

視覚協創学(5): マッピング技術 ～特に時系列データ可視化技術の体系化に向けて～

Visual Synergetics (5): Mapping Techniques
... especially towards systemization of time-varying data visualization techniques

伊藤貴之¹⁾

Takayuki Itoh

1) 博(工) お茶の水女子大学 教授 (〒112-8610 東京都文京区大塚2-1-1 理学部情報科学科, E-mail: itot@is.ocha.ac.jp)

Mapping techniques calculates visual properties (colors, positions, shapes, ...) for scientific and information visualization. This paper surveys mapping techniques mainly in information visualization, and focus on systemization of time-varying data visualization techniques.

Key Words : Mapping techniques, Information visualization, Time-varying data visualization

1. はじめに

可視化技術においてマッピングは一つのコア技術である。マッピングとは一般に、入力情報を色・位置・形などの視覚的特性に変換することを指す。可視化技術はしばしば、科学技術系情報の可視化(Scientific Visualization)と、一般的な情報の可視化(Information Visualization: 以下「情報可視化」と称する)に大別され、両者ともにマッピングの処理工程は大きなコア技術となる。

科学技術系情報の可視化において多くの比率を占める「ボリューム可視化」では、物理空間において計測あるいは計算された数値がスカラ場であるか、あるいはベクタ場やテンサ場であるかによって、マッピング技術が大別される。また与えられた物理空間が2次元の場合には、その空間の2軸をそのままディスプレイの2軸に割り当てれば容易に可視化が実現できるが、物理空間が3次元の場合には、ディスプレイ空間へのマッピングには一定の工夫が必要である。ボリューム可視化では多くの場合において、点、線、面などの中間図形を可視化に用いる。この考え方を一般的にインダイレクト・アプローチと呼ぶ。等値面、流線、粒子アニメーション、断面上の各種表現(等高線、矢印、2次元LICなど)の多くの手法がこの考え方に基づいて考案されている。それに対して、中間図形に数値を割り当てずにボリューム全体を可視化する考えをダイレクト・アプローチと呼び、ボリュームレンダリングがその代表的手法として研究が進められている。このようにボリューム可視化においては、入力数値の形式と、ダイレクトか否かという計算手順によって、明確にマッピング技術が体系化されていると考える[1]。

一方で情報可視化の分野では、入力情報のデータ構造に基づいて個別に手法の体系化が進んでおり、その内容は多岐に渡っていて把握しづらい。本報告では情報可視化のマッピング技術について体系的に概観し、特にその中でも議論の余地が大きく残っている時系列データ可視

化技術の体系化について重点的に議論する。

2. 情報可視化の諸手法の分類

情報可視化が対象とするデータ構造は、1次元、2次元、3次元、多次元、時系列、木構造、グラフの7種類であることがShneiderman [2]によって1996年に提唱されており、その考え方は現在もほとんど変化していない。情報可視化の技術の多くは、データ構造ごとに個別に開発されてきた歴史を有する。表1は、情報可視化の分野を代表する国際会議IEEE Information Visualizationにおいて、多次元、時系列、木構造、グラフの各データ構造を対象とした可視化手法の発表が、各年においてセッションとして集約されているか否かをまとめたものである。表1において◎は単独のセッション、○はそれを含む複数の話題を組み合わせる1セッションとしたものである。この表からも、情報可視化では、入力データ構造を出発点として多くの手法が開発されてきたことがわかる。

表1 国際会議IEEE Information Visualizationにおける多次元、時系列、木構造、グラフの各セッションの存在

	多次元	時系列	木構造	グラフ
2002			◎	◎
2003	◎			◎
2004	◎			◎
2005	◎	◎	○	○
2006			◎	◎
2007			○	○
2008				◎
2009	◎	○		◎
2010	◎			◎
2011	◎	○	○	◎

また情報可視化において、入力情報のデータ構造と、

ユーザの操作手順などを指定すれば、それに合致した可視化手法の自動選出[3]がある程度は可能である。このような考えを有効にするためにも、既存の情報可視化手法の分類や体系化は重要であるといえる。

