

横断面に配置した 2 スピーカによる 3 次元音像制御*

飯田 一博*¹ 石井 孝*^{2,†} 石井 要次*² 池見 隆史*²

【要旨】 2 個のスピーカにより 3 次元音像制御をする方法の一つとして、トランスオーラルシステムが提案されている。しかし、受聴時の位置や姿勢のわずかなずれによって再生系伝達関数が増減するため、受聴者の頭部を固定する必要があった。本研究では頭部を固定せずとも精度高く音像制御できるスピーカ配置を検討した。横断面内 15 方向のスピーカ配置に対して、音像定位実験と再生系伝達関数の分析を行った結果、側方から上方にスピーカを配置した場合に音像制御精度が高くなることを示した。特に、天頂から $\pm 70^\circ$ に配置した場合は、水平面定位においても上半球正中面定位においても、実音源と有意差のない定位精度となることを示した。

キーワード トランスオーラルシステム, 音像制御, 頭部伝達関数, 横断面, スピーカ
Transaural system, Sound image control, Head-related transfer function, The transverse plane, Loudspeaker

1. はじめに

受聴者の外耳道入口で頭部伝達関数 (HRTF: Head-Related Transfer Function) を再現し、任意の方向に音像を制御する方法の一つとして、トランスオーラルシステムが提案されている [1–4]。このシステムは、2 個のスピーカ (図-1) から式 (1) に示す刺激を放射することにより、原理的にはスピーカから外耳道入口までの伝達関数 (以降、再生系伝達関数と呼ぶ) をキャンセルすると同時に、目標とする方向の HRTF を外耳道入口で再現することが可能である。

$$\left. \begin{aligned} X_L(\omega) &= S(\omega) \times \frac{H_L(\omega) \times C_{R,r}(\omega) - H_R(\omega) \times C_{L,l}(\omega)}{C_{L,l}(\omega) \times C_{R,r}(\omega) - C_{L,r}(\omega) \times C_{R,l}(\omega)} \\ X_R(\omega) &= S(\omega) \times \frac{H_R(\omega) \times C_{L,l}(\omega) - H_L(\omega) \times C_{L,r}(\omega)}{C_{L,l}(\omega) \times C_{R,r}(\omega) - C_{L,r}(\omega) \times C_{R,l}(\omega)} \end{aligned} \right\} \quad (1)$$

ここで、 $X_i(\omega)$ はスピーカから放射される刺激、 $S(\omega)$ は音源信号、 $H_j(\omega)$ は目標方向に対する耳 j の HRTF、 $C_{i,j}(\omega)$ はスピーカ i から耳 j への再生系伝達関数であり、 i は左右のスピーカ (L, R)、 j は左右の耳 (l, r) を表す。

2 個のスピーカの配置については、水平面内の前方 $\pm 30^\circ$ が標準的であるが、 $\pm 5^\circ$ に近接配置する Stereo Dipole [5] も提案されている。しかし、受聴時の位置のわずかなずれや頭部の動きにより、再生系伝達関数は必ずしも一定ではないため、実際には 3 次元音像制御は達成できているとは言えない。

Morimoto and Ando [6] は、被験者の頭部を固定し、再生系伝達関数に顕著なスペクトルノッチがみられないことから 2 個のスピーカを横断面の天頂から $\pm 30^\circ$ に配置したトランスオーラルシステムで音像定位実験を行った。その結果、受聴者本人の HRTF を使えば、実音源と同等の精度で音像を制御できることを示した。Takeuchi *et al.* [7] は Stereo Dipole を用いて、頭部を固定せずに水平面内の音像制御実験を行った。その結果、比較的高い精度で制御できた被験者のグループと、制御精度が低く水平面内の後方には定位しない被験者のグループに分かれたと報告している。

頭部の固定は実用化の観点からは許容し難い拘束条件であり、定められた位置に設置された椅子に座れば頭部を固定せずとも任意の 3 次元方向に音像を制御できることが望まれる。制御領域を拡張するために、3 個以上のスピーカを用いる方法も検討されているが [8, 9]、本研究では、最小構成である 2 個のスピーカを用いたトランスオーラルシステムを取り上げる。2 個のスピーカを横断面上の様々な方向に配置して、頭部を固定せずに、音像定位実験及び再生系伝達関数の分析を行って、3 次元音像制御に適したスピーカ配置を検討した。

* Three-dimensional sound image control by two loudspeakers located in the transverse plane, by Kazuhiro Iida, Takashi Ishii, Yohji Ishii and Takashi Ikemi.

¹ 千葉工業大学工学部

² 千葉工業大学大学院工学研究科

[†] 現在、アイホン(株)

(問合先: 飯田一博 〒275-0016 習志野市津田沼 2-17-1 千葉工業大学工学部)

(2011 年 5 月 30 日受付, 2012 年 2 月 20 日採録決定)

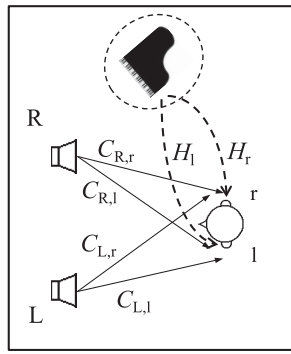


図-1 トランスオーラルシステムの伝達関数

2. 音像定位実験

2.1 スピーカ配置

本研究で用いたスピーカ配置は、図-2 に示すように横断面の天頂を 0° として、 $\pm 20^\circ$ から $\pm 160^\circ$ まで 10° 間隔の 15 種類 (T20~T160)、及び比較のために用意した従来のトランスオーラルシステムの水平面の前方 $\pm 30^\circ$ (H30) と [5] を参考に本実験システムにおいて最もスピーカを近接配置できる開き角とした $\pm 6^\circ$ (H6) の計 17 種類である。

2.2 再生系伝達関数と頭部伝達関数の測定

音像定位実験に先立ち、再生系伝達関数及び頭部伝達関数を無響室で測定した。測定システムは、ノートパソコン (Hewlett-Packard Compaq 6710b)、D/A コンバータ (RME Fireface 400)、アンプ (YAMAHA HC1500)、スピーカ (FOSTEX FE83E)、スピーカキャビネット (ダイトーボックス SV-70)、イヤマイクホン [10]、A/D コンバータ (Roland M-10MX) で構成した。被験者は 20 代男性 2 名 (A, B) である。音源信号には swept-sine 信号を用いた。音源信号のサンプリング周波数は 48,000 Hz、サンプル長は 2^{18} である。被験者の頭部は固定せず、背筋を伸ばして正面を向いて静止して座るよう指示した。

再生系伝達関数は、図-2 の 17 方向のスピーカから放射した音源信号を被験者の外耳道入口に装着したイヤマイクホンで収録することにより測定した。

一方、頭部伝達関数は、水平面の 12 方向 ($0 \sim 330^\circ$, 30° 間隔)、及び上半球正中面の 7 方向 ($0 \sim 180^\circ$, 30° 間隔) について、式 (2) に示すように、被験者の外耳道入口に装着したイヤマイクホンで収録した伝達関数 G_j を受聴者がいない状態で受聴者の頭部中心に相当する位置までの伝達関数 F で除することにより求めた。添え字 j は左右の耳 (l, r) を示す。

$$HRTF_j(\omega) = G_j(\omega)/F(\omega) \quad (2)$$

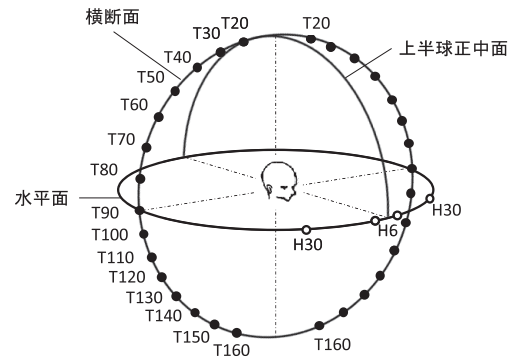


図-2 再生系伝達関数の分析及び音像定位実験で用いたスピーカ配置
●：横断面上のスピーカ配置，○：従来の水平面上のスピーカ配置

2.3 実験方法

17 種類のスピーカ配置について音像定位実験を行った。実験は無響室で行った。実験システムは、ノートパソコン (Dell xps M1330)、D/A コンバータ (RME Fireface 400)、アンプ (marantz PM4001)、スピーカ (FOSTEX FE83E)、スピーカキャビネット (ダイトーボックス SV-70) で構成した。被験者は再生系伝達関数を測定した 20 代男性 2 名 (A, B) である。音源信号には広帯域ホワイトノイズ ($200 \sim 17,000$ Hz) を用いた。音源信号のサンプリング周波数は 48,000 Hz である。刺激は式 (1) によって作成した。HRTF は被験者本人の水平面の 12 方向 ($0 \sim 330^\circ$, 30° 間隔)、及び上半球正中面の 7 方向 ($0 \sim 180^\circ$, 30° 間隔) である。刺激の音圧レベルは正面方向の HRTF を用いた場合に被験者の左外耳道入口において A 特性音圧レベルが 60 dB となるように調整した。他の方向の場合の音圧レベルは、実音源の場合と同様に、HRTF の方向依存性で決まる値とした。刺激の継続時間は 1 秒、回答時間は 6 秒である。

