주고 선호도를 조사하였으며, 그림 9와 같은 결과를 얻었다. 실험 결과에 따르면 3D 영상을 보고 있을 때에는 3D 영상을 위해서 준비된 음향에 대한 선호도가 50%로 가장 높았다. 2D 영상을 보고 있을 때에는 2D 영상이 아니라 3D 영상을 위해 준비된 음향에 대한 선호도가 36%로 가장 높게 나왔다. 이는 2D 영상을 이용하여 선정된 음원 위치가 높은 분산을 가지고 있어 실험 참가자는 특정한 위치를 크게 선호하지 않는다고 해석할 수 있다. 그러므로 3D 영상을 보고 선정했던 음원의 위치가 주어지더라도 큰 거부감을 느끼지 못하는 것으로 해석된다. 그러므로 2D와 3D를 모두 지원하는 방송을 하는 경우, 하나의 음향만을 전송해야한다면 3D 영상을 위해 만들어진 음향을 전송하는 것이 보다 많은 시청자들이 선호할 것이라고 예상할 수 있다.

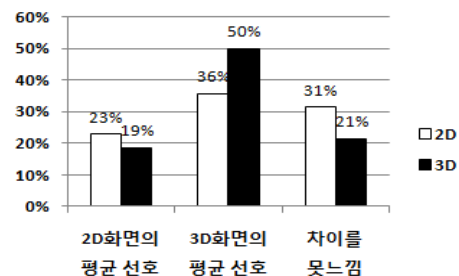


그림 9. 실험 결과를 이용한 선호도 조사

Fig. 9. Preference for the experimental results

영상의 차이가 장면의 공간감에 미치는 영향에 대한 실험의 결과를 종합적으로 살펴보면 2D 영상을 보았을 때 보다 3D 영상을 보았을 때 좀 더 잔향의 크기가 큰 소리를 선호했음을 알 수 있다. 이는 장면의 종류에 관계없이 일관적으로 보이는 결과이다. 그러므로 2D와 3D를 모두 지원하는 방송을 하는 경우, 하나의 음향만을 전송해야한다면 잔향은 2D 영상을 위한 음향을 전송하고 3D 영상을 위해서는 영상을 재생할 때 TV단에서 화면에 부합하는 만큼의 잔향을 만들어 음향을 재생하는 방법을 사용할 수 있다. 이와 같은 재생을 위하여 특정한 시간에 얼마만큼의 잔향이 필요한지 정보를 추가하여 전송하면 2D TV와 3D TV에서 모두 볼 수 있는 방송이 가능하다.

이 두 실험의 결과를 통합해서 생각해 본다면 2D TV와 3D TV를 모두 지원하는 방송을 할 때에는 3D 영상에 부합하는 위치 정보를 가지는 음향과, 3D TV로 시청할 때 추가할 잔향 정보를 송출하는 것이 적절하다고 볼 수 있다.

5. 결 론

본 연구에서는 동일한 내용의 영상을 2D로 시청할 때와 3D로 시청할 때 함께 재생되는 음향이 어떻게 달라져야 하는지를 알기 위한 실험을 수행하였다. 실험을 위하여 음향 정보를 음원이 제공하는 정보인 음원의 거리와 방위각(위치) 그리고 음원의 환경 혹은 장면이 제공하는 정보인 공간감으로 분리하여 2D 영상과 3D 영상이 음향의 인지에 얼마나 다른 영향을 미치는지를 실험으로 확인하였다. 동일 내용의 2D/3D 영상이 음원의 위치 선정에 미치는 영향 평가 실험 결과, 3D 영상을 감상하는 경우 2D 영상을 감상할 때보다 음원의 위치 선정에 있어서 거리는 스크린에서 멀수록 더 멀어지고 방위각은 기준 위치에서 멀수록 더 커지는 결과를 얻을 수 있었다. 이는 2D 영상용 소리보다 거리와 방위각이 큰 3D 영상용 소리를 만들어야 한다는 것을 알 수 있다. 또한 3D 영상용 소리는 3D 영상 뿐만 아니라 2D 영상과도 잘 어울린다는 결과를 얻었다. 동일한 내용의 2D/3D 장면이 음향 공간감에 주는 영향 평가 실험에서는 3D 영상을 감상하는 경우 공간감이 강화되어 2D 영상을 감상할 때 보다 잔향이 더 많은 소리를 선호함을 알 수 있었다. 이와 같은 결과를 기반으로 2D와 3D를 동시에 지원하는 방송에서 하나의 음향만을 전송한다면, 이러한 방송용 데이터는 3D 영상에 부합하는 위치 정보를 가지는 음향 하나와 단말기에서 영상 재생 모드에 맞게 장면에 따른 잔향을 조절할 수 있는 형식으로 정의하는 것이 적절할 것이다. 또한 본 연구의 결과를 2D-to-3D 음향자동변환 기술을 개발하는데 도움이 될 것으로 기대된다.

