

多変量データを SPJ 質問により統合する 平行座標プロット型情報可視化システム

濱崎 裕太[†] 植村 智明^{††} 富井 尚志^{†††}

[†] 横浜国立大学大学院 環境情報学府 情報メディア環境学専攻 〒240-8501 横浜市保土ヶ谷区常盤台 79-7

^{††} 横浜国立大学大学院 環境情報学府 情報環境専攻 〒240-8501 横浜市保土ヶ谷区常盤台 79-7

^{†††} 横浜国立大学大学院 環境情報研究院 〒240-8501 横浜市保土ヶ谷区常盤台 79-7

E-mail: [†]hamasaki-yuta-cz@ynu.jp, ^{††}uemura-toshiaki-xh@ynu.jp, ^{†††}tommy@ynu.ac.jp

あらまし 現在、センサ・ストレージ技術の発達やオープンデータの公開により、多くの属性を含むデータをすべて蓄積可能になった。それに伴い、複数の属性からなる多変量データを分析し、新たな知見を獲得可能となった。しかしながら、多変量データの分析は一般に困難であり、可視化システムによる支援が不可欠である。これに対し本研究では、可視化手法の 1 つである平行座標プロット (PCP) を関係代数における SPJ(Selection-Projection-Join) 質問に相当する表現力を持つように拡張した情報可視化システムを提案する。加えて、PCP の操作結果を保存・再現するための仕組みも提案する。また、文化財の 1 つである小袖屏風と、自動車走行ログという異なる 2 つのコンテンツについて情報可視化を行うことにより本システムの有用性を示す。

キーワード 情報可視化, 平行座標プロット, 多変量データ分析, データベース

1 はじめに

センサ・ストレージ技術の発達やオープンデータの公開により、多くの属性を含むデータをすべて蓄積可能になった。それに伴い、複数の属性からなる多変量データを分析することで、新たな知見を獲得可能となった。総務省の「情報通信白書」によれば、世界のデータトラフィックは 2020 年には 1 か月あたり 228 エクサバイトに達すると予想されており、それらデータ利活用に向けた取り組みの重要性を述べている [1]。

しかしながら、多変量データの分析は一般に困難であり、可視化システムなどによる技術的支援が不可欠である。この技術の中でも、多変量データ可視化手法の 1 つとして、平行座標プロット (PCP: Parallel Coordinate Plots) [2] が知られている。これは、多次元データの属性を軸として描き、それに対してデータの 1 タブルを 1 つの線で表す手法である。これにより、データ全体のクラスタの偏りや、属性間の相関を俯瞰的に把握することができる。

この PCP に関しては、これまで様々な論文で議論がされてきた [3]。しかし、PCP の軸の見せ方やクラスタの見せ方に関して議論がされているものがほとんどであり、操作の過程に着目した議論はされていない。また、PCP に対し関係代数における SPJ(Selection-Projection-Join) のような演算を加えながら操作するものも議論されていない。ところが、実際のデータ分析においては試行錯誤によって所望のデータ可視化を目指していくものであり、その試行錯誤の過程を保存できることは情報可視化において不可欠なものである。また、リレーションスキーマを適切に設計した上で多変量データを蓄積したからこそ、リレーショナルデータベース上でうまくマネジメント可能になり、分析者が求めるアドホックなクエリが記述できる。そのため、多変量データを利用

する上では関係代数に対して親和性が高く、分析過程のスナップショットを残せる可視化システムを構築することが望ましい。

そこで本研究では、可視化手法の 1 つである PCP に対し、PCP の操作過程を保存・再現する操作言語を組み込んだ情報可視化システムを提案する。本言語は関係代数における SPJ(Selection-Projection-Join) 質問に相当する記述が可能である。これにより、データ分析者が PCP 上での操作と言語を相互にやり取りすることを可能にし、データ分析の試行錯誤を支援する。

本研究の目標を次のように定める。

- (1) SPJ 質問機能を統合した PCP を構築することによって、可視化システム上で体系的な操作を可能にする。
- (2) 操作結果の保存・再現をする言語を統合した PCP を構築することによって、データ分析の試行錯誤を支援する。
- (3) 服飾文化財と自動車走行ログという、趣向が全く異なる 2 つのコンテンツについて情報提示を行うことにより、本システムの汎用性と有用性を示す。

2 関連研究

2.1 平行座標プロット (PCP)

1985 年、Inselberg によって初めて PCP の概念が定義された [2]。それ以来、PCP に関する様々な議論がされてきており、情報可視化の分野において重要なトピックの 1 つとなっている [3]。Johansson らによれば、PCP の研究は次の 4 つに分類される [3]: (1) PCP の (属性) 軸レイアウト, (2) PCP の Clutter 軽減方法, (3) PCP の実応用例の提示, (4) PCP と他のデータ解析手法との比較。上記の通り、ほとんどは PCP の見せ方に関して議論がされているものがほとんどであり、操作の過程に着目した議論はされてい

い。また、操作の過程でデータ自体に SPJ のような演算を加えながら操作するものは議論されていない。

また、Bouali らは、対話型遺伝的アルゴリズムを使用することで、可視化手法の推薦を行うシステムを構築した [4]。これは、データや利用者の要求に応じてより適切な可視化手法 (散布図行列や PCP など) の選択を支援するものである。我々の提案システムは関係代数における SPJ 質問に相当する表現力を持つような可視化システムを構築するため、タプルが 1 つの線で明示され、詳細に参照・分析可能である PCP が適切である。

一方で、インタラクティブに操作しながら PCP による分析を支援するシステムの提案もされている。Itoh らは、属性軸間の相関に基づいてインタラクティブに次元削減を行い、PCP から所望する情報の発見を支援するシステムを構築した [5]。Zhou らは、エントロピーの概念を導入することで、PCP の属性軸の整列順序をクラスタに基づいて決定する手法の提案をした [6]。これらの提案システムは、属性間の相関に基づいて操作したまでにとどまっており、データに対する操作方法を十分に提供しているとは言い難い。また、インタラクティブに操作が可能ではあるものの、操作の過程について主題に置いた議論はされていない。それに対し我々が提案するシステムは、SPJ に相当する表現力を持つ自由度の高い操作と、その操作過程のスナップショットを言語の形で保存することが可能である。

