

I-1) 背景情報処理を導入した図面情報数値化システムの構築

DEVELOPMENT OF DIGITAL DATA PROCESSING SYSTEM FOR MAP
INTRODUCING BACKGROUND INFORMATION PROCESSING桑原 祐史*・野北 舜介**
Y. Kuwahara and S. Nogita

抄録：本研究では、各種図面情報の数値化作業の効率化を目指し、スキャナを用いた数値化システムを構築した。画像データとしてコンピュータに取り込まれた図面情報は、凡例項目毎に指定された色情報に加えて、各種の背景情報を有している。本研究では、凡例項目毎に着色された領域以外の情報（等高線、構造物の記号等）を背景情報と定義し、画像の閾値処理により背景情報を抽出した。作成された基準データを基にして、画像平滑化処理に工夫を加えた背景情報低減処理を提案した。領域の輪郭線の抽出処理に際しては微分処理を、凡例項目毎に着色された領域の抽出時には2次元ヒストグラムを使用し、繰り返し試行を容易に進めることができるシステムを構築した。本システムは、背景情報抽出・低減サブシステムおよび領域抽出サブシステムの2つのサブシステムから構成される。一般に販売されている土地利用図を対象とし、実際に数値化処理を進めた結果、効率的な数値化作業が行えることが判った。

Abstract: A data processing system is proposed in this paper for obtaining geographic information from a map, the information of which is gathered by an image scanner. As the scanner collects all types of background data, in the first step, the proposed system purifies the image of map, that is, decreases the influence of background data (All characters, lines, and symbols are defined as background data in this paper.) by applying the threshold method using 1 or 2-dimensional histogram. Then, the coloured regions with minimized background data are classified by applying the threshold method. The application study to a land use map shows that the proposed system improves the processing time considerably.

キーワード：地理情報、画像データ、地理情報システム

Keywords: geographic information, image data, geographic information system

1. はじめに

近年、官庁や企業等では地理情報システム (GIS : Geographic Information System) を導入し、様々な計画策定作業の効率化を図っている。目的とする主題図および評価図の作成に際しては、予め数値化された各種のデータが必要となる。現在、国土地理院で販売されている数値地形モデルや官公庁で数値化された各種の情報を必要に応じて購入・収集し、解析者の目的とする主題に応じた数値情報を加えて様々な分野における研究事例が数多く公表されている^{1)~2)}。既に図面化された情報をラスタ形式で数値化する際には、①原図を目的とするメッシュサイズで分割し属性情報を読み取る、②ディジタイザを用いて領域毎に属性を指定する、③スキャナを用いて画像として取り込み、各種デジタル処理を通して数値化する、等の方法が考えられる。この内、①および②の方法は、作業に携わる解析者の入念な確認のもとに作業が進められることから、精度の点では安定した結果を得ることができる。しかし、図面の属性情報の数や領域、さらにはメッシュサイズによってはかなりの作業時間を要する場

合もある。そこで、本研究では特に作業効率の点を考慮し、③の方法に着目した。

本研究では、まず、各種の図面で使用されている色や背景情報等の表現形態を整理し、数値化作業を進める際に対象とする項目を明確にする。次に、システム構築の前段階として、輪郭線および色分けされた閉領域の抽出に関する基礎検討を進める。具体的には、既に数値化されたデータを使用し、導入した要素技術毎に輪郭線および領域の抽出精度を把握する。最後に、一連の検討結果を基にして、パーソナルコンピュータを用いたシステムを構築し、実際の図面を対象とした作業例を整理する。

2. 研究の目的

本研究の目的は、以下の4点である。

①各種の図面は凡例項目毎に様々な色や表現方法を用いて作成されている。そこで、まず、一般に使用されている代表的な表現方法を整理するとともに、本研究で対象とする表現項目を明確にする。

