



地図情報を利用した情報検索

■ 情報のビジュアル化

今日の情報システムは、いうまでもなくコンピュータの処理能力・記憶容量の増大とともに著しい進歩をとげてきた。そして、1980年代後半から爆発的に普及を始めた情報通信環境、そしてインターネットによって、さらにその進歩は加速を早めている。これら情報システムの初期の段階においては、処理容量の制限から処理対象のデータは一般的に文字・数値からなるテキストデータであった。したがって、従来の情報検索は常にテキスト情報を対象に行われていた。

しかし最近になって、情報システムでビジュアルデータなどの非テキストのデータが取り扱われるようになり、ビジュアルな情報検索が必要となってきた。非テキストの情報は従来の情報システムと比べて、検索技術に限らず、利用環境、ユーザインタフェース環境、そしてアプリケーションに至るまで大きく異なる。そこで本稿では、近年実用的になってきた非テキスト情報の1つであるグラフィックスの代表的な用途である電子地図メディアを利用した情報システムと、そこでの新しいビジュアルな情報検索技術の姿について述べてみる。

(株) KDD 研究所

高木 悟・松本 一則

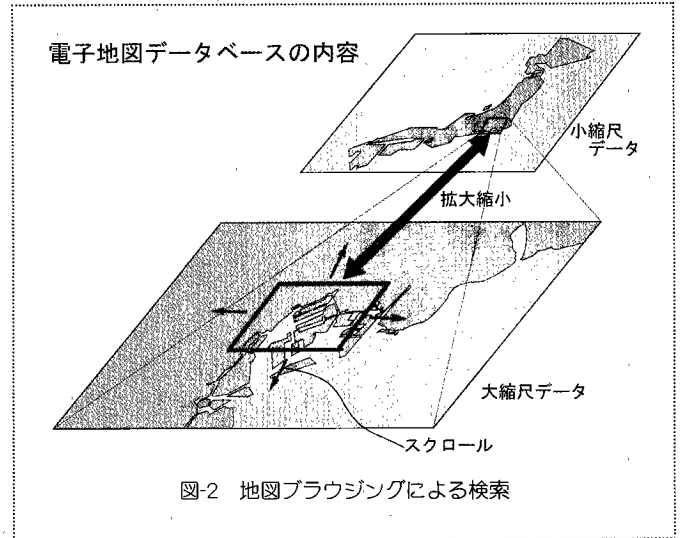
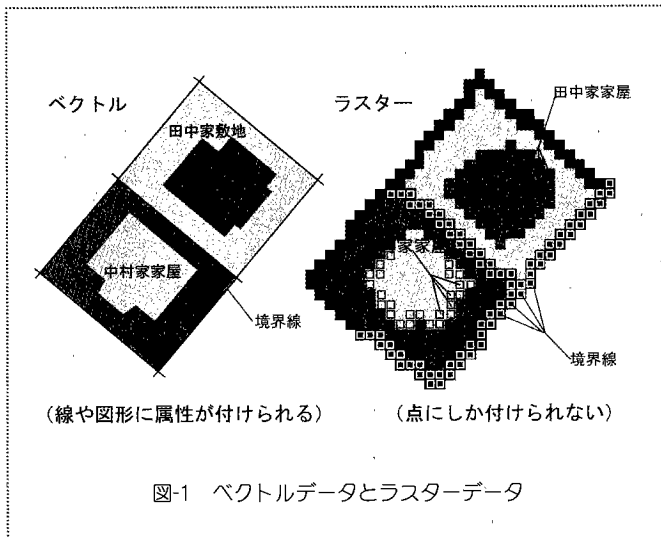
■ 電子地図

頻繁に活用されているグラフィックス情報である電子地図とは、古来一般に紙の上に線・面・文字そして記号などの図形を使用して描くことで地形の形状や町並みなど表現していた地図を、数値情報としてコンピュータに蓄積し、それをコンピュータのグラフィックス機能を利用して表示・利用する情報メディアである。表示機能の進歩によって、電子地図では二次元のスクリーン上に擬似的に三次元の形状も表現可能となっている。

