

図・本館

知識情報処理による
地図情報システムと
文字認識後処理に関する研究

名古屋大学図書	
和	1144654

丹羽 寿男

目次

1	序論	1
1.1	はじめに	1
1.2	背景	2
1.2.1	地図情報システム	2
1.2.2	文字認識	5
1.3	従来の研究	8
1.3.1	地図情報システム	8
	(1) 地図表現のデータ構造	8
	(2) 経路探索方法	9
	(3) 地図情報の提示方法	9
1.3.2	文字認識後処理	10
	(1) 候補外文字	10
	(2) 言語処理の精度	11
	(3) 文字認識後処理の評価	12
1.4	本研究の考え方	13
1.4.1	地図情報システム	13
	(1) 地図データの表現	13
	(2) 経路探索	13
	(3) 地図情報の提示	14
1.4.2	文字認識後処理	14
	(1) 文書内容知識	14
	(2) 最適文節選択	15
	(3) 文字認識後処理の評価法	15
1.5	本論文の構成	16
2	道路網の階層的表現	19
2.1	まえがき	19
2.2	道路網の階層的表現と経路探索	21
2.2.1	道路網の階層化	21
2.2.2	経路探索の問題	21

2.2.3	道路網の階層的表現	24
2.3	経路の探索アルゴリズムとその評価	28
2.3.1	経路の探索アルゴリズム	28
2.3.2	アルゴリズムの評価	32
2.4	階層型モデルにもとづく地図情報システムの実現	34
2.4.1	地図情報とその実現	34
2.4.2	システムの機能	36
2.5	まとめ	38
3	略図と文章による地図情報の提示	41
3.1	まえがき	41
3.2	略図と道路網	42
3.2.1	略図の特性	42
3.2.2	道路網の階層表現	43
3.3	略図の構成法	44
3.3.1	略図構成の概要	44
3.3.2	経路を表す略図	46
	(1) 下位層における道路選択	46
	(2) 上位層における道路選択	48
	(3) 略図の構成	50
3.3.3	目的地を表す略図	52
3.4	説明文の作成	55
3.5	実験結果と考察	58
3.6	まとめ	59
4	文字認識誤りモデルを用いた文字認識後処理機能の分析と評価	61
4.1	まえがき	61
4.2	文字認識誤りモデル	63
4.2.1	候補外文字モデル	63
4.2.2	候補内文字モデル	66
4.3	文字認識後処理の評価法	69
4.3.1	評価用データの作成	69
4.3.2	文字認識部の認識率に対する評価	71
4.3.3	文字認識部の認識処理方法に対する評価	71
4.4	文字認識後処理の分析評価実験	73
4.4.1	文字認識部の認識率に対する評価	74
4.4.2	文字認識部の認識処理方法に対する評価	76
4.5	まとめ	78

5	キーワード情報を用いた文字認識後処理	79
5.1	まえがき	79
5.2	文字認識後処理法の基本構成	81
5.3	文字認識後処理法	83
5.3.1	キーワード情報の利用	83
	(1) Zipfの法則	83
	(2) キーワード情報の抽出	85
	(3) キーワード情報を利用した後処理	86
5.3.2	最適文節選択法	87
	(1) 句読点の抽出	87
	(2) 最適文節選択	89
5.3.3	実験結果と考察	91
	(1) 後処理法の基本効果	91
	(2) キーワード情報利用の効果	94
	(3) 最適文節選択法の効果	96
	(4) 後処理の例	98
5.3.4	後処理の問題点の分析	101
5.4	まとめ	102
6	結論	103
6.1	本研究のまとめ	103
6.2	今後の研究課題	105
	謝辞	107
	参考文献	109

表 目 次

2.1	地図情報システムの応答時間	37
4.1	候補内モデルにおける定数 b の実際のデータとの比較	68
5.1	句読点の統計データ	88
5.2	後処理後の文字認識率	92
5.3	キーワード抽出の例	94
5.4	後処理の誤り分析	101

目次

1.1	地図情報システム	3
1.2	文字認識後処理	6
2.1	階層的表現	22
2.2	道路網の例	24
2.3	交差点隣接関係データ	25
2.4	交差点位置データ	27
2.5	道路データ	27
2.6	階層間交差点对应データ	27
2.7	最短経路の検索	31
2.8	道路網モデル	33
2.9	検索システムの構成	35
2.10	最短経路検索の例	37
3.1	道路網の階層	43
3.2	特徴交差点	45
3.3	下位層で選択される道路	47
3.4	上位層で選択される道路	49
3.5	経路を示す略図	51
3.6	一般の地図	51
3.7	上位層で選択される道路	53
3.8	目的地を示す略図	54
3.9	特徴交差点情報リスト	56
3.10	文の生成規則	57
3.11	説明文の例	57
4.1	実際のデータにおける候補外文字確率	65
4.2	候補内文字モデルと実際のデータの比較	67
4.3	評価用データの例	70
4.4	文字認識部の認識率に対する特性	75
4.5	文字認識部の認識率に対する評価	75
4.6	文字認識部の認識処理に対する評価	77

5.1	文字認識後処理の構成	82
5.2	Z i p f の法則	84
5.3	キーワードの抽出	85
5.4	接続評価値	90
5.5	第1候補認識率と後処理後の認識率	92
5.6	文字認識部の認識率と後処理の修復率	93
5.7	キーワード情報を利用した後処理	95
5.8	本手法の文節選択の効果	97
5.9	文字認識部の出力（後処理前）	98
5.10	後処理結果	99
5.11	後処理に用いた文書画像	100

第 1 章

序論

1.1 はじめに

コンピュータ技術の著しい発展に伴い、様々な分野でコンピュータが用いられるようになってきた。それに従い、より複雑な処理が要求されるようになってきている。このような要求に応えるために、知識を利用して極めて複雑な処理を可能とする知識情報処理の研究が盛んである。

知識情報処理の歴史は、1960年代に様々な基礎的な技術が確立されたことに始まる。1960年にマサチューセッツ工科大学のマッカーシーによりリスト処理言語 LISP が開発され、記号論理学、知的探索手法などの各種基礎技術の研究が行われた。1971年にスタンフォード大学のファイゲンバウムらの DENDRAL (質量スペクトル分析や核磁気共鳴などのデータから未知の化合物の構造を推論するシステム)、1971年にマサチューセッツ工科大学のモーゼスらの MACSYMA (数式処理システム)、1976年にスタンフォード大学のショートリッフェらの MYCIN (医療診断システム) が開発され、エキスパートシステムが実用化の見地から知識情報処理の一つの大きな成果として注目を浴びるようになった。現在では、エキスパートシステムは多くの分野において実用化されている。また、エキスパートシステム以外にも知識を利用した様々なシステムの研究が行われている。

本論文では、地図情報システムと文字認識後処理¹に知識情報処理を適用し、人間の行っている直感的な思考を模倣して知識情報処理を行うことにより、柔軟な処理をめざす。そのために、人間がどのようにして問題を解決しているかを調べ、その処理過程を導入した知識情報処理を実現する。

¹文字認識後処理とは、文字認識において1文字ごとの処理では認識できない文字を文法や語彙などの知識を用いて訂正する処理である。

このような知識情報処理を実現するために、本論文では次のことがらについて考える。

- 知識情報処理を行うための知識のデータベースを人間が処理対象に対して持っている概念にもとづいた構造で表現する。
- 知識情報処理を行う場合に、知識情報処理を行う対象の状態を抽出し、その状態にしたがって処理を行う。

これらの処理を実現することにより、柔軟な処理をめざす。

1.2 背景

人が地図を見て経路を決定するとき、単に距離の情報だけではなく、道路の状況や目標物などの種々の情報から経路を決める。それで、このような経路の決定には知識情報を用いた処理が必要である。

一方、人が文書を読むとき、文書が汚れていて判読できない文字があっても、人は文章の前後の関係から文字を推定することができる。それで、コンピュータによる文字認識においても、種々の知識情報を用いることにより、このような文字の推定が可能となる。

そこで、本論文では、地図情報システムと文字認識を知識情報処理の対象とした。以下に、本論文で取り上げるこの二つの問題の背景について述べる。

1.2.1 地図情報システム

近年、地図に含まれている様々な情報をデータベース化して、各種管路やケーブル（水道、ガス、電力など）の管理、工事の支援、あるいは地図情報のサービスを効果的にする試みが盛んである。

地図情報システム（図 1.1）の構築のためには、次のような技術が必要である [1]。

- 地図データの取得
- 地図データベースの構築
- 地図データベースの検索

地図データの取得は、地図情報システムで用いるデータを獲得する技術である。地図データの取得において、地図情報のデータ入力 of 効率的な手段の実現、地図データの標準化、

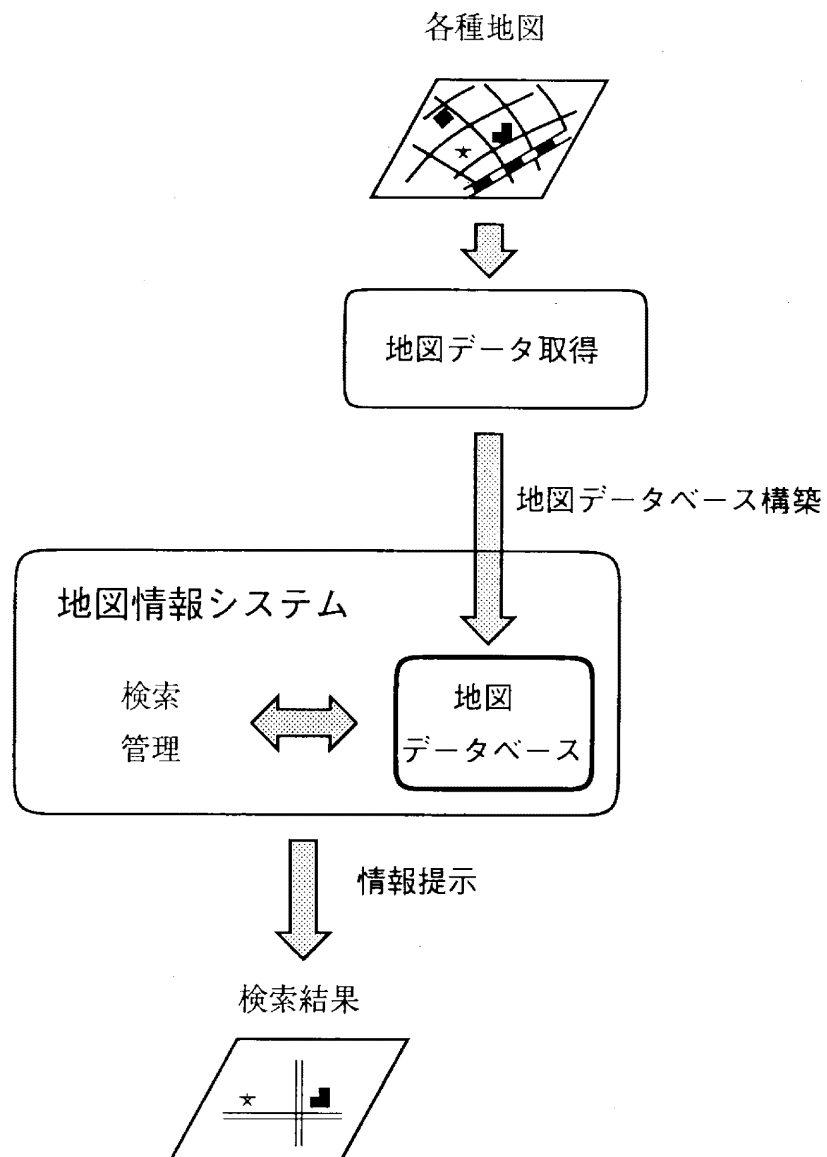


図 1.1: 地図情報システム

実際に有用な地図データの整備などの問題を解決する必要がある。地図データが十分に整備され、標準化されていない現状では、既存の地図図面や航空写真などから、図面の自動的な認識あるいは人手によりデータを入力しなければならない。しかし、現状の図面の自動認識技術では、すべての情報を完全に認識することは、実現できていない。それで、自動認識においても人手による修正作業が必要となっている。したがって、データ入力効率化のために、地図データのすぐれた自動入力が必要であり、その方面の様々な研究が続けられている。

