Vol.2017-MUS-115 No.29
2017/6/17

音源特性と両耳差に着目した音場情報のパラメトリック表現

長瀬翔斗[†1]　福島学[†1]　近藤善隆[†2]　松本光雄[†3]　風間道子[†4]　柳川博文[†5]

概要：音場情報をコンパクトかつ高い再構築精度で記録することは記録容量的にも低遅延伝送的にも重要な技術課題である．本稿では，１）音場情報を再構築する電気音響変換器であるラウドスピーカに関して実測インパルス応答がバッフル板上ピストン運動の呼吸球近似式モデルで表現できるかについて，２）ランダム信号系列の両耳間位相差で音場空間印象が表現できるかについて，調査した結果を報告する．

キーワード：音源特性，両耳差，音場情報，パラメトリック表現，呼吸球，位相

Parametric representation for acoustic field information focused on sound source characteristics and binaural information

SYOTO NAGASE[†1]　MANABU FUKUSHIMA[†1]　YOSHITAKA KONDO[†2]
MITSUO MATSUMOTO[†3]　MICHIKO KAZAMA[†4]　HIROFUMI YANAGAWA[†5]

Abstract: Parametric representation is an important aim not only for signal understanding, but also for immersive and interactive audio technology through network.　In this paper, parametric representation on, 1)modeling of a loud speaker that is a source of acoustic phenomena, 2)sound perception modeling focused on binaural information, is discussed.　As a result, 1)possibility of a loud speaker modeling with spherical source and vibration, 2)possibility of phase information between 2 channel, are confirmed.

Keywords: Sound Source Characteristics, Binaural Information, Acoustic Field Information, Parametric Representation, Spherical Source, Phase

1. はじめに

通信帯域幅の増加に伴うリッチコンテンツ化が可能となり，またデバイス環境の普及に伴い仮想空間を共有したシステム[1]のニーズが高くなっている．仮想空間の一要素である音場の共有を実現するには，音場の計測と再構築，移動検知と変化のリアルタイムでの反映が必要である．変化に対応した物理的に厳密な音場変化には，逐次対応したインパルス応答を用いた畳み込み演算が必要となり，レスポンスの遅延ならびに同期の問題が発生する．音場共有として，仮想会議室をバイノーラル再生により共有するシステム[2]のように，音場内での徘徊や耳打ちを実現するシステムがある．このシステムは共有する部屋を固定し，移動で生じる変化量のうち幾何情報だけを伝送し音場を生成することで少ない情報伝送量での音場共有を実現している．しかし，音場の汎用モデル化とパラメータ表現までは至っていない．このような情報記述として，情報構造に着目した情報圧縮を行う MPEG-4[3]がある．MPEG-4 は，音源信号を構造的に表現し符号化する Structured Audio[4]を MPEG-4 Structured Audio[5]として採用している．そのツールの１つである SAOL(the Structured Audio Orchestra Language)は音源を記述する言語であり，その制御に標準 MIDI(Musical Instrument Digital Interface)[6]を採用している．また MPEG-4 は，空間に音源を配置するため，インターネットの仮想空間記述言語 VRML(Virtual Reality Modeling Language)[7]を基に，音に関する記述を拡張した AudioBIFS (BIFS : Binary Format for Scene Description)が定義されている．このように MPEG-4 では音源の構造と配置に着目し符号化することで，効率的な情報伝送を実現しようとしている．しかし空間知覚に基づく記述に至っているとは言い難い．

そこで，本研究では音場共有に必要な音場情報のパラメトリック表現の確立を目的とする．このためには，物理的現象である音響現象（Acoustic Phenomena）と，知覚事象である音響事象（Acoustic Events）の側面からモデル化およびパラメトリック記述を検討しなければならない．

これまでに著者らは双方向音場共有ネットワーク（ISFN: Interactive Sound Field Network）と音響事象記述言語（AEML: Acoustic Events Modeling Language）を提案している[8-17]. そこでは音場を表現する言葉と，それに対応する物理特性を明らかにし，音響刺激で想起される空間印象「広さ」感に関して実測インパルス応答を所望の音場印象となるように信号処理することが可能であることを明らかにし

†1 日本文理大学
Nippon Bunri Univ.
†2 （株）ジェイテック
J-Tech. Co., Ltd.
†3
†4 NH ラボ（株）
NH Lab. Inc.
†5 千葉工業大学
Chiba Inst. of Tech.