実験は、スピーカ配置ごとに行った。また、目標方向を水平面においた場合と正中面においた場合で分けて行った。各刺激をランダムな順に 10 回呈示した。被験者は、知覚した音像の方位角と仰角をマッピング法で回答した。被験者の頭部は固定せず、背筋を伸ばして正面を向いて静止して座るよう指示した。実験時は、被験者にスピーカを視認させないため無響室を消灯し、小型ライトで手元の回答用紙のみ照らした。

2.4 実験結果

2.4.1 目標方向を水平面においた場合の回答分布

被験者の回答を 1° の精度で読み取り、 5° で丸めてプロットした (図-3~10)。円の直径は回答頻度に比例している。

まず、方位角の回答結果について述べる。被験者 A (図-3) では T30, T40, T60~T110, T130, 及び T140

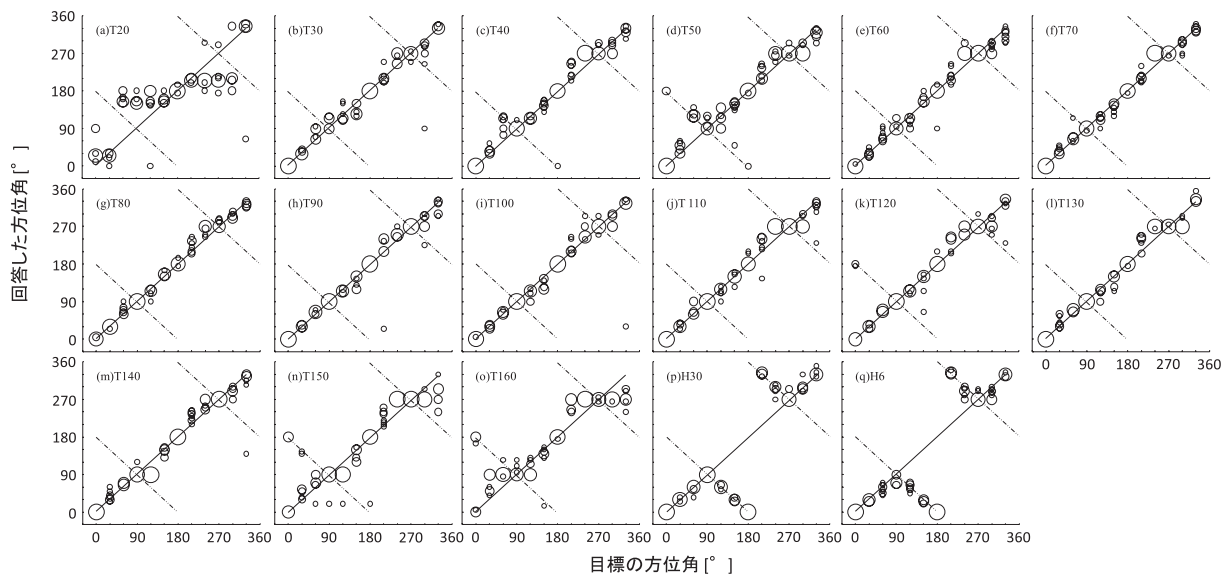


図-3 目標方向を水平面においた場合の方位角の回答 (被験者 A)

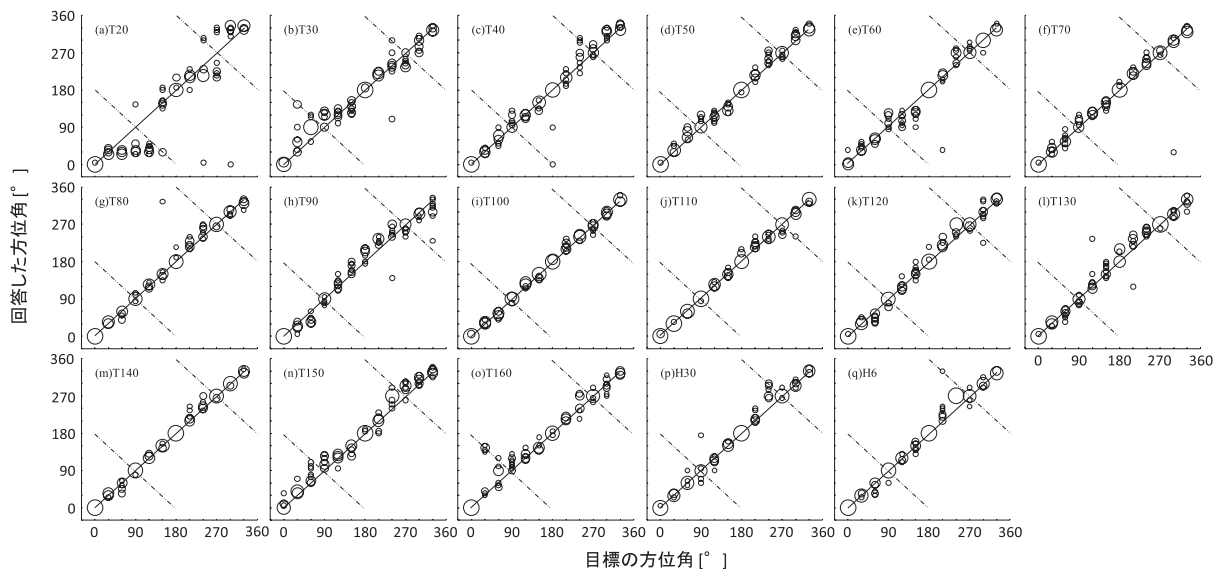


図-4 目標方向を水平面においた場合の方位角の回答 (被験者 B)

のスピーカ配置では、概ね目標方向に知覚していた。T20では前方もしくは後方にのみ回答する傾向がある。T50では、正面と後方の刺激に対して前後誤判定が生じた。T120, T150, 及び T160では、正面方向を後方に知覚する傾向がある。T150と T160では、更に目標方向 60°と 120°を 90°に、目標方向 240°と 300°を 270°に知覚し、回答が側方に偏る傾向がある。H30と H6では目標方向が後方の刺激をすべて前方に誤判定していた。

被験者 B (図-4) では T40~T150, H30, 及び H6のスピーカ配置で概ね目標方向に知覚していた。T20では側方に知覚せず、回答が前方もしくは後方に偏る傾向がある。T30と T160では、目標方向 30°を側方もしくは後方に知覚していた。

次に、仰角について述べる。被験者 A (図-5) では、

T30, T60~T110, H30, 及び H6ではすべての刺激を 0° (水平面上) に知覚した。T20では概ね 0°に知覚していたが、わずかながら上方に知覚していた場合があった。T40と T50では、水平面より上昇して知覚される場合があった。T120~T140では水平面よりも下方に知覚されることが多いが、上昇して知覚される場合もあった。水平面より下方への知覚については、スピーカの方向と一致しているが、方位角の回答 (図-3) をみると、スピーカの方向 (90°) には知覚していない。T150と T160では水平面よりも下方に知覚される場合が非常に多い。方位角の回答 (図-3) と併せると、T150と T160ではスピーカの方向に音像を知覚していることが多いと考えられる。