②システム構築の基礎検討として、既に数値化され

*正会員 工修 茨城大学 助手 工学部 都市システム工学科
**正会員 工博 茨城大学 教授 工学部 都市システム工学科

凡例項目毎に色分けされている情報を教師データとして選定し（以下、教師データと呼ぶ）、属性情報毎に輪郭線および領域の抽出精度を確認する。

③一連の処理手順を取りまとめ、数値化アルゴリズムとして整理する。加えて、誰しもが入手しやすいハードウェアを用いてシステムを構築する。

④実際の図面を対象として数値化作業を進める。この過程において、数値化の精度を確認するとともに、追加・検討すべき今後の課題を整理する。

3. 研究開発の着目点

3.1 図面の表現形態と研究対象とする項目

凡例項目毎に色分け、着色された図面の代表例として、一般に購入することができる土地利用図（国土地理院）および現存植生図（環境庁）や、省庁、地方公共団体にて作成している管内図、都市計画図、さらに総理府令にて定められている作業規定準則に基づいた土地分類基本調査の成果図面等があげられる³⁾。各種の図面は、通常、紙面に印刷されたアナログ情報であることが多い。しかし、現在、各種計画作業の合理化および効率化を目指し、地理情報システム（GIS）の普及が進むにつれて、従来、紙面に印刷していた情報（アナログ形式）をデジタルの形で整備するケースが増えている。情報の新規更新や維持・管理に加えて、目的に応じた代替案の作成等、寄与する所は大きいと言える。このような中、本研究では、既に図面化された情報を対象として、凡例項目毎に色分けされた項目を画像処理を通じて抽出、数値の属性付けを施すことを試みた。このことから、図面に記述されている着色領域以外の情報（等高線、構造物等の記号）を本研究では背景情報と定義し、以降の検討を進めた。

図面に印刷されている各種の表現形態を取りまとめると、表-1に示す通りとなる。本研究では、作業の効率化を目指し、スキャナを用いた数値化を試みる。この場合、図面はRGB3色の画像としてコンピュータ内に取り込まれる。このことから、線や点さらにはハッチング部分の情報を取り込むことは困難である。このような図面への対処法としては、予めハッチング部分を着色し、領域を明確にすることにより着色領域と同様の取り扱いが可能となる。以上の観点から、本研究では、表-1のハッチングを付与した項目に絞り、研究を進めることとした。

3.2 本研究で指向する数値化システム

凡例項目毎に数値の属性付けをする際には、以下の処理項目が必要となる。

- ①背景情報の低減処理
- ②閉領域毎の輪郭線の抽出
- ③閉領域の抽出

凡例項目毎に着色された各種の図面は、位置を把握するために地形図を背景情報として印刷していることが多い。人間が目視判読する際には重要な情報であるが、コンピュータ処理を通して閉領域毎の輪郭線および領域を抽出する際には計算結果に悪影響（混色等）を及ぼすこととなる。そこで、②および③の処理を施すに先立ち、背景情報の低減処理を試みる。また、この研究で指向する数値化処理システムでは、閉領域毎の輪郭線の抽出処理を加えることとした。抽出された輪郭線は、閉領域の抽出を進める際に、結果の良否を判断するものとして利用する。

4. システム構築に関する基礎的検討

4.1 ハードウェア構成

ハードウェアの選定に際しては、誰しもが容易に入手することのできる一般機器を選定した。具体的に、パーソナルコンピュータは、32ビットCPUを有する機種を、スキャナは光学系移動方式のスキャナ、さらに、成果品の出力にはインクジェット方式のカラープリンタを選定した。また、スキャナで取得した画像をフルカラーでカラーディスプレイ装置に表示するため、専用フルカラーフレームボードを使用した。

4.2 導入した要素技術

(1) 領域の境界線の抽出

輪郭線の抽出に際しては、計算の対象とするデータが少なく処理速度が早いことを考慮し、式-(1)に示すRobertsの方法（微分法による）を採用した。さ

表-1 図面の代表的な表現形態

領域・位置	背景情報(地図等)	領域の表現方法と表示内容	
	図面全体的な背景	表現方法	種別を表現する背景情報
閉領域の情報	全体的背景情報無し	ハッチングによる領域分け	ハッチング内の背景情報
		塗色された領域分け	塗色域内の背景情報
	全体的背景情報有り	ハッチングによる領域分け	ハッチング内の背景情報
		塗色された領域分け	塗色域内の背景情報
点と線の情報			

$$\begin{cases} f_x = f(i, j) - f(i+1, j+1) \\ f_y = f(i+1, j) - f(i, j+1) \end{cases} \dots\dots\dots \text{式} - (1)$$

