

通信会議における遠方音声収録のための残響抑圧方式の研究

古家, 賢一

<https://doi.org/10.15017/459055>

出版情報 : Kyushu University, 2005, 博士 (芸術工学), 論文博士
バージョン :
権利関係 :

第2章

残響に関する予備知識と従来の残響抑圧技術

2.1 残響の表現

残響とは、室内音場において音源から放射された音が室内の壁・床・天井などで反射を繰り返し、受音点（聴取者、マイクロホン）において直接音に続いて到達し、音源が停止した後にもしばらく残る音である。ここでは、室内の直接音と残響音を線形システムでモデル化し、音源から受音点までの音響伝達特性すなわちインパルス応答を用いて残響特性を表現する。

2.1.1 線形システムによるモデル化

通常、可聴周波数帯 20~20kHz、音圧 0~100dB 程度の範囲において、ここで扱う家庭のリビングや会議室の容積程度の室内音場は、重畳の原理が成り立つ線形システムであると考えられる。つまり、音源から入力信号 $s_1(k)$ を放射したときの受音点での出力信号を $x_1(k)$ 、入力信号 $s_2(k)$ を放射したときの出力信号を $x_2(k)$ としたとき、入力信号 $As_1(k) + Bs_2(k)$ を放射したときの出力信号が $Ax_1(k) + Bx_2(k)$ となることを意味する。ただし、 A, B は任意定数である。このような、線形システムでは入力として δ 関数（インパルス）を用いた場合の

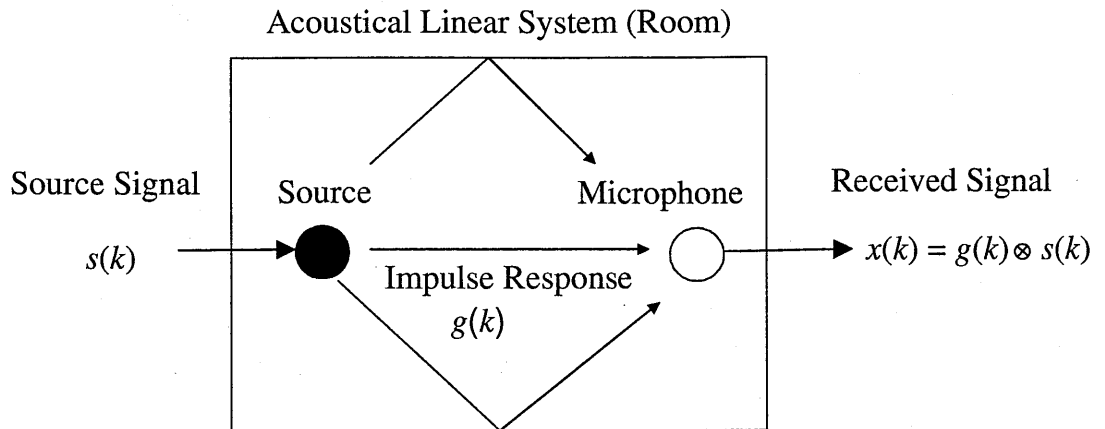


図 2.1: 残響の線形システムによるモデル化

出力（インパルス応答）が測定できれば，どのような入力信号に対する出力信号も，インパルス応答との畳み込みにより表現できる．室内インパルス応答を $g(k)$ ，音源から放射される入力信号を $s(k)$ ，受音点で観測される出力信号を $x(k)$ とした場合，次式が成り立つ．

$$x(k) = g(k) \otimes s(k) \quad (2.1)$$

但し， \otimes は畳み込み演算を表す．

2.1.2 伝達関数

インパルス応答を z 変換すれば，音源と受音点間の伝達関数が得られる．

$$G(z) = \sum_{n=0}^{\infty} g(n)z^{-n} \quad (2.2)$$

ここで， $z = re^{j\omega}$ であり， $|z| = 1$ における伝達関数の特性が周波数特性であり，その絶対値が振幅周波数特性（周波数スペクトル特性），位相が位相周波数特性である．

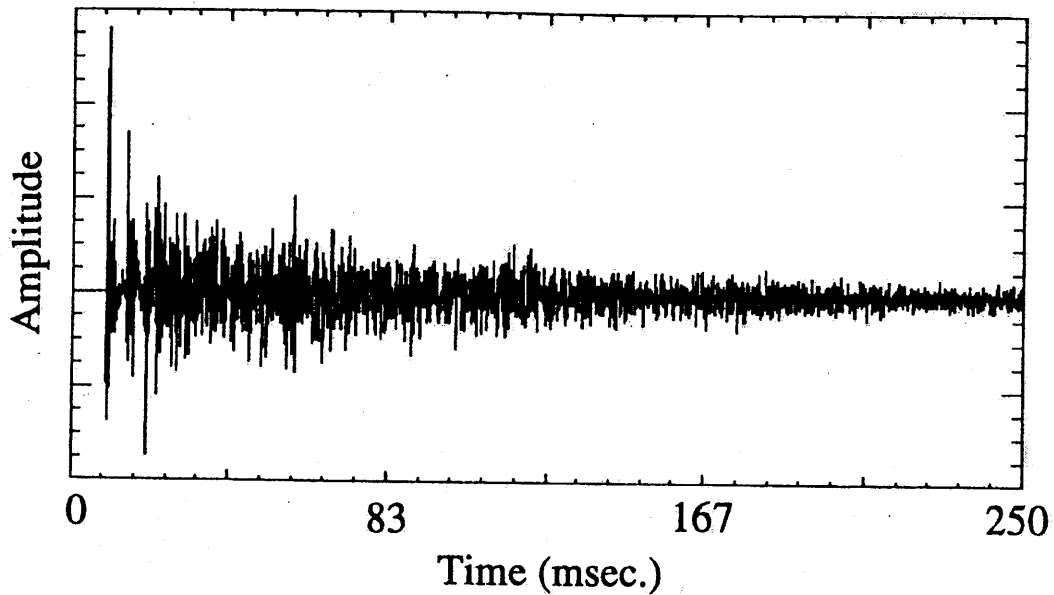


図 2.2: 室内インパルス応答の測定例

2.1.3 最小位相特性と非最小位相特性

逆システムを用いて残響抑圧を行う上で、伝達関数の重要な性質として最小位相特性、非最小位相特性がある [49]。最小位相であるとは、伝達関数のすべての極と零点が単位円の内部 $|z| \leq 1$ に存在していることである。このとき同じ振幅周波数を持つ伝達関数のうち最も位相遅れが小さくなることから最小位相特性と呼ぶ。伝達関数が最小位相であれば、次式を満たす因果的で安定な逆関数 $G(z)^{-1}$ が存在する。

