

地図情報の要素間の自動関連付け技術とその共有に関する研究

大谷英之

A Study on Automatic Association of GIS Data Elements and Data Processing Platform for Technology Sharing Hideyuki O-TANI

Abstract: Large-scale disaster simulations for earthquake and tsunami have been developed in the field of high performance computing, and the product programs require the building-resolved models about a target city. However, data resources for such models are limited and fragmented, and many types of them are originally designed to be understood with human eyes, lack the descriptions of relationships between data elements, and thus we need to compensate the relationships automatically to integrate a large-scale urban model. In this study, we have developed a method to associate GIS data elements with each other based on their locations, and proposed a data processing platform to share the technology for such data integration.

Keywords: 大規模災害シミュレーション (large-scale disaster simulation), 自動関連付け (automatic association), データ処理プラットフォーム (data processing platform)

1. はじめに

スーパーコンピュータ「京」に代表される大型並列計算機により、地震・津波などの災害のシミュレーションが大規模化・高解像度化され、建物一棟一棟まで詳細にモデル化された都市モデルに対して実行可能となっている。

このシミュレーションの入力として必要な都市モデルの構築には、シミュレーションの大規模化・高解像度化に応じた広域かつ詳細な都市の情報が求められ、また、膨大な数の建物に対応するため自動化が重要である。しかし、有力な情報源となる2次元及び3次元の地図情報の多くは、線分や多角形面、文字列などの地図情報の要素の配置を人間が目で見て理解するものであり、要素間の関連付けをデータとして持たない断片として記録されているためにシミュレーションのモデルの情報源として直接利用することができない。

大谷英之

〒650-0047 兵庫県神戸市中央区港島南町 7-1-26

国立研究開発法人 理化学研究所 計算科学研究機構

Phone: 078-940-5781

E-mail: h.o-tani@riken.jp

著者ら (2014) は、地図情報の断片化された要素の間の関連付けを位置情報に基づいて自動構築する技術の一例として、市販の3D地図データに記録された構造物の3D形状と、行政データに記録された構造物の構造種別や建築年等の属性情報を含む構造物個々のデータとを自動的に関連付ける手法・プログラムを開発しているが、特定の地図会社のデータ、特定都市の行政データに合わせたプログラムとなっているために、手法自体は他都市に適用可能であっても、プログラムは他都市でのモデル構築で共有・再利用可能となっていない。

そこで、本研究では地図情報の断片化された要素の間の関連付けを位置情報に基づいて自動構築する技術の開発とともに、この技術に基づくプログラムを効率的に共有・再利用するためのシステムの開発を目的とする。

本稿では、地図情報の要素間の自動関連付けの一例として、2.に、地番参考図を構成する線分データと文字列データから地番情報付きの多角形データを抽出する手法を提示する。また、3.では、異種データの本質的でない差異を吸収してプログラムの共有・再利用性を高めるデータ処理

プラットフォームの概要について報告する。

2. 地番参考図の機械判読技術

地番参考図は、図-1(上)のように、土地の識別のためにつけられた地番と土地区画の関係が地図として表現されたものであり、緯度・経度の情報を直接もたない行政データに位置情報を付加するために有用な情報となるが、多くの場合には、人間が目で見て理解するための図画情報として記録されているため、地番と土地区画を自動的に関連付けて抽出することは容易ではない。

本節では、地番参考図を構成する線分データと文字列データをそれぞれ抽出して、線分データから土地区画ポリゴンを構築し、さらに、引き出し線データの位置情報に基づいて地番文字列と土地区画を関連付ける手法を提示する。

2.1 線分データと文字列データの抽出

PDF データとして記録された地番参考図からの線分データの抽出は、フリーソフトウェアである pdf2svg によって PDF データを XML で記述されたベクターイメージ用画像形式である SVG データへ変換することで行う。これにより、線種や線の太さ等の属性情報を含む線分データが得られる。また、文字列データは、市販のソフトウェアである PDFlib TET を利用することで、PDF データから文字列とその位置と傾き、文字列を囲む文字枠の大きさが抽出可能である。

2.2 土地区画多角形データの構築

2.1 で抽出された線分データからの土地の区画を表す多角形データの構築は次の手順で行う。

- (1) 線分データの細分割
- (2) 端点が孤立した線分の削除
- (3) 左右両側が同一の区画である線分の削除
- (4) 多角形の抽出
- (5) 穴の開いた多角形の構成

(1)の処理では、一定距離以内にある点を同一視するとともに、図-2 のように、線分が交差する点ですべての線分を分割する。この処理により、土



図-1 地図情報の要素間の関連付けの例：地番参考図では、土地区画と地番が関連付けられる。

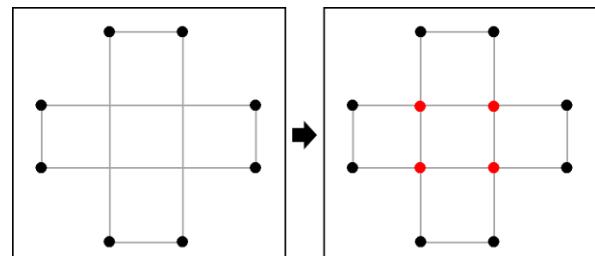


図-2 線分データの細分割の例。線分が交差する点（赤点の位置）ですべての線分を分割する。ただし、黒点は再分割前の端点を表す。



図-3 端点が孤立している線分の例（左）と左右両側が同一の区画である線分の例

地区画の多角形を構成する辺の候補が生成される。また、各線分の端点に接続している他の線分の点が明確となる。

(2)及び(3)の処理では、図-3に示す土地区画の辺とならない線分の削除を行う。(2)の処理では、特に他の線分と接続されず孤立している端点（孤立端点）を削除する。ただし、この削除により孤立端点が新たに生じることがあることに注意する。(2)までの処理により、図-4に示すように隣接する線分を左手に領域をみるように辿ることで多角形を構成することが可能となっている。(3)の処理は、左右両側が同一の区画にある線分の削除であるが、これは構成した多角形分に同じ線分が二度含まれている場合に、その線分を削除することで可能である。

