

スペクトルによる音声解析

2007MI215 清水理貴

指導教員：杉浦洋

1 はじめに

本研究では、楽音・音声の音質の違いが周波数構成にどのように反映されるかを調べるため、スペクトルを調べる。短いデータからもスペクトルの推定が可能であり、またスペクトルの分解能が極めて高いMEM スペクトルを使う。そして、MEM スペクトルの計算法には Burg アルゴリズムの計算方法を用いる。

2 フーリエ解析

フーリエ解析とは、「周期的な波として捉えられる現象は全て単純な正弦波 (sine, cosine) の重ね合わせで表現することができる」という理論の下、積分などの計算を行うことで振動している音などを解析する手法である。関数 $x(t)$ ($-\infty < x < \infty$) に対して、そのフーリエ変換は、

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-i2\pi ft} dt$$

とすると、

$$x(t) = \int_{-\infty}^{\infty} X(f)e^{i2\pi ft} df$$

と書ける。よって、フーリエ成分 $X(f)$ は $x(t)$ に含まれる周期 f の波の振幅であり、 $|X(f)|^2$ はその強さ・エネルギーを表している。しかし、有限区間での周期関数などでは、エネルギー $|X(f)|^2$ も有限と考えられ、エネルギーでスペクトルを定義してもよいが、区間が無限の場合では、単位時間あたりの平均エネルギーをとったパワースペクトル密度関数 $P(f)$ を

$$P(f) = \lim_{T \rightarrow \infty} \frac{1}{T} \left| \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)e^{-i2\pi ft} dt \right|^2$$

と定義する。これによって、音声信号の中に、周期 f の成分がどの程度含まれているかがわかる。

3 Yule-Walker 方程式

自己回帰式の係数を決定する Yule-Walker の方程式について述べる。音を不規則性を含む時間変数 $x(t)(x_i = x(i\Delta t))$ と考え、微分方程式

$$k_m \frac{d^m x}{dt^m} + k_{m-1} \frac{d^{m-1} x}{dt^{m-1}} + \dots + k_0 x = n(t), \quad (1)$$

あるいは、離散表示により

$$x_i + \gamma_1 x_{i-1} + \dots + \gamma_m x_{i-m} = n_i \quad (2)$$

($n(t), n_i$: ランダム入力) で表されるとする。この式は x_i の予測値 (\hat{x}_i) を求める回帰予測式で、

$$\hat{x}_i = -\gamma_1 x_{i-1} - \gamma_2 x_{i-2} - \dots - \gamma_m x_{i-m} \quad (3)$$

であり、係数 $(1, \gamma_1, \gamma_2, \dots)$ は予測誤差 $\varepsilon_i = x_i - \hat{x}_i$ を白色雑音 $\{n_i\}$ にするための数値フィルターとみなせる。式 (2) は再帰型の数値フィルターである。

上記に、順に x_{i-k} を掛けて期待値

$$C_k = E(x_{i-k}x_i) = \lim_{n \rightarrow \infty} \frac{1}{2n+1} \sum_{i=-n}^n x_{i-k}x_i \quad (4)$$

をとれば、 $E(x_{i-k}n_i) = 0$ であることを考慮し、 $P_m = E(n_i n_i)$ とおけば、

$$\begin{aligned} C_0 + \gamma_1 C_{-1} + \gamma_2 C_{-2} + \dots + \gamma_m C_{-m} &= P_m, \\ C_k + \gamma_1 C_{k-1} + \gamma_2 C_{k-2} + \dots + \gamma_m C_{k-m} &= 0 \\ (k &= 1, 2, \dots, m+1) \end{aligned} \quad (5)$$

の関係 (Yule-Walker 方程式) が得られる。 $C_j = C_{-j}$ であるから、式 (4) は、

$$\begin{bmatrix} C_0 & C_1 & \dots & C_m \\ C_1 & C_0 & \dots & C_{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ C_m & C_{m-1} & \dots & C_0 \end{bmatrix} \begin{bmatrix} 1 \\ \gamma_1 \\ \vdots \\ \gamma_m \end{bmatrix} = \begin{bmatrix} P_m \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (6)$$

と書ける。この方程式を解いて、 $\gamma_1, \gamma_2, \dots, \gamma_m$ と P_m が得られたとする。このとき式 (2) に $x(n\Delta t) = \int X(f)e^{i2\pi fn\Delta t} df$ を代入して変形すれば、容易に

$$P(f) = \Delta t P_m \left/ \left| 1 - \sum_{k=1}^m a_k \exp(i2\pi fk\Delta t) \right|^2 \right. \quad (7)$$

が得られる。これが、MEM で推定したスペクトルである。

4 MEM のアルゴリズム

方程式 (6) の解は m に依存するので、改めて

$$(\gamma_1, \gamma_2, \dots, \gamma_m) = (\gamma_{m1}, \gamma_{m2}, \dots, \gamma_{mm}),$$

と書く。

Burg 法では、 $m = 1$ からはじめて γ_{mm} および P_m のみならず自己相関関数 C_m も順次漸化関係式により決定される。 C_m は x_i から直接計算されるのではないことに注意する。

また、Levinson の漸化式によって、フィルター係数 γ_{mk} には、次の漸化関係

$$\begin{bmatrix} 1 \\ \gamma_{m1} \\ \gamma_{m2} \\ \vdots \\ \vdots \\ \gamma_{mm} \end{bmatrix} = \begin{bmatrix} 1 \\ \gamma_{m-1,1} \\ \gamma_{m-1,2} \\ \vdots \\ \gamma_{m-1,m-1} \\ 0 \end{bmatrix} + \gamma_{mm} \begin{bmatrix} 0 \\ \gamma_{m-1,m-1} \\ \gamma_{m-2,m-2} \\ \vdots \\ \gamma_{m-1,1} \\ 1 \end{bmatrix} \quad (8)$$

が成立することによって,

$$\gamma_{mk} = \gamma_{m-1,k} + \gamma_{m,m}\gamma_{m-1,m-k} \quad (9)$$

と書ける. したがって, $\gamma_{m,m}$ が求められると, 式 (9) により, $\gamma_{mk}(1 \leq k < m)$ が計算できる. また,

$$P_m = C_0 + \gamma_{m1}C_1 + \dots + \gamma_{mm}C_m$$

について, 漸化式

$$P_m = P_{m-1}(1 - \gamma_{mm}^2) \quad (10)$$

が成立する.

