

野外調査のためのワークベンチ

北川文夫

岡山理科大学

kitagawa@mis.ous.ac.jp

概要

本稿では、岡山理科大学総合情報学部複数の学科による「岡山学研究会」における研究を支援するシステムについて報告する。岡山学研究会では、岡山を共通テーマに考古学、植物学、地理学などの研究者が野外調査を元に、これらいくつかの研究を相互に利用した研究をしていく目的で発足した。このような研究をするには、調査地点の個々のデータと共に位置情報が重要な役割を果たす。従って、これら異なる野外調査データを同時に検索し、地図上の位置や領域と共に表示できるシステムが相互の関連性を調べる上で有効と考えられる。本稿では、まず、システムの要求分析をし、野外調査データの種類を示し、それらに共通なデータモデルの表現を提案する。次に、このデータ表現を用いてデータの相互関係を視覚的に発見するための支援システムを提案する。

Workbench for field research

Fumio KITAGAWA

Okayama University of Science

abstract

In this paper, we propose an assistant system for field research, especially "Okayamaology". Okayamaology is one of a field research on Okayama prefecture with archaeological, biological and geographical point of view. For such a research, each location data is so important as well as each field data. Therefore, it is usefull to develop an assinstant system in which we can retrieve information from some cross data of hetelogeneous field and can view the result on a map. In this paper, we analyse some requirements of the system. Next, we propose a data model which can present various field data. At last, we propose an assinant system to discover some interaction of each field data by seeing the result of this sysytem.

1 はじめに

岡山学研究会 (Okayamaology) は、岡山理科大学総合情報学部の複数の学科のメンバーによる共同研究グループであり、1999年3月に発足した。岡山学研究会は、岡山地域での特色ある歴史、文化、自然を総合して研究することで、地域的な特徴の原因を探ることを目的とする。つまり、岡山を共通テーマに、考古学、植物学、地理学などの研究者が野外調査を元に、これらいくつかの研究を相互に利用した研究を行う。この研究の特徴は、あるテーマに対して、このような異なるいくつかの分野の野外調査を元に研究することにある。例えば、気象と植生の関係や地理的な特徴と遺跡の関係などである。

このような研究を進めるには、それぞれの野外調査の個々のデータと共に位置情報が重要な役割を果たす。また、調査データも考古学、植物学、地理学では、その性質が異なる。従って、これら異なる野外調査データを共通の地図の上に表示できることが必要である。

地図を用いた情報システムはGISと呼ばれ、多種発表/研究されている。これらの多くのシステムでは地図上の全てのオブジェクトをベクトルデータで表現することが多い。これは地図拡大縮小が自由であるとか、全てのオブジェクトが検索対象になるなどの利点はある。しかし、われわれが必要としている地図データとしては、データ入力にコストがかかり、また野外調査データの性質上それほど重要でないと思われるもの(人口構造物など)までがオブジェクトになっていると捉えられる。したがって、ここで提案するシステムは、野外調査で調査者が用いるいつもの地図の上で位置指定ができ、同時に調査データが入力でき、検索条件を複数の入力フォームの組み合わせで表すなどの特徴がある。

本稿では、野外調査で得られたデータをその調査地点・区域のデータと共に登録でき、登録したデータを色々な条件で検索し地図上に表示するシステムの提案をする。まず、個々のデータの特徴を調べ、それらが登録できるデータモデルを提案する。次に、調査データの入力システムを検討し、最後に検索システムの要件を検討する。なお、このシステムは一般的な野外調査データの登録と検索にも利用できるものを目指して開発する予定である。

2 要求分析

2.1 データの分類

まず、どのようなデータを扱い、どのような結果が欲しいかを検討した。例えば次のような問合せを試みたいという要求があった。

- 「カヤ系の須恵器を出土した5世紀前期の遺跡の分布」
- 「地図中の任意の3つの神社が特定の角度をなしているものを探す」
- 「特定の植物の分布と地質/気候との関係を見たい」
- 「備前焼の窯の位置と周辺の森林の関連を見たい」

上記の検索に更に地質や気象、標高などに関連して調べたい要求もある。これらは、最初の漠然とした要求であるので、より要求を明確にするために、まず野外調査データやその他に利用できるデータにどのようなものがあるか検討した。その結果、目的を実現するために必要で入手可能なデータには、(1) 野外調査データと(2) 公的機関などによる広域の調査データを加工したデータの2つがあることが分かった。これらのデータは、単に自前で野外調査を行ったかが異なるだけではなく、データの構造や意味的な違いがある。

