

# 波面合成法を用いた立体音場再生における合成音像定位モデル

木村 敏幸<sup>1</sup> 勝本 道哲<sup>1</sup> 笥 一彦<sup>2</sup>

<sup>1</sup>独立行政法人情報通信研究機構ユニバーサルメディア研究センター

<sup>2</sup>中京大学情報理工学部

波面合成法による立体音場再生システムを実用化するためには定位実験を実施することが非常に重要である。しかし、定位実験を実施するには多大な時間と労力を要する。本論文では定位実験に要する時間と手間を削減するために、定位実験の結果を予測することができる定位モデルを提案する。定位モデルを実装したところ、定位実験における知覚方向を十分に予測できることが分かった。

## Localization Model of Synthesized Sound Image in Sound Field Reproduction Using Wave Field Synthesis

Toshiyuki KIMURA<sup>1</sup> Michiaki KATSUMOTO<sup>1</sup> and Kazuhiko KAKEHI<sup>2</sup>

<sup>1</sup>Universal Media Research Center, National Institute of Information and Communications Technology

<sup>2</sup>School of Information Science and Technology, Chukyo University

Although it is very important to perform localization tests in order to construct the practical sound field reproduction system based on wave field synthesis, it takes a lot of troubles to perform localization tests. In this paper, the localization model of synthesized sound images which predicts the result of localization tests is proposed in order to reduce the trouble of the performance of listening tests. It shows from the result between the proposed model and the localization test that the proposed localization model can predict the results of localization tests.

### 1. はじめに

近年、聴覚バーチャルリアリティシステムのための立体音場再生技術が盛んに研究されている。この技術が実用化されると、遠隔通信会議システムや遠隔アンサンブルシステムのように、離れた場所にいる人々があたかも同じ会議室内で会議を行ったり、同じコンサートホール内で一緒に演奏したりしているように思わせることが可能になる。これらのシステムはテレビ電話のような従来のコミュニケーション手段とは比べ物にならないほどの臨場感を与えることができるため、普及すれば社会全体において遠隔コミュニケーションがより一層活発になっていくことが期待される。

立体音場再生技術はマイクロホンで音を収録し、ヘッドホンもしくはスピーカで収録した音をそのまま再生するだけで聴取者に臨場感を感じさせることのできる技術である。中でも、波面合成法[1-6]は原音場に配置したマイクロホンアレーで音を収録し、再生音場に配置したスピーカアレーで収録した音をそのまま再生するもので、

Huygens の原理によって原音場の制御領域内の波面が再生音場の聴取領域内に忠実に再現される。この技術はバイノーラル[7]やトランスオーラル[8]とは異なり、聴取者が聴取領域内のどこに居ても構わないし、頭部を自由に動かしても自然な音像を得られるというところに特徴がある。

波面を物理的に合成するためには空間サンプリング定理[9]に基づきマイクロホンやスピーカを波長の 1/2 以下の間隔で配置する必要があるため、今まで波面合成法による立体音場再生システムを実用化することは困難であるとされてきた。しかし、聴取者の聴感に基づけばシステムの構築に必要なマイクロホンやスピーカの数を少なくすることができるので[10,11]、聴取実験を実施することで実用的なシステムを構築することができると考えられる。

しかし、実際にはシステムの使用用途に応じた様々な領域が存在し、その度にマイクロホンアレーやスピーカアレーを配置して聴取実験を実施する必要があるため、聴取実験に多大な時間と労力

を要していた．そこで，マイクロホンアレーやスピーカアレーの配置を指定しただけで聴取実験の結果を予測することができれば，実験に要する時間や労力を大幅に削減できるようになる．特に，臨場感を方向感，距離感，空間印象に大別して[12]聴取実験を行った場合には，空間印象よりも方向感を再現するのに要するマイクロホンとスピーカの数の方が多くなるので[10,11]，定位実験の結果を予測するモデルを構築することが重要であると考えられる．

従来提案されている音像定位モデル[13-16]は聴取者の両耳に入力される信号から両耳間時間差を算出し，それから定位方向を推定するものである．ゆえに，本論文で構築する定位モデルもこれらの従来モデルと同じように両耳間時間差を定位方向の推定基準とする．また，過去の定位実験においてマイクロホンやスピーカの数が少なすぎると先行音効果[17]が発生するために定位方向に偏りが生じていたことから[10,11]，定位モデルに先行音効果を組み込む必要がある．

しかし，Lindemann のモデル[14,15]は定位モデルに先行音効果を組み込んでいるが，ヘッドホンで音を提示した場合を想定しているため，両耳への入力信号にスピーカから聴取者への音響伝達特性を考慮していない．一方，黒住らのモデル[16]は2個のスピーカから音を提示した場合に合成される音像の方向を推定するモデルであるが，先行音効果を組み込んでいない．ゆえに，スピーカから聴取者までの音響伝達特性を考慮し，なおかつ先行音効果を組み込んだ定位モデルはまだ提案されていない．

そこで，本論文では波面合成法による立体音場再生において聴取実験に要する時間や労力を削減するために，先行音効果を組み込んだ合成音像定位モデルを提案する．2章では提案したモデルの原理について説明する．3章ではモデルの構築のために必要なモデルパラメータの設定について説明する．4章では先行研究における定位実験の結果[10,11]とモデルの結果を比較し，提案した定位モデルの有効性を検討する．

## 2. 合成音像定位モデルの原理

定位には音源からの直接音が主に作用している．ゆえに，原音場は反射音のない自由音場とする．図1の左側に示すように  $M$  個のマイクロホンを制御領域の境界上に配置した場合，音源から  $i$  番目

