

fNIRS データの Smith Waterman 法を利用した類似部分抽出システムの提案と評価

西井 琢真

Takuma NISHII

1 はじめに

近年、脳血流の変化を測定することにより非侵襲的に脳機能イメージングを行う装置である fNIRS(functional Near-Infrared Spectroscopy) や fMRI(functional Magnetic Resonance Imaging) が注目を集めている。これらの装置は、脳機能の解明に役立ち、種々の病理の判定や生体信号によるコンピュータ操作などに利用されている。fNIRS は、脳の計測部位 (チャンネル) ごとの脳血流の増減を測定することで脳の活性度を把握する装置である。この装置の性能は向上しつつあり、出力される時系列データ量が增大している。そのため、解析者が効率的にデータを解析できないという問題が生じている。その 1 つに解析者が実験結果を検討する際、出力された多数の時系列データの中で、どの時系列データに着目するか解析者に依存してしまうという問題がある。この問題を解決するためには、多チャンネルの時系列データ間において類似部分を探索し、解析者の注目する特徴的なデータを他の脳部位データから探索することが必要となる。そこで本論文では、画期的な類似部分探索アルゴリズムを提案し、それを実装したインターフェースを紹介する。

2 fNIRS と出力データ

fNIRS は、近赤外線分光法を用いた非侵襲の脳機能イメージング装置である。近赤外光を用いて脳血流の相対的変化量を多点で計測する。大脳皮質には脳機能の局在性があること¹⁾、神経活動による酸素代謝の亢進に伴い脳血管が拡張し脳血流が上昇する現象²⁾ (神経血管カップリング) が確認されていることにより、fNIRS を用いた脳機能の解析が可能となる。本論文では約 120 箇所を計測する日立メディコ社製の fNIRS である ETG-7100 を用いた。各計測部位はチャンネル (CH) と呼ばれ、0.1 秒ごとに血流の相対的変化量が出力される。本装置は、脳を前頭部 (24CH)、右側頭部 (22CH)、左側頭部 (22CH)、後頭部 (24CH)、頭頂部 (22CH) の 5 プローブに分けて扱う。前頭部のプローブの設置と CH の関係を Fig.1 に示す。fNIRS 出力データにおいて横軸は時間 [100ms]、縦軸は血流の相対的変化量 [mM・mm] を示す。

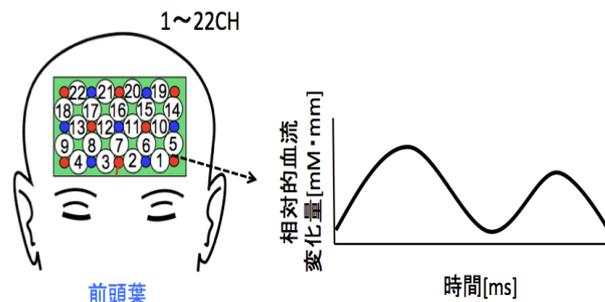


Fig. 1 前頭部のプローブ設置位置と出力データ

3 類似部分の抽出アルゴリズム

本論文では、fNIRS に適した前処理と Smith Waterman 法 (SW 法) を用いた 2 本の時系列データからの類似部分の抽出手法を提案する^{3) 4)}。Fig.2 に提案手法の流れを示す。提案手法は、fNIRS に適した前処理、時系列データの再量子化、類似部分の抽出の 3 つに分けられる。まず、fNIRS に適した前処理では時系列データの山と谷を伸張する。理由は後述する。次に、伸長された実数値の時系列データを文字列に変換する。例えば、Fig.2 の上部の時系列データは “BAABCCB” に、下部の時系列データは “ABCCBAA” に変換される。これを再量子化と定義する。時系列データを再量子化することによって、SW 法を適用することが可能となる。最後に、SW 法を用いて 2 つの文字列から類似部分の抽出を行う。SW 法による類似部分抽出により “BAABCCB” と “ABCCBAA” から類似部分として “ABCCB” を取り出すことができる。

以上の提案手法により、fNIRS から出力された 2 つの時系列データから類似部分を抽出することが可能となる。

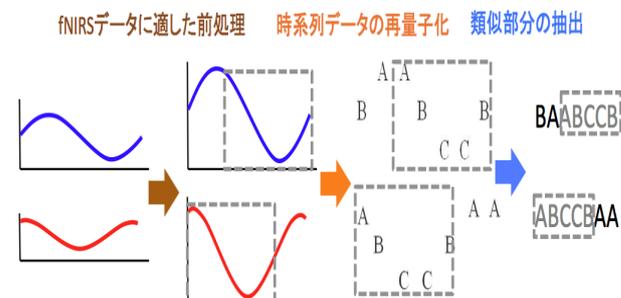


Fig. 2 類似部分の抽出アルゴリズムの流れ

3.1 fNIRS 出力データに適した前処理手法

fNIRS 出力データに適した前処理手法とは、大きさの度合いが異なる山の頂点を引き伸ばし、山を最大値に向けて伸長、最小値に向けて伸長する手法である。Fig.3 に伸長手法を示す。fNIRS が計測する血流の増減データは絶対的变化量ではなく相対的变化量であるため相対的なパターンを定性的に解析する必要がある。fNIRS の出力データは、被験者の生理的变化や慣れの影響により、ある刺激に反応を示すが山の頂点の増減の大きさが異なる場合がある。山の頂点の増減の大きさに関わらず“反応を示す”部分として扱うため、伸長手法を適用した。

