

修士論文

複数くぼみつきマイクロホンバッフルを用いた
360°映像のための立体音生成に関する研究



平成 29 年度

三重大学 工学研究科

情報工学専攻

今村健人

2018 年 3 月

目次

第 1 章	序論	7
1.1	研究背景	7
1.2	本稿の構成	9
第 2 章	マイクロホンバッフル	13
2.1	左右前後に対称なくぼみつきマイクロホンバッフル	13
2.2	聴点を増やした場合のくぼみ付きマイクロホンバッフルの設計	17
2.3	マイクロホンバッフルの音響特性	19
第 3 章	任意視線に対応する立体音の生成	23
3.1	音響信号の統合手法	23
3.2	実験・結果	26
第 4 章	結論	33
	参考文献	35
	謝辞	37

要旨

実空間の音や映像を記録し再現するための技術として、音響では立体音響技術、映像では全天球映像技術がある。立体音響は3次元的な音響空間を収録・再現する技術であり、音の方向感や距離感などが再現できる。立体音収録の代表技術として、人間の頭部を模したダミーヘッドを用いたバイノーラル録音があり、高臨場感音響再生が可能となっている。また、全天球映像技術により、観光名所や街並みを任意の視点で眺めることが可能となった。この全天球映像と立体音とを同時に用いることができれば、より臨場感の高い「場」の収録・再生が可能となると考えられる。しかし、従来のバイノーラル録音においては、一度の録音では単一の視線方向のみ録音されるため、全天球映像におけるユーザの視線方向に対応した立体音収録は非常に困難である。

本研究では、複数くぼみを持つ球状マイクロホンバツフルを用いた立体音収録システムによる、任意視線に対応した立体音響信号の生成法について述べる。先行研究で設計・評価された、複数のくぼみ内の受音点において数値解析手法により求められた音響伝達関数を用い、任意視線に対応する立体音響信号の生成方法の提案ならびに評価を行った。本研究では、ユーザの視線方向に対応するように、収録信号に重みづけし、線形和をとることで、立体音を生成する手法について検討した。実験では、一次関数、三角関数、スプライン補間による曲線による重み関数を検討し、得られた信号をスペクトル歪評価尺度により客観評価した。

実験の結果、一次関数、三角関数やスプライン補間を用いた重み関数を用いた場合にスペクトル歪の平均が3 dB以下となり、音響信号のみで音源方向が判別できる可能性が示唆された。またスペクトル歪の最大値は約7 dBであったが、このような大きな歪となる条件であっても、映像と音響とを組み合わせることで、音源方向を正しく知覚させることが可能となると考える。また三角関数を用いた場合では音源と視線方向が異なる場合においても歪の平均値が他の重み関数より小さい結果となった。以上より、本論文では、三角関数を重み関数に用いる立体音響信号の統合が有効であることが示された。

第1章

序論

1.1 研究背景

我々は常にたくさんの音に囲まれて生活している。人間が耳にする音は、音源から鼓膜へ到達する間の伝達経路において、両鼓膜間での伝達時間差が生じるとともに、音圧、周波数特性、位相特性が変化する。音波の到来方向や空間の広がり、これらの情報によって知覚されている。オーディオ機材による録音再生において、人間が実環境で聴取し、知覚するような方向感や距離感を感じさせる技術を立体音響技術と呼ぶ。

立体音響は、音の収録、再生において、音源から聴取者の鼓膜までの伝搬経路を再現することで実現されている。立体音収録の代表的技術として、Fig.1.1 に示されるようなダミーヘッドを用いたバイノーラル録音・再生があり、多くの立体音収録がなされている。ダミーヘッドは人間の頭部を模擬した形状をしており、人間の外耳道入口や鼓膜位置に相当する場所にマイクロホンが設置されている。この2チャンネルのマイクロホンを用いてバイノーラル録音をすることにより立体音の収録を実現している。立体音響は鼓膜までの伝搬経路が考慮され、現実で聴取するような感覚を与えることができるため、ステレオ録音よりも優位な音響情報として研究されている。

一方、映像分野においては、全天球カメラなどを用いることで任意の視線に対応可能な動画や静止画像の撮影が可能である [1]。この全天球映像に対して立体音を付与し、より臨場感のある空間情報の記録と再生が可能となると注目されている。映像分野ではユーザの視線に対応できる様な多方向で動的な映像の撮影が可能である。しかし、立体音響の問題点として、Fig.1.2 に示す様な通常のバイノーラル録音では、単一聴点の音しか収録できないことが挙げられる。ダミーヘッドは前後左右に非対称の形状であるため、これを用いた収録法では収録時のダミーヘッドの向きに対応した聴点のみの音が録音される。信号処理による音響伝達関数の補正によって正面方向を疑似的に変更することは可能であるが [2, 3]、動的な収録再生を行う場合や、同一の音場を再現できない場合、Fig.1.3 に示す様に一度の収録で多方向の視線に対応した録音ができることが望ましい。このような任意のユーザの向きに対応した立体音収録において、マイクを複数使用した多方向収録について、球形バッフルを用いた技術が提案されている [4, 5]。Fig.1.4 に示すように球形のバッフルに複数のマイクロホンを設置し

多方向の収録を行い、収録信号に対して聴取者の頭部伝達関数 (HRTF; Head Related Transfer Function) を用いることで立体感を付与している。このシステムの問題点として、非常に高価なシステムであり、かつ信号処理技術が複雑であることが挙げられる。また、複数の耳介を持つ収録システムや、耳介形状を模擬したくぼみを複数有するマイクロホンバッフルも提案 [4] されている。これらのシステムでは、安価で扱いやすい形状のバッフルを作成し、ユーザの視線方向に対応するように耳介やくぼみを選択することによって、所望の立体音収録を可能としている。しかし、音源と視線方向に対応する適切な耳介やくぼみを選択ができない場合 (Fig.1.5) には、構造的に立体音の収録がしづらい聴点が存在してしまう。そこで、本論文では先行研究 [4, 5] で使用されたマイクロホンバッフルを用いて、任意の視線方向の立体音を生成することが可能になるよう、各くぼみで収録された信号を統合する手法について述べる。従来方法では対応不可能な視線方向の場合については、ダミーヘッドやバッフルの向きを合わせて複数回の収録を行わなければならなかったが、音響信号を合成し、聴点に対応する信号を生成することが実現すれば、一度の収録で複数方向の立体音を収録可能になり、より高度な空間情報の取得が実現可能となると考えられる。

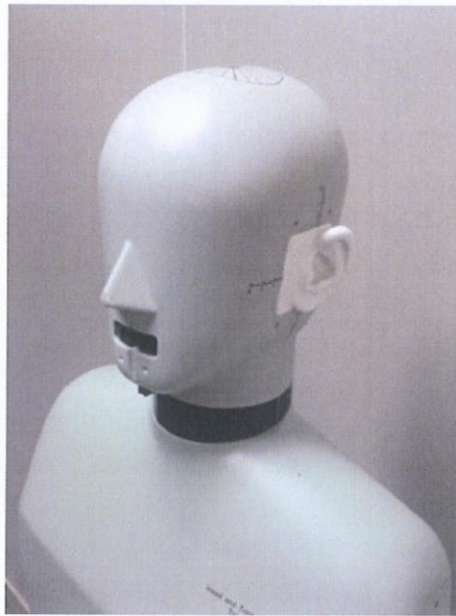
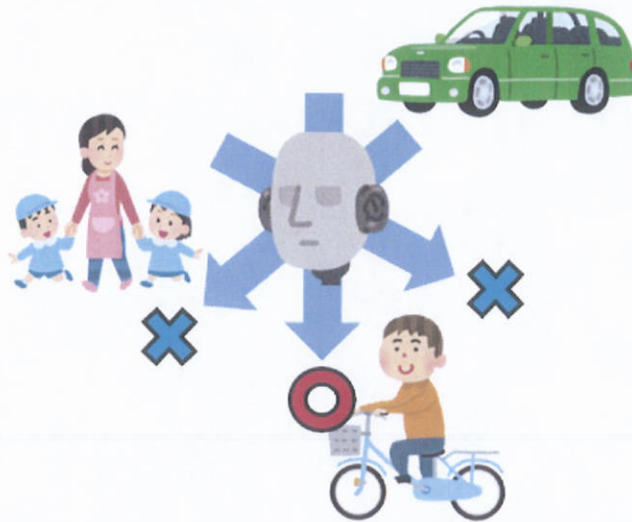


Fig.1.1: ダミーヘッド [6]