ⓒ 2017 Information Processing Society of Japan

ている．しかし，音場を再構築する装置にまでは踏み込んでおらずまた空間知覚に関して未解明な点を残している．

本稿では，１）音響現象を発生する電気音響変換器（ラウドスピーカ），２）両耳差情報の一つである両耳間位相差，に着目し，モデル化とパラメトリック表現の可能性を調べた結果を報告する．

2. 電気音響変換器のパラメトリック表現

音場情報を空間に再構築するには，電気音響変換器が必要不可欠である．電気音響変換器には駆動方式やキャビネット等により伝達特性が異なる．装置が共通であれば概ね同等の物理現象を発生させることが可能である．しかし，不特定多数の利用者で仮想空間を共有することを考えると，装置の共通化は現実的ではない．

電気音響変換器を 1 つのモデルを基に特性に応じたパラメータで記述できれば，装置を共通化しなくても，そのパラメータに応じた変換により仮想空間の共有が可能であると考えた．このアプローチは電気音響変換器の逆特性で補正を行うことと思われがちだが，物理的な逆特性の厳密解を求めることも，一般家屋内で補正して空間を再構築するのも難しい．本研究のアプローチは，提示された刺激により想起された空間印象が同等であることを目指すため，逆問題に踏み込まず，主となるモデルを探索することとする．

そのための第一ステップとして，ピストン運動する平板と無限バッフル板でモデル化された電気音響変換器が，呼吸球で近似されることに着目し，実測伝達特性をモデル化できるかについて検討する．

半径 a (m)の呼吸球の体積速度 v は体積を x とすると

$$v = \frac{d}{d\xi}x(\xi) = (1-\xi)e^{-\xi} \qquad (1)$$

但し，$\xi = \frac{ct}{a}$，a は呼吸球の半径，c は音速，t は時間，

となり，振動変位 y は式（1）の積分値となる．呼吸球の半径 a= 0.1 (m) で計算した体積速度 v と振動変位 y の例を Fig.1 に示す．図は横軸に時間を秒で示し，破線が v，実践が y を示す．

一方，電気音響変換器の伝達特性は，時間領域でインパルス応答と言い，音圧波形として計測できる．音圧波形に対応するのは振動加速度であり，その積分値が振動速度であり，さらに積分値が振動変位となる[18]．Fig.2 に実測インパルス応答（measured Impulse Response）から振動変位を求めた例を示す．図は横軸に時間，縦軸に振幅を示しており，右上に計測対象の電気音響変換器の図を示す．使用した電気音響変換器はパイオニア製 S-ST5 であり，直径 0.12 (m)のリニアパワー方式のユニットと，直径 0.025 (m)のコーン型ユニットを搭載している．Fig.2 に示した，音圧波形に対応した振動加速度を 2 回積分して求めた振動変位推定値（図中一点鎖線）に半径 a=0.07(m),0.0125(m)とした呼吸球理論式から求めた振動変位（図中破線）を Fig.3 に示す．