ただし、 f_x : X方向の微分
 f_y : Y方向の微分

らに、式-(2)に示す方法でエッジ強度を計算し、結果を8(bit)に量子化した値を用いて閾値処理にて輪郭線を抽出した。

$$g(i, j) = \sqrt{f_x^2 + f_y^2} \dots\dots\dots \text{式} - (2)$$

(2) 領域の抽出

図面を作成する際には、技術者が判読しやすく、かつ誤判読を防ぐこと等を目的として、以下の点が配慮されていることが多い。

- ①技術者が判読しやすい明確な色を使用する。
- ②凡例項目毎に設定する色は、彩度が異なるように配色する。

自然画、航空写真および衛星リモートセンシングデータと比較すると、図面の色区分は極めて明確である。加えて、本研究ではなるべく一連の演算処理を簡便にするという観点から、後述する2次元ヒストグラムを領域の抽出方法として採用することとした。

具体的には、図-1に示す2次元のヒストグラムを作成した。本図は、縦軸にR(赤)の画像データの濃度値を、横軸にはG(緑)の画像データの濃度値を取り、データの散布状態を図化したものである(以下、2次元ヒストグラムと呼ぶ)。黒色が濃い部分ほど該当する色のデータ数が多いことを示している。本研究では、図に示すように、縦軸および横軸毎に閾値を設定し、簡単な条件分けに該当するデータ群つまり図面に着色された領域を抽出していく。

4. 3 サンプル図面を使用した基礎検討

(1) 教師データ

本節では、輪郭線と2次元ヒストグラムを用いた領域抽出、両精度を確認することを目的に検討した結果を述べる。図-2に教師データを示す。データサイズは横方向320(pixel)、縦方向320(pixel)であり、6色でカラー印刷されている。このデータを解像度300(dpi)のスキヤナを用いてピクセルサイズが多少大きくなるように取得した。取得された教師データは、8(bit)で表現されるRGB各色成分より構成される画像データとしてコンピュータに記憶される。

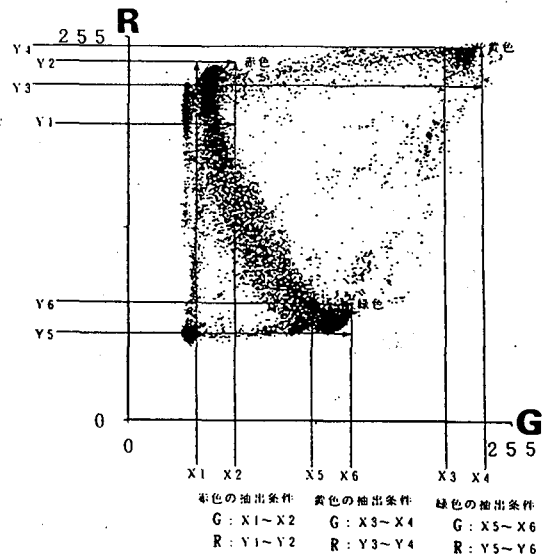


図-1 2次元ヒストグラムと閾値設定の概念

(2) 境界線の抽出とその精度

作成した教師データを対象として、Robertsのオペレータを用いて輪郭線を抽出した。具体的には、RGB各色データの有する情報を最大限に生かすことを目的として、データ毎に輪郭線の抽出処理を施し、個別の処理結果を統合した。また、輪郭線の強度の計算時には式-(2)に係数を乗じて(本検討では一律5とした)値の変動を強調した。以上の処理の後に、強度分布をヒストグラムを用いて確認し、閾値を設定、2値化処理を通して輪郭線を抽出した(輪郭線データと呼ぶ)。図-3に抽出結果を示す。教師データの背景は「赤」で表現されている。図に示す輪郭線のうち、領域が「橙」で表現されている領域に対する個所は、局所的に非連結個所が見受けられた。作業に際しては、係数の値を工夫する必要性が伺えた。本検討における処理結果の精度を検証するため、教師データより輪郭線をトレース、スキヤナで数値化したデータ(輪郭線教師データと呼ぶ)との適合率を計算した。なお、処理に先立ち、輪郭線データに対して内挿処理(最近隣内挿)を施し、輪郭線教師データとのサイズを統一した。適合率を計算した結果、約71(%)となった。

