

比例シンボルマップによる高精度なデータ可視化

永田 彩季[†] 斎藤 隆文[‡]

東京農工大学 工学部情報工学科[†]

東京農工大学 大学院生物システム応用科学府[‡]

1. 背景と目的

データの可視化は、分析において非常に有効である。特に、データの値が一見して分かることや、大小関係を視覚的に把握できるような可視化方法が良いとされる。本研究では、データの可視化の中でも比例シンボルマップについて言及する。これは、地図上に量的な値を持つデータを表示する際に、円などの図形を用いて、その大きさで値の大小を表現する可視化手法である。しかし、従来のこの手法では、図形からは大雑把な大小関係しか読み取れず、正確な数値を読み取ることも困難である。また、値の大小比(ダイナミックレンジ)が大きいデータ群へも対応していない。そこで本研究では、比例シンボルマップを用いたデータ可視化手法の改善を試みる。特に、数値そのものや数値同士の大小比の正確な読み取りや、ダイナミックレンジの大きいデータ群の扱いが容易となるような手法の提示を目的とする。

2. 先行研究

ここでは、本研究に関連するデータ可視化手法として、最近提案された2つの研究について紹介する。

まず、Daniel らによる非線形ドットプロット[1]である。これは、棒グラフ状に点をプロットすることでデータを可視化する方法である。点の数が少ない場合は点を大きくし、点の数が多場合は点を小さくする。例を図1に示す。



図1 非線形ドットプロットの例

このグラフでは、点の数が少ない場合に数え上げによって正確な値が把握できる。また、数が多い場合は点を小さくしているため、グラフの高さを抑えることができる。これにより、ダイナミックレンジが大きいデータ群の可視化が可能である。

次に、Halwatsch らによる積み上げ型棒グラフ[2]である。これは、データ範囲が異なる複数の線形棒グラフを積み上げることによってデータを可視化する方法である。例を図2に示す。

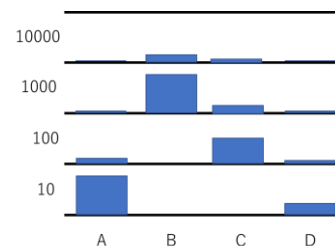


図2 積み上げ型棒グラフの例

このグラフは、積み上げられたそれぞれの棒グラフは線形であるため、視覚的なデータ大小関係と実際の大小関係に矛盾が出ない。また、様々な範囲のグラフを利用するため、ダイナミックレンジが大きいデータ群に対応している。さらに、グラフの範囲がそれぞれ異なるため、桁数に関する情報を加えることができる。

3. 可視化

本研究では、3種類の可視化方法を提案する。また、大小関係を示す図形の大きさの決定方法や、桁が増えた際の表示方法についても述べる。従来の手法と同様に、直感的な大小関係が把握できるように図形の大きさを数値によって変更する。それに加えて、正確な数値を読み取ることができる手法を提案する。

3.1 可視化方法

3種類の可視化方法について述べる。

手法Aとして、円を分割する手法を挙げる。これは、分割数によって、最上位の桁の値と、その1つ下の桁の値を示す方法である。例を図3に示す。

Precise Data Visualization with Proportional Symbol Map
[†]Aki Nagata, Department of Computer and Information Sciences, Tokyo University of Agriculture and Technology.
[‡]Takafumi Saito, Graduate School of Bio-Applications and Systems Engineering, Tokyo University of Agriculture and Technology.

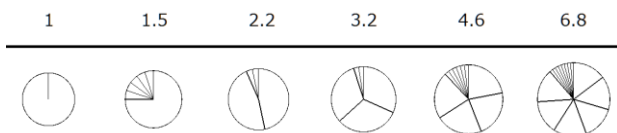


図3 手法 A の例

この手法は、Daniel らの手法[1]より、数え上げによるデータの可視化を利用している。

手法 B として、アナログ時計の表示方法を利用した手法を挙げる。これは、最上位の桁の値を短針で示し、それ以下の数値を、長針によって示す方法である。例を図4に示す。

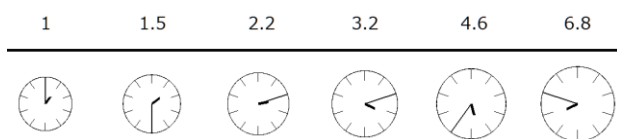


図4 手法 B の例

手法 C として、円を塗りつぶす手法を挙げる。これは、データの値によって円グラフのように円を塗りつぶすことで数値を示す方法である。円を 10 等分する線を利用し、数値を示す。例えば 2 という値を示す際は、円全体で 10 とみなし、2 の値まで塗りつぶすといった方法である。例を図5に示す。

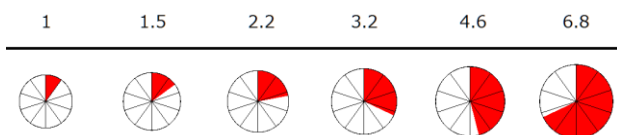


図5 手法 C の例

手法 B 及び C は、Halwatsch らの手法[2]より、離散的な目盛りによるデータの可視化を利用している。

3.2 桁の明示

ダイナミックレンジが広いデータ群を扱う場合、桁数の変化が大きくなる。そこで、桁数が変化した場合の表示方法について述べる。桁が増えた際、外円を増やす方法を提案する。増やした場合について、案 1 を利用して例を図6に示す。

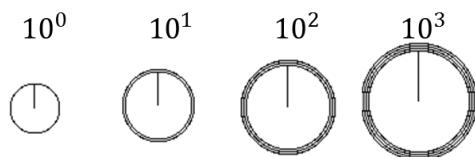


図6 桁によって外円を増やした例

これにより、データの桁数(指数部)を明示することができる。

3.3 円の半径設定

本手法では、データの大小関係を表すため、基準となる円の拡大を提案する。ひとつは面積を比例関係で拡大する方法であり、もうひとつは数値の対数を取ることで半径を決定し円を描画する方法である。面積を比例関係で拡大する方法では、視覚的な比較の容易さは向上する。しかし、ダイナミックレンジの広いデータ群に対応することが難しい。その場合、対数を取る方法を用いる。対数を用いた変換方法では、データの値が 1 の場合に、0 に変換されてしまう。この問題の解決のため、変換後の半径の値に 1 加算するなどして対応する。また、単純に対数を取った場合の大小差が微小になることの調整方法としても、変換後に一定の値を加減算する方法を取る。

ダイナミックレンジの広いデータで対数を用いて変換した場合、一見しただけでは桁数の違いが分かりづらくなるため、3.2 節で述べた桁の明示方法が特に有効である。

4.今後の課題

本研究では、直感的に大小関係を把握し、より正確に数値を読み取る可視化手法の提案を行った。提案手法は、値の大小関係を図形の大小関係と対応付けることで直感的な読み取りを、数え上げや目盛りを利用してより正確な読み取りを実現した。

今後の課題として、図形同士の重なりが挙げられる。現在、描画した図形が重なった場合を想定せず、あくまで図形全体が見えるような場合に関して比較を行っている。図形が重なることで、線や目盛の数え上げが困難になる。その場合の改善方法や、より分かりやすい可視化方法の提案が今後の課題である。

参考文献

- [1] Nils Rodrigues, Daniel Weiskopf, Nonlinear Dot Plots, IEEE Transactions on Visualization and Computer Graphics Vol.24, No.1, pp.616-625 (2018).
- [2] M. Halwatsch, F. Sadlo, M. Bunch, D. Weiskopf, Scale-Stack Bar Charts, Computer Graphics Forum, Vol.32, No.3, pp.181-190 (2013).