

A Study on Recording Techniques for Describing
Frontal Images with High Separation Between
Sound Sources in Multichannel Audio-Focusing
on the OCT (Optimized Cardioid Triangle)
Method

I. Introduction

1.1. Research Background

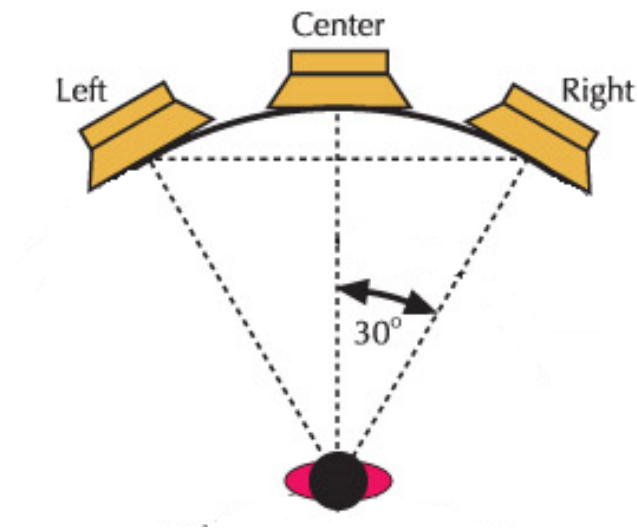
1.1.1. The Role and Correlation of Front Channels in Surround Sound

The range in which a person can accept information composed of sound is from sound coming from all directions surrounding the two ears. Today's advances in audio technology have taken a leap to provide realistic sounds that people can practically accommodate beyond gradual development of Two-channel audio system. With introduction of various playback technologies, and surround and 3D recording technologies that can be integrated with them, there have been sound receiving and methodological researches to realistically contain spatial information without distortion. [1][2]

In 1880s, the Two Channel Audio System were first demonstrated by Clément Ader, and since the announcement of the fact by Carl Stumpf, German philosopher and psychologist in 1916, that both ears of human perceive different information, respectively, the auditory image based on the 'psychological sound' of how human perceives the sound information from today's surround and immersive audio systems has been an important criterion for the development of audio technology to reenact this. According to the R&D engineers of BBC (British Broadcasting Corporation), if there is any omitted or distorted part of the audio during playback that is regarded as important information for human perception of space, it can difficult to experience realistic virtual reality as much as the loss due to this [3]. As such, for surround and 3D audio it is important for playback system and recording to provide auditory experiences similar to actual spaces based on 'psychological sound' beyond two channel audio system.

In a realistic audio playback system, in order to describe the results of

recording and playback similar to the auditory experience in real space, the criteria of correlation between the playback system and recording technology must be properly defined and matched. [4] [5] The front channel of the standard monitor of the multi-channel audio according to the ITU-R BS 774-1 scheme, which is the speaker array of the surround system, which is the background of the present study, is building a system with continuous arrangement of three speakers at 30° intervals [Figure 1]. It quotes the stereo main pair angle of a two-channel audio system as it is, but differs from stereophonic due to the addition of a center speaker, and fundamentally changes the perceived characteristics of the frontal sound region, so it should be understood as a continuous space divided by Front L and Center, Center and Front R. Therefore, in order to experience in-phase listening to the listener through the surround system, the microphone array for recording the information on the front face must also be a technique in which the perceived characteristics of the sound region on the front face are considered.



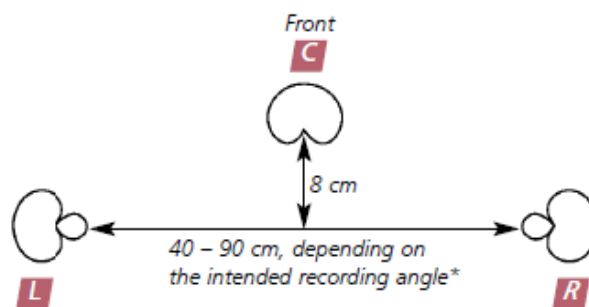
[Figure 1] Front Channel Configuration of ITU-R BS 774-1

1.1.2. Characteristics Discrete of Surround Recording Method OCT

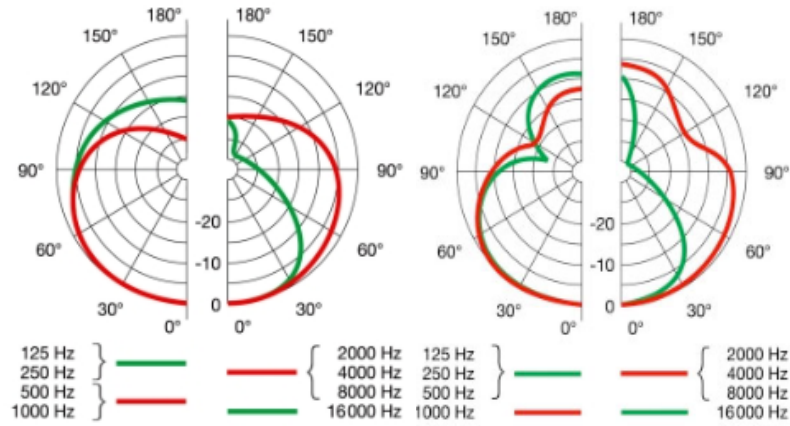
(Optimized Cardioid Triangle)

Using conventional OCT (Optimized Cardioid Triangle) [6] [7] recording methods, the time difference and level difference between the front L and center microphones, and the center and front R microphones can be derived continuously, so this is the ideal recording method for a surround playback system [Figure 2].

