

# 学生実験の紹介

ケプストラム法による音声識別

千葉大学西千葉地区事務部理工系総務課技術グループ  
工学部総合工学科情報工学コース担当  
技術専門職員 和田 淳

---



# 自己紹介

---

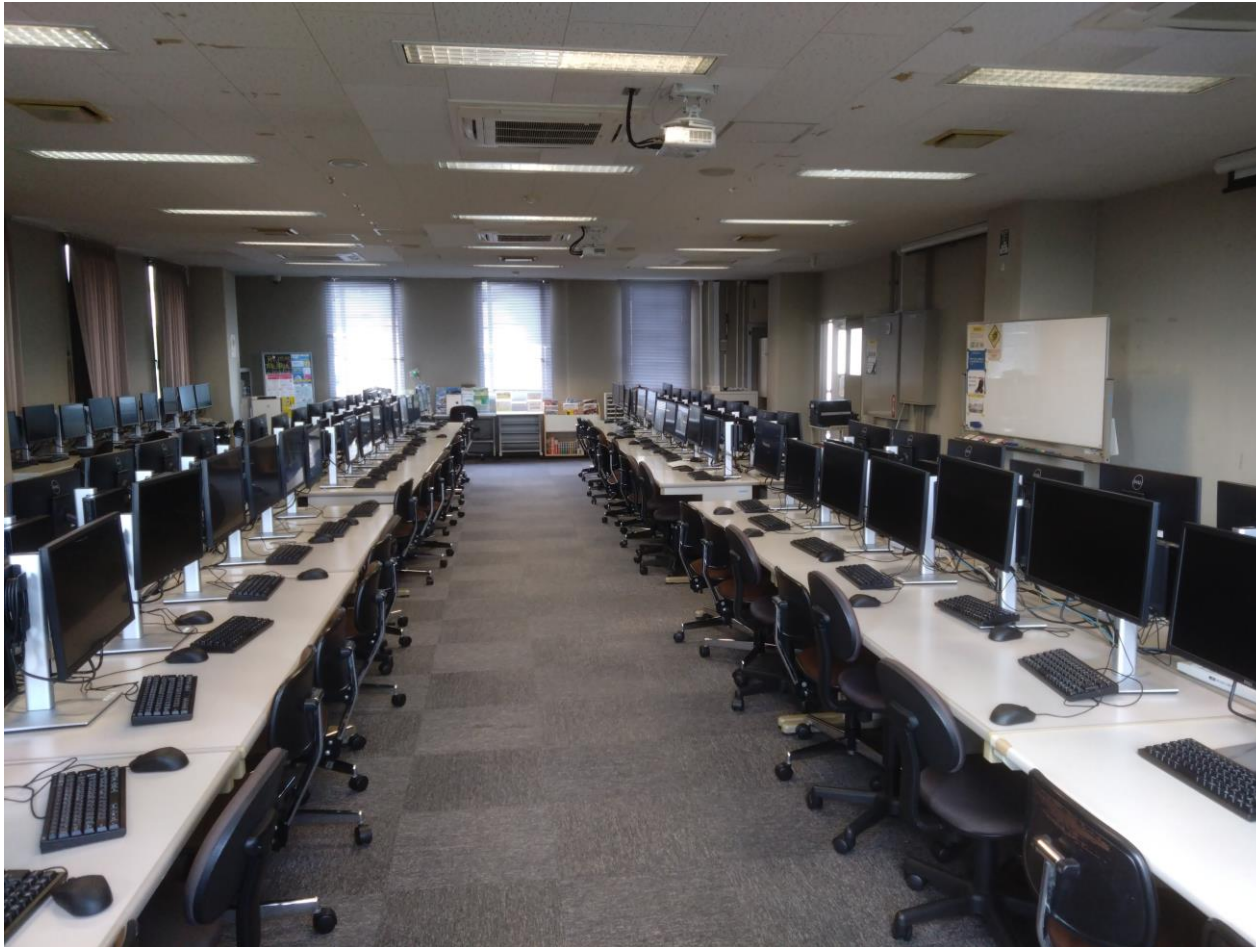
- ▶ 前職はシステムエンジニア（3年間）
- ▶ 2011年4月 千葉大学事務部教室系技術職員・工学部情報画像学科担当として着任
- ▶ 現在、工学部総合工学科情報工学コースを担当



# 通常の業務

---

## ▶ 情報工学実験室の管理



# 通常の業務

---

- ▶ 情報工学実験室の管理

（講義で使用するソフトのインストール、備品の管理、防犯カメラ映像の確認、講義時トラブル対応、コロナ対策等）

- ▶ Webサーバの管理

- ▶ メールングリストの管理

- ▶ その他



# 音声識別実験の紹介

---

- ▶ 目標：学生が自身の声を録音し、それが「あ・い・う・え・お」のどれであるかを判別するプログラムをC言語で作成する。
- ▶ 子音は難しいので母音のみ。
  
- ▶ 実験は4回にわたって実施される
  - ▶ 1：AD変換による音声のデータ化
  - ▶ 2：フィルタとフーリエ変換
  - ▶ 3：ケプストラム法
  - ▶ 4：音声識別の実践