地図データベースの構築では、効率良く検索でき、さらに効率良く蓄積できるようにデータを表現する必要がある。地図情報には、道路、鉄道、等高線、ガス、電力、河川、地名、建物などの1次元および2次元の情報があり、これらの情報をその地理的關係も含めて蓄積することが必要である。また、データ表現は汎用的な地理幾何検索が効率良く処理できるデータベース表現でなければならない。

地図情報システムにおいては、様々な検索が考えられる。地図データベースの単純なデータ検索から、複雑な地理幾何演算を伴う検索まで種々の検索に関する手法が研究されている。また、地図以外にも関連する他のデータを組み合わせ、システムを形成する研究 [2, 3, 4] も試みられている。さらに、地図以外の画像や音声などのマルチメディア・データを用いて、システムを形成する方法も報告されている [5]。このようなシステムでは、地図情報と他のメディアとの相互のデータ検索も必要になる。地図情報は日々変化していくので、地図の変化に対応してデータベースを修正する処理も必要である。

本論文では、地図情報システムの対象として市街地地図を取り上げる。市街地地図には、複雑な網目状に道路が示されていて、地名、町名、建物、鉄道、河川などの情報も示されている。これらの情報を効率良く蓄積し、さらに、蓄積された情報から迅速にかつ正確に必要な情報を抽出する手法を提案する。ここで提案する経路探索の手法は、単に距離的に最短な経路を求めるのではなく、人が道路に関して持っている概念の知識情報を用いることにより、人にとって分かりやすい経路を求める。また、今日までに開発された多くの地図情報システムでは、地図データベースとして構築された情報を、検索指示にしたがってそのままの状態を表示しており、利用目的に応じて情報を的確に表示しているとはいえない。地図に含まれている様々な情報をそのまま表示することは、地図情報システムの応用目的によっては、必ずしも有効な情報サービスではない。そのために、利用目的に応じて地図情報システムで提示される情報を取捨選択する機能が大切である。このような機能を

実現するには知識情報処理が有効であり、本論文では、地図情報システムにおける情報の提示方法についての手法を提案する。

1.2.2 文字認識

コンピュータの普及とそれに伴う情報の価値の高まりにより、高速で簡単なデータ入力の要求が高まっている。この要求に応えるもののひとつとして文字認識が注目されている[24]。日本語文書入力用のOCRの認識精度は年々向上してきており、印刷文書に関しては、文字品質の良い文書であれば、認識率99%以上の装置が開発されている。[25]。しかし、品質の悪い文字や手書き文字の場合には、まだ実用的な認識率のレベルに達していない。

文字認識処理は、以下の処理から構成される[26]。

- 紙の上に書かれた文字をスキャナーにより画像として入力する。(画像入力)
- 1文字ごとに文字パターンを切り出し、雑音除去や大きさの正規化などのパターン整形を行う。(文字切り出し)
- 文字パターンから文字固有の特徴を抽出する。(特徴抽出)
- 抽出された特徴とあらかじめ辞書として登録されている文字の特徴とを照合して、最も類似したものを認識結果とする。(文字識別)
- 語彙・文法などの知識を用いて認識結果の妥当性を調べ、これを修正する。(後処理)

これまでの文字認識の研究では、特徴抽出と文字識別に力点が置かれていた。しかし、認識対象がOCR帳票から一般文書へ拡大するにつれて、自由書式や手書き文字を処理するために、文字切り出しや後処理が以前に増して重要となっている。

特に、日本語情報を扱う場合に、漢字は文字種が多く、類似文字も多いため、低品質文字や手書き文字の場合、1文字ごとの認識処理だけでは、認識率の向上に限界がある。そこで、上述の文字認識後処理(図1.2)が重要である。これは、人間が文書を読むときには、文法や語彙などの知識を用いることにより、少々のノイズがのった文書画像でも認識することが可能であることから、このような知識を用いて文字認識の処理を行うものである。知識情報処理を用いることにより、1文字ごとの処理では認識できなかった文字を訂正することができる。

文字認識後処理の対象は、

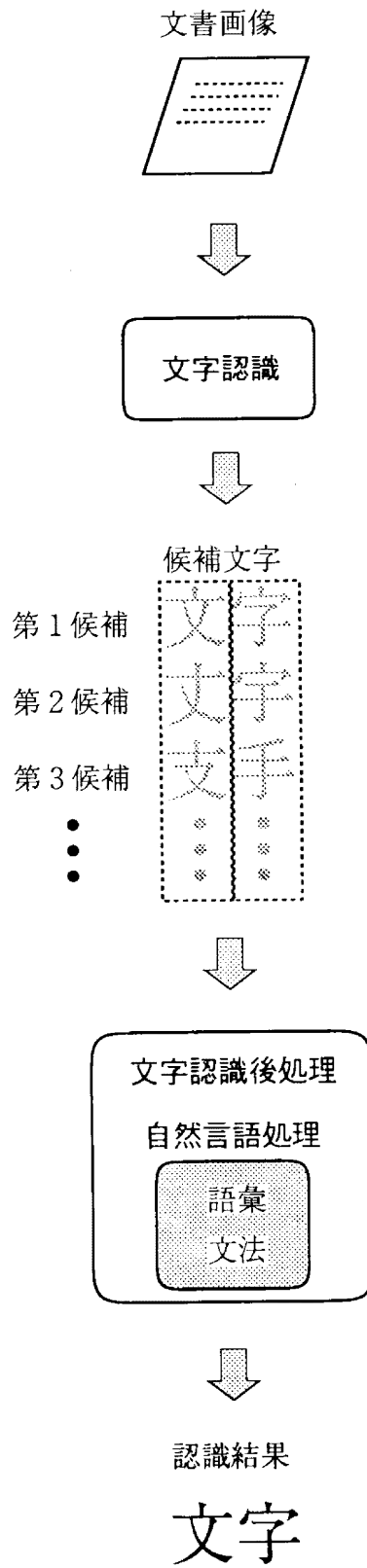


図 1.2: 文字認識後処理

- 決められた欄に項目ごとに記入されている住所や商品名を入力する単語入力
- 一般の文章を入力する文章入力

に分けられる。文字認識後処理では、処理対象にもとづいた種々の知識が用いられ、それぞれの処理対象に合わせた処理方法が研究されている。

単語入力に関しては、入力対象の単語を蓄積した単語辞書との照合を行う単語レベルの処理で、大きな効果が得られる。さらに住所についての階層関係の利用など、入力対象に特定した知識情報処理により、単語入力に関しては実用化のレベルに達している [27, 28, 29]。

一方、文章入力に関しては、一般的な文法の知識や語彙の知識を用いて後処理を行う [30, 31, 32]。しかし、手書き文字の認識など文字認識部²での文字認識率が低い場合には、訂正文字列の候補が複数存在するために、誤認識文字の訂正は、まだ実用化のレベルに達していない。さらに、文字認識部での文字認識率がある値以下の場合には後処理を行うことにより、逆に、認識率を低下させるといわれている。これは、訂正文字列の候補が多く、文法や語彙の知識だけでは正解を判断できないために、逆に正解文字列を誤訂正してしまうことがあるからである。

本論文では、文書の内容を理解して、文書内容に沿ったより正確な文章に訂正する手法を提案する。また、文字認識部の文字認識率が低い場合においても、誤訂正の少ない訂正文字列の選択法を提案する。

ところで、従来文字認識後処理の性能評価としては、文字認識後処理後の文字認識率や訂正率（文字認識部が間違えた文字の中で訂正できた率）を用いることが一般的である。しかし、文字認識後処理の入力データである文字認識部出力結果の認識率により、文字認識後処理後の認識率や訂正率は大きく変化する。そのため、これらで評価できることは文字認識後処理のある条件における特性だけであり、これだけでは十分な評価ができない。つまり、文字認識後処理の性能特性を分析するためには、文字認識部が出力する種々の認識率のデータに対する特性を分析する必要がある。しかし、認識対象画像の品質と処理結果の認識率との関係を正確に決定することはできないので、文字認識部から任意の認識率のデータを得ることは容易ではない。さらに、低認識率のデータを収集するためには、画像にノイズをかけたり、画像を回転させたりして故意に画像を劣化させる必要がある。また、文字認識後処理の性能評価では、文字認識後処理とシステムを構成している文字認識部から出力される候補文字がテスト・データとして使われる。しかし、候補文字集合の中

²本論文では、後処理を含まない文字認識処理部分を文字認識部と呼ぶ。

に正解文字が含まれている確率や各候補文字に正解が存在する確率は、文字認識部の性能によって変化する。このように、文字認識系の条件によっても、文字認識後処理の結果は変動するので、文字認識後処理の性能を評価するためには、種々の文字認識部の出力データに対する特性も分析する必要がある。したがって、文字認識後処理の性能を評価するためには、文字認識部出力結果がどのような特性を持っているかを調べておく必要がある。

本論文では、文字認識後処理を評価するための評価用データを容易に得るために、文字認識部が出力する候補文字データをシミュレートする文字認識誤りモデルを提案する。

1.3 従来の研究

1.3.1 地図情報システム

今日まで、地図情報システムに関して様々な研究が行われているが、本節では、本論文で扱う次の3つの分野について従来の研究をまとめる。

- 地図表現のデータ構造
- 経路探索方法
- 地図情報の提示方法

(1) 地図表現のデータ構造

2次元平面上に描かれた地図に含まれている情報をデジタル・データとして取り扱うためには、地図に含まれる「点」、「線分」、「面分」の0次元、1次元、2次元の情報を表現する効率の良いデータ構造が必要である。

地図に含まれるこれらの情報を表現するために、地図の持つ位相幾何学的な情報に注目して、「点」をノード、「線分」をエッジ、「面分」をエッジで囲まれた領域とするグラフで表現する手法 [6, 7] が提案されている。また、同次元のデータと異次元のデータとの接続関係により表現する手法 [8, 9] が提案されている。

また、地図情報に含まれる情報が多量な場合には、情報を扱いやすくするために、地図を何らかの単位図形に分割して表現する必要がある。地図の分割法には2種類 [9] ある。ひとつは2次元平面をグリッドと呼ばれる互いに合同な長方形に分割する方法 [10] で、「グリッド分割」、「メッシュ分割」などと呼ばれる。グリッド分割は容易に実現することがで

きるが、各グリッドを区画する境界は全く人為的なもので、実際の地図上の境界とは一致しない。もう一つの分割法は、「多角形分割」と呼ばれ、道路、鉄道、河川、行政区界などを地図の境界とするもので、地図をこれらの境界線で多角形に分割する³。多角形分割は境界線が複雑であるために、データとして扱いにくい。

(2) 経路探索方法

地図情報システムにおける検索要求のひとつに経路探索がある。経路の探索は、道路網を有向グラフと考えた場合、グラフ上の最短パスを求める問題と同等であると考えられる。さらに、グラフの最短パス探索問題に交通渋滞度のような様々なコストを付加し、このコストを最小にする最適経路の探索問題が研究されている [11]。また、各経路を木によって体系的に表現することによって、 k 個の最短経路を求める手法 [12, 13] が提案されている。