$$G(z)G(z)^{-1} = 1 \quad (2.3)$$

また、逆に伝達関数が最小位相でない場合を非最小位相特性と呼ぶ。伝達関数が非最小位相特性を持つ場合の因果的で安定な逆関数 $G(z)^{-1}$ が存在しない。

2.1.4 室内インパルス応答の測定法

音源から受音点までの室内インパルス応答を測定するには、持続時間が短いパルス信号を近似的なインパルスとして用いて測定すればよい。しかし、このようなパルスは波高値の割にエネルギーが小さく S/N の良い測定を行うためには長時間の同期加算が必要である。そこで、測定用信号として時間引き延ばしパルス (TSP 信号) [50] や M 系列信号 [51] のような時間的にエネルギーを分散し波高値を小さくし継続時間が長い信号が一般に用いられる。さらに TSP 信号や M 系列信号は自己相関関数が δ 関数であり、観測信号と測定信号の相関を取るだけでインパルス応答を計算できる特徴がある。図 2.2 に、M 系列法で測定した室内インパルス応答の一例を示す。

2.2 残響特性を表す物理指標

2.2.1 残響曲線と残響時間

残響曲線とは、残響のある室内で広帯域白色雑音などをスピーカから放射し、定常状態になった時点で音源を停止したあとの残響エネルギーの減衰過程を示したものである。残響の多い部屋では、時間が過ぎても残響エネルギーが大きい。Schroeder[54]によれば、室内インパルス応答 $g(t)$ と残響曲線 $D(t)$ は、次の関係がある。

$$D(t) = 20 \log_{10} \left(\frac{\int_t^{\infty} g^2(\tau) d\tau}{\int_0^{\infty} g^2(\tau) d\tau} \right) \quad (\text{dB}) \quad (2.4)$$

図 2.2 のインパルス応答の残響曲線を図 2.3 に示す。

Sabin[55]による残響時間はこの残響曲線を用いて、室内の残響エネルギーが定常の値から 60dB 減衰するまでの時間として算出される。また、聴感上の残響感において重要と言われている V.Jordan[55]が提案した初期残響時間 (EDT)

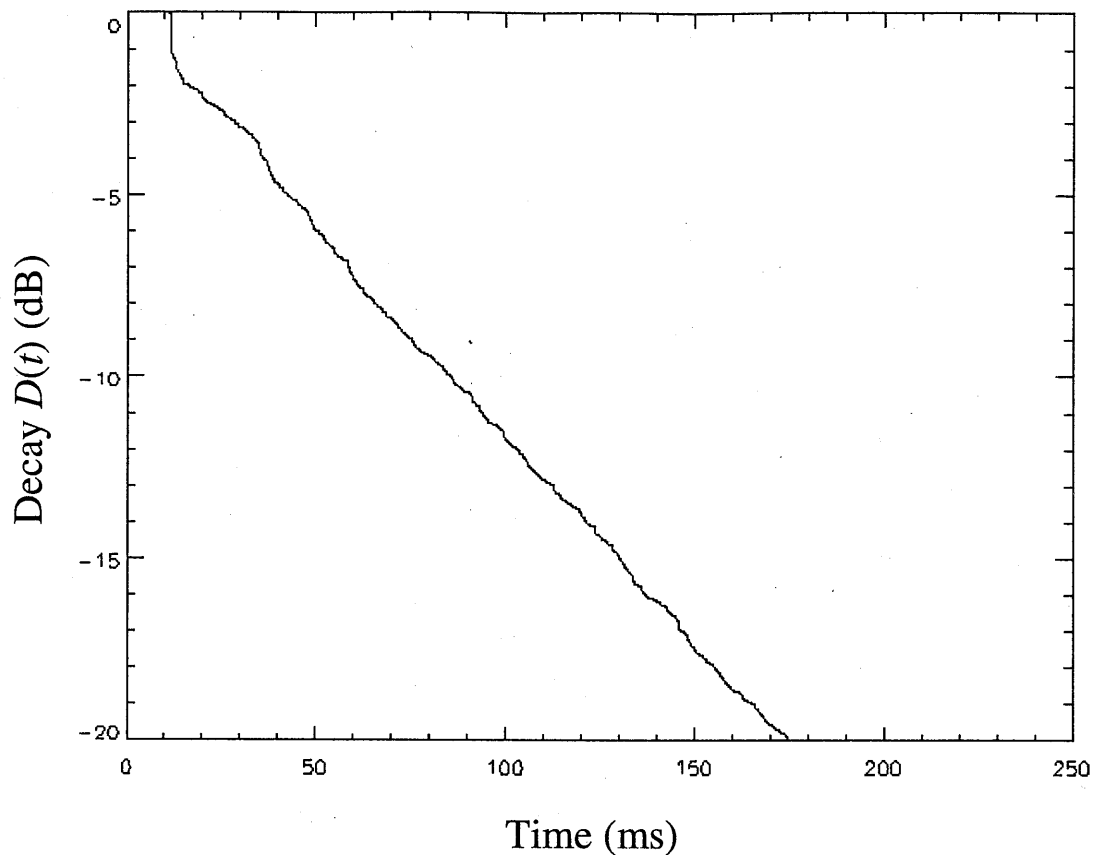


図 2.3: 残響曲線

は定常状態から 10dB 減衰する時間を 6 倍した値で計算される。EDT は、必ずしも Sabin の残響時間と一致はしないが、聴感実験によれば主観的な残響感と強い相関があることが確認されている [55]。

2.2.2 D 値（音の明確性）

Thiele [55] は、室内の残響による音の明確性（distinctness）に関する音響品質に対応した物理指標として、インパルス応答から D 値を求めた。音源が音声の場合、残響は基本的には音声の時間構造を不鮮明にし、連続する音素や音節を混濁させ、聞き取り難くする。しかし、ハース効果として知られているよう

に、反射音の遅延時間と相対振幅が特定の限界値以下であれば反射音は直接音と分離して聞こえず、直接音を補強する効果がある。このD値は、この人間の聴覚特性を考慮して、直接音から50msより短い遅延の残響音は直接音に有用に作用するエネルギーとして直接音エネルギーと和を取り、全体のエネルギーとの比をとったものである。

$$D_{50} = \frac{\int_0^{50\text{ms}} g^2(\tau) d\tau}{\int_0^{\infty} g^2(\tau) d\tau} \times 100 \quad (\%) \quad (2.5)$$

