信号処理論第一
第1回：はじめに

東京大学 大学院情報理工学系研究科 システム情報学専攻
小山 翔一

講義の目的と概要

- 講義の目的
  - 信号処理の基礎を習得する。

- 講義の概要
  - 決定論的信号（確定的信号）に関する信号処理について，基礎事項を理解する。本講義では，できるだけ特定の応用分野に限らない，一般的な信号処理論について講義する。ただし，統計的信号処理については扱わず，信号処理論第二の範囲となる。

- キーワード
  - Fourier級数, Fourier変換, 離散時間Fourier変換, 離散Fourier変換, FFT, 窓Fourier変換, 線形時不変システム, サンプリング定理, インパルス応答, 伝達関数, 周波数応答, 差分方程式, ブロック線図, 因果性, 安定性, FIRフィルタ, IIRフィルタ

講義スケジュール

- Sセメスター 金曜１限 @~~63講義室~~ zoom

- 日程（暫定版）
  - 4/10　第１回
  - 4/17　第２回
  - 4/24　第３回
  - 5/ 1　第４回
  - 5/ 8　**休講**
  - 5/15　第５回
  - 5/29　**補講** - 第６回
  - 6/ 5　第７回
  - 6/12　第８回
  - 6/19　第９回
  - 6/26　第１０回
  - 7/3　第１１回
  - 7/10　第１２回
  - 7/17　学期末試験 (?)

オンライン講義の実施方法

- zoom（https://zoom.us/ ）を用いる。設定方法等については ポータルサイト（ https://utelecon.github.io/oc/ ）を参照すること。
- 講義のURLは毎回同じにする予定だが，念のため毎回UTASをチェックすること。
- 質問がある場合は手を挙げるボタンを押すこと。タイミングをみてこちらからマイクのミュートを解除する。
- チャットにも適宜質問やコメントを書いてもらって構わない。ただし，こちらからはリアルタイムで内容を追うことは難しいと思われる。
- もしオンラインでの受講に不都合が生じた場合は，メール等で連絡すること。

shoichi_koyama@ipc.i.u-tokyo.ac.jp

講義資料と成績評価

- 講義資料
  - https://www.sh01.org/ja/teaching/
  - システム１研のウェブサイトからたどることもできる
  - できるだけ講義前日までに資料をアップロードするようにします。

- 成績評価
  - 学期末試験（変更の可能性が高いが暫定として）

参考文献

- 眞溪歩, "ディジタル信号処理工学," 昭晃堂, 2004.
  - 信号処理の基礎事項を学ぶ上で良書。残念ながら廃刊。
- 原島博, “信号解析教科書 ー信号とシステムー,” コロナ社, 2018.
  - 現在入手可能な，決定論的信号処理の基礎事項に関する教科書。
- A. V. Oppenheim and R. W. Schafer, "Digital Signal Processing," Pearson, 1975.
  - 信号処理教科書の古典。内容としては古めだが，基本的な内容は網羅されている。
- M. Vetterli, J. Kovačević, and V. K. Goyal, "Foundations of Signal Processing," Cambridge University Press, 2014.
  - 関数解析的な視点からの信号処理の教科書。無料でダウンロード可能。

講義の全体像

1. 数学的準備
2. Fourier級数展開
3. Fourier変換
4. サンプリング定理
5. 離散Fourier変換
6. z変換
7. 離散時間線形時不変システム
8. フィルタの設計

本日の目次

- 本格的に講義に入る前に，信号処理とはどのようなものかイメージを持ってもらうため，信号処理の概要や応用例などをいくつか紹介する。
  - 講義では，できるだけ特定の応用範囲に限定されないように，一般的な信号を対象として説明を行うが，今回は私の専門分野の関係上，音声・音響信号に対する応用の話が多い。
  - 計数工学科システム１研のウェブサイト：
    http://www.sp.ipc.i.u-tokyo.ac.jp/

信号処理とは？

➢ IEEE Signal Processing Societyによる信号処理の紹介動画

https://www.youtube.com/watch?v=EErkgr1MWw0

信号処理とは？

- IEEE (米国電気電子学会) は，アメリカ合衆国に本部を置く，電気・情報工学分野における世界最大規模の学術研究団体
- その中で，信号処理分野のコミュニティであるSignal Processing Societyは最初 (1948年) に設立され，現在は4番目に大きいソサイエティ（Computer, Power and Energy, Communications, Signal Processing)

https://signalprocessingsociety.org/

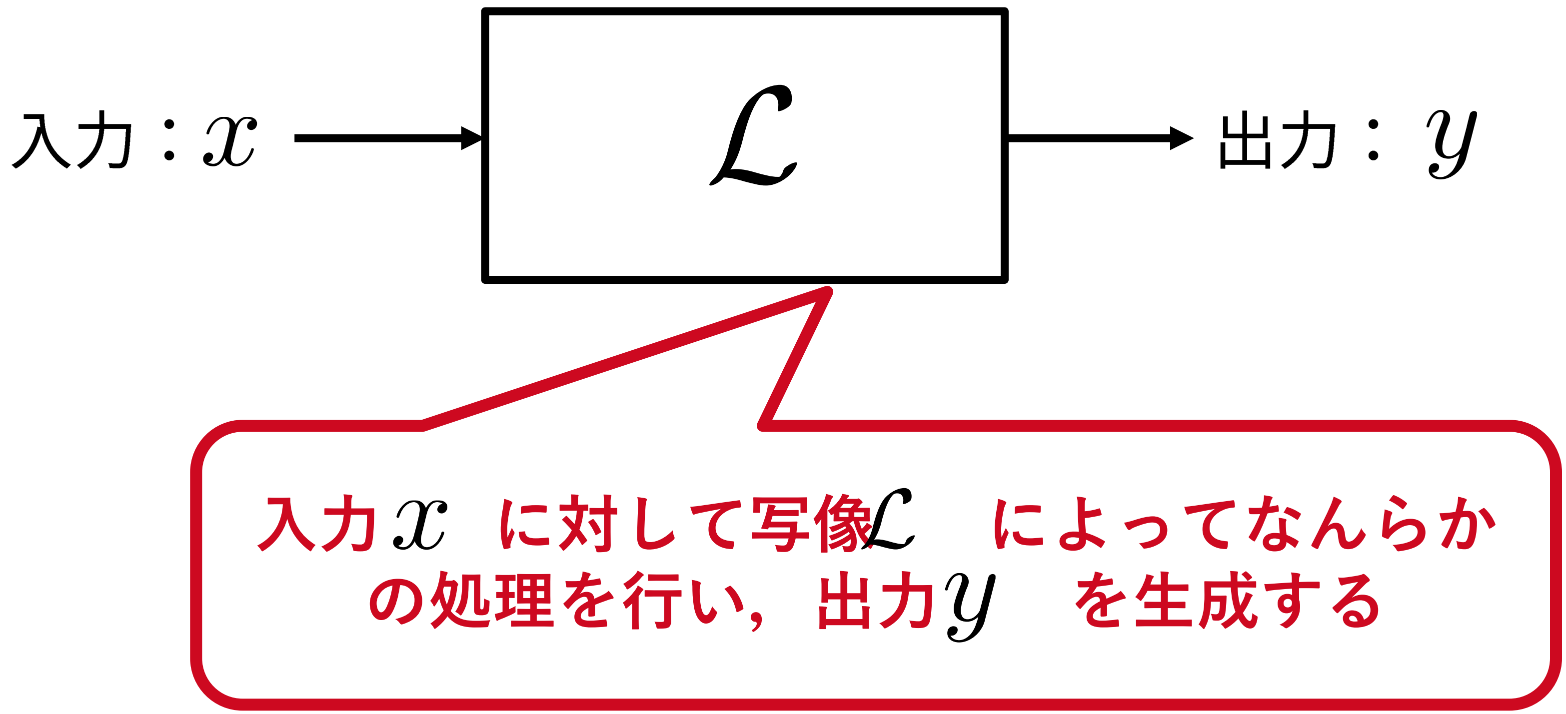
信号処理とは？

- IEEE Signal Processing Society – Technical Committees
  - Audio and Acoustic Signal Processing
  - Bio Imaging and Signal Processing
  - Computational Imaging
  - Design and Implementation of Signal Processing Systems
  - Image, Video, and Multidimensional Signal Processing
  - Industry DSP Technology
  - Information Forensics and Security
  - Machine Learning for Signal Processing
  - Multimedia Signal Processing
  - Sensor Array and Multichannel
  - Signal Processing for Communications and Networking
  - Signal Processing Theory and Methods
  - Speech and Language Processing

