

〔論 文〕

## AVS方式に着目した音像定位距離制御と両耳信号の リサージュ図形に着目した想起される印象変化

沖田 和久\*, 梨子木 快晴\*<sup>2</sup>, 大里 一矢\*<sup>3</sup>, 河納 隼一\*<sup>4</sup>, 近藤 善隆\*<sup>5</sup>, 福島 学\*<sup>2</sup>

\*日本文理大学大学院工学研究科環境情報学専攻

\*<sup>2</sup>日本文理大学工学部情報メディア学科

\*<sup>3</sup>日本文理大学大学院工学研究科環境情報学専攻 (2021年度卒業)

\*<sup>4</sup>株式会社アーネット

\*<sup>5</sup>株式会社ジェイテック

### Control of Sound Image Distance by the AVS System and Recalled Impression Change Focused on the Lissajous Figure of the Binaural Signal

Kazuhisa OKITA\*, Kaisei NASHIKI\*<sup>2</sup>, Kazuya OSATO\*<sup>3</sup>, Syun'ichi KAWANO\*<sup>4</sup>  
Yoshitaka KONDO\*<sup>5</sup> and Manabu FUKUSHIMA\*<sup>2</sup>

\*Department of Environmental Engineering and Applied Information Science,  
Graduate School of Engineering, Nippon Bunri University

\*<sup>2</sup>Department of Media Technologies, School of Engineering, Nippon Bunri University

\*<sup>3</sup>Department of Environmental Engineering and Applied Information Science,  
Graduate School of Engineering, Nippon Bunri University (Graduate, AY2021)

\*<sup>4</sup>Artnet Co., Ltd.

\*<sup>5</sup>J-TEC Co., Ltd.

### Abstract

This paper describes the possibility of controlling the sense of distance and the possibility of creating changes in spatial impressions by imposing certain rules on the relationship between instantaneous amplitudes. For distance perception control, we extended the AVS method, which is a distance localization control method, and investigated whether localization distance control is possible. As a result, the possibility of control by the amount of control signals was confirmed by subjective evaluation experiments. In terms of spatial impression, subjective evaluation experiments confirmed that introducing a certain rule to the instantaneous amplitude relationship can change the spatial impression of a sound stimulus that is perceived as random noise by a single ear with equal time waveform envelopment information and frequency amplitude spectrum. These findings provide the basis for new media processing technology.

キーワード：AVS方式, 音像定位, 両耳信号, リサージュ図形, 心象印象

**Keywords** : AVS system, sound image, binaural signal, Lissajous figure, sound impression

## 1. はじめに

人が情報を受け取る五感において、視覚情報と聴覚情報は重要である<sup>1)</sup>。空間を知覚する際もこの情報に基づいて行われる。劇場演出<sup>2)</sup>等のエンタテインメント産業においても、創出される空間に対する取り組みが行われている<sup>3)</sup>。劇場や舞台等でのイベントはDVDコンテンツ化されるだけでなくライブビューイングとして展開されている。視覚情報は人が受け取る情報の多くを占めると言われているが、死角や視界によって情報が得られなかったり、減少したりすることがある。このため、死角が少なく伝搬媒体があれば受け取ることで聴覚情報も重要である。

音刺激により想起される空間の印象は視覚刺激と統合することにより印象を補強するだけでなく変化を与えることもできる。このためこの分野において空間記述という表現力の拡張は今後ますますその重要性を増すことが予想される。災害対策においても、避難警告等のサイン音<sup>4)</sup>を避難誘導方向から提示する等の防災や、見守り<sup>5)</sup>にも展開されている。これは遠隔教育<sup>6)</sup>での利活用や、ヘルスケアにおける睡眠の質<sup>7)8)</sup>においても重要視されている。

しかし、コンピュータグラフィック（CG：Computer Graphic）が実写と隣り合わせた時に空間的な差を感じるように、音刺激により任意の空間を創出するには未だ技術的課題が残されている。音刺激で想起される印象は、音源に関する事柄とそれが伝送される室に関する事柄に依存する。このため、音源だけでは空間的な印象を感じないが、音源が無ければ刺激が存在できない。

そこで、本研究は音刺激により形成される空間印象について知覚メカニズムの解明とそれに基づく制御に取り組んでいる。本稿では、その一部として、音像定位に関して距離感のある音像定位が可能な AVS 方式に着目した距離制御と、両耳信号においてリサージュ図形のみが異なり他の時間包絡情報および振幅スペクトルが等しい信号により想起される印象について調べた結果を報告する。

