

遅延和アレーに基づく音源方向推定の研究

指導教授 金田 豊 教授

98kc132 森山 佳奈

目次

1 . はじめに	...2
2 . 時間差に基づく方向推定	...3
2 - 1 . 音の到来方向を推定する	...3
2 - 2 . 正弦波を用いた音源方向検出シミュレーション	...5
2 - 3 . シミュレーション結果	...8
2 - 4 . 雑音を用いた音源方向検出シミュレーション	...11
2 - 5 . シミュレーション結果	...13
3 . 遅延和アレー(超指向性マイクロホンアレー)に基づく方向推定	...14
3 - 1 . 音の到来方向を推定する	...14
3 - 2 . 使用データの説明	...16
3 - 3 . 二つのマイクロホンで受音したRWD Bデータの 信号の時間差検出	...21
3 - 4 . RWD Bデータの評価	...26
3 - 5 . RWD Bデータに対する補正	...34
3 - 6 . 遅延和アレーの構成と、音源方向推定の評価 (インパルス応答を使用した推定)	...37
3 - 6 - 1 . 遅延和アレープログラムで使用した マイクロホン(#17,21,25,29)のインパルス応答 の波形について調べる。	...48
3 - 6 - 2 . 音源 - マイクロホンの距離の関係	...49
3 - 7 . 遅延和アレーの構成と、音源方向推定の評価 (インパルス応答 + 音声信号による方向推定)	...51
4 . まとめ	...61
参考文献	...62
謝辞	...62

1 . はじめに

音源方向の推定は、音の位置情報を用いた音源分離等の他、テレビ会議システム等での話者方向の特定、移動ロボット等の環境認識システム、監視システム、騒音の分析など幅広い分野で求められ、研究が行なわれている。

従来の技術には、相関関数、遅延和アレー、高分解能法などがあり、それらの問題点として、目的音の他に、周囲の雑音や室内の反射音などが存在する実環境下では、この時間差の検出が困難となるケースが生じ、推定精度が劣化するという問題点があり、実用化に向け耐雑音・耐残響性能の向上が必要とされている。さらに、研究上のテクニカルな課題として、共通なデータベースが存在しないため、各研究者がいろいろな手法を別個の室内で検討しているため、手法の優劣の比較がしにくいことが問題となっている。このことを解決するためには、共通のデータを使用し、手法の効果を比較できるようなデータベースの構築が必要であると考え。

の問題解決のため構築されたものとしてRWC P実環境音声・音響データベース マイクロホンアレーデータベース(以下、RWDBとする)があるが、このデータベースは構築されたばかりであり、データの評価が十分に行われていないという問題がある。

本研究の第一の目的として、このRWDBの実用性評価を行うことである。

第二の目的としては、このRWDBを基にして、遅延和アレー法による音源方向推定法の実用性評価を行うことである。

実用性評価とは

-) 残響のある室内における推定性能評価
-) 音声信号を使用した推定性能評価

である。

2 . 時間差に基づく方向推定

2 - 1 . 音の到来方向を推定する。

s 方向から到来する平面波を距離 d だけ離れて設置された二つのマイクロホンで受音することを考える。この様子を図 1 に示す。

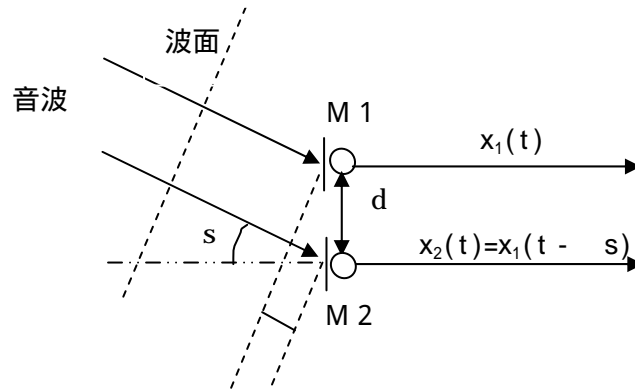


図 1 . 音の到来方向推定モデル

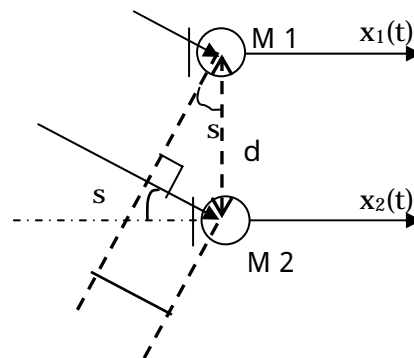


図 2 . マイクロホンの部分の拡大図

図 1 の s から到来した音波は、まず M1 において受音される。次に、音波は図 1 に示した距離 d だけ進んで M2 に到達する。図 2 より、この距離 d は、

$$d = d \sin s \quad \dots (1)$$

と表される。したがって、M2 での受音信号 $x_2(t)$ は M1 での受音信号 $x_1(t)$ と比べて音波が距離 d だけ進行するのに要する時間 s だけ遅れた信号となっている。即ち

$$x_2(t) = x_1(t - s) \quad \dots (2)$$

$$s = d \sin s / c = (d \sin s) / c \quad \dots (3) \quad c : \text{音速}$$

の関係が成立する。したがって逆に、信号 $x_1(t)$ と $x_2(t)$ の間の時間差 s がわかれば、音波の到来方向 s は次式から求められる。

$$s = \sin^{-1} (c \cdot s / d) \quad \dots (4)$$

時間差 s は、 $x_1(t)$ 、 $x_2(t)$ との相互相関関数 $r_{12}(\tau)$ から求めることができる。

$$\begin{aligned} r_{12}(\tau) &= 1/N \sum \{x_1(t) \cdot x_2(t + \tau)\} \\ &= 1/N \sum \{x_1(t) \cdot x_1(t + \tau - s)\} \quad \dots (5) \end{aligned}$$

$r_{12}(\tau)$ の最大値を与える τ を求めれば s が得られ ($\tau = s$ で $r_{12}(\tau)$ は最大値をとる。
(* 1)) 式(4)に代入して、音波の到来方向 s が推定できる。

* 1)

はじめに、自己相関関数と相互相関関数について述べる。

自己相関関数とは一つの信号波形に対して相関処理を行い、一つの信号波形中に繰り返し(周期性)があるかどうかを調べることである。

信号波形 $x_1(t)$ の自己相関関数は

$$r_{11}(\tau) = 1/N \sum \{x_1(t) \cdot x_1(t+\tau)\} \dots (6)$$

で表される。

この式は、N個のサンプリング値からなる一つの信号波形 $x_1(t)$ と、それを遅れ時間 τ だけずらした信号波形 $x_1(t+\tau)$ に対して、それぞれを掛け合わせて累積し、平均化するという処理を行っている。この場合、遅れ時間がない($\tau = 0$)ときには、 $x_1(t)$ と $x_1(t+\tau)$ とは同じ波形になるのでぴったりと重なり、自己相関関数 $r_{11}(\tau)$ は最大値となる。

相互相関関数とは異なる二つの信号波形に対して相関処理を行い、二つの類似度を調べることである。

信号波形 $x_1(t)$ と信号波形 $x_2(t)$ の相互相関関数は

$$r_{12}(\tau) = 1/N \sum \{x_1(t) \cdot x_2(t+\tau)\} \dots (7)$$

で表される。基本的には、自己相関処理と同じように信号波形 $x_1(t)$ と遅れ時間 τ だけずらした信号波形 $x_2(t+\tau)$ において、それぞれ対応する部分を掛け合わせて累積し、平均化する処理を行う。

今回の時間差に基づく方向推定では一つの音源をマイクロホン M1、M2 で受音するものなので、上に記したように、M2 での受音信号 $x_2(t)$ は M1 での受音信号 $x_1(t)$ と比べて音波が距離 s だけ進行するのに要する時間 s/c だけ遅れた信号となっている。よって、式(7)は式(5)となり、さらに自己相関関数 $r_{11}(\tau)$ で表すと

$$\begin{aligned} r_{12}(\tau) &= 1/N \sum \{x_1(t) \cdot x_1(t+\tau - s/c)\} \\ &= r_{11}(\tau - s/c) \dots (8) \end{aligned}$$

となる。式(8)からも分かるように、相互相関関数 $r_{12}(\tau)$ は自己相関関数 $r_{11}(\tau)$ を s/c だけずらしたものであることが分かる。さらに、自己相関関数 $r_{11}(\tau)$ では $\tau = 0$ で最大値をとることが分かっている。したがって、相互相関関数 $r_{12}(\tau)$ では、 $\tau = 0$ より s/c だけずらした $\tau = s/c$ で最大値をとる。

2 - 2 . 正弦波を用いた音源方向検出シミュレーション

正弦波の到来・収音・方向検出のシミュレーションを以下のように行った。

< 方法と手順 >

) はじめに、パラメータの設定をする。

パラメータ $d = 0.17$ [m]

$c = 340$ [m/s]

$F_s = 48000$ [Hz]

$s = 50$ [°] 単位を[ラジアン]に変換する。

$\theta = 50 * \pi / 180$ [ラジアン]

) 正弦波の設定をする。(* 2)

) 信号の時間差の最大値を決める。(* 3)

) $s = 50$ [°]の時の信号の時間差 τ を求め、y 式をその時間差分ずらす。

式(3)より、信号の時間差 s は

$$s = (d \sin \theta) / c$$

$$= 0.17 * \sin 50 / 340$$

$$= 0.000383022$$

s はアナログ表示なのでデジタル表示の τ にするためには標本化周期

$T_s (= 1/F_s)$ で割る必要がある。よって

$$\tau = s * F_s$$

$$= 0.000383022 * 48000$$

$$= 18.385066$$

) 相互相関関数の計算、表示。

) 表示されたグラフの最大値を示す x 軸の値(時間差)を求める。その値(時間差)から角度を計算する。

以上のことを行い、はじめに設定した角度 $s = 50$ [°]と、) で求められる角度が同じになるかを確認する。

以下に、プログラムを示す。

```

% seigenha1

% 正弦波の音源方向探査 ~時間差に基づく方向推定~
% 2001.9.28 moriyama

% パラメータの設定
Fs=48000;          %サンプリング周波数
x=(1:32000)';     %時間(単位:サンプル)
d=0.17;           %マイクロホン間の距離
c=340;            %音速
thita=50*pi/180; % s:50°と仮定。それを単位[ラジアン]に変換。

% 正弦波の設定
y=0.5*sin(2*pi*(x-1)/50); %正弦波信号, 1周期Tは50サンプルなので、
                          %T=50*(1/Fs)=16/48000[s](*2)

tmax=round(d/c*Fs) %信号の時間差の最大値(*3)

Ccor=zeros(2*tmax,1);

xt=y(1001:31000); %xtはy式の1001番目~31000番目を切り取った式。
disp(xt(1:10)); %y(1001)~y(1010)(xt(1)~xt(10))を表示。

taus=round(d*(sin(thita))/c*Fs) %実際の信号の時間差

yt=y(1001+taus:31000+taus); %ytはy式の1001+taus番目~31000+taus番目を
                             切り取った式。
disp(yt(1:10)); %y(1001+taus)~y(1010+taus)(yt(1)~yt(10))を表示。

% 相互相関関数の計算
for tau=-tmax:tmax
    Ccor(tau+tmax+1)=xt(1+tau+tmax:10000+tau+tmax)'*yt(1+tmax:10000+tmax)/10000;
end

```

```
xx=-tmax:tmax;  
plot(xx,Ccor);
```

```
% x 軸、 y 軸のラベル表示  
xlabel('時間 [サンプル]')  
ylabel('相互相関関数')
```

```
[yjiku xjiku]=max(Ccor); %表示されたグラフの最大値を示す x 軸の値を xjiku,  
                        y 軸の値を yjiku とする。
```

```
xjiku=xjiku-tmax-1 % x 軸の値 0 が tmax + 1 番目であるため。
```

```
hoko=asin(c*xjiku/Fs/d); %時間差から角度[ラジアン]を求める。  
kaku=hoko/pi*180 % 音の方向を単位[ラジアン]から[°]に変換。
```


2 - 3 . シミュレーション結果

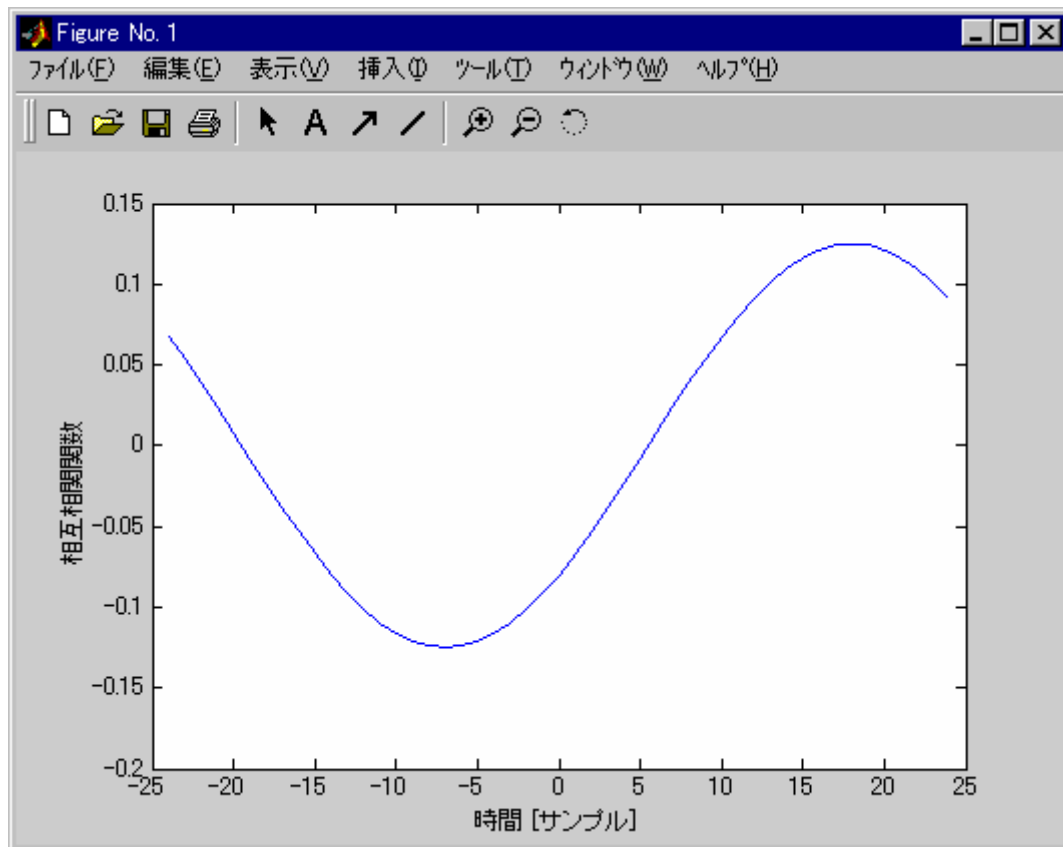


図 3 . 正弦波の相互相関関数

～コマンド ウィンドウの表示～

tmax =24 音波が 90° で到来したときの時間差。
 taus =18 相互相関関数が最大になるときの X 軸の値。(計算結果より)
 xjiku =18 相互相関関数が最大になるときの X 軸の値。(図 3 より)
 kaku = 48.6

以上より、はじめに設定した $s = 50^\circ$ とほぼ等しい値となったためプログラムは正しいといえる。

* 2)

