

聴覚情報処理過程を考慮した高次音場情報の再生手法

鈴木陽一

東北大学電気通信研究所

1 はじめに

臨場感通信を実現するためには、臨場感を感じる主体である人間における情報処理過程を理解し、その理解に基づいて適切な情報の伝達と臨場感の再現を行うことが重要である。音響通信に関する高精度の臨場感通信を考えると、音質に関する高忠実度伝送だけではなく、音が聞こえてくる方向や音の広がり感など音場に関する情報—音空間情報—を伝達・再生することが極めて重要である。また、もし任意の音空間情報を高精度で伝送・再生することができれば、臨場感通信だけではなく、音楽ホールのような既存の空間や、設計途上のホールのようにまだ存在しない空間、更には、現実には存在しえないような空間の「音場」を、仮想的に実現することが可能となり、工学的な意義は大きい。

音空間情報の知覚については、音源から聴取者の外耳までの伝達関数—音像定位伝達関数—が重要であることが知られている。また、この伝達関数を実現するためのデジタル信号処理技術も発展しつつある。しかし、音像定位伝達関数と音空間情報の関係については不明な点が多く、このため、音像定位伝達関数の実現方法についても課題が多く残されている。

本研究は、このような観点から、音響通信における臨場感通信を行う上で極めて重要な役割を担っている音空間情報について、まず人間の聴覚における情報処理過程について考察し、合わせて音場情報の高精度な再生を目指したシステムの構築について述べる。

2 音場情報知覚の手がかりと制御

我々がある音を聞いたとき、水平方向と上下方向、それに距離を知覚することが可能であり、このため、その音が空間内のある場所に存在しているように感じられる。このため、視覚における像になぞらえて、聴覚的な像が生じていると考え、このような「音の聞こえてくる場所や範囲」を「音像」と呼ぶ。また、音像がある場所に存在するように知覚されることを、「音像定位」、あるいは単に「定位」という。3次元的な定位の結果知覚される空間のことを、音空間と呼ぶ [1]。一方、音像は定位知覚だけではなく、反射音や残響など音場の条件に応じて、広がり感や残響感の知覚をもたらす。音像定位に加え音場に関する知覚も総合したものを、音場知覚や音空間知覚などと呼ぶ。

2.1 音場情報知覚の手がかり

音場知覚・音空間知覚の手がかりには、骨導音や視覚情報など [2, 3]、両耳に入力される音とは無関係の要因もある。しかし本研究では、聴取者の両耳音圧に関連するものだけに限定して考察を進める。これらの要因には、両耳間音圧信号差、頭部の移動・回転による両耳音圧の変化、反射音や残響による情報などがあげられる。音源が正中面にある場合には、両耳間信号差は全く無いか、あってもごくわずかである。したがって、正中面における音像定位を説明するためには、両耳間信号差ではなく、音源の方向によって両耳音圧のスペクトルに現われる特徴的なパターンが重要である [4, 5]。

以上に述べた様々の両耳間差や、周波数スペクトル上の特徴は、空間内のある点から耳道までの伝達関数—音像定位伝達関数—という考え方を導入すれば統一的に表現できる。音像定位伝達関数は、反射音や残響など音場に起因する室内伝達関数と、頭部や外耳道などの回折や共振に起因する頭部伝

達関数 [6, 7, 8, 9] の従属接続で表現できる (図 1) . 音像定位伝達関数は, 音場知覚に関する音響的な手がかりを総合的に含んでいることから, これを模擬することによって音場知覚を制御することが可能である [10, 11, 12, 13, 14].

以上の研究は, 頭部が動かない場合の静的な頭部伝達関数の模擬による定位の制御を行ったものである. しかし後述するように, 音像定位には, 聴取者が頭を動かしたときの両耳の音圧変化 (音質変化) — すなわち頭部伝達関数の動的変化も重要な要因であることが明らかになってきた [13, 15, 16, 17, 18, 19].

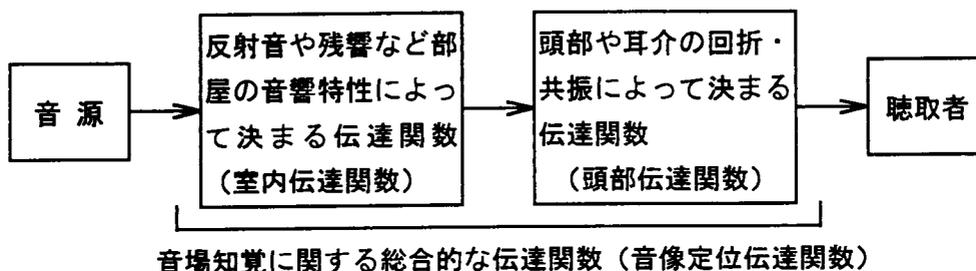


図 1: 音像定位伝達関数の 2 段モデル

2.2 音空間情報と音像定位伝達関数

任意の音像定位伝達関数を左右両耳に自由に再生・合成して与えることができれば, 原理的に任意の音場の再生が実現できることとなる. このため, ヘッドホンや少数 (通常 2 つ) のラウドスピーカによって任意の伝達関数を合成する試みがこれまで行われてきた. これらの手法は, 頭部の動きが無視できれば, 有効な音場情報の再生手法である. そこで, まず, これらの方法の原理について述べる. 実際の測定法, 及び伝達関数の模擬の手法の詳細は, 文献 [20, 21, 22] に述べてある.

2.2.1 ヘッドホン再生による音像定位伝達関数の合成

いま, 図 2 のようにヘッドホン再生によって, ある音場に置かれた 1 音源 (図ではラウドスピーカ) から耳までの伝達関数を仮想的に実現することを考える.

任意の位置に置かれたラウドスピーカの電気入力端子から, 聴取者の左右の鼓膜面上に置かれたマイクロホンの出力端までの伝達関数をそれぞれ $H_L(j\omega)$, $H_R(j\omega)$ とする. また, ヘッドホンを装着した聴取者の左右の変換器の電気入力端から, 左右の鼓膜面上に置かれたマイクロホンの出力端までの伝達関数をそれぞれ $G_L(j\omega)$, $G_R(j\omega)$ とする. ソース信号 $S(j\omega)$ をラウドスピーカによって再生したとき, 聴取者の両耳の鼓膜面上には, それぞれ次式で与えられる音圧が生じる.

$$\begin{aligned} P_L^{SP}(j\omega) &= H_L(j\omega)S(j\omega) \\ P_R^{SP}(j\omega) &= H_R(j\omega)S(j\omega) \end{aligned} \quad (1)$$

一方, ヘッドホンを装着した聴取者の左右の変換器に, それぞれ信号 $S_L(j\omega)$, $S_R(j\omega)$ を与えれば, 聴取者の両耳の鼓膜面上には, それぞれ次式で与えられる音圧が生じる.

$$\begin{aligned} P_L^{HP}(j\omega) &= G_L(j\omega)S_L(j\omega) \\ P_R^{HP}(j\omega) &= G_R(j\omega)S_R(j\omega) \end{aligned} \quad (2)$$

したがって、ヘッドホン再生時に、ラウドスピーカ再生時と等しい両耳音圧を与えるためには、ヘッドホンの左右の変換器入力端に、次式で表わされる信号を入力すればよい。

$$\begin{aligned} S_L(j\omega) &= G_L^{-1}(j\omega)H_L(j\omega)S(j\omega) \\ S_R(j\omega) &= G_R^{-1}(j\omega)H_R(j\omega)S(j\omega) \end{aligned} \quad (3)$$

ただし、 $G_L^{-1}(j\omega)$ 、 $G_R^{-1}(j\omega)$ は、ヘッドホンの変換器入力端から、鼓膜面上のマイクロホンまでの伝達関数の逆特性である。