電子地図のデータ形式には二次元・三次元によらず、大別して2つの形式がある。1つはインターネットのWebブラウザでも頻繁に利用されているGIFやJPEG形式などのようにグラフィックスを点の集合として表現するビットマップまたはラスター形式と呼ばれるもので、航空写真・衛星画像・紙地図をイメージスキャナにより電子化したものなどがその具体例である。一方で、建設図面や電子回路図面などの作成に利用されるCAD (Computer Aided Design) システムやイラストレーシ

ョンで主に利用される、グラフィックスを線やポリゴンといった図形で表現するベクトルデータを利用した電子地図がある。ベクトルデータの電子地図はラスターデータに比べて作成が困難であり、一般にCADシステムによって、オペレータが手作業で図形を入力してデータを作成する。一方データ利用に関してはベクトルデータの方が容易で、応用性も高く必要な処理能力も相対的に低くてすむ。

電子地図はすでにさまざまな場所で利用されている。直接的には、パソコン用の地図帳ソフトウェアをよく目にするが、さらに身近なものは、実は紙の上に印刷された一般の地図である。旧来の紙地図はコンピュータを利用せずに製版していたが、近年ほとんどの地図作成過程は電子地図を扱うCADシステムに代わっている。そのため、手書きでは困難であった地形の形状を反映した陰影を付けた地図や、衛星画像を下地にした地図など、非常に多彩な地図が出版されている。



■ 意味を持つ図形と地理情報システム ——

このような電子地図メディアを扱う情報システムは、電子地図をコンピュータのグラフィックススクリーン上に表示する地図帳に相当する最も基本的なシステムや、前述のような電子地図の作成・管理を行うためのCADシステムだけに限られたものではない。地図のグラフィックスの各要素にたとえば機材・建築物の機能や名称、土地の地価や所有者などの意味を持った属性情報を持たせ、それを検索キーとして利用するデータベースシステム、さらに属性情報と図形情報を用いた空間的な論理演算を行うことで、設備管理・都市計画・エリアマーケティングなどに利用できる情報システムなどへの応用が活発になされている。このような情報システムは一般にGIS (Geographical Information System: 地理情報システム) と呼ばれており、その処理結果を表示する情報メディアとしても、もちろん電子地図が広く用いられている。

地図グラフィックスの要素に属性を持たせる手法は、ベクトルデータの方が多彩である。ラスターデータは、グラフィックスを点の集まりとしてしか表現できないため、たとえば人工衛星で撮影した画像から海面の温度分布を出力した結果のように、基本的に一点一点に対して個別の属性を付けることしか行えない。それに対しベクトルデータの場合は、グラフィックスが点だけでなく線やポリゴン、そしてそれらをグループ化したものなどから構成されるため、それらの図形に対して属性を付加することができる。たとえば、家屋の形状のポリゴンに対して、その家屋に関する属性を持たせるといったことが可能になる (図-1)。

このように一般にベクトルデータへの属性づけの方がより高度な意味づけを容易に行うことが可能である

ため、GISではベクトルデータが多用される傾向にある。

■ 地理情報システムの検索スタイル ——

電子地図にかかわる情報検索は地理情報システムの基本機能であるが、その検索スタイル・ユーザインタフェースは主に次のようなものがある。

まず、キーワード検索がある。これは、住所・地名・駅名検索、道路、路線、交通機関検索、建物・ランドマーク検索、電話番号検索など、空間的な情報を文字情報に置き換えたものを使用して地図を検索するタイプである。

次に、よりビジュアルな検索スタイルを与えるアイコン検索がある。キーワードに相当するものをアイコンとして用意し、それを使って情報検索する。

そして最後に地図独特のビジュアルな検索がある。電子地図にはコンテンツの地理的な場所という情報が必ず存在する。加えて、情報の地理的な分布範囲すなわち地図の縮尺も存在する。これら2つの要素は、電子地図特有の情報検索キーとして与えられるべきものである。これらの検索条件は、地図の場合ブラウジング行為と連携したビジュアルな検索スタイルで与えられることが多い。たとえば広い地域が見られる (小縮尺の) インデックスの地図を利用して、領域を視覚的に指示することが可能である。これは、従来の紙地図でもインデックス地図から詳細な大縮尺の地図を検索することがしばしば行われており、自然な方法である。このとき、地図の利用者は単に地図を拡大・縮小・スクロールしたり、場合によってはより詳細な地図がある矩形領域をクリックしただけで、特に検索を大きく意識することはない。すなわち、利用者が縮尺や地域の変更を繰り返しながら情報をブラウジングしていく行為が自然に情報検索に繋