2.2 デジタルアーカイブに関連する可視化システム

近年、情報可視化に注目した取り組みが分野を問わず様々に展開されている。情報技術を実際の展示で活用する取り組みに着目すると、次のようなものがある。岡本らは、絵画作品の一部を切り出し、その素材要素群をデジタルミュージアムに展示することで興味喚起を行うインタフェースの開発を行った [7]。Wakita らは、立体的に織られた大型の織物を計測し、実物大展示システムや視触覚展示システムを構築した [8]。安達らは、大量の写真資料の中から所望する画像の候補を効果的に提示するシステムの検討を行った [9]。Yoshino らは、日本の伝統的な書籍に対し直感的に操作可能な体験型鑑賞支援システムを構築した [10]。これらのシステムは、利用者に対しコンテンツへの興味喚起を行うことを目的としている。一方で、我々はデータ分析を支援することを目的としており、その点で我々の研究とは異なる。

また、利用者へ様々な視点を提供し分析を行えるような枠組みの提案もされている。Thudt らは、デジタルブックコレクションに対して Serendipity に着目した情報可視化システムを構築した [11]。An らは、自己組織化マップを応用することで、複数博物館の類似点や相違点を視覚的に分析し、観光客のための包括的な博物館ツアーの提案に関する仕組みを提案した [12]。Whitelaw は、従来からあるキーワード検索による情報検索に代わり、検索対象全体の概観と個々の詳細とを往来することで検索を行うインタフェースを構築した [13]。これらのシステムは、コンテンツについてあまり詳しくない利用者を想定したものである。一方で、我々の提案システムはコンテンツに対してある程度の知見があるエキスパートを対象としており、その点で我々の研究とは異なる。

さらに、コンテンツについて詳しくない人だけではなく、専門

家に対しても有用な情報可視化の検討も行われている。Odijk らは、著者の異なる複数のデータを統合し、文化財に対する様々な視点を提供するシステムを構築した [14]。Hinrichs らは、文学研究を支援する上で重要な 4 つの観点を発見し、それに基づいて未知の大規模文学コレクションの可視化システムを構築した [15]。Glinka らは、作品集の全体を把握しながら個々の作品に着目可能にすることで、文化財の定性分析と定量分析を行うことができる情報提示システムを検討した [16]。これらの提案手法では、それぞれのコンテンツに特化した枠組みの提案にとどまっている。それに対し我々は、コンテンツを問わない汎用的な枠組みの提案をするものであり、その点で我々の研究とは異なる。

2.3 自動車ログに関連する可視化システム

近年、自動車より収集した走行ログを活用する取り組みが展開されている。飯島らは、プローブカーで収集した道路状況を元に、渋滞状況を地図の色分けで可視化するシステムを構築した [17]。平岡らは、運転者に自発的なエコドライブを促すためのエコドライブ支援システムを提案し、情報提示の重要性を示した [18]。Wilmink らは、情報提示による運転の改善が二酸化炭素の排出を削減することをシミュレートした [19]。Komamizu らは、EV の運転軌跡をドライバの属性情報と共に可視化するフレームワークを提案した [20]。これらの提案手法は、コンテンツに特化した枠組みとなっている。一方我々は、コンテンツを問わない汎用的な枠組みの提案をするものであり、その点で我々の研究とは異なる。

2.4 提案システム概要

本節で設計する提案システムの概要を図 1 に示す。

提案システムは、データ分析者が 2.5 節で後述する PCP に対しインタラクションを行いつつ、その操作結果を 2.6 節で後述する (PC)²L という言語で保存・再現することで試行錯誤を行いながら所望の可視化を獲得することを想定する。以下では、想定する操作の流れについて示す。

- (1) 1 つのリレーションを PCP に適用し可視化する。
- (2) 表示結果をもとに、PCP 上で操作を行い試行錯誤を行う。
- (3) 分析者が所望するときに、(2) で操作した結果のスナップショットを (PC)²L で保存する。
- (4) (2)-(3) を繰り返す。その際、過去のスナップショットに戻る必要がある場合は該当する (PC)²L を入力し、PCP 上に操作結果を再出力する。
- (5) データ分析者が所望の可視化結果を獲得する。

2.5 PCP に対するインタラクション

本節では、提案システムの PCP における関係代数表現と、PCP 上で利用可能な SPJ 質問について述べる。ここで用いる記号の定義は表 1 に示す。なお、表 1 の記号及び「選択 (Selection)」、 「射影 (Projection)」、 「結合 (Join)」に関する定義については、文献 [21] を参考にした。

