

課題 1 音響信号可視化 GUI 作成

音響信号ファイルを読み込み、音響信号のさまざまな情報を表示するグラフィカルユーザーインターフェースを作成せよ。少なくとも以下の 3 つを同時に表示させること。

1. 音響信号のスペクトログラム
2. 音響信号の基本周波数
3. 母音などの何らかの識別を行った結果

加えて、このインターフェースがより便利なものになるように改良せよ。例えば以下のような改良が考えられるが、もちろんこれ以外の改良でも構わない。創意工夫すること。

- 音響信号の区間を選択し、その区間のスペクトルを表示する。
- 音響信号を再生し、その再生位置をアニメーションで示す。
- 音楽音響信号のコードとその区間を認識し表示する。
- NMF を用いて、音声と音楽を分離し、選択的に再生する。

実装した項目を以下に挙げる。また、それぞれについて実装方法を説明する。

まず、図 1-1 では、以下を実装した。

1. 音響信号のスペクトログラム (左図 背景)
2. 音響信号の基本周波数 (左図 赤線)
3. 母音の識別を行った結果 (左図 白線)
4. 音響信号の音量 (左図 黄線)
5. 音響信号の区間を選択し、その区間のスペクトルを表示 (右図)

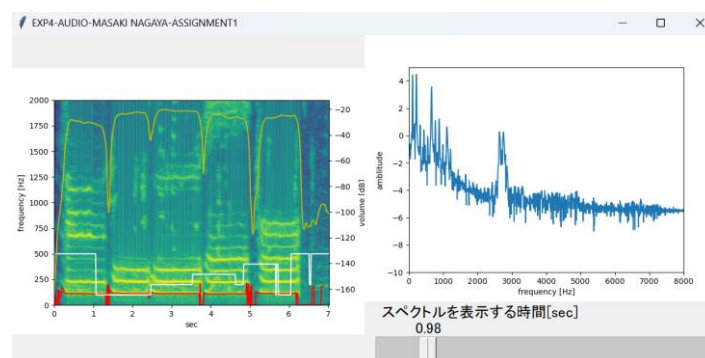


図 1-1 表示画面

次に、図 1-2 では以下を実装した。図 1-1 と同様の実装は省略する。

6. 音響信号のクロマグラム（左上図）
7. コードの識別を行った結果
8. 音響信号の再生ボタン（左上中央ボタン「Play Audio」）
9. NMF のよるスペクトログラムの分解の結果（左下図 2 つ）
10. 音響信号の波形（右下図）

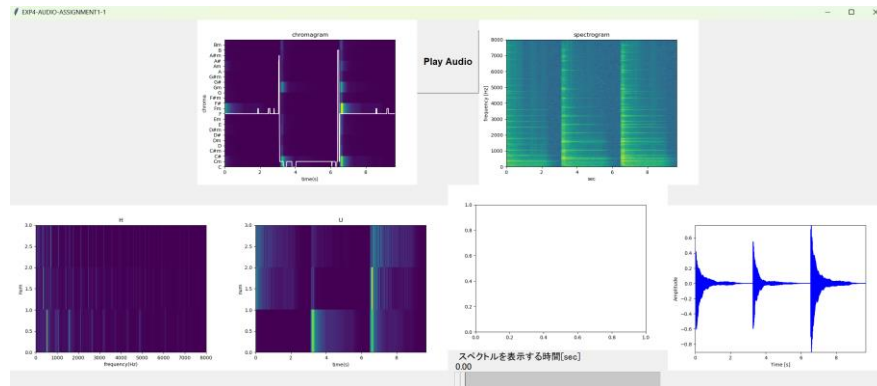


図 1-2 表示画面

1. 音響信号のスペクトログラム

音声データをフーリエ変換した振幅スペクトルを各フレームに対して求め、縦軸を周波数、横軸を時間として画像として表示している。

2. 音響信号の基本周波数（赤線）

自己相関を用いて、各フレームに対して基本周波数を求め、(1) のスペクトログラムに重ねている。また、ここでは、ゼロ交差数を用いて、有声音と判断された部分のみ表示している。

3. 母音などの何らかの識別を行った結果（白線）

まず、学習フェーズで、20 次までのケプストラム係数を用いて、20 次元正規分布の確率モデルを構築する。モデルでは「あ」、「い」、「う」、「え」、「お」の 5 つの母音のそれぞれに対して対数尤度を計算し、対数尤度が最大となるような母音を返り値とする。音素の認識フェーズでは、過去 30 フレームのケプストラム係数の行列をモデルの入力としている。出力は、(1) の図に重ねて表示している。あ、い、う、え、おの順に値が等間隔に大きくなるように定義した。

4. 音響信号の音量

音量の定義式に従って計算した音量を (1) の図に重ねて表示している。音量の縦軸は、ボリュームなので、図に新しく軸を追加している。

5. 音響信号の区間を選択し、その区間のスペクトルを表示（右図）

(1) で計算したスペクトルの行列に対して、スライダーで決定した一つのベク

トルを取得しスペクトルとして表示している。横軸に周波数、縦軸にその大きさをとっている。

6. 音響信号のクロマグラム (左上図)

クロマベクトルを計算し、縦軸をクロマベクトルのインデックス、横軸を時間として画像として表示している。

6. コードの識別を行った結果

(6) に重ねる形で、推定したコードを表示している。ここでは、コードの名前を縦軸のラベルに取っている。

8. 音響信号の再生ボタン (左上中央ボタン「Play Audio」)

ボタンを押すと、音声再生される。

9. NMF のよるスペクトログラムの分解の結果 (左下図2つ)

NMF を行い、 H と U を表示する。 H はスペクトル情報、 U は時間情報である。

10. 音響信号の波形 (右下図)