Vol.2016-HCI-170 No.1
Vol.2016-UBI-52 No.1
2016/10/24

閲覧者の関心に対応できる視覚的情報の提示手法に関する研究

渡邊玲[1,a)]　三末 和男[2]

概要：多くの視覚的表現が情報提供側の意図に基づいて作成されているため，閲覧者は知りたい情報を必ずしも入手することができない．その問題を解決するには，閲覧者が元のデータから新たな視覚的表現を作成できるようにすればよいが，閲覧者は必ずしも可視化の専門家ではないため視覚的表現を設計したり選択したりすることが難しい．そこで本研究では，閲覧者にとって表現が容易であると考えられる「関心」を入力とした情報提供の枠組みを構築することを考えた．そのために，思考実験を通して閲覧者が関心を持つと考えられるデータの側面を洗い出し，それらの形式的な表現を提案する．さらに，構築した枠組みの普及を容易にするためのプログラム及びツールキットの開発構想について述べる．

1. はじめに

Web ページや書籍，雑誌などにおいて，データを分かりやすく表現するために視覚的な情報表現をすることが多い．例えば，オンラインの New York Times[*1]は様々な視覚的表現及びインタラクティブな操作を用いることで，記事内で用いるデータを分かりやすくしている Web ページの好例である．

既存の視覚的な情報提示手法は，情報を提供する側が視覚的な情報の提示方法を設計しているため，閲覧者は必ずしも欲しい情報を入手することができない．また，情報を提供する側が無知や悪意により適当でない視覚的表現を提示してしまう場合もある．例えば，例えば量的データとカテゴリデータからなる二次元データに対して，棒グラフではなく折れ線グラフを使ってしまい，関係の無い項目同士に量的変化があるように見せてしまうような場合である（図 1）．

さらに，閲覧者が提供された情報を元に自由に視覚的表現を作成できたとしても，閲覧者は必ずしも可視化の専門家ではないため，自分の欲しい情報を読み取るために適切な視覚的表現を設計できるとは限らない．

そこで本研究では，閲覧者の関心を入力とした視覚的情報を提供できる枠組みを構築する．具体的には以下のことを行う．

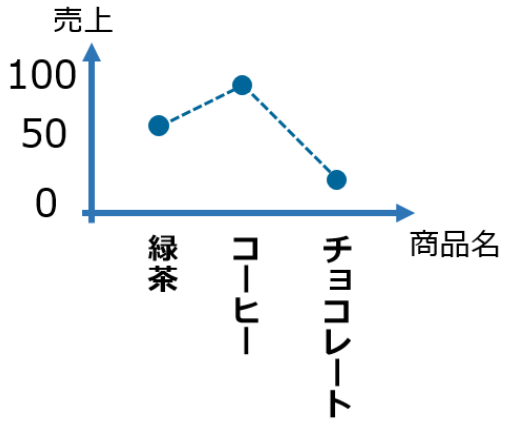


図 1　誤った可視化の例

- 閲覧者の関心の対象をまとめる
- 関心の対象を表現できる可視化手法をまとめる
- 枠組みを実現するプログラムを作成する
- 関心の入力に適した UI を考案する
- 枠組みの普及のためにツールキットを作成する

関心を入力としたのは，閲覧者は自分の関心であれば把握できると考え，関心と視覚的表現を結びつけることで，閲覧者が可視化分野への知識を持っていなくても自分の欲しい情報を入手できるようになると考えたためである．また，そのような設計に基づいたプログラム及び汎用的なツールキットを作成することで，実際に閲覧者の関心に応じた可視化手法の提示手法を実現する．利用シーンとしては，設計者が HTML などに本研究の提示手法を用いた可視化手

[1] システム情報工学研究科 コンピュータサイエンス専攻
[2] 筑波大学 システム情報系
[a)] watanabe@vislab.cs.tsukuba.ac.jp
[*1] http://www.nytimes.com/