(1) 野外調査データ

文字通り野外調査したデータであり、考古学調査では遺跡の地図上での位置、標高、形、方角、時代(年代)といった遺跡自体のデータに加え、出土物ごとに、形、材質、形質、年代、用途、個数というデータがある。また、植生調査では、調査エリア(通常10m×10m)の地図上での位置、調査年、調査者、区分(赤松林、草原など)のエリア自体のデータに対して、そのエリア内に生えている植物の占有面積がそれぞれの植物ごとにデータとして発生する。また、それぞれの調査ともに写真も重要なデータになっている。

ここに示したとおり、野外調査データの構造は、ベースとなる遺跡や土地の情報に対して、たくさんデータが付属しているツリー構造をしている。また、個々のデータの種類も数値、文字、年代、写真などたくさんある。データの例を図1に示す。

位置情報は、点、線、領域(ポリゴンで表現)の3種で十分と思われる。

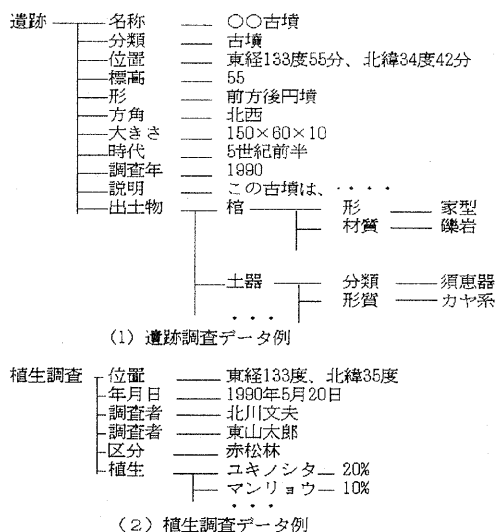


図1 調査データ例の木による表現

(2) 広域調査データ

利用できる広域調査データには、アメダスによる降水量、日照時間、年間／月間平均気温、国土地理院による標高などのメッシュデータ、更に土地利用図から得られる領域データがある。

これらのデータは、コンピュータ可読になっているものもあるが、多くは印刷された数字の表か、地図上への表現になっているので、加工して入力しないおさなければならぬ。¹ 加工の際にメッシュとすることで他のデータと同様の操作が可能になる。

広域調査データは、基本的にはメッシュ型データであり、地理的に等間隔の地点のデータを連続した数字列として表現したものである。また、メッシュの幅は1km,500m,50m,10m など様々である。

2.2 地図とデータ入力

上記のデータのうち(1)の位置情報の精度は、これらの調査を行う際に利用する国土地理院の1:25000地形図上で示せばよい精度になる。この地図では1mmが25mであるから、その程度の誤差は許容する。

調査データのデータ入力(デジタル化)は、可能ならば調査地点ごとにGPSの位置情報と調査データを逐一入力できれば効率的と考えられる。しかし、そのようなシステムが無かったり、現地では調査が優先してPCなどが利用しにくいことや、また既に調

¹ 土地利用図は10mメッシュのデジタルデータが大都市圏には発行されている。

査してある大量のデータの位置情報は紙に印刷された地図の上に示されているということもある。このようなことを考慮すると、野外調査のデータ入力をする際に、次の点が要求される。

- 入手が容易ということや、普段利用しているということから紙に印刷された地図をデータ入力時に利用したい。

一方、メッシュデータは最初から位置情報を含んでいるので、地図とは無関係に入力可能である。ただし、メッシュデータのなかでもメッシュのサイズや全体の領域が異なるので、全体の領域と個々のメッシュの領域を地図から対応できるようにしておく必要がある。

2.3 情報検索と演算

前述の2種類のデータのいくつかを組み合わせた情報検索を行うことが必要である。

(1) 位置データと位置データ

個々の位置情報を持った調査データではそれら同士の位置的な演算が必要と必要と思われる。ここで位置演算は、その結果の精度が高くなってもよいので、一般的なGISの持つもの又はそのサブセットで良いと考えられる。しかし、先の検索質問例にもあるように3点のなす角度がある値のものを求めたいという特殊な演算の要求もある。

(2) 位置データとメッシュデータ

次に、これら調査データとメッシュデータとの演算を行う必要がある。基本的にはメッシュデータ同士だけの演算は無いので、調査データとメッシュデータの演算が必要になる。先の問合せ例では特定の植物と降水量などがこれにあたる。

検索結果の表示

検索結果は地図上に表示されることが望まれるが、表示にはデータ入力に用いた精度が必要ない場合が多く、より小縮尺な地図で広域を見られることが要求される。また、いくつかの条件の違いを地図上で色や形を変えて表示することも必要である。

