

解説



音響と情報

6. 人間の聴覚と音場制御†

山崎 芳 男†

1. まえがき

乳幼児でさえ母親の声は確実に識別することができるし、合奏の中から特定の楽器の音を聞き分けることも可能である。また、人勢の人の集まったパーティ会場などで遠くの気になる人の会話が耳にとまり、注意を向けるとその会話が次第に明瞭にきこえてくる。この現象は文字どおりカクテルパーティ効果と呼ばれている。

人間はたった二つの耳で巧みに音源の方向や距離を捉えている。二つのマイクロホンでは前後や上下の方向判定は原理的に不可能である。頭を動かすことにより距離や方向を判断しているといわれているが、頭を動かさなくとも、たとえ片耳を塞いでも音源の方向や距離の判別は可能である。これは人間の聴覚能力のすばらしいところであり、到来方向や距離による微妙な音質の違いを学習記憶しているからに他ならない。コンピュータが異質的に発展した現代といえども、機械に人間のような高度な処理をさせることは不可能であるが、一歩でもこれに近づけるべく様々な研究が行われている。ここでは人間の聴覚に基づく音場の能動制御について考察する。

2. 音場の再現

2.1 バイノーラルシステム

図-1 に示すように原音場において実際の間人あるいはダミーヘッド、HATS (Head And Torso Simulator) と呼ばれる擬似頭や人間の上半身を模した人形の耳の位置に2つのマイクロホンを設置して2チャンネル収録を行い、2チャンネル

ヘッドホンを用いて聴取すると原音場にいるのと同じように感じる。これをバイノーラルシステムと呼ぶ。

ヘッドホン受聴は電気音響変換器の振動板と鼓膜による反射、装着がうっとうしい等の問題はあるが、自分とそっくりな胸部、頭部を使い自分の動きに合わせてHATSを動かせば現場にいるのと同じ音響空間を感ずることができるはずである。

ところで現行のCD等のステレオ音楽のほとんどがスピーカ受聴を前提として作られている。しかるにこれだけヘッドホンステレオが普及している現状を考えると、ヘッドホン受聴を対象としたバイノーラル録音のソフトウェアがほとんど作られていないのは不思議である。

2.2 トランスオーラルシステム

バイノーラルシステムにおいてヘッドホン受聴をスピーカ再生に置き換えたものがトランスオーラルシステムである。バイノーラルシステムのダミーヘッド等で収録された音をそのまま2つのスピーカから再生したのでは左耳のみに与えるべき音が右耳にも、右耳への音が左耳にも聞こえてしまうクロストークが発生する。そこで図-2 に示すようにスピーカに与える信号に制御系を設けこのクロストーク分をキャンセルすることにより厳密な音場の再現が可能となる。再生音場の響きが少ない場合にはM.R.Schroederや三浦種敏、浜田晴夫、岡部肇らが提唱しているように2つのスピーカでトランスオーラルシステムの実現が可能である¹⁾。

さらに部屋に響きがあっても実現可能な逆フィルタシステムMINT (Multiple-input/output Inverse Theorem) を金田豊、三好正人らが提案している²⁾。図-3 に示すように制御系を3系統にすることにより正確な制御が可能である。一般に

† Human Auditory Characteristics and Sound Field Control by Yoshio YAMASAKI (Department of Computer Science, Chiba Institute of Technology).

‡ 千葉工業大学情報工学科

$n+1$ 個の制御系を設けることにより響きのある部屋の n 点を精度良く制御することが可能となる。

2.3 キルヒホッフの積分公式に基づく音場制御

トランスオーラル系を用いてスピーカ再生による正確な音場の再現が可能である。しかし制御さ

れるのは耳の位置だけであるから、頭を動かすと動制御系は崩れてしまう。人の動きを許容するには、耳の位置の動きを検出してそれに応じて伝送系を適応制御しなければならない。さらに部屋の音響条件、窓や戸の開閉、あるいは人間の数が変わったり、家具の配置が変わったりすればそれに適応した制御を行わなければならない。これは現実には大変難しい。

キルヒホッフの積分公式によれば、図-4 に示すように、ある閉曲面内に音源がない場合、たとえばホールで音楽を聞くようなモデルに対しては、原理的には閉曲面の表面の音圧と法線方向の粒子速度を目的の音場のそれと一致させることにより所望の音場が得られる³⁾。法線方向の粒子速度は法線方向に隣接した2点の音圧から計算することが可能であるから、境界面を N 点に離散化する場合 $2N$ 点の音圧を制御すればよいことに