Bore[55]は、様々な部屋においてD値と単音節明瞭度との関係を求めた。その結果を図2.4に示すが、単音節明瞭度とD値の間には良好な相関関係があることが分かる。したがって、物理指標としてD値が高ければ音声が聞き取りやすいと考えられる。

2.2.3 残響抑圧改善量

残響抑圧効果のひとつの指標として、どの程度、全体のエネルギーの中で残響エネルギーを小さくできたかをdB表示する残響抑圧改善量がある。残響エネルギーは、時間と共に減衰していくので、D値に習って直接音から50msの時点での残響エネルギーで評価する。残響曲線の直接音から50msまでに減衰した残響エネルギー LD_{50} を次式で定義する。

$$LD_{50} = 10 \log_{10} \frac{\int_0^{\infty} g^2(\tau) d\tau - \int_0^{50\text{ms}} g^2(\tau) d\tau}{\int_0^{\infty} g^2(\tau) d\tau} \quad (\text{dB}) \quad (2.6)$$

LD_{50} の残響抑圧処理前と処理後の差分を残響抑圧改善量 (Reverberation Reduction Rate: RRR) として、次式で定義する。

$$\text{RRR} = LD_{50:\text{before}} - LD_{50:\text{after}} \quad (\text{dB}) \quad (2.7)$$

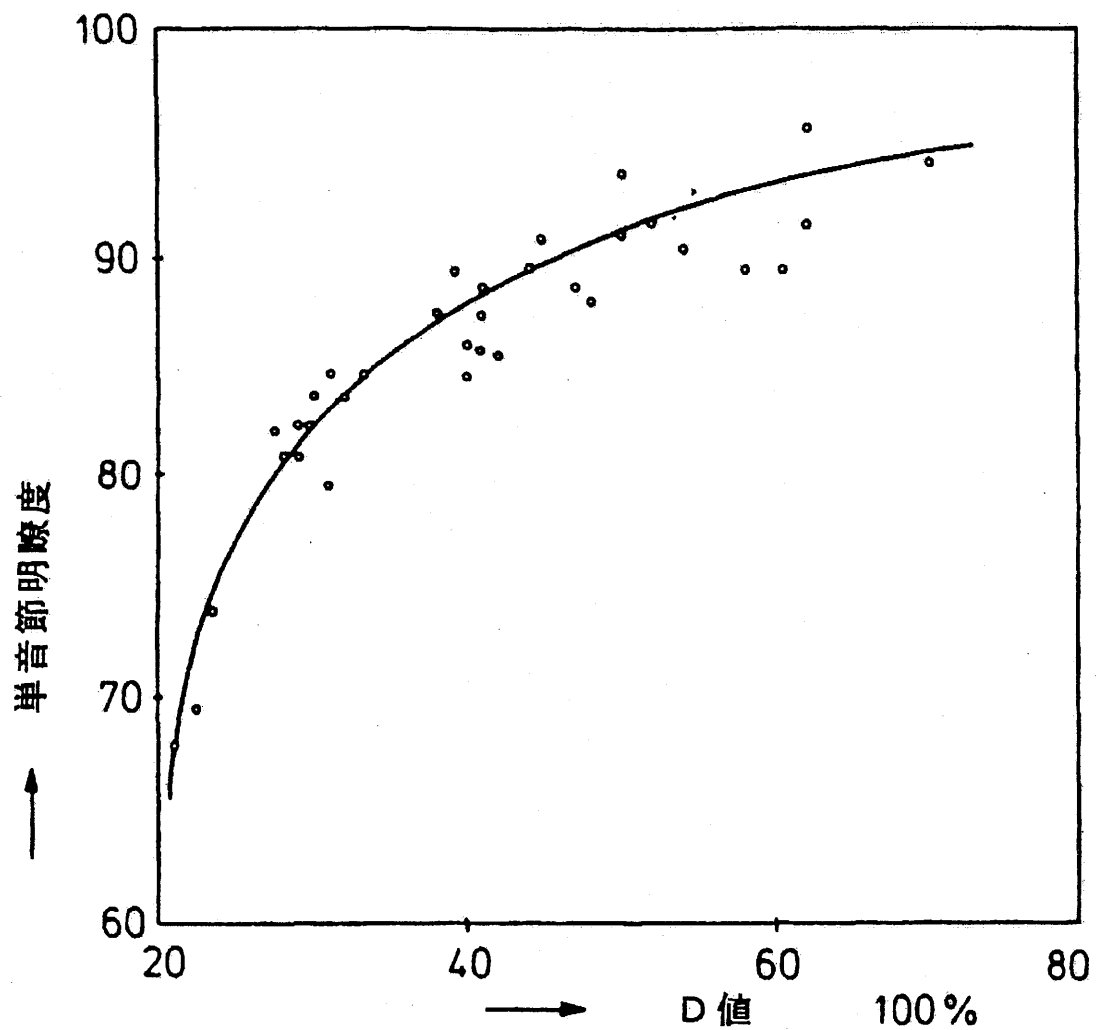


図 2.4: 単音節明瞭度と D 値の関係 (Bore[55] による)

ここで、 $LD_{50.before}$ は残響抑圧処理前の LD_{50} を示し、 $LD_{50.after}$ は残響抑圧処理後の LD_{50} を示す。これは、D 値が直接音エネルギーが全体に占める割合を示めているのに対して、残響抑圧改善量では残響エネルギーを基準にどの程度小さくなったかを示す指標である。

2.3 音声の長時間平均スペクトル

音源信号とインパルス応答が両方未知の場合、残響抑圧問題では何らかの手掛かりを用いて音源信号とインパルス応答を分離し、音源信号を抽出しなければならない。手掛かりの一つとして、音源信号の統計的な性質である長時間平均スペクトルがある。単独に発生される音声においては、音節中の音韻や会話などのような連続音声の中の音韻は、時間と共に変化しており、いわゆる非定常信号であるが、十分な長さの時間範囲では周波数スペクトルはエルゴード過程¹に従って与えられるとみなせる [56]。男性、女性の会話音声の15分間の長時間平均スペクトルを図 2.5 に、無意味音節と会話音声の10~15分の長時間平均スペクトルを図 2.6 に示す。三浦ら [56] によると、