以下に、多次元、木構造、グラフを対象とした各手法に関するサーベイや体系化の動きを論じる。

(1) 多次元データ可視化手法

多次元データ可視化手法はデータマイニング分野における一覧表示や意思決定の目的において旧来から多用されており、その視点からのサーベイがいくつか発表されている[4][5]。いずれのサーベイ文献も、可視化手法を概ね以下のように分類している。

- 散布図による任意の2/3次元の可視化、およびそのマトリックス表現。
- 次元削減手法や自己組織化マップなどを適用した散布図。
- ヒートマップ、ハイトマップ、ピクセルマップなどによるヒストグラム型の表現。
- レーダーチャートやParallel Coordinatesなど、多次元に対応する多数の座標軸を生成する手法。
- アイコンやグリフの適用。
- その他の手法。(Worlds within Worlds, VisDBなど)

(2) 木構造データ可視化手法

木構造データ可視化手法の多くは、ノード=リンク型の表現、あるいは空間充填型・入れ子型の表現を採用している。前者については後述するグラフデータ可視化手法のサーベイ論文に詳しく解説されている。後者については[6]に詳しく解説されている。

(3) グラフデータ可視化手法

グラフデータ可視化手法はグラフ描画(Graph Drawing)という伝統的な研究分野を踏襲する面があり、そのサーベイや教科書は旧来から多く出版されている[7][8]。グラフデータ可視化の代表的な問題として以下があげられる。

- ノード=リンク型表現におけるノード配置の最適化。平面化モデル、力学モデル、次元削減モデル、空間充填モデルなどの各種手法が議論されている。
- ノード=リンク型表現における視認性向上のためのデータ構築手法と描画・操作手法。例えばノードのクラスタリングやサンプリング、エッジの再配置や束化、ズームやフィルタリングなどの対話操作手法があげられる。
- 一方で、ノード=リンク型表現の代用としてのマトリックス表現も提案されている。

3. 時系列データ可視化手法に関する議論

表1からわかるように、多次元、木構造、グラフの各データ構造と比較して、時系列データ可視化手法の発表はあまり多くない。また時系列データ可視化手法のサーベイ論文も、著者が検索した限りではあまり見当たらない。また著者自身の研究の過程においても、時系列データ可

視化手法への着手は、多次元、木構造、グラフの各可視化手法と比べても遅れていた。これには以下の背景があると考える。

- 時系列データは非時系列データに比べて極端に容量が大きいので、研究が流行するのに時間がかかった。
- 時系列を表現する大半の可視化手法には「画面の横軸を時刻とする」という強い拘束があり、描画手法の研究としての新規性を追うための自由度に欠ける。
- 時系列データは他のデータ構造との複合的なデータである場合が多く、時系列性の表現に着目される機会が少なかった。

一方で、昨今のIT業界のBig Dataブームにて、センサ情報やシステムログがもたらす時系列情報が主役の一端を担っている。またデータ工学の研究動向を眺めても、例えば国内最大級の会議であるDEIM(データ工学と情報マネジメントに関するフォーラム)においては、時系列という単語を含む発表が2011年11件、2012年6件と活況である。このことから今後、時系列データの可視化が情報可視化において重要視される可能性が考えられる。

時系列データ可視化のための主要なマッピング手段について、以下に論じる。

(1) 折れ線グラフ型可視化手法

日常生活において、時系列に沿って変化する数値の多くは、折れ線グラフで表示されている。情報可視化においても、時系列データ可視化手法には折れ線グラフを用いたものが最も多い。例として図1に示すように、一画面に大量に描かれる折れ線を制御するためのクラスタリングとスケッチインタフェースが提案されている[9]。

折れ線グラフ型手法は、後述するヒートマップ・ピクセルマップ型手法とともに、主として画面の水平方向に時間軸を割り当てる。ヒートマップ・ピクセルマップ型手法と比べると、以下の優位性が考えられる。

- 日常的に見慣れている。
- 自然科学系の計測値など、連続性の高い数値の表現に向いている。
- 色を用いたヒートマップ・ピクセルマップと比べて、正確に数値を読み取れる。

(2) ヒートマップ・ピクセルマップ型可視化手法

折れ線グラフが高さ方向で数値を表現するのに対して、ヒートマップやピクセルマップでは色で数値を表現する。この考え方をういた時系列データ可視化手法の例として図2に示すように、ピクセルマップから興味深い部分を切り取って折れ線表示する手法が提案されている[10]。

