

Web 上でのグラフ表現

山本 義郎

多摩大学 経営情報学部

E-mail: yama@tama.ac.jp

1 はじめに

インターネットの普及により、多くの統計情報が身近に扱えるようになるとともに、様々な統計サービスがインターネットを介して利用可能となっている。WWW(World Wide Web) で利用できる統計サービスについては、WWW が普及しはじめた初期は単なる情報の提供がほとんどであったが、マルチメディア、インターネットテクノロジーの発達・普及に伴いさまざまなアプリケーションによる対話的なサービスやマルチメディアを利用したサービスも徐々にではあるが公開されるようになってきている。そのようなマルチメディアの要件を満たすアプリケーションは統計解析システムにおいても有効に働くものと考えられ、Mitag(2000) で行なわれているようなマルチメディアを統計教育に活用する取り組みが国内外において重要視され実際に利用され始めている。また、Web 上での対話的な解析を可能にするための方法も、CGI による方法や Java アプレットによる方法だけでなく、クライアントサーバ型のアプリケーションを構築した XploRe (Härdle et al., 1999) などのアプリケーションも見られるようになってきている。本論文は、ブラウザのプラグインを利用することにより対話的な機能をもつ統計グラフを Web 上に表現できる方法を紹介する。その実装としては、3D モデリング可能な VRML(Virtual Reality Modeling Language) という言語を利用して、その特徴を効果的に利用できる統計グラフを考察する。実際に実用例を示すことにより、このような新たなテクノロジーを利用することによる統計グラフ表現の可能性について考察を与える。

2 インターネットテクノロジーを活用した統計グラフ表現

ホームページを利用して統計情報を提供したり、統計処理を行うサイトが一般的になっているばかりでなく、多くのソフトウェア同様に、統計解析ソフトも Web ブラウザから利用できるようになってきている。しかしながら、統計グラフィックスにおいては、一般の統計ソフトウェアでは非常に多彩なグラフ表現が可能であり、インタラクティブにグラフ表現を変更することが可能であるにもかかわらず、Web 上で見られるグラフ表現にはそのような高性能なものはあまり見られない。

これは、HTML の仕様起因するものである。すなわち、Web ブラウザにおいて表示できるイメージは GIF, JPEG といった静止画のフォーマットでなければならない (GIF はアニメーションも可能ではあるが) という制約のためであり、通常は Web 上のグラフ表示はこれらのフォーマットの静止画で表示されている。山本・垂水 (1998) にあるように CGI を利用して、統計解析に対応するグラフ表現を与える場合にも表示されたグラフは GIF などの静止画であり、インタラクティブに扱うことは不可能である。その対処方法として、Java を用いてダイナミックグラフィックスを実現するものも増えてはいるが、高度なプログラミングテクニックが必要となるため簡単には利用できないのが現状である。

Web 上でのマルチメディアを実現するための方法としては、Macromedia Flash などのようにプラグインを利用して多彩な表現力を実現する方法も一般的となり、Web 上でもさまざまな方法によりインタラクティブな表現が可能となってきた。そのようなテクノロジーとして、VRML や Alice などの 3D モデリン

グ言語があり、この言語を利用すると容易に 3D グラフィックスを構成でき、プラグインを利用してインタラクティブに操作することが可能となる。本論文では、VRML の特徴やそれにより実現されるグラフ表現を紹介することにより、この種のテクノロジーを取り入れることによりもたらされる Web 上の統計グラフ表現の可能性について考察を与える。

2.1 VRML とは

VRML とは 3 次元グラフィックスをインターネットにおいて利用可能にするため、転送するデータ量が少なくすむように考案されたモデリング言語である。HTML 同様にテキストベースで記述し、VRML ブラウザや Web ブラウザのプラグインにより 3 次元仮想空間を実現するものである。また、多くの VRML ブラウザでは単に 3 次元の空間を構成するだけでなく、視点を変更したり、その空間内を自由に動き回ることが可能である。