最短経路の探索では、道路の数が多い場合に、計算量が膨大となる問題がある。この問題に対処するために、ニューラルネットワークを用いて、経路探索を並列的に計算する手法 [14] が提案されている。また、計算量を減らすために、最短経路の探索に用いる道路の数を制限する方法が研究されている。道路の数を制限するために、道路網をグリッド分割して表現し、最短経路が経由すると考えられるグリッド部分を推定する。そして、この推定されたグリッド部分の道路だけを用いて最短経路を探索する手法 [10] が提案されている。この手法を用いれば、探索に用いる道路を少なくできるので、計算量が減る。しかし、この手法を用いても、出発地と目的地との間の距離が大きいときには、探索に用いる道路の数は多くなる。

また、従来の経路探索で検索される経路は、最短経路であるため、人が地図を見て選択する経路とは異なり、分かりにくい経路になりやすいという問題がある。

(3) 地図情報の提示方法

地図情報システムにより検索された情報を提示する方法として、地図データベースに構築された情報をそのまま表示する場合が多い [15]。利用目的によってはすべての地図情報が表示されるより、地図情報を加工して提示した方が望ましい。情報を加工するために、地図情報をカテゴリごとに分けて蓄積し、利用目的に応じて必要なカテゴリの情報を提示する方法 [16] がある。しかし、検索された情報を分かりやすく提示するためには、必要でな

³実際の境界線は曲線であるが、通常これを折れ線近似するので多角形に分割される。

い情報はできるだけ省いた方がよく、必要最小限の提示する情報を選択する必要がある。

また、電話やカー・ナビゲーションにおいて、検索された情報を音声により提示するために、経路情報を文章に変換する研究 [17, 18, 19] が行われている。しかし、図で提示する場合に比べて、経路情報を文章だけで表現した場合には、情報を文章により厳密に伝える必要があるため、長く分かりにくい文章になることが多い。

1.3.2 文字認識後処理

文字認識後処理では、入力文字列の言語情報を利用して文字の読み取り精度を向上させている。帳票など定型の書類以外の文字認識後処理においては、語彙や文法などの一般的な知識により、候補文字集合の中から正しい文字列を選び出す。文字認識後処理で問題となるのは、

- 候補外文字の処理
- 言語処理の精度
- 文字認識後処理の評価

である。

(1) 候補外文字

認識結果の候補文字集合の中に正解の文字が含まれていない文字を候補外文字と呼ぶことにする。候補外文字がある場合には、候補外文字の訂正が困難であるだけでなく、候補外文字を含む単語が単語辞書を用いて検索できないために、候補外文字の周辺に誤りが広がってしまうことがある。候補外文字が発生する原因として、次の2つがある。

- 文書画像のつぶれやかすれ
- 文字の切り出し間違い

文書画像のつぶれやかすれに対しては、候補文字補完 [33] が提案されている。これは、つぶれやかすれが発生した文書画像において、文字認識部がある文字をどの文字に誤るかを類似文字テーブルとして、あらかじめ記憶しておき、実際の文字認識時に候補文字集合の文字を調べて、類似文字テーブルから正解の可能性のある文字を取り出し、候補文字集

合に補完する手法である。これにより、候補外文字を推定することが可能である。候補文字補完は平仮名に関して有効である。しかし、漢字に関しては文字数が多く類似文字も多いためあまり有効ではない。また、文字間の接続情報に着目して、文字接続表を用いて正解を判定する方法 [34] も研究されている。これは、文字間の接続関係を統計的に調べ、統計的な接続関係の確率によって候補外文字を推定するものである。しかし、統計的確率によって文字を推定するので、誤訂正がよく起こる。さらに、実際にOCR装置を人が使用するとき、人間が文字認識結果を訂正した操作履歴を認識誤りマトリクスとして記憶し、このマトリクスにより正解文字を推定する手法が研究されている [35]。この手法を用いることにより、活字漢字フォントの違いによる文字認識部の誤りを後処理で訂正することができる。

文字の切り出し間違いに対しては、文字切り出しにおいて、切り出し位置として可能性のある位置をすべて出力させ、その切り出し文字の組み合わせの中から正しい文字を選び出す方法が提案されている [36]。ただし、すべての切り出し文字候補の組み合わせを計算すると計算量が膨大となる。

単語辞書に未登録の単語が文書に含まれる場合にも、単語辞書による検索ができないために候補外文字と同様な処理が必要である。単語辞書の未登録語に対しては、単語であると思われる文字の並びを未登録語テンプレートとして登録し、この未登録語テンプレートと照合する文字列を検索することにより、未登録語の単語範囲を推定する手法 [37] が提案されている。

(2) 言語処理の精度

文字認識後処理では、文法などの言語知識により正解文字列を決定している。文書画像の品質が低く、文字認識部における文字認識率が低い場合には、選択すべき文字列の候補が多数になる。後処理において、一般の自然言語処理に用いられていると同様な言語処理を行うと良い性能を得ることができる [38] が、計算量が非常に多く、計算時間が膨大になり、実用的ではない。

そこで、文字認識後処理では構文解析などは行わずに、文節レベルでの言語処理の結果と認識結果の距離情報とを総合的に判断して、最適な候補を決定する手法 [39, 40, 41] が多く用いられている。

最適な候補を選択するための手法としては、グラフ探索の手法にもとづいて、前の文節

から順に接続できる文節を探索していき、句読点を含む文節まで探索できたら、探索された文節を出力文字列として選択する手法 [42, 43, 44] が提案されている。グラフの探索手法を用いることにより、容易に文節を探索できるが、前から順に探索するため必ずしも最適解を見つけることができない。また、候補外文字が途中にあるとそれ以降文字列の探索に失敗する。これらを解決するために、正解である確率の高い文節に対するパスのみを初めに求め、それらのパスの中から最適なパスを求める手法 [45] も提案されている。

最適な候補の評価の基準として、認識結果の距離情報、単語の長さ、文字単位での接続情報、単語間の接続規則、単語出現頻度などを用いる研究 [37] が行われている。

(3) 文字認識後処理の評価

文字認識後処理の評価法は、従来提案されていない。そのため、従来の文字認識後処理の評価は、文字認識システム全体としての評価、あるいは、ある文字認識部を用いた一定条件下での性能評価によって行われていた [33, 37]。しかし、これらの手法では異なる文字認識部に接続する文字認識後処理の性能を比較することは不可能である。また、文字認識後処理の性能を分析するためには、いろいろな品質の文書を用意して、それを文字認識部で認識し、種々の認識率の評価データを作成して、それぞれの後処理結果を求める必要があった。この場合、文書によって後処理性能が変化するため、正確に性能を分析するためには、非常に多くの文書を処理する必要がある。

1.4 本研究の考え方

人間の持つ柔軟な思考を知識情報処理を用いることによってコンピュータ上に実現することが本研究の目的である。本論文では、地図情報システムと文字認識後処理を適用先とし、それぞれの問題において、人間がどのようにして問題を解決しているかを調べ、その処理過程を導入した知識情報処理を実現する。前節で述べた従来の研究状況に対して、本研究では以下のように考える。

1.4.1 地図情報システム

地図情報として市街地地図を用いて、道案内を目的とする地図情報システムを考える。そのために、市街地地図情報システムにおいて最も重要な要素技術と考えられる道路網の効率的な表現方法と各種の経路探索問題について考える。さらに、検索された地図情報の提示方法についても考える。

(1) 地図データの表現

市街地地図に含まれるすべての道路網のデータを一様に表現すると、データを検索するための計算量が膨大となる。また、計算量を減らすために道路網をメッシュ分割すると、検索のための計算が複雑となり、検索された経路も必ずしも分かりやすい経路ではない。そこで、人間が地図の道路網を把握する場合に、地図データの大局的な構造を見るときと、局所的な構造を見るときでは、見ている道路のレベル（生活道路、幹線道路、高速道路など）が違うことに注目する。すなわち、本研究では地図データをそのレベルに応じて階層化を図り、下位の道路網は上位の道路によって多角形分割する。このようにデータを人間が地図を見るときに知識を用いてデータを蓄積することにより、データの検索や更新を効率化できる。

(2) 経路探索

経路探索の問題に対して、単純な最短経路ではなく道案内のために最適な経路を考える。地図情報システムが提供する経路情報は、対象とする領域の地理に詳しくない人にも分かりやすい経路である必要がある。そこで、検索される経路は、分かりやすさに重点をおいた経路とした。すなわち、近距離の移動は生活道路を利用し、遠距離の移動ほど広い道路

(地図上で大きく描かれている道路)を利用する経路である。このような経路は必ずしも距離的に最短な経路ではないが、人間の直感に一致し、分かりやすい経路である。

(3) 地図情報の提示

地図情報システムで検索された情報を、分かりやすく提示する方法を考える。地図データベースのデータをそのまま提示すると、必要のない情報も提示されるため分かりにくい図となる。そこで、道案内のための情報提示では、必要な情報だけを選択して提示する略図の方が実際の地図より有効である。略図は、検索された情報の中の特徴的な情報が誇張されている。そのため、実際の地図に比べて、人間にとって必要な情報を得ることが容易である。また、略図では広い範囲の地図と狭い範囲の詳しい地図を同一紙面上に表すことができ、非常にコンパクトに情報を提示することができる。そこで、このような特徴を持った略図を生成する方法について考える。さらに、分かりやすく情報を提示するために、略図において表現できない内容や図で表しにくい情報について説明文を生成して、略図と説明文の両方を用いて検索結果の情報提示を図る。

1.4.2 文字認識後処理

文字認識後処理に用いる知識として文書の内容を取り入れ、文書内容に沿って文字認識することを考える。さらに、候補外文字の含まれる文章における後処理による誤訂正の低減を考える。また、文字認識後処理の性能評価方法についても考える。

(1) 文書内容知識

人が文書を読む場合に、文書の文字が汚れなどにより一部分読み取れなくても、正しい文字列を判断することができる。これは、人間が文書に書かれている内容を把握しながら文章を読んでいるからである。文字認識後処理において、文書画像が低品質な場合には訂正文字列候補が多くなり、文法や語彙の知識だけでは正解を判断することはできない。このような場合に、文書の内容を把握することにより、正しい判断を行い、文字認識率の向上を図る。文書の内容として、文書のキーワードを取り上げ、文書中からキーワードを自動的に抽出する。そして、このキーワードの情報を用いて文字認識後処理を行う。

(2) 最適文節選択

候補外文字による誤訂正の低減を図る。人が文章を読む場合に、読みにくい文字を飛ばして文章全体を読み、前後の関係から読みにくい文字を推定していると考えられる。そこで、訂正文字列候補から正解文字列を選択するときには、文章の前から順に選択するのではなく、最も確からしい文節から先に選択確定していくことを考える。これにより、不確かな文節の選択確定は最後に処理され、候補外文字が存在する場合においても、候補外文字による誤訂正が周辺に広がることを防ぐ。

(3) 文字認識後処理の評価法

文字認識後処理の性能を比較できるように、一般に文字認識部がどのような候補文字集合を出力するかを調べ、この候補文字集合をシミュレートする方法を考える。これにより、文字認識部の性能や文書画像の品質に左右されずに文字認識後処理の性能を分析・評価することができる。

1.5 本論文の構成

本論文は、以下の各章から構成される。

第1章では、研究の背景、研究の考え方を述べ、本研究の目的と方法の概略を整理した。

第2章と第3章では、地図情報システムについて述べる。第2章で、地図情報システムにおける地図データの表現と経路探索方法を、第3章で、地図情報システムにおける情報の提示方法について述べる。