When the location information of sound sources located in front of a certain space can be recorded and played back with high separation, a multi-channel audio monitor environment can be provided with an auditory experience similar to that in a real space. Recording and playback using the OCT method create high separation through two main features. Firstly, different polar pattern cardioid and super cardioid microphones are arranged in the center and L-R [Figure 3]. This produces a distinct and continuous level change between center and L or between center and R. Secondly, set the offset angle between the L and R microphones to 180° . This is to record the location information of the sound source in front by using the level attenuation characteristic by the polar pattern from the front lobe to the rear lobe from the super cardioid microphone of L-R forming the level difference with the cardioid microphone of the center [Figure 3], allowing you to clearly record and describe a wider range of content than the normal angle for the front channel.



[Figure 2] OCT setup



[Figure 3] Polar diagram characteristics of AKG C414 xl ii microphone by frequency (left, cardioid right, super cardioid)

1.2. Purpose of Study

The purpose of this study is to propose a recording technique for high-definition sound source description between front channels (L-C-R) in relation to the surround microphone technique for multichannel audio. To do this, the polar pattern and 0° direction of the L and R microphones are changed from the optimized cardioid triangle (OCT) microphone array of existing Scheeps company. After that, the level difference between the L and the center channel in each method is analyzed and the hearing evaluation is conducted to evaluate the objectified superiority.

II. Main Subject

2.1. Experiment

2.1.1 Experiment Overview

This study tried to find an excellent method in terms of the separation of sound sources among front channels (L-C-R) of multi-channel audio standard monitors, which is the recommended method of ITU-R BS 774-1 with additionally recording through total of three modified methods together with the OCT method as shown in [2]. The method used here is as follows. First, it changed the polar pattern of the L and R microphones of the OCT method from super cardioid to cardioid, and second, it changed the 0° direction of the first method of the L-R microphone to 0° at $\pm 90^\circ$ respectively. AKG C414 ULS cardioids were used for the center microphone, and AKG C414 xl ii cardioid and super cardioid polar patterns were recorded for the L-R microphones.

The sound source used in the experiment was a sine wave of 1 kHz, a clock second hand sound, and a recorded vocal having different characteristics and patterns [Table 1]. These sources were reproduced with 4W output speakers, amplified by Micstasy of RME company, and recorded at AD (Analog to Digital) conversion at 16Bit with a sampling frequency of 44.1kHz through the Protocols HD Native audio interface.

	Type of sound source	Pattern change	Recording length
1.	Sine wave	none	10s
2	Clock second hand	regular	10s
3	Vocal	Irregular	10s

[Table 1] Types and Features of Sound Sources Used in Experiments

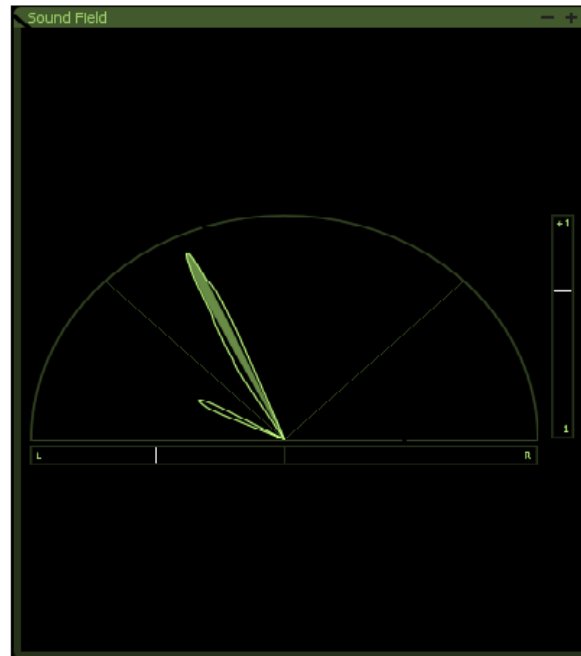
2.1.1 Experimental sequence

Measure the size of the space to avoid highlighting a particular wave reflected by the recording environment as a stationary wave. Several test recordings were made to determine the proper microphone setup position. The microphones were set up at 2.8m width, 3.5m length, and 1.7m in a space of 3.301m width, 6.947m length and 2.508m height.

② Determine the amplification value of each microphone. The microphones used in the experiments were cardioid and super cardioid microphones, and the sensitivity was measured by radiating white noise at the size of 96dB SPL at 0° direction, 1m distance of each microphone. Then, determine the amplification value of the preamplifier so that the measured result becomes the output level of -18dB SF of each microphone.

③ Select the position to play the sound sources. In this experiment, the distance between the L-R microphones was set to 40cm, and the position of 45° from the center microphone was selected as the playback position of the sound source. On the same line of 45°, selected the initial point of about 1.2m at the monitor speaker where the sound image becomes 45° by moving the speaker backward as shown in [Figure 4], through the phase scope.

④ The sine wave 1kHz was played and recorded, and the clock second hand sound and the recorded vocal sound were reproduced and recorded at the same time. For each sound source, a total of three modified methods were additionally recorded along with the OCT method. The structure of each method is shown in [Table 2]. First, it is the same as OCT method of [Figure 2] and implemented as [Figure 5] for experiment. Second, the 0° direction of the L and R microphones was kept the same as the OCT method, and the polar pattern was recorded by changing from cardioid to super cardioid. Third, the structure of the existing OCT microphones was maintained, and the 0° direction of the L and R microphones were recorded so as to face the front at +/- 90° [Figure 6]. Fourth, the 0° direction of the L and R microphones was changed to face the front, and the polar pattern was changed from super cardioid to cardioid for recording.