応用分野は多岐にわたる

信号処理とは？

- 音声や画像などの信号を，数理的な手法で分析／加工／合成する技術



信号処理技術の例

- 雑音を含む音声を入力し，雑音を除去した音声を出力

➡ 雑音除去

雑音を含む
音声信号

$$\mathcal{L}$$

雑音が除去された
音声信号

音声や雑音の性質をうまく利用して音声のみを抽出するような処理を行う

信号処理技術の例

- 音声を入力し，その内容をテキストとして出力

➡ 音声認識

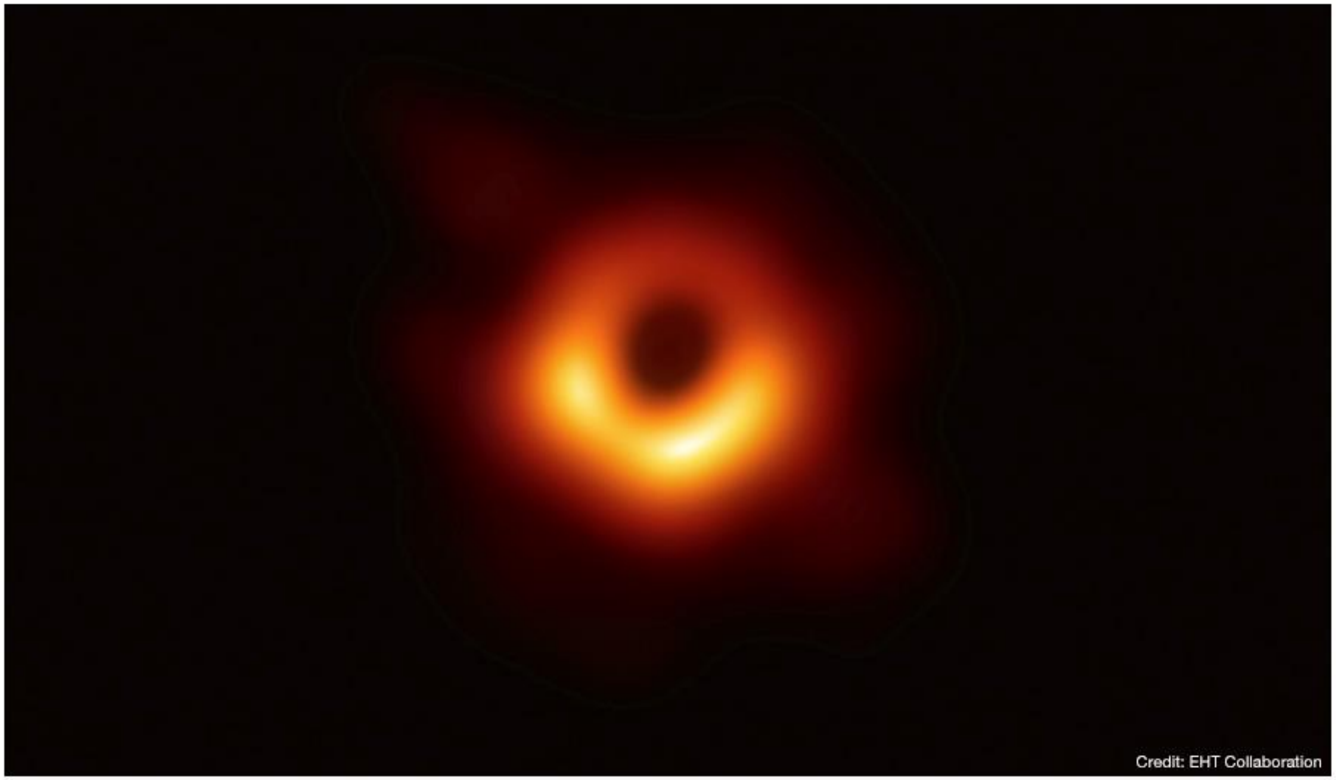
音声信号 → $\mathcal{L}$ → テキスト

音声信号のパターンに対して，対応する
テキストを出力するような関数を教師
データを用いて学習する

信号処理技術の例

史上初、ブラックホールの撮影に成功 — 地球サイズの電波望遠鏡で、楕円銀河M87に潜む巨大ブラックホールに迫る

2019年4月10日 | 研究成果

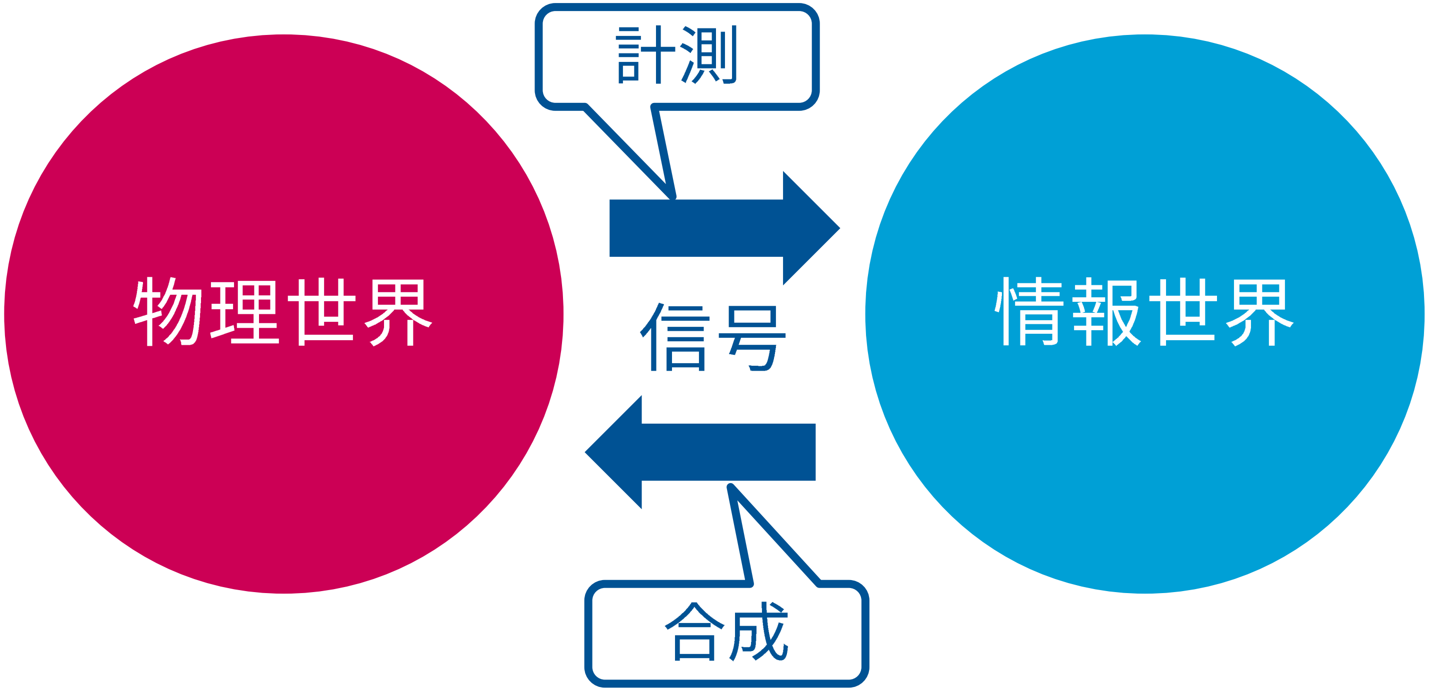


イベント・ホライズン・テレスコープで撮影された、銀河M87中心の巨大ブラックホールシャドウ。リング状の明るい部分の大きさはおよそ42マイクロ秒角であり、月面に置いた野球のボールを地球から見た時の大きさに相当します。（Credit: EHT Collaboration）オリジナルサイズ（643KB）

https://www.nao.ac.jp/news/science/2019/20190410-eht.html

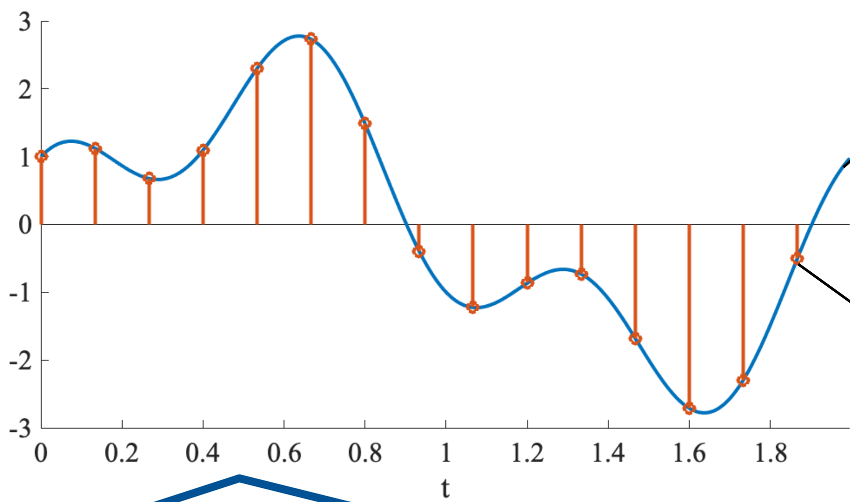
信号

- 信号とはセンサ等で計測した物理量の時間的／空間的な変動，あるいはそれを記号として表したもの
  - 音声，音楽，画像，動画，超音波ソナー，電波，脳波，筋電位，地震波，株価 etc…