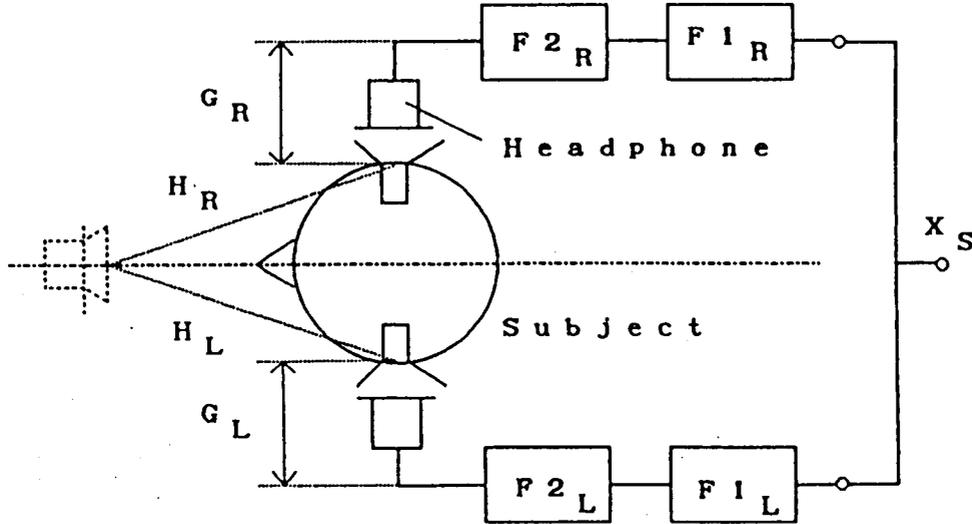


図 2: ステレオヘッドホン再生による任意の伝達関数合成の概念図

2.2.2 ラウドスピーカ再生による音像定位伝達関数の合成

ヘッドホン再生ではなく、図 3 のように空間内に配置した複数のラウドスピーカによって、別の位置の音源から両耳までの伝達関数を合成することを考える。この場合には、1つのラウドスピーカからの音が両方の耳に達するクロストークが存在するため、ヘッドホン再生の場合より複雑な処理が必要である。以下では、伝達関数合成用のラウドスピーカの個数が2の場合について述べる。

任意の位置に置かれた再生対象の音源としてのラウドスピーカの電気入力端子から、聴取者の左右の鼓膜面上に置かれたマイクロホンの出力端までの伝達関数をそれぞれ $H_L(j\omega)$ 、 $H_R(j\omega)$ とする。また、この音源による音像を再生するために用いる2つのラウドスピーカの電気入力端子から、聴取者の左右の鼓膜面上に置かれたマイクロホンの出力端までの伝達関数をそれぞれ $G_L^1(j\omega)$ 、 $G_R^1(j\omega)$ 、 $G_L^2(j\omega)$ 、 $G_R^2(j\omega)$ とする。

再生の対象として考えている音源が、実際に存在するときと等しい両耳音圧を、伝達関数合成用の2つのラウドスピーカによって与えるためには、ヘッドホン再生の場合と同様に考えることにより、変換器入力端に次式で表わされる信号を入力すればよいことが分る（ヘッドホン再生に比べて式が複雑になっているのは左右耳へのクロストーク成分が存在するためである）。

$$\begin{aligned} S^1(j\omega) &= (G_R^2(j\omega)H_L(j\omega) - G_L^2(j\omega)H_R(j\omega))S(j\omega)/X^{-1}(j\omega) \\ S^2(j\omega) &= (-G_R^1(j\omega)H_L(j\omega) + G_L^1(j\omega)H_R(j\omega))S(j\omega)/X^{-1}(j\omega) \end{aligned} \quad (4)$$

ただし、 $X^{-1}(j\omega) = 1/(G_L^1(j\omega)G_R^2(j\omega) - G_R^1(j\omega)G_L^2(j\omega))$

3 水平面音像定位制御の実現

考えている音場がデッドで、反射音が無視できる場合には、この音場の室内伝達関数がほぼインパルスであると仮定することができるので、音場情報の知覚は、頭部伝達関数による音像定位の問題に

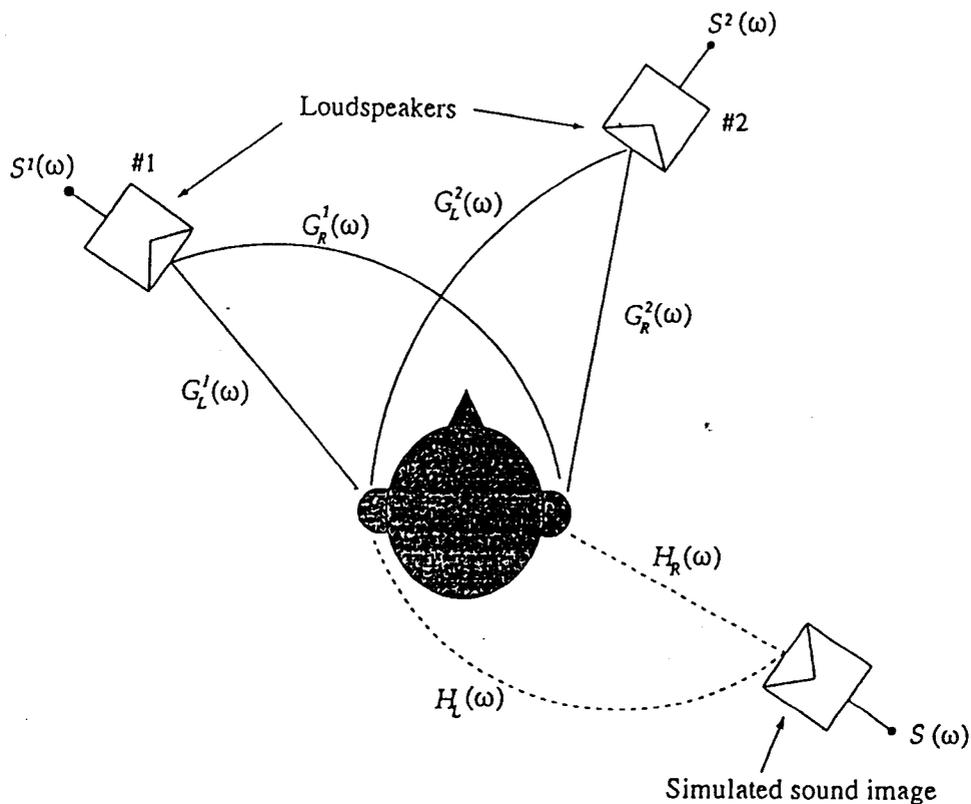


図 3: 複数スピーカ再生による任意の伝達関数合成の概念図

帰着される。

そこで本節では、前節で述べた伝達関数合成の手法を用いて水平面内の音像を作成し、静的な頭部伝達関数の合成のみを行った場合と、それに加えて頭部回転にともなう頭部伝達関数の動的な変化まで考慮した場合の2通りについて、頭部伝達関数の合成による音像定位に関する聴取実験を行った [13].

3.1 静的な伝達関数の合成による音像の定位

3.1.1 実験方法

被験者は、正常な聴力を有する男子3名である。合成した音源の方向は、被験者から見て正面を0度とする30度おきの12方向である。被験者の頭を中心から音源までの距離は1.5 mとし、音源と被験者の耳の高さが一致するように、被験者の座るいすの高さを調節した。実験に先立って、被験者全員について、ヘッドホン再生状態と合成音源として用いるラウドスピーカ再生状態における電気音響変換器の電気入力端から、外耳道入口におかれたプローブマイクロホンの出力端までの伝達関数を測定した。この伝達関数を、512タップのFIRデジタルフィルタによってヘッドホン経由で再現した。刺激音作成のソースとしては、ジャズ音楽とピンクノイズを用意した。

静的な伝達関数の合成のみを行う実験では、3人の伝達関数、2種類のソースの全てを使用して、刺激音を作成した。それぞれのソース音源について、12方向、伝達関数が3人分、繰り返し数が3で、刺激数は合計108ずつとなる。これらの刺激音をランダムな順で提示して、知覚された音像の位置を回答用紙に記入するよう被験者に求めた。なお、被験者の頭部は、ヘッドレストにより軽く固定した。

求めた回答の内容は、

- 音像の方向を、正面を0度とし、右側方を90度とする角度で記入。
- 音像の距離を、次に従って番号で記入。

1. 頭の中心付近

2. 頭の表面付近だが頭内
3. 頭外だが頭の付近
4. 手をのばせば届くくらいの距離
5. 模擬したラウドスピーカのあった程度 (1.5 m)
6. 模擬したラウドスピーカより遠い

• 音像の上昇の度合を、次に従って番号で記入

1. ほぼ水平面
2. やや上昇
3. かなり上昇

の3点である。

3.1.2 実験結果

ソースが音楽の場合について、全被験者の回答をまとめたものを、図4(a)～(d)に示す。ソースにピンクノイズを用いた場合も、結果に大きな差は見られなかった。

まず、図4(a), (b)は、方向の回答のみをまとめた結果である。回答は、 ± 15 度の範囲のものをまとめて集計してある。(a)図は、模擬に用いた伝達関数が自分自身のものである場合で、(b)図は他人の伝達関数を用いた場合である。ヘッドホン再生時に、ラウドスピーカ再生状態を模擬することによって、ほぼ全方向にわたっての音像定位が実現されているが、前半面のラウドスピーカを模擬した場合に、両耳を結ぶ線に対して対称な後方に音像が定位してしまう誤りが頻発している。また、この実験では、方向定位に関して、模擬に用いる伝達関数が自分自身のものか否かは、それほど大きな要因とはなっていない。