2.5.1 PCP と関係代数

本小節では、提案システムの PCP における関係代数表現について述べる。

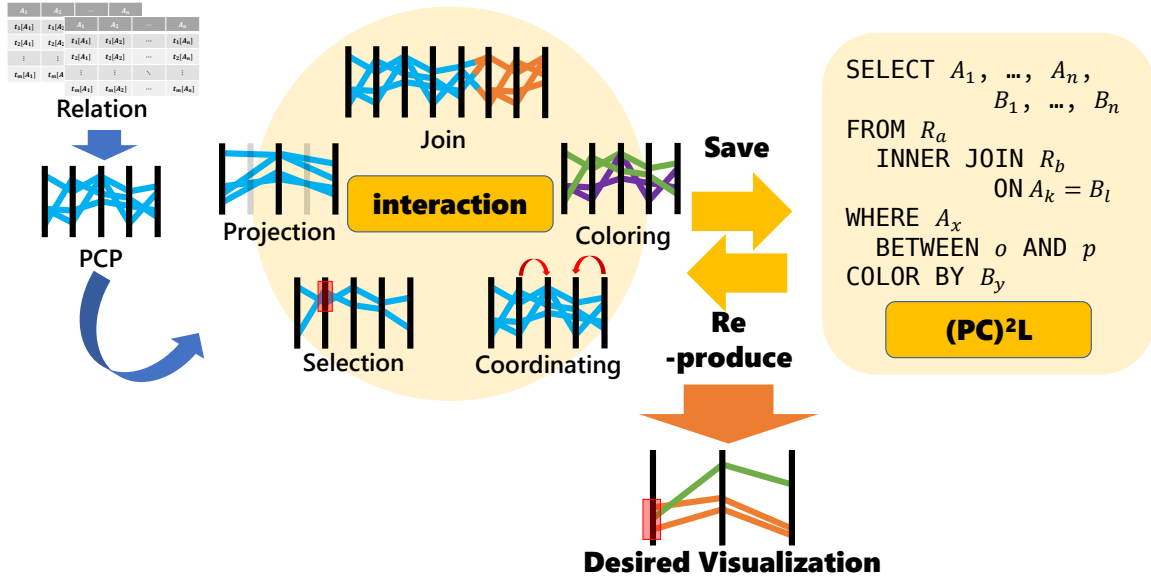


図1 提案システムの概要図

表1 記号の定義

記号	概要
A_i	属性
$R(A_1, A_2, \dots, A_n)$	リレーション
$t_j = (a_1, a_2, \dots, a_n)$	タプル
$a_i = t_j[A_i]$	タプル t_j の属性 A_i 上の属性値
$dom : A_i \rightarrow D_i$	属性 A_i のドメイン D_i を導出する関数
$D_i = dom(A_i) = \{\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{im}\}$	属性 A_i のドメイン (定義域)
$R[A_i \theta A_j]$	R の A_i と A_j 上の θ -選択演算
$R[A_{i1}, A_{i2}, \dots, A_{ik}]$ ($1 \leq i_1 < i_2 < \dots < i_k \leq n$)	R の $A_{i1}, A_{i2}, \dots, A_{ik}$ 上の射影演算
$R[A_i \theta B_j]S$	R と S の A_i と B_j 上の θ -結合演算

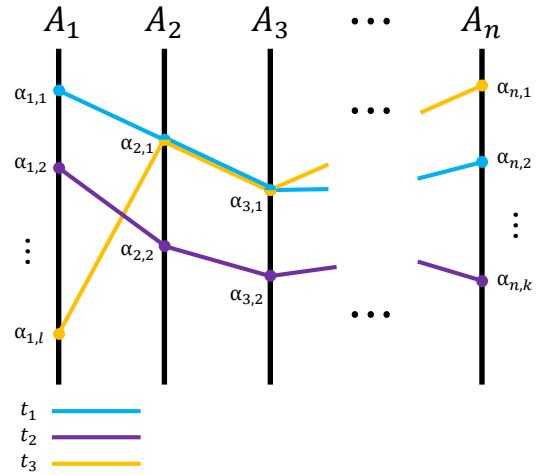


図2 PCP上にリレーション R を表した例

本稿では、リレーションの持つ属性を軸として複数並列し、タプル1つが持つ「属性値」を線で結ぶことでデータ1タプルを表現する。この線を複数描き重ねることで線の疎密や偏りが現れるので、ある属性同士がどのような関係にあるのか、全体を見通してどのような傾向にあるのかを一目で把握することができる。

以下では、PCPとリレーションの関係について定義する。リレーション R の属性 A_i の名称を図2の軸名とする。また、 $dom(A_i)$ の要素、つまり属性 A_i の属性値を図2の A_i 軸の中に配置する。その後、タプル t_j を図2の線1本に相当するような形で描くことで、リレーション R のインスタンスをすべてPCP上で表す。

また、タプルを表す線の色によってデータ全体の見え方が変化することがある。そこで、任意の属性 A を基準に線の色を塗り分けられるようにする。この機能をColoringとして定義する。これにより、線の疎密や偏りだけでなく、色によってもデータ全体の傾向を把握することが可能となる。

2.5.2 PCPとSPJ質問

本小節では、PCPに対するインタラクションとして利用可能なSPJ質問について述べる。2.5.1小節で示したPCPに対し、関

係代数における「選択 (Selection)」, 「射影 (Projection)」, 「結合 (Join)」の機能を忠実に実装し、それら演算を可視化システム上で簡単に利用可能な仕組みを備える。すなわち、「選択 (Selection)」により属性値による絞込みを行い、「射影 (Projection)」により表示する属性を切り出し、「結合 (Join)」により属性の追加を可能とする。以下では、SPJ質問とPCPの関係について定義する。

選択 (Selection)

選択操作は、軸に対し図3の赤枠のような形で範囲選択をすることで、表示するタプルを絞り込む機能を以て実現する。すなわち、本稿では図4のように $R[A_i \geq y_1] \wedge R[A_i \leq y_2]$ を行う操作として定義する。

射影 (Projection)

射影操作は、図5のように軸を切り出す機能を以て実現する。すなわち、 R の全属性集合 $\{A_1, A_2, \dots, A_n\}$ の部分集合 $X = \{A_{i1}, A_{i2}, \dots, A_{ik}\}$ に該当する属性のみを切り出す操作

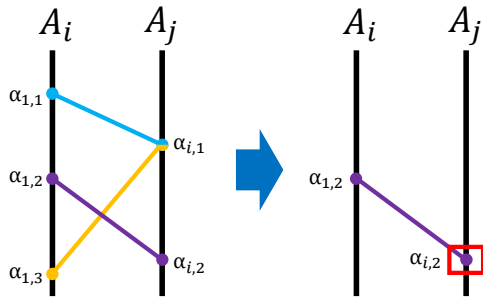


図3 選択 (Selection) の例

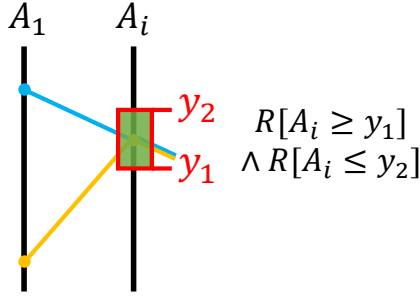


図4 選択 ($R[A_i \geq y_1] \wedge R[A_i \leq y_2]$)

