

色合成を用いた WWW からの数値データの視覚化

小桜 健, 笹倉 万里子

sasakura@momo.it.okayama-u.ac.jp

岡山大学 工学部 情報工学科

概要

一般に、数値データは、実際の現象のある側面をある基準で抜き出したものに過ぎない。特に観測などで得られた数値データの場合、観測したごく少数の地点での数値データは得られるが、その他の大部分を占める観測しない地点での数値データは得られない。しかし、これは観測していないところにはデータがないことを意味するわけではない。一般に二次元のデータの場合、観測していない場所でのデータを推測して見せるひとつの方法として、データをグラフにプロットして近似線を描くという方法がある。同様に三次元のデータについても同じような近似を目で見えるように表現する方法はないだろうか。今回我々は、地図上に円でデータを表示しその色をぼかすことで三次元データの近似表示を試みた。色が重なるところでは色合成して二つ以上のデータの影響を見ることができるようにした。本発表では研究の現状と今後の課題について述べる。

1 はじめに

情報化社会と言われている近年、我々が生活する中で必要とする大抵の情報はインターネット、特に WWW (World Wide Web) を情報源として利用することで得ることができる。WWW には様々な情報が存在し、また多くの人が情報を提供している。しかし、その情報の質はさまざまであり、偏りがあることも多い。このような偏った情報から全体をどのように推測するのは今日的な課題であると考える。

さて、一般に、さまざまな現象を表現するのに数値データがよく使われる。数値データは、実際の現象のある側面をある基準で抜き出したものに過ぎない。特に観測などで得られた数値データの場合、観測したごく少数の地点での数値データは得られるが、その他の大部分を占める観測しない地点での数値データは得られない。しかし、これは観測していないところにはデータがないことを意味するわけではない。一般に二次元のデータの場合、観測していない場所でのデータを推測して見せるひとつの方法として、データをグラフにプロットして近似線を描くという方法がある。同様に三次元のデータについても同じような近似を目で見えるように表現する方法はないだろうか。またこの方法は WWW から得られる偏りのある情報を表現するために応用できないだろうか。

今回我々は、地図上に円でデータを表示しその色をぼかすことで三次元データの近似表示を試みた。WWW の国土交通省「防災情報提供センター」[1]

のウェブページから岡山県の降雨量の数値表から数値データを抽出し、それを視覚化する。今回は岡山県の地図上に数値データを円で表し、数値データの大小を色によって表す。円を描く際に透明度を変化させてぼかし、地図上で重なった円と円の色を合成することで、具体的な数値データがない場所の降雨量に関しても表示を行った。

2 表示データの取得

本研究では、国土交通省の「防災情報提供センター」のウェブサイトにある、1時間ごとの降雨量を数値表にまとめられた数値データを使用する。これらのデータは岡山県の観測所（水文水質観測所）の数値データであり、今回は 25カ所の観測所の数値データを使用する。ここから取得できるデータは、「観測所記号」、「観測所名」、「水系」、「河川名」、「日付」、「1時間ごとの降雨量」である。これらは1ヶ月ごとにまとめられた数値表が一つのウェブページとなっている。今回は、ウェブページのフォーマットがあらかじめわかっているので、それを利用して必要なデータを取得した。

「防災情報提供センター」のウェブサイトでは、1時間ごとの降雨量が1ヶ月ごと一つの観測所につき一つのウェブページにまとめられている。すなわち、25カ所の観測所のデータを取り込もうとすると、1ヶ月分だけで25のウェブページにアクセスする必要がある。そこで、今回は、あらかじめすべての数

値データを取り込む作業をし、その結果をテキストデータとしてまとめ、そのファイルをデータベースとして置いておくという方法をとる。テキストデータを読み込む際にはそのデータベースから読み込むことにし、表示の際の時間短縮をすることにする。

