2001年　3月
APA No. 78-4
財団法人 日本測量調査技術協会

ベクトルデータの効率的な更新及び利用に関する研究
～25,000 レベル GIS におけるデータ入力、及び更新について～

小野木　康介

はじめに

国土地理院では、国土の基本的な地理情報の電子情報化を目的として、平成10年度から、「数値地図25,000（地図画像）をベースとした構造化ベクトルデータ（以後、25,000 レベル GIS 基盤情報と呼ぶ）」の作成を行っている。対象は、道路・鉄道・水系・行政界、および基準点の５項目である。

GIS 基盤情報は、国土交通省が目指す「電子国土」構築の骨格情報として位置づけられており、今後の地形図や、数値地図の更新にあたっての、基礎データとして利用される予定である。

25,000 レベル GIS 基盤情報と、既にかなり広範囲にわたって整備されている、2,500 レベル GIS 基盤情報（空間データ基盤）の仕様を、表１に示す。国土全域を対象としていることが、25,000 レベル GIS 基盤情報の、最大の特色である。

本研究は、25,000 レベル GIS 基盤情報の作成におけるデータ入力、および更新業務の効率化に寄与することを目的とした。

１．研究概要

データ作成、および更新業務の効率化、と言っても、対象となる範囲は非常に広い。本研究では、以下の２項目を対象に行った。

（１）効率的なベクトルデータ取得手法の検討

ベクトルデータの取得手法としては、ラスター／ベクトル自動変換とディジタイズの２通りがある。それぞれの特徴を比較した上で、効率的と考えられるデータ取得手法の検証を行った。

（２）データの更新、および提供手法の検討

GIS 基盤情報は、「時間」を１つの属性要素と捉えているのが特徴である。この性質を利用した更新管理、また、データ提供のあり方について、考察を行った。

表１　25,000・2,500 レベル GIS 基盤情報の仕様

	25,000 レベル	2,500 レベル
範囲	国土全域が対象	都市計画区域が対象
単位	1/25,000 地形図図葉	1/2,500 国土基本図図葉
データモデル	構造化ベクトルデータ	構造化ベクトルデータ
位置精度	許容誤差は図上 0.1mm 以内	1/2,500 縮尺相当
項目	行政界・道路・鉄道・水系	行政界・道路・街区・直轄国道・直轄河川・鉄道・場地・建物・基準点

２．効率的なベクトルデータ取得手法の検討

25,000 レベル GIS 基盤情報におけるデータは、構造化されたベクトルデータである。ベクトルデータの取得には、その元となる資料から、地物の形状、および属性を取得する必要がある。本章では、一般的なベクトルデータ取得手法の比較を行い、それぞれの特徴をふまえた上で、効率的な取得手法の検討を行った。

2.1　ベクトルデータ取得手法の概要と比較

ベクトルデータの取得手法は、ラスター／ベクトル自動変換とディジタイズの２通りに分けられる。それぞれの特徴を以下に示す。

１）ラスター／ベクトル自動変換

コンピュータ上での一括処理で、ラスターデータをベクトルデータに変換する方式である。変換処理は、２万５千分の１地形図１枚あたり数分程度の時間で行うことが可能である。

変換効率（データに含まれる図形が、正しい形状に変換されている度合）は、データ取得元となる図面の状態に左右される。現在のところ、100%正確な変換は存在せず、変換処理後に変換漏れや変換誤りなどの修正が必要になる。修正は、オペレータが対象箇所を指定した上での自動処理、もしくは手作業で行うことになる。変換誤りの修正作業には、新規ディジタイズ以上に手間が必要となる場合も多い。また、構造化を行う場合は、それに応じた手作業が生じる。

２）ディジタイズ

データ取得元の資料・図面から、オペレータとコンピュータとの対話的処理により、ベクトルデータを取得する方式である。ディジタイザを使用して直接図面からベクトルデータを取得する方式や、ディスプレイ上で背景にラスターデータを表示して、ベクトルデータを取得する､ラスタートレースと呼ばれる方式がある。