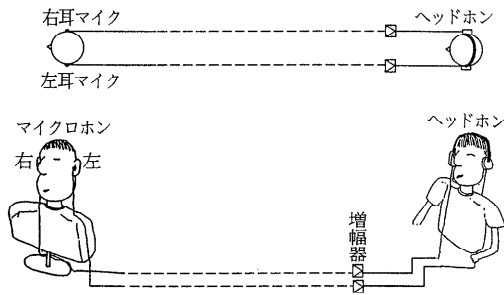


図-1 バイノーラルシステム

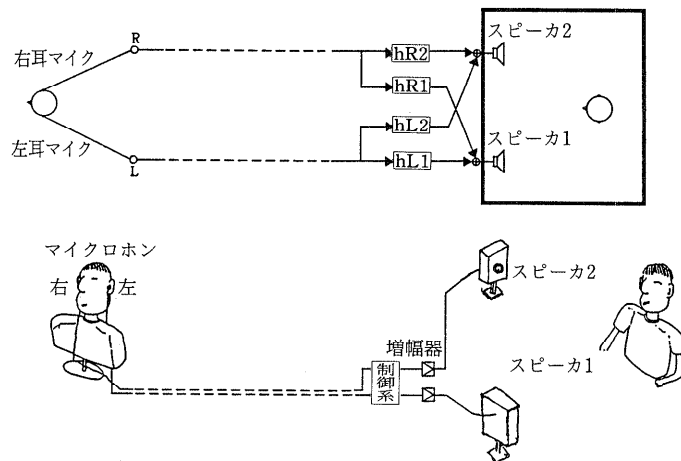


図-2 トランスオーラルシステム

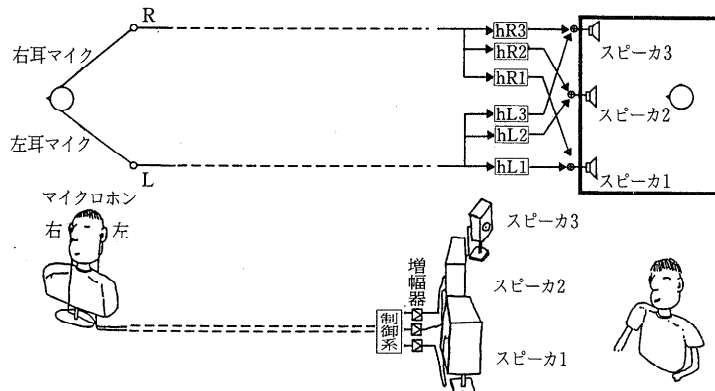


図-3 MINT 法による響きのある部屋でのトランスオーラルシステム

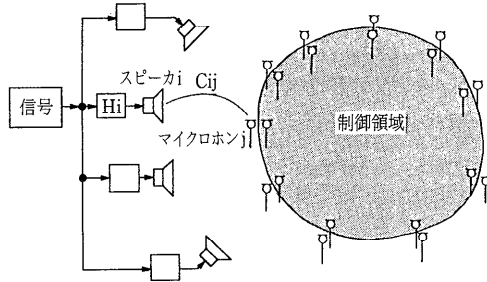
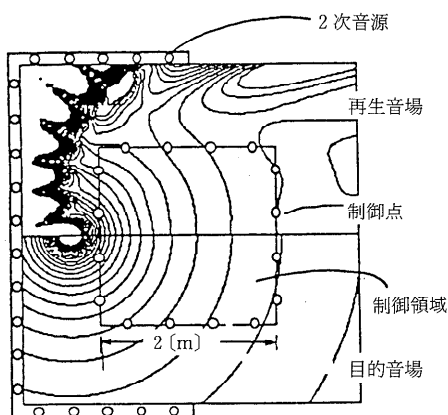
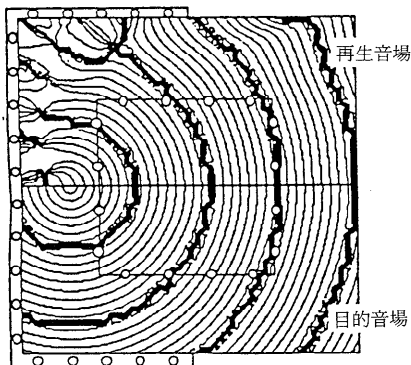


図-4 再生音場におけるシステムの構成



(a) 振幅



(b) 位相

図-5 キルヒホッフの積分公式に基づく制御結果

なる。

ただし音波の波長と人間の可聴領域とを考えると制御点を境界面上1~2cmの間隔で網羅する必要があり、現時点ではとうてい実現不可能である。

マイクロホンやスピーカの数がある程度減らしてどこまで音場を再現できるかという研究も

試みられている。図-5に周波数400 Hz(波長0.85m)、制御点の数16、間隔0.5mの場合の位相と振幅の計算結果を示す。空間の標本化定理は満足していないが、原音場と比較的よく一致している。