VRML のソースファイルはテキストであり、そのファイルを VRML プラグインが導入済みのブラウザ（ここでは Cosmo Player プラグインを導入済みのインターネットエクスプローラ）で開くと仮想空間が実現される。実現された空間内では、ブラウザの下に現れているコックピットのような部分で動作を設定し、マウストラッグにより設定された動きで移動・回転できる。

VRML では配置する物体は大きさや材質（色や透明度など）が指定でき、表面にテクスチャーを貼ることもできる。また物体に対してリンク（HTML の<A>タグの HREF パラメータによる指定に相当）を指定でき、説明（ALT パラメータでの指定に相当）を記述することが可能である。VRML2.0 では動きやサウンドへの対応も行われ、アニメーションを設定することも可能であり、大きなファイルに対する zip 圧縮もサポートされネットワークを経由した利用も考慮されている。VRML のオーサリングツールも存在するが、VRML の作成は HTML 同様にテキストエディタで可能であり、確認のための VRML ブラウザがあればよい。そのため、誰でも手軽に作成・利用でき、作成された VRML ファイルは各種 OS において VRML ブラウザやプラグインのもとで利用できる。

2.2 VRML による統計グラフの構成

前述の特徴をもつ VRML により統計グラフを作成する利点として、(i)3 次元空間のモデリングのみ行えばよく、空間内の移動等の処理をプログラミングする必要がない、(ii) ブラウザの利用によりインターネットにおいて利用可能、(iii) プラットフォームに依存しない、などがあげられる。更に、VRML の特徴である物体の透明度の指定が可能であることやタイムセンサーによるアニメーションやタッチセンサーによるインタラクティブな効果が実現可能である。

VRML の記述は、テキスト形式であることから、VRML により表現されたグラフがあれば、そのソースファイルのデータ部分を編集するだけで、同様のグラフ表現を他のデータに対して行うことが可能である。このことから、プログラムの出力結果として VRML 形式のグラフを記述することは、ある程度のテンプレートがあれば容易に実現できる。

VRML ではユーザ定義ノードを利用してプロトタイプを作成できるため、同様の形状の物体に関してはデフォルトの性質を定め、変更点のみを記述することにより複製が可能である。統計データにおけるデータ値の表示など、位置や色のみを変更したい場合には、プロトタイプの利用によりコードの記述が簡単となる。プログラムにより作成する際にも、利用しやすいことは容易にわかる。散布図のようなグラフの場合には特にプロトタイプが有効である。

