8　到来方向推定

(Direction of Arrival (DOA) Estimation)

到来音の空間的情報（音源方向、位置）は、雑音抑圧処理の性能向上や自律システムの状況判断のために重要。ヒトやロボットの危険察知にも。

The spatial information of arriving signals (source direction, source position) is important for improving performance of noise reduction systems and autonomous systems.

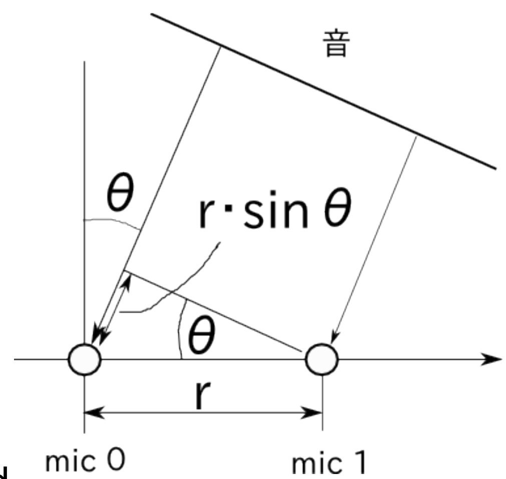
8-1. チャネル間時間差からの方向推定（相互相関関数）

DOA estimation based on time difference (cross-correlation function)

音速を c 、到来方向を θ 、観測信号に関するチャネル間の到達時間差を τ とすると、

Let c as sound velocity, θ as DOA, τ as the arrival time difference between channels, then,

$$\tau = r \sin\theta / c$$



であるため、到来方向 θ は、

is hold. Therefore, the arrival angle θ is obtained from time difference τ as,

$$\theta = \sin^{-1} (c\tau / r) .$$

となり、時間差 τ から求まる。

時間差検出には、相互相関関数を白色化によって分解能を向上させた一般化相互相関関数(Generalized Cross Correlation Function (**GCC**)) がよく使われる。

The Generalized Cross Correlation Function (GCC) is often used for time difference estimation.

■GCC

通常の相互相関関数:

The ordinary cross correlation function:

$$r_{\tau,\mathrm{ORG}} = \mathrm{E}[\ x_n\, y_{n+\tau}\ ] = \mathrm{IDFT}[\ \overline{X_k\, Y^*_k}\ ]$$

GCCは、重み関数によって特性が変わる。

GCC characteristics differs depending on the weighting function.

$$r_{\tau,\mathrm{GCC}} = \mathrm{IDFT}[\ \overline{X_k\, Y^*_k}\, \Psi_k] \qquad \Psi_k : \text{重み関数}$$

$\Psi_k = 1/|X_k\, Y^*_k|$ のとき　:　Phase Transfer（PHAT）

$\Psi_k = |\overline{X_k\, Y^*_k}|^2 / |\overline{X_k}|^2 |\overline{Y_k}|^2$　: (2乗コヒーレンス関数）ML 最適重み

$\Psi_k = 1/|\overline{X_k - Y^*_k}|^2$ のとき　:　差信号GCC

8-2. ビームフォーマ法(beamformer method)

ビームフォーマの指向性を回転させ、メインローブ（＝拘束方向）が音源方向と一致したときにビームフォーマ出力が大きくなることを利用する。
Exploiting the characteristics of the beamformer output which becomes large when the directivity of the beamformer mainlobe is rotated to the DOA.

メインローブの方向を look-direction と呼ぶ。
The direction of the beamforemr's mainlobe is called as “look-direction”.

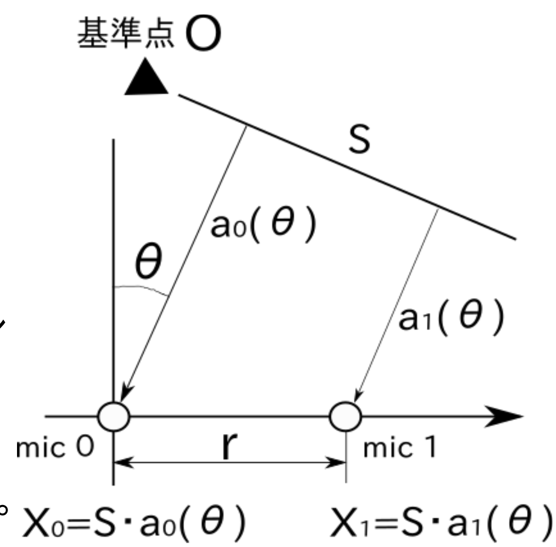
方向とビームフォーマ出力の対応を空間スペクトルと呼ぶ。
The beamformer output as the function of the direction is called “**spatial spectrum**”.