信号

- 本講義では特に時間とともに振幅が変動する一次元信号である，時系列信号を主に扱う。
  - 連続時間信号／アナログ信号：
    時間・振幅両方に関して連続な値を取る信号の表現。
  - 離散時間信号／デジタル信号：
    コンピュータで扱うため，離散化された時刻における量子化された振幅値によって表現。



アナログ信号 $x(t)$ ➡ 関数

デジタル信号 $x[n]$ ➡ 数列／ベクトル

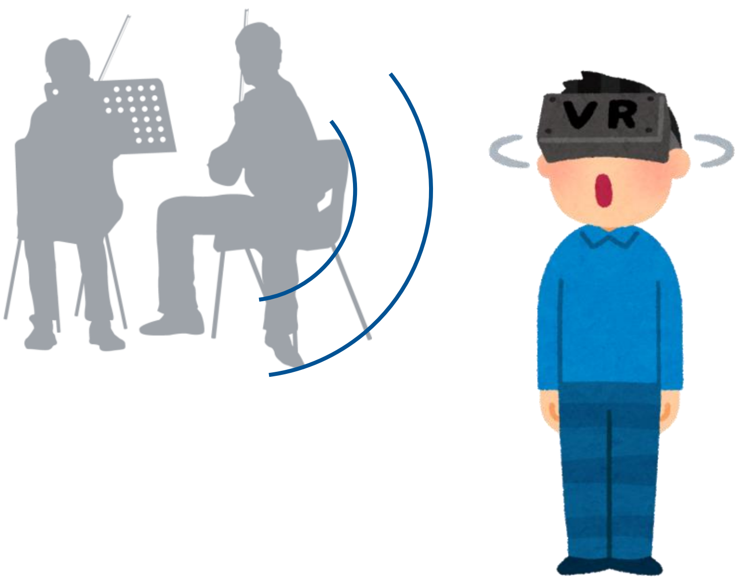
信号処理理論は広範囲な数理的基盤の上に成り立っている

応用例：バイノーラル合成

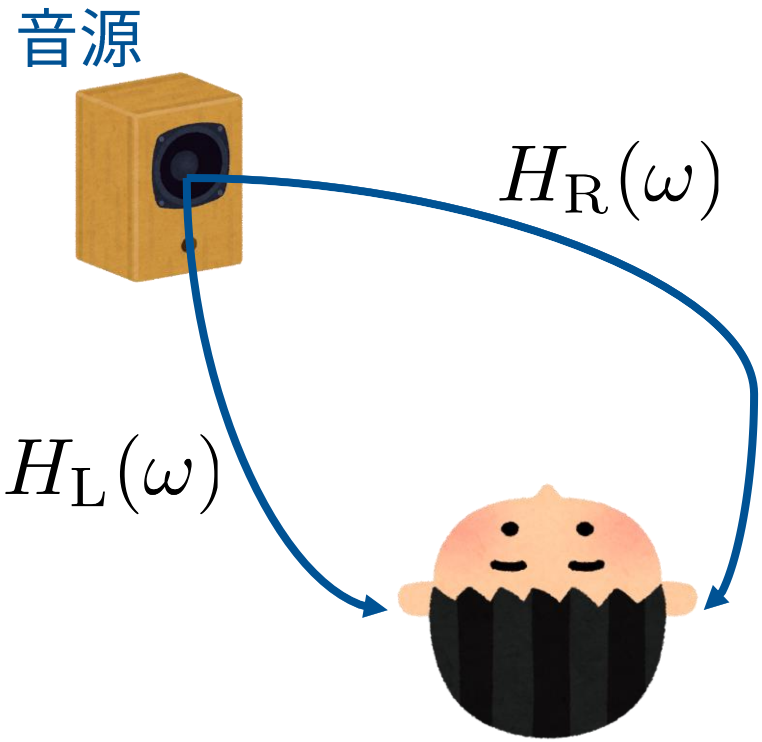
- ヘッドフォンで受聴するVRオーディオシステムを実現するには？
  1. 両耳位置での音をそのまま録音する方法。

     ➡ バイノーラルマイクなどによる収音
  2. あらかじめ用意した音源から両耳までの伝達特性を用いる方法。

     ➡ バイノーラル合成



- 頭部伝達関数（HRTF）
  - 音源から放射された音がヒトの鼓膜に到達するまでの伝達特性であり，音の空間知覚に重要な役割を持つ。



音源から外耳道入口/鼓膜までの伝達特性

ある音源位置から右耳までのHRTF

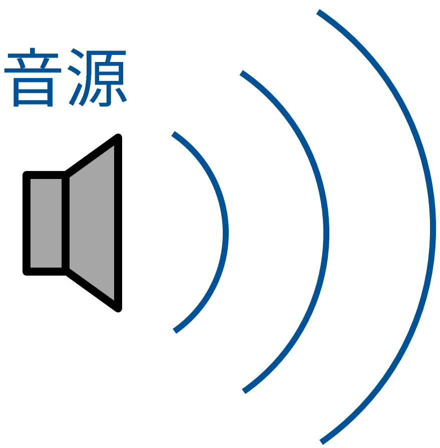
$$\mathrm{HRTF}_{\mathrm{R}}(\omega) = \frac{H_{\mathrm{R}}(\omega)}{H_0(\omega)}$$