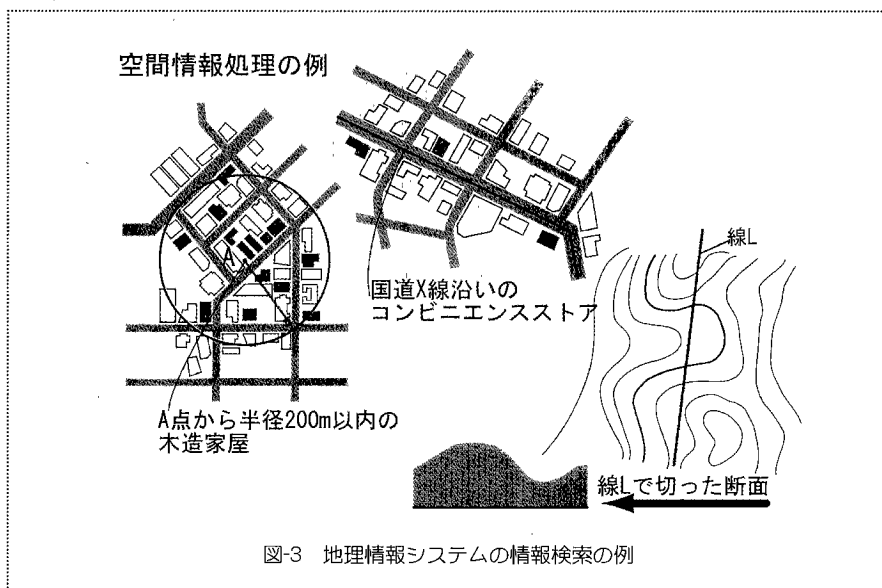


図-3 地理情報システムの情報検索の例

がっているのである (図-2)。

こうした地図ブラウジングによる自然な検索は、表示領域を検索キーとしてデータベースに問合せをすることに相当している。

この機能は表示領域の移動機能と、表示領域の広さに応じてデータの詳細度を調整する機能によって具体化されている。表示領域の移動機能は、ユーザがスクロールボタンなどで表示領域の変更を指示すると、要求された領域の情報を検索し、表示するものである。拡大縮小機能は、ユーザによる拡大・縮小の指示に従って、縮尺の違うすなわち詳細度の違う地図を段階的に切り替えながら表示するものである。拡大縮小機能は、より高度な機能も伴うこともある。それは情報の要約機能であり、縮尺に見合った情報がなかった場合、より大縮尺のための情報からその縮尺に見合う密度の情報に間引き、すなわち要約を行った上で電子地図を生成する。間引き機能の実装は、情報がさまざまな意味合いを持った図形の形で保存されているベクトルデータの方が行いやすいのはいうまでもない。

地図情報の検索は、地図ブラウジング以外にも、検索対象の領域・線・点などを視覚的にユーザが指示するスタイルがある。このような検索の基本的な例は、WWWのクリックブルマップによる情報検索にもみられるが、さらにテキストやアイコンベースの検索と組み合わせられた例もある。たとえば〇〇町の中にある公園を検索といったときに、〇〇町の領域図形の指示をしたり、国道〇号線沿いのコンビニエンスストアを検索といったときの国道の線図形を指示したりといった具合である。さらに、自分で描画した線や面などの図形を検索条件として利用する情報検索も業務システムなどでは頻繁に利用され、たとえば自分で描画した線分に沿った断面を表示する (図-3)。

以上のような情報検索技術は、スタイルこそ大きく

違うが実はすべて二次元・三次元空間に配置された情報を、たとえば地点や領域などの空間的な検索条件によって検索する技術を基礎としている。

■ コンシューマ市場への展開とWeb

他の情報メディアと同様に、電子地図や、それを直接的に利用する情報システムも特殊業務用途からコンシューマ用途へ展開し始めている。その最も代表的な例は1990年頃から普及を始めたカーナビゲーションシステムである。カーナビゲーションシステムのアプリケーションは、電子地図帳として機能する基本的な表示機能に加えて、GPS (Global Positioning System) などの位置測定技術を利用した自車位置の自動表示、そして目的地までの経路検索など、地理情報システムの解析機能が一般的に盛り込まれている。さらに1997年頃から、電子地図を配信するWebサイトも複数出現するなど、電子地図は新たな情報メディアとして広く活用され始めている。特に現在大変注目されているモバイルコンピューティング分野では、デスクトップ環境以上に電子地図メディアが重要になると考えられている (表-1)。

