計算機科学実習Ⅳ

2024/11/8

長谷真暉

1029341181

課題 2 音響信号操作 GUI 作成

音響信号ファイルを読み込み，音響信号を操作するグラフィカルユーザーインタフェースを作成せよ．少なくとも以下の操作を可能にすること．

1. ボイスチェンジ

2. トレモロ

加えて，このインタ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．

• エフェクトのパラメータを GUI 上で変更できる．

• 選択した区間のみにエフェクトを施す．

• ビブラート，エコー，コンプレッサー，ディレイ，LPF（ローパスフィルタ）など他のエフェクタを実装する．

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

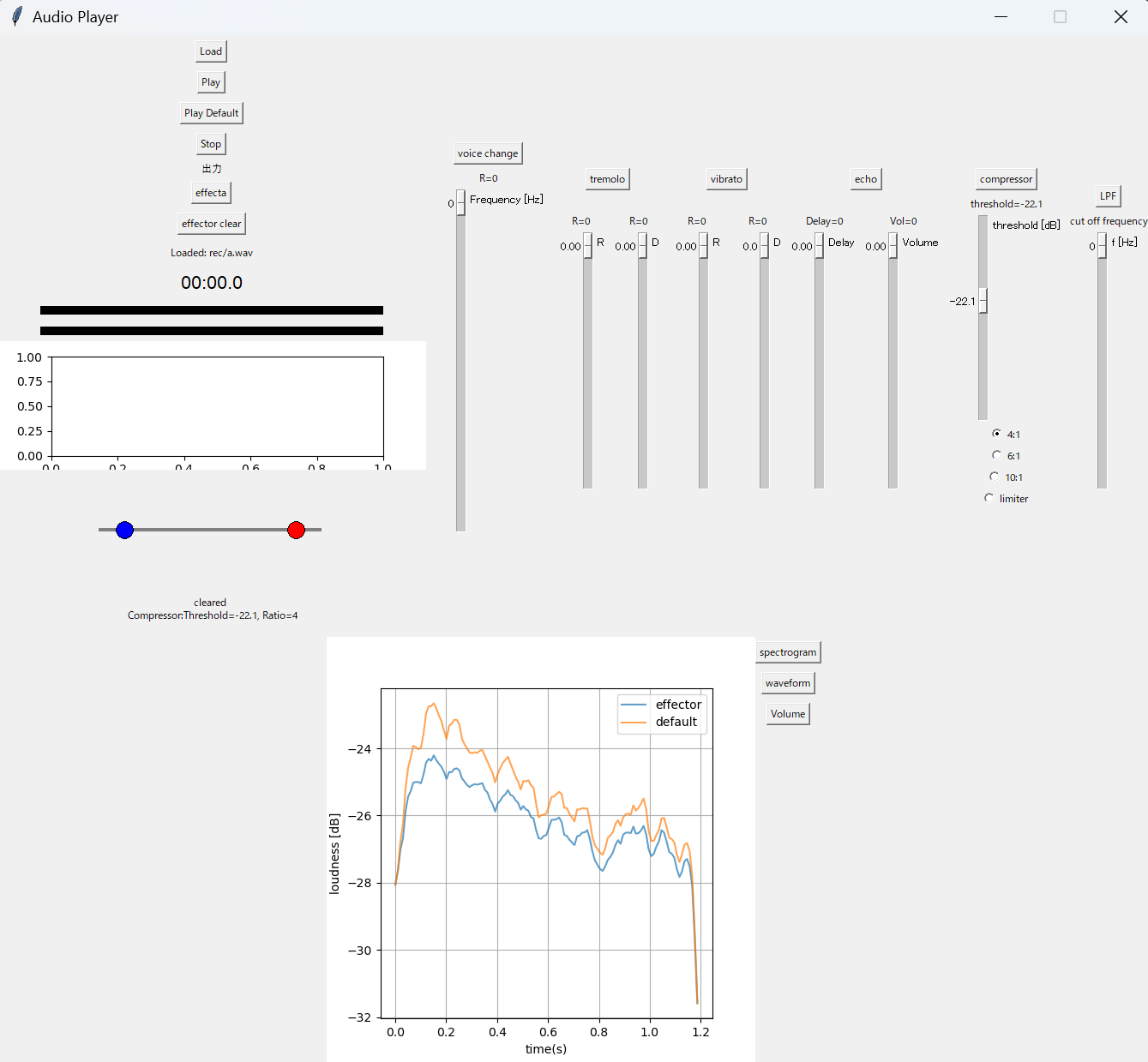


図1-1　表示画面

　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）

