

映像を伴う多チャネルステレオに関する研究

小宮山 撰



①

映像を伴う多チャンネルステレオに関する研究

小宮山 撰

1章 序論	1
1.1 本研究の位置づけ	1
1.2 論文の構成	2
2章 テレビ視聴時における視覚と聴覚の相互作用に関する考察	4
2.1 緒言	4
2.2 視覚刺激と音像定位の関係に関する従来の知見	5
2.2.1 感覚間相互作用の実験例	5
2.2.2 従来の実験の問題点と本研究の意義	6
2.3 テレビ映像による音像方向の偏移に関する実験的考察	7
2.3.1 実験的考察の概要	7
2.3.2 実験装置の配置と実験方法	7
2.3.3 映像による音像方向の偏移の実験結果	10
2.3.4 実験結果に関する考察	10
2.4 映像がサウンドステージの広さに及ぼす影響に関する実験的考察	11
2.4.1 主観評価実験の概要	11
2.4.2 主観評価実験の方法	11
2.4.3 実験結果及びその考察	14
2.4.4 テレビ受像機のスピーカ間隔に関する考察	17
2.5 映像と音像の方向差の許容値に関する実験的考察その1	17
2.5.1 実験の概要	17
2.5.2 主観評価実験の方法	18
2.5.3 映像と音像の方向差の検知限と許容限	20
2.6 映像と音像の方向差の許容値に関する実験的考察その2	22
2.6.1 実験条件の概要	22
2.6.2 主観評価実験の方法	22
2.6.3 評価結果に関する考察	25
2.6.4 実際の番組による許容限の検証	29
2.7 映像と音の時間的な“ずれ”の影響について	30
2.8 結言	31
3章 HDTV用音声システムに関する考察	32
3.1 緒言	32

3. 2	広画角の映像と2チャンネルステレオの組み合わせの問題点	3 2
3. 3	最適な多チャンネルステレオ音声方式の検討	3 4
3. 4	結言	3 6
4 章	多チャンネルステレオ音場の波面解析とその応用	3 8
4. 1	緒言	3 8
4. 2	波面による音場解析の特長	3 9
4. 3	線形の収音再生系における波面の導出法	4 0
4. 4	複数音源により合成される波面の解析	4 1
4. 4. 1	2チャンネルステレオシステムの場合	4 1
4. 4. 2	多チャンネルステレオシステムの場合	4 4
4. 5	広帯域を有する信号への波面の概念の拡張	4 8
4. 5. 1.	等群遅延面に関する考察	4 8
4. 5. 2.	入出力相互相関関数による等遅延面に関する考察	4 9
4. 6	多チャンネルステレオ用音像モニター装置の開発	5 1
4. 6. 1	音像方向の視覚的表示の開発の背景	5 1
4. 6. 2	波面の法線方向の検出法	5 1
4. 6. 3	波面の法線方向のモニター装置	5 3
4. 6. 4	試作機による表示の例	5 4
4. 7	結言	5 5
5 章	音像の遠近制御法の開発	5 6
5. 1	緒言	5 6
5. 2	スピーカアレイを用いた音像の遠近制御法に関する仮説	5 7
5. 2. 1	基本原理	5 7
5. 2. 2	スピーカアレイの焦点近傍の音場解析	5 8
5. 3	スピーカアレイによる音像の距離感に関する主観評価実験 (主観評価実験1)	6 1
5. 3. 1	主観評価実験の目的	6 1
5. 3. 2	評価実験用スピーカアレイの構造	6 1
5. 3. 3	主観評価実験の方法	6 3
5. 3. 4	音像の距離感の実験結果	6 4
5. 4	スピーカアレイの音像制御に寄与する周波数帯域(主観評価実験2)	6 5

5. 4. 1	主観評価実験の概要	6 5
5. 4. 2	主観評価実験の方法	6 6
5. 4. 3	実験結果に関する考察	6 7
5. 5	スピーカアレイの大きさの効果に関する実験的考察(主観評価実験3)	6 7
5. 5. 1	主観評価実験の概要	6 7
5. 5. 2	主観評価実験の方法	6 8
5. 5. 3	実験結果に関する考察	6 8
5. 6	聴取範囲に関する考察	6 9
5. 7	外耳道入り口の音圧の両耳間コヒーレンスと音像の距離感に関する考察	7 0
5. 7. 1	コヒーレンスを用いる理由	7 0
5. 7. 2	外耳道入り口の音圧の両耳間コヒーレンスの測定結果結果	7 1
5. 7. 3	スピーカアレイによる音像の距離感の要因に関する考察	7 3
5. 8	既存の音場制御法との比較検討	7 4
5. 8. 1.	頭部伝達関数による方法	7 4
5. 8. 2.	マルチチャンネル再生による音場制御	7 5
5. 9.	スピーカアレイによる音像の遠近制御システムの実用化に関する考察	7 6
5. 10.	結言	7 7
6 章	番組音の相関を利用したVTRとテープレコーダーの同期運転装置の開発	7 9
6. 1	同期運転装置開発の背景	7 9
6. 2	同期運転装置の仕様に関する考察	8 0
6. 3	時間差検出の方法に関する考察	8 0
6. 4	相関関数検出のための回路構成	8 2
6. 5	VTRとATRの速度偏差の影響	8 3
6. 6	VTRとATRの音質差の影響	8 4
6. 7	ATRの速度制御法	8 4
6. 8	実際の使用例	8 5
7 章	結論	8 7
謝辞		8 9
参考文献		9 0

## 映像を伴う多チャンネルステレオに関する研究

### 1 章 序論

#### 1. 1. 本研究の位置づけ

本研究は、HDTVや立体テレビ等の将来のテレビ方式への応用を目的として、高い臨場感を持つ映像に視聴覚心理の面で適合する音声の再生方式のあり方を考察している。特にHDTVに対しては音声方式を設計するための基本的な指針を与えている。ここでは本研究の背景と特徴を概説する。

長い間、日本のテレビ方式はNTSC方式に限られていたが、最近になってHDTV (High Definition Television) の実用化等の新しい動きが見られる。また、次世代のメディアをめざした超高精細度テレビや立体テレビ等の研究も行われている。これらの新しいテレビ方式は広い視野や立体感など、視覚的な臨場感が従来のテレビ方式に比べ格段に優れている。しかし、それと同時に、その音声方式のあり方に大きな影響を与える可能性を持っている。例えば、HDTVについていえば、従来のテレビでほとんど問題視されなかった映像と音像の方向差が、映像が大画面であるがゆえに無視できなくなり、違和感を引き起こす恐れがある。また、立体テレビにおいて映像と音像の方向感だけでなく距離感まで一致させるためには、音像の遠近感を自由に制御する技術が必要になるであろう。逆に言えば、HDTVや立体テレビ等の新しい映像システムは適切な音声システムと組み合わせられることによって、一層効果的になると考えられる。

本研究では、上記のような新しいテレビ方式の音声システムとして、多チャンネルステレオを提案するが、実は、音響システムの側から見ても、映像を伴うことには大きな意味がある。音声の多チャンネル再生方式としては、すでに1970年代に4チャンネルステレオレコードが開発されているが、必ずしも成功したとは言えない。当時の4チャンネルステレオにも音響的には相当の効果があつたにもかかわらず、普及に至らず市場からほとんど姿を消している。本研究で敢えて再び多チャンネルステレオを検討する理由は、ひとえに映像を伴うシステムを対象としているためである。

4チャンネルレコードが成功しなかった理由は当時は4チャンネルを2チャンネルに変換する方法の不統一性にあると思われていたが、実は視覚情報の役割を考慮しなかったことも失敗の理由の一つと考えられる。音楽再生を例に取ってみると、後方に楽器を配置する特殊

な演奏形態があったとしても、生演奏ならば聴衆は視覚的に楽器の配置を確認することができるので、後方から楽器音が聞こえてくる必然性を理解できる。しかし、レコードで同じ曲を再現する場合は視覚情報が欠落しているため聴取者にその意図が伝わらず、4チャンネル音声の効果が理解されない恐れがある。

この問題点は演奏の映像が同時に提示される場合には軽減される。映像から楽器配置等の情報を得られるので音声のミクシング意図を理解しやすくなり、多チャンネルステレオの効果がより一層増すと思われる。このように音響システムの側からみても、多チャンネルの方式を検討する際には、視覚情報を含めた形で検討する必要があると考えられる。

上記のような背景をふまえ、本研究はHDTVを中心とする新しいテレビ方式に対して、それにふさわしい音声方式の基本指針を得ることを目標としている。特に、映像を伴う場合の音像の性質を考察し、大画面に見合う多チャンネルステレオ方式を提案する。

また、映像と音声の相互関係に加えて、本研究は音場の解析にも新しい手法を用いている。従来、専ら両耳の入り口における音圧の再現に関心が払われてきたステレオ音場の解析に、波面の再現能力という新しい要素を導入し、それによって、聴取範囲の推定、音像の遠近制御など、従来のステレオ理論で扱いにくかった分野に新しい解析手法を提案する。

さらに、映像と音声の関係や波面の再現性の検討をもとに、実際の放送技術に直結した機器の開発を行っている。

## 1. 2. 論文の構成

まず、次の第2章で映像が音像定位や音の広がり感に与える効果について実験的考察を行う。そして、その実験結果をもとにAVシステムにおける映像と音像の方向のずれの許容範囲を明らかにする。3章では2章の結果を受けて、映像を伴う多チャンネルステレオシステムの設計例として、HDTVの音声方式について考察する。ここでは、センターチャンネル（主として正面方向の音を再生するために左右のスピーカの中央に置かれたスピーカ用のチャンネル）が必要である点について述べる。

4章では波面の可視化によるステレオ音場の評価法を提案し、それによって多チャンネル再生の音場再現性について考察する。さらに、実際の多チャンネルシステムによる波面の方向を実時間で観測する方法を考案し、「音像方向モニター装置」を設計する。

5章では4章の方法をさらに発展させ、音像の遠近制御の新しい方法について述べる。この章では、スピーカと聴取者の間に音像を提示することがスピーカアレイによって可能であることを、理論的、実験的に示す。

6章で映像を伴う音声の番組制作に効果的な、VTRとテープレコーダの同期運転システムについて述べる。ここで述べる方法は、本研究で初めて開発したもので番組音の相関を利用した全く新しい方法による。最後の7章で全体を総括する。

なお、研究の分野が多岐にわたっているので、過去の研究に関する検討は各章の冒頭でそのつど行うこととする。

## 概要

本章ではHDTV（ハイビジョン）のための音声方式の設計指針を得ることを目的として、テレビ視聴時に映像が音像定位に与える効果について実験的考察を行う。次に、その結果を基に、HDTV等の大画面映像に適した音声システムの条件を明らかにする。

## 2. 1. 緒言

情報の表示手段として、コンピュータのCRT画面とテレビ映像とを比較すると、音声の役割に大きな違いがある。テレビ番組には音声が必要不可欠で、音声がないければ情報の欠落が大きく、番組として成立しない場合が多い。一方、CRT画面はマルチメディアが叫ばれる今日となっても、音声は補助的な出力手段である。我々は日常生活では外からの刺激に対し、無意識の内に視覚と聴覚を同時に働かせている。日常的な映像を扱うテレビや映画の場合、両方の情報の一体感が要求されるのは自然であろう。

実物とは明らかに違うとわかる平面的な映像と、実物とほとんど区別のつかない再生音を我々は一体不可分のものとして知覚できる。もし人間にこのような複数の感覚を統合する能力が無ければ、映像は映像、音は音として知覚され、両者が一体の情報としては認識されず、テレビも映画も娯楽メディアとしては成立しなかったかも知れない。

しかし、もし、テレビの画面の位置とスピーカの位置が極端に離れていても、映像と音声の一体感は得られるであろうか。残念ながらこのような映像と音声の関係、言い換えれば視覚と聴覚の関係は、心理学の研究対象となることはあっても、技術の問題としてはほとんど検討されていない。わずかに番組制作の場で心理効果として論じられることがあるにすぎず、システム設計のための議論に生かされてきたとはいえない。

しかし、一旦消えかけた4チャンネルステレオが映像を伴うことによって再び浮上してきたように、この相互関係がシステムの総合品質に大きな影響を与えるようになってきている。従って、HDTVや3DTV（3 Dimensional TV）等の将来のメディアも映像、音声単独の品質だけでなく、人間のこのような感覚の性質を有効に活かすべきである。

そこで本章では、将来のテレビシステムの設計に活用するという観点から、映像と音声の相互作用について考察を行う。特に、相互作用のうち最も顕著な現象の一つである、映像が音像定位に与える効果、及び両者のずれの許容値について実験的な考察を行う。

## 2. 2. 視覚刺激と音像定位の関係に関する従来の知見

## 2. 2. 1. 感覚間相互作用の実験例

視覚と聴覚の相互作用のうち最も顕著なものは、音像の方向が視覚刺激の影響を受けて変化する現象であろう。心理学的に見るとこのような感覚間の相互作用の現象は視覚と聴覚の間の特殊な現象ではなく、多くの感覚の間で存在する。音像定位のような「空間定位」に関しては、視覚、聴覚、自己受容感覚（自己の筋や腱で自分の姿勢を感じる感覚）の3つの感覚が関与し、互いに影響しあう。<sup>(1)</sup>

心理学の分野では視覚と自己受容感覚の相互作用の研究が最も進んでいる。人間は物体の位置関係を視覚だけでなく触覚や自己受容感覚でも判断する。例えば、ものの大きさは見なくても手で触れれば見当がつくし、自分の姿勢は目を閉じていても分かる。日常生活では2つの感覚の間に矛盾が生じることは希であるが、レンズ等を用いれば2つの感覚の間に矛盾を生じさせることができる。例えば、目に見える形と手で触った形が違う、あるいは視覚像の方向と実際の物の方向が異なる状態を作ることが可能である<sup>(2)</sup>。このような人工的な方向差を用いた実験がStratton(1897)<sup>(3)</sup>以降いくつか行われている。

Strattonは視覚像が網膜上で上下逆転しているにもかかわらず、そう感じないのはなぜかという疑問から、上下左右が反転する眼鏡を装着したときの感覚の変化を実験している。それによると、最初はめまいなど著しい違和感を感じるが、徐々に慣れて、ついには普通に見えるようになるという。当初は視感覚の側が何らかの修正を受けると思われていたが、修正を受けるのは視感覚の方ではなく、自己の姿勢を筋肉等で感じ取る自己受容感覚のほうであることが分かっている。すなわち、視覚と自己受容感覚の感覚のあいだに矛盾が生じた場合には視覚が自己受容感覚に影響を及ぼし、変化させてしまうことが多い。

音像の方向が視覚刺激の提示された方向に引かれる現象も、この視覚優位の現象が視覚と聴覚の間で起こったものと考えられる。しかし、視覚の影響力は絶対的なものとは考えられない。四角い形のものが丸く見えたとしても、触覚で丸く感じることがないように、視覚刺激と音源が大きくずれば、それらの方向のずれは知覚される。例えば、Jackson(1953)<sup>(4)</sup>はベルの音とライトの点滅等を使って実験し、音の方向判断の正答率から音像と視覚像は20度から30度ずれると分離するとしている。

スピーカ内蔵の現行テレビ受像機が映像と音像の方向差が気にならない理由は、標準的な視距離である7H（H：Picture height）の距離から視聴した場合、画面の見込み角が

高々10度であるためである。しかし、ステレオ音声の場合、音響的にはスピーカは60度の開き角としたい。60度配置とすると映像と音像の方向差は非対称聴取位置で相当大きくなることから、音声多重放送の開始前後にテレビの音をステレオ化する意味について様々な議論が行われている。八木<sup>(5)</sup>は20インチのモニターと実音源を用いて評価実験を行い、ニュースアナウンスの場合、映像と音像が8度以上ずれると半数以上の人が違和感を覚えるとしている。また、中林ら<sup>(6)</sup>は、20インチのモニターとステレオ再生による合成音像を用い、映像を注視したときの実際の音像定位方向を主観評価実験により求め、映像と音像の方向差が10度以内であれば音像は映像に強く引かれ、20度以上ずれば映像の影響が小さくなることを明らかにしている。これらの実験結果は画面の見込み角が高々10度程度のNTSC方式のテレビ受像器では、スピーカを画面に近接して配置すれば定位の方向差はほとんど気にならないことを裏付けている。

## 2. 2. 2. 従来の実験の問題点と本研究の意義

心理学の分野では音像が視覚刺激の方向に引かれる現象はよく知られていたが、それは心理学の分野にとどまっておらず、そのままではHDTVの音声システムを決めるための知見とはならない。例えば、映像と音像のずれが知覚された場合、それが違和感となるかどうか、なるとすればどの程度か、などについては未知である。一方、音声多重放送開始時に行われた工学的立場からの少数の実験も、それらの結果をそのままHDTVにあてはめるには無理がある。例えば、(1)画面の大きさが異なること(音像が映像に引かれる程度は映像の大きさに依存する<sup>(6)</sup>)、(2)評定者の分類(専門家か非専門家か)が行われておらず、信頼性に欠けること、(3)ごく単純な構図の映像と音声を用いており、一般的なテレビ番組に当てはまる結果かどうか明らかでない、の3点から十分とは言えない。また、画面が大きくなると、広い視野にわたる視覚刺激は音像定位の精度を高めるという報告もあり<sup>(7)-(8)</sup>、20インチ大の画面の結果をHDTVに適用することは危険である。従って、正確な結果を得るためには、大画面による映像を用いて実験を行い、的確な統計処理を行うことが必要である。そこで、以下の節では、大画面(ただし、家庭用の範囲内)の映像と音像の相互作用について実験的考察を行う。さらに、単純な方向感に関する実験だけでなく、音の広がり感と映像の相互作用のように番組効果に重大な影響を与える現象の存在についても明らかにする。

## 2. 3. テレビ映像による音像方向の偏移に関する実験的考察<sup>(10)</sup>

### 2. 3. 1. 実験的考察の概要

大画面に提示された人物の映像を注視しながら音像方向を判定したときの、音像方向の本来の提示方向からの偏移量を主観評価実験によって測定し、その性質に定量的考察を加える。

### 2. 3. 2. 実験装置の配置と実験方法

(1) 実験装置の配置 実験装置の配置は以下の通りである。12畳程度の部屋でHDTVを視聴することを想定して、72インチ大の投射型スクリーン、スピーカ、評定者をFig. 2.1のように配置する。スピーカは視覚情報とならぬよう不透明なカーテンで遮蔽されている。ただし、音波の透過損失は無視し得るほど小さい。評定者は映像を見ながら音像の方向を判定するが、2つのスピーカによる合成音像の方向の判断であるため、頭の動きによる定位方向の変化を避けるために聴取者の頭はヘッドレストにより軽く固定されている(2チャンネルステレオのような2つのスピーカにより合成される音像の方向は頭の向きに依存する)。

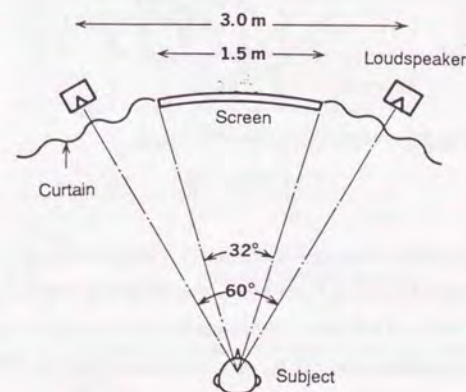


Fig.2.1 Loudspeaker layout for evaluation of visual effects on sound localization.

音像方向に与える視覚刺激の効果を評価するための実験配置

映像提示装置はいわゆる前方投射型CRTプロジェクターと呼ばれるものを用いた。スクリーンの下部にあるCRT本体から出た光が、評定者の膝の前にある反射鏡で一度反射して、スクリーンにビデオ映像を映し出す。使用したスピーカは16cm径のフルレンジスピーカユニットを密閉箱に取り付けたもので、音声の再生には十分な周波数帯域を有する。スピーカの床からの高さ、映像の口の高さ、評定者の耳の位置の高さをほぼ同じ水平面上に配置している。

(2) 実験手順 次に実験の手順を述べる。実験は「実験者」と呼ばれるオペレーターが刺激の提示順序を管理し、評定者の回答を記録する方法で行う。実験者がVTRを再生し、Fig. 2.2に示すような映像を約20秒間提示する。同時にその音声信号を2分岐し、レベル差を与えて、左右のスピーカに加える。レベル差の調節により、音像を左右に配置した2つのスピーカの間の任意の方向に提示することができる。

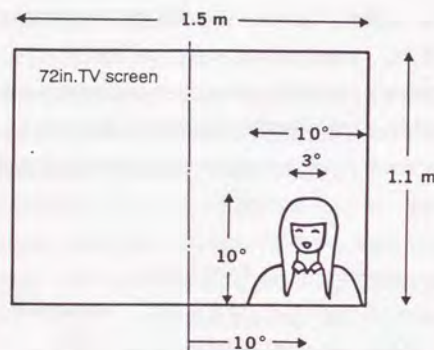


Fig.2.2 Position of picture on TV screen.

画面上の映像の位置

評定者は映像を注視しながら、手に持ったペンライトで音像の方向を指示する。支持した場所と画面の下辺に付けられた目盛りとを照合して音像方向を読み取り、実験者に報告する。視覚刺激により変化した音像方向の抽出が実験の目的であるため、評定者は音像の方向を意識して見ないように指示されている。なお、角度は正面を0度、右方向を正としている。

(3) 刺激の種類 ここで実験に用いた映像と音声について述べる。本実験では実際のテレビ番組を視聴する状態に近い条件にするために、人物の映像とその音声というテレビ

番組で最も一般的な被写体と音を用いている。実験条件を明確に規定するためにはライトの点滅と純音の断続のような単純な刺激の組み合わせが望ましいが、本研究の目的には実際のテレビ視聴により近い刺激が適合すると判断した。

映像と音声の収録はNHK中央研修所のテレビスタジオにて行った。椅子に着座して朗読している女性をビデオカメラで正面から撮影、音声も同時に収録し、VTRに記録した。映像は画面上で右に約10度寄るように撮影した。これは映像の音像方向に与える影響の非対象性の存在の有無を確認するためである。また、人物の背景は Horizont と呼ばれるテレビスタジオの壁で一般的なグレー色をしている。

(4) 実験条件 実験の評定者は視力、聴力とも正常な成人男女11名で、一人当たりの実験の繰り返し回数は2回である。この種の主観評価実験では評定者数が10人以上、できれば20人程度が望ましいとされているが(CCIR Recommendation 562)、繰り返しを行うことで、のべ22人を確保した。音声の聴取レベルは映像のモデルとなった人物の肉声の大きさにほぼ等しい約65dB Aに設定した。実験に用いたプロジェクターは通常の部屋の明るさでは輝度が不足するため、部屋の照明は消灯した。

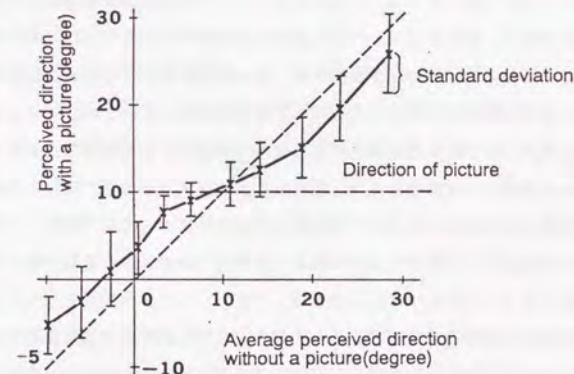


Fig.2.3 Displacement of sound-image direction by picture on TV screen (for 11 subjects and 2 times, female voice, 65dBA.)