のマイクロホン( $i=1\dots M$ )までの室内インパルス応答  $g_i(t)$ は以下のように定義される．

$$g_i(t) = a_i \delta(t - t_i) \quad (1)$$

但し， $a_i (=1/d_i)$ ， $t_i (=d_i/c)$ は音源から  $i$  番目のマイクロホンまでの距離  $d_i$  に伴う減衰係数と遅延時間， $c$  は音速， $\delta(t)$  は Dirac のデルタ関数である．音源信号を  $s(t)$  とすると，各マイクロホンによって収録されるチャンネル信号  $x_i(t)$  は以下のように表わされる．

$$x_i(t) = D_{im} \{g_i(t) * s(t)\} = D_{im} a_i s(t - t_i) \quad (2)$$

但し， $*$  は畳み込み演算を表わす．また， $D_{im}$  は  $i$  番目のマイクロホンの指向特性係数で，制御領域の外側からの音波のみ収録するように設定する[18]．

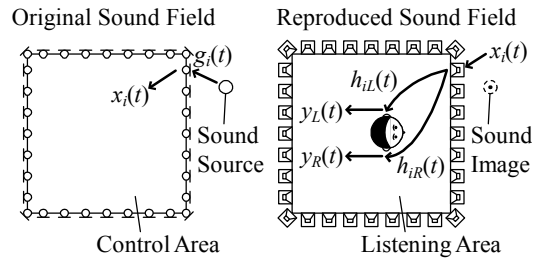


図1 定位モデルにおけるチャンネル信号及び両耳信号の作成

図1の右側に示すように再生音場において聴取者と聴取領域の境界上にスピーカアレーを配置し， $i$  番目のスピーカから両耳までの頭部伝達関数  $h_{iL}(t)$ ， $h_{iR}(t)$  を黒住らのモデル[16]と同じく以下の式のように近似する．

$$h_{iL}(t) \approx a_{iL} \delta(t - t_{iL}) \quad (3)$$

$$h_{iR}(t) \approx a_{iR} \delta(t - t_{iR})$$

但し， $a_{iL}$ ， $a_{iR}$  は  $i$  番目のスピーカから左耳及び右耳までの振幅レベル， $t_{iL}$ ， $t_{iR}$  は  $i$  番目のスピーカから左耳及び右耳までの初期遅延時間である．各チャンネル信号と頭部伝達関数から，両耳信号  $y_L(t)$ ， $y_R(t)$  は以下のように表わされる．

$$\begin{aligned} y_L(t) &= \sum_{i=1}^M p_i D_{is} \{h_{iL}(t) * x_i(t)\} \\ &= \sum_{i=1}^M p_i D_{im} D_{is} a_i a_{iL} s(t - T_{iL}) \\ y_R(t) &= \sum_{i=1}^M p_i D_{is} \{h_{iR}(t) * x_i(t)\} \\ &= \sum_{i=1}^M p_i D_{im} D_{is} a_i a_{iR} s(t - T_{iR}) \end{aligned} \quad (4)$$

但し， $T_{iL} = t_i + t_{iL}$ ， $T_{iR} = t_i + t_{iR}$ ， $D_{is}$  は  $i$  番目のスピー

カの指向特性係数で、聴取領域の内側に向けて設定する[18]. また、 $p_i$ は先行音効果を表す係数で、Lindemann のモデル[14,15]と同じく以下の式のように設定する.

$$p_i = \exp\{\alpha(t_{\min} - t_i - t'_i)\} \quad (5)$$

$$t_{\min} = \min_i(t_i + t'_i)$$

但し、 $t'_i(=d'_i/c)$ は  $i$  番目のスピーカから聴取者までの距離  $d'_i$  に伴う遅延時間である. つまり、音源から最も早く聴取者に到達する音の経路を基準とし、その他の経路による音がそれから遅れて聴取者に到達するに従って、定位への影響が小さくなるように重み付けを行う. その際、適切な重み付けになるように  $\alpha(>0)$  を設定する.

両耳信号から両耳間相関関数  $R(\tau)$  を算出すると、以下の式ようになる.

$$R(\tau) = E\{y_L(t)y_R(t-\tau)\} \quad (6)$$

$$= \sum_{i=1}^M \sum_{j=1}^M P_{ij} D_{ij} A_{ij} E\{s(t-T_{iL})s(t-T_{jR}-\tau)\}$$

但し、 $P_{ij}=p_i p_j$ ,  $D_{ij}=D_{im} D_{jm} D_{is} D_{js}$ ,  $A_{ij}=a_i a_j a_{iL} a_{jR}$  である. ゆえに、音源信号の統計的性質が両耳間相関関数の概形に反映されることになる. 本論文では両耳間時間差を算出したいので、必要な情報は両耳間相関関数のピークの時間だけである. そこで、ピークを強調するために音源信号は自己相関性を持たないことにすると、(6)式は以下のように書き換えられる.

$$R(\tau) = \sum_{i=1}^M \sum_{j=1}^M P_{ij} D_{ij} A_{ij} \delta(\tau - T_{ij}) \quad (7)$$

但し、 $T_{ij}=T_{iL}-T_{jR}=(t_i-t_j)+(t_{iL}-t_{jR})$  である. ゆえに、 $T_{ij}$  の位置に大きさ  $P_{ij} D_{ij} A_{ij}$  のピークが合計  $M^2$  個発生するので、黒住らのモデル[16]と同じように実効的な両耳間時間差  $\tau_E$  を以下の式のように算出する.