被験者 B (図-6) では、T30と T40を除いたスピーカ配置のすべての刺激を 0° (水平面上) に知覚した。

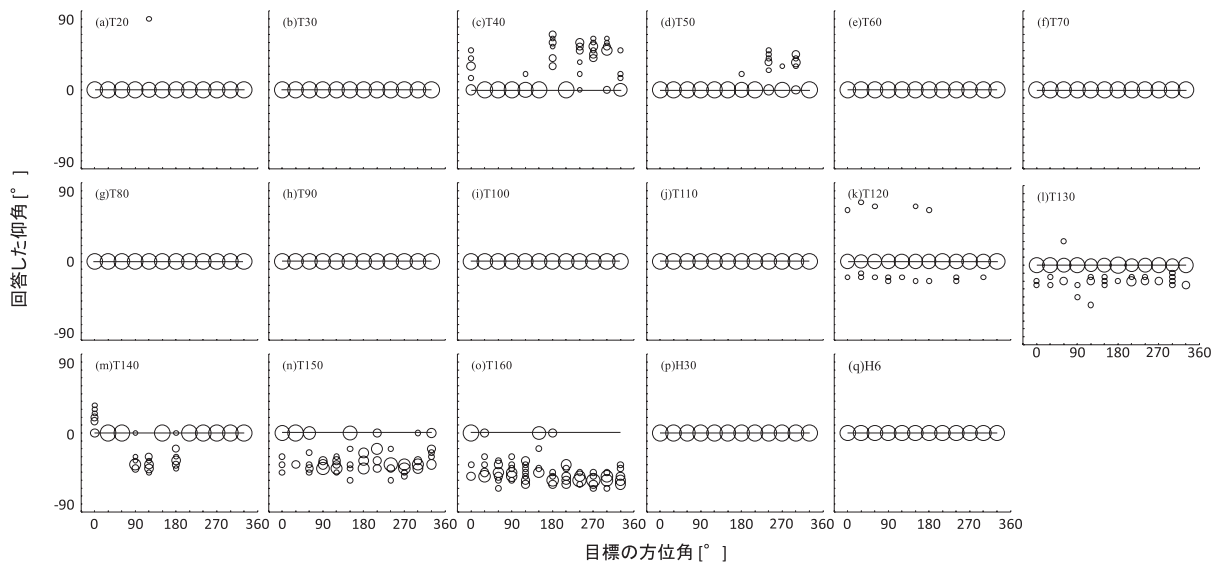


図-5 目標方向を水平面においた場合の仰角の回答 (被験者 A)

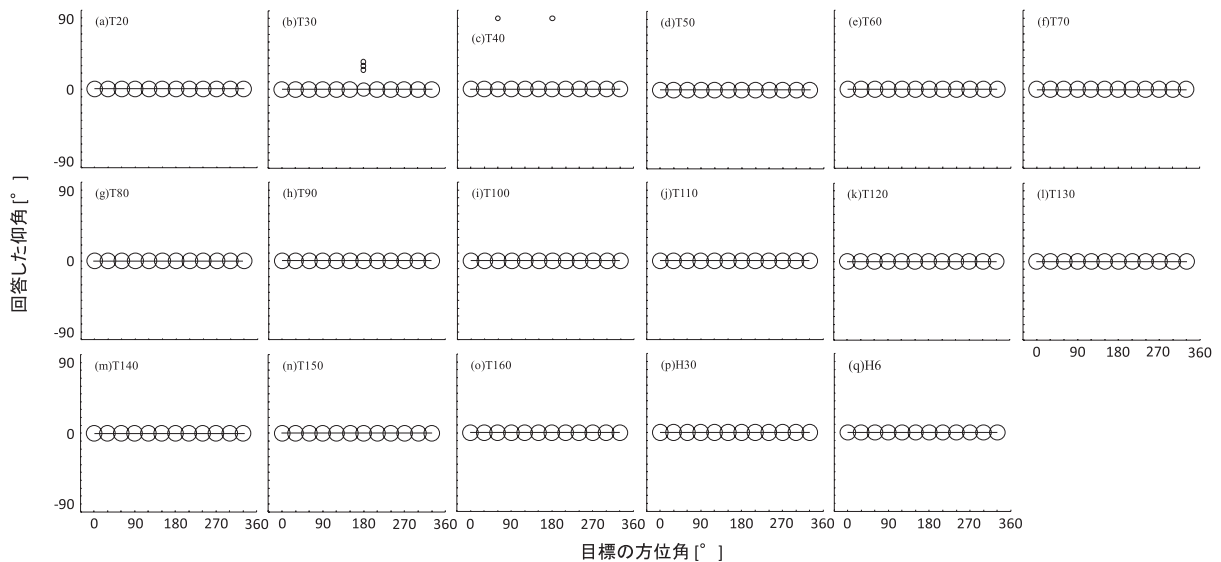


図-6 目標方向を水平面においた場合の仰角の回答 (被験者 B)

T30 と T40 では概ね 0°に知覚していたが、少数ではあるが水平面より上昇した音像を知覚する場合があった。

以上より、被験者 2 名に共通して概ね目標の方位角及び仰角に定位したスピーカ配置は T60～T110 とと言える。

2.4.2 目標方向を正中面においた場合の回答分布

目標方向を正中面においた場合の方位角の回答結果を図-7, 8 に示す。被験者 A では、H30 と H6 を除いたスピーカ配置では、概ね 0° (正中面上) に知覚していた。H30 と H6 では、正中面から左右にずれて知覚した場合があった。一方、被験者 B ではすべてのスピーカ配置のすべての刺激を 0° (正中面上) に知覚していた。

次に、仰角の回答結果であるが、上で述べたように、被験者 A の H30 と H6 の一部の刺激を除き、いずれの

刺激に対しても正中面内に知覚していたので、ここでは、球座標系において各方位角に対応して ±90°をとる仰角ではなく、図-11 に示すように矢状面座標系 [11-13] において各矢状面内で 0～360°の範囲をとる上昇角を用いて表した (図-9, 10)。上昇角では 0°が前、90°が上、180°が後ろとなり、正中面内での定位方向が一目瞭然となる。なお、上昇角は被験者が回答した方位角 (φ) と仰角 (θ) から式 (3) により算出した。

$$\beta = \sin^{-1} \frac{\sin \theta}{(\sin^2 \theta + \cos^2 \phi \cos^2 \theta)^{\frac{1}{2}}} \quad (3)$$

被験者 A (図-9) では、T60～T80 では概ね目標方向に知覚していた。T20, T150, H30, 及び H6 ではすべての刺激を前方に知覚していた。T30 と T40 では回答がややばらついた。T50 では前方と後方の刺激を前後誤判定する場合があった。T90～T130 では後

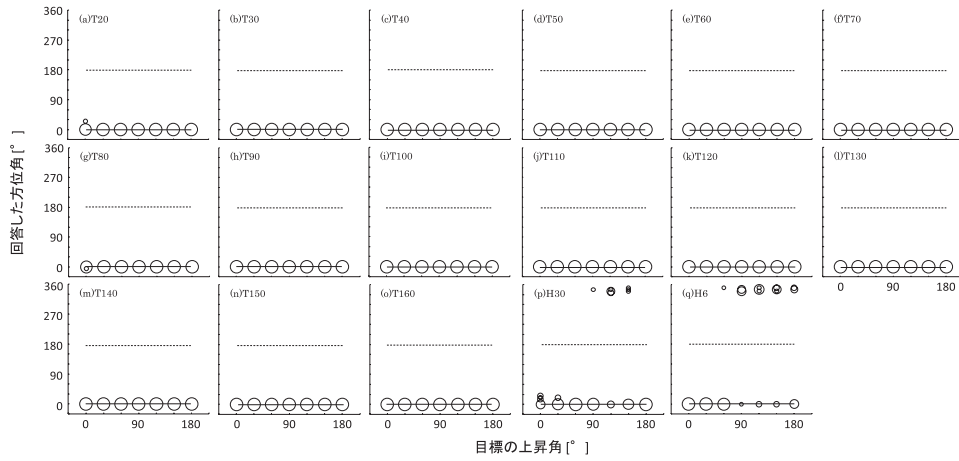


図-7 目標方向を正中面においた場合の方位角の回答 (被験者 A)

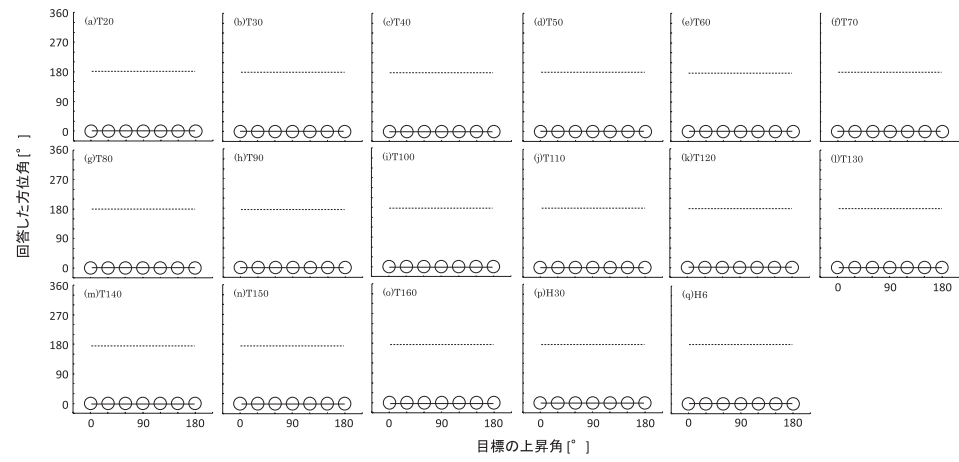


図-8 目標方向を正中面においた場合の方位角の回答 (被験者 B)