[Figure 4] Check that the sound image of sine wave is 45° through phase scope

	Direction of 0°	Center	L, R	Setup
Method 1	+ /- 90°	Cardioid	Super Cardioid	
Method 2	+ /- 90°	Cardioid	Cardioid	
Method 3	0°	Cardioid	Super Cardioid	
Method 4	0°	Cardioid	Cardioid	

[Table 2] Characteristics of Modified Method for Experiment



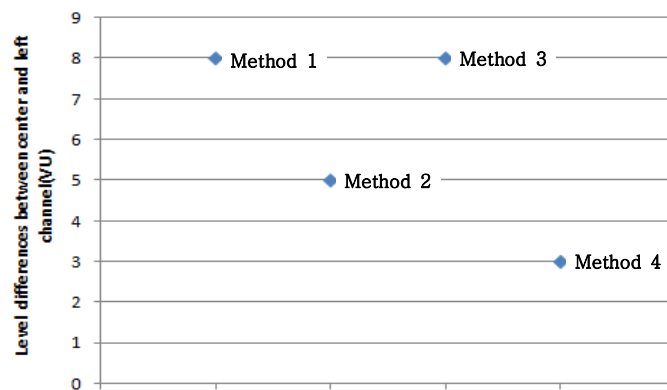
[Figure 5] Implementation of OCT method



[Figure 6] OCT type Super cardioid is transformed to face front at $\pm 90^\circ$,
and Super cardioid is transformed to Cardioid in the fourth type

2.1.2. Experiment result

For example, this study compared and analyzed the level difference of the recorded results among L-C channels in total of 4 kinds recording methods, including OCT based on the fact that in the case of stereo microphone technique XY using only the level difference, the greater the level difference between the channels for a sound source, the higher the separation between the sound sources. The level of each channel was measured by VU meter, and the result of level difference between OCT and L-C of each recording method is shown in [Figure 7]. For the OCT method, the sine wave level measured at the center channel is -18dB, the L channel level is -26dB, and the level difference between the center and L channels is 8dB. The levels of the center channels from modified modes 2 to 4 are -16 dB, -16 dB, and -18 dB, respectively, and the levels of the L channels are measured as -21 dB, -24 dB, and -21 dB, respectively, and the level difference due to this between the center and L channel is 5dB, 8dB and 3dB, respectively. Method 1, which is an OCT microphone array type, and Method 3, which changed only the 0° direction of the microphone, showed the same level difference, and Method 2 had a relatively low level difference of 3 dB.



[Figure 7] Variation of level difference of sine wave 1kHz between center and L channel for each method in [Table 2]

2.2. Hearing evaluation

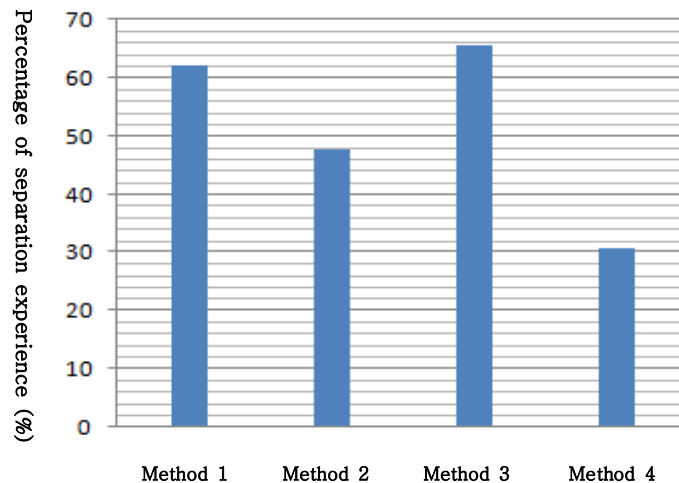
2.2.1 Evaluation method

As for the total of four recording methods of this experiment to describe the separation between the sound sources in the front channels of the surround monitor system hearing evaluation were conducted for the objective evaluation of the superiority. This study used recorded sound source by playing back the speakers at 45° and about 1.2m from the center microphone of the OCT microphone array for the vocal sound source clock second sound whose rhythm and pitch pattern change of sound level is regular and irregular vocal sound source. The sound source was edited with the results of the four recording methods of the experiment in 7-second intervals and arranged at 0.5-second intervals, and the order of each edited sound source was randomly arranged and a total of four sets were used for evaluation.

To assess the degree of separation between sound sources, the subjects were 10 persons who major music. In addition, before the evaluation progressed, the degree of separation of the sound source was fully explained, and the evaluation was conducted to increase the objectivity of the evaluation result. Experiments were carried out to listen to the results of four randomly reproduced recordings of each set and number them in the order of perceived high separation between sources. In addition, the level of individual feeling of the degree of separation of the sound sources was evaluated from 1 to 5, and an absolute evaluation between the recording methods applied in the experiment was attempted.

2.2.2. Evaluation results

[Figure 8] shows the results of evaluation of individual sensory levels as a percentage of the results of the four recording methods applied in the experiment. A relatively large number of subjects experienced a high degree of separation through the method 1 corresponding to the OCT microphone array and the method 3 in which the configuration of the microphone was maintained while the 0° direction was faced in the OCT method. On the other hand, although the sensory level of separation between the results of Method 2 and Method 4 using only cardioid microphones showed a relatively low change of less than 15%, in absolute evaluation, it was confirmed that the sound sources can be experienced with a maximum of 47.7% of separation. In addition, it was confirmed through an interview that three of the subjects who evaluated that the results of Method 1 and Method 3 provided the highest degree of separation were evaluation for the degree of separation experience from 1 to 2 out of 5, which was evaluated by distinguishing only slight differences between the results of each method.



