

軌跡の直接操作に基づく時系列データの視覚的分析支援 インタフェースの設計と評価

Design and Evaluation of Visual Analytics Interface for Time Series Data based on Direct Manipulation of Trajectory

高見 玲 高間 康史
Rei Takami Yasufumi Takama

首都大学東京大学院 システムデザイン研究科
Graduate School of System Design, Tokyo Metropolitan University

In recent years, time series data have been collected in many fields, and visual analytics interface is expected to be useful for utilizing such data. However, several issues arising when applying it to time series data should be considered. For example, when time series data is visualized using animation, collision would occur between movement of time series data itself and movement caused by interaction with users. In order to solve those issues, this paper proposes a visual analytics interface for time series data based on trajectory manipulation, which can handle temporal and spatial changes uniformly. The results of user experiment show the effectiveness of the proposed interface.

1. はじめに

近年、様々な分野において大規模かつ高次元の特徴を持つデータが蓄積されている。中でも、時系列データは、センサログデータや SNS ストリームなど、様々な分野で普及している。これらのデータから有用な知識を発見するため、視覚的分析インタフェースが研究されている [Keim 08]。本稿では、可視化されたオブジェクト自体を操作し、操作結果に基づき背後に存在するモデルやパラメータを操作し、ユーザにフィードバックする直接操作に基づくアプローチに着目する。時系列データに対して直接操作的アプローチを適用する場合の問題点として、(1) データ自身の変化とユーザ操作による変化の時間軸上での衝突や、(2) 時間軸と空間軸の間におけるインタラクションの衝突、(3) 異なる可視化手法間のトレードオフの存在が想定される。これらの問題点を解決するために、本稿では、データの時間的変化を空間上に表現し、時間および空間軸を統一して扱うことで、(1) を回避できる軌跡表現に着目する。そして、アニメーションで時系列を表現する散布図上にて軌跡を直接操作することで、データの時空間的变化の特徴把握を促進し、知識形成を支援する視覚的分析インタフェースを提案する。提案インタフェースではアニメーションの再生/静止モードに基づく適用操作の分岐を採用する。これにより、(2) に対応し、両モードの利点を生かすことで (3) にも対応する。さらに、軌跡表現に対する選択操作といった基本的なインタラクションに加えて、複数オブジェクト/軌跡のグルーピング機能も採用することで、視覚的分析プロセス全体の支援を行う。プロトタイプインタフェースを実装し、ユーザ実験を行った結果より、上述の問題点の解消や、提案インタフェースの視覚的分析における有効性を示す。

2. 関連研究

2.1 視覚的分析

一般的に、ドメイン専門家(分析者)は必ずしもデータマイニングの専門家ではない。そのため、可視化の背後に存在するモデルやパラメータの理解が困難であることが指摘されている [Sacha 17]。この問題に対処するため、ドメイン知識による

連絡先: 高間康史, 首都大学東京大学院システムデザイン研究科,
〒191-0065 東京都日野市旭が丘 6-6, ytakama@tmu.ac.jp

データの解釈と、可視化を介した洞察の獲得支援を目的とする視覚的分析が研究されてきた [Keim 08]。

視覚的分析において、可視化オブジェクトの直接操作とフィードバックにより、データに対する理解を促進し、知識形成を支援する手法が研究されている [Endert 13]。代表的なアプローチとして、直接操作に基づき、可視化の背後に存在する次元削減などのモデルやパラメータを推定・修正することで知識形成を促進する semantic interaction のコンセプトが提案されている [Endert 12]。このコンセプトでは、システムが可視化オブジェクトへの操作内容を解釈し、結果に基づき背後に存在するデータマイニングアルゴリズムのモデルやパラメータを調整する。調整結果は可視化表現に反映され、分析者に即時フィードバックされる。これによって、分析者が可視化の背後に存在するモデルやパラメータに関する知識を有しない場合でも、それらを調整可能となる。また、ドメイン知識を調整結果に適用することで、探索と知識形成の間の直感的な移行が可能となる。そのため、コンピュータの計算処理能力と、可視化されたデータに対する人間の認知能力や知識の双方を活用した探索が行える [Sacha 17]。