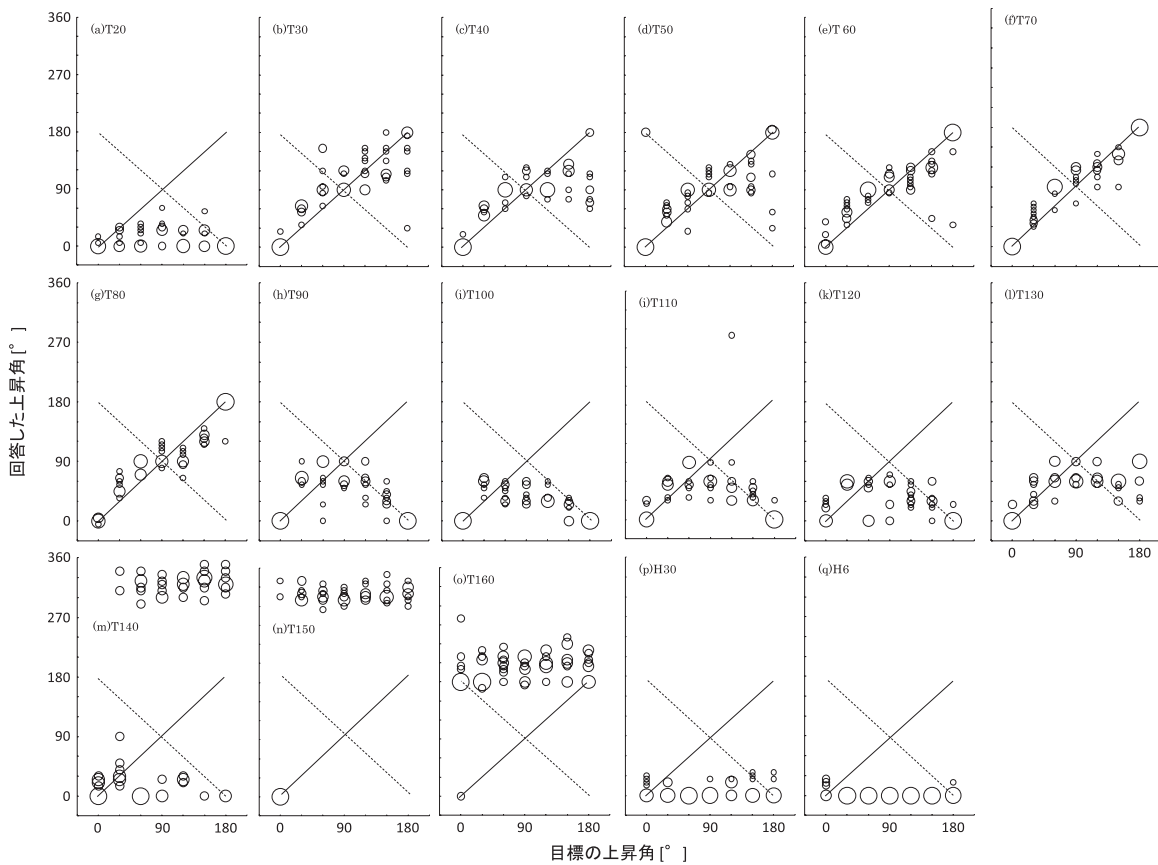


図-9 目標方向を正中面においた場合の上昇角の回答 (被験者 A)

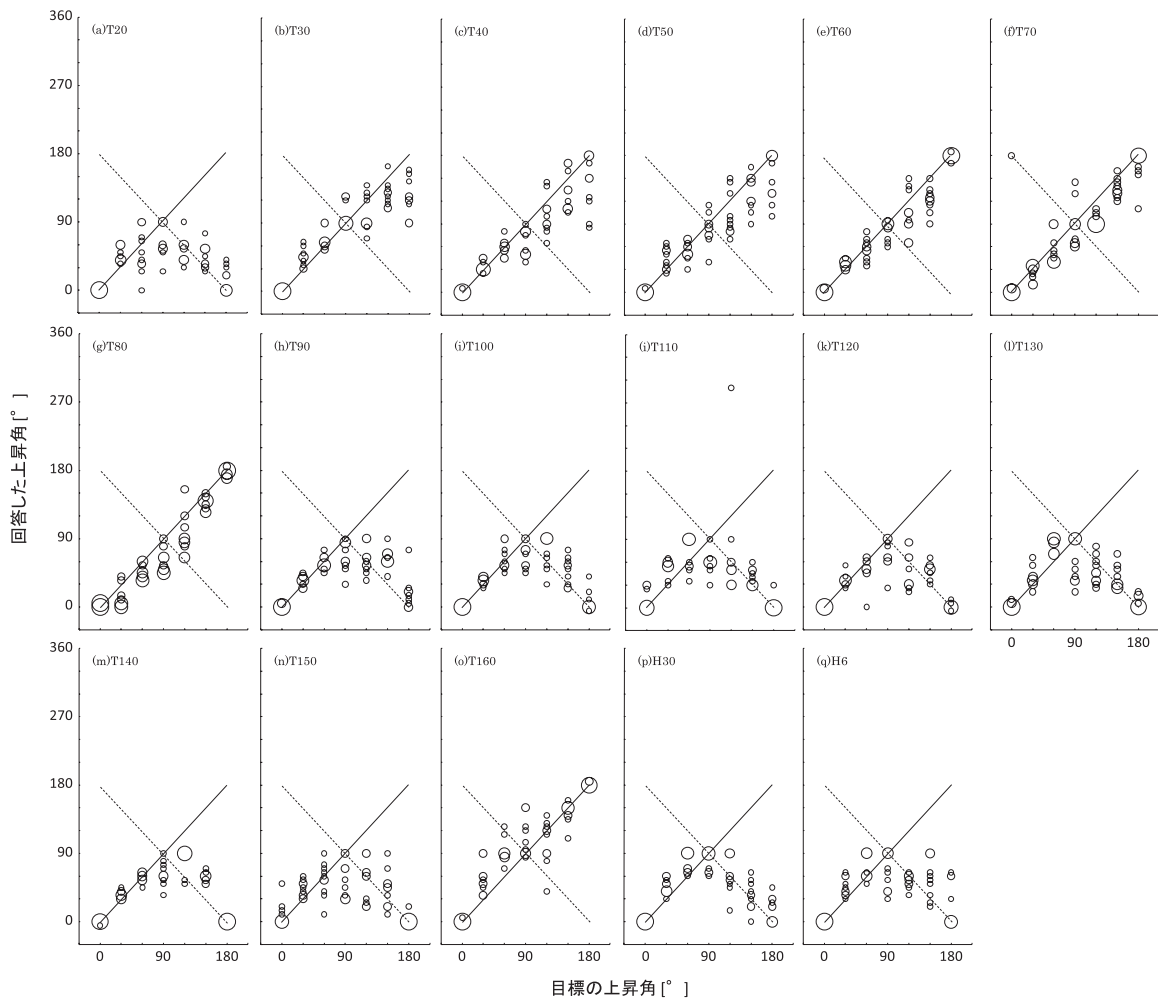


図-10 目標方向を正中面においた場合の上昇角の回答 (被験者 B)

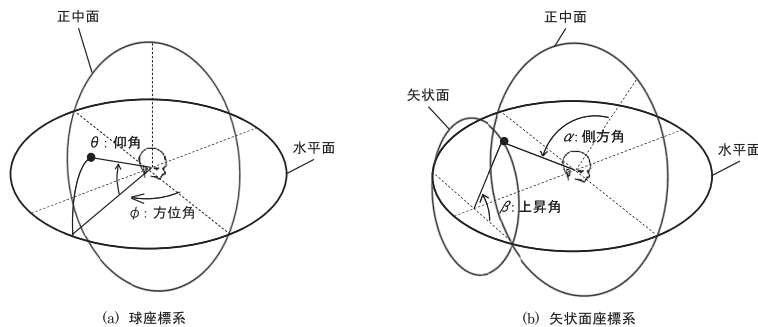


図-11 頭を中心とする座標系 (a) 球座標系, (b) 矢状面座標系

方の刺激を前後誤判定していた。T140 と T150 ではほとんどの刺激を前方もしくは前斜め下方向に知覚していた。T160 ではほとんどの刺激を後ろ斜め下方向に知覚していた。

被験者 B (図-10) では、T60 と T80 で概ね目標方向に知覚していた。T70 では正面方向の刺激に対して少数の前後誤判定がみられたが、その他の刺激に対しては概ね目標方向に知覚していた。T20, T90~T150, H30, 及び H6 では後方の刺激をすべて前後誤判定していた。T30~T50, 及び T160 では回答がややばら