# 共振周波数と基本周波数

---

声の高さが違っていても同じ音素だと認識できる

- ▶ 音素の音声波形には、**共振周波数**と**基本周波数**がある。声の高さは**基本周波数**で決まる。
- ▶ **共振周波数**により音素を識別できる。

共振周波数を調べる方法は？

---



# フーリエ変換の式

---

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \underline{F(\omega)} e^{i\omega t} d\omega$$



$f(t)$  のフーリエ変換

$$\underline{F(\omega)} = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$$

$F(\omega)$  は周波数  $\omega$  の成分  $e^{i\omega t}$  の大きさを表す。

---



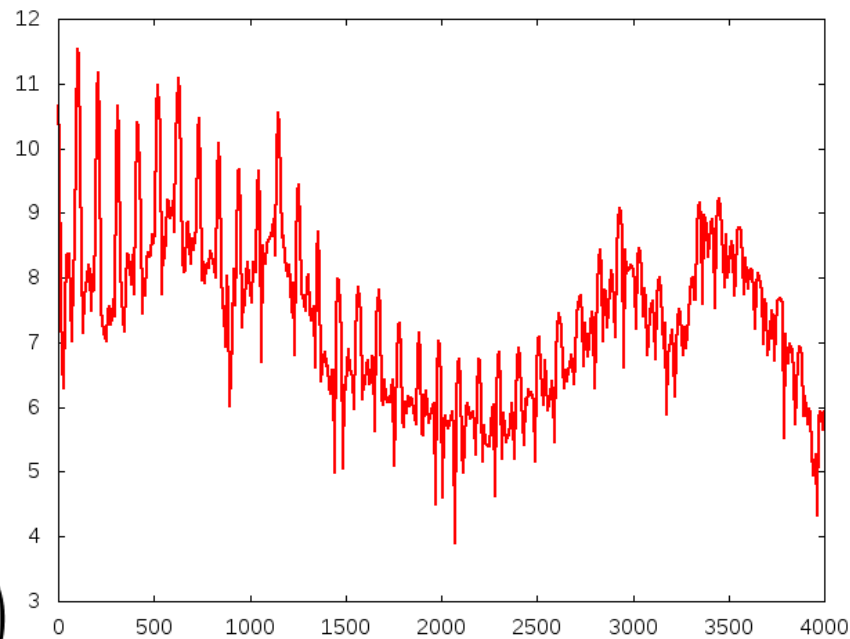
# 音声のスペクトル

- ▶ フーリエ変換 :  $F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt$
- ▶ 信号  $F(\omega)$  をスペクトルという。
- ▶ 共振周波数と基本周波数はパワースペクトル  $|F(\omega_k)|^2$  の概形に現れる。

- ▶ 扱いやすくするため、通常は対数パワースペクトルで表示 (底が10の対数で表示する)

$$\log\left(|F(\omega_k)|^2\right)$$

$$= \log\left(\operatorname{Re}(F(\omega_k))^2 + \operatorname{Im}(F(\omega_k))^2\right)$$



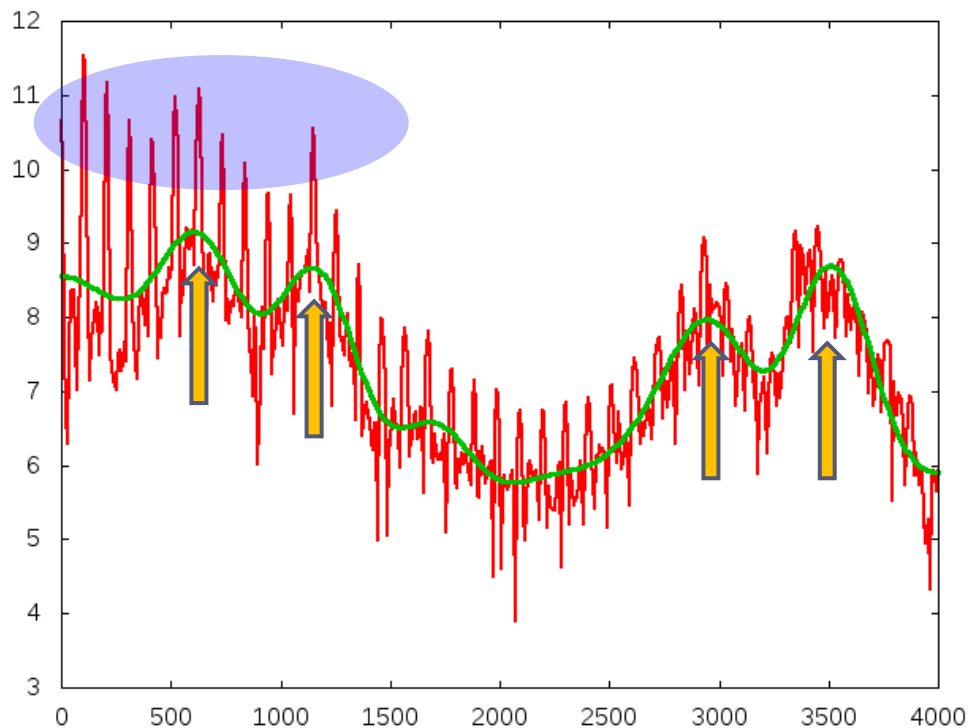
/a/の対数パワースペクトル



# /a/のスペクトル包絡

ぎざぎざした部分が  
基本周波数とその倍音  
(声の高さ)

