D207

情報デザインのための三次元時系列データ可視化

橘 完太[○] (工学院大学)

Visualization of 3D Time-Series Data for Information Design

Kanta TACHIBANA (Kogakuin University)

Abstract: Nowadays, many 3D time-series phenomena become measurable or reproducible in sufficient resolution of time and space owing to hardware and software development of measurement and simulation. Coordinate data and grid data are easy to handle with computers and such data are accumulated for quantitative understanding of phenomena. But, human cannot understand huge bit pattern of such data as it is. Visualization is necessary for human to figure out the essence of phenomena. This article shows visualization of 3D time-series data, which are represented by coordinate data or grid data, with self-organizing map. This article also discusses visualization of 3D time-series data from the viewpoint of information design.

Keywords: Information Design, Self-Organizing Map, Molecular Dynamics Data, Atmospheric Data

1 緒論

現代人にとって、桁違いに大量なデータの取得と保有が可能なコンピュータは非常に有用な道具である。特に、三次元の現象を定量的に表現する座標データ（興味ある点の三次元座標）と格子データ（格子状に配置した固定点における物理量）に対して、コンピュータはその威力を最大限に発揮する。幸いなことに、計測と通信技術の発達により、時間、空間ともに十分に高い解像度のデータが入手できるようになり、その蓄積が進んでいる。しかし、コンピュータには、情報を取捨選択し物理現象の本質に迫る推論を行う能力はない。必ず人間の判断が必要となる。人間が三次元現象を定量的に把握するためには、多次元データの可視化が不可欠となる。

2 本論

自己組織化マップ (Self-Organizing Map, SOM) [2] は、多次元のデータ空間から2次元の可視空間へのトポロジー保存写像を獲得する技術であり、多次元データの類似関係の可視化に有効である。トポロジー保存とは、データ空間で近い要素が低次元の可視空間でも近く表示されることである。トポロジー保存の低次元化写像は、ニューロンと呼ばれる素子の集合によって獲得される。各ニューロンは可視空間内に可視位置を持ち、データ空

間内に参照点を持つ。以下の、データの対応付けと参照点の更新を収束するまで繰り返すことで、類似したデータが可視空間で近くに配置され、多次元空間におけるデータの類似関係を把握できるようになる。データ集合の各要素を、最も近い参照点を持つニューロンに対応付ける。各ニューロンは、自分、および、可視位置が近いニューロンに対応したデータの代表点（平均など）へ参照点を更新する。

データから参照点までの「近さ」を測る距離尺度は、データの特性を考慮して定義する。以下では、生体内分子データと全球気象データのそれぞれの特性を捉えた距離尺度を用い、各応用分野へSOMを適用し三次元時系列データを可視化した例を示す。

2.1 生体内分子データの可視化

いくつかのアミノ酸の直列によって構成されるペプチドは、水溶液中で様々な立体構造をとり、その動的な挙動により生体に影響を与えると考えられている。異なる状態の水溶液中で、ペプチドがどのような時間経過でどのような立体構造をとるのかは興味深いテーマであり、分子動力学シミュレーションによる理解が試みられている。著者が携わった研究 [1] では、SOMを用いて立体構造の時系列を可視化した。ペプチド中には、ペプチド結合と呼ばれるアミノ酸の結合部分が数個存在し、それ

ら相互の立体的な向きにより、ペプチドの立体構造をほぼ表現できる。そのため、ペプチド結合の相対的な向きを四元数で表し、四元数配列間の距離を定義して SOM の距離尺度に用いた。なお [1] では、将来の拡張としてペプチドの動的挙動をさらに詳細に示す変数も組み込むことができるよう、四元数よりも広いクラスの代数であるクリフォード代数を用いてアルゴリズムを統一的に記述した。四元数 SOM により獲得した可視空間を図 1 に示す。可視空間内で各ニューロンは六角形のセルとして描かれる。各ニューロンの参照点が表す立体構造のシミュレーション中の頻度を黒丸で、立体構造の遷移を黒丸から出る線で表し、立体構造の時系列を二次元平面内の軌跡として表示した。

図 1: Visualization of peptide conformations by quaternion SOM

2.2 全球気象データの可視化

全地球を覆うような気象観測システムが整備されている。この気象観測システムに基づき、経度緯度 2.5° 間隔のグリッド点での 6 時間おきの等気圧高度、気温、湿度、風ベクトルが、再解析データとしてほぼリアルタイムに報告される。また、地球シミュレータなどを用いるシミュレーションにより、観測と同程度に高解像度の予想データも蓄積されている。これら気象データの類似関係可視化は気候変動を理解するための強力な手段となる。著者が携わった研究 [3, 4] では、 20°N 以北の等気圧高度に着目し、過去 57 年の冬季（11 月から翌年 4 月）におけるデータを可視化した。観測地点は、28 レベルの各緯度に 144 地点あり、それに北極点を加えた $4033 (= 144 \times 28 + 1)$ 地点となる。各地点での 500hPa 高度を要素に持つ 4033 次元ベクトルによって各時点の気圧状況を表現する。緯度により経線間隔が異なるため、

それを加味して重み付けた 4033 次元ベクトル間の距離を定義して SOM に用いた。可視空間を 2 次元球面（3 次元球体の表面）とし、ニューロンを逐次増やす高速化アルゴリズムを適用した。図 2 に可視空間を示す。各ニューロンは球面多角形で描かれ、出現頻度が多角形の中心に黒丸で表示される。

図 2: Visualization of atmospheric data by high speed spherical SOM

3 結論

人間が直接知覚することが出来ない座標データと格子データの可視化手法として、SOM を適用した。高解像度の三次元データが大量に入手可能となるなか、現象の本質に迫るための人間の判断を助ける多次元データ可視化の技術は今後ますます重要となるであろう。

参考文献

- [1] E. M. S. Hitzer, K. Tachibana, S. Buchholz, and I. Yu. Carrier method for the general evaluation and control of pose, molecular conformation, tracking and the like. *Advances in Applied Clifford Algebras*, 2009. (published online first).
- [2] T. Kohonen. *Self-Organizing Maps*. Springer, Berlin Heidelberg, 3rd edition, 2001.
- [3] N. Sugimoto and K. Tachibana. A first attempt to apply high speed spherical self-organizing map to huge climate datasets. *Science Online Letters on the Atmosphere (SOLA)*, 4:41–44, May 24 2008.
- [4] K. Tachibana, N. Sugimoto, H. Shioyama, and T. Nozawa. Visualization of huge climate data with high speed spherical self-organizing map. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 13(3):210–216, May 2009.