ついた。

以上より、被験者 2 名に共通して概ね目標の方位角及び上昇角に定位したスピーカ配置は T60~T80 と言える。

2.4.3 平均定位誤差

被験者の回答を定量的に分析するため、式 (4) により各スピーカ配置における平均定位誤差 E を求めた。

$$E = \frac{1}{J} \times \frac{1}{K} \times \sum_{j=1}^J \sum_{k=1}^K |S_{j,k} - R_{j,k}| \quad (4)$$

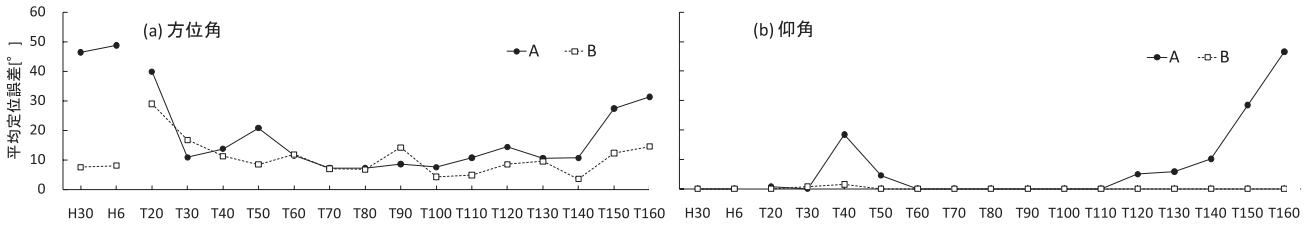


図-12 目標方向を水平面においた場合の平均定位誤差
(a) 方位角, (b) 仰角

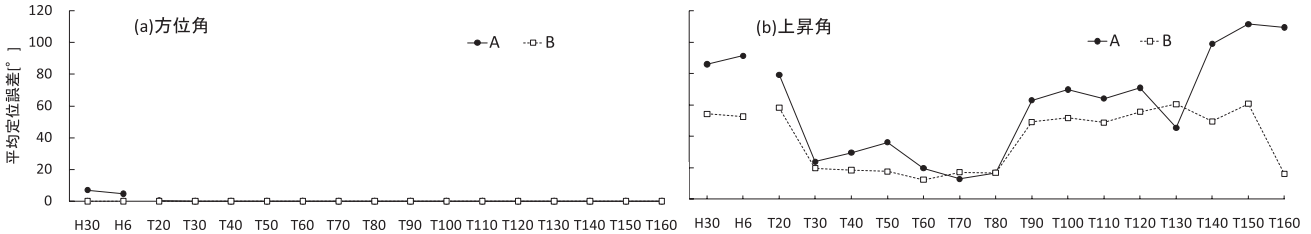


図-13 目標方向を正中面においた場合の平均定位誤差
(a) 方位角, (b) 上昇角

ここで、 $S_{j,k}$ は目標方向、 $R_{j,k}$ は知覚した音像の方向、 J は目標方向の数、 K は呈示回数である。なお、 $S_{j,k}$ 及び $R_{j,k}$ は、目標方向を水平面においた場合は方位角と仰角で、正中面の場合は方位角と上昇角でそれぞれ表した。つまり、平均定位誤差は、前後誤判定も含めた値として算出した。

まず、目標方向を水平面においた場合について述べる。方位角の平均定位誤差 (図-12(a)) は、被験者 A では、従来のスピーカ配置の H30 で 46.4°、H6 で 48.7°であったのに対し、横断面配置 (T20~T160) においてはいずれもそれより小さかった。被験者 B では、H30 で 7.5°、H6 で 8.0°であったが、T70、T80、T100、T110、及び T140 ではそれより小さな値が得られた。

一方、仰角 (図-12(b)) については、どちらの被験者も従来の水平面配置での平均定位誤差は 0.0°であった。横断面配置では、被験者 A では T30 及び T60~T110 で 0.0°であった。T40 と T50 で、平均定位誤差が大きいの、水平面より上昇した音像を知覚したためである (図-5)。同様に T120~T160 で平均定位誤差が大きいの、水平面より下降した音像を知覚したためである。被験者 B では T30 と T40 を除くすべての横断面配置で 0.0°であった。

以上より、2名の被験者に共通して、方位角も仰角も従来のスピーカ配置と同等以下の平均定位誤差となったのは、T70、T80、T100、及び T110 であった。

次に、目標方向を正中面においた場合について述べる。方位角の平均定位誤差 (図-13(a)) は、被験者 A では、従来の配置の H30 で 7.0°、H6 で 4.7°であったのに対し、横断面配置においてはいずれもそれより小さかった。また、T30~T160 では 0.0°であった。被験

者 B では、従来配置及び横断面配置のすべてにおいて 0.0°であった。

上昇角の平均定位誤差 (図-13(b)) については、被験者 A では、従来のスピーカ配置の H30 で 86.1°、H6 で 91.4°であったのに対し、T20~T130 ではそれより小さくなり、T60~T80 では 20°以下となった。被験者 B では、H30 で 54.3°、H6 で 52.6°であったのに対し、T30~T110、T140、及び T160 でそれより小さくなり、T30~T80 では 20°以下となった。

以上より、2名の被験者に共通して、方位角も上昇角も従来のスピーカ配置と同等以下となったのは、T30~T110 であった。

2.4.4 実音源とトランスオーラルシステムの平均定位誤差の差の検定

更に、実音源とトランスオーラルシステムの平均定位誤差に統計的に有意な差があるか否かを検証した。実音源を用いた音像定位実験は、今回のトランスオーラルシステムを用いた音像定位実験以前に実施した。無響室内に水平面 12 方向 (0~330°, 30°間隔)、及び上半球正中面 7 方向 (0~180°, 30°間隔) にスピーカを設置し、水平面の場合と正中面の場合に分けて行った。音源信号は 20~20,000 Hz の広帯域ホワイトノイズを用いた。その他の実験条件はトランスオーラルシステムの場合と同様である。ただし、水平面の場合は方位角のみ、正中面の場合は上昇角のみ回答した。回答結果を図-14、15 に示す。平均定位誤差は、水平面においては被験者 A が 7.2°、被験者 B が 5.1°、正中面においては被験者 A が 8.8°、被験者 B が 12.2°であった。実音源とトランスオーラルシステムの平均定位誤差について t 検定を行った結果を表-1 に示す。

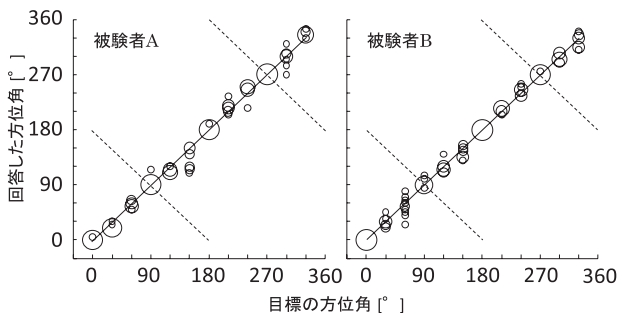


図-14 水平面内の実音源に対する方位角の回答

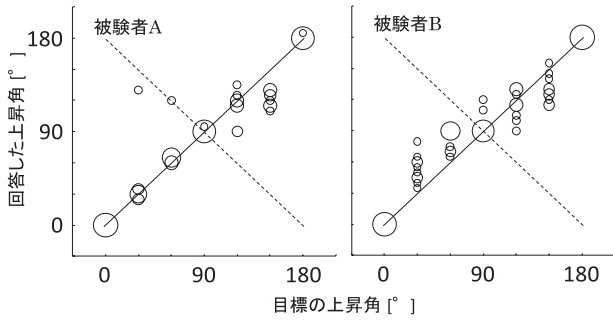


図-15 正中面内の実音源に対する上昇角の回答

表-1 実音源とトランスオーラルシステムの平均定位誤差の差の検定結果

スピーカ配置	水平面		正中面	
	被験者		被験者	
	A	B	A	B
T20	**	**	**	**
T30	*	**	**	*
T40	**	**	**	*
T50	**	**	**	*
T60	**	**	**	
T70				
T80			**	
T90		**	**	**
T100			**	**
T110	*		**	**
T120	*	**	**	**
T130	**	**	**	**
T140	*		**	**
T150	**	**	**	**
T160	**	**	**	
H30	**		**	**
H6	**	*	**	**