対数パワースペクトルの  
概形で山になっている  
部分が共振周波数。



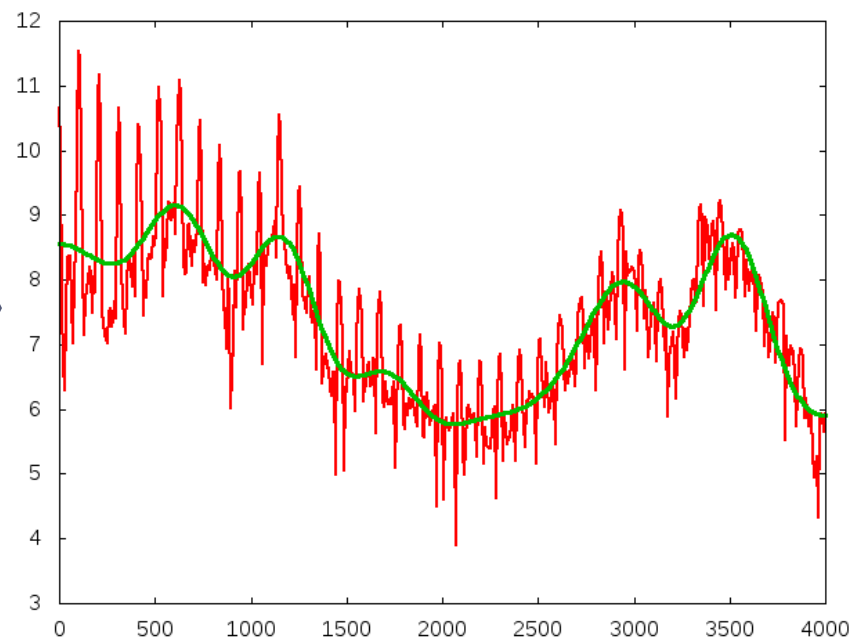
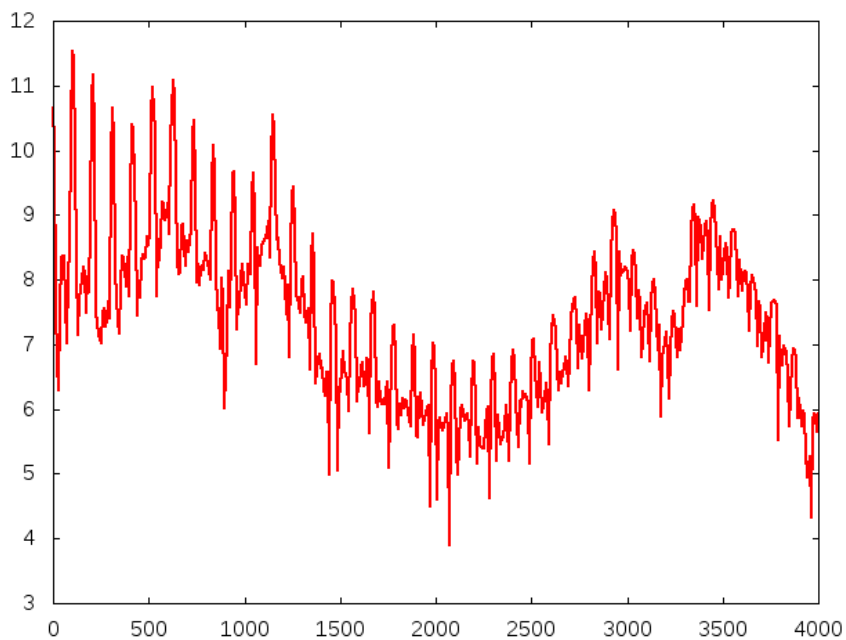
/a/の対数パワースペクトル

- ▶ 全体的な概形をスペクトル包絡と呼ぶ(緑の曲線).