映像による音像方向の偏移：本来であれば45度の直線に乗るはずの音像方向が、10度方向の映像のためにS字型に変化している。

### 2. 3. 3. 映像による音像方向の偏移の実験結果

前述の実験により判定された音像方向の平均と標準偏差をFig. 2. 3に示す。この図において横軸は音だけを提示したときの音像方向、縦軸は映像を注視したときの音像方向である。この図から映像を注視した場合の音像方向は約10度の方向を中心にS字形となっていることが分かる。これは音像が10度の方向、即ち映像の方向に引かれていることを示している。特に音像と映像の角度差が5度以下では音像の方向が映像の方向にきわめて近くなっていることが分かる。

この結果は全般的には20インチモニターを用いた中林<sup>(4)</sup>らの結果と同じ傾向であるが、標準偏差に注目するとやや異なる。中林らの結果では映像と音像が10度から15度程度離れたときに標準偏差が増大しているが、今回はそのような傾向がみられない。そのかわり20度離れたとき映像の影響が残っている。次にこれらの結果について考察を行う。

### 2. 3. 4. 実験結果に関する考察

前述の実験により映像と音像の提示方向のずれが5度以内の場合、音像が映像の方向に強く引かれることが明かとなった。これは20インチ大のモニターを用いて行われた過去の結果と同様である。逆に言えば、映像と音像の方向差を知覚させないためには両者を5度以内の範囲に提示すれば良いと考えられる。この問題については2. 5節で映像と音像の方向のずれの検知限、許容限という形で再び取り上げる。

一方、方向差が10～15度の場合に標準偏差の増大があまり顕著でないのは、スピーカが評定者から見えないためであろう。中林らの実験はスピーカが評定者から見えるため、スピーカそのものに音像が引かれた可能性がある。そのため、テレビ映像とスピーカの間位置である10から15度でテレビ映像と“スピーカの実像”の影響がきつ抗したものと考えられる。

今回の実験結果を注意深く観察すると、映像の左右で音像方向の変位量がわずかに違っている。このような左右非対称が生じた理由は2つ考えられる。ひとつは映像が右に寄っているため、画面の枠までの距離が異なること。もうひとつは音像の質で、右によるほど音像は実音像（単一スピーカによる音像）に近づき、いわゆる“締まった”音像になることである。どちらも右側の音像の方が映像の影響を受けにくいことと矛盾しない。この結果だけではどちらの効果が支配的かは断定できないが、映像と音像の方向差を知覚させないためには中央寄りに音像を提示するほうが安全であるといえるであろう。

### 2. 4. 映像がサウンドステージの広さに及ぼす影響に関する実験的考察<sup>(10)</sup>

#### 2. 4. 1. 主観評価実験の目的

2. 3節で映像が音像の方向を引きつける効果を実験的に確認したが、ステレオ音響の効果としては音像の方向性だけでなく音の広がり感も重要な要素である。しかし、2. 3節の実験は映像、音声とも単一の方向であるため、その実験結果は広がりを持つ映像や音声には適用できないと考えられる。映像や音像が広がりを持つ場合には、視聴者は一点を注視するよりは全体を眺める傾向が強くなるであろう。しかし、音の広がり感に及ぼす映像の影響についての実験例はほとんど見あたらない。そこで、本節では大画面映像がサウンドステージの広さ（音像が分布する範囲）に及ぼす影響について検討する。

ここで想定している場面は個々の音像の発音源が映像のどの部分であるか判定できないが、全体として映像と音声に相関がある場面である。具体的な例をあげれば、オーケストラ全体を映した映像、雑踏の映像、スポーツやコンサートにおける観客の映像等である。このような映像は番組としてごく一般的なものである。

#### 2. 4. 2. 主観評価実験の方法

本節の実験は画面全体に広がる映像がステレオによる音声の広がり感に及ぼす影響を明らかにすることが目的である。実験条件を明確にするために、映像、音声として、その広がり幅が明確に規定できるものを用いる必要がある。そこで映像は7人の人間が横に並んで拍手をしているところをビデオカメラで収録したものをを用いた。撮影された映像の概略をFig. 2. 4に示す。一方、音声の収録にはMSステレオマイク（後述）を用い、Midマイク信号とSideマイク信号のまま録音した。MSマイクを用いることにより、音源とマイクの位置関係からサウンドステージの幅を計算で求めることができる。

ここでMSマイクの出力からサウンドステージの幅を求める方法について述べる。MSマイクはmiddle-sideマイクの略称で、単一指向性のMマイクと双指向性のSマイクを同軸上に配置したステレオ用マイクである。通常はMマイクの出力とSマイクの出力の和信号と差信号をそれぞれ左右のチャンネルの信号とする。MマイクとSマイクの混合比を変えることでサウンドステージの広さを変えることができる。本実験ではFig. 2. 5に示す逆連動減衰器でM信号とS信号の混合比を調節した。逆連動減衰器は2つの信号のパワーの合計を一定に保ったまま、2信号の比を変化させることができる2連の減衰器である。

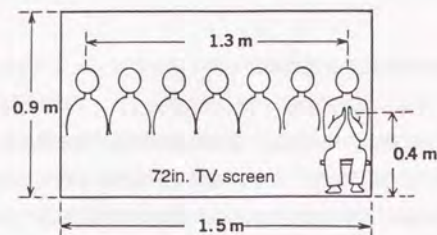


Fig.2.4 Picture for evaluation of visual effects on sound stage.  
サウンドステージの広さに与える映像の影響を評価するための映像

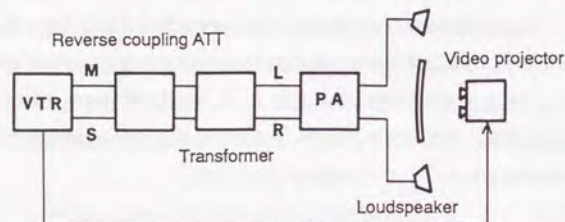


Fig.2.5 Setup for changing width of sound stage.  
サウンドステージの幅を調節するための回路

この方法によれば音の大きさをほぼ一定に保つことができることが知られている。

サウンドステージの幅とMSマイクの混合比との関係は次の通りである。MSマイクのそれぞれの利得を  $I_m$ 、 $I_s$  とし、和信号を左信号  $S_l$ 、差信号を右信号  $S_r$  とすると、音源の方向が  $\theta$  のときはそれぞれ次のようになる。

$$S_l = I_m \left( \frac{1 + \cos \theta}{2} \right) + I_s \sin \theta \quad (2.1)$$

$$S_r = I_m \left( \frac{1 - \cos \theta}{2} \right) - I_s \sin \theta$$

この信号をスピーカの見込み角 60 度の 2 チャネルステレオ再生システムで聴取したときの正面から測った音像方向  $\alpha$  は次の式で与えられる。

$$\tan \alpha = \tan 30^\circ \frac{S_l - S_r}{S_l + S_r} \quad (2.2)$$

$$= \tan 30^\circ \frac{2K \sin \theta}{1 + \cos \theta} \quad : K = \frac{I_s}{I_m}$$

この式を用いて音源の 7 人の右端の人から左端の人までの幅と MS マイクの M 信号と S 信号の混合比から再生音場でのサウンドステージの幅を算出する。

家庭での HDTV 視聴を想定して、72 インチ投射形スクリーン、スピーカ、評定者を Fig. 2. 1 のように配置し、評価実験を実施した。聴取レベルは実際の発音体を映像と同じ見込み角から聴取する時のレベルに近い約 70 dBA とした。また、映像を提示する実験では実験室の照明は映像を見やすくするため消灯した。

実験は次に述べる恒常法と極限法の 2 つの手順で行い、実験方法に依存する要素も抽出することとした。

恒常法は実験者が刺激をランダムな順番で多数回提示する方法で、実験時間が長くなるが精度が高い性質を持つ。極限法は実験者が刺激をわずかず連続的に変化させ、評定者に判断の転換点を答えさせる方法で、学習や疲労等による判断の時間変化を抽出しやすい。

#### a) 恒常法による実験手順

実験者はサウンドステージの幅をある値になるように逆連動減衰器を設定した後、VTR を再生し、映像と音を約 5 秒間提示した。評定者は提示された音のサウンドステージの幅をスクリーンの幅と比較し、「狭い」、「同じ」、「広い」のいずれであるかを判断した。また、映像を提示せず、音だけを提示する実験も行った。評定者は健全な視力聴力を有する成人男女 12 名で、繰り返しを 4 回行った。

#### b) 極限法による実験手順

実験者は映像を提示すると共に、サウンドステージの幅を狭くから広くへ、次に逆に広くから狭くへ単調にゆっくりと変化させた。評定者は提示されたサウンドステージの広さを映像の幅と比較して、判断の転換した時点で合図した。例えばサウンドステージが広がっていく上昇系列では「サウンドステージの幅が映像の幅と比べて狭い」が「等しい」に変化したときと、「等しい」が「狭い」に変化したときに合図した。実験者はそのときの逆連動 ATT の目盛りを読み取った。

評定者は成人男子 5 名で、上下各系列を交互に合計 20 回実施した。また、音だけを提示する実験も同様に行ったが、比較の対象は画面上の両端の人の手の位置に相当する場所につけたマーカー間の幅とした。

## 2. 4. 3. 実験結果及びその考察

### a) 恒常法による実験結果に関する考察

恒常法による結果をFig. 2. 6に示す。図の横軸はサウンドステージの中心から端までの幅の理論値である。Midマイクの信号とSideマイクの利得差から算出したものである（算出法については前節参照）。また、横軸は判断の出現比率を示す。「サウンドステージの幅がスクリーンより狭い」の判断出現比率が50%になる刺激値が下弁別刺激、「広い」の判断出現比率が50%になる刺激値が上弁別刺激になる。

この図は、「狭い」の出現比率は映像の有無に依存しないが、「広い」の出現比率は映像を提示すると減少する傾向にあることを示している。しかし、角度の差にしてわずか2度であり、映像の効果が大きいとは言えないことがわかる。

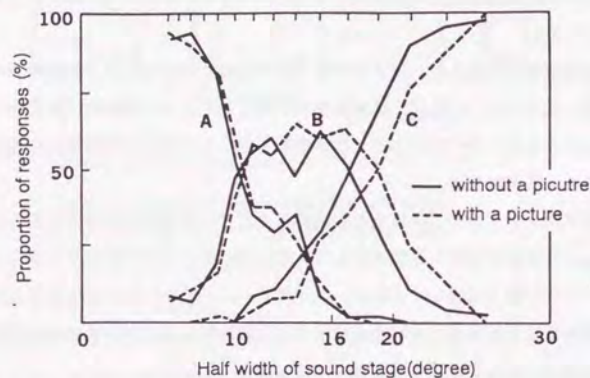


Fig.2.6 Proportion of three responses to sound stage (for 12 subjects and 4 times, studio room, sound of clapping.) Solid lines: Case of sound images without a picture. Dotted lines: Case of sound images with a picture. A: Response indicating that sound stage is narrower than the screen. B: Sound stage is the same size as the screen. C: Sound stage is wider than the screen.

3つの判断の出現比率：サウンドステージがスクリーンの幅よりも広いと感じる確率（C）が映像の提示によって減少している。

### b) 極限法による実験結果

極限法による実験の結果をTable 2. 1とTable 2. 2に示す。「狭い」と「等しい」の境界のサウンドステージの幅の1/2（下弁別刺激）をTable 2. 1に、「等しい」と「広い」の境界のサウンドステージの幅の1/2（上弁別刺激）をTa-

ble 2. 2に示している。この表でも恒常法による結果と同じく、映像によって変化するのは上弁別刺激だけであることがわかる。平均値の差の検定でも上弁別刺激のみ映像の影響が1%で有意であり、上弁別刺激の映像の有無による角度の差は2.3度と恒常法による結果と同じく小さな値である。

しかし、Fig. 2. 7に示すように、回ごとの平均を見ると実験初期には視覚の影響がきわめて大きいことがわかる。それも下降系列に顕著で、約5度に及ぶ。すなわち、視聴条件によっては映像がサウンドステージの幅（音の広がり感）を狭く感じさせることを示している。この実験結果の持つ意味については次節で考察する。

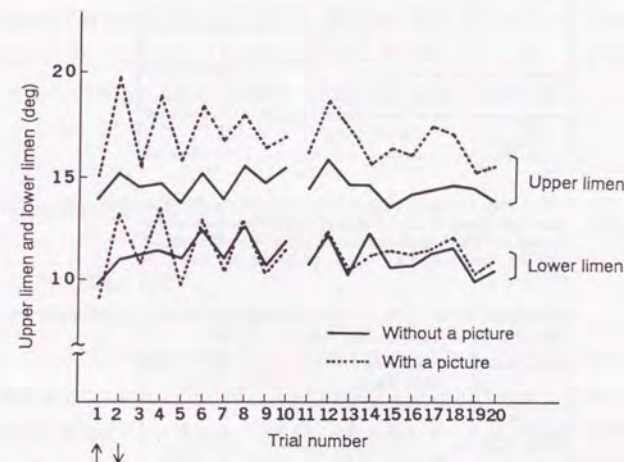


Fig.2.7 Changes in the limens for the 20 trials (for 5 subjects.) Odd: ascending series, Even: descending series. Subjects take a five-minute recess from the 10th trial to the 11th one.

上弁別刺激と下弁別刺激の実験回数による変化

Table 2.1 Average lower limens: Average thresholds below which the sound stage is perceived as narrower than the screen (for 5 subjects and 20 times, studio room, sound of clapping.)

下弁別刺激の平均値：サウンドステージが映像の幅よりも狭いと判断される域値  
(評定者5名、繰り返し20回、スタジオ、拍手音)

series subj.	with image		without image	
	↑	↓	↑	↓
C <sub>1</sub>	11.7	13.8	11.9	13.9
C <sub>2</sub>	10.8	13.8	10.2	12.6
C <sub>3</sub>	9.7	7.3	9.4	7.7
C <sub>4</sub>	10.5	14.5	12.2	14.2
C <sub>5</sub>	9.2	11.3	9.2	9.5
average	10.4	12.1	10.6	11.6
total average	11.2 (deg)		11.1	

↑上昇系列  
↓下降系列

Table 2.2 Average upper limens: Average thresholds over which the sound stage is perceived as wider than the screen (for 5 subjects and 20 times, studio room, sound of clapping.)

上弁別刺激の平均値：サウンドステージが映像の幅よりも広いと判断される域値  
(評定者5名、繰り返し20回、スタジオ、拍手音)

series subj.	with image		without image	
	↑	↓	↑	↓
C <sub>1</sub>	17.1	17.7	14.5	15.8
C <sub>2</sub>	16.4	19.4	14.1	16.7
C <sub>3</sub>	16.2	13.1	13.2	10.9
C <sub>4</sub>	18.0	21.9	17.1	18.1
C <sub>5</sub>	12.9	15.2	12.0	12.6
average	16.1	17.5	14.2	14.8
total average	16.8 (deg)		14.5	

#### 2. 4. 4. テレビ受像機のスピーカ間隔に関する考察

2. 4. 3 節の実験結果は、映像の影響によってサウンドステージの幅が狭くなる現象の存在を示している。特に極限法の実験初期ではその量は約5度に及ぶ。この5度という値は実際のテレビ受像機のスピーカ配置について以下のことを示唆している。

ハイビジョン（見込み角30度程）の画面の両サイドに近接してスピーカを設置した場合を想定してみる。スピーカの大きさにもよるがスピーカの方向は画面の端より5度以上外側になることは少ない。このため、このスピーカ間隔ではサウンドステージが聴感上画面の内側にしか存在しない印象を与える恐れがある。これは、画面の外側から背景音等を再生し、空間的な広がりを表現することが困難になるという意味で好ましくない現象である。スピーカの外部に音が広がるようにチャンネル間のクロストークを制御する手法もいくつかあるが、第4章で述べるように、2つのスピーカから等距離の聴取位置でないと効果が半減する。広い聴取範囲で映像の影響に抗して広いサウンドステージを再現するためには、スピーカは画面から5度以上外側に設置する必要があるといえる。

#### 2. 5. 映像と音像の方向差の許容値に関する実験的考察その1<sup>(19)</sup>

##### 2. 5. 1. 実験の概要

2. 3 および 2. 4 節の実験で音像の方向とサウンドステージの幅が映像の影響を受ける現象を確認したが、この結果はテレビの音声システムの設計に対し重要な示唆を与える。すなわち、音像は対応する映像の位置に正確に提示しなくとも、ある範囲内に提示すれば方向差が検知されない、言い換えると、音像定位は精密に映像の方向に設定する必要はないことを意味する。そして、その範囲は映像に対し約5度以内である。5度を越えれば、このような音像の引き込み現象が生じにくくなり、方向差が検知される。

しかしながら、方向差が検知されることと、それを不自然と感じるか否かは別の問題である。そこで、本節と次節では、ハイビジョンの音声システムのスピーカ配置に関する基礎資料を得る目的で、方向差の許容限に着目した主観評価実験を行う。映像が音像方向に与える影響の大きさには様々な心理的要因が関与すると考えられており、一般性のあるデータを得るためにはいくつかの条件で実験を行うことが望ましい。この立場に立ち、本節では、心理測定法に基づいた実験を行い、次の2. 6 節で家庭での視聴条件を加味した実

験を行う。

## 2. 5. 2. 主観評価実験の方法

(1) 実験条件 本節ではまず視覚の引き込み効果があまり期待できない条件下での実験的考察を行うために、次のような条件を設定する。

1) 刺激の提示時間を約8秒と短い条件とし、長時間映像を注視することによる音像方向の変化が起こらないようにする。

2) 映像の提示位置をランダムにし、映像の提示方向に関する先入観を取り除く。

上記の条件を実際のテレビ番組に当てはめれば、頻繁に場面の切り替えが行われる場合に相当する。

(2) 実験装置の配置 実験に用いたスクリーン、スピーカおよび、評定者の配置をFig. 2. 8に示す。この実験では映像の提示位置が変わるので、それを見ようとする評定者の頭の位置が動くことは避けられない。そこで、音像の方向が評定者の頭の動きに左右されないようにするために、実音源を用いて音像を提示した。実音源用のスピーカ(F社FE103)はスクリーンのすぐ下に3度おきに配置した。スピーカの“実像”そのものの影響を除くために、スピーカは不透明なカーテンで覆った。評定者の位置からみたスクリーン、スピーカ及び、映像の位置関係をFig. 2. 9に示す。

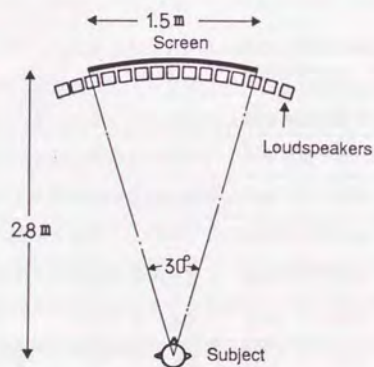


Fig.2.8 Loudspeaker layout for evaluation of discrepancy between picture and sound-image direction.

映像と音像の方向差を評価するためのスピーカ配置

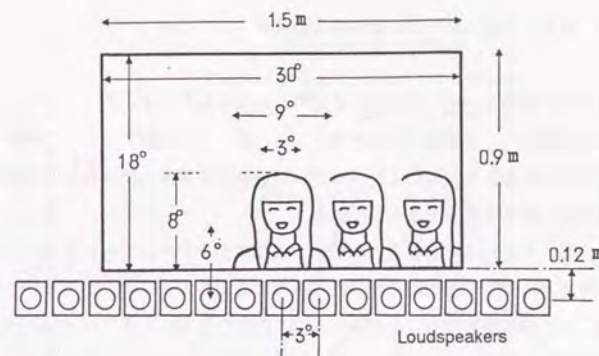


Fig.2.9 Arrangement of loudspeakers, screen and pictures (a view from the subject.)

評定者からみた映像とスピーカの位置

(3) 提示刺激 本条件の目的に適合するように撮影した映像と音声を用いた。被写体そのものは2. 3節と同一であるが、画面上の映像の位置を変えて3種類撮影し、それをランダムな順番で提示することとした。

(4) 実験の手順 被写体の映像は正面、右6度、右12度のいずれかの位置にランダムに提示し、音像はその映像の左右9度、6度、3度、0度の位置にある、いずれかのスピーカで提示した。すなわち、合計で一つの方向の映像に対し7方向の音像の組合せを提示したことになる。

評定者は映像を注視しながら音聞き、映像と音像の方向のずれをTable 2. 3に示す5段階のカテゴリーで判断した。

聴取レベルは被写体の肉声の大きさに近い65 dBAとした。評定者は音響技術者を中心とした成年男子10名、繰り返しは5回行った。

Table 2.3 Categories for judgments of discrepancy between picture and sound-image directions.

音像方向が映像方向とくらべて
1. 右側で不自然である
2. 右側だが許容範囲内である
3. 一致している
4. 左側だが許容範囲内である
5. 左側で不自然である

### 2. 5. 3. 映像と音像の方向差の検知限と許容限

「音像の方向が映像と一致している」と判断した比率をFig. 2. 10に、「方向差がわかるが許容範囲内」と判断した比率をFig. 2. 11に示す。ここで横軸は音像の方向を示す（右方向を正としている）。図中の映像提示位置は人物像の口の位置の水平方向角を唇の横幅も含めて表示したものである。

一致している、あるいは許容範囲内と判定した比率が50%となる角度差がそれぞれ映像と音像の方向差に対する検知限と許容限に相当する。Fig. 2. 10と2. 11からそれぞれ約4度と約8度となることがわかる。8度という値は八木<sup>(6)</sup>の結果とも一致する。

この値を実際のスピーカ内蔵型のテレビ受像機に適応すると次のことが言える。まず、NTSC方式の20インチ程度のテレビモニターでは、適視距離からの画面の見込み角（画角）が約10度である。この場合、スピーカが画面の中心を通る垂直面内に設置されていれば、映像の提示位置によらず水平方向の角度差は5度以内であり、映像と音像の方向差はほとんど検知されない。また、スピーカが受像機の側面に設置されている場合も、方向差は10度以下であり、許容値に近い。

一方、HDTVの場合は画角が30度以上あり、スピーカが中央に設置されていたとしても、映像と音像の方向差が許容値を越えることがありうる。さらに、ステレオ音声に対応したHDTV受像機の場合は条件がさらに厳しくなる。すなわち、ステレオ再生による合成音像は非対称聴取位置で聴取した場合、近いスピーカの方に定位するので、映像の提示位置によっては最大30度の方向差が生じることになり、許容値を大きく越える。従って、本実験で得られた映像と音像の方向差に関する許容限から判断すると、HDTV受像機のスピーカを単純に画面のすぐ横あるいは側面に配置するのは好ましくないといえる。

しかしながら、最初に述べたようにここでの実験条件は、映像の提示時間が短い、評定者に音響技術者が多い、等の理由から家庭におけるテレビの視聴実態からはやや遊離している恐れがある。それが許容値を必要以上に小さな値にした原因とも考えられる。そこで、次節では一般の聴取者が家庭でテレビを見るときの状態により近い条件で実験を行い、許容限に関して再度考察を行う。

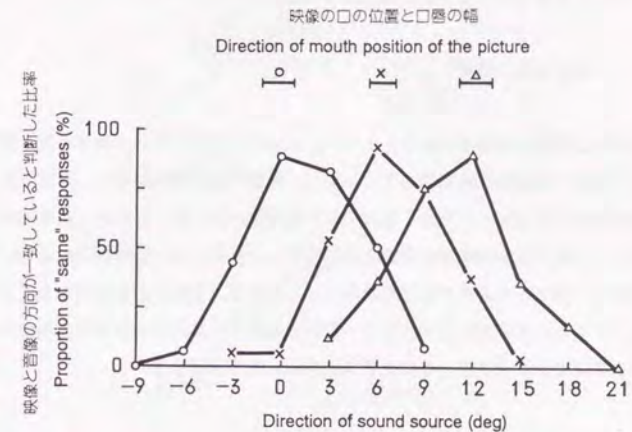


Fig.2.10 Proportion of "same" responses to displacement between picture and sound directions (for 10 subjects and 5 times, studio room, female voice, 65dBA.)