ところで、Webの情報メディアは従来の情報メディアと大きく異なり、情報を発信する立場と受信する立場に明確に分かれることはなく、一般ユーザでも自由に情報発信が可能であるという大きな特徴を持っている。情報発信の自由が広く一般に得られることは、従来の情報流通形態のような情報の流れを一方向的に制限されてしまうメディアと比べてさまざまなメリットが得られ、これがWebの爆発的な普及を促進した1つの理由であり、Webの先進性の本質でもある¹⁾。

特に自由に情報発信できる環境はコミュニティのための基盤としては不可欠なものであり、同じく自由な情報発信の場を提供した電子掲示板 (BBS) とともにネット

電子地図提供サービス

Mapion	http://www.mapion.co.jp/
MapFan	http://www.mapfan.com/
AtlasMate	http://www.sei.co.jp/atlasmate/
MAPOO	http://www.mapoo.or.jp/
Yahoo Map	http://maps.yahoo.com/yahoo/

その他のモバイル情報サービス

乗換案内	http://www.jorudan.co.jp/index.html
ハイパーダイヤ	http://www.hyperdia.com/
駅前探検倶楽部	http://ekimae.toshiba.co.jp/
i-Point network	http://www.i-point.ne.jp/pc/index.html

表-1 電子地図関連Webサイト

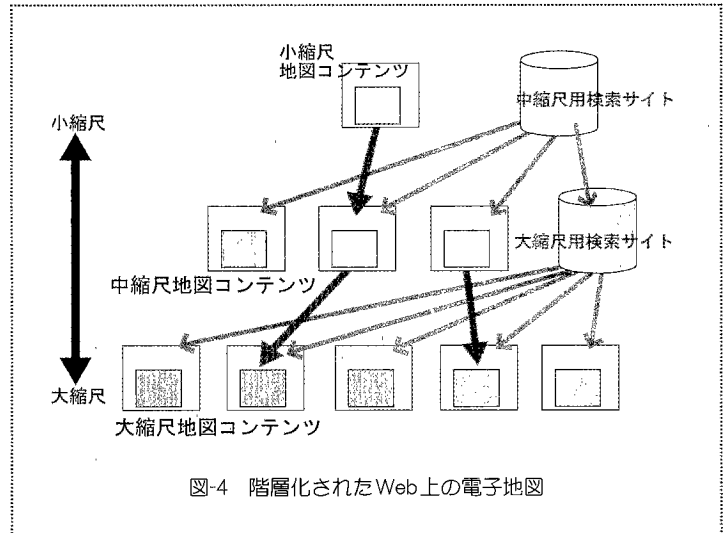


図-4 階層化されたWeb上の電子地図

ワーク上のコミュニティであるネチズンを生み出した。このWebの自由な情報発信環境は、インターネット上の誰でも利用できる情報発信サーバ（Webサーバ）と、容易に情報を表現できる情報メディアフォーマット（HTML）、すなわち情報リテラシが容易に獲得できること；そして情報同士の連携を作りコミュニティを作ることができるハイパーリンクがあることから成り立っている。

さて、Webの情報メディアとして電子地図を考えてみた場合、やはり情報を発信する立場としても自由に使えることが大切なことになる。ところが、電子地図による情報発信は一般に情報発信側と受信側が明確に分断されており、これはWeb上でも同じである。残念ながらその意味で現状の電子地図はWeb情報メディアとしてまだ効果的に機能しておらず、改善の余地のある発展途上の情報メディアだといえる。

■ Web上の電子地図の検索

仮に電子地図が誰にでも発信可能なWeb情報メディアとして成立した場合には、現在のWebコンテンツと同じようないくつかの問題点が生じてくる。その1つとして情報検索の困難さが挙げられる。