第2章では、市街地地図における道案内のための地図情報システムの実現方法を提案する。市街地地図に含まれる道路網を道路のレベル（生活道路、幹線道路、高速道路など）にしたがって、階層化を行うことにより、地図データを効率的に表現し、道案内のための最適な経路探索を実現する。まず、道路網の階層的表現方法と、道路・交差点などのデータ構造を提案する。次に、道案内のための最適経路として、経路探索の問題を定義する。この経路探索問題を解くアルゴリズムを提案し、アルゴリズムを評価する。さらに、これらの階層化された地図データ表現と最適経路探索手法にもとづいてワークステーション上に実現した地図情報システムについて述べる。そして、道案内のための経路が実用的な応答速度で検索でき、検索された経路も人間が地図を見て選択する経路のような分かりやすい経路であることを示す。

第3章では、地図情報システムから検索された道案内のための情報を、略図とその説明文を用いて提示する方法を報告する。略図は地図に含まれている様々な情報から有用な情報を取り出して構成する必要がある。そのため、出発地から目的地までを一様に表現するのではなく、場所によって提示する情報のレベルを変化させる。このような略図を構成するために、道路網の階層性に着目し、道路網を道路のレベルによって階層化して、略図および説明文を構成する。階層の上位層からは道路網の大局的な情報を取り出し、下位層からは道路網の局所的で詳しい情報を取り出す。これらの情報から略図を構成し、同時に略図を補完する道案内の文章を生成する。

第4章と第5章では、文字認識後処理について述べる。第4章で、文字認識後処理の分析評価法を、第5章で、文字認識後処理手法について述べる。

第4章では、文字認識後処理における性能を評価するために、文字認識誤りモデルを提案する。このモデルでは、文字認識部出力の第1候補認識率と文字認識部の性能に関するパラメータから、各候補文字に正解文字が存在する確率を推定する。モデルにしたがって評価用データを作成することにより、任意の認識率や任意の認識系における文字認識部の

出力データを作成でき、文字認識部認識率依存特性や文字認識系依存特性を容易に分析することが可能である。これらの特性を分析することにより、文字認識後処理を評価することができることを示す。さらに、この評価法を実際の文字認識後処理法に適用した評価実験を行い、この評価法の有効性を示す。

第5章では、手書き文書などの低品質の文書に対しても対応できる文字認識の後処理法を提案する。訂正文字列の候補が多数存在する場合にも対処できるように、従来から利用している文法や語彙の知識情報に加え、認識対象の文書の内容に関する知識情報も用いて後処理を行う。文書の内容に関する知識としてキーワードを取り上げ、これを候補文字集合から自動的に抽出し、キーワードの情報を用いて後処理する方法について述べる。さらに、後処理では、正解文字が候補文字に入らない場合に、その文字の前後の文字に対して、誤訂正が行われる問題点に関して、確からしさの高い文節から訂正文字列を確定する最適文節選択法を提案する。これらの手法を用いて後処理の実験を行い、後処理後の認識率が向上することを示す。

第6章では、本研究で得られた研究の成果についてまとめ、次に、本研究に関する今後の課題について述べる。

第 2 章

道路網の階層的表現

2.1 まえがき

地図情報のデータベースを人間の持っている道路の概念に合った構造で表現することによって、より人間らしい処理を容易に行えるようにして、効率の良い知識情報処理をめざす。本章では、地図に含まれる道路網を例として、データベースの表現について述べる。

地図には、道路網、地名、建物、鉄道など種々の情報が示されている。これらの情報を有効に利用するために、各種の管路あるいはケーブル（水道、ガス、電力等）の管理、工事計画支援、あるいは地理情報サービスを目的としてデータベース化が進められている [1, 2, 8, 13]。

市街地地図には、網目状に道路が示され、さらに地名、町名、建物、鉄道、河川などの情報も示されている。これらの情報を管理するためには、その地理的關係も含めて情報を蓄積することが必要である。また、各種の地理情報を検索するシステムでは、蓄積された情報から迅速に、かつ正確に必要な情報を抽出することが重要となる。

本章では、市街地地図のデータベース化において、市街地における道路の役割が階層的であることに着目して、道路網を表現する方法と、それにもとづいて実用的な意味で最適な経路を求めるアルゴリズムを提案する。また、この方法を応用して、実際にワークステーション上に実現した地図情報システムについて述べる。

市街地地図のデータベース化において、地図表現の様々なデータ構造が提案されている [6, 9]。本章で示す道路網の表現方法は、市街地の道路がバス道路などの幹線道路と、それらに囲まれた領域内部の生活道路とに分類されることにもとづいている。また、経路の探索方法 [10, 12] は、道路の役割にもとづき、我々人間が経路を定めるときに、単純な最短

経路にはよらず、例えば、分かりやすい経路、車が走りやすい幹線を主とする経路を選択することなどに着目している。このような最適な経路を求めるアルゴリズムを、最短経路を求める Dijkstra のアルゴリズム [20] を基本にして構成した。また、このアルゴリズムをもとに、道路に一方通行のような制約条件のある場合に対しても適用可能なアルゴリズムを与える。

以下、2.2 節で道路網情報の記述について述べ、2.3 節ではこの記述を用いて表現された道路網上での最適な経路の探索アルゴリズムを示す。さらに、2.4 節でこれらの記述とアルゴリズムを適用して実現した地図情報システムについて述べ、2.5 節で、成果のまとめについて触れる。

2.2 道路網の階層的表現と経路探索

本節では、経路の探索問題と本章で提案する道路網情報を階層的に表現する方法について述べる。

2.2.1 道路網の階層化

人が地図から種々の情報を抽出する場合には、局所的な情報から大局的な情報まで、いろいろなレベルで地図を参照する。よって、広い範囲にわたる地図全体を一律に扱うのではなく、地図全体の領域をいくつかの小領域に分割し、階層化すると効果的である。階層を定義する基準として道路の役割による階層（路地、バス道路、高速道路など）による方法、行政区画の階層を用いる方法が考えられる。後者は必ずしも道路をその基準として区画化されているとは限らないので、道路網を中心とする種々の情報の検索のためには、前者を基準とするのが妥当である。後者は、実際にデータとして道路網とそれに付随する情報を蓄積する場合に、各階層に含まれる道路網の規模を定める基準とするために有効と考えられる。

階層は道路の役割を基準として定義されるので、上位層の道路ほど、広く長い道路で、長距離の移動に適している。例えば、最下層が生活道路、それから幹線道路（バス道路）、国道、高速道路というような階層が考えられる。

図 2.1 は、2 階層の場合で、幹線道路とそれらに囲まれた領域内の道路網というように階層化されている。一般に、上位層は地図の大局的構造を幹線道路図にもとづいて表現している。一方、下位層は上位層が表現する領域を幹線道路で区切り、それで得られる各多角形領域（これをブロックと呼ぶ）内の道路網の構造を表現している。

2.2.2 経路探索の問題

ここで考える問題は、対象とする領域の地理に詳しくない人が、道案内のための経路を探索する場合（徒歩および自動車）を想定している。すなわち、探索する経路の出発地と目的地が与えられて、出発地から目的地までの経路を求める問題である。

一般に、最短経路は、Dijkstra のアルゴリズム [20] で解くことができる。しかし、ここで求める経路は、必ずしも最短経路ではなく、初めてその土地を訪れる人が地図を見て選択する経路を考える。すなわち、近距離の移動は生活道路を利用し、遠距離の移動ほど広

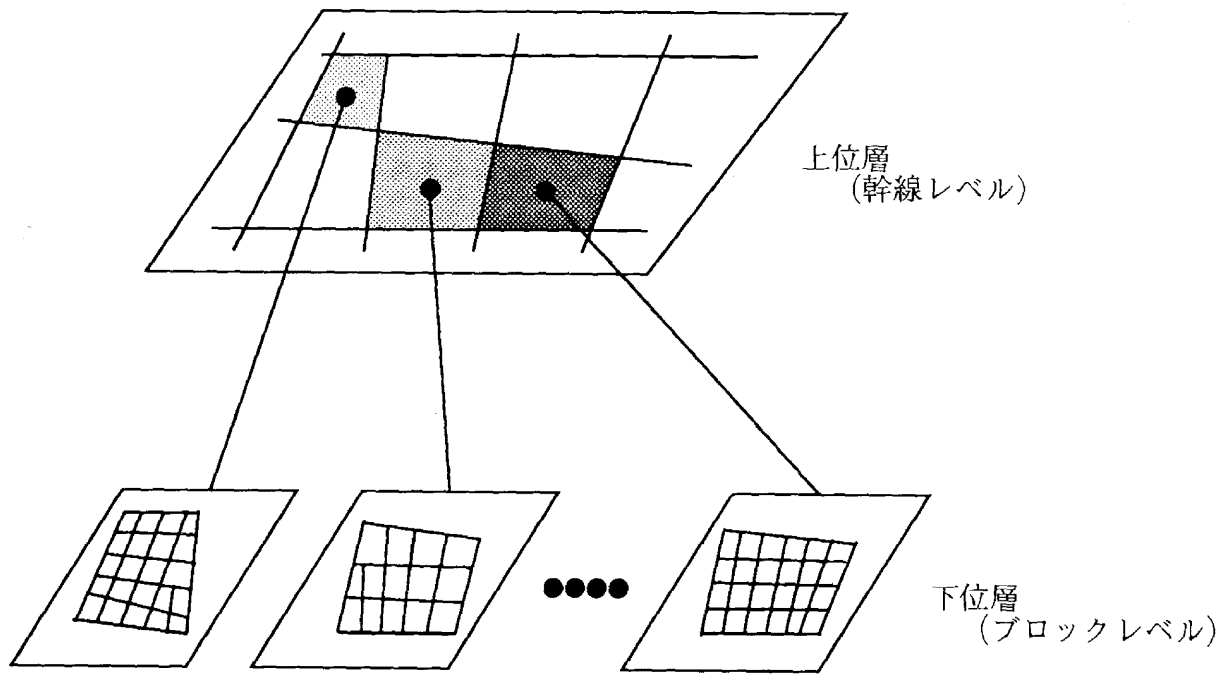


図 2.1: 階層的表現

い道路（地図上で大きく描かれている道路）を利用する経路である。これは、遠距離の移動ほど上位層の道路を利用することを意味している。

このような、道路の階層性を考慮した経路の探索は、経路の分かりやすさ、鉄道、バスなどの各種の交通手段を利用する場合の経路をモデル化している。ここで述べるアルゴリズムにより求められる経路は、すべての道路の階層を考慮せずに、一様に扱って Dijkstra のアルゴリズムを適用して得られる経路とは異なる。この問題（これを基本問題と呼ぶ）は次のように定式化される。

【経路探索の基本問題】

2つの地点（出発地と目的地）が与えられる。出発地および目的地の属する階層から、出発地と目的地を同時に含むブロックの属する階層までのそれぞれの階層において、出発地および目的地の属するブロック内の道路のみを経由する経路で最短なものを求めよ。

また、経路の選択に対して一方通行・右折禁止などの交通規制に対する制約条件がある場合には、その条件に従う経路を探索する制約条件付きの問題（変形問題）がある。この問題は、あらかじめ交差点、あるいは道路の制約に対する付加情報が与えられていれば、選択させる経路を構成する道路の隣接関係に制約を加え、そのもとで基本問題を解くことに相当する。したがって、変形問題の定式化は、基本問題のそれとほとんど変わらない。