## 2. 音が想起させる空間

音刺激は、音のみなもと（音源）と、その音が耳やマイクにどう届くのか（音の伝わり）がある。耳に届いた音刺激は、聴知覚を工学的に表現すると復元（現象の主要な要素で構築する）から信号生成（認知情報）そして

情報提示（行動に必要な情報となる）することとなる。耳の代わりにマイク等の信号計測となると、聴知覚の神経発火パルスを送信と見立てると、それ以降の過程は復元（デコード等の信号復元）から信号生成（メディア処理）そして情報提示（システムで利用可能な情報とするまたは伝送先で再構成すなわち現象復元も含まれる）となる。この関係を整理すると図1に示すように、現象発生から計測までが、音源／音像のブロックと音の伝わりとなり、聴知覚または計測された信号が伝わる部分が、信号の伝送となり、復元から情報提示までが、音から得るもの、と捉えることができる。これは、ネットワークという枠組みにシステムと人を統合しうる枠組みであり、古くはバーチャルリアリティの走りである蓄音機（信号伝送が記録メディアを介することで時間が経過しても現音場（元の音源と伝送）が再現できる）や電話（信号伝送が電気信号とすることで地理的に離れた場所でも相手が耳元にいるように再現できる）から、音刺激の理解または音声認識と音声合成を使った ChatGPT 等の生成 AI サービスさらにはヘルスケアやパブリックビューイングや XR といった DX への展開までも同一のフレームワークで捉えることができる。

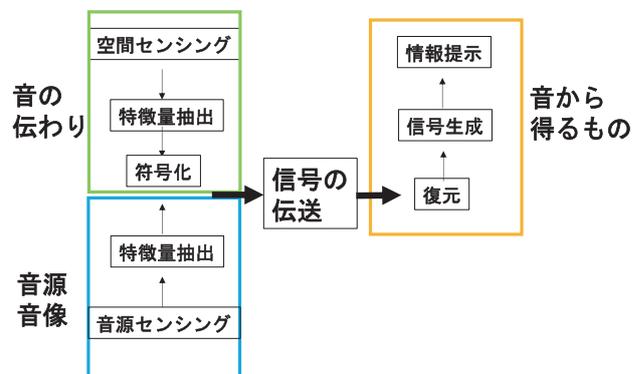


図1. 音で想起する空間の要素と信号の流れ

本稿ではこのフレームワークの内、1) 音源／音像の「音から得るもの」において必要不可欠な音像定位、2) 音の伝わりの「音から得るもの」において必要不可欠な空間印象の制御、に着目する。

1) 音源／音像の情報提示において、定位が重要な要素である。定位に関して様々な研究が取り込まれ多くの成果が出ているが、距離感に関して工学的に利活用するには未だ十分とは言えない状況である。音像定位技術の1つである AVS 方式は、狭帯域毎に1組の時間遅延とレベル差のみで音再生空間の近傍反射等が存在しても定位制御できる技術であり「距離感を感じる音像定位」がその特徴である。しかし、なぜ「距離感」を生じるのか

そのメカニズムは解明されておらず「距離感制御」には至っていない。そこでここでは AVS 方式に着目し、距離感制御の可能性を検討する。

2) 音の伝わりは「空間印象の想起」すなわち聴知覚による空間把握につながる要素である。音に含まれる感覚に関して、周波数振幅スペクトルのキューが関係するとの研究成果や、両耳時間波形包絡線の相関が関係するとの研究成果がある。しかし、振幅スペクトルや時間波形包絡線が等しくても位相が異なれば包絡線が変化し、特に両耳受聴では空間的印象に変化を生じる。そこでここでは、リサージュ図形のみが異なる単耳では区別つかない音源を刺激音とし、変化知覚、印象変化、を調べる。

### 3. 狭帯域毎の時間・レベル差による音像定位制御方式

ここでは、音像定位に関して距離感のある音像定位が可能な AVS 方式に着目し、側方遠方定位する音像の距離感制御を試みる。

AVS 方式は、図 2 に示すように 1/4 oct. 分割した信号を、スピーカ (SP<sub>R</sub>) からそのまま、制御スピーカ (SP<sub>L</sub>) から ICTD と ICLD のみをパラメータとし符号反転して放射する<sup>9)10)11)</sup>。ここでダイレクトトークを  $h_{RR}$ ,  $h_{LL}$  (赤実線)、クロストークを  $h_{RL}$ ,  $h_{LR}$  (赤破線) とする。このとき使用するパラメータによって図 1 の sound image に音像を定位することのできる手法である。任意位置のパラメータを探索すれば sound image の位置を変えられる。