映像と音像方向が一致していると答えた割合

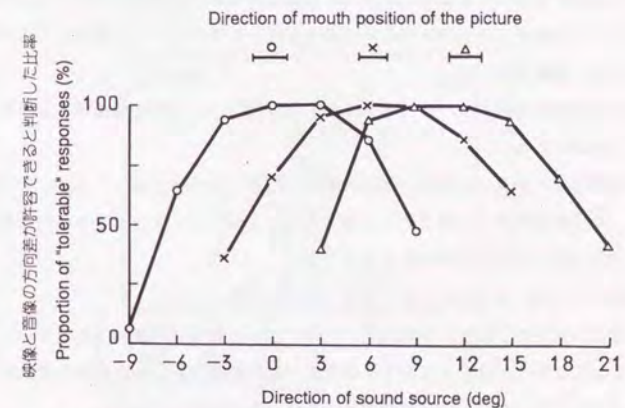


Fig.2.11 Proportion of "tolerable" responses to displacement between picture and sound directions (for 10 subjects and 5 times, studio room, female voice, 65dBA.)

映像と音像の方向差が許容できると答えた割合

## 2. 6. 映像と音像の方向差の許容値に関する実験的考察その2<sup>(11)(12)</sup>

### 2. 6. 1. 実験条件の概要

前節の実験は映像と音像の位置をともにランダムに提示しており、評価に先入観を与えないという意味では厳密な実験である。しかし、映像の提示時間が短く、評定者に音響技術者が多い条件で行っているため、家庭での視聴条件としては、必要以上に評価が厳しい恐れがある。本節では実験条件を家庭におけるテレビ視聴により近づけた場合の、映像と音像の方向差に対する許容限の推定を試みた。その結果、音響技術者で約11度、一般視聴者で約20度という結果を得た。以下、実験の経緯と前節の実験結果との関係について述べ、HDTVの音声システムが満たすべき条件を明らかにする。

### 2. 6. 2. 主観評価実験の方法

実験に使用したスクリーン、スピーカ及び評定者の配置をFig. 2. 12に示す。2. 6節の実験方法と異なる点を以下に列挙する。

(1) スピーカの配置が正面から背面まで広い範囲にわたっている。これは極端に映像と音像が離れた条件も提示することによって、評価の下限の目安を評定者に与え、評価を適正に行わせようと意図したものである。なお、左右30度のところに設置したスピーカは2. 6. 4節の実験で用いる。

(2) 提示する映像はFig. 2. 13に示すものを用いた。これは実際のニュース番組等を想定した構図である。

(3) 評定者は成人男女44名で、その内訳は専門家(音響技術者)14名、男子非専門家19名、女子非専門家11名である。繰り返しは5回行い、専門家と非専門家を分けて統計処理するために十分な実験回数を確保した。

音声の聴取レベル、その他の条件は前節の実験と同様とした。

映像と音像の定位の方向差の評価法は、放送に関する国際標準機関であるCCIRの勧告に従って、Table 2. 4に示す5段階の尺度で実施した。また、評定者は家でテレビを見るようなリラックスした気持ちで評価を行うように指示された。

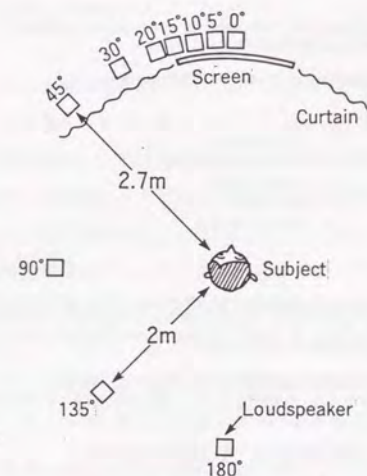


Fig.2.12 Loudspeaker layout for evaluation of the discrepancy between video and sound images.

映像と音像の方向差を評価するためのスピーカ配置

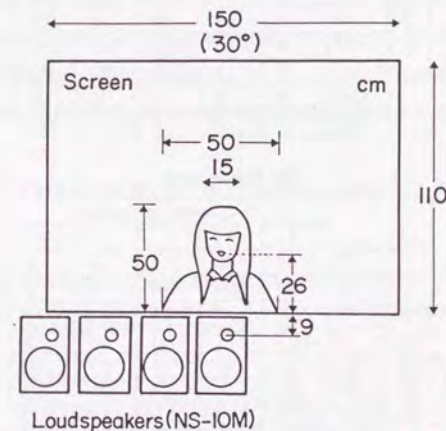


Fig.2.13 Placement of picture on TV screen and frontal loudspeakers (1°=5cm).

映像とスピーカの位置関係

Table 2.4 Categories for evaluation of discrepancy between picture and sound-image directions.

映像と音像の方向差の評価のためのカテゴリー

- A. 差がわからない
- B. 分かるが気にならない
- C. やや気になる
- D. 気になる
- E. 非常に気になる

Table 2.5 Evaluation of discrepancy between video and sound images by the five categories (for 14 experts and 30 non-experts, 5 times, studio room, female voice, 65dBA.)

映像と音像の方向差判断のカテゴリーの出現比率  
(専門家14名、非専門家30名、繰り返し5回、女声65dBA)

(a) Acoustic research engineers

		Successive categories				
		A	B	C	D	E
Discrepancy	0°	85.7	14.3	0	0	0
	5°	70	24.3	1.4	4.3	0
	10°	15.7	44.3	24.3	14.3	1.4
	15°	4.3	24.3	35.7	20	15.7
	20°	1.4	15.7	34.3	25.7	22.9
	30°	0	5.7	30	20	44.3
	45°	0	0	11.4	22.9	65.7
	90°	0	0	0	7.1	92.9
	180°	1.4	0	0	21.4	77.1

(b) Non-experts

		Successive categories				
		A	B	C	D	E
Discrepancy	0°	86	13.3	0.7	0	0
	5°	78	18.7	3.3	0	0
	10°	48	42.7	8.7	0.7	0
	15°	20.7	54.7	21.3	3.3	0
	20°	5.3	52	34	8.7	0
	30°	3.3	24	50	19.3	3.3
	45°	0	6	32.7	42.7	18.7
	90°	0	0	10.7	28	61.3
	180°	2	18.7	14	32.7	32.7

2. 6. 3. 評価結果に関する考察

本実験の結果から、映像と音像の方向差に対する許容限を求めるためには若干煩雑な手続きを要するが、結論をさきに述べると、専門家に対しては約11度、非専門家に対しては約20度という値を得た。これは、一般の視聴者がテレビを見るときに、映像と音像の方向差が20度以内であれば、50%以上の確率で許容することを示している。以下、この値を得る手続きについて述べる。

実験の結果、音響技術者のグループと非専門家のグループでかなりの評価傾向の違いが見られたので、2つのグループごとに結果を分けて処理を行った。それぞれのグループの評価の相対度数をTable 2.5に示す。この表から音像と映像の方向差(Discrepancy)が大きくなるにつれて評価がAからEに下降していることが分かる。なお、専門家、非専門家を問わず、側方(90度)に対する評価が最も低く、後方になると若干評価が上昇する傾向が見られる。これは聴覚の性質上、音像方向の前後の判断が曖昧であることに起因していると思われる。この部分の実験結果は正規分布を期待できないので、以後の分析の対象から除外する。

Table 2.5の度数分布から許容限を求める方法として最も簡単なやり方は、評価A～Eを点数5から1に置き換え、方向差ごとに平均点を算出して3.5点(評価BとCの中間)に相当する方向差を求めるものである。しかし、この方法は評価A～Eが心理尺度上で等距離にあることを仮定しており、本実験のように評価A～Eが単なる順位、すなわちカテゴリーを示す場合には適用出来ない。そこで、ここでは系列範ちゅう法を用いて、評価A～EすなわちカテゴリーA～Eの間の距離を求め、評価の平均値を求めることとした。

系列範ちゅう法の手続きの詳細は専門書<sup>(13)</sup>に譲り、結果だけをTable 2.6に示

カテゴリーEを原点とし、判断の標準偏差を単位とした各カテゴリーまでの距離

Table 2.6 Category mean values measured in standard deviation units. (Category E is forced to 0.)

Category	A	B	C	D	E
Expert	4.171	2.818	1.886	1.128	0
Non-expert	4.89	3.442	2.272	1.184	0
Total	4.22	2.838	1.788	0.983	0

す。これが評価Eを起点としたときの各評価までの距離である。この表から、評価CとDの距離よりも評価AとBの距離の方が大きいことが分かるが、これは評価AとBの差の方がCとDの差よりも心理的に大きいことを示している。なお、表の数字の単位は評価の標準偏差を1としたものである。

次に各刺激からカテゴリーAからEまでの距離を求め、そのTable 2. 6のカテゴリー中心値への回帰をFig. 2. 14に示す。この図において、回帰がほぼ直線に乗っているが、これは刺激に対する判断が正規型であり、系列範ちゅう法の適用が妥当であることを示している。

この回帰直線が横軸と交わるときの尺度からその刺激の平均尺度の推定値が、また傾きの逆数から標準偏差の推定値が得られる。それをFig. 2. 15 a, bに示す。

評価BとCの境界を許容限とすれば、方向差の大きさの許容限は音響技術者で11程度、一般評定者で20程度であることが分かる。以上が系列範ちゅう法による結果である。

ここで得られた一般評定者の許容限(20度)は音響技術者の許容限(前節の8度及び今回の11度)と比べてかなり大きい。すなわち、映像と音像の方向差に対する反応は評定者による違いが極めて大きいといえる。ほとんどのテレビ視聴者は音響技術者ではないので両者の中間の約15度という値が実質的な許容限であると考えられる。15度であれば一般評定者は75%の確率で許容し、音響技術者でも”やや気になる”程度である。

ところで、本節の実験で用いた評価用映像と音声は、映像、音像とも一種類で、かつ静止しており、番組の形態としては最も単純なものの一つである。実際の番組においては映像、音声ともより複雑である。例えば音楽番組を例に取れば、歌手の映像は移動することがあり、伴奏音のような画面上に現れない音源の音声も加わることもある。そこで次の2. 6. 4節では、実際の番組への上記許容値の適用の妥当性について考察する。

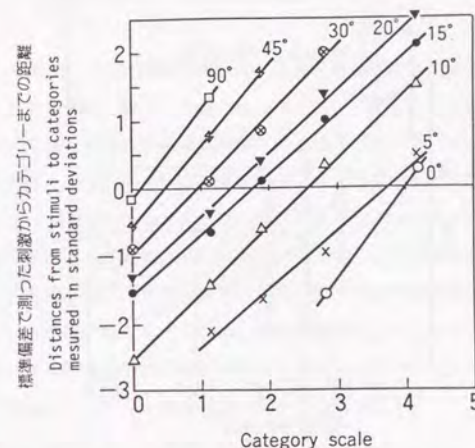


Fig.2.14a Regression of distances from stimuli to categories on the common category scale.(Experts)

刺激からカテゴリーまでの距離のカテゴリー尺度への回帰(専門家)  
(横軸と交わる位置のカテゴリー値が平均カテゴリー値を、また、傾きの逆数が標準偏差を表す)

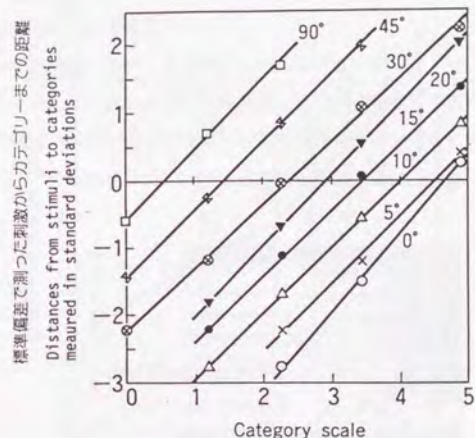


Fig.2.14b Regression of distances from stimuli to categories on the common category scale.(Non-experts)

刺激からカテゴリーまでの距離のカテゴリー尺度への回帰(非専門家)  
(横軸と交わる位置のカテゴリー値が平均カテゴリー値を、また、傾きの逆数が標準偏差を表す)

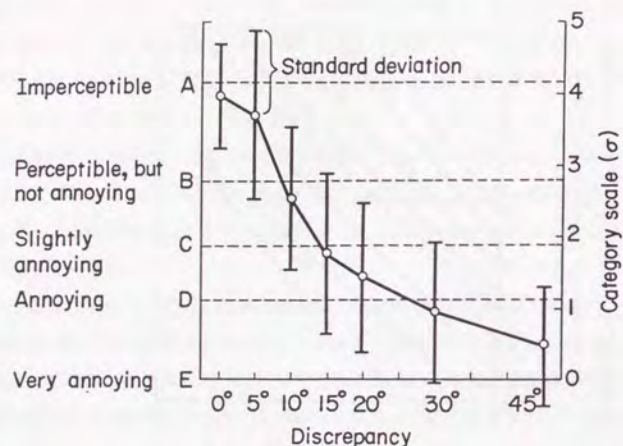


Fig. 2.15a Evaluation of discrepancy between video and sound images. (Experts)

映像と音像の方向差に対する評価

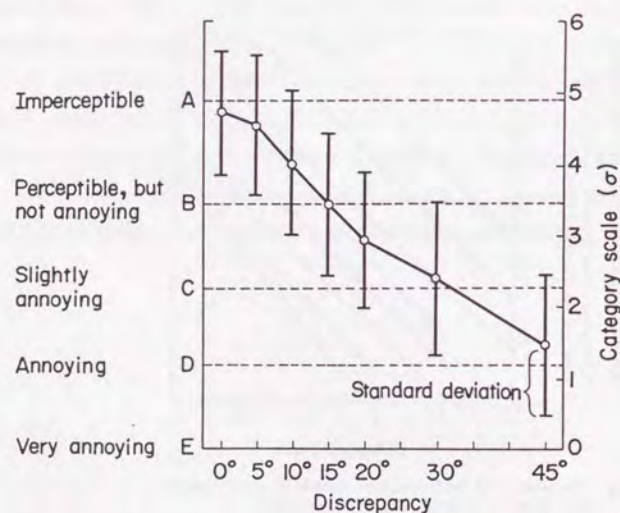


Fig. 2.15b Evaluation of discrepancy between video and sound images. (Non-experts)

映像と音像の方向差に対する評価

## 2. 6. 4. 実際の番組による許容限の検証

前述したようにテレビ番組の音声は、画面外の音源の音声を含む方が一般的である。また、映像も同じ位置に静止しているとは限らない。そこで、本節では前節で得られた映像と音像の方向差の許容限の実際の番組への適用の可否について実験的考察を行う。

評価の対象として、ハイビジョンで制作された歌謡番組の一部をNTSC方式に変換した映像と3-1方式4チャンネル(次章参照)で制作された音声を用いた。これは実際にハイビジョン番組として放送されたものの一部であり、評価実験を意図して制作されたものではない。映像は歌手の上半身を写したもので、画面のほぼ中央に位置しているが、歌手が歌に合わせて軽い動きをするため、常に画面中央というわけではない。また、音声としてボーカル以外に伴奏音が存在する。

音声の各再生チャンネルへの割当は以下のとおりである。

右および左チャンネル：伴奏、エコー

センターチャンネル：女性ボーカル

実験配置はFig. 2. 12に示したものと同一である。左右30度配置されたスピーカで伴奏とエコーを再生し、ボーカルを他の各方向のスピーカで再生する。評定者は歌手の映像を見ながら、その映像とボーカルの音像の方向の差を判断する。聴取レベルはピークで約80 dBAである。

映像の構図は刻々と変化するが、その代表的な絵柄をFig. 2. 16に示す。評価はTable 2. 4に示した5段階の尺度で行うが、映像の構図が時間とともに変化するため、歌手の口の動きがわかる場面で判断を行うように評定者に指示した。評定者は音響技術者7名、一般評定者5名で、繰り返しは5回である。



Fig. 2.16 A scene of the evaluated program.

評価用番組の一場面

次に実験結果について考察する。2. 6. 3節において44名の実験結果から求めた範ちゅうの中心値 (Table 2. 6 参照) を用いて今回の結果を尺度化したものを Fig. 2. 17 に実線で示す。ここで、破線は同じ評定者による Fig. 2. 2 に示した映像に対する結果である。この図から、実際の番組についても方向差の評価結果は0度方向を除いて前節の結果とほとんど差がないことが分かる。0度方向でやや評価が低いのは歌手の映像が画面中央に静止していないことによって生じる映像と音像の方向差のためと考えられるので、本質的な差ではないといえる。

以上のことから、画面に提示されない音源の音声为重畳された場合も、2. 5. 3で述べた専門家に対して11度、非専門家に対して20度という許容限はそのまま適用できると考えられる。

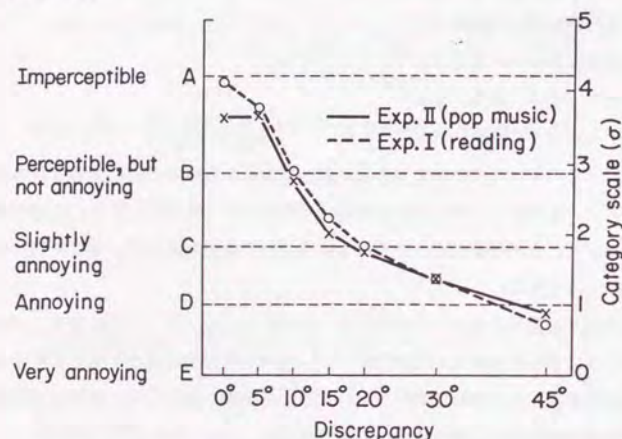


Fig.2.17 Evaluation of the discrepancy for an actual program (for 12 subjects and 5 times, studio room, female vocal, 80dBA.)

実際の番組を使用した場合の映像と音像方向のずれに対する評価  
(破線で示された実験用映像(音読)に対する評価と大差ないことがわかる)

## 2. 7. 映像と音の時間的な“ずれ”の影響について

テレビ番組制作時においては映像と音声の空間的なずれとともに、時間的なずれについても留意する必要がある。近年、フレームメモリー等の時間遅延を伴う機器の使用や、音声の後処理で制作するポストプロダクション方式の番組制作が一般的になるにつれ、映像と音の同期をいかにして保持するかが重要な課題となりつつある。

時間的なずれが過大になると、音像が映像に引きつけられる効果が減少することが知られている。Jack 他<sup>(14)</sup> (1973)によれば、テレビモニターとスピーカを20度離して置き、音声を選延させていった場合には、200 msec以上音声が遅れると音像が映像に引かれなくなるといふ。また、時間的なずれそのものも大きな違和感を生じさせることが知られている<sup>(15)</sup>。

このように映像と音像の方向感を一致させるためには、時間的な同期を保つことも重要な要素である。特に番組制作上ではVTRとATRの同期運転装置が有用である。そこで、本研究ではきわめて簡易な方法で両者を同期させることのできる手法を開発した。その詳細については6章で述べる。

## 2. 8. 結言

本章では大画面テレビを視聴する視聴の際に問題になると思われる映像と音像の方向差について一連の評価実験を行った。その結果をまとめると次のようになる。

- 1) 刺激の提示時間の長い、通常の視聴状態での方向差の許容限は一般視聴者で約20度、音響の専門家で約11度である。
- 2) 刺激の提示時間が短い条件での、音響の専門家による評価では、映像と音像の方向差の検知限は約4度、許容限は約8度である。
- 3) 音像と映像の方向差が検知できる状態においても、映像を注視しながら判断する音像の方向は映像の方向に偏移する。
- 4) 画面よりも広い音の広がり感を提示した場合でも、映像の提示によって狭まり、画面内に限定される場合がある。

これらの実験結果から判断すると、映像と音像の方向差はおおむね15度以内に抑えるべきである。現行の20インチ程度のテレビ受像機(画角約10度)では、音に特別の注意を払わない限り映像と音像の方向のずれが問題になることはないが、これから展開が予想されるHDTVの場合、画角が30度に達するので通常の2チャンネルステレオでは、方向差が気になる恐れがあることがわかる。

本章ではこの問題点を詳細に検討し、その解決方法としてHDTVの音声システムの多チャンネル化を提案する。

## 概要

本章では2章の結果を受けて、HDTVの音声システムのチャンネル数について考察を行う

## 3. 1. 緒言

2章では実験的考察によって映像と音像の方向の差は約15度以下に抑えるべきであることを示した。本章ではこれを受けてHDTVの音声システムのあり方について考察する。

HDTVと現行のNTSC方式の観視条件の大きな違いは視距離にある。NTSC方式のテレビはアスペクト比3対4の画面を6Hから7H(H: Picture height)の距離から観るのが標準であるのに対し、HDTVはアスペクト比は9対16の画面を3Hの距離から観るのが標準観視条件である<sup>(16)</sup>。これはHDTVが、視覚的な臨場感の向上が顕著となり始める見込み角30度程度を標準的な観視条件としているからである。このような広画角を持つHDTVの画面に従来の2チャンネルステレオ音声を組み合わせた場合、次節で述べるように映像と音像の方向の一致の点で問題がある。

映像と音像の方向差を必要な範囲内に抑え、かつ、音声による高い臨場感を実現する方式としてNHKはHDTV(ハイビジョン)の音声システムとして3-1方式4チャンネルステレオ(前方に3つのスピーカ、後方に2つ以上のスピーカを配し、前方3、後方1の独立4チャンネル信号で駆動する方式)を提案している<sup>(16)(17)</sup>。この方式の決定にあたっては本研究の結果が反映されている。ここでは、この方式の前方のチャンネル数について、2章で述べた映像と音像の相互作用の面から考察し、前方3チャンネルの妥当性について述べる。

なお、HDTVの音声方式に関しては、ITU(放送に関する国際推奨規格を審議する団体)によって審議されており、マルチチャンネルオーディオとして、3-1、3-0方式を含む3-2方式を基準として推奨することが決定される模様である。

## 3. 2. 広画角の映像と2チャンネルステレオの組み合わせの問題点

最初に現行のNTSC方式の場合について考察する。標準的な視距離といわれている7Hから視聴した場合、Fig. 3. 1に示すように画角が約10度となる。一方、2チャンネルステレオのステレオ効果はスピーカ開き角が60度位の時、最も良好と言われており<sup>(18)</sup>、現実に放送局や多くのスタジオが60度前後に配置している。しかし、スピーカを音響的に最適な60度の開き角に配置した場合には、映像と音像のずれが過大になる場合がある。例えば、画面の中央に映像があり、音像を2つのスピーカから同じ信号を再生して中央に定位させた場合には、中心からずれた位置で視聴すると音像だけが近いスピーカに片寄り、映像と音像の方向は30度ずれ、許容限を越える。実際にはスペースの問題もあって音声多重放送対応の受信機の多くはスピーカは画面の両サイドに内蔵あるいは近接させて配置されている。スピーカ開き角も必然的に狭くなる。この場合には映像と音像の方向のずれは10度以内で問題とならない。サウンドステージの狭さは、番組制作時に音場拡大装置等によってスピーカの外側にもサウンドステージが広がるような加工を施すことである程度救うことができる。

次に、HDTVの場合について考察する。HDTVの場合、標準的な聴取位置からの画面の見込み角は約33度になる。従って、画面の両側にスピーカを近接させて配置した場合でも、映像と音像の方向のずれが許容範囲を越える恐れがある。例えば、映像が画面の左端にあり、右に寄って視聴した場合、最大約30度ずれる。しかも、ステレオ効果の点からはこの開き角でも不足である。このようにHDTVの場合は、従来の2チャンネルステレオでは映像と音像の方向の一致とサウンドステージの広さのどちらの面からみても不十分である。

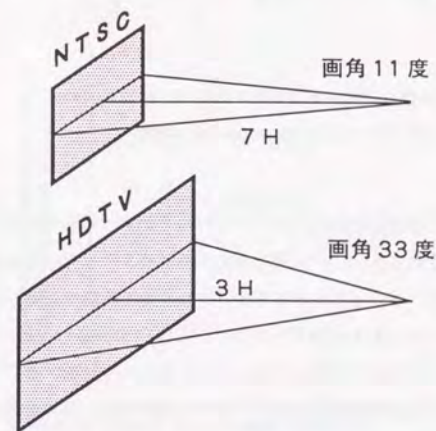


Fig.3.1 Viewing angle of TV screens.