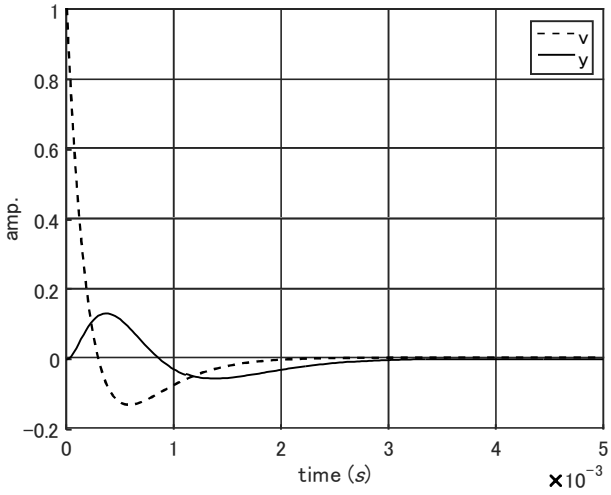


Fig.1 呼吸球半径が 0.1 (m)の体積速度 v と振動変位 y

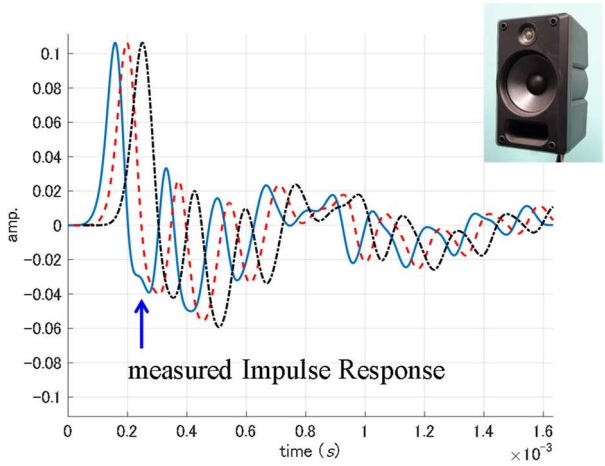


Fig.2 電気音響変換器（パイオニア製 S-ST5）実測インパルス応答（実線）積分値（破線）2 回積分値（一点鎖線）

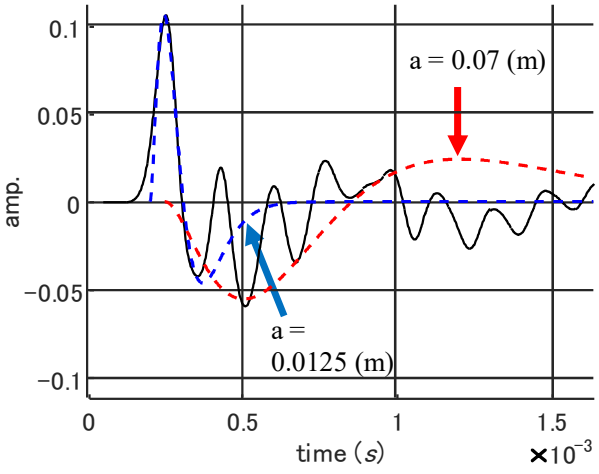


Fig.3 実測インパルス応答の 2 回積分値と
呼吸球の振動変位（a=0.07(m), a=0.0125(m)）の対応関係

Fig.3 は 2 種類の呼吸球モデルで近似されていることを

ⓒ 2017 Information Processing Society of Japan

示している．そこで，使用した電気音響変換器のツイータユニットを外し，バスレフを閉じることでリニアパワー方式のユニットのみが駆動する状態でインパルス応答を計測し，Fig.3 と同様に振動変位を求める．その結果を Fig.4 に示す．Fig.4 の横軸は Fig.2 と同様に横軸に時間，縦軸に振幅を示す. Fig.4 の一点鎖線で示したインパルス応答の 2 回積分により求めた振動変位に Fig.3 と同様に呼吸球の振動変位を Fig.5 に示す．図の実線はインパルス応答の 2 回積分値で求めた振動変位，破線は呼吸球理論式から求めた振動変位（半径 a = 0.35 (m)）を示している．インパルス応答の積分を行う際に積分定数の処理を含むため，ここではこの処理に相当する減衰補正（decay correction）を加えた値を一点鎖線で示す．