3. 音場の能動制御

3.1 聴取音の聴感特性や好みを考慮した音場制御

トランスオーラル系の手法に基づき、響きのある室内にいる2人の聴取者が自分の好みに応じて響きや音量、音質を独立に変えることも可能である⁴⁾。前述のように信号源、制御系、観測点の数が、自由音場や響きの少ない室内では (N, N, N) 系、ある程度響きのある一般の室では $(N, N+1, N)$ 系で正確な N 点の制御が可能である。つまり2人の4つの耳に対して $(4, 4, 4)$ あるいは $(4, 5, 4)$ 系を構成すればよいことになる⁴⁾。

容積 150 m^3 、残響時間 0.5 s の室内で $(4, 4, 4)$ の制御系を構成したところ、各スピーカから各耳までのインパルス応答は $16,384$ 点(標本化周波数 48 kHz 、約 0.3 s)で十分収束した。これらを $524,288$ 点DFT(離散的フーリエ変換)したうえで $100 \text{ Hz} \sim 15 \text{ kHz}$ の帯域制限を施し、制御用各逆フィルタの応用を計算したところ、逆フィルタが十分に収束するには約 $262,144$ 点(約 5.46 s)と室の残響時間の約16倍の長さが必要であることが分かった。図-6に、近くの人には小音量で平坦な特性を、遠くの人には大音量で高域上りの音を与える制御例を示す。このように好みの聴取位置において、任意の音量、音質あるいは響きを選ぶことができる。ただし、テレビの場合には画像信号にも音場制御に要する遅れに等しい遅延処理を施さなければならない。

3.2 騒音源が電気信号として得られる場合の能動騒音制御

最近能動騒音制御が注目を集めているが、能動騒音制御は決して万能な騒音対策ではない。能動騒音制御が有効な点は騒音を消すための付加音源が騒音源あるいは被制御点に近い場合に限られる。窓から侵入する不特定の騒音を軽減することは困難であるが、騒音源が電気信号として観察できるような場合には比較的容易に音場波長に見合った範囲の騒音の能動制御が可能である⁵⁾。