聴取実験の結果、ヘッドホン再生によって、ラウドスピーカ再生状態の伝達関数を模擬することにより、ほぼ全方向にわたっての音像定位が実現されることが示された。ただし、前半面のラウドスピーカを模擬した場合には、両耳を結ぶ線に対して対称な後方に音像が定位してしまう誤りが少なからず生じている。また、図4(a)と(b)を比較すれば、方向定位に関しては、模擬に用いる伝達関数が自分自身のものか否かが、それほど大きな要因とはなっていないことがわかる。

音像の距離感と、上昇の度合の回答をまとめた図4(c), (d)を見ると、頭外への定位はかなり達成されているものの、他人の伝達関数を用いた場合や、本人の伝達関数を用いても正中面付近の音像を模擬した場合には、距離感の減少と、音像上昇が頻発している。この距離感不足、音像上昇等の現象は、ダミーヘッドで収録して、ヘッドホン再生を行なうバイノーラル方式の実験などにおいても、しばしば報告されている。本人の頭部伝達関数が正確に与えられていないことが原因であるとする意見が多い。

3.2 頭部の動きに伴う頭部伝達関数の動的変化の模擬の効果

3.2.1 頭部の移動による頭部伝達関数の動的変化への対応

耳が左右に1対2個あることから、左右の判断は比較的容易である。これに対し、両耳を結ぶ線に対して対象な前後の位置の判定は、難しい。このため、我々は必要に応じて、頭部を動かして情報を得ている。頭部の運動は、回転・かしげ・うなずきに分解することができる。過去の研究によれば、実音源聴取時に聴取者の頭部の運動を許すと、前後誤判定率は有意に改善される。この場合、かしげ等の運動よりも、特に回転の効果が大きい[17]とされている。

そこで、頭部伝達関数の静的な模擬を正確に行うこととあわせて、頭部回転に伴う動的な変化をも模擬することの効果について検討した。このため、まず頭部の回転による頭部伝達関数の動的な変化を模擬するシステムを構築した。

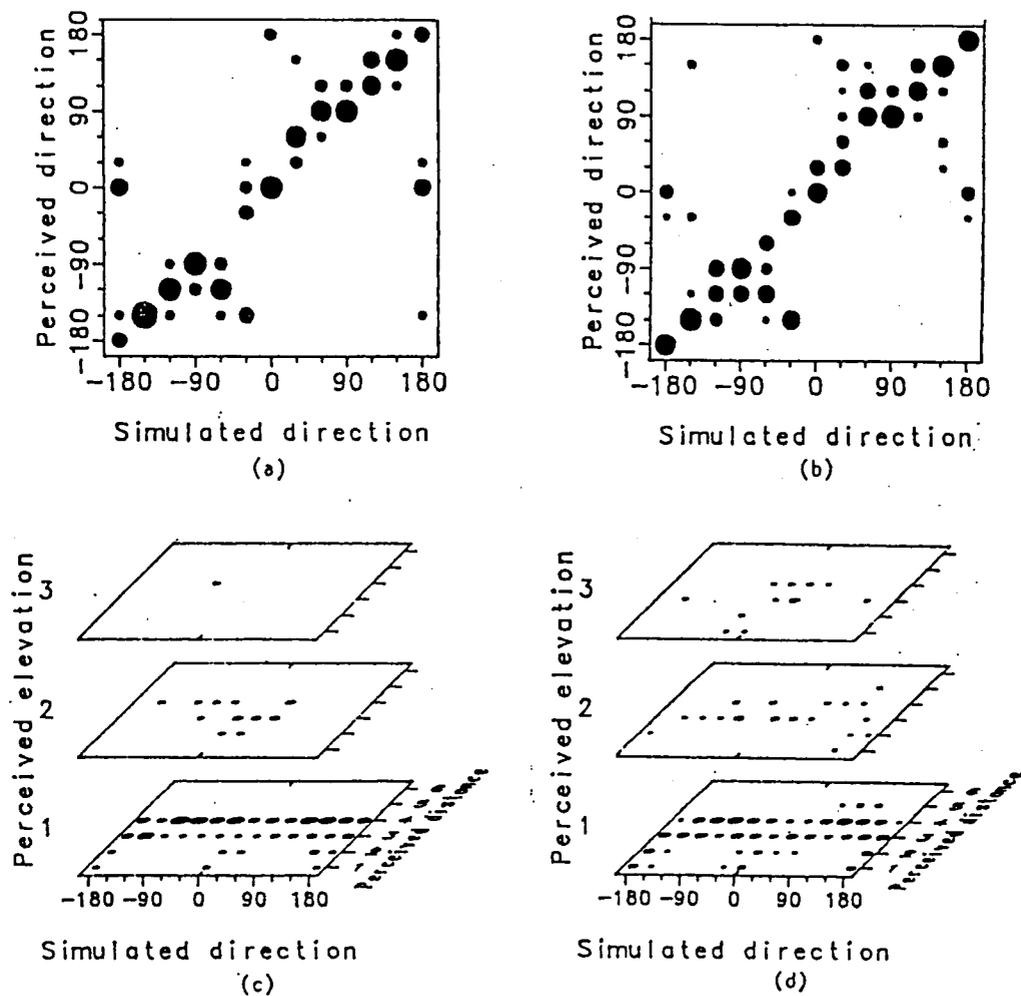


図 4: 音像から外耳までの静的な伝達関数を、ステレオヘッドホン再生によって合成したときの聴取実験結果. パネル (a) と (b) は知覚方向を, (c) と (f) は音像の上昇の程度と距離を示している.

本節で述べる検討では、回転に伴う頭部伝達関数の変化は、回転角度が微小であるとして、位相項のみを考えることにし、両耳間到達時間差の変化を考慮の対象とした。そこで、図5のような実験装置を製作した[13]。聴取者の頭部は、いすにとりつけた治具に固定され、水平回転運動のみが可能になっている。頭部の方向角は、治具のポテンショメータによって比例する電圧に変換され、さらに、アナログ演算回路、V/F変換器などによって、方向角に応じた両耳間時間差に変換される。ここで問題となるのは、音源位置が聴取者の前方か後方かによって、時間差変化率の符号が異なることである。スイッチSw1は、頭部の回転方向と、各チャンネルの到達時間遅れの増減の関係を切り替えるものである。スイッチSw1を「前方モード」に設定すると、聴取者の頭部が時計回りに回転したときに、左チャンネルの時間遅れが減少し、右チャンネルの時間遅れが増加する。したがって、「前方モード」では、聴取者の頭部の回転にともない、音源方向が聴取者前方にあるという「動的」情報が与えられる。逆に、Sw1を「後方モード」に設定したときは、両チャンネルの時間遅れの増減の関係が逆転し、音源方向が後方であるという情報が与えられる。

到達時間変化を直線近似しているため、聴取者側方に音源がある場合には、両耳間到達時間差の模擬はできないが、音源方向が正中面付近のときには実際の到達時間変化に近い模擬が行える。

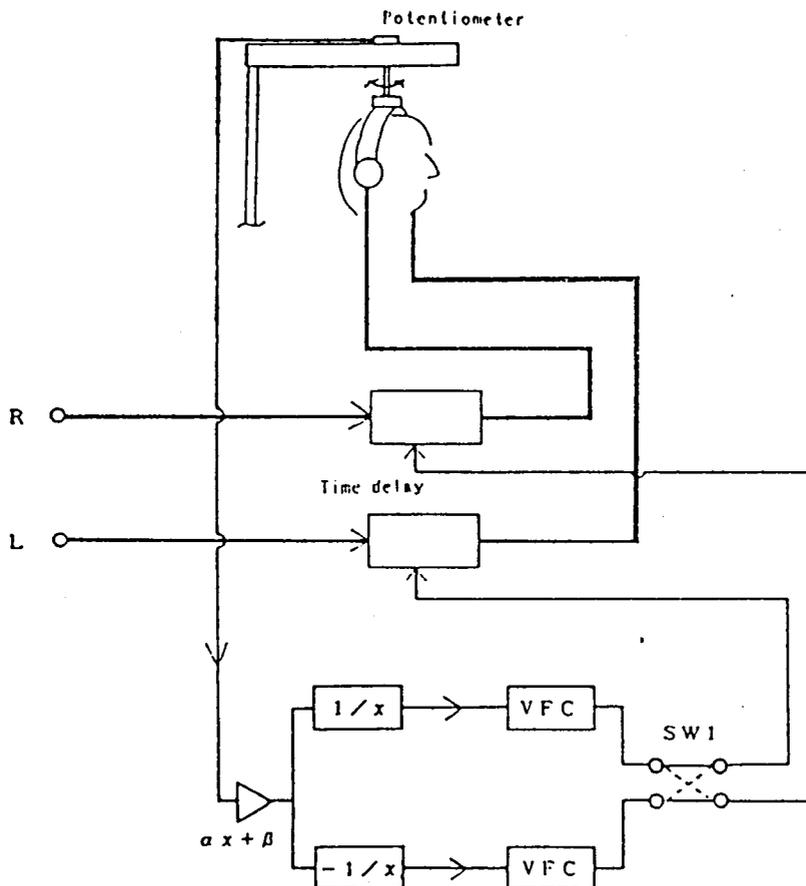


図 5: 頭部の回転に伴う頭部伝達関数の動的変化を、両耳間到達時間差の変化として模擬するための装置の概念図

3.2.2 実験方法

前述のように頭部の動きの中でも動的変化の効果が大きい回転の効果に関する実験を行った。用いた伝達関数は、被験者3名のうち1名のもののみで、被験者2のものである。ソース信号は、ジャズ音楽だけを用いた。