正弦波の周波数 f の決め方

マイクロホン間に、正弦波の最大値が一つになるようにすればよい。

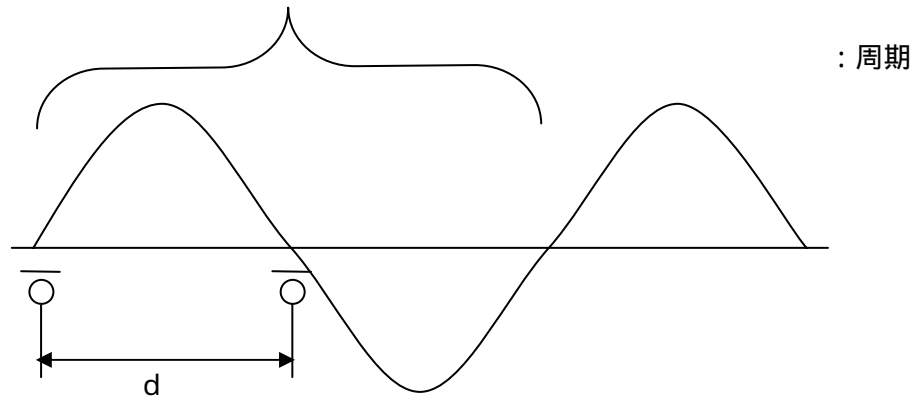


図4 . 正弦波のモデル

図4のように、マイクロホン間 d を $\lambda / 2$ より狭くすればよい。

$$c / f = \lambda$$

$$c / (2 * f) = \lambda / 2 > d$$

よって

$$f < c / (2 * d) \quad \text{とすればよい。}$$

上記のプログラムの例を取ると

パラメータ

$$d = 0.17 \text{ [m]}$$

$$c = 340 \text{ [m/s]}$$

$$F_s = 48000 \text{ [Hz]}$$

であるので、

$$f < 340 / (2 * 0.17)$$

$$f < 1000$$

f は単位が [Hz] なので、デジタル表示の F [サンプル] にすると、

$$F_s / F < 1000$$

$$F > 48000 / 1000 = 48$$

よって、正弦波の周波数は 48 [サンプル] より、大きくなる必要がある。

上記のプログラムでは周波数を 50 [サンプル] に設定してある。

* 3)

tau の範囲の決め方。

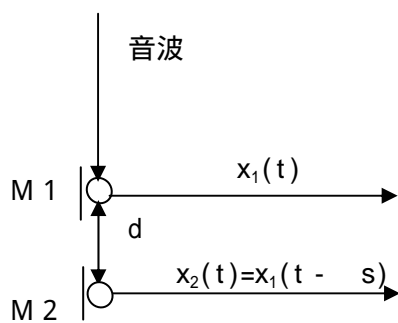


図5 . 音の到来方向推定モデル
(到来方向直角の場合)

M 1 での受信信号 $x_1(t)$ と M 2 での受信信号 $x_2(t)$ の間の時間差 s の最大値 (s_{\max}) は音波が図 4 のように、直角に到来してきた場合である。

$$s_{\max} = d / c \quad (c : \text{音波}) \dots (9)$$

式(9)は、単位が[秒]であるので、[サンプル]にするには、サンプリング周波数 F_s をかければよい。式(5)の相互相関関数 $r_{12}(\tau)$ は、 s で最大値となることがわかっているので、 $r_{12}(\tau)$ が最大になる値は、 s_{\max} より大きくなることはない。よって、 τ に s_{\max} を含むような範囲をとると、 $r_{12}(\tau)$ が最大になる値を含むため、正常な結果を得ることができる。

2 - 4 . 雑音を用いた音源方向検出シミュレーション

雑音の到来・収音・方向検出のシミュレーションを以下のように行った。

<方法と手順>

) はじめに、パラメータの設定をする。

パラメータ $d = 3.4$ [m]

$c = 340$ [m/s]

$F_s = 8000$ [Hz]

$s = 40$ [°] 正弦波と同様に単位を[ラジアン]に変換する。

$y = \text{randn}(32000, 1)$ 雑音信号

) 信号の時間差の最大値を求める。

) $s = 40$ [°]の時の信号の時間差 τ_s を求め、 y 式をその時間差分ずらす。

式(3)より、信号の時間差 s は

$$\begin{aligned} s &= (d \sin s) / c \\ &= 3.4 * \sin 40 / 340 \\ &= 0.00642787 \end{aligned}$$

s はアナログ表示なのでデジタル表示の τ_s にするためには標本化周期

$T_s (= 1 / F_s)$ で割る必要がある。よって

$$\begin{aligned} \tau_s &= s * F_s \quad \dots(10) \\ &= 0.00642784 * 8000 \\ &= 51.42272 \end{aligned}$$

) 相互相関関数の計算、表示。

) 表示されたグラフの最大値を示す x 軸の値(時間差)を求める。その値(時間差)から角度を計算する。

以上のことを行い、はじめに設定した角度 $s = 40$ [°]と、) で求められる角度が同じになるかを確認する。

以下にプログラムを示す。

```

% zatuon

% 雑音の音源方向探査 ~時間差に基づく方向推定~
% 2001.9.28 moriyama

% パラメータ設定
Fs=8000;
y=randn(32000,1); %雑音信号
c=340;          %音速
d=3.4;         %マイクロホン間の距離
thi ta=40*pi/180; % s : 40° と仮定。それを単位[ラジアン]に変換。

taus=round(d*(sin(thi ta))/c*Fs)
tmax=round(d/c*Fs)

Ccor=zeros(tmax,1);

xt=y(1001:31000);
yt=y(1001+taus:31000+taus);

% 相互相関関数の計算
for tau=-tmax:tmax
    Ccor(tau+tmax+1)=xt(1+tau+tmax:10000+tau+tmax)'*yt(1+tmax:10000+tmax)/10000;
end

xx=-tmax:tmax;
plot(xx,Ccor);

% x軸、y軸のラベル表示
xlabel('時間[サンプル]')
ylabel('相互相関関数')

[yjiku xjiku]=max(Ccor);
xjiku=xjiku-tmax-1
hoko=asin(c*xjiku/Fs/d);
kaku=hoko/pi*180          %角度の単位[ラジアン] [°]

```

2 - 5 . シミュレーション結果

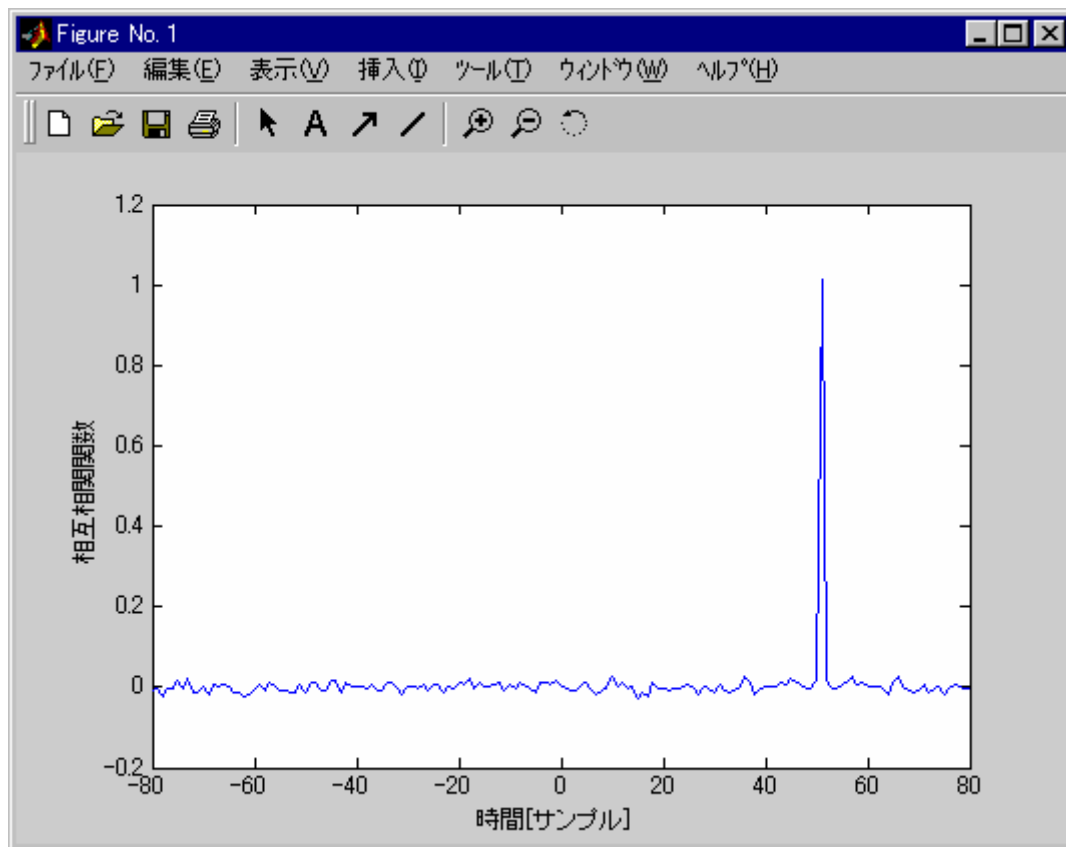


図 6 . 雑音の相互相関関数

～コマンド ウィンドウの表示～

tmax = 80

taus = 51

xjiku = 51

kaku = 39.6057

以上より、はじめに設定した $s = 40^\circ$ とほぼ等しい値となったためプログラムは正しいといえる。

3 . 遅延和アレー(超指向性マイクロホンアレー)に基づく方向推定

3 - 1 . 音の到来方向を推定する。

<原理>

条件：一般性を損なわない限り、等間隔 d で配置された直線配列マイクロホンアレーを受音器とする。

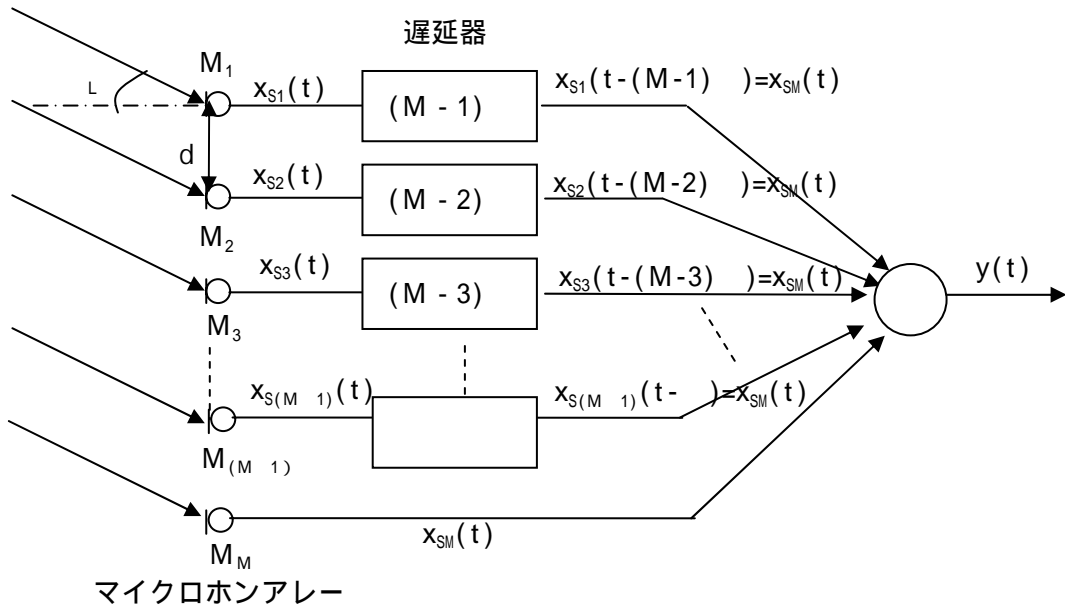


図7 . 遅延和アレー処理

図7は遅延和アレーの原理を表したものである。音源が θ の方向から到来しているとき、まず、マイクロホン M_1 で受音され、次に時間だけ遅れてマイクロホン M_2 に受音される。マイクロホン M_1 と M_2 で受音される信号の間には、式(11)の遅れが発生する。

$$= (d \sin \theta) / c \dots (11)$$

等間隔 d でマイクロホンが配置されているため、マイクロホン M_1 で受音される信号とマイクロホン $M_2 \sim M_M$ で受音される信号の間には $\sim (M - 1)$ の遅れが発生する。 M_M の受音信号の波形と同じにするため(同相化)、図7のように、遅延器でマイクロホン $M_1 \sim M_{(M-1)}$ にそれぞれ $(M - 1) \sim$ の遅れを付加すると、 θ 方向から到来する信号の時間差が補正され、同相化される。この同相化した信号を加算することで、目的方向 θ から到来する音は強調される。

$$y(t) = \sum_{i=1}^M x_i(t - i \cdot \tau) \dots (12) \quad (M: \text{マイクロホン数})$$

一方、 θ 以外の方向から到来する音は、 θ 方向のときの時間差 $M \tau$ とは異なり、 $M \tau$ を付加する遅延操作では信号が同相化されないため加算しても強調されることはない。その結果、図8のような指向性を実現する。

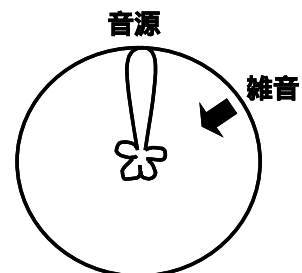


図8 . 遅延和アレーの指向特性

この実験では、 θ 方向に指向性ビームを向けたときのアレー出力パワー $P_y(\theta)$ を、 θ 方向から到来する音波のパワーの推定値とする。そして、ビーム方向 θ を変化させれば、方向ごとの到来音パワーの推定値が得られる。

パワーが最大となった角度 θ_{MAX} 方向が音源方向であると考える。

$P_y(\theta)$ は以下のように求める。

$$P_y(\theta) = \sum_k y(k, \theta)^2 \quad \dots(13)$$

3 - 2 . 使用データの説明

<使用データ>

RWC P実環境音声・音響データベース マイクロホンアレーデータベースの

cirline (円形+直線マイクロホンアレー) : 図9, 10, 11 参照

今回の実験では、直線部分のみ(マイク番号17~30)使用している。

<データの種類>

- ・インパルス応答・・・float 型、リトルエンディアン
- ・サンプリング周波数・・・48[kHz]
- ・マイクロホンアレーを中心に2mの距離にスピーカーを配置した。
- ・直線マイクロホンアレーの素子間隔・・・2.83[cm]

cirline	部屋	RT	CH	Ang	Len	Byte	Total
E2A/imp 000	残変室(パネル)A : 図12	0.30	30	9	30000	4	31.0MB
E2B/imp 000	残変室(パネル)B : 図12	1.30	30	9	125000	4	128.8MB
JR1/imp 000	畳部屋(大) : 図13	0.60	30	9	60000	4	61.9MB
JR2/imp 000	畳部屋(小) : 図14	0.47	30	7	45000	4	36.1MB
OFC/imp 000	会議室 : 図15	0.78	30	9	75000	4	77.3MB

表1、データの一覧表

RT: 室内の残響時間 , CH: マイクロホンの素子数
 Ang: 収録データの方向数 , Len: 1素子あたりのデータの長さ
 Byte: ファイルのバイトオーダー(全てリトルインディアン表記)
 Total: 全容量