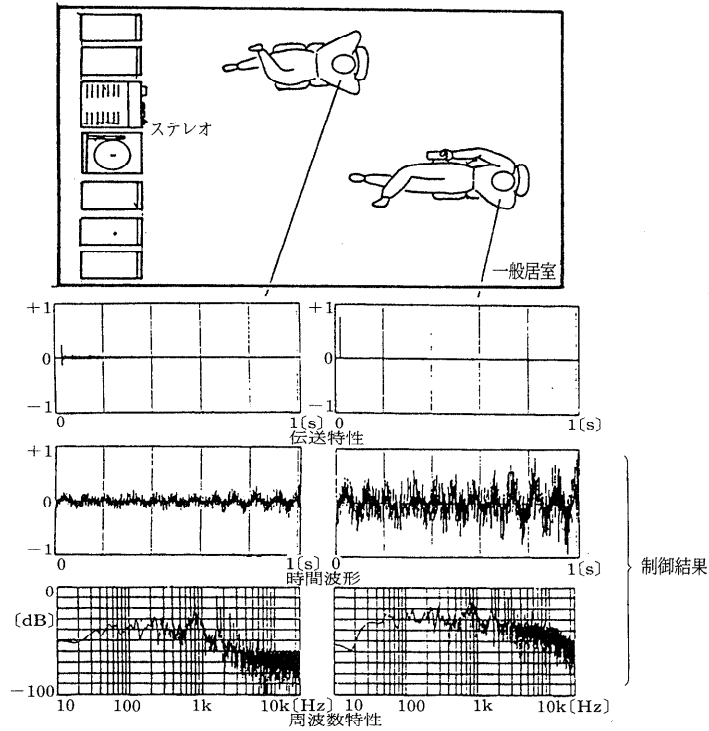


図-6 聴取者の聴感特性や好みに応じた音場制御

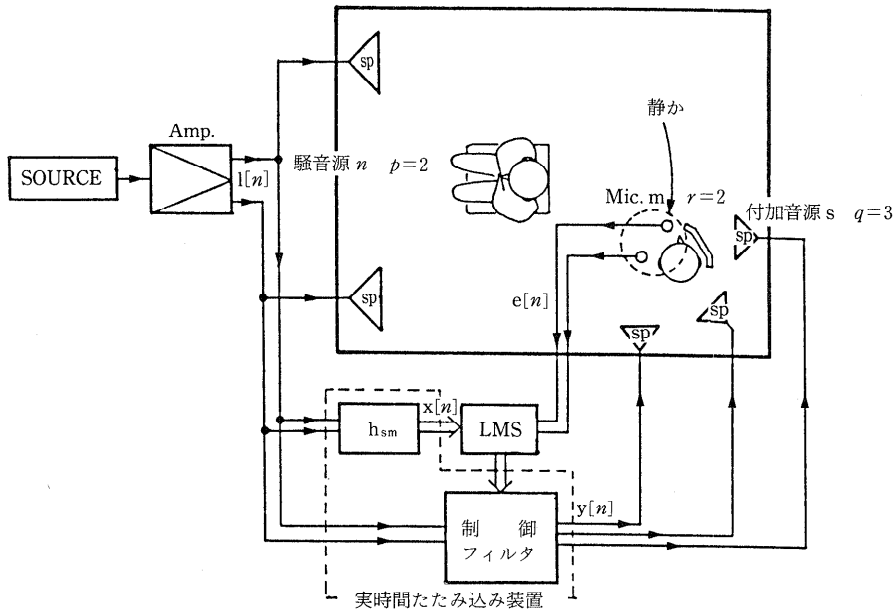


図-7 騒音源が電気信号で得られる場合の能動制御

図-7 に響きのある室内に p 個の騒音源があり、 q 個の付加音源により r 個のマイクロホンへの騒音信号をできるだけ小さくするよう制御の例を示す。

$q=r+1$ であれば MINT により音場の逆フィルタが計算可能である。前項同様、フィルタは音場の響きが増すと急激に長くなる傾向があるが、騒音源が電気信号として得られれば受音点に所望

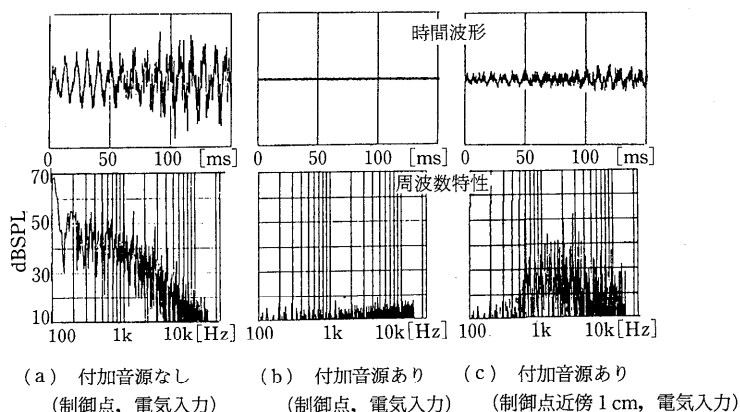


図-8 制御結果

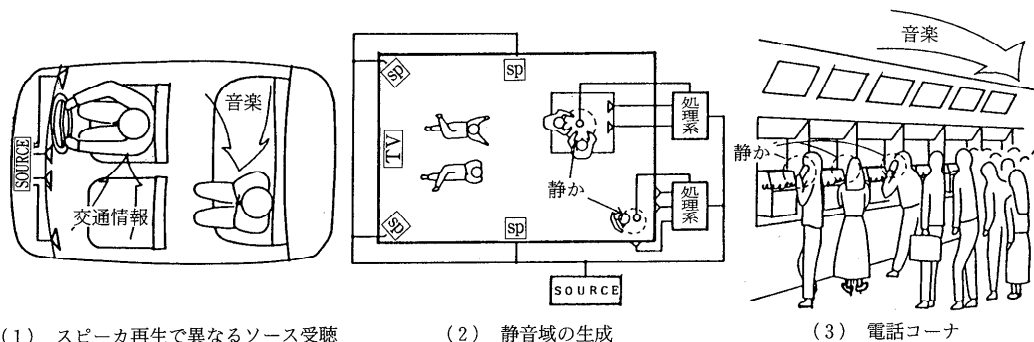


図-9 騒音源が電気信号で得られる場合の能動制御の応用例

の音を与えたり、無音になるように制御することも可能である。

容積約 115 m³、残響時間約 0.3 s の居室において、2 個のスピーカから再生される騒音に対して 3 個の制御用スピーカを使って机の上にある 1 点の騒音を制御する実験を行った⁹⁾。図-8 に制御結果を示す。