(i) γ_{mm} の決定

Burg 法では, γ_{mm} は次のように計算される.

$$\gamma_{mm} = -2 \frac{\sum_{i=1}^{N-m} b_{mi}b'_{mi}}{\sum_{i=1}^{N-m} (b_{mi}^2 + b'_{mi}^2)} \quad (11)$$

(ii) 係数 $\gamma_{mk}(1 \leq k < m)$ の計算

γ_{mm} が求まったので, γ_{mk} は Levinson の漸化式

$$\gamma_{mk} = \gamma_{m-1,k} + \gamma_{mm}\gamma_{m-1,m-k} \quad (1 \leq k < m) \quad (12)$$

より計算する.

(iii) P_m の計算

Levinson の漸化式の所で導いた関係式から

$$P_m = P_{m-1}(1 - \gamma_{mm}^2) \quad (13)$$

と計算する.

(iv) 自己相関関数 C_m の計算

方程式 (6) の最下行の関係から, ラグ m の自己相関係数を,

$$C_m = -(\gamma_{m1}C_{m-1} + \gamma_{m2}C_{m-2} + \dots + \gamma_{mm}C_0) \quad (14)$$

で計算する. 式 (14) から C_m は係数 $\gamma_{mi}(i = 1, 2, \dots, m)$ に関係する. したがって, C_{m-k} は係数 $\gamma_{m-k,i}(i = 1, 2, \dots, m-k)$ のみに関係し, 新たな係数 γ_{mm} には無関係であるから再び求め直す必要はない.

以上 (i),(ii),(iii),(iv) により, すでに計算された, $\gamma_{m-1,i}(1 \leq k \leq m-1), P_{m-1}, C_{m-1}$ から, $\gamma_{m,i}(1 \leq k \leq m), P_m, C_m$ が計算できる. M の設定は実験的に行う. M の値が小さすぎると精度が出ず, 大きすぎるとデータの誤差に敏感に反応してしまう.

5 実験

Mathematica で, 以下の実験を行った. データ長は短く, また, M は少ない方がより計算時間がより小さいので, その中で音程の判別できるものを探す. それぞれ考慮して実験を行い, データ長と M の値を考えた.

5.1 実験 1 単音による実験

電子ピアノから録音した A(ラ)(440Hz) の音を, プログラムで解析した. その結果が図 1 である.

図 2 に, M の値が小さすぎる場合を示す. M の値が小さすぎると, 主要な音程は判別できるが, 同時に含まれる他の音程を判別することが難しい.

5.2 3つの音程の場合

電子ピアノから C(ド)(261.6Hz), E(ミ)(329.6Hz), G(ソ)(392Hz) の3音を同時に鳴らしたものを録音し, 調査した. 図 3 の結果より, 主要な音程は確認できた. 主音となるのが, 262Hz, 331Hz, 394Hz である. また, それぞれに倍音の周波数があらわれている.

図は横軸を周波数, 縦軸をパワースペクトルにとった片対数グラフである.

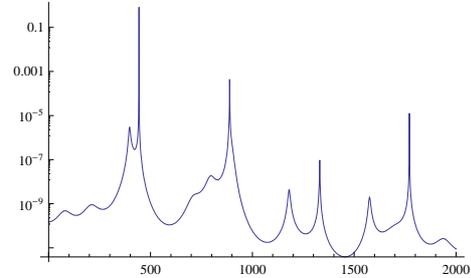


図 1 0.08 秒, M=350

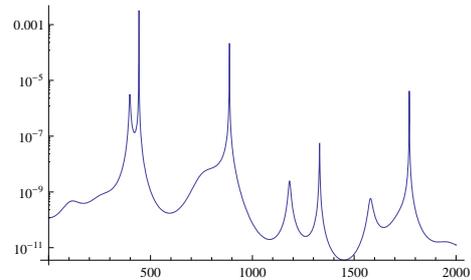


図 2 0.16 秒, M=300

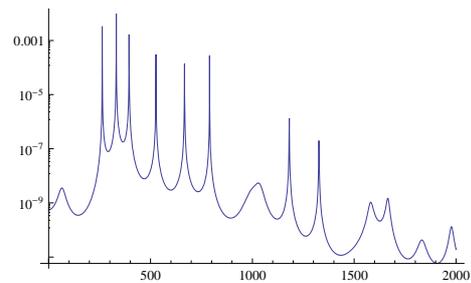


図 3 0.08 秒, M=350

6 おわりに

MEM スペクトルによって, 楽音・音声かどのような周波数であるか, 理解することができた. またそれらが様々な要素を含み, 音色として表現されているかが, グラフとして確認することができた.

これからの課題として, より複雑な楽音・音声に対しても調べていきたい.

参考文献

[1] 日野幹雄:『スペクトル解析』. 朝倉書店, 1977.