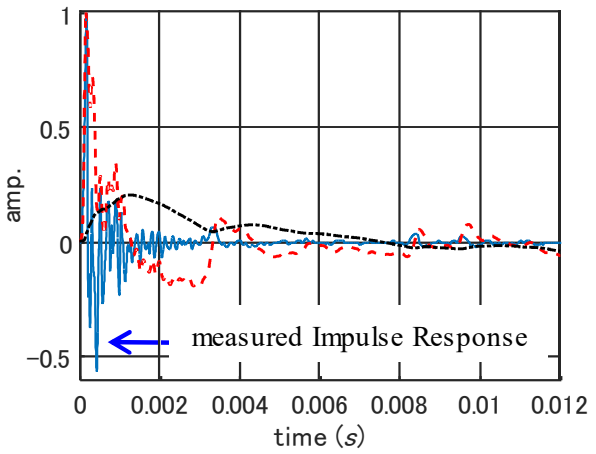


Fig.4　リニアパワー方式のユニットのみとした電気音響変換器（パイオニア製 S-ST5）の実測インパルス応答（実線）積分値（破線）2 回積分値（一点鎖線）

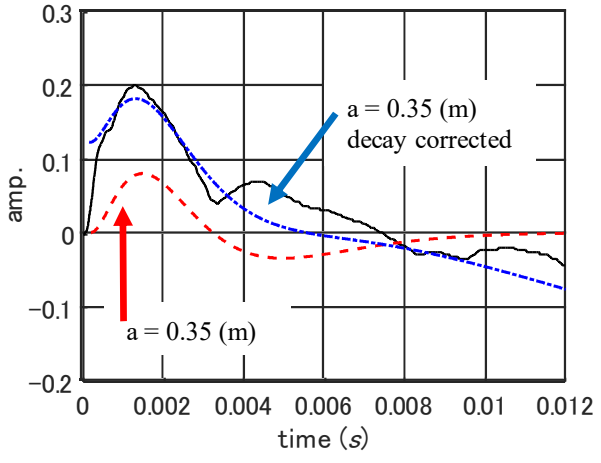


Fig.5　リニアパワー方式ユニットのみでのインパルス応答 2 回積分値で求めた振動変位と呼吸球の振動変位（a = 0.35 (m)）の対応関係

Fig.5 は実寸の直径である 0.06 (m) の 5 倍以上となっているが，使用されているリニアパワー方式は，一般にコーン紙を固定するためのエッジがないエッジレスとなっており，振動版がピストン運動する．このため，実寸よりも直径が大きなユニット相当の低音が再生できるとされているものである．このため最もフィッティングする半径が実寸よりも大きな数字となったものと判断する. Fig.3 に示した半径よりも大きくなっている理由として, Fig.3 に示す通り，振動変位の主要部分がツイータ相当の呼吸球モデルによる近似となっていること，バスレフによる影響によるものと判断する．

以上のことから，検討が必要であるものの，電気音響変換器のインパルス応答を呼吸球でモデル化し，個々の特性をパラメトリックに表現することの可能性が確認できたと判断する．音響事象の元となる音響現象を産む電気音響変換器特性のパラメトリック記述が可能になることで，音場情報の再構築に必要な音響現象の再生環境への依存度が下がることが期待できる．

3. 両耳間位相差と想起される空間印象のパラメトリック表現

本章では音場情報の聴知覚に関して，両耳差の中でも重要な両耳間位相差に関し，両耳に提示される 2ch 信号間の位相差による音響空間知覚との関係を調べる．両耳の到来時間差と空間印象に関しては両耳間相関関数（IACC）及び過渡的両耳間相関関数（TRICC）[19-22]により明らかにされている．しかし，両耳間位相差のみで生じる広さに関する印象のパラメトリック表現にまでは至っていない．

そこで，ここでは 2ch の位相差により想起される印象をモデル化およびパラメトリック表現可能であるかについて調査する．

位相差が瞬時値で一定の条件を保つために，ここでは基準信号 x(t)を

$$x(t) = \sin(\theta_t) = \sin(2\pi ft) \quad (2)$$

とし，対となる信号 y(t)を

$$y(t) = \sin(\theta_t + \varphi) = \sin(2\pi ft + \varphi) \quad (3)$$

とする．この 2 つの信号の瞬時値対の関係の例として φ= π /2 すなわち 1 周期で直交となる場合を Fig.6 に示す．図は横軸に t に相当する離散番号（sample No.），縦軸に振幅値を示す．図には 3 対を信号対の例として縦線で 3 種類の θ （θ_1, θ_2, θ_3）を示している．図中 x(t)に相当する離散値を「×」，y(t)に相当する離散値を「＋」で示す．この信号のリサージュ図形は Fig.7 となる．Fig.6 の縦線で示した 3 種類の θ が Fig.7 で確認できる．純音の位相差は群遅延では基本周波数に依存し，また純音の位相差は物理的には 1 周期以内，聴覚的には知覚可能な条件を満たさなければならない．ここでは信号対の関係が一定である場合かつ 2 信号の位相差 φ と想起される空間印象の関係を調べることが目的である．そこで，純音を基にランダムノイズを作