[Figure 8] Percentage of separation experience for the four methods used in the experiment

III. Conclusion

This study was studied for recording techniques in which sound sources are described with high separation when surround sound, which is realistic audio, is played back through multichannel audio. The experiment was conducted by changing the polar pattern of the L-R microphone and the direction of the central axis around the OCT microphone array where the front channels were recorded with high separation, and it is confirmed that the change of the axial direction to the front of the Method 1, OCT microphone array and Method 3, OCT type L-R microphone shows excellent separation in the front channels of the multi-channel audio standard monitor. Method 1 and Method 3 use cardioid microphones for C and super cardioid microphones for LR. This result shows that the level difference between L-C channel and C-R channel can be made more than 3dB than that of using cardioid microphone for L-R, and it is confirmed that the sensory level for the separation in multi-channel audio is also excellent as the level difference is big between channels.

Unlike the OCT method, the actual experiment result showed the same level difference of 8dB in method 3 for the sine wave 1kHz different from the inference that the level difference between L-C or C-R will be smaller than that of the OCT method due to the change in the position of the lobe where the sound source is mainly absorbed because the center axis of the L-R microphone faces the front. On the other hand, despite having the same level difference, the perceived level of separation in multi-channel audio was about 3% higher than that of the OCT method. This result is presumably due to the phase distortion caused by the rear lobe of the super cardioid microphone having a phase characteristic opposite to that of the front lobe.

Based on the above results, it is thought that the recording technique for describing high separation in multi-channel audio is affected not only by the level difference between channels but also by the phase characteristics of the microphone and psychological factors related to hearing, thus the recording technique considering this should be done continuously.

References

- [1] Michael Williams, UNIFIED THEORY OF MICROPHONE SYSTEMS for STEREOPHONIC SOUND RECORDING, Presented at AUDIO the 82nd Convention 1987 March 10-13 London
- [2] Michael Williams and Guillaume Le Du, MICROPHONE ARRAY ANALYSIS for MULTICHANNEL SOUND RECORDING, Presented at AUDIO the 107th Convention 1999 September 24-27 New York
- [3] Spatial Audio for Broadcast.
Helmut Witt Gunther Theile, The recording angle-based on localization curves, Presented at the 112th Convetion 2002 May 10-13 Munich, Germany
- [5] Michael Williams, The stereophonic zoom, "Sounds of sootland", BP50, 94364 BRY SUR MARNE CEDEX, France
- [6] SCHOEPS_surround_brochure (2006-11)
- [7] Schoeps Manual MAB 1000

멀티채널 오디오에서 음원간의 분리도가
높은 프론탈 이미지 묘사를 위한 레코딩
테크닉 연구 - OCT(Optimized Cardioid
Triangle) 방식을 중심으로

I. 서론

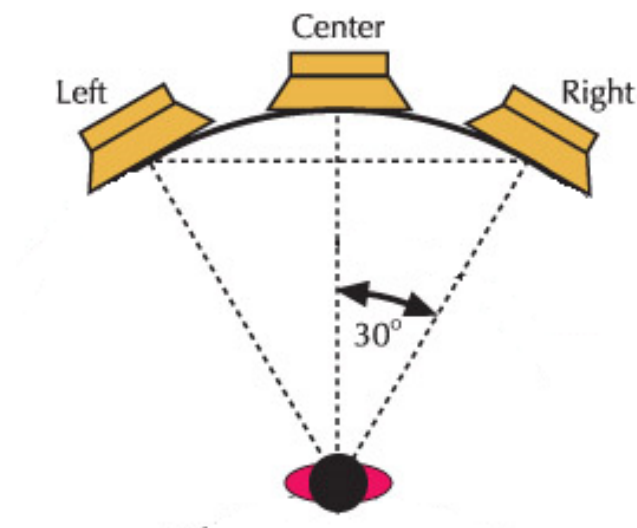
1.1. 연구 배경

1.1.1. 서라운드 사운드에서 프론트 채널들의 역할과 상호관계

사람이 소리로 이루어진 정보를 받아들일 수 있는 위치의 범위는 두 개의 귀를 둘러싼 모든 방향으로부터 오는 소리이다. 오늘날의 오디오 기술의 발전은 두 채널 오디오 시스템(Two-channel audio system)이 점차 발전되며 이를 넘어선 사람이 실질적으로 수용할 수 있는 만큼의 현실감 있는 소리를 제공하기까지 도약하였다. 다양한 재생기술과 이와 융합될 수 있는 서라운드(Surround) 및 3D 녹음 기술들이 소개되며 실감적으로 공간의 정보를 소리로 담아내기 위한 수음과 이를 왜곡 없이 재생시키기 위한 방법론적 연구가 이루어져 오고 있다.[1][2]

1880년대, Clément Ader에 의해 두 채널 오디오 시스템이 처음 시연되었고 1916년도 독일의 철학자이자 심리학자 Carl Stumpf에 의해 사람의 양쪽 귀는 각기 다른 정보를 인지한다는 사실이 발표된 이후 오늘날의 서라운드 및 실감형 오디오 시스템에 이르기까지 사람이 소리의 정보를 어떠한 이미지로 인지하는지에 대한 '심리음향'에 의거한 청각적 심상은 이를 재현하기 위한 것으로서 오디오 기술의 발전에 중요한 기준이 되어왔다. 영국의 공영 방송 BBC의 R&D 엔지니어들에 따르면 사람이 공간을 인지하기 위해 주요한 정보로 여기는 오디오(Spatial audio)가 재생과정에서 생략 및 왜곡되는 부분이 있다면 이로 인해 손실된 만큼의 실감형 가상현실을 경험하는 것이 어려울 수 있다는 사실을 밝힌 바 있다.[3] 그런 만큼 '심리음향'에 근거하여 재생시스템과 녹음이 실제 공간에서와 유사한 청각적 체험을 제공하는 것은 두 채널 오디오 시스템을 넘어 서라운드 및 3D 오디오를 위해 중요하다.