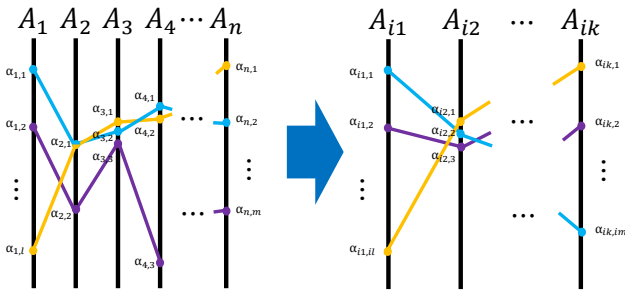


図5 射影 ($R[X], X = \{A_{i1}, A_{i2}, \dots, A_{ik}\}$)

を $R[X]$ として定義する。また、軸を切り出した後は軸間を対応する各タプル同士で結合しなおすものとする。

加えて、軸の順番を決定する操作を Coordinating として定義する。これは本来関係代数にはない定義ではあるが、PCP においては軸の順番によってクラスタの見え方や軸間の相関 [22] が現れるため重要な定義である。

結合 (Join)

結合操作は、図 6 のように軸を追加する機能を以て実現する。すなわち、リレーション $R(A_1, A_2, \dots, A_n)$ と、リレーション $S(B_1, B_2, \dots, B_m)$ それぞれが持つタプルを、 $A_i \theta B_j$ の条件で繋ぎ、軸の追加を行う操作を $R[A_i \theta B_j]S$ として定義する。なお、本稿で扱うことができる θ は、 $=$ のみとする。

2.6 (PC)²L

本節では、我々の提案システムにおいて重要な構成要素の 1 つである (PC)²L¹ について述べる。

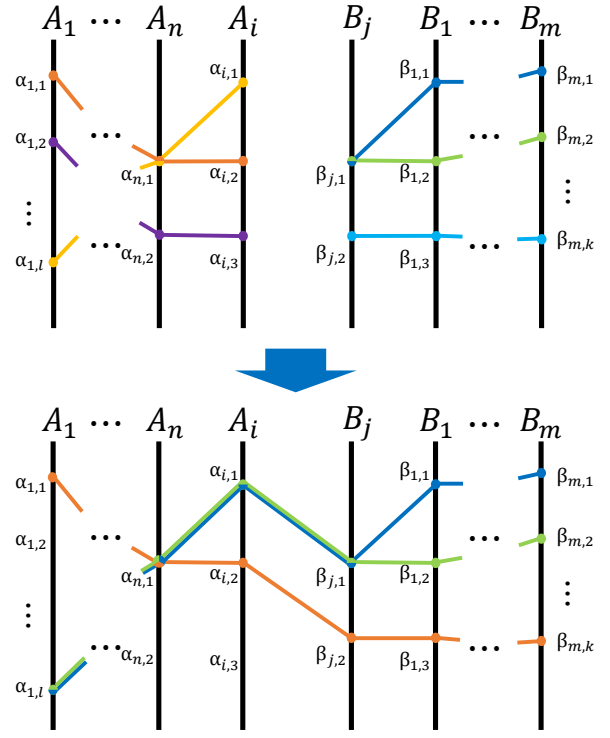


図6 結合 ($R[A_i \theta B_j]S$)

本節では、2.5 節での定義をもとに、PCP に対するインタラクションの過程を保存・再現するための言語である (PC)²L を定義する。すなわち本機能は、PCP における操作結果を言語化し保存する仕組みと、逆にその言語をもとに過去の PCP の操作結果を再現する仕組みを備える機能である。

以下では、(PC)²L の言語仕様の概要を述べる。なお、提案システムが生成および受理すべき (PC)²L の文法の詳細については、付録 1 に示す。本言語については、SQL の SELECT, FROM, WHERE 句に加えて COLOR BY 句を新たに利用可能とする。但し、使用しているライブラリの制約および実装の都合上、実装システムにおいては以下の制約が存在する。

- WHERE 句において、論理演算は AND および OR のみ使用可能である。
- WHERE 句において、同一軸上の論理演算は OR のみ使用可能である。
- WHERE 句において、軸間の論理演算 (AND/OR) は一括で同一のものだけ使用可能である。
- JOIN 句において、INNER JOIN と LEFT OUTER JOIN, RIGHT OUTER JOIN, FULL OUTER JOIN のみ使用可能であり、直積や自然結合などは使用できない。
- JOIN 句において、条件式として記述可能なものは 1 つのみ許容する。

2.7 実装方法

提案システムの画面の例を図 7 に示す。本画面では、関係代数演算の一部である、SPJ 質問の機能をシステム上で実現可能とした。また、提案システムは環境を問わず利用できるようにするため、Web ブラウザを通して多くの端末から利用できるように構築

1 : Parallel Coordinate Plots Commutative Language の頭文字より。

した。開発言語としては、サーバーサイド処理に PHP、クライアント処理に HTML, CSS, JavaScript を使用した。

画面上部 A では PCP を表示している。画面下部 C では、画面上部で操作した結果に対応するデータが抽出および表示されている。なお、画面上部 A の PCP 部分に関しては、「D3.js²」ならびに「Parallel Coordinates³」を使用した。

2.7.1 実装システムにおける SPJ 質問の実行

本画面における操作と 2.5.2 小節で述べた選択 (Selection)、射影 (Projection)、結合 (Join) の各操作について述べる。

図 7 において、画面上部 A の PCP 各軸上をマウスカーソルでドラッグすることで範囲選択が可能となっており、これにより選択 (Selection) が利用可能である。

また、画面中央 B にあるトグルボタンは、横 1 行がリレーション 1 つと対応しており、射影 (Projection) と結合 (Join) が利用可能である。ボタンの最左部には、結合 (Join) のためのリレーションボタンが配置されており、クリックすることで図 8 の画面が表示され、リレーションの結合条件を選ぶことができる。なお、すでに結合 (Join) されているリレーションボタンをもう一度クリックすると、結合 (Join) を解除できる。

