

多変量データの可視化における影の利用法の提案とその評価

鄭中翔 齋藤豪

東京工業大学大学院 情報理工学研究科

1 はじめに

データの可視化を行う際には、その概要を一目で読み取れるように表示することが望まれる。特に多変量データの可視化を行う際に、それぞれの変量を明確に区別し値を読み取れるように表示することはデータを解析する上で重要なことである。データの可視化を行う上で使用できる要素には色、模様、形がある[1]。影は色の変化の一種であると言えるが、光学現象として現れる影を意図的に可視化に使用した例は見られない。そこで我々は色とは独立かつ並列に認識することが容易であるという特徴を持つと考えられる影に注目し、影を利用した可視化手法を提案した[2, 3]。本稿ではこれらの手法中で用いられている影として表示した模様の可読性に関する評価実験について述べる。

2 評価対象の手法

提案手法[2, 3]は円柱の影を用いてデータを表示する。本稿では次に説明する簡略化した提案手法に関して評価を行う。表示するデータは平面上にマッピングすることができる二変量データとし、それぞれ var_1, var_2 とする。 var_1, var_2 は 0 から 1 までの値を取る。 var_1 は色、 var_2 は円柱の影の大きさに線形的に対応させる。 var_1 により決定された色を地の色と呼び、地の色の变化は赤から緑へのトーンを用いる。地の色及び影による明度の低下に関する色の計算は CIELAB 色空間で行なわれ表示される色 C は次の式で表される。

$$C = LabtoRGB(l, a, b) \quad (1)$$

$$l = L - t * s \quad (2)$$

$$a = k(var_1 - 0.5) \quad (3)$$

ここで、 L, k, t, b は定数、影を表す変数 s は 1 で完全に影、0 で影の外部的になっている状態とする。影同士の重なりをなるべく少なくするために、影はポアソンディスクサンプリングに従った配置とする。ただしポアソンディスクサンプリングにおける最小距離はその地点での値の大きさに応じたものとする。そのため、影の密度は変数の値が小さい領域では高く、値が大きい領域では低くなる。また、図 1 に表示される影と、結果の例を示す。

3 評価実験

3.1 影の読み取りに関する実験

影として表示した模様から変数の値を正確に読み取れるのかを評価する実験を行った。実験では $var_2 = 1$

Evaluation of visualization method using shadows for multivariate data

† Zhongxiang ZHENG

† Suguru SAITO

Graduate School of Information Science & Engineering, Tokyo Institute of Technology (†)



(a) 表示される影の模様 (b) 提案手法による結果の例

図 1: 提案手法で使用する影の模様と可視化結果

であるときの模様のみが凡例として表示される。被験者は可視化結果画像中に存在する模様のうちの一つを指定され、凡例と大きさ比較することで値を推測し、その値をスライダーにより解答する(図??)。解答は 0.01 単位で答えることができる。解答時間に制限はない。影の模様との比較対象として境界がはっきりとした黒い円の模様 A(図 2(a))、境界がぼやけた黒い円の模様 B(図 2(b))、ぼやけていない影の模様である模様 C(図 2(c)) を使用した。模様 B は模様 A をぼかして作成したものであり、図 1(a) に示される影の模様は、模様 C の上部を弱く、下部を強くぼかすことにより作られたものである。4 つの模様は表示面積が等しくなるように表示される。影の模様における最もぼやけている部分のぼやけ具合と模様 B のぼやけ具合は同等である。4 つの模様の配置点と大きさは、先に述べたアルゴリズムにより決まる。また、それぞれの模様は影と同様に明度の変化により表示される。実験時のパラメータとして $L = 69, t = 10, k = 100, b = 4$ を使用した。可視化の対象は気象データとし、気温を色、風速を模様割り当てた。気温についてはすべてのデータ中の最小、最大値がそれぞれ $[0, 1]$ に収まるように正規化した。風速については、様々なサイズの模様を表示するために、各時間における最小、最大値が $0 \sim 1$ に収まるように正規化した。4 人の被験者に対して実験を行い、各模様につき 60 問出題した。図 3 は解答の平均絶対誤差をグラフとして表示したものである。4 つの模様による結果を比較すると、模様 A, B の方が精度よく読み取ることができる傾向があるが、影の模様もほぼ同等の精度で読み取ることが可能であった。また、模様がぼやけたことにより可読性が大きく変化することはないことがわかる。

3.2 一覧性を評価する実験

この実験では可視化結果で表示されたデータの概要を短時間で読み取る際に、模様の種類と生成法により差があるのかを評価する。次の手順で実験を行う。最初に 2 組の 2 変量データ (A, B) と (A', B') を用意する。次に (A, B) をそれぞれ色と模様で可視化した結果を被験者に一瞬間提示し(図 5(b))、非表示にする

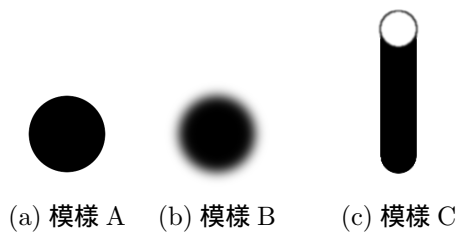


図 2: 比較対象として表示される模様.