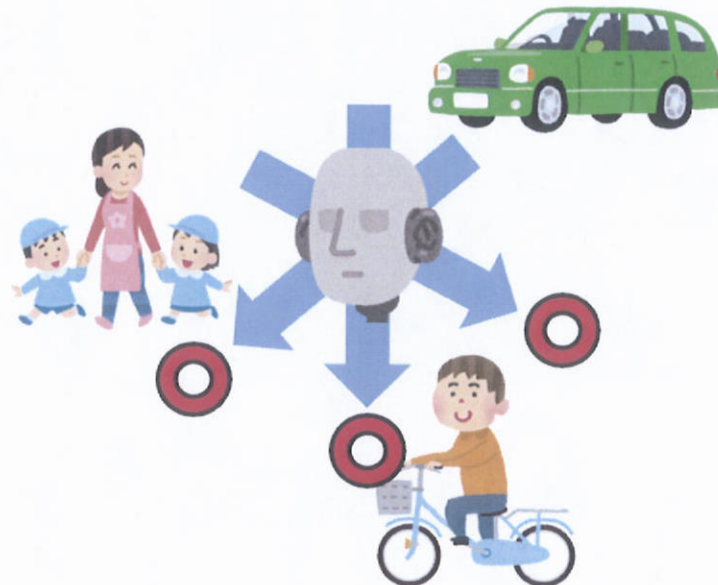
1.2 本稿の構成

本論文の構成は次の通りである。第2章で本論文で用いるマイクロホンバッフルの仕様について述べる。第3章では、音響信号の合成法と実験や結果について述べ、第4章で結論を述べる。



同時刻での複数聴取方向の立体音収録に不向き
同時刻の収録では単一方向のみの収録に限られる

Fig.1.2: 多聴点の収録に対応していない場合



同時刻で複数聴取方向の立体音収録が可能

Fig.1.3: 多聴点の収録に対応させた場合

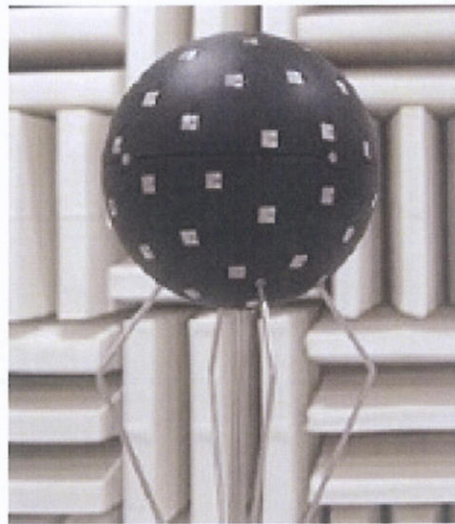
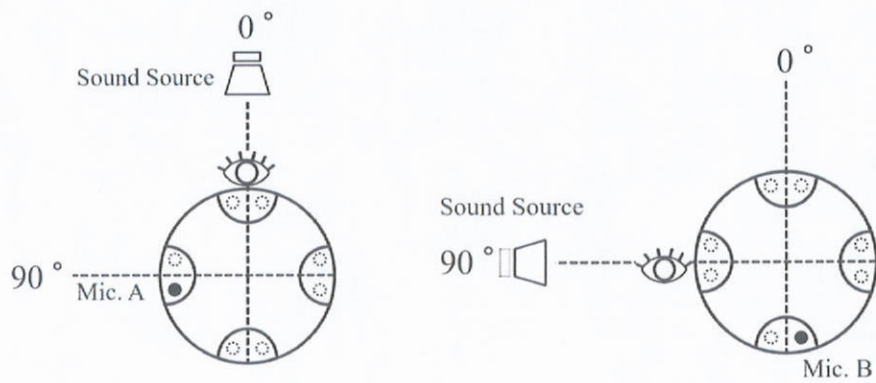


Fig.1.4: 252ch 等密度球状アレイ [3]

収録可能

音源位置: 正面(0°), 視線方向: 正面(0°)
 音源位置: 左(90°), 視線方向: 左(90°)



収録困難

音源位置: 左前(45°), 視線方向: 左前(45°)

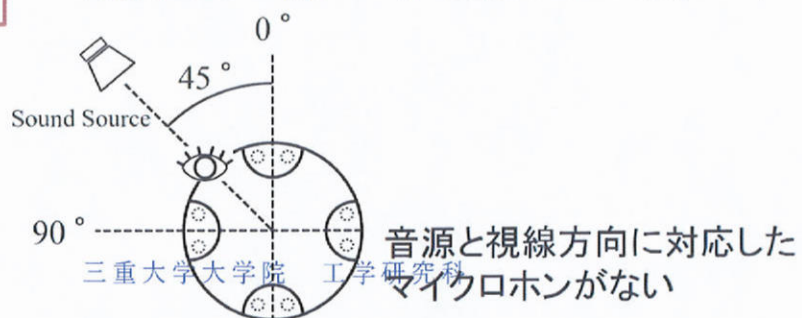


Fig.1.5: 立体音収録が可能な視線方向と音源の方向

第2章

マイクロホンバッフル

マイクロホンバッフルとはマイクロホンに装着し、音波の回折や反射を利用し収録される音響信号に音響特性を付与する機材である。主に形状や材質によって音響特性が異なる。人間の頭部を模擬したダミーヘッドも立体音を収録可能なマイクロホンバッフルと言える。先行研究 [4] では、立体音を付与するため、球状のバッフルの外側にくぼみとマイクロホンを設置したマイクロホンバッフルを設計し評価している。また、全天球映像に対応させるため複数のくぼみを持つマイクロホンバッフルを提案している。本論文では、先行研究 [4] で提案されたマイクロホンバッフルについて数値解析をし求められた伝達関数を用いて、音響信号の統合について検討した。そのため本章では、先行研究で設計されたマイクロホンバッフルの仕様や音響特性について述べる。

2.1 左右前後に対称なくぼみつきマイクロホンバッフル

先行研究 [4] では、全天球バイノーラル録音のためのマイクロホンバッフルの作成に先立ち、前後、左右に対象なマイクロホンバッフルの設計と音響特性の調査を行っている。先行研究で作成されたマイクロホンバッフルは、システムの扱い易さを考慮し、人間の頭部の半分ほどの大きさになるよう直径 10 cm の球状をしている (Fig.2.1)。球状バッフルは、耳甲介腔を模擬した 2 つのくぼみが設置され、それらのくぼみは半球状で深さは 1 cm である。両くぼみの最深部は、マイクロホンバッフルの両端にそれぞれ位置する。受信位置は、Fig.2.2 に示すように人間の外耳道入口の位置に倣い、前後・上下が非対称となるように考慮した位置に設定されている。これらの仕様に基づき、数値解析による音響特性の算出がされている。

先行研究で設計されたマイクロホンバッフルは、3DCAD(PTC Creo Elements/Direct Modeling Express) により設計された。このマイクロホンバッフルの三次元データに対して数値解析的に音響特性を求めるため、メッシュ分割ソフトウェア (Gmsh) によるメッシュ処理が施され、4832 個の三角形のメッシュに分割されている。このメッシュデータを用いて、音響解析ソフトウェア (サイバネット, WAON) により、100~16000 Hz (100 Hz 刻み) の帯域における音響解析が数値解析的に行われて

いる。なお、この数値解析では、境界要素法が使用されている。解析条件は、音源に 1 Pa の音圧の点音源を、中心からの距離 1 m の位置に設置した。この点音源を、水平面上を反時計回りに 5° 間隔で移動させて、音響特性を求めた。ただし、マイクロホンバッフルの正面が 0° である。音源位置の詳細について Fig.2.3 に示す。この数値解析により 72 方向の音源からの伝達関数を得ている。音源と左マイクとの間で求められた伝達関数を Fig.2.4 に示す。この図では、音響特性を見やすくするため、音圧レベルの最大値を 0 dB とし、-60 dB 以下の値は-60 dB が代入されている。横軸が音源の方位角、縦軸が周波数、色が音圧レベルとなっている。方位角ごとに、周波数特性が変動していることが確認できる。方位角 90° と 270° は、両くぼみの方向に対応しており、得られた伝達関数から明確な音圧差があることが確認できる。また、前後方向である 0° と 180° においても周波数特性に違いが現れ、特に 8 kHz 以上の高音域に強く見られた。これらの点から先行研究では、設計されたマイクロホンバッフルを用いて立体音収録が可能であると述べられている。



Fig.2.1: くぼみ付き球状マイクロホンパッフル [4]