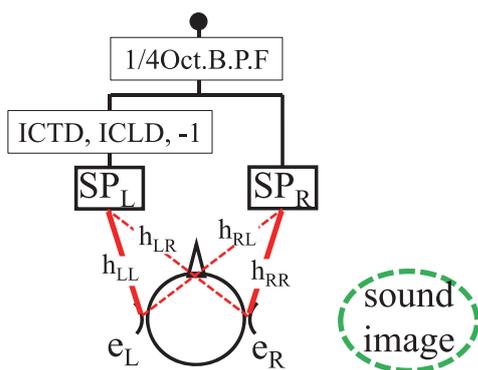


図 2. AVS システムの構成

実際に 2 個のスピーカと両耳一のマイクを用いて IR ( $h_{LL}$ ,  $h_{LR}$ ,  $h_{RL}$ ,  $h_{RR}$ ) を計測する。ここではスピーカと両耳位置のマイクを図 3 に示す配置にした。計測した結果を図 4 に示す。図 4 の 1 段目にダイレクトトーク、2 段目にクロストーク、のインパルス応答 (IR)

を示し、3 段目にインパルス応答の和を示す。図は横軸に時間、縦軸に振幅を示している。図中矢印で示した波形到来時刻から実験時のスピーカと耳までの距離が概ね一致していることから計測値が妥当であると判断する。

この環境において図 5 に示す音像位置 (最近傍、最遠方、中間) となる AVS パラメータ (ICTD と ICLD) を探索した結果、表 1 に示す値となった。

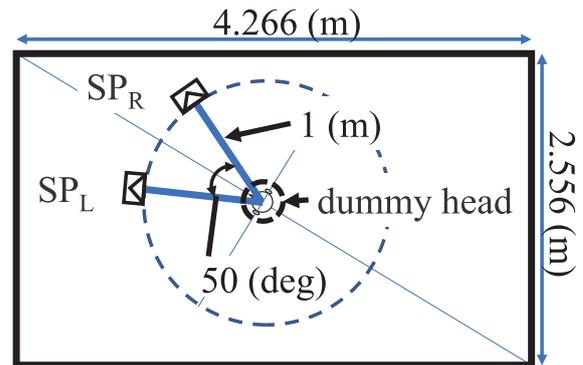


図 3. インパルス応答計測時の室寸法および音源・観測点の位置関係

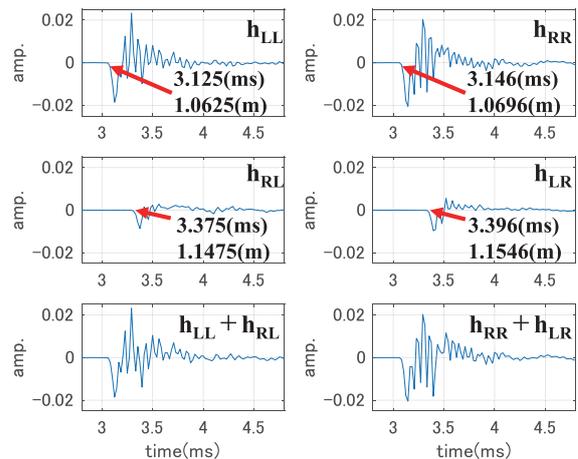


図 4. ダイレクトトークとクロストークの IR

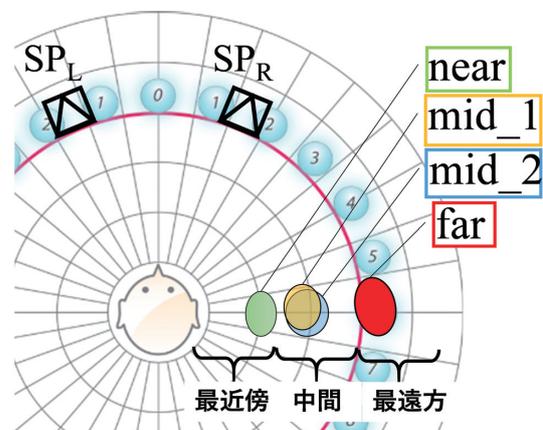


図 5. 音像探索位置と定位音像のイメージ

表 1. 図 4 に示す音像位置となるパラメータ (ICTD は sample, ICLD は dB)

distance	ICTD	ICLD
far	12	-3.9
mid_1	9	-3.9
mid_2	16	-2.9
near	20	-2.0

これらのパラメータを図 4 に示した制御スピーカに関連するインパルス応答に反映した結果が図 6 から図 9 である。図には図 4 の到来時刻を赤縦線で示し、制御前の振幅を赤横線で示している。図の並びと軸は図 4 と同じである。図は、ICTD と ICLD により、遅延 (赤縦線から ICTD 分の時間遅れ) と減衰 (赤横線から ICLD 分の振幅減少) を生じていることを示している。