2.2 時系列データの可視化手法

時系列データを視覚的分析する場合、データの時間的特徴や、分析対象のデータ特性、データの表現手法などの様々な考慮すべき特性が存在し、多くの可視化手法が研究されている [Aigner 07]。代表的な手法として、動きで時間的変化を表現するアニメーションのような動的可視化や、複数のグラフの併置で表現する small multiples、データの変化を空間上に重畳するトレース(軌跡)のような静的可視化が存在する [Robertson 08]。

本稿で着目する軌跡表現は、空間的座標の変化を同一ビュー上に表示する可視化手法である。視覚的分析における軌跡表現の利用例として、Kondo と Collins は軌跡の操作に基づく時間的ナビゲーション手法 DimpVis を提案している [Kondo 14]。DimpVis ではデータの変化軌跡をなぞる操作に基づき、時系列データの散布図表現内の時間的遷移を行える。Bach らは、時系列データの視覚的概要の提示手段として軌跡に着目し、データの各時点での自己類似性に基づく軌跡の描画手法を提案している [Bach 16]。Rind らは、医療における患者の経過観測や治療の選択のためのインタフェースにおいて、散布図上のオブジェク

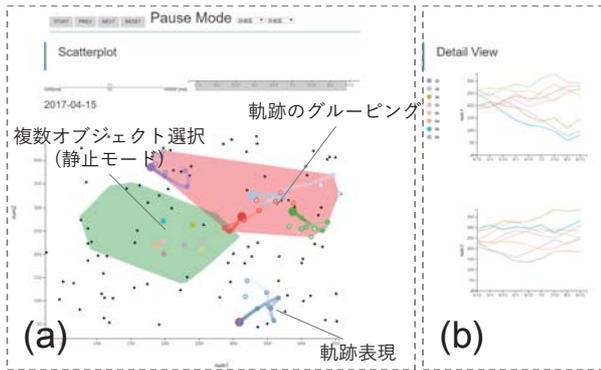


図1: 提案インタフェース: (a) 散布図ビュー (b) 詳細ビュー

トの時間的変化を軌跡で表現している [Rind 11]. これらの既存手法では、時間的ナビゲーションや時系列データの特性を表現するために軌跡が活用されている。しかし、視覚的分析の包括的なプロセスへの統合に関しては議論されていない。

3. 時系列データの視覚的分析における問題点

時系列データに対して直接操作に基づく視覚的分析を適用する場合に、以下のような問題点が想定される。

3.1 時間軸上におけるインタラクションの衝突

時系列データの可視化手法の一つであるアニメーションは、時間的変化を実時間上の変化として表現する。そのため、直接操作に基づく視覚的分析にアニメーションを用いた場合、再生操作に伴いオブジェクト配置の変化が生じる。また、オブジェクトに対してユーザーがドラッグ操作などの直接操作を適用した場合にも、オブジェクト配置の変化が起こる。このとき、2つの配置の変化はどちらも時間的なものとなる。本稿では、この状況を時間軸上の衝突とみなし、時系列データの視覚的分析における主要な問題点の一つとして捉える。

3.2 時間軸と空間軸の間でのインタラクションの衝突

時系列データの分析に際して必要な操作として、アニメーションの制御といった時間軸への操作が挙げられる。加えて、視覚的分析における、可視化が配置されたインタフェースの画面(空間座標)を空間軸と定義すると、直接操作によるオブジェクトの配置変更などの操作は空間軸に対するインタラクションとみなせる。結果として、同一操作の適用先として2つの軸が同時に存在しうる。そのため、ユーザーの直接操作の解釈において、対象となる軸に関して衝突が生じると考える。例えば、時系列データに対して詳細を確認するために拡大(ズームアップ)操作を行った場合、空間的に拡大表示したいのか、再生速度を遅くしたいのか、二通りの解釈がありうる。このように、採用された解釈がユーザーの意図したものとは異なる場合には分析の妨げとなるため、時系列データの視覚的分析における考慮すべき問題点の一つと考える。