ⓒ 2016 Information Processing Society of Japan

法を埋め込み，閲覧者に簡単な操作によって視覚的表現を変更できるようにさせる，ということを想定している．

本論文では，まず既存の研究を元にまとめた閲覧者の関心の対象及び関心の対象を表現することができる可視化手法について述べる．また，関心の対象及びパラメータが入力されると，それらに対応した視覚的情報を提示するプログラムのプロトタイプについても述べる．

2. 関連研究

2.1 可視化手法の体系化

今までに可視化手法は数多く開発されており，それらに対して様々な体系化がなされてきた．Shneiderman[1] はデータの種類を 7 つに分類し，その分類に基づいてタスクを 7 つに分類することで，可視化手法を理解することを助けている．Chi ら [2] は，可視化手法の参照モデルというモデルを用いることで可視化手法を分類し，そのことによって設計者の設計の理解を助けている．これらは開発者の立場から可視化手法を体系化しているが，本研究では閲覧者の立場での可視化手法を分類する．

Elmqvist[3] は，対話的な集約的可視化のためのモデルを提供するために，集約を伝えることができる情報や集約の度合いを変えるためのインタラクティブな操作についてまとめている．しかし，具体的にどの可視化手法がどの程度集約した情報を表しているのかなどについて明記していない．

2.2 グラフ作成のためのソフトウェア

所持しているデータを簡単にグラフにすることができるソフトウェアが数多く開発されている．Excel[*2]は，所持しているデータを簡単に可視化するための機能を持っている．また，Wolfram Research 社は，インタラクティブ操作が可能な可視化手法を PDF に埋め込んだ，計算可能ドキュメント[*3]を開発した．これらは手軽に視覚的表現を作成できる一方，可視化する目的に応じて閲覧者が可視化手法や表示するデータを変更することが難しい．

2.3 グラフ作成のためのツールキット

また，インタラクティブな可視化ツールの作成を助けるために，様々な可視化ツールキットが作成されている．Flare[*4]及び Prefuse[4] は，可視化の作成及び分野固有のカスタマイズを容易にするための，Flash 及び Java のライブラリである．D^3[5] は，HTML や SVG，CSS を使用した，データ駆動の JavaScript ライブラリである．これは，インタラクティブな操作，また可視化手法のアニメーションの作成を容易にしている．しかしこれらは可視化手法単体の作成を助けることを目的としており，可視化手法の変更に対応した設計になっていない．

[*2] https://products.office.com/ja-jp/excel
[*3] http://www.wolfram.com/cdf/
[*4] http://flare.prefuse.org/

3. 閲覧者の関心

閲覧者の関心の対象及びパラメータをまとめ，関心の対象を表現することができる可視化手法について述べる．また，それぞれの関係を 3.4 節及び表 1 にまとめた．

3.1 関心の対象

一般的な視覚的表現の閲覧者がどのようなタスクを実行するのかをまとめた低レベルタスク [6] を元に，7 つに定義した．

説明のために，例としてレコードにタイムスタンプ，購入者，商品 ID，商品名，値段，支払い方法，JAN コード，英語商品名，商品カテゴリ等が記載されている購買データセットを挙げる．

レコードの値
例:購買データセットにおける値段の値

集約値
例: 購買データセットの値段の値の中央値

ある条件を満たすレコード
例:購買データセット内で値段が 3000 円以上のレコード

ランキング
例:購買データセットの値段のランキング

分布
例:購買データセットの値段の分布

クラスタ
例:購買データセットの値段のクラスタ

相関
例:購買データセットの発売日と値段の相関

3.2 パラメータ

関心の対象を詳細に定義するためのパラメータをまとめた．パラメータを更に増やすことも可能であるが，閲覧者が関心を入力とする際にはパラメータの数が少ない方が閲覧者の負担が少なくなると考え，最小限にまとめた．