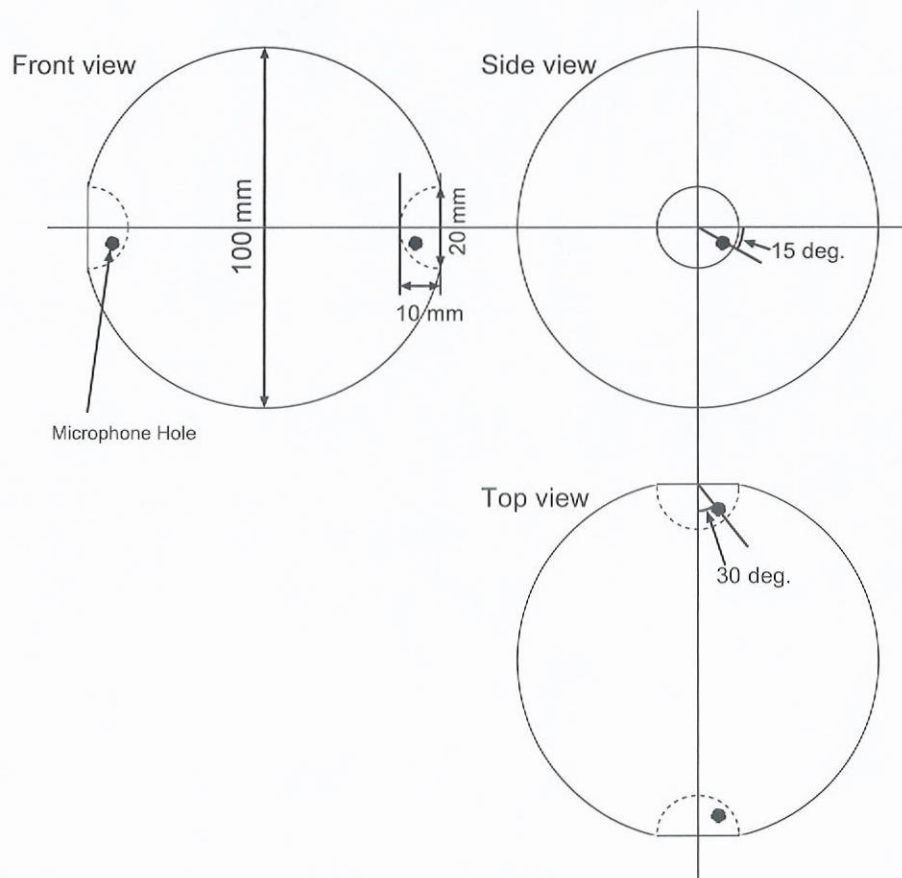


Fig.2.2: くぼみ付き球状マイクロホンパッフルの仕様 [4]

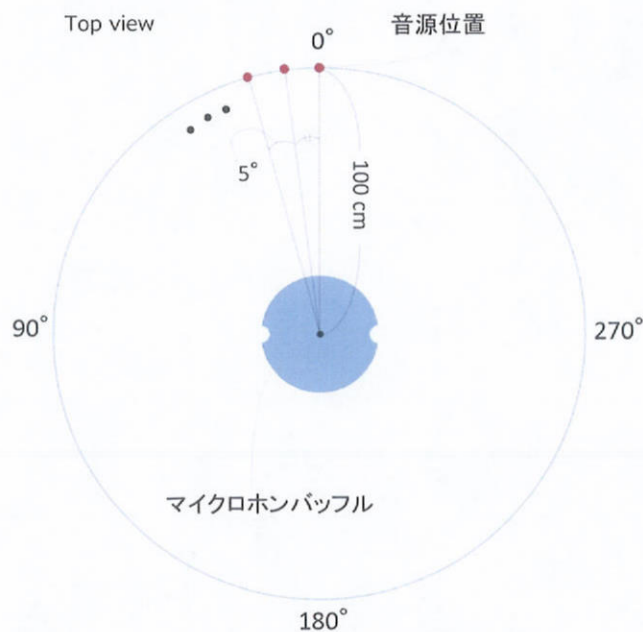


Fig.2.3: 解析における音源の位置設定 [4]

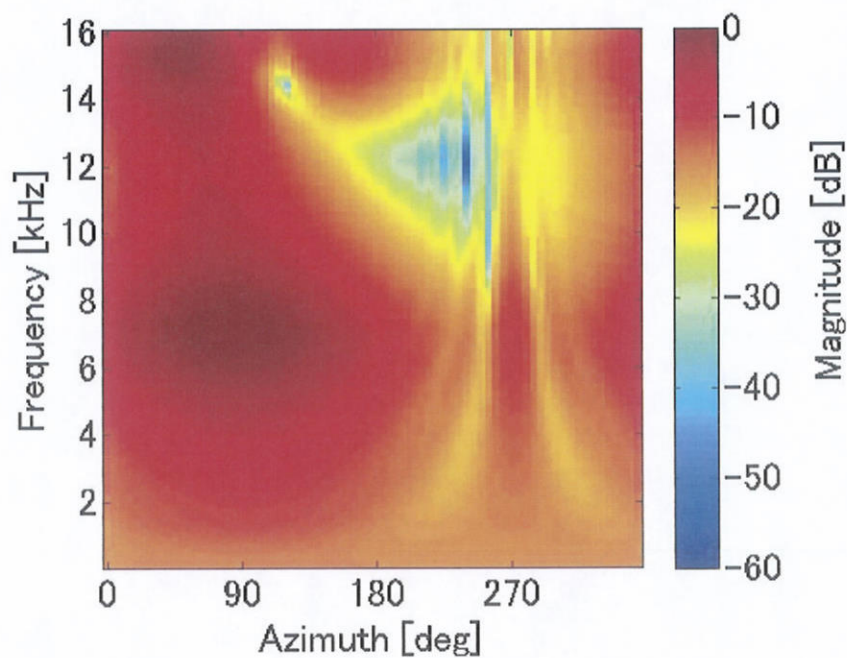


Fig.2.4: 数値解析によるバツフルの左マイクでの伝達関数 [4]

2.2 聴点を増やした場合のくぼみ付きマイクロホンバッフルの設計

先行研究では、左右前後に対称なくぼみつき球状マイクロホンバッフルについて立体音の収録が可能であることが示された。この結果を踏まえて、先行研究では全天球バイノーラル収録に対応するバッフルも設計、評価されている。二つのくぼみを有するマイクロホンバッフルでは、従来技術と同様、単一の聴点でしか立体音の収録ができない。そこで、全天球バイノーラル収録に対応するために、耳に相当するくぼみの数を増加させることにより、多聴点に対応させることを行っている。くぼみを左右、前後、上下の六ヶ所に設置した場合の音響特性の変動について示す。10 cm のくぼみ付き球状マイクロホンバッフルのくぼみの数を6個にし、それ以外の条件は同一とし、Fig.2.5 に示す 3DCAD データを作成した、メッシュ処理を施し、7392 個の三角形で構成した。このメッシュデータに対して、数値解析による音響特性の算出を行った。ここで、解析される受信位置については、二つのくぼみのマイクロホンバッフルと同一の位置である。音源位置ごとの伝達関数を Fig.2.6 に、10 cm のくぼみ付き球状マイクロホンバッフルとの差分を Fig.2.7 に示す。この結果より、くぼみの増加によって変化した特性はほぼ ± 2 dB と僅かであり、二つのくぼみのマイクロホンバッフルの立体音響の特徴を保っていることが確認できる。また先行研究 [4] では、両耳間相互相関係数 (IACC ; InterAural Cross-Correlation) と両耳間時間差 (ITD ; Interaural Time Difference)[7, 8, 9] を用いた立体感の評価を行っている。IACC はそれぞれの耳で聴取する音の相互相関係数から得られる相関度であり、ITD はそれぞれの耳で聴取する音の時間差を表す。これらの値を算出することで、音の立体感の評価ができ、人頭の IACC, ITD に近いほど、適切な立体感を付与できると考えられている。実頭 10 人分の IACC, ITD を平均したものと、くぼみ付き球状マイクロホンバッフルの IACC, ITD との比較を行い、IACC, ITD の両値とも実頭の値に沿った増減をしていることが記述されている。

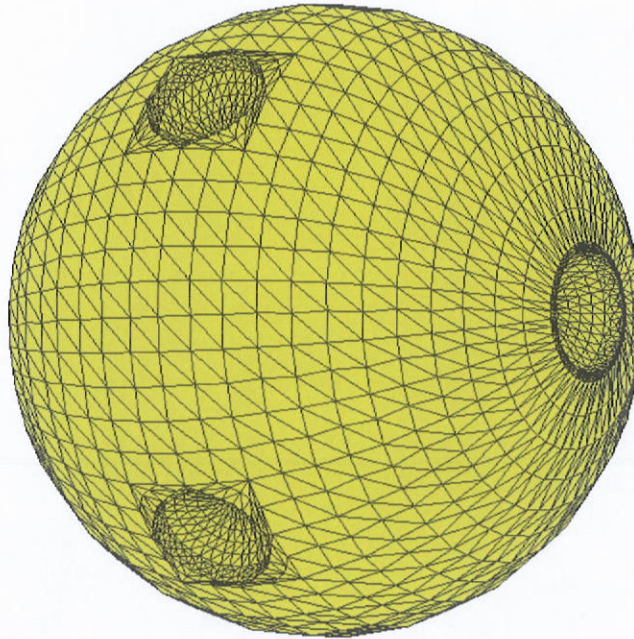


Fig.2.5: くぼみを6個とした球状マイクロホンバツフル [4]