リレーションボタンより右側には、射影 (Projection) のための軸ボタンが配置されており、クリックすることで軸の表示/非表示を切り替えることができる。なお、これらトグルボタンに表示する要素の設定については別途設定ファイルにて管理しており、データセットに特化した実装にはしていない。

また、画面上部 A の軸名部分をクリックしながら左右にドラッグすると軸の Coordinating が行うことができる。

2.7.2 実装システムにおける (PC)²L の実行

2.6 節で述べた (PC)²L について説明する。本機能は、図 7 の画面 C の (PC)²L タブから利用可能である。ページ下部に表示されるテキストフィールドに (PC)²L を記述し、PCP 上に反映が可能である。一方で、PCP 上の操作結果に関しても、(PC)²L によって記述されたテキストファイルとしてダウンロード可能である。なお、(PC)²L を解析するパーザに関しては、「PEG.js⁴」を用いて作成した。

画面上部 A の PCP 軸上をクリックすることで、その属性を基準に Coloring が可能である。(PC)²L においては COLOR BY 句に相当する。また、今回は Coloring の色の順番に関しては、属性値が数値の場合は昇順に「青、緑、黄、赤」が滑らかに変化するように配色し、文字列の場合は固定色で配色する。

3 情報可視化例

3.1 情報可視化例 1: 小袖屏風

これまでに我々は、多様で散在するデジタルアーカイブや書籍をデータベースに統合した「小袖屏風⁵データベース」の構築を

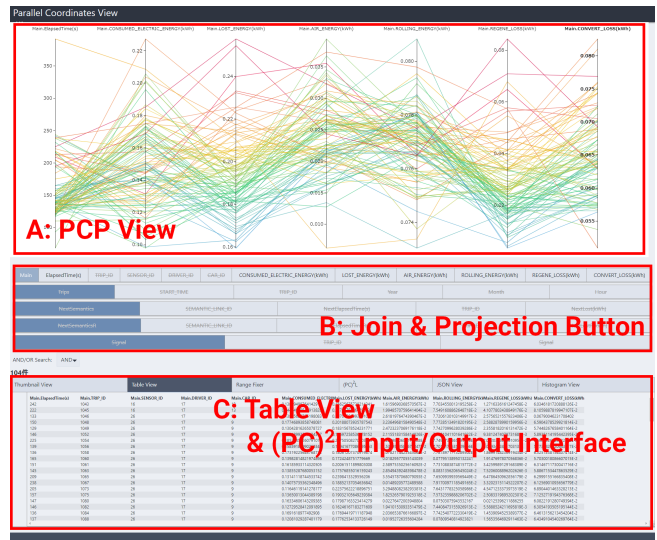


図 7 提案システムの表示例

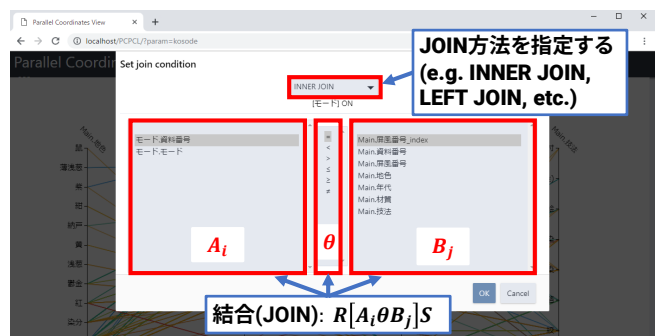


図 8 結合 (Join) の条件を指定する画面例

行ってきた [24] [25]。本データベースにおいては、服飾文化財である小袖屏風 100 隻および小袖 111 着について検索が可能である。本節では、このデータベースを使用して情報可視化例を示す。具体的には、文献 [24] [25] で述べている服飾様式の流行の変遷に基づいて可視化を行う。

ここでは、服飾様式の流行 (以下、「モード」という) に着目した組合せを明示する操作について説明する。なお、本節で扱うリレーションおよび属性は表 2 に示す。操作開始時の PCP と適用する (PC)²L、適用結果の PCP について図 10 に示す。この結果を導出する上では、2.1 節で示したような PCP 上での操作を試行錯誤しながら所望の結果を獲得した。その手順について以下で説明する。

まず、PCP に対し R_1 を適用することで、小袖屏風に関するデータの全体像を PCP で表示することができる。しかしながら、この状態ではどこに着目すればよいのかが不明瞭である。

そこで今回は、専門家の知見の一例として、小袖の流行を示す「モード」に関するリレーション R_2 を結合 (Join) する。すなわち、 R_1 に対し、 R_2 を小袖 1 枚に付与されている「資料番号」をキーとして結合 (Join) を行う。これにより、図 9 に示すような表示となる。なお、図 9 の赤丸は絞りの流行を示す要素であり、青丸は染めの流行を示す要素である。この図から、およそ「モー

