音声データを扱う！

文科省『数理・データサイエンス・AI教育プログラム』の応用基礎レベルに認定された「実践的な演習を含む科目」である ＜**応用データ処理技術**＞より

千葉大学　情報・データサイエンス学部（2024年4月新学部）

黒岩眞吾

音声の全３回の Google Colab.ノートブック＆データをWSで配布しました.
数理・データサイエンス・AI教育強化拠点コンソーシアムにて公開を検討いただいておりますが，今すぐ欲しい方は kuroiwa@faculty.chiba-u.jp にご連絡ください.

CHIBA UNIVERSITY

CHIBA UNIVERSITY

数理・データサイエンス・AI教育の概要

千葉大学について

- 国立総合大学（10学部、学部定員 9,674名）
- 2024年度「情報・データサイエンス学部」新設

数理・データサイエンス科目の必修化

2020年度より全学必修化（３単位）

- 数理・データサイエンス科目（基礎） ２単位
  情報リテラシー　31クラス
  リテラシーレベル・モデルカリキュラムの「導入」、「心得」、「基礎」の範囲をカバー
- 数理・データサイエンス科目（展開） １単位
  モデルカリキュラムの「選択」に該当
  データサイエンスA-Dなど31クラス
- 上記３単位のプログラムが、
  2021年度に文部科学省
  数理・データサイエンス・AI教育プログラム
  認定制度　リテラシーレベルに認定、
  リテラシーレベルプラスにも選定された

全学実施体制

- 情報戦略機構のデータサイエンス部門が、全学教育センター、スマートラーニングセンター、アカデミックリンクセンター等と連携して運営
- 各学部の常勤教員約50名から構成される数理・データサイエンス教員集団が講義・演習を担当

応用基礎レベルのプログラム

- 情報リテラシー（２単位）と応用基礎レベルモデルカリキュラムのデータサイエンス基礎、データエンジニアリング基礎、AI基礎をカバーするデータサイエンスB（１単位）が必修。
  実践的な演習を含む数理・データサイエンス科目（展開）の指定科目１単位。微積分学、線形代数学、
  統計学　各２単位　　合計10単位
- 2023年度に応用基礎レベルに認定された
- 本日紹介する「応用データ処理技術」は、実践的な演習を含む指定科目のひとつ

「応用データ処理技術」 概要(シラバス)

音声、**画像**、映像、**主観的データ**などのデータ収集や前処理では、物理学や認知科学的な知識が必要である。この授業ではプログラミング演習を交えながらこれらのデータの扱い方を学ぶ。具体的には、連続信号のディジタル化の原理と手法、メディアごとに異なるデータ表現や特徴抽出手法、教師なし学習と教師あり学習、主観的データの取得方法や分析手法、仮説検定である。プログラミングにはPythonを用い、Google Colaboratoryを用いて各種プログラムの作成を通じ，中級レベルのスキルを習得する。

「応用データ処理技術」 目標(シラバス)

A) **音声データを情報処理できる技術を習得する.** 具体的には, (1)音声データ(物理信号)をディジタル化するまでの手順を理解し, パソコン上で実行できる. Pythonを利用して(2)音声データの可視化, (3)音声データから特徴ベクトルへの変換および可視化. (4)音声データをクラスタリングして可視化, できるようになる.

B) **画像データを情報処理できるAI技術を習得する.** 具体的には, 深層学習フレームワークであるTensorFlow(Python言語)を用いて, 深層畳み込みニューラルネットワークによる特徴抽出と画像分類の実装と評価ができるようになる.

C) **主観評価実験の実施および結果を科学的に分析する技術を習得する.** 具体的には, 簡単な視覚評価実験を行い, 得られた結果を統計分析ツール(R等)を用いて統計的に分析することにより, 主観評価データを定量的・科学的に扱うことができるようになる.

CHIBA UNIVERSITY

注)評価は全てレポートによる

「応用データ処理技術」 履修要件（シラバス）

抽選に参加する前に下記を確認してください：

① Pythonのnumpyを使って3行3列の行列の固有値を求めることができること，および，固有ベクトルの意味を理解していること.

② MacOS/Linuxのターミナルの利用，WindowsのPowershellの利用等，CUI (character-based user interface)でのコンピュータ利用に興味をもっていて勉強したいという意欲をもっていること.