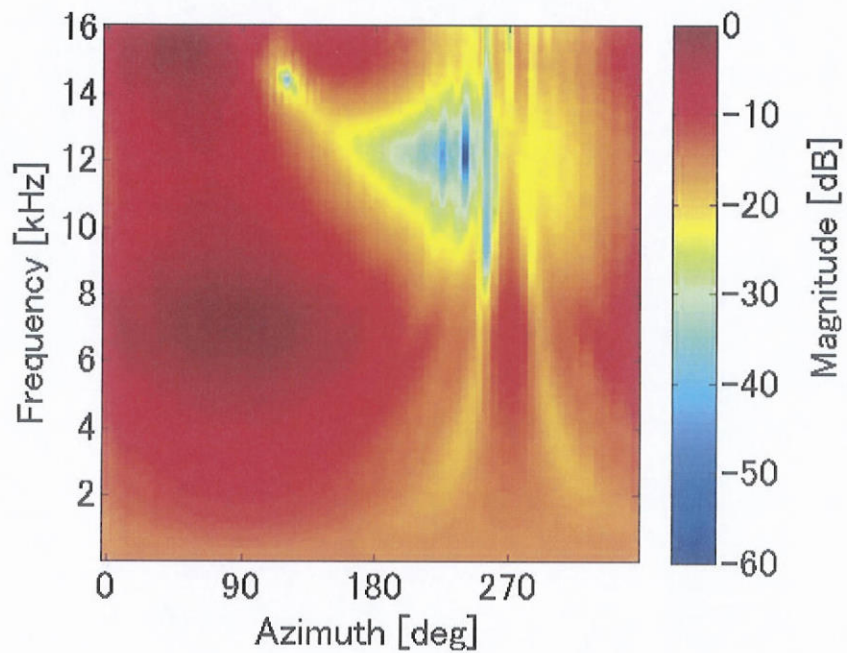


Fig.2.6: くぼみを6個とした球状マイクロホンバツフルの伝達関数 [4]

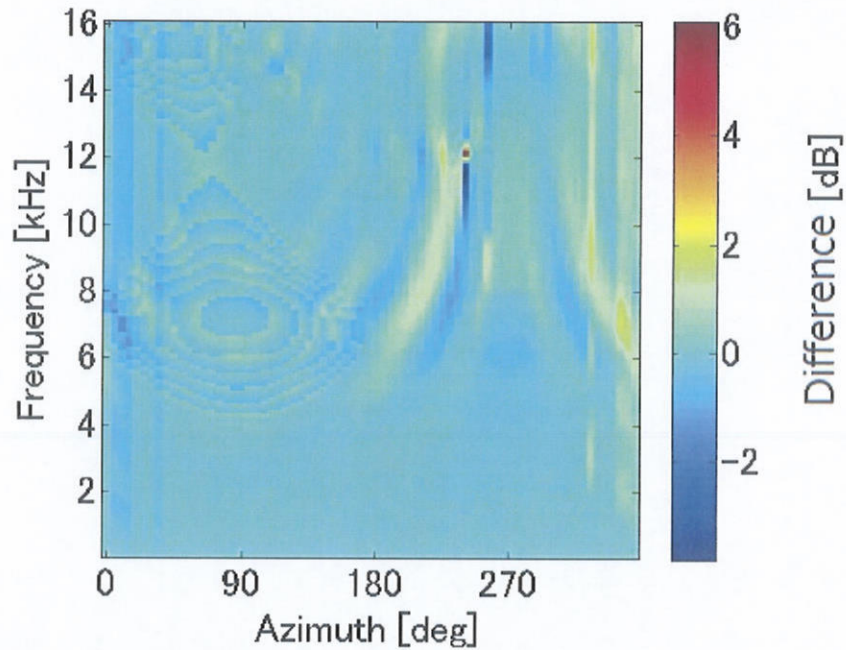


Fig.2.7: くぼみを6個とした球状マイクロホンバツフルと10 cmのくぼみ付き球状マイクロホンバツフルとの特性差分 [4]

2.3 マイクロホンバツフルの音響特性

本節では、先行研究で調査されていない垂直方向の音響特性について解析した結果について述べる。解析には、先行研究と同じ仕様のバツフルデータを用いた。また、先行研究同様に音源には音圧1 Paの点音源を用い、音源とバツフルの中心との距離を1 mとした。音源方向は、Fig.2.8に示すように鉛直方向を -60° ~ 60° の間を 30° 間隔とする5方向とした。なお、解析対象となるバツフルと受信位置は正中面ならびに前額面に対して対称形であるため、水平方向を 0° ~ 90° の間を 5° 間隔で19方向の受信位置の解析結果を統合することで、水平方向 360° 分の解析結果を得ている。Fig.2.9~Fig.2.13に、音源位置を鉛直方向 60° 、 30° 、 0° 、 -30° 、 -60° とした場合の解析結果をそれぞれ示す。解析の結果、音源の角度を鉛直方向に変化させたとしても、音源方位ごとに異なる両耳間音圧差が確認できた。また、垂直方向で比較した場合も異なる音圧差が得られることが確認できた。