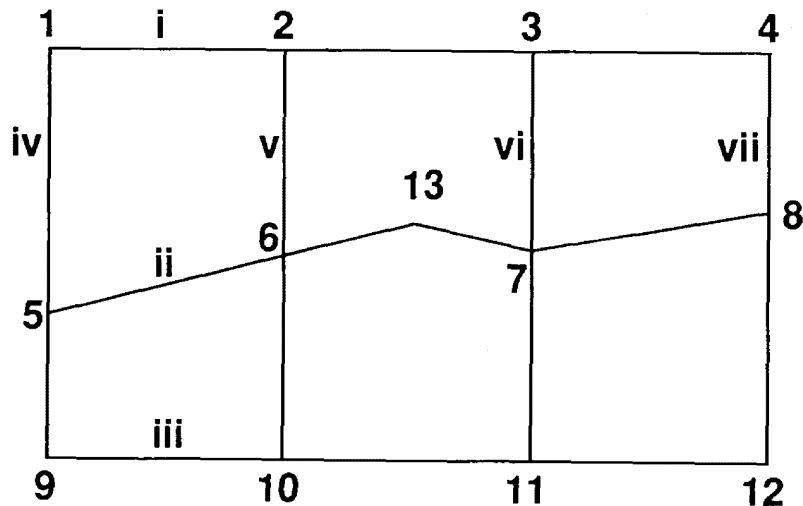
2.2.3 道路網の階層的表現

各階層に固有な情報は、それぞれのブロックのデータとして記憶される。また、各階層のデータのほかに上位層と下位層の間関係（例えば、対応する地点、下位層の境界を与える幹線道路など）を表すデータが必要となる。

ブロックに含まれる道路網の表現として、グラフを用いる。ここでは、道路網に関するデータだけを考え、領域に関するデータ（例えば、建物や公園などの領域）は考えない。以下では、2.3節で経路の探索アルゴリズムを表すために、道路網データの表現について述べる。図2.2に道路網の例を示す。道路網の接続関係は、交差点をノード、道路をエッジとするグラフとなる。これを交差点隣接関係データと呼ぶ。また、道路および交差点の位置を表すデータを、それぞれ道路データ、交差点位置データと呼ぶ。具体的には、以下のようである。

(1) 交差点隣接関係データ

それぞれの交差点（番号で識別される）について、隣接交差点番号、隣接交差点までの経由道路と距離である。図2.3に交差点隣接関係データの例を示す。



1,2,...,12: 交差点

13 : 屈曲点

i,ii,...,vii: 道路

図 2.2: 道路網の例

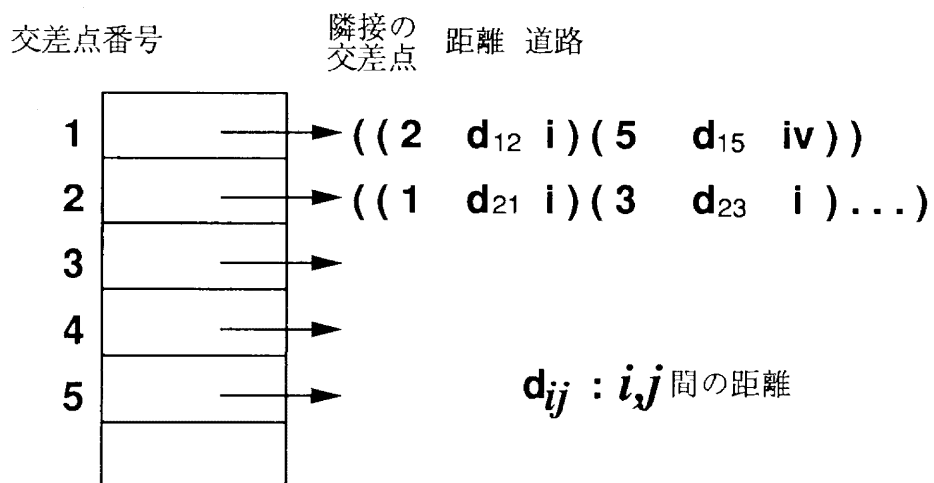


図 2.3: 交差点隣接関係データ

(2) 交差点位置データ

交差点の位置 (X座標、Y座標) で、座標系は、交差点間の相対的な位置を定めるのに用いられる。図 2.4 に交差点位置データの例を示す。

(3) 道路データ

道路の形状を折れ線で表現する。道路の始点 (交差点)、経由屈曲点、終点 (交差点) からなる。図 2.5 に道路データの例を示す。

さらに、下位層と上位層の関係を次に示す 2 種類のデータで表現する。

(4) 階層間交差点对应データ

下位層と上位層の対応している交差点の番号を表す。図 2.6 に階層間交差点对应データの例を示す。

(5) 階層間領域対応データ

上位層における道路で囲まれた領域と、下位層のブロック内データとの対応関係を表す。

(2) の交差点位置、(3) の経由屈曲点などは、2.3 節で述べる最適経路の探索問題を解くためには不要であるが、地図情報システムを実現する場合に経路の表示や説明のために、あるいはより複雑な経路の探索問題を解くために付与した。

交差点等 番号	X座標	Y座標
1	0	200
2	300	200
3	600	200
4	900	200
5		

図 2.4: 交差点位置データ

道路番号	道路の		
	始点	屈曲点	終点
i	(1		4)
ii	(5	13 7	8)
iii	(9		12)
iv			
v			

図 2.5: 道路データ

ブロック番号	ブロックでの 番号	幹線での 番号
1	((15 27)	(27 2)...
2	((16 37)	(11 38)...
3		
4		
5		

図 2.6: 階層間交差点对应データ

2.3 経路の探索アルゴリズムとその評価

ここでは、前節で与えた問題を解くアルゴリズム、およびアルゴリズムの評価について述べる。

2.3.1 経路の探索アルゴリズム

はじめに、階層化された道路網データにもとづいて前章の基本問題を解くアルゴリズムを示す。

【基本問題の解法アルゴリズム】

(入力) 出発地と目的地

(出力) 出発地から目的地への経路

(方法)

- Step 1. 出発地および目的地が属するブロックを各階層ごとに求める。対象レベルを最下位層にする。
- Step 2. 現在対象としているレベルにおいて、出発地および目的地が同じブロックにあれば、Step 6 へ。
- Step 3. 階層間交差点对应データにより、出発地の属しているブロックの交差点の中で、上位の階層の交差点にもなっている交差点をすべて求める。
- Step 4. 出発地と Step 3 で求めた交差点との間の最短経路とその距離を、Dijkstra のアルゴリズムにより求める。これには、出発地の属しているブロックの交差点隣接関係データを用いる。
- Step 5. 出発地から Step 3 で求めた各交差点までの、Step 4 で求めた距離を長さとする仮想的な道路を設定し、これらの道路を上位層の道路網に追加する。対象レベルを上位層にして、Step 2 へ。
- Step 6. 対象レベルが最下位層ならば、Step 10 へ。
- Step 7. 対象レベルの交差点の中で、下位層中で目的地の属しているブロックにも含まれる交差点を求める。

- Step 8. 出発地と Step 7 で求めた交差点との間の最短経路とその距離を Dijkstra のアルゴリズムにより求める。
- Step 9. 出発地から Step 7 で求めた交差点までの、仮想的な道路を設定し、下位層の道路網に追加する。これらの道路の長さは、Step 8 で求めた距離とする。対象レベルを目的地の含まれる下位層にして、Step 6 へ。
- Step 10. 出発地と目的地間の最短経路とその距離を Dijkstra のアルゴリズムを使って求める。
- Step 11. Step 10 で求めた経路に含まれている仮想的な道路を、それぞれに対応する経路に戻して得られる経路全体が求める経路である。 ■

図2.7は、道路網を2段階に階層化したときに、上のアルゴリズムを適用した例である。図2.7(a)の出発地○から目的地●までの経路を探索する。図2.7(a)の▲は、Step 3で求めた交差点であり、Step 4の出発地○から▲までの最短経路を求めている。図2.7(b)の▲は、Step 7で求めた交差点であり、Step 8の出発地○から▲までの最短経路を求めている。図2.7(c)は、Step 10の出発地○から目的地●までの経路を求めている。

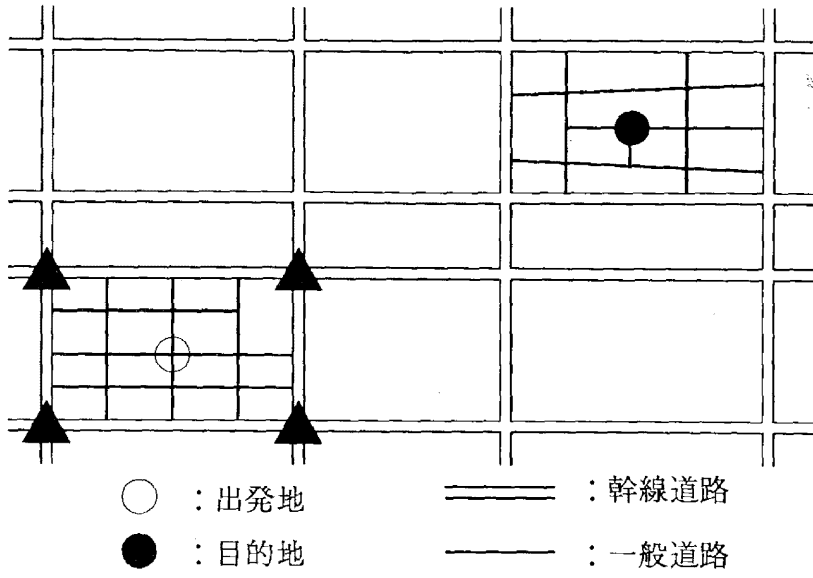
次に、制約条件のある場合の問題の解法について考える。ここでは、一方通行および右折禁止の制約条件について考える。ただし、右折禁止の規制は、幹線道路についてのみ扱う。

まず、制約条件に関するデータを付与する。すなわち、道路に関する属性データとして、一方通行の方向属性を、交差点に関する属性データとして右折禁止の属性を付与する。これらのデータが与えられると、制約条件付き問題のアルゴリズムは、基本問題のアルゴリズムを修正することで得られる。ここでは、アルゴリズムの修正内容を示す。

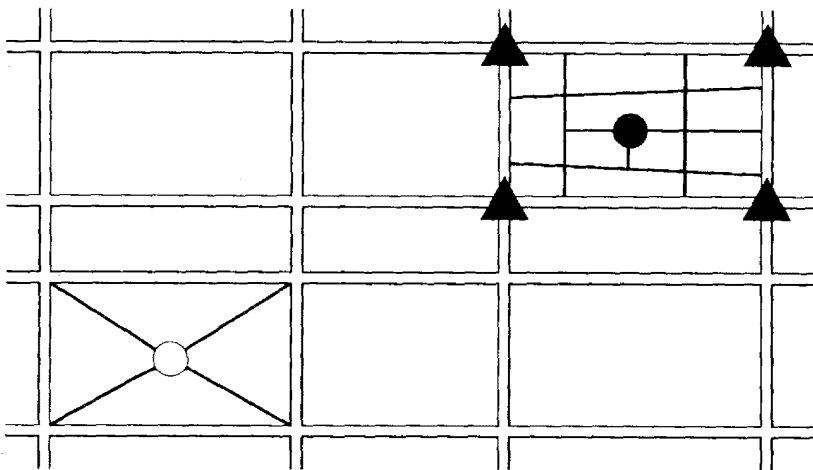
【変形問題の解法アルゴリズム】

一方通行の制約条件問題では、基本アルゴリズムのStep 4,8,10において、Dijkstraのアルゴリズムを適用するときに、道路の属性を参照して一方通行の交通規制がある場合は、その方向に道路がないものとして計算する。