に貼り付けることで、衣服本来の印象を再現したものである [23]。

2 : <https://d3js.org/>

3 : <https://syntagmatic.github.io/parallel-coordinates/>

4 : <https://pegjs.org/>

5 : 小袖屏風とは、安土桃山時代から江戸時代ごろに実在した小袖 (現在の着物の原型) のうち、状態の芳しくない小袖の裂の実物を金箔地や金砂子地の二曲一隻屏風

表2 小袖屏風に関するリレーションとそこにある属性

リレーション	属性
R_1	屏風番号, 資料番号, 年代, 材質, 技法
R_2	資料番号, モード

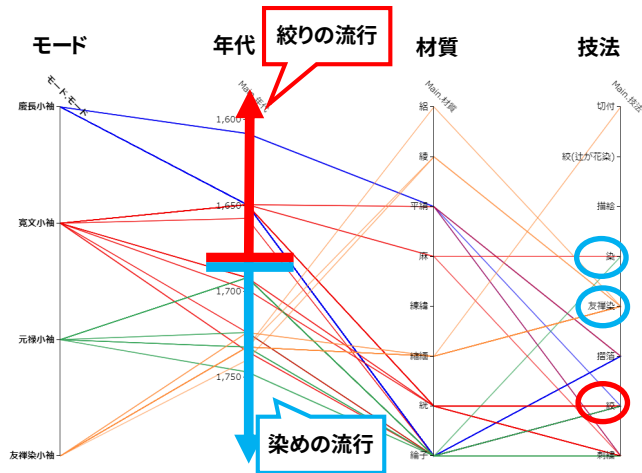


図9 小袖屏風の流行の変遷とPCPにおける可視化

ド」というものは「年代」の変遷と関係があり、「モード」とは小袖の流行と関係があるものとみることができる。

加えて、「モード」軸内にある要素「慶長小袖」は、「技法」軸内にある要素「絞」に対して多くのクラスタが存在することがわかる。それに対して、「モード」軸内にある要素「寛文小袖」と「元禄小袖」は、「技法」軸内にある要素「絞」と「染」に対してクラスタが存在することがわかる。一方で、「モード」軸内にある要素「友禅染小袖」は、「技法」軸内にある要素「友禅染」に対して多くのクラスタが存在することがわかる。これらのことから、以下のことが可視化により確認できる。

- ・「慶長小袖」は絞りを中心とした流行がある
- ・「寛文小袖」「元禄小袖」は絞りを中心としつつ染めの流行がある
- ・「友禅染小袖」は染めを中心とした流行がある

すなわち、「モード」の変遷に伴い、小袖が持つ「技法」の要素が変化の様子を確認できる。

また、上記操作の過程では試行錯誤を行いながら操作しているため、分析者はその過程を保存しながら進める。本システムでは、2.6節で示した $(PC)^2L$ によって逐一操作結果を保存可能であり、後戻りが発生した際にも $(PC)^2L$ クエリを呼び出すことで再現が可能となっている。

以上から、PCPに対する操作および $(PC)^2L$ による操作を行うことで、小袖や服飾の専門家が持つ「モード」と「技法」との関係可視化により明示できた。

3.2 情報可視化例2: 自動車走行ログ

これまでに我々は、自動車走行ログから「その人がEV(Electric Vehicle)に乗り換えて同じ運転をしたときの消費エネルギー」を推定しデータベースに蓄積しておくことで、EVの電力消費に関



図10 PCPに小袖屏風の「モード」の変遷に基づいた可視化を行う $(PC)^2L$ を適用した例

する多様な検索を実現する ECOLOG(Energy Consumption LOG) システムを提案してきた [26]. 本節では、この ECOLOG システムにおける情報可視化例を示す。具体的には、表4の区間Aを対象とした損失エネルギーの要因分析を行うことを考える。なお、詳細な分析に関しては文献 [26] で述べているため、ここでは所望の可視化がどのような手順によって達成したかについて述べる。

ここでは、区間Aに存在する信号Xの影響による、後続する区間Bの損失エネルギーの影響を明示するまでの操作について説明する。本節で扱う属性を表4に示す。なお、「Signal」は区間Aに存在する信号Xの直前に滞在する時間により「通過」、「停止」をラベル付けしたもの、「ElapsedTime(s)」はある区間を通過するためにかかった時間、「LostEnergy(kWh)」はある区間を通過するために損失した電力量を表す。また、分析対象の条件を表3に示す。なお、ここでの1トリップとは1回の運転のことであり、PCPにおける1タプルはこの1トリップを単位とした。

操作開始時のPCPと適用する $(PC)^2L$ 、適用結果のPCPについて図11に示す。この結果を導出する上では、試行錯誤しながら提案システム上で操作を行い、所望の結果を獲得した。その手順について以下で説明する。

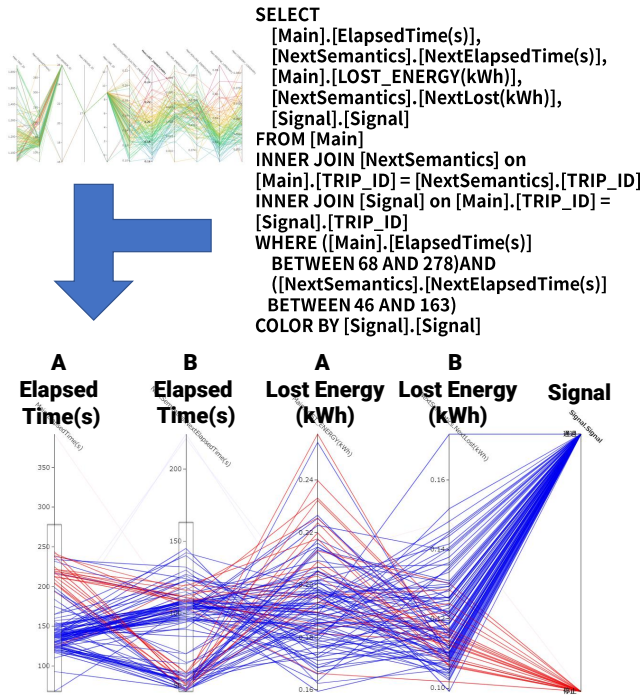
まず、PCPに対し S_1 を適用することで、区間Aに関するデータの全体像をPCPで表示することができる。しかしながら、この状態ではタプル線が複雑に交差しており、分析が難しい。

そこで今回は、区間Aに新たに設置された信号Xの影響を考えるために、リレーション S_2 を結合(Join)する。すなわち、 S_1 に対し、 S_4 をトリップに付与されている「TRIP_ID」をキーとして結合(Join)を行う。これにより、信号Xの「通過」、「停止」を区別しながら、区間Aにおける「ElapsedTime(s)」「LostEnergy(kWh)」の相関を可視化できる。

さらに分析者は、区間Aに隣接する区間Bや区間B'と、信号Xの関係の分析を行いたくなった。このような場合でも、提案システム上ではアドホックに操作を実行することができる。すなわ

表3 集計条件

集計期間	2017/8/4 – 2018/12/18
区間 A	岸谷 2 丁目～新子安
区間 B	新子安～七島町
区間 B'	宮ノ下～岸谷 2 丁目
進行方向	横浜方面 (上り)
被験者数	1 名
センサ	Zenfone2
トリップ数	98 件