ⓒ 2017 Information Processing Society of Japan

Vol.2017-MUS-115 No.29
2017/6/17

成する．ここでは信号対の関係を維持するために，時間情報をランダム化する．これにより生成された 2 信号を Fig.8 に示す．Fig.8 の縦線で示す 3 種類の θ に着目すると Fig.6 で示した信号対の関係が保たれていることと，θ_t の t が変わったことがわかる．

ここでは 2 信号間の位相差 φ を 0（同相）から π（逆相）までを

$$\varphi = 0\frac{\pi}{6}, 1\frac{\pi}{6}, 2\frac{\pi}{6}, 3\frac{\pi}{6}, 4\frac{\pi}{6}, 5\frac{\pi}{6}, 6\frac{\pi}{6} \quad (4)$$

の 7 種類とし，刺激音を作成する．作成条件は，サンプリング周波数 fs=44100 (Hz)，量子化ビット数 16bit，式（2）（3）の f=440 Hz とする．作成した信号のリサージュ図形を Fig.9 に示す．ここでは刺激音の長さを 2 秒とし，主観評価実験で使用する刺激音は立ち上がりおよび立下りが滑らかな振幅変形である必要があるため，ここでは Fig.10 に示す sin 二乗窓を信号対の包絡線として使用する．

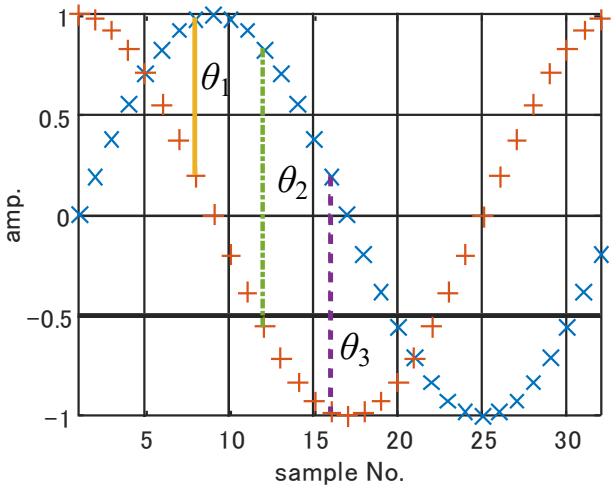


Fig.6　信号間の関係が一定条件となる信号対の 1 周期
（縦実線：θ_1，縦一点鎖線：θ_2，縦破線：θ_3）

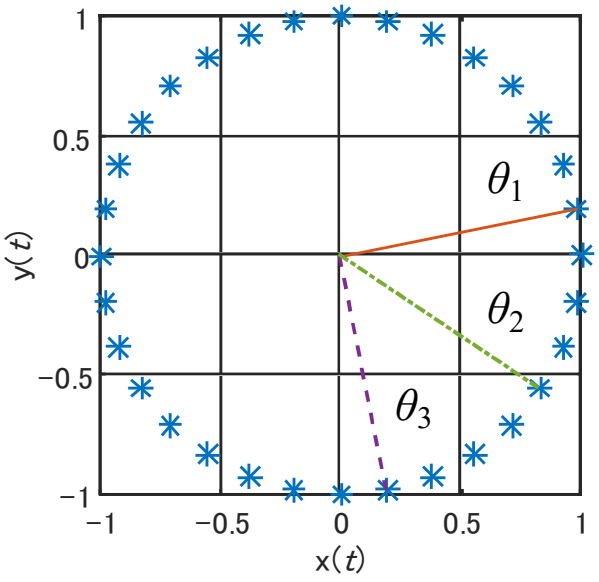


Fig.7　2 信号のリサージュ図形
（縦実線：θ_1，縦一点鎖線：θ_2，縦破線：θ_3）