現在は、ラスタートレース方式が主流である。

３）適切なベクトルデータ取得手法の考察

速度・効率等の面から、自動変換方式とディジタイズ方式の比較を行った。比較結果を、表２に示す。

先にも述べた通り、現在のところ、自動変換で構造化まで行うことは、極めて困難である。そのため、特に大縮尺の場合には、ディジタイズ方式が使用されることも多い。また、はじめに自動変換でベクター化を行い､修正および構造化をディジタイズで行う手法も、多く採られている。本研究においても、まず全体のベクター化を行い、その上で構造化や属性付与を個別処理、またはディジタイズで行っていく、という考え方を採用した。

2.2　検証作業

2.1での比較結果に基づき、ベクトルデータを効率的に取得するための検証を行った。取得対象は、25,000 レベル GIS 基盤情報において、特にデータ量が多くなる１条道路・２条道路（中心線）・水涯線とした。

表２　自動変換方式とディジタイズ方式

	自動変換方式	ディジタイズ方式
速度	使用するソフトウェア・ハードウェアに依存する。非常に高速である。	オペレータによる手作業のため、高速化にはおのずから限界がある。
効率	100%の変換は不可能である。一般的に90～95%といわれている。	オペレータの熟練度などにより格差が生じる。
修正量	変換漏れや変換誤りの部分について、手作業による修正が必ず生じる。	人的な判断ミスによるものが大半を占める。オペレータの熟練度などにより異なる。
総評	大量なベクトルデータの取得に、特に適する。構造化には不適な部分もある。	構造化されたベクトルデータの取得に適する。

表３　２万５千分の１地形図　版分け一覧

版	表記内容
墨版	行政界，道路，鉄道，建物，境界等
藍版	水涯線，水田記号
褐版	等高線，崖，土堤，地下鉄
墨マスク版	樹木に囲まれた住宅地のマスク
藍マスク版	水マスク
褐マスク版	国道マスク
注記・記号版	注記，記号
注記マスク版	注記周囲のマスク

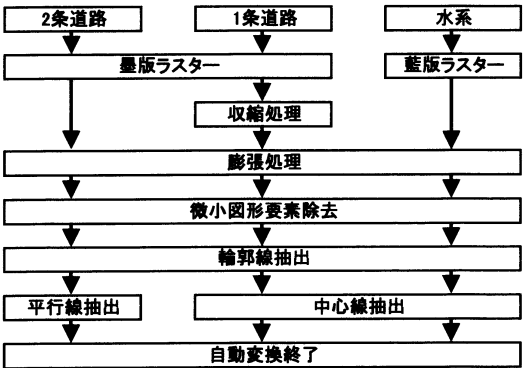

図１　自動変換処理フロー

１）使用した地形図

２万５千分の１地形図「上郷」　（実質1.00面）

国土地理院の２万５千分の１地形図（以下「地形図」と呼ぶ）は、版ごとに区別して管理されている。その一覧を表３に示す。地形図のひとつの版には、数十種類に及ぶ項目が収められており、ベクター化対象外の項目も数多く含まれている。効率を高めるためには、ベクター化対象項目のデータ取得を行いやすくするため、ベクター化前にラスターに対して、幾通りかの処理を行う必要がある。

地形図ラスターは、25μm（40dot/mm）で取得されたものを使用した。データの取得精度としては、かなり高精度である。地形図の場合、精度が高いことに加えて、表３に示した通り版分けもなされていることから、ラスターとしては、かなり高品質な状態で管理されていると言える。

25,000レベルGIS基盤情報の取得において、直接的に取得資料として使用する地形図は、墨版（行政界・道路・鉄道）、藍版（水涯線・湖沼）、褐版（地下鉄）の３種類である。「上郷」は鉄道・地下鉄を含まないため、本検証では、墨版と藍版の２種類を使用した。