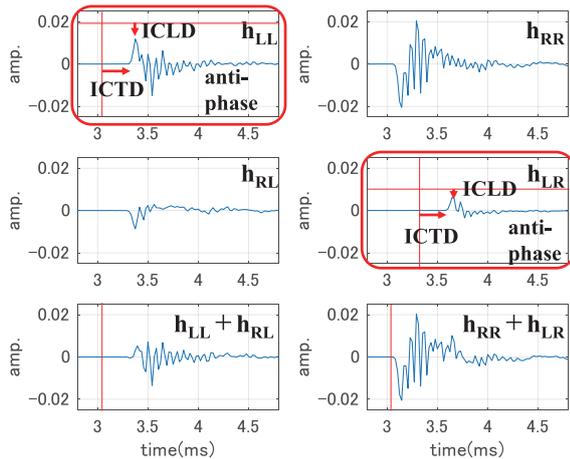


図 6. far の ICTD と ICLD を反映させた IR

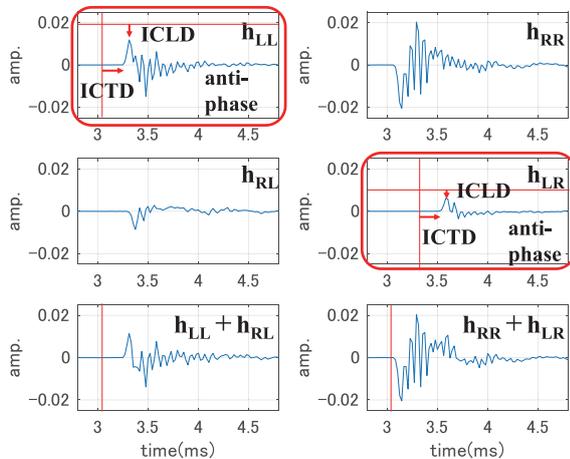


図 7. mid\_1 の ICTD と ICLD を反映させた IR

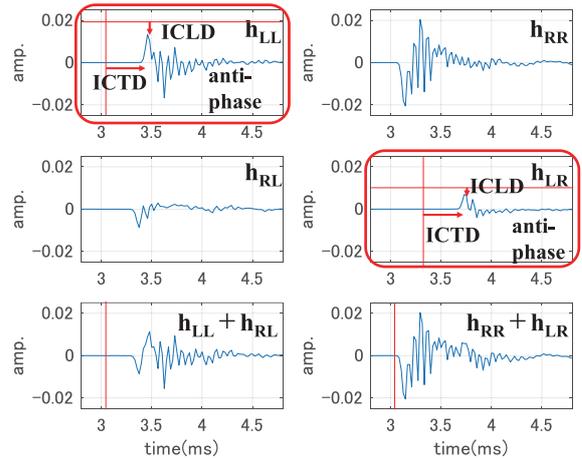


図 8. mid\_2 の ICTD と ICLD を反映させた IR

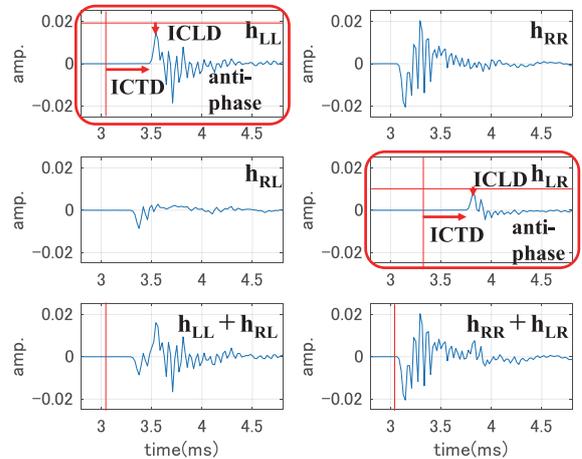


図 9. near の ICTD と ICLD を反映させた IR

比較の為に図 5 から図 8 の 3 段目を図 9 に再掲する。図の左側、左耳に到来する波形は 4 つ全て冒頭の到来時刻が遅れている。また、図の 1 段目と 2 段目すなわち far と mid\_1 はクロストークがキャンセルされているが 3 段目と 4 段目すなわち mid\_2 と near についてはクロストークがキャンセルされていない。

図 9 に示したインパルス応答は全帯域の応答である。これに対して図 5 に示した音像探索は  $f_c = 500\text{Hz}$  の  $1/4$  oct. バンドパルスで実施している。そこで、探索に使用したバンドパルスを図 10 に示したインパルス応答に畳み込み、距離のある音像定位現象を生じた際の両耳に提示された信号を導出した。結果を図 11 に示す。図は、図 10 よりも音像近傍時にクロストークがキャンセルされていること、全体を通して音像定位側の耳で波形に違いが無いことを示している。AVS 方式は  $1/4$  oct. バンド毎に ICTD と ICLD を探索する理由がこのデータから確かめられた。