3 データモデルと演算

3.1 データ入力用地図と位置データ

野外調査での調査地点が印刷された地図上に示されている状況を考慮し、印刷地図をスキャナーで入力した画像を位置入力の原因にする。例えば1:25000

の地形図は日本全体をカバーするメッシュになっているので、その管理番号を地図IDにしておけば、小縮尺の全体図との対応をとることができる。この地図の周辺の緯度経度をスキャンした地図の属性情報にしておくことで、地図中の任意の位置の緯度経度は計算で求められる。実際には、地形図は岡山付近で横が約46cm、縦が約37cmと大きいので1/4の領域ごとに地図をスキャンしサブIDを付けることにする。また、国土地理院の地形図は正確には台形をしている。しかし、例えば「岡山北部」では、上辺と下辺の差が0.5mmというふうに小さいので、誤差の範囲と考えられる。

ここで、緯度経度の基準が近く大きく変わることも考えておかないといけない。それは、国土地理院の「測地成果2000」[5]に示されるように、日本の地図の緯度経度の基準を明治以来の日本測地系から世界測地系に変えることが予告されているからである。この変更はGPSにより得られる緯度経度の値が国内で利用されている地図と異なるので、それを解消するために行うのが主な目的ということである。これによる同一地点間のずれは全国で400mから450mにもなるので、配布が予告されている変換プログラムで変換することが必要なことも考慮する必要がある。

経度、緯度のデータは度分秒で表現するが、位置の演算などに都合の良いように内部表現は浮動小数点を用いる。値の精度は「岡山北」の地形図では次に示すように、それぞれ1秒あたり約1mmに相当することから、1/10秒までで十分と考えられる。

緯度差 5' = 36.96cm(岡山北部地形図, 右辺)
 ⇒ 1" = 1.23mm
 経度差 7'30" = 45.81cm(岡山北部地形図, 底辺)
 ⇒ 1" = 1.02mm

位置データには、緯度経度で表現される点を使って、線、領域(ポリゴン)も表現する。

3.2 調査データ

野外での調査データは図1に示したように本構造をしているので、このようなデータを表現するモデルとしてXMLが利用できる。ただし、全てのデータをXMLで表現したのでは位置情報特有の演算が利用しにくいことや、木を何層もたどる検索も無いと思われるので、データ構造を表現するためや、外部とのデータ交換等にXML表現を使うことにする。XML文書をRDBに格納することに関しては吉川ら[6]の研究にあるので、参考にしたい。

XMLでデータ構造を表現することの利点は、次のものがあげられる。

1. いろいろなデータ項目(属性)があるので、それらを柔軟に入力/蓄積したい。
2. 検索での項目設定に利用する意味で利用したい。
3. 検索条件を内部表現に変換するために利用したい。

図2に遺跡データのXML表現例を示す。この例ではDTDは定義していない。データの表現で注意が必要なものに時代表現がある。考古学での時代表現は、何世紀前半、中頃、後半などと紀元年代で表現するが、これはある程度の範囲を持った時間データと考えなければいけない。また、データ自体の幅とは別に、集落など長期にわたり存在した遺跡もあるので、時代を表すには(開始, 終了)という表現が必要である。そのほか、検索時には古墳時代前期などという時代区分表現にも対応が必要なので変換テーブルを用意する必要もある。

```
<?xml version="1.0" encoding="Shift_JIS" ?>
<考古学データ>
  <遺跡 ID="001">
    <名称>〇〇古墳</名称>
    <分類>古墳</分類>
    <位置>
      <点>
        <経度>133 度 55 分</経度>
        <緯度>34 度 42 分</緯度>
      </点>
    </位置>
    <標高>55m</標高>
    <形>前方後円墳</形>
    <方角>北西</方角>
    <大きさ>
      <長さ>150m</長さ>
      <幅>60m</幅>
      <高さ>10m</高さ>
    </大きさ>
    <時代>5 世紀前半</時代>
    <調査年>1990</調査年>
    <説明>この古墳は、...</説明>
    <出土物 type="棺">
      <形>家型</形>
      <材質>砂岩</材質>
    </出土物>
    <出土物 type="土器">
      <分類>須恵器</分類>
      <材質>カヤ系</材質>
    </出土物>
  </遺跡>
</考古学データ>
```

図2 遺跡データのXML表現例

3.3 メッシュデータ

メッシュデータは、広い矩形領域を小さな矩形領域に分割して、その1つ1つの領域を代表する値を

表現している。メッシュデータの構成要素は次のように示される。

要素	表現	例
基準点 (緯度経度)	(x_0, y_0)	$(34^{\circ}40', 133^{\circ}0')$
領域の高さと幅	(V, H)	$(1^{\circ}0', 0^{\circ}40')$
1 メッシュのサイズ	(v, h)	$(5'0'', 7'30'')$
各メッシュの属性値	a	値