### 3. 3. 最適な多チャンネルステレオ音声方式の検討

NHKではHDTVの高精細度と大画面に見合う音声方式に関する検討を4チャンネル以下のシステムを対象に行ってきた。その結果、伝送路の効率的な使用、広がり感等の音響心理効果、映像と音像の方向のずれなどの考察から、Fig. 3. 2に示す3-1方式の4チャンネルステレオがふさわしいとの結論にいたっている。この結論は以下に述べる実験的考察等を考慮して決定されたものである。ここで行われた一連の実験のうち、映像と音像の定位方向のずれに対する評価は本論文の2節で行った実験の結果を裏付けるものになっている。

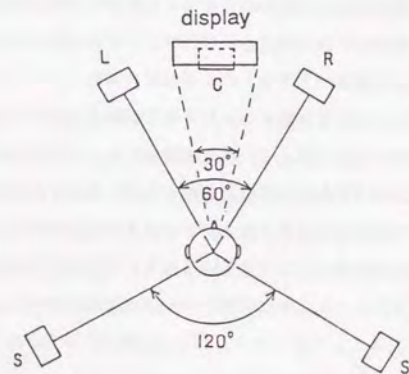


Fig.3.2 Loudspeaker layout of the 3-1 quadrasonic system.  
3-1方式4チャンネルステレオのスピーカ配置

NHKの検討において評価の対象となった7つのステレオ方式をFig. 3. 3のa~gに示す。aは従来の2チャンネルステレオ、bはそれにセンターチャンネル用スピーカを加えた3チャンネルステレオ、c~gが4チャンネルステレオ用の配置である。このうちgの方式は5つスピーカがあるが後方用の2つのスピーカは同一の信号を再生している。映像を再生するのに用いたディスプレイはNTSC方式の投射型プロジェクターであるが、映像はHDTVで収録した素材を9対16のアスペクト比を保ったまま、NTSCにダウンコンバートしたものを用いている。実験用プログラムはテレビ番組としての要素を

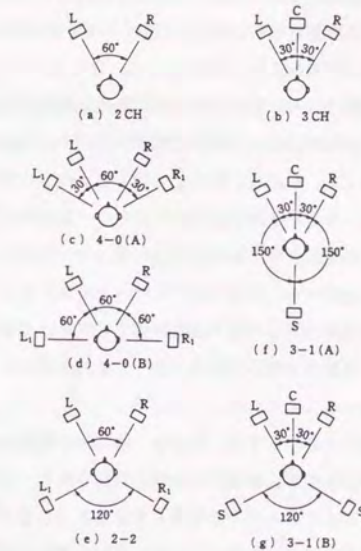


Fig.3.3 Experimental arrangements of loudspeakers for sound reproduction systems of HDTV.  
評価の対象となったHDTV用音声システムのスピーカ配置

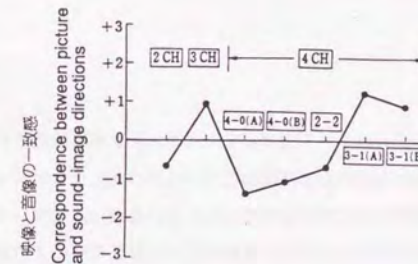


Fig.3.4 Result of the assessment of the correspondence between picture and sound-image directions.  
横にずれた聴取位置での映像と音像方向のずれに対する評価

数多く持つと考えられる軽音楽系の音楽番組を用いている。映像の内容はソロの歌手が踊り手を背景にして歌う場面である。音声のミクシングは各ステレオ方式ごとに行い、それぞれの方式で最良となるように調整されている。ただし、いずれの方式でもヴォーカルの音像は前方中央に定位するように配置されている。

上記7つのステレオ方式で制作された番組に対して主観評価実験を行った結果、評定者が中央で聴取した場合には、臨場感に優れた2-2方式、3-1(B)方式、4-0(B)方式が評価が高いことがわかった。しかし、中心から約70cm左にずれた位置で聴取した場合には、2-2方式と4-0方式は映像と音像の方向の一致感の点で3-1方式よりも劣ることが分かった。評価項目のうち「映像と音像の定位の一致度」に関する結果をFig. 3. 4に示す。この結果から、前方中央にスピーカのある3-0、3-1(A)、3-1(B)方式がこの評価項目に関しては7段階評価で2ポイント優れていることが明かである。これらの結果に、映画の音響システム(3-1方式が主流)との互換性を加味して3-1(B)方式を採用することとした。

HDTVの音声方式を決めるにあたっては、定位の一致だけが議論されたのではなく、臨場感や既存のシステムとの互換性等、総合的な品質が議論された。定位の一致だけを問題にするならば、前方にさらに多くのスピーカを置く方式もありうる<sup>(19)</sup>。しかし、2節の実験結果が示すように、一般の視聴者の映像と音像の定位のずれの許容値は20度である。画面に表示される映像に対する音はすべてセンタースピーカから再生するようにすれば、映像が端に寄った場合でも映像と音像の方向のずれは15度以下であり、許容範囲内である。従って、定位の一致に関してはセンターチャンネルを再生するセンタースピーカを加えるだけで必要十分であるといえる<sup>(20)</sup>。

### 3. 4. 結言

本章ではHDTVの音声システムが3-1方式4チャンネルステレオに決定されるに際し、映像と音像の相互作用に関する知見が果たした役割について述べた。テレビジョン音声システムの仕様に初めて映像と音声の相互作用が考慮されたといえよう。なお、4チャンネル音声の伝送については4チャンネルのPCM音声を準瞬時圧伸差分PCMにより帯域圧縮し、ハイビジョンの伝送方式であるMUSE信号に多重する方式が開発されている<sup>(21)</sup>。MUSE信号を伝送あるいは記録することができれば音声は常に4チャンネルが確保される。

3-1方式の特長は必要最小限のチャンネル数で、映像と音像の方向の一致、広い聴取範

囲及び高い臨場感が得られることである。2チャンネルステレオと3-1方式4チャンネルステレオを聴取範囲の面から比較すると、左右のスピーカ間隔が3m程度の場合、3-1方式を右または左に2座席分ずれた位置で聴取した時の評価が2チャンネルステレオを中央で聴取するよりも優れていることがわかっている。

一方、チャンネル数が増加すると番組制作が煩雑になることが懸念される。そこで、次章では番組制作を能率的に行うための機器の開発について述べることにする。

## 概要

複数のコヒーレントな音源により合成される波面の解析をもとに、3-1方式等の多チャンネルステレオによる聴取範囲について考察する。さらに、波面の観測を利用した番組制作用機器の開発を行う。

## 4. 1. 緒言

前章ではテレビ試験時における映像と音像の方向差の許容限の実験的考察から、H D - T Vの音声システムにはセンターチャンネルが必要であることを述べた。前方に3チャンネル(左右30度と正面方向)にスピーカを配し、独立な音声を再生することにより、映像と音像の方向差を許容限以下に抑えると同時に、広いサウンドステージを実現することができる。

しかし、多チャンネルステレオで番組制作を行う際には2チャンネルステレオにはない問題点が発生する。その一つが同じ音像方向を与えるために何種類もの方法があること、そして、もう一つが360度全周にわたる音像の方向を監視しなければならないミキシング技術者の負担の増大である。

最初の問題について一例をあげる。右とセンタースピーカーの中間に音像を提示する場合に、次の3種類の信号分配法がある。

- (ア) センターと右のチャンネルだけを使用する、
- (イ) 3つのチャンネルを使用する、
- (ウ) 右と左のチャンネルだけを使用する。

2チャンネルステレオでは当然のことながら(ウ)の方法しかない。どの方法でも中央の聴取位置における聴感上の印象は似ており、差を検知するのは容易でない。ところが、中央以外の聴取位置では上記3種類の方法の差が音像方向の偏りの差というかたちで現れ、聴取範囲の広さが異なる。

次に、ミキシング技術者の負担増大の例では、音像の前後判断が挙げられよう。後方にスピーカを配置した多チャンネルステレオではミキサーは音像方向の前後の識別を行わなければならない。良く知られているように、人は音像の前後判断を誤りやすいので<sup>(28)</sup>、これはかなりの負担となる。

そこで、本章では合成音場の波面の解析によって、信号の分配法と聴取範囲の関係について考察する。あわせて、音像方向の監視に関するミキサーの負担を軽減するための視覚的モニター装置を開発する。

## 4. 2. 波面による音場解析の特長

聴取範囲を予測する手法は多数存在するが、ほとんどが両耳の外耳道入り口の音圧に着目した議論を行うものである。これは厳密な方法であるが、聴取範囲を推測するためには、聴取位置ごとに外耳道入り口の音圧を評価する必要があり、膨大な計算が必要になる。一方、聴感実験による方法は多数の聴取位置で行う必要があり時間がかかるうえに、適用範囲の広い一般的な議論を展開しにくい。

そこで本章では、複数のスピーカによって合成される波面を用いて音像の方向を予測し、聴取範囲を推定する手法をとる。この方法は合成された波面(以後、合成波面)から、両耳に発生する位相差を近似的に求め、音像方向を予測するものである。頭部による音波の回折を無視するので低い周波数にのみ適用できる近似的な方法であるが、波面は音場全体にわたって計算することが容易なので、本章の目的に良く合致する。

聴取範囲を決定する要素として音像の方向を取り上げる理由は、2章で述べたように、映像を伴う音声システムにとって映像と音像の方向の一致がきわめて重要であるからである。

ここで簡単に波面の理論について述べる。

「単一音源による音波は波面の法線方向に伝搬し、かつ音像もその方向に知覚される。従って、頭の大きさに比べ十分長い波長の音波では、頭による波面の擾乱が無視できると考えられるので、複数音源による音場においても合成波面の放線方向に音像が知覚されるであろう」という仮説が波面の理論の論拠である。実際、2チャンネルステレオで左右のスピーカに加える信号の間に位相差がなく、かつ、聴取者が波面に正対している場合は500 Hz位まで実際の音像方向とよく一致することが確かめられている<sup>(22)</sup>。しかし、これ以外の場合、すなわち聴取者が波面に正対していないときやチャンネル間に位相差が存在する場合は波法線の方向と音像の方向は必ずしも一致しない<sup>(23) (24)</sup>。この不一致は主に次の要因に起因する。

(1) 波面の進行方向が同一であっても、位相速度によって両耳に生じる位相差が異なり、音像方向が違ってくる<sup>(19)</sup>。

(2) 波面の形状が頭の大きさに比べて急激な変化をする。

波面による音像方向の予測には上記のような問題点が存在するものの、これらは波面の法線方向を計算する代わりに波面そのものを音場全体にわたって表示すれば、波面の形状や、波面の密度すなわち音波の位相速度も明らかになるので、すべて予測が可能である。また、1.3 kHz以上の周波数の音像の方向は位相差よりも両耳の外耳道入り口に生じる音圧のレベル差によって決定されるため、波面の理論は成立しないが、通常の番組音のエネルギーは500 Hz程度の中低域に集中するので、低い周波数のみの解析で実用上十分である。

そこで本章では、まず、2チャンネルステレオの聴取範囲の推定に波面に注目した解析法が有効であることを述べた上で、各種の信号配分をした場合の前方3チャンネルシステムにおける聴取範囲について考察する。

さらに、波面の概念を等時間遅延の面に発展させる方法について考察し、帯域幅を持つ信号に適用可能な表示法を提案する。

#### 4. 3. 線形の収音再生系における波面の導出法

ここでは波面の求め方について述べる。波面を計算するために音源信号から音場までの伝達関数を導入する。収音音場の音源信号  $s(t)$  と再生音場の一点  $\vec{r}$  における音圧  $p(\vec{r}, t)$  は収音再生及び伝送の過程で非線形処理をしなければ、チャンネル数にかかわらず線形な関係にある。従って、系のインパルス応答  $h(\vec{r}, t)$  で音場を表現することができる。すなわち

$$p(\vec{r}, t) = h(\vec{r}, t) * s(t) \quad (4.1)$$

ここで\*は畳み込み積分を表す。同じことを周波数領域で表現すれば次式のようになる。

$$P(\vec{r}, \omega) = H(\vec{r}, \omega) S(\omega) \\ = |H(\vec{r}, \omega)| e^{j\theta(\vec{r}, \omega)} S(\omega) \quad (4.2)$$

ここで、 $H(\vec{r}, \omega)$  は  $h(\vec{r}, \omega)$  のフーリエ変換、 $\theta(\vec{r}, \omega)$  は  $H(\vec{r}, \omega)$  の位

相である。再生空間における波面構造は位相に関する項の  $\vec{r}$  に関する依存性を計算することによって明らかにすることができる。 $H(\vec{r}, \omega)$  の位相を  $\theta(\vec{r}, \omega)$  とすると、位相に関して次の3つの遅延量が定義されている。

$$(1) \text{ 位相遅延 } \tau_{ph} = \frac{\theta(\omega)}{\omega} \quad (4.3)$$

$$(2) \text{ 群遅延 } \tau_{gr} = \frac{d\theta(\omega)}{d\omega} \quad (4.4)$$

$$(3) \text{ 波頭遅延 } \tau_{tr} = \lim_{\omega \rightarrow \infty} \frac{\theta(\omega)}{\omega} \quad (4.5)$$

$\tau_{ph} = \text{const}$  の線の集合が等位相面すなわち波面となる。また、 $\tau_{gr} = \text{const}$  が等群遅延面を表す。

以下の節で複数の音源による合成波面を実際に計算し、図式化する。まず、2チャンネルステレオシステムでチャンネル間にレベル差を与えて音像方向を設定する場合、合成波面が中央聴取位置で一様になることを明らかにする。次に、チャンネル間に位相差が存在する場合、および、多チャンネルステレオシステムの場合の波面について考察する。

#### 4. 4. 複数音源により合成される波面の解析

##### 4. 4. 1. 2チャンネルステレオシステムの場合

入力信号  $S(\omega)$  が伝達関数  $L(\omega)$  と  $R(\omega)$  を介して左右のスピーカに供給される2チャンネルステレオシステムを考える。スピーカが平坦な周波数特性と全指向特性を持ち、自由音場で再生すると仮定した場合には、空間の一点  $(x, y)$  の速度ポテンシャル  $\Phi(x, y, \omega)$  は次式で与えられる。

$$\Phi(x, y, \omega) = \left\{ \frac{L(\omega) \exp(-jk r_l)}{r_l} + \frac{R(\omega) \exp(-jk r_r)}{r_r} \right\} S(\omega) \quad (4.6)$$

ここで  $r_l, r_r$  は左右のスピーカから再生音場における点  $(x, y)$  までの距離である。

また、 $k=2\pi/\lambda$  である。なお、速度ポテンシャルとは音圧や粒子速度と次の式で結びついた仮想的なポテンシャルであり、音場の解析に常用される。

$$p = \rho \frac{\partial \phi}{\partial t} \quad (4.7)$$

$$\vec{u} = -\text{grad} \phi$$

ここで  $\rho$  は媒質密度を表す。上式を用いて各種の条件での波面を計算する。

(1) 左右のチャンネル間にレベル差だけが存在する場合

この場合、 $L(\omega)$ 、 $R(\omega)$  は定数となるので、次のように置くことができる。

$$L(\omega) = A, R(\omega) = B$$

このとき、伝達関数  $H(x, y, \omega)$  は次のようになる。

$$H(x, y, \omega) = \left[ \frac{A \cos(kr_l)}{r_l} + \frac{B \cos(kr_r)}{r_r} \right] - j \left[ \frac{A \sin(kr_l)}{r_l} + \frac{B \sin(kr_r)}{r_r} \right] \quad (4.8)$$

この場合の波面は上式で位相項一定となる点の集合で与えられるので、次のように表すことができる。

$$\arctan \left[ \frac{\frac{A \sin(kr_l)}{r_l} + \frac{B \sin(kr_r)}{r_r}}{\frac{A \cos(kr_l)}{r_l} + \frac{B \cos(kr_r)}{r_r}} \right] \quad (4.9)$$

スピーカ間隔 3 m のステレオ再生用配置で、 $A=B$  の場合と  $A/B=0$ 、32 の場合の 500 Hz における上式の計算結果を Fig. 4.1 と 4.2 に示す。これは位相が  $\pi/4$  変化するごとに白黒を反転させて表示したものである。

これらの図より次のことが分かる。

(ア) 2つのスピーカからの距離の差が半波長になる場所に、スピーカを焦点とする双曲線上に波面が急激に曲がる“うねり”が生じている。

(イ) また、そのうねりとうねりの間には比較的一様な(法線方向が一定の)波面が形成されている。

(ウ) チャンネル間にレベル差を与えた場合も、一様な波面のできる領域はレベル差の無い場合とほとんど同じであり、波面の法線方向の向きだけが変化する。

一様な波面の領域の幅は、500 Hz の場合、頭の大きさよりもかなり広く、波面の法線方向の向きもかなり広い幅で一定であることが分かる。

ところで、一様な波面の形成される領域は周波数に依存するが、スピーカから等距離の場所には周波数に無関係に一様な波面が形成される。これは言い替えると、「チャンネル間にレベル差だけが存在する2チャンネルステレオ再生音場では、スピーカからほぼ等距離の領域に、波面の法線方向の向きに関して平面波と等価な音場が再現されている。」と考えてよいことを示唆している。別の言い方をすれば、2チャンネルステレオは、「チャンネル間にレベル差を与えることにより、スピーカからほぼ等距離にある、ある幅を持った領域の波面の向きを制御できる、」という意味で音場制御システムである。

もちろん理想的な平面波ではないことは明かで、音圧分布や位相速度については異なる。例えば、スピーカを結ぶ線に近いほど波面の間隔が実音源の場合よりも広がっており、その場所の波の位相速度が通常の音速よりも速いことが分かる。音波の位相速度が速くなると、両耳の外耳道入り口に生じる音圧間の位相差が通常よりも小さくなる。そのため、実際の波面の法線方向よりも顔の正面方向に音像が寄ることになる。事実、左右のスピーカから同じレベルで音を再生しても、聴取者がどちらかのスピーカの方を向いた場合には、音像が顔を向けたほうのスピーカ寄りに定位することが知られている。

上記に関する補正は必要であるが、チャンネル間レベル差を用いる2チャンネルステレオは単純に両耳の外耳道入り口の音圧だけを制御するシステムではなく、波面の法線方向に関して音場を制御するシステムと考えることができる。

(2) チャンネル間に位相差がある場合

左チャンネルが  $\pi/2$  進んだ場合の波面を Fig. 4.3 に示す。この図が示すように位相差がある場合は波面のうねりが位相遅れをもつスピーカ側に移動するものの、波面の法線方向はあまり変化していない。しかし、実際の聴感実験によれば、位相の先行するスピーカの方に音像方向が寄ることが知られている。これは左右の耳が空間的に離れており、片方の耳が波面のうねりにかかる為、両耳間に位相差が生じたものと思われる。もちろん、

両耳に生じる音圧差も無視できない。このようにチャンネル間に位相差が存在する場合は、一点における法線方向でなく、両耳の間隔を考慮した平均的な法線方向を扱う必要がある。

### (3) レベル差と位相差がある場合

チャンネル間にレベル差と位相差を適当に与えると、スピーカの外側に音像が定位することが知られている。これは両耳間のレベル差、位相差を音像の方向に音源がある場合と等しくすることができるからである。例えば、左右のスピーカに

$$L(\omega) = 0.3e^{j0.25\pi} R(\omega) \quad (4.10)$$

の条件の信号を供給し、スピーカを見込む角度が90度の中央聴取位置で聴取した場合、右側方に音像が定位するという<sup>(28)</sup>。このときの波面の計算結果をFig. 4.4に示す。この図からスピーカを見込む角度が90度の場所における波面の法線方向は右スピーカの方に近く、右側方ではないことが分かる。それにもかかわらず音像が右側方に定位する理由は次のように波面の間隔で説明できる。Fig. 4.4は聴取位置での波面の間隔が実音源のそれ（例えばスピーカのごく近傍）に比べ非常に狭くなっていることを示している。波面の間隔が狭いために、両耳に生じる音圧の位相差が大きくなり、音像が顔の横方向に寄る（実音源の場合、真横から来る音波が両耳に最も大きな位相差を発生させる）。

ただし、この条件での両スピーカから等距離の線上の音場は波面がうねる場所でもある。Fig. 4.4からも中央の線上の聴取位置の波面の様子はFig. 4.1, Fig. 4.2と比較して狭い範囲であることが分かる。これは聴取範囲が狭く限られることを意味しており、経験上の聴感的印象と符合する。

以上3種類の条件について、2チャンネルステレオによる波面を計算し、その特徴を述べた。その結果、聴取範囲の広さに限って言えば、チャンネル間に位相差を与えず、レベル差だけで音像方向を制御する方法が優れていることが示唆された。

#### 4.4.2. 多チャンネルステレオシステムの場合

前節の計算により、2チャンネル再生でチャンネル間に位相差がない場合、スピーカから等距離の場所の波面の法線方向は比較的広い範囲で一定であることが分かった。ここでは、前方3チャンネルのシステムにおける波面を計算し、信号配分と聴取範囲の関係について考察する。

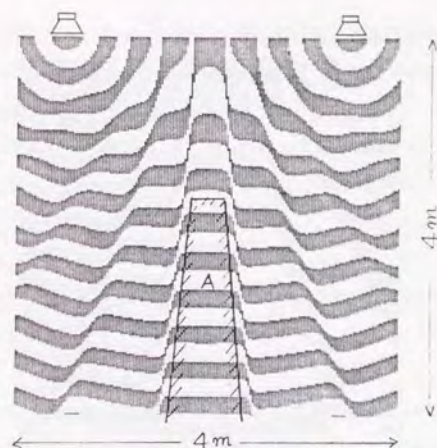


Fig.4.1 Contours of constant phase for  $L=R$  (freq:500Hz, sound velocity: 341m/s, free field, omni-directional loudspeakers). The black and white areas appear alternately at intervals of  $\pi/4$ .