２）使用した機器・ソフトウェア

本検証に使用した機器、およびソフトウェアは以下の通りである。

【使用した機器】

自動変換

Sun ultra5 Model 270 Workstation

128MB　SunOS 5.6

ベクトル検証・出力処理

COMPAQ AP400　　512MB　WindowsNT 4.0

【使用したソフトウェア】

自動変換

Wit/Map ver.3.0　（社会調査研究所）

ベクトル検証・出力処理

SIS ver.4.1J　（インフォマティクス）

３）作業内容

本検証において、実際に行った変換処理の流れを図１のフローに示す。

ラスター／ベクトル自動変換前に、処理対象であるラスターに対して、膨張処理（線幅を太くする処理）を行う。ラスター上、線幅が細い箇所で、ベクター化の時に線分が寸断されるのを防ぐためである。ただし、膨張量が大きくなりすぎると、平行する線要素の間が潰れ、たとえば、本来２条線であったものが、１条線になってしまう現象が生じる。したがって、図面の状態等を判断して、適切な設定を行う必要がある。

地形図上では多数の記号シンボルが表示されているが、そのままベクター化を行うと、短い線分が大量に発生することになり、作業の妨げになる。ベクター化

図２　２条道路中心線の抽出結果

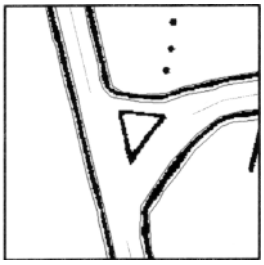

図３　変換漏れの例（ロータリー部）

前に微小図形要素除去（図形のピクセル数が、一定値未満のものを除去する処理）を行うことで、その大半を除去することが可能である。

２条道路における変換処理手順を、以下に示す。

①２万５千分の１地形図墨版ラスターデータ（以下「墨版ラスター」）を取得する。
②墨版ラスターに対して膨張処理（４ピクセル分）を行う。
③膨張処理済み墨版ラスターに対して微小図形要素除去（1,000ピクセル未満）を行う。
④輪郭線ベクトルを抽出する。
⑤輪郭線に対する平行線ベクトルを抽出する。

処理の結果、取得されたベクトルデータを、図２に示す（細線が、ベクトルデータである）。

４）検証結果

自動変換を行ったそれぞれの要素について、複数の観点から検討を行った。その結果を以下に列記する。

①処理時間

一連の処理に要した時間は、地形図１枚当たり、最もデータ量の少ない水系で約20秒，以下１条道路で50秒，２条道路で110秒であった。ディジタイズで行えば、数時間～２日ほどに及ぶ作業になると想定される。

②位置正確度

GIS基盤情報の製品仕様書では、ベクトルデータの位置正確度は、２万５千分の１地形図と比較して0.1mm以内の標準偏差に収まるよう、定められている。変換後のベクトルデータについて、地形図ラスターとの重ね合わせ確認を行ったところ、変換が行われている箇所については、概ね位置正確度を満たしていた。

また、曲線部など、ディジタイズでは作業効率、および位置正確度が低下しやすい部分においても、特に低下が見られなかった。

しかし、不定形の交差部や狭小な街区等において、変換誤りや変換漏れが散見された。しかし、整形が必要となる箇所は、交差部（特に三叉路等、形状不定のもの）や大規模な構造物との交点等、位置的にある程度限定されるため、一括処理による整形を行った上で、個別の修正対象を絞り込むことが可能と考えられる。

２条道路中心線の取得における、変換漏れの例を図3に示す。

2.3　まとめ

今回の検証によって確認された内容、ならびに、効率化のために必要と考える事項を、以下に記す。

（１）曲線部の取得における利点

位置正確度のところで述べた通り、曲線部のベクトルデータ取得に関して、自動変換の優位性が確認された。あらかじめ補間点の間隔等を設定することで、平滑化の程度を指定できるため、項目ごとに適切な値を

設定することで、さらに効率の向上が見込まれる。

（2）変換後作業の自動・半自動化

自動変換直後のベクトルデータを、そのまま構造化することは困難である。変換の後作業として、交差部等の結合処理や整形作業が生じる。この工程の自動、または半自動化を図ることが全体的な作業効率の向上につながる。