メッシュの領域は標準地域メッシュ・コード体系に基づいて作成されたものならば、それぞれの矩形にコードが定まっている。このコード体系は、図3に示すように領域ごとに立て横の座標が割り当てられているので、任意の緯度経度が与えられたとき、どのメッシュのなかに入るかは簡単に計算で求められる。つまり、まず求めたい座標点の緯度経度 (x, y) からメッシュの基点座標の緯度経度 (x_0, y_0) をそれぞれ引き $(x - x_0, y - y_0)$ 、メッシュの高さと幅の度数で割り算した商を並べたもの $[(x - x_0)/v], [(y - y_0)/h]$ がメッシュIDになる。このメッシュ・コード体系は地勢図や地形図の参照番号に用いられており、先の1:25000 地勢図「岡山北部」は“5233-07”というコードが振られている。これは日本全土をカバーする20万分の1地勢図のメッシュで南北52、東西33の地図のなかの南北0番、東西7番のメッシュの領域を示している。この1:25000 地図をさらに 9×9 に分割したものが1km メッシュである。

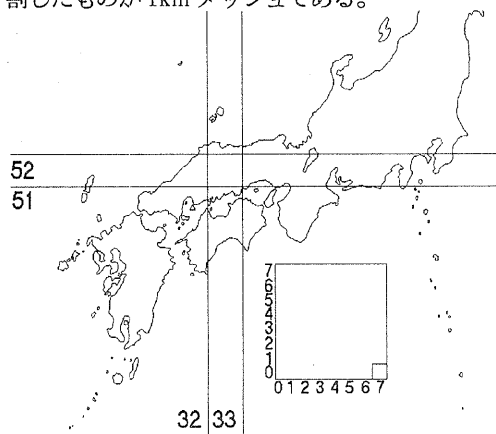


図3 標準地域メッシュ
検索順序と効率: 検索に関しては、調査データが位置の特定に比較的成本がかかるのに対して、メッシュデータはその領域内の座標点からメッシュIDが計算で求められるので、メッシュ領域の属性値が求めやすい。したがって、調査データとメッシュデータの組み合わせられた検索では、先に調査データから条件による検索をおこない、その位置情報を元にメッシュデータを結合させ、次にメッシュデータの条件を検索すると検索効率が良くなると考えられる。

3.4 結果表示地図

対象にしている野外調査では、検索結果の地図はデータ入力で用いる地図よりも広範囲を見られるほうが都合が良い場合が多い。その理由は、遺跡や植物の分布の調査が広範囲にわたっているためであり、それらの相互の関連を見たい場合には全体を見渡せる方がよい。安価な地図データを購入して拡大縮小を自由にできるものも候補の一つにあげられる。もちろん、データ入力に用いた地図の上に検索結果を表示しても良い。これらの地図に対して、検索結果は点や線、ポリゴンといった図形的な特徴を持つので、地図と重ね合わせて表示することや、メッシュデータに対しては値を半透明の色の濃さなどで表すことが必要と思われる。

4 実装

システムの実装にはJavaとPostgreSQLを用いる。JavaではXMLのパarserや地図画像と検索結果の位置情報の表示などを記述し、PostgreSQLはデータの蓄積に用いる。位置データの演算はPostgreSQLが標準で持っているものを利用するが、不足の機能はユーザ定義関数で追加する。システム構成の概略を図4に示す。

