生体情報システム
2006 年 6 月 12 日

課題 6: 聴覚系の情報処理と音声

（提出締切 7 月 3 日）

目的： コンピュータで周波数分析（フーリエ変換）をおこない，その意味を感じ取り，フーリエ変換のイメージを確固たるものにする．octave の練習も兼ねる．例えば以下のようなことを確認する．1) コンピュータで扱える情報は離散的な情報のため，手計算で変換したときほどきれいな結果は得られない．どのくらい正確にもとの周波数成分をとり戻せるか．2) 位相を変化させても $|F(\omega)|$ の値は変化しない（波形は変わっても人間の耳にも同じ音に聞こえる）．

課題（必須）：

以下の演習をおこなうと，不思議な点，おかしな点など，疑問点がたくさんでてくるはずである．それを一つ一つ記述し，フーリエ解析の教科書を読むなどして，できるだけ解決せよ．フーリエ変換は奥が深い．フーリエ変換の離散版は講義では説明していないが，フーリエ変換を離散化したものと思ってよい．フーリエ変換の性質などは応用数学 2 で使った教科書や google で探す．octave は使ったことがないかもしれない．octave の使い方は使いながら学び，やはり google で探す．凝りだすときりがないので，生活に支障のない適当なところで切り上げる（A4 1 枚のレポートではでは少なすぎる）．

課題（自由）：

C 言語等でさまざまな波データを作成し，それをフーリエ変換してみる．

評価基準：

まず疑問点の量（この演習が何をしているか大ざっぱにでも分かって演習をしていれば多くの疑問点がでてくる）．次に，疑問点の質，実際に解決できたかどうかは，その後.

補足：

サンプルデータを講義の web ページよりダウンロードする．

`http://www.cs.miyazaki-u.ac.jp/~date/lectures/2006bis/`

レポートの LaTeX を使った簡単な書き方はリンクでたどれるようにしている．

1 周波数分析

1.1 離散フーリエ変換とパワースペクトル

実際に具体的な信号の周波数分析をおこなってみよう．周波数分析は信号をフーリエ変換することであるが，以下に示す手順に従えば，なんとなくフーリエ変換のイメージがつかめるようになる．細かい数学的な性質については，各自，必要に応じて応用数学 2 で使用したフーリエ解析の教科書を読み返して理解を深めよう．まずは octave を走らし，以下を実行してみる．

```
% octave
octave:1> t = 0: 1/44100 : 5;         % 時間軸の作成
octave:2> y = 0.9 * sin(2*pi*440*t); % 440Hz の正弦波を作成
octave:3> plot(y);
octave:4> plot(y(1:1000);
octave:5> Y = fft(y, 1024);           % 最初の 1024 個のデータをフーリエ変換
octave:6> f = (0:512)/1024*44100;     % 周波数軸を作成
octave:7> plot(f,20*log10(abs(Y(1:513)))); % フーリエ変換後のデータを表示（絶対値）
```

解説：

$y(t) = 0.9\sin(2\pi f t)$ という周波数 $f = 440$ Hz の正弦波信号を作成．コンピュータでは信号は離散的にしか保持できない．ここでは信号 $y(t)$ を 1 秒間に $f_s = 44100$ 個サンプリングしたデータを 5 秒間記録したデータを作成したと思っていい．次に，最初の 1024 個のデータを（離散）フーリエ変換している．これにより 1024 個のデータは 1024 個のデータに変換される．Y は 1024 個のデータからなるベクトルである（それぞれの要素は複素数）．6 行目では周波数の軸のスケールを真の周波数にプロットするために変換している（だけ）．7 行目では，フーリエ変換した結果 $Y(f)$ を表示（正確には絶対値 $20\log_{10}|Y(f)|$ dB を表示）．実数値の信号 $y(t)$ をフーリエ変換した結果 $Y(f)$ に対しては $Y(-f) = \overline{Y(f)}$ が成り立つ．したがって $|Y(-f)| = |Y(f)|$ であり 1024 個のうち半分は同じ値なので，512 個表示しておけば十分である．表示されるグラフの横軸は周波数で，440 Hz にピークがあることを確かめる．440Hz にピークがあるのは分かりにくいかもしれない．そのときは

```
octave:8> plot(f(1:50),20*log10(abs(Y(1:50))));
```