Webの情報は、ハイパーリンクを活用して情報を関連づけることで情報の検索を可能とするが、情報発信の側が必ず関連づけを作らなければならないわけではなく、実際にコンテンツ同士の結びつきは現在のWebではあまり密ではない。このことは情報の作成を容易にしているのだが、このままでは使いづらい情報システムになってしまう。そこで情報同士のいわば仲介するインデックスの役割を果たすYahooやInfoseekのような検索サイトが出現し、重要な役割を果たしている。これら検索サイトはエンドユーザが与えた検索条件に合致する情報に関して、ハイパーテキストによる情報

一覧（インデックス）を提供する。

しかし、Web上の電子地図メディアを考えた場合、Webコンテンツ同士の関連づけを行う検索エンジンには、さらに異なる役割が追加される必要があるだろう。先に述べたように、地図の場合は同じ道路地図情報だが、領域や縮尺の異なる地図が存在する。そして、それらの地図の検索・切替えは自然なブラウジング行為によってビジュアルに行われていくのが自然である。これを先のWebの構成に実装すると図-4のような構成が想定される。インデックス地図と実際の地図には基本的に大きな区別はなく、取り扱う縮尺・地域が異なる電子地図コンテンツが階層的に分散し、リンクされたシステムである。

中・大縮尺用検索サイトにおけるコンテンツの関連づけは不十分なことが多く、一般的に小縮尺のインデックス地図を提供する検索サイトが関連づけを行うことになる。たとえば、米国連邦地理データ委員会（FGDC）では、どのような地理情報がどこにあるかを記述したメタデータをClearinghouseと呼ぶデータベースで管理しようとしている⁹⁾。また、日本でも同様に国土空間データ基盤推進協議会（NSDIPA）が国土空間データの整備推進を関係省庁に働きかけている。

■ 電子地図のリテラシと新たな情報流通スタイル

Web上での電子地図発信の姿としては先のような例が考えられるのだが、電子地図はまだ誰もが情報発信できるメディアとして確立していないという根本的な問題は残されたままである。本章では、その問題を解決すべく筆者らが開発しているJaMaPS²⁾と呼ぶ新しい情報流通スタイルを持ったWebプラットフォームについて述べる。

電子地図を利用した情報発信で大きな問題は地図の作

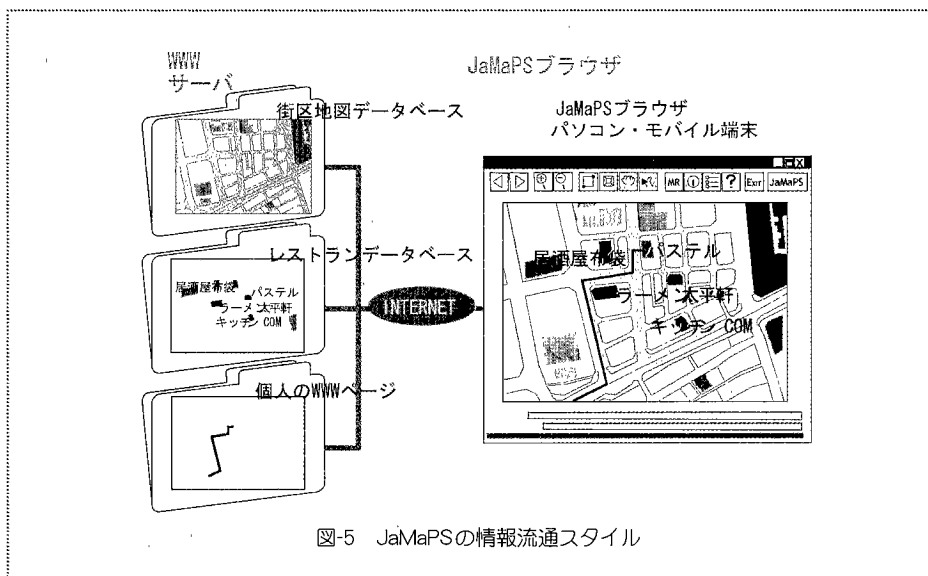


図-5 JaMaPSの情報流通スタイル

製が容易ではないことである。個人で地図を利用する場合、市販の地図を購入して、その地図の上に各自が追加したい情報を書き加えるという使い方がしばしば行われている。これはほんのわずかであるが、ベースとなる地図の上に各自が独自の地図のコンテンツを作成し追加したことに相当し、個人による数少ない地図作製例である。この方法ならば、かなり容易に地図情報を作成することができる。そこで、筆者らのシステムはこの概念をWebに適した方法で導入する。