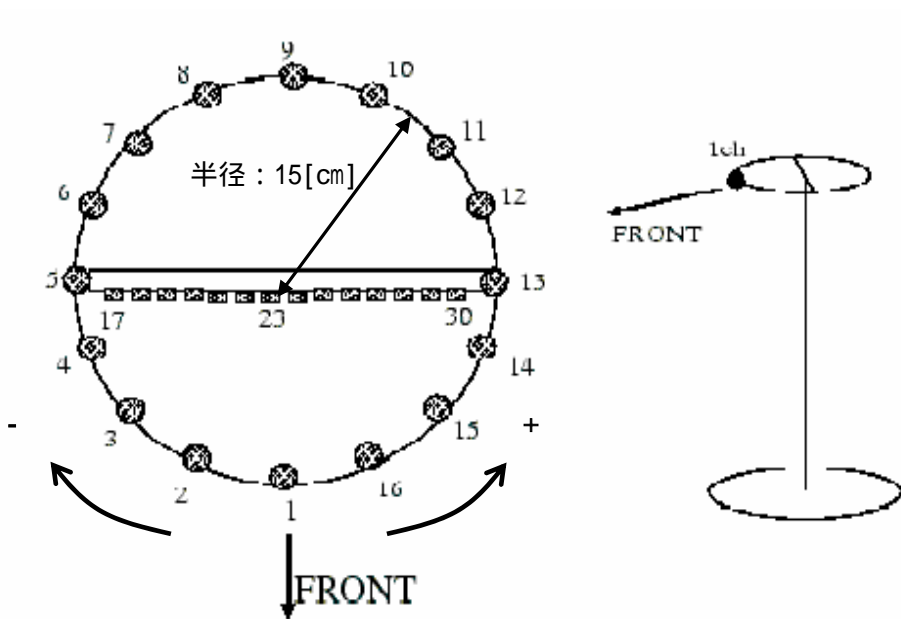


図9 . マイクロホンアレーの素子配置

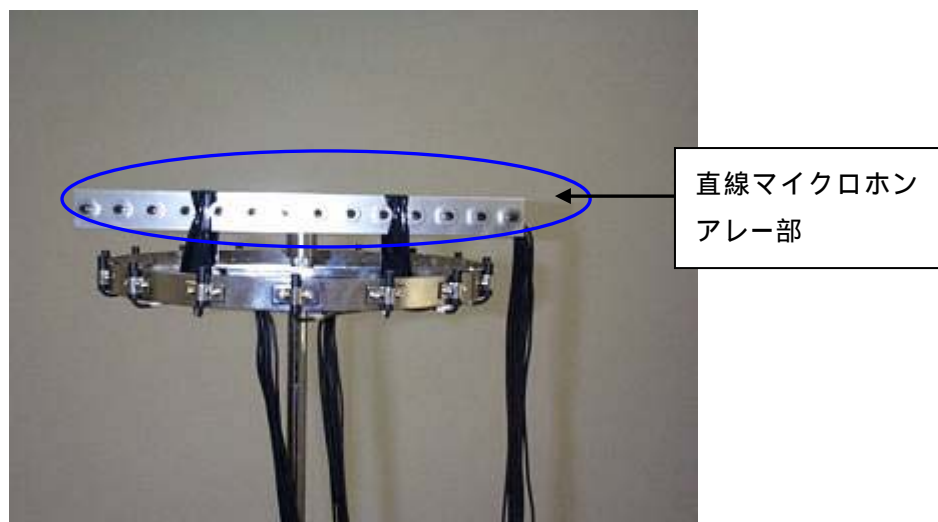


図10 . 円形+直線マイクロホンアレー

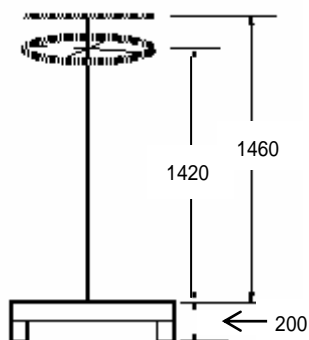


図 1 1 . 円形 + 直線マイクロホンアレーの詳細

< 音源方向(公称角度)とファイル名の関係 >

ファイル名	imp010	imp030	imp050	imp070	imp090	imp110	imp130	imp150	imp170
公称角度 [°]	-80	-60	-40	-20	0	+20	+40	+60	+80

表 2 . 公称角度とファイル名の関係

上記は、E2A,E2B の場合。

Jr1,0fc: imp090 を中心に 10[°]おきに ±40[°]

Jr2: imp090 を中心に 10[°]おきに ±30[°]

<収録風景>



図 1 2 . 残響可変室 (E2A, E2B) の収録風景



図 1 3 . 畳部屋・大 (Jr1) の収録風景



図14 . 畳部屋・小(Jr2)の収録風景



図15 . 会議室(0fc)の収録風景

3 - 3 . 二つのマイクロホンで受音したRWD Bデータの信号の時間差検出

音源 マイクロホンのインパルス応答を使用し、二つのマイクロホンで受音したインパルス応答信号の最大値の差 t_{ij} を求め、これを二つの受音信号の時間差とする。

< 方法と手順 >

) パラメータの設定

$$d = 0.0283 * 5 \text{ [m]}$$

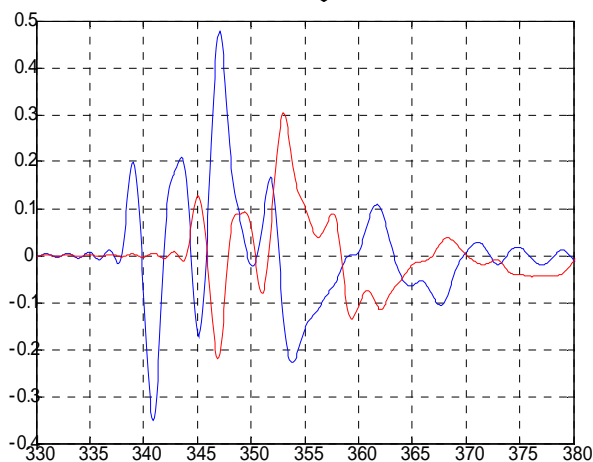
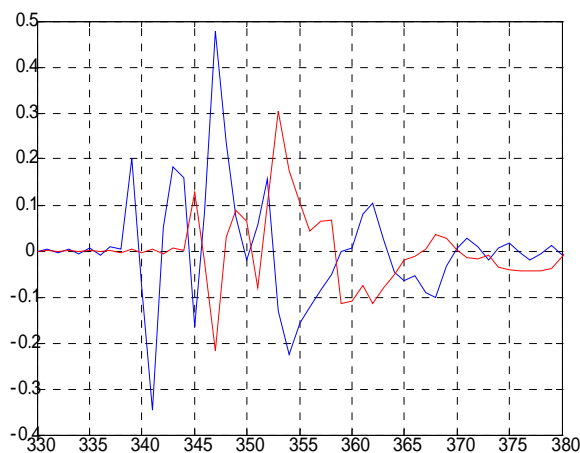
$$c = 340 \text{ [m/s]}$$

$$F_s = 48000 \text{ [Hz]}$$

) ファイルの読み込み

読み込むデータはマイクロホン素子数 $i = 21$ 、 $j = 26$ のデータとする。

) 読み込んだファイルのデータを 0.1 サンプル刻みで内挿する(細かく取る)。



-) 二つの内挿したデータの最大値を求め、その値から時間差を求める。
(時間差はコマンドウィンドウに表示される。)

以下に、二つの受信信号の時間差を求めるプログラムを示す。

< マイクロホン素子数 2 1 と 2 6 における波形から、時間差を求めるプログラム >

```
% tienwa2
```

```
% マイクロホン素子数 2 1 と 2 6 における波形から、時間差を求める。
```

```
% 2001.10.18 moriyama
```

```
%パラメータの設定
```

```
c=340
```

```
d=0.0283*5
```

```
Fs=48000
```

```
% ファイルの読み込み
```

```
fid = fopen('d:\array\E2a\imp070\imp070.21','r')
[FILENAME,PERMISSION,MACHINEFORMAT] = fopen(fid)
data1=fread(fid,'float32');
```

} マイクロホン素子数 2 1 のインパルス
応答信号の読み込み(¥E2a¥imp070)

```
fid = fopen('d:\array\E2a\imp070\imp070.26','r')
[FILENAME,PERMISSION,MACHINEFORMAT] = fopen(fid)
data2=fread(fid,'float32');
```

} マイクロホン素子数 2 6 のインパルス
応答信号の読み込み(¥E2a¥imp070)

```
fclose(fid);
```

```
cc1=interp1(data1,[300:0.1:450],'spline'); %データを内挿して値を細かく取る。
```

(300 ~ 400 を 0.1 ずつ)

```
plot([300:0.1:450],cc1,'b');
```

```
hold on
```

```
cc2=interp1(data2,[300:0.1:450],'spline');
```

```
plot([300:0.1:450],cc2,'r');
```

```
grid;
```

```
% x 軸、 y 軸のラベル表示
```

```
xlabel('時間 [サンプル]')
```

```
ylabel('振幅')
```


[yjiku1 xjiku1]=max(cc1); %cc1 のグラフから最大値を求める。

[yjiku2 xjiku2]=max(cc2); %cc2 のグラフから最大値を求める。

xjiku=(xjiku1-xjiku2)/10 %cc1 最大値を示す xjiku1 と cc2 の最大値を示す
xjiku2 を引き時間差を求める。

<例．二つの受信信号の時間差検出実行結果(E2a/imp070)>

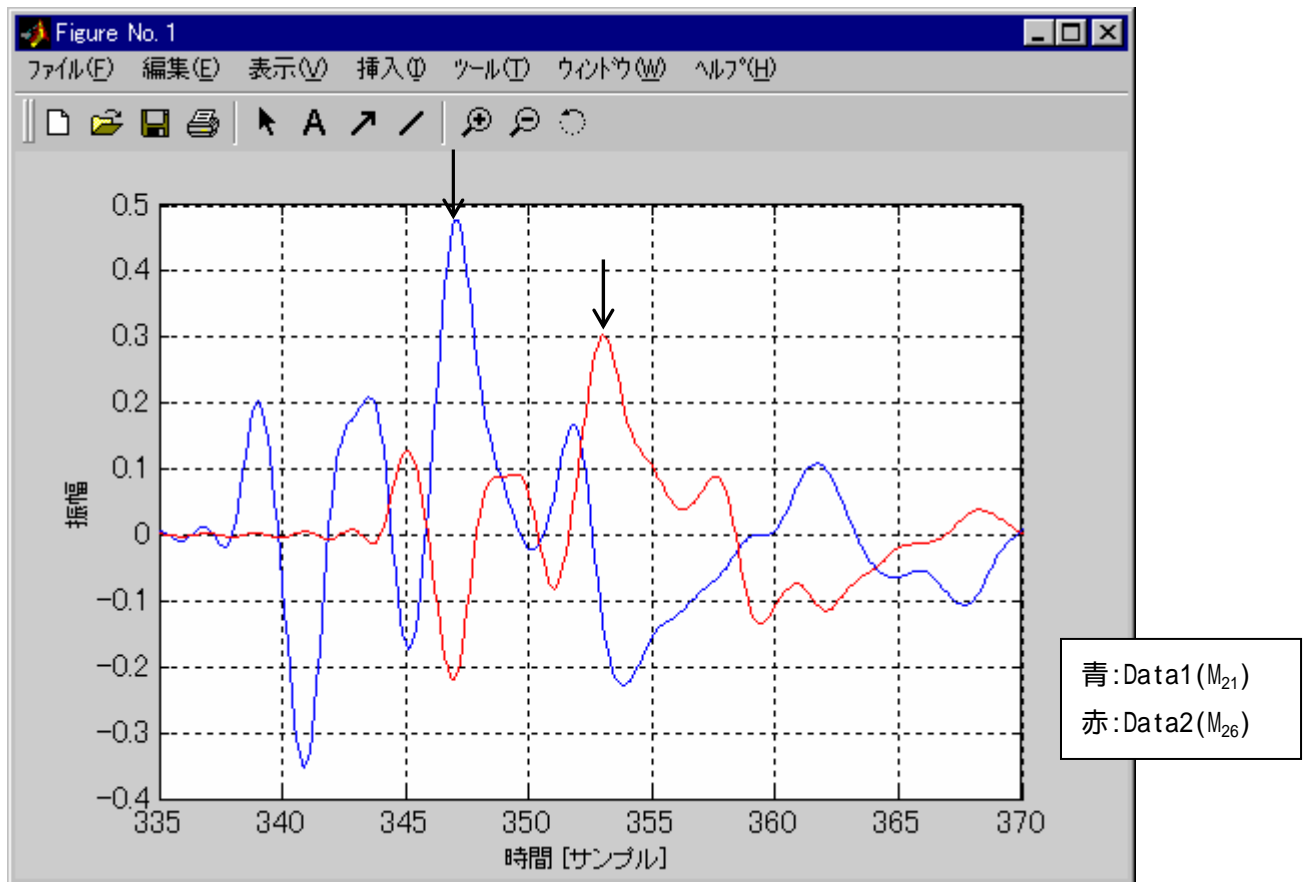


図 1 6 . Imp070 の時のマイクロホン素子数 2 1 と 2 6 における波形

矢印()で示した最大値をとる x 軸の値から時間差 t_{ij} (コマンドウィンドウに計算結果が表される)を求める。

3 - 4 . RWDBデータの評価

二つのマイクロホン(#21,26)から受音した信号より音源方向 θ を求める。

式(4)より

$$\theta = \sin^{-1}(c * \tau_{ij} / d) \dots(14)$$

上式の τ_{ij} は単位[秒]であり、それを[サンプル]で表すと、

$$\theta = \sin^{-1}(c * (\tau_{ij} / Fs) / d) \dots(15)$$

となる。

$$c=340$$

$$d=0.0283*5(=0.1415)$$

$$Fs=48000$$

3 - 3でも述べた通り、図16とコマンドウィンドウから、Data1 と Data2 の最大()
を取る値の間隔(時間差) τ は、-5.9 [サンプル]であることがわかる。式(14)より

$$\begin{aligned} \theta &= \sin^{-1}(c * (\tau / Fs) / d) \\ &= \sin^{-1}\{340 * ((-5.9) / 48000) / 0.1415\} \\ &= -17 [^\circ] \end{aligned}$$

imp070 の実際の角度は $\theta = -20 [^\circ]$ である。(imp090 は $\theta = 0$)

この時 τ を計算すると、式(10)より

$$\begin{aligned} \tau &= Fs * d * \sin(\theta) / c \\ &= 48000 * 0.1415 * \sin(-20) / 340 \\ &= -6.83 [サンプル] \end{aligned}$$

となる。

同様に、他のRWDBデータについても行い、DBの評価をした。

以下、表3～7に各部屋のデータをまとめた。

< E2A(残変室 A) >

(a) データベース の公称角度 [°]	(b) (a)と式 4 より求 められる時間差 [サンプル]	(c) データより求め た時間差 t_{ij} [サンプル]	(d) (c)と式 1 4 よ り求められる 角度 [°]
-80	-19.7	-17.0	-58.3
-60	-17.3	-14.9	-48.2
-40	-12.8	-10.9	-33.0
-20	-6.8	-5.9	-17.2
0	0	0	0
+20	6.8	6.1	17.8
+40	12.8	11.4	34.8
+60	17.3	15.1	49.1
+80	19.7	16.5	55.7

表 3 . Cirline/E2A/imp000(残変室 A)の結果

< 表の説明 >

(a) ファイル名 imp090 を公称角度 0° (図 9 の F R O N T)として 20° おきに ± 80° の範囲で音源を移動させている。(E2A, E2B の場合。 Jr1, 0fc: 10° おきに ± 40° 、 Jr2: 10° おきに ± 30°)