騒音源が電気信号として得られる場合は制御点の近傍 1/4 波長程度の範囲では大幅に騒音レベルが低下している。

ところで原理的には図-9(a)に示すようにスピーカ再生で 2 人に別の音、たとえば前席には交通情報、後席には音楽を聞かせることも可能である。しかし耳の位置を少しでも動かすと両方の音が聞こえてしまい実用的なシステムとはいえない。これに対し、TV やステレオを聴いている部屋に電話がかかってきたような場合に電話に回答する人の耳の位置だけを静かにするような制御系では、回答する人は積極的に静かな場所を探す、すなわち人間もフィードバックグループに入っているうへ TV やステレオの音が聞かれては困る

通話に対してマスキングの効果を生み実用的なシステムが構成できる。

同様に図-9(b)に示すようにディスコやカラオケの騒々しい空間で特定の机の上や、電話機の近傍など、ある領域を静かにするという制御も可能である。

また、電話の話声を他の人の迷惑にならないように制御したり、(c)で示すように屋外の電話コーナーで騒音制御とマスキングの効果を合わせ持たせることも原理的に可能である。

3.3 マルチスピーカ再生システム

トランスオーラル系は音場再生の有効な道具である。しかし、音場の評価や電気音響による音場補正の検討等にはスピーカを使った多人数での頭の動きを許容するシステムが望まれる場合も多い。キルヒホッフ積分に基づく制御は現段階ではシステムがあまりにも複雑になり実現困難である。

ここでは近傍 4 点法により求めた反射音位置に基づく比較的響きの少ないデッドな室内で複数のスピーカを使った再生システムを紹介する⁹⁾。近

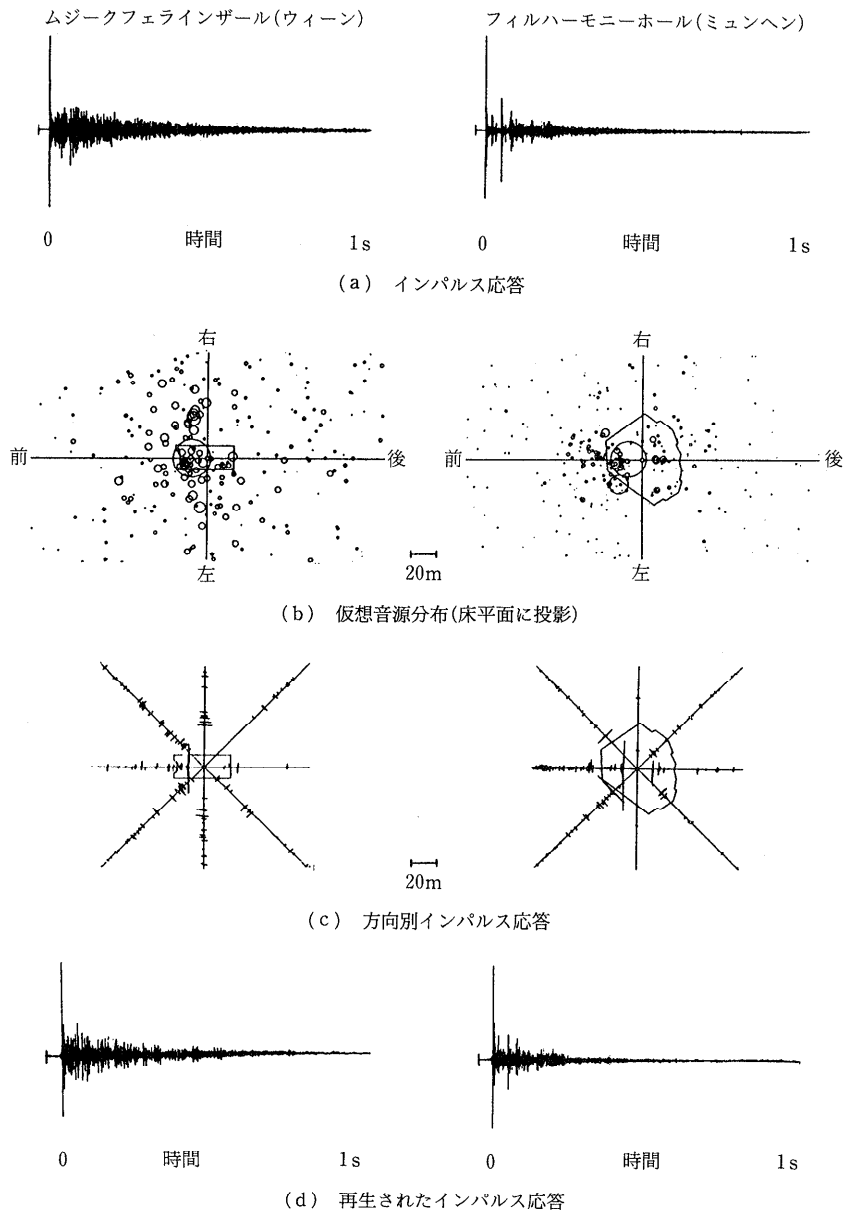


図-10 近傍4点法に基づくマルチスピーカ再生

傍4点法は同一平面上にない4点、たとえば原点と各直交軸上5cmの各点でインパルス応答を測定し、相関やインテンシティにより反射音の音場の空間情報を抽出する測定方法である⁷⁾。