(3) 領域の抽出とその精度

RGB各色毎のデータを2つずつ組み合わせ、2次元ヒストグラムを作成した。領域の抽出に際しては、凡例項目1つに対して縦軸および横軸の閾値を二つずつ設定する必要がある。また、処理作業は凡例項目の数だけ繰り返す必要がある。図-4に抽出結果を示す。教師データと、内挿処理(最近隣内挿法)を用いてデ

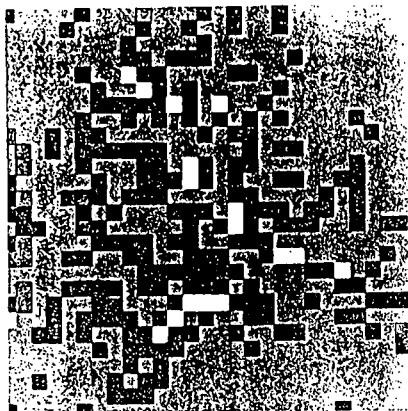


図-2 教師データ

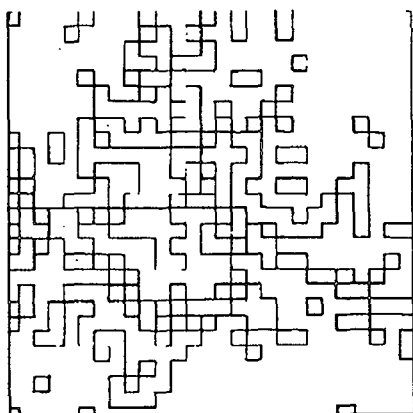


図-3 輪郭線の抽出結果

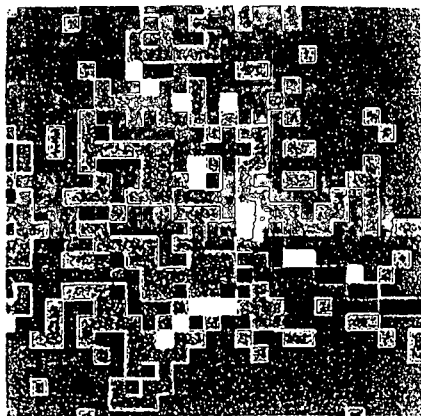


図-4 領域の抽出結果

ータサイズを合わせた抽出結果との適合率を計算した結果、約89(%)を得た。図より明らかなように、非適合領域の大部分は、隣接する色間のエッジ部分となっていることが判った。ピクセル数にすると、約1~2(pixel)のずれとなる。この点については、システム構築時にマウスで調整する機能を付与し、対応していく必要がある。

5. 数値化システムの構築

5.1 システムの要件

前節までの検討結果に基づき、実際の地図や図面の数値化システムを構築していく際には、以下の5点の機能を持つことが望ましい。これらの点を、システム構築の要件とした。

- ①汎用性を考慮し、広く一般に利用されているハードウェアを使用してシステムを構築する。
- ②背景情報を抽出し、低減する機能を持たせる。
- ③凡例項目毎の領域抽出作業過程における繰り返し試行検討が容易なシステム構成とする。
- ④作成された数値データの詳細部分を修正することができる機能を持たせる。