2スピーカを同レベル同位相の500Hzの純音で駆動したとき（ $L=R$ ）に自由音場に生じる波面の計算値（ $\pi/4$ ごとに白黒を反転させて表示）  
両スピーカからほぼ等距離の場所（図の領域A）に法線方向が2スピーカ  
の中間に向く一様な波面が形成される



Fig.4.2 Contours of constant phase for  $L=0.32R$  (freq:500Hz, sound velocity: 341m/s, free field, omni-directional loudspeakers). The black and white areas appear alternately at intervals of  $\pi/4$ .

$L=0.32R$ となるレベル差を与えて駆動した場合の波面の計算値  
（500Hz純音、 $\pi/4$ ごとに白黒を反転させて表示）  
Fig. 4.1の領域Aに相当する場所の波面が一様に傾斜を持つ

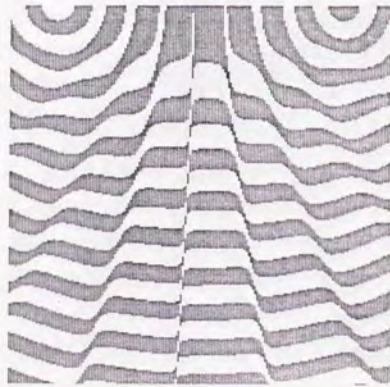


Fig.4.3 Contours of constant phase for  $L=e^{-j0.5\pi}R$  (freq: 500Hz, sound velocity: 341m/s, free field, omnidirectional loudspeakers). The black and white areas appear alternately at intervals of  $\pi/4$ .

チャンネル間に  $\pi/2$  の位相差を与えた場合の波面の計算値  
(500Hz 純音、 $\pi/4$ ごとに白黒を反転させて表示)  
Fig. 4. 1の領域Aに相当する領域が右に移動する



Fig.4.4 Contours of constant phase for  $L=0.3e^{j0.95\pi}R$  (freq:500Hz, sound velocity: 341m/s, free field, omnidirectional loudspeakers). The black and white areas appear alternately at intervals of  $\pi/4$ .

中央聴取位置で音像が右横に生じるようにレベル差と位相差を与えた場合の波面 (500Hz 純音、 $\pi/4$ ごとに白黒を反転させて表示)  
Fig. 4. 1の領域Aに相当する領域は形成されない

4. 1の緒言で述べたように、3チャンネルシステムでは同じ音像方向を与える複数の信号配分法がある。ここでは正面方向に音像を提示する次の4種類の方法について波面を比較し、それらの聴取範囲を比較する。

- (1)  $L:C:R=2:1:2$
- (2)  $L:C:R=1:1:1$
- (3)  $L:C:R=1:2:1$
- (4)  $L:C:R=1:-1:1$

これらは中央聴取位置では同じ物理条件となり、両耳の外耳道入り口の音圧に着目する方法では区別できないものである。

計算結果をFig. 4. 5に示す。この図から、 $L:C:R=1:2:1$ のようにセンタースピーカのレベルが大きい場合を除き、一様な波面の形成される領域は拡大されないことが分かる。これは、2つのスピーカから等距離の場所の集合が直線をなすのに対し、3つのスピーカから等距離の場所は一点に限られるためである。従って、音像をある方向に定位させる場合に、3個のスピーカを用いることは、聴取範囲を狭める結果になりうる。ことが分かる。



Fig.4.5 Contours of constant phase delay for 3 front loudspeakers.

前方3スピーカによる等位相面の例 ( $L:C:R=1:2:1$ の場合を除くと各スピーカから等距離の聴取場所周辺の波面の変化が激しい、すなわち聴取範囲を広げる効果はないといえる)

#### 4. 5. 広帯域信号への波面の概念の拡張

波面は広い範囲にわたって可視化し、位相速度も考慮すれば、音像の方向を予測することができ、聴取範囲の推定に有効な手段である。しかし、波面は等位相の面を図式化しているため、その間隔は周波数に依存する。そのため、異なる周波数の波面を重ね合わせて比較することはできない。それを可能にするためには、位相ではなく時間を単位にした表示が必要である。時間を単位とすれば、異なる周波数の波面を重ね合わせることができ、波面の形状の周波数依存性等を明示することができる。そこで本節では、波面の概念を拡張し、時間を単位にする方法を検討する。まず、4. 2節で示した群遅延を用いる方法について検討し、次に、音源信号と音圧の相互相関関数を用いる方法を検討する。

##### 4. 5. 1. 等群遅延面に関する考察

Mertens<sup>(28)(29)</sup>によれば2チャンネルステレオ再生によって両耳に生じる音圧間の群遅延時間差と音像方向がよく対応するという。もしこれが正しければ、群遅延の等しい点を結んだ“等群遅延面”によって音像方向が推定できるはずである。しかし、彼の聴感実験は、中央の聴取位置で、しかも左右のチャンネルの間に位相差の無い場合に限られている。そこで実際に等群遅延面を描いて波面との比較を行う。

群遅延は位相 $\theta(\vec{r}, \omega)$ を $\omega$ で微分すれば得られるが、 $\Delta\omega$ が十分小さな値であれば次式を直接数値計算をしても得られる。

$$\frac{\Delta\theta}{\Delta\omega} = \frac{\theta(\omega + \Delta\omega) - \theta(\omega)}{\Delta\omega} \quad (4.11)$$

上式から求めた $L=R$ における計算結果をFig. 4. 6に示す。これを見ると、スピーカから等距離にある場所の近傍ではFig. 4. 1の波面とよく似ていることがわかる。図面は省略するが、チャンネル間にレベル差がついた場合も、スピーカから等距離の場所では波面と等群遅延面は良く似ている。

一方、中央以外の場所では両者は異なっており、特に2つのスピーカからの距離差が半波長の整数倍の場所では等群遅延時間面の変化の仕方は経験的な音像方向とは対応しない。Mertensは両耳における外耳道入り口の音圧の時間重心の差(群遅延時間差と等価)にもとづく予測式と聴感実験の結果が一致したことを述べているが、境ら<sup>(30)</sup>のAM変調音を用

いた音像定位の実験によれば、低い周波数の定位方向と対応する物理量は信号のキャリアの位相差であり、エンヴェロープの時間差すなわち群遅延時間差ではない。中央の聴取位置では時間重心による予測式も位相差による予測式もほとんど同じ式となるので、結果的に同じ予測値を与えるが、一般の聴取位置に群遅延時間差を適応することは適当でないと考えられる。



Fig.4.6 Contours of constant group delay for  $L=R$  (freq: 500Hz, sound velocity: 341m/s, free field, omni-directional loudspeakers). The black and white areas appear alternately at intervals of 0.5ms.

等群遅延面

##### 4. 5. 2. 入出力相互相関関数による等遅延面に関する考察

波面を時間の尺度で表現するもう一つの方法は音源信号と再生音場の一点の音圧の相互相関関数を用いる方法である。等群遅延面が聴感と対応しない理由は、信号のエンヴェロープの遅延時間に着目するためであり、聴感と対応づけるためには、瞬時波形の遅延時間を計算する必要がある。そこで、ここでは波面に等群遅延面の長所を組み合わせるために、音源信号 $s(t)$ と再生音場の音圧 $p(\vec{r}, t)$ の相互相関関数

$$R(\vec{r}, \tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T s(t) p(\vec{r}, t + \tau) dt \quad (4.12)$$

の最大値を与える時間を系の遅延時間と見なし、そのコンターを描く。ここで $T$ は積分時

間を表す。この方法によれば位相情報も含んだ等遅延時間面が求められる。

この方法による波面は信号の系の伝達関数だけでなく自己相関関数にも依存する。一例として、音像方向の評価実験に良く用いられる狭帯域雑音を音源信号とし、相互相関関数によって系の遅延時間を求めた。Fig. 4. 7に中心周波数500Hz、帯域幅1/3オクターブの狭帯域雑音による等遅延面を示す。この結果はFig. 4. 1の波面と良く似ており、波面と同等の扱いができる。

異なる周波数の波面を重ね合わせた表示の例として、2チャンネルステレオにおいて、200Hzから800Hzまで重ね合わせたものをFig. 4. 8に示す。明確な縞模様の生じている場所は波面の形状が周波数にほとんど依存しない場所である。両スピーカからほぼ等距離の場所と、スピーカの近傍において、波面が揃っており、明確な音像方向が知覚できる場所であると推測できる。



Fig.4.7 Contours of constant time delay by cross-correlation for L=R (freq:500Hz, sound velocity: 341m/s, free field, omni-directional loudspeakers). The black and white areas appear alternately at intervals of 0.5ms.

音源と聴取位置の音圧の相互相関関数で規定した等時間遅延面

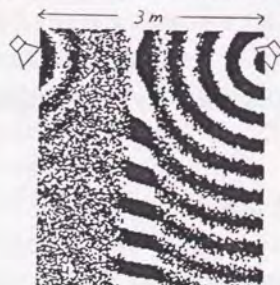


Fig.4.8 Statistical contours of constant time delay for L=0.536R (f=200 to 800Hz).

統計的等遅延面（音場の各点の遅延時間を200Hzから800Hzまでの間のランダムな周波数で計算したもの）

#### 4. 6. 多チャンネルステレオ用音像モニター装置の開発<sup>(27)</sup>

##### 4. 6. 1. 音像方向の視覚的表示装置の開発の背景

前節までに述べてきた波面の理論を応用して、多チャンネルステレオ再生による音像の方向をリアルタイムで監視するための装置を開発した。実際の番組制作や番組送出の監視に使用することができる。以下、その概要を述べる。

3章で触れたように、近年、HDTVの音声システムとして3-1方式4チャンネルステレオを含めていくつかの多チャンネルステレオが提案されている。共通する特徴は前方に3チャンネル以上有すること、後方用のチャンネルを持つことである。それらは映像と音像の方向の一致、高い臨場感、広い聴取範囲を目的としている。しかしながら、多チャンネルステレオは番組制作を複雑化し、音声ミクサーに煩雑な作業を要求する恐れがある。耳だけを頼りに多チャンネルステレオの複雑なミキシングをするのは容易なことではない。2チャンネルステレオであれば前方60度の範囲に注意を集中していれば済むが、多チャンネルステレオでは360度全周に注意を払わなければならない。従って、音声の視覚的な表示があれば、ミキシング作業の補助手段として大いに役立つであろう。

現在使われている視覚的な表示法として、オシロスコープによるリサーチの観測がある。これは2チャンネルステレオでは非常に有用であるが、多チャンネルステレオに対しては不十分である。その理由は前方からの音と後方からの音の方向の違いを表示できないからである。また、他にも2-2方式の4チャンネルステレオのために、いくつかの表示法が考案されている<sup>(53)</sup>。それらの中には前後の音像の方向を区別できる方式もあるが、聴感的な音像方向と必ずしも対応しない。また、2-2方式のスピーカ配置だけに適応し、スピーカ配置が異なる場合や、さらにチャンネル数が増加した場合には使用できない。

ここで述べる方法は、任意の多チャンネルシステムで音像の方向を正しく表示することを意図したものである。4. 3で述べた音場の可視化技術を応用し、音場における波面の法線方向を音圧と粒子速度の積によって計算し、音像の定位方向を視覚化するものである。この方法によれば、従来のリサーチによる方法では難しかった、360度すべての方向の音像表示が可能になる。

##### 4. 6. 2. 波面の法線方向の検出法

本装置は4. 3で述べた「低い周波数における音像の定位方向は合成波面の法線方向に

近似的に等しい。」という理論に基づいている。この波面の法線方向は次に示すように実際の音場では音圧と粒子速度から検出することができる。

周波数応答関数が次式で与えられる再生システムを考える。

$$\Phi(x, y, \omega) = A(x, y, \omega) e^{-j\theta(x, y, \omega)} S(\omega) \quad (4.13)$$

ここで $\Phi$ は音場の一点 $(x, y)$ における速度ポテンシャルで $A$ と $\theta$ は周波数応答関数の振幅と位相である。一方、 $S(\omega)$ は音源信号である。 $(x, y)$ における音圧 $P$ と粒子速度 $\bar{U}$ は定義により次式で与えられる。

$$\begin{aligned} P(x, y, \omega) &= -\rho \frac{\partial \Phi}{\partial t} = j\omega\rho\Phi(x, y, \omega) \\ &= -j\omega\rho A(x, y, \omega) e^{-j\theta(x, y, \omega)} S(\omega) \\ \bar{U}(x, y, \omega) &= -\text{grad} \Phi(x, y, \omega) \\ &= (-\text{grad} A + jA \text{grad} \theta) e^{-j\theta(x, y, \omega)} S(\omega) \end{aligned} \quad (4.14)$$

ここで $\rho$ は媒質の密度である。

波面の法線方向は $\text{grad} \theta$ で与えられる。この勾配は次式で示されるように音圧と粒子速度のクロススペクトル $P^* \bar{U}$ の実部に現れる。

$$\text{Real}[P^* \bar{U}] = \omega \rho A^2 \text{grad} \theta |S(\omega)|^2 \quad (4.15)$$

このクロススペクトルの実部は周波数 $\omega$ における音響インテンシティ（音圧と粒子速度の積の時間積分）と等しいので、音響インテンシティの計測手段と同一の方法で波面の法線方向を測定できることになる。すなわち音圧 $p(t)$ と粒子速度ベクトル $\bar{u}(t)$ を狭帯域バンドパスフィルタを通した後、積を取り時間積分を行えばよい。全周波数の平均的な波面の向きを観測する場合には、次式に示すインテンシティ $\bar{I}(t)$ の $x$ 成分と $y$ 成分をオシロスコープの水平、垂直軸に入力すれば法線方向を観察できる。

$$\bar{I}(t) = \int_{t-T}^t p(\xi) \bar{u}(\xi) d\xi \quad (4.16)$$

ここで、 $T$ は積分の時間窓長である。

#### 4. 6. 3. 波面の法線方向のモニター装置

スピーカの配置と聴取位置があらかじめ決まっていれば、聴取位置における直接音による音圧と粒子速度はスピーカの入力信号から直接計算できる。

一例として Fig. 3. 1 に示した 3-1 方式の 4 チャンネルステレオのスピーカ配置について考察する。中心の聴取位置の音圧 $p(t)$ は次式で与えられる。

$$p(t) = k_1 \{L(t) + R(t) + C(t) + 1.4 S(t)\} \quad (4.17)$$

ここで $L, R, C, S$ はスピーカの入力信号である。 $S$ の係数が 1.4 になっているのは 3-1 方式においては $S$ 信号のエネルギーを他の各チャンネルからのエネルギーと同一とするため、後方 2 つのスピーカから各々 3 dB 減衰させて再生するように決められているためである。

Fig. 3. 1 のスピーカ配置に対応する粒子速度ベクトル $\bar{u}(t)$ は次のようになる。

$$\bar{u}(t) = k_2 \{0.5(R(t) - L(t)), 0.866(L(t) + R(t)) + C(t) - 0.7S(t)\} \quad (4.18)$$

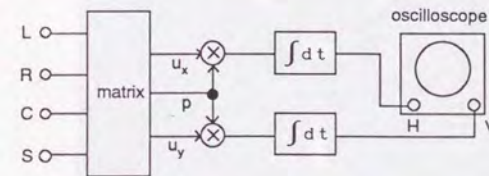


Fig.4.9 Schematic diagram for the monitor scope to display sound image direction in quadraphonic stereo.

4チャンネルステレオ用音像方向表示装置の構成

純音に対しては音圧と粒子速度の積の時間積分が聴取位置における波面の法線方向に一致する。従って、この  $x$  成分と  $y$  成分を各々水平、垂直成分としてオシロスコープに入力すると波面の法線方向がリアルタイムで観察できる。広帯域音の場合には全周波数の平均的な方向になる。また、積分前の瞬時値をオシロスコープに加えると、チャンネル間の位相差を観察することができる。本装置の系統図を Fig. 4. 9 に示す。

上式は 3-1 方式に対応しているが、他の多チャンネル方式についても同じ考え方で  $p$  と  $\bar{u}$  を計算することにより対応可能である。

#### 4. 6. 4. 試作機による表示の例

現在、本方法によるモニターの試作機が制作現場で試験運用中である。実際の番組制作においては音像方向だけでなく、チャンネル間の位相差も監視する必要があるため、音圧と粒子速度の積の瞬時値を表示できる構成としている。各チャンネルのレベルを同時に表示することにより、レベルの監視を可能とする設計とした。表示の一例を Fig. 4. 10 に示す。

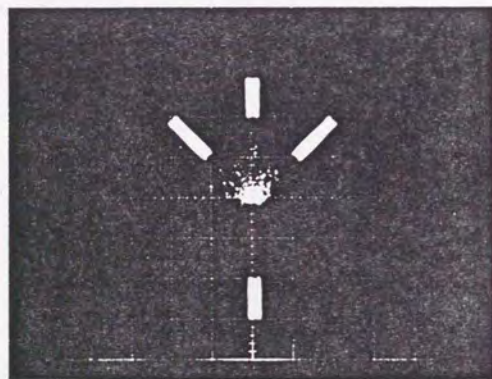


Fig.4.10 An example of display.

#### 4. 7. 結言

本章では波面の可視化によって多チャンネルステレオの聴取範囲を推測する方法を考察した。まず、2チャンネルステレオでチャンネル間に位相差のない場合、2つのスピーカから等距離の場所の近傍に法線方向の様な波面が形成されることを明らかにした。また、チャンネル間に位相差を与えて、スピーカの外側に音像を定位させる方法は、様な波面のできる幅が狭い傾向にあることがわかった。

センタースピーカを付加した3チャンネルステレオでは、3個のスピーカを用いて音像を提示するよりも2個だけを用いて提示する法が様な波面のできる範囲がむしろ広くなる。従って、聴取範囲を広げる目的で多チャンネル再生を行うときは、音像定位に関しては少ない数の近接したスピーカで提示すべきであることが分かった。

さらに、帯域幅のある信号に波面の概念を拡張する方法として音源と音場における音圧の相互相関関数を用いた等時間遅延面を用いる方法があり、それによって異なる周波数の波面を比較することが可能になることを示した。

また、波面理論の応用として、波面の放線方向をリアルタイムで観測する装置を開発した。この装置により、多チャンネルステレオ番組制作において、360度全周にわたる音像のおよその方向を監視することを可能とした。

以上のように波面を用いた解析は、従来困難とされた多チャンネルステレオシステムの聴取範囲の評価や番組制作機器の開発に有効な道具であることが明かとなった。次章では音像の遠近制御法について述べるが、本章で展開した波面の理論が音像の遠近制御法の解析にもきわめて有効であることを示す。

## 概要

立体テレビの映像の遠近感に対応した音像の遠近法を開発した。これはアレイ状に並べたスピーカで音波の焦点を作り、その場所に音像を知覚させるものである。ここでは、その原理を中心に考察する。

## 5. 1. 緒言

立体テレビは将来の有力なメディアの一つである。距離感と立体感のある映像は見る人に映像との一体感を与える。さらに、映像と音像の相乗効果から考えて、立体テレビの映像の遠近表現に対応させて、音像を3次的に自在に合成させることができれば、その臨場感は一層高いものになると思われる。

しかし、音像の遠近感を自在に制御することは従来の音声システムでは困難である。立体テレビは映像を画面の手前に提示することができるのに対し、従来の音響再生システムは音像をスピーカから飛び出させ、聴取者の近くに定位させることがほとんどできない。現行の音声システムでも音の大きさや直接音と間接音のエネルギー比率を変えることで、音像の遠近感ある程度制御できるが<sup>(31)</sup>、スピーカから音像を飛び出させ、聴取者のすぐ近くに提示することは難しい。音源から外耳道入口までの頭部伝達関数を、近距離の音源と等しくなるようにフィルターで補正すれば、任意の位置に音像を提示できるという考えもあるが<sup>(32)</sup>、聴取位置が限定されるという欠点がある。また、この方法は聴取する聴取する部屋の反射音の影響を考慮にいれる必要があるため、部屋ごとに制御が必要になる。

本章で考察する音像の遠近制御の方法は上記のものと異なり、近距離に音源が存在する場合の音場そのものを模倣するものである。聴取者の近くに実音源が存在し、急峻な音圧勾配や同心円状の波面が存在する音場を再現すれば、音像が聴取者の近傍に定位すると考え、スピーカアレイによって実音源に類似した音場を合成する方法を試みた。本章では、試作したスピーカアレイによる実験をもとに、音像の遠近制御の可能性について述べる。

## 5. 2. スピーカアレイによる音像の遠近制御法に関する仮説

## 5. 2. 1. 基本原理

本章で述べる音像の遠近制御の基本的な原理は次のようなものである。凹面鏡によって光の焦点ができるように、多数のスピーカを球面の上にアレイ状に並べて同じ信号で駆動すれば、空間の一点にきわめて音圧の高い点が出現すると考えられる。この発想については、すでにオルソンが空間の限られた場所に拡声する目的で、同一の球面上に多数のスピーカを配置し、同一の信号で駆動する装置を提案している<sup>(33)</sup>。この球の中心で高い音圧が得られることは明かであるが、その場合、音像が知覚される位置については必ずしも明かではない。オルソン自身も音像については触れていない。しかし、もし、その場所が音圧が高いだけでなく、周辺に球面波的な音場が合成されていれば、音像もその位置に知覚されると思われる。

遅延装置を使用すればスピーカの位置は球面上に限定されない。Fig. 5. 1 にデジタルディレイを使用した平面状のスピーカアレイの概略図を示す。

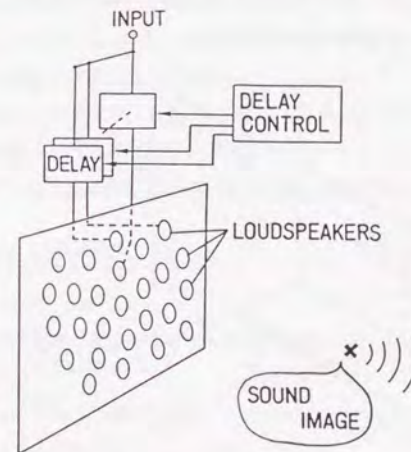


Fig.5.1 Construction of the sound-image distant-control system using a two-dimensional loudspeaker array.