③ 下記のURLの内容に対し「自分でやってみよう」という意欲をもっていること.
https://colab.research.google.com/notebooks/welcome.ipynb?hl=ja

④ 様々なクラウドプラットフォーム(moodle，Google Classroom，Teams，Zoom など)に興味をもち，「使ってみよう」という意欲をもっていること.

履修登録を行うと，数日内にGoogle Classroom「応用データ処理技術」が見られるようになるので，「授業」タブの**「事前準備」を実施してください**（**履修要件テストもありますので第1回目までに実施しておいてください**）

履修要件を明確に： やる気のある２割の学生の履修を想定

「応用データ処理技術」 講義内容

① 第1回　Google Colab. 入門（黒岩）

② 第2回　音声の生成と音響特徴量（黒岩）

③ 第3回　クラスタリングと音声認識（黒岩）

④ 第4回　深層学習による画像分類1（川本）

⑤ 第5回　深層学習による画像分類2（川本）

⑥ 第6回　心理物理データの分析、検定方法1（溝上）

⑦ 第7回　心理物理データの分析、検定方法2（溝上）

⑧ 第8回（オンデマンド）研究所のオープンハウスの視聴

複数人で分担すると
開講しやすい

↓

やり始めると、
各担当者が7回でき
そうな内容に。

第1回　Google Colab. 入門

「Colaboratoryへようこそ」はやっていることを前提

1. Colabとのデータのやりとり
   ① Google Driveのマウント
   ② ローカルPCとGoogle Colab.とのデータのやり取り
   ③ おまけ：VMの説明，linuxのコマンドとディレクトリ構造など
2. 音声の録音と再生
   ① PC上の録音再生アプリ（Wavesurfer, Audacity）で録音
   ② 音声データ(Wavファイル）の読み書き
   ③ Google Colab上での音声再生
   ④ 波形の描画
   ⑤ 音声データの加工（課題　赤：AKA, 汗：ASE, 足：ASHI を録音＆逆再生で考察）