以上のことから、キャンセル量の変化によって距離感

が変化すると考えられる。

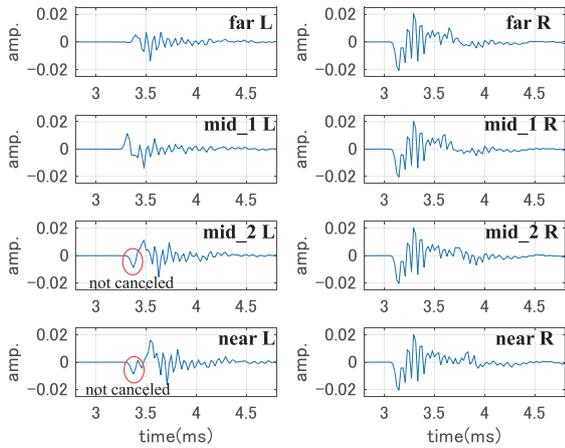


図10. 各距離の ICTD と ICLD を反映した IR

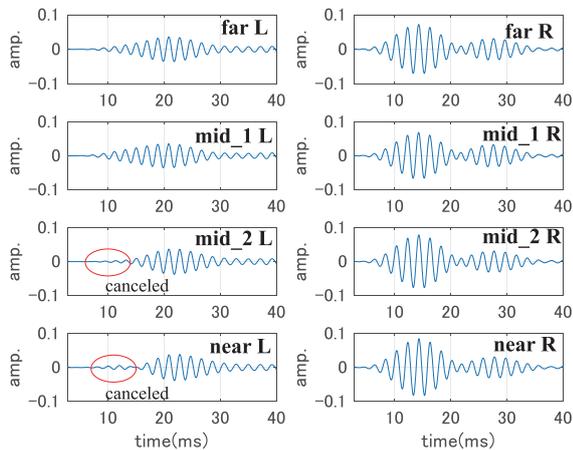


図11. 観測信号 (fc=500Hz, 1/4 oct. BP)

#### 4. リサージュ図形のみが異なる信号で想起する印象

ここでは、振幅・位相スペクトルおよび時間波形の包絡が等しい両耳信号において、リサージュ図形のみを変化させた際に想起される心象印象について着目し、聴取実験により心象印象の変化を調査する。

ここでは刺激音を4種類用意した。

信号(1): 位相角 $0^\circ$  (同相) の白色雑音

信号(2): 無相関の白色雑音 (内積0 / 直交)

信号(3): 位相角 $90^\circ$  (sin と cos) の時間置換信号

信号(4): 位相角 $180^\circ$  (逆相) の白色雑音

信号(1) ~ (4) をそれぞれ11秒の長さで作成し連結させた合計44秒の刺激音を作成した。

図12に信号(1) ~ (4) の時間波形の先頭を拡大表示し、図13に刺激音の時間波形全体を示す。図はそれぞれ

横軸に時間、縦軸に振幅を示し、図12は左上に信号(1)、右上に信号(2)、左下に信号(3)、右下に信号(4)を表示し、青印( $\times$ )がLch(左耳)、赤印( $+$ )がRch(右耳)を示している。図13は上段にLch、下段にRchを示し、どの信号がいつ再生されるかを図上部に番号で示している。

図12から、図左上の信号(1)は同相のため両chで値が同じになっていることが描画された点の重なっていることからわかる。対して図右下の信号(4)では逆相となりLchに対してRchの符号が反転している。図右上の信号(2)と図左下の信号(3)は図から一見すると違いは判断ができないが、信号の構成に違いがある。信号(2)は無相関な白色雑音2つによって構成されているのに対し、信号(3)はsin信号とcos信号(ともに440Hz)のため片側chの値が0のとき必ず別chが1または-1を示している。

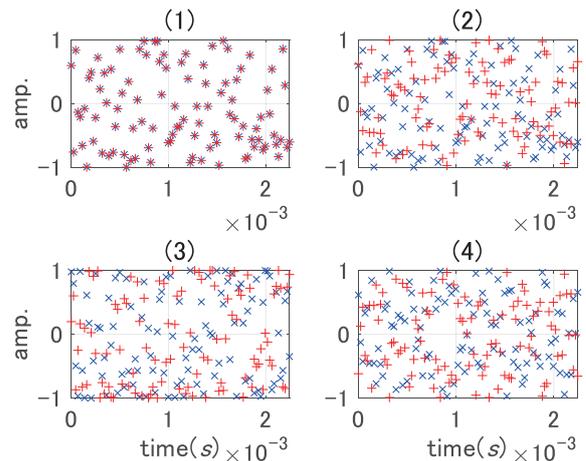


図12. 両耳時間波形先頭部分の拡大図

(青: L ch, 赤: R ch)