⑤データの内挿処理を導入し、様々なピクセルサイズに対応した数値データを作成する機能を持たせる。

なお、本論文では紙面の都合上、構築したシステムの画面構成については割愛する。

5.2 システムの構成

図-5にシステムの全体構成を示す。構築したシステムでは、スキャナを用いて取得された図面や地図情

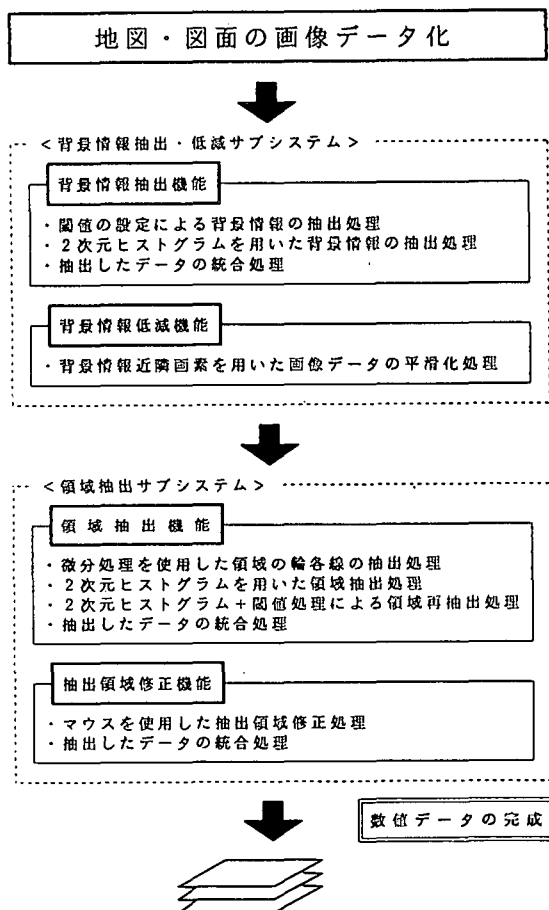


図-5 図面情報数値化システムの全体構成

報のRGBで構成される画像データを入力データとする。最終的には、凡例項目毎に数字で属性付けされた数値データが出力結果となる。図に示すように、システムは「背景情報抽出・低減サブシステム」および「領域抽出サブシステム」の2つのサブシステムから構成される。図面や地図の種別によっては、必ずしも背景情報の抽出と低減処理を要しない場合もある。この点を全体構成作成時に考慮している。以下、サブシステム毎に、処理機能の詳細を含めて説明を進める。

(1) 背景情報抽出・低減サブシステム

本サブシステムでは、情報の抽出と低減の2つの機能を整備した。図-6にシステムにおける処理の流れを示す。以下、各機能毎にその詳細を説明する。

①背景情報抽出機能

各種の図面や地図の背景情報は、明確な黒色や薄い色(茶色等)で印刷されていることが多い。本機能では、背景色が「黒」の場合と「その他」の場合をメニュー選択形式で選定できるようにした。具体的に、背景色が「黒」の場合には画像データの閾値処理、「その他」の場合には2次元ヒストグラムを用いた閾値処理にて背景情報を抽出する。また、背景色が複数に跨る場合もあることから、抽出した個別の背景情報を統合する処理項目についても整備した。

②背景情報低減機能

背景情報抽出機能で作成された背景情報を入力データとして、図面や地図の画像データの背景情報を低減する。具体的には、図-7に示す平滑化処理を施す。図の中心に示すXの画素は、背景情報として抽出された画素である。この背景画素を中心とした5×5のウィンド内で、背景情報以外の画素を抽出し、その平均値で中心画素Xを置き換える処理を施した。ウィンドサイズの設定問題さらには背景情報の幾何構造を考慮に入れたウィンド形状の工夫等、処理効果の向上についてはさらに検討の余地があるが、第一次の改善は達成されたと考えられる。