空間スペクトル上のピークにより音源方向を判定する。
Peaks in the spatial spectrum indicate DOAs.

以降は周波数の番号 k を省略する。
frequency number *k* is omitted in the description after here.

■ステアリングベクトル／遅延ベクトル
steering vector／delay vector

信号は複数のマイクロホンに時間差（ときに振幅差）を伴って到達するが、ある基準点で観測される信号値から各マイクロホンへの伝達関数を
Signals arrive at microphones with time delays. Let the transfer function between the observed signal at some fixed point O and that at each microphone as



$$a_m(\theta) = A_m e^{j 2\pi f \tau_m} \quad (m = 0, ..., M\text{-}1),$$

とし、全マイクロホンについてこれを並べたベクトル
and the vector consist of the above elements

$$\boldsymbol{a}(\theta) = \{ a_0(\theta), a_1(\theta), \ldots , a_{M-1}(\theta) \}$$

を、ステアリングベクトルまたは遅延ベクトルとよぶ。ここで、A と τ は伝達に伴う振幅比と遅延時間である。
is called as steering vector or delay vector. A is amplitude ratio and τ is time delay according to the signal propagation.

ステアリングベクトルは、観測基準点、マイクロホン配置、方向 θ によって変化する関数である。例えば、基準点を上の図の mic1 にしたとき、ステアリングベクトルは、
Steering vector is a function of observation reference point, microphone arrangement and signal arriving direction θ. For an example, the steering vector shown in the above figure with the reference point of microphone 1 is

$$\{ a_0(\theta), a_1(\theta) \} = \{ \exp(-j\ 2\pi f \cdot r \sin(\theta)/c), 1 \}$$

である。ここで、$r\sin(\theta)/c$ はmic0,1間の到達時間差である。
where, $r\sin(\theta)/c$ is arrival time difference between microphone 0 and 1.

ステアリングベクトルは、到来方向に応じた観測信号を計算するために用いることができる。例えば、基準点の観測信号Sにステアリングベクトル $\boldsymbol{a}(\theta)$ をかけた
Steering vector can be used to calculate observed signal depending on the DOA. For example, multiplication of reference signal S with the steering vector $\boldsymbol{a}(\theta)$ results the below vector

$$\{Sa_0(\theta),\ Sa_1(\theta),\ \ldots,\ Sa_{M-1}(\theta)\}^T$$

は、信号 S が θ 方向から到来したときの観測信号ベクトルである。
which is signal vector observed when the signal S arrives from the angle θ.

また、ステアリングベクトル成分各々の逆数を並べたベクトル
And the vector consist of inverse of the steering vector component

$$\boldsymbol{a}^{-1}(\theta) = \{\ a^{-1}{}_0(\theta),\ a^{-1}{}_1(\theta),\ \ldots,\ a^{-1}{}_{M-1}(\theta)\ \}$$

を補正ベクトルと呼ぶこともある。
is often called "compensation vector".

遅延ベクトルや補正ベクトルは周波数で異なるので、周波数ごとに独立した処理を行うのが普通である。
Because the steering vector and/or the compensation vector differs by frequency, the processing is generally performed at each frequency.

■固定ビームフォーマによる方法
DOA estimation by using fixed beamformer

遅延和アレーなどの固定ビームフォーマの指向性を回転させる。
The directivity of a fix beamformer , e.g., DAS beamformer is rotated.

遅延和ビームフォーマの出力は、look-direction θ と、θ に対応した補正ベクトル $\boldsymbol{a}^{-1}(\theta)$ を使って
The output of DAS beamformer is

$$Y_{DS}\ (\theta) = \Sigma_{m=0,M-1}\ a^{-1}{}_m\ (\theta)\ X_m = \boldsymbol{a}^{-1}(\theta)\cdot \boldsymbol{X}$$

と表される。
using look-direction θ and compensation vector $\boldsymbol{a}^{-1}(\theta)$.