(4)の処理では、前述の手法で多角形を構成する。構成される多角形は、線分を反時計回りに辿ってできる多角形（正の多角形）と、時計回りに辿ってできる多角形（負の多角形）の2種類あるが、負の多角形のうち他のすべての多角形を包含する最大の多角形が存在するので削除する。その結果として、正の多角形が土地区画を表し、負の多角形が土地区画の穴を表す。

(5)の処理で(4)までの処理で得られた正負の多角形を組み合わせて土地区画を構成する。ある正の多角形に対してそれが包含する負の多角形のうち他の負の多角形に包含されないものは全てその正の多角形の穴を表すので、これにより、穴の開いた多角形が構成可能である。この手法によって構成される多角形については、一つの正の多角形に穴を表す負の多角形が0個以上存在し、負の多角形同士に包含関係がない。

2.3 土地区画と文字列データの関連付け

地区画と文字列データの関連付けは、引き出し線がない場合に、文字列データを囲む文字枠が一定の面積割合以上で重なる土地区画を選択する。引き出し線がある場合には、図-5のように引き出し線の矢印のある側の端点を抽出して、その

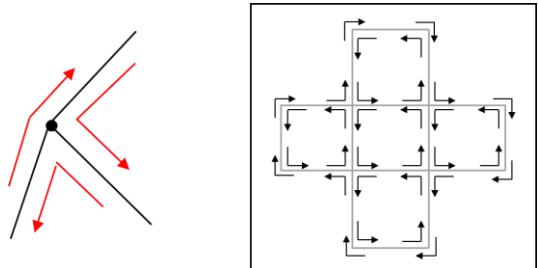


図-4 隣接する線分を左手に領域をみるように辿ることで得られる多角形の例。



図-5 引き出し線の矢印側の（赤点）と文字列を囲む文字枠（赤枠）の例。



図-6 地図情報の要素間の関連付け成功例：地番参考図中の矢印の位置を基に、地番区画と地番文字列が自動的に関連付けられている。

端点が含まれる土地区画と矢印のない側の端点を延長した線と交差する文字枠に対応する文字列データを関連付ける。図-6に土地区画と文字列データの関連付けに成功した例を示す。

3. 技術の共有

高性能計算の分野で開発された大規模計算が可能なプログラムが広く利用されるためには、その入力となる詳細な都市のモデルを構築するためのプログラムも共有・再利用される必要がある。

3.1 入力データの異種性

災害自体は日常的に起きることではないため、災害シミュレーションのための都市モデル構築ではデータの2次利用が主となりやすく、災害シミュレーションの対象都市ごとに都市モデルの情報源が異なる可能性がある。このため、入力となる情報源の異種性を吸収する仕組みがプログラムの共有・再利用のために必要である。

3.2 コンポーネント指向に基づく中間データ

個々のデータは、同一の内容をもつデータであっても、データ提供者によって構成ファイルの数や配置、データ形式が異なる。異種データのこうした本質的でない差異を吸収するために、異種データのそれから中間データを生成し、処理プログラムはこの中間データに適用することでプログラムの共有・再利用性が高まる。この中間データは珍しいものではなく、多様なデータ形式が扱える市販のGISソフトウェアのメモリ上でのデータも中間データの一種とみることができる。

本研究では、日進月歩の災害シミュレーション技術に合わせたモデル構築の高度化を想定しているため、中間データは容易に高度化が可能である必要がある。そこで、中間データとその中間データに関連する処理を動的ライブラリとしてまとめるコンポーネント指向を採用し、ライブラリ群の核となるデータ処理プラットフォームを開発した。

3.3 データ処理プラットフォーム

データ処理プラットフォームは、ライブラリに含まれる個々の中間データをそれぞれプラットフォームが扱えるデータ型として認識し、ユーザが入力するスクリプトに従った処理を実行するインターフェースの機能をもつ。

インターフェース言語の機能として、あるデータ型のデータが別のデータ型のデータに変換可能である場合に、データ型の変換を自動で行う仕組みを導入している。これにより、変換先のデータ型に対して定義されている処理が変換前のデータ型に対しても適用可能となった。したがって、個々の異種データに対して、その情報を完全に保持した中間データを定義すれば、そのデータから異種性を吸収するための中間データへの変換は自動で実行することが可能となった。

4. おわりに

本研究では地図情報の断片化された要素の間の関連付けを位置情報に基づいて自動構築する技術の開発とともに、この技術に基づくプログラムを効率的に共有・再利用するためのシステムであるデータ処理プラットフォームを開発した。

謝辞

本研究は公益財団法人計算科学振興財団研究教育拠点（COE）形成推進事業の助成を受けたものです。本研究はJSPS科研費26820188の助成を受けたものです。本研究の一部は、総合科学技術・イノベーション会議のSIP（戦略的イノベーション創造プログラム）「レジリエントな防災・減災機能の強化」（管理法人：JST）によって実施されました。

参考文献

- 大谷英之・陳健・堀宗朗（2014）：異種GISデータに記録された構造物の3D形状と属性情報の自動関連付け、土木学会論文集A2(応用力学), Vol. 70, No. 2, pp. I-631-I-639.