$$\tau_E = \frac{\sum_{i=1}^M \sum_{j=1}^M P_{ij} D_{ij} A_{ij} T_{ij}}{\sum_{i=1}^M \sum_{j=1}^M P_{ij} D_{ij} A_{ij}} \quad (8)$$

もし、定位方向  $\phi$  と両耳間時間差  $\tau$  との間に以下のような関係が成り立つならば、

$$\tau = f(\phi) \quad (9)$$

合成音像の方向  $\phi$  は実効的な両耳間時間差  $\tau_E$  より以下の式のように算出される.

$$\phi = f^{-1}(\tau_E) \quad (10)$$

$$= f^{-1}\left(\frac{\sum_{i=1}^M \sum_{j=1}^M P_{ij} D_{ij} A_{ij} T_{ij}}{\sum_{i=1}^M \sum_{j=1}^M P_{ij} D_{ij} A_{ij}}\right)$$

### 3. モデルパラメータの設定

#### 3.1. チャネル信号の作成

先の定位実験[10,11]の結果と比較するために、図2に示すように制御領域を半径  $r=2$  m の円とし、4箇所聴取位置(中心、前方、後方、側方)を設定した. このとき、中心から正面方向への聴取位置の移動距離を  $\Delta_x$ 、左側方向への移動距離を  $\Delta_y$  とすると、4箇所聴取位置の座標は以下のように表わされる. 但し、単位は m である.

$$(\Delta_x, \Delta_y) = \begin{cases} (0,0) & \text{(中心)} \\ (0.5,0) & \text{(前方)} \\ (-0.5,0) & \text{(後方)} \\ (0,0.5) & \text{(側方)} \end{cases} \quad (11)$$

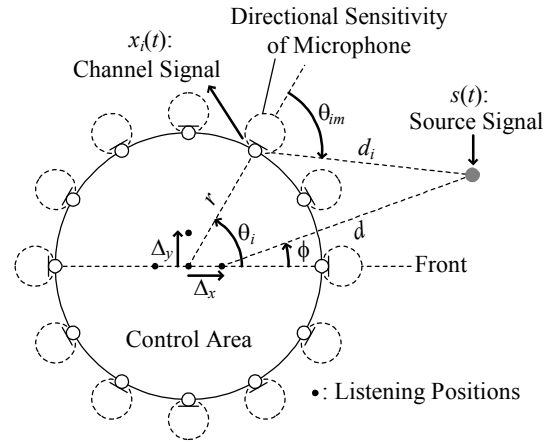


図2 モデル構築におけるチャネル信号の作成

音源から聴取位置までの距離を  $d$  m とすると、音源から  $i$  番目のマイクロホンまでの距離  $d_i$  は以下ようになる.

$$d_i = \sqrt{(d \cos \phi + \Delta_x - r \cos \theta_i)^2 + (d \sin \phi + \Delta_y - r \sin \theta_i)^2} \quad (12)$$

但し、 $\phi$  は聴取位置における音源の方位角、 $\theta_i$  は  $i$  番目のマイクロホンの方位角である.  $d_i$  から  $x_i(t)$  は(2)式から以下のように作成される.

$$x_i(t) = D_{im} \frac{d-r}{d_i} s\left(t - \frac{d_i}{c}\right) \quad (13)$$

但し、 $c$  ( $=340$  m/s) は音速である. また、 $D_{im}$  は過去の研究成果[18]に従い以下のように超指向性とした.

$$D_{im} = \begin{cases} \cos \theta_{im} & (|\theta_{im}| \leq 90^\circ) \\ 0 & (|\theta_{im}| > 90^\circ) \end{cases} \quad (14)$$

但し、 $\theta_{im}$  は図2に示すように  $i$  番目のマイクロホンにおける音源の入射角を表す. 従って、モデルの構築に必要な  $a_i, t_i$  は以下のように表わされる.

$$a_i = \frac{d-r}{d_i}, \quad t_i = \frac{d_i}{c} \quad (15)$$

### 3.2. 頭部伝達関数の測定

モデルの構築に際して，両耳信号を作成するのに必要な  $a_{iL}$ ,  $a_{iR}$ ,  $t_{iL}$ ,  $t_{iR}$  は測定した頭部伝達関数から推定した値を用いる．そこで，実際に頭部伝達関数の測定を行った．

低残響室内において半径 2 m の円周上に 15 度間隔で計 24 個のスピーカ(Emic: Soundevice)を配置し，円の中心に HATS(Head and Torso Simulator)を配置した．その際，HATS の耳の高さはスピーカの高さと同じにした．また，HATS の向いている方向を正面(0 度)とした．つまり，各スピーカの方位角は  $-165^\circ$ ,  $-150^\circ$ , ...,  $0^\circ$ ,  $15^\circ$ , ...,  $165^\circ$ ,  $180^\circ$  となる．各スピーカからサンプリング周波数 48 kHz で 65536 点の TSP 信号[19]を再生して頭部伝達関数を測定した．同期加算は 8 回とした．部屋の残響の影響を除去するため，頭部伝達関数はスピーカからの直接音だけが到達する最初の 440 点で打ち切った．

### 3.3. 両耳間時間差，初期遅延時間の推定

音像定位の際に両耳間時間差が作用するのは主に 1.6kHz 以下である[20]．ゆえに，両耳間時間差及び初期遅延時間は 1.6kHz のローパスフィルタで処理した頭部伝達関数から推定を行う．