などとすれば判別できるであろう．

1.2　復習

波をフーリエ変換した結果はどうなるか，典型的な波データに対してだけは記憶しておこう．

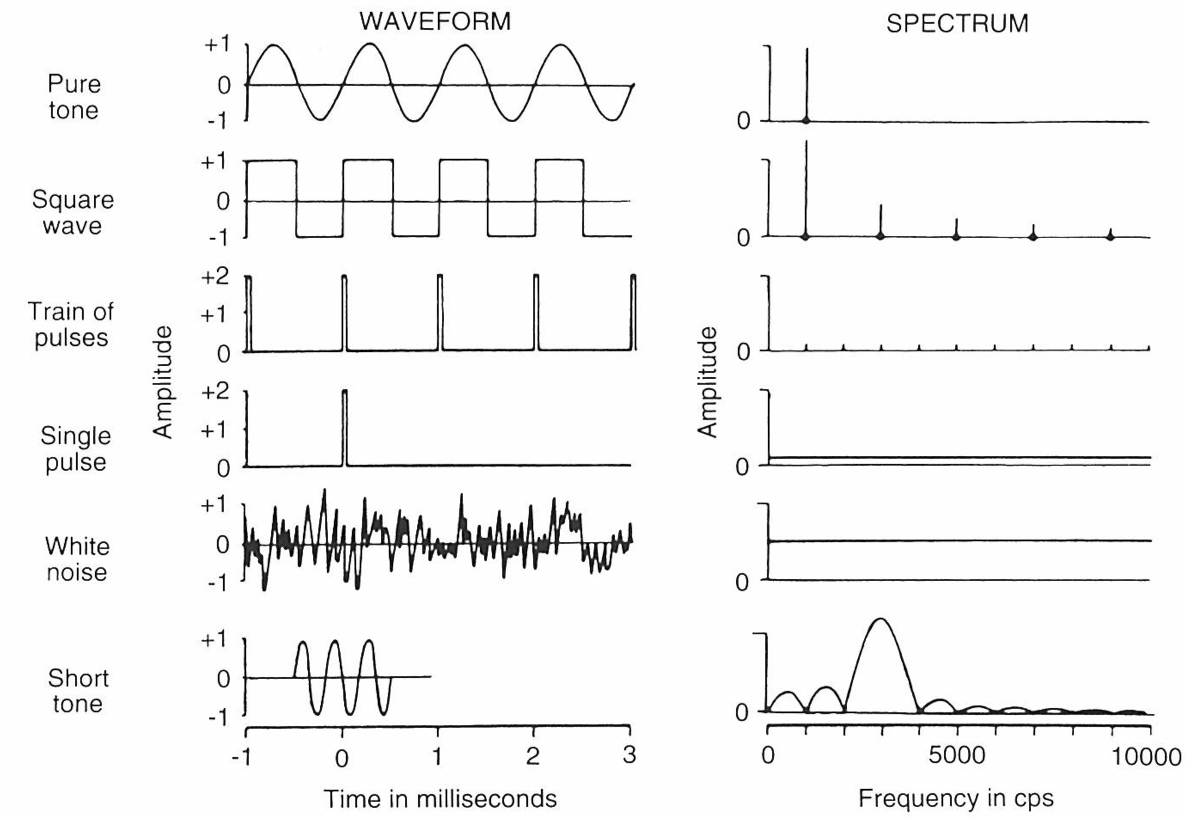

FIGURE 1.3　On the left are shown the waveforms of some common auditory stimuli, and on the right are the corresponding spectra. The periodic stimuli (pure tone, square wave, and train of pulses) have line spectra, while the nonperiodic stimuli (single pulse, white noise, and short tone burst) have continuous spectra.

1.3　複合波

配布資料（6 月 5 日）の⑩ を実行してみよう．

```
octave: 9> y1 = sin(2*pi*250*t) + sin(2*pi*500*t) + sin(2*pi*750*t);
octave:10> y2 = sin(2*pi*250*t) + sin(2*pi*500*t + pi/2) + sin(2*pi*750*t);
octave:11> plot(y1(1:600));
octave:12> plot(y2(1:600));
octave:13> Y1 = fft(y1, 32768);   % 32768 = 2^15
octave:14> Y2 = fft(y2, 32768);
octave:15> f = (0:16384)/32768*44100;
octave:16> plot(f,20*log10(abs(Y1(1:16385))));
octave:17> plot(f(1:1000),20*log10(abs(Y1(1:1000))));
octave:18> hold
octave:19> plot(f(1:1000),20*log10(abs(Y2(1:1000))));
```

※ 1.1 ではデータの 1024 点をフーリエ変換したが，ここでは 32768 点に対してフーリエ変換をおこなった．何が変わったか，最初の結果と比較してよう．

1.4　白色ノイズ

```
octave:20>  y3 = rand(5*44100,1);
octave:21>  Y3 =  fft(y3);
octave:22>  plot(f(1:1000),20*log10(abs(Y3(1:1000))));
```