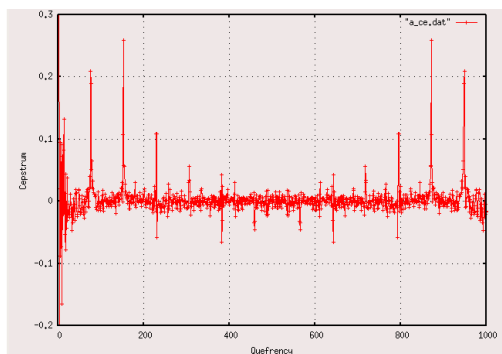
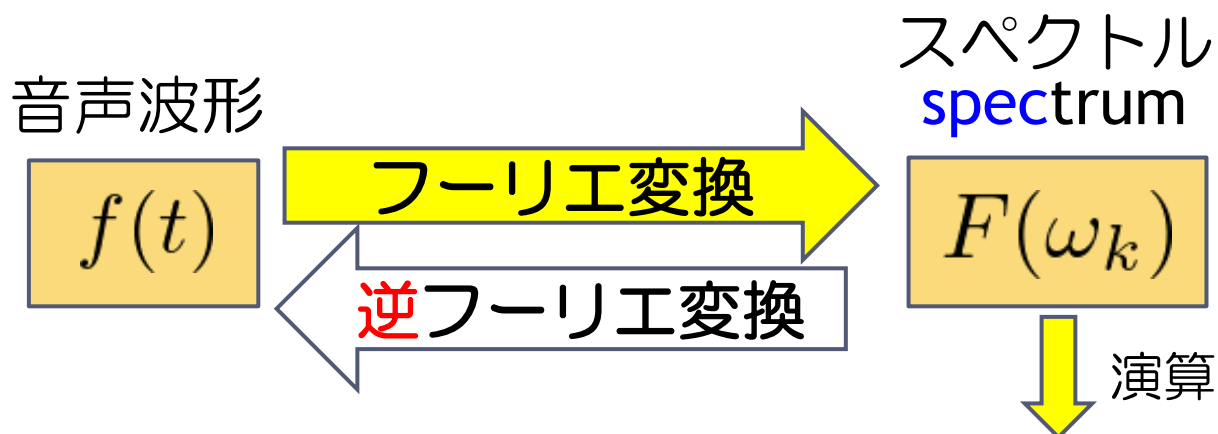
# ケプストラム法

- ▶ 音素を識別するため、対数パワースペクトルをスペクトル包絡と微細構造(基本周波数)に分離したい。

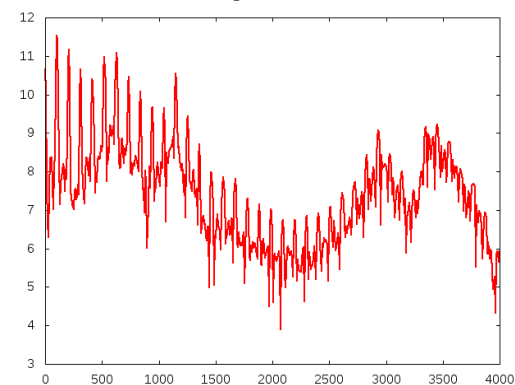
## ケプストラム法



# ケプストラム



ケプストラム  $C_m$   
cepstrum



対数パワースペクトル  
 $\log(|F(\omega_k)|^2)$

本実験においては.....

---

離散フーリエ変換のプログラムを  
作成するのは時間がかかりすぎる



本実験においては.....