(2) 領域抽出サブシステム

領域抽出サブシステムでは、領域の抽出と修正の2つの機能を整備した。図-8にシステムにおける処理の流れを示す。以下、各機能毎にその詳細を説明する。

①領域抽出機能

ここでは、まず微分処理(Robertsのオペレータ)を導入した輪郭線の抽出処理を進める。閾値の設定を

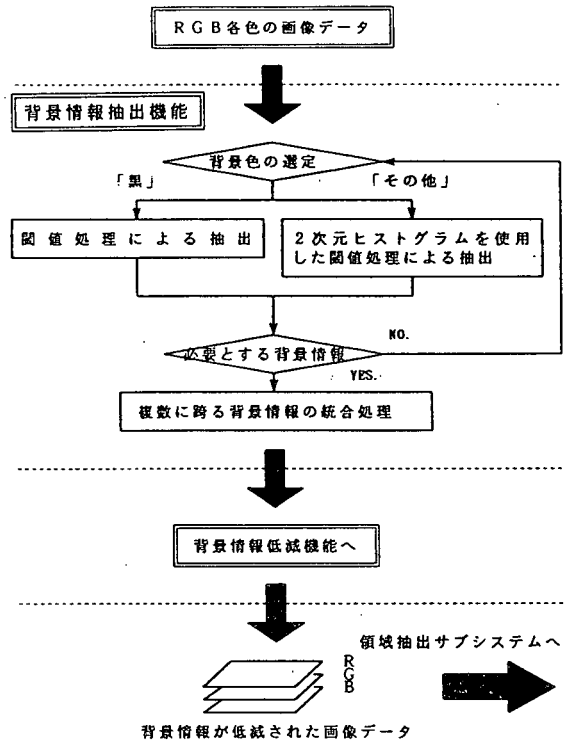
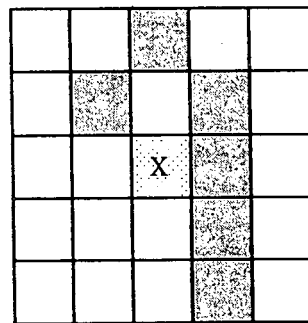


図-6 背景情報抽出・低減サブシステム



■: 背景情報として抽出された画素
 着目画素の周辺24画素に注目し、背景情報として抽出されていない画素の平均値を中心画素とする。

図-7 背景情報低減処理の概念

通して輪郭線が抽出される。本サブシステムでは、「画像のヒストグラムの表示→閾値の選定処理」が繰り返し進められるように工夫している。次に、2次元ヒストグラムを用いて凡例項目毎に着色されている色領域を抽出する。処理の過程では、「画像の組み合わせ(R-G、R-B、G-B)→2次元ヒストグラムの表示→閾値の設定→抽出領域の表示」を繰り返し行い易いように工夫した。橙色等の中間色の場合には、

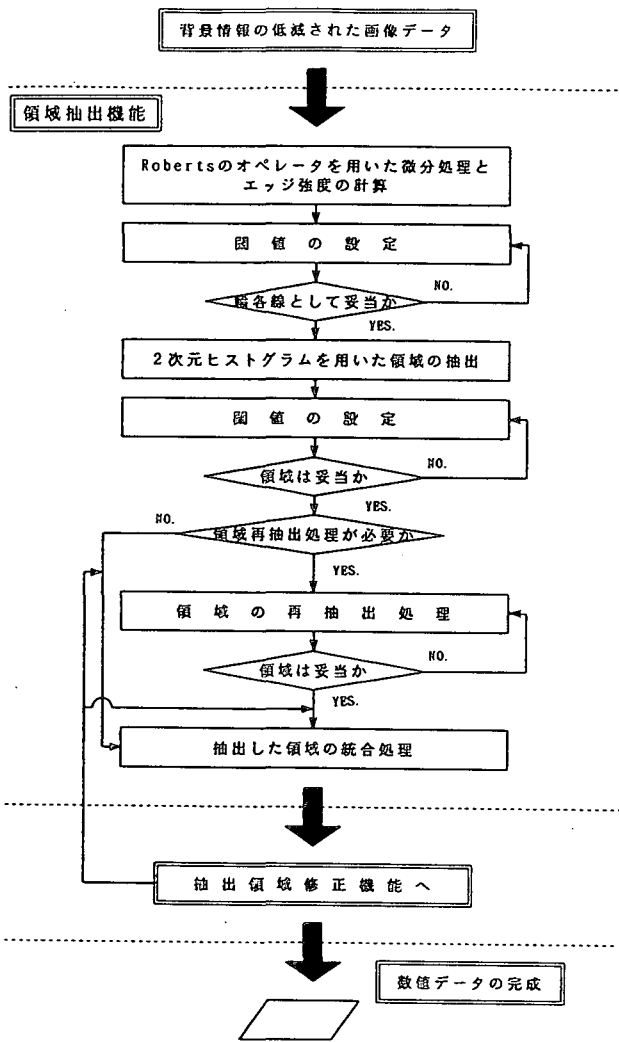


図-8 領域抽出サブシステム

必ずしも2つの画像データの組み合わせでは領域を抽出できないこともある。図-8の下部に示す領域の再抽出処理は、このような場合の処理過程である。具体的には、2次元ヒストグラムを用いて抽出された領域に該当する残りの画像データのヒストグラムを参照し、閾値の設定を変えて領域を再抽出する。この処理は、画像データを3次元的に捉える処理に該当する。本システムでは、データ分布と閾値を容易に把握・設定できることを考慮し、2次元表現を採用した。