1.5　窓掛け

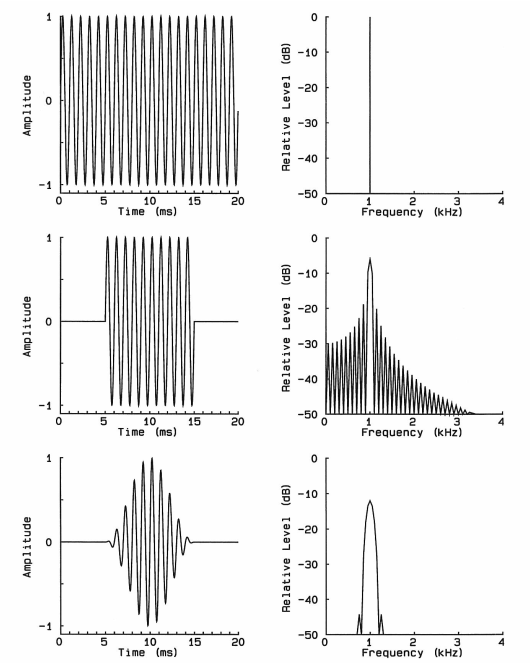

FIGURE 1.4　The top left panel shows part of the waveform of a 1-kHz sinusoid. The waveform should be pictured as continuing indefinitely. The spectrum, shown in the top right panel, contains a single line at 1 kHz. The middle left panel shows a segment of the 1-kHz sinusoid obtained by applying a 10-ms-long rectangular window. The middle right panel shows the spectrum of the windowed sample of the sinusoid. It contains multiple peaks and dips spread over a wide frequency range, although the biggest peak is still at 1 kHz. The bottom left panel shows the effect of applying a weighting function to the sinusoid which tapers off at the edges. The bottom right panel shows the spectrum of the waveform in the bottom left.

実際に解析するデータは，ある短い時間では周期信号になっている場合が多いが，通常，有限で周期もない．フーリエ変換は対象とするデータが，ある周期をもって無限に繰り返すことを前提としているので，データの両端が不連続であると，本来は存在しない周波数成分が現れる．本来，存在しない周波数成分はできるだけ取り除きたい．そのためには両端がゼロになる「窓」をかけてからフーリエ変換をおこなう．配布資料（6 月 5 日）の⑨ 参照．

```
octave:23> hold off
octave:24> load 'short_tone.dat';
octave:25> y4 = short_tone;
octave:26> plot(hamming(1024)); % 窓関数
octave:27> plot (y4);
octave:28> hold
octave:29> plot (y4(1:1024).*hamming(1024));
octave:30> hold off
octave:31> Y4 = fft(y4(1:1024));
octave:32> Y5 = fft(y4(1:1024).*hamming(1024));
octave:33> f = (0:512)/1024*44100;
octave:34> plot(f(1:100),20*log10(abs(Y4(1:100))));
octave:35> hold
octave:36> plot(f(1:100),20*log10(abs(Y5(1:100))));
```

1.6 音声波形とスペクトログラム

```
octave:37> load  'ohayo57.txt'  % 変数 sig に音声データを入力
octave:38> plot (sig); % 音声波形の表示
octave:39> plot (sig(17000:20000)); % 音声信号の一部を表示
octave:40> plot (sig(16000:17024));
octave:41> sp = fft(sig(16001:17024).*hamming(1024));
octave:42> f = (0:512)/1024*44100;
octave:43> plot(f,20*log10(abs(sp(1:513))));
```

フーリエ変換にかけるサンプル点の数を増やせば，きれいなグラフが描ける．ただし，現実の信号は時々刻々変化するので，それほど多くのサンプル点はとれないかもしれない．1024 点（サンプル周波数 44100 Hz とすると，1024/44100 sec）は 0.023 秒間の信号を解析していることになる．この解析を時間をずらして何度もやるとスペクトログラムが描ける．

補足：

```
octave:44> whos
```

とすれば，いまどんな変数がどんな型で使われているか表示される．また abs という関数の使い方がわからなければ

```
octave:45> help abs
```

とすれば，その使い方を教えてくれる．レポートを LaTeX で書く場合，グラフを eps ファイルに出力しておくと直接張り付けることができてとても便利である．octave でプロットした結果を eps ファイルとして出力するには，例えば

```
octave:46> __gnuplot_set__ term postscript
octave:47> __gnuplot_set__ output "hoge.eps"
octave:48> plot (y4);
```

とすればいい．もとにもどすには

```
octave:49> __gnuplot_set__ term X11
```

としておく．

```
octave:50> exit  # octave の終了
% ggv file.eps   # グラフを見る
```