属性 A　データの属性
集約方法 X　中央値，平均値，四分位など
条件 Y　ある値を持つ，極値を持つ，例外を持つなど

3.3 可視化手法

可視化手法は数多く開発されているが，本研究はその中でも関心の対象を表現することができて，かつ一般的な閲覧者が読み取れるものを対象とする．

- Strip Chart
- 集合棒グラフ
- 積み上げ棒グラフ
- 円グラフ

ⓒ 2016 Information Processing Society of Japan

- 散布図
- ヒストグラム
- 折れ線グラフ

3.4 関心の対象とパラメータと可視化手法

関心の対象，パラメータ，可視化手法の関係を表 1 にまとめた．

閲覧者の関心の対象に対して，それを詳細に定義するパラメータと，関心の対象とパラメータから定義される関心に対して適切に情報を表現できる可視化手法を記載している．関心に対して情報を適切に表現できる可視化手法が複数ある場合は，データの特徴に沿った可視化手法を提示することが望ましいと考えている．

これらの対応は，5 節に述べる実験で適切であるかどうかを確認する．

4. 枠組みを実現するプログラム

3 節のような閲覧者の関心を入力として視覚的情報を提供できる枠組みを実現するためのプログラムを作成する．

4.1 プロトタイプ

3 節で定義した関心の対象及びパラメータを入力すると，それに応じた視覚的表現を出力するプログラムを作成した (図 2)．使用言語は JavaScript である．

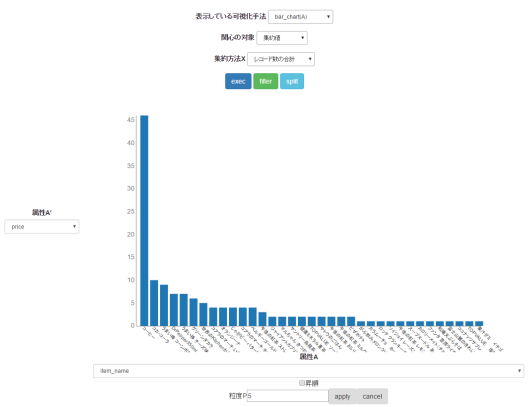


図 2 プロトタイプの概観である．関心の対象やパラメータを入力するためのセレクトボックスが配置されている．

出力された視覚的表現の調整のために，3 節で定義したパラメータの他に，粒度や並び順を指定できるようにしている．閲覧者はセレクトボックスによって関心の対象，集約方法，パラメータを選択することができ，実行ボタンによって入力を終了する．また，グラフ上でレコードを表す矩形を選択し分割ボタン，フィルタボタンを押すことでデータセットを分割したりフィルタリングしたりすることができる．

4.2 利用例

プロトタイプの利用例を述べる．利用しているデータは研究室内で収集した購買データである．

閲覧者が購買データに対し最初に関心を持つのは商品の分布であると考えたとする．その場合，関心の対象に「分布」を，パラメータに「商品名」を入力し，実行ボタンを押すと，図 3 のような棒グラフが出力される． 左端の 2

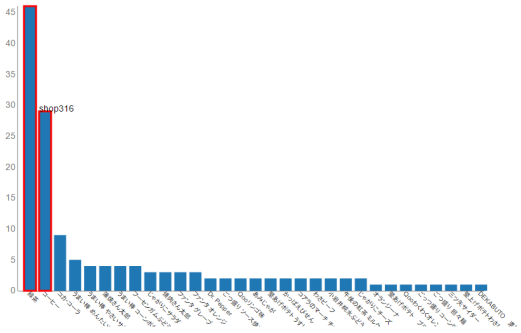


図 3 商品の分布

つの商品が他の商品に比べ著しく購入回数が高いことが分かる．

閲覧者が左端 2 つの商品それぞれについて，購入者の分布に関心を持ったとする．その場合，まず 2 つの商品を示す棒をクリックし，分割ボタンを押し，データセットを分割する．その後，関心の対象に「分布」を，パラメータに「購入者」を入力し，実行ボタンを押すと，図 4 のような円グラフが出力される．