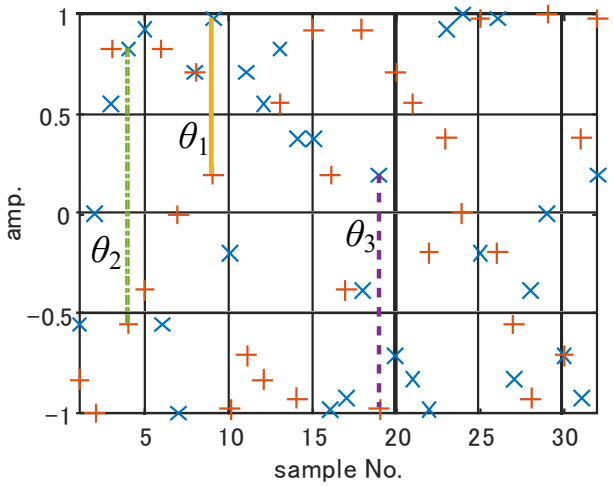


Fig.8　信号対の時間情報をランダム化して作成したランダムノイズ
（縦実線：θ_1，縦一点鎖線：θ_2，縦破線：θ_3）

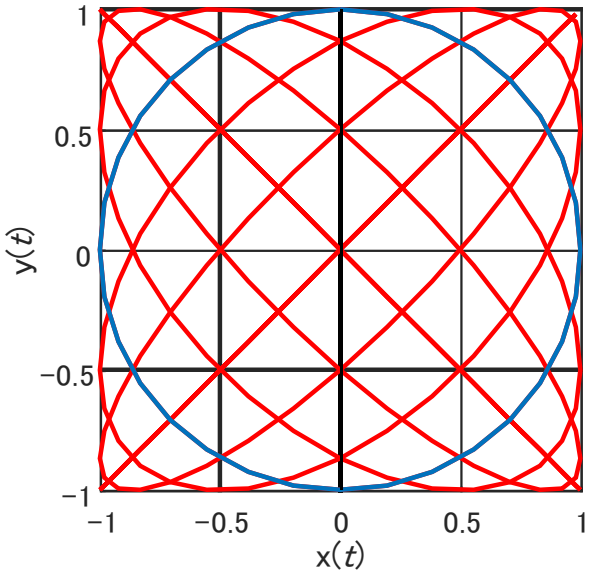


Fig.9　φ が同相から逆相まで 7 種類として作成した信号のリサージュ図形

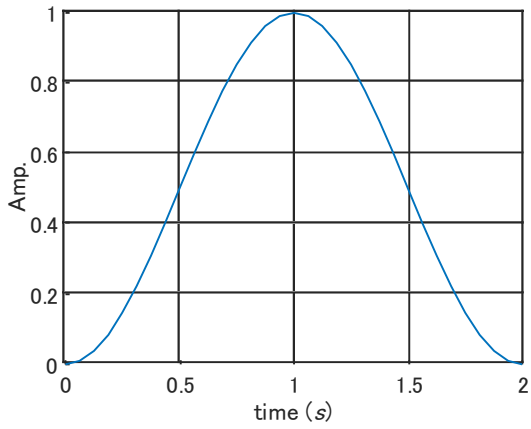


Fig.10　主観評価実験で使用する刺激音とするために使用する包絡線

ⓒ 2017 Information Processing Society of Japan

主観評価実験は，一対比較法・二者強制選択法を用いて行うこととする．ここでは，刺激音提示をヘッドフォンで行い，1 つの刺激音対を，Fig.11 に示すように，マーカ信号，無音（1 秒），刺激音 1（2 秒），無音（1.5 秒），刺激音 2（2 秒），とする．実験協力者は刺激音対を聴き終わった時点で，1 つ目の刺激音と 2 つ目の刺激音のいずれの刺激で想起した空間印象が広いかを回答する．回答が終了した時点で次の刺激音対を提示する．

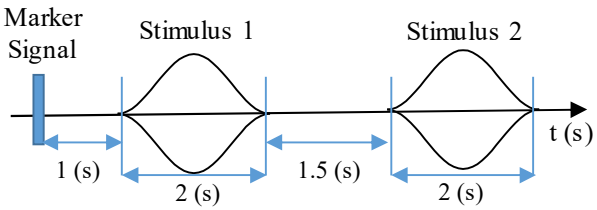


Fig.11　刺激音対の提示パタン

実験サンプル数 11 で得られた結果を Fig.12 に示す．ここでは最も広いと感じなかった刺激音がマイナス最大値，最も広いと感じた刺激音がプラス最大値となるように，広いと判断された刺激音のスコアを 1，判断されなかった刺激音のスコアを −1 とした．Fig.12 の横軸は 2 信号間位相差 φ，縦軸は選択度数（selectivity）を示す．縦軸の最大値は 66（11 サンプル×6 対（同一対を除く））である．図には実験結果から求めた近似曲線も併せて示す．