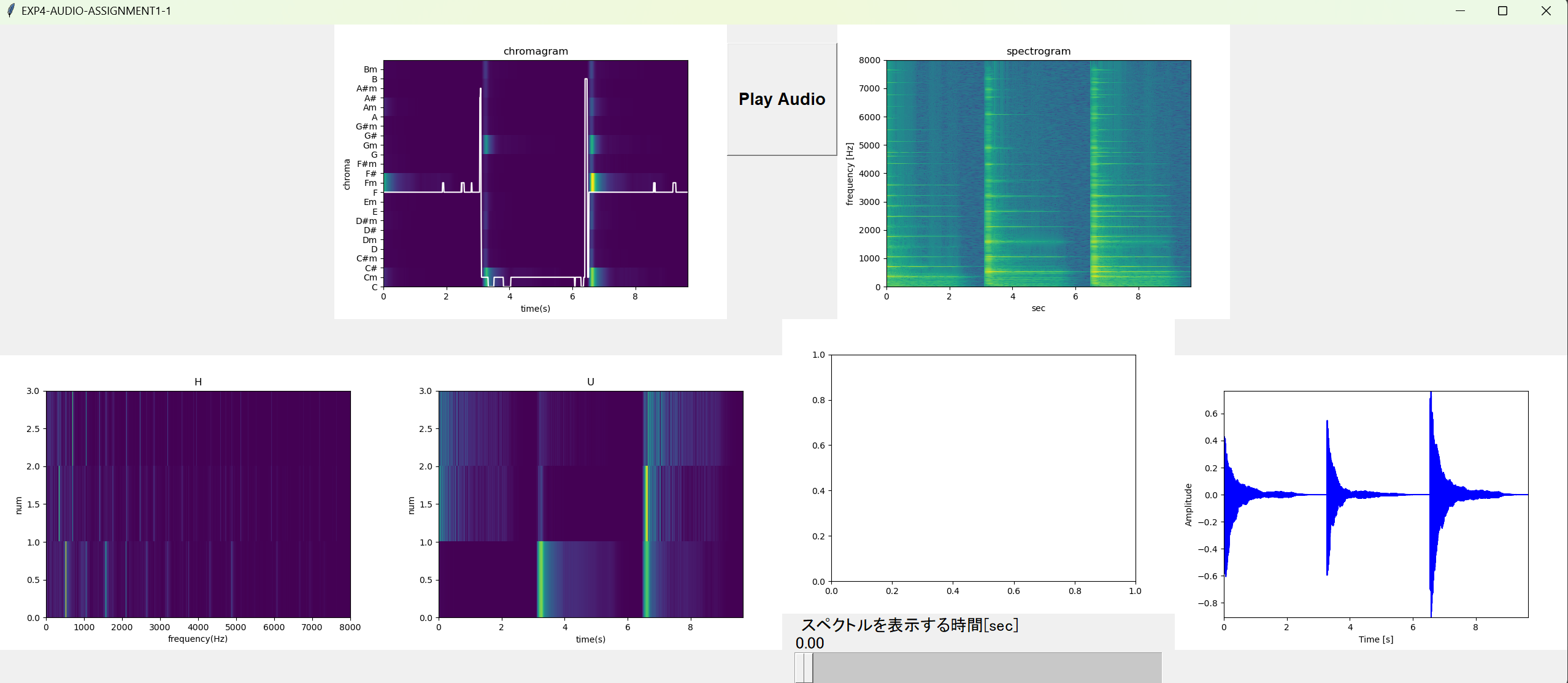


図1-2　表示画面

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。

1. 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。

1. 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。

1. 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。

1. 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、スライドバ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。

1. 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、音声が再生される。

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）