近傍4点法で求められた直接音と反射音の座標と大きさから、各方向のインパルス応答を計算することができる。図-10(a)にウィーンのムジークフェラインザールとミュンヘンのフィルハーモニーホールのインパルス応答、(b)に近傍4点法

で測定した仮想音源分布、(c)にこの仮想音源分布から計算した床平面8方向のインパルス応答を示す。

スピーカの数、配置や種類(点音源、平面音源等)には様々な方法がある。ここでは受信点から2mの距離の高さ1.5mの水平面に配置した8個の小型全帯域スピーカにより再生を試みた。図-10(d)は一般居室において8スピーカで再生したインパルス応答を各スピーカの中心点で測定

した結果である。原音場のおおよその傾向が再現されている。

3.4 実時間たみ込み装置を用いた音響模型実験

音響設計に際して行われる模型実験には無響室でテープレコーダ等で録音した響きのない楽音を模型内で縮尺に応じ高速再生し、その応答を高速で録音、低速で再生するという手法が使われてきた。しかしこの方法ではマイクロホン、スピーカや記録系の性能の制約等から音質には限界があった。

これに対し、最近デジタル信号処理を導入して、模型内ではインパルス応答のみを測定し、後処理でこのインパルス応答と響きのない信号とをたみ込み演算するという手法が導入されている。この結果尺縮等の自由度が上るとともに飛躍的に音質が向上し、模型実験は実用に供されるに至った。

最近、デジタル信号処理(DSP)技術の進歩

により4秒以上のインパルス応答と実時間たみ込みが可能となった。図-11に示すように、この実時間たみ込み装置を導入することにより、響きの演奏者への影響まで含めた音場予測が可能となり、筆者らは実際の劇場等の音響設計に役立てている⁸⁾。

3.5 拡声系を持つ室内の音場制御

建築音響設計と電気音響設計はまったく別個に、それも建築音響設計が終了後あるいは、建物完成後に、電気音響設計が行われているのが現状である。しかし本来両者は密接に関連しており、本来同時に進められるべきものである。

前述のように波動理論に基づくキルヒホッフヘルムホルツ(Kirchhoff Helmholtz)の積分方程式により境界要素法を使って、音場の数値計算が可能である。

室内の聴衆にはできるだけ明瞭で均一な拡声音を供給し、外部にはできるだけ内部の拡声音が遮音の悪い窓を通して外に漏れないような音場設計

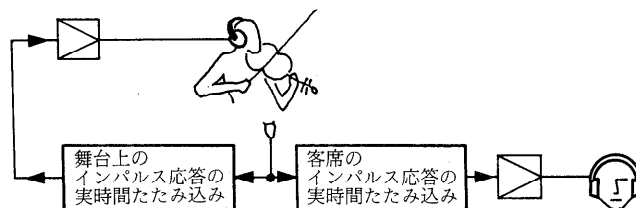
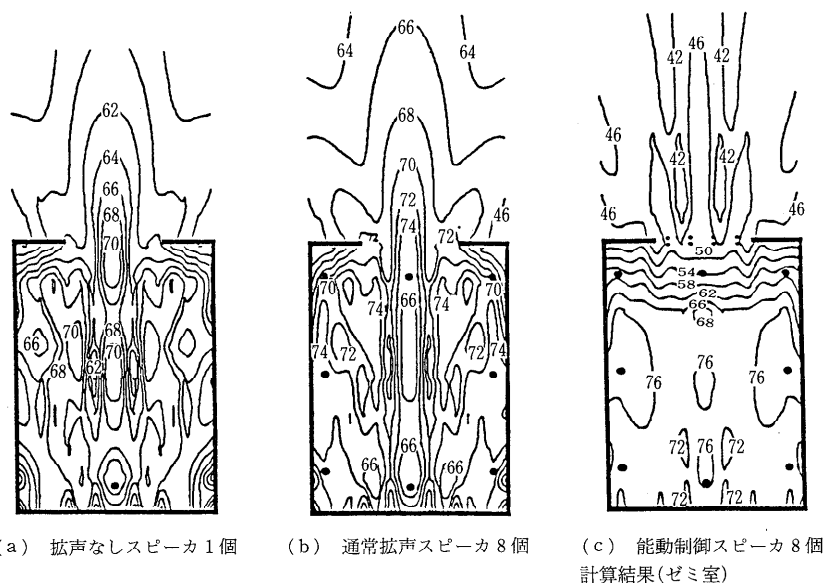