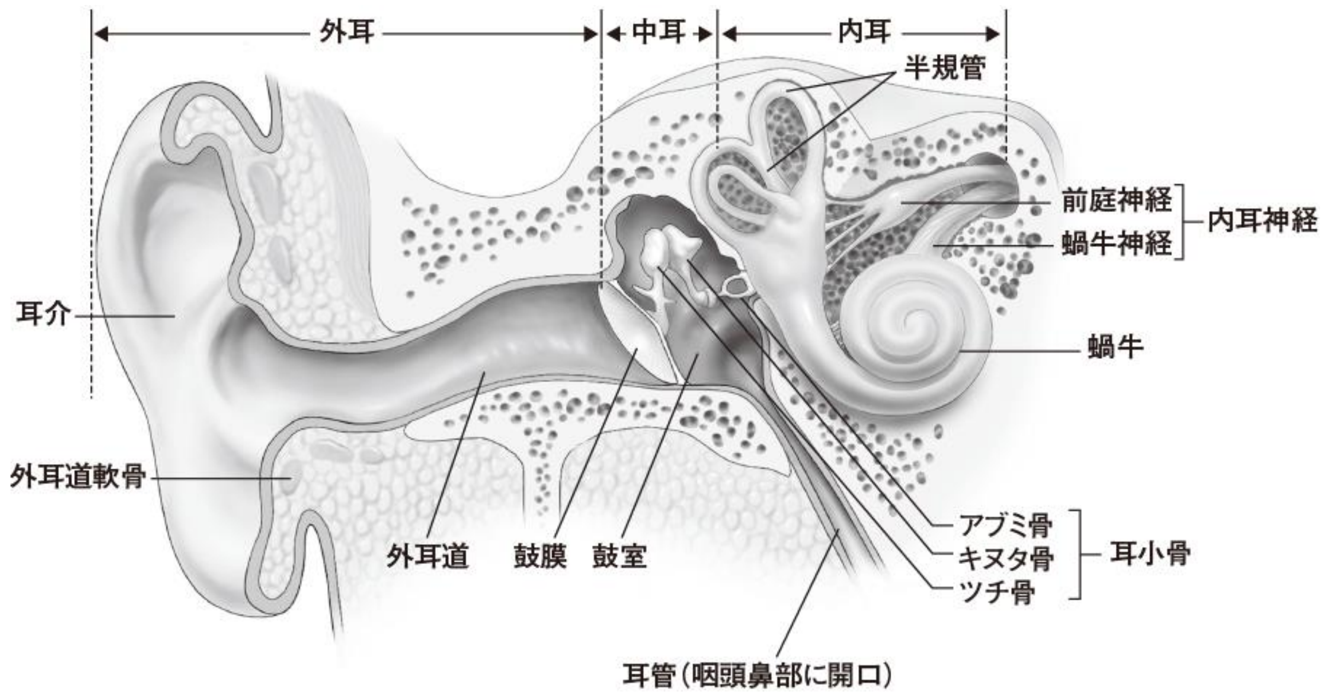
受聴者がいない状態での頭部中心までの伝達特性

音の空間知覚

- 頭部伝達関数（HRTF）
  - 音源から放射された音がヒトの鼓膜に到達するまでの伝達特性であり，音の空間知覚に重要な役割を持つ。

音源



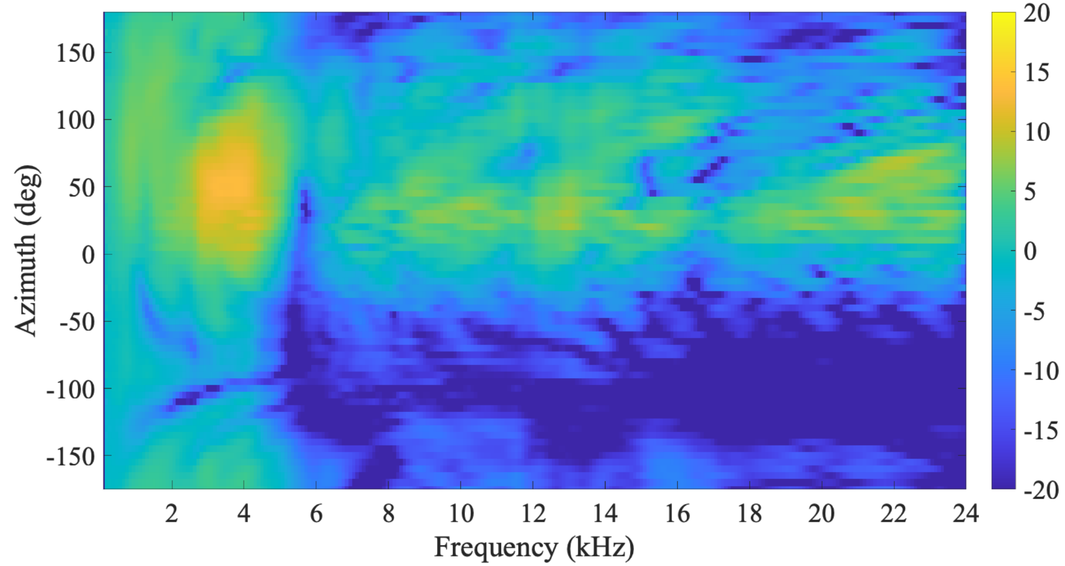


https://www.kango-roo.com/sn/k/view/1720

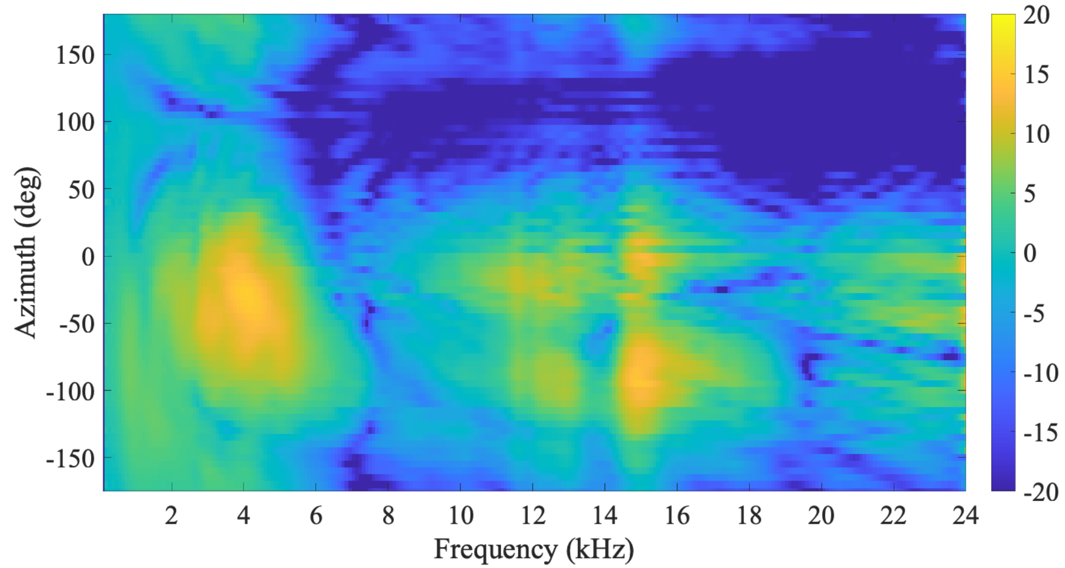
➡ HRTFのどの要素が音の空間知覚をもたらしている？

- HRTFの例

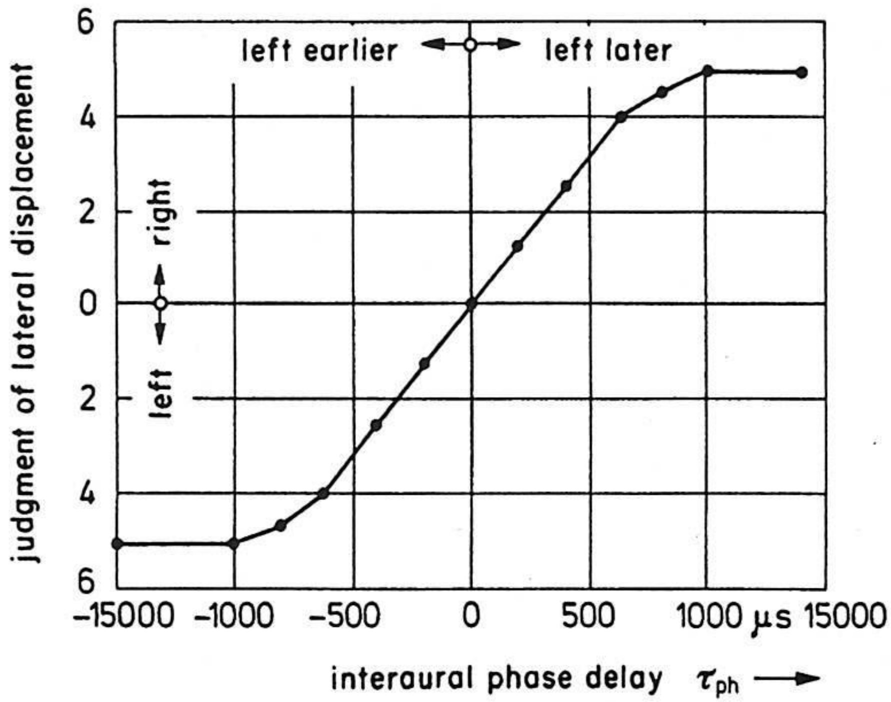
左耳


150
100
50
0
-50
-100
-150
Azimuth (deg)
2
4
6
8
10
12
14
16
18
20
22
24
Frequency (kHz)
20
15
10
5
0
-5
-10
-15
-20

右耳


150
100
50
0
-50
-100
-150
Azimuth (deg)
2
4
6
8
10
12
14
16
18
20
22
24
Frequency (kHz)
20
15
10
5
0
-5
-10
-15
-20