실감형 오디오 재생 시스템에서 녹음과 재생이 실제 공간에서의 청각적 체험과 유사한 결과를 묘사하기 위해선 재생 시스템과 녹음 기술 사이에 상관관계의 기준을 적절히 정한 후 매치시켜야 한다.[4][5] 이번 연구의 배경이 되는 서라운드 시스템의 스피커 어레이(Speaker array)인 ITU-R BS 774-1 방안에 따른 멀티채널오디오의 표준모니터의 프론트 채널(Front channel)은 3개의 스피커를 30°의 간격으로 연속적으로 배열하여 시스템을 구축하고 있다 [그림 1]. 이는 두 채널 오디오 시스템의 스테레오 메인 페어(Stereo main pair)의 각도를 그대로 인용하고 있지만 센터 스피커(Center speaker)가 추가됨으로 인해 스테레오포닉과는 다르며, 정면에 소리의 영역에 대해 인지되는 특성을 근본적으로 바꾸어 놓기 때문에 프론트 L과 센터, 센터와 프론트 R로 나눈 연속적인 공간으로서 이해해야 한다. 그러므로 서라운드 시스템을 통해 청자에게 인 페이즈한 청음을 경험시키기 위해선 정면의 정보를 녹음하기 위한 마이크로폰 어레이(Microphone array) 또한 필수적으로 정면에 소리의 영역에 대해 인지되는 특성이 고려되어야 한다.

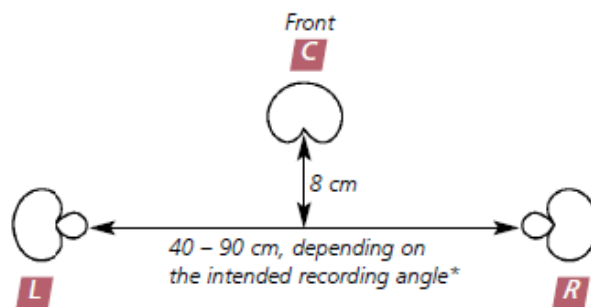


[그림 1] ITU-R BS 774-1 의 Front Channel Configuration

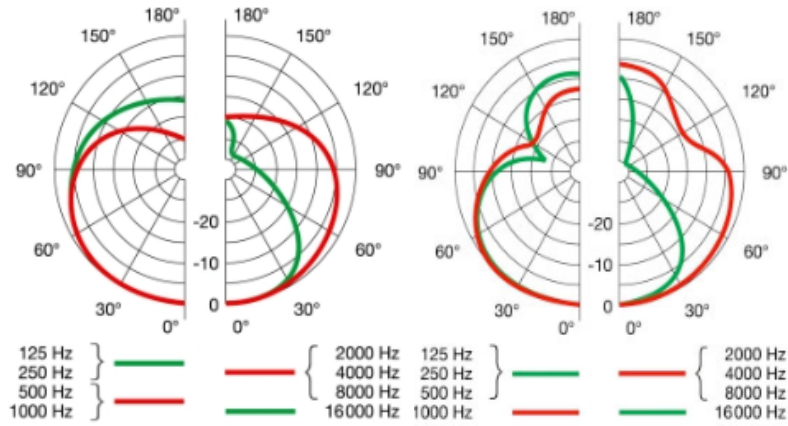
1.1.2. 서라운드 녹음 방식 OCT(Optimized Cardioid Triangle)의 특성 Discrete

기존의 OCT(Optimized Cardioid Triangle)[6][7] 녹음 방식을 사용하면 프론트 L과 센터 마이크로폰, 센터와 프론트 R 마이크로폰 간의 시간차(Time difference) 및 레벨차(Level difference)를 연속적으로 도출해 낼 수 있으므로 서라운드 재생 시스템을 위한 이상적인 녹음 방식이다 [그림 2].

임의의 공간에서 정면에 위치한 음원들의 위치 정보를 높은 분리도로 녹음 및 재생할 수 있게 되면 실제 공간에서 느끼는 청각적 경험을 멀티채널 오디오의 모니터 환경을 통해 보다 실제 공간에서와 유사한 청각적 체험을 제공할 수 있다. OCT 방식을 통한 녹음 및 재생은 크게 2가지 특징을 통해 높은 분리도를 만든다. 첫째로 센터와 L-R에 각기 다른 폴라 패턴의 카디오이드 마이크로폰과 슈퍼 카디오이드 마이크로폰을 배열하여 사용하는 것이다 [그림 3]. 이를 통해 센터와 L 또는 센터와 R사이에서 뚜렷하고 연속적인 레벨의 변화를 만들어 낸다. 둘째로 L과 R 마이크로폰 간의 Offset angle을 180° 로 두는 것이다. 이는 정면에 있는 음원의 위치정보를 센터의 카디오이드 마이크로폰과 레벨차를 형성하는 L-R의 슈퍼 카디오이드 마이크로폰으로부터 Front lobe부터 Rear lobe까지의 폴라 패턴에 의한 레벨 감쇠 특성[그림 3]을 함께 활용하여 녹음하기 위한 방법으로 프론트 채널에 대하여 일반각보다 광범위한 내용을 선명하게 녹음 및 묘사할 수 있도록 한다.