音源方向が被験者の側方となる場合には、頭部回転の効果が小さくなるうえに、到達時間差変化と

頭部回転角が直線関係でなくなるため、真正面とその ± 30 度、真後とその ± 30 度の計6方向についてのみ模擬を行なった。両耳到達時間変化模擬装置の、前述した「前方モード」、「後方モード」のそれぞれについて、合計24の刺激音を、ランダムな順序で3回ずつ提示し、前記実験と同様に、知覚された音像の方向、距離、上昇の度合について回答を求めた。被験者には、刺激音提示中は目をつぶり、頭を最大で ± 30 度程度まで自由に回転させ、音像の位置が安定したときの回答を記入するように指示を与えた。実験は同じ条件について、4回繰り返して行った。

3.2.3 実験結果

図6(a)～(c)に、それぞれ、被験者1～3の回答をまとめたものを示す。図の、縦の4段のうち、上2段が前述の「前方モード」に、下2段が「後方モード」に設定した場合の回答に対応する。

それぞれのモードの回答のうち、上段は、被験者前方のラウドスピーカ再生を模擬した場合で、下段は、後方を模擬した場合である。図中の右下の数字、及び、頭の中心から伸びている一点鎖線は、模擬したラウドスピーカの方向を表す。黒丸は、知覚された音像の方向及び距離を表し、その面積は度数に比例している。また、右上の数値は、知覚された音像の上昇度合の平均を示したもので、1が「ほぼ水平」、2が「やや上昇」、3が「かなり上昇」を表す。

この実験においては、音源位置が被験者の前方にあるか後方にあるかの情報が、頭部伝達関数による静的な情報と、頭部の回転に伴って生じる動的な情報の2種類の形で、聴取者に与えられる。この2種類の情報は、独立に制御することが可能である。したがって、図6の最上段および再下段では、音源位置が前方か後方かの情報は、静的なものとの間で一致しているが、2段目と3段目では、両者は矛盾した設定となっている。

2種類の前後情報が一致して与えられた場合の結果を見ると、静的な伝達関数のみを模擬したときに比べて、前後の誤判定がなくなり、音像の上昇の頻度も大きく減少していることがわかる。

一方、静的な伝達関数からの前後情報と、頭部回転に伴う動的な伝達関数の変化からの前後情報が一致せず、矛盾を生じている場合の結果を見ると、どの被験者の場合も、頭内定位や、音像上昇が頻発していることがわかる。被験者からの報告では、この場合の音像は、頭にへばりつくように定位することが多い。また、音源方向の判断の仕方には、個人差が見られる。被験者1、3では、音源が前方か後方かの決定が、ほぼ両耳間到達時間差によって行なわれているのに対し、被験者2の場合は、常に後方に定位してしまっている。

この結果は、音源位置が前方か後方かを判定するためには、頭部伝達関数等の静的な両耳音圧の特徴からの情報のみでは不十分であり、頭部を回転・移動したときの両耳音圧の動的な変化の様子からの情報が、静的な情報以上に重要であることを示している。このような動的な変化を考慮しなければ、前後の反転、音像上昇などの不自然な現象は避けられない。したがって、高精度の音場合成・再生システムを構築してゆくためには、頭部の回転・移動に伴う伝達関数の動的な変化までを考慮することが必須であることが分かる。

4 正中面音像定位制御の実現

4.1 正中面定位におけるスペクトラルキューの重要性

3次元の音場知覚を考えてゆくうえで、水平面に次いで重要なのは、仰角にかかわる正中面（真正面と真後を結ぶ鉛直面）の音像定位である。この正中面定位においては、水平面定位で重要な役割を果たす頭部伝達関数の両耳間差がほとんど存在しない。このため、伝達関数の両耳間差より、伝達関数のそのものの特性が重要な役割を果たしている。これは、たとえば上方からの伝達関数には3 kHz付近にピークがあり、このようなピークやディップを手がかりに音像が情報であるという定位を行っているというようなことである。これをスペクトラルキューという。

仰角について所望の定位を実現するためには、スペクトラルキューを正しく再生できるよう、頭部伝達関数を全帯域について十分な精度で合成する必要がある。ここで、スペクトルの形状の弁別につい

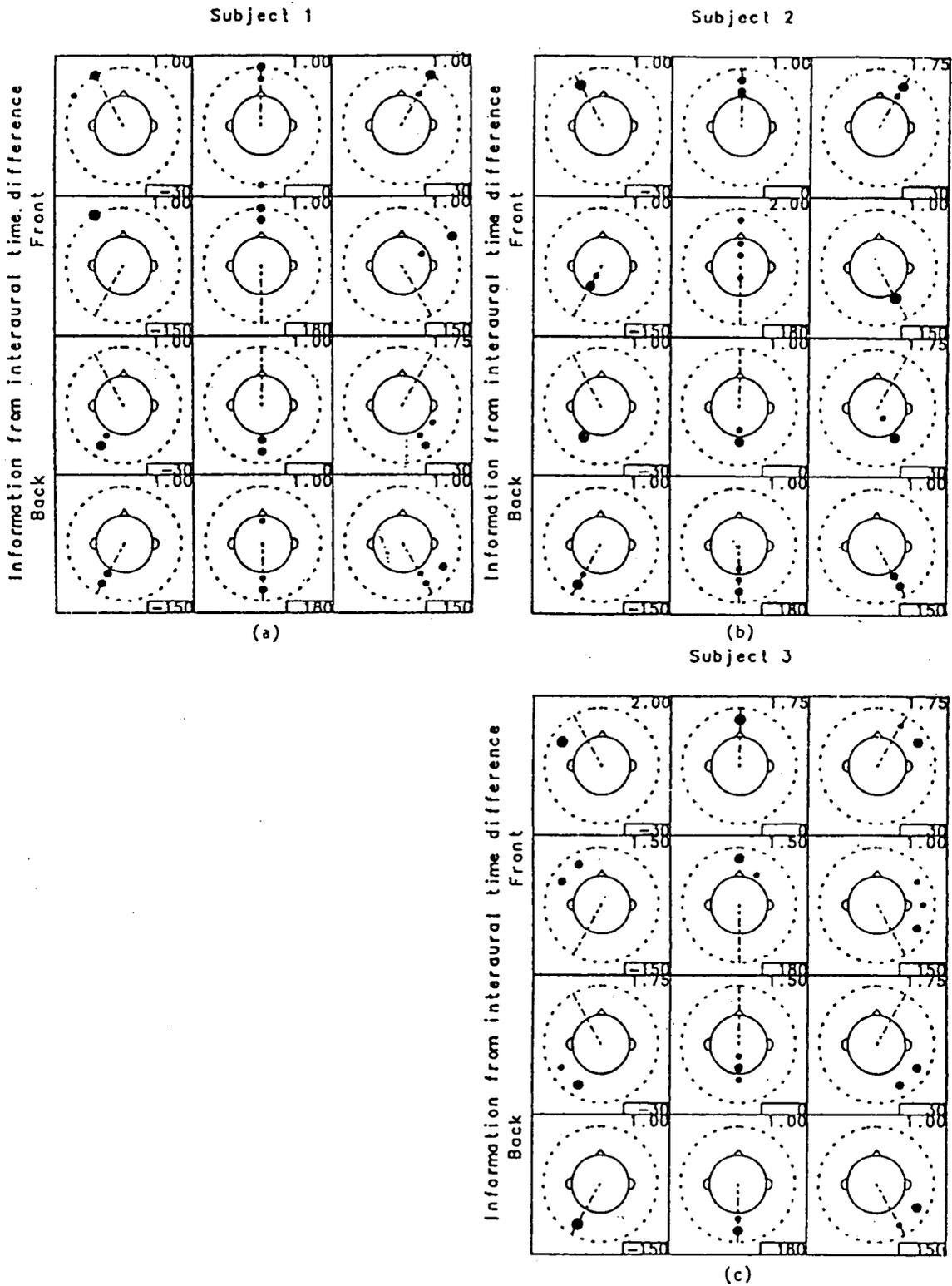


図 6: 音像から外耳までの静的な伝達関数に加えて、頭部の回転による伝達関数の動的な変化までを考慮して、ステレオヘッドホン再生によって頭部伝達関数を合成したときの聴取実験結果。黒丸は音像の定位方向と距離を示している。丸の面積は被験者の判断の頻度に比例している。各パネルの右上の数字は、音像の上昇の程度を示している。