(b) 公称角度から式(4) ($\theta = (d \sin \theta_L) / c$) を用いて求めた時間差

(c) 二つのマイクロホン(#21, 26)で受音したインパルス応答信号の時間差 t_{ij} の検出値

(d) 時間差 t_{ij} と式(15) ($\theta_L = \sin^{-1}(c * (t_{ij} / Fs) / d)$) から求めた角度

・表 3 (b), (c) と表 3 (a), (d) の値が同じような値になれば、 R W D B データは完成度の高いものといえる。

次に、表 3 (b)、 (c) についてグラフにする。

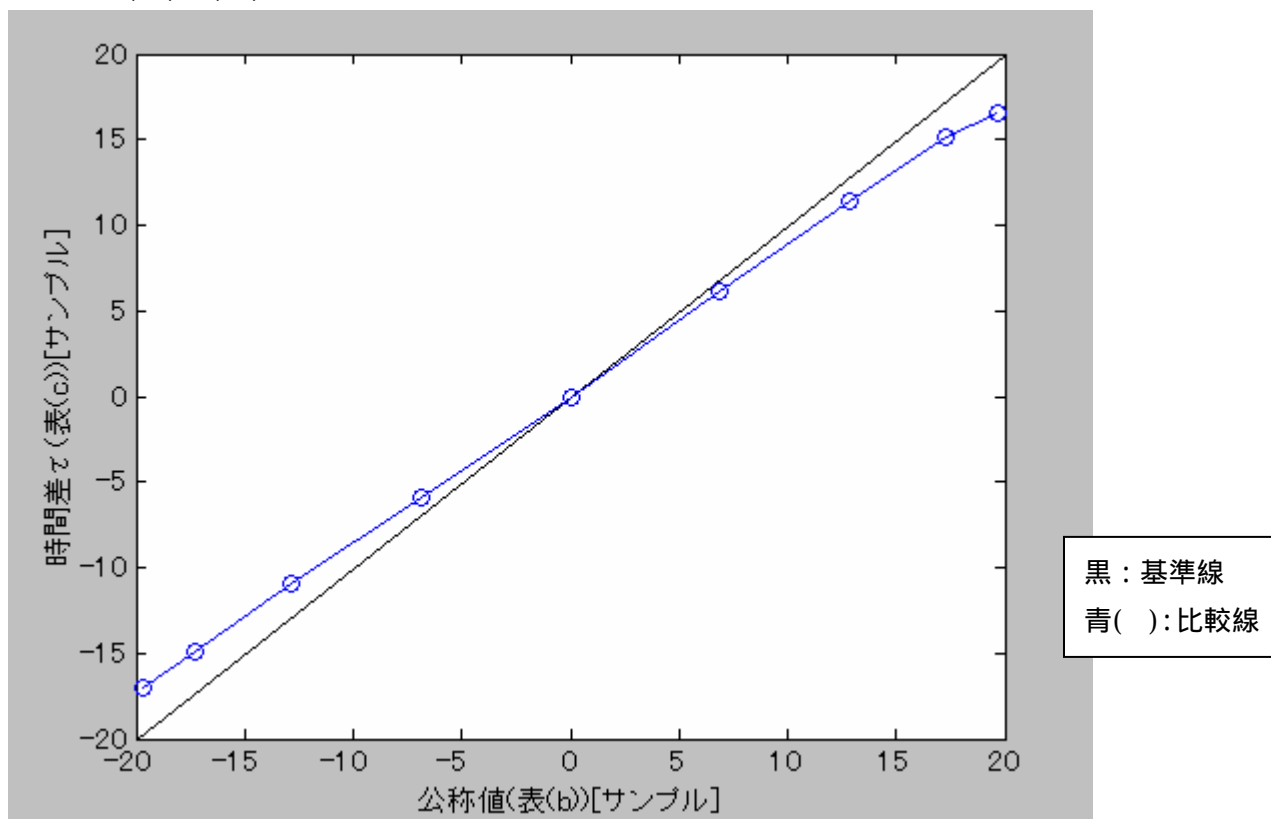


図 1 7 . 時間差の比較 (E2A)

< 図の説明 >

x 軸：表 3 (a) の公称角度から求められた時間差 (公称値) (表 3 (b))

y 軸：マイクロホン (#21, 26) で受信したインパルス応答信号の時間差 t_{ij} (表 3 (c))

黒い実線 (マーカなし) : 基準線

青い実線 (マーカ :) : 公称値と時間差 t_{ij} の交点 (印) を結んだ線 (比較線とする) 。

・ 公称値と時間差 t_{ij} の値が等しい場合に、比較線が基準線と一致する。

よって、比較線が基準線に重なると、R W D B データは完成度の高いものといえる。

< E2B(残変室 B) >

(a) データベース の公称角度 [°]	(b) (a)と式 4 より求 められる時間差 [サンプル]	(c) データより求め た時間差 t_{ij} [サンプル]	(d) (c)と式 1 4 よ り求められる 角度 [°]
-80	-19.7	-16.8	-57.2
-60	-17.3	-14.9	-48.2
-40	-12.8	-10.9	-33.1
-20	-6.8	-5.4	-15.7
0	0	0.3	0.9
+20	6.8	6.5	19.0
+40	12.8	11.3	34.4
+60	17.3	15.1	49.1
+80	19.7	16.3	54.7

表 4 . Cirline/E2B/imp000(残変室 B)の結果

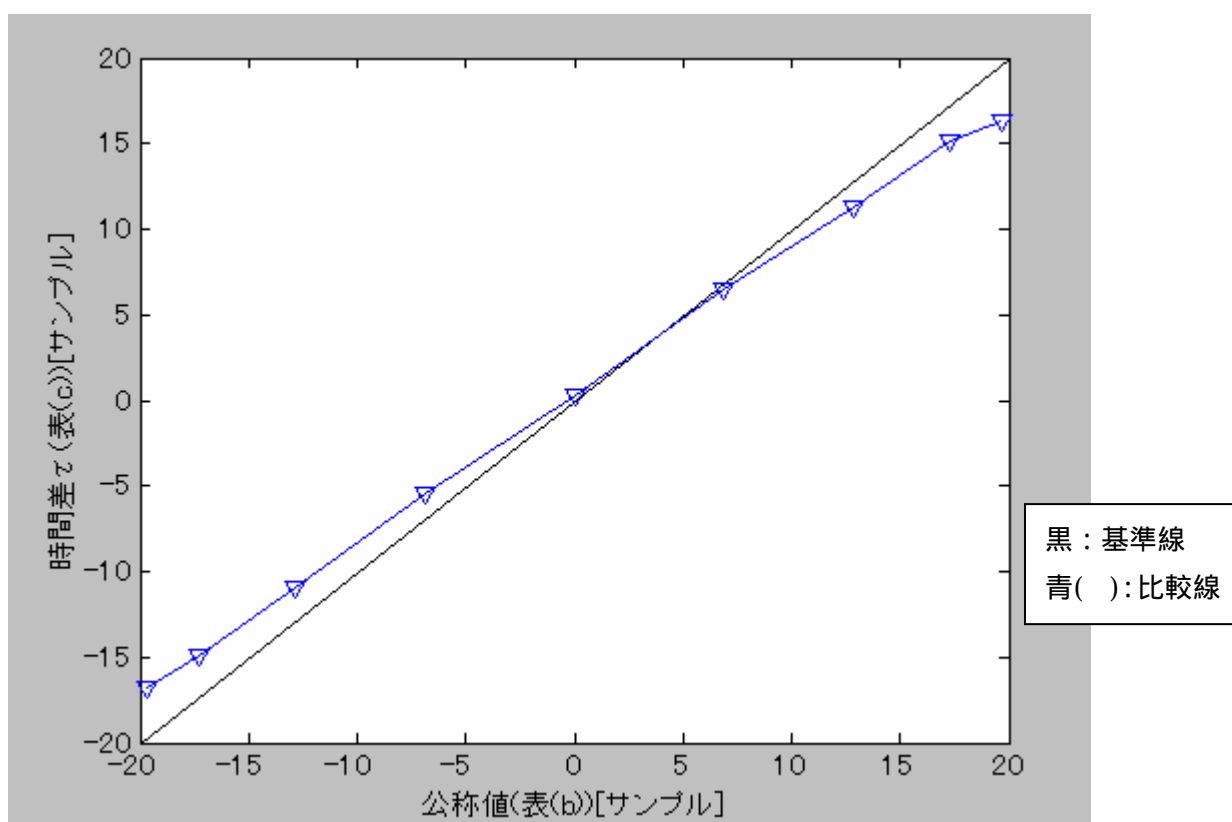


図 1 8 . 時間差の比較(E2B)

< Jr1(畳部屋・大) >

(a) データベース の公称角度 [°]	(b) (a)と式 4 より求 められる時間差 [サンプル]	(c) データより求め た時間差 t_{ij} [サンプル]	(d) (c)と式 1 4 よ り求められる 角度 [°]
-40	-12.8	-13.3	-41.7
-30	-10.0	-9.4	-28.1
-20	-6.8	-6.9	-20.2
-10	-3.5	-4.2	-12.1
0	0	-0.8	-2.3
+10	3.5	2.3	6.6
+20	6.8	5.5	16.0
+30	10.0	8.6	25.5
+40	12.8	11.1	33.8

表 5 . Cirline/Jr1/imp000(畳部屋・大)の結果

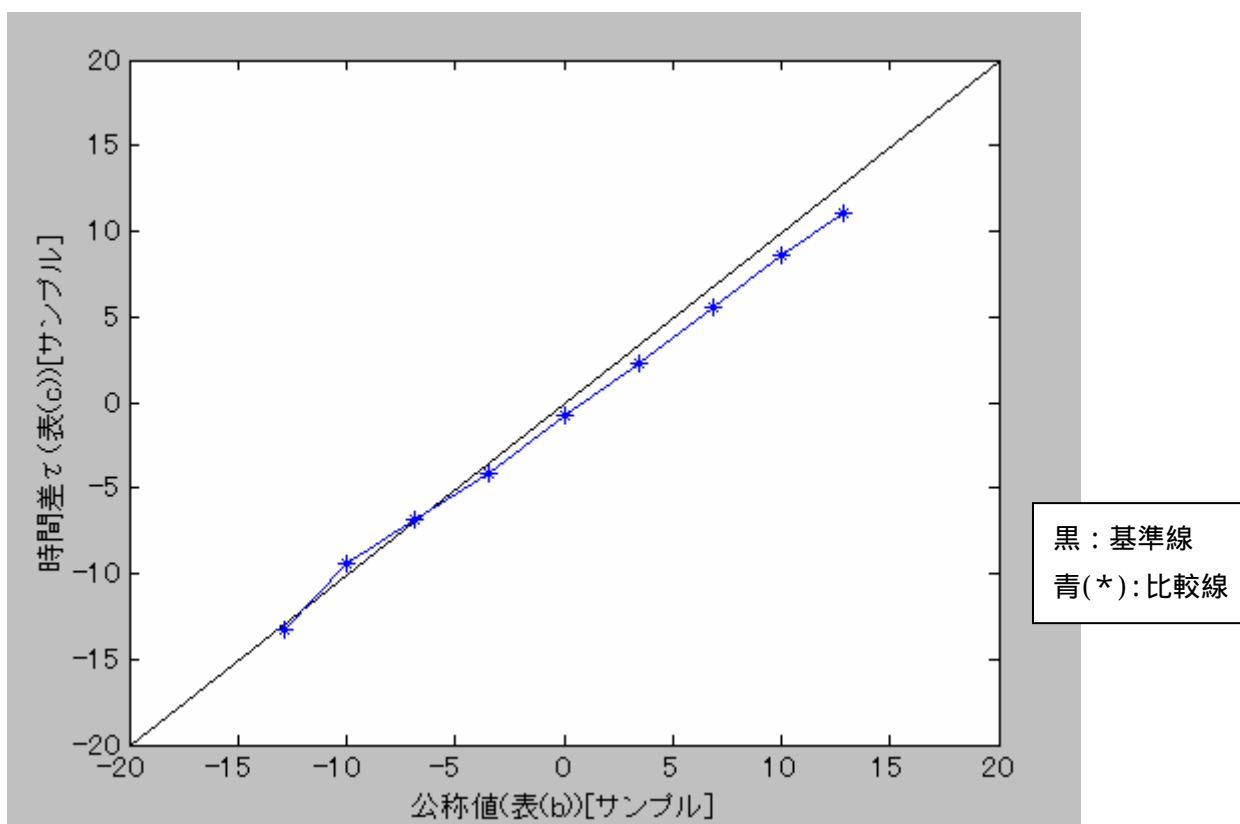


図 1 9 . 時間差の比較(Jr1)

< Jr2(畳部屋・小) >

(a) データベース の公称角度 [°]	(b) (a)と式4より求 められる時間差 [サンプル]	(c) データより求 めた時間差 t_{ij} [サンプル]	(d) (c)と式14よ り求められる角 度 [°]
-30	-10.0	-9.3	-27.8
-20	-6.8	-6.1	-17.8
-10	-3.5	-3.4	-9.7
0	0	-0.6	-1.7
+10	3.5	2.6	7.5
+20	6.8	5.8	16.9
+30	10.0	8.4	24.9

表6 . Cirline/Jr2/imp000(畳部屋・小)の結果

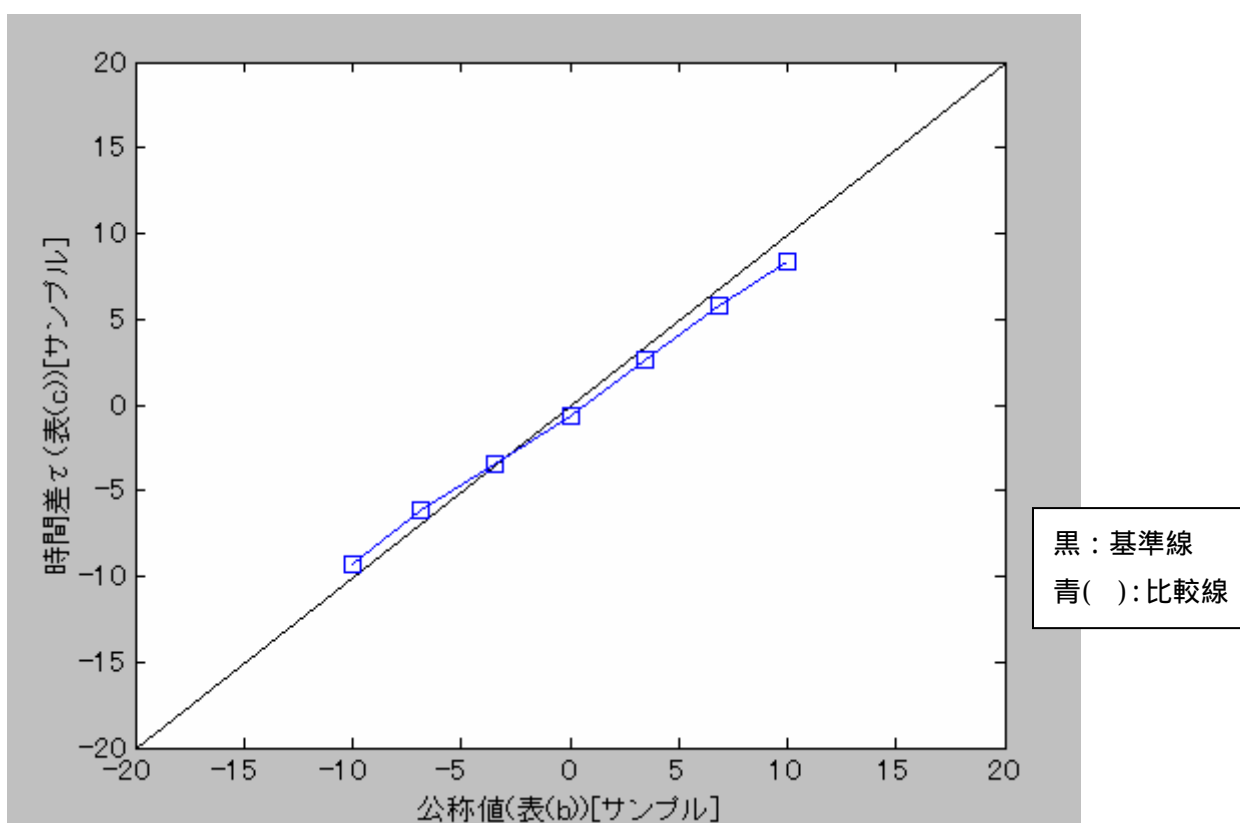


図20 . 時間差の比較(Jr2)

<0fc(会議室)>

(a) データベース の公称角度 [°]	(b) (a)と式 4 より求 められる時間差 [サンプル]	(c) データより求め た時間差 t_{ij} [サンプル]	(d) (c)と式 1 4 よ り求められる 角度 [°]
-40	-12.8	-11.4	-34.8
-30	-10.0	-8.1	-23.9
-20	-6.8	-5.2	-15.1
-10	-3.5	-2.4	-6.9
0	0	0.4	1.2
+10	3.5	3.8	11.0
+20	6.8	6.8	19.9
+30	10.0	9.3	27.8
+40	12.8	13.0	40.6