*: $p < 0.05$ **: $p < 0.01$

目標方向を水平面においた場合、方位角の平均定位誤差が実音源のそれと有意な差がないと見なせるスピーカ配置は、被験者 A では T70~T100, 被験者 B では T70, T80, T100, T110, T140, 及び H30 であった。一方、目標方向を正中面においた場合、上昇角の平均定位誤差が実音源のそれと有意な差がないと見なせる

スピーカ配置は、被験者 A では T70, 被験者 B では T60~T80, 及び T160 であった。

以上より、2名の被験者に共通して、水平面の方位角においても正中面の上昇角においても平均定位誤差が実音源のそれと有意な差がないと見なせるスピーカ配置は T70 だけであった。

2.4.5 前後誤判定率

最後に、前後誤判定率 FBER (Front-Back Error Ratio) を求めた。前後誤判定率は耳軸を基準として目標方向が前方もしくは後方の刺激に対して、後方もしくは前方へ知覚した率であり、式 (5) より求めた。

$$FBER = \frac{M}{L \times K} \times 100[\%] \quad (5)$$

ここで、 L は目標方向の数 (横断面上を除く)、 K は呈示回数、 M は前後誤判定した数である。

算出結果を図-16 に示す。目標方向を水平面においた場合、どちらの被験者でも実音源での前後誤判定率は 0.0% であった。トランスオーラルシステムで前後誤判定率が 0.0% となったのは、被験者 A の T80 と T130, 被験者 B の T70, T100, T130, 及び T140 であった。また、従来のスピーカ配置より前後誤判定率が低いスピーカ配置は、被験者 A ではすべての横断面配置、被験者 B では T70~T100, 及び T120~T140 であった。2名の被験者に共通して実音源と同等以下となるスピーカ配置は T130 であった。また、2名の被験者に共通して従来のスピーカ配置よりも前後誤判定率が低くなるのは、T70~T100, 及び T120~T140 であった。

目標方向を正中面においた場合は、実音源での前後誤判定率は、被験者 A で 1.6%, 被験者 B で 0.0% であった。これと同等以下の値が得られたスピーカ配置は、被験者 A では T70, 被験者 B では T90 と T140 であった。一方、従来のスピーカ配置より前後誤判定率が低いスピーカ配置は、被験者 A ではすべての横断面配置、被験者 B では T20 を除くすべての横断面配置であった。2名の被験者に共通して実音源と同等以下となるスピーカ配置は存在しなかった。しかし、T20 を除くすべての横断面配置で2名の被験者に共通して従来のスピーカ配置よりも前後誤判定率が低くなった。

以上より、2名の被験者に共通して、水平面においても正中面においても前後誤判定率が実音源のそれと同等以下になるスピーカ配置はなかったが、T70~T100, 及び T120~T140 では従来のスピーカ配置と比較して低くなった。

2.5 総合的評価

ここまで検討してきた音像定位実験の結果をまとめて表-2 に示す。各評価項目において、基準を満たすス

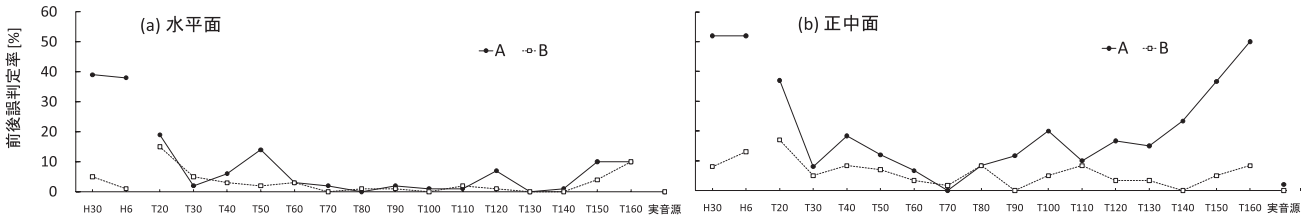


図-16 前後誤判定率
(a) 目標方向を水平面においた場合, (b) 目標方向を正中面においた場合

表-2 各評価項目を満たす被験者
●被験者 A, □被験者 B

評価項目	スピーカ配置																
	T20	T30	T40	T50	T60	T70	T80	T90	T100	T110	T120	T130	T140	T150	T160	H30	H6
目標方向を水平面においた場合																	
回答の分布が目標方向と概ね一致(方位角)		●	●□	□	●□	●□	●□	●□	●□	●□	□	●□	●□	□		□	□
回答の分布が目標方向と概ね一致(仰角)	●□	●□	□	□	●□	●□	●□	●□	●□	●□	□	□	□	□	□	●□	●□
平均定位誤差が従来配置と同等以下(方位角)	●	●	●	●	●	●□	●□	●	●□	●□	●	●	●□	●	●	—	—
平均定位誤差が従来配置と同等以下(仰角)	□	●		□	●□	●□	●□	●□	●□	●□	□	□	□	□	□	—	—
t検定で実音源と有意差なし(方位角)						●□	●□	●	●□	□			□			□	
前後誤判定が実音源と同等以下						□	●		□		●□	□					
前後誤判定が従来配置と同等以下	●	●	●	●	●	●□	●□	●□	●□	●	●□	●□	●□	●	●	—	—
目標方向を正中面においた場合																	
回答の分布が目標方向と概ね一致(方位角)	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	□	□
回答の分布が目標方向と概ね一致(上昇角)					●□	●□	●□										
平均定位誤差が従来配置と同等以下(方位角)	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	—	—
平均定位誤差が従来配置と同等以下(上昇角)	●	●□	●□	●□	●□	●□	●□	●□	●□	●□	●	●	□	□	□	—	—
t検定で実音源と有意差なし(上昇角)					□	●□	□									□	
前後誤判定が実音源と同等以下						●		□					□				
前後誤判定が従来配置と同等以下	●	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	●□	—	—

スピーカ配置と被験者の組み合わせを記号で示している。
 目標方向を水平面においた場合、2名の被験者に共通して多くの基準を満たすスピーカ配置は、T70, T80, 及び T100 であることが分かる。これら三つのスピーカ配置では、方位角、仰角とも回答の分布が目標方向と概ね一致している。また、平均定位誤差は、従来配置より小さく、実音源の方位角とも有意な差がない。前後誤判定率も、従来配置より低く、実音源との比較では、それぞれ2名中1名が同等以下となった。
 一方、目標方向を正中面においた場合は、2名の被験者に共通して最も多くの基準を満たすスピーカ配置は T70 であり、次いで T60 と T80 であった。これら三つのスピーカ配置では、方位角、上昇角とも回答の分布が目標方向と概ね一致している。また、平均定位誤差は、従来配置より小さい。ただし、実音源の上昇角との比較においては、T70 では2名とも有意な差がなかったが、T60 と T80 で有意な差がなかったのは1名だけであった。前後誤判定率は、従来配置より低かったが、実音源との比較においては、T70 で1名が同等以下となったが、T60 と T80 では2名とも同等以下ではなかった。
 以上より、T70 では2名の被験者に共通して、目標

方向を水平面においた場合も正中面においた場合も、回答の方向は目標方向に概ね一致し、水平面における方位角及び正中面における上昇角の定位精度は、それぞれ実音源と有意な差がないと見なせる。
3. 考察:再生系伝達関数と音像定位精度の関係
3.1 目標の伝達関数の再現精度
 音像制御精度がスピーカ配置に依存する理由を考察するために、目標とする伝達関数の再現精度を分析した。音像定位実験の直前に再生系伝達関数を再度測定した。これを $C'_{i,j}(\omega)$ と表す。 i は左右のスピーカ (L, R), j は左右の耳 (l, r) である。
 あらかじめ測定した再生系伝達関数を用いて作成した信号 $X_i(\omega)$ (式(1)) を用いると、実験時の受聴者の外耳道入口における音圧 $P_j(\omega)$ は式(6)のように表せる。

$$\left. \begin{aligned} P_l(\omega) &= X_L(\omega) \times C'_{L,l}(\omega) + X_R(\omega) \times C'_{R,l}(\omega) \\ P_r(\omega) &= X_R(\omega) \times C'_{R,r}(\omega) + X_L(\omega) \times C'_{L,r}(\omega) \end{aligned} \right\} \quad (6)$$