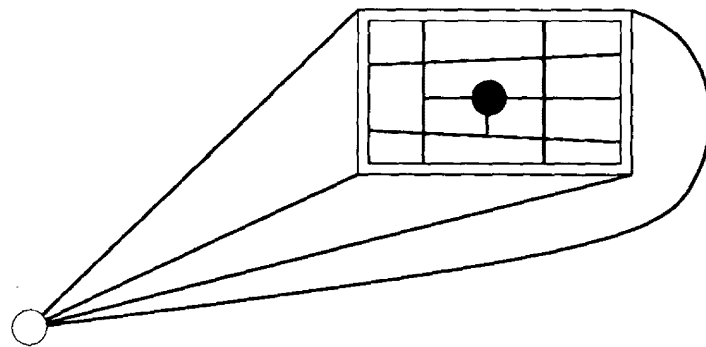
また、右折禁止の制約条件問題では、中央分離帯のある幹線道路は下位層における道路の属性として、ブロック側から見て中央分離帯を越えないで走行できる車線についての一方通行属性を付与する。実際に解を求めるためには、基本問題のアルゴリズムのStep 4,8,10において、Dijkstraのアルゴリズムを適用するときに、出発地から幹線道路のある交差点までの経路が求まる。その幹線道路に中央分離帯がある場合には、隣接する交差点までの距離を求めるときに、次のようにする。一方通行禁止の方向に対しては、その交差点に最も近い右折できる交差点を調べ、その交差点までの距離を求める。そして、その右折可能な交差点が隣接する交差点であるように扱う。



(a) 第一段階



(b) 第二段階



(c) 第三段階

図 2.7: 最短経路の検索

2.3.2 アルゴリズムの評価

前節で述べたアルゴリズムの計算量を評価する。はじめに、図2.8に示すような道路網のモデルを考える。簡単のためにそれぞれの階層レベルにおける各ブロックには n 個の交差点が含まれ、 k 段階の階層化が行われているとする。また、あるレベルのブロックは、そのひとつ下位の階層のブロックを m 個含んでいるとする (n, m は階層によらないとする)。この道路網モデルで経路の探索問題を考える。Dijkstra のアルゴリズムでは、 i 個のノードのグラフで最短経路を見つけるのに必要な計算時間は $O(i^2)$ である。ここで提案したアルゴリズムでは、Dijkstra のアルゴリズムが適用される回数は最大で $(2k - 1)$ 回である。それぞれの計算時間 (Dijkstra のアルゴリズムの計算時間) は、ブロック内の交差点の数が n であることから $O(n^2)$ となる。したがって、全体の計算時間は $O((2k - 1)n^2) = O(n^2)$ である。一方、階層化を行わない場合には、道路網全体で交差点の数が nm^{k-1} であることから、最短経路を検索する計算時間は $O(n^2 m^{2(k-1)})$ となる。

本章で提案したアルゴリズムは、上のモデルでは道路網全体の交差点の数が増えても、 $k \ll n$ であれば、計算時間は $O(n^2)$ で変わらない。例えば、上のモデルで2段階の階層化を行い、 $m = n$ の場合、全体の交差点の数は、 $nm = n^2$ であるから、本アルゴリズムでは、全体の交差点の数に比例する計算時間で経路を探索できる。これは、道路網全体に対して一様に Dijkstra のアルゴリズムを適用した場合の計算時間 $O(n^4)$ に比べ、はるかに効果的である。

一方、このアルゴリズムを使って探索される経路は、Dijkstra のアルゴリズムで求められる最短経路を探索するのではなく、長距離の移動ほどより上位の階層に含まれる道路を使う経路を選択する。この考え方は、一般的に人が選択する経路によく似ていると思われ、道案内のための経路を求めるのに適している。

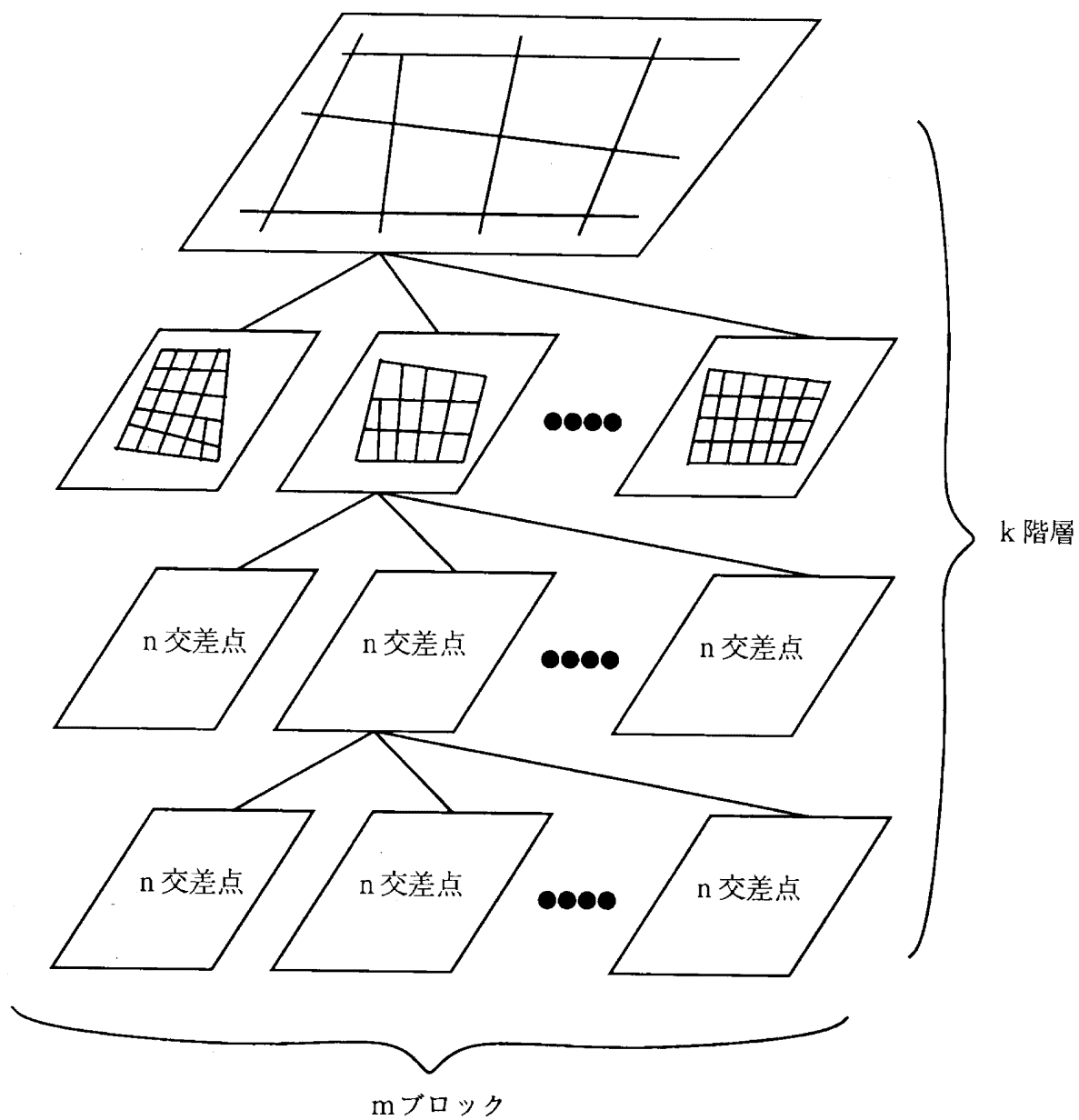


図 2.8: 道路網モデル

2.4 階層型モデルにもとづく地図情報システムの実現

2.2節、2.3節で述べた階層型道路網モデルにもとづいてワークステーション上に実現した地図情報システムについて述べる。

2.4.1 地図情報とその実現

2.2節、2.3節で述べた階層型道路網モデルにもとづいた地図情報システムをワークステーションに実現した。地図情報として、名古屋市千種区(約6km×約5km)の地図[21]を使用した。この地図に含まれているデータは、交差点数4701、道路数1732、地名・施設などの数664である¹。これらのデータを2.2節で述べた方法にしたがって、2段階に階層化して表現した。上位層の幹線道路図に対して下位層は57枚のブロック地図で構成されている。システムの実現には、ワークステーション XEROX 1100SIP を使用した。

このシステムでは、道路網情報のほかに地名・施設などのデータを追加している。これらのデータを用いて、道案内のための種々の地図情報をマウスによる簡単な入力とマルチウィンドウを使った分かりやすい表示により検索できる。このシステムは、さらに道路網が変化したときに、階層間の関係も含めてデータの一貫性を保ってデータを更新する機能も備えている。

地図情報システムの構成を図2.9に示す。検索の場合、まずマウスよりウィンドウ管理部へ検索要求が出され、検索コマンドが選択される。次に検索コマンドが地図情報アクセス部に送られ、入力が必要な場合は入力部へ入力要求が出され、入力データを得る。検索結果は画面表示部へ送られ、表示のためのウィンドウが開かれ、ディスプレイに表示される。

システムは、Interlisp-Dで記述され、プログラム・サイズは約5000行から成り、地図情報データが占めるファイルの容量は約350kBである。

¹地図のデータは、名古屋市千種区の地図[21]をもとに手作業(マウス)により入力したものである。

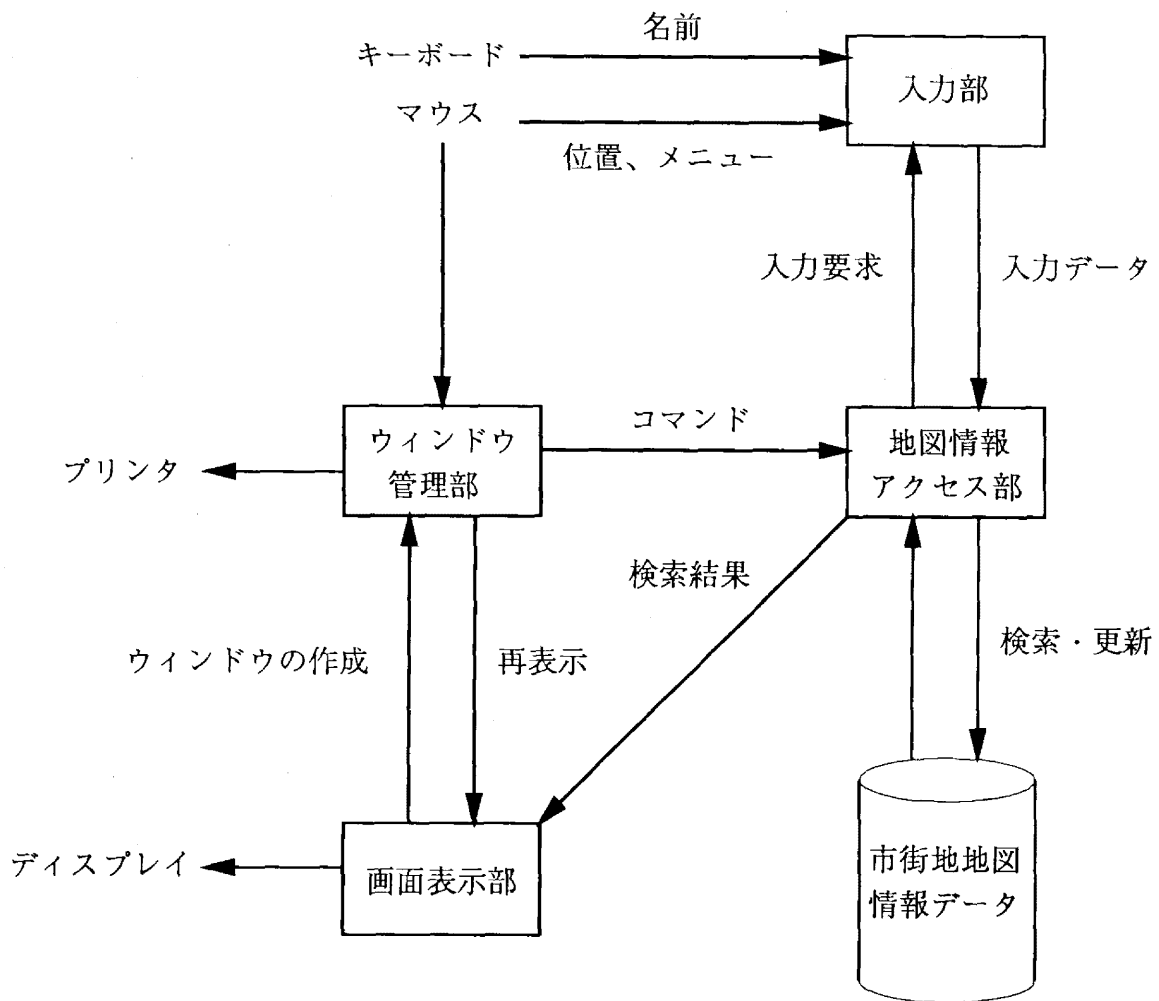


図 2.9: 検索システムの構成

2.4.2 システムの機能

今回実現した地図情報システムでは、次のような情報を検索できる。

1. ある領域の詳しい地図の表示
2. 町名に対応する領域の表示
3. 地名・施設などの所在地の検索
4. 出発地から目的地までの経路の探索
5. ある種類の施設（例えば、銀行、小学校など）の中で出発地からの経路が2.3節で述べた意味で最適であるものの検索²