両耳間時間差  $\Delta t_{LR}(=t_{iL}-t_{iR})$  は以下の式のようにローパスフィルタで処理した両耳の頭部伝達関数から算出した両耳間相関関数のピーク値とした．

$$\Delta t_{LR} = \arg \max_{\tau} \{h'_{iL}(t)h'_{iR}(t-\tau)\} \quad (16)$$

但し， $h'_{iL}(t)$ ,  $h'_{iR}(t)$  は方位角  $\theta_i$  のスピーカから左耳及び右耳までの頭部伝達関数である． $\Delta t_{LR}$  の測定結果を図 3 中の左上に丸点で示す．

測定結果から左右対称性を考慮して両耳間時間差  $\Delta t_{LR}$  を以下のように推定する．

$$\Delta t_{LR}(\theta_i)[\text{ms}] = -0.64 \sin \theta_i \quad (17)$$

推定結果を図 3 中の左上に実線で示す．測定結果と推定結果の平均二乗誤差は 0.046 ms であり，十分に精度良く両耳間時間差を推定していると考えられる．

ローパスフィルタリング処理を行った頭部伝達関数の波形において最初のピークが来るまでの時間を左耳及び右耳の初期遅延時間  $t_{iL}$ ,  $t_{iR}$  とした． $t_{iL}$ ,  $t_{iR}$  の測定結果を図 3 中の左下に丸点及び三角点で示す．測定結果から左右対称性を考慮して左

耳及び右耳の初期遅延時間を以下のように推定する．

$$t_{iL}(d'_i, \theta_i)[\text{ms}] = \begin{cases} \frac{1000d'_i}{c} - 0.16 \sin \theta_i & (\theta_i \geq 0^\circ) \\ \frac{1000d'_i}{c} + 0.16 \sin \theta_i & (\theta_i < 0^\circ) \end{cases} \quad (18)$$

$$t_{iR}(d'_i, \theta_i)[\text{ms}] = -t_{iL}(d'_i, -\theta_i)$$

但し， $d'_i$  は図 2 より以下のように求められる．

$$d'_i = \sqrt{(r \cos \theta_i - \Delta_x)^2 + (r \sin \theta_i - \Delta_y)^2} \quad (19)$$

左耳及び右耳について推定した結果を図 3 中の左下に実線及び点線で示す．測定結果と推定結果の平均二乗誤差は 0.043 ms であり，十分に精度良く初期遅延時間を推定していると考えられる．

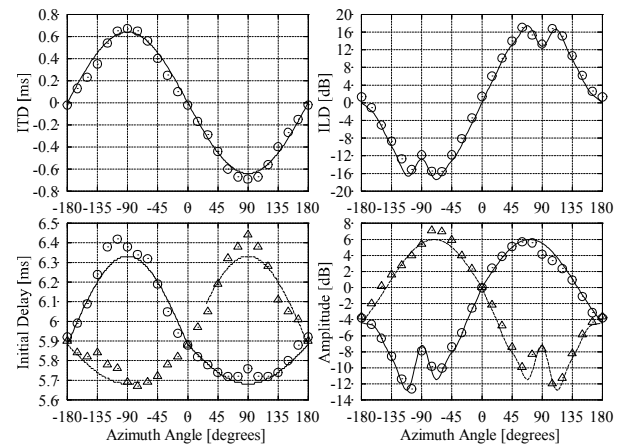


図 3 頭部伝達関数パラメータの測定及び推定結果

### 3.4. 両耳間レベル差，振幅レベルの推定

音像定位の際，両耳間レベル差は全帯域において作用している[20]．ゆえに，両耳間レベル差及び振幅レベルは測定した頭部伝達関数をそのまま用いて推定を行う．

左右それぞれの頭部伝達関数の平均パワーを左耳及び右耳の振幅レベル  $a_{iL}$ ,  $a_{iR}$  とした．その際，0 度方向の振幅レベルを基準にして dB 値に換算した．さらに左耳及び右耳の振幅レベルから両耳間レベル差  $\Delta a_{LR}(=a_{iL}/a_{iR})$  を算出した． $\Delta a_{LR}$  の測定結果を図 3 中の右上に丸点で示す．測定結果から左右対称性を考慮して両耳間レベル差を以下のように推定する．

$$\Delta a_{LR}(\theta_i)[\text{dB}] = \sum_{k=1}^7 Q_k \sin q_k \theta_i \quad (20)$$

但し， $q_k$ ,  $Q_k$  は以下のように設定されたパラメー

タである。

$$\{q_k\} = 1, 2, 4, 5, 7, 9, 11 \quad (21)$$

$$\{Q_k\} = 16.06, 1.64, 0.70, -1.36, 0.88, -0.70, 0.37$$

推定結果を図3中の右上に実線で示す。測定結果と推定結果の平均二乗誤差は1.035 dBであり、十分に精度良く両耳間レベル差を推定していると考えられる。

$a_{iL}, a_{iR}$  の測定結果を図3中の右下に丸点及び三角点で示す。測定結果から左右対称性を考慮して左耳及び右耳の振幅レベルを以下のように推定する。

$$a_{iL}(d'_i, \theta_i)[\text{dB}] = \begin{cases} 20 \log_{10} \frac{2}{d'_i} + 5.99 \sin\left(\frac{180^\circ}{143^\circ} \theta_i\right) & (\theta_i \geq 0^\circ) \\ 20 \log_{10} \frac{2}{d'_i} - 5.99 \sin\left(\frac{180^\circ}{143^\circ} \theta_i\right) & (\theta_i < 0^\circ) \end{cases} \quad (22)$$

$$+ \Delta a_{LR}(\theta_i)$$

$$a_{iR}(d'_i, \theta_i)[\text{dB}] = -a_{iL}(d'_i, -\theta_i)$$

左耳及び右耳について推定した結果を図3中の右下に実線及び点線で示す。測定結果と推定結果の平均二乗誤差は0.676 dBであり、十分に精度良く振幅レベルを推定していると考えられる。