図 11 A 区間にある信号 X による損失エネルギーの影響を表す可視化を行う (PC)²L を適用した例

ち、 S_1 と S_2, S_3, S_4 のリレーションを「TRIP_ID」をキーとして結合 (Join) を行うことで、PCP 上でそれぞれ属性間の関係を分析できる。これらの操作により分析者は、『信号 X の影響を分析するためには C(=A 区間+B 区間) という区間に着目するとよい』、という新たな知見をもたらした。

また、3.1 小節と同様に、分析者は試行錯誤しながら操作結果を (PC)²L で保存しながら進めことができる。後戻りが発生した際にも (PC)²L クエリを呼び出すことで再現が可能である。

以上から、PCP に対する操作および (PC)²L による操作を行うことで信号 X と各区間との関係を分析できた。

4 まとめと今後の課題

本研究では、可視化手法の 1 つである平行座標プロットを関係度数における SPJ に相当する表現力を持つように拡張した情報可視化システムを提案した。加えて、PCP の操作結果を保存・再現するための言語として (PC)²L の提案も行った。また、本システムを服飾文化財である小袖屏風と、自動車走行ログという全く異

なる 2 つのコンテンツに対して適用し、小袖屏風では服飾のモードの変遷を示し、自動車走行ログでは分析対象の区間 A について新たな知見をもたらした。

今後は、(PC)²L の受理できる文法の拡張を行い、PCP との可換性を高めていく必要がある。また、可視化事例をさらに積み重ね、より有益な情報可視化を目指す。

謝辞 本研究の一部は、JSPS 科研費 (課題番号 18K11750) の助成による。また、一部は国立歴史民俗博物館共同研究 (平成 25 年度～平成 27 年度)「歴史資料デジタルアーカイブデータを用いた知的構造の創生に関する研究—小袖屏風を対象として⁶」より着想を得た。

文 献

- [1] 総務省. 平成 30 年度版情報通信白書. <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h30/pdf/index.html>. Last accessed: 12 February 2019.
- [2] Alfred Inselberg. The plane with parallel coordinates. *The Visual Computer*, Vol. 1, No. 2, pp. 69–91, 1985.
- [3] Jimmy Johansson and Camilla Forsell. Evaluation of parallel coordinates: Overview, categorization and guidelines for future research. *IEEE Trans. on Visualization and Computer Graphics (TVCG)*, Vol. 22, No. 1, pp. 579–588, 2016.
- [4] Fatma Bouali, Abdelheq Guettala, and Gilles Venturini. VizAssist: An interactive user assistant for visual data mining. *The Visual Computer: Int'l Journal of Computer Graphics*, Vol. 32, No. 11, pp. 1447–1463, 2016.
- [5] Takayuki Itoh, Ashnil Kumar, Karsten Klein, and Jinman Kim. High-dimensional data visualization by interactive construction of low-dimensional parallel coordinate plots. *Journal of Visual Languages & Computing*, Vol. 43, pp. 1–13, 2017.
- [6] Zhiguang Zhou, Zhifei Ye, Jiajun Yu, and Weifeng Chen. Cluster-aware arrangement of the parallel coordinate plots. *Journal of Visual Languages & Computing*, Vol. 46, pp. 43–52, 2018.
- [7] 岡本辰夫, 小山嘉紀, 松田敏之, 池田隼, 古川文, 横田一正. 美術作品の素材要素検索による興味喚起と鑑賞を支援するパーツミュージアムの開発と評価. 日本データベース学会論文誌, Vol. 7, No. 4, pp. 19–24, 2009.
- [8] Wataru Wakita and Hiromi T. Tanaka. A digital archiving for large 3D woven cultural artifacts exhibition. *ITE Trans. on Media Technology and Applications*, Vol. 2, No. 3, pp. 236–247, 2014.
- [9] 安達文夫, 青山宏夫, 田中紀之, 徳永幸生. 時間と場所の情報を有する大量の写真資料の提示法. 国立歴史民俗博物館研究報告, Vol. 189, pp. 41–77, 2015.
- [10] Koichi Yoshino, Koichi Obata, and Satoru Tokuhisa. FLIPPIN': Exploring a paper-based book UI design in a public space. In *Proc. of the 2017 CHI Conf. on Human Factors in Computing Systems (CHI '17)*, pp. 1508–1517. ACM, 2017.
- [11] Alice Thudt, Uta Hinrichs, and Sheelagh Carpendale. The bohemian bookshelf: Supporting serendipitous book discoveries through information visualization. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (CHI '12)*, pp. 1461–1470. ACM, 2012.
- [12] Lu An, Liqin Zhou, Xia Lin, and Chuanming Yu. Visual topical analysis of museum collections. In *Proc. of the 17th Int'l Conf. on Asia-Pacific Digital Libraries (ICADL2015)*, pp. 1–11. Springer-Verlag, 2015.
- [13] Mitchell Whitelaw. Generous interfaces for digital cultural collections. *DHQ: Digital Humanities Quarterly*, Vol. 9, No. 1, 2015.
- [14] Daan Odijk, Cristina Gărbacea, Thomas Schoegje, Laura Hollink,

6: <http://www.rekihaku.ac.jp/research/list/joint/2013/digitalarchive.html>

表4 ECOLOG システムに関するリレーションとその中にある属性

リレーション	属性	説明
S_1	TRIP_ID, ElapsedTime(s), LostEnergy(kWh)	区間 A における自動車走行ログ
S_2	TRIP_ID, ElapsedTime(s), LostEnergy(kWh)	区間 B における自動車走行ログ
S_3	TRIP_ID, ElapsedTime(s), LostEnergy(kWh)	区間 B' における自動車走行ログ
S_4	TRIP_ID, Signal	区間 A にある信号 X について「通過」または「停止」を TRIP_ID ごとに付与したラベル