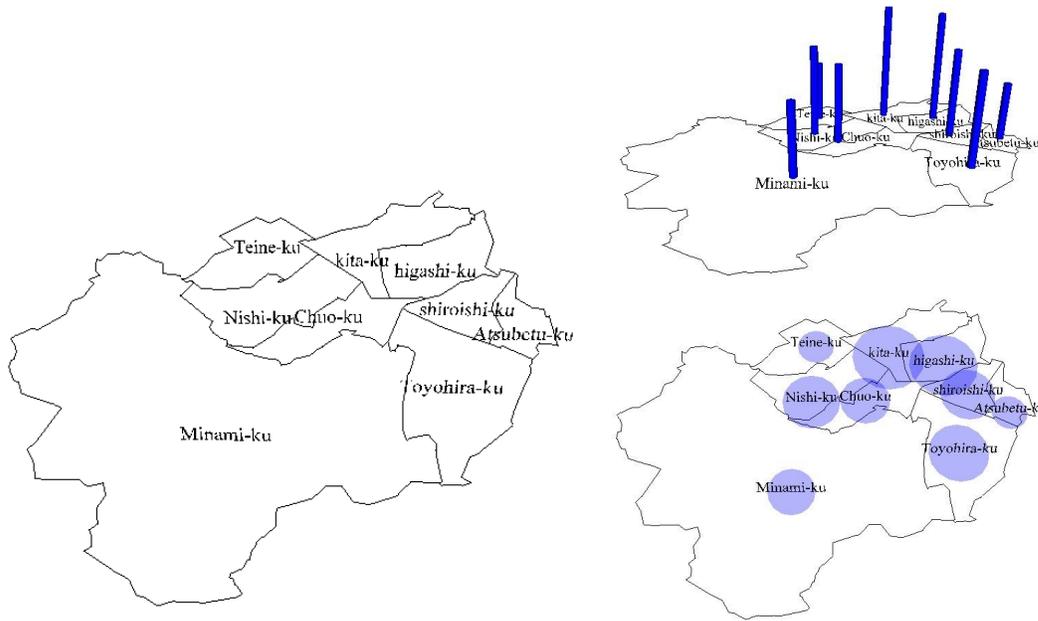


図 1: VRML による地図 (左) を使った 2 種類の地図グラフ (札幌市の区民人口)

これらの特徴を活用して、再利用の利便性を考慮し、プロトタイプを利用して構成した統計グラフのテンプレートをホームページに公開している (<http://dataminig.timis.ac.jp/~yama/vrml/>)。それらは VRML グラフ作成の補助になると考えられたため、ソースコードも公開している。現在のところ 3D 散布図と 3D 棒グラフ、3D 折れ線グラフの簡単なテンプレートと地図グラフ作成のヒントとなる地図や地図グラフのテンプレートを与えている。更に高度なプロトタイプを作成する場合には田中・小林 (1999) などを参考にするとよいが、基本的なものはこのホームページに準備した部品により構成が可能である。

3 VRML 統計グラフの実用例

前節により VRML を使って対話的な機能をもつ Web 上で利用可能な統計グラフが実現できることがわかっただろう。本節では、VRML により実現することにより効果的に利用できると思われる 2 つのグラフ表現と、Web 上で 3D 散布図を作成するアプリケーションについて紹介する。

3.1 VRML による地図グラフ

VRML により実現される地図グラフは、3D グラフ同様に効果的に活用できると考えられる。というのも、VRML による地図グラフでは面倒な移動などのプログラムをすることなく、自由に興味のある地点に移動することが可能であるからである。また、グラフ要素 (棒やデータ点など) にリンクを張り、グラフに関する説明とリンクすることが可能である。

VRML はイメージファイル (GIF, JPEG, PNG 形式が利用可能) を利用できるため、地図のイメージファイル上に VRML オブジェクトを配置し地図グラフを実現することもできる (この例では地図イメージは GIF

形式である)。この方法は非常に簡単であるが、現状ではイメージファイルを利用した場合、移動の際の書き換えに多くのメモリを費やし CPU 性能を要求する。これらは PC の高性能化により解決されていく問題であると考えられる。

別の地図グラフの表示法として、VRML で地図を作成してグラフを構成することも当然可能である。その場合、数値地図をもとに VRML の地図ファイルを作成し、それを 1 つの VRML オブジェクトとして保存しておくといふ。というのも、VRML は他の VRML オブジェクトをファイルとして読み込めるため、そこにグラフ要素を追加的に配置することにより利便性が向上する。図 1 の右の 2 つの図は札幌市の地図を作成した VRML 地図ファイル(左図: sapporo.wrl)を読み込み作成したものである。

これらの地図グラフは、近年の情報公開の名の下に公官庁などがホームページ上に公開するに至った地域に関する情報(多くは Excel ファイルで公開されている)を地図上にグラフ化する際に効果的である。