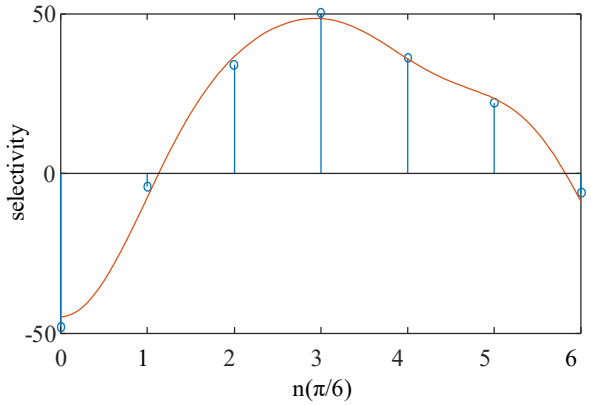


Fig.12　主観評価実験結果と近似曲線
（横軸：2 信号間位相差 φ，縦軸：選択度数）

Fig.12 は，同相（モノラル）で最も空間印象が広くないと判断している，直交（π/2 ）で空間印象が広いと判断している，逆相は φ=π/6 と同程度の空間印象と判断している，ことを示している．また，φ が小さい範囲で変化が大きく，φ が π/2 よりも大きな範囲で変化が緩やかであることを示している．これは最小可聴角（MAA: Minimum Audible Angle）[23]を考えると妥当な結果であると判断する．

心理モデル化するには主観評価実験結果を尺度化するべきであるが，そのためには実験結果を統計処理で信頼性を確保できるだけのサンプル数が必要である．このことから，さらなる検討が必要であるが，モデル化およびパラメトリック表現の可能性が確認されたと判断する．

4. おわりに

本研究は音場共有に必要な音場情報のパラメトリック表現の確立を目的としている．本稿では，1）音響現象を発生する電気音響変換器（ラウドスピーカ），2）両耳差情報の一つである両耳間位相差，に着目し，モデル化とパラメトリック表現の可能性を調べ，以下の項目を確認した．

1）電気音響変換器（ラウドスピーカ）

実測のインパルス応答から求めた振動変位と呼吸球モデルから導出した振動変位が検討は必要ではあるものの，整合性が確認できた．このことから，電気音響変換器の呼吸球モデルとそのパラメトリックな表現により記述できる可能性があると判断する．

2）両耳間位相差

両耳間位相差のみで想起される印象に違いを生じ，同相が最も広くなく，直交で最も広く，逆相で同相よりも広く感じる，という既知の事象が確認されたことから適切な主観評価実験が実施できたと判断する．心理モデル化するには尺度化が必要である．統計的な信頼性確保には十分なサンプル数が必要である．これらの検討が必要であるが，主観評価実験で得られた選択度数から，モデル化およびパラメトリックな表現により記述できる可能性が確認されたと判断する．

以上のことから，1）音場を再構築する装置である電気音響変換器（ラウドスピーカ），2）空間知覚の一つとして両耳間位相差と想起印象，のモデル化及びパラメトリックな表現による記述の可能性が確認できたと判断する．

謝辞

時間領域での考え方およびパラメトリック表現に多大なご協力を賜りました WSS（Wave Science Study）所長・東山三樹夫先生に感謝いたします．

ⓒ 2017 Information Processing Society of Japan

参考文献

[1] R.Braham, R.Comerford, "Sharing Virtual Worlds", IEEE SPECTRUM, pp.18-51, vol.34, No.3, 1997

[2] M.Cohen and N.Koizumi, "Exocentric control of audio imaging in binaural telecommunication", IEICE Trans. E75-A,pp.164-170,1992

[3] R.Koenen, "MPEG-4:multimedia for our time", IEEE SPECTRUM, pp.26-33, vol.36, No.2, 1999

[4] Barry L. Vercoe, William G. Gardner and Eric D. Scheirer, "Structured Audio : Creation, Transmission, and Rendering of Parametric Sound Representations", Proc. of the IEEE, pp.922-939, Vol. 86, No.5, May, 1998