- (a) 個人による長時間平均スペクトルの有意差はない。
- (b) 性別による 160Hz 以上の帯域では有意差はない。
- (c) 発声内容による差はない。

と示されている。したがって、長時間平均スペクトルは、話者や発話内容に比較的ロバストな音声の統計的な特徴量として用いることができる。

¹一般に、ある確率過程に関する集合平均と時間平均とが一致する場合の確率過程をエルゴード過程と呼ぶ。

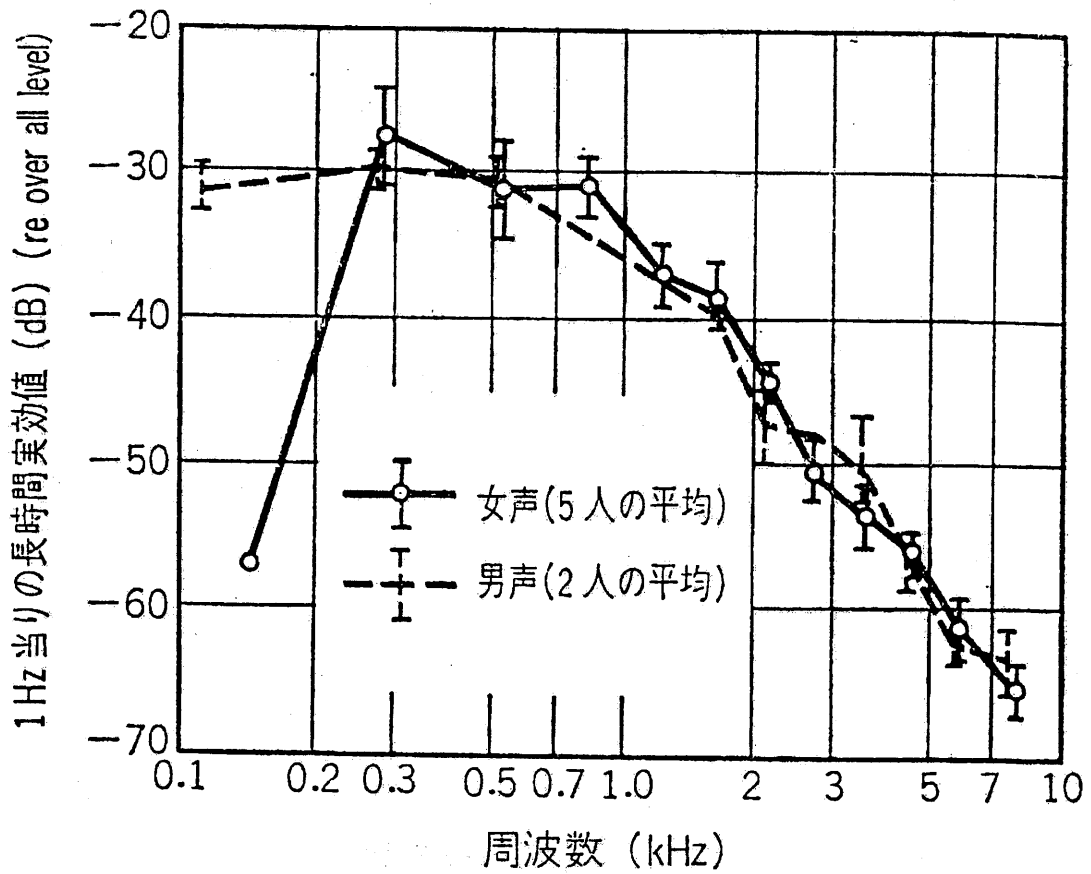


図 2.5: 人の会話音声長時間平均スペクトル: 平均値と標準偏差 (三浦 [56] による)

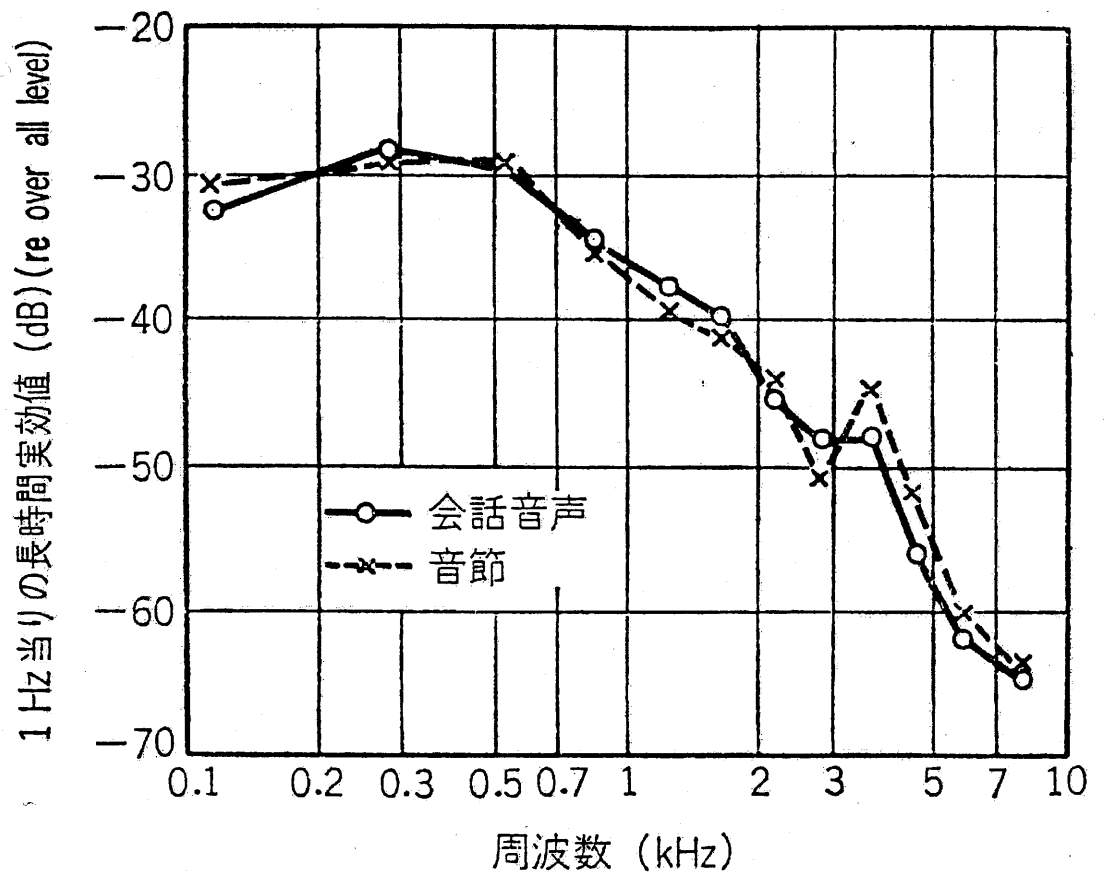


図 2.6: 男性話者による会話音声と音節の長時間平均スペクトル (三浦 [56] による)