て人間の聴覚系はさほど狭帯域な周波数分解能を持っていないことを考えれば、「十分な精度」がどの程度の精度なのかを検討することが工学的に重要である。

そこで第2節で述べた伝達関数模擬の手法を用いて、正中面内音像の刺激音を作成し、聴取実験を行った。ここでは特に、伝達関数をできる限り精密に合成した場合と、ARMAモデルを用いて伝達関数を簡略化し概形だけの合成を行った場合の比較によって、正中面内の定位に果たすスペクトラルキューの役割について考察した [14]。

4.2 実験方法

合成した音源の方向は、被験者から見て正面(0度)から真後(180度)まで10度おきに上方の18度方向である。被験者の頭の中心から音源までの距離は1.5 mとした。刺激音には、持続時間1秒の白色雑音を用いた。被験者は、正常な聴力を有する男子3名である。刺激音をランダムな順で7秒間隔で提示した。これ以外の実験方法は、前節の水平面定位に関する実験と同様である。

頭部伝達関数の測定方法自体は、水平面の実験の場合と同様である。ただし、この伝達関数をヘッドホン経由で再現する場合に、伝達関数を次の4通りの方法で合成した。

1. 512タップのFIRフィルタでできる限り精密に合成。
2. 全帯域をARMAモデルで簡略化して合成。
3. 5 kHz以上の周波数だけをARMAモデルで簡略化して合成。
4. ARMAモデルを適用する周波数を、5 kHz以下500 Hzまで下げて合成。

4.3 頭部伝達関数の合成精度と正中面定位

4.3.1 精密な合成と、全帯域簡略化

図7は、頭部伝達関数をできる限り精密に合成した場合と、全帯域をARMAモデルで簡略化した場合の頭部伝達関数の1例である。ここで40/40、30/30のように書かれているのは、ARMAモデルの次数である。たとえば40/40とは、伝達関数を40個の極と40個の零点を用いて表現したことを示している。図8は、頭部伝達関数をできる限り精密に合成した場合と、全帯域をARMAモデルで簡略化した場合の定位実験の結果である。図で、一番上の行が伝達関数を精密に合成した場合、2～5行はARMAモデルで簡略化した場合の結果である。左の2列は、被験者1の伝達関数を合成したときの実験結果、右の2列は、被験者2の伝達関数を合成したときの結果である。図の最上部には聴取した被験者が記されている。各パネルの横軸は意図した定位方向、縦軸は実際に知覚された方向である。

図8をみると、まず最上行に示されている簡略化しない場合には、ほぼ意図どおりの正中面音像の定位制御が行われていることが分る。ただし、被験者2の伝達関数を用いた場合には、特に被験者1の聴取結果に、前後判定誤りが生じている。ARMAモデルを用いて伝達関数の簡略化を行った場合には、ほぼ30/30程度以下の次数になると、前後判定誤りが頻発する。

図7をみると、ARMAモデルによる伝達関数簡略化の影響は特に低周波数領域で顕著であることが見て取れる。そこで、ある境界周波数を設けて、それより高周波数の領域だけにARMAモデルを適用して伝達関数の合成を行った場合の正中面内音像定位について検討した。

4.3.2 高周波数領域だけの伝達関数簡略化

図9は、5 kHzを境界周波数とし、これ以上の周波数領域だけを簡略化したときの聴取実験の結果である。この図から、簡略化を高周波数領域に限った場合には、ARMAモデルの次数を10/10まで減らしても極めて高精度な正中面音像の定位制御が可能であることが分る。

では、この5 kHzという境界周波数はどこまで下げることができるのだろうか。

図 10は、ARMA モデルの次数は 10/10 に保ったまま、境界周波数を 500 Hz まで下げてみたときの聴取実験の結果である。この図から、境界周波数がおよそ 2 kHz 付近の場合が臨界領域となっていることが分る。境界周波数がこれより低いと前後の判定誤りが頻発し、高ければ高精度の定位制御が行われている。

以上の検討結果から、正中面定位におけるスペクトルキューは、およそ 2 kHz 以上ではおおまかな山谷の知覚に基づいており、低域ではスペクトルの微細構造にも基づいていることが推定される。

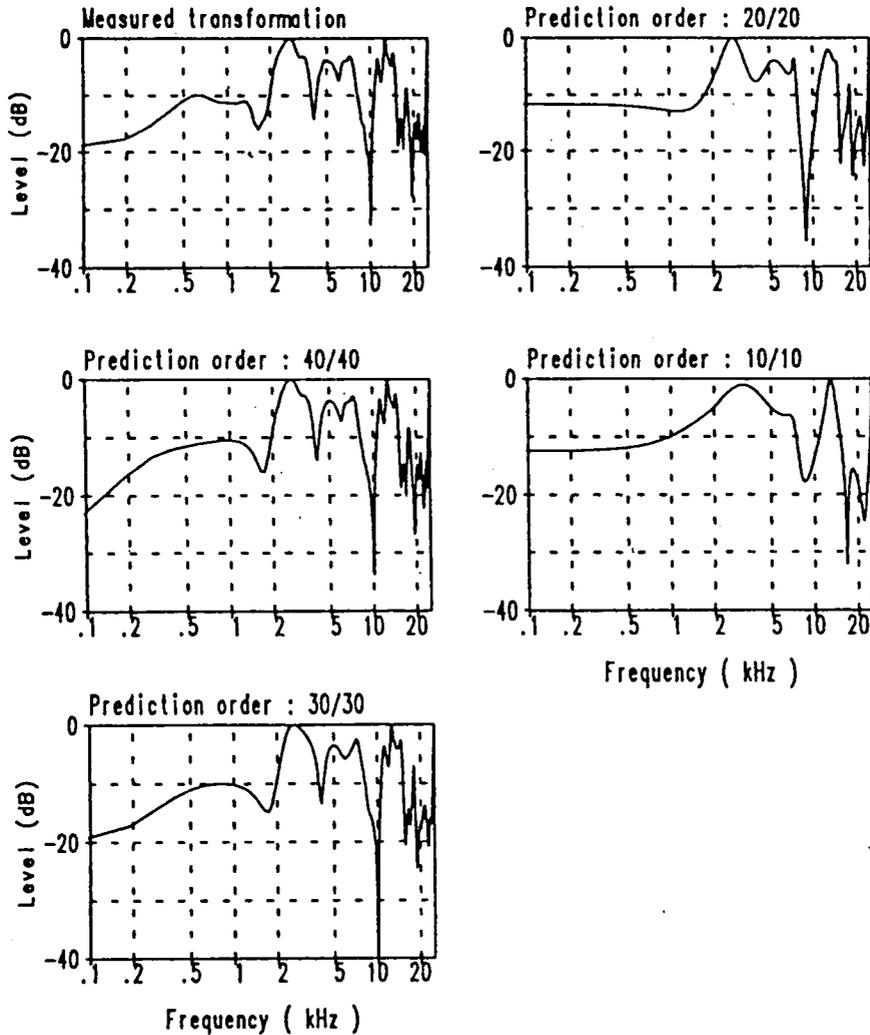
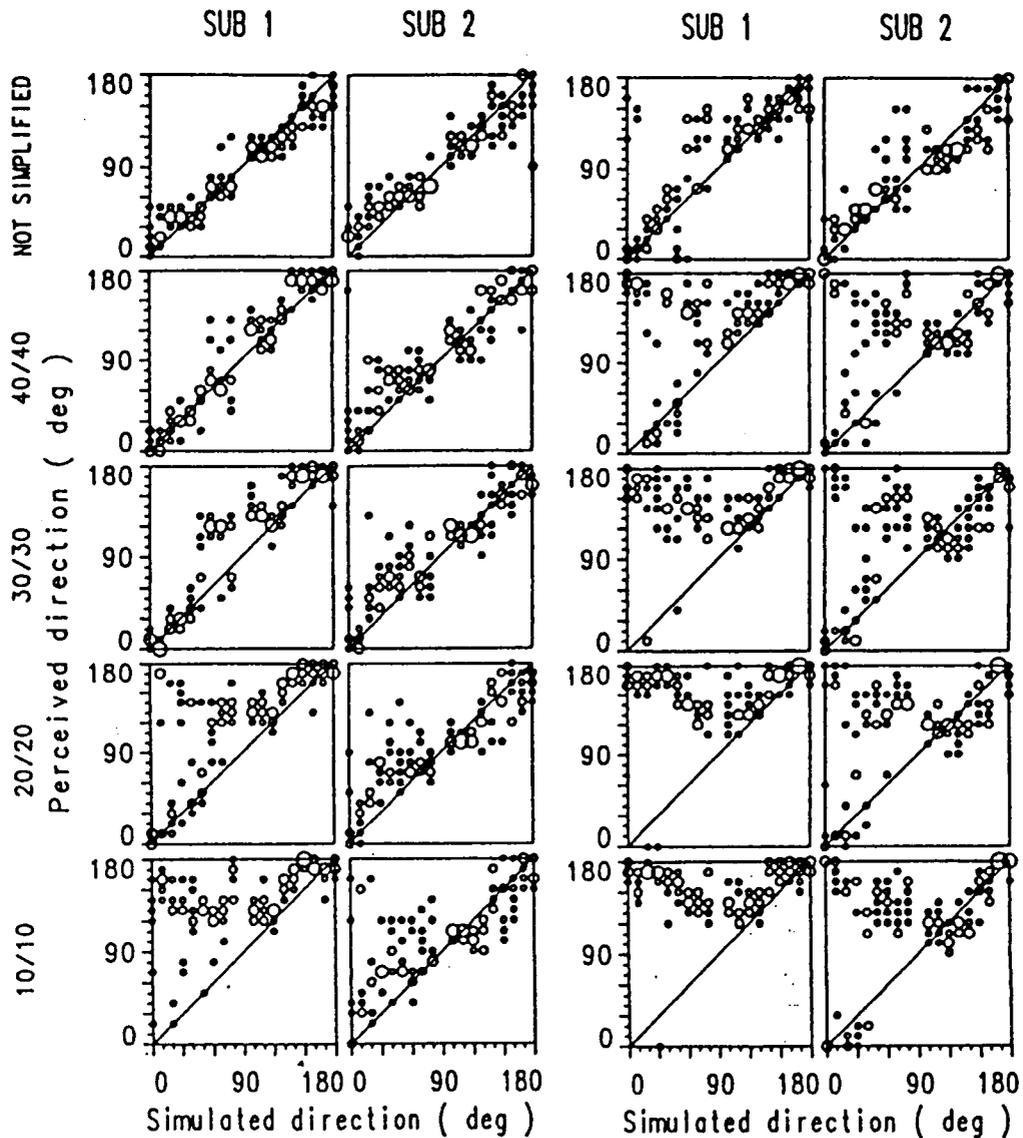