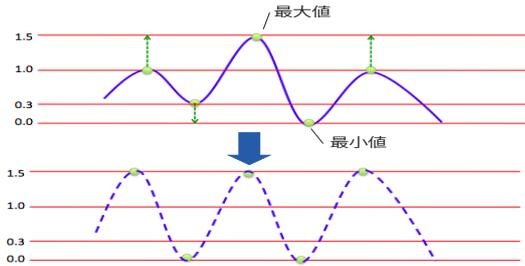


Fig. 3 前処理手法の流れ

3.2 時系列データの再量子化手法

再量子化とは、実数値の時系列データを文字列に変換することである。再量子化のためには実数値を変換する際に基準となる分割線を定める必要がある。再量子化手法としては等間隔領域分割と SAX(Symbolic Aggregation approXimation)⁵⁾ が挙げられる。等間隔領域分割は時系列データの最大値と最小値の間を等間隔に分割する方法である。SAX では時系列データが正規分布に従うことを仮定し、時系列データの平均が 0、分散が 1 となるように標準化を行う。標準化の後、各文字の出現率が同じになる分割線を定める。SAX による再量子化の流れを Fig.4 に示す。Fig.4 では文字 a が 33% 文字 b が 33%、文字 c が 33% となるような分割線が定められている。

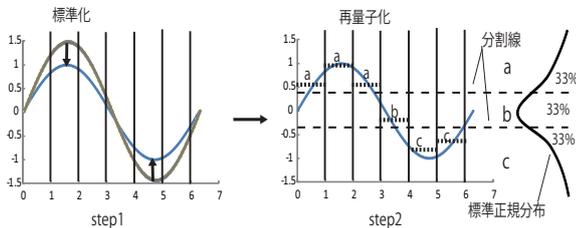


Fig. 4 SAX による再量子化の流れ

3.3 Smith Waterman 法

SW 法は 2 つの文字列から類似部分を抽出するアルゴリズムである⁶⁾。例えば、2 つの文字列“PELICAN”と“COELACANTH”からは“ELICAN”と“ELACAN”

が類似部分として抽出される。SW 法はバイオインフォマティクスの分野で並列アルゴリズムの開発が盛んに行われており、高速処理が可能である。fNIRS の多チャンネルの時系列データは大規模であり、その計算コストが問題となるが並列アルゴリズムを用いることで解決が可能となると考える。

4 可視化インターフェース

探索アルゴリズムを用いて抽出した類似部分を解析者に表示するインターフェースを Fig.5 に示す。このインターフェースはユーザとして fNIRS のデータ解析者を想定しており、ある特徴的な波形に類似した波形が他の脳部位のどこにあるかを特定し提示する。Fig.5 は複数のプローブにおいてある CH データに類似するデータを探索する機能を示している。左側頭葉のある CH を基準として他の CH との類似部分を抽出している。右側頭葉の右上に基準 CH と類似した部分が青色で表示されている。

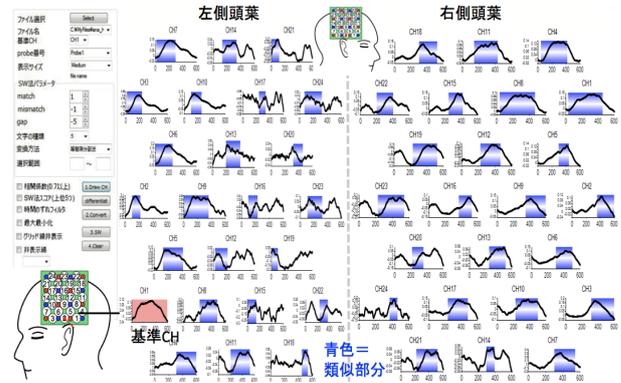


Fig. 5 可視化インターフェース

参考文献

- 1) 新井康允. 脳のしくみ. 日本実業出版社, 1997.
- 2) 福田正人. 精神疾患と NIRS-光トポグラフィー検査による脳機能イメージング. 中山書店, 2009.
- 3) Takuma Nishii; Tomoyuki Hiroyasu; Masato Yoshimi; Mitsunori Miki; Hisatake Yokouchi. Similar subsequence retrieval from two time series data using homology search. *Systems Man and Cybernetics*, pp. 1062–1067, 2010.
- 4) 西井琢真; 廣安知之; 吉見真聡; 三木光範; 横内久猛. Smith Waterman 法を利用した fNIRS データの類似部分抽出アルゴリズムの提案. 同志社大学理工学研究報告, Vol. 52, No. 4, pp. 37–45, 2012.
- 5) Abdullah Mueen; Eamonn Keogh; Qiang Zhu; Sydney Cash; Brandon Westover. Exact Discovery of time-series Motifs. *SDM*., pp. 473–484, 2009.
- 6) T F Smith; M S Waterman. Identification of common molecular subsequences. *Journal of Molecular Biology*, Vol. 147, No. 1, pp. 195–197, March 1981.