3.2 その他のグラフ表示例

データマイニング手法のうち、POS データの同時購買の傾向を解析する手法は、マーケットバスケット分析と呼ばれている。この手法では商品の同時購買のされやすさが相関ルール (association rule) として得られる。一般に相関ルールは $X \Rightarrow Y$ という形式で構成されており、ルールの前件部をルールヘッド、後件部をルールボディと呼ぶ。相関ルールを抽出する方法は、項目の可能なすべての組合せを作ることが基本になる。そして、解析者が設定したレベルにより抽出された結果から意味のあるルールを確認する。ルールを絞る際に、解析者が設定するパラメータとして、支持度 (support) と信頼度 (confidence) がある。支持度は、トランザクション T のデータベース全体の中から検出される相関ルール ($X \Rightarrow Y$) の出現度を示し、信頼度は相関ルールを含むようなトランザクション内でルールがどれだけ当てはまるかを示す尺度である。以上の 2 つのパラメータが相関ルールの評価基準として使用される。

相関ルールの視覚的表現として代表的な IBM アルマデン研究所の Quest というシステムでは、ルールヘッドとルールボディの項目番号を (x, y) 座標とし、すべての x, y の組に対して棒の色と高さでそれぞれ支持度と信頼度を表現する方法が用いられている。このような表現は、SGI などの高性能なハードウェアを用いて実現されているが、同様の表現が VRML により可能である。

図 2(左)では、最低信頼度を 0.55 としてそれを示す半透明の面をグラフに与えている。マウスポインタを合わせた棒が $53 \Rightarrow 8$ というルールで、そのルールのその支持率が 0.522、信頼度が 0.705882 であることがブラウザ左下のステータスバーに表示されるよう工夫している。また、回転して上から眺めることにより(右図)、数量化 III 類に基づく配置により信頼度が高いもの(長い棒で示されている)が左下から右上への対角線にいくつかのクラスターとしてまとまっていることが確認できる。透明な物体が配置できると、物体に説明を記述できる性質を利用している点が VRML を利用したメリットとなっている。また、このグラフは、VRML ファイルであるため VRML プラグインを導入したブラウザ上で自由に利用できる。

3.3 VRML 統計グラフ作成アプリケーション

グラフの作成方法としては、所有しているデータに対して簡単に VRML のグラフを描画できるアプリケーションがあることが望ましいため、その実現方法の一例として、CGI を利用してホームページにおいて、フォームへグラフのタイプやデータを指定することにより VRML グラフを表示する Web アプリケーションを構築した。その結果として得られる VRML ファイルはテキストファイルであることから、当然のことながら VRML ファイルとして保存することにより別途利用も可能である。

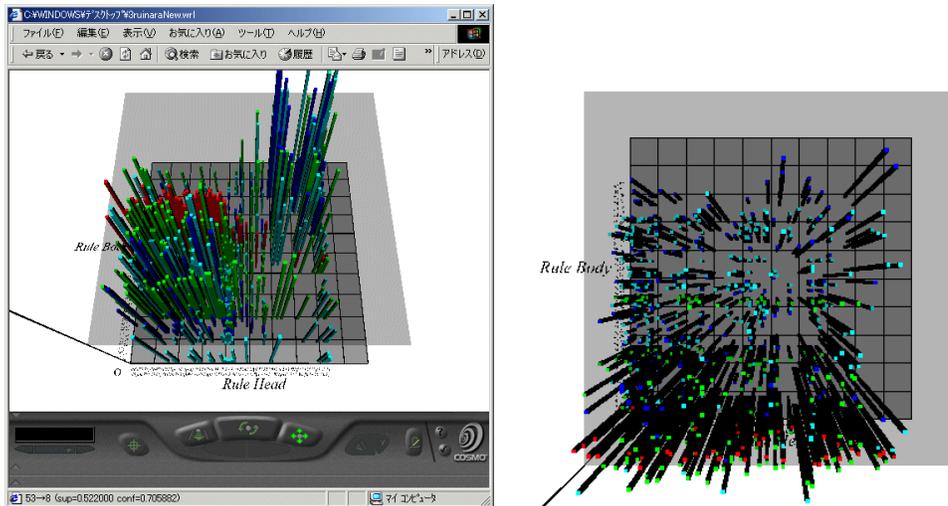


図 2: 数量化 III 類にもとづく相関ルールの表示

またテンプレートを利用することにより、S などのプログラム可能な統計ソフトウェアにおいてもテンプレートを利用することにより容易に関数として作成できる。S と XploRe で作成した関数を紹介する。