---

実用的な速度が出るような  
アルゴリズムを用いないと  
動作が遅すぎて実験に使えない



# 高速フーリエ変換 (FFT)

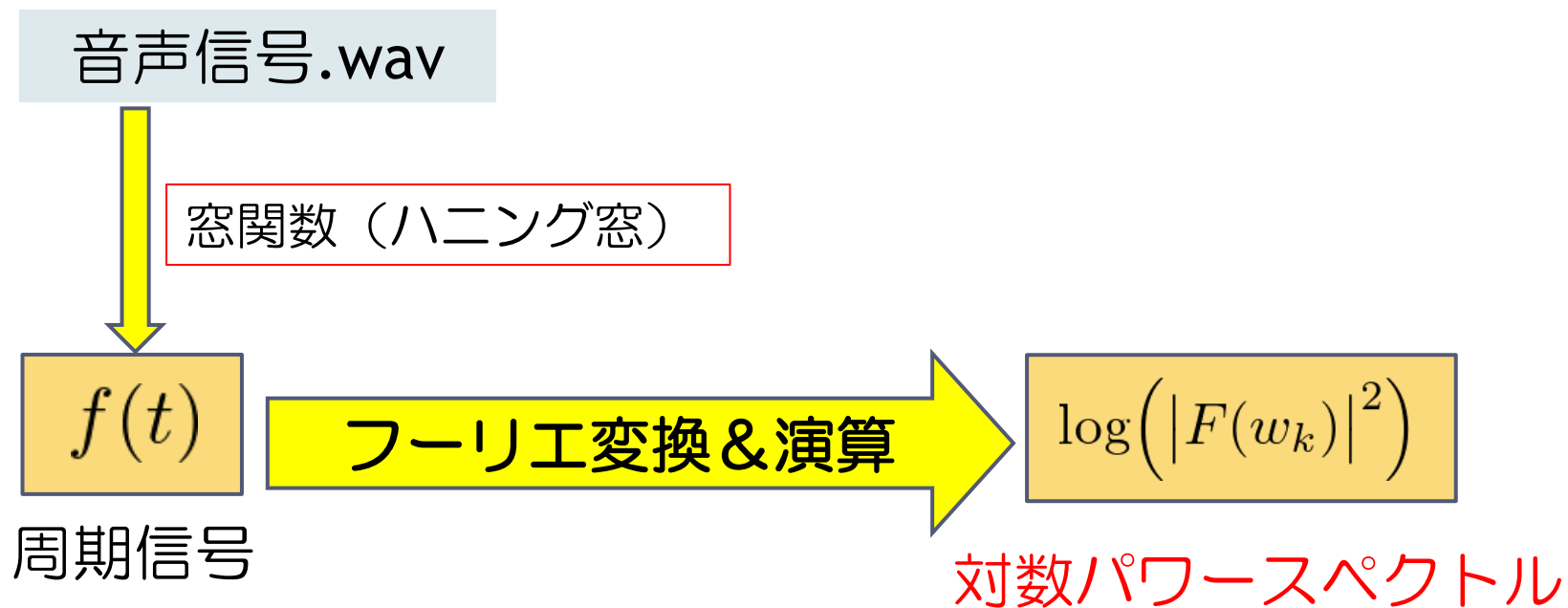
---

- ▶ **Fast Fourier Transform**
- ▶ 離散フーリエ変換を効率よく計算するアルゴリズム
- ▶  $N$  (窓幅) が2のべき乗でないと使えない  
(本実験では  $N=1024$ )



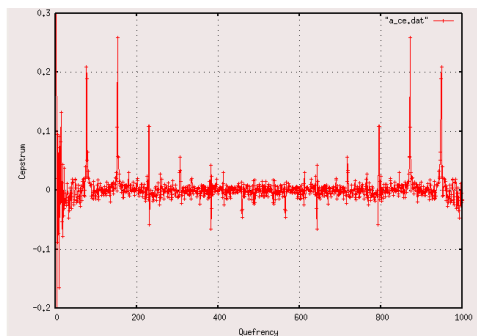
## ケプストラム法・実際の手順

---



# ケプストラム法・実際の手順

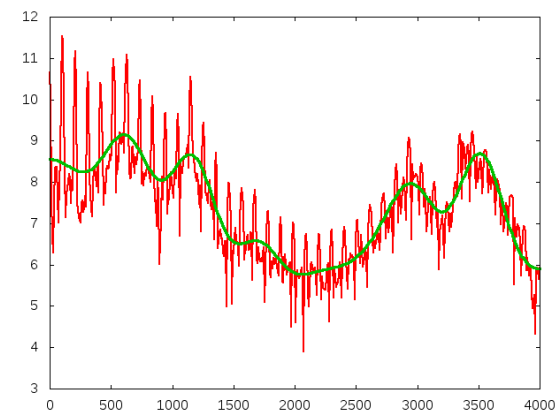
ケプストラム  $C_m$



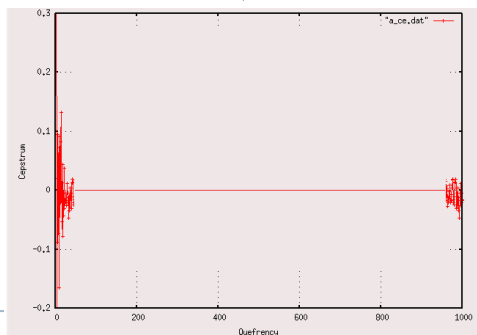
逆フーリエ変換

$$\log\left(|F(w_k)|^2\right)$$

対数パワースペクトル



高次元成分  
を除去



フーリエ変換

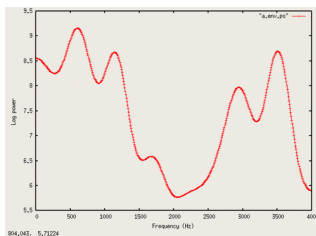
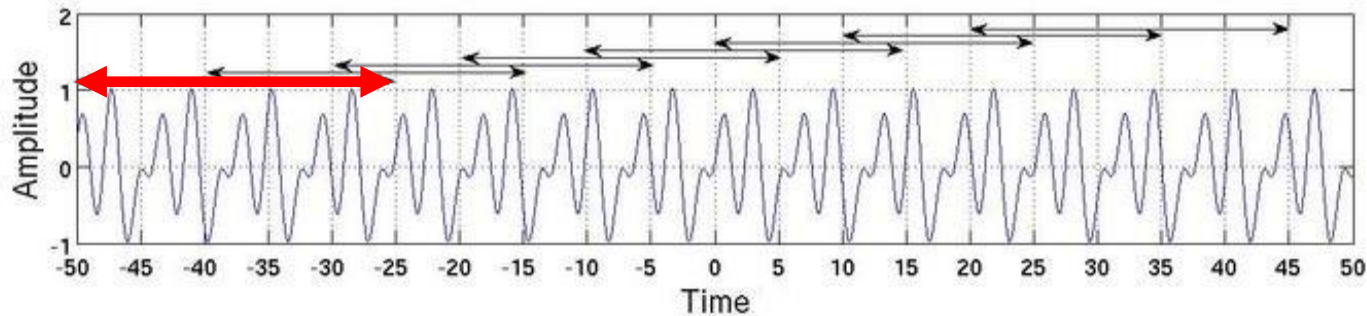
スペクトル包絡