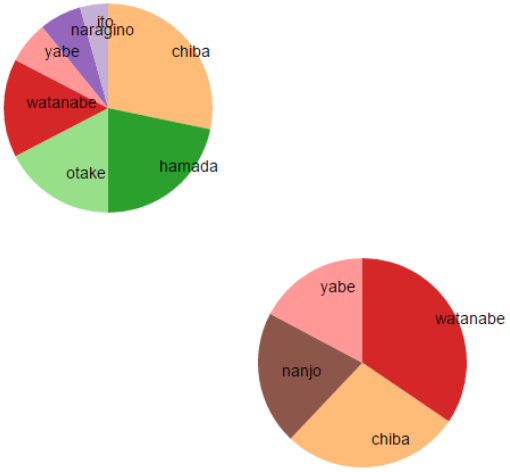


図 4 購入者の分布

閲覧者が 2 つの商品の日時と売上の関係に関心を持ったとする．その場合，関心の対象に「相関」を，パラメータに「タイムスタンプ」「値段」を入力し，実行ボタンを押すと，図 5 のような折れ線グラフが出力される．

以上のように，関心の対象とパラメータを入力することで，適切な可視化手法を得ることができる．

ⓒ 2016 Information Processing Society of Japan

表 1　関心の対象とパラメータと可視化手法．括弧内は可視化手法に適したデータの種類．

関心の対象	パラメータ	可視化手法
レコードの値	属性 A_1，属性 A_2	Strip Chart(カテゴリ×量的)，散布図 (量的×量的)，折れ線グラフ (時間を含む)
集約値	属性 A，集約方法 X	棒グラフ (1 つの集約値)，四分位を表記した棒グラフ
ある条件を満たすレコード	属性 A_1，条件 Y	Strip Chart(カテゴリ×量的)，散布図 (量的×量的)，折れ線グラフ (時間を含む)
ランキング	属性 A_1，属性 A_2	Strip Chart(カテゴリ×量的)，散布図 (量的×量的)，折れ線グラフ (時間を含む)
分布	属性 A	ヒストグラム (量的)，円グラフ (カテゴリ，項目数 8 以下)，棒グラフ (カテゴリ，項目数 9 以上)
クラスタ	属性 A	ヒストグラム (量的)，円グラフ (カテゴリ，項目数 8 以下)，棒グラフ (カテゴリ，項目数 9 以上)
相関	属性 A_1，属性 A_2	散布図 (量的×量的)，折れ線グラフ (時間を含む)

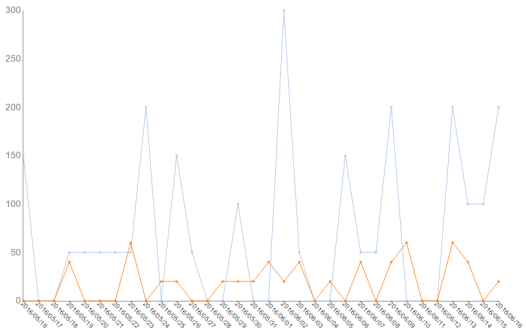


図 5　商品と売上の関係

4.3　ツールキット

本研究で構築した枠組みを普及させるために，プログラムをツールキットとして整備する予定である．ツールキット作成にあたり，既存の可視化ツールキットはどの可視化手法を用いるのか，またデータのどの属性を使用するかなど，記述しなければならない情報が多く気軽に使うには適していないのではないかと考えた．そこで，本研究作成するツールキットは以下のような簡単な記述で本研究の提案する枠組みを利用できるようにする予定である．

HTML ファイル

```
<div id='container'></div>
```

JavaScript ファイル

```
visdoi.draw(data); //JSON ファイルを指定
```

5.　被験者実験

3.4 項の可視化手法の提示方法が適切であるかどうかの確認及び 4 節のプログラムの UI 改善のために，閲覧者があるデータセットに対しどのようなデータをどのような可視化手法で視覚的に表現するのか，また関心の遷移にはどのようなものがあるのかを調べた．

5.1　関心の遷移

購買データなど，時間データ，カテゴリデータ，量的データを含むようなデータを閲覧者が分析する際，「時間データと量的データの相関を見た後，カテゴリデータと量的データの相関を見る」，「カテゴリデータと量的データの相関を見た後，カテゴリデータを細分化して量的データとの関係を見る」など，閲覧者の関心には遷移のパターンがあるのではないかと我々は考えた．また，そのような関心の遷移のパターンを発見することによって，枠組みを実現するプログラムを作成する際，UI を簡略化することに役立つのではないかと考えた．

