

事例空間を用いた学習型地図認識システムの構成

2M-2

魯偉† 奥橋俊之† 坂内正夫†
 東京大学生産技術研究所†

1 はじめに

図面入力システムの構成に際して、完全自動化はまだ難しく、人間の介入に頼らなければいけないのは実情である。そのため、筆者らは図面入力の効率及び信頼性を向上させる方式を提案した [1]。この方式では、自動処理できなかった部分を解析し、可能な回答をユーザに提示することによって、ユーザインタフェースの負担を減少し、処理の信頼性を改善した。一方、このシステムでは似たパターンに対しては、提案が間違っても、ユーザの正解入力によらずにいつも同じである。

この問題を解決するために、対話によって入力された正解をシステムに学習させ、システムの提案をその学習結果を用いて修正してゆく方式を提案する。実験結果ではこの方式の有効性を証明した。

では図目データを二値化画像データに変換し、その結果から細線化の結果に相当する芯線データ及びその接続情報を抽出する [2]。「認識システム」では、平行線抽出によって、道路、ハッチング、鉄道などを認識し、ラベル付けを行なう。また、認識できないオブジェクトに関する記述用パラメータを求める。認識過程を支配するのは、「認識知識」部分である。「会話システム」では、自動認識できなかったオブジェクトをユーザへの提示、ユーザの回答を学習システムへのフィードバックす、提示された回答に基づいて分類の提案等を行なう。「学習システム」では、教示された事例を「事例データベース」に蓄積し、新しい未知オブジェクトを獲得した事例データベースに基づいて分類し、その分類案をユーザに提示する。次の節では、各部分について説明する。

2 システム構成

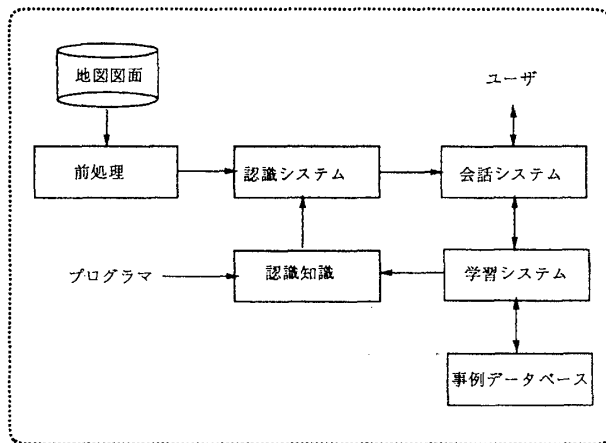


図 1: Configuration of the Proposed System

図 1はシステムの全体構成を示す。「前処理部」

Construction of Map Recognition System by Learning Through Example Space

Wei LU†, Toshiyuki OKUHASHI†, Masao SAKAUCHI†

†Institute of Industrial Science, University of Tokyo

‡currently with MIC Pte. Ltd.

3 前処理

地図図面から必要となる情報を抽出するために、その図面を構成する要素の抽出を行なわなければいけない。本システムは筆者が提案した改良した芯線化方式を用いて、図面要素の抽出を行なう。詳細は参考文献 [2] で述べられているので、ここでその概略を紹介する。

芯線化は紙面上に描かれた図面をスキャナによるイメージ画像への変換から始まる。次に、イメージ画像を二値化し、その輪郭を追跡し、折れ線によって近似する。それから、隣接した平行輪郭線のペアから芯線ベクトルを引く。最後に、芯線は交差点で途切れるので、途切れたところの輪郭情報を芯線データに追加する。また、図形シンボルに対応する部分は面積が小さくて、他の図形オブジェクトと接続しないので、芯線化が行なわれない。結果としては、独立成分と呼ばれる輪郭線のループの形で残される。この方式では、線図形情報を高速にベクトル化でき、芯線間の接続情報も正確に把握できるし、接続成分がある時でも、元の情報を保持できる。図 2に前処理の一例を示す。

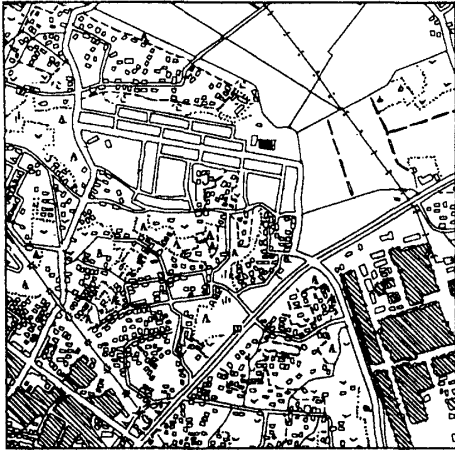


図 2: Example of Vectorization of Map Drawings

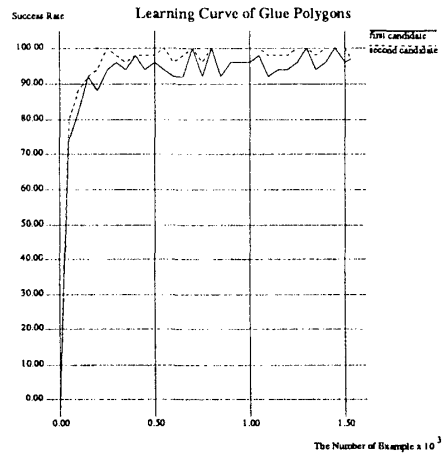


図 3: Experimental Result

4 図形オブジェクトの表現

図形オブジェクトは芯線、孤立成分と接続成分(グルー)の特徴量によって記述されている。具体的に、下記のような特徴量が使用されている。

- 形状：線幅、長さ等
- 芯線 ○ 接続：連結する芯線の数、その属性
- 形状：面積、周囲長、円形度、凸性、細長さ、突起の数等
- グルー ○ 接続：連結する芯線の数、その属性
- 形状：面積、周囲長、円形度、凸性、細長さ等
- 孤立成分 ○ 接続：連結する芯線の数、その属性

但し、全てのパラメータは正規化されている。

5 事例空間及びその学習

対話システムによって、教示されたオブジェクトのパラメータ及びその分類は事例データとして、事例空間に蓄積されてゆく。単純なオブジェクトなら、事例空間は簡潔なパラメータ分布にまとめられるが、ノイズ、変形などの影響でそうならないものも存在する。このような対象に対しては、

事例データそのものを用いて学習を行なうほかはない。

ニューラルネットワークによる学習が研究されているが、その内部構造の不透明による制御の難しさがある。そのため、本システムは、事例空間の更新による学習を行なう。新しい事例が入力される度に、そのパラメータは各概念のサブ空間に蓄積される。

未知のオブジェクトが与えられた時に、対応するパラメータの各サブ事例空間との距離が求められる。その距離によって、未知オブジェクトの分類提案が決められる。

6 実験結果

節5で述べた方式に基づいて実験を行なった。提示した事例数とその時点のシステムの提案の正解率との関係を図3に示す。実験結果から分かるように、やく200事例を与えた後、システムの提案の正解率が90%以上になるようになる。第二候補を入れると正解率は97%以上になる。

7 おわりに

事例空間を用いた学習対話型地図図面入力システムの構成を提案した。この方式ではより明瞭な学習が実現されている。実際の地図図面データを対象にして行なった実験では、提案方式の有効性を示した。