- Victor de Boer, Kees Ribbens, and Jacco van Ossenbruggen. Supporting exploration of historical perspectives across collections. In *Proc. of the 19th Int'l Conf. on Theory and Practice of Digital Libraries (TPDL2015)*, pp. 238–251. Springer-Verlag, 2015.
- [15] Uta Hinrichs, Stefania Forlini, and Bridget Moynihan. Speculative practices: Utilizing InfoVis to explore untapped literary collections. *IEEE Trans. on Visualization and Computer Graphics (TVCG)*, Vol. 22, No. 1, pp. 429–438, 2016.
- [16] Katrin Glinka, Christopher Pietsch, and Marian Dörk. Past visions and reconciling views: Visualizing time, texture and themes in cultural collections. *DHQ: Digital Humanities Quarterly*, Vol. 11, No. 2, 2017.
- [17] 飯島護久, 堀口良太. ブロープデータに基づくエリア流動性情報提供に関する研究. 第9回 ITS シンポジウム 2010 論文集, pp. 1–4. 2010.
- [18] 敏洋平岡, 聖明西川, 浩司川上, 隆之塩瀬. 自発的な省燃費運転行動を促すエコドライブ支援システム. 計測自動制御学会論文集 = Transactions of the Society of Instrument and Control Engineers, Vol. 48, No. 11, pp. 754–763, 2012.
- [19] Isabel Wilmlink, Johan Olstam, Martijn van Noort, and Charisma F Choudhury. Microscopic simulation of eco-driving systems using real-world vehicle models. In *22nd ITS World Congress, Bordeaux*, pp. 1–12, 2015.
- [20] Takahiro Komamizu, Toshiyuki Amagasa, and Hiroyuki Kitagawa. Visual spatial-OLAP for vehicle recorder data on micro-sized electric vehicles. In *Proc. of the 20th Int'l Database Engineering & Applications Symposium (IDEAS '16)*, pp. 358–363. ACM, 2016.
- [21] 増永良文. リレーショナルデータベース入門—データモデル・SQL・管理システム. サイエンス社, 2003.
- [22] Mats Lind, Jimmy Johansson, and Matthew Cooper. Many-to-many relational parallel coordinates displays. In *13th Int'l Conf. on Information Visualisation*, pp. 25–31, 2009.
- [23] 国立歴史民俗博物館 (編). 国立歴史民俗博物館資料図録 2 野村コレクション 小袖屏風. 国立歴史民俗博物館, 2002.
- [24] 萩生田明徳, 田中友章, 大門利都子, 富井尚志. 専門家の知見を用いた小袖屏風閲覧システムの構築と展示のストーリーに沿った情報提示. 第8回データ工学と情報マネジメントに関するフォーラム (DEIM2016), A4-1, pp. 1–8. 2016.
- [25] 濱崎裕太, 大門利都子, 田中友章, 富井尚志. デジタルアーカイブをSPJ 質問により統合する情報可視化システム. 第10回データ工学と情報マネジメントに関するフォーラム (DEIM2018), E5-3, pp. 1–8. 2018.
- [26] 植村智明, 濱崎裕太, 齊藤祐亮, 吉瀬雄大, 磯部康太, 有長拓海, 富井尚志. EV 消費エネルギーデータベースの構築と連続する道路間の相関を考慮したデータ分析手法. 第11回データ工学と情報マネジメントに関するフォーラム (DEIM2019), E7-5, pp. 1–8. 2019.

付 録

1 BNF 記法による (PC)²L の文法定義

```
<Query> ::= <SelectClause> <FromClause> <WhereClause>
          | <SelectClause> <FromClause> <ColorByClause>
          | <SelectClause> <FromClause> <WhereClause>
          | <SelectClause> <FromClause>
```

```
<SelectClause> ::= 'SELECT' <ColumnList>
<FromClause> ::= 'FROM' <JoinList>
<WhereClause> ::= 'WHERE' <ConditionList>
<ColorByClause> ::= 'COLOR' 'BY' <ColumnName>
<ColumnList> ::= <ColumnName> ',' <ColumnList>
                | <ColumnName>
<JoinList> ::= <RelationName> <JoinClause>
                | <RelationName>
<ConditionList> ::= '(' <AxisCondition> ')'
                  | <LogicalOperator> <ConditionList>
                  | '(' <AxisCondition> ')'
<JoinClause> ::= <JoinStatement> <JoinClause>
                | <JoinStatement>
<AxisCondition> ::= <Condition> <AxisCondition>
                  | <Condition>
<LogicalOperator> ::= 'AND' | 'OR'
<JoinStatement> ::= <JoinMethod> <RelationName>
                  | 'ON' <JoinCondition>
<Condition> ::= <ColumnName> <Range>
<JoinMethod> ::= 'INNER' 'JOIN'
                | 'LEFT' 'JOIN'
                | 'LEFT' 'OUTER' 'JOIN'
                | 'RIGHT' 'JOIN'
                | 'RIGHT' 'OUTER' 'JOIN'
                | 'FULL' 'JOIN'
                | 'FULL' 'OUTER' 'JOIN'
<JoinCondition> ::= <LeftCondition> <Theta>
                  | <RightCondition>
<Range> ::= 'BETWEEN' <NumberLiteral> 'AND'
            | <NumberLiteral>
<LeftCondition> ::= <ColumnName>
<Theta> ::= '=' | '>' | '<' | '>=' | '<=' | '<>' | '!= '
<RightCondition> ::= <ColumnName>
<NumberLiteral> ::= DecimalLiteral | FloatLiteral
<ColumnName> ::= <RelationName> '.' <AxisName>
<RelationName> ::= '[' UnescapedIdentifier ']' | Identifier
<AxisName> ::= '[' UnescapedIdentifier ']' | Identifier
```

但し, DecimalLiteral, FloatLiteral, UnescapedIdentifier, Identifier はそれぞれ終端記号を表す. これら終端記号の定義は, 以下の正規表現で表す.

```
DecimalLiteral = [+]?([1-9][0-9]*0)
FloatLiteral = [+]?([1-9][0-9]*0)'.'[0-9]+
UnescapedIdentifier = [^\[\]x00-xfx7f0-9]
                    [^\[\]x00-xfx7f]*
Identifier = [^!,<,>=().\[\]'\\x00-20x7f0-9]
            [^!,<,>=().\[\]'\\x00-20x7f]*
```