6-17

波形解析における変数削減量の検討

― 波形データの軽量化 ―

日大生産工(学部)　○則武　秀昌

日大生産工　　矢野　耕也

１．はじめに

情報技術の発達に伴いコンピューターを用いたデータ解析が一般的となってきたが、肥大したデータの処理にかかる処理時間とコストの問題は依然として解決されていない。

現在、高性能なコンピューターを手軽に入手できるほど価格も下がったことから、コンピューターを使ったデータ解析は個人ユーザーレベルでも日常的に行われている。今の技術ではコンピューターは 2020 年頃に性能限界を迎えることが言われているが、新しい技術として量子コンピューターなどが研究されており、その性能限界はまだ先の話となりそうである。

さらにコンピューターの処理能力が増えるにつれ、情報量の多いデータでも対応できるようになってきたことから、データに対して質の良さも求められるようになった。特に御、音圧や電圧、静止画・動画などのデータに関しては、より高精度・高解像度を求められ、今まで発見されなかった微量な情報までも取り扱われている。

しかしコンピューターの処理性能が高度なものになったとはいえ、膨大な情報量を持つデータをそのまま処理させるには時間がかかる。また、これらを満足させるコンピューターはコスト面において多大な負担になる。そのため、今のコンピューターはデータを処理するのに必要な性能を満たしていないという問題がある。

そこで、当研究では波形データを元に任意のデータ軽量化を行い、どの程度の軽量化が最適なのかをＭＴシステムを用いた解析による検証を試みた。

２．研究目的

消費電流の波形に限らず、波形データは特徴的な部分が特別な情報を持っている場合が多いことから、その特徴部分のみを抽出し、他のデータを削減していくことができれば、情報処理の手間を軽減しつつ、元データと同等の再現率が得られる可能性がある。

データを軽量化した場合、どの程度の再現率があるかを確認しながら、最適な軽量化の方法について検討していく。

３．解析方法

今回実験に用いるデータは、CD ドライブの動作時による消費電流値を波形データにしたものである。(図 1)

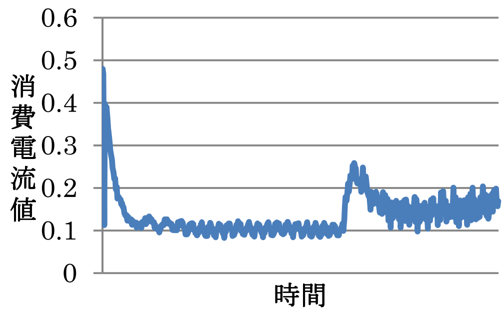

図 1　元データの消費電流値

このデータを元に等間隔にデータを削除し軽量化を行う。間隔については整数倍とし、それぞれについて MT システムによる評価を行っていく。

一例として、以下にデータ量を 1/2 にしたものと 1/3 にしたものを示す。

A study on the verification of data reduction in waveform analysislightweight waveform data