# 母音識別プログラムを作る

## ▶ 今までやっていたこと

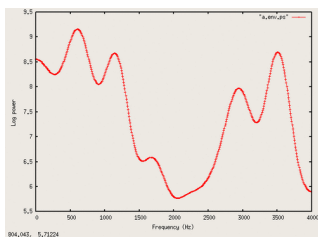
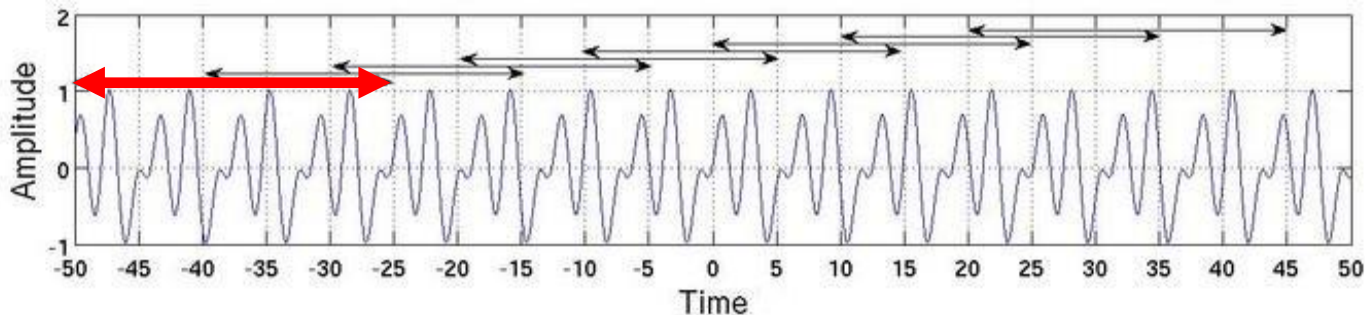
録音開始から一定時間のデータだけ使用して  
スペクトル包絡を求める。



# 母音識別プログラムを作る

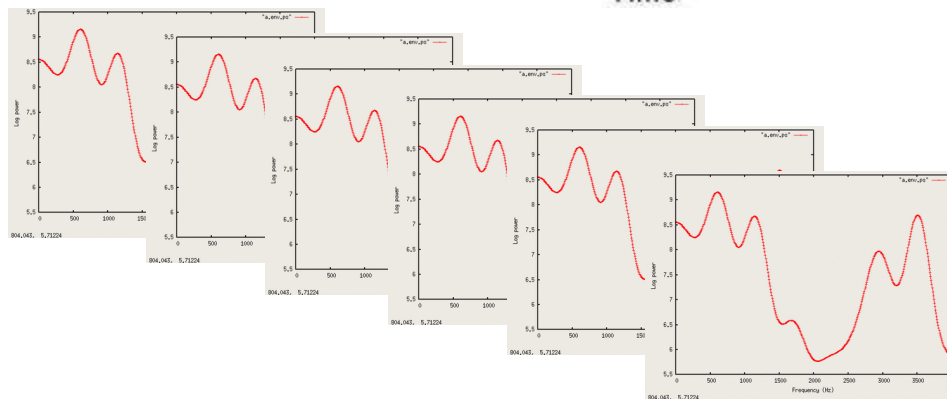
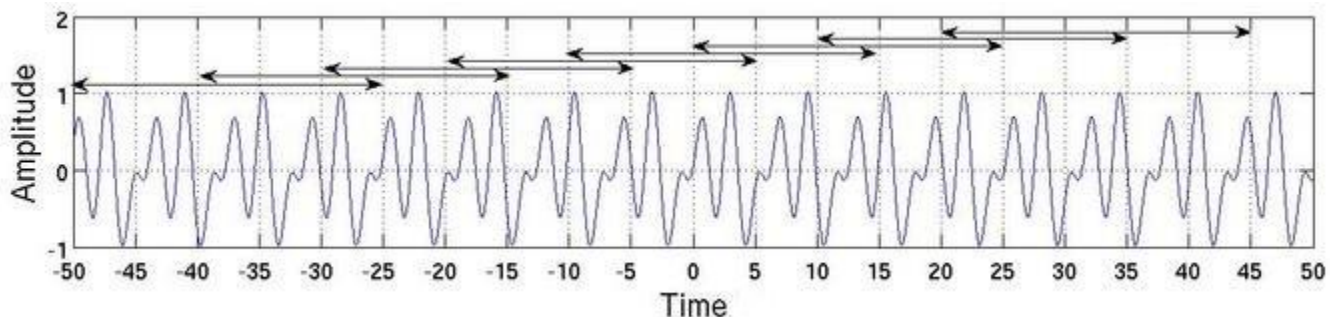
## ▶ 今までやっていたこと