2.4 相関行列の固有値と固有ベクトル

マイクロホンでの観測信号のみから計算できる特徴量として相関行列がある。相関行列には、音源の統計的な特徴、インパルス応答の特徴が含まれており、残響抑圧を行う上では重要な役割を果たす。

2.4.1 相関行列

N チャンネルの入力信号 $\mathbf{X} = [X_1 X_2 \cdots X_N]$ に対して入力相関行列 \mathbf{R} を次式のように定義する。

$$\begin{aligned} \mathbf{R} &= E\{\mathbf{X}^T \mathbf{X}\} \\ &= \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \cdots & \mathbf{R}_{1N} \\ \mathbf{R}_{21} & \mathbf{R}_{22} & \cdots & \mathbf{R}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{N1} & \mathbf{R}_{N2} & \cdots & \mathbf{R}_{NN} \end{bmatrix} \end{aligned} \quad (2.8)$$

但し、 $\mathbf{X}_j = [x_j(k) \ x_j(k-1) \ \cdots \ x_j(k-(L-1))]$,

$$\mathbf{R}_{ij} = \begin{bmatrix} r_{ij}(0) & r_{ij}(-1) & \cdots & r_{ij}(-L+1) \\ r_{ij}(1) & r_{ij}(0) & \cdots & r_{ij}(-L+2) \\ \vdots & \vdots & \ddots & \vdots \\ r_{ij}(L-1) & r_{ij}(L-2) & \cdots & r_{ij}(0) \end{bmatrix},$$

$r_{ij}(n)$ は i 番目入力信号 $x_i(k)$ と j 番目入力信号 $x_j(k)$ の相関関数を表し、

$r_{ij}(n) = E\{x_i(k)x_j(k+n)\}$ ($i = 1, \dots, N, j = 1, \dots, N$) であり、 $E\{\cdot\}$ は期待値を表す。実際の応用では、期待値は次式のような標本平均により計算する。

$$r_{ij}(n) = \frac{1}{M} \sum_{k=0}^{M-1} x_i(k)x_j(k+n) \quad (2.9)$$

但し、 M は標本数である。

2.4.2 固有値問題

N 次正方行列 \mathbf{R} に関する次の方程式を満たすベクトル \mathbf{q} を求める問題を固有値問題と呼ぶ..

$$\mathbf{R}\mathbf{q} = \lambda\mathbf{q} \quad (2.10)$$

ここで、 λ は未定助変数である。これは、正方行列 \mathbf{R} にベクトル \mathbf{q} を掛けたベクトルが \mathbf{q} と平行であるような \mathbf{q} を求める問題である。 N 次正方行列 \mathbf{R} は、次式の固有方程式の根として、ちょうど N 個の固有値 λ_i を持つ。

$$\det(\mathbf{R} - \lambda\mathbf{I}) = 0 \quad (2.11)$$

助変数 λ の λ_i に対してのみ式 (2.10) を満たす固有ベクトル \mathbf{q}_i が存在する。固有値のうち、最大の値のものを最大固有値、最小のものを最小固有値、それぞれに対する固有ベクトルを最大固有ベクトル、最小固有ベクトルと呼ぶ。

2.4.3 相関行列の固有ベクトルの性質

相関行列 \mathbf{R} の固有ベクトルは次の特殊な性質を持つ。

(a) 固有ベクトルは互いに直交する。つまり、任意の i, j の組に対して、

$$\mathbf{q}_i^T \mathbf{q}_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \quad (2.12)$$

(b) 固有値は、 $\lambda_i \geq 0$ である

(c) 固有ベクトルを並べた固有行列 $\mathbf{Q} = [\mathbf{q}_0 \mathbf{q}_1 \cdots \mathbf{q}_{N-1}]^T$ は、実直交行列である。つまり、

$$\mathbf{Q}^T \mathbf{Q} = \mathbf{I} \quad (2.13)$$

である。

2.5 FFT(高速フーリエ変換)による畳み込み演算の高速化

逆フィルタや室内インパルス応答との畳み込み演算は、多くの演算量が掛かる。この畳み込み演算を高速に計算する方法としてFFT（高速フーリエ変換）[60]を利用した計算法がある。

長さ L の $x_1(n)$ と $x_2(n)$ の畳み込み演算は、次式で書ける。

$$x_3(n) = \sum_{m=0}^{L-1} x_1(m)x_2(n-m) \quad (2.14)$$

$x_1(n), x_2(n)$ の M 点 DFT（離散的フーリエ変換）は、それぞれ次式で書ける。

$$\begin{aligned} X_1(k) &= \sum_{m=0}^{M-1} x_1(m)e^{-j(2\pi/M)mk} \\ X_2(k) &= \sum_{r=0}^{M-1} x_2(r)e^{-j(2\pi/M)rk} \end{aligned} \quad (2.15)$$

但し、 $M \geq 2L - 1$ である。

ここで、 $x_1(n), x_2(n)$ の M 点 DFT の乗算は、

$$X_1(k)X_2(k) = \sum_{m=0}^{M-1} \sum_{r=0}^{M-1} x_1(m)x_2(r)e^{-j(2\pi/M)mk} e^{-j(2\pi/M)rk} \quad (2.16)$$

となる。その逆変換は、

$$\begin{aligned} \text{IDFT}\{X_1(k)X_2(k)\} &= \frac{1}{M} \sum_{k=0}^{M-1} X_1(k)X_2(k)e^{j(2\pi/M)nk} \\ &= \sum_{\substack{m=0 \\ L-1}}^{M-1} x_1(m) \sum_{r=0}^{M-1} x_2(r) \left\{ \frac{1}{M} \sum_{k=0}^{M-1} e^{-j(2\pi/M)k(n-m-r)} \right\} \\ &= \sum_{m=0}^{L-1} x_1(m)x_2(n-m) \\ &= x_3(n) \end{aligned} \quad (2.17)$$