これらの機能のうちで、4と5に対して、前節で述べたアルゴリズムが用いられている。また、それぞれの検索に対して、交通規制にしたがった経路を探索するか交通規制に関係しない経路を探索するかを選ぶことができる。図2.10は、指定した出発地から目的地までの経路を探索した例である。左上の大きいウィンドウが上位層の幹線道路全体図を表示しており、下の2つのウィンドウがそれぞれ下位層のブロック内の地図を表示している。これらの3つのウィンドウにそれぞれの階層における経路が表示され、全体の経路を知ることができる。

それぞれの検索に要する時間を、表2.1に示す。各項目に対する値は、30種類の異なるキーを指定して検索し、時間を測定した。検索時間には、表示時間およびガーベッジ・コレクションの時間は含まれていない。検索システムに登録されている建物の数は、大学6、公園23、銀行32、派出所12である。ほとんどの検索要求に対して、数秒以内で検索することができ、実用的な応答速度を持っていると言える。最寄りの建物の検索では、その検索方法（ある範囲内にあるものに候補を絞ってからすべての経路を検索する）の特性から、検索時間が登録されている建物の数とその散らばり方に依存する。

²このシステムでは施設の所在地で検索するのではなく、建物に最も近い交差点を求め、この交差点を使って検索する。実用化に際しては、建物が接する道路および隣接交差点をデータとして持ち、検索時にはその地点をグラフに組み込んで検索を行えばよい。

表 2.1: 地図情報システムの応答時間

検索コマンド	検索時間 (秒)		
	最小	最大	平均
最短経路検索	0.593	4.895	2.797
町の検索	0.022	0.117	0.053
地名・建物の検索	0.001	0.003	0.002
最寄りの建物の検索 (大学)	2.225	9.983	6.394
最寄りの建物の検索 (公園)	0.728	13.547	5.139
最寄りの建物の検索 (銀行)	0.561	12.189	3.944
最寄りの建物の検索 (派出所)	0.360	7.167	2.846

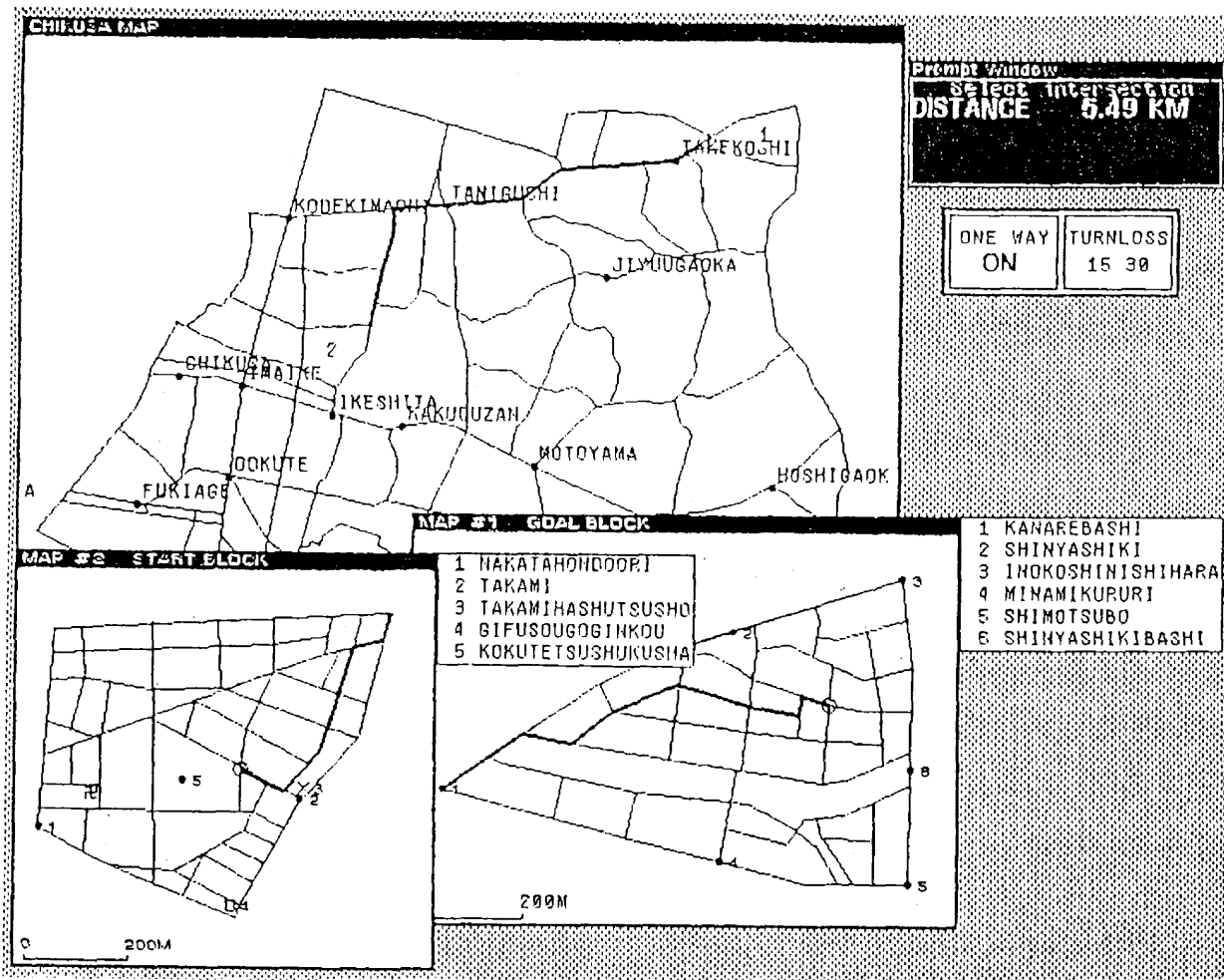


図 2.10: 最短経路検索の例

2.5 まとめ

本章では、人間の思考にもとづいて、道路網情報を道路のレベルによって階層的に表現する方法を提案した。また、道案内のための最適な経路を探索するために、その階層型道路網モデルにもとづく経路の探索アルゴリズムを提案した。階層型道路網モデルにもとづく経路の探索アルゴリズムは計算時間においてすぐれており、探索される経路も人が地図を見て選択する経路に似ている。また、階層型道路網モデルにもとづいてワークステーション上に実現された地図情報システムでは、経路探索を十分に実用的な速度で計算することができた。

地図情報システムでは、名古屋市千種区を例として挙げたが、階層化するときの各ブロックの大きさは均一なほうがよいことや、あまり遠距離の移動に関しては距離以外の要素で経路が決ってくることを考えると、本アルゴリズムは市街地地図を使った地理不案内の人への道案内に適しており、実用的にはひとつの都市の範囲に適用することが妥当である。この方法の問題点として、階層間にまたがる道路網データの更新問題がある。すなわち、データを更新する場合、そのデータが最下位層レベルにおけるブロック内だけの変更であるならば、データを簡単に更新できる。しかし、更新の内容が複数の階層に影響を及ぼすとき（例えば、幹線道路が新しくできた場合）には、データの更新は単純ではない。最下位の階層レベル以外で道路が追加・削除された場合、更新された道路があったブロックの下位に属するすべての階層レベルのブロックに対して、データを更新する必要がある。道路が追加された場合は、対応する下位の階層レベルに属するブロックがその道路によって分割されるのでブロックを再構成しなければならない。また、道路が削除された場合は、対応する下位の階層レベルに属するブロックを併合しなければならない。特に、階層化の段階が多く、最上位の階層レベルに属する道路に対して更新がなされたときには、これらの処理はかなりの手間を要する。今回実現したシステムでは、2階層の場合について、この機能を実現しているが、より一般的な更新問題の解決は今後の課題である。

さらに、次のことがらが今後の課題として挙げられる。

1. 道路の混雑など動的に変化する情報をデータに取り入れ、それらの変化を含めて経路の知的探索や動的な変更ができるようにする。
2. 道路階層の概念を各種の交通網（鉄道、バス、航空機など）に拡張し、より汎用的な地図情報システムに応用する。

3. より広範囲の地図に対して、人間が選択するような経路が検索できるように、種々の情報を考慮して経路を検索できるようにする。

第 3 章

略図と文章による地図情報の提示

3.1 まえがき

地図に含まれている様々な情報をデータベース化して、各種管路やケーブル（水道、ガス、電力など）の管理、工事計画の支援、あるいは地図情報のサービスを効果的にする試みが盛んである [1, 2, 8]。しかし、今日までに開発された多くの地図情報システムでは、地図データベースとして構築された情報を検索指示にしたがってそのまま表示しており、利用目的に応じた的確に情報を表示する機能を提供していない。地図には種々様々な情報が複雑に含まれ、それをそのまま表示することは、地図情報システムの応用目的によっては必ずしも有効な情報サービスを実現していない。例えば、道案内を目的とする地図情報では実際の地図よりも略図の方が有効である。さらに、提示情報を利用目的に応じて取捨選択する機能も重要である。

本章では、第 2 章で実現した地図情報システムにもとづいて、道案内のための情報提示機能を議論する。すなわち、地図情報システムで検索された情報を、道案内のために取捨選択、加工・変換し、利用目的に合致した略図と説明文として構成する方法を提案する。

略図は実際の地図に比べて、特徴的な情報が誇張され、主要な目印・道路の位相的な関係だけが保存されている場合が多い。また、同一紙面上に目的地付近の詳細な局所情報と、おおまかな大域情報が均衡を保って描かれている。したがって、道案内のための略図の構成には、検索された情報からそれぞれの重要度を判断して、制約された紙面上に配置することが必要になる。また、略図では明示できない情報を説明文として付加することも道案内に対して重要である [22]。

3.2 略図と道路網

道案内のための略図では、道路図が基本となる。したがって、道路網を地図データベースにどのように構造化するかが、略図を構成する場合に重要となる [6, 9, 10]。本節では、道案内のための略図について検討し、その下で道路網の構造化について言及する。

3.2.1 略図の特性

道案内のための略図には、主要な道路網を基本として、目的地の相対的な特徴を表す鉄道、河川、施設などの記号や名称が記載されている。これらの情報は実際の地図から一様に抽出されるのではなく、経路を効果的に表現する情報によって構成されている。略図には、次のような特徴がある。