### 3.5. 先行音効果の設定

$p_i$  は3.1節で得られた  $t_i$  と  $t'_i (=d'_i/c)$  から設定した。また、 $\alpha$  は  $5.25 \times 10^3$  とした。この値の場合、先行音効果が発生する下限時間(0.63 ms-1 ms[20])における  $p_i$  は0.0366-0.0052となり、先行音効果によって定位への影響が小さくなるのを十分に表していると考えられる。

### 3.6. 音像方向の予測

3.1節で得られた  $a_i, t_i, D_{im}$ , 3.3節で得られた  $t_{iL}, t_{iR}$ , 3.4節で得られた  $a_{iL}, a_{iR}$ , 3.5節で得られた  $p_i$  を(8)式に代入して実効的な両耳間時間差  $\tau_E$  を算出した。また、 $D_{is}$  は過去の研究成果[18]に従い無指向性( $D_{is}=1$ )とした。

(17)式から、両耳間時間差  $\tau$  と定位方向  $\phi$  の間には以下に示すような関係が導かれる。

$$\tau = -0.64 \sin \phi \quad (-90^\circ \leq \phi \leq 90^\circ) \quad (23)$$

ゆえに、(23)式を(10)式に適用すると、音像方向  $\phi$  は以下のように予測される。

$$\phi = \sin^{-1}\left(-\frac{\tau_E}{0.64}\right) \quad (24)$$

但し、予測できる方向は  $-90^\circ \leq \phi \leq 90^\circ$  とし、 $-\tau_E/0.64 < -1, -\tau_E/0.64 > 1$  のときは強制的に  $\phi = -90^\circ, 90^\circ$  とした。

## 4. 定位実験結果との比較

第3章にて構築したモデルの有効性を検討するために、先に実施した定位実験[10,11]の結果との比較を行った。定位実験の実施手順は文献[10,11]に記述している。

提案モデルによる音像の定位方向は(24)式によって予測した。その際、提示方向  $\phi$  は  $-20^\circ$  から  $20^\circ$  までとした。また、提案モデルに対する比較対象として、従来モデルによる音像方向の予測も行った。これは(5)式で  $\alpha=0$  にすることで先行音効果のパラメータを  $p_i=1$  にしたものであり、黒住らのモデル[16]に相当する。

従来モデル、提案モデルによって得られた音像の予測方向  $\phi'$  と定位実験において音源に白色雑音、音声を用いた場合の回答方向をチャンネル数条件ごとに図4-8に示す。チャンネル数が8, 12, 18のときには定位実験における回答方向の偏りが生じている。これは先の研究でも言及したように先行音効果が生じたためであると考えられることができる[10,11]。

従来モデルは先行音効果を組み込んでいないモデルである。そのため、回答方向が偏っている場合には音像の予測方向が回答方向と著しく異なっている。また、聴取位置が側方の場合には、すべてのチャンネル数条件において音像の予測方向が回答方向と大きく異なっている。それに比べて、提案モデルは先行音効果を組み込んだモデルであるため、回答方向が偏ったり聴取位置が側方である場合においてもモデルによる予測方向は回答方向とほぼ一致している。ゆえに、波面合成法による立体音場再生システムにおいて定位方向を正確に予測するためには、定位モデルに先行音効果を組み込むことが非常に重要であると考えられる。

さらに定量的な考察を行うため、以下の式で定義される平均二乗誤差を算出する。

$$\text{MSE}[\text{degrees}] = \sqrt{\frac{\sum_{\phi} (\phi' - \phi_0)^2}{L \times S}} \quad (25)$$

但し、 $\phi', \phi_0$  は提示方向を  $\phi$  とした場合のモデルの予測方向と定位実験における回答方向、 $L(=7)$  は提示方向の総数、 $S(=2)$  は使用したドライソースの総数である。モデルの種類(従来、提案)及び音源距離(3, 4 m)ごとに得られた平均二乗誤差を聴取位置ごとに図9に示す。