この方法だと、たまたまあるタイミングで採取したデータを識別の基準に用いることになってしまう。



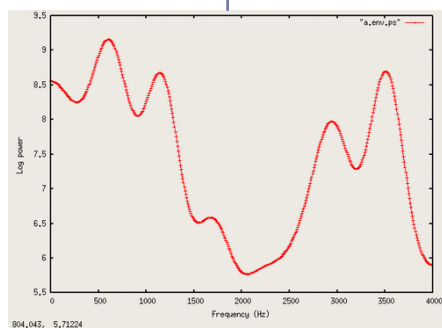
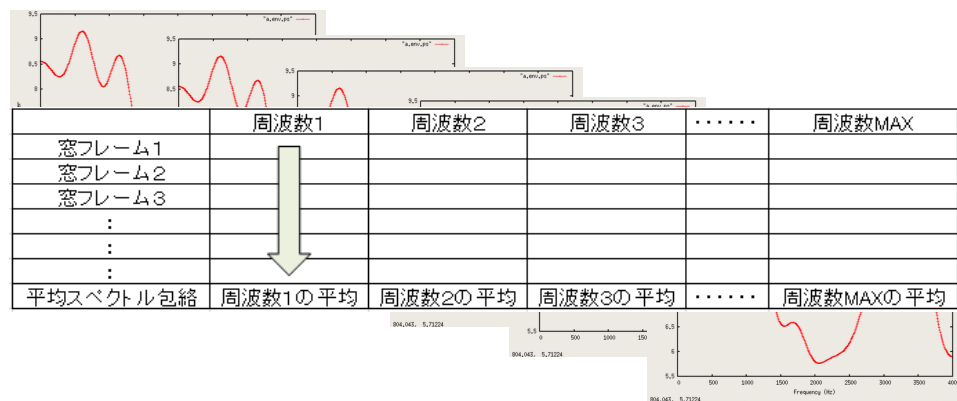
# 母音識別プログラムを作る

- ▶ 等間隔で複数の時間窓を設定し、それぞれの時間窓についてスペクトル包絡を求める。



# 平均スペクトル包絡

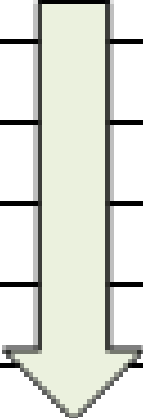
- ▶ 各周波数(0Hz~MAX)について平均を求める



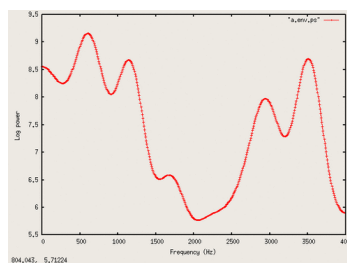
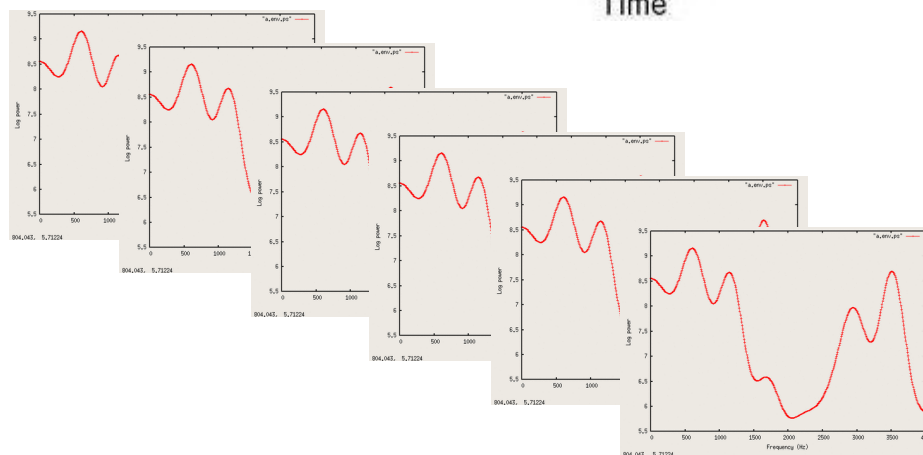
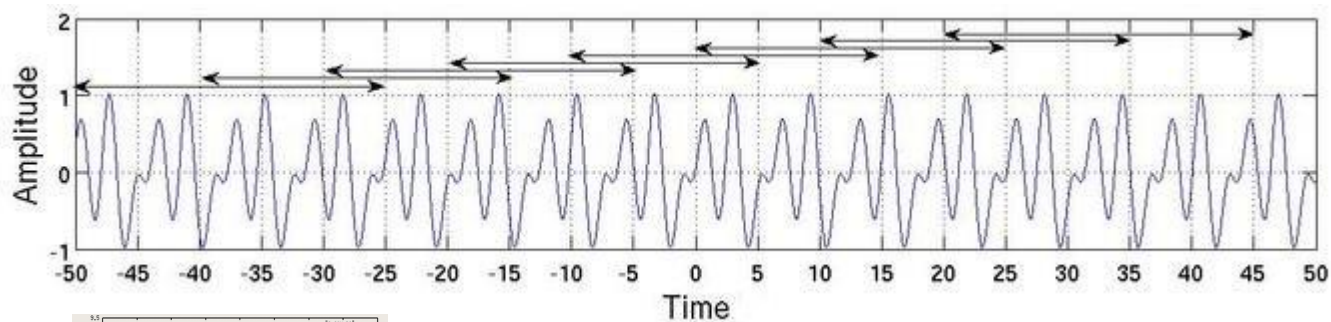
平均スペクトル包絡