5.2　実験に使用するツール

画面左側には可視化手法を指定するセレクトボックスと，データセット内の属性を指定するセレクトボックスが配置されている．また，属性の値をフィルタするためのセレクトボックス及びボタンが配置されている．ここで選択できる可視化手法は，3.3 項に述べた通りである．集合棒グラフ，積み上げ棒グラフに限り，中央値を表示できるようになっている．このツールでは，関心を持った属性及び可視化手法を指定すると，それらを表現するグラフが出力されるようになっている．

ツールの動作はすべて記録されるようになっている．記録する情報は以下の通りである．

- 動作の時間
- 現在表示されている可視化手法
- 現在表示されている属性

実験に使用するデータは，購買データセットである．

5.3　実験設計

実験では，被験者にツールを用いて自由にデータの分析を行ってもらう．しかしデータに対して関心を持っている状態のデータを取る必要があるため，以下の必要最低限の

ⓒ 2016 Information Processing Society of Japan

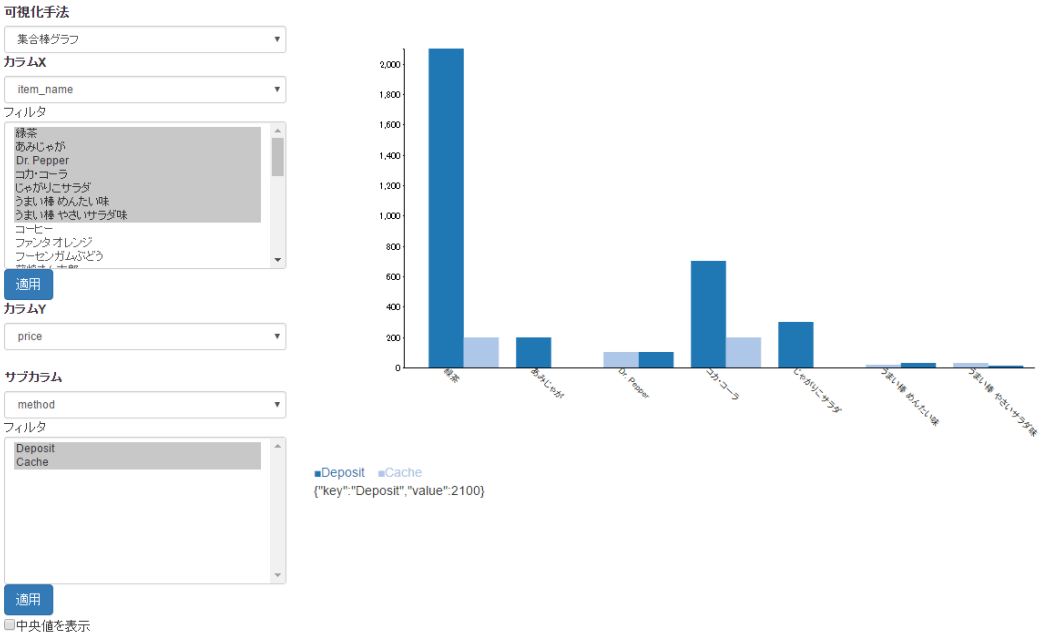


図 6　実験ツールの概観

タスクを課す．

- 日毎の特徴を調べる
- 購入者の特徴を調べる
- 値段の特徴を調べる
- 商品の特徴を調べる
- カテゴリの特徴を調べる

なおデータの各属性については実験前に説明する．

また，動作によってどのような情報を得ようとしているのかなど，考えていることを可能な限り発話してもらう．被験者がツールを使って関心を実現する視覚的表現を作ることができない場合は，どのような関心だったか，またどのような機能があればそれを実現できたかを発話してもらう．

