

デバッグのための二次元配列データの可視化

3E-01

渡部秀文 斎藤隆文[†]

東京農工大学大学院工学研究科電子情報工学専攻

[†] 東京農工大学工学部情報コミュニケーション工学科

1. はじめに

本研究では、画像処理等のプログラミングのデバッグ支援手法としての二次元配列データの可視化法について提案する。一般的にプログラミングのデバッグにはデータの値をコンソールやファイルに数値で出力する方法が取られる。また、二次元配列を使ったプログラムでは画像として出力する事も考えられる。数値ベースの方法ではデータの読み取りの精度は保証される。しかし全体を表示するには表示スペースを浪費してしまう。一方数値を画素値にあてはめて画像で表示すれば、表示スペースはコンパクトにできる。しかし、データの読み取り精度がなくなる。また、画像として表示できるのは $[0, 255]$ の範囲であり、この範囲を大きく超えるデータの表現は難しい。この問題を解決する手法として、疑似カラーを使用することが考えられる。しかし、一般的な連続色では細かいデータの読み取りができず、離散的に色を塗り分ける手法では精度が取れない。本研究では、二次元配列を二色塗り分け疑似カラー表示を用いて可視化する手法を提案する。この手法は、ある程度の数値読み取り精度を持ち、コンパクトな表示が可能である。本手法は、上で述べた数値ベースの表示と画像による表示の中間的な性格を持つ。

2. 二色塗り分け疑似カラー表示による可視化

2.1. 一次元連続データ

まず、二色塗り分け疑似カラー表示[1]について述べる。ここでは前提として、変数 x を横軸にとったときの一変数関数の値 $f(x)$ を、縦軸 m 画素の带状領域に疑似カラー表示することを考える。疑似カラーとして、 $k+1$ 個の離散色を基本とし、これらの色番号を順に $0, 1, \dots, k$ とする。さらに、色番号 c の示す関数値を、 $Ac+B$ (A, B は定数) とする。関数値 $f(x)$ は、表示したい全ての x に対して、 $B \leq f(x) \leq B+Ak$ の範囲にあるものとする。

いま、上記の数値範囲を A ごとに区切ったうちの i 番目の範囲 $[B+Ai, B+A(i+1))$ に着目する。通常の離散

色による表示では、この範囲の全ての値に対して色 i が用いられる。また連続色による表示では、二つの色 $i, i+1$ を

$$(f(x) - (Ai+B)) : ((A(i+1)+B) - f(x))$$

に内分する中間色が用いられる。

二色塗り分け法では、これを二つの離散色 $i, i+1$ で塗り分ける。すなわち、帯の上部を色 i で、下部を色 $i+1$ で塗る。二色の境界位置は、関数 $f(x)$ の値によって定め、

$$f(x) = B+Ai \quad \text{のとき、帯の下部}$$

$$f(x) = B+A(i+1) \quad \text{のとき、帯の上部}$$

となるように、線形補間によって求める。

図 1 に、グラフ、擬似カラー表示(離散色、連続色)、提案方法の比較の例を示す。

提案手法により可視化された図から、数値を読み取るには次のようにする。帯の任意の位置 x において、まず上下二色の組み合わせを見る。これにより $f(x)$ の値の範囲がわかる。次に、二色の境界位置(上下方向)を読み取る。そこから、より詳細な値がわかる。また、 $f(x)$ の値の変化を見る場合、まず帯の色の变化から、大まかな値の変化は判別できる。値の細かい変化は、二色の境界線の変化として表現されるので、グラフのごとく読み取ることができる。結果として、提案手法では、狭い帯の中でも関数値とその変化を精度よく読み取ることが可能となる。

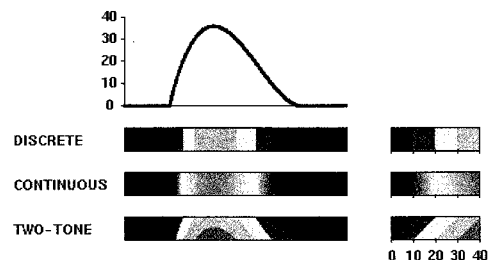


図 1: 各種の疑似カラー表示の比較

2.2. 二次元配列データ

本研究の提案手法では、二次元配列に対して可視化を行う。まず、可視化の基本的手法 [2] について述べる。この場合、隣接する二色の面積比で数値を表現することとなるが、面積比の制御には縦方向のみならず、横方向も利用できる。分かりやすさと数値読み取り精度を両立させるために、次の方法を用いる。

1) 配列要素一つは正方形領域で表現する

2) 正方形を縦に 4 (3~5) 等分し、小ブロックに分ける。

数値が $B+A(i+j/4)$ のとき、上の $(4-j)$ 個の小ブロックを色 j で、下の j 個の小ブロックを色 $j+1$ で塗り分ける

Two-Dimensional Array Visualization for Debugging
Hidefumi WATANABE, Takafumi SAITO
Division of Electronic and Information Engineering,
Graduate School of Technology, Tokyo University of
Agriculture and Technology

[†] Department of Computer, Information and
Communication Sciences, Faculty of Technology, Tokyo
University of Agriculture and Technology