θが真の到来方向である場合、$\boldsymbol{a}^{-1}(\theta)$によって各chの時間遅れを補正すると観測信号ベクトルの位相が揃うため、和のパワーが極大となる。したがって、
When θ is true arrival direction, if delay time of the each channel is unified by the compensation vector, the power of the summed signal is maximized. Therefore, plot of

$$|Y_{DS}\ (\theta)|^2 = |\boldsymbol{a}^{-1}(\theta)\cdot \boldsymbol{X}\ |^2$$

を空間スペクトルとしてプロットすれば、音源方向に極大が現れることになる。
shows a peak at the arrival direction.

雑音抑圧処理と同様、高性能化（高分解能化）には多数のマイクロホンが必要である。
Same as noise reduction, many microphones are necessary to obtain high performance (high resolution).

■高分解能法

High resolution method

スペクトル推定で出てきたパラメトリック法、あるいは、多チャネル雑音抑圧の Null ステアリングビームフォーマと同様の処理原理。

The processing principle of the high resolution method is the same as that in parametric spectrum estimation or null steering beamformer.

入力信号をチャネルご とにフィルタリングし、その和を出力する処理を仮定する。

It is assumed that the input signal is filtered by each channel and summed signal over channels is derived.

出力最小化フィルタにより、look direction 以外の音源方向に指向性の谷を形成する。指向性は、look-direction ごとに異なるので、look-direction ごとに計算する。

Directivity nulls in the sources' directions excluding the look-direction are produced by the filter which minimize the output power. Because the directivity differs depending on the look-direction, the filter is calculated by changing the look-direction.

少数のマイクロホンで高分解能を達成することが可能であり、代表的な手法として、最小分散法(MV法)とMUSIC法がある。

High resolution DOA can be attained with small number of microphones. Minimum Variance method (MV) and MUSIC are the representative methods.

・ <u>最小分散法**(MV)**</u>

<u>Minimum Variance method (MV)</u>

周波数 k について、m チャネル目の信号のDFTを X_m とし、それらの線形和

Let X_m as the DFT of m-th channel signal at frequency number k, and consider multi-channel FIR filtering to produce the weighted sum over X_m.

$$Y = h^*_0 X_0 + h^*_1 X_1 + \dots + h^*_{M-1} X_{M-1} = \Sigma_{m=0,M-1}\, h_m X_m = \boldsymbol{h}^{\mathrm{H}} \boldsymbol{X}$$

$$\boldsymbol{X} = \{X_0, X_1, \dots, X_{M-1}\}^{\mathrm{T}}$$

$$\boldsymbol{h} = \{h_0, h_1, \dots, h_{M-1}\}^{\mathrm{T}}$$

を出力する多chフィルタ処理を考える。このとき、look-direction θ からの到来音を変化させずに出力信号パワー $\overline{|Y|^2}$ を最小化するようなフィルタ $\boldsymbol{h}_{\mathrm{MV}}$ を θ ごとに求め、$\boldsymbol{h}_{\mathrm{MV}}$ を使ったときの出力パワー $\overline{|Y|^2}$ をその θ における空間スペクトルとする。

Then, we calculate filter $\boldsymbol{h}_{\mathrm{MV}}$ which minimizes output power $\overline{|Y|^2}$ while maintaining the amplitude of signal arrives from look-direction θ to be unaltered, and let the $\overline{|Y|^2}$ as the value of spatial spectrum at direction θ.

フィルタ出力パワー $\rho(\theta)$ は、

Filter output power $\rho(\theta)$ is,

$$\rho(\theta) = \overline{|Y(\theta)|^2} = \overline{|\boldsymbol{h}^{\mathrm{H}}(\theta)\boldsymbol{X}|^2} = \boldsymbol{h}^{\mathrm{H}}(\theta)\overline{\boldsymbol{X}\boldsymbol{X}^{\mathrm{H}}}\boldsymbol{h}(\theta) = \boldsymbol{h}^{\mathrm{H}}(\theta)\boldsymbol{R}\,\boldsymbol{h}(\theta)$$

$$\boldsymbol{R} = \boldsymbol{X}\,\boldsymbol{X}^{\mathrm{H}} = \begin{bmatrix} \overline{X_0 X^*_0} & \cdots & \overline{X_0 X^*_{M-1}} \\ \cdots & \cdots & \cdots \\ \overline{X_{M-1} X^*_0} & \cdots & \overline{X_{M-1} X^*_{M-1}} \end{bmatrix}$$

と表せる。ここで、$\boldsymbol{R}$ は周波数 k についての空間的な相関行列である。空間的とは、DFTに関してチャネル間の積を行列要素とすることを意味する。（スペクトル推定のときは、サンプル間の積である。）

where, $\boldsymbol{R}$ is the spatial correlation matrix in frequency number k. “spatial” means that products of DFTs in different and/or same channels are used as the matrix components. (products of different and/or same samples in the same channel are used in spectrum estimation)