表 7 . Cirline/Ofc/imp000(会議室)の結果

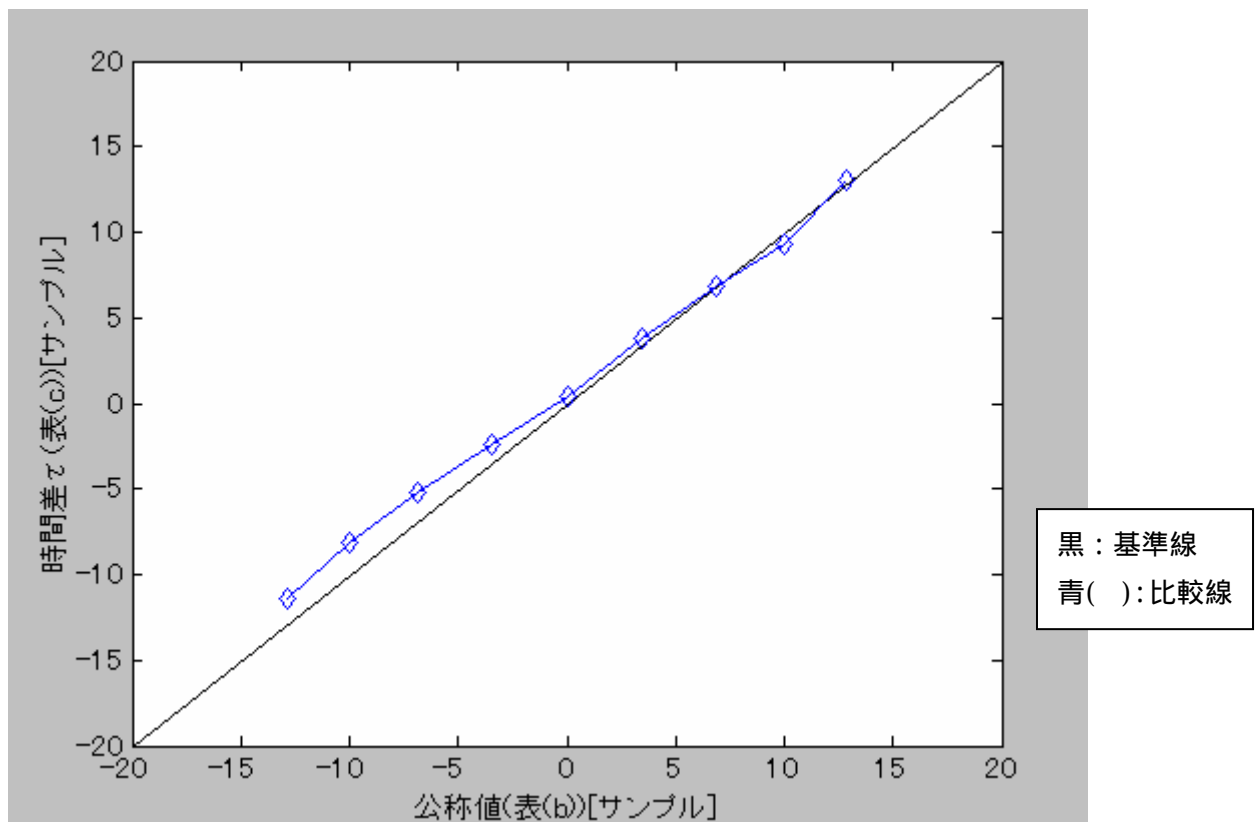


図 2 1 . 時間差の比較(0fc)

<結果>

E2A、E2B は、図 1 1 . 1 2 より基準線(黒い実線)と比較線(青い実線)は、一致していないことがわかる。

その他の Jr1, Jr2, Ofc は、ほぼ基準線と重なっていることがわかる。

以下に、比較線が基準線とずれている E2A、E2B に対してデータベースの実験条件に修正を加えることを検討した。

3 - 5 . RWDBデータに対する補正

E2A,E2B のプログラムで、表 2 , 3 (a)の角度：-80,-60,60,80[°]のように角度が鋭角に入ってくる場合に、遅延和アレーより求められた角度と異なっている。(表 3、4 より、(a)より求められる時間差とマイクロホン 2 1 , 2 6 の時間差が異なっている。)

このことについて検討する。

最初に与えられている設定について考える。

<パラメータ> c=340 不変

d=0.0283

Fs=48000

よって、d と Fs をそれぞれ変えてみる。

<復習>

$$\begin{aligned} \text{taus} &= s * Fs \\ &= (d \sin s) / c * Fs \end{aligned}$$

(a) ファイル名 [imp000] (角度[°])	(b) d : 1225 (0.0245)	(c) Fs : 41500	(d) データより求めた 時間差 _{ij} (E2A)[サンプル]	(e) データより求め た時間差 _{ij} (E2B)[サンプル]
010(-80)	-17.0	-17.0	-17.0	-16.8
030(-60)	-15.0	-15.0	-14.9	-14.9
050(-40)	-11.1	-11.1	-10.9	-10.9
070(-20)	-5.9	-5.9	-5.9	-5.4
090(0)	0	0	0	0.3
110(+20)	5.9	5.9	6.1	6.5
130(+40)	11.1	11.1	11.4	11.3
150(+60)	15.0	15.0	15.1	15.1
170(+80)	17.0	17.0	16.5	16.3

表 8 . Cirline/E2A・E2B の d、Fs 変更結果(時間差)

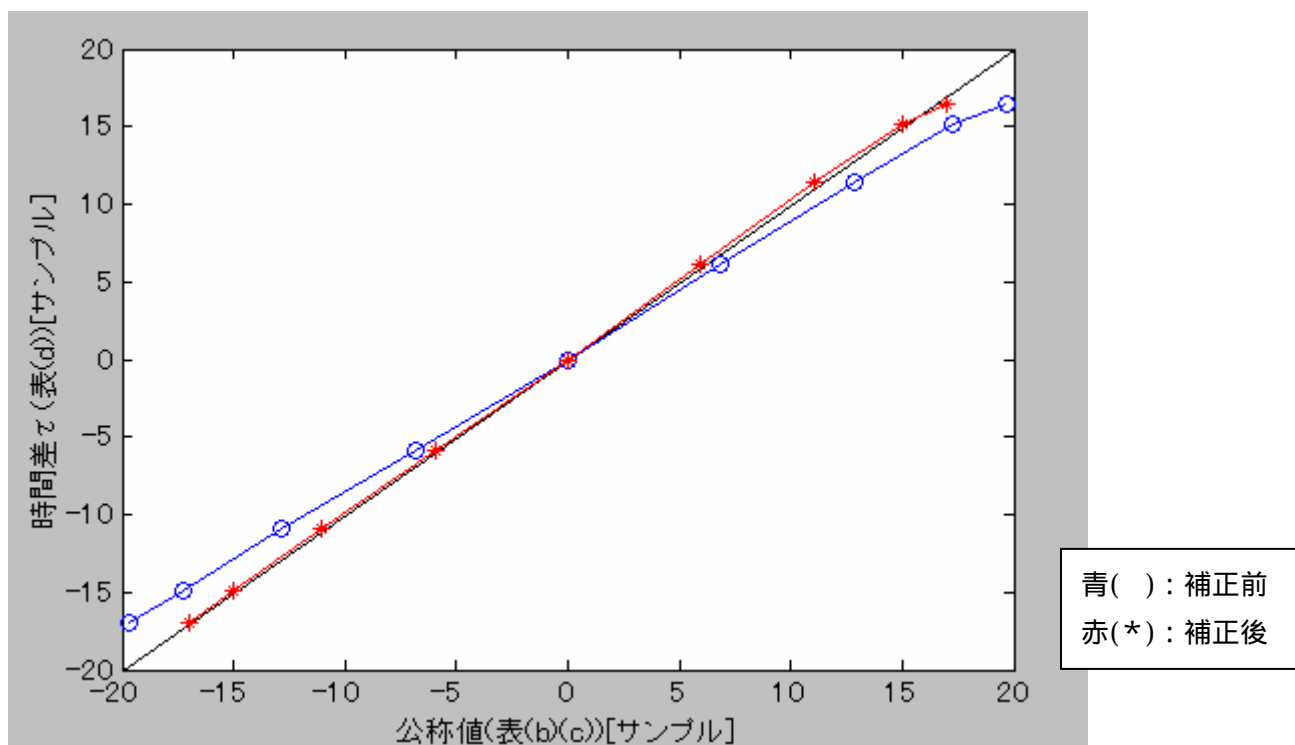


図 2 2 . 時間差の比較と補正(E2A)

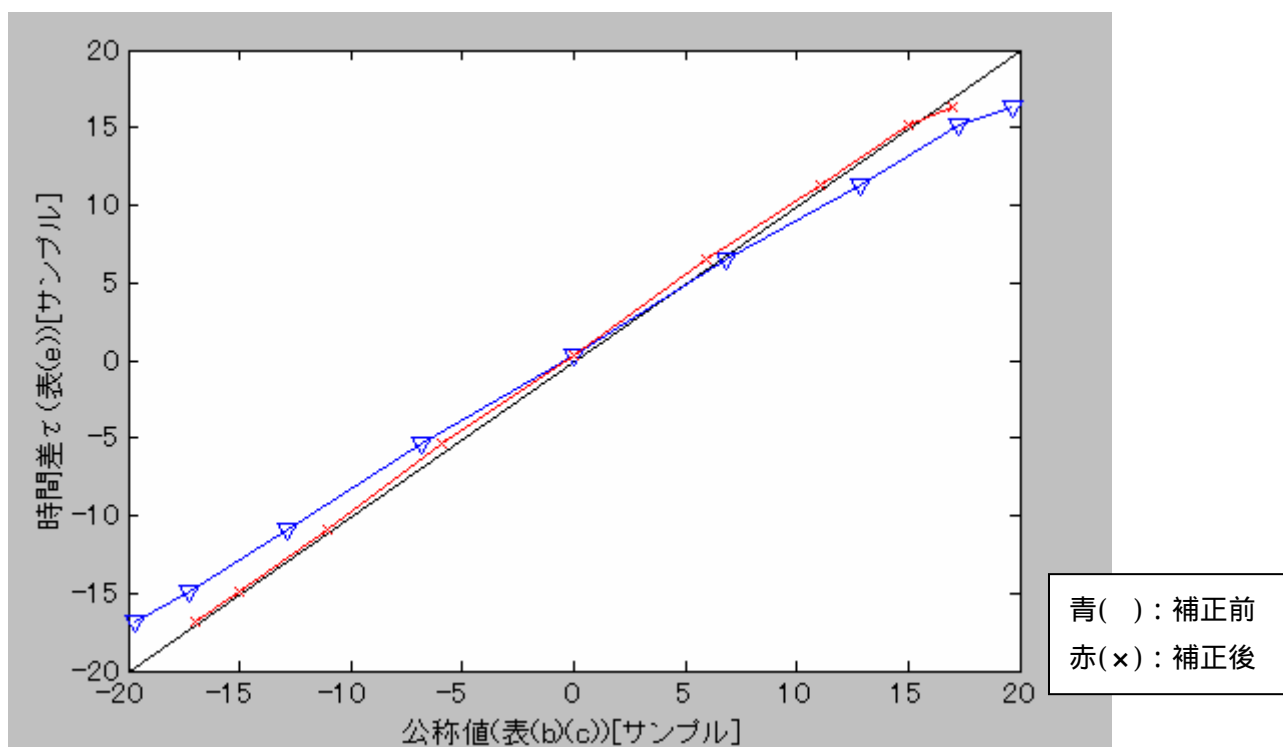


図 2 3 . 時間差の比較と補正(E2A)

図 2 2 , 2 3 より、 $d = 0.0245$ ($F_s=48000$ のまま)で計算した時、または、 $F_s=41500$ ($d = 0.0283$ のまま)で計算した時に、データより求めた時間差 t_{ij} の値に近くなることがわかる。

(より、黒い実線に重なっている)

よって、今回使用したデータには、適切な補正が有効であると考えられる。

3 - 6 . 遅延和アレーの構成と、音源方向推定の評価

(インパルス応答を使用した推定)

4つのマイクロホン(17,21,25,29)を利用した遅延和アレーでインパルス応答データに基づいて音源方向 θ を推定する。

<方法と手順>

)パラメータの設定

$$d = 0.0283 * 4 \text{ [m]}$$

$$c = 340 \text{ [m/s]}$$

$$F_s = 48000 \text{ [Hz]}$$

)ファイルを読み込む(4つのマイクロホン(#17,21,25,29)のインパルス応答)。

)遅延和アレーを用いて、方向ごとのアレー出力パワーの計算。

)データを内挿する。

)表示されたグラフから、最大値をとるx軸の値(音源方向の推定角度)を読み取る。

(コマンドウィンドウに最大値をとるx軸の値を表示する。)

以下に、遅延和アレーのプログラム(音源到来方向を推定するプログラム)を示す。

```
% tienwa1

% 遅延和アレーによる音源方向探査(このプログラムでは、ファイル名 imp090 は thita=0
である。)
% 2001.11.06 moriyama

% パラメータの設定
c=340
d=0.1132
Fs=48000

% ファイル読み込み
fid = fopen('d:\array%E2b%imp010%imp010.17','r')
[FILENAME,PERMISSION,MACHINEFORMAT] = fopen(fid)
data1=fread(fid,'float32');

fid = fopen('d:\array%E2b%imp010%imp010.21','r')
[FILENAME,PERMISSION,MACHINEFORMAT] = fopen(fid)
data2=fread(fid,'float32');

fid = fopen('d:\array%E2b%imp010%imp010.25','r')
[FILENAME,PERMISSION,MACHINEFORMAT] = fopen(fid)
data3=fread(fid,'float32');

fid = fopen('d:\array%E2b%imp010%imp010.29','r')
[FILENAME,PERMISSION,MACHINEFORMAT] = fopen(fid)
data4=fread(fid,'float32');

fclose(fid);
```

```

% 遅延和アレーの計算
for thita=-90:10:90          %-90~90[°]を10「°」ずつ
    tau=round(Fs*d*(sin(thita*pi/180))/c);
    y=data1(101-tau:1500-tau)+data2(101-2*tau:1500-2*tau)+data3(101-3*tau:1500-3
        *tau)+data4(101-4*tau:1500-4*tau);
    py(thita/10+10)=y*y;
end

cc1=interp1(py,[1:0.1:19],'spline'); %1度ずつ内挿する。
x=[-90:1:90];          %x軸の目盛ラベルを-90~90[°]にする。
plot(x,cc1,'b');
grid;
[yjiku xjiku]=max(cc1);
xjiku1=(xjiku/10-9)*10

% x軸、y軸のラベル表示
xlabel('角度 ')
ylabel('パワー ')

```


<例．遅延和アレープログラムの実行結果(E2a/imp070) >

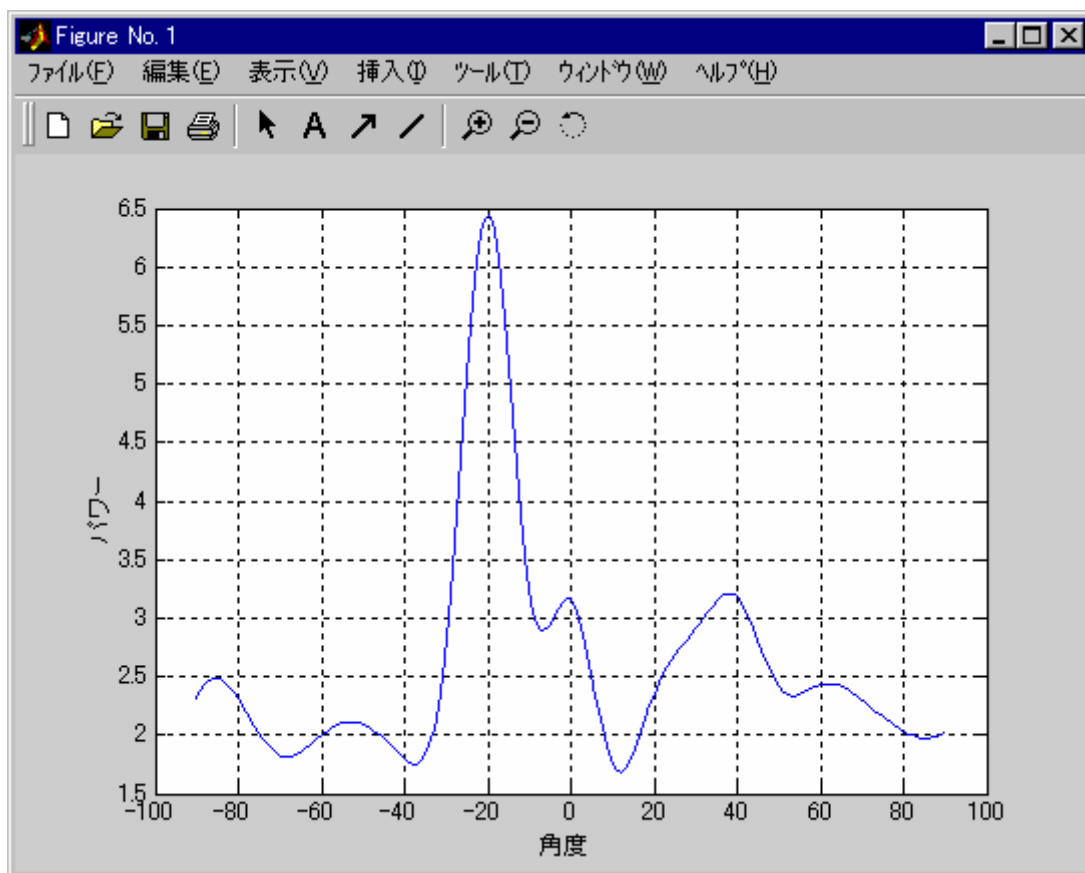


図 2 4 . imp070 の時の、アレー出力パワー $P_y(\theta)$

上図のような結果が表れる。

図 2 4 とコマンドウィンドウより、 $\theta = -20$ であることが分かる。

同様に、遅延和アレー(インパルス応答のみ)で音源方向 θ をもとめ、時間差 τ_{ij} ($i = 2, 1$ 、 $j = 2, 6$) から推定した角度と比較する。