Webでは、通常他人のコンテンツに情報を加えることはできない。仮に可能であっても所有権の問題が大きい。一方、少し工夫した方法で、地図上に情報を書き加える作業が業務用途で広く行われている。元の地図の上に透明なフィルムを重ね、そのフィルム上に地図を書き加えていく方法である。コンピュータ上のCADやドローソフトではこの方法はレイヤ処理と呼ばれ、元の図面を変更することなく新しい情報を追加表示することが可能とする。JaMaPSではこのレイヤ処理をWeb上の分散したサーバに置いた各レイヤに相当する地図コンテンツと、それらにアクセスして画面上にそれぞれのサーバの地図を1つのレイヤとして処理し、重ね合わせて表示する新しいWebブラウザを用いて再現した(図-5)。

このとき問題となるのは、各レイヤが同じ場所同士で重ね合わさるべきではないことである。通常のWeb上の電子地図はGIFやJPEGといったグラフィックスで表現されているが、それらには明確に場所を示すような情報は載っていない。グラフィックスで表現された地図を同じ場所同士重ね合わせる簡単な手法は、それぞれのグラフィックスにたとえば緯度経度座標などの共通のグローバル座標を与えておき、そのグローバル座標系ですべてのグラフィックスを描画すればよい。

この手法は、グラフィックスとして表現された電子地

図に一般的に利用可能な方法であり、従来の1つのWebサーバから取得したコンテンツのみの表示を行うものから、複数のいわばモジュール化されたコンテンツを同時に利用する、従来のWebにはなかった新しい情報流通スタイルを提供する。しかもそれらのコンテンツはビジュアルな重ね合わせによる相乗効果で価値の高い情報を生み出す。

本章の冒頭で述べたように、レイヤ処理を利用した同手法は元の図面の上に書き加える情報だけを別のサーバから独立した情報として発信可能であるため、それだけならば情報作成が容易である。もちろん元データの上に重ねて使用することへの合意がネットワーク上のコミュニティで形成されなければならないが、元データを複製してさらに情報を書き加えたものを発信することに比べれば合意は得やすく、しかもこの手法によって元データを複製使用される危険性も低減させることができる。

■ 新たな情報検索スタイル

以上のような情報流通スタイルでは地図情報検索にはさらに新たな機能を必要とする。従来の地図情報はそれぞれモジュール化されているため、個別ではあまり役に立たず、1つのコンテンツだけを検索できるような従来の検索サーバでは使いづらい。

そこで、モジュール化された情報をビジュアルに重ね合わせることで、ユーザの要求を満足できるような情報の組合せを検索できるサーバが必要となる。たとえば新宿駅周辺の駐車場の空車情報を探する場合、新宿駅周辺の道路地図サイトと、駐車場の情報を提供する地図サイトを探し出し、これらを重ねて使うようにユーザに案内する必要がある。このためには、まず各モジュール化された情報の特性をインデックスとして蓄積する必要がある。

1つはYahooなどのように、その情報がどんな情報で

あるのかといったインデックス情報をあらかじめ情報発信者に登録してもらう方法があり得る。先の Clearinghouse では、コンテンツ作成者がメタデータを作成し、それを自ら登録するようになっている。

それに対して Infoseek のような情報収集ロボットと呼ばれる自動的にインターネット上の Web コンテンツの内容を巡回し収集する仕組み（ロボットと呼ばれることも多い）とインデックス情報の自動生成技術を統合することも考えられる。NTT が行ったモバイルインフォサーチ実験⁶⁾では、LIG と呼ぶロボットが使用され、位置情報を示す言語表現が含まれるアンカーを検出し、その先のコンテンツを優先的に収集する戦略をとっている^{10), 11)}。ただし、インデックスの生成に際し、コンテンツの内容自体を解釈する必要が出てくるため、登録方式に比べ、インデックスの精度は良くない。インデックス生成の性能は、自然言語処理技術、とりわけ HTML などのコンテンツ解析技術に負うことが多いと考えられる。