スピーカアレイによる音像遠近制御法の概念図

光の波長とレンズの大きさの関係から類推できるように、音波のように波長の長い波の焦点を作るには、一般にはほとんど実現不可能な巨大なスピーカアレイが必要になる。そこで、現実的な大きさのスピーカアレイでの聴感上の効果の有無を確認するために、まず、数十個の密閉型の小型スピーカを球面上に配置し、予備実験を行った。その結果、球面の中心付近に音像が知覚され、この方法が効果があることが分かった。そこで、約200個のスピーカとデジタルディレイラインを組み合わせたスピーカアレイを試作し、計算機シミュレーションと主観評価実験によって本方法の検討を行った。

## 5. 2. 2. スピーカアレイの焦点近傍の音場解析

スピーカアレイによってできる音波の「焦点」近傍の音場を数値計算によって分析する。

Fig. 5. 2に示すように、200個のスピーカを $3.6 \times 1.8$  mの板に縦10個、横20個を格子状に配置したスピーカアレイを考える。この配置は試作したスピーカアレイの大きさに近いものである。このスピーカアレイの軸上の2.5 mの位置に焦点がくるように、各スピーカの駆動信号にディレイをかけることを仮定し、音圧の理論計算を行う。アレイに用いたスピーカは全指向性（無指向性）であると仮定し、反射等の存在しない自由音場で等音圧面と等位相面を計算した。

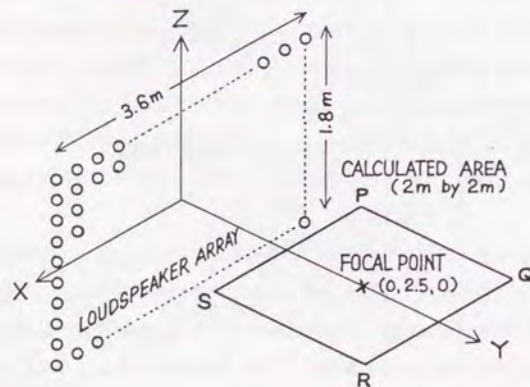


Fig.5.2 Structure of the loudspeaker array for calculation of sound field.

音場の音圧と位相を計算するためのスピーカアレイの構成

Fig. 5. 3に計算した等音圧面と等位相面を示す。これは焦点の近傍 $2 \text{ m} \times 2 \text{ m}$ の範囲を計算したものである。等音圧面は $80 \times 80$ の格子状の点の音圧の計算値から等高線プログラムを用いて作成したものである。一方、等位相面については $391 \times 197$ の格子状の点の位相を計算し、それを周波数で除算して時間遅延に変換し、 $0.5 \text{ msec}$ おきに白黒を反転させることによって表示している。

Fig. 5. 3の $500 \text{ Hz}$ から $2 \text{ kHz}$ までの結果をみると、次のような特徴がある。

(1) 焦点近傍において音圧が極大値を持つ。ただし、低い周波数ほどピークがスピーカアレイの側にやや寄る傾向がある。

(2) 焦点を中心とする円弧状の等位相面が形成されている。

これらの特徴は、スピーカアレイによる焦点近傍の音場には球面波音場と類似性があることを示している。もちろん、アレイと焦点に挟まれた領域（図のP-S側）では音波はアレイから焦点の方向に進行するので、実際の球面波と進行方向が逆であるが、図のQ-R側では音波の進行方向も一致する。これは、Q-R側で聴取したときに音像が焦点近傍に定位する可能性を示しているといえる。

一方、 $500 \text{ Hz}$ 以下の低い周波数では球面波音場との類似性は薄れている。音圧最大のピークはスピーカアレイの側に寄り、等位相面も焦点を中心とする円弧状ではない。また、図示していないが、 $2 \text{ kHz}$ 以上の高い周波数では音場がきわめて複雑な様相を呈し、やはり、球面波音場との類似性が薄れることが確かめられている。

以上の音場の計算結果から判断して、 $500 \text{ Hz}$ から $2 \text{ kHz}$ 程度の周波数帯域において、焦点近傍に音像が知覚されうる可能性が十分あることが分かる。そこで、以降の節で実際に音像が知覚される位置を聴感実験によって確認する。聴感実験は次の4種類の項目に着目して実施する。

- (1) 音像の距離感の音の大きさに対する依存性
- (2) 距離感の周波数依存性
- (3) スピーカアレイの大きさが音像の距離感に及ぼす影響
- (4) 聴取範囲

また、後述するように、これらの実験により、スピーカアレイによる音像の距離感は球面波音場との類似性だけでは説明できないことも明らかになる。

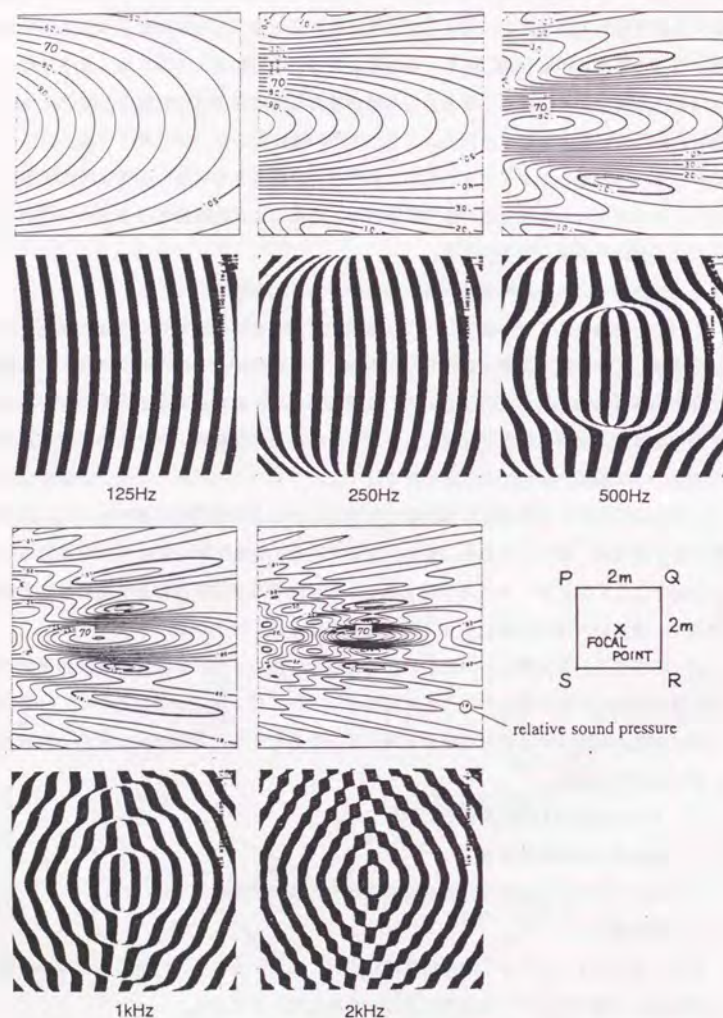


Fig.5.3 Contours of constant sound pressure and phase for the loudspeaker array.

Fig.5.2の領域PQRSの中の等音圧線(上)と等位相線(下)の計算結果  
等音圧線は相対音圧(リニア)で表示、等位相線は0.5msecおきに白黒を  
反転させて表示している)

### 5. 3. スピーカアレイによる音像の距離感に関する主観評価実験(主観評価実験1)

#### 5. 3. 1. 主観評価実験の目的

すでに述べたとおり、数10個の密閉型のスピーカを同一球面上に配置し、同一の信号で駆動した予備実験では、確かに球面の中心あたりに音像が定位しているように感じられた。しかし、音像の距離感には音の大きさが強く影響するといわれており、単に焦点近傍の音圧上昇が音像の距離感を変化させたにすぎない可能性もある。もし、そうであれば音の大きさが小さい場合には音像を近距離に定位させることができず、立体テレビ用のシステムとはなり得ない。そこで、約200個のスピーカを用いたスピーカアレイとデジタルディレイからなるシステムを試作し、それを様々な音圧レベルで駆動することによって、本方法が音の大きさによらず、音像の距離感の制御が可能であることを実験的に明らかにする。

#### 5. 3. 2. 評価実験用スピーカアレイの構造

試作したスピーカアレイの外観をFig.5.4に示す。このスピーカアレイは約200個のスピーカを11重の同心円に配置したものである。各同心円上でのスピーカ間隔はおおよそ15cmである。各同心円ごとにスピーカは直列に接続され、一つのアンプで駆動されるようになっている。各アンプの入力信号にそれぞれ遅延を与え、スピーカアレイの中心軸上の一点で各同心円からの音波の位相を等しくできるようにしている。

位相の合わせ方は次の通りである。焦点を設定したい場所にマイクロホンを設置し、ある一つの同心円上のすべてのスピーカからパルス状の信号を再生する。マイクの出力信号とスピーカ入力信号をシンクロスコープで観察し、マイク出力信号のスピーカ入力信号に対する遅れ時間が一定になるようにデジタル遅延装置の設定値を調節する。すべての同心円に対し同様の遅延時間の調節を行えばマイク位置に焦点が形成される。ここで用いたディレイの精度は20μsである。従って、音波の波長の精度にして6mm程度であり、可聴範囲の音波に対しては十分精密な調整が行える。

アレイに使用したスピーカとディレイ装置の仕様をTable 5.1に示す。

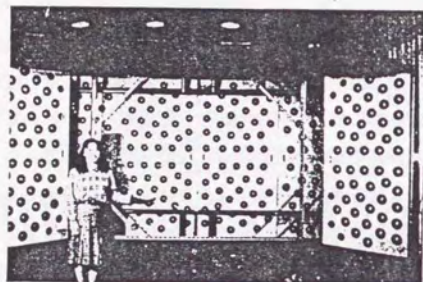


Fig.5.4 Experimental loudspeaker array.

実験用スピーカアレイの概要（実験では中央部の同心円状のアレイのみ使用）

Table 5.1 Specifications of loudspeakers and delay line.

The loudspeakers of the array
$\phi 8\text{cm}$ , 100-18000Hz(-10dB)
Loudspeakers A-D
Fostex FE103
$\phi 10\text{cm}$ , 80-18000Hz(-10dB)
Closed box: 14cm(H), 14cm(W), 12cm(D)
The digital delay lines
YAMAHA DDL-3
20-20000Hz
fs: 50kHz S/N: 100dB
Delay time: 0-1300msec
Accuracy: 20 $\mu\text{sec}$

実験に使用したスピーカと遅延装置の仕様  
（スピーカの周波数特性は-10dB減衰する周波数で表示、 $\phi$ は直径の意味、  
Closed boxとはいわゆる密閉箱のことを指す）

### 5. 3. 3. 主観評価実験の方法

試作したスピーカアレイを用いて、音像の距離感の音の大きさに対する依存性を主観評価実験により明らかにする。スピーカアレイと評定者の位置関係をFig. 5. 5に示す。評定者とスピーカアレイの間隔は2チャンネルステレオの場合の標準的な距離である3mとした。また、0.5m, 1m, 1.5m, 3mのところに評定者からみて重ならぬように単一スピーカA, B, C, Dを配置した。これらの単一スピーカは基準の距離となる音像を提示するために設置した。

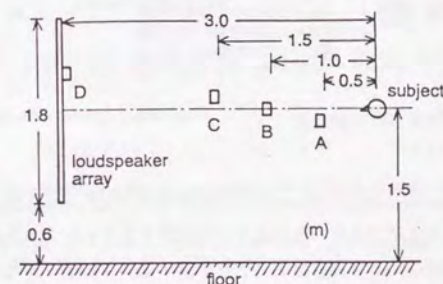
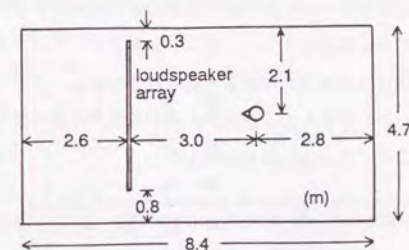


Fig.5.5 Loudspeaker layout for subjective evaluation of sound image distance. Plan view(above) and side view (below).

スピーカアレイの音像の距離感に関する主観評価実験のための実験配置

スピーカアレイの焦点の位置は評定者から0.5 mの位置に設定した。これはスピーカAの位置と距離的には同じであるが、Fig. 5. 5に示すように上下にわずかにずれており、焦点がスピーカAの陰になることはない。

実験はスピーカアレイまたはスピーカDをランダムな音圧レベルで再生して、評定者に音像の距離を判断させる方法で行った。ただし、距離知覚の性質上、絶対距離を答えさせるのは困難であるので、スピーカA, B, Cの音像と比較した相対的な距離をTable 5. 2に示す7つのカテゴリ（例えば、BとCの間、Aと同じ、など）で答えさせた。

基準となるスピーカA, B, Cを駆動する信号レベルは、それらの音像の距離感を一定に保つために、実験を通じて一定とした。その結果、聴取位置での音圧レベルは、それぞれ70, 65, 60 dB SPLであった。実験室は音響スタジオを用い、照明は評定者の手元が見える程度にまで減じた。試験音は予備実験で最も効果のあった無響室録音の女性ボーカルと、番組音の帯域を代表する中心周波数1 kHzの1/3オクターブバンドノイズを用いた。

評定者の音像の距離判断への視覚的な影響を取り除くために、スピーカAと評定者との間に音響的に透明なカーテンを配した。評定者の耳の位置が着座状態でスピーカアレイの中心の高さと一致するように椅子の座高を調節した。

Table 5.2 Categories for judgments of sound image distance (for loudness dependence experiment.)

1. スピーカAの音像より近い
2. スピーカAの音像と同じ距離
3. スピーカAとBの音像の間
4. スピーカBの音像と同じ距離
5. スピーカBとCの音像の間
6. スピーカCの音像と同じ距離
7. スピーカCの音像より遠い

#### 5. 3. 4. 音像の距離感の実験結果

実験結果をFig. 5. 6に示す。この図の縦軸は音像の相対的距離、横軸は音圧レベルを表している。●がスピーカアレイによる音像に対するTable 5. 2にもとづく判断結果、○が同じ距離のスピーカDによる音像に対する判断結果で、丸の面積が判断の出現頻度を表している。この図から以下のことが分かる。

- 1) 音圧レベルに関係なく、アレイによる音像は同じ位置にある単独のスピーカより近距離に判断される。
- 2) アレイによる音像の距離感は音圧レベルに関係なく1 mの距離のスピーカBよりも近距離に判断される確率が高い。

3) 3 mの距離にある単一スピーカDの音像は、音圧レベルに関係なく、1 mの距離のスピーカBより遠方に判断される。

4) スピーカアレイ、単一スピーカを問わず、音圧レベルが大きいほど音像が近くに判断される傾向がある。

5) 1/3オクターブバンドノイズのような狭帯域の信号に対しても、スピーカアレイによる音像はスピーカBによる音像より近い場所に判断される。

以上、総合するとスピーカアレイによる音像は音圧レベル45 dBA以上であれば聴取者から1 mの距離に置いたスピーカによる音像より近距離に判断されるといえる。もちろん音圧レベルが小さい場合には音像が遠ざかる傾向にあるが、1 mの距離に置いたスピーカによる音像より遠距離に知覚されることは希である。従って、試作したスピーカアレイは広い音圧レベルの範囲で近距離に音像を提示することが可能であるといえる。

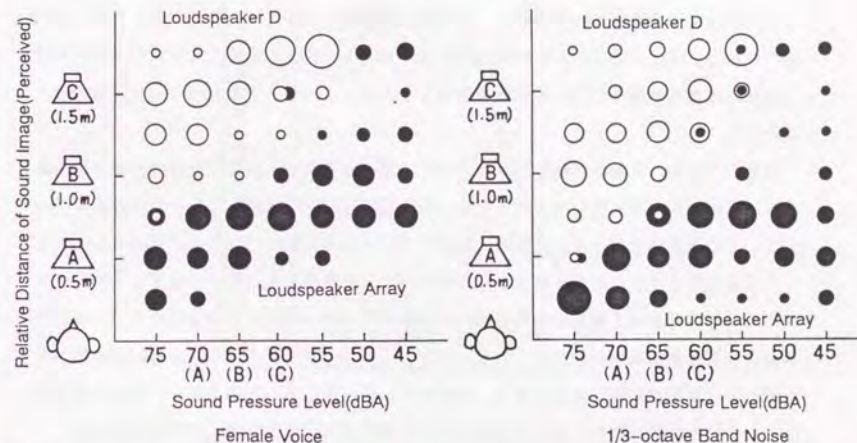


Fig. 5.6 Perceived sound image distance as a function of A-weighted sound pressure level (for 18 subjects, studio room, darkened.) The area of the circles means frequency of judgments.

実験結果（●がアレイ、○が単一スピーカの結果を示す。丸の面積は判断の出現比率を表す。）アレイによる音像は音圧によらず1 m以内の距離に知覚されていることが分かる。

#### 5. 4. スピーカアレイの音像制御に寄与する周波数帯域（主観評価実験2）

##### 5. 4. 1. 主観評価実験の概要

5. 2節において行った音場の理論計算の結果で明らかにしたように、スピーカアレイの焦点近傍の音場の形状は周波数に大きく依存する。それによれば、焦点を中心とした同心円的な波面が形成されるのは500 Hzから2 kHz程度の周波数帯域であった。そこで、ここではさまざまな周波数の1オクターブバンドノイズを用いて、距離感の周波数依存性を評価し、5. 2節の計算結果との対比を行う。実験の結果は、スピーカアレイは125 Hzから1.5 kHzの範囲で音像を近距離に提示できることを示しており、音場の計算結果だけではスピーカアレイの性能を説明できないことを明かにする。

#### 5. 4. 2. 主観評価実験の方法

前節の実験結果で、スピーカアレイの音像は音圧レベルによらず1 mの距離に置いたスピーカBの音像の前後に判断された。そこで、本実験ではスピーカBの音像より近くに判断されれば、スピーカアレイの効果があると考え、その位置を基準としてスピーカアレイの音像の距離感を種々の中心周波数を持つ1オクターブバンドノイズについて評価することとした。

理想的には5. 3節の実験で行ったように、様々な音圧レベルで刺激を提示すべきであるが、実験回数が膨大なものとなるため、評定者の疲労等を考慮して、一つの音圧レベルについてのみ実験を行った。具体的にはすべての中心周波数のバンドノイズに対し、スピーカBを約63 dB、スピーカアレイを約60 dBに設定した。スピーカアレイのレベルをあえて約3 dB低く設定した理由は、この条件でスピーカアレイの音像がスピーカBの音像より近距離に判断されれば、スピーカアレイの効果が音圧レベルによる効果を確実に上回ることが示すからである。また、全ての中心周波数において両者のレベルを一致させることが困難であるため、微妙な音圧差による判断を避ける意味でも有効と考えた。

距離判断のカテゴリーをTable 5. 3に示す。評定者は成人男女10名で、繰り返しは2回である。そのほかの実験条件は実験1と同一である。

Table 5.3 Categories for judgment of sound image distance (for frequency dependence.)

1. スピーカBの音像より近い
2. スピーカBの音像と同じ距離
3. スピーカBの音像より遠い

\* スピーカB: 1 mの距離にあるスピーカ (Fig. 5.2参照)

#### 5. 4. 3. 実験結果に関する考察

実験結果をFig. 5. 7に示す。図の縦軸は「スピーカアレイによる音像の方がスピーカBの音像よりも近い」と判断された割合に、「スピーカBの音像と同じ」と判断された割合の1/2を加えた比率である。これは「同じ」という判断は「近い」と「遠い」が半々の確率で生じたものとする統計的手法に従ったものである。この図を見ると、125 Hzから1.5 kHzの範囲で75%以上の値を示しており、5. 2. 2節で行った計算結果から予想される200から2000 Hzという範囲よりも低い周波数まで音像が近距離に定位していることが分かる。すなわち、焦点近傍の音場が同心円的な形状をしていない低い周波数帯域においても音像が近距離に判断されており、音像の距離感に関して別の要因が働いたと考えられる。この要因については5. 7節で考察する。

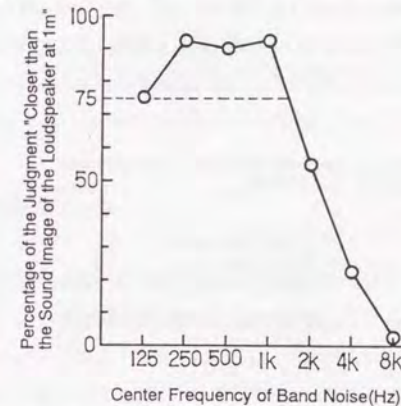


Fig.5.7 Percentage of near sound localization as a function of frequency (percentage for 10 subjects and 2 times, studio room, darkened, 60dBA, one-octave band noise.)

オクターブバンドノイズを音源としたときの音像の距離感（1 mのスピーカBの音像よりも近いと判断された比率）

#### 5. 5. スピーカアレイの大きさの効果に関する実験的考察（主観評価実験3）

##### 5. 5. 1. 主観評価実験の概要

これまでの実験で、試作したスピーカアレイ（3.6 m x 1.8 m）は、その焦点の近傍

に音像を生じさせることが確認された。このアレイの寸法は家庭用としては実用的でないが、アレイの大きさが小さくなれば、当然のことながら、距離感の制御が困難になることが予想される。そこで、ここではアレイの小型化が音像の遠近制御の効果に及ぼす影響を実験的に考察する。

#### 5. 5. 2. 主観評価実験の方法

試作したスピーカアレイは11重の同心円の構造をなしており、再生に使用する同心円の数を変えることによってアレイの面積を調節できる。ここでは再生に使用する同心円の数をランダムに設定し、その音像の距離をスピーカBによる音像と比較する方法で行った。

評価に用いた音源は実験1と同じ女性ボーカルである。判断のカテゴリーをTable 5. 4に示す。また、聴取レベルは65 dB AでスピーカBと同じ値とした。(女性ボーカルのように帯域幅の広い音源に対しては音圧レベルを揃えることが可能である。) 評定者は成人男女4人、繰り返しは5回である。他の実験条件は5. 4節の実験2と同一とした。

Table 5.4 Categories for judgment of sound image distance (for array-size dependence)

- |                 |
|-----------------|
| 1. スピーカBの音像より近い |
| 2. スピーカBの音像より遠い |

※ スピーカB: 1 mの距離にあるスピーカ (Fig. 5.2参照)

#### 5. 5. 3. 実験結果に関する考察

実験結果をFig. 5. 8に示す。縦軸は「スピーカアレイによる音像がスピーカBの音像より近い」とする判断の出現比率、横軸はスピーカアレイの最も外側の同心円の直径である。これより、スピーカアレイの直径が1. 3 m以上あれば75 %以上の判断出現比率になることが分かる。この理由については5. 7節で考察する。

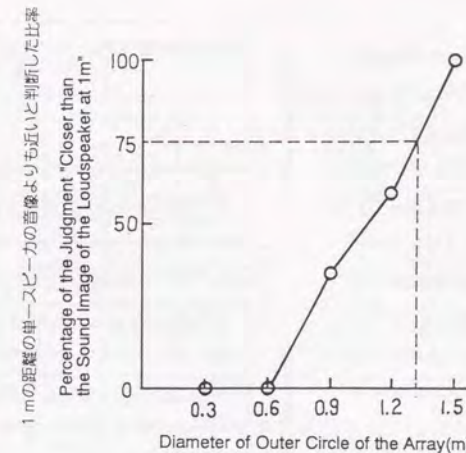


Fig. 5.8 Percentage of near sound localization for small arrays (percentage for 4 subjects and 5 times, studio room, darkened, female voice, 65dB A.)