被験者は，情報系を専攻している 21-23 歳の大学生及び大学院生 5 名である．

5.4　実験結果

動作記録から，誤動作や途中だと思われる動作を削除するために，2 秒以内に次の操作に進んだ動作は削除した．さらに，発話記録から分析に用いられていないと思われる動作も削除した．作成された視覚的表現におけるデータ型と可視化手法の関係を表 2 にまとめた．

5.5　考察

5.5.1　用いられる可視化手法の傾向

3.4 項では時間を用いるデータには折れ線グラフが適切であるとしたが，実験の結果，時間的データであっても積み上げ棒グラフも使用されていることが分かった．発話内容から，被験者は分布の時間的変化を知りたいことが多く，折れ線グラフでは分布を把握できないからだと考えられる．

また，被験者によって積み上げ棒グラフを使うか集合棒グラフを使うかに偏りがあった．そのため，4 節のプログラムでは，被験者がどちらを先に使用したかによって提示する可視化手法を変えるのが望ましいと考えられる．

この実験では散布図や Strip Chart，ヒストグラムはあまり使用されなかった．これは，今回使用したデータセットには量的データの属性が少なかったためであると思われる．そのため，量的データの属性の多いデータセットを使った実験を追加で行う必要があると考えられる．

5.5.2　関心の推移の傾向

関心の推移にはいくつか傾向が見られた．

5 人中 3 人は，同一のデータに対し割合と絶対値の両方を知りたいと考え，円グラフ及び積み上げ棒グラフ，あるいは円グラフ及び集合棒グラフを使用した (図 7)． このことから，4 節のプログラムでは，円グラフと積み上げ棒グラフ，集合棒グラフの切り替えを容易にすると良いと考えられる．

また 5 人中 4 人は，突出したデータだけに着目したいと考え，突出したデータのみを抜き取り新たな視覚的表現を作成した (図 8)． またそのうち 1 人は突出したデータのせいで他のデータが見にくくなるため，突出したデータを削除して新たな視覚的表現を作成した（図 9)．

このことから，4 節のプログラムでは，突出したデータだけを抜き出し新たな視覚的表現を作成することを容易にする操作が必要であると考えられる．

6.　まとめと今後の課題

本研究では，閲覧者が欲しい情報を入手できるようにす

表 2　データ型に対して可視化手法が使用された回数

可視化手法	集合棒グラフ	円グラフ	積上げ棒グラフ	ヒストグラム	折れ線グラフ	Strip Chart	散布図
時間-カテゴリ-量的	3	3	6	0	5	0	1
時間-カテゴリ-カテゴリ	1	0	0	0	0	0	0
時間-時間-量的	0	2	0	0	0	0	0
カテゴリ-カテゴリ-量的	5	13	5	0	0	0	0
カテゴリ-量的-量的	3	4	2	0	0	0	1
カテゴリ-量的	0	1	2	2	0	0	0
量的-量的	1	1	0	1	0	1	0