（3）最適な変換設定

地形図上における地物の存在密度等により、自動変換時の最適値は異なる。これに関しては、複数の異なる図郭を対象に検証を行って設定値ごとの結果を比較することで、ある程度の基準となる設定値をストックしておく必要がある。

（4）パターン認識の採用

本検証で採用した、ラスター全体に対して、一律に演算処理を行う方法では、取得可能項目の細分化には限界がある。パターン認識を処理に取り入れることで、記号化された図形要素の取得を効率化できると期待される。道路を例に取ると、幅員ごとに区別してのデータ取得、ならびに建設中道路の認識による区別が可能と考えられる。

（5）フィルタ機能の追加

今回使用したソフトウェアでは、微小図形要素の除去機能について、ピクセル数の下限のみが設定可能であるが、上限も設定可能とすることで利便性が高まると考えられる。たとえば、ラスター上の図形を500ピクセル未満、500～3000ピクセル、3000ピクセル以上の3種に分類して、それぞれ除去、手作業、自動変換と、異なる処理を行うことが可能である。

３．データの更新、および提供手法の検討

25,000レベルGIS基盤情報上のすべての地物は、その存在意義に関連した属性を保有する。属性は、幾何（空間）・主題・時間の３要素より構成されている。

データの更新とは、それぞれのデータが、その存在意義に関連して持つ属性要素に対して、新しく変更を行うことを指す。

本章では、GIS基盤情報の存在意義を考慮した上での、効率的なデータ更新管理方法の検討を行った。

3.1　更新データ管理について

更新データの管理とは、具体的に言えば、データをバージョン化して管理することである。バージョン化とは、あるデータの変更時点ごとの属性を、ひとつの属性として保有することであり、変更時点ごとにデータの複製を取ることとは異なる。

バージョン化の単位には、大きく分けて２通りが考えられる。１つは、データを一定の集合体ごとに管理する方式である。集合体としては、データの作成単位をそのまま適用するのが一般的である。もう１つは、実際に変更を行ったオブジェクトごとに管理する方式である。

１）集合体でのデータ管理

空間データの場合、データ作成時には、各図葉をひとつの単位として管理していることが多い。したがって、この場合の集合体とは、それぞれの図葉ということになる。

図葉をひとつのオブジェクトとみなす場合には、各図葉は、図葉内に存在する各地物に対する変更時点ごとの属性を保有することになるが、現状では、各変更時点のデータを、バックアップとして保管する方式が一般的である。

この方式では、データの管理単位が初期データの作成単位と等しいため、地物が分割された場合等を除き、更新時にデータが細分化されない。したがって、更新時に、地物間の関係維持に特別な注意を払う必要がない点がメリットである。

しかし、特別に管理を行わない限り、「いつ、どの地物を更新したか」という記録が残らない。そのため、更新履歴の把握が困難である。

２）差分によるデータ管理

全域、もしくは図葉単位で更新を行う方式と対照的なものとして、差分ごとにデータ管理を行う方式がある。たとえば、ある道路において線形が変化した場合に、変化した部分のみを、独立したデータとして管理する方式である。

この場合のデータ管理単位としては、セグメント（図形を構成する辺）単位と地物単位の、２通りが考

えられる。

（1）セグメント単位での管理

セグメント単位でデータを管理する場合は、実際に形状変化が生じた部分のみを、更新データとして扱うことになる。この方式のメリットは、更新データの量を最小限にできることである。

その反面、データを地物以下に細かく分割するため、データ構造が壊れやすく、以後のメンテナンス作業が煩雑化しやすい。GIS基盤情報においては、作成したデータの保有している構造を、更新後もなお維持する必要があるため、特別な注意を払う必要がある。

（2）地物単位での管理

地物単位でデータを管理する場合は、変更が発生したセグメントを含む地物自体を、更新データとして扱う。そのため、データの管理単位は、セグメント単位の場合と比べると大きくなる。

この方式の場合、データを地物以下に細かく分割することはないので、データ構造は壊れにくい。また、データ仕様上、各種の属性要素は地物単位で保有されるため、地物をひとつのオブジェクトとして取り扱う上でも好都合である。