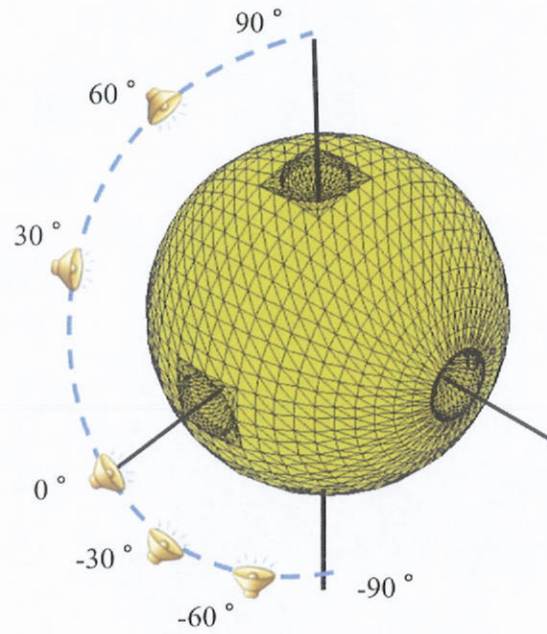


Fig.2.8: 垂直方向での解析位置

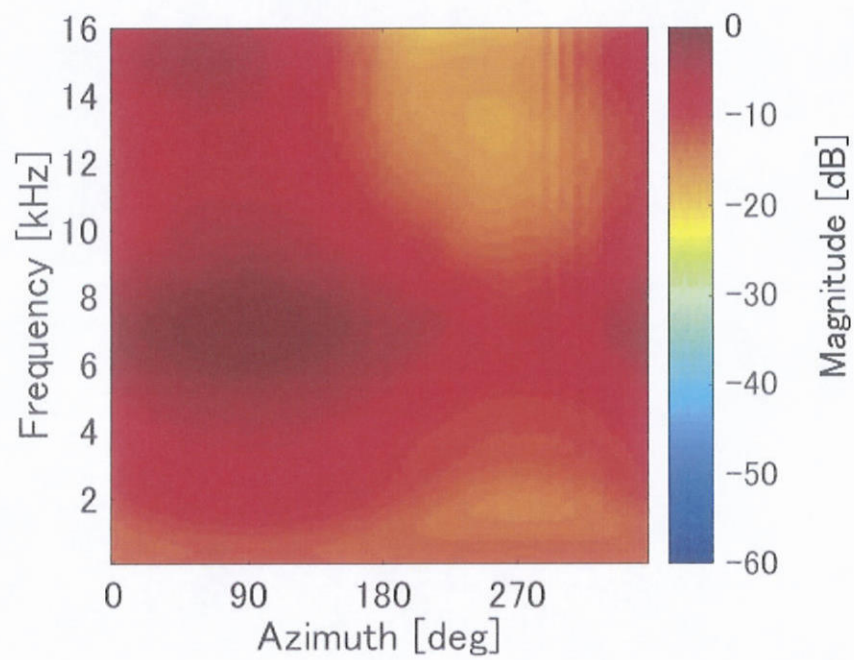


Fig.2.9: 垂直方向 60°での解析結果

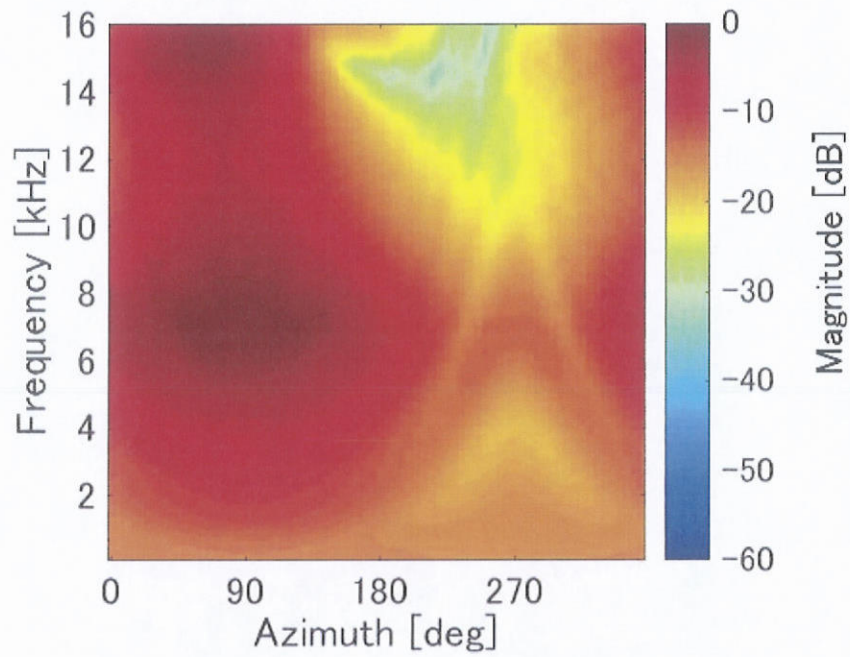


Fig.2.10: 垂直方向 30°での解析結果

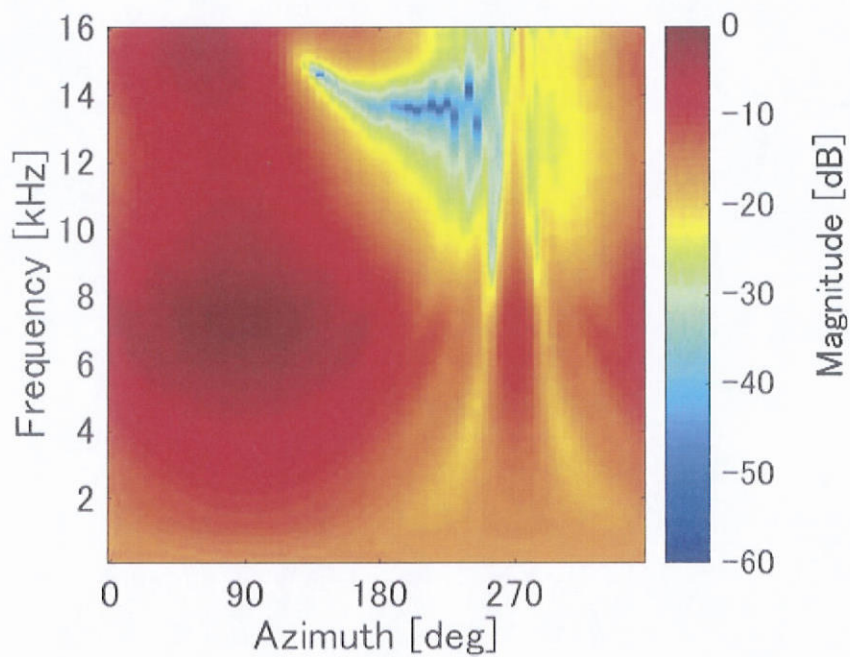


Fig.2.11: 垂直方向 0°での解析結果

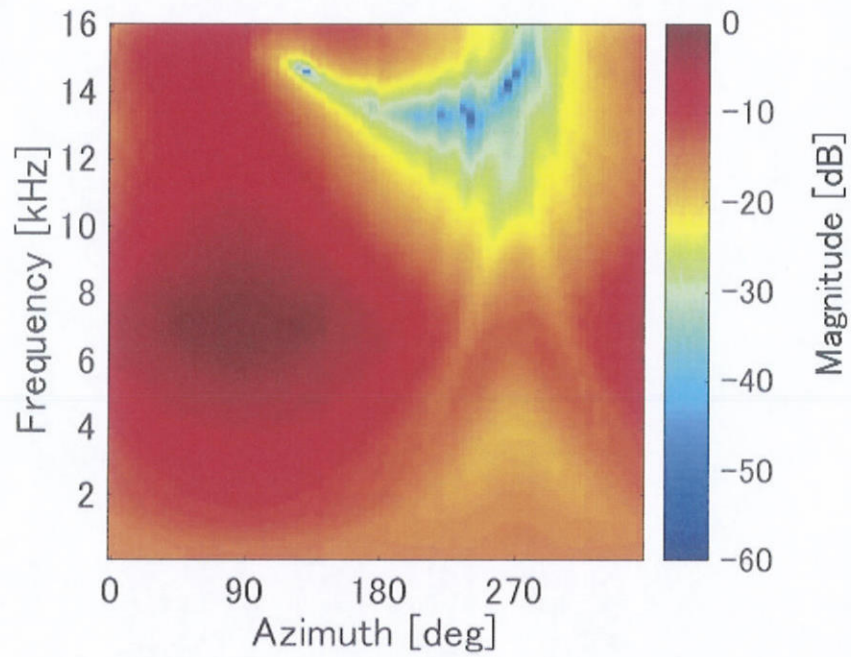


Fig.2.12: 垂直方向-30°での解析結果

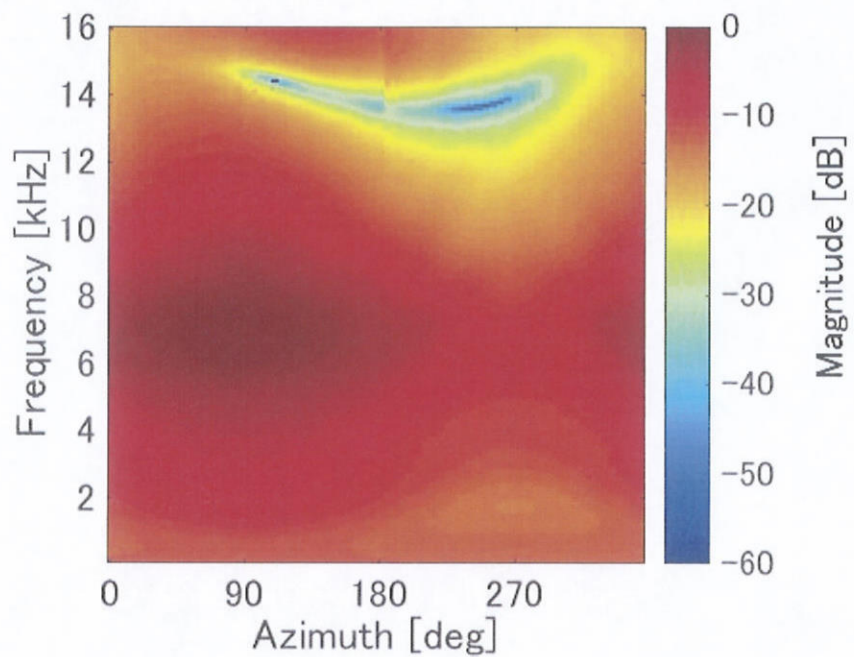


Fig.2.13: 垂直方向-60°での解析結果

第3章

任意視線に対応する立体音の生成

先行研究で設計・評価されたマイクロホンバツフルを使用することにより、ある程度の立体感が得られる事が数値解析的に確認された。すなわち、このマイクロホンバツフルでは、ユーザの視線方向に対応するように耳介やくぼみを選択することができれば、所望の立体音収録が可能となると考えられる。しかし、ユーザの視線方向は任意の方向に変えることができるが、設計されたマイクロホンバツフルでは一度の収録で得られる立体音は前後左右方向の4方向に限られる。そのため、視線方向と対応する適切な耳介やくぼみを選択できない場合、新たにくぼみを増やして集音する方法や、収録された音信号を用いて、視線方向に対応した立体音を疑似的に生成する方法が考えられる。しかし、くぼみを増やして收音点を増やす方法ではくぼみの数を増やしていくと元の形である球が変形してしまい、球状バツフルが持つ音響特性が失われてしまう可能性がある。また、くぼみを増やし続けるのも物理的に限界がある。そのため本研究では、任意視線に対応した立体音の生成手法について考える。本研究では、各くぼみ内で収録された信号統合することで任意視線への対応を図る。

3.1 音響信号の統合手法

各くぼみ内で収録された信号を統合する方法について述べる。ここでは、視線方向に常に音源が位置するような状況を考える。Fig.3.1に先行研究で使用されたバツフルの収録の様子を上から示した図を示す。図に示されるように、くぼみの内部にはマイクロホンが設置されており、音源と視線方向が 0° の場合はMic.Aにより収録される信号を左耳信号として使用する。次にFig.3.2に示すように視線が 90° の場合はMic.Bにより収録される信号を左耳信号として使用する。この2つの例の場合は視線方向に対応するくぼみとマイクが存在するため直接的に立体音を収録することが可能である。

一方、Fig.3.3のように視線が 45° となった場合は、視線方向に対応するくぼみとマイクが存在しないため、直接的に立体音を収録することが困難である。そのためMic.Aの収録信号と、Mic.Bの収録信号を用いて、新たに左耳信号を生成することを考える。統合方法は各マイクで収録された信号に対して、視線方向に対応した重み係数をそれぞれ掛け、これら重み付けされた信号を加算する。新たに統合した信号を

$X_L(t)$, Mic.A で収録された信号を $S_A(t)$, Mic.B で収録された信号を $S_B(t)$ とする。これらの信号の関係を式 (3.1) に示す。

$$X_L(t) = (1 - r)S_A(t) + rS_B(t) \quad (3.1)$$

ここで, r は重み係数であり, この重み係数 r を視線方向の角度に応じて 0 から 1 の間で変化させる。本実験では, 重み係数を式 (3.2) のように視線方向 θ° の累乗 ($n = 1, 2, 3, 4$) で与えた場合, 式 (3.3), 式 (3.4) の三角関数を用いた場合, およびスプライン補間を用いた場合について検討した。