ヒートマップ・ピクセルマップ型手法は、折れ線型手法の適用に重大な問題点があるときに、代用的手法として重要な役割を担える。例えば以下の点において、ヒートマップ・ピクセルマップ型手法に優位性が考えられる。

- 折れ線グラフにおいて、大量の数値群を一領域に描いたときのクラッタリングの問題が発生しない。
- 尖ったピークなどの不連続性を有するデータにおい

て、折れ線グラフよりも視認性が高いことが多い。

- パターン検出手法やSAX法などを用いて非数値情報（文字列やメタ情報）として表現された時系列データの可視化に向いている。典型的な例として、楽譜情報からの旋律パターン検出結果をピクセルマップ表示した例がある[11].

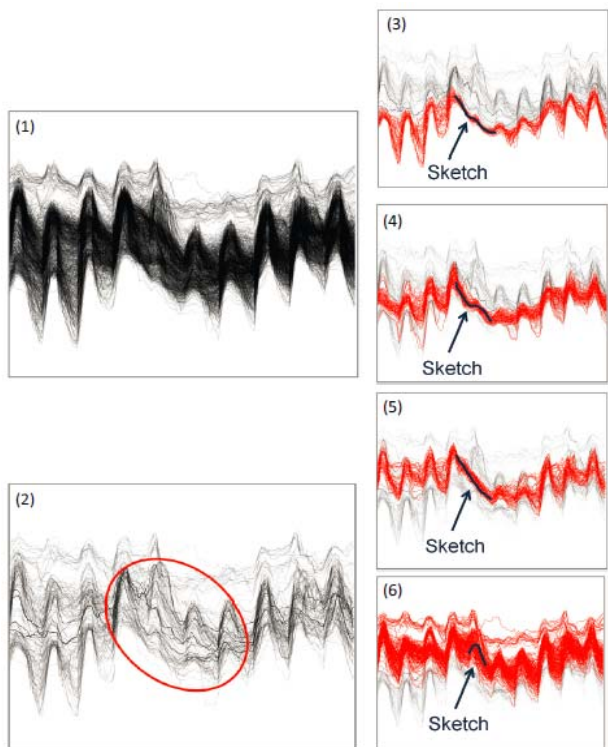


図 1. 折れ線グラフ型時系列データ可視化手法の例。(文献[9]から転載)

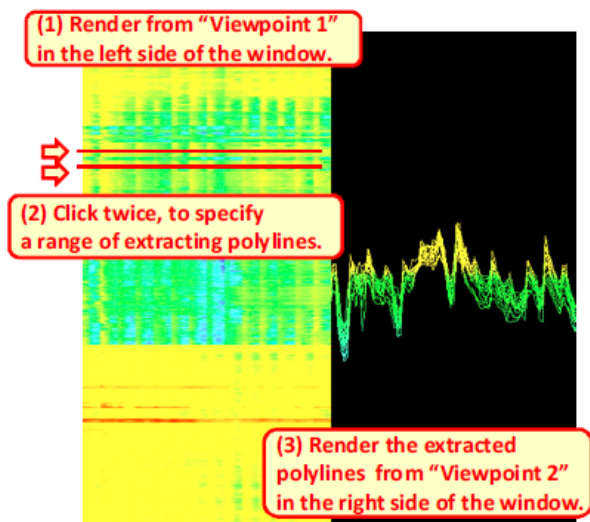


図 2. ヒートマップ型時系列データ可視化手法の例。(文献[10]から転載)

(3) 3次元型可視化手法

非時系列なデータ構造（例えば多次元・木構造・グラフ）との複合的な時系列データにおいて、非時系列なデータ構造を2次元空間に配置し、時刻を3つめの次元に割り当てて可視化する。ここで「3つめの次元」とは多くの場合において、画面の奥行き方向であるとして立体的に

情報を可視化するか、あるいは時間であるとしてアニメーション表示を適用する。例として図1に示すように、時系列情報をもつグラフデータに対して、ある時刻におけるグラフを2次元平面に描き、3つ目の座標軸に時刻を割り当てることで任意時刻におけるグラフを表示できるようなインターフェースが提案されている[12].