3.3 可視化手法間のトレードオフ

アニメーションなどの動的な可視化手法は、時系列データが有する意味のある変化の発見や、聴衆に理解させるプレゼンテーションのような場面では有効である。一方で、変化の見落としの発生や、スライドなどによるアニメーションの制御に伴う操作回数の増加などの要因により、詳細なデータ分析には適していない [Robertson 08]. 一方で、トレースや small multiple などの静的な可視化手法は効率と正確性の面で詳細な探索に適

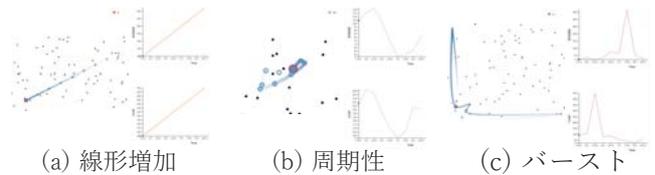


図2: 属性値の時系列的変化特性と軌跡の形状の関係

しているが、スケーラビリティや視覚的混乱について問題がある [Robertson 08]. このように、時系列データに対して一意に有効な可視化手法は存在しない。そのため、分析目的や対象に応じて複数の手法を組み合わせる必要がある。

4. 提案インタフェース

3節で述べた問題点に基づき、本稿では、軌跡表現と2次元散布図を組み合わせた時系列データの視覚的分析支援インタフェースを提案する。インタフェースは、Ruby on Rails 5を用いてWebアプリケーションとして実装した。可視化部とフロントエンド処理の実装にはD3.jsおよびjQueryを用いた。

図1に提案インタフェースのスクリーンショットを示す。提案インタフェースは散布図ビュー(図1(a))と詳細ビュー(図1(b))から構成される。散布図ビューでは、選択されたオブジェクトに対応するデータから2属性を選択し、X、Y軸に対応付ける。そして、各時点での属性値を散布図上に可視化し、アニメーションによる位置の変更でデータ分布の時間的変化を表現する。ユーザーは関心を持った散布図上のオブジェクトを選択し、軌跡の表示や直接操作を行える。詳細ビューでは、散布図上で選択されたオブジェクトの各属性値の時間的変化を示す線グラフや、メタデータなどの詳細情報を表示する。

ユーザーは散布図ビューで得た視覚的傾向に対して、詳細ビューを用いて詳細な探索や仮説検証を行う。両ビュー間の移行の繰り返しにより、視覚的分析が行われることを想定する。一方のビューへの操作は他方へも適用され、インタフェースの一貫性を保持する。提案インタフェースの特徴的な機能として、軌跡による可視化や、探索モード、オブジェクトのグルーピングがある。以降に、それぞれの機能の説明を示す。

4.1 軌跡による可視化

散布図上でオブジェクトにマウスオーバーすると、属性値の時間的変化が軌跡として可視化される。軌跡は各時点でのオブジェクトの空間座標を表すノードと、各ノードを結ぶ曲線であるパスから構成される。2.2節で述べたように、軌跡表現はデータの視覚的概要として用いられる。例として、属性値が線形増加した場合には、図2(a)のような直線状の軌跡が描かれる。周期的に変化した場合には、図2(b)のように、丸まった形状の軌跡が描かれる。属性値が特定の期間内でバーストする場合には、図2(c)のように、広域に渡り急峻な変化をする軌跡が描かれる。軌跡は属性値の変化を空間上に重畳して表現するため、対象の時系列的特性を空間上のみで判断できる。そのため、ユーザー操作に応じて軌跡を表示・操作することで、視覚的混乱を抑制しつつ、3.1節で述べた時間軸上でのインタラクションの衝突を回避できると考える。また、軌跡表現では、オブジェクトの時間的側面(軌跡)と空間的な側面(他のオブジェクトとの関係)の両方が同時に散布図上で空間的に表現される。そのため、3.2節で述べた時間軸と空間軸間のインタラクションの衝突を回避できると考える。

