

十二単ビューによる階層型多変数データの可視化と詳細度制御

山澤舞子 伊藤貴之

お茶の水女子大学 理学部情報科学科

1. 概要

情報可視化とは、物理的な形を持たない情報を、CGによって視覚的に表現する技術である。大規模なデータを可視化することで、ユーザは複雑な情報への理解を深めることができる。また可視化することで、ユーザは情報の中の特に注目すべき部分を見つけ出せるようになる。

階層型多変数データを情報可視化する一手法として、「十二単ビュー」[1]という手法が提案されている。十二単ビューは、大規模な階層型多変数データ全体を一画面に表示できるという特徴がある。個々のノードの詳細情報を見たいときはビューを拡大し、データ全体を見渡したいときはビューを縮小することで、ユーザのニーズに合った情報を取り出すことができる。

本報告では、十二単ビューの機能拡張として、拡大表示時にはデータの葉ノードを単位として変数値を表示し、縮小表示時には下位階層の変数値を統合し、上位階層のみを表示する、というような詳細度制御を実現する手法を提案する。この機能拡張により、画面のサイズやユーザの操作に合った最適な情報量で、階層型多変数データを可視化できると考えられる。

2. 関連研究

提案手法が前提とする可視化手法「十二単ビュー」は、大規模階層型データ可視化手法「平安京ビュー」[2]の拡張手法である。平安京ビューは、階層型データの葉ノードをアイコンで表示し、枝ノードを長方形の枠で表示することで、階層型データの全体を一画面に表現することができる手法である。その他の階層型データの可視化手法として、階層型データを木構造により可視化する Hyperbolic Tree[3] や2次元平面を長方形で分割していく Treemaps[4] などがある。

また、多変数データの可視化手法として、変数の本数の平行軸を用いる Parallel Coordinates[5] などがある。しかし階層構造と多変数情報を同時に表現した可視化手法はまだ少なく、議論の余地が残っていると考えられる。

3. 提案内容

十二単ビューで多変数データを表示する際は、アイコンを以下のように様々な視覚的特性を用いて表示する。

◆ データの変数の個数を n としたとき、アイコンを n

“Visualization and the level of detail control of hierarchical multi-variate data using JunihitoeView“

Maiko Yamazawa, Takayuki Itoh, Ochanomizu University

{maiko, itot}@itolab.is.ocha.ac.jp

個の領域に分割し、 i 番目 ($0 \leq i < n$) の領域を色相

$$H = \frac{2\pi i}{n} \quad (0 \leq H < 2\pi)$$
 で塗り分けて表示する。

◆ i 番目の領域の変数値 t_i ($0 \leq t_i \leq 1$ の値に正規化されている)によって、 i 番目の領域を、以下の算出方法で求められる明度 S ($0 \leq S \leq 1$)・彩度 I ($0 \leq I \leq 1$) で表示する。

$$S = I = 0.2 + 0.8t_i$$

十二単ビューを用いて階層型多変数データを可視化した例を、図1に示す。

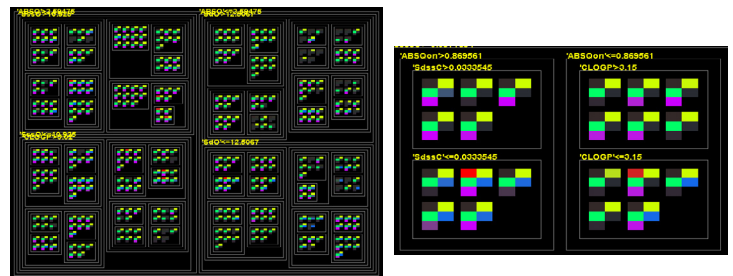


図1: (左)十二単ビューによる多変数データの可視化結果 (右) (左)を拡大した図

図1(左)は十二単ビューで階層型多変数データを可視化し、データ全体を表示した結果であり、図1(右)は図1(左)を部分的に拡大した図である。図1(右)からわかるように、十二単ビューでは、複数の色相で塗り分けた長方形領域によって、葉ノードが有する多変数を表現する。個々のノードの詳細を見るためには、図1(右)のように拡大することで、必要な情報を得ることができる。しかし、図1(左)のようにデータ全体を一画面に表示させたときには、それぞれのノードは小さすぎて判別しにくいことがある。

そこで同じ階層にある葉ノードを全て統合し、この階層を表すノードとして表示することで、画面上に表示されるアイコンの数とサイズを調節することを考える。

図2は、4個の5変数の葉ノードがある階層を図示した例である。図2(左)ではこの階層を4個の葉ノードをアイコンで表現しているが、提案手法における詳細度制御では図2(右)のように、これら4個の葉ノードの変数値から算出した数値を、この階層を示すノードとして1つのアイコンで表現する。ここで、同一階層に属するノード変数値の傾向を表す数値には、平均値・最大値・最小値・分散などがあげられる。これらの数値をどのように複合的に、かつ効果的に表現するか、今後検討したい。