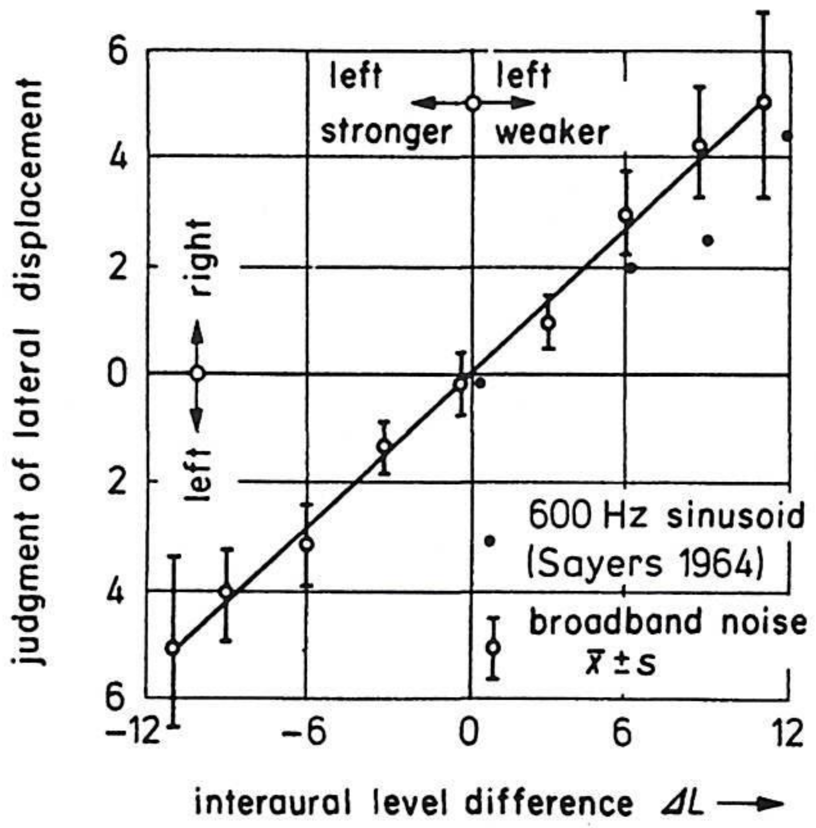
- 左右方向の知覚
  - 両耳間時間差（interaural time difference: ITD）
    - 左右の耳に到達する音の時間差



[Blauert 1997]

  - ITDが手がかりになっているのは約1.6kHz以下で，それ以上では両耳入力信号の包絡線の時間差が手がかり

- 左右方向の知覚
  - 両耳間レベル差（interaural level difference: **ILD**）
    - 左右の耳に到達する音の音圧差

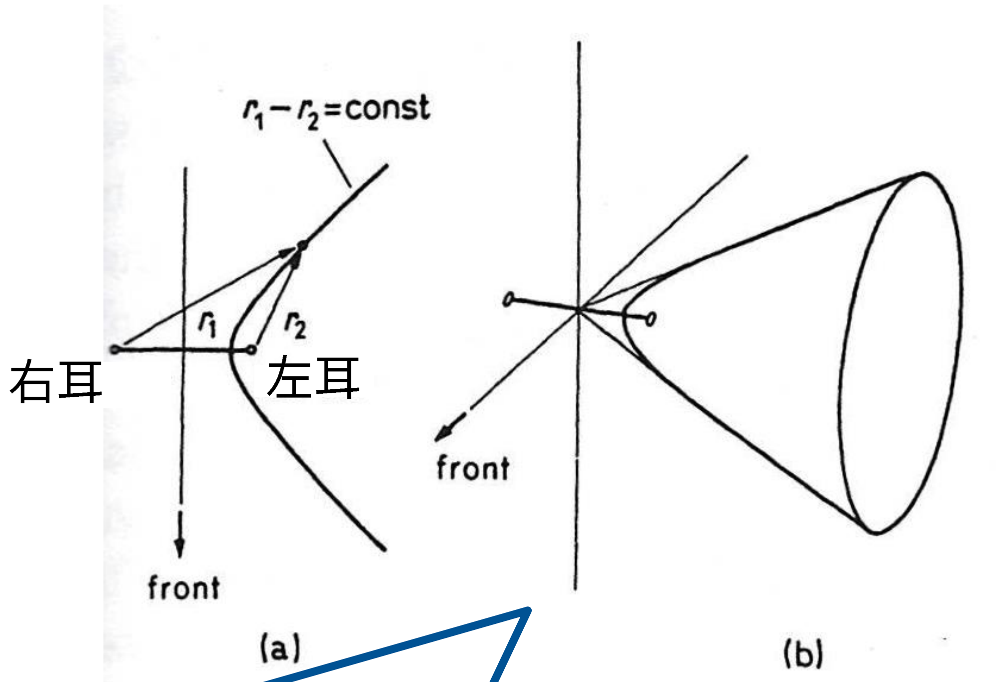


[Blauert 1997]

音の空間知覚

➢ コーン状の混同（cone of confusion）

– 頭部や外耳が前後・上下で対象な形状だと仮定すると，音源から両耳までの距離の差は円錐台上で一定となる。

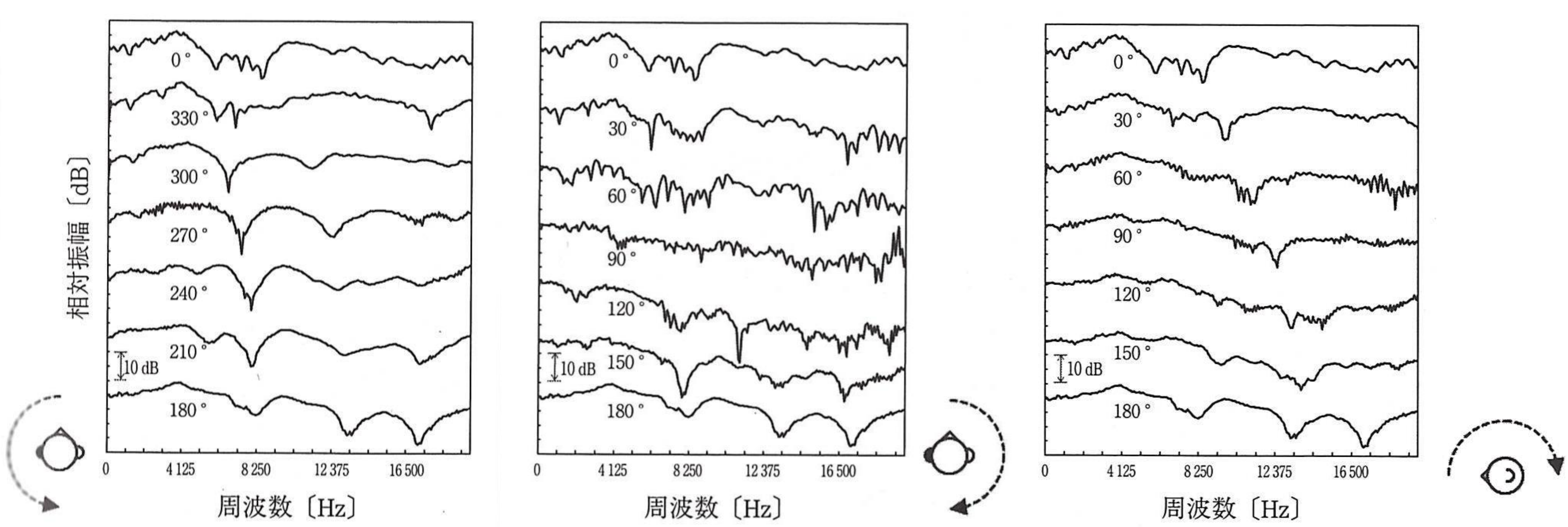


[Blauert 1997]

ITD・ILDだけでは前後・上下方向を同定することはできない

➢ 前後・上下方向の知覚

– HRTFの振幅スペクトルが重要な役割を果たしており，これをスペクトラルキュー（spectral cue）と呼ぶ。



[飯田+ 2010]

音源の方向によってHRTFの振幅スペクトルは大きく変化

バイノーラル合成

- ➢ 音の空間知覚において，HRTFが重要な役割を持つ
  - ➡ HRTFを直接用いて両耳の信号を合成すればよい
- ➢ HRTFを線形時不変システムとしてみなす