方向 θ から音が到来したとして、基準点における到来音ス ペクトルを Xo とし、θ 方向のステアリングベクトル $\boldsymbol{\alpha}(\theta)$ を用いると、$\boldsymbol{X} = \boldsymbol{\alpha}(\theta)X$o であるので、$\boldsymbol{h}^{\mathrm{H}}(\theta)$によるフィルタリングによって Xo を変化させずに出力する条件は、

Let the signal arrive from direction θ, and let the spectrum of arrived signal at reference point as Xo. Then, using steering vector $\boldsymbol{\alpha}(\theta)$ for direction θ, the condition to maintain Xo as unaltered through the filtering by $\boldsymbol{h}^{\mathrm{H}}(\theta)$ is obtained from the below relation.

$$Y = \boldsymbol{h}^{\mathrm{H}}(\theta)\boldsymbol{X} = \boldsymbol{h}^{\mathrm{H}}(\theta)\boldsymbol{\alpha}(\theta)\, X\mathrm{o} = X\mathrm{o}\text{、}$$

より、

and is

$$\boldsymbol{h}^{\mathrm{H}}(\theta)\boldsymbol{\alpha}(\theta) = 1$$

である。

この拘束条件の下で $\rho(\theta)$を最小化するような $\boldsymbol{h}(\theta)$を求めて $\rho(\theta)$ を計算すれば、それが最小分散法のスペクトル（θ におけるスペクトル値）である。

The filter which minimizes $\rho(\theta)$ under this constrain gives the value of minimum variance spectrum at θ.

スペクトル推定のときと全く同じ拘束条件付き最小化問題であるので、ラグランジュの未定係数法を用いて解く。

The same constrained minimization as in the MV spectrum estimation, the optimum filter is obtained to apply Lagrange multiplier method.

・スペクトル推定との対応

(correspondence between spectrum estimation and spatial spectrum estimation)

スペクトル推定			空間スペクトル推定（周波数ごと）
正弦波ベクトル	$\boldsymbol{e}(f)$	$\boldsymbol{\alpha}(\theta)$	ステアリングベクトル
自己相関行列	$\boldsymbol{R}_{\mathrm{p}}$	$\boldsymbol{R}$	空間相関行列
フィルタ	$\boldsymbol{a}$	$\boldsymbol{h}(\theta)$	フィルタ

ラグラジアン、

Lagrangean

$$L(\boldsymbol{h}, \lambda) = \boldsymbol{h}^{\mathrm{H}}(\theta)\boldsymbol{R}\,\boldsymbol{h}(\theta) - \lambda\,(\boldsymbol{h}^{H}(\theta)\,\boldsymbol{\alpha}(\theta) - 1),$$

をで微分して0とおいて解くと、

is partially differenciated by $\boldsymbol{h}^H$ and λ, then equated to 0, we obtain the optimum filter

$$\boldsymbol{h}(\theta)_{MV} = \lambda \boldsymbol{R}^{-1}\boldsymbol{a}(\theta) = \boldsymbol{R}^{-1}\boldsymbol{a}(\theta) / (\boldsymbol{a}(\theta)^H \boldsymbol{R}^{-1}\boldsymbol{a}(\theta))$$

となり、これを $\rho(\theta) = \boldsymbol{h}(\theta)^{\mathbf{H}} \boldsymbol{R}\, \boldsymbol{h}(\theta)$ に代入してパワー $\rho(\theta)$ を求めると、
$\boldsymbol{h}(\theta)$ in the above expression is substituted with $\boldsymbol{h}(\theta)_{MV}$, and the filter output power $\rho(\theta)$ is obtained as,

$$\rho(\theta)_{MV} = 1 / (\boldsymbol{a}(\theta)^H \boldsymbol{R}^{-1}\boldsymbol{a}(\theta)).$$

となり、これが最小分散法による空間スペクトルである。
This is the MV spatial spectrum.