[5] ISO/IEC JTC 1/SC 29/WG 11,N2503-sec5, 1999

[6] MIDI Manufactures Association, "The Complete Detailed MIDI 1.0 Specification", 1996

[7] "The Virtual Reality Modeling Language", International Standard ISO/IEC 14772-1:1997

[8] Manabu Fukushima, Hirofumi Yanagawa, "A Perceptional Modeling of Acoustic Events Focused on Spatial Impression", Proc. of The 17th International Congress on Acoustics, vol.4, 2001

[9] Manabu Fukushima, Hirofumi Yanagawa, "A Perception based approach for Acoustic Events Modeling in Interactive Sound Field Network", Proc. of The 8th International Conference on Auditory Display, pp.387-393, 2002

[10] Kaoru Miyazaki, Kiyoshi Nakayama, Ken'itiro Kamura, Manabu Fukushima, Hirofumi Yanagawa, " Evaluation of Sound Field Perception by Psychological Scale based on Sound Field Expressing Words for Virtual Audio Space Sharing ", AES 11th Regional Convenstion, pp.120-123, 2003

[11] Kiyoshi Nakayama, Ken'itiro Kamura, Isao Yako, Manabu Fukushima, Hirofumi Yanagawa, "An investigation for rendering 'Spacious' by EDT with measured Impulse Response", Proc. of ASJ-ASK Kyushu Youngnam Joint Conference on Acoustics 2003, pp.69－72, 2003

[12] Manabu Fukushima, Kiyoshi Nakayama, Hirofumi Yanagawa, "'Spatial' impression enhancement by changing the initial envelope of impulse responses", The Third International Workshop on Spatial Media '02-'03, pp.17-20, 2003

[13] Manabu Fukushima, Kiyoshi Nakayama, Isao Yako, Hirofumi Yanagawa, "Effect Of The Envelope In Impulse Response For Changing The Acoustic Scene Of 'Spacious'", The 8th Western Pacific Acoustics Conference, WE13 (CD-ROM), 2003

[14] Manabu Fukushima, Hiroto Inoue, Ken'itiro Kamura, Hirofumi Yanagawa, Ken'iti Kido, "A method for the determination of noise factor in estimated transfer function - cross spectral technique by use of 1-0 and 1-000 windows -", Proc. of The 18th International Congress on Acoustics, pp.166-169, vol.25, no.2, 2004

[15]Manabu Fukushima , Hirofumi Yanagawa, " An investigation of the disturbance factor of 'Spacious' impression and its control in a re-produced sound filed", AES 12th Regional Convention Tokyo, pp.188-pp.191, 2005

[16]Manabu Fukushima , Takeshi Shiraishi, Takayuki Hoshishiba, Tetsuya Akahoshi, Ken Ito, Hisao Okamoto, and Hirofumi Yanagawa, " An investigation on the changes of the spatial impression in reproduced sound fields according to the height of the loudspeakers ", AES Conference Fukuoka 2006,P24 , 2006

[17]Takeshi Shiraishi, Manabu Fukushima, Hirofumi Yanagawa, " An investigation of the relation between audio signal compression and “spacious”", AES Conference Fukuoka 2006,P02 ,2006

[18] 東山三樹夫,”音の物理”，コロナ社，pp. 172−pp. 173，2015 年

[19]Yanagawa et al.,"Effect of transient signal length on cross-correlation functions in a room", J. Acoust. Soc. Am. 84, 1728–1733 (1988)

[20]Hirofumi Yanagawa, Music and Concert Hall Acoustics (Academic Press, London, 1997), pp. 251–260.

[21]Hirofumi Yanagawa, et al., "Interaural Correlation Coefficients and their Relation to the Perception of Subjective Diffuseness", ACUSTICA, Vol.71, pp.230-pp.232,1990

[22]Hirofumi Yanagawa, and Mikio Tohyama, "Sound Image Broadening by a Single Reflection Considering Temporal Change of Interaural Cross-Correlation", Acta Acustica united with Acustica 87(2), pp.247-pp.252, March 2001

[23]A. W. Mills, "On the Minimum Audible Angle", The Journal of the Acoustical Society of America 30, 237 (1958)

ⓒ 2017 Information Processing Society of Japan