これらの点を考慮して、本考察では、「差分」とは各々の地物を指すものとして考える。

差分単位でのデータ管理の流れを、図４に示す。更新により生じた「消滅データ」および「発生データ」は、それぞれ、履歴として保管される。

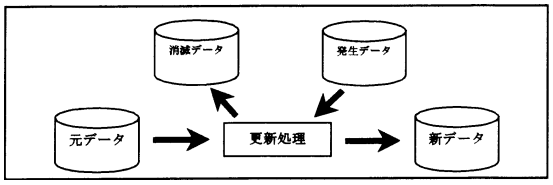

図４　差分単位でのデータ管理の流れ

3.2　差分データの消滅・発生

差分単位でのデータ管理を行う場合、更新時点ごとに生じる差分を更新履歴として有効に利用するため、地物の消滅・発生、及び実際にデータの更新を行った時間を記録しておく必要がある。これらの時間を、地物の時間属性として記録する。

消滅・発生の時間については２通りの考え方がある。１つは、現実世界に存在する地物自体の消滅・発生現象が発生した時間を、そのままデータの消滅・発生時間として記録する考え方である。もう１つは、更新システム上で消滅・発生作業を行った時間を記録する考え方である。

GIS基盤情報に限らず、空間データ作成の目的は、現実世界のデータ化であることから、現実世界における現象の発生時間を記録するのが妥当と考える。しかし、厳密にそれらの時間を把握することは、特別に観測を行わない限り不可能と想定される。したがって、実際の運用においては、消滅・発生が「確認」された時間を取得することになる。

更新によるデータの消滅、および発生の単純な例として、図５に、地物同士が併合された場合における、時系列的変化を示す。

図５において、t1は現実の地物が消滅・発生した瞬間、t2は更新システムにおいて実際に更新が行われた瞬間を示す。実線の矢印は、更新システム上でのデータの存在期間、点線の矢印は、現実世界上でのデータ（地物）の存在期間を示す。

この２つの時間の関係は、t1≦t2が成り立つ。これは、更新が、必ず当該地物の消滅・発生以後に行われることを示す。また、データの統合及び分割におい

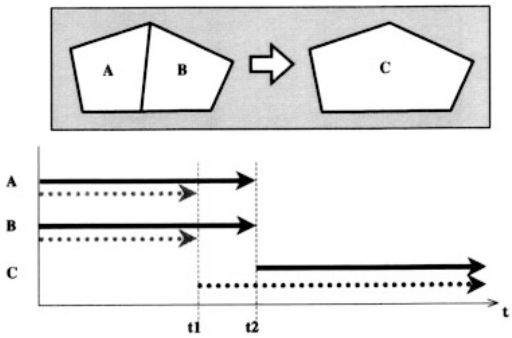

図５　データの時系列的変化

ては、元となる地物の消滅時間と、新たな地物の発生時間は必ず等しくなる。この時間が異なる場合は、単に「地物の消滅以後に、別の地物が発生した」ということであり、統合、または分割とは異なる。

消滅・発生時間、及び更新時間を時間属性として扱うことで、地物内部には「(当該地物の) 現実世界上の動きに関する時間」と、「更新システム上で処理が行われた時間」の２種類の時間を持つことになる。なお、ここで言う「時間」とは、あくまで地物の状態に関わるものを指している。

したがって、特に地物の状態と関係のない日付項目(たとえば、登記日など) を属性として付与する必要がある場合は、時間属性としてではなく、主題属性として取り扱うべきである。

3.3　データ提供について

更新したデータを効率よくユーザに提供を行うには、その内容を明確に定めておく必要がある。ここでは、提供体制、ならびに提供単位について、考察を行った。

１）データの提供体制

更新済みデータをユーザに提供する間隔は、２通りの方式が考えられる。１つは、更新ごとに随時提供を行う方式である。この場合には、ユーザは常にデータを最新に近い状況に保つことが可能である。もう１つは、従来の地形図と同様、定期的な提供に加えて記載内容に重大な変更が加えられた時点で、差分の集合を提供する方式である。