②抽出領域修正機能

本処理機能では、マウスを用いて抽出領域の修正処理を施すことを可能とした。背景情報の完全な除去は困難であることから、背景情報周辺では誤った領域抽出が生じ易い。また、凡例項目毎に設定された色が類似している場合やスキャナによる画像取得精度の問題も生じる。このような場合に柔軟に対処できるように、

本機能を付加することとした。具体的には、画像データと抽出領域をオーバーレイ表示し、マウスで指定した地点を任意のデータに変換または除去できるようにした。

6. 実際の図面を対象とした検討

6.1 図面の選定

本研究では、以下の4点を考慮し、図-9に示す国土地理院発行1/25,000土地利用図を選定した。

- ①背景色が複数（黒色、薄茶色）に跨っている。
- ②加えて、背景情報の種類（建物、鉄道、公共機関の表示等）が豊富であり、背景情報の低減処理効果を論じる上で妥当な図面、領域である。
- ③凡例項目数が豊富であり、類似した着色領域が散在している。色毎の領域の抽出効果について検討する



図-9 選定した図面



図-10 スキャナで取得した土地利用図

上で十分な検討ケースを設定することができる。

④図面が一般的なものであり、本検討結果を実際の作業過程に反映することが可能である。

6. 2 処理結果

図-10にスキャナで取得した土地利用図を示す。この画像を対象として、以降の数値化処理を進めた。

(1) 背景情報抽出・低減サブシステムによる作業

①背景情報抽出処理

図-11(1)および図-11(2)に背景情報の抽出結果を示す。本出力は、RとGおよびGとBの画像データを使用した各2次元ヒストグラムを作成し、閾値処理を通して抽出したものである。(1)では主として黒色の背景情報(林道やJ Rの線、住宅地等の記号)が、(2)では主として薄茶色の背景情報(主として等高線)が抽出されている。黒色で表現されている背景情報は比較的多数抽出されているが、薄茶色で表現されている背景情報(主として等高線)は必ずしも明確に抽出されていない箇所も見受けられる。着色領域と背景情報との混色が大きく影響している。本検討では、この2つの出力結果を統合し、背景情報低減処理における基準データとした。

②背景情報低減処理

図-12に、前項で作成した基準データを基にし、背景情報の低減処理を施した結果を示す。図-10に示す元データと比較すると、鉄道、構造物等の背景情報が低減されていることが判る。低減処理方法の一層の工夫によりさらに高い低減効果が得られることも考えられるが、本図面については相当の低減効果が得られたと解釈できる。本画像を用い、領域の抽出処理へと作業を展開した。

(2) 領域抽出サブシステムによる作業

①領域抽出処理

図-12に示す画像データを用いて領域の抽出作業を進めた。凡例項目として「9項目」を設定した。作業の過程では、着色領域の判断に使用する輪郭線の抽出処理や、2次元ヒストグラムを作成する際の色の組み合わせと閾値選定作業等、煩雑な繰り返し処理が生じる。この点については、メニュー選択形式で容易に繰り返し試行ができるようにシステム構築時に工夫した。繰り返し試行を進めた結果、表-2に示す閾値を設定、後述する修正処理の前段階までの作業時間は約1時間半程度であった。なお、薄橙色の抽出に際して

は領域の再抽出処理を施した。この処理に際しては、RGB3原色の組み合わせによる色表現に関する事前知識の有無が作業時間に大きく影響する。この点については、別途色表示の表を準備しておく等の対応を取ることが有効であろう。