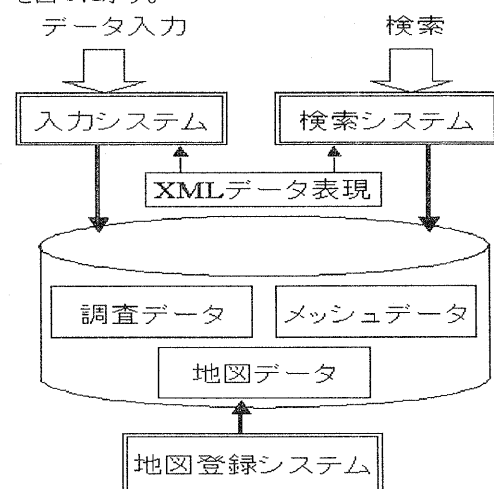


図4 システム構成

次に、上記に示したデータの操作に対して、地図の登録、データ入力、情報検索でのユーザインタフェースを検討しておく必要があるので、次に示す。

4.1 地図の登録

位置入力 の 原地図を 1:25000 地形図にする場合、1/4 の領域 (縦約 18cm, 横約 23cm) を単位として画面全体に表示する程度のサイズが好ましい。画面の横幅を 1024 ドットとすると、1 ドットあたり約 $230 \div 1024 = 0.22\text{mm}$ の精度が必要なので、100dpi の入力精度でスキャンする。この地図の属性情報に地図周辺の緯度経度を入れることで、画面上の相対ドットと緯度経度が計算で求められる。この地図自身をユーザに入力させる予定なので、スキャンした画像 (ファイル名) とそのサイズと緯度経度を合わせたデータが入力できる画面構成をしたインタフェースを作成する。確認のために、入力画像上でポイントを支持してその緯度経度を表示させることをする。

4.2 調査データの入力

上記に示した方法で登録した地図に対して、調査データの入力インタフェースを提供する必要がある。調査データには調査対象によってデータ項目が変わるものがあるが、いちいちそれらに対応して入力画面を設計していたのでは効率が悪いので、自動生成させる。まず、基本的なデータ項目を XML で定義しておき、この XML から構造を取り出して、データ入力画面を構成する仕組みにする。1 つのデータに同じ項目が複数個ある場合も考慮すると DTD を定義しておくことが必要と思われる。この XML は DB へのデータ格納スキーマの決定にも利用する。位置データは PostgreSQL の位置データ型の point, path, polygon に変換して定義する。その他の項目は個々の属性で定義すればよいが、階層が深いものは別のテーブルにすると効率がよい場合もあると考えられる。データの更新と削除の画面構成も同様でよいと考えている。

4.3 検索方法と画面構成

いくつかの調査データを検索するための条件入力 は、XML で定義した構造から入力ボックスが生成可能と考えられる。この入力ボックスを使った検索は次のように行う予定である。(1) 定義してある XML にはそれぞれ名前がついているので、ユーザはこの名前をまず選択する (テーブルの選択)。(2) XML に対応した入力ボックスが表示されるので、必要な項目の条件を入力する。(3) 繰り返して、この条件ボックスを必要だけ入力する。(4) 実行すると各条件を満たしたものが地図上に (点や線や領域) 表示される。

各条件ボックスにラベルを付けて括弧や論理演算子でそれぞれの条件の関係を指定できるようにすると複雑な条件も記述可能になると思われる。

また、検索結果の表示には色の指定と点の図形形状、線の太さも指定できるようにしておく。こうして得られた表示地図は、検索結果ごとに表示をリセットしないようにしておく。こうすることで、例えば年代順の検索結果を異なる色で同一地図上で見ることができ結果の解釈に役立つと思われる。また、検索条件を個別に指定できるのでシステムを複雑にすることなく、また問合せ条件の解析処理を簡単にする利点がある。視点を換え、表示画面上の AND 演算と考えても良いが、色をそれぞれの検索結果で返られることで単なる AND では表現しにくい結果を得ることができると考えられる。

5 終わりに

この論文では、野外調査で収集したデータを検討し、そのデータをどのように蓄積しておく、研究者が必要とする検索が行えるかを論じた。特に、地図上に条件を満たしたデータ位置を表示したいということに関して、データ構造、データ入力、データ検索に関するシステムの概略設計を示した。今後は、詳細な設計をしてシステム構築する予定である。謝辞 この論文を作成するにあたり、岡山理科大学総合情報学部の岡山学研究会の方々に協力をいただきました。ここに感謝します。特に、社会情報学科の志野敏夫先生、生物地球学科の波田善夫先生には野外調査データについて多くの助言をいただきました。ここに、感謝します。

参考文献

- [1] 久保幸夫, 巖網林, 「地理情報科学の新展開」, 日科技連 (1996)
- [2] 高木悟, 松本一則, “地図情報を利用した情報検索”, 情報処理 Vol.41, No.4, 357-362 (2000)
- [3] 松本圭司, 他, “地図データベースへの応用を考えた多次元データアクセス:ビット埋め込み R 木”, 情報研報 Vol.2000, No.10, 169-176 (2000)
- [4] G-XML, <http://gisclh01.dpc.or.jp/gxml/contents/>
- [5] 国土地理院, 測地成果 2000, <http://www.gsi-mc.go.jp/LAW/G2000/g2000.htm>
- [6] 吉川正俊, 志村壮是, 植村俊亮, “オブジェクト関係データベースを用いた XML 文書の格納と検索”, 情報処理学会論文誌, Vol.40, No. SIG6(TOD3), 115-131 (1999)