データの提供者側は、ユーザのニーズに対して、できる限り広範囲で対応できる体制を確保しておくことが重要である。したがって、更新ごとの差分提供にも修正時の差分提供にも、対処できる提供体制を確保する必要がある。

データ刊行から改版までを１サイクルとみなした場合、更新に誤りがあった場合を除き、更新時に提供される差分データの和は、修正時に提供される差分量と等しくなる。各ユーザは、自らが運用するシステムで必要とされるデータの鮮度に応じた形で、更新を行うことが可能である。

ユーザ側の運用システムは、用途や重要とされる機能等によって様々であり、データのフォーマットも異なる。このようなデータの交換の場合には、あらかじめ標準フォーマットを定めておき、異なるシステム同士のやり取りは、標準フォーマットを通して行う方式が広く採用されている。

GIS基盤情報においても、各々の属性要素に対する標準フォーマットが定義されている。標準フォーマット以後の変換及び取り込み処理は、運用システム側で対処する。

２）データの提供方式

データの提供方式を、その提供単位から考えた場合、更新及び修正により発生した差分のみを提供する方式と、差分を含む図葉のデータを提供する方式の２通りが考えられる。前者の場合、具体的な更新内容を、データ内に持たせる必要がある。後者の場合は、単純にデータを上書きして差し替えればよい。後者の方式は、2,500レベルGIS基盤情報（空間データ基盤）における更新データ提供に採用されている。

差分のみ提供する場合について、提供する差分情報の内部構造を考察する。差分にはデータ更新の内容により、それぞれ追加・削除・変更に関する情報が含まれる。

差分情報の内部構造についてのイメージを、図６に示す。差分情報自体を１つのオブジェクトと見なして、追加（add)，削除（delete)，変更（change）の各種クラスを設けている。

ユーザ側の運用システムに、差分情報を読み込む際は、更新内容に関する情報も、あわせて読み込む必要がある。したがって、データ提供時に差分情報の構造を公開するか、読み込み時にモジュール的な役割を果たす機能を、差分情報内部クラスの１つとして持たせる必要がある。

モジュール的機能とは、差分情報のユーザ側システ

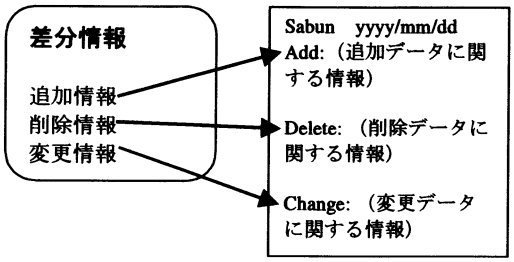

図６　差分情報の内部構造イメージ

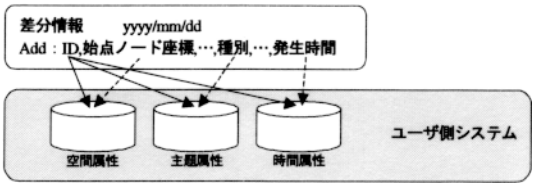

図７　ユーザ側システムでの追加処理

ム読み込みに、ユーザ側システム内における対象データの各属性要素に対して適切な更新を行うための機能である。追加・削除・変更のそれぞれについて、その概念を以下に記す。

（１）新規地物の追加に関する情報

はじめに、差分情報内に含まれる地物のIDが、既存のデータ内に存在していないことを確認する。同一のIDが存在する場合、地物の新規追加とは呼べないためである。属性の一部のみ追加（たとえば、名称の追加など）を行うのは、新規追加ではなく属性内容の変更に該当する。

確認の結果、同一のIDが存在しなければ、追加の対象となる属性項目をユーザ側システム内のデータに追加する。

図７に、ユーザ側システムでの追加処理イメージを示す。

（２）地物の削除に関する情報

はじめに、差分情報内に含まれる地物のIDが、既存のデータ内に存在していることを確認する。同一のIDが存在しない場合は、提供側とユーザとでデータの内容に不一致が生じていることになる。この場合、不一致が存在する旨をユーザに知らせる必要がある。また、属性の一部のみ削除を行うのは、削除ではなく属性内容の変更に該当する。