- – 任意の音源信号に対して，あらかじめ測定したHRTFを畳み込むことで，そのHRTFを測定した音源位置からのバイノーラル信号を合成することができる

頭部伝達関数の測定

- 任意のバイノーラル信号を合成するには，あらゆる方向からのHRTFを測定しておくことが必要

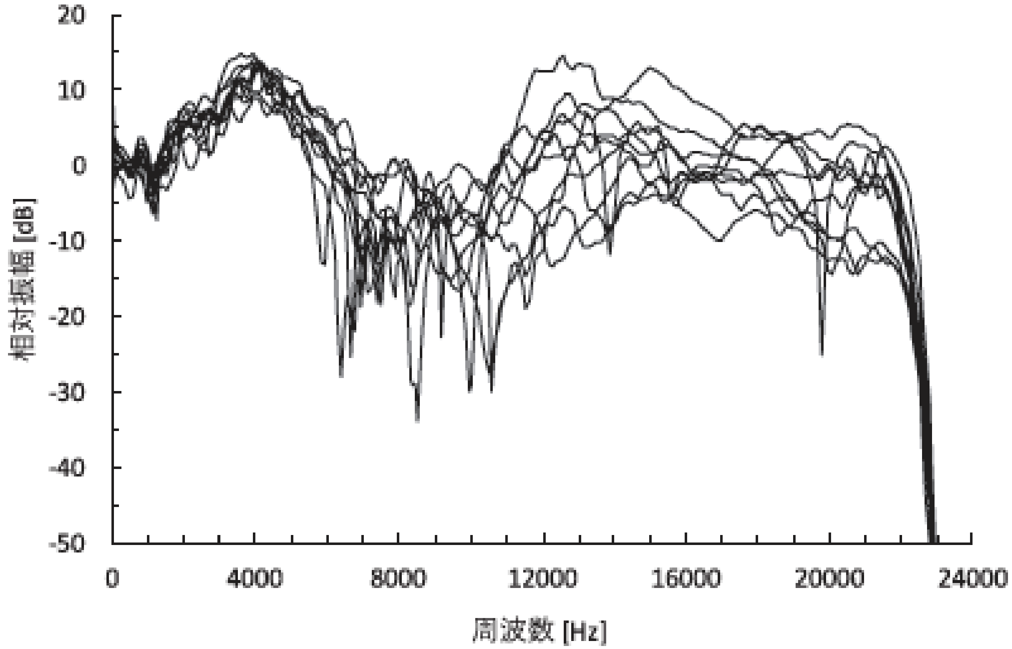
> - 無響室などで外耳道入り口に設置したマイクを用いて各スピーカからのインパルス応答を計測
> - 実際のスピーカでデルタ関数（パルス波）を出力することは難しいので，swept sine信号やM系列信号などを用いることが多い

東北大学のHRTF測定システム
[坂本+ 2016]

- HRTFのデータベースが様々な機関から公開されている
  - 東北大学: http://www.riec.tohoku.ac.jp/pub/hrtf/
  - UC Davis: https://www.ece.ucdavis.edu/cipic/spatial-sound/hrtf-data/

頭部伝達関数の個人性

- HRTFは個人差が大きい
  - 受聴者本人のHRTF以外で合成されたバイノーラル信号では，うまく音像定位ができないことが知られている



成人10名の正面方向のHRTF

[飯田 2017]

➡ バイノーラル合成ではHRTFの個人適応が大きな問題となる

応用例：ブラインド音源分離

ブラインド音源分離とは？

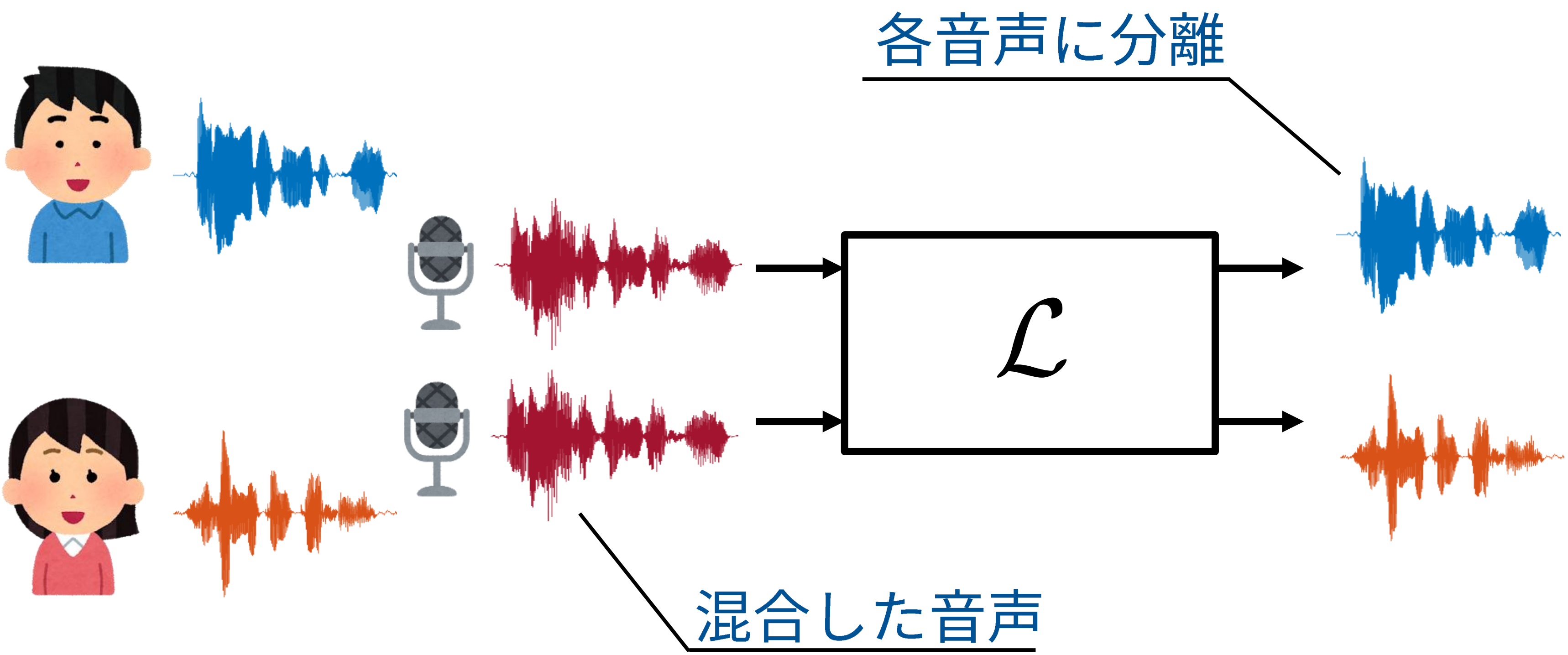
- カクテルパーティー効果
  - カクテルパーティのような騒がしい人混みの中でも，自分の名前や自分が興味のある会話は自然と聞き取ることができる現象。

人間はたくさんの音が混ざり合っている状況でも
音声を選択的に聴取できる

ブラインド音源分離とは？

- 音の選択的な聴取をコンピュータで実現したい！

➡ ブラインド音源分離（Blind Source Separation: BSS）



音源やマイクの位置関係などの情報が未知の状態で，（複数の）マイク信号のみから各音源の信号を分離

問題設定

- J 個の音源信号とM 個のマイクによる観測信号が，時間周波数領域で以下のように関係付けられるとする。

音源信号

$$\underbrace{x_{ft,m}}_{\text{観測信号}} = \sum_{j=1}^{J} \underbrace{a_{f,mj}}_{\text{伝達関数}} s_{ft,j}$$

観測信号　伝達関数

$$\begin{cases} m & \text{: マイク} \\ j & \text{: 音源} \\ t & \text{: 時間} \\ f & \text{: 周波数} \end{cases}$$