結構盛り上がる

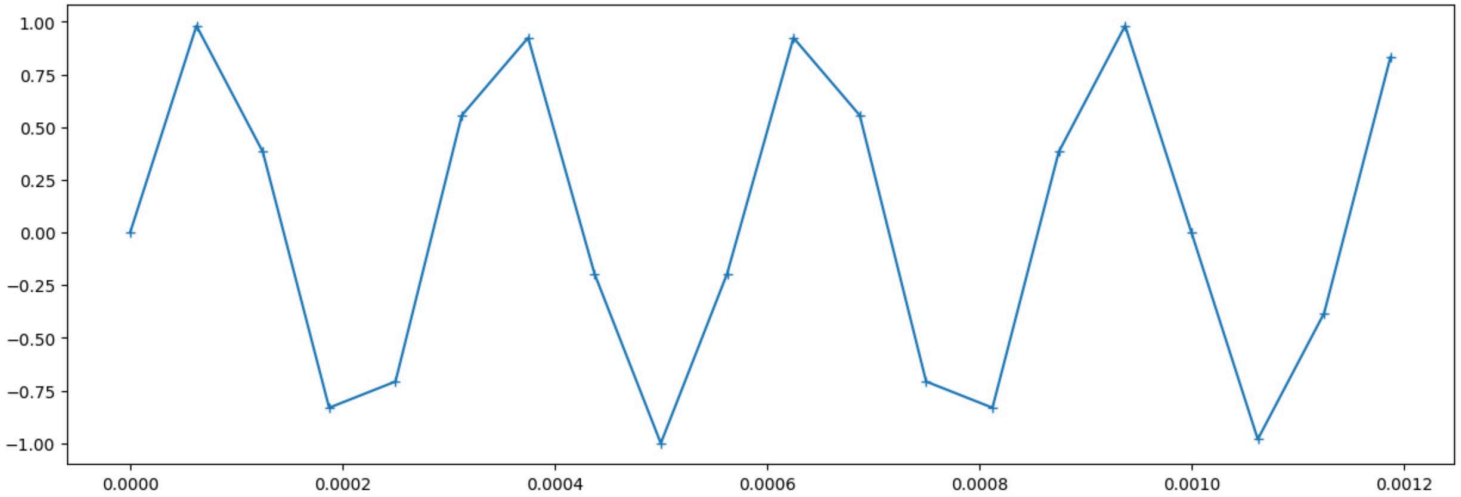
第2回　音声の生成と音響特徴量

宿題を演習にすると
90分x3回分程度のボリューム

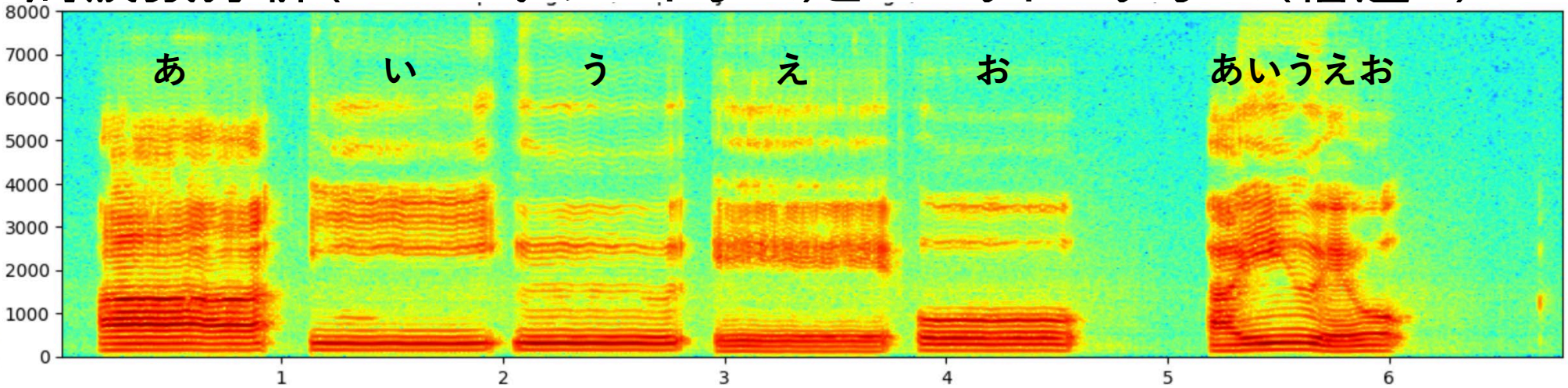
1. AD変換とサンプリング定理
2. 波形の正しい描画（宿題１）
   - 3500Hzの正弦波の波形 →

     （サンプリング周波数16kHz）

     アナログ信号に直すコードを書く



3. 音声の生成過程と聴覚の特性を反映してた音響特徴量の説明
4. 周波数分析（FFT～ケプストラム)とスペクトログラム（宿題２）



第3回　クラスタリングと音声認識

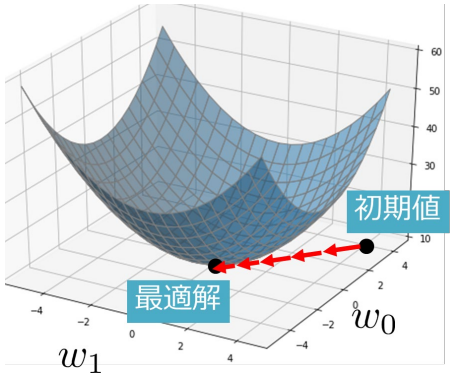
予習・宿題を演習にすると
90分x3回程度のボリューム

k-meansクラスタリング と 主成分分析 を youtube動画で予習

1. クラスタリングと可視化
   ① 音声「あ，い，う，え，お，あいうえお」の音響特徴ベクトルの時系列を抽出
   ② k-meansクラスタリンとPCAで可視化 → 概ね母音ごとにクラス化されることを確認
   ③ 宿題：「あいうえお」が正しく分離できる方法の模索
2. 単語音声認識：DTWでテンプレートマッチング
   ① Dynamic Time Warping (DP: 動的計画法)の説明
   ② 「はい」「いいえ」をテンプレートにして，「はーい」を認識
   ③ 宿題： 自身で録音した「はい」「いいえ」を認識
3. Google音声認識APIを使ったクラウド音声認識

第4・5回　深層学習による画像分類

- 第4回：機械学習の基礎に関する座学
  - 回帰や分類の具体例を通して、機械学習を解説。細かな数式導出はせずに、機械学習が何をやっているのかを直感的に解説。
  - 次回の深層学習につながるような内容構成。
    線形回帰>ロジスティック回帰>多層ニューラルネットワーク