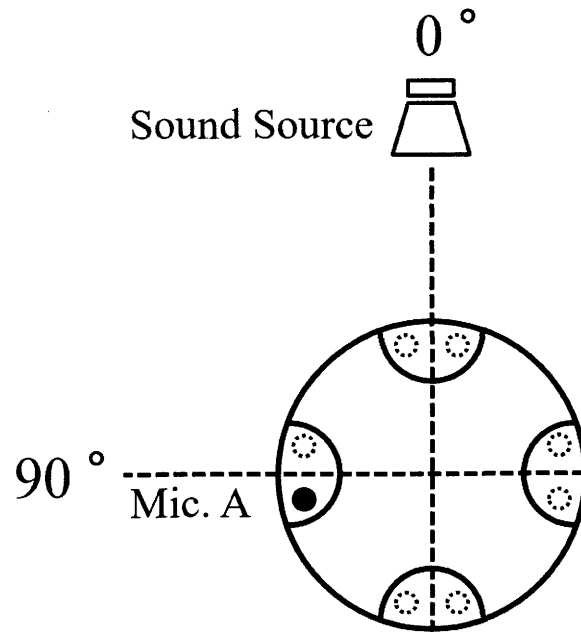
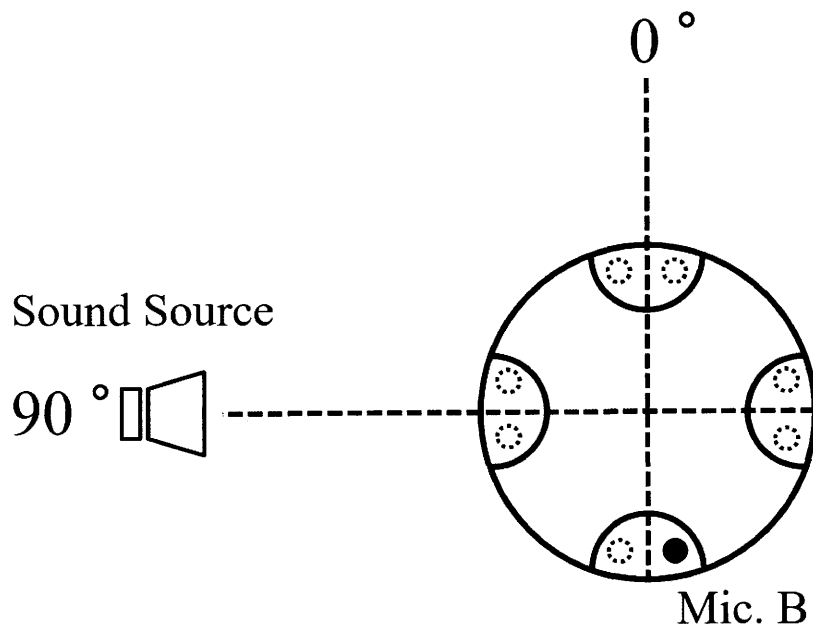
$$r = \left(\frac{\theta}{90}\right)^n \quad \theta = (0, 15, 30, \dots, 90) \quad (3.2)$$

$$r = \sin \theta \quad \theta = (0, 15, 30, \dots, 90) \quad (3.3)$$

$$r = \cos^2\left(\frac{\pi + \theta}{2}\right) \quad (3.4)$$

ここで, θ° は Fig.3.1 に示したような正面から反時計周りの角度を表す。Fig.3.4, Fig.3.5 に式 (3.2), 式 (3.3), 式 (3.4) の重みについて図示する。横軸は音源の存在する角度を示し, 縦軸は r の値を示している。

スプライン補間は, 複数個の制御点を設定し, その制御点を通るような滑らかな曲線を決定する手法である。本実験で検討した重み関数を, Fig.3.6 に示す。音源が $0^\circ \sim 45^\circ$ の範囲では一次関数となるよう, $45^\circ \sim 90^\circ$ 以降は Mic.B の重みが大きくなるように制御点 ((15, 0.16), (30, 0.33), (45, 0.5), (60, 0.75), (75, 0.9), (90, 0.97)) を設定した。これは $0^\circ \sim 45^\circ$ までは一次関数を使用し重み関数を設定した場合の結果が最も良く, $45^\circ \sim 90^\circ$ では音源がマイク A に近づいてくるためマイク A の特性が強く影響すると考えたため, この範囲ではマイク B の特性が強くなるように重み付を行った。また r の値が 1 を超えないといった制約を課した概形となっている。実験では, これら 7 種類の関数により求まる重み係数 r について統合の性能評価を行った。

Fig.3.1: 音源が 0° にある場合収録の様子Fig.3.2: 音源が 90° にある場合収録の様子

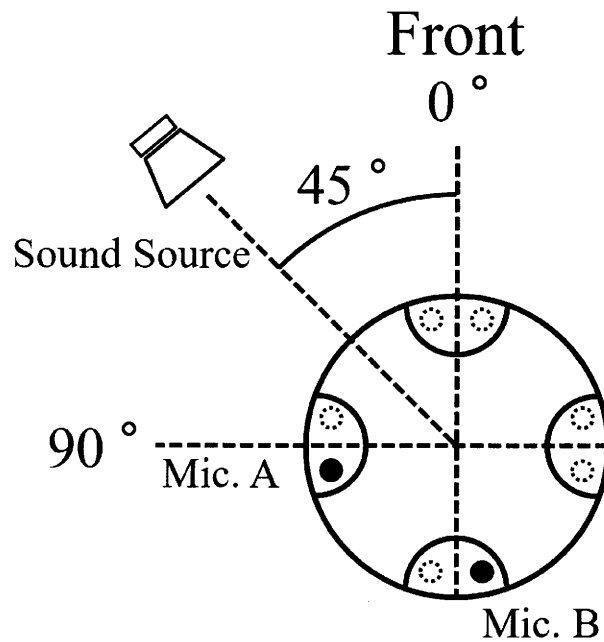


Fig.3.3: 立体音が直接収録しづらい場合

3.2 実験・結果

3.2.1 実験条件

実験では、先行研究で求められた伝達関数を用いた。受信位置については、Fig.3.1, Fig.3.2 に示すように音源が 0° の位置にある際に左耳の位置に相当する Mic.A と、音源が 90° の位置にある際に左耳に相当する Mic.B を使用した。統合した信号は式 (3.5) のスペクトル歪により評価した。 $X_L(f_n)$ は統合した信号の周波数振幅、 $S_{A0}(f_n)$ は正面 0° に音源がある場合の Mic.A における伝達特性の周波数振幅、 f_n は周波数、 N は周波数ビンの総数ある。なお、本実験では $N = 512$ である。また、本実験では常に視線方向に音源があることを仮定した場合と、視線方向から音源が $15^\circ \sim 90^\circ$ まで 15° 間隔でずれていく場合について調査を行った。スペクトル歪を算出するために用いた信号は、視線方向が 0° の時、各角度の条件に合うように Mic.A で求められた信号 $S_A(t)$ を用いた。この使用した信号が目標信号となり、目標信号との差が少なくなるよう信号を統合をはかる。

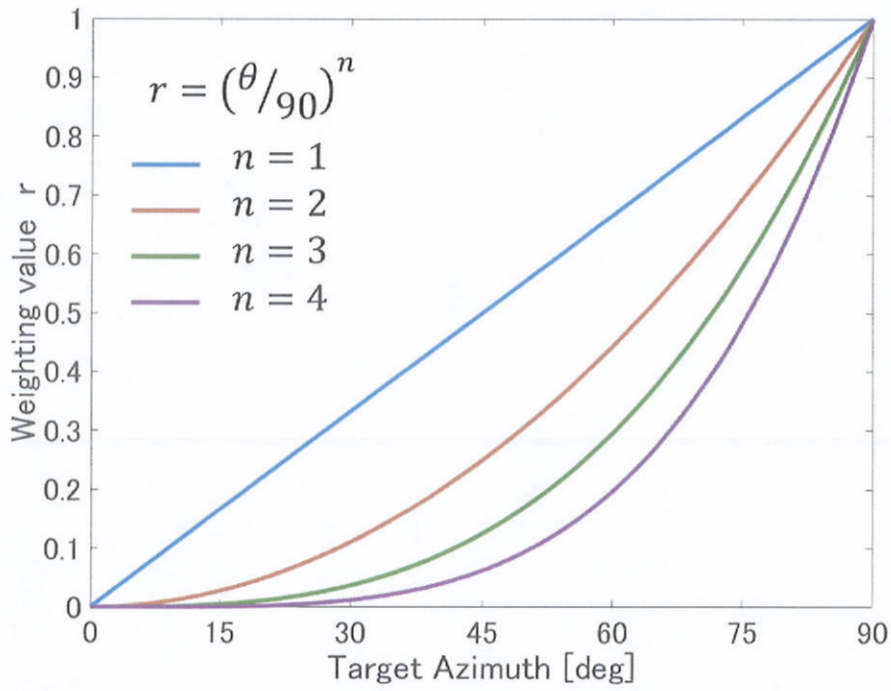
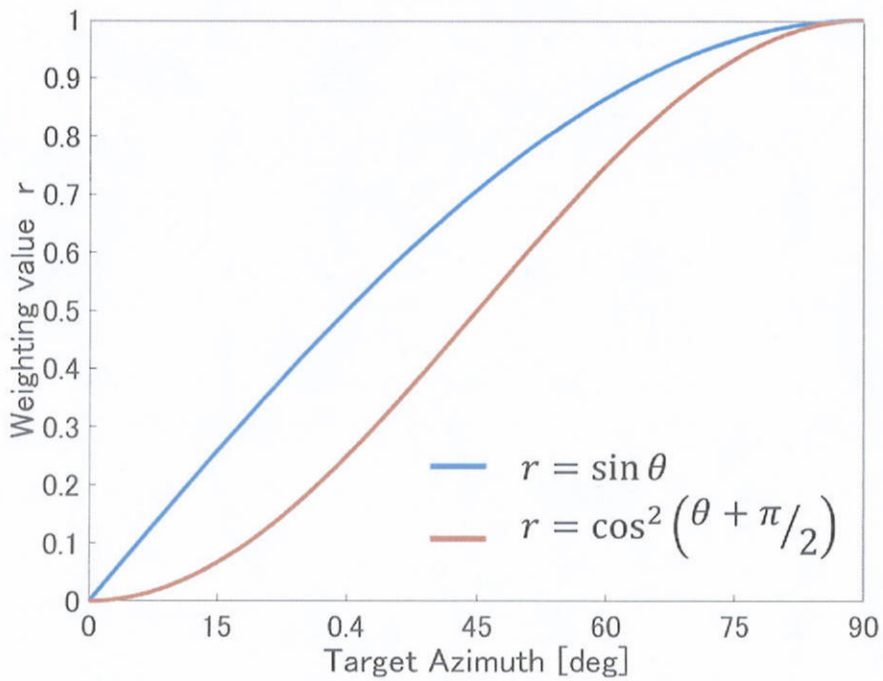
Fig.3.4: n 次関数を用いた重み関数

