

計算機科学実験 4 音声 レポート 1

安済翔真

2023 年 12 月 26 日

目次

1	機能・プログラム説明	2
1.1	ファイル選択	2
1.2	スペクトログラムの表示	2
1.3	基本周波数の表示	3
1.4	音程の表示	4
1.5	音声の再生・停止	6
1.6	表示区間・再生区間の限定	6
2	実行例とテスト	6
2.1	ファイル選択	6
2.2	スペクトログラムの表示	6
2.3	基本周波数の表示	6
2.4	音程の表示	6
2.5	音声の再生・停止	6
2.6	表示区間・再生区間の限定	6
3	工夫点	6
3.1	ファイル選択機能	6
3.2	区間選択機能	6
3.3	音程の表示	6
3.4	音声の再生・停止機能	6
4	考察	6

1 機能・プログラム説明

課題 1 の演習として複数の機能を実装した。この章ではそれらの機能について説明する。

1.1 ファイル選択

ファイル選択機能を実装した。ファイル選択ボタンを押すと、ファイル選択ダイアログが表示される。選択したファイルを読み取り、表示対象の音声データとして扱う。

ファイル読み取りのプログラムはコード 1, 2 の通りである。コード 1 は tkinter の filedialog を用いてファイル選択ダイアログを表示し、選択したファイル名を取得する。コード 2 は librosa を用いて音声データを読み取る。

1 ファイル読み取り 1

```
1  commentstylefilename = commentstyletk. commentstylefiledialog.  
    commentstyleaskopenfilename()  
2  commentstyleself. commentstyle_c. commentstyleload_file(  
    commentstylefilename)
```

2 ファイル読み取り 2

```
1  commentstyledef commentstyleload_waveform( commentstylefilename):  
2  commentstylex, commentstyle_ = commentstylelibrosa.  
    commentstyleload( commentstylefilename, commentstyleSR=  
    commentstyleSR)  
3  commentstylereturn commentstylex
```

1.2 スペクトログラムの表示

選択したファイルの音声のスペクトログラムを表示する機能を実装した。コード 3 はスペクトログラムを計算するプログラムである。numpy の hamming 関数を用いて窓掛けを行い、FFT を行う。その後、対数振幅スペクトルを計算し、配列に保存する。

3 スペクトログラム計算

```
1  commentstyledef commentstylespectrogram( commentstylewaveform,  
    commentstylesize_frame, commentstylesize_shift):  
2  commentstylespectrogram = []  
3  commentstylehamming_window = commentstylenp. commentstylehamming(  
    commentstylesize_frame)  
4  
5  commentstylefor commentstylei commentstylein commentstylenp.  
    commentstylearange(0, commentstylelen( commentstylewaveform) -  
    commentstylesize_frame, commentstylesize_shift):  
6  commentstyleidx = commentstyleint( commentstylei)
```

```

7      commentstylex_frame = commentstylewaveform[ commentstyleidx:
          commentstyleidx + commentstylesize_frame]
8
9      # 窓掛けしたデータを commentstyleFFT
10     commentstylefft_spec = commentstylenp. commentstylefft.
        commentstylerfft( commentstylex_frame *
        commentstylehamming_window)
11
12     # 振幅スペクトルを対数化
13     commentstylefft_log_abs_spec = commentstylenp. commentstylelog(
        commentstylenp. commentstyleabs( commentstylefft_spec))
14
15     # 配列に保存
16     commentstylespectrogram. commentstyleappend(
        commentstylefft_log_abs_spec)
17     commentstylereturn commentstylespectrogram

```

コード 4 はスペクトログラムを表示するプログラムである。matplotlib の imshow 関数を用いてスペクトログラムを表示する。

4 スペクトログラム表示

```

1  commentstyleself. commentstyle__ax. commentstyleimshow(
2      commentstylenp. commentstyleflipud( commentstylenp.
        commentstylearray( commentstylespectrogram). commentstyleT),
3      commentstyleextent=[0, commentstylelen(
        commentstylespectrogram), 0, commentstyleSR / 2],
4      commentstyleaspect=' commentstyleauto',
5      commentstyleinterpolation=' commentstylenearest',
6  )
7  commentstyleself. commentstyle__ax. commentstyleset_ylim(0, 3000)

```

1.3 基本周波数の表示

選択したファイルの音声の基本周波数を表示する機能を実装した。コード 5 は基本周波数を計算するプログラムである。まず、numpy の correlate 関数を用いて自己相関係数を計算する。その後、ピークを検出し、ピークのインデックスを取得する。最後に、ピークのインデックスから基本周波数を計算する。

5 基本周波数計算

```

1  commentstyledef commentstyleget_f0( commentstylewaveform,
        commentstylesampling_rate):
2      commentstyleautocorr = commentstylenp. commentstylecorrelate(
        commentstylewaveform, commentstylewaveform, ' commentstylefull')
3      commentstyleautocorr = commentstyleautocorr[ commentstylelen(
        commentstyleautocorr) // 2:] # 不要な前半を捨てる