참 고 문 헌

[1] Ben Drawbaugh, "3D stole the show at CES 2010", engadgetHD, Available: <http://hd.engadget.com/2010/01/21/3d-stole-the-show-at-ces-2010/> Jan 21, 2010.

[2] 정선영, "헐리우드, 3D영화관에 거액 몰린다, 아시아경제, Available: <http://www.asiae.co.kr/news/view.htm?idxno=2009091202011730600>, 2009년 3월 12일.

[3] Blu-ray Association, *White Paper Blu-ray Disc Format 2.B Audio Visual Application Format Specifications for BD-ROM Version 2.4*, Available: http://www.blu-raydisc.com/assets/Downloadablefile/BD-ROM-AV-WhitePaper_100423-17830.pdf, April, 2010.

[4] 신흥창, 정광희, 정원식, 허남호, 이수인, "3DTV 방송기술 개발 및 산업화 동향", *TTA Journal*, no. 127, pp. 53-60, 2010.

[5] 박홍식, 김규현, 이권희, 윤국진, 서덕영, 박광훈, "스테레오스코픽 비디오 방송 서비스를 위한 MPEG-2 전송스트림 구성 방안," *방송공학회논문지*, 제14권, 6호, pp. 1-14, 2009.

[6] Cepic Andre, Jean-Jacques Embrechts, and Jacques G. Verly, "Adding 3D Sound to 3D Cinema: Identification and Evaluation of Different Reproduction Techniques," *ICALIP*, pp. 130-137, 2010.

[7] Setsu Komiyama, "Subjective Evaluation of

Angular Displacement between Picture and Sound Directions for HDTV Sound Systems," *J. Audio Eng. Soc.*, vol. 37, no. 4, pp. 210-214, April 1989.

[8] Werner P.J. de Bruijn and Marnus M. Boone, "Application of Wave Field Synthesis in life-size videoconferencing," *114th AES convention(paper 5801)*, March 22-15, 2003, Amsterdam, The Netherlands.

[9] Werner P.J. de Bruijn and Marnus M. Boone, "Subjective Experiments on the effects of combining spatialized audio and 2D Video projection in audio-visual systems," *112th AES convention(paper 5582)*, May 10-13, 2002, Munich, The Netherlands.

[10] Helmut Wittek, "Perceptual differences between wavefield synthesis and stereophony," *Ph.D Thesis, University of Surrey*, pp. 11-12, October 2007.

[11] Francis Rumsey, *SPATIAL AUDIO*, Focal Press, 2001.

[12] Setsu Komiyama, "Subjective Evaluation of Angular Displacement between Picture and Sound Directions for HDTV Sound Systems," *J. Audio Eng. Soc.*, vol. 37, no. 4, pp. 210-214, April 1989.

[13] Jens Blauert, *Spatial Hearing: the psychophysics of human sound localization*, MIT Press, 1997.

저 자 소 개



박성욱(Sung-Wook Park)

1995년 2월 : 연세대 전자공학과 석사
 1998년 9월 : 연세대 전자공학과 박사
 2009년 2월 : 삼성전자 DM연구소 수석
 2009년 ~ 현재 : 강릉원주대 전자공학과 조교수

관심분야 : 3D오디오/비디오, 스마트그리드, 센서네트워크
 Phone : +82-33-640-2388
 Fax : +82-33-646-0740
 E-mail : swpark@gwnu.ac.kr



정태윤(Tae-Yun Chung)

1989년 2월 : 연세대 전기공학과 석사
 2000년 2월 : 연세대전기컴퓨터공학과박사
 1989년 3월 ~ 2001년 2월 : 삼성전자 중앙연구소 책임
 2001년 3월 ~ 현재 : 강릉원주대 전자공학과 부교수
 2004년 9월 ~ 현재 : 강원임베디드소프트웨어 연구센터 센터장

관심분야 : 임베디드 시스템, 영상부호화, 센서네트워크
 Phone : +82-33-640-2425
 Fax : +82-33-646-0740
 E-mail : tychung@gwnu.ac.kr