 ここで、式(6)の $X_i(\omega)$ に含まれる音源信号を $S(\omega) = 1$ とおくと、音像定位実験時の伝達関数が求められる。

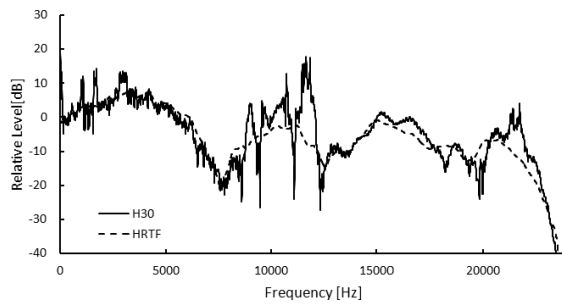


図-17 H30 配置による伝達関数制御
実線：制御後の伝達関数，破線：目標の伝達関数（被験者 A）

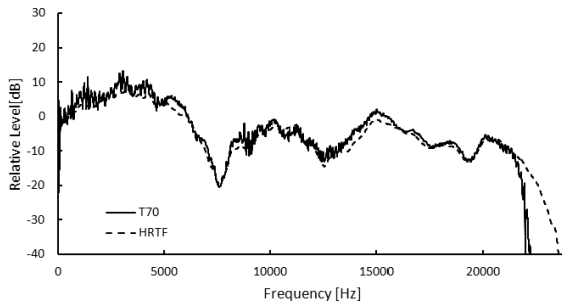


図-18 T70 配置による伝達関数制御
実線：制御後の伝達関数，破線：目標の伝達関数（被験者 A）

H30 で定位精度が低い例として，目標方向が真後ろの場合をとりあげ，H30 及び T70 における左耳での音像定位実験時の伝達関数と目標の伝達関数（頭部伝達関数）を比較して示す（図-17, 18）。

H30（図-17）では，制御後の伝達関数には，約 8,500～11,000 Hz に目標の伝達関数には存在しない顕著なスペクトラルピーク・ノッチが多数発生している。このスピーカ配置では，着席時の受聴点のずれによって生じる再生系伝達関数の変化の影響で，式 (1) で求めた信号を放射しても，目標とする伝達関数を外耳道入口で再現することができなくなったと考えられる。

一方，T70（図-18）では，ほぼ目標どおりの伝達関数が再現されている。従って，T70 では，実際の耳の位置が制御点から多少ずれたとしても，再生系伝達関数の変化が小さく，目標とする伝達関数が精度良く再現されることが考えられる。

3.2 再生系伝達関数の分析

次に，伝達関数の再現精度とスピーカ配置の関係を検討するため，各スピーカ配置における再生系伝達関数を左（又は右）のスピーカから左（又は右）耳に到達するダイレクト成分，及び左（又は右）のスピーカから右（又は左）耳に到達するクロストーク成分に分けて分析した。

3.2.1 ダイレクト成分の振幅スペクトル

図-19 に被験者 A の左スピーカから左耳への伝達関

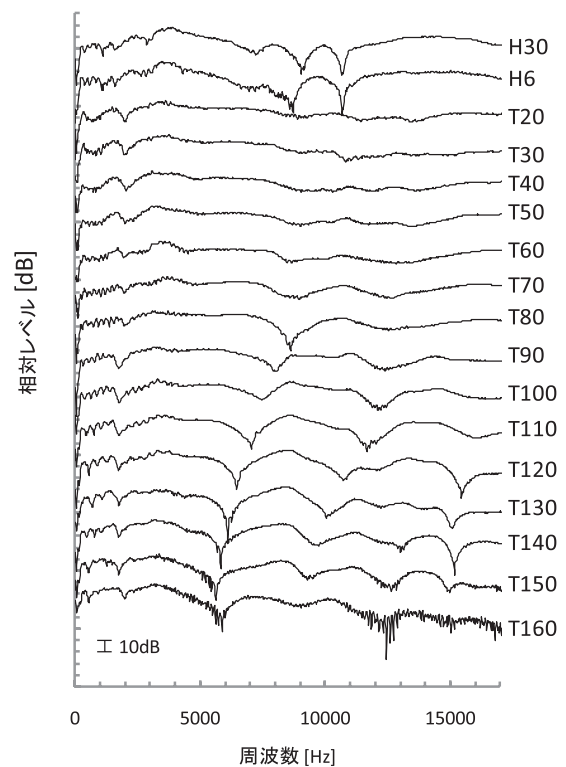


図-19 再生系伝達関数のダイレクト成分の振幅スペクトルの一例（被験者 A，左スピーカから左耳）

数の振幅スペクトルを示す。従来の水平面上のスピーカ配置 H30 と H6 では，8,500 Hz 付近及び 10,000 Hz 付近に顕著なスペクトラルノッチが生じている。これらのノッチは，音像の上昇角知覚のスペクトラルキュー（N1 と N2）であり [10]，前方の水平面上に音像を知覚する重要な手がかりである。従って，上方や後方へ音像制御するには，これらのスペクトラルノッチを補正する必要があり，前述の着席時の位置の違いによる影響を考えると，3次元音像制御にとって有利なスピーカ配置とは考えにくい。

一方，横断面上のスピーカ配置のうち上方から側方の T20～T70 では顕著なノッチが生じておらず，比較的平坦なスペクトルを有している。T80 では 8,500 Hz 付近にノッチが生じており T90 では 7,700 Hz 付近及び 12,000 Hz 付近にノッチが生じている。スピーカ配置が下方になるに従ってノッチ周波数が低くなる傾向がみられる。なお，ノッチ周波数には個人差がみられたが，横断面上の上方から側方でスペクトルが比較的平坦であること，及びスピーカが下方になるとノッチ周波数が低くなるという傾向は，2名の被験者の両耳に共通してみられた。

更に，ダイレクト成分の平坦性について定量的な評価を試みる。いま，ダイレクト成分及びクロストーク成分の相対音圧レベルを $L_{i,j}(\omega)$ で表す。ここで， i は左右のスピーカ (L, R)， j は左右の耳 (l, r) を表

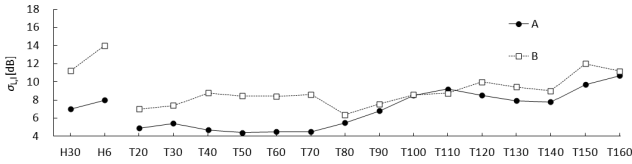


図-20 ダイレクト成分の標準偏差 $\sigma_{L,1}$ (200~17,000 Hz)

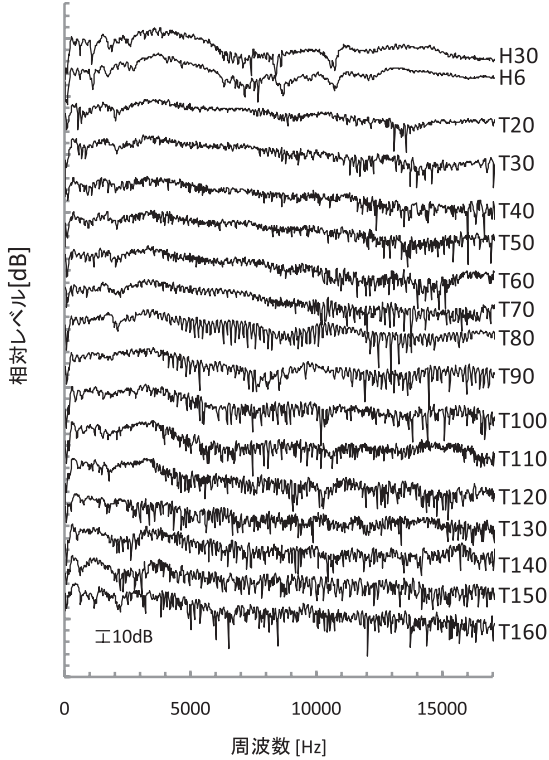


図-21 再生系伝達関数のクロストーク成分の振幅スペクトルの一例 (被験者 A, 右スピーカから左耳)

す。ダイレクト成分の平坦性を定量化するために、相対音圧レベル $L_{L,1}(\omega)$, $L_{R,r}(\omega)$ の周波数軸上 (200~17,000 Hz) での標準偏差 $\sigma_{L,1}$, $\sigma_{R,r}$ を求めた。図-20 に2名の被験者の $\sigma_{L,1}$ を示す。

$\sigma_{L,1}$ の値は被験者間で差異があるが、H30とH6を比較すると、どちらの被験者においてもH30の方が低い。H30よりも $\sigma_{L,1}$ が低くなる横断面のスピーカ配置は、被験者AではT20~T90、被験者BではT150以外であった。

3.2.2 クロストーク成分の振幅スペクトル

被験者Aの右スピーカから左耳への伝達関数の振幅スペクトルを図-21に示す。従来の水平面上のスピーカ配置と比較すると、横断面上のスピーカ配置のクロストーク成分は高周波数帯域でのレベルが低いことが分かる。また、スピーカ配置が下方になるに従って、より低い周波数からレベルが低下する傾向がみられる。この傾向は全被験者の両耳に共通してみられた。