[그림 2] OCT setup



[그림 3] AKG C414 xl ii 마이크로폰의 주파수 별 폴라 다이어그램 특성
(좌, 카디오이드 우, 수퍼 카디오이드)

1.2. 연구 목적

본 연구는 멀티채널 오디오를 위한 서라운드 마이크로폰 테크닉과 관련하여 프론트 채널들(L-C-R) 사이에서 분리도(Discrete)가 높은 음원 묘사를 위한 레코딩 테크닉을 제시하는 것을 목적으로 한다. 이를 위해 기존 Scheops사의 OCT(Optimized Cardioid Triangle) 마이크로폰 어레이로부터 L과 R 마이크로폰의 폴라 패턴과 0°방향을 변경한다. 이후 각 방식들에서 나타나는 L과 센터 채널 간의 레벨차를 대조 분석하고 객관화된 우수성을 평가하기 위해 청음 평가를 진행한다.

II. 본론

2.1. 실험

2.1.1 실험 개요

본 연구에서는 [5]에서 권고한 [그림 2]와 같은 OCT 방식과 함께 총 3가지의 변형된 방식을 통해 추가적으로 녹음하여 ITU-R BS 774-1의 권고 방안인 멀티채널오디오 표준모니터 중 프론트 채널들(L-C-R)에서의 음원의 분리도와 관련하여 우수한 방식을 찾고자 하였다. 여기에 사용된 방식은 첫 번째, OCT 방식의 L과 R 마이크로폰의 폴라 패턴을 수퍼 카디오이드에서 카디오이드로 바꾼 것이며, 두 번째, 첫 번째 방식의 L-R 마이크로폰의 0° 방향을 각각 $\pm 90^\circ$ 에서 0°로 변형한 것이다. 센터 마이크로폰에는 AKG C414 ULS 카디오이드를 사용하였으며 L-R 마이크로폰으로는 AKG C414 xl ii의 카디오이드와 수퍼 카디오이드 폴라 패턴을 사용하여 녹음하였다.

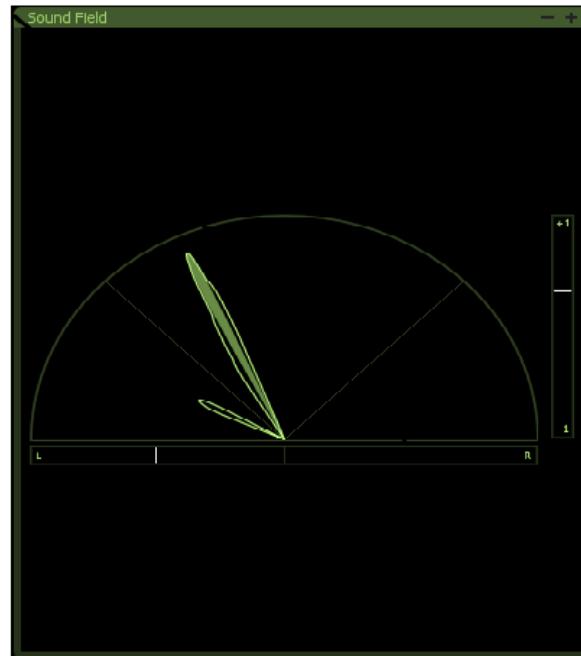
실험에 사용된 음원에는 1kHz의 사인파와 특성 및 패턴이 상이한 시계 초침 소리와 녹음된 보컬이 활용되었다 [표 1]. 이 음원들은 4W 출력의 스피커로 재생되었으며 이를 RME사의 Micastasy를 통해 증폭하여 Protools HD Native 오디오 인터페이스를 통해 44.1kHz의 샘플링 주파수와 16Bit로 AD(Analog to Digital) 변환하여 녹음하였다.

	음원의 종류	패턴 변화	녹음 길이
1	사인파	없음	10s
2	시계 초침	규칙적	10s
3	보컬	불규칙적	10s

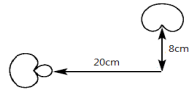
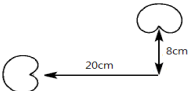
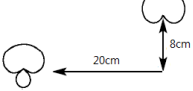
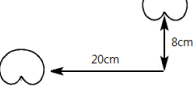
[표 1] 실험에 사용된 음원의 종류와 특징

2.1.1 실험 순서

- ① 녹음 환경에 의해 반사된 특정 웨이브가 이동이 고정(Stationary wave)되어 강조되는 것을 피하기 위해 공간의 규모를 측정한다. 몇 차례의 테스트 녹음을 통해 적절한 마이크로폰 셋업 위치를 결정하였다. 가로 3.301m, 세로 6.947m, 높이 2.508m인 공간에서 가로 2.8m, 세로 3.5m, 1.7m에 마이크로폰을 셋업 하였다.
- ② 각 마이크로폰들의 증폭 값을 결정한다. 실험에 사용 된 마이크로폰은 카디오이드와 수퍼 카디오이드 마이크로폰이며 각 마이크로폰의 0°방향, 1m의 거리에서 화이트 노이즈를 96dB SPL의 크기로 방사하며 감도를 측정하였다. 이후 측정된 결과 값이 각 마이크로폰의 아웃풋 레벨 -18dBFS이 되도록 Preamplifier의 증폭 값을 결정한다.
- ③ 음원들을 재생할 위치를 선택한다. 본 실험에서는 L-R 마이크로폰 간의 거리를 40cm로 두었으며 센터 마이크로폰을 기준으로 45°가 되는 지점을 음원의 재생 위치로 선택하였다. 45°의 동일 선상에서 스피커를 점점 뒤로 이동시켜 [그림 4]와 같이 페이즈 스코프(Phase scope)를 통해 모니터 스피커에서의 음상이 45°가 되는 최초의 기점, 약 1.2m를 음원을 재생시킬 위치로 선정하였다.
- ④ 사인과 1kHz를 재생하여 녹음하였고 시계 초침소리와 녹음된 보컬 음원은 동시에 재생하여 기록하였다. 각 음원에 대하여 OCT 방식과 함께 총 3가지의 변형된 방식을 추가적으로 녹음하였다. 각 방식의 구조는 [표 2]와 같다. 이는 첫 번째, [그림 2]의 OCT 방식과 동일하며 실험을 위해 [그림 5]와 같이 구현하였다. 두 번째, L과 R 마이크로폰의 0°방향은 OCT 방식과 동일하게 유지하였으며 폴라 패턴을 수퍼 카디오이드에서 카디오이드로 변경하여 녹음하였다. 세 번째, 기존 OCT 방식의 마이크로폰의 구성은 유지하였으며 L과 R 마이크로폰의 0°방향은 $\pm 90^\circ$ 에서 정면을 바라보도록 변경하여 녹음하였다 [그림 6]. 네 번째, L과 R 마이크로폰의 0°방향은 정면을 바라보도록 변경하였으며 폴라 패턴은 수퍼 카디오이드에서 카디오이드로 변경하여 녹음을 진행하였다.