(1) R-Gデータ (2) G-Bデータ

図-11 背景情報の抽出結果



図-12 背景情報の低減処理結果

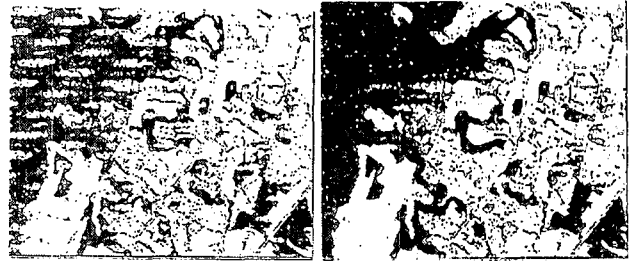
表-2 閾値の設定結果

色項目	属性番号	設定した閾値		
		R	G	B
青	1	150~50	—	250~150
緑	2	200~150	255~190	
橙	3	250~220	130~100	—
茶	4	250~230	240~220	—
ピンク	5	254~220	—	254~220
黄	6	255~240	255~240	—
灰	7	230~210	220~190	—
薄 橙	8	註) 245	250~200	150~100
濃 橙	9	—	200~150	100~50

註): 領域の再抽出処理を施した。

②領域修正処理

図-13(1)に修正前のデータ(領域抽出処理結果)を、図-13(2)に修正後のデータを示す。図-13(1)の左側における横長な楕円形状のデータ欠落領域は、スキャナによる画像取得時の“濃度ムラ”に起因している。このような領域を修正する際に、本システムでは画像データに着色領域をオーバーレイ表示した画面上にてマウスで位置を指定、着色または除去処理を容易に進めることができる。本処理を通じて成果品としての数値データが作成される。



(1) 修正前のデータ (2) 修正後のデータ

図-13 領域修正機能による処理

(3) 最終成果品としての数値データ

図-14に最終成果品としての数値データを示す。本図では、抽出領域を明確に示すことを目的として、土地利用図における着色と若干異なる色表現をしている。図の中心部分に位置するピンクで表現した市街地については、領域の抽出結果に問題が残された。この凡例項目の着色個所については、背景情報が多数重なっており、抽出結果が不安定になる原因があった。一方、青で表現した工場領域は、背景情報の低減効果が高く良好な抽出結果になっていることが確認された。



図-14 数値化データ(最終成果品)

作成した数値データの加工と利用に際しては、以下に示す2つのニーズが考えられる。

①図面に着色された全凡例項目を必要とする。

②利用目的に必要な凡例項目のみ必要とする。

このようなニーズに対しては、必要とする凡例項目に付与されている数値属性値のみ抽出する処理は容易である。また、要求されるピクセルサイズは多様である。本成果品データは、1ピクセルあたり約8(m)のサイズを有する詳細な数値化データである。このようなニーズに対しては、内挿処理(最近隣内挿法)を用いて数値データを加工、対応していくことが可能である。

7. 研究の成果と今後の課題

7.1 研究の成果

本研究の成果は、以下の3点である。

①既に数値化されたデータを対象とした検討を通して、導入した要素技術の精度確認を行った。検討結果に基づき、図面情報数値化システムを構築していく要件を整理した。

②図面情報数値化システムを構成するサブシステムとして「背景情報抽出・低減サブシステム」および「領域抽出サブシステム」を設定し、各々に含まれる

機能を実現するソフトウェアと処理手順を整備した。

③市販されている図面を対象として、構築したシステムを使用した数値化作業を進めた。検討の結果、色分布に着目した領域の抽出処理と修正作業を通して数値化作業の効率が向上することを確認した。

7.2 今後の課題

今後の課題は以下の2点である。

①背景情報の低減処理に関して、背景情報の分布状態を考慮に入れたウィンド形状の工夫をする。

②RGB3原色の組み合わせによる色表現に関する情報をヘルプ機能としてシステムに組み込み、閾値設定を進める上での作業効率の向上を目指す。

【参考文献】

- 1)小島尚人・大林成行・村上信・内堀富喜：土地分級評価支援システムの構築とデータセット管理について、土木情報システム論文集Vol. 3, pp. 15~22, 1994年
- 2)町田聡・大池浩司：触って覚えるGIS、山海堂、1995年
- 3)東京都・国土庁：土地分類基本調査(秩父、五日市、三峰、丹波)、1994年