以下、表 9 ~ 1 3 に各部屋のデータをまとめた。

< E2A(残変室 A) >

(a) データベース の公称角度 [°]	(b) データより求 めた時間差 t_{ij} [サンプル]	(c) (b)と式 1 4 よ り求められる 角度 θ_L [°]	(d) 遅延和アレー より求められ た角度 θ_s [°]
-80	-17.0	-58.3	-60
-60	-14.9	-48.2	-49
-40	-10.9	-33.0	-35
-20	-5.9	-17.2	-19
0	0	0	1
+20	6.1	18.7	21
+40	11.4	34.8	39
+60	15.1	49.1	51
+80	16.5	55.7	61

表 9 . 音源方向 θ_s と時間差 t_{ij} から推定した角度 θ_L の比較(E2A)

< 表の説明 >

- (a) ファイル名 imp090 を公称角度 0 ° (図 9 の F R O N T) として 20 ° おきに ± 80 ° (の範囲で音源を移動させている。(E2A, E2B の場合。 Jr1, 0fc: 10 ° おきに ± 40 ° 、 Jr2: 10 ° おきに ± 30 °)
- (b) 二つのマイクロホン (#21, 26) で受信したインパルス応答信号の時間差 t_{ij} の検出値
- (c) 時間差 t_{ij} と式 (15) ($\theta_L = \sin^{-1}(c * (t_{ij} / F_s) / d)$) から求めた角度
- (d) 遅延和アレープログラム (マイクロホン #17, 21, 25, 29 を使用、インパルス応答データ使用) を実行して求められる音源方向 θ_s

- ・表 9 (c), (d) の値が同じような値になれば、遅延和アレープログラムは正常に動作していると考えられる。

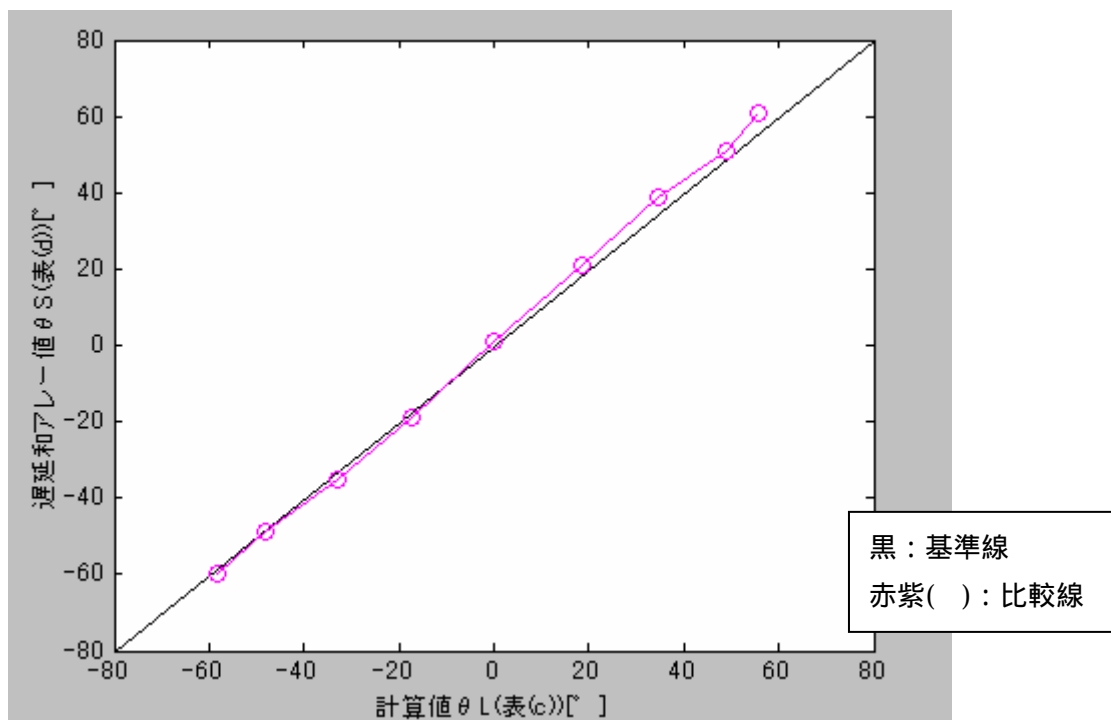


図 2.5 . 角度の比較(E2A)

< 図の説明 >

x 軸：表 9 (b) と式 (15) から求めた角度 θ_L (計算値) (表 9 (c))

y 軸：遅延和アレープログラム (マイクロホン #17, 21, 25, 29 を使用) を実行して求められる音源方向 θ_S (表 9 (d))

黒い実線 (マーカなし)：基準線

赤紫の実線 (マーカ：)：計算値 θ_L と音源方向 θ_S の交点 (印) を結んだ線 (比較線とする)。

- ・ 計算値 θ_L と音源方向 θ_S の値が等しい場合に、比較線が基準線と一致する。
よって、比較線が基準線に重なると、遅延和アレープログラムは正常に動作していると考えられる。

< E2B(残変室 B) >

(a) データベース の公称角度 [°]	(b) データより求 めた時間差 t_{ij} [サンプル]	(c) (b)と式 1 4 よ り求められる 角度 θ_L [°]	(d) 遅延和アレー より求められ た角度 θ_S [°]
-80	-16.8	-57.2	-60
-60	-14.9	-48.2	-49
-40	-10.9	-33.1	-29
-20	-5.4	-15.7	-17
0	0.3	0.9	1
+20	6.5	19.0	21
+40	11.3	34.4	40
+60	15.1	49.1	51
+80	16.3	54.7	61

表 1 0 . 音源方向 θ_S と時間差 t_{ij} から推定した角度 θ_L の比較 (E2B)

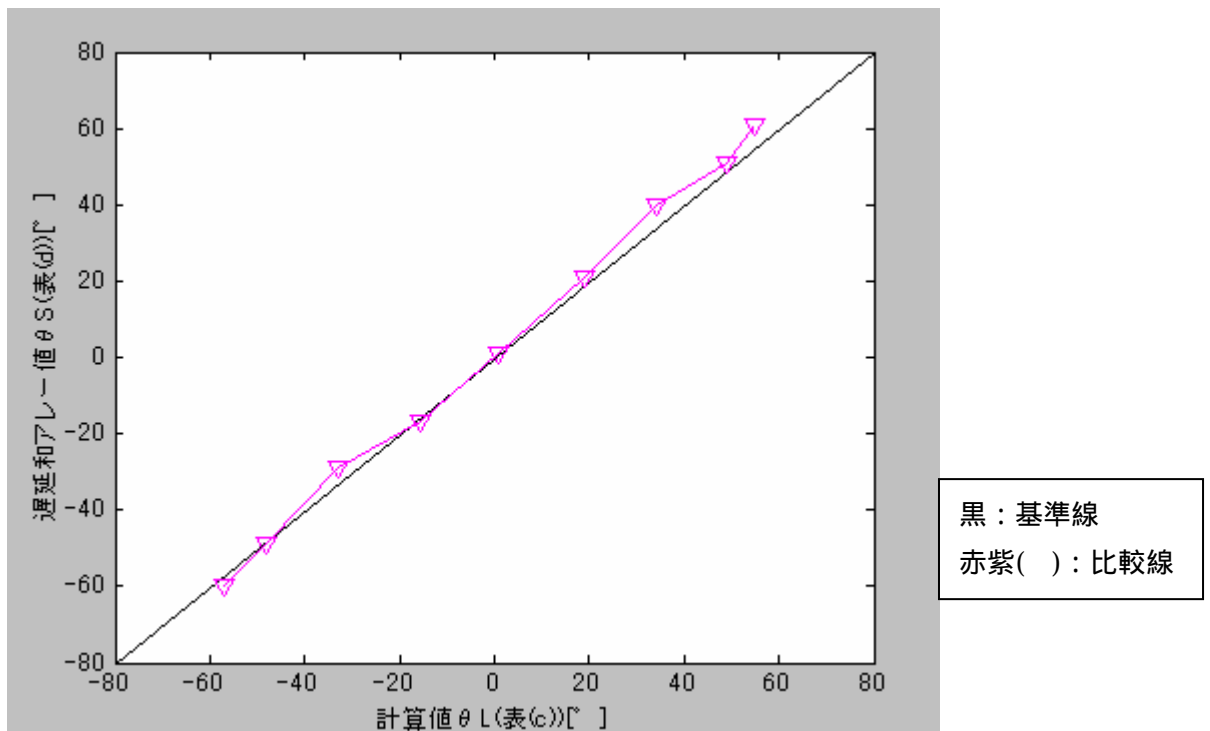


図 2 6 . 角度の比較 (E2B)

< Jr1(畳部屋・大) >

(a) データベース の公称角度 [°]	(b) データより求 めた時間差 t_{ij} [サンプル]	(c) (b)と式 1 4 よ り求められる 角度 θ_L [°]	(d) 遅延和アレー より求められ た角度 θ_S [°]
-40	-13.3	-41.7	-39
-30	-9.4	-28.1	-29
-20	-6.9	-20.2	-19
-10	-4.2	-12.1	-9
0	-0.8	-2.3	1
+10	2.3	6.6	11
+20	5.5	16.0	20
+30	8.6	25.5	30
+40	11.1	33.8	38

表 1 1 . 音源方向 θ_S と時間差 t_{ij} から推定した角度 θ_L の比較 (Jr1)

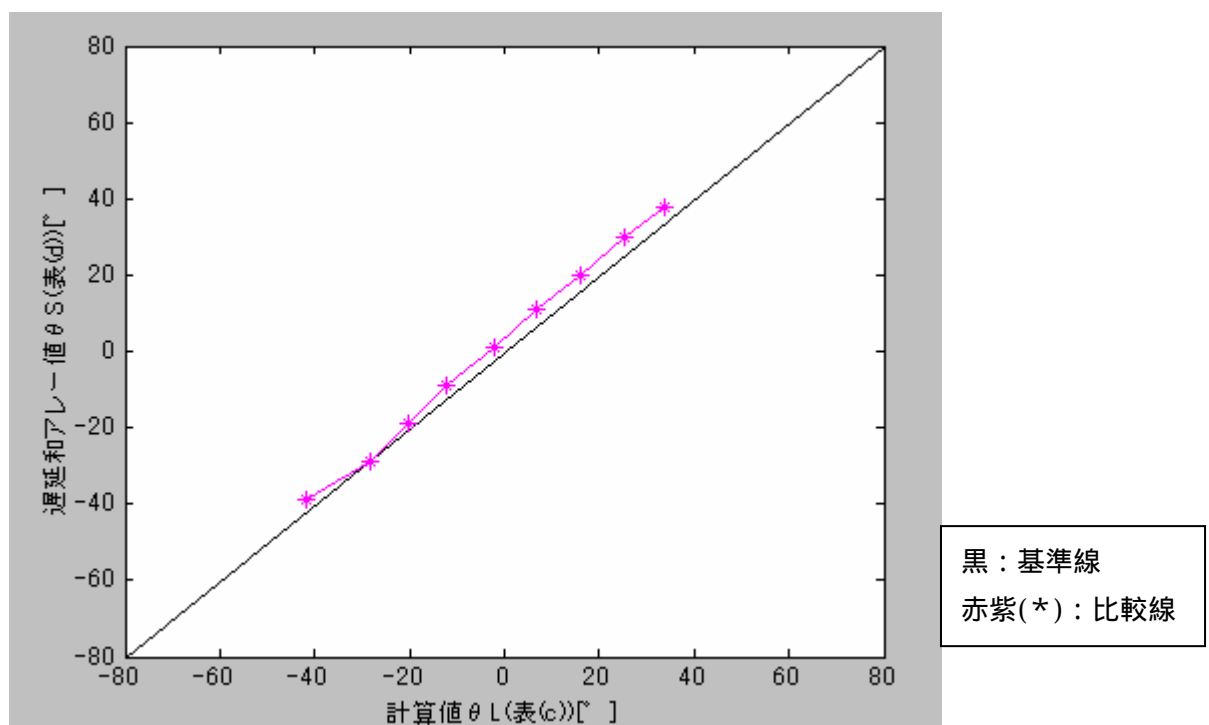


図 2 7 . 角度の比較 (Jr1)

< Jr2(畳部屋・小) >

(a) データベース の公称角度 [°]	(b) データより求め た時間差 t_{ij} [サンプル]	(c) (b)と式14よ り求められる 角度 θ_L [°]	(d) 遅延和アレー より求められ た角度 θ_s [°]
-30	-9.3	-27.8	-29
-20	-6.1	-17.8	-19
-10	-3.4	-9.7	-9
0	-0.6	-1.7	1
+10	2.6	7.5	11
+20	5.8	16.9	20
+30	8.4	24.9	30

表12. 音源方向 θ_s と時間差 t_{ij} から推定した角度 θ_L の比較(Jr2)

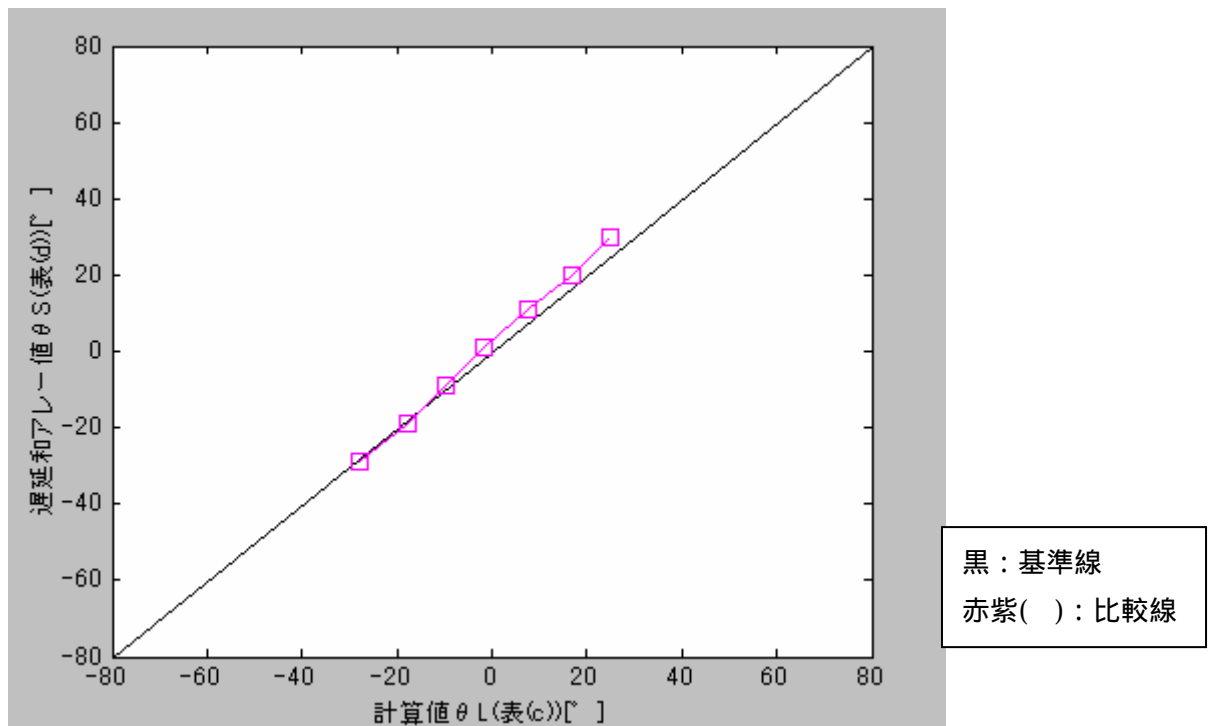


図28. 角度の比較(Jr2)