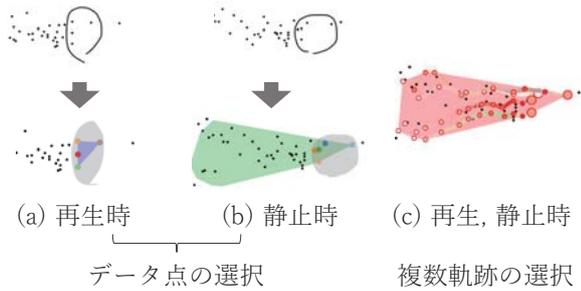


図 3: 提案インタフェースにおけるオブジェクトのグルーピング

4.2 探索モード

3.3 節で述べた可視化手法間のトレードオフへ対処するため、探索モードとしてアニメーションの再生と静止をインタフェースに導入する。探索モードに基づき、ユーザ入力の解釈を分離することで、動的な可視化と静的な可視化手法の両者の利点を活かし、トレードオフの解消を図る。再生モードではオブジェクト全体の傾向を把握するための利用を想定し、オブジェクトをインタラクティブなアニメーションで可視化する。分析者は、データ全体の傾向に基づき、興味を持ったオブジェクトの変化を追跡できる。静止モードでは、静的な散布図上で、軌跡を直接操作したり、再生モードで得た視覚的傾向を詳細ビューで検証したりしながら詳細な分析を行う。分析者は再生モードで得た視覚的概要に基づき、静止モードにて詳細な探索や仮説の検証を行い、両モード間の遷移により視覚的分析を行うことを想定する。

4.3 オブジェクトのグルーピング

提案インタフェースでは、オブジェクトをグルーピングするためのスケッチベース入力を採用する。散布図ビュー上でオブジェクト群を囲うような線を描画すると、図 3(a)(b)のようにそれに含まれる全てのオブジェクトを含む凸包が散布図上に描画される。凸包は選択されたオブジェクト群の空間的な広がりを表現可能なため、類似性の判断に対して有効と考える。

4.2 節で述べた探索モードに基づき、異なる形状の凸包が描画される。再生モードでは、特定の時点における全選択オブジェクトを含む凸包が描画される(図 3(a))。これによって、複数オブジェクトの大局的な時間的変化の概要の把握を行える。静止モードでは全選択オブジェクトの軌跡を含む凸包が描画される(図 3(b))。また、複数の軌跡をマウスクリックにより選択した場合には、探索モードによらず図 3(b)と同様の凸包が描画される(図 3(c))。スケッチベース入力による選択では、範囲に基づく対象の暗黙的な指定により、被選択オブジェクト群の時間的な変化傾向の把握が促進されると考える。一方で、クリックによる選択では、対象の明示的な指定による複数オブジェクト時間的変化の比較のための利用を想定する。

5. 評価実験

提案インタフェースの有効性を検証するため、20 歳から 24 歳までの 32 名の工学系大学生および大学院生(男性 30 名、女性 2 名)を対象としたユーザ実験を行った。分析対象のデータには、日本プロ野球の投手データを使用した。また、実験で行うタスクへの習熟による影響を考慮するため、セリーグとパリーグの 2 つの異なるデータセットを採用した。

5.1 実験内容

実験開始時に、協力者はチュートリアルに従ってインタフェースの各機能を試行した。次に、特定のオブジェクトを探索し、詳細ビューへ表示する可視化生成タスクを行い、その後、特定のオブジェクト群に対する視覚的傾向を発見し、説明を記述する説明タスクの 2 つのタスクセットを順に行った。実験協力者は、提案インタフェースと、2 つの比較インタフェース(アニメーション、もしくは軌跡表現を提案インタフェースから除去)のうちどちらか一つを用いて、両タスクを実行した。実験終了後、主観的フィードバック取得のため、提案インタフェースの各機能を 5 段階リッカート尺度で評価するアンケートに回答してもらった。一部の設問については、その理由を文章で記述してもらった。タスク毎のデータセット、インタフェースの組み合わせ、提案・比較インタフェースにおけるタスクの実行順序は、順序効果などを考慮して決定している。また、定量的な考察のためにインタフェースの各機能の使用回数を記録し、定性的な考察のために実験中の画面を録画した。