スピーカアレイの大きさを小さくした場合の音像の距離感(1 mのスピーカBの音像より近いと判断された比率)

#### 5. 6. 聴取範囲に関する考察

スピーカアレイの焦点近傍に生じる音像は聴取位置が多少横にずれても動かないという特徴を持っている。2、3の評定者に焦点の周りを移動しながら聴取した時の印象を聞いたところ、正面から左右に数十cm動く程度では、音像は動かないという回答を得た。これは5. 2. 2節で計算した等位相面が、焦点を中心とした球面波に類似していることから、500 Hzから1 kHz程度の信号についてはある程度予測されたことであるが、それ以外の帯域を含む信号についても音像が移動しにくい。そこで、ここでは帯域幅の広い信号に対しても音像が動きにくいことを説明するために、ダミーヘッドによる両耳間相関関数の測定を行い、その最大値の位置から聴取者の両耳に発生する時間差を評価する。

両耳間相関関数とは両耳の外耳道入り口に生じる音圧の間の相互相関関数をいう。そのピークの位置から両耳間の時間差を推定できる。聴取位置が左右にずれたときのピークの位置の移動が、スピーカアレイと焦点近傍の実音源とで同じ傾向を持てば、スピーカアレイによる音像は聴取者の左右の動きに対し実音源と同等であることが期待できる。

中心軸上でアレイから3 m離れた位置から25 cm右にずれた位置にダミーヘッドを置き、ホワイトノイズを再生したときの両耳間相関関数をFig. 5. 9に示す。図にはス

ピーカBからホワイトノイズを再生したときの結果をも示した。この2つの両耳間相関関数の最大ピーク的位置を比較すると、スピーカレイの方がピーク的位置が右に寄っている。これはスピーカBよりもスピーカレイの方が大きな両耳間の時間差を発生させることを示しており、スピーカレイの音像の方向がスピーカBより左に寄ることの裏付けとなる。

右にずれた位置で聴取した場合に、音像の方向がスピーカBよりも左になるということは、音像方向が焦点の位置にとどまる傾向を持つことを意味している。この関係をFig.

5. 10に図示する。

## 5. 7. 外耳道入り口の音圧の両耳間コヒーレンスと音像の距離感に関する考察

### 5. 7. 1. コヒーレンスを用いる理由

5. 3から5. 6節において、スピーカレイによる合成される音像の距離感の性質を聴感実験に基づいて述べた。このうち実験2で示した周波数依存性については5. 2節で計算した音場の形状による説明ができないことを5. 4. 3ですでに述べた。音像の距離感最終的には両耳に加わる音刺激の性質によって説明されるべきであろう。そこで、本節では両耳に加わる音圧間のコヒーレンスを用いた説明を試みる。その結果、コヒーレンスがスピーカレイによる音像の距離感を極めて良く説明することを明らかにする。

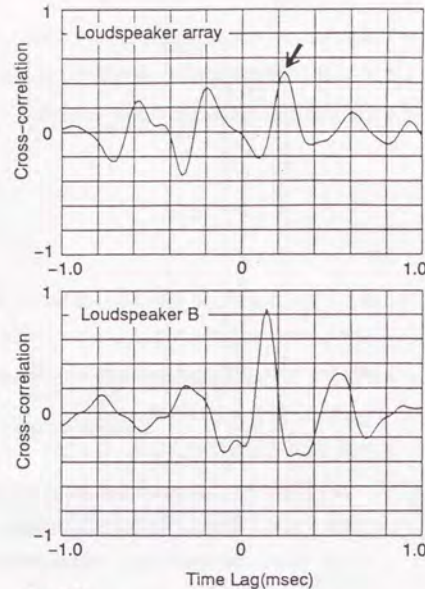


Fig.5.9

Interaural cross-correlation functions measured at a 25cm right shifted position (with an artificial head, white noise.)

25 cm右にずれた位置で測定した両耳間相互相関関数

### 5. 7. 2. 外耳道入り口の音圧の両耳間コヒーレンスの測定

音像の距離感に関する過去のいくつかの研究は音像の距離感は直接音対反射音の比率に大きく左右されることを示している<sup>(32)</sup>。すなわち、反射音の比率が高ければ音像は遠方に、また、低ければ近くに感じられるという。スピーカレイの音像の距離感にも実験に用いた試験室の壁面による反射音が寄与していることが予想される。

直接音と反射音の比率を正確に測定することは困難であるが、反射音の経路が十分多く、その方向性が失われているとすれば、次の理由により2点間の相互相関係数またはコヒーレンス関数で評価することが出来る。すなわち、反射音の存在しない無響室等の空間でスピーカレイを再生した場合は、左右対称の音場が形成されるので、中心軸上に設置したダミーヘッドの両耳の位置に生じる信号は等しくなる。従って、コヒーレンスは1となる。一方、完全に拡散性の音場（例えば残響室）では、コヒーレンスは拡散音場の相関係数の絶対値

$$|\sin kd / kd|$$

$k$ : 波数、 $d$ : 2点間の距離

に近づく<sup>(39)</sup>。一般の部屋ではこの中間となると考えられる。

実際の測定は、主観評価実験の評定者の位置にダミーヘッドを置き、両耳の位

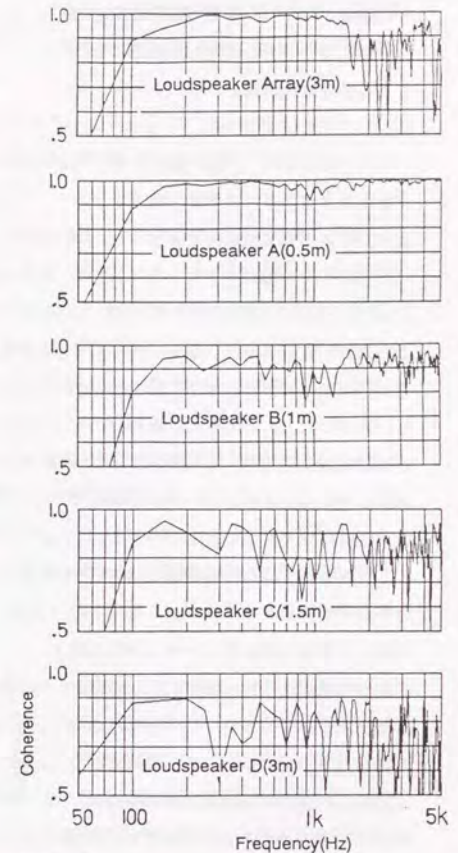


Fig.5.10

Interaural coherence for the loudspeaker array and loudspeaker A-D (with an artificial head, white noise, window:20ms, fs:12.8kHz.)

スピーカレイ、及びスピーカA-Dによる両耳間コヒーレンス

置に配置したマイクロホンで収音した信号の間のコヒーレンスを対象にした。もし、ダミーヘッドを設置した位置において直接音の比率が高ければコヒーレンスは1に近づくはずである。

スピーカアレイ、スピーカA、B、C、Dをホワイトノイズで駆動したときのコヒーレンスの測定結果をFig. 5.10に示す。この結果からスピーカアレイによるコヒーレンスは同じ距離に設置した単一スピーカDよりも大きく、1.5kHz以下ではダミーヘッドに最も近いスピーカAとほとんど等しいことが分かる。これは実験2の結果、すなわち、スピーカアレイによる音像は1.5kHz以下で近距離に定位することと符合する。

1.5kHz以上の周波数ではスピーカアレイのコヒーレンスが低くなっているが、その理由は、ダミーヘッドの位置が焦点の位置と50cmずれていることに起因する。Fig. 5.3の等音圧面の計算結果から分かるように、周波数が高くなるほど音圧が極大を示す領域は狭く鋭くなり、ダミーヘッドの位置では直接音による音圧の上昇が見られなくなる。

次に、スピーカアレイの直径を変化させたときのコヒーレンスをFig. 5.11に示す。この図から直径が1.2mあれば1.5kHz以下ではスピーカBよりも十分コヒーレンスが高いことが分かる。これは1.3m以上の直径があれば、音像は75%以上の確率で1m以内

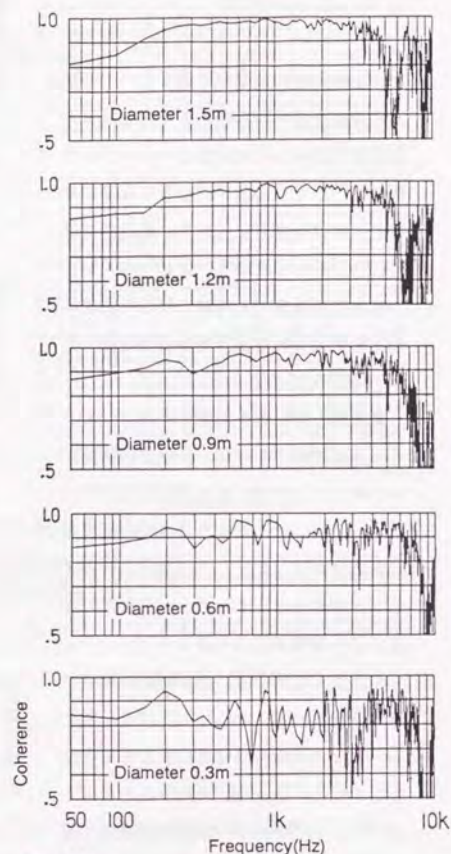


Fig.5.11

Decrease of interaural coherence with small arrays (with an artificial head, white noise, window:20ms, fs:25.6kHz.)

スピーカアレイの大きさを小さくした場合の両耳間コヒーレンスの減少

に定位するという実験3の結果とほぼ一致する。

このように両耳間のコヒーレンスはスピーカアレイによる音像の距離感の性質（周波数依存性とサイズの影響）を良く説明することが分かる。この結果を基に音像の距離感の要因に関して次節で詳細に考察する。

### 5. 7. 3. スピーカアレイによる音像の距離感の要因に関する考察

前節において、ダミーヘッドによる両耳の外耳道入り口に生じる音圧信号間のコヒーレンスの測定により、近距離に音像が知覚される場合は常にコヒーレンスが高いことが示された。スピーカアレイの焦点近傍で起きている現象は以下のように推測される。まず、焦点近傍ではアレイの各スピーカからの直接音が、ほぼ同時に到達するため、それらの直接音は同相で加算される。従って、直接音による音圧レベルはスピーカの数を $n$ とすると単一のスピーカの $n$ 倍に近くなる。一方、各スピーカからの音波の反射音は、多数の異なる位置のスピーカからの音波がさまざまな反射経路を経て合成されるので、各スピーカからの音波が焦点近傍に到達する時の位相遅延がそろふ確率はきわめて低い。従って、反射音同士の加算は無相関信号同士の加算と見なせるので、反射音による音圧レベルは1倍となる。結局、アレイで再生することによる直接音の相対的比率は単一スピーカの約 $n$ 倍となる。直接音の比率が高ければ音像の距離が近くなることは過去のいくつかの実験で示されている。

ここで注意すべきことは、聴覚の性質における一般論として、コヒーレンスの大小が音像の距離感を完全に説明するわけではないことである。例えば、左右の耳に響きのない独立な音を聞かせた場合には、音像は左右に分離し、しかも両耳の位置に知覚される。このときコヒーレンスは明らかに0である。また、それ自身に響きを含んだ音を単一のスピーカで無響室で再生した場合、コヒーレンスが1であるにもかかわらず、音像は遠ざかる傾向を持つ。従って、音像の距離感とコヒーレンスに関する議論は通常の部屋で響きを含まない音源を用いる場合に限定される。しかし、現実の聴取環境で上記のような状況が起こることは希であり、事実上考慮しなくても良いので、コヒーレンスはスピーカアレイを設計するためのきわめて有用な指針となりうる。

聴覚心理の観点から見ると、音像の距離知覚の要因に関しては諸説がある。Gardner<sup>(40)</sup>によれば無響室における正面方向の音像の距離感は音の大きさが第一要因であるという。中林ら<sup>(41)</sup>によれば、無響室でも正面以外の方向の近距離の音像については両耳間に生じるレベル差がキューになるという。また、後藤ら<sup>(42)</sup>によれば直接音を正面から提示し、反

射音を様々な方向から付加した場合、60度方向からの反射音が音像を遠方に知覚させる効果大きいという。同時に、反射音の方向が直接音と同一の場合には、その効果が最小であるという。しかし、複数の要因が関与する場合はどの要因が主要な働きをするか不明であった。今回のスピーカアレイの音像の距離感に関する実験結果から推測すると、適度な反射音のある通常の聴取環境では、直接音と間接音の比率が音像の距離感に対し最も寄与が大であると結論づけられよう。

## 5. 8. 既存の音場制御法との比較検討

ここでは、他の音響システムあるいは音場制御法について触れ、本論文の方法がそれらに比べ立体テレビのための音像の遠近制御法として優れていることを示す。

### 5. 8. 1. 頭部伝達関数による方法

#### (1) 概要

2チャンネルステレオで音像を3次元的に制御する場合は、頭部伝達関数による方法が一般的である。

ダミーヘッドで収音した音をヘッドホンで聴取するときわめて原音場に近い聴感覚が得られることはよく知られている。頭部伝達関数の補償という手法を用いれば、スピーカによってダミーヘッドとヘッドホンの組み合わせと同等の効果を実現することができる。頭部伝達関数の模擬による音像制御の原理は1960年代から70年代にかけての研究<sup>(43)</sup><sup>(44)</sup>によってすでに確立されており、現在は所望の特性のフィルタをいかに少ないハードウェアで実現するかに研究の重心が移っている。しかし、今のところ正中面内の音像の不安定さや聴取位置が限定されるという問題点が解決されていない。

頭部伝達関数を用いる制御の原理は、左のスピーカから右耳への、また、右スピーカから右耳へのクロストーク成分を耳の入り口で打ち消す信号を、それぞれ右信号と左信号に加えておくものである。このクロストークキャンセル回路を応用すれば、ある信号を右耳だけ、あるいは左耳だけに提示することが可能となる。従って、モノの音声を2つのスピーカで任意の位置に定位させることができる。

以上が頭部伝達関数による音像制御の概要である。すべての系が理想的に働くならば、少なくとも頭を固定した状態で得られる音の空間的印象は完全に再現できると考えられる。

ただし、この方法で任意の方向に音像を定位させるためには、すべての音源の方向に関する頭部伝達関数のデータが必要になる。そのうえ、正面方向の音像の上下感は再現が難しく、聴取者に依存した補正が必要であるという<sup>(45)</sup>。

#### (2) 頭部伝達関数による手法の問題点

頭部伝達関数の形状は個人ごとに異なる。従って、正確な音像定位を得ようとすれば聴取者本人の頭部伝達関数をあらかじめ測定しておく必要がある。これは音場の測定や評価など、系を管理できる場合には不可能ではないが、不特定多数の聴取者を対象とするオーディオでは、事実上、不可能である。さらにオーディオの場合、固定した位置で聴取されるとは限らないので、厳密なクロストークキャンセルは期待できない。

さらに、本質的な問題点は、聴取場所において加わった間接音が音像の距離感を変えるために、音像の距離感を正確に再現できない点である。

以上のように頭部伝達関数による手法は、立体映像システムと組み合わせるためには、聴取位置が限定される、音像の距離感の正確な再現は原理的に困難である、という2つの欠点がある。

### 5. 8. 2. マルチチャンネル再生による音場制御

2チャンネルシステムが両耳の音圧に着目する手法であるのに対し、マルチチャンネルのスピーカを用いて、音場そのものを制御する考え方もある。しかし、多くはまだ理論の段階であるうえ、音像の遠近制御を可能とするシステムとなっていない。

#### (1) 面積分による方法

閉空間内の音場の完全な再現を目指すものとして、キルヒホッフの積分公式に基づくものがある。音場におけるキルヒホッフの積分公式とは、「内部に音源の存在しない閉空間の内部の音場はその境界面における音圧と音圧傾度が決まれば、一意に定まる。」というものである。この考え方に従えば、ある空間を取り囲む面の音圧と粒子速度を多数のスピーカによって原音場と等しくなるよう制御すれば、空間内については原音場が再現されると考えられる。しかし、この考え方も閉空間内に音源や吸音体がある場合に適用できないので、近い音像を合成させることには無理がある。また、一般の部屋では壁の反射音を打ち消す必要があり、システムの複雑化は避けられない。

キルヒホッフの積分公式の考え方を半球型の半自由空間に適用すると、ホイヘンスの原理となる。この考え方に基づき、Berkhout<sup>(46)</sup>らは平面状のスピーカアレイを用いて波面

を合成する手法を提案している。ある平面の音圧と音圧傾度を制御する代わりに、双指向性スピーカを配置するので、システムとしてやや簡便となるが、音源と受音点が面の反対側になければならない制限は変わらない。また、双指向性スピーカの必要性など現実的でない部分がある。

#### (2) 方向別反射音による方法

前2項よりも現実的な考え方として、ある範囲を取り囲むようにスピーカを多数配置し、方向別の反射音を再生するという手法がある。原音場における各方向からの反射音を再生音場で多数のスピーカから再生すれば、頭部伝達関数によらなくても中央の聴取位置では原音場とほぼ等しい反射音の時系列を聴取できるとする考え方である。これはオルソン<sup>(47)</sup>が提唱した考えで、現在も多チャンネル再生の基本的な考え方である。しかし、この方法も聴取する部屋の反射音が余分に付加されるため、音像の遠近感を制御する目的には合致しない。

以上のように、従来の音場制御の手法は無響室再生で忠実再現を原則としているため、一般の反射のある部屋では付加される反射音が音像の距離感に悪影響を及ぼす。付加される反射音は音像を遠ざける効果を持つので、聴取者のごく近傍に音像を定位させることが困難になる。これに対し、スピーカアレイによる焦点の合成による手法は、直接音の比率が制御されるため、音像の距離感が効果的に変化するのである。

### 5. 9. スピーカアレイによる音像の遠近制御システムの実用化に関する考察

スピーカアレイのすべてのスピーカが独立な遅延装置を持っていれば、理論的には空間の任意の場所に焦点を作ることができる。もし、その焦点の位置に音像を知覚させることができれば、3次元的な音像制御が実現する。本方法を実際の立体テレビ用音声システムとして用いるためには、いくつかの課題がある。

#### (1) 装置の小型化

今回試作したスピーカアレイは3.6m x 1.8mの大きさがあり、家庭や小さなスタジオに設置できるものではない。5.5節において、アレイの直径は1.3mあれば十分と述べたが、実はアレイが小さいほど聴取範囲が狭くなることが、簡単な試聴の結果、明らかとなっている。立体テレビとの組み合わせを考えた場合、聴取形態にもよるが聴取範囲の広さは重要である。しかし、今回、試作したスピーカアレイは聴取範囲について特に考慮した設計となっていないので、スピーカ単体の指向性を考慮し、各周波数ごとに波面の

形状の最適化を行うことができれば、聴取範囲をさらに拡大できる可能性がある。また、将来的にはスピーカの形態そのものを検討し、本方法に適した方式を開発する必要がある。

#### (2) 視覚との相乗効果

音像の距離感には視覚刺激の影響が大きいといわれている<sup>(48)</sup>。今回も試験室を明るくし、焦点の近傍にダミーのスピーカを設置すると、多くの評定者があたかもダミーのスピーカから音が出ているように感じられると述べた。ダミーのスピーカを置くことによって、何もない空間から音が出ている不自然さが解消され、音像が「安定」する。立体テレビの映像が提示された場合も、それが音像定位を安定させる効果を持つと期待できる。

#### (3) 周波数帯域の拡大

今後の課題で最も大きなものは周波数帯域の拡大である。今回、主観評価実験に用いたスピーカアレイでは、制御できる帯域はおおよそ125Hzから1.5kHzまでに限られる。幸い多くの番組音のエネルギーは中域に集中しているので、音像の距離感に関しては実用上問題ないが、音質に関して高品質とは言い難く、改善の余地がある。

#### (4) 伝送法

スピーカアレイを駆動する100チャンネルもの信号をそのまま伝送するとすれば、非常に多くの伝送チャンネル数が必要となる。また、音声信号に与える遅延量はスピーカアレイの構造に依存するため、再生用スピーカアレイの配置の規格化が必要となり、現実的でない。チャンネル数を削減する方法としては、音像の距離感の第一要因が直接音と間接音の比率であることを利用することが考えられる。前節で述べたように音像の距離感は直接音と間接音の比率に支配されるので、音波の焦点を移動させることは必ずしも必要ない。何らかの方法で直接音と間接音の比率が制御できればよい。間接音を直接音と別チャンネルで伝送し、スピーカアレイ側で再構成するのも一つの方法であろう。この方法を使えば通常の2倍のチャンネル数で伝送可能である。

### 5. 10. 結言

本章では立体テレビの音声システムへの応用を意図して、スピーカアレイによる音像の遠近制御法について考察した。その結果、この方法は音像を近距離に定位させることができるばかりではなく、聴取者が多少動いても音像は動かないという、従来の音声システムにない著しい特長があることが分かった。また、本方法の設計にあたって、聴取範囲の拡大という点では音場に合成される波面による解析が、また、音像の距離感の評価に関して

は両耳間のコヒーレンスが有効な指標であることが明かとなった。

評価実験のために今回試作したスピーカアレイは3.6m x 1.8mの大きさがあり、現段階では実用的と言えないが、音像の距離感を明確に制御する基本技術を開発した点で意義あるものとする。装置の小型化、複数音像の制御、距離感の連続的な制御などの研究をさらに進め、立体テレビの音声システムの実現をめざしたい。

## 6章 番組音の相関を利用したVTRとテープレコーダーの同期運転装置の開発 (49) (50)

### 概要

番組音の音声の相関を利用したVTRとテープレコーダーの同期運転装置について述べる。これは、これはタイムコード等のキュー信号を使用せず、VTRの音声トラックとテープレコーダーに録音された番組音の相互相関関数から両者の時間差を検出し、その時間差が0になるようにテープレコーダーの走行速度を制御する装置である。映像を伴う音声の効率的な制作を行うために開発したものである。

### 6.1. 同期運転装置開発の背景

一般にステレオの音声の制作はモノフォニックの場合よりも複雑な作業を要するため、音声と映像の編集を同時に行うことが困難な場合がある。特にドラマや軽音楽系の音楽番組では、映像に合ったきめ細かなミキシングをするために、映像を先に完成させ、それに合わせて音声を制作することがしばしばある。また、映像の編集時に繰り返し行われるVTRテープのダビングによる音声の品質劣化を避けるためにも、音声を映像編集とは別系統で処理することは有効である。ポストプロダクションとよばれるこの制作方法では、編集済みの映像に合わせて合わせて音声をあとから制作するので、VTRと音声用テープレコーダー（以下ATRと略す）を同期運転させる装置が必要である。

一般的には番組収録時にタイムコードと呼ばれる同期信号をVTRとATRの双方に記録し、再生時に専用の制御装置がタイムコードを読み取って同期をとる。しかし、この方式ではタイムコード用の専用トラックを有しない通常の2チャンネルATRは同期させることができない。また、VTR側のタイムコードは編集の段階でつけかえられることが多いので、タイムコード差と実時間差のずれを補正するオフセットを探す作業が必要になる。

タイムコード等の同期信号を必要としない同期運転装置ができれば、どのような種類のATRで録音した音声も同期させることができ、汎用性が高いシステムが構成できる。また、タイムコード方式のバックアップとしての利用も期待できる。同期信号を使わないで同期をとる方法としては、電源周波数を使う方法が考えられるが、電源周波数による方法ではDCサーボモーターのATRは制御できない上、時間差そのものが検出できない。そこでここでは音声信号そのものを使う方法について述べる。

音声をVTRとは別のATRに録音する場合でも、通常はVTRの音声トラックにガイ

ド用としてATRと同じ音声録音される。この場合、VTRとATRの音声信号はSN比、歪率、周波数特性等、音質に関する要素を除けばほぼ同一の音声信号である。そこで、この2つの音声信号を比較してATRを制御すれば同期運転が成立する。比較には2つの音声信号の相互相関関数を用いる。その最大値の位置からVTRとATRの時間差を推定できる。この方式は特殊な信号を使用しないため、現場でタイムコードの録音に気を使う必要がなく、どのようなATRで録った音声信号も同期させることができる。