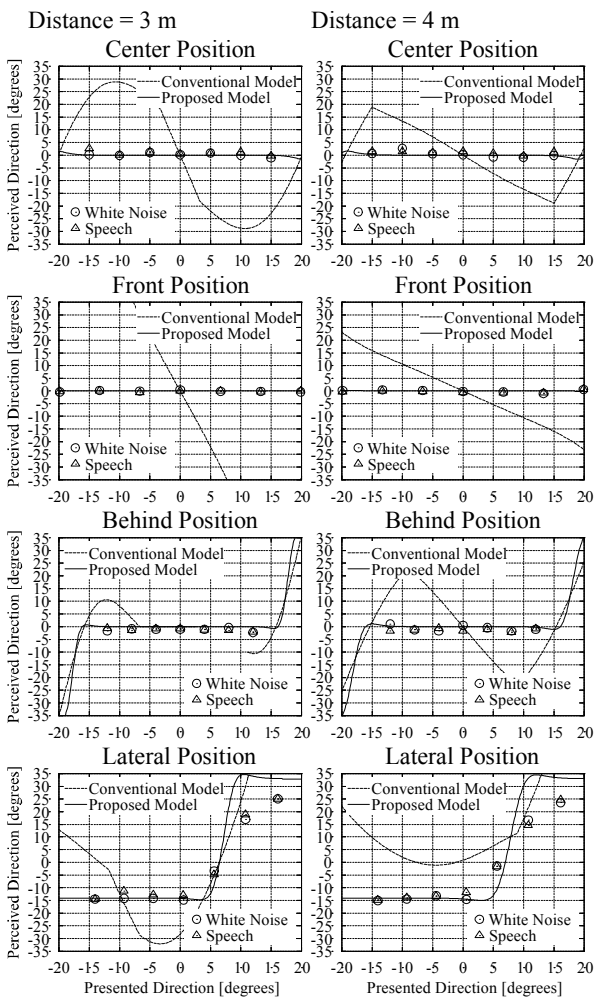


図 4 定位実験と定位モデルの結果：チャンネル数 8

従来モデルにおける平均二乗誤差は全体的に大きな値を示している．特に，チャンネル数が 8 の時や聴取位置が側方の場合にはたいていの場合において平均二乗誤差が 10 度を超えている．一方，提案モデルにおける平均二乗誤差はチャンネル数が 8 の時や聴取位置が側方の場合には総じて従来モデルの平均二乗誤差よりも小さい．また，それ以外の場合でも平均二乗誤差はおおむね 5 度以下である．この値は白色雑音のような広帯域音における正面方向の最小弁別角度(約 3 度[20])と比べると大きいので，聴覚のみの観点からは十分な精度とはいえないが，腹話術効果における正面方向の最小弁別角度(少なくとも約 11 度[21])と比べると十分に小さいので，視覚と聴覚を組み合わせた視聴覚バーチャリアリティシステムを構築する場合には提案した定位モデルは定位実験の結果を予測して実験に要する時間や手間を削減するのに十分効果的であると言える．

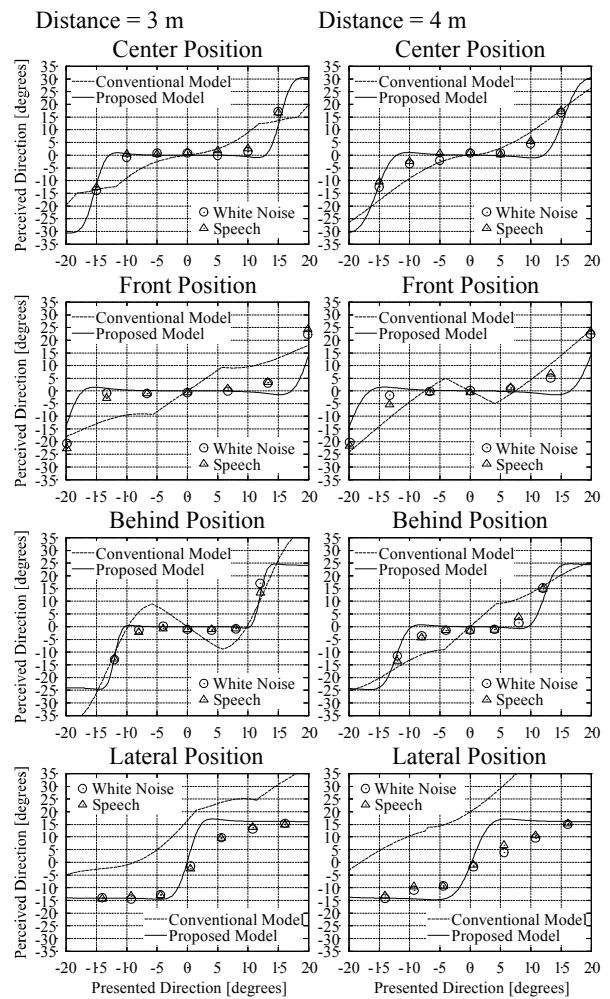


図 5 定位実験と定位モデルの結果：チャンネル数 12

## 5. まとめ

本論文では波面合成法による立体音場再生システムを実用化するにあたって，方向感を再現するのに必要なマイクロホンやスピーカの数を検討するための聴取実験に要する時間と手間を削減するために，聴取実験の結果を予測することができる合成音像定位モデルを提案した．提案モデルは先行音効果を組み込んだうえで両耳間時間差を基に定位方向を予測している．定位モデルを実装し，先に実施した聴取実験における知覚方向とモデルによる予測方向を比較したところ，十分な精度で聴取実験の結果を予測できることが分かった．

本論文で提案した定位モデルは両耳間時間差に基づいたものであるため，予測できる方向は水平面前方に限定した．しかし，一般的に臨場感を伴うシステムは前方のみならず後方や上方にも音を提示できるシステムである．そこで，今後は後方