- 行列形式で書けば，

伝達関数行列（混合行列）

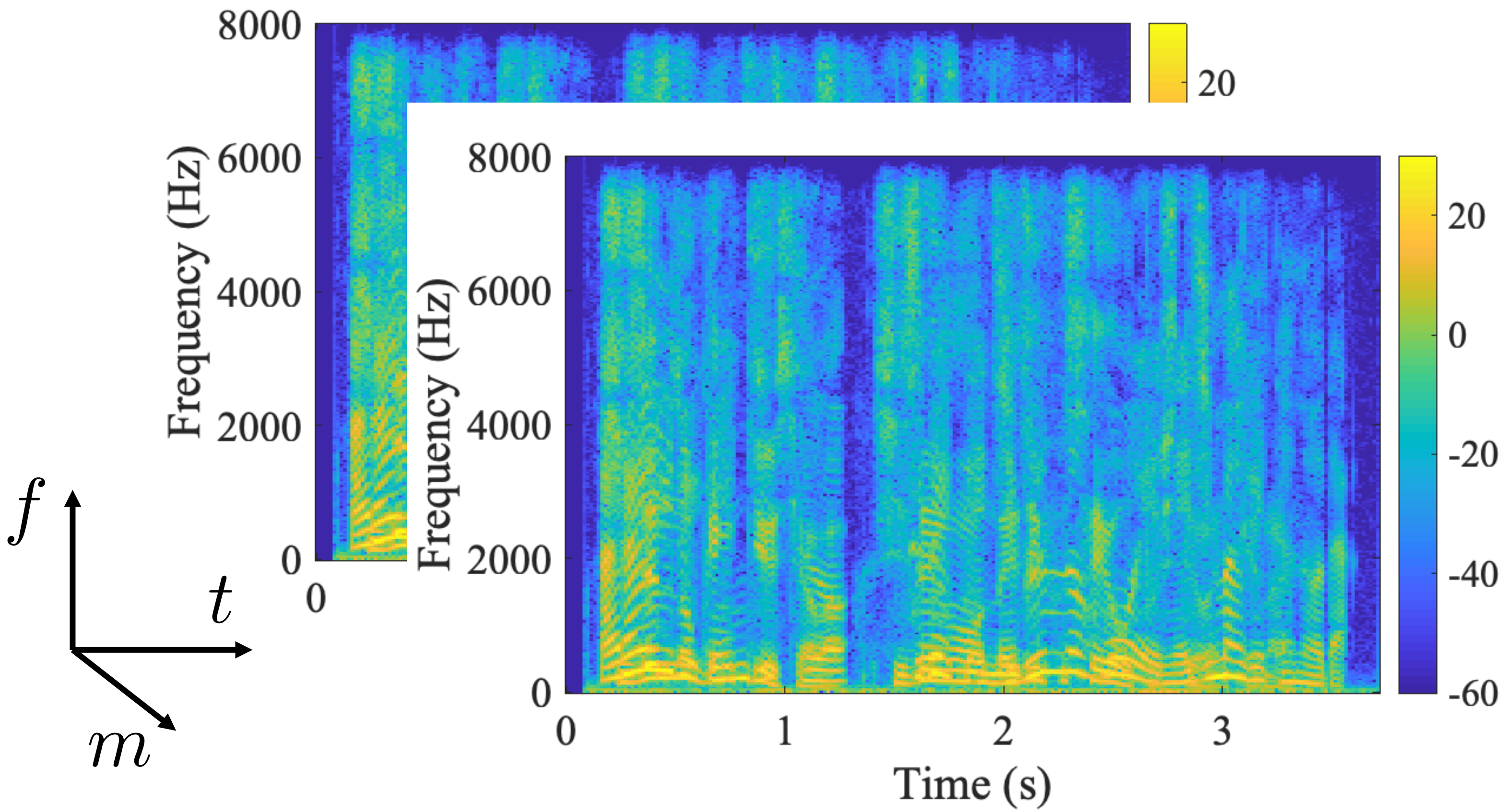
$$\boldsymbol{x}_{ft} = \boldsymbol{A}_f \boldsymbol{s}_{ft}$$

観測信号ベクトル　音源信号ベクトル

時間周波数領域表現

➢ 時間信号を短い時間フレームに区切ってフーリエ変換を行うことで，時間変化する信号の周波数分析を行う。

➡ 短時間フーリエ変換：Short-Time Fourier Transform (STFT)



問題設定

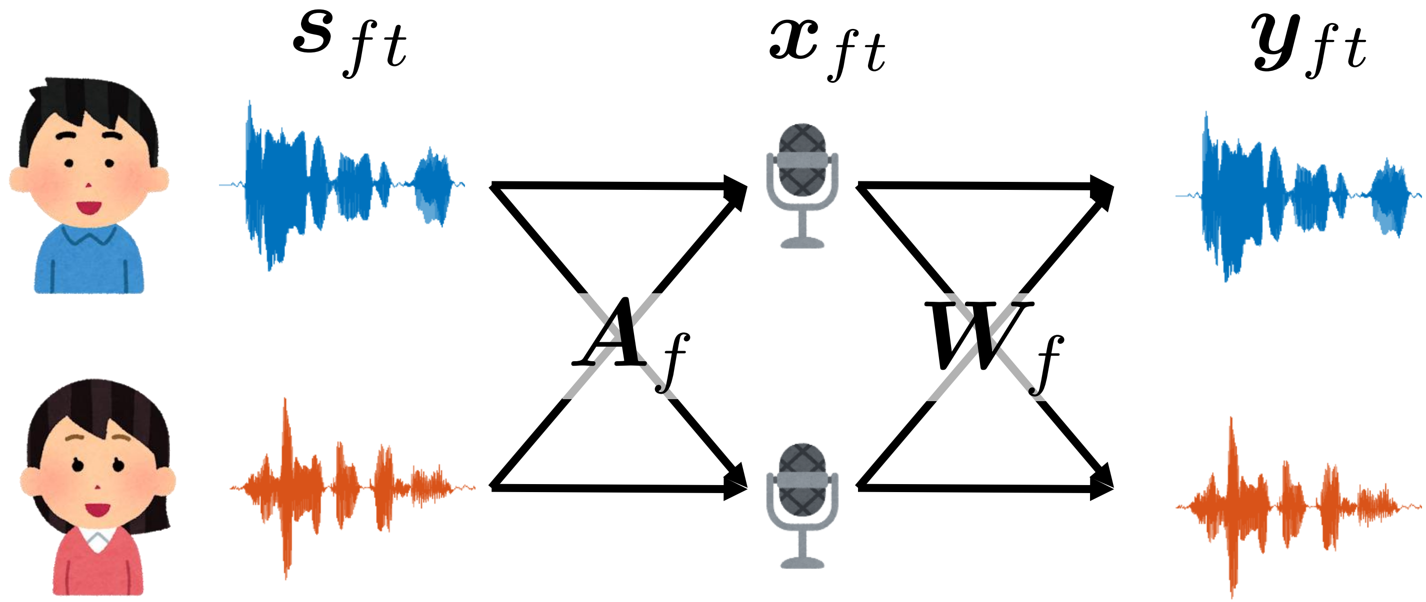
- 混合した観測信号 $\boldsymbol{x}_{ft} \in \mathbb{C}^M$ を各音源の信号 $\boldsymbol{y}_{ft} \in \mathbb{C}^J$ に分離するための，分離行列 $\boldsymbol{W}_f \in \mathbb{C}^{J \times M}$ を求めたい。

$$\underbrace{\boldsymbol{y}_{ft}}_{\text{分離信号}} = \overbrace{\boldsymbol{W}_f}^{\text{分離行列}} \underbrace{\boldsymbol{x}_{ft}}_{\text{観測信号}}$$

- 簡単のため，$J = M$ の場合のみを考えることとし，分離行列 $\boldsymbol{W}_f$ は正方行列とする。

問題設定

- 混合行列$\boldsymbol{A}_f$が未知の状況で，マイク信号$\boldsymbol{x}_{ft}$のみから各音源を分離する分離行列$\boldsymbol{W}_f$を推定する。



独立成分分析（ICA）

- 独立成分分析（Independent Component Analysis: ICA）では，混合された音源が**統計的に独立**であるという仮定の下で分離行列を推定する。

$$p(\boldsymbol{y}_{ft}) = p(y_{ft,1}, \ldots, y_{ft,J}) = \prod_{j=1}^{J} p(y_{ft,j})$$

- 統計的な独立性は無相関性/白色性よりも強い仮定であり，優ガウスな分布に従う信号源の混合を分離することを可能にする。
- ICAに基づくBSSは様々な評価尺度を用いて実現されているが，ここでは音源分離においてよく用いられる最尤推定によるICAを紹介する。

最尤推定によるICA

- 分離行列 $\boldsymbol{W}_f$ の尤度関数を考えると，

$$\mathcal{L}(\boldsymbol{W}_f) = p(\boldsymbol{x}_{f,1}, \ldots, \boldsymbol{x}_{f,T} | \boldsymbol{W}_f)$$
$$= \prod_{t=1}^{T} p(\boldsymbol{x}_{ft} | \boldsymbol{W}_f)$$
$$= \prod_{t=1}^{T} |\mathrm{det} \boldsymbol{W}_f|^2 \, p(\boldsymbol{y}_{ft})$$

- ここで，線形変換 $\boldsymbol{y}_{ft} = \boldsymbol{W}_f \boldsymbol{x}_{ft}$ の確率密度に関する以下の関係式を用いた。

$$p(\boldsymbol{x}_{ft} | \boldsymbol{W}_f) = |\mathrm{det} \boldsymbol{W}_f|^2 \, p(\boldsymbol{y}_{ft})$$