## 6. 2. 同期運転装置の仕様に関する考察

番組制作用としてのVTRとATRの同期運転装置には次のような性能が必要である。

### (1) 1フレーム以内の同期精度

VTRとATRの同期精度は映像と音声のずれ(リップシンク)の検知限が約1フレーム(1/30秒)といわれているので<sup>(8)</sup>、その範囲内に常に収まるようにATRを制御できなければならない。

### (2) 初期時間差の吸収

起動時のVTRとATRの初期時間差は、現行のVTRの起動特性を考慮すると2~3秒程度あるものと考えなければならない。従って、最大、2~3秒の時間差を検出できる機能を有している必要がある。

### (3) 聴感上の音質変化がない。

VTRに追従させてATRの速度制御を行うが、制御中の音声も番組に使用可能な品質を有していなければならない。従って、制御量はワウフラッターやピッチ変化が検知限以下の範囲内で行う必要がある。

## 6. 3. 時間差検出の方法に関する考察

この同期運転装置はVTR及びATRに記録された番組音の相互相関関数を計算し、時間差を検出する方式をとっている。ただし、音声信号そのものではなく、それを包絡検波した後、相関を計算している。その理由を以下に示す。

VTRの音声信号 $x(t)$ とATRの音声信号 $y(t)$ が時間差を除いてほぼ同じ信号であるとすれば、それらの相互相関関数 $R(\tau)$ のピークの位置から時間差が求められる。この $R(\tau)$ は次式で与えられる。

$$R(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t) y(t + \tau) dt$$

サンプル値を用いると次のように表される。

$$R(m\Delta t) \approx \frac{1}{N} \sum_{n=1}^N x(n\Delta t) y(n\Delta t + m\Delta t)$$

ここで $\Delta t$ はサンプリング周期である。 $m$ の範囲を $-M$ から $M$ まで計算することで、 $-M\Delta t$ から $M\Delta t$ の範囲の相互相関関数を計算できる。時間差がその範囲内であれば最大値の位置から時間差を特定できる。従って、大きな時間差まで検出可能とするためには $M$ または $\Delta t$ の値を大きくしなければならない。しかし、番組音の相互相関関数のピークを検出するためには、標準化定理により信号の最高周波数の2倍以上の周波数でサンプリングする必要がある。従って、信号の帯域幅を $B$ とすると次の不等式が成り立つ。

$$1/\Delta t > 2B$$

検出可能な最大時間差 $T_{max}$ は

$$T_{max} = M\Delta t$$

で与えられるので $M$ は次式を満たさなければならない。

$$2T_{max}B < M$$

従って、音声信号の帯域を20kHzとすれば、40kHz以上のサンプリング周波数が必要になり、+-2秒以上の時間差を検出するためには8万点以上の相互相関係数を計算しなければならない。これは番組制作用機器としては現実的でない。

計算点数を減らすために信号に低域遮断フィルタをかけ、狭帯域化する方法もあるが、一般に低域遮断フィルタをかけると、楽音のように調波構造をもつ信号は周期性が際立つようになる。その結果、信号の相互相関関数も周期構造を持ち、最大のピークを特定す

るのが困難になる。最大のピークの位置を誤り無く特定するためには相関関数は単峰性であることがのぞましい。

同期運転の目的に限れば、VTRとATRの時間差の検出精度は0.5フレーム程度でよい。このため本装置では音声信号そのものの相関を計算するのではなく、包絡検波したのち相互相関をとる方式を採用している。低域遮断と異なり、包絡検波された音声信号は周期構造を持つ可能性が低いので、相互相関関数の単峰性を保つために有効な帯域制限である。

なお、サブサンプリングの技術を用いれば、サンプリング周期以上の精度で相関関数の形状を推測することができ、時間差推定の精度も向上する。しかし、現状ではシステムの複雑化、高コスト化は避けられず、サンプリング周期を短くする方法に比べて利点がないと判断し、今回は用いないこととした。

#### 6.4. 相関関数検出のための回路構成

信号処理部の構成をFig. 6.1に示す。VTRおよびATRの音声は、まず録音系統による音質の差を取り除くために100Hz以下を遮断する。つぎに時定数約0.1秒で包絡検波する。これにより信号の帯域が20Hz程度と大幅に落ちるので、マイクロコンピュータを使って実時間の処理が可能になる。さらに、相関処理に不要な直流成分を除いてA/D変換する。

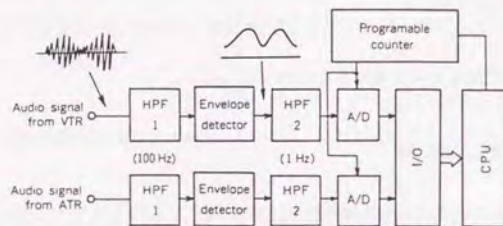


Fig.6.1 Block diagram for signal pre-processing of the tape-lock system.

相関関数を計算するための信号の前処理

実際の相互相関の計算はソフトウェアによって行われる。サンプリング周期毎に入力ポートからCPUに割り込みがかかり、A/D変換されたデータがメモリーに取り込まれる。

メモリーのうち256\*2バイトをシフトレジスタとして扱い、新しいデータが最も古いデータのアドレスに書き込まれる。

このデータを用いて相互相関関数の計算を定義に従って実時間で実行するためには乗算が必要だが、本装置の使用目的ではピークの位置だけがわかればよいので、データの最上位ビット(符号ビット)の相関を計算している。この方法は極性相関法と呼ばれ、積の演算を、論理積で代用できるので計算速度が速い利点がある。

この方法で歌謡番組の音のある部分3秒間についてVTRとATRの音声の相互相関関数を計算した結果をFig. 6.2に示す。これはサンプリング周波数が80Hz、累積回数256回の結果で、各点の間隔は12.5msである。ピークの位置よりATRがVTRよりも0.16秒遅れているのが分かる。

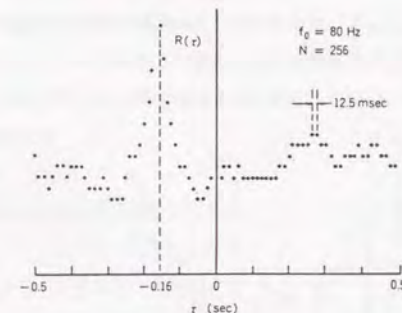


Fig.6.2 An example of the computed cross-correlation function (for a Japanese pop song, fs:80Hz.)

測定された相互相関関数の一例

#### 6.5. VTRとATRの速度偏差の影響

相互相関から時間差を検出するときに、ATRとVTRの速度偏差が誤差の原因の一つとなる。

VTRとATRに速度差がなく時間差のみ存在する場合は、 $x(t)$ と $y(t)$ の時間差の検出精度を上げるためには積分時間を長くすれば良い。しかし、ATRとVTRで速度偏差がある場合、時間差が時間と共に変化してしまう。従って、積分時間の長さには限度がある。時間差が積分の間に0.5フレーム以上変化しないことを目安とすれば、速度偏差 $\Delta v$ と積分時間 $T$ の関係は次のようになる。

$$\Delta v \cdot T < 1/60$$

放送用機器であっても  $\Delta v$  は通常 0.5% 程度は存在しうるので、 $T$  は 3 秒程度が限度であるといえる。

#### 6. 6. VTR と ATR の音質差の影響

ここでは VTR の音質が極端に悪い場合の時間差検出の可能性について述べる。

VTR のアナログトラックの音質が ATR に比べて劣る場合を想定し、相関をとる片方の信号に LPF または HPF をかけた場合の相互相関関数の最大値の位置と真の時間差との対応を調べた。Fig. 6. 3 に結果を示す。これは真の時間差と相互相関関数の最大値の位置が 100 ms 以上ずれる確率を示している。この結果から、300 Hz から 1 kHz までの帯域があれば正しい時間差を検出する確率が高いことがわかる。

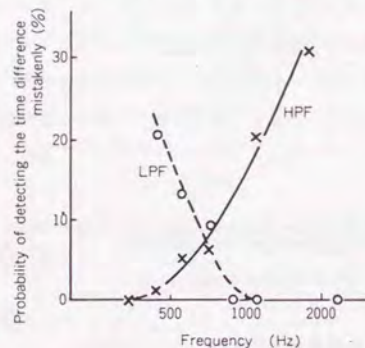


Fig. 6.3 Incorrect detection of time difference for band-limited signal.

VTR の信号に帯域制限をかけた場合の影響

#### 6. 7. ATR の速度制御法

ATR の速度を制御し VTR との時間差を縮めるにあたって、次のことを考慮する必要がある。

- (1) ATR の速度制御は再生音のワウフラッターが検知されない範囲で行う。

(2) 相互相関関数のピークの位置が真の時間差に対応しない場合にそれに追従しないこと。例えば、周期的な音、エンベロープが変化しない定常音、無音等の場合には相互相関関数のピークの位置が時間差に対応しないことがある。

(3) ATR と VTR の速度差が大きい場合には相互相関関数から時間差を検出することができない。

(4) 相互相関関数で計測される時間差は積分時間内で平均化された時間差であり、瞬時の値ではない。従って観測に数秒の時間遅れを伴う。

これらのことを考慮した本装置の制御系統図を Fig. 6. 4 に示す。相互相関関数の計算と ATR の制御はすべてソフトウェアで行っている。観測による時間遅れを吸収するために、測定された時間差の履歴を記憶しておき、予測制御を行っている。なお、具体的なプログラムは制御対象となる ATR の機種に依存する。

#### 6. 8. 実際の使用例

実際の番組制作に使う装置には VTR のタイムコードを読んで、ATR の起動をかける機能を付加しており、一度リハーサルを行えば ATR が起動した時点で 1 フレーム以内の時間差にすることができる。また、リハーサル時に ATR の速度偏差をも検出し、補正する機能もつけ加えた。

本装置は現在、音楽番組を中心に現用に供されている。最近では VTR がモノ音声であった時代の番組の映像に、FM 放送用として同時に収録されていたステレオ音声テープを同期させ、過去の名演を映像付きでよみがえらせることに用いられた。

また、同期運転装置以外の使い方として、

衛星回線の遅延時間の測定に使用されている<sup>(5,2)</sup>。現在、海外の音楽番組を高品質で日本に伝送する場合、「映像+アナログ音声」を伝送する回線とは別の回線により PCM 音声

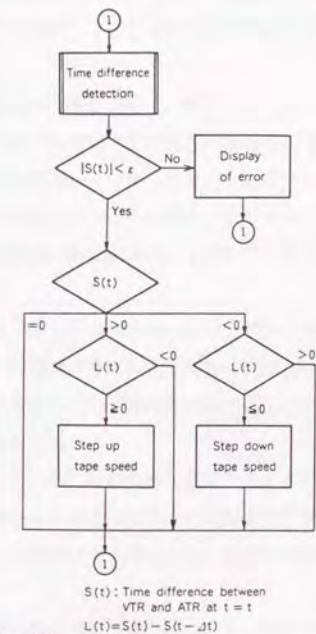


Fig. 6.4

Flow-chart for speed control program.

制御用フローチャート

が送られる。回線ルートの直前の変更等があった場合、受け手の放送局側で事前に把握できないことが多いので、放送直前まで映像と音声の間の時間差を確認できないことがある。生放送の場合は時間差の補正を速やかに行わなければならないが、本同期運転装置はこのような場合に威力を発揮する。演奏の始まる前の会場の聴衆ノイズなどで時間差を検出し、速やかに必要な音声の遅延量を決定することができる。

## 7章 結論

本研究で得られた結果を総括する。

本研究は映像を伴う音声システムの満たすべき条件を求めるという主題で行った。

まず第2章において、映像が音像定位と広がり感に与える影響について実験し、映像が音像を引きつける効果があること、サウンドステージを狭める効果があることを明らかにした。さらに、テレビ視聴時の映像と音像のずれの検知限、許容限を求め、非専門家で約20度、専門家で約11度であることを示した。次の第3章においては、これらの結果をHDTVのシステム設計にあてはめ、HDTVの音声チャンネルの数は2では不足で、少なくとも前方に3チャンネル必要であることを述べた。日本のハイビジョンの音声方式はこの結果をもとに3-1方式に決定されている。

続く第4章では、多チャンネルステレオの聴取範囲を推定する目的で、複数スピーカ再生により合成される波面を計算した。その結果、多チャンネル再生で広い聴取範囲を得る再生法は、なるべく少数の近接したスピーカの組み合わせを用いるべきであることを述べた。また、波面の理論を応用して、多チャンネルステレオにおける音像の方向をリアルタイムで監視する音像モニター装置を開発した。なお、この装置は製品化され、マルチチャンネル音声制作用のスタジオ等に配備された。

第5章では、将来の立体テレビのための音声システムを念頭におき、新しい音像の遠近制御手法を開発した。これはスピーカアレイを用いて音波の焦点を作るもので、焦点の近傍に球面波に類似した音場を再現する。この結果、焦点近傍に音像を知覚させることができるばかりか、広い聴取範囲で音像が所定の場所に知覚される。

最後の第6章では、映像を伴う音声の番組制作に用いるために開発したVTRとATRの同期運転装置の開発について述べた。VTRとATRの双方に記録された番組音声の相互相関を用いた同期運転装置で、タイムコード等の同期信号を必要としない簡便な装置である。

なお、音像モニター装置、音像の遠近制御法と同期運転装置の3つについては特許申請中である。

映像との関係を考慮にいれることはこれからの音響システムに必須の事柄になるであろう。HDTVや立体TVだけでなく、いわゆるヴァーチャルリアリティ（仮想現実感）の目指す究極の映像音声提示システムには絶対必要な条件である。映像と音声それぞれの単独で高品質化を進めることも重要であるが、高い臨場感や現実感を得るためには、映像と

音像の位置の一致感や、視聴者の身体の動きに対して映像と音像が同方向に動くことがきわめて重要な要素となろう。本研究は、このような感覚間の相互作用に初めて本格的に着目したものである。まだ第一段階ではあるが、本研究の結果及び方法がこの分野の発展に寄与するところが大であると確信している。

## 謝辞

本論文をまとめるにあたり、東京大学工学部・羽鳥光俊教授に終始、御指導、御鞭撻を頂いた。また、東京大学工学部・高木幹雄教授、原島博教授、相澤清晴助教授、宇宙科学研究所・廣澤春任教授には有効適切なお指導と助言を賜った。心からの感謝の意を表する。

本研究はNHK放送技術研究所において1980年から1992年の間に実施したものである。研究の実施にあたっては、NHK放送技術研究所次長・中林克己氏から常に有益なご助言をいただいた。

また、本研究の実施にあたり、元音響聴覚研究部長・氏原淳一氏、同、若栗尚氏、同元副部長・大串健吾氏、元音響研究部主任研究員・二階堂誠也氏に多大のご便宜を計っていただいた。大阪放送局チーフエンジニア・辻本廉氏、音響聴覚研究部長・宮坂栄一氏、同主任研究員・黒住幸一氏、盛田章氏には、多くの助言と有益な議論をしていただいた。

さらに、番組制作用機器の試作にあたっては、元制作技術局副部長・大見規太郎氏をはじめとする多くの方々に御協力頂いた。

以上の方々にこの場を借りて謝意を表する。

参考文献

- 1) Welch, R.B. and Warren, D.H., "Immediate perceptual response to intersensory discrepancy," Psychol. Bull., 88, 3, 638-667 (1980)
- 2) Pick, H.L. and Warren, D.H., "Sensory conflict in judgments of spacial direction," Perception & Psychophys., 6, 4, 203-205 (1969)
- 3) Stratton, G.M., "Vision without inversion of the retinal image," Psychological Review, 4, 341-360, 463-481 (1897).
- 4) Jackson, C.V., "Visual factors in auditory localization," Quarterly Journal of Experimental Psychology, 5, 52-65 (1953).
- 5) 八木信忠, "音声多重放送とその問題点," 映画テレビ技術 315, 24-28 (1978)
- 6) 中林克己, 辻本康, 二階堂誠也, "ステレオ音像とテレビ映像の相互作用に関する基礎実験," 音講論集2-5-10 (1979.6)
- 7) Jones, B., "Visual facilitation of auditory localization in school children: A signal detection analysis," Percept. Psychophys., 17, 217-220 (1975)
- 8) Shelton, B. and Serrle, C., "The influence of vision on absolute identification of sound-source position," Percept. Psychophys., 28, 589-596 (1980)
- 9) Warren, D.H. and Platt, B., "Auditory localization: The importance of eye movements and a textual visual environment," Percept. Psychophys., 12, 245-248 (1972)
- 10) 小宮山拱, "大画面テレビ視聴時における音像定位," 音響誌, 43, 9, 664-669 (1987).
- 11) 小宮山拱, "大画面テレビにおける映像と音像定位のずれの許容値について," 電子情報通信学会技術研究報告 IE87-109, 41-46 (1988)
- 12) Komiyama, S., "Subjective evaluation of angular displacement between picture and sound directions for HDTV sound systems," J. Audio Eng. Soc., 37, 4, 210-214 (1989)
- 13) Guilford, J.P., (秋重義治監訳): "精神測定法", 276-301, 培風館 (1959).
- 14) Jack, C.E. & Thurlow, W.R., "Effects of degree of visual association and angle of displacement on the 'ventriloquism' effect," Perceptual and Motor Skills, 37, 967-979 (1973)
- 15) 湯山一郎, "高品位テレビジョンの画面方式、走査方式," N H K 技研月報, 昭和56年11・12号, 441-446 (1981)
- 16) 黒住幸一, 辻本康, 小宮山拱, 盛田章, 大串健吾, 氏原淳一, "ハイビジョン用ステレオ音声方式," テレビ誌, 42, 6, 579-587 (1988)
- 17) Ohgushi, K., Komiyama, S., Kurozumi, K., Morita, A., Ujihara, J. and Tsujimoto, K., "Subjective evaluation of multi-channel stereophony for HDTV," IEEE Trans., Broadcast. BC-33, 4, 197-202 (1987)
- 18) 松岡進, "ステレオを聞く場合の最適聴取条件," 音響技術, 18, 34-38, (1969.5)
- 19) Theile, G., "HDTV sound systems: How many channels?," The proceedings of the AES 9th international conference, 217-232 (1991.2)
- 20) Komiyama, S., "An application of the 3-1 sound system to HDTV home receiver," Digest of technical paper in IEEE International conference on consumer electronics, 300-301 (1990)
- 21) Takegahara, T., Tanabe, H. and Suganami, H., "Sound transmission for HDTV using baseband multiplexing into MUSE video signal," IEEE Trans. Broadcast. BC-33, 4, 188-196 (1987).
- 22) Makita, Y., "On the directional localisation of sound in the stereophonic sound field," E. B. U. Review, No. 73-A (1962)
- 23) 二階堂誠也, "立体再生音場における方向定位についての一考察," 音響学会講演論文集2-1-16 (1961.10)
- 24) 遠藤俊郎, "非対称聴取位置におけるステレオ音の方向定位," 音響学会講演論文集2-3-15 (1964.5)
- 25) Nakabayashi, K., "A method of analyzing the quadrasonic sound field," J. Audio Eng. Soc., 23, 3, p. 192 (1975)
- 26) 小宮山拱, "複数音源による合成波面について," 電子情報通信学会技術研究報告, EA88-76,

- 5-10 (1989.3)
- 27) 橋川俊哉, 小宮山拱, 亀川徹, 富田弘之, "サラウンドオーディオモニターの開発", 放送技術, 46, 8 (1993)
- 28) Mertens, H., "An energy theory of directional hearing and its application in stereophony," E. B. U. Review, No. 59-A (1960)
- 29) Mertens, H., "Directional hearing in stereophony - Theory and experimental verification," E. B. U. Review, No. 92-A (1965)
- 30) 境久雄, 井上恒夫, "振幅変調音による音像定位の考察," N H K 技術研究, 23, 6 (1971)
- 31) Blauert, J., "Spatial Hearing," The MIT press, 116-137 (1983)
- 32) 森本政之, "聴覚ハンドブック 第5章," ナカニシヤ出版, 難波精一郎編, 209-217 (1984)
- 33) オルソン, "音響工学," 上巻, 近代科学社, 221 (1959)
- 34) 小宮山拱, 盛田章, 黒住幸一, 中林克己, "スピーカアレイによる音像の遠近制御," 日本音響学会講演論文集3-3-9 (1990.9)
- 35) 小宮山拱, 盛田章, 黒住幸一, 中林克己, "映像と音像の相乗効果を利用した音像の遠近制御に関する一考察," 日本音響学会講演論文集2-7-16 (1990.3)
- 36) 小宮山拱, 盛田章, 黒住幸一, 中林克己, "立体テレビのための音像遠近制御法," 1991年テレビジョン学会年次大会講演予稿集, 399-400 (1991)
- 37) Komiyama, S., Morita, A., Kurozumi, K. and Nakabayashi, K., "Distance control of sound images by a two-dimensional loudspeaker array," J. Acoust. Soc. Japan, (E), 13, 3, 171-180 (1992)
- 38) Komiyama, S., Morita, A., Kurozumi, K. and Nakabayashi, K., "Distance control system for a sound image," The proceedings of the AES 9th international conference, 233-239 (1991)
- 39) 子安勝, 山下充康, "空間相関による残響室内音場の拡散度評価," 音響誌, 26, 3, 132-143 (1971)
- 40) Gardner, M.B., "Distance estimation of 0° or 0°-oriented speech signal in anechoic space," J. Acoust. Soc. Am., 45, 47-53 (1969)
- 41) 中林克己, 二階堂誠也, 森秀一, "音源の距離の弁別能力についての一実験," 日本音響学会講演論文集2-6-10, 435-436 (1975.10)
- 42) Gotoh, T., Kimura, Y., Kurahashi, A. and Yamada, A., "A consideration of distance perception in binaural hearing," J. Acoust. Soc. Japan, 33, 667-671 (1977)
- 43) Schroeder, M.R. and Atal, B.S., "Computer Simulation of Sound Transmission in Rooms," IEEE, Conv. Rec., pt7, 150-155 (1963)
- 44) Damaske, P., "Head-related Two-channel Stereophony with Loudspeaker Reproduction," J. Acoust. Soc. Am., 50, 4(part 2), 1109-1115 (1971)
- 45) 岡部啓, "ダミーヘッドを用いた音場再生," 音響誌, 46, 8, 650-656 (1990)
- 46) Berkhout, A.J., de Vries, D., Vogel, P., "Acoustic control by wave field synthesis," J. Acoust. Soc. Am. 93 (5), 2764-2778 (1993)
- 47) Olson, H.F., "Home entertainment: Audio 1988," J. Audio Eng. Soc., 17, 4, 390-404 (1968)
- 48) Gardner, M.B., "Proximity image effect in sound localization," J. Acoust. Soc. Am. 43, 163 (1968)
- 49) 小宮山拱, "番組音の相関を用いたテープレコーダーとVTRの同期運転," N H K 技研月報 27, 3, (1983)
- 50) Komiyama, S., "Tape-lock system without time code," NHK Labs. Note, 305, (1984)
- 51) 柳町昭夫, 鈴木浩, "番組補完型テレミュージック放送の番組映像との同期ずれに関する一検討," テレビ誌, 43, 1, 78-80 (1989)
- 52) 辻本康, "海外からの衛星中継によるBモード・ステレオ放送," 映画テレビ技術 1991年10月号 32-36 (1991.10)
- 53) 太田一穂, "4チャンネルステレオ," 日刊工業新聞社, 150-160 (1971)