今回は、2002年1月から2006年12月までの5年間のデータを使用したので、データを取り込むために、 $25(\text{カ所}) \times 5(\text{年}) \times 12(\text{ヶ月}) = 1500(\text{回})$ のアクセスを要した。すべての数値データを取り込むために実際にかかった時間は、CPU: Intel(R) Pentium(R) 4 CPU 2.80GHz, メモリ: 約1GB, HDD: 約50GB, 回線速度: 約21Mbpsのマシンで20分程度であった。

3 色合成を用いた円による視覚化

2節でまとめられたテキストデータを用い、色合成を用いた円による視覚化方法を提案する。今回は岡山県の地図 [2] 上に降雨量のデータを表示する。

3.1 円を用いた表現

本研究では円を使って岡山県の降雨量を表現する。すなわち、岡山県内25カ所ある観測所それぞれを中心とした円を描き、色によってその地域の降雨量を表す。円の大きさはすべての観測所で同じとする。人は色を見ただけでもその色に対するイメージ [3] を持っているもので、これを利用しさらに情報を理解し易い形にする。

今回は「雨」のイメージもあり「安全」のイメージもある「青」を降雨量の少ないときに当てはめる。また降雨量が多いときは「危険」や「怒り」という危ないイメージのある「赤」が当てはめる。「赤」の色が強ければ降雨量が多く、「青」の色が強ければ降雨量は少なく、中間であれば「赤」と「青」の合成で「紫」となる。降雨量の大小の基準 [4] を、表1で示す。

3.2 透明度の変化による信頼度の表現

本研究では降雨量を扱っているが、降雨量は実際は観測所で観測された値であり、それ以外の場所の降雨量は厳密に言えばわからない。しかし、実際には観測所のデータを元に我々は観測所でない場所の降雨量を以下のように推測している。

表 1: 降雨量の基準

日数	大小	降雨量	備考
1日	大	70mm	大雨注意報が発表される
	小	0mm	晴れの日
30日	大	300mm	梅雨で降雨量の多い地域
	小	30mm	雨の少ない季節で少ない地域
365日	大	4000mm	年間降雨量が多い地域
	小	1000mm	年間降雨量が少ない地域

- 一般的には観測所の近くでは観測所と同じように雨が降っていると考えてよい。

- 観測所から遠く離れるほど観測所とは雨の降り方が違う可能性が高くなると推測される。

そこで本研究ではデータに対し「信頼度」という概念を導入することにより、普段我々が暗黙のうちに行っている推測を明確に視覚化する。

信頼度は、次のように定義する。

- 観測所ではデータの信頼度はもっとも高い。

- 観測所から離れるほど信頼度は低くなる。

- 複数の観測所のデータを参考にできる場合は一つの観測所のデータから推測するより信頼度は高くなる。

そして、信頼度を円の中心すなわち観測所の位置から遠ざかるにつれて、円の透明度を高くする方法で表現する。ここでいう透明度とは色付きの円を描く際に、その色の「透明さ」を表すもので、高ければ透明に近く、低ければ不透明に近くなる。

その透明度の変化を用いることで各観測所を中心にして円が外へ徐々に広がっているように見せると共に、各観測所の数値データは、その観測所に近い地域では観測所と同じようなデータが表れ、観測所から遠い地域になるにつれて、その観測所の数値データとは誤差が大きくなり、信頼度が低くなることを表現する。

