

閲覧者の関心度に対応できる視覚的情報提示手法に関する研究

渡邊 玲 三末 和男*

概要. 電子書籍や Web において視覚的情報を提示する際、閲覧者自身がデータのどのような面に関心があるのかに応じて、可視化手法を変更できたり、可視化の詳細度を変更できたりすれば便利である。しかし、従来の視覚的な情報の提示手法では、情報提供側が提示する可視化手法を指定しているため、閲覧者の関心に応じた視覚的な情報の提示になっていない。本研究の目的は、閲覧者の関心に応じて可視化するデータの詳細度を変更できる視覚的な情報の提示手法を構築することである。そのために我々は、データの種類、関心の種類、詳細度の高低によって可視化手法を分類した。また、そのような提示手法を実現するプログラムを開発した。

1 はじめに

書籍や雑誌などで見かける視覚的な情報表現であるグラフやチャートなどは、情報発信者が自分の発信意図に沿って設計したものである。Web ではインタラクティブな情報表現も一般的になってきている(たとえば New York Times¹ などに見られるインフォグラフィックスなど)が、そこで採用されるインタラクションも情報発信者の意図の範囲内に留まっている。

その一方で閲覧者の情報に対する要望は、発信者の想定範囲内に留まらないことも多い。例えば何かに対する 100 人分の評価が平均値という一つの値で提示されていた時に、評価値の分布を知りたいとか、100 人分の個々の評価を知りたいと思うことは少なくない。逆に、100 人分の評価値が個々に提示されていた時に、要するにみんなはどう評価しているのかといった要約を知りたいこともある。

このような要望に対応するためには、閲覧者自身が視覚的表現を選んだり、表現のパラメータを変更したりできる枠組みが必要である。しかしながら、閲覧者は一般に情報可視化の専門家ではないため、適切な視覚的表現を選んだり、適切なパラメータを設定したりすることは難しい。そこで、我々は閲覧者の「関心」に着目し、関心に応じた視覚的な情報の提示手法を構築することを目指している。

我々は、まず閲覧者がどのような関心を持つかを調べるために、よく用いられる可視化手法を、それらが表すデータの種別及びどのような情報をどれだけ表示するのかという度合いによって分類した。さらに、その分類の妥当性を検証するために、分類に沿った視覚的表現を提供するプログラムを作成した。

2 関心と詳細度による体系化

閲覧者がどのような関心を持つかを調べるために、既存の可視化手法を体系化する。

2.1 データの種類と対象とする可視化手法

可視化手法を分類するにあたり、まず可視化の研究分野でも多く用いられている Shneiderman の分類 [2] を参考に、データの種別を 1 次元、2 次元、3 次元、多次元、時間、ネットワーク、木の 7 つに絞った。

また、扱う可視化手法は、可視化の研究分野で広く使われていたり、一般的に使用されている可視化ソフトウェアで扱われていたりしているもののみとした。これは、高度な可視化手法は手法に不慣れな閲覧者にとっては情報が読み取りにくく、一般的な使用には不向きと考えたためである。具体的には、1 次元のデータに対しては StripChart、箱ひげ図、棒グラフを、2 次元のデータに対しては、円グラフ、散布図、ヒストグラム、棒グラフを、3 次元のデータに対しては、バブルチャートを、多次元のデータに対しては、レーダーチャートを、時間のデータに対しては、折れ線グラフを、ネットワークのデータに対しては、ノードリンクダイアグラムを、木のデータに対しては、ノードリンクダイアグラムを使用する。

2.2 関心の種類と詳細度

2.1 節で対象とする可視化手法が表すことができる情報を整理し、そこから「現在提示されている可視化手法が表しているデータから、一部について詳しく知りたい」「要約統計量を知りたい」「集約の粒度を変更した情報を知りたい」という 3 つの関心の種類があると考えた。そこから、詳細度の種類を 3 つ設定した。

詳細度の種類 (a) 現在提示されている可視化手法が表しているデータから、一部についてを表す可視化手法

Copyright is held by the author(s).