最尤推定によるICA

- 音源信号が互いに独立であると仮定すれば，

$$p(\boldsymbol{y}_{ft}) = \prod_{j=1}^{J} p(y_{ft,j})$$

- $\boldsymbol{W}_f$ の負の対数尤度を考えると，

$$\begin{aligned}
\mathcal{J}(\boldsymbol{W}_f) &= -\log \mathcal{L}(\boldsymbol{W}_f) \\
&= -\log \left\{ \prod_{t=1}^{T} |\mathrm{det}\boldsymbol{W}_f|^2 \prod_{j=1}^{J} p(y_{ft,j}) \right\} \\
&= \sum_{t=1}^{T} \sum_{j=1}^{J} G(y_{ft,j}) - 2T \log |\mathrm{det}\boldsymbol{W}_f|
\end{aligned}$$

ここで，$G(y_{ft,j}) = -\log p(y_{ft,j})$ とした。

最尤推定によるICA

- $G(y_{ft,j})$ はコントラスト関数と呼ばれ，音源信号が従うと仮定できる確率密度関数に基づいて設定する必要がある。
- 音声・音響信号では，優ガウスな分布として以下のような分布を用いる場合が多い。

$$p(y_{ft,j}) \propto \exp\left(-\frac{\sqrt{|y_{ft,j}|^2+\alpha}}{\beta}\right)$$

ここで，α，β は非負のパラメータ。

- コスト関数 $\mathcal{J}(\boldsymbol{W}_f)$を最小化する $\boldsymbol{W}_f$を求めることで分離が達成できると考えられるが，分離行列 $\boldsymbol{W}_f$ は周波数　ごとに別々に求まるため，音源の順序に関する任意性が残ることになる。（パーミュテーション問題）

独立ベクトル分析（IVA）

- パーミュテーション問題を解決する方法の一つとして，音源信号の各要素 $y_{ft,j}$ ではなく，全周波数の要素を並べたベクトル $\boldsymbol{y}_{t,j} = \left[y_{1t,j}, \ldots, y_{Ft,j}\right]^{\mathsf{T}}$ の独立性を仮定する，独立ベクトル分析（Independent Vector Analysis: IVA）が知られている。[Hiroe+ 2006, Kim+ 2006]

- 音源信号間の独立性の仮定より，$\{\boldsymbol{y}_{t,j}\}_{j=1}^{J}$ の確率密度関数は，以下のように書ける。

$$p(\boldsymbol{y}_{t,1}, \ldots, \boldsymbol{y}_{t,J}) = \prod_{j=1}^{J} p(\boldsymbol{y}_{t,j})$$

独立ベクトル分析（IVA）

- 分離行列 $\boldsymbol{W}_f$ の尤度関数を考えると，ICAの場合と同様の定式化により，

$$\begin{aligned}\mathcal{L}(\boldsymbol{W}_1,\ldots,\boldsymbol{W}_F) &= \prod_{t=1}^{T} p(\boldsymbol{x}_{t,1},\ldots,\boldsymbol{x}_{t,M}|\boldsymbol{W}_1,\ldots,\boldsymbol{W}_F) \\ &= \prod_{t=1}^{T} p(\boldsymbol{y}_{t,1},\ldots,\boldsymbol{y}_{t,J}) \prod_{f=1}^{F} |\det \boldsymbol{W}_f|^2 \\ &= \prod_{t=1}^{T} \left\{ \prod_{j=1}^{J} p(\boldsymbol{y}_{t,j}) \right\} \prod_{f=1}^{F} |\det \boldsymbol{W}_f|^2\end{aligned}$$

独立性の仮定

独立ベクトル分析（IVA）

- 分離行列 $\boldsymbol{W}_f$ の負の対数尤度は，

$$\begin{aligned}\mathcal{J}(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_F) &= -\log \mathcal{L}(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_F) \\ &= \sum_{t=1}^{T} \sum_{j=1}^{J} G(\boldsymbol{y}_{t,j}) - 2T \sum_{f=1}^{F} \log |\det \boldsymbol{W}_f|\end{aligned}$$

ここで，$G(\boldsymbol{y}_{t,j}) = -\log p(\boldsymbol{y}_{t,j})$ はコントラスト関数である。

- IVAにおいても，音源信号の確率密度関数には，優ガウスな分布として以下のような分布を用いる場合が多い。

$$p(\boldsymbol{y}_{t,j}) \propto \exp\left(-\frac{\sqrt{\sum_{f=1}^{F} |y_{ft,j}|^2 + \alpha}}{\beta}\right)$$

独立ベクトル分析（IVA）

- ➢ コスト関数 $\mathcal{J}(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_F)$ を最小化する $\{\boldsymbol{W}_f\}_{f=1}^{F}$ を求めればよい。

$$\underset{\boldsymbol{W}_1, \ldots, \boldsymbol{W}_F}{\text{minimize}} \, \mathcal{J}(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_F)$$

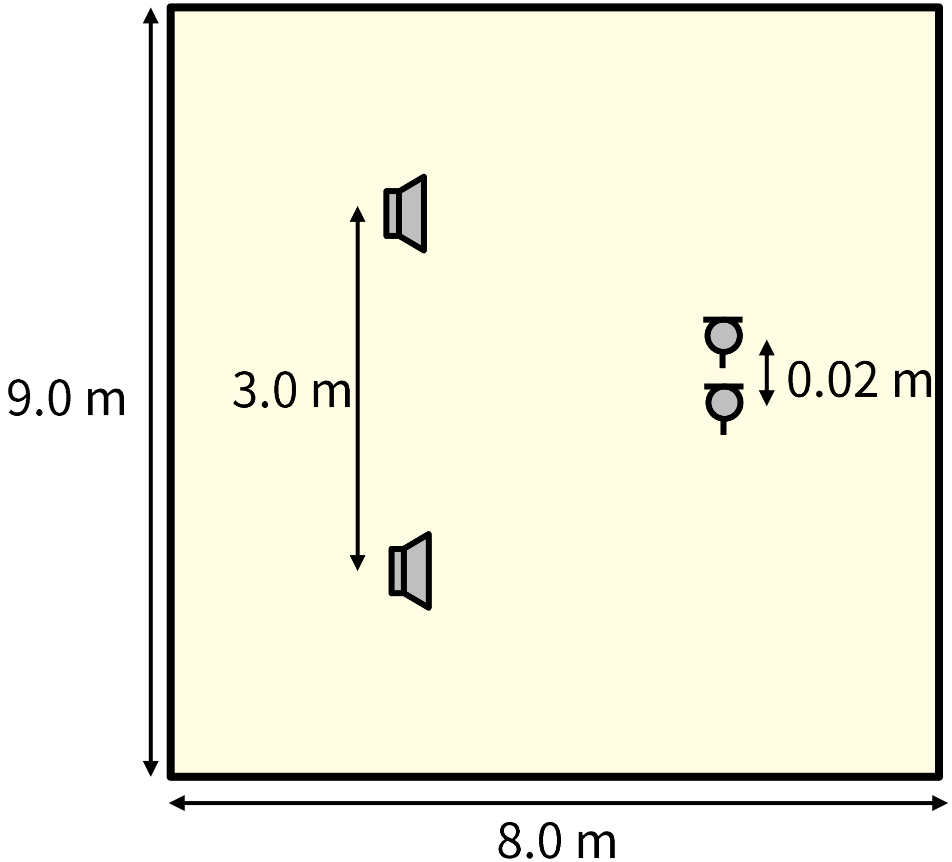
$$= \sum_{t=1}^{T} \sum_{j=1}^{J} G(\boldsymbol{y}_{t,j}) - 2T \sum_{f=1}^{F} \log |\det \boldsymbol{W}_f|$$

- ➢ この最適化問題を解くための様々な最適化アルゴリズムが提案されているが，Majorization-minimization (MM) アルゴリズムと呼ばれる，非線形最適化問題を効率的に解くためのアルゴリズムがよく用いられる。

[Ono+ 2011]

シミュレーション実験例

➢ 残響環境下でのIVAによる２音源の分離



混合音声：

分離音声１：

分離音声２：

参考文献

1. 日本音響学会編, “音響学入門ペディア,” コロナ社, 2017.
2. 飯田, 森本, “空間音響学,” コロナ社, 2010.
3. Blauert, “Spatial Hearing: The Psychophysics of Human Sound Localization,” MIT Press, 1997.
4. Hyvärinen, et al., “独立成分分析 – 信号解析の新しい世界,” 東京電機大学出版局, 2007.
5. Sawada, et al. “A review of blind source separation methods: two converging routes to ILRMA originating from ICA and NMF,” *APSIPA Trans. SIP*, 2019.