- 第5回：深層学習に関する座学＋実習、生成AIの活用
  - 画像分類を例に、深層学習（畳み込みニューラルネットワーク）を解説。
  - Google Colab. 上での実習。画像分類のPythonサンプルコードを用意。
  - **Pythonサンプルコードの不明点に対して、生成AI（ChatGPTなど）の利用を許可**
    生成AIによるプログラムの説明は大変優秀。受講生一人ひとりにTAをつけるのと同等

第4･5回　深層学習による画像分類

- 課題：画像分類Pythonコードを改良し、分類精度を上げよ
  - 2023年度では試行として、生成AIを積極的に利用してよいことを明示
  - 入力プロンプトも書くように指示

- 結果：
  - 生成AIが利用できなかった一昨年度に比べて、全体的に分類精度が向上
  - レポートの書き方の指示については要改善の印象。なぜ精度が向上したかきちんと分析してあるレポートもある一方で、「入力プロンプトとその出力」のペアだけを並べたレポートもあった。

CHIBA UNIVERSITY

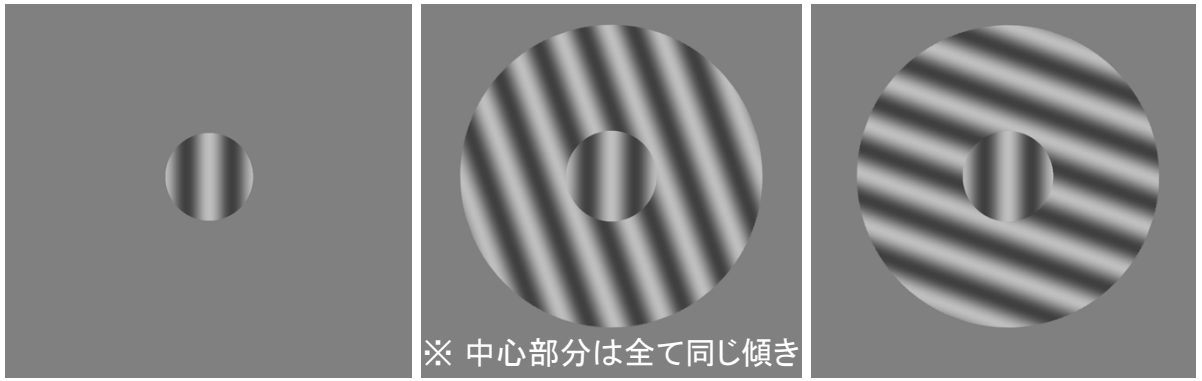
第6・7回　心理物理データの分析、検定方法

1. 心理物理測定の手法と視覚評価実験

- Question: 人の感覚は測定できるのか？
  → 心理学で扱われる感覚量を数値で表すことにより可能 → 心理物理測定
- 基本的な心理物理測定手法の説明
  - 様々な刺激を与えたときの応答を記録 → 知覚特性を調べる
    - 調整法、極限法、階段法・上下法、恒常法 etc.
- 心理物理実験の体験
  - Psychopy（Pythonベースの心理実験環境構築用アプリ）
  - 傾きの対比現象