となり、式 (2.15) の畳み込み演算と等しくなる、つまり DFT の乗算により畳み

込み演算が計算できる。但し、

$$\frac{1}{M} \sum_{k=0}^{M-1} e^{-j(2\pi/M)k(n-m-r)} = \begin{cases} 1 & r = (n-m) + lM \\ 0 & \text{otherwise} \end{cases}$$

である。

通常、式 (2.14) の計算には L^2 の乗算が必要である。一方、FFT を利用して式 (2.15) を計算すれば、 $2M \log_2 M$ の乗算で計算できしたがって、式 (2.17) の計算も FFT を利用して $3M \log_2 M + M$ で計算できる。ここで、 $M \geq 2L - 1$ なので簡単のため $M = 2L$ とすると、FFT を利用した畳み込み演算の計算には $6L \log_2 2L + 2L$ の乗算で計算できる。これは、 L が大きくなるにしたがって、式 (2.14) を直接計算する場合に比べ、大幅な高速化が図れる。

2.6 従来の残響抑圧技術とその問題点

従来より様々な残響抑圧手法が提案されている。大別して、事前にインパルス応答を測定する方法と事前測定が不要な方法がある。

前者の事前にインパルス応答を測定する方法としては、MINT 法 [44]、帯域分割 MINT 法 [45] などがある。これらの方法は、インパルス応答が測定できればかなり高い残響抑圧効果が期待できるが、一般的な家庭での使用や日常の会議などにおいては事前にインパルス応答を測定するのは難しい。また、一般に時間とともにインパルス応答は変動するので、利用の度にインパルス応答を測定することは現実的でない。

後者の事前測定が不要な方法としては、最小位相逆フィルタを構成する方法 [46]、位相特性の補正も含めた波形回復を目的とした Blind Deconvolution 法 [47]、[48]、マイクロホンアレーを用いて、目的音方向に指向性を向けそれ以外の方向からの残響成分を抑圧する遅延和アレー法 [24],[57] などがある。最小位相逆フィ

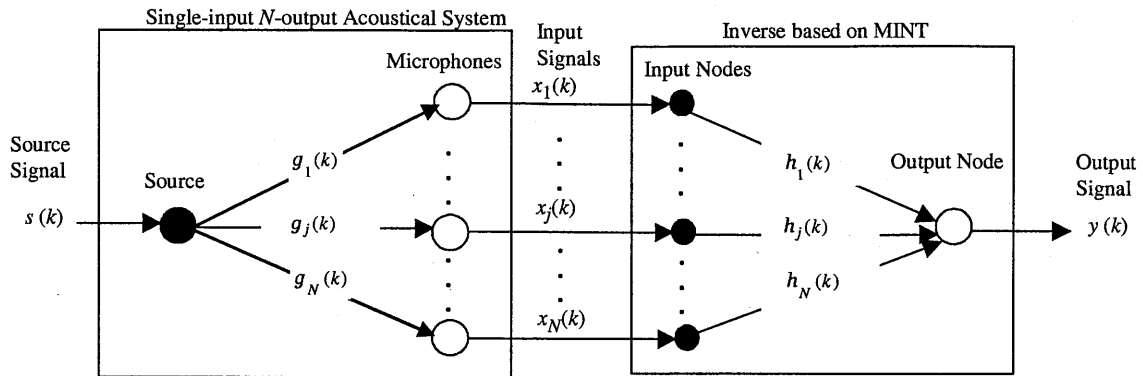


図 2.7: 1 入力 N 出力音響システムに対する MINT 法による残響抑圧

ルタを構成する方法では、非最小位相特性を持つ室内インパルス応答の最小位相成分のみ用いて逆フィルタを構成し、周波数振幅特性のみ補正しており位相特性の補正を行っていないため音源波形の回復は困難であり残響抑圧量も小さい。Blind Deconvolution 法では、逆フィルタを構成するためにはインパルス応答長を与えなければならず、適切に与えられない場合は誤差が大きく不安定であるという問題があった。遅延和アレー法は、最も簡単で容易に実現できるが抑圧量を大きくするためには多数のマイクロホンが必要となる問題点があった。

ここでは、事前測定が必要な方法として MINT 法 [44] を、事前測定が必要でない方法として遅延和アレー法 [24],[57] を説明する。

2.6.1 MINT 法

図 2.1 のような 1 入力 1 出力系の音響システムでは、伝達関数が最小位相であれば、式 (2.3) の逆関数つまり逆の特性を持つ逆フィルタが存在し、残響を除去できる。しかし、残響時間が 0.2 秒以上長い室内では、一般に伝達関数は単位円外にも零点を持ち、非最小位相となることが多い [52]。Neely ら [53] のシミュレーション実験によると通常の大きさの部屋 (W3.96m × D3.35m × H2.1m) に

において、伝達関数が最小位相になるための壁の反射係数の閾値は37%であった。実際の室内の場合は、反射係数は70%~90%程度であることから、実環境においてはほとんどインパルス応答は非最小位相であることが考えられ、安定な逆フィルタを構成することは困難となる。

この室内伝達関数の非最小位相性により逆フィルタが構成できない問題を解決した方法がMINT法である。MINT法では、室内伝達関数が非最小位相である場合でも、音源数より多い複数のマイクロホンを用いて安定な逆フィルタを計算できることを理論的に示している。MINT法は、多入力多出力理論であり複数音源にも対応しているが、ここでは単一音源の残響抑圧に問題を限定して述べる。

図2.7に示す1入力N出力の音響システムを考える。ここでNはマイクロホン数で $N > 1$ である。 $s(k)$ は音源信号、 $x_j(k) (j = 1, \dots, N)$ はj番目のマイクロホンでの受信信号、 k は離散時間を表す。さらに $y(k)$ は逆フィルタ出力であり、 $s(k)$ の回復信号を表す。 $g_j(k)$ は音源からj番目のマイクロホンまでの長さ K の室内インパルス応答、 $h_j(k)$ はj番目のマイクロホン出力に繋がる逆フィルタを構成する長さ L のFIRフィルタである。ここで簡単のため、マイクロホンは完全に平坦な理想特性を持つものとする。伝達関数 $G_j(z), H_j(z)$ をそれぞれ $g_j(k), h_j(k)$ の z 変換とすると、この1入力N出力音響システムに対するMINT逆フィルタシステムは次式のディオファントス方程式を満たす。

$$G_1(z)H_1(z) + G_2(z)H_2(z) + \dots + G_N(z)H_N(z) = 1 \quad (2.18)$$

$G_j(z)$ が非最小位相であっても共通零点を持たない限り、MINT逆フィルタ $H_j(z)$ は存在することが、Miyoshiらによって示された[44]。