この地図上に近くの観測所のデータを元にした降雨量を円で表すということで、正確ではないが観測所から遠い地域の降雨量を推測することが可能となる。

3.3 点による円の表示

円を点で表示する方法として提案するのは、点つまり1ドットごとに色と透明度を考え、その点の集

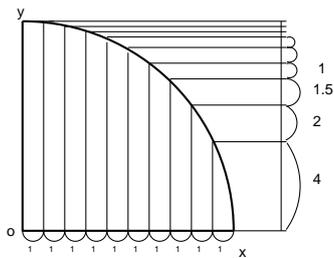


図 1: X の変化量と Y の変化量の差

合によって円が中心から外側に広がっているように見える図に近づける方法である．それには次の手順で行う．

1. 表示させたい円の半径を決める．
決めた半径を最大半径 (Max) とし，半径を 1 から Max まで順に変化させながら次の作業を行う．
2. ある半径 $R (1 \leq R \leq \text{Max})$ のときを考える．
 x 座標の値 (X) を 1 から $(R \times 1/\sqrt{2})$ まで順に変化させ，そのとき半径 R の円になるための y 座標の値 (Y) を求める．求める y 座標の値は次の (1) 式で求まる．

$$Y = \sqrt{R^2 - X^2} \quad (1)$$

このとき解は 2 つ求まるが正の数を扱う．

$X = (R \times 1/\sqrt{2})$ まで変化させると，そのときの y 座標は $Y = (R \times 1/\sqrt{2})$ であり，次に y 座標の値 (Y) を $(R \times 1/\sqrt{2})$ から 0 まで順に変化させる．今度は半径 R の円になるための x 座標の値 (X) を求める．求める x 座標の値は次の (2) 式で求まる．

$$X = \sqrt{R^2 - Y^2} \quad (2)$$

このときも同様に解が 2 つ求まるが正の数を扱う．

また上記の作業で変化させるものを途中で X から Y へ変更した理由は，変更せずにすべて X だけの变化で Y を求めていくと，X が半径の最大 Max に近づくとつれて，X の変化量と Y の変化量の差が大きくなるからである．図 1 に示す．

3. 透明度を決定する．

2 の作業を全ての $R (1 \leq R \leq \text{Max})$ について行い，そのデータを半径の大きさごとにまとめる．半径の大きさにより透明度を決める．半径が大きくなるにつれて，透明度は高くなる．

4. 完全な円にする．

1~3 の作業では図 1 のように円の 1/4 の点しか求まっていない状態であるので，これらを完全な円にするためにいま求まっている点を x 軸， y 軸をそれぞれ対称軸として対称な点を求め，中心 o を中心に 180 度回転させた点を求める．求まっている点を (x_1, y_1) とし，求める点をそれぞれ (x_2, y_2) ， (x_3, y_3) ， (x_4, y_4) とすると以下の式で求まる．

$$x_2 = x_1 \quad y_2 = (-1) \times y_1 \quad (3)$$

$$x_3 = (-1) \times x_1 \quad y_3 = y_1 \quad (4)$$

$$x_4 = (-1) \times x_1 \quad y_4 = (-1) \times y_1 \quad (5)$$

これらをすべての点に対して行うことで，円のすべての点が求まる．

5. 観測所ごとに表示する．

1~4 の作業により中心が $(0, 0)$ のとき，点によって円を描くためのすべての点が求まる．これを各観測所ごとに対応させ，各点を観測所の座標に合わせて平行移動させることで表示が可能である．

6. 円の色を決定する．

色は各観測所の降雨量に基づいて決定する．

3.4 色の合成

表示させる円は点により構成されているものであるので，1 つ 1 つは独立しており各点で色や透明度を変化させることが可能となる．

重なり合っている点では，先に書き込まれた色情報と透明度がある．色の合成については，先に書き込まれている色情報と後から書き込もうとしている色情報を，それぞれの透明度を使い合成する．透明度はそもそも観測所の数値データとの誤差が，観測所から遠ければ遠いほどあり，信頼度が低いと考え設定したものである．よって観測所に近いものほど透明度が低く，その数値データは信頼度が高いとい