左上: 信号(1), 右上: 信号(2),  
左下: 信号(3), 右下: 信号(4)

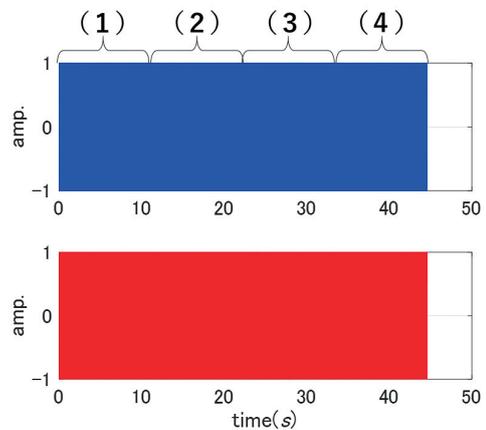


図13. 刺激音の時間波形  
上段: L ch, 下段: R ch

図13からは、刺激音の時間包絡が4種類全て同じであることがわかる。また4つの信号の切り替わりで無音区間などがない事を示している。

次に、周波数特性において刺激音に差がないことを確認するため、刺激音の周波数スペクトルを時間軸方向に並べたソノグラムを図14に示す。図は上段にLch, 下段にRch, 横軸に時間, 縦軸に周波数を対数で示し, 値の強度 (dB) を色で表現している。色が明るい (黄色) ほど値が大きく, 暗い (青色) ほど値は小さくなる。また, 4つの信号の時間範囲を示すために黒の補助線を引き, 図13同様に図上部に番号で信号の種類を示している。図はサンプリング周波数 (44.1kHz) と分析区間長の関係により, 値が計算されていない箇所が白く表示されている。

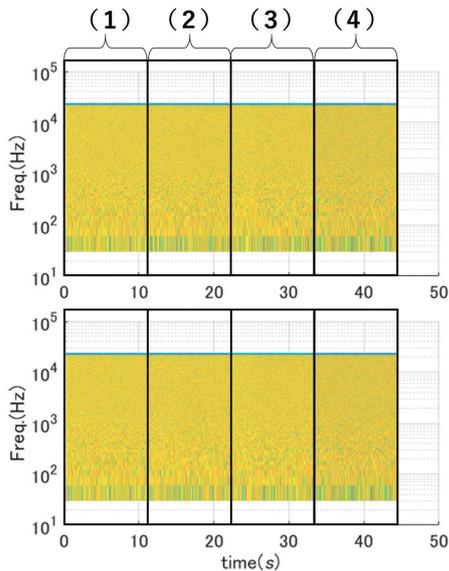


図14. 刺激音のソノグラム  
上段: L ch, 下段: R ch

図14から, 4つの信号において周波数特性に違いがないこと, 全帯域に渡って一様に値が分布していることがわかる。またこれらのことから, 刺激音が単耳聴 (片側 ch のみの受聴) においてはほとんど違いを聞き分けることができない信号の組み合わせであることが確認できた。

以上のことから, スペクトルと時間波形の違いが無いことを確認した。次に, チャンネル間の関係について確認する。

チャンネル間の関係を見るために, 刺激音のチャンネル間相関係数を確認する。刺激音のチャンネル間相関係数の時間変動を図15に示す。図は横軸に時間, 縦軸にチャンネル間相関係数を示している。また, 図13同様, 4つの信号

の時間範囲を示すために黒の補助線を引き図上部に番号で信号の種類を示している。

図14から, 同相である信号 (1) は相関が1を示し, 逆相である信号 (4) は相関が-1を示している。無相関である信号 (2) と信号 (3) については, 2信号ともに相関が0付近に収束している。これらから, 設計した相関の値と刺激音として作成した音信号の相関の値が許容範囲内であることが確認できた。また, 相関係数においては信号 (2) と信号 (3) で明確に違いがないことが確認できた。