図 7: 頭部伝達関数をできる限り精密に模擬した場合と、全帯域を ARMA モデルで簡略化した場合の正中面頭部伝達関数の 1 例。

5 音場情報の高精度再生手法

以上の考察から、音源から耳までの伝達関数—音像定位伝達関数—を用いて音場情報の知覚を制御するためには、頭部の動きに伴う音像定位伝達関数の動的な変動を正確に再生することと、伝達関数を特に低周波数域において精密に再生することが重要であることが明らかになった。

そこで、本節ではこれらの知見に基づいて、音像定位伝達関数を精密に再生することにより、高度な音空間再生能力を持つシステムの構成について考察する。



(a) Transfer function of SUB1 (b) Transfer function of SUB2

図 8: 頭部伝達関数をできる限り精密に模擬した場合と、全帯域を ARMA モデルで簡略化した場合の頭部伝達関数を用いた、正中面方向音像定位の結果. 各パネルの横軸は意図した定位方向で、縦軸は実際に知覚された方向. 40/40, 30/30 等は、ARMA モデルの次数.

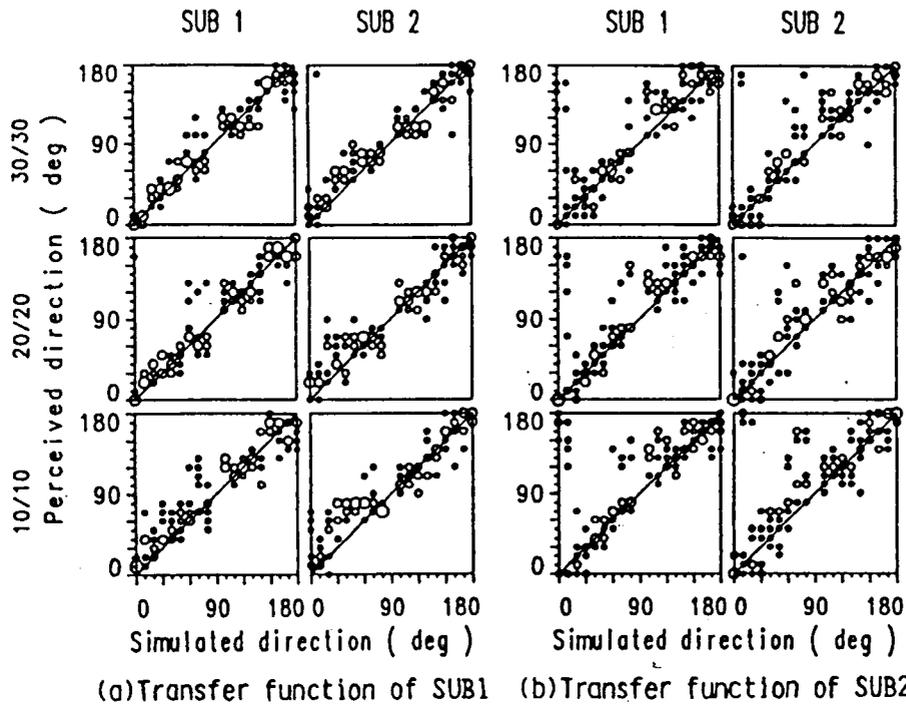


図 9: 5 kHz を境界周波数とし、これ以上の周波数領域だけを簡略化して伝達関数を合成したときの正中面合成音像聴取実験の結果。

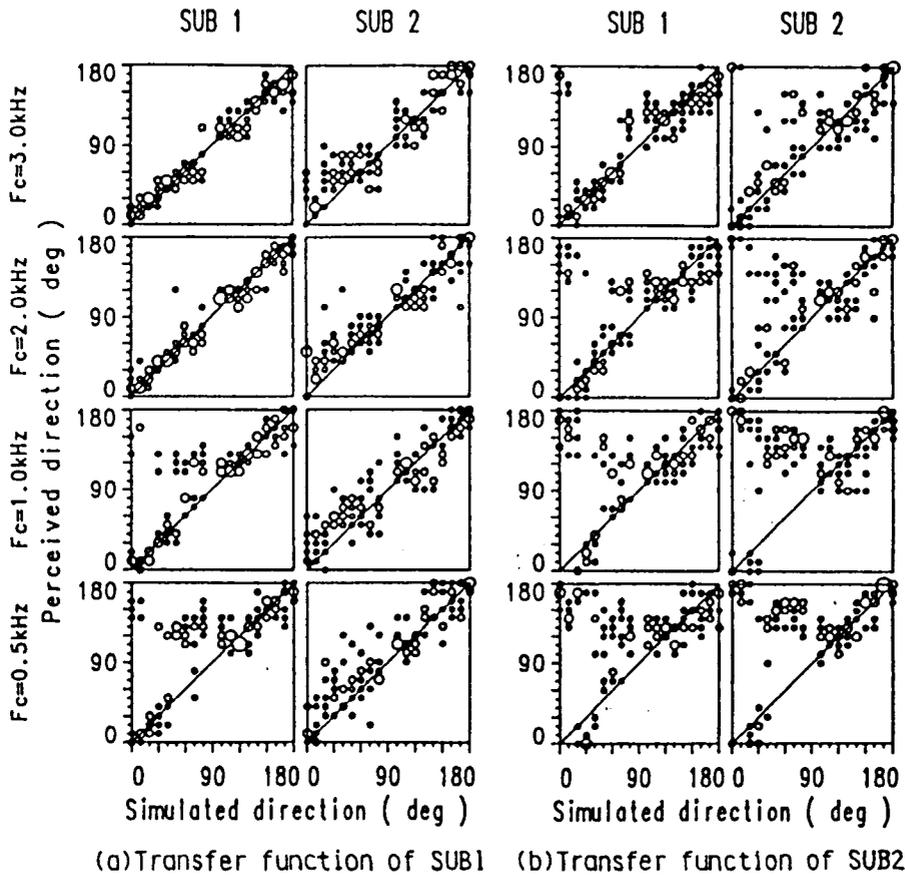


図 10: ARMAモデルの次数は 10/10 に保ったまま、境界周波数を 500 Hz まで下げて伝達関数を合成したときの正中面合成音像聴取実験の結果。

5.1 従来提案された方法

通常のステレオ方式や、映画などで用いられているドルビーサラウンドは、音場情報の概要を伝送・再生するための手法であり、音場や音空間を精密に再生することを意図したものではない。精密に音場を再生することを目指した音場合成・再生方法は、次の2種類に大別される。