* Watanabe Rei, 筑波大学大学院 システム情報工学研究科 コンピュータサイエンス専攻, Kazuo Misue, 筑波大学 システム情報系

¹ <http://www.nytimes.com/>

詳細度の種類 (b) 現在提示されている可視化手法が表しているデータから、要約統計量を表す可視化手法

詳細度の種類 (c) 現在提示されている可視化手法が表しているデータから、集約の粒度を変更した情報を表す可視化手法

ここでは、生データにより近い情報を得たいという場合に関心度を高いとし、生データにより近い情報を表す場合を詳細度が高いとしている。

3 作成したプログラム

2節の分類がうまく機能するかを確認するために、その分類を実現するプロトタイプを作成した(図1)。

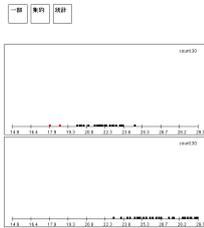


図 1. 作成したプログラムの概観

ユーザは、左上のボタン(一部、集約、統計)を押すことで関心の種類を入力し、表示する可視化手法を変更できる。そして、同じボタンを複数回押すことによって関心の度合いをシステムに伝え、視覚的な情報の詳細度を変更できるようにしている。ここでは、1次元(量)について作成したプログラムを紹介する。

詳細の種類(a)に関しては、すべてのデータを表す StripChart(図2(a)) 及びデータをフィルタリングした StripChart を実装することで、詳細度を変更できるようにしている。詳細の種類(b)に関しては、四分位を表す箱ひげ図(図2(b))と平均値を表す棒グラフを実装することで、詳細度を変更できるようにしている。詳細の種類(c)に関しては、ヒストグラムのビンのサイズを変更できるようにする(図2(c), 図2(d)) ことで、詳細度を変更できるようにしている。

4 ユースケース

2節の分類及び3節のプログラムが妥当であるかを検証するために、低レベルタスクを実行できることを示す。低レベルタスクとは、データ理解のための可視化ツールを使用した際によく持ちだされる質問を、タスクとして分類したものである[1]。具体的には、(1)データの属性を見つける、(2)ある条件を満たすデータを見つける、(3)集約的表現を見つける、(4)極値データを見つける、(5)ソートする、(6)

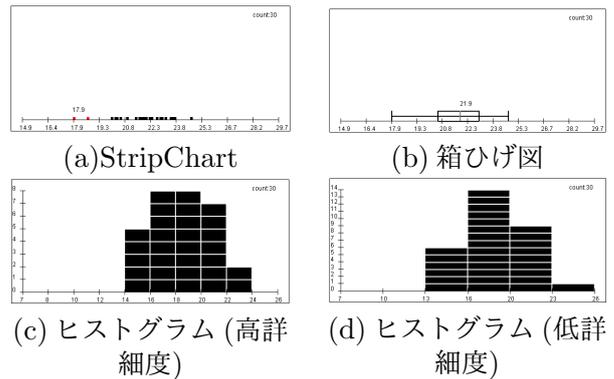


図 2. 1次元(量)の可視化手法

データセットのサイズを特定する、(7)分布を把握する、(8)例外を探すの8つのタスクのことである。

タスク(1)については、データを表す図形の上にマウスカーソルを置くことで、データの数値的属性を見ることができる。タスク(5)、(6)、(8)については、最初に提示される視覚的情報において既にデータがソートされており、外れ値を赤く表示しているため、右上にデータセットのサイズを書いているため、何も操作をしなくてもタスクを実行することができる。外れ値は平均±標準偏差×1の範囲から外れたものとした。タスク(2)については、データを表す図形の上をクリックすることによってデータをフィルタリングすることができる。また、フィルタリングの結果残ったデータの数も確認することができる。タスク(3)、(4)については、統計ボタンを一度押すことによって箱ひげ図が表示され、集約的データ(中央値)や極値データ(最大値や最小値)を見つけることができる。タスク(7)については、集約のボタンを一度だけ押すことによって、ヒストグラムが表示され、分布を把握することができる。

5 まとめ

閲覧者の関心に応じて可視化するデータの詳細度を変更できるようにするために、視覚的な情報の提示手法を構築した。また、そのような提示手法を実現するプログラムを作成した。今後は、作成したプログラムでの関心の入力方法について再考し、さらにツールキットを整備する予定である。

参考文献

- [1] R. Amar, J. Eagan, and J. Stasko. Low-level components of analytic activity in information visualization. In *Proceedings of IEEE Symposium on Information Visualization*, pp. 111–117, 2005.
- [2] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of Visual Languages*, pp. 336–343, 1996.