Hidemasa NORITAKE and Koya YANO

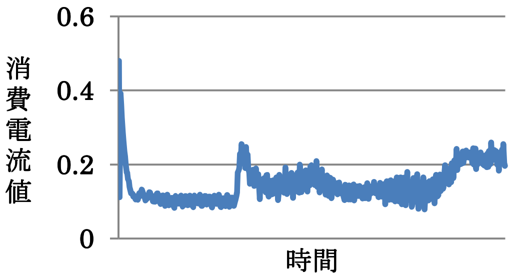

図 2　データ量 1/2

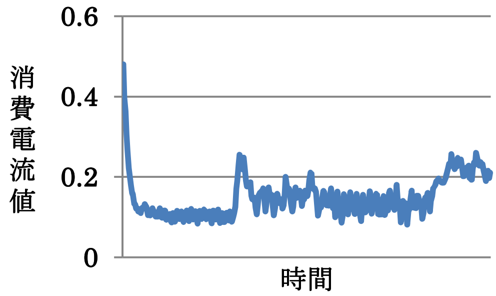

図 3　データ量 1/3

次に、MT システムによる評価方法の計算手順を示す。手順(1)から(12)のにより波形のパターンを距離という尺度評価することができる。

有効除数　$r = {m_1}^2 + {m_2}^2 \cdots + {m_{1000}}^2 \cdots (1)$

線形式

$L = m_1 X_{1.1} + m_2 X_{1.1} \cdots m_{1000} X_{1000.1} \cdots (2)$

感度　$\boldsymbol{\beta} = \frac{L}{r} \cdots (3)$

全変動　$S_T = {X_{1.1}}^2 + {X_{1.2}}^2 \cdots + {X_{1000.1}}^2 \cdots (4)$

比例項の変動　$S_\beta = \frac{L_2}{r} \cdots (5)$

誤差変動　$S_e = S_T - S_\beta \cdots (6)$

誤差分散　$V_e = \frac{S_e}{1000-1} \cdots (7)$

SN 比　$\boldsymbol{\eta} = \frac{1}{V_e} \cdots (8)$

$Y_1 = \beta \cdots (9)$　$Y_2 = \frac{1}{\sqrt{\eta}} \cdots (10)$

余因子行列　$A = \begin{bmatrix} V_{22} & -V_{21} \\ -V_{12} & V_{11} \end{bmatrix} \cdots (11)$

$$V_{11} = \frac{1}{n-1}\sum_{i=1}^{n}(Y_{1i} - \bar{Y}_1)^2$$

$$V_{12}(V_{21}) = \frac{1}{n-1}\sum_{i=1}^{n}(Y_{1i} - \bar{Y}_1)(Y_{2i} - \bar{Y}_2)$$

$$V_{22} = \frac{1}{n-1}\sum_{i=1}^{n}(Y_{2i} - \bar{Y}_2)^2$$

余因子行列距離

$$D = \frac{1}{2}\begin{bmatrix} V_{22}(Y_{11} - \bar{Y}_1)^2 \\ -V_{21}(Y_{11} - \bar{Y}_1)(Y_{21} - \bar{Y}_2) \\ -V_{12}(Y_{11} - \bar{Y}_1)(Y_{21} - \bar{Y}_2) \\ +V_{11}(Y_{22} - \bar{Y}_2)^2 \end{bmatrix} \cdots (12)$$

4．解析結果

データ数を減少させながら余因子行列の距離からＳＮ比を求めた結果、ＳＮ比の推移は図 3 のようになった。

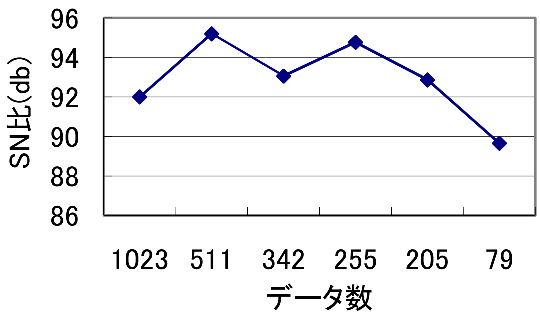

図 4　データ数の減少に伴う解析精度

5．考察とまとめ

ＳＮ比は、データ数を減少させたときの波形の情報の精度を示すものであるが、ある特定の減少領域を過ぎると、情報の損失、すなわちＳＮ比の低下が急激に発生する。このことは、データ数の省略にある種の限界点があることを示唆するものである。図 4 から判断して、データの省略は 1/4 程度が限界であることが考えられるが、逆に元情報を 1/4 程度(255 項目)まで落としても精度はそれほど変わらないことがわかる。すなわち、波形データの多変量解析において、項目数を 1/4 程度まで省略しても同等の結果が得られることが推測される。

なお今回は減少間隔を等間隔にして行ったが、不等間隔に減少間隔を取り特徴をより強く残す検証が今後の課題である。