多数音源合成法 一つは多数音源音場合成法とよばれるもので、この方法は原音場の一点における直接音および全ての反射音、残響音を多数のスピーカを用いて模擬する方法である [23]。制御用スピーカは受音者にとって最も重要な反射音の入力方向に配置する必要がある、そのほかに多くのスピーカを拡散、残響音場用として配置しなければならないので、忠実な音場の再生には比較的多数の電気音響のチャンネル数が必要である。また、原音場を、直接音、反射音、残響音として表現するため、それぞれを分離しての測定が困難であるなどの問題点があり、音像定位伝達関数として統一的な取り扱いには不向きである。

伝達関数合成法 もう一つのアプローチとして、音場の合成・再生を、単一の音像の定位制御の拡張であるにとらえ、比較的小数のスピーカやヘッドホンを用いて音像定位伝達関数（言い換えれば人間の鼓膜位置の音）をできるだけ忠実に合成・再生する方法がある [10]。

この手法では、反射音や残響といった音場のパラメータが、音源から聴取位置までの伝達関数で与えられる。このため、反射音を明確には分離できない実際の測定からでも、計算機シミュレーションの結果からでも、統一的に与えることが可能である。特に、近年大きな進展を見せている、低周波数域における音の波動性も考慮した精密な音場解析手法とも相性のよい方法であり、現状の性能や今後の発展の点から優れた方法である。

しかし、この方法では音像定位伝達関数が反射音と残響を含んだものとなるため、そのインパルス応答が長くなる。例えば、サンプリング周波数が (CDと同じ)44.1 kHz で、残響時間が2秒の空間の情報を FIR 型のデジタルフィルタで再生するためには、9万点近いタップ数が必要となる。頭部の動きに対応して伝達関数を動的に制御するためには、このようにタップ数が極めて長いデジタルフィルタを多数備える必要があるため、システムとして実現するのは極めて困難である。

5.2 提案する方法

そこで、本研究では、後者の伝達関数合成法の長所を残しながら、頭部の動きに追隨して音源定位伝達関数を動的に再生することが可能なシステムを、合理的なシステムサイズで実現する手法を考案した。

このシステムの基本的な考え方は、音像定位伝達関数が、図 1 のように室内の音場条件による室内音響伝達関数と、聴取者の頭部や耳介によって決定される頭部伝達関数の縦続接続で表現されることに基づいている。すなわち、図 11 のように、聴取者を中心とする仮想的な球を考え、実現すべき音像定位伝達関数を、この仮想球の外側と内側とに分割して考える。

キルヒホッフの積分定理から、ある閉曲面上の音圧と音圧傾度が定まれば、この閉曲面内の任意の位置の音圧を規定することができる。したがって、上記の仮想球の上で十分な密度で音圧と音圧傾度を求めてやれば、この球の内側の全ての点における音圧が定まることになる。

例えば音楽ホールの場合について考えてみよう。ステージから仮想球の球面上の多数の点までの伝達関数（室内伝達関数）を予め測定あるいは算出しておく。この伝達関数を球面上に十分な密度で再生することができれば、球内の全ての点で所望の音場が得られる。しかし、実際に球面を作成し、そこに多数の音源を配置すれば、これは前述の多数音源合成法とほぼ同様のシステムとなり、システムの規模も極めて大きくなってしまふ。

そこで、この仮想球までの伝達関数に、仮想球上の点から左右両耳までの伝達関数（頭部伝達関数）を縦続接続する。このような構造を取ることにより、仮想球の範囲であれば頭部が移動しても、複雑な構造を持つ室内伝達関数は変化せず、構造が単純な頭部伝達関数の変化だけに追隨すればよいこと

になる。また、このときの頭部伝達関数の再生には、第2節で述べたように、ヘッドホンを用いる方法と、少数のスピーカを用いる方法を用いることができるので、システムの規模を小さく抑えることが可能となる。

我々は、このような伝達関数の合成法を、仮想球モデルと名付け、図12に示すシステムとして実現した。サンプリング周波数は32 kHzである。したがって、ステージから仮想球上の各点までの室内伝達関数は、例えば残響時間が1秒の部屋の伝達関数を表現する場合には、32000点という長いインパルス応答で表現される。このインパルス応答を、FIRフィルタによって予めソース信号(音楽など)と畳み込み演算によって合成しておくこと。次に、仮想球上の各点における音に頭部伝達関数を合成すれば、この出力が最終的な出力音となる。頭部伝達関数の合成後の音の提示は、左右のクロストークがなく高精度の再生が可能であることから、ヘッドホンを用いて行うこととした。頭部伝達関数は128点のインパルス応答(時間長4 ms)で表わした。頭の運動を検出するため、ヘッドホンのヘッドバンドに小型のジャイロを装着した。現在は、水平運動は考慮せず、回転、うなづき、かしげの3自由度に対応している。頭部が動くと、左右の両耳から仮想球上のある点を見込んだ角度が変化する。そこで、この運動に対応して、使用する頭部伝達関数を動的に変化させる。このため、仮想球上から左右両耳までの伝達関数は、予め水平方向については5度おきに、仰角方向については10度おきに測定して、測定点の間では直線補間を行っている。

今後は、このシステムを用いて、臨場感通信のための高次音場情報の再生手法の検討を行うと共に、頭部の動的な動きが音空間知覚に及ぼす効果・影響について更に詳しく検討してゆく予定である。

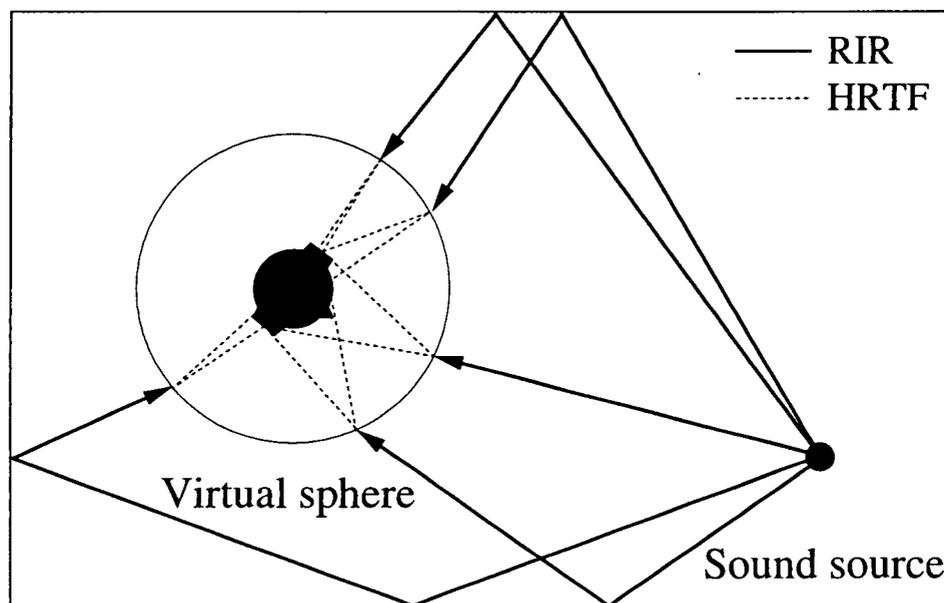


図 11: 聴取者のまわりに仮想球を置き、音像定位伝達関数を2段階でモデル化する概念図。

6 まとめ

音源から聴取者までの伝達関数を合成・再生することによって、高精度の音場知覚制御が可能となる。特に、静的な伝達関数の合成だけでなく、頭部の移動を考慮した伝達関数の動的な変化に追従することが、自然な音場情報の再生に極めて重要であることを明らかにした。また、正中面の定位については、頭部伝達関数の両耳間差がほとんど存在しないため、伝達関数のそのものに含まれる特徴(スペクトラルキュー)の特性が重要な手がかりの役割を果たしている。正中面定位におけるスペクトラルキューは、およそ2 kHz以上ではおおまかな山谷の知覚に基づいており、低域ではスペクトルの微細構造にも基づいていることを示した。