PsychoPy3
Now running studies online
https://www.psychopy.org

PsychoPy Builderで作る心理学実験
（愛媛大学・十河宏行）
http://www.s12600.net/psy/python/ppb/



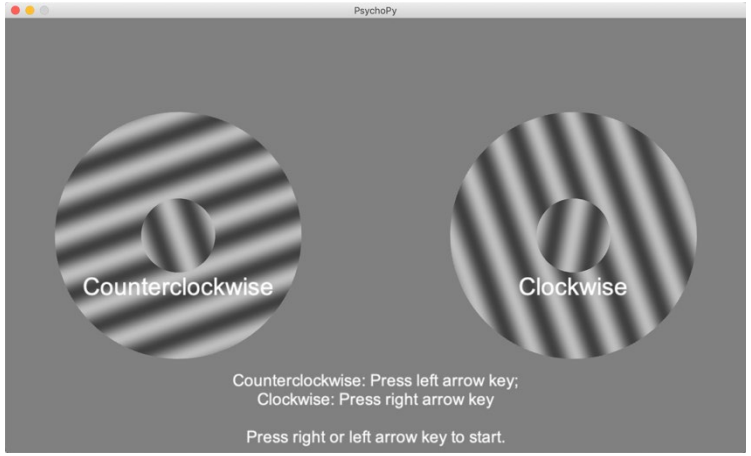
周囲の縞の傾きが
中心の縞の傾き知覚
に与える影響を調べる

CHIBA UNIVERSITY

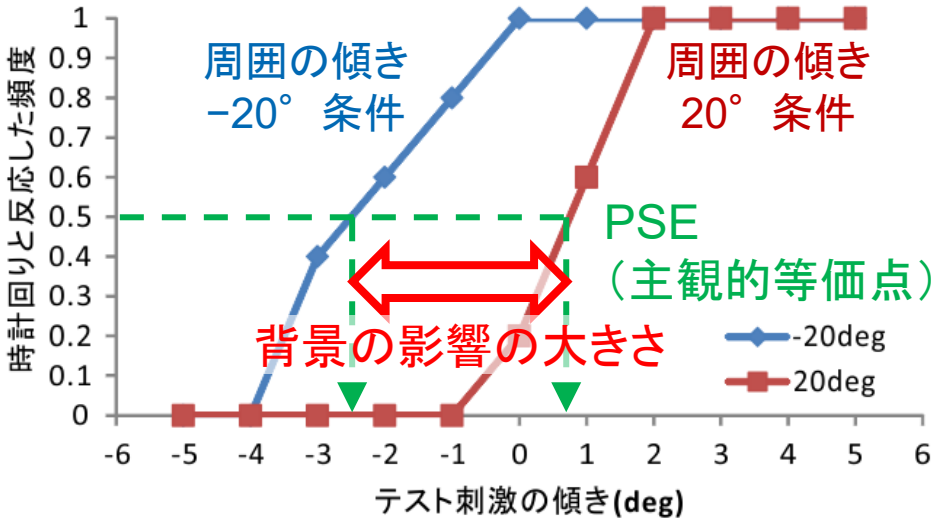
第6・7回　心理物理データの分析、検定方法

2. 心理物理データの分析、検定方法

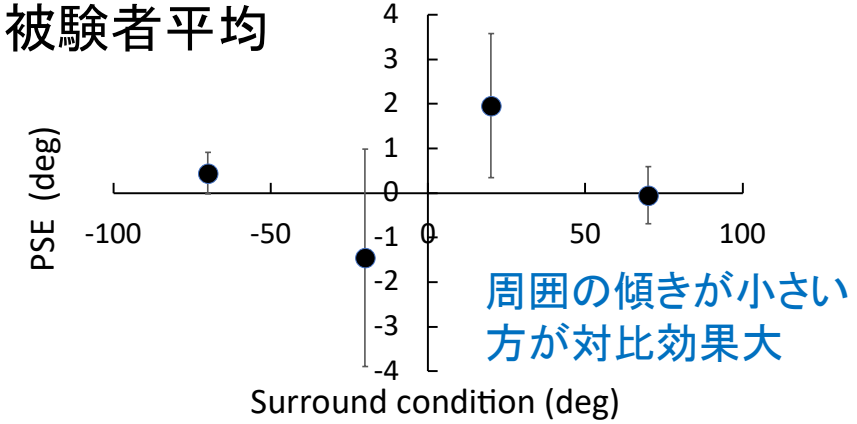
- 自ら実験して、結果を解析（宿題）
  - 条件間の比較、平均、標準偏差 etc.



結果の例



被験者平均



- <u>有意差検定</u>：統計的に意味のある差かどうか？
  - t検定（対応のある）：2群間の比較
  - 分散分析（対応のある）：3群間の比較
  - 多重比較：各条件間の比較

→ 心理物理データ（感覚量）の定量的分析方法の概要を理解

残りの時間で配布 Google Colab を解説

1-1 Colabとのデータのやりとり

1-2 音声の録音と再生

2-2 波形の正しい描画（宿題1）& 解答

2-2 周波数分析（FFT～ケプストラム)とスペクトログラム（宿題2）& 解答

3-1 クラスタリングと可視化

3-2 単語音声認識：DTWでテンプレートマッチング

3-3 Google音声認識APIを使ったクラウド音声認識