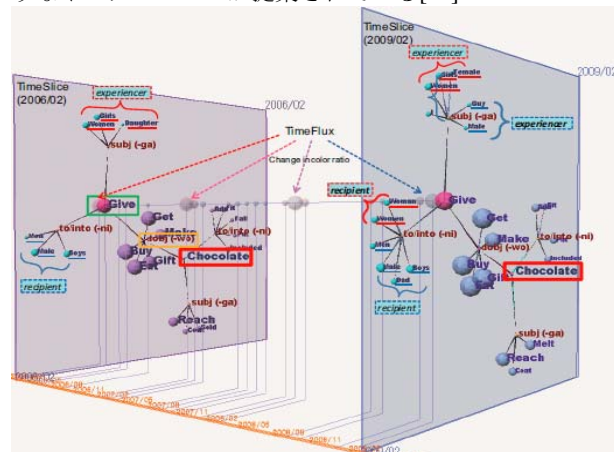


図 3. 3次元型時系列データ可視化手法の例。(文献[12]から転載)

(4) 連携型可視化手法

非時系列なデータ構造（例えば多次元・木構造・グラフ）との複合的な時系列データにおいて、Linked Viewの考え方をを用いて、1つ目の描画領域にはある時刻における非時系列なデータ構造を描き、2つ目の描画領域には時系列情報を描き、互いにそれらを連携操作させる。例として図2に示すように、画面の左側にある縦長のピクセルマップで時系列情報を描画し、その上で特定の時刻を指定すると画面の右側でその時刻における木構造情報を描画するようなインターフェースが提案されている[13].

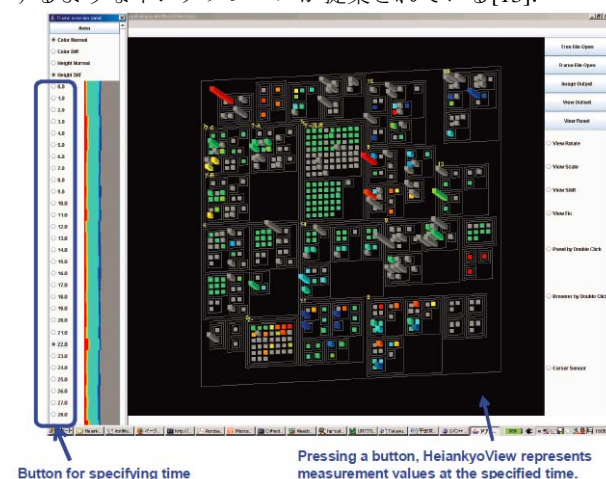


図 4. 連携型時系列データ可視化手法の例。(文献[13]から転載)

以上の4種類の時系列データ可視化手法を使い分ける手段として、例えば以下のような条件分岐による手法選択が考えられる。ただし、このような単純なフレームワークが適用できない手法や事例は多々あると考えられる。今後さらにサーベイを進め、より網羅的な体系化を進め

たい。また、この条件分岐の有効性についてユーザテストを実施したい。

[Q1] 入力時系列データは非時系列な構造を有し、その表現が重要であるか？ (→Yesであれば[Q2]へ、Noであれば[Q3]へ)

[Q2] 3次元表示やアニメーション表示を利用して差し支えないか？ (→Yesであれば手法(3)を適用、Noであれば手法(4)を適用)

[Q3] 以下の(a)~(c)を優先度の高い順に並べ替えた上で、手法(1)(2)のいずれかを選べ。

(a) 対象となる時系列情報は数値よりもメタ情報や文字列を中心とするものか？ (→Yesであれば手法(2)を適用)

(b) 対象となる時系列情報は連続数値であり、それを正確に読み取りたいか？ (→Yesであれば手法(1)を適用)

(c) 時系列情報の概観のためにクラッタリングを極力避けたいか？ (→Yesであれば手法(2)を適用)

4. 今後の展望

本報告では情報可視化の諸技術を分類・体系化するための議論として、多次元・木構造・グラフの可視化手法のサーベイを紹介し、比較的サーベイの進んでいない時系列データ可視化手法の体系化について検討した。