以上の表示方法に加えて提案手法では、ユーザによる拡大縮小操作に伴って詳細度を自動制御する機能を実装する。この機能では縮小操作に伴い、その縮小率に合わせて、自動的に下位階層のノードを統合して上位階層のアイコンを表示する。また拡大操作に伴い、拡大率に合わせて自動的に下位階層のノードを表示する。これにより、拡大操作によって個々の葉ノードの変数値を、また縮小操作によってデータの全体像を見渡しながら、下位階層の変数値の傾向を読み取れるようになる。

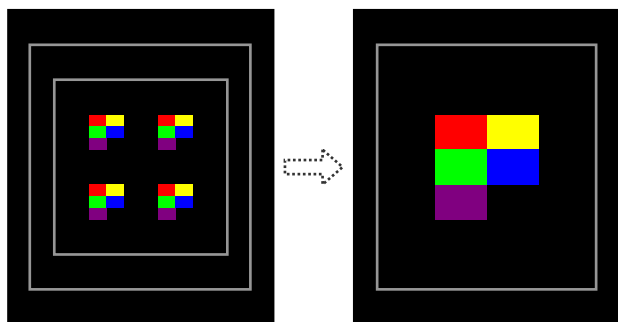


図 2: 同一階層のノードを統合して表示した例

4. 実行結果と課題

図 3 は、図 1(左)の可視化結果の下位階層にある葉ノードの変数値を統合し、データの階層の最上位から、任意に指定した階層までを表示した例である。今後の課題として、ユーザによる拡大操作または縮小操作に伴った対話的な詳細度制御を実装する予定である。

図 4 は、階層のバランスが悪いデータを、下位階層を統合して可視化を行った図である。赤色の円で囲まれた部分では、同一階層内にある 2 つのノードが、異なるサイズのアイコンで表示されている。また、青色の円で囲まれた部分では、アイコンの縦横比が崩れ、変形してしまっている。このような、アイコンの表示方法について、今後検討していく予定である。

5. 適用事例

提案手法の適用事例として我々は、以下に示す薬物データを用いて可視化を試みている。

まず、多数の薬物を、分子構造上の特徴によって階層型に分類したデータを用意する。次に、各薬物に対する数種類の実験値(本適用事例では数種類の酵素に対する代謝感受性)を、各薬物に対して与える。以上により、各々の薬物が数種類の代謝感受性の値をもつ、という階層型多変数データが得られる。

現実の製薬開発では、数万から数十万個の薬物データを保有しているが、我々は現在、161 個の薬物データが 5 種類、または 6 種類の代謝感受性を有する階層型多変数データを、適用事例として用いている。

6. まとめ

本報告では、階層型多変数データ可視化手法「十二単ビュー」において、詳細度制御を実現する手法を提案した。今後の課題として、ユーザの操作に対する対話的な詳細度制御を実装したい。また、図 4 に示したような好ましくな

い可視化結果を生む場合への対策を検討したい。また、提案手法を用いて薬物データを可視化した結果から、薬物の分析などの研究に対して提案手法がどのように貢献できるか検討したい。



図 3: 下位階層を統合し、上位階層のみを表示した例

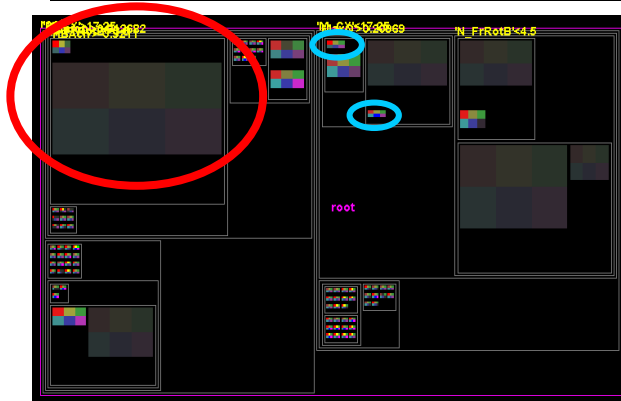


図 4: 階層のバランスが悪いデータの例

謝辞

薬物データを提供して頂いた京都大学薬学部の山下 義助教授に感謝いたします。

参考文献

- [1] 伊藤, 山下, 十二単ビュー:大規模階層型多変数データの可視化の一手法, 情報処理学会データベース と Web 情報システムに関するシンポジウム(DBWeb 2006), 2006.
- [2] 伊藤, 山口, 小山田, 長方形の入れ子構造による階層型データ視覚化手法の計算時間および画面占有面積の改善, 可視化情報学会論文集, Vol. 26, No. 6, pp. 51-61, 2006.
- [3] Lamping J., Rao R., Pirolli P., A Focus+Context Technique Based on Hyperbolic Geometry for Visualizing Large Hierarchies, Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'95), pp. 401-408, 1995.
- [4] Johnson B., Shneiderman B., Treemaps: a space-filling approach to the visualization of hierarchical information structures, Proceedings of IEEE Visualization'91 Conference, San Diego, pp. 284-291, 1991.
- [5] Johansson J., Ljung P., Jern M., Cooper M., Revealing Structure within Clustered Parallel Coordinates Displays, Proceedings of IEEE Symposium on Information Visualization, (InfoVis 2005), pp. 125-132, 2005.