図-11 実時間たみ込み装置を使ったシミュレーション



(a) 拡声なしスピーカ 1個 (b) 通常拡声スピーカ 8個 (c) 能動制御スピーカ 8個 計算結果(ゼミ室)

図-12 拡声系を持つ室内の音場制御

例を以下に示す。

開口部と剛壁からなる2次元音場を仮定して8個の拡声装置を含めた音場の数値解析を試みた。数値計算の結果を図-12に示す。拡声を行わない(a)や制御を行わないすべてのスピーカを同時に駆動した場合(b)に比較して制御することにより(c)は内部の音圧分布が比較的均一となり、また外部への音の漏れも拡声のみものものと比較して20 dB程度低下している⁹⁾。

同様の手法による野外劇場の形状と電気音響システムの最適化、すなわち拡声された音を外部に漏らさず、内部の人には明瞭にかつ均一な聴取が可能な音場の実現が試みられている。また廊下やアトリウムを対象に、催し物の退場時や災害時の先行音定位を利用した音による誘導システムの実現を検討されている。

4. 人間の聴覚と VR

最近バーチャルリアリティ (VR) という言葉を頻繁に耳にするようになった。仮想現実感あるいは人工現実感という訳は VR を必ずしも的確に捉えていないのではないか。そもそも現実に本物も仮想もないのではないか。ここでは VR を時間・空間の移動技術と捉えることにする。

4.1 蓄音機・電話・ラジオと VR

レコード再生装置は古くは蓄音機と呼ばれ文字どおり音を蓄えることを目的としていた。昭和初期は単にポータブルといえは蓄音機のことであったように、1950年代はハイファイ (HiFi) がレコード再生装置のあるいは高級ラジオの呼称として定着していた。HiFi は High-Fidelity (高忠実度) の略であって、何に対して忠実かは多少議論があるが、演奏会場の音をそのまま自宅に持ち込もうという願望の込められた呼称である。

立体音響 (ステレオ) の登場とともに臨場感という言葉が盛んに使われはじめた。この臨場感こそ VR を最も的確に捉えた表現といえるのではないか。日本においても伊藤¹⁰⁾、牧田¹¹⁾らにより立体音響の理論検討や実験が行われている。

電話もまたその原語 Telephone が示すとおり VR のテレグジスタンス (tele-existence) を志向したシステムであることは間違いない。1933年にはワシントン、フィラデルフィア間で電話回線を3通話路使った遠距離立体音響伝送の実験も

試みられていた¹²⁾。電話の通話品質は決して高いものではないが、電話を多くの人々がそれを VR と意識するか否かは別として自然にテレグジスタンスの道具として利用している。

4.2 前より後ろの音に敏感

人間にとって当然視覚が重要な情報摂取手段である。しかし目で捉えられる範囲は前方に限られ、それほど広いものではない。したがって人間は元来後ろからの情報は多くは聴覚、気配というものに頼っている。実際、聴覚は後ろに対して驚くほど鋭い識別能力を持っている。音については後ろからの情報が大切であるということを常に認識する必要がある。

3.3 で述べた我々が行った実験においても制御点の数が同じ場合、その配置は常に前方よりも後ろの数が多の場合に自然に感ずるという結果が得られている。これは簡単な定位の実験をすれば明らかである。すなわち眼をつぶって前方に人が立ってしゃべった場合、定位は比較的曖昧であるが、後方については的確に位置や距離を判断できる。

4.3 片目・片耳で VR 実現

視覚の VR は一般に HMD や偏光眼鏡等を使った2眼視で行われているが、1987年9月のテレビジョン学会誌にテレビジョンを片目で見れば自然で臨場感があるという木内雄二氏の大変興味深い寄稿があった¹³⁾。実際、TV画面や写真を片目で見ると非常にリアルである。これは一台のカメラでとったものを両目で見ると無理があるのであって、片目で見れば自然であるという理になつたものである。

耳についても音質評価等ではステレオよりもモノラル音源の方が的確な情報を得られることが知られている。案外割り切って通常の画面とモノラル收音ソースにより片目片耳で実用上十分な VR が実現できるのかもしれない。この場合 HMD は片目片耳を覆うだけの簡単な道具ですむことになる。