<0fc(会議室)>

(a) データベース の公称角度 [°]	(b) データより求 めた時間差 τ_{ij} [サンプル]	(c) (b)と式14よ り求められる 角度 θ_L [°]	(d) 遅延和アレー より求められ た角度 θ_S [°]
-40	-11.4	-34.8	-29
-30	-8.1	-23.9	-20
-20	-5.2	-15.1	-14
-10	-2.4	-6.9	-10
0	0.4	1.2	0
+10	3.8	11.0	11
+20	6.8	19.9	20
+30	9.3	27.8	31
+40	13.0	40.6	41

表13. 音源方向 θ_S と時間差 τ_{ij} から推定した角度 θ_L の比較(Jr2)

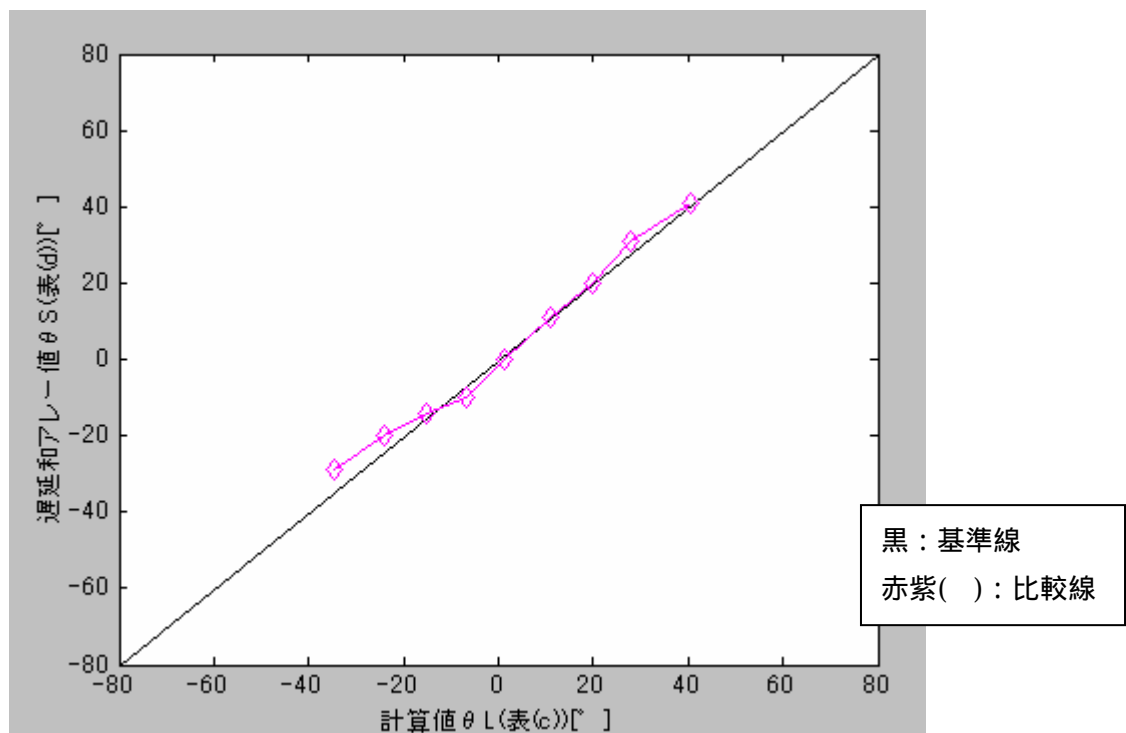


図29. 角度の比較(0fc)

<結果>

図 2 5 , 2 6 , 2 7 , 2 8 , 2 9 より、多くの比較値(マーカ部分)は基準線に、近い値をとっている事がわかる。このことから、遅延和アレーのプログラムは正常に動いていると考えてよい。

しかし、その中でも計算値 L (表(c))と遅延和アレーのプログラムから求められる音源方向 s (表(d))の値が大きくなってしまったデータについて検討した。

3 - 6 - 1 . 遅延和アレープログラムで使したマイクロホン(#17,21,25,29)のインパルス応答の波形について調べる。

残響可変室 A(ファイル名 : E2A)では、表9、図25からわかるように、データベース公称角度 +40[°]の時に、計算値 θ_L (表(c))と遅延和アレーのプログラムから求められる音源方向 θ_s (表(d))が、 $\theta_L=34.8^\circ$ 、 $\theta_s=39^\circ$ と大きく異なってしまっている。この時の、インパルス応答の波形(#17,21,25,29)を調べてみた。

以下に、インパルス応答波形の図を示す。

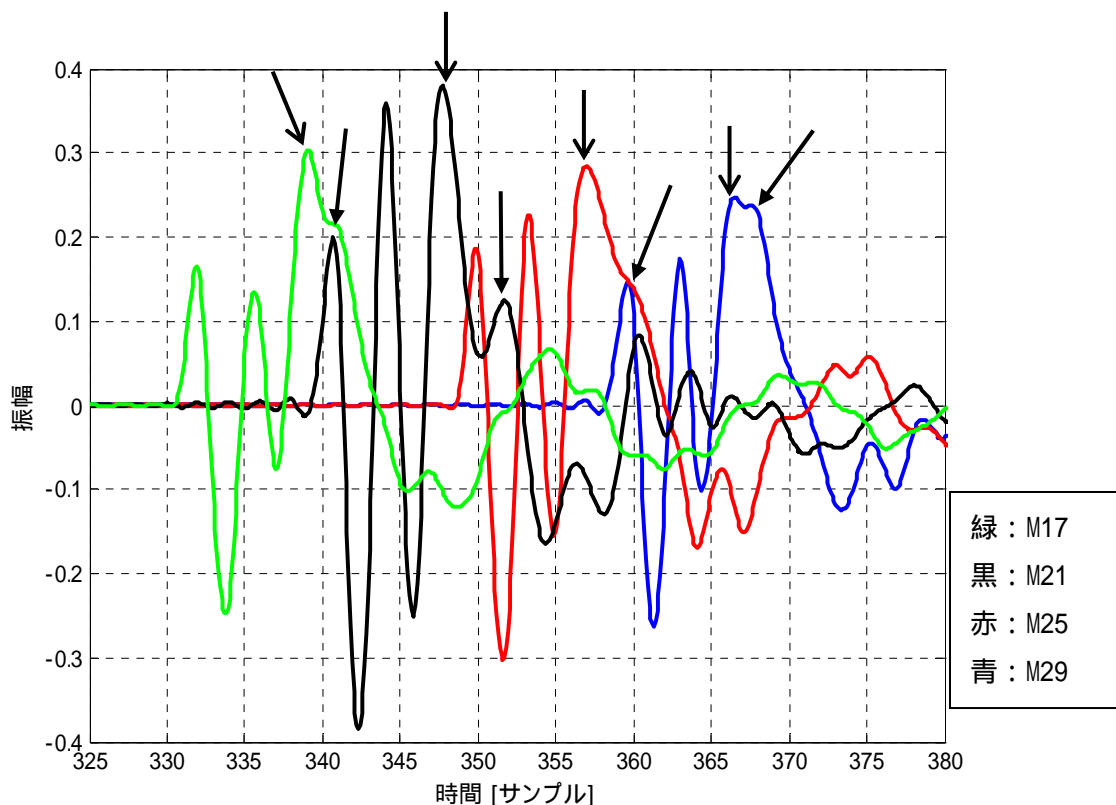


図30 . インパルス応答の波形(公称角度 +40°)

振幅の最大値をとる直接音(\searrow)の他に、反射音(\swarrow)が明確に見ることができる。この反射音が測定値に影響を与えた要因と考えられる。E2Aのデータベース公称角度 +40[°]の他にも、計算値 θ_L (表(c))と音源方向 θ_s (表(d))が大きく異なってしまっているところがある。それらについても、反射音に影響を与えたと考えられる。

3 - 6 - 2 . 音源 - マイクロホンの距離の関係

3 - 1 の遅延和アレーの原理で示した音源方向 θ は無限遠方から到来していることを前提として考えられている。しかし、今回使用した RW データベースは、マイクロホンアレーを中心に 2[m] の距離にスピーカー (音源) が配置されている。

この場合、等間隔 d (2.83[cm]) で配置されたマイクロホンで受信される信号に、 $\theta \sim M$ の遅れを付加することで、本当に同相化されるか検討してみる。(遅延和アレープログラムで使用した 4 つのマイクロホン (#17, 21, 25, 29) で検討する)

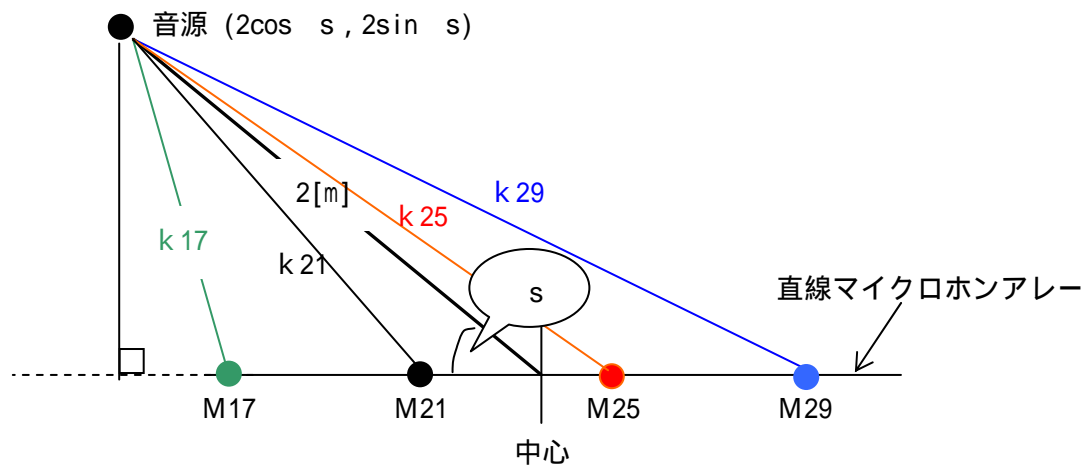


図 3 1 . 音源 - 各々のマイクロホン間の距離

音源 - 中心間の距離 2[m] と音源方向 s より、音源の座標は

$$(x, y) = (2\cos s, 2\sin s) \cdots (16)$$

で求めることができる。そして、音源 M17, 21, 25, 29 のマイクロホン間の距離を上図のように $k_{17, 21, 25, 29}$ とする。距離 $k_{17, 21, 25, 29}$ は以下のような式で求められる。

$$k_{17} = \sqrt{(2\cos s - d_{17})^2 + y^2} \cdots (17)$$

(d_{17} : 中心から M17 までの距離 (= -6.5 * d))

$$k_{21} = \sqrt{(2\cos s - d_{21})^2 + y^2} \cdots (18)$$

(d_{21} : 中心から M21 までの距離 (= -2.5 * d))

$$k_{25} = \sqrt{(2\cos s - d_{25})^2 + y^2} \cdots (19)$$

(d_{25} : 中心から M25 までの距離 (= 1.5 * d))

$$k_{29} = \sqrt{(2\cos s - d_{29})^2 + y^2} \cdots (20)$$

(d_{29} : 中心から M29 までの距離 (= 5.5 * d))

ただし、-6.5、-2.5 などの数字は、マイクロホン間隔 d を単位としたものである。マイクロホン #23, 24 との間にマイクロホンアレー中心があるため、例えば、#25 までの距離は $1.5 * d$ となる。

式(17)(18)(19)(20)から求められる実際の距離から、時間差を求める。

k 17,21 から求められる時間差： τ_1 、k 21,25 から求められる時間差： τ_2 、k 25,29 から求められる時間差： τ_3 とする。

10[°]おきに公称角度-80~0[°]の範囲で計算した時間差 τ_1 ・ τ_2 ・ τ_3 を表にまとめる。

データベース 公称角度[°]	k 17,21から求め られる時間差 τ_1 [サンプル]	k 21,25から求め られる時間差 τ_2 [サンプル]	k 25,29から求め られる時間差 τ_3 [サンプル]
-80	- 15.8	- 15.7	- 15.7
-70	- 15.1	- 15.0	- 14.9
-60	- 14.0	- 13.9	- 13.6
-50	- 12.6	- 12.3	- 11.9
-40	- 10.8	- 10.3	- 9.8
-30	- 8.7	- 8.1	- 7.4
-20	- 6.3	- 5.6	- 4.7
-10	- 3.7	- 2.9	- 2.0
0	- 1.0	- 0.1	+ 0.8

表 1 4 . 時間差 τ_1 ・ τ_2 ・ τ_3 の比較

表 1 4 からわかるように、時間差 τ_1 ・ τ_2 ・ τ_3 は公称角度が同じ場合でも、同じ時間差で到来していない。よって、遅延和アレーのプログラム上で、時間差 τ を公称角度とマイクロホン間の距離から計算し、その倍数で遅延していく方法では、完全な同相化ができていないことがわかった。

3 - 7 . 遅延和アレーの構成と、音源方向推定の評価

(インパルス応答 + 音声信号による方向推定)

実際の問題として、音源方向を推定する場合、音源からの音はインパルスではなく、より一般的な音である場合が多い。たとえば、人が呼ぶ方向を検出する場合などでは、対象が音声信号となる。

そこで、4つのマイクロホン(17,21,25,29)を利用した遅延和アレーでインパルス応答に音声をたたみ込んだデータに基づいて音源方向 s を推定して、評価を行った。

<方法と手順>

)MATLABにおいて音声を取り込み、保存する。(下記のプログラムで行った。)

```
%onsei1
```

```
% 音声を取り込み、保存する。
```

```
% 2001.12.10 moriyama
```

```
h1=wavrecord(48000,48000); %h1 に音声を取り込む。
```

```
plot(h1); %取り込んだ音声をプロットする。
```

```
wavwrite(h1,48000,'d:¥検出¥otomo'); %“ d:¥検出¥otomo ”に保存する。
```

```
pause %一時停止(“ enter ” で再開)
```

```
sound(h1,48000) %取り込み音声のチェック
```

) 遅延和アレープログラムのファイルの読み込み部分を、取り込んだ音声データとRWB Dのデータをたたみ込んだデータを読み込むようにする。

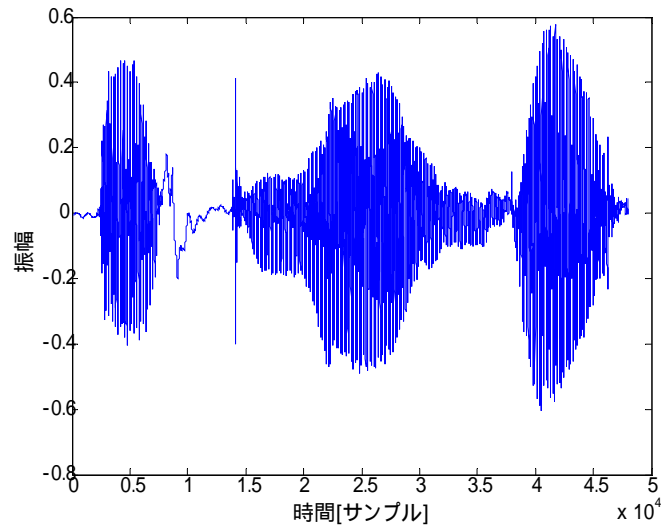
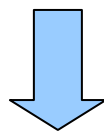