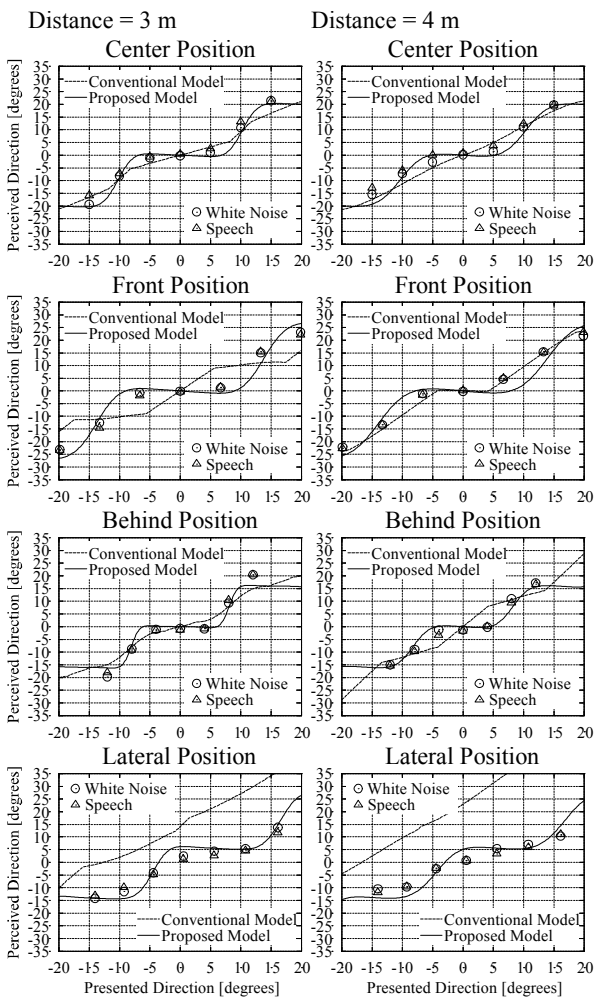


図 6 定位実験と定位モデルの結果:チャンネル数 18

や上方に音を提示した場合の聴取実験の結果を予測できるような定位モデルを検討する必要がある。

## 6. 謝辞

本研究の遂行にあたりご支援下さった名城大学理工学部板倉文忠教授, 名古屋大学大学院情報科学研究科武田一哉教授, 東京農工大学大学院工学府中川正樹教授諸先生方に感謝の意を申し上げる。

## 文 献

[1] H. Fletcher, "Symposium on wire transmission of symphonic music and its reproduction on auditory perspective: Basic requirement," *Bell System Technical Journal*, vol.13, no.2, pp.239-244, April 1934.  
 [2] 伊藤毅, 音響工学原論, コロナ社, 東京, 1973.  
 [3] M. Camras, "Approach to recreating a sound field," *J. Acoust. Soc. Am.*, vol.43, no.6, pp.1425-1431, June 1968.

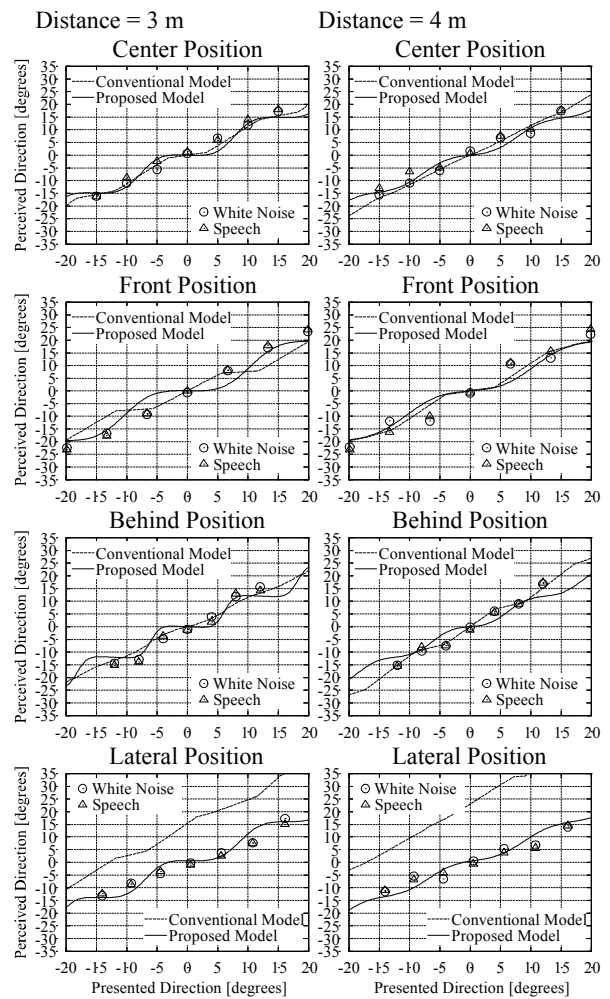


図 7 定位実験と定位モデルの結果:チャンネル数 24

[4] A.J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Am.*, vol.93, no.5, pp.2764-2778, May 1993.  
 [5] 及川靖広, 馬場貞如, 山崎芳男, "人間の聴覚を考慮した波面合成法による音場制御と音場再現," *日本バーチャルリアリティ学会論文誌*, vol.5, no.3, pp.919-926, September 2000.  
 [6] 平井徹, 本地由和, 池田雅弘, 清水寧, 川上福司, 東山三樹夫, "遠隔地音楽レッスン用 AVW システムとその原理," *電子情報通信学会技術研究報告*, no.EA2000-17, pp.25-32, June 2000.  
 [7] 岡部馨, "ダミーヘッドを用いた音場再生," *日本音響学会誌*, vol.46, no.8, pp.650-656, August 1990.  
 [8] 安藤四一, 設楽貞樹, 前川純一, 城戸健一, "電子計算機による室内音響設計法に関する二, 三の基礎的考察," *日本音響学会誌*, vol.29, no.3, pp.151-159, March 1973.  
 [9] 大賀寿郎, 山崎芳男, 金田豊, 音響システムとデジタル処理, 電子情報通信学会, 東京, 1995.