```

```

4
5  # ピークを検出
6  commentstylepeak_indices = [ commentstylei  commentstylefor
    commentstylei  commentstylein  commentstylerange( commentstylelen(
    commentstyleautocorr))  commentstyleif  commentstyleis_peak(
    commentstyleautocorr,  commentstylei)]
7  commentstylepeak_indices = [ commentstylei  commentstylefor
    commentstylei  commentstylein  commentstylepeak_indices
    commentstyleif  commentstylei != 0] # 最初のピークは除く
8
9  commentstyleif  commentstylelen( commentstylepeak_indices) == 0:
10     commentstylereturn 0
11
12  commentstylemax_peak_index =  commentstylemax(
    commentstylepeak_indices,  commentstylekey= commentstylelambda
    commentstyleindex:  commentstyleautocorr[ commentstyleindex])
13
14  # 基本周波数を推定
15  commentstylef0 =  commentstylesampling_rate /
    commentstylemax_peak_index
16  commentstylereturn  commentstylef0

```

コード 6 は基本周波数を表示するプログラムである。matplotlib の plot 関数を用いて基本周波数を表示する。スペクトログラムと同じ x 軸を用いるため、x 軸のデータはスペクトログラムのデータを用いる。

6 基本周波数表示

```

1  commentstylex_data =  commentstylenp. commentstylelinspace(0,
    commentstylelen( commentstylespectrogram),  commentstylelen(
    commentstylef0s))
2  commentstyleself. commentstyle__ax. commentstyleplot(
    commentstylex_data,  commentstylef0s)

```

1.4 音程の表示

選択したファイルの音声の音程を表示する機能を実装した。コード 7 は音程を計算するプログラムである。nn2hz 関数は MIDI ノートナンバーを周波数に変換する関数である。shs 関数はスペクトルを用いて音程を計算する関数である。候補の音程 (NOTES) の各周波数について、スペクトルの対数振幅スペクトルを用いて尤度を計算する。尤度を計算する際は、対象の音程の「倍音」の強さも考慮した。特に、倍音の強さは 0.8 の指数関数的な減衰を考慮することで精度が大幅に向上した。

7 音程計算

```

1  commentstyledef  commentstylenn2hz( commentstylenn):
2      commentstylereturn 440.0 * 2 ** (( commentstylenn - 69) / 12.0)

```

```

3
4  commentstyledef commentstyleshs( commentstylespectrum,
    commentstylesample_rate, commentstylesize_frame):
5  commentstylelikelihood = commentstylenp. commentstylezeros(
    commentstylelen( commentstyleNOTES))
6  commentstylefor commentstylei commentstylein commentstylerange(
    commentstylelen( commentstylelikelihood)):
7  commentstylebase_freq = commentstylenn2hz( commentstyleNOTES[
    commentstylei])
8  commentstylefor commentstylej commentstylein
    commentstylerange(1, 16):
9  commentstylefreq = commentstylebase_freq * commentstylej
10  commentstylefft_idx = commentstyleint( commentstylefreq *
    commentstylesize_frame / commentstylesample_rate)
11  commentstylelikelihood[ commentstylei] += 0.8** commentstylej *
    commentstylenp. commentstyleexp( commentstylespectrum[
    commentstylefft_idx])
12  commentstylereturn commentstyleNOTES[ commentstylenp.
    commentstyleargmax( commentstylelikelihood)]

```

コード 8 は音程を表示するプログラムである。matplotlib の plot 関数を用いて音程を表示する。yticks で y 軸のラベルを設定している。ラベルの内容は C3 から B4 までの音程を表示することで、表示される音程がわかりやすくなるようにした。

8 音程表示

```

1  commentstyleplt. commentstyleplot( commentstylelist( commentstylemap(
    commentstylelambda commentstylex: commentstylex -
    commentstyleNOTES[0], commentstylemelody)))
2  commentstyleplt. commentstyleyticks( commentstylenp.
    commentstylearange(24),
3      commentstylelist([" commentstyleC3", " commentstyleC#3", "
    commentstyleD3", " commentstyleD#3", " commentstyleE3", "
    commentstyleF3", " commentstyleF#3", " commentstyleG3",
4      " commentstyleG#3", " commentstyleA3", "
    commentstyleA#3", " commentstyleB3", "
    commentstyleC4", " commentstyleC#4", "
    commentstyleD4", " commentstyleD#4",
5      " commentstyleE4", " commentstyleF4", "
    commentstyleF#4", " commentstyleG4", "
    commentstyleG#4", " commentstyleA4", "
    commentstyleA#4", " commentstyleB4"]])
6  )

```

1.5 音声の再生・停止

1.6 表示区間・再生区間の限定

2 実行例とテスト

2.1 ファイル選択

2.2 スペクトログラムの表示

2.3 基本周波数の表示

2.4 音程の表示

2.5 音声の再生・停止

2.6 表示区間・再生区間の限定

3 工夫点

3.1 ファイル選択機能

3.2 区間選択機能

start の値が end の値より大きくなならないようバリデーションを行った。

3.3 音程の表示

3.4 音声の再生・停止機能

4 考察