# 平均スペクトル包絡

- ▶ 各周波数(0Hz~MAX)について平均を求める

	周波数1	周波数2	
窓フレーム1			
窓フレーム2			
窓フレーム3			
⋮			
⋮			
⋮			
平均スペクトル包絡	周波数1の平均	周波数2の平均	周波

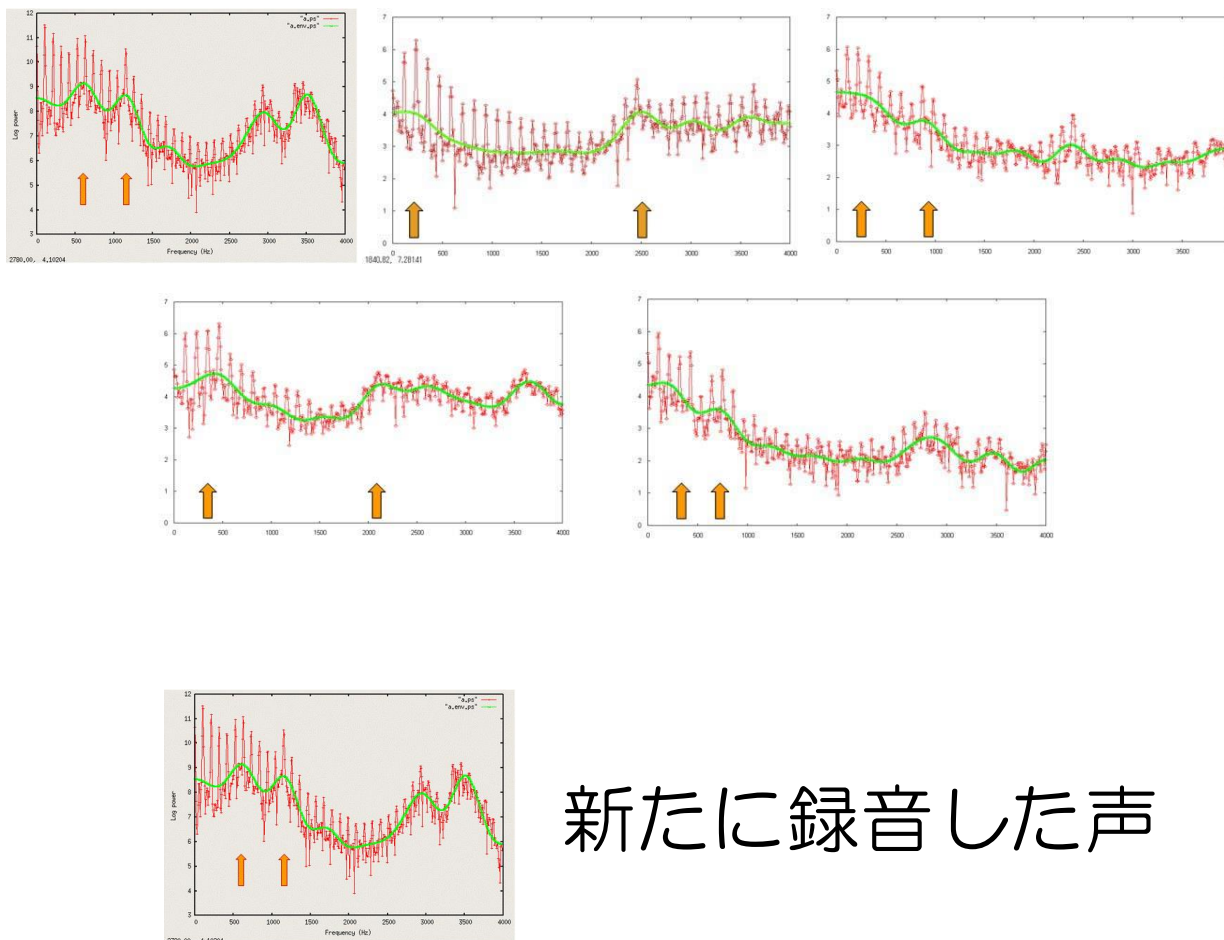
# 平均スペクトル包絡



平均スペクトル包絡

# 本実験の目標を達成

- ▶ 平均スペクトル包絡を比較して、母音を識別する



新たに録音した声

## さいごに

---

- ▶ 千葉大学は令和6年度に改組を予定しており、それに伴うカリキュラムの見直しにより本実験は新しい内容に置き換えられる。  
(情報・データサイエンス学部／学科を新設)
  - ▶ 機械学習やPythonなど、より時流に合わせた内容を取り入れたものになる？ (検討中)
- 
- 