5.2 実験結果

5.2.1 可視化生成タスク

可視化生成タスクにおける実行時間について、 t 検定による有意性の検証を行ったところ、提案・比較インタフェースの間に有意な差は見られなかった。一方で、主観的評価では提案インタフェースの習熟コストに対して問題点が指摘されていた。また、多くのタスクについて、実行時間の標準偏差は提案インタフェースで最も大きかった。この結果も、提案インタフェースへの習熟コストが影響したものと考える。

5.2.2 説明タスク

説明タスクでは、提案インタフェースにおける各機能の操作回数が可視化生成タスクと比較して少ない傾向にあることが確認された。特に、オブジェクトの選択操作や詳細ビューでの属性値の確認などの視覚的分析の基本的なタスクにおいて、 t 検定の結果、有意差が観測された。記述された説明内容について、提案・比較インタフェース間のキーワードの含有数についてカイ二乗検定を用いた有意性の検証を行った。その結果、全体的なデータ傾向を問う問題については局所的な傾向や根拠を表すキーワードについて有意差が認められた。また、特定オブジェクト群の傾向を問うタスクでは、実際の値を表すキーワードについて有意差が確認された。すなわち、提案インタフェースにおいて、探索対象の傾向を実際の値を参照し具体的かつ詳細に記述できたことが確認できる。そのため、提案インタフェースを用いることで、複数オブジェクトの傾向を把握するタスクにて、効率的な探索が可能であるといえる。

5.2.3 軌跡表現と探索モードの有効性

表 1 に、提案インタフェースの各機能に対するアンケート結果を示す。表内の数値は、インタフェースの各機能(行)に対する、5 段階リッカート尺度の各段階(列)ごとの協力者数を表している。表 1 より、軌跡表現に対して高い評価が得られていることが確認できる。理由として、軌跡表現は直感的に理解しやすく、視覚的混乱が少ないことが挙げられていた。また、提案インタフェース全体について、多くの実験協力者は、異なる種類の可視化手法の組み合わせによりトレードオフが解消され、探索目的ごとにモードを変えながら効率的な探索を行えたと回答した。これは、5.2.2 節における操作回数の軽減からも確認できる。このように、軌跡表現による時間的変化の可視化により、散布図上の興味のあるオブジェクトに対して視覚的混乱を防止しつつ、時間的変化の把握が可能であるため、時間軸上におけるインタラクションの衝突の問題に対し有効と考える。

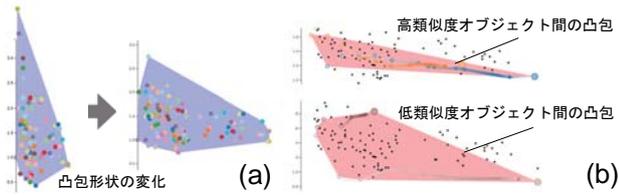


図 4: グルーピング操作における探索モード毎の実験協力者の分析行動例: (a) 再生モード (b) 静止モード

5.2.4 可視化手法毎の利点を生かした探索

ユーザの主観的評価と操作ログの集計結果より、グルーピング操作がタスクごとに異なる用途で用いられる傾向が確認された。図 4 に、実験協力者の分析行動の例を示す。例えば、データ全体の傾向を問うタスクでは、協力者は再生モードでオブジェクト全体をグルーピングし、アニメーションによる凸包形状の変化に基づき、全体的な傾向を判断していた (図 4(a))。一方で、特定オブジェクト群の傾向を問うタスクでは、静止モードで該当するオブジェクトの軌跡を複数選択することで、データ間の類似性や関係を判断していた (図 4(b))。

比較対象に対する提案インタフェースの有効性について、アニメーションは表示が簡潔である一方、データの時間的変化を把握しづらいと回答した協力者が複数存在した。一方で、トレースはデータ全体の時間的変化の把握が容易である一方、データの重なりによって、選択操作や値の把握が困難な場合があるという意見が確認された。両者に対する問題点は相補的なものとなっている。そのため、軌跡表現や、探索モードによる操作の解釈の分離により、可視化手法間のトレードオフの問題が解消されると考える。すなわち、提案インタフェースにより、動的・静的可視化手法の利点を活かした分析作業が行えるといえる。