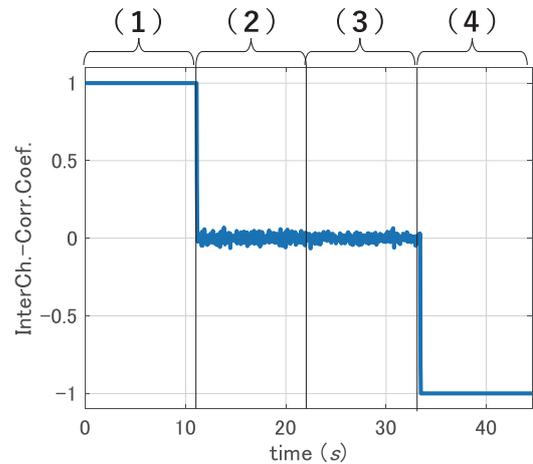


図15. 刺激音のチャンネル間相関係数の時間変動

次にチャンネル間の関係をリサージュ図形から確認する。信号 (1) ~ (4) のリサージュ図形を図16に示す。図は青色の点でリサージュ図形を描いており, 左上に信号 (1), 右上に信号 (2), 左下に信号 (3), 右下に信号 (4), のリサージュ図形を示している。

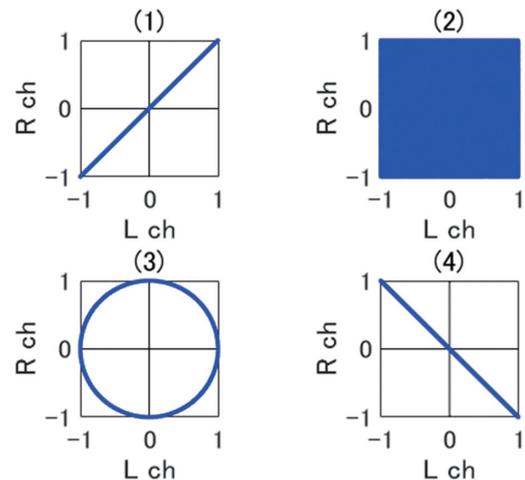


図16. 刺激音のリサージュ図形  
左上: 信号 (1), 右上: 信号 (2),  
左下: 信号 (3), 右下: 信号 (4)

図16から、同相である信号(1)のリサージュ図形は図左上のように傾き1の直線が描かれ、逆相である信号(4)のリサージュ図形は図右下のように傾き-1の直線が描かれる。図右上の信号(2)のリサージュ図形は、一様分布ランダム信号のため点が各軸1から-1の範囲を埋め尽くす図が描かれている。図左下の信号(3)のリサージュ図形には時間置換しても同一時刻における $\sin(t)$ と $\cos(t)$ の関係から円が描かれている。信号(3)が瞬時値において $\sin$ と $\cos$ の関係を保つ、すなわち瞬時振幅の関係に一定の規則として位相角 $90^\circ$ の関係を与えることができていることがわかる。図15で示した相関係数においては違いが見られなかった信号(2)と信号(3)は、リサージュ図形においては違いが生じていることが確認できる。

以上のことから、4つの信号および刺激音の構成について振幅・位相スペクトルおよび時間波形の包絡が等しく、リサージュ図形のみが変化した両耳信号であることを確認した。

これらの信号に基づいて、リサージュ図形のみ異なる両耳信号において印象が変化するか、作成した刺激音を用いて聴取実験を行った。

実験は、実験協力者に刺激音の構成については説明せずヘッドホンにて両耳で受聴してもらい、いくつのパターンがあったかを質問する。聞き分けることができたパターンについて、刺激音の何秒からなっていたか、空間的印象はどうであったか、を図17に示す記入シートで回答する形式とした。実験協力者は20代男性17名を対象とした。実験の結果を表2に示す。

#### 実験1

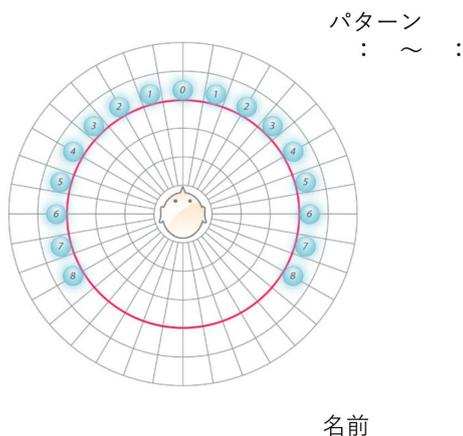


図17. 聴取実験で使用した記入シート

表2. 聴取実験の結果

subject No	4種類を 区別できる	広がり に 差を感じた	距離 に 差を感じた
1	○	○	
2	○	○	○
3	○	○	○
4	○	○	
5	○	○	
6	○	○	
7	○		○
8	○	○	○
9	○		○
10	○		○
11	○	○	○
12	○		○
13	○		○
14		○	
15			
16		○	
17			