確認の結果、該当のIDが存在すれば、対象となる属性項目をユーザ側システム内のデータより削除する。該当のIDを持つ地物の全属性要素が削除対象となるため、個々の属性項目に関する記載は不要である。

ただし、当該データを抹消してしまうのは、ユーザ側での利用形態によっては問題となる。したがって、削除フラグを設けて、更新時には対象データのフラグをONにすることで、見かけ上の削除を行うのが適切である。

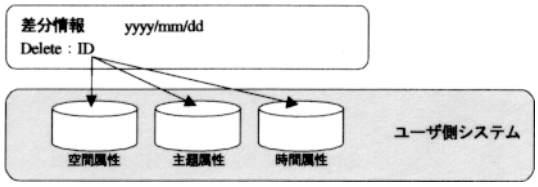

図８　ユーザ側システムでの削除処理

図８に、ユーザ側システムでの削除処理イメージを示す。

定時修正の際、削除対象分をすべて抹消したデータがユーザに提供される。ユーザにおいて、該当分が不要であれば、データの差し替えを行えばよい。

（３）地物の属性内容変更に関する情報

はじめに、差分情報内に含まれる地物のIDが、既存のデータ内に存在していることを確認する。同一のIDが存在しない場合の対処に関しては、（２）と同様である。

確認の結果、該当のIDが存在すれば、変更の対象となる属性項目をユーザ側システム内のデータに上書きする。差分情報に含まれていない属性項目は、更新を行わない。

図９に、ユーザ側システムでの変更処理イメージを示す。

3.4　まとめ

１）地物の更新履歴管理について

本考察では、地物を、単独で存在する図形要素とみなして、更新の考察を行った。しかし、実際には、それぞれの地物が単独で成立していることは少なく、複数の地物が密接に関連づけられていることが多い。

地物どうしの関連まで考慮に入れ検討を行うには、データの応用スキーマ（当該データの内容や構造を、定められた記法で明確に記述したもの）を正しく定義する必要がある。この定義は、データの提供時に、異

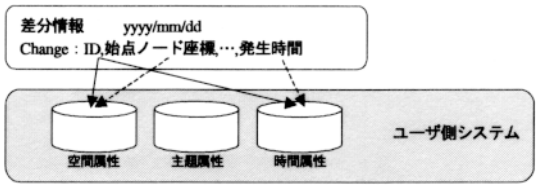

図９　ユーザ側システムでの変更処理

なるシステム間で正しくデータの交換を行うためにも必要なものである。

２）データの提供について

更新済みデータの提供に関しては、提供側の管理システムにおける迅速な更新体制と、ユーザへの確実な差分情報の提供が重要である。

今回はデータの流れに的を絞って検討を行ったが、実際に運用の可能性を評価するためには、今回の検討結果を反映した機能を備えたツールの、試験的な実装に基づく運用テストが必要であろう。運用の試行を通して、発生が予想される課題を抽出し、それぞれに対策を立てる必要がある。

４．所感

本研究では、ベクトルデータ作成と、データの更新・提供の２項目に対して考察及び検証を行った。これらは、決して互いに無関係な要素ではなく、25,000レベルGIS基盤情報の効率的な整備、および運用のために、双方とも不可欠なものである。

今後、さらに進展した形での研究が行われることを期待する。

謝辞

本研究を進めるにあたり、吉成課長補佐をはじめ、地図部地図情報課の皆様には、業務ご多忙中にもかかわらず、貴重な指導、資料・設備の提供をいただき、大変お世話になりました。ここに、あらためて感謝の意を表します。

参考文献

１）坂内正夫・大沢裕：「画像データベース」（株）昭晃堂、1987

２）石塚圭樹：「オブジェクト指向データベース」アスキー出版局、1996

３）建設省国土地理院：「2万5千分の1レベル基本情報に関する調査・研究作業報告書」、1999

４）建設省国土地理院：「GIS基盤情報の時系列差分情報管理手法に関する研究報告書」、2000

（国際航業株式会社）