■ SYG (Scalable Vector Graphics) —

電子地図を表現するには、ベクトルデータとラスターデータの2つの形式が存在し、ベクトルデータの方が利用しやすいことは先に述べた。さらに、前章の最後に記したロボット型の検索システムでは、なおさらにベクトルデータの方が利用しやすい。それは、ベクトル型のデータではグラフィクスが細かな図形データ（グラフィクスオブジェクト）の集まりとして構成され、さらにたとえば地名や駅名などの文字列情報も図形情報の一種として処理しやすい形で記述されているからである。したがって文字列を対象とした現状の検索システムでも容易にインデックス情報を生成可能であるし、グラフィクスオブジェクトをインデックスとした検索方法も実現できる。

それに対して、ラスターデータでは利用者には文字として認識できる情報も実際には点の集まりとしてしか記述されていないため、画像認識技術などの高度な技術を元に検索サーバを再構築しなければならない。

ところが、Web 上で流通させることができるグラフィクスデータは今のところ GIF、JPEG の2種類であり、それらはすべてラスターデータである。それに対して現在 W3C ではインターネット上の標準ベクターデータフォーマットとして SVG (Scalable Vector Graphics)³⁾を策定中である。本フォーマットは XML⁴⁾に準拠しているため、HTML などの Web コンテンツとの親和性が非常に高く、しかも XML の特性を生かしてグラフィクスオブジェクトに地理情報システムで不可欠な属性データを柔軟に追加可能でもある。

■ 異種情報メディアの電子地図上への統合 —

異種メディア間の変換技術は、近年大変注目されるようになっており、大画面用 HTML・極小画面用テキスト間変換技術としての文書要約技術⁷⁾や、HTML・音声間の変換技術⁸⁾など、特にモバイルコンピューティングサービスを1つのターゲットとして、サービスが開始されてきている。

電子地図は、実空間上に散らばる物体や情報などを視覚的に表現するのに優れた表示方法をユーザ端末上で提供する。しかも、ベクトルグラフィクスデータ利用すれば、拡大縮小しても画像が崩れず、間引きも容易であるため、特に PDA などの携帯情報機器でも画面のサイズや伝送速度に応じた柔軟なサービスを提供することが可能になる。

そこで、地図として表現されていない空間情報、たとえば HTML コンテンツ中の電話番号や住所、さらにもっと抽象的な表現などで記述した情報をメディア変換し、地図上に表示することが可能にすれば、電子地図を利用した検索に有効活用できる^{5), 6)}。「モバイルサーチ実験」では、住所が記述されたコンテンツに対し緯度と経度の情報を付与している。こうした位置情報の付与と地図上のオブジェクトとの関連づけによって、各種の空間情報の検索に用いることができる情報の量が急速に増えていくことが期待されている。

■ まとめ —

以上のように、特に Web 上での情報メディアとしての電子地図は、現時点では未熟なメディアであるが、Web の利用範囲や情報流通のスタイルを大きく変える可能性も持つなど、今後大いに発展する可能性を持つ情報メディアである。21世紀へ向けた新しい社会基盤、情報システム基盤として、電子地図が広く貢献できることを期待してまとめとしたい。

参考文献

- 1) 吉見俊哉, 水越 伸: メディア論, 放送大学教材, 53456-1-9711.
- 2) 高木 悟: インターネット対応分散型地図表示システム, 第8回機能図形情報システムシンポジウム論文集, pp.45-49 (1997).
- 3) SVG: <http://www.w3.org/TR/SVG>
- 4) XML: <http://www.w3.org/TR/REC-xml>
- 5) 相良 毅, 有川正俊, 坂内正夫: ネットワーク上各種情報源からの地理情報抽出収集手法, Procs of GISA, Vol.8, pp.331-334 (1999).
- 6) モバイルインフォサーチ2: <http://www.kokono.net/>
- 7) 要約技術 (I-gate): <http://www.i-greet.com/i-gate/index.html>
- 8) 音声記述言語 (VoiceXML): <http://www.voicexml.org/>
- 9) FGDC Geospatial Data Clearinghouse Activity: <http://www.fgdc.gov/clearinghouse/>
- 10) 横路誠司, 高橋克巳, 島 健一: 情報内容を考慮した情報収集方法, 情報処理学会第56回全国大会 (1998).
- 11) 横路誠司, 三浦信幸, 高橋克巳, 島 健一: 特定分野のリソース収集を行う WWW ロボットの評価, 情報処理学会第57回全国大会 (1998). (平成12年2月28日受付)