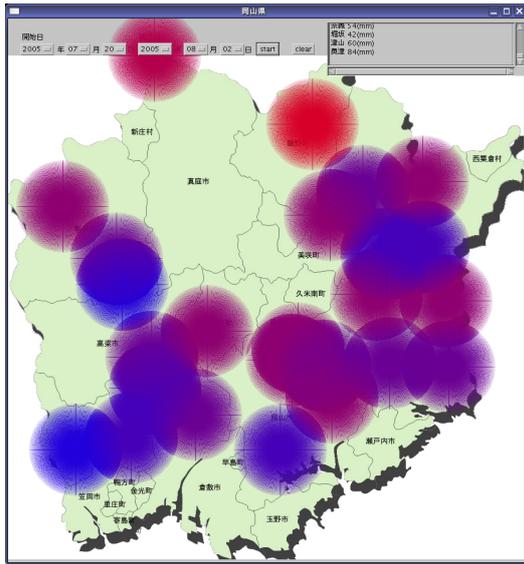


図 2: 色合成を用いた図

える。これにより色の合成をする際、数値データにより色は決定しているの、信頼度が高いものすなわち透明度が低いものと、信頼度が低いものすなわち透明度が高いものとを同等に扱うべきではない。それを考慮し、色の合成をした後の数値データを求める方法として次のことをした。

重なり合った 2 つの点のそれぞれの透明度を a_1 , a_2 とし、色すなわち数値データを d_1 , d_2 とし、求める数値データを d_0 とすると次の式で表せる。

$$d_0 = \left(d_1 \times \frac{a_2}{a_1 + a_2} \right) + \left(d_2 \times \frac{a_1}{a_1 + a_2} \right) \quad (6)$$

上の式は色の合成を行う際に透明度が低いものはその数値データを多く反映され、透明度が高いものはその数値データはあまり反映されないことを示している。

透明度については、重なっている点というのは 2 つの観測所の数値データを元に計算された数値データが求まるということなので、信頼度は 1 つの観測所だけのものより高くなっていることから、透明度は低くなると考える。よって 2 つの透明度を足し合わせ、透明度の最大の値（今回は 255）から引き、その値をその点の透明度とした。これにより重なった部分の透明度は低くなる。

上で提案した方法生成した図を図 2 に示す。色合成を用いた図は、重なった部分の色が合成され重なった部分の色が変化していることが分かる。透明度を見ても重なった部分は低くなっていることが分かる。

4 おわりに

本研究では現在、情報取得のために良く使われている情報源である WWW から数値データを抽出し、それを透明度の変化を用いた円で表し、円が重なる部分は色合成を用いて表す方法で視覚化を行った。今回の視覚化では円の中心、つまり観測所から遠ざかるにつれて透明度を上げることで、観測所から円が外に広がっているように見せると共に透明度の高いところに曖昧性を持たせることで、ユーザは観測所がない地域でも降雨量の推測が可能となった。

また円を点で表すことにより、点 1 つ 1 つで色の設定、透明度の設定が可能となり、円が重なった部分の色合成を可能にした。その結果、円と円の重なった部分すなわち 2 つ以上の観測所が近くにある地域では、2 つ以上の観測所のデータを元に降雨量を計算することができ、それに応じた色を設定することができた。よってその地域の降雨量データの信頼度を上げることができたということになる。これは観測所がない地域でも信頼度の高い降雨量データを得ることができたということになり、視覚化により信頼度の高い推測ができたと考えられる。

この方法は、基本的にデータが空間的に連続していると仮定できるものであれば適用可能であると考えられる。今回の方法では、色の合成はできたが、表示の形が円に固定されていること、円の大きさや色の使い方にまだ工夫の余地があることなどが課題として残されている。今後は他のデータを用いて改良を行っていきたい。

参考文献

- [1] 国土交通省「防災情報提供センター」,
<http://www.bosaijoho.go.jp/index.html>
- [2] “WD CD-ROM FREE DATA 2006”, 毎日コミュニケーションズ株式会社, 平成 18 年 4 月発行通巻 58 号付録。
- [3] 色のイメージと感情効果,
<http://www.aichi-iic.or.jp/co/sunlight/kanjyou.html>
- [4] お天気プラザ: <http://tenki.wet.co.jp/>