5. むすび

人間の五感のうちで聴覚、視覚と触覚はそれぞれ音声、文字と手話という言葉を持ち、情報のやりとりを司る重要な感覚である。聴覚は視覚と補完しあいながら重要な情報処理を行っている。養

老孟司氏によれば視覚は空間、聴覚は時間に関連し、視覚は直感、静的であるのに対し聴覚は論理的、動的に情報を捉えるというようにきわめて対照的であり、その両者を結ぶものが言語であるそう¹⁴⁾。そして氏の専門である解剖学的に脳を見れば、視覚と聴覚・運動に係わる部分は、それぞれ30%程度であるそう¹⁴⁾。しかるに百聞は一見にしかずなどと言われるように情報処理において映像に比較して音響信号に割り当てられる割合は非常に小さいのが現状のよう¹⁴⁾。「脳と心」を実体と考えると音の割合をもう少し増やす必要があるのではないか。

参 考 文 献

- 1) 浜田晴夫：バイノーラル音場再生系について，音響学会誌，Vol. 48，No. 4，pp. 250-257 (1992)。
- 2) 三好正人，金田 豊：音場の逆フィルタ処理に基づく能動騒音制御，音響学会誌，46，pp. 3-10 (1990)。
- 3) 伊勢史郎，鮫島俊哉，山崎芳男：キルヒホッフ積分公式に基づく音場制御理論とその応用，音響学会建築音響研究会資料 AA 94 5 (1994.1)。
- 4) 鮫島俊哉，白石吾朗，前田英邦，松本文孝，山崎芳男：聴取者の聴感特性や好みを考慮した室内の複数点の音場制御，音講論集，pp. 837-838 (1992.10)。
- 5) 岡田俊哉，木村建一，金子元司，白石吾朗，前田英邦，山崎芳男：騒音源が電気信号として得られる場合の多点能動制御，音講論集，pp. 591-592 (1991.3)。
- 6) 山崎芳男，清山信正，吉原和弘，村田清孝，伊藤 毅，橘 秀樹，前川純一，森本政之，平澤佳男：仮想音源分布に基づくマルチスピーカによる音場の評価システム，音講論集，pp. 559-560 (1988.3)。
- 7) Yamasaki Y. and Itow T.: "Measurement of Spatial Information in Sound Fields by Closely Located Four Point Microphone Method, J. Acoust. Soc. Jpn. (E) Vol. 10, No. 2, pp. 101-110 (1989)。
- 8) 鶴田泰弘，木村建一，岡田俊哉，金子元司，高林和彦，山崎芳男：実時間たみ込み装置の音響模型実験への導入，音講論集，pp. 697-698 (1989.10)。
- 9) 鮫島俊哉，伊勢史郎，工藤修二，山崎芳男，山本貴紀，小沢 仁：数値計算による拡声系を持つ室内の音響設計，音講論集，pp. 839-840 (1994.3)。
- 10) 伊藤 毅：音響工学原論(下)，コロナ社。
- 11) Makita Y.: On the Direction Localization of Sound in a Stereophonic Sound Field, European Broadcasting Union Review, 37 A, pp. 102-108 (1962.6)。
- 12) Symposium on Wire Transmission of Symphonic Music and Reproduction in Auditory Perspective, E. E., Vol. 53 (1934).
Fletcher H., "Basic Requirement", pp. 9-11, J. C. Steinberg and W. Snow, "Physical Factors", pp. 12-17.
E. C. Wente and A. L. Thuras, "Loudspeaker and Microphones", pp. 17-24.
E. O. Scriven, "Amplifiers", pp. 25-28.
H. F. Afel, R. W. Cjesnut and R. F. Mills, "Transmission lines", pp. 216-219.
- 13) 本内雄二：テレビ画像の立体感を強める一方法，テレビジョン学会誌，41，pp. 834-835 (1994.3)。
- 14) たとえば養老孟司，吉田直哉：目から脳に抜ける話，筑摩書房。

(平成6年12月27日受付)

山崎 芳男

1968年早稲田大学理工学部電子通信工学科卒業，1970年同修士課程修了，以降早稲田大学理工学研究所，千葉工業大学情報工学科，早稲田大学理工学総合研究センターにおいて音響学，デジタル信号処理，建築音響等に関する研究に従事，現在千葉工業大学情報工学科・早稲田大学理工学総合研究センター教授。1984年，1990年日本音響学会佐藤論文賞受賞，工学博士（早稲田大学）
著書に「デジタルオーディオ」「音響システムとデジタル処理」（通信学会）等。