[그림 4] 페이즈 스코프를 통해 사인 웨이브의 스피커에서의 음상이 45°임을 확인

	Direction of 0°	Center	L,R	Setup
방식 1	+/-90°	Cardioid	Super Cardioid	
방식 2	+/-90°	Cardioid	Cardioid	
방식 3	0°	Cardioid	Super Cardioid	
방식 4	0°	Cardioid	Cardioid	

[표 2] 실험을 위해 변형시킨 방식의 특징



[그림 5] OCT 방식 구현



[그림 6] OCT 방식의 수퍼 카디오이드의 0° 방향을 $\pm 90^\circ$ 에서 정면을 바라보도록 변형, 네 번째 방식에서는 수퍼 카디오이드를 카디오이드로 변형

2.1.2. 실험 결과

예를 들어 레벨차만을 이용한 스테레오 마이크로폰 테크닉 XY의 경우, 한 음원에 대한 채널간의 레벨차가 클수록 음원들 간의 분리도가 높아지는 사실들에 입각하여 OCT를 포함한 총 4가지 녹음 방식들의 L-C채널 간의 녹음된 결과물들의 레벨차를 대조 분석하였으며 명확한 대조를 위해 1kHz의 사인파로 분석하였다. 각 채널의 레벨은 VU미터를 통해 측정하였고 OCT 및 각 녹음 방식의 L-C간의 레벨차의 결과는 [그림 7]과 같다. OCT 방식에 대하여 센터 채널에서 측정된 사인파의 레벨은 -18dB이며 L 채널의 레벨은 -26dB로 센터와 L채널 간의 레벨차는 8dB이다. 변경된 방식들의 2부터 방식 4까지의 센터 채널들의 레벨은 각각 -16dB, -16dB, -18dB이며 L 채널들의 레벨은 각각 -21dB, -24dB, -21dB로 측정되어 이로 인한 센터와 L 채널 간의 레벨차는 각각 5dB, 8dB, 3dB이다. OCT 마이크로폰 어레이인 방식 1과 마이크로폰의 0°방향만을 변경한 방식 3은 동일한 레벨차를 보였으며 방식 2는 상대적으로 3dB 낮은 레벨차를 형성했다.



[그림 7] [표 2]의 각 방식들에 대한 센터와 L 채널 간의 사인파 1kHz의 레벨차 변화도

2.2. 청음 평가

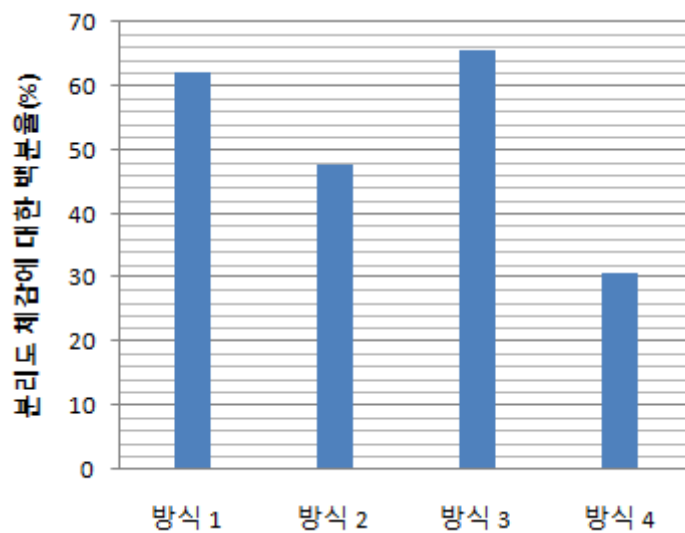
2.2.1 평가 방법

서라운드 모니터 시스템의 프론트 채널들에서 음원들 간의 분리도 묘사를 위한 이번 실험의 총 4가지 녹음 방식들에 대하여 객관화된 우수성 평가를 위해 청음 평가를 진행하였다. 리듬과 음 높이의 패턴 변화가 규칙적인 시계 초침소리와 불규칙적인 보컬 음원을 OCT 마이크로폰 어레이의 센터 마이크로폰을 기준으로 45°, 약 1.2m인 지점에서 스피커를 재생하여 녹음된 소스 음원을 사용하였다. 음원은 실험의 총 4가지 녹음 방식들의 결과물을 7초 단위로 편집하여 0.5초 간격으로 배열했으며 각 편집된 음원의 순서는 무작위로 나열하여 총 4개의 세트를 만들어 평가에 사용하였다.