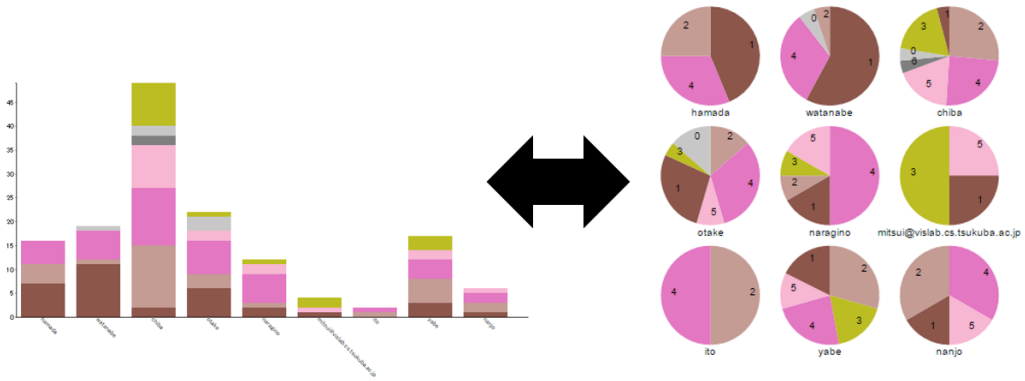


図 7　同一データに対し，円グラフと積み上げ棒グラフで可視化した図

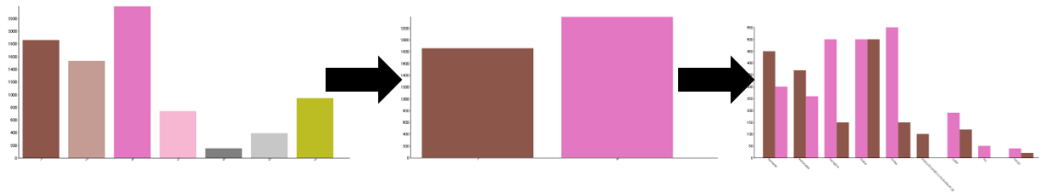


図 8　突出したデータだけを抜き取り可視化した図

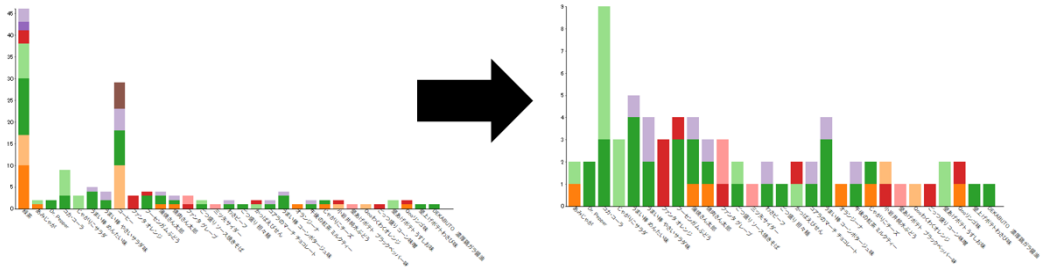


図 9　突出したデータを削除し可視化した図

るために，閲覧者の関心に沿った視覚的情報を提示するための枠組みを構築した．そのために，既存の研究を元に関心の対象及びパラメータを定義し，関心の対象を表すことができる可視化手法をまとめた．また，関心 (関心の対象とパラメータ) を入力すると，関心に対応する視覚的表現が提示されるようなプログラムのプロトタイプを作成した．

そして，提示する可視化手法の適切さの確認及びプログラムの UI 改善のために実験を行った．実験の結果，提示する可視化手法はデータ型だけでなく閲覧者の行動も踏まえて決定すると良いことが分かった．また，いくつか関心の遷移の傾向が見られた．使用したデータセットの不備によりいくつか十分に検査できなかった可視化手法があったため，追加で実験を行い，その後 3 節の関心と可視化手法とパラメータの対応及び 4 節のプログラムを修正する予定である．プログラム修正後，4.3 項のツールキットを実装する予定である．

ⓒ 2016 Information Processing Society of Japan

参考文献

[1] Shneiderman, B.: The eyes have it: A task by data type taxonomy for information visualizations, *Proceedings of Visual Languages*, pp. 336–343 (1996).
[2] Chi, E. H.: A taxonomy of visualization techniques using the data state reference model, *Proceedings of IEEE Symposium on Information Visualization*, pp. 69–75 (2000).
[3] Elmqvist, N. and Fekete, J.-D.: Hierarchical aggregation for information visualization: Overview, techniques, and design guidelines, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 16, No. 3, pp. 439–454 (2010).
[4] Heer, J., Card, S. K. and Landay, J. A.: Prefuse: a toolkit for interactive information visualization, *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, pp. 421–430 (2005).
[5] Bostock, M., Ogievetsky, V. and Heer, J.: D^3 data-driven documents, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 17, No. 12, pp. 2301–2309 (2011).
[6] Amar, R., Eagan, J. and Stasko, J.: Low-level components of analytic activity in information visualization, *Proceedings of IEEE Symposium on Information Visualization*, pp. 111–117 (2005).

ⓒ 2016 Information Processing Society of Japan