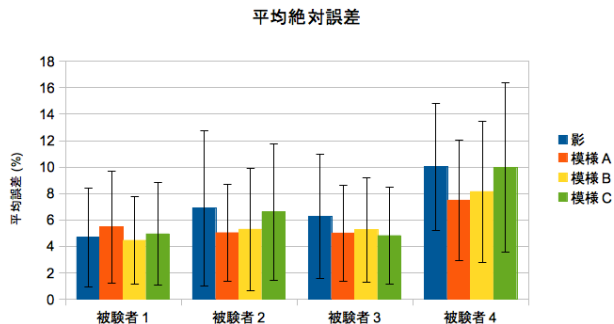


図 3: 模様から数値を読み取る実験の平均絶対誤差を被験者, 模様ごとに示した結果.

と同時に (A, B) , (A', B) , (A, B') , (A', B') の 4 種類の可視化結果を提示する (図 5(c)). このとき, 色によるデータの表示は先に表示されるものと全く同じであるが, 模様の表示位置は新たに生成される. これにより模様の位置を覚えて解答することを防ぐ. 被験者はこの中から (A, B) を可視化した結果をキーボードにより選択する. 選択後にはブランクスクリーンが 1 秒間表示され, 次の問題に進む. この実験では色の区別は容易に行えるようにデータを選択した. 使用する模様は先の実験と同様の 4 種であり, それぞれの模様をアルファブレンディングによる合成と明度の低下により表示する. アルファブレンディングによる模様の表示では, $L = 69, a = 0, b = 0$ の無彩色に模様を表示した際の明度の低下が 10 になるようにアルファ値を調整した. 式 (1) ~ (3) 中のパラメータとデータは先の実験と同一である. 4 人の被験者に対して実験を行い, 各模様, 生成法につき 45 問出題した. 図 4 は各模様ごとの正答率のグラフである. いずれの場合でも高い正答率であることがわかる. このことから影の模様を用いた場合でもデータの概要は十分に読み取ることができる. また, この実験からは模様の生成法による違いは見られなかった. つまりアルファブレンディングと明度の低下によるどちらの生成法でも模様は十分に読み取れると言える.

4 考察

影の模様は円の模様比べて複雑な形状をしているが, 円のように単純な模様と同じように大きさから値の推定が可能であると言える. 円の模様の場合, 円の直径を比較することで値の推定が行われると考えられる. これに対して影の模様と模様 C の場合には長さを比較することで推定が行われると考えられる. 影の模様

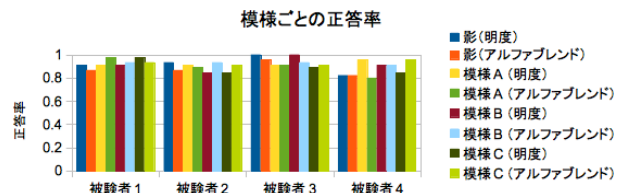


図 4: 一覧性に関する実験の正答率を模様, 生成法ごとに示した結果.

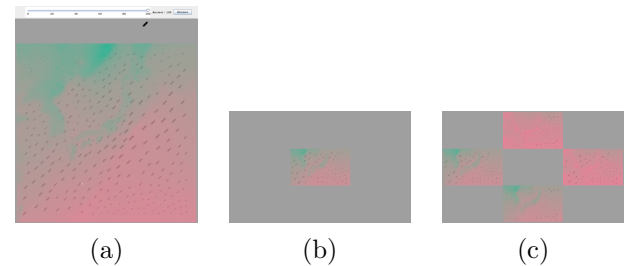


図 5: (a) 影の読み取りに関する実験で被験者に表示される画面. 上部に設置されたスライダーによって解答する.(b) 一覧性を評価する実験で被験者に始めに表示される画面.(c)(b) の表示後に表示される画面.

の方が値を推定する際に用いる基準を大きく表示できるが今回の実験ではその優位性はみられなかった. これは模様が単純に縮小されることで, 幅も変化するため比較が行いにくくなっていることが一因であると考えられる.

5 まとめと今後の課題

本稿では影の模様の可読性に関する実験を行った. その結果, 影の模様のように複雑な模様を用いた場合でも, 模様から値を読み取ることが可能であることがわかった. 今後の課題としてアルファブレンディングにより模様を表示した場合と, 明度を低下により模様を表示した場合で, 色の読み取りへの影響があるのかを評価する必要がある. また, 模様がぼやけていることなど影の特徴を取り入れた模様と, そのような特徴を持たない模様を明度の低下により表示した際に, 模様と色の読み取りやすさに違いがあるのかも評価する必要がある. これらの評価により, 影に注目した理由として挙げた影は色の読み取りに対して影響が少ないということの評価する予定である.

参考文献

- [1] H. Senay and E. Ignatius. A knowledge-based system for visualization design. *Computer Graphics and Applications, IEEE*, Vol. 14, No. 6, pp. 36–47, nov 1994.
- [2] 鄭中翔, 齋藤豪. 多変量データの可視化における影の利用に関する研究. 情報処理学会 グラフィックスと CAD 研究会, June 2012.
- [3] Z. Zheng and S. Saito. Multivariate data visualization using shadows. In *IEVC2012*, nov 2012.