Fig.3.5: 三角関数を用いた重み関数

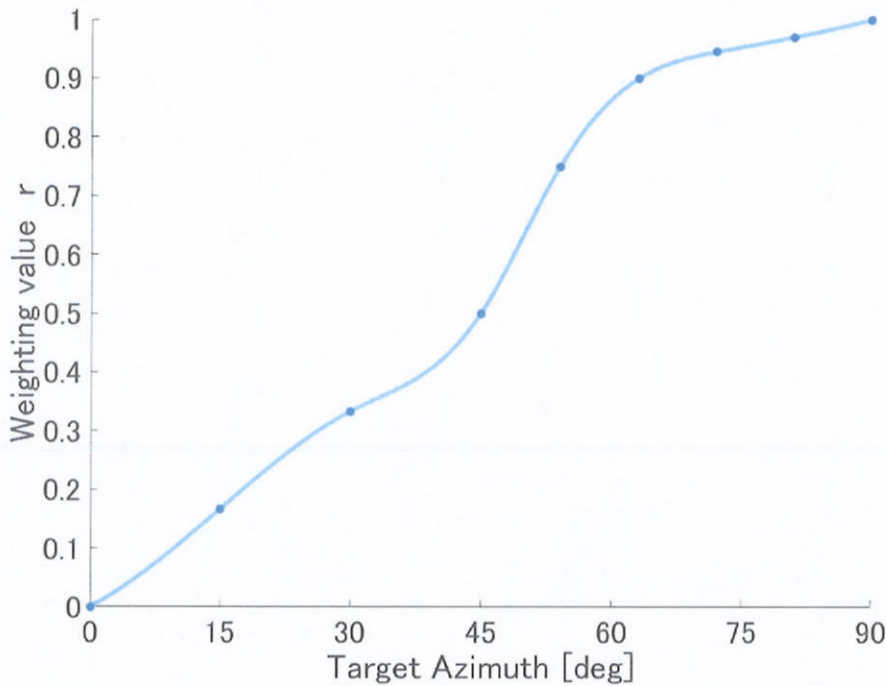


Fig.3.6: スプライン補間を用いた重み関数

$$\mathbf{SD} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} \left(20 \log_{10} \frac{|X_L(f_n)|}{|S_A(f_n)|}\right)^2} \quad [\text{dB}] \quad (3.5)$$

3.2.2 結果

まず、音源が常に視線方向にあると仮定した場合について結果を示す。Fig.3.7, Fig.3.8 に統合により生成した信号と音源が正面にある時の信号のスペクトル歪値を示す。横軸は正面を 0° とした場合の視線方向の角度を示している。Fig.3.7 は式 (3.2) による重み係数を用いた場合、Fig.3.8 はスプライン補間による重み付係数を用いた場合と式 (3.3)、式 (3.4) による三角関数の重み係数の場合、Fig.3.7, Fig.3.8 の結果において平均を計算したところ、一次関数と式 (3.4) で重み付した場合とスプライン補間で重み付した場合でスペクトルの差が小さくなった。検討した 7 種類の重み係数の中で、スプライン補間による重み係数を用いた場合が最小のスペクトル歪であった。これは、視線方向および音源方向が 90° に近づくにつれて、Mic.A の伝達特性は大きな振幅特性を持つようになる。この変化は、統合の性能劣化につながるため、Mic.A には小さな重み、Mic.B には大きな重みとなることが望ましいと考えられる。スプライン補間による重み関数が最も良い結果となった原因は、この重みの関係が他の関数よりも適

切に表現できたためであると考え。次に、視線方向と音源が $15^\circ \sim 90^\circ$ まで 15° 間隔でずれていった場合の結果を示す。また、これらの重み係数を用いた場合のスペクトル歪の平均値を Fig.3.12 に示す。結果として三角関数を用いた場合では音源と視線方向が異なる場合においても歪の平均値が他の重み関数より小さい結果となった。

一次関数、三角関数やスプライン補間を用いた重み関数を用いた場合にスペクトル歪の平均が 3 dB 以下となり、音響信号のみで音源方向が判別できる可能性が示唆された。またスペクトル歪の最大値は約 7 dB であったが、このような大きな歪となる条件であっても、映像と音響とを組み合わせることで、音源方向を正しく知覚させることが可能となると考える。スプライン補間を用いた場合、最適な値を求められれば、スペクトル歪の値を小さくすることが可能であるが、対象とする音源が視線方向から外れることがあるような場合は三角関数を用いた重み付関数を設定するとスペクトル歪の平均値は小さくなることが明らかになった。