6. おわりに

本稿では、多次元時系列データの分析支援を目的として、軌跡の直接操作に基づく視覚的分析インタフェースを提案した。プロトタイプインタフェースを実装し、評価実験を行った。結果より、軌跡表現の直感性や、異なる可視化手法の利点の組み合わせによる視覚的分析全体の支援に関して、有効性が示された。

一方で、アンケートでは重なりなどについて視覚的混乱が報告された。また、探索モードに関するインタラクションについてわかりにくいと回答した協力者もみられた。これらの問題点は、今後の提案インタフェースのユーザビリティの向上により対応できると考える。また、探索モードに基づく操作解釈のマッピングについて、実験結果を元に再検討する必要があると考える。インタフェースの機能の改善に加えて、ドメイン専門家による視覚的分析を支援するため、semantic interaction 的アプローチの導入も有効と考える。また、知識形成プロセス全体における有効性検証のため、ケーススタディによる長期間かつ定性的な評価を行うことも有用と考える。

参考文献

[Aigner 07] Aigner, W., Miksch, S., Müller, W., Schumann, H., and Tominski, C.: Visualizing time-oriented data—a systematic view, *Computers & Graphics*, Vol. 31, No. 3, pp. 401–409 (2007)

表 1: 提案インタフェースの各機能に対する主観的評価

評価対象	5 段階評価				
	1	2	3	4	5
視覚的分析における全体的な有効性	0	3	0	23	6
比較インタフェースと比べた有効性	0	1	5	10	16
散布図ビューと詳細ビューの連携	0	0	1	7	24
軌跡による可視化	1	2	2	10	17
軌跡の直接操作	0	5	5	16	6
オブジェクトのグルーピング	1	1	2	10	18
再生モード	2	3	7	13	7
静止モード	0	2	4	7	19
全体的なユーザビリティ	0	3	6	21	2

[Bach 16] Bach, B., Shi, C., Heulot, N., Madhyastha, T., Grabowski, T., and Dragicevic, P.: Time curves: Folding time to visualize patterns of temporal evolution in data, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 22, No. 1, pp. 559–568 (2016)

[Endert 12] Endert, A., Fiaux, P., and North, C.: Semantic interaction for visual text analytics, in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 473–482 (2012)

[Endert 13] Endert, A., Bradel, L., and North, C.: Beyond control panels: Direct manipulation for visual analytics, *IEEE Computer Graphics and Applications*, Vol. 33, No. 4, pp. 6–13 (2013)

[Keim 08] Keim, D. A., Mansmann, F., Schneidewind, J., Thomas, J., and Ziegler, H.: Visual Analytics: Scope and Challenges, in Simoff, S. J., Böhlen, M. H., and Mazeika, A. eds., *Visual Data Mining: Theory, Techniques and Tools for Visual Analytics*, pp. 76–90, Springer (2008)

[Kondo 14] Kondo, B. and Collins, C.: DimpVis: Exploring time-varying information visualizations by direct manipulation, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 20, No. 12, pp. 2003–2012 (2014)

[Rind 11] Rind, A., Aigner, W., Miksch, S., Wiltner, S., Pohl, M., Drexler, F., Neubauer, B., and Suchy, N.: Visually exploring multivariate trends in patient cohorts using animated scatter plots, in *International Conference on Ergonomics and Health Aspects of Work with Computers*, Vol. 6779 LNCS, pp. 139–148 (2011)

[Robertson 08] Robertson, G., Fernandez, R., Fisher, D., Lee, B., and Stasko, J.: Effectiveness of animation in trend visualization, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 14, No. 6, pp. 1325–1332 (2008)

[Sacha 17] Sacha, D., Sedlmair, M., Zhang, L., Lee, J. A., Peltonen, J., Weiskopf, D., North, S. C., and Keim, D. A.: What you see is what you can change: Human-centered machine learning by interactive visualization, *Neurocomputing*, Vol. 268, pp. 164–175 (2017)