3) 数値が $[B+A(i+j/4), B+A(i+(j+1)/4)]$ に含まれるとき、下から j ブロック目を 2.1 と同様に横方向に塗り分ける。

図 2 に $[0, 255]$ の範囲の数値を四等分し、五色を使って表示したときの例を示す。図 2 の値は数値 94H を可視化したものである。

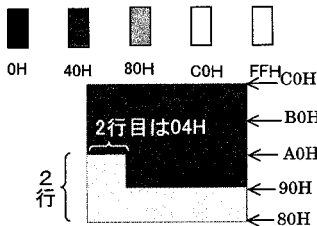


図 2: 濃淡画素の可視化手法

2.3. Log スケールへの対応

次に、データ分布の上限、下限が未知のデータを可視化する手法について述べる。データ分布が未知であるため、可視化可能な数値の範囲を十分大きく取る必要がある。また、疑似カラーに使用する色数が限られるため、塗り分けを線形スケールでなく Log スケールで行う。こうする事で広範囲のデータを可視化することができる。

具体的な可視化手法は次のとおりである。ここでは、底が 10 のときの例(図 3 左)で説明する。Log スケールの表示は、縦方向に 1 段埋まると数値が一定の比を持って等倍されていくように塗る。横方向には線形スケールで塗る。まず、正方形領域を三つに分ける。次に可視化する値を $A \times 10^n$ の形に直し、 A の値によって二色の色を塗り分ける。使用する色は n の値から $(n, n+1)$ と決まる。塗り分けは、 A が 1 以上 2 未満のときは正方形領域の下段を色 $n+1$ で横方向に塗り分ける。残りを色 n で塗りつぶす。2 以上のとき縦方向は下段全部を $n+1$ で埋め、その中段を数値によって横方向に塗る。5 以上では中段まで塗り、上段を横方向に塗る。残りは n で塗る。10 のときは上段まで塗るので、全て $n+1$ で塗りつぶされる。すなわち、 A が 10 のときは、可視化する数値は $1 \times 10^{n+1}$ であるので、 $A=1$ と同様の塗り分け方になる。このようにすれば底を 10 に取った数値の可視化が可能である。底を 16 にしても同様に可視化が可能である(図 3 右)。この場合、正方形領域は縦に 4 分割する。図 4 に、数値 161 を、底を 10 に取ったときと 16 に取ったときの可視化例を示す。

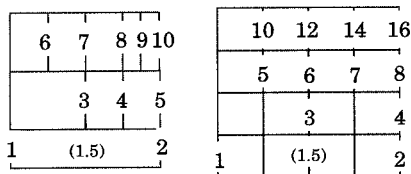


図 3: 底が 10(左)の時のと 16(右)の時の表示方法

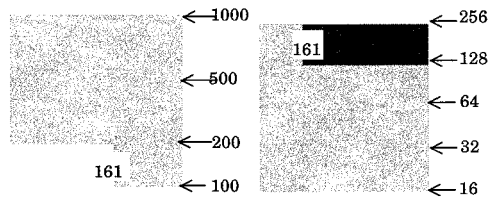


図 4: 161 を可視化した例。左が底 10 右が底 16
それぞれ、行が埋まると右に書かれた数値になる

3. 実験結果

本手法の適用例を図 5 に挙げる。本例は、底を 16 に取った時のある配列のデータの一部分を可視化したものである。図 6 はこのデータを可視化するときを使用した色のパターンである。可視化結果から、このデータが最小 1 以下、最大 8192 以上の値を持っていることが分かる。適宜この小さな正方形の値をコンソールに出力して読むことでさらに細かく値を把握することができる。実際に値を読むと、この部分での最小値は 0、最大値は 10757 であった。

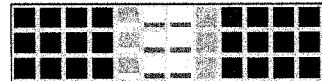


図 5: 可視化結果例

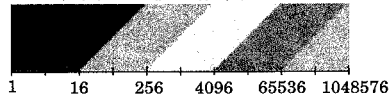


図 6: カラーパターン

また、本手法を簡単な画像処理プログラムに適用したところ、アクセスする配列の添字の書き間違いによるバグを発見できた。画像処理プログラムでは典型的で、出力画像を見ただけでは発見しづらいバグの一つである。また、アルゴリズム上絶対に取りえない値が出てしまった場合のバグも発見できている。

4. おわりに

本研究は、デバッグ手法として二次元配列の可視化手法を提案した。データの可視化には二色塗り分け法を使用した。データ分布が広い場合には Log スケールの表示法を取ることで、データ分布の広い配列にも対応できた。本可視化手法は、データ表示スペースをコンパクトにでき、ある程度の数値読み取り精度を持っている。

本手法を適用することで、すでにいくつかの典型的なバグを発見できた。今後は、本手法を様々なデバッグ作業に役立てる方法を考えていく。

参考文献

- [1] 斎藤隆文, "二色塗り分けによる疑似カラー表示", Visual Computing グラフィクスと CAD 合同シンポジウム 2000 予稿集, pp.119-123, 画像電子学会, 情報処理学会, 2000.
- [2] 渡部秀文, 斎藤隆文, "二色塗り分けによる疑似カラー表示を用いた情報可視化", 夏のプログラミング・シンポジウム「計算機と表現」報告集, pp.17-26, 情報処理学会, 2001.