4 VRML の問題点と今後の統計グラフ表現

VRML において、より対話的な処理を行う (選択したデータ点の色を変更するなど) ためには Java を利用することにより対処できる。単純なプログラムはスクリプトノードにおいて Java スクリプトで記述でき、複雑な処理や複雑なコミュニケーションを実装するためには Java クラスファイルを読み込むことにより実装できる。VRML ではファイル入出力などの処理も行えないため、データファイルの読み込みなどを行うためには Java を利用することとなる。しかし、高度な対話的処理を伴う処理を必要とする場合には Java だけでアプリケーションを構成した方がよいだろう。したがって、VRML により構成するグラフは静的表示では不満だが、高度な対話的処理を必要としない 3 節に与えたようなグラフが適しているといえる。

VRML の現状としては、VRML ブラウザの主流は Cosmo Player と Community Place であるが、どちらも Windows プラットフォーム用のプラグインがあるだけである。Flash などに比べ、VRML そのものがいまひとつ普及していない点が一番の問題点であるといえよう。実際、インターネットエクスプローラには IE4.0 までは VRML2.0 ビューアがオプションで用意されていたが、それ以降のバージョン用には用意されていない。

イメージデータの描画性能やデータサイズが計算機性能に依存することも現状では問題点となっている。地図グラフを VRML で作成する際にも、詳細な地図の場合には非常に大きな VRML ファイルとなり描画速度が低下してしまうため、少々粗い数値地図をもとに地図を構成することが好ましい。しかしながら、この速度の問題は計算機の性能が日々向上していることから考えるとそれほど大きな問題ではない。実際、数年前には描画に長い時間を要したのも最新の PC では難なく描画できるようになっている。

本論文中で紹介した VRML で実現可能なグラフ表現は、近年 VRML 同様に仮想空間をモデリングするために作られた Alice などの言語や環境において、同様の手法により構築可能であることは想像に難くない。他のマルチメディアを利用可能なプラグインなどにより、新しいグラフ表示法が利用できると思われる。し

たがって、状況に応じて利用できる方法を使い分け、インターネットの普及により始まった新たなインターネットテクノロジーをいかに効率よく利用していくかが今後のグラフ表現の幅を広げるカギになると考えられる。

参考文献

- [1] Härdle, W., Slinke, S. and Müller, M. (1999). *XploRe Learning Guide*, Springer. Germany.
- [2] Mittag, H.-J. (2000). Learning and teaching statistics in a networked world, 第 68 回日本統計学会講演報告集, 269–269.
- [3] Rickk, C. and Bell, G. (1997). *The Annotated VRML 2.0 Reference Manual*. Addison-Wesley. America.
- [4] 田中成典・小林孝史編 (1999). Web 工房シリーズ VRML の達人, 森北出版. 東京.
- [5] 松田晃一・宮下健二 (1999). JAVA+VRML Java と VRML2.0 で作るインタラクティブ 3D ワールド, プレンティスホールジャパン, 東京.
- [6] 山本義郎・垂水共之 (1998). Web 上の統計解析システムの構築 - CGI による統計処理とグラフ描画の実装 -, 計算機統計学, 11(1), 45–50.
- [7] 米森 力・山本義郎・佐藤義治 (2000). データマイニングにおける相関ルールの視覚的表示, 日本計算機統計学会第 14 回シンポジウム論文集, 2–5.

参考 URL

- Macromedia Flash (<http://www.macromedia.com/jp/software/flash/>)
- VRML (<http://www.vrml.org/>)
- Alice (<http://www.alice.org/>)
- Cosmo Player (<http://www.cai.com/cosmo/>)
- Cosmo Player (<http://www.cai.com/cosmo/>)
- Community Place (<http://www.sony.co.jp/Products/CommunityPlace/>)