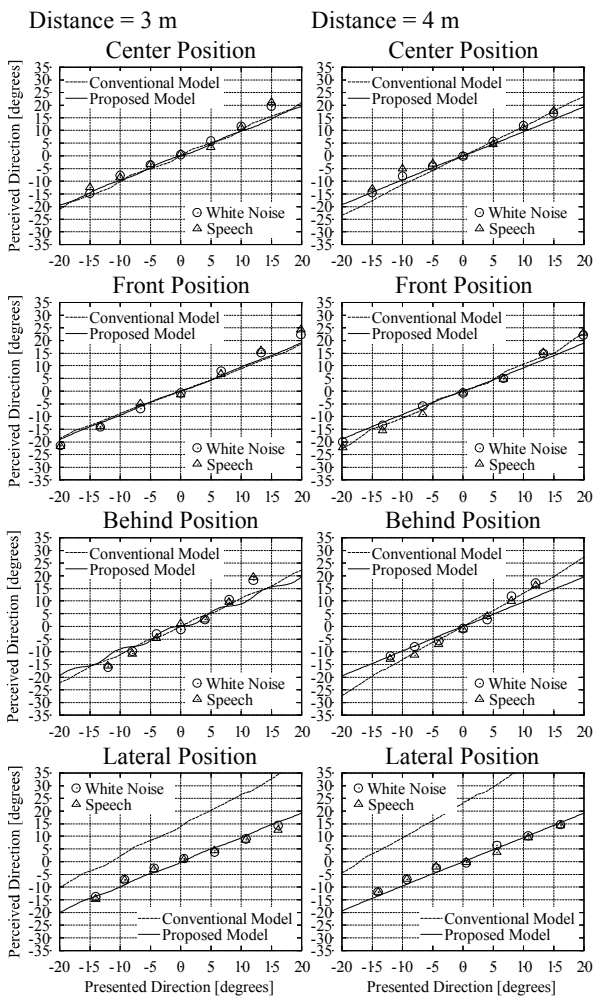


図 8 定位実験と定位モデルの結果:チャンネル数 36

[10] 木村敏幸, 箕一彦, 武田一哉, 板倉文忠, “波面合成におけるチャンネル数の主観的影響—音源が正面付近にある場合—,” 日本バーチャルリアリティ学会論文誌, vol.10, no.2, pp.257-266, June 2005.

[11] 木村敏幸, 波面合成法による立体音場再生における情報量削減手法に関する研究, 博士論文, 名古屋大学大学院人間情報学研究科, December 2005.

[12] M. Morimoto, “The relation between spatial impression and the precedence effect,” Proceedings of International Conference on Auditory Display, no.SS#2-1, pp.297-306, Kyoto, Japan, July 2002.

[13] L.A. Jeffress, “A place theory of sound localization,” J. Comp. Physiol. Psychol., vol.61, pp.468-486, 1948.

[14] W. Lindemann, “Extension of a binaural crosscorrelation model by contralateral inhibition. I. Simulation of lateralization for stationary signals,” J. Acoust. Soc. Am., vol.80, no.6, pp.1608-1622, December 1986.

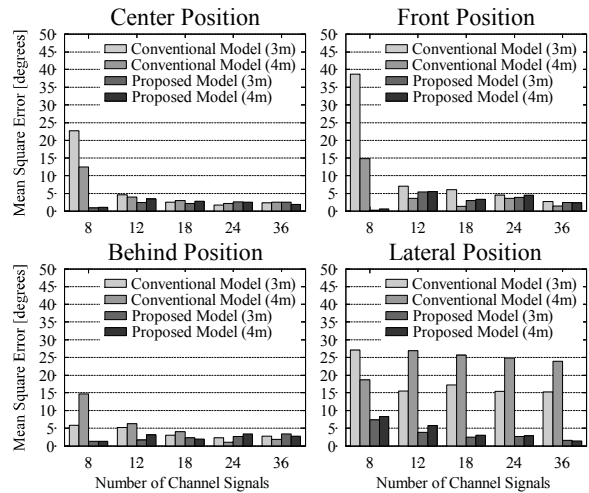


図 9 定位モデルの平均二乗誤差

[15] W. Lindemann, “Extension of a binaural crosscorrelation model by contralateral inhibition. II. The law of the first wave front,” J. Acoust. Soc. Am., vol.80, no.6, pp.1623-1630, December 1986.

[16] 黒住幸一, 大串健吾, “音響信号の両耳間相関係数に基づく音像定位の予測モデル,” 日本音響学会誌, vol.44, no.10, pp.726-734, October 1988.

[17] H. Wallach, E.B. Newman, and M.R. Rosenzweig, “The precedence effect in sound localization,” Am. J. Psycho., vol.62, no.3, pp.315-336, July 1949.

[18] 木村敏幸, 箕一彦, “波面合成法による三次元音場再現におけるマイクロホン及びスピーカの指向特性による影響,” 日本音響学会研究発表会講演論文集, no.3-Q-20, pp.689-690, March 2007.

[19] Y. Suzuki, F. Asano, H.Y. Kim, and T. Sone, “An optimum computer-generated pulse signal suitable for the measurement of very long impulse responses,” J. Acoust. Soc. Am., vol.97, no.2, pp.1119-1123, February 1995.

[20] イェンス・ブラウエルト, 森本政之, 後藤敏幸, 空間音響, 鹿島出版会, 東京, 1986.

[21] 小宮山撰, “視覚と聴覚による音像知覚,” 日本音響学会誌, vol.52, no.1, pp.46-50, January 1996.