トランスオーラルシステムでは、クロストーク成分は消去されるべき成分であり、もともとこの成分が少

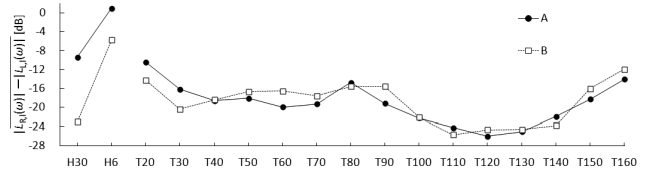


図-22 ダイレクト成分に対するクロストーク成分の平均相対レベル $|L_{R,1}(\omega)| - |L_{L,1}(\omega)|$ (200~17,000 Hz)

表-3 再生系伝達関数と平均定位誤差の相関係数 (被験者 A)

再生系伝達関数 パラメータ	平均定位誤差			
	水平面		正中面	
	方位角	仰角	方位角	上昇角
$\sigma_{L,1}$	0.19	0.53	0.07	0.78
$ L_{R,1}(\omega) - L_{L,1}(\omega) $	0.81	-0.01	0.69	0.25

表-4 再生系伝達関数と平均定位誤差の相関係数 (被験者 B)

再生系伝達関数 パラメータ	平均定位誤差			
	水平面		正中面	
	方位角	仰角	方位角	上昇角
$\sigma_{L,1}$	-0.23	-0.17	—	0.35
$ L_{R,1}(\omega) - L_{L,1}(\omega) $	0.38	-0.04	—	-0.26

ないスピーカ配置を選べば音像制御精度が向上すると推測される。また、もしクロストーク成分がダイレクト成分と同等の振幅を有していれば、式(1)の分母がゼロに近づいて発散する恐れがある。系の安定性という面からもダイレクト成分に対するクロストーク成分の相対レベルが低いことが望ましい。そこで、ダイレクト成分に対するクロストーク成分の相対的なレベルの低さを定量的に評価するために、200~17,000 Hzでのクロストーク成分の相対レベルとダイレクト成分の相対レベルの差の周波数平均 $|L_{R,1}(\omega)| - |L_{L,1}(\omega)|$ 及び $|L_{L,r}(\omega)| - |L_{R,r}(\omega)|$ を求めた。図-22に2名の被験者の $|L_{R,1}(\omega)| - |L_{L,1}(\omega)|$ を示す。

どちらの被験者についても、T80付近からレベルが減少し始め、T110~T130で低くなり、スピーカ配置が更に下方になるとT160まで再び上昇している。H30とH6を比較すると、どちらの被験者においてもH30の方が低い。H30よりも $|L_{R,1}(\omega)| - |L_{L,1}(\omega)|$ が低くなる横断面のスピーカ配置は、被験者Aでは全配置、被験者BではT110~T140であった。

3.3 再生系伝達関数と音像定位精度の相関分析

各スピーカ配置における再生系伝達関数の分析結果と音像定位精度の関係を考察するために相関分析を行った。表-3, 4にダイレクト成分の平坦性 $\sigma_{L,1}$ 、及びクロストーク成分の相対レベル $|L_{R,1}(\omega)| - |L_{L,1}(\omega)|$ と平均定位誤差の相関係数を示す。ただし、被験者Bで

目標方向を正中面においた場合の方位角の平均定位誤差は、すべてのスピーカ配置で 0.0° であったため、相関関係は議論できない。

有意水準 5% で相関があると見なせるのは、自由度 16 では相関係数の絶対値が 0.468 以上であることから、被験者 A における以下の四つの関係である。

1) ダイレクト成分の平坦性 ($\sigma_{L,1}$) と水平面における仰角の平均定位誤差

2) ダイレクト成分の平坦性 ($\sigma_{L,1}$) と正中面における上昇角の平均定位誤差

3) クロストーク成分の相対レベル ($|L_{R,1}(\omega)| - |L_{L,1}(\omega)|$) と水平面における方位角の平均定位誤差

4) クロストーク成分の相対レベル ($|L_{R,1}(\omega)| - |L_{L,1}(\omega)|$) と正中面における方位角の平均定位誤差

これより、被験者 A の場合は、目標方向が水平面、正中面に関わらず、ダイレクト成分の平坦性は仰角及び上昇角の制御精度と相関があり、クロストーク成分の相対レベルは方位角の制御精度と相関があると言える。

一方、被験者 B については、ダイレクト成分の平坦性と正中面における上昇角の平均定位誤差の関係、及びクロストーク成分の相対レベルと水平面における方位角の平均定位誤差の関係において他の関係と比較して高い相関係数が得られたが、有意な相関であるとは言えない。

ただし、ダイレクト成分の平坦性及びクロストーク成分の相対レベルがトランスオーラルシステムの音像制御精度と因果関係があるか否かを論じるには、更に多くの被験者を用いて研究を進める必要がある。

4. 結 論

本研究では、横断面上の $\pm 20^\circ$ から $\pm 160^\circ$ までの 15 種類 (T20~T160)、及び従来提案されている水平面の前方 $\pm 30^\circ$ (H30) と $\pm 6^\circ$ (H6) の計 17 種類のスピーカ配置において、受聴者の頭部を固定しない状態で、トランスオーラル方式による音像定位実験を行った。その結果、3次元音像制御に適したスピーカ配置について以下の結論を得た。

1) 目標方向を水平面においた場合、被験者 2 名に共通して概ね目標の方位角及び仰角に定位したスピーカ配置は T60~T110 であった。また、被験者 2 名に共通して方位角の平均定位誤差が実音源のそれと有意な差がないと見なせるスピーカ配置は T70, T80, 及び T100 であった。

2) 目標方向を正中面においた場合、被験者 2 名に共

通して概ね目標の方位角及び上昇角に定位したスピーカ配置は T60~T80 であった。また、被験者 2 名に共通して上昇角の平均定位誤差が実音源と有意な差がないと見なせるスピーカ配置は T70 であった。

3) 従って、T70 では 2 名の被験者に共通して、定められた位置に設置された椅子に座れば頭部を固定せずとも、目標方向が水平面の場合も正中面の場合も、回答の方向は目標方向に概ね一致し、水平面における方位角及び正中面における上昇角の定位精度は、実音源と有意な差がないと見なせる。

4) 以上より、トランスオーラルシステムのスピーカは、従来のように水平面前方に配置するよりも、横断面の斜め上方に配置する方が 3次元音像再生において有利であると言える。

文 献

- [1] M.R. Schroeder and B.S. Atal, "Computer simulation of sound transmission in rooms," *IEEE Int. Conv. Rec.*, Part 7, 150-155 (1963).
- [2] P. Damaske and V. Mellert, "Ein Verfahren zur richtungstreu Schallabbildung des oberen Halbraumes über zwei Lautsprecher," *Acustica*, 22, 153-162 (1969/70).
- [3] P. Damaske, "Head-related two-channel stereophony with loudspeaker reproduction," *J. Acoust. Soc. Am.*, 50, 1109-1115 (1971)
- [4] 浜田晴夫, "基準的收音・再生を目的とする Orthostereophonic System の構成," *音響学会誌*, 39, 337-348 (1983)
- [5] O. Kirkeby, P.A. Nelson and H. Hamada, "Local sound field reproduction using two closely spaced loudspeakers," *J. Acoust. Soc. Am.*, 104, 1973-1981 (1998)
- [6] M. Morimoto and Y. Ando, "On the simulation of sound localization," *J. Acoust. Soc. Jpn. (E)*, 1, 167-174 (1980).
- [7] T. Takeuchi, P.A. Nelson, O. Kirkeby and H. Hamada, "Robustness of the performance of the "Stereo Dipole" to misalignment of head position," *AES 102nd Convention*, Preprint 4464, Munich, Germany (1997).
- [8] 飯田一博, 森本政之, *空間音響学* (コロナ社, 東京, 2010), pp. 118-122.
- [9] T. Takeuchi and P.A. Nelson, "Optimal source distribution for binaural synthesis over loudspeakers," *J. Acoust. Soc. Am.*, 112, 2786-2797 (2002).
- [10] K. Iida, M. Itoh, A. Itagaki and M. Morimoto, "Median plane localization using a parametric model of the head-related transfer function based on spectral cues," *Appl. Acoust.*, 68, 835-850 (2007).
- [11] M. Morimoto and H. Aokata, "Localization cues in the upper hemisphere," *J. Acoust. Soc. Jpn. (E)*, 5, 165-173 (1984).
- [12] J.C. Middlebrooks, "Narrow-band sound localization related to external ear acoustics," *J. Acoust. Soc. Am.*, 92, 2607-2624 (1992).
- [13] 飯田一博, 森本政之, *空間音響学* (コロナ社, 東京, 2010), p. 5.