実験の結果、刺激音に含まれる4つの信号を聞き分けることができたと回答した人数は17人中13人であった。残りの4人についても、2種類以上の信号があると回答した。空間的な広がり差を感じたのは17人中10人、距離に差を感じたのは17人中9人であり、広がり差と距離のどちらも差を感じたのは17人中4人であった。

実験協力者からは、「音質や音色の違いは区別できなかったが、空間的な差は感じることができた」というインプレッションが挙げられた。

以上のことから、リサージュ図形のみ異なる4つの両耳信号において空間的印象の変化が生じることの可能性が確認された。

#### 5. おわりに

音刺激により想起される空間的印象について知覚メカニズムの解明とそれに基づく制御に取り組んでいる。本稿では、その一部として、定位させた音像の距離感に関してAVS方式を手掛かりに制御の可能性を調べた。また、1チャンネルでは白色雑音としか知覚できない雑音性音刺激において、想起される印象について調べた。その結果、AVSにおいて制御耳側信号の振幅量すなわちキャンセル注入量により距離感が制御できる可能性が確認できた。また、リサージュ図形が特定形状となるすなわち瞬時振幅の関係に一定の規則を与えることで空間的

印象に変化を生じることの可能性が確認された。これらは、従来距離感が直接音と部屋の響きである間接音の比率により生じるとされていたのに対し、響きによらず距離感と空間印象を変えることの可能性を示すものであり、新たなメディア処理技術の基礎となる現象の発見である。

#### 参考文献

- (1) 姜大基, “視聴覚情報による空間知覚に関する研究”, 東北大学, 学位論文, <http://hdl.handle.net/10097/51152> (2023年6月14日アクセス)
- (2) 公益社団法人劇場演出空間技術協会, <https://www.jatet.or.jp/> (2023年6月14日アクセス)
- (3) 福島学, 沖田和久, 林友哉, 伊藤毅登, 河野拓海, “実測データを用いた音空間再現精度の評価方法に関する基礎的検討 - 音響現象 (Acoustic Phenomena) と音響事象 (Acoustic Event) -”, 日本文理大学紀要, 第50巻, 第1号, pp. 9-18, 2022
- (4) 橘秀樹, 飯田一博, 大川茂樹, 木幡稔, 佐藤史明, 須田宇宙, 世木秀明, 柳川博文, 矢野博夫, 山崎治, 横山栄, “公共空間における安全確保のための音響情報伝達に関する研究”, 科研費報告書, <https://kaken.nii.ac.jp/ja/grant/KAKENHI-PROJECT-22241040/> (2023年6月14日アクセス)
- (5) 福島学, 大里一矢, 山下涼介, 樋口幸, 市田秀樹, 石井秀樹, 森竹隆広, “音響技術による見守りシステム”, 日本文理大学紀要, 第49巻, 第1号, pp. 87-96, 2021年
- (6) 柳川博文, 三井田惇郎, 鈴木英男, 山崎芳男, 東山三樹夫, 福島学, 中村直人, “一体感のある合同学習を可能にする遠隔双方向教育システムのための音響系の開発”, 科研費報告書, <https://kaken.nii.ac.jp/ja/grant/KAKENHI-PROJECT-14380084/> (2023年6月14日アクセス)
- (7) Manabu Fukushima, “Characteristic and function of the newly developed articulatory panels (Aural Sonic)”, *Glycative Stress Research* 2019 : 6 (2), 103-112, (doi : 10.24659/gsr6.2\_103) (<http://www.toukastress.jp/webj/article/2019/GS19-05.pdf>)
- (8) Manabu Fukushima, Shiori Uenaka, Masayuki Yagi, Wakako Takabe, Yoshikazu Yonei, “Effect of the newly developed articulatory panels (Aural Sonic) : A pilot clinical trial”, *Glycative Stress Research* 7 (2) : 123-131, 2020 (doi : 10.24659/gsr.7.2\_123) (<https://www.toukastress.jp/webj/article/2020/GS19-11.pdf>)
- (9) 沖田和久, 林友哉, 福島学, 松本光雄, 風間道子, 柳川博文, “AVSにおける横方向定位時の両耳波形について”, 音講論集(秋), 1-R-17, 2022.
- (10) 沖田和久, 林友哉, 福島学, 松本光雄, 柳川博文, “AVS方式における遠方および耳近傍音像定位現象のITDとILDを指標とする調査と評価”, 第16回学生のための研究発表会, 13, 2022
- (11) 沖田和久, 林友哉, 福島学, 松本光雄, 柳川博文, “AVS方式による距離のある音像定位現象に着目した音像定位距離制御の試み”, 日本音響学会, 日本音響学会2023年春季研究発表会講演論文集, 2-1P-5, 2023

(2023年6月14日受理)