･MUSIC法

スペクトル推定のときと同様、最小分散法における拘束条件 $\boldsymbol{h}^H(\theta)\boldsymbol{a}(\theta) = 1$ をフィルタの大きさを1に保つ拘束条件
As in the case of spectrum estimation, the constrain of MV method $\boldsymbol{h}^H(\theta)\boldsymbol{a}(\theta) = 1$ is replaced with that of filter norm constraint as shown below.

$$|\boldsymbol{h}^H(\theta)| = \boldsymbol{h}^H(\theta)\boldsymbol{h}(\theta) = 1$$

に置き換える。
この結果、拘束条件付き最小化は、
As the result, the constrained minimization problem turns to the eigenvalue problem.

$$\boldsymbol{R}\, \boldsymbol{h}(\theta) = \lambda \boldsymbol{h}(\theta)$$

の固有値問題になる。
固有値展開を行って得られる $\boldsymbol{R}$ の固有ベクトルが求めるフィルタ係数である。
The eigen vectors obtained from eigen decomposition of R are the desired filters.

$\boldsymbol{R}$ の固有ベクトルからなる行列を $\boldsymbol{U}$ 、固有値を対角成分にもつ対角行列を $\boldsymbol{\Lambda}$ とすると、
Let the matrix consist of eigenvectors of $\boldsymbol{R}$ as $\boldsymbol{U}$, the diagonal matrix where the diagonal components are the eigenvalues of $\boldsymbol{R}$ as $\boldsymbol{\Lambda}$,

$$\boldsymbol{R} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H$$

$$\boldsymbol{U} = [\boldsymbol{u}_0, \boldsymbol{u}_1, \dots, \boldsymbol{u}_{M-1}]^T$$

$$\boldsymbol{\Lambda} = \mathrm{diag}[\lambda_0, \lambda_1, \dots, \lambda_{M-1}]$$ （$\lambda_0 > \lambda_1 > \dots, > \lambda_{M-1}$ とする）

である。固有値は降順にソートされているものとする。
The eigenvalues are assumed to be sorted in descending order.

ここで、最小固有値に対応した固有ベクトル $\boldsymbol{u}_{M-1}$ を用いて次式を計算したものが、最小ノルム法と呼ばれる方法である。
The minimum norm method as shown below calculates the spatial spectrum using the eigenvector corresponding to the minimum eigenvalue $\boldsymbol{u}_{M-1}$.

$$S(\theta)_{min} = 1 / |\boldsymbol{u}_{M-1}^H\, \boldsymbol{a}(\theta)|^2$$

分母は$\boldsymbol{u}_{M-1}$をフィルタとしたときの出力パワーであり、その逆数をとる演算によって、音源方向に鋭い正のピークが生じる。
The denominator in the above expression is the power of filtering with filter $\boldsymbol{u}_{M-1}$.

$\boldsymbol{R}$ の固有値は,通常,背景雑音に対応するものと、到来 信号に対応するものに分けることができ、雑音と信号に対 応する固有ベクトルの張る空間を、各々、雑音空間、信号 空間とよぶ。MUSIC 法は、$\boldsymbol{u}_m$ のうち、雑音空間に属する 複数の固有ベクトルを用いて上式の分母を平均化する。
Eigenvalues of R are generally categorized into two groups where one corresponding to noise and the other corresponding to signal. The subspace spanned by the eigenvectors for signal is called “signal subspace”, and the other is called “noise subspace”. MUSIC method averages the inverse of the filtered power in the above expression over the multiple eigenvectors belong to noise subspace among $\boldsymbol{u}_m$ as shown below.

$$S(\theta)_{\mathrm{MUSIC}} = 1 / \Sigma_{i=m,,M-1}| \boldsymbol{u}^H_i \boldsymbol{a}(\theta)|^2$$

$\boldsymbol{a}(\theta)$ が信号到来方向に対応したステアリングベクトルであるとき、$\boldsymbol{a}(\theta)$ は信号空間に含まれ、また、信号空間は雑音空間を張る固有ベクトルすべてと直交するため、$\boldsymbol{a}(\theta)$ と雑音空間の固有ベクトルとは直交し、分母のフィルタ出力パワーはそれらの内積であるため小さい値となる。これにより、$S(\theta)_{\mathrm{MUSIC}}$ 上に到来方向において鋭いピークが生じる。
If $\boldsymbol{a}(\theta)$ is a steering vector corresponding to the arrival direction, $\boldsymbol{a}(\theta)$ belongs to the signal subspace and is diagonal to all the eigenvectors spanning noise subspace. Therefore, $\boldsymbol{a}(\theta)$ is diagonal to the eigenvectors spanning noise subspace and the denominator in the above expression becomes small value. This results the sharp peaks at direction θ in $S(\theta)_{\mathrm{MUSIC}}$.