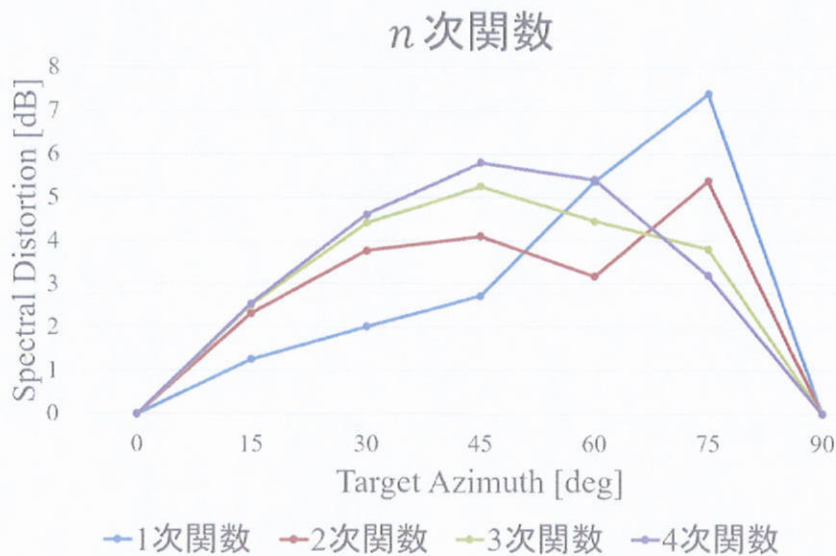


Fig.3.7: n 次関数により重み付けした結果

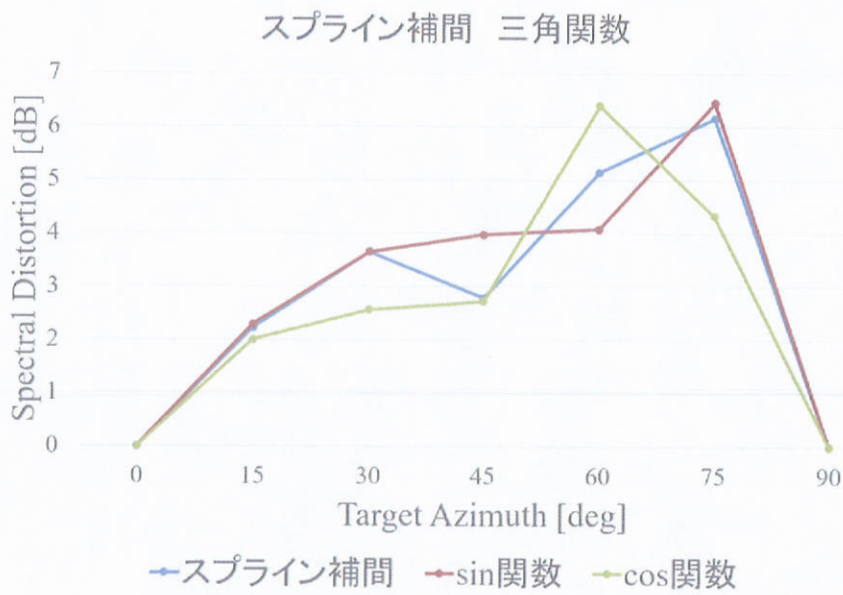


Fig.3.8: スプライン補間・三角関数により重み付けした結果

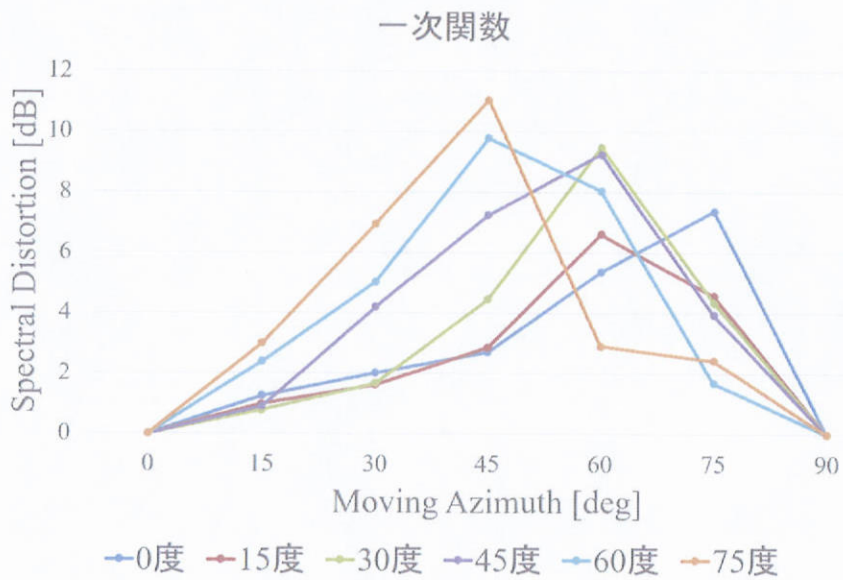


Fig.3.9: 音源と視線方向がずれた場合に一次関数で重み付けした結果

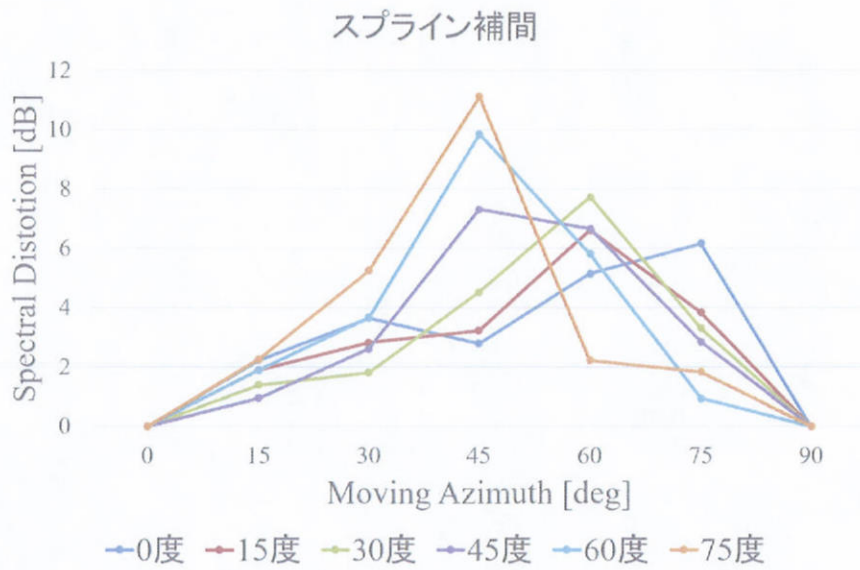


Fig.3.10: 音源と視線方向がずれた場合にスプライン補間で重み付けした結果

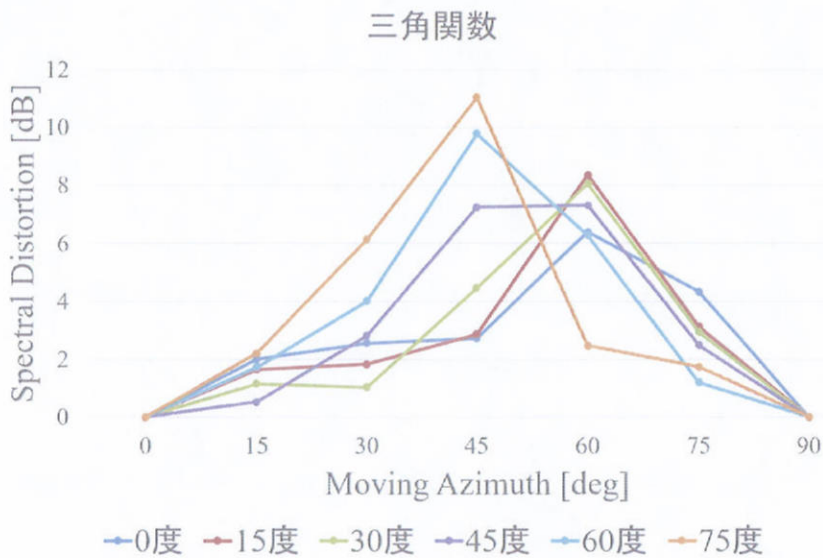


Fig.3.11: 音源と視線方向がずれた場合に三角関数で重み付けした結果

歪の平均値	
$y = \frac{\theta}{90}$	4.50 [dB]
$y = \cos^2\left(\frac{\theta + \pi}{2}\right)$	4.01 [dB]
スプライン	4.07 [dB]

Fig.3.12: 音源と視線方向がずれた場合に算出した歪の平均値

第4章

結論

複数くぼみを持つ球状の立体音収録システムを用いて、全天球映像に立体音を付与するため、複数のくぼみ内の受音点において収録した信号の重みづけ和をとることで、任意の視線方向に対応する立体音を生成する手法について検討を行った。先行研究では全天球バイノーラル録音のための複数くぼみを持つマイクロホンバッフルは6方向同時に立体音の収録可能である。全天球映像に対応させる場合さらに同時収録可能な方向を増加する必要がある。先行研究同様に方向に対応したくぼみを新たに増設する方法も考えられるが、あまりくぼみを増やすと球体のバッフルが持つ球状の音響特性を失ってしまう可能性があるため、本論文では、収録した信号の重みづけ和をとり、重み関数を変えた場合について検討した。

実験の結果、一次関数、三角関数やスプライン補間を用いた重み関数を用いた場合にはスペクトル歪の平均が3 dB以下となり、音響信号のみで音源方向が判別できる可能性が示唆された。またスペクトル歪の最大値は約7 dBであったが、このような大きな歪となる条件であっても、映像と音響とを組み合わせることで、音源方向を正しく知覚させることが可能となると考える。また三角関数を用いた場合では音源と視線方向が異なる場合においても歪の平均値が他の重み関数より小さい結果となった。以上より、本論文では、三角関数を重み関数に用いる立体音響信号の統合が有効であることが示された。

参考文献

- [1] 寺本渉, 吉田和博, 坂本修一, 浅井暢子, 岩谷幸雄, 日高聡太, 鈴木陽一, “バーチャル・リアリティの”リアリティ”, 映像情報メディア学会技術報告, 36(12), pp.37-44, 2012.
- [2] 坂本修一, ”SENZI: 球状マイクロホンアレイを用いた3次元音空間収録再生,” 日本音響学会誌, vol.70, no.7, pp.379-384, 2014.
- [3] 坂本修一, 小玉純一, 本郷哲, 岡本拓磨, 岩谷幸雄, 鈴木陽一 “252ch 等密度球状アレイを用いた3次元音場情報取得システム”, 電子情報通信学会技術研究報告. EA, 応用音響 110(171), pp.19-23, 2010-08-02 .
- [4] 中桐大志, “全天球バイノーラル音響収録のためのくぼみ付き球状マイクロホンバッフル,” 三重大学工学研究科修士学位論文, 2016.
- [5] 今村健人, 西野隆典, 成瀬央, “全方位立体集音のためのマイクロホンバッフルの設計と評価”, 日本音響学会講演論文集 (春季), pp.631-632, 2017.
- [6] 三重大学パターン情報理研究室, 西野グループ, <http://www.pa.info.mie-u.ac.jp/nishinogroup/hist2014.shtml>(2018.2.6 access)
- [7] 松井浩二, 工藤彰洋, 穂刈治英, 島田正治, “近接音場における単一音源を用いた方向定位知覚に関する検討,” 電子情報通信学会技術研究報告 (応用音響), 106(607), pp.23-28, 2007.
- [8] 渡邊一樹, 澤谷郁子, 大出訓史, 小野一穂, 安藤彰男, 小澤賢司, “両耳情報に基づく音の包み込まれ感の推定モデルに関する一考察,” 電子情報通信学会技術研究報告 (応用音響), 113(177), pp.63-68, 2013.
- [9] 森川大輔, “両耳間時間差が音像の分離知覚に与える影響,” 研究報告音楽情報科学 (MUS) , 2014-MUS-103(46), pp.1-3, 2014.

謝辞

本研究を進めるにあたり、終始適切なご指導とご助言をいただきました、名城大学都市情報学部 都市情報学科教授 西野 隆典 先生に深く感謝します。本研究に対し、貴重な助言を頂きました、本学工学部 教授 成瀬 央 先生に厚くお礼申し上げます。また、学生生活を共にした、パターン情報処理研究室の皆様感謝します。