음원들 간의 분리도 평가를 위해 피험자는 음악 전공자 10명을 대상으로 하였다. 또 평가 진행에 앞서 음원의 분리도에 대하여 충분히 설명한 뒤 평가를 진행하여 평가 결과의 객관성을 높이하고자 하였다. 각 세트의 무작위로 재생되는 4가지 녹음 방식의 결과물을 듣고 음원들 간의 분리도가 높게 지각되는 순서로 번호를 매기는 실험이 진행되었다. 또 음원들에 대하여 분리도에 대한 개별적인 체감 수준을 1부터 5로 평가하도록 하여 실험에 적용된 각 녹음 방식 간의 절대적인 평가를 이루고자 하였다.

2.2.2. 평가 결과

[그림 8]은 실험에 적용된 4가지 녹음 방식의 결과물들에 대하여 개별적인 감각 수준의 평가 결과를 백분율로 나타낸 것이다. 비교적 많은 피험자가 OCT 마이크로폰 어레이에 해당하는 방식 1과 OCT 방식에서 마이크로폰의 구성은 유지하고 0°방향은 정면을 향하도록 변경한 방식 3을 통해 높은 분리도를 경험하였다. 반면 카디오이드 마이크로폰만을 활용한 방식 2와 방식 4의 결과물 간 분리도의 감각 수준이 상대적으로는 15% 이하로 낮은 변화 추이를 보이지만 절대적 평가에 있어서는 최대 47.7%의 분리도로 음원들을 체험시킬 수 있음을 확인했다. 또 방식 1과 방식 3의 결과물이 가장 높은 분리도를 제공한다고 평가한 피험자 중 3명의 피험자는 분리도 체감에 대한 수준을 1부터 5중 2로 평가하였으며 이는 각 방식의 결과물 간의 근소한 차이만을 구분하여 평가한 것임을 인터뷰를 통해 확인했다.



[그림 8] 실험에 적용된 4가지 방식들에 대하여 분리도 체감에 대한 백분율

III. 결론

이 연구는 실감형 오디오인 서라운드 사운드가 멀티채널 오디오를 통해 재생될 때 음원들이 높은 분리도로 묘사되는 것을 목적으로 레코딩 테크닉이 연구되었다. 프론트 채널들이 높은 분리도로 녹음되는 OCT 마이크로폰 어레이를 중심으로 L-R 마이크로폰의 폴라 패턴 및 중심축의 방향을 변경하며 비교 방식의 실험을 진행하였고 방식 1, OCT 마이크로폰 어레이와 방식 3, OCT 방식의 L-R 마이크로폰의 중심축 방향을 정면으로 변경한 방식이 멀티 채널 오디오 표준모니터의 프론트 채널들에서 분리도 묘사가 우수함을 확인 하였다. 방식 1과 방식 3은 C에 카디오이드 마이크로폰을 사용하고 L-R에는 슈퍼 카디오이드 마이크로폰을 사용한다. 이는 L-R에 카디오이드 마이크로폰을 사용했을 때보다 L-C채널 혹은 C-R채널 간의 3dB 이상의 레벨차를 만들 수 있으며 각 채널 사이의 레벨차가 클수록 멀티 채널 오디오에서의 분리도에 대한 감각 수준 역시 우수함을 청음평가를 통해 확인하였다.

OCT 방식과 달리 방식 3은, L-R 마이크로폰의 중심축이 정면을 향했기 때문에 음원이 중점적으로 수음되는 로브(Lobe)의 위치가 변화됨으로 L-C 혹은 C-R 간의 레벨차가 OCT 방식보다 작을 것이라고 추론한 것과 다르게 실제 실험 결과는 사인과 1kHz에 대하여 OCT 방식과 방식 3이 8dB의 동일한 레벨차를 보였다. 반면 동일한 레벨차를 가짐에도 멀티 채널 오디오에서의 분리도에 대한 체감 수준은 방식 3이 OCT 방식보다도 약 3% 높게 조사 되었다. 이러한 결과는 슈퍼 카디오이드 마이크로폰의 리어 로브가 프론트 로브와 반대되는 위상 특성을 가짐으로 야기되는 위상 왜곡 현상 때문일 것으로 추정된다.

위 결과를 통해 멀티 채널 오디오에서 높은 분리도 묘사를 위한 레코딩 테크닉은 채널 간의 레벨차 뿐만 아니라 마이크로폰의 위상 특성, 청각과 관련된 심리적 요인에도 영향을 받는 것으로 생각되며 이를 고려한 레코딩 테크닉이 앞으로 지속적으로 연구 되어져야 할 것으로 사료된다.

참고문헌

- [1] Michael Williams, UNIFIED THEORY OF MICROPHONE SYSTEMS for STEREOPHONIC SOUND RECORDING, Presented at AUDIO the 82nd Convention 1987 March 10-13 London
- [2] Michael Williams and Guillaume Le Du, MICROPHONE ARRAY ANALYSIS for MULTICHANNEL SOUND RECORDING, Presented at AUDIO the 107th Convention 1999 September 24-27 New York
- [3] Spatial Audio for Broadcast. www.bbc.co.uk/rd/projects/periphony-for-broadcast
- [4] Helmut Witt과 Gunther Theile, The recording angle - based on localisation curves, Presented at the 112th Convetion 2002 May 10-13 Munich, Germany
- [5] Michael Williams, The stereophonic zoom, "Sounds of sootland", BP50, 94364 BRY SUR MARNE CEDEX, France
- [6] SCHOEPS_surround_brochure(2006-11)
- [7] Schoeps Manual MAB 1000