具体的な逆フィルタ係数計算のために式(2.18)をフィルタ係数の行列表現する。つまり、室内インパルス応答行列 \mathbf{G} と逆フィルタ \mathbf{H} の行列掛け算をした結果が、音源から逆フィルタ出力までのインパルス応答を表し、それがデルタ関

数を表す目標ベクトル \mathbf{B} となるように逆フィルタ \mathbf{H} は計算される.

$$\mathbf{B} = \mathbf{GH} \quad (2.19)$$

ここで, \mathbf{B} は NL 行の列ベクトルであり,

$$\mathbf{B} = \begin{bmatrix} 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

\mathbf{G} は $NL \times NL$ の行列であり,

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_1 & \mathbf{G}_2 & \cdots & \mathbf{G}_N \end{bmatrix},$$

$$\mathbf{G}_j = \begin{bmatrix} g_j(0) & 0 & \cdots & 0 \\ g_j(1) & g_j(0) & \cdots & 0 \\ \vdots & g_j(1) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ g_j(K-1) & \cdots & \cdots & g_j(K-L) \\ 0 & g_j(K-1) & \cdots & g_j(K-L+1) \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & g_j(K-1) \end{bmatrix}$$

\mathbf{H} は NL 行の列ベクトルであり,

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \vdots \\ \mathbf{h}_j \\ \vdots \\ \mathbf{h}_N \end{bmatrix}, \mathbf{h}_j = \begin{bmatrix} h_j(0) \\ h_j(1) \\ \vdots \\ h_j(L-1) \end{bmatrix}$$

である.

逆フィルタ長 L が次式を満たすとき, 行列 \mathbf{G} は正方行列となる.

$$L = \left\lceil \frac{K-1}{N-1} \right\rceil \quad (2.20)$$

ここで, $\lceil \cdot \rceil$ は引数より大きな最小の整数を表す切り上げ記号である. 室内インパルス応答の伝達関数の零点が重なっておらず, 尚且つ式 (2.20) を満たすとすると, \mathbf{G} は正則行列であり, 逆フィルタ \mathbf{H} は次式で計算される.

$$\mathbf{H} = \mathbf{G}^{-1}\mathbf{B} \quad (2.21)$$

つまり, 室内インパルス応答行列 \mathbf{G} が事前に分かっているならば, その逆行列を計算することにより逆フィルタ \mathbf{H} を求めることが可能である.

MINT 法の問題点は, 事前にインパルス応答を測定しなければ適用できない点である. 実際の家庭での通信や日常の会社での通信会議では事前に室内インパルス応答の測定は困難である.

2.6.2 遅延和アレー法

残響や雑音の存在する環境下で目的信号を SN 比高く收音するための方法として, 目的信号方向に鋭い指向性を向けそれ以外の方向からの音は收音しないよ

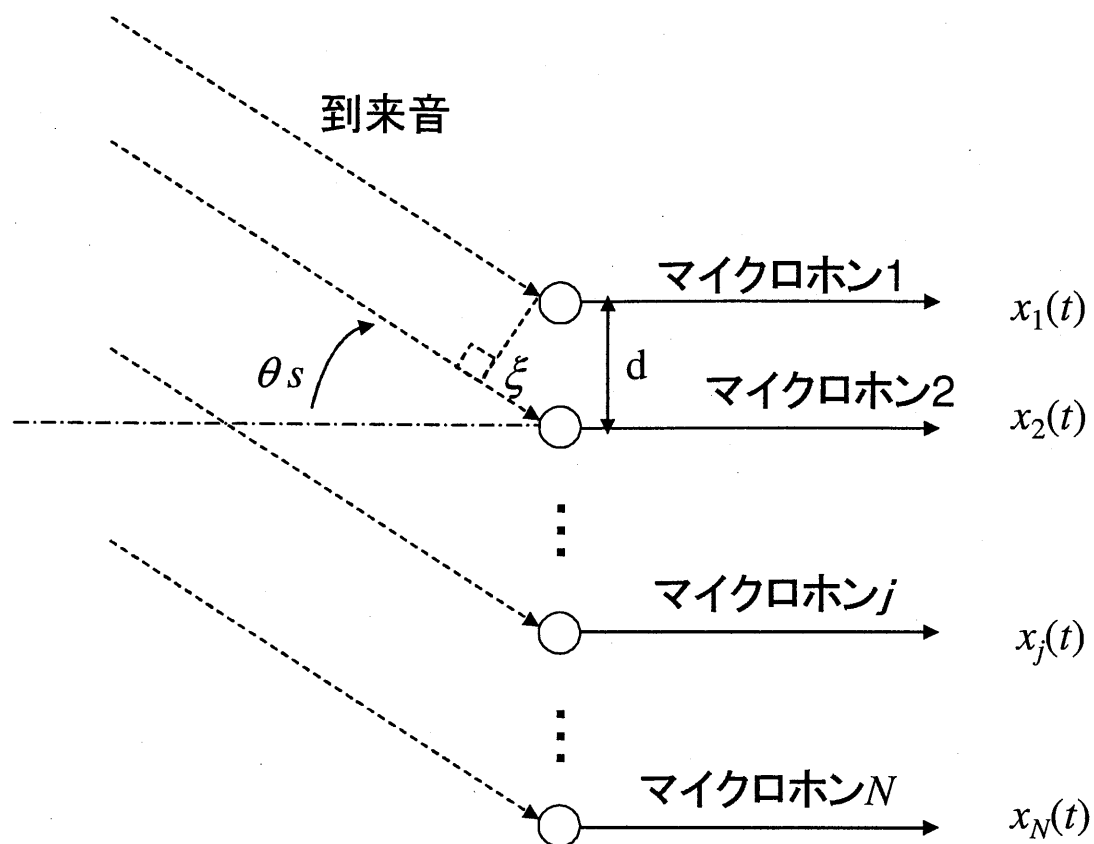


図 2.8: 等間隔配列のマイクロホンアレーによる受信

うにする超指向性マイクロホンがある。これは、通常は無指向性や単一指向性マイクロホンなどでの受信に比べ、大きなSN比の改善が期待できる。従来、超指向性マイクロホンは、パラボラ受信器やラインマイクロホンなどが知られているが、目的方向を変化させたい場合には、物理的にマイクロホンを移動させる必要があった。遅延和アレー法では、機器を固定したまま電氣的に指向性の向きを制御することが可能となり、使いやすさの点で優れている。