今後の展開として、以下を考えている。まずは3章の(1)~(4)に該当しない時系列データ可視化手法を網羅的にサーベイしたい。一例として著者らは、数値情報とメタ情報の両方を有する時系列データの可視化に着手している[14]が、これは3章の(1)(2)の中間的性質をもつ新しい試みである。このような拡張手法を包括した大きなフレームワークを検討したい。続いて、空間情報と時系列情報の複合型データの可視化について検討したい。例えば地理情報を扱った時系列データでは、地図上にグリフを配置する[15]などの手段で時系列情報を可視化する場合があるが、これは3章の(1)~(4)に該当しない。同様なことはボリュームデータの時系列情報の可視化にも該当する。今後の展開として、これらを包括した議論を進めたい。

著者の主観として、情報可視化における手法の体系化は、以下の点において難易度の高い問題であると考えられる。

- 1) 空間情報を持たない入力値を扱うことで、情報配置の問題を伴うため、表現上の自由度が高い。
- 2) 情報産業や社会科学における突発的な現象や流行を扱うことが多いため、適用事例が発散しやすい傾向がある。
- 3) データ構造は複合的なものになることが多く、そのための可視化技術も複合的になりやすい。

情報可視化の分類や体系化を議論するにあたり、このような点も意識して検討を進めていきたい。

参考文献

- [1] 中嶋, 藤代編著, コンピュータビジュアルゼーション, 共立出版, 2000.
- [2] B. Shneiderman, The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations, IEEE Symposium on Visual Languages, 336-343, 1996.
- [3] I. Fujishiro, Y. Ichikawa, R. Furuhashi, Y. Takeshima, GADGET/IV: A Taxonomic Approach to Semi Automatic Design of Information Visualization Applications Using Modular Visualization Environment, IEEE Symposium on Information Visualization, 77-83, 2000.
- [4] P. C. Wong, R. D. Bergeron, 30 Years of Multidimensional Multivariate Visualization, Scientific Visualization: Overviews Methodologies and Techniques, IEEE Computer Society Press, 3-33, 1997.
- [5] G. Grinstein, M. Trutschl, U. Cvek, High-Dimensional Visualizations, KDD Workshop on Visual Data Mining, 2001.
- [6] H.-J. Schulz, S. Hadlak, H. Schumann, The Design Space of Implicit Hierarchy Visualization: A Survey, IEEE Transactions on Visualization and Computer Graphics, 17(4), 393-411, 2011.
- [7] I. Herman, G. Melancon, M. S. Marshall, Graph Visualization and Navigation in Information Visualization: A Survey, IEEE Transactions on Visualization and Computer Graphics, 6(1), 24-43, 2000.
- [8] W. Cui, A Survey on Graph Visualization, Hong Kong University of Science and Technology Technical Report, 2007.
- [9] Y. Uchida, T. Itoh, A Visualization and Level-of-Detail Control Technique for Large Scale Time Series Data, 13th International Conference on Information Visualisation (IV09), 80-85, 2009.
- [10] M. Imoto, T. Itoh, A 3D Visualization Technique for Large Scale Time-Varying Data, 14th International Conference on Information Visualisation (IV10), 17-22, 2010.
- [11] A. Hayashi, T. Itoh, M. Matsubara, Colorscore - Visualization and Condensation of Structure of Classical Music, 15th International Conference on Information Visualization (IV2011), 420-425, 2011.
- [12] M. Itoh, N. Yoshinaga, M. Toyoda, M. Kitsuregawa, Analysis and Visualization of Temporal Changes in Bloggers' Activities, IEEE Pacific Visualization, 57-64, 2012.
- [13] T. Itoh, S. Furuya, H. Ohshima, K. Okamoto, Hierarchical Data Visualization for Atomic Plant Data, Journal of Fluid Science and Technology, 3(4), 553-562, 2008.
- [14] S. Yagi, Y. Uchida, T. Itoh, A Polyline-based Visualization Technique for Tagged Time-Varying Data, In review.
- [15] S. Yagi, T. Itoh, M. Kurokawa, Y. Izu, T. Yoneyama, T. Kohara, Data Layout and Level-of-Detail Control for Flood Data Visualization upon Maps, IEEE Pacific Visualization, Poster Session, 2012.