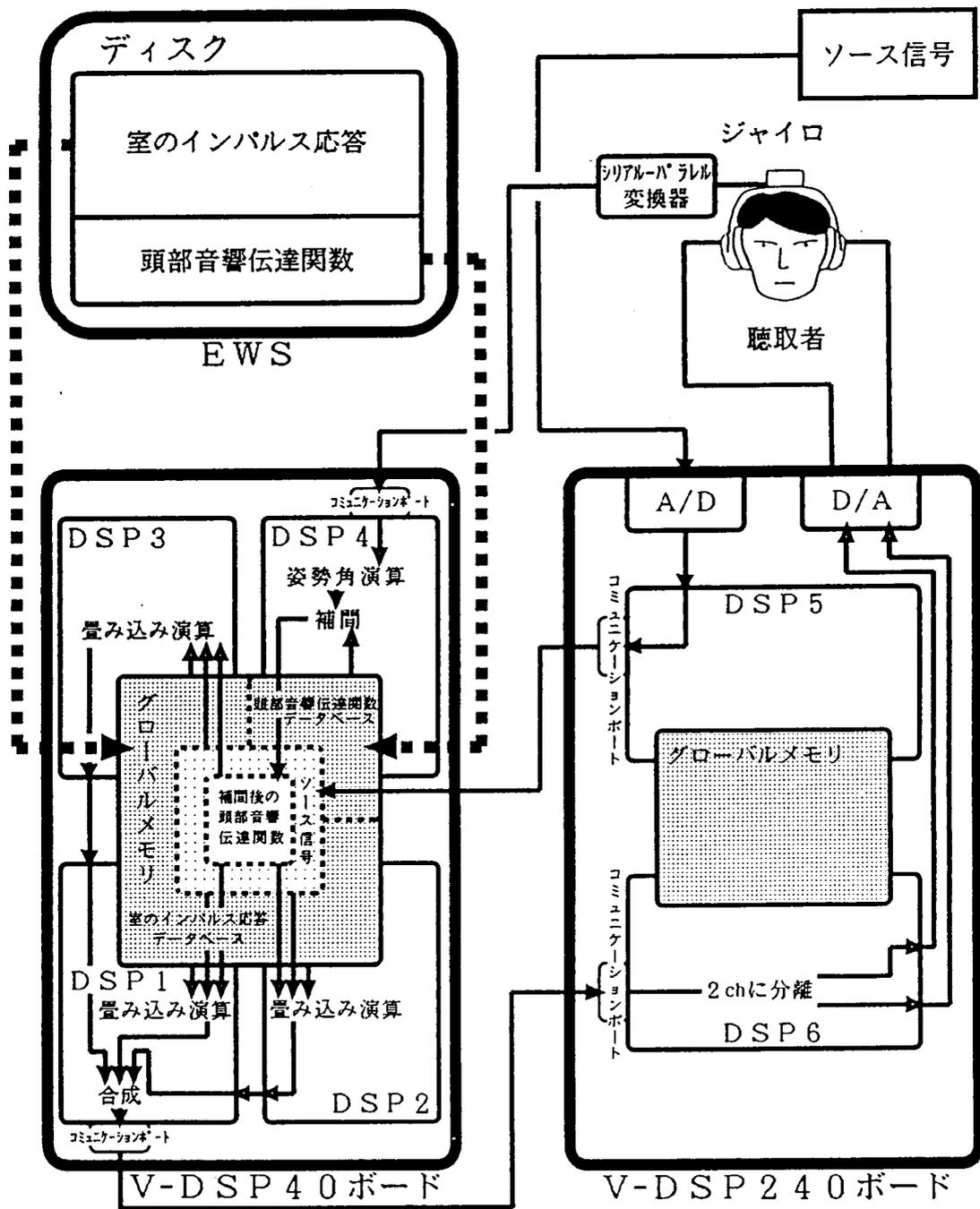


図 12: 仮想球モデルに基づく音像定位伝達関数の合成を行なう、音空間情報再生システムのブロック図。

以上の検討を元に、任意の音場についての音像定位伝達関数を、頭部の動きも考慮して動的に再生することが可能な仮想球モデルに基づくシステムの構成法について述べた。この方法は、空間中に聴取者を取り囲む仮想的な球を設け、この外側ではキルヒホッフの積分公式に基づく静的な伝達関数の合成を行い、内側では頭部の動きを考慮した動的な合成を行うというものである。

謝辞 一連の実験は、東北大学大学院工学研究科の大学院生であった川浦淳一氏（現新日鐵）、浅野太氏（現電総研）の協力の下で行われたものである。また、音空間再生システムの構築と本稿の準備に当たっては、東北大学電気通信研究所曾根研究室の高根昭一氏、新井大輔氏、清水建設宮島徹氏の協力を得た。ここに記して、深く感謝する。

参考文献

- [1] J. Blauert, 後藤, 森本, "空間音響", 鹿島出版会
- [2] T. Sone, M. Ebata, and T. Nimura, "On the Difference Between Localization and Lateralization", 6th Intl. Cong. Acoust., A-3-6, (1968).
- [3] 中林克己, 二階堂誠也, "音像, 映像の相互作用に関する基礎実験", 聴覚研資, H-57-1, (1979).
- [4] S. K. Roffler and R. A. Butler, "Factors that Influence the localization of Sound in the Vertical Plane", J. Acoust. Soc. Am. 43, 1255-1259, (1968).
- [5] J. Blauert, "Sound Localization in the Median Plane", *Acustica* 22, 205-213, (1970).
- [6] P. Damaske and B. Wagener, "Richtungshorversuche uber einen nachgebildeten Kopf", *Acustica* 21, 30-35, (1969).
- [7] E. A. G. Shaw, "Transformation of Sound Pressure Level from the Free Field to the Eardrum in the Horizontal Plane", J. Acoust. Soc. Am. 56, 1848-1861, (1974).
- [8] S. Mehrgardt and V. Mellert, "Transformation Characteristics of the External Human Ear", J. Acoust. Soc. Am. 61, 1567-1576, (1977).
- [9] 黒沢明, 都木徹, 山口善司, "頭部伝達関数と方向弁別能力について", 音学誌 38, 145-149 (1981).
- [10] M. R. Schroeder and B.S. Atal, "Computer Simulation of Sound Transmission in Rooms", IEEE Int. Conv. Rec. 7, 150-155, (1963).
- [11] 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", 音学誌 29, 151-159, (1973).
- [12] M. Morimoto and Y. Ando, "On the simulation of sound localization", J. Acoust. Soc. Jpn. (E) 1, 167-174, (1980).
- [13] 川浦淳一, 鈴木陽一, 浅野太, 曾根敏夫, "頭部伝達関数の模擬によるヘッドホン再生音場の定位", 音学誌, 45, 756-766, (1989).
- [14] F. Asano, Y. Suzuki, and T. Sone, "Role of spectral cues in median plane localization," J. Acoust. Soc. Am. 88, 159-168, (1990).
- [15] K. de Boer and A. TH. van Urk., "Some Particulars of Directional Hearing", Philips Tech. Rev. 6, 12, 359-364, (1941).

- [16] D. M. Leakey, "Some Measurement on the Effects of Interchannel Intensity and Time Differences in Two Channel Sound Systems", *J. Acoust. Soc. Am.* 31, 977-986, (1959).
- [17] W. R. Thurlow and P.S. Runge, "Effect of Induced Head Movements on Localization of Direction of Sounds", *J. Acoust. Soc. Am.* 42, 480-488, (1967).
- [18] G. Boerger, P.Laws and J.Blauert "Stereophone Kopfhörerwiedergabe mit Steuerung bestimmter Übertragungsfaktoren durch Kopfdrehbewegungen", *Acustica* 39, 22-26, (1977).
- [19] 朝日伸光, 青山寛, 松岡進, "自然な音像定位を再現するヘッドホン受聴システム", *信学技報*, EA79-24, (1979).
- [20] 川浦淳一, 鈴木陽一, 曾根敏夫, 相馬次郎, "デジタル信号処理を用いた音響系の模擬法について", *音学誌* 42, 774-779, (1986).
- [21] 鈴木陽一, 浅野太, 曾根敏夫, "音響系の伝達関数の模擬をめぐって(その1)," *音学誌* 44, 936-942(1988)
- [22] 鈴木陽一, 浅野太, 曾根敏夫, "音響系の伝達関数の模擬をめぐって(その2)," *音学誌* 45, 44-50(1989)
- [23] E. Meyer, W. Burgtorf, and P. Damaske "Eine Apparatur zur elektroakustischen Nachbildung von Schallfeldern. Subjective Hörwirkungen beim Übergang Kohärenz-Inkohärenz," *Acustica* 15, 339-344, (1965)