図 3 2 . 使用した音声のグラフ ('d:¥検出¥otomo3)



インパルス応答に
たたみ込む。

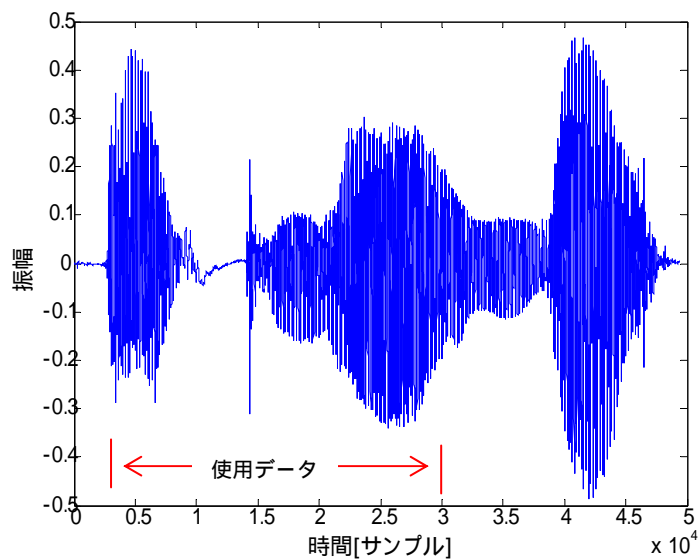


図 3 3 . インパルス応答 + 音声波形 (imp070, M17)

-) 遅延和アレーの計算。
-) データを内挿する。
-) 表示されたグラフから、最大値をとる x 軸の値を読み取る。
(コマンドウィンドウに最大値をとる x 軸の値を表示する。)

以下に、インパルス応答 + 音声のデータを使用した遅延和アレーのプログラムを示す。

```

% tienwa1otomo3

% 遅延和アレーによる音源方向探査(データはインパルス応答 + 音声・otomo3)
% 2002.2.21 moriyama

% パラメータの設定
c=340
d=0.1132
Fs=48000

% 音声の読み込み
data=wavread('d:¥検出¥otomo3');
%plot(data)

% x 軸、y 軸のラベル表示
xlabel('時間[サンプル]')
ylabel('振幅 ')
%return

% ファイル読み込み
fid = fopen('d:¥array¥E2a¥imp070¥imp070.17','r')
[FILENAME,PERMISSION,MACHINEFORMAT] = fopen(fid)
data1=fread(fid,'float32');
data1=conv(data1(100:1500),data);

fid = fopen('d:¥array¥E2a¥imp070¥imp070.21','r')
[FILENAME,PERMISSION,MACHINEFORMAT] = fopen(fid)
data2=fread(fid,'float32');
data2=conv(data2(100:1500),data);

fid = fopen('d:¥array¥E2a¥imp070¥imp070.25','r')
[FILENAME,PERMISSION,MACHINEFORMAT] = fopen(fid)
data3=fread(fid,'float32');
data3=conv(data3(100:1500),data);

```

```

fid = fopen('d:\array\E2a\imp070\imp070.29', 'r')
[FILENAME, PERMISSION, MACHINEFORMAT] = fopen(fid)
data4=fread(fid, 'float32');
data4=conv(data4(100:1500), data);

fclose(fid);

% 遅延和アレーの計算
for thita=-90:10:90          %-90 ~ 90 を 10 ずつ
    tau=round(Fs*d*(sin(thita*pi/180))/c);
    y=data1(3000-tau:35000-tau)+data2(3000-2*tau:35000-2*tau)+data3(3000-3*tau
    :35000-3*tau)+data4(3000-4*tau:35000-4*tau);
    py(thita/10+10)=y*y;
end

cc1=interp1(py, [1:0.1:19], 'spline');
x=[-90:1:90];
plot(x, cc1, 'b');
grid;
[yjiku xjiku]=max(cc1);
xjiku1=(xjiku/10-9)*10

% x 軸、 y 軸のラベル表示
xlabel('角度 ')
ylabel('パワー ')

```


< 例. インパルス応答 + 音声データの遅延和アレープログラム実行結果 (E2A/imp070) >

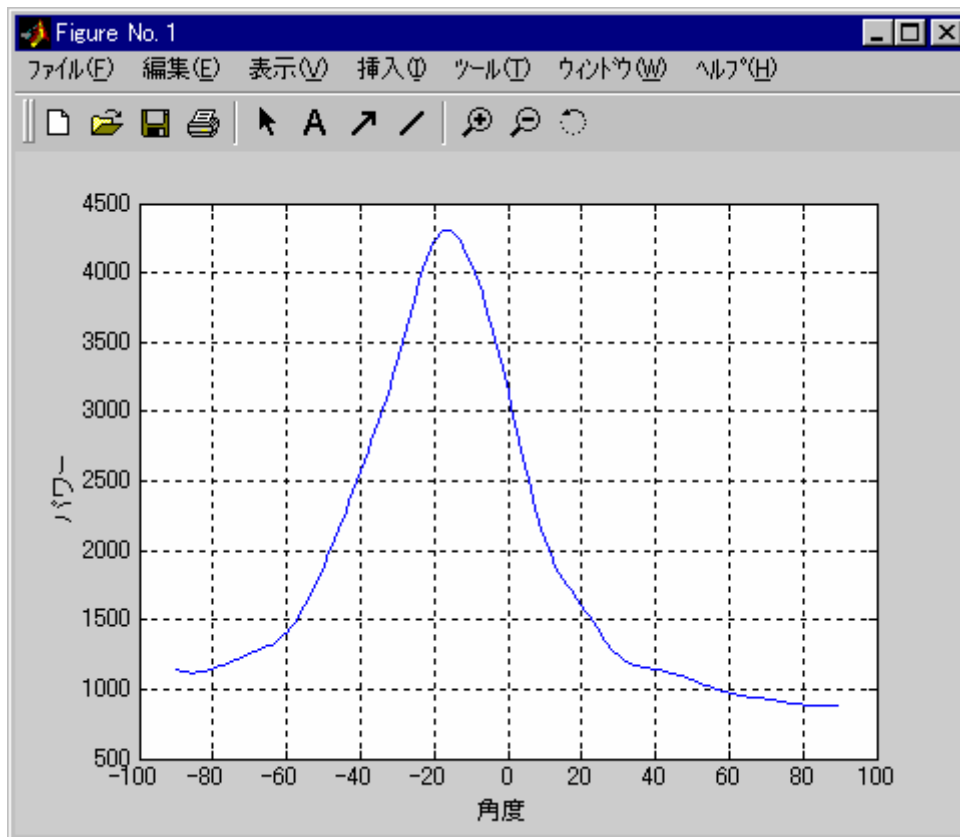


図 3 4 . Imp070 の時の、アレー出力パワー $P_y (s)$

上図のような結果が表れる。

図 3 4 とコマンドウィンドウより、 $s = -16.0$ であることが分かる。

同様に、音源方向や室条件の異なったインパルス応答を音声とたたみ込み、これを用いて遅延和アレーによる音源方向推定を繰り返した。音声データは以下を使用した。

音声データ : 'd:¥検出¥otomo3' 音声... “バクオンガ”

音声データ : 'd:¥検出¥otomo' 音声... “ザジズゼゾ”

インパルス応答に音声データをたたみ込んだデータを使用して求められる音源方向を s_1 、音声データを使用して求めた音源方向を s_2 とする。

また、3 - 6, 3 - 7 で説明した方法で、2 つのマイクロホン (#21, 26) のインパルス応答の最大値の差 (時間差 t_{ij}) を用いて推定した角度 θ_L と表し、これらと比較した。

以下、表 1 5 ~ 1 9 に各部屋のデータをまとめた。

< E2A(残変室 A) >

(a) データベース の公称角度 [°]	(b) 時間差 t_{ij} か ら求められる 角度 θ_L [°]	(c) 遅延和アレー より求められ た角度 s_1 [°]	(d) 遅延和アレー より求められ た角度 s_2 [°]
-80	-58.3	-59	-59
-60	-48.2	-47	-49
-40	-33.0	-33	-34
-20	-17.2	-16	-17
0	0	1	1
+20	18.7	19	20
+40	34.8	35	37
+60	49.1	52	51
+80	55.7	59	58

表 1 5 . 音源方向 s_1 , s_2 と時間差 t_{ij} から推定した角度 θ_L の比較(E2A)

< 表の説明 >

(a) ファイル名 imp090 を公称角度 $0[^\circ]$ (図 9 の FRONT) として $20[^\circ]$ (おきに $\pm 80[^\circ]$) (の範囲で音源を移動させている。(E2A, E2B の場合。 Jr1, 0fc: $10[^\circ]$ おきに $\pm 40[^\circ]$ 、 Jr2: $10[^\circ]$ おきに $\pm 30[^\circ]$)

(b) 二つのマイクロホン (#21, 26) で受信したインパルス応答信号の時間差 t_{ij} と式(15) ($\theta_L = \sin^{-1}(c * (t_{ij} / Fs) / d)$) から求めた角度(計算値)

(c) 遅延和アレープログラム(マイクロホン#17, 21, 25, 29 を使用、インパルス応答 + 音声データ使用) を実行して求められる音源方向 s_1

音声データ: 'd:¥検出¥otomo3' 音声... “バクオンガ”

(d) 遅延和アレープログラム(マイクロホン#17, 21, 25, 29 を使用、インパルス応答 + 音声データ使用) を実行して求められる音源方向 s_2

音声データ: 'd:¥検出¥otomo' 音声... “ザジズゼゾ”

・表 1 5 (b) が (c) (d) の値と同じような値になれば、遅延和アレープログラムは正常に動作していると考えられる。

< E2B(残変室 B) >

(a) データベース の公称角度 [°]	(b) 時間差 t_{ij} か ら求められる 角度 θ_L [°]	(c) 遅延和アレー より求められ た角度 s_1 [°]	(d) 遅延和アレー より求められ た角度 s_2 [°]
-80	-57.2	-45 ★	-59
-60	-48.2	-45	-48
-40	-33.1	-26	-28
-20	-15.7	-16	-16
0	0.9	2	1
+20	19.0	20	20
+40	34.4	25 ★	35
+60	49.1	51	51
+80	54.7	27 ★	59

表 1 6 . 音源方向 s_1 , s_2 と時間差 t_{ij} から推定した角度 θ_L の比較(E2B)
(★: θ_L と s_1 の違いが大きなデータ)

< Jr1(畳部屋・大) >

(a) データベース の公称角度 [°]	(b) 時間差 t_{ij} か ら求められる 角度 θ_L [°]	(c) 遅延和アレー より求められ た角度 s_1 [°]	(d) 遅延和アレー より求められ た角度 s_2 [°]
-40	-41.7	-19 ★	-34
-30	-28.1	-16 ★	-27
-20	-20.2	-16	-17
-10	-12.1	-13	-10
0	-2.3	-3	-1
+10	6.6	6	9
+20	16.0	14	19
+30	25.5	18	25
+40	33.8	25	33

表 1 7 . 音源方向 s_1 , s_2 と時間差 t_{ij} から推定した角度 θ_L の比較(Jr1)
(★: θ_L と s_1 の違いが大きなデータ)

< Jr2(畳部屋・小) >

(a) データベース の公称角度 [°]	(b) 時間差 t_{ij} か ら求められる 角度 θ_L [°]	(c) 遅延和アレー より求められ た角度 s_1 [°]	(d) 遅延和アレー より求められ た角度 s_2 [°]
-30	-27.8	-24	-28
-20	-17.8	-16	-17
-10	-9.7	-10	-10
0	-1.7	-1	-4
+10	7.5	10	11
+20	16.9	19	16
+30	24.9	31	25

表 1 8 . 音源方向 s_1 , s_2 と時間差 t_{ij} から推定した角度 θ_L の比較(Jr2)

< 0fc(会議室) >

(a) データベース の公称角度 [°]	(b) 時間差 t_{ij} から 求められる角 度 θ_L [°]	(c) 遅延和アレー より求められ た角度 s_1 [°]	(d) 遅延和アレー より求められ た角度 s_2 [°]
-40	-34.8	-33	-31
-30	-23.9	-27	-21
-20	-15.1	-16	-13
-10	-6.9	-3	-3
0	1.2	6	2
+10	11.0	16	15
+20	19.9	18	18
+30	27.8	32	29
+40	40.6	38	39

表 1 9 . 音源方向 s_1 , s_2 と時間差 t_{ij} から推定した角度 θ_L の比較(0fc)

< 結果・吟味 >

音源方向 s_1 について(音声データ: 'd:¥検出¥otomo3' 音声... “バクオンガ”)

E2A, Jr2, 0fc については、音源方向 s_1 (表(c))の値が計算値 L (表(b))の値とほぼ同じになった。E2B, Jr1 では、音源方向 s_1 (表(c))の値が計算値 L (表(b))の値と大きく異なってしまうデータがあった(表16, 17の★参照)。

これは、各々の室内の残響時間が影響を及ぼしたのではないかと考えられる。

残響可変室 A(E2A)の室内の残響時間は 0.30、これに対して、残響可変室 B(E2B)は、1.30 となっている。また、畳部屋・小(Jr2)の室内の残響時間は 0.47、畳部屋・大の室内の残響時間は 0.60 である。(表1のデータの一覧表参照)

残響時間が長く、残響の多い室内での推定誤りが多く見られたことがわかる。

音源方向 s_2 について(音声データ: 'd:¥検出¥otomo' 音声... “ザジズゼゾ”)

全てのデータ(E2A, E2B, Jr1, Jr2, 0fc)において、音源方向 s_1 (表(c))の値が計算値 L (表(b))の値とほぼ同じになった。

インパルス応答のみのデータで行った遅延和アレーで比較・検討してわかったような、反射音や、音源 マイクロホン間の距離が近いことも、同様に、インパルス応答 + 音声データで行った遅延和アレーの方向推定に影響を与えたと考えられる。

4 . まとめ

本稿では、初めに、時間差に基づく方向推定について学び、正弦波と雑音を用いて音源方向検出シミュレーションを行った。次に遅延和アレーに基づく方向推定を学び、本研究の目的である、共通データベースの評価および、遅延和アレーの性能検討を行った。

共通データベースの評価研究の結果では、データベースの実験条件に修正を加えることで、公称角度から求められる時間差と、二つのマイクロホン(#21, 26)で受信されるインパルス応答から求めた時間差 τ_{ij} の値がより近くなった。このことより、今回使用したデータには、適切な補正が有効であることがわかった。

遅延和アレーの性能検討研究の結果では、インパルス応答の波形より、直接音のほかに反射音があること、また、音源 - マイクロホンの距離が近いため、遅延和アレーのプログラム上で先に時間差を公称角度とマイクロホン間の距離から計算し、その倍数で遅延していく方法では、完全な同相化ができないことが、音源方向推定結果に影響を及ぼしているとわかった。さらに、インパルス応答に音声をたたみ込んだデータを使用して、遅延和アレープログラムより音源方向を推定した場合、データによって、残響時間の長く、残響の多い室内での音源方向推定には誤りが多く見られることもわかった。

今後は、遅延和アレーのプログラムで使用するマイクロホンの個数を増加させたり、音源とマイクロホンの距離が近いことによる影響を反映させるなどして、反射音の影響の克服することが課題となる。また、インパルス応答に音声をたたみ込んだデータを使用する場合、たたみ込む音声の種類により、音源方向推定の結果が異なってしまうため、よりさまざまな音声データを集め、研究を進めていく必要がある。

参考文献

-)坂巻 佳美,「見てわかる デジタル信号処理」, 工業調査会 (1998 年)
-) R W C P 実環境音声・音響データベース
<http://www.slt.atr.co.jp/~tnishi/DB/micarray/index.htm>
-)大賀、山崎、金田 :「音響システムとデジタル処理」電子情報通信学会,
pp.181 ~ 186(1995 年)

謝辞

本研究を進めるにあたり、広範なご指導をいただいた金田 豊教授、および卒業研究生各位に深い感謝の意を示します。