図 2.8 は、遅延和アレー法の原理を説明する図である。空間的に配置された複数のマイクロホンで音を受信すると各信号の間には時間差や振幅差が生じる。ここでは簡単のため自由空間に平面波が到達している状況を仮定する。

マイクロホンは間隔 d で等間隔に配置されているとすると、 θ_s 方向から到来

した音は、マイクロホン1で受信信号 $x_1(t)$ となり、次に距離 ξ 進んでマイクロホン2で受信信号 $x_2(t)$ となる。この距離 ξ は図 2.8 の幾何学的関係から次式で表せる。

$$\xi = d \sin \theta_s \quad (2.22)$$

したがって、マイクロホン2での受信信号 $x_2(t)$ はマイクロホン1での受信信号 $x_1(t)$ に比べ、音が距離 ξ だけ進むのに要する時間 $\tau_s = \xi/c = (d \sin \theta_s)/c$ だけ遅れた信号となる。

$$x_2(t) = x_1(t - \tau_s) \quad (2.23)$$

ここで、 c は音速である。同様に、 j 番目のマイクロホンでの受信信号 $x_j(t)$, ($j = 1, 2, \dots, N$) は時間 $(j-1)\tau_s$ 遅れた信号、

$$x_j(t) = x_1(t - (j-1)\tau_s) \quad (2.24)$$

と表せる。

遅延和アレー法では、図 2.9 に示すように各マイクロホンの後段に遅延器を接続し、その各遅延器の出力信号の和を出力 $y(t)$ する。

$$y(t) = \sum_{j=1}^N x_j(t - D_j) \quad (2.25)$$

ここで、遅延量 D_j は次式で与えられる。

$$D_j = D_0 - (j-1)\tau_L \quad (j = 1, 2, \dots, N) \quad (2.26)$$

$$\tau_L = (d \sin \theta_L)/c \quad (2.27)$$

但し、 θ_L は指向性を向ける目的方向を表し、 D_0 は固定遅延量である。

目的信号が、指向性を向けている目的方向 θ_L から到来したと仮定すると、各マイクロホンで受信される目的信号 $x_{sj}(t)$ は式 (2.24) で示されるように、

$$x_{sj}(t) = x_{s1}(t - (j-1)\tau_L) \quad (2.28)$$

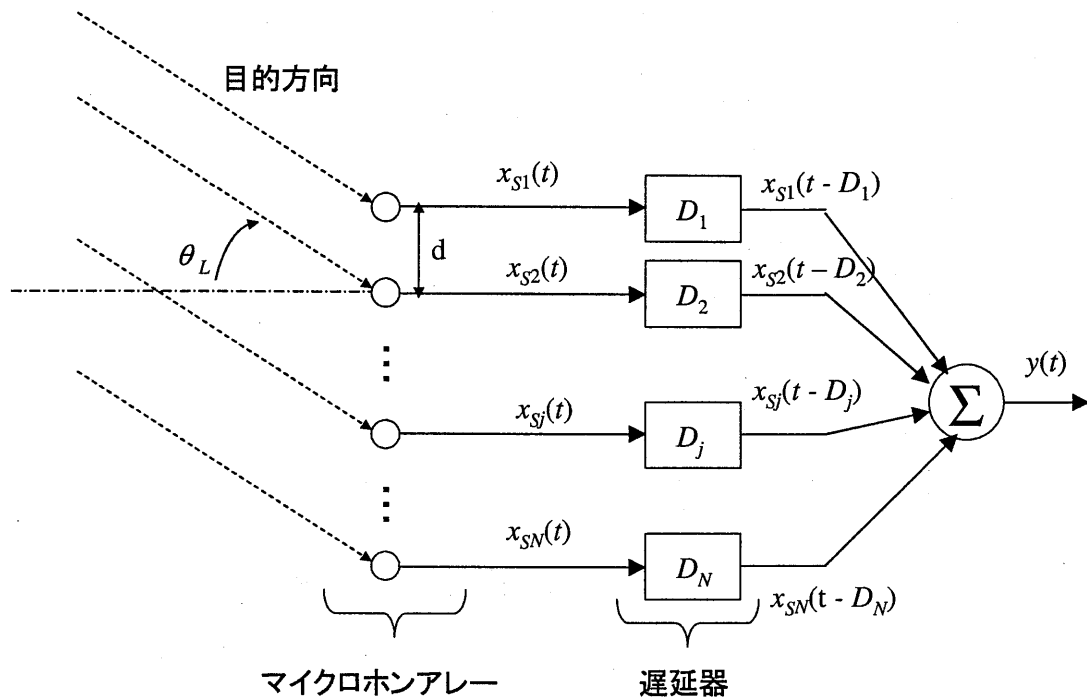


図 2.9: 遅延和アレー法の構成図

となる. 式 (2.28) を代入すると,

$$\begin{aligned}
 y(t) &= \sum_{j=1}^N x_{sj}(t - D_j) \\
 &= \sum_{j=1}^N x_{s1}(t - (j-1)\tau_L - D_j) \\
 &= \sum_{j=1}^N x_{s1}(t - D_0) \\
 &= Nx_{s1}(t - D_0)
 \end{aligned} \tag{2.29}$$

となる. この式からも分かるように, 各遅延器からの出力 $x_{sj}(t - D_j)$ はチャンネルによらず同一信号 $x_{s1}(t - D_0)$ となっていることが分かる. つまり, 遅延器により θ_L 方向から到来した信号の各マイクロホン間の遅延差が補正され, 同相化されている. この同相化された信号を加算することにより, 目的方向 θ_L からの到来音は同期加算の原理で強調される. 一方, θ_L と異なる方向から到来する音

は、 τ_L と異なる時間差で受音されるため、各遅延器によっては同相化されず、加算された信号は時間的にずれた波形となり強調されない。したがって、目的音の到来方向に目的方向 θ_L を設定することにより、目的音の方向に指向性が向けられ強調され、他の方向から到来する残響音や雑音は抑圧される。

遅延和アレーの問題点としては、(i) 空間的に方向性の強い音に対しては効果が高いが、拡散性の高い音の場合は効果が低くなる点、(ii) SN比を上げるためにはマイクロホンの数が多数必要になり、アレーサイズが大きくなる点、(iii) 残響音のように目的音と相関のある妨害音に対しては効果が低い点がある。