1. 実際の地理に対して、位相的な位置関係のみを保存し、変形、強調操作が施されている。
2. 目的地に到る道路図を中心に、特徴的な施設、鉄道、場所などの指示記号が明示されている。
3. 利用目的に応じて、経路図の表現、使用される指示記号が異なっている。
4. 出発地、目的地に隣接する地域ほど詳細に描かれている。
5. 目印となる道路、建物、河川、地域などを明記するために、文字が用いられている。

すなわち、略図は実際の地図に比べて、多くの情報が排除され、図形表現として単純化されている。しかし、略図の描写法に必ずしも一定の規則があるわけではなく、様々な構成図が存在している。

本章では、地図情報システムから検索して得られた道案内のための経路図にもとづいて、上記 1 から 5 の性質を満たす略図の構成法を検討する [64]。この地図データベースには [57, 59]、一般の市街地地図が表すすべての情報を格納しているのではないが、道路、建物、地域名などが蓄積されている。したがって、略図の構成には、道案内のための経路にもとづいて、出発地・目的地の隣接域の同定と、それによって取捨選択すべき情報の認識が重要になる。

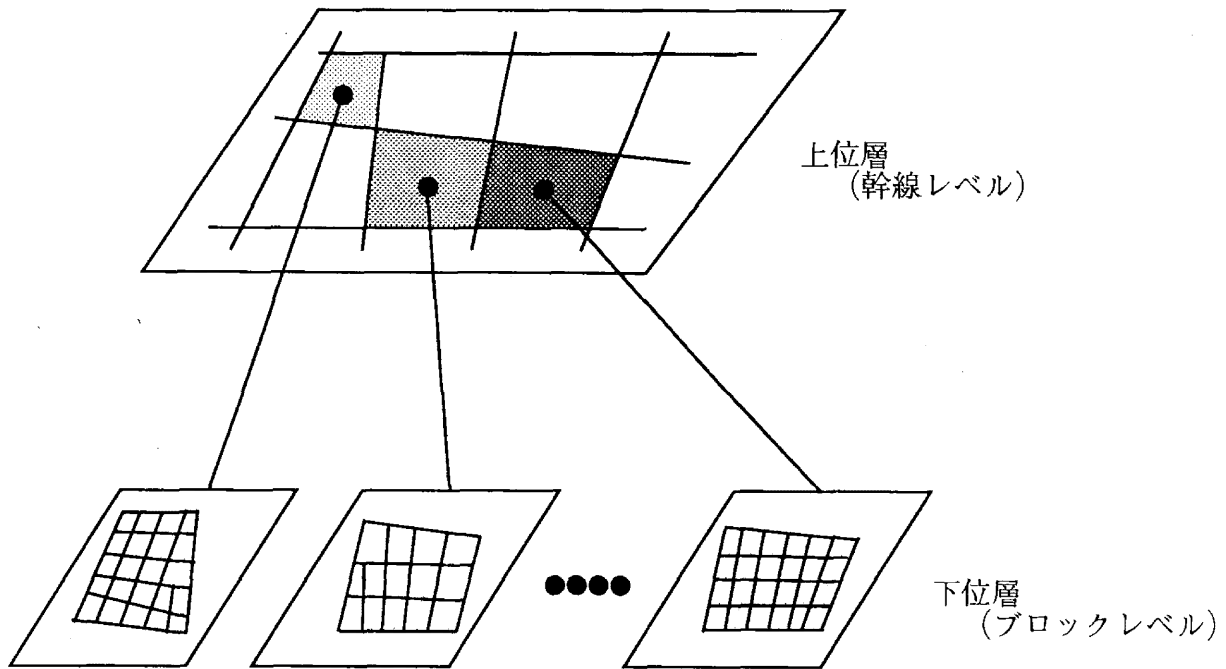


図 3.1: 道路網の階層

3.2.2 道路網の階層表現

道案内のためには、経路は最も重要な情報であり、略図も道路情報を中心に描かれる [23]。したがって、道路網を地図情報データベースとして、いかに表現するかが略図の構成法に大きく作用する。前提とする地図情報システムでは、2.2節で示したように、多階層で道路網を表現可能である [57]。すなわち、市街地地図の道路は路地、バス道路、高速道路などの役割を基準として道路網を階層化できる。本研究では、市内などの比較的狭い地域を対象としているために、生活道路、幹線道路を基準として道路網を2層の構造とした。図 3.1 は、その例で、道路網の大域情報と局所情報を段階的に表現している。

道路網の階層表現により、出発地および目的地に隣接する地域の詳細情報を下位層で、その間の経路に関する情報を上位層で表現でき、略図の性質 4 を容易に実現可能とする構造となっている。

3.3 略図の構成法

略図の構成に先立って、略図を成す道路網の構成成分を次のように定義する。

(1) 道路：

交差点で右左折しないで通行できる一続きの道路。ただし、階層や隣接するブロック（下位層の道路網で構成される多角形領域をブロックと呼ぶ）にまたがる道路は、別々の道路とする。

(2) 交差点：

2本以上の道路が交わる地点、および道路の行き止まりの地点。

以下、本節では具体的に略図の構成手続きを述べる。

3.3.1 略図構成の概要

道案内のための地図情報システムでは、次の2種類の情報が提示可能である。

- 経路情報
- 目的地の指示情報

これらの道案内情報では、提示すべき内容に多少の相違があるものの、基本的な視点は同じである。したがって、それぞれに対応する略図の構成においてもほぼ同様に処理できる。基本的に道路網の下位層と上位層に分けて情報を抽出し、それらの情報を統合する手順で実現する。

下位層では、出発地および目的地に隣接する地域に関して、特徴交差点を求めて、略図の描写に必要な情報を抽出する。ここで、特徴交差点とは、経路上にあって、次の条件のいずれかを満たす交差点である。

1. 経路が右左折する交差点
2. 経路の途中に信号機および施設などがある場合、これらの信号機および施設の間に存在する交差点を除いた交差点

特徴交差点の例を図3.2に示した。したがって、特徴交差点に接続している道路を抽出し、出発地および目的地付近の局所情報を選択できる。また、特徴交差点だけに注目することにより、経路上の重要と思われる情報を指示できる。

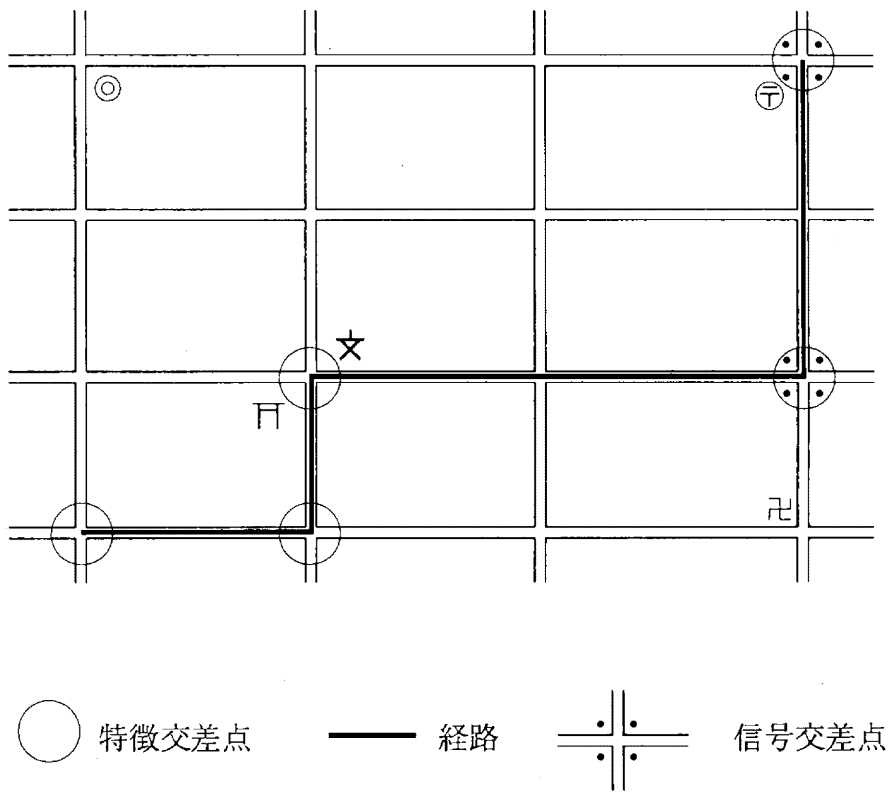


図 3.2: 特徴交差点

一方、上位層では、下位層間の大局的な位置関係、経路を特定できるように、検索・抽出された経路から一定範囲に存在する道路を求める。

3.3.2 経路を表す略図

経路を表す略図には、出発地から目的地への道順情報が示されている。入力された出発地と目的地によって、この間の経路を第2章で述べた地図情報システムから求める [57]。この経路は、長距離間の移動には、すべて上位層を通る経路となる。この経路に対して、下位層（生活道路レベル）と上位層（幹線道路レベル）で、それぞれに略図を構成する道路を選択する。そして、これら選択された道路にもとづいて略図を構成する。

(1) 下位層における道路選択

下位層における道路選択は、出発地のブロックと目的地のブロックで実行される。基本的に、それぞれのブロックで同じ選択手続きが実行されるので、ここでは目的地のブロックに対する道路選択の方法を示す。

【下位層における道路選択アルゴリズム】

(入力) 目的地

(出力) 下位層で略図を構成する道路

(方法)

- Step 1. 目的地が属するブロック内の主要な交差点を求める。
- Step 2. Step 1 で求めた交差点の数が3以下ならば、ブロックと幹線道路の接続交差点（ブロックの四隅の交差点となる場合が多い）を主要な交差点に加える。
- Step 3. 主要な交差点から目的地までの最短経路をそれぞれ求める。
- Step 4. 各経路上を成す道路部分を抽出する。
- Step 5. 各経路に存在する特徴交差点を求める。
- Step 6. Step 5 で求めた特徴交差点に接続する道路を抽出する。
- Step 7. 信号がある交差点に接続している道路全体を抽出する。 ■

ただし、Step 6において選択される道路については、その道路が特徴交差点に接続する回数によって、その道路部分が決められる。すなわち、次の基準に従う。

(1) 一カ所で特徴交差点と接続

経路と接続する交差点から隣接する交差点までの道路部分

(2) 二カ所以上で特徴交差点と接続

道路全体 (3.2.2 節の定義で示された始点から終点まで)

図 3.3は、このアルゴリズムで選択された道路である。中央の○が目的地、四隅の○が Step 1 で求められた交差点であり、2本線で描かれた道路が選択された道路である。

出発地のブロックにおける道路選択では、このアルゴリズムの Step 3 を「出発地から主要な交差点までの経路を求める。」と変更する。

(2) 上位層における道路選択

上位層における道路選択のアルゴリズムを次に示す。

【上位層における道路選択アルゴリズム】

(入力) 出発地と目的地が属するブロックおよび上位層における出発地から目的地までの経路

(出力) 上位層で略図を構成する道路

(方法)

Step 1. 出発地および目的地が属するブロックに接続する交差点と、経路上に存在する交差点を求める。

Step 2. Step 1 で求めた交差点に接続する道路を抽出する。

Step 3. Step 2 で抽出されたそれぞれの道路について、相互に接続する道路と交差点を求める。

Step 4. Step 2 で抽出されたそれぞれの道路に対して、Step 3 で求めた交差点の中で、道路の始点側に最も近い交差点と終点側に最も近い交差点を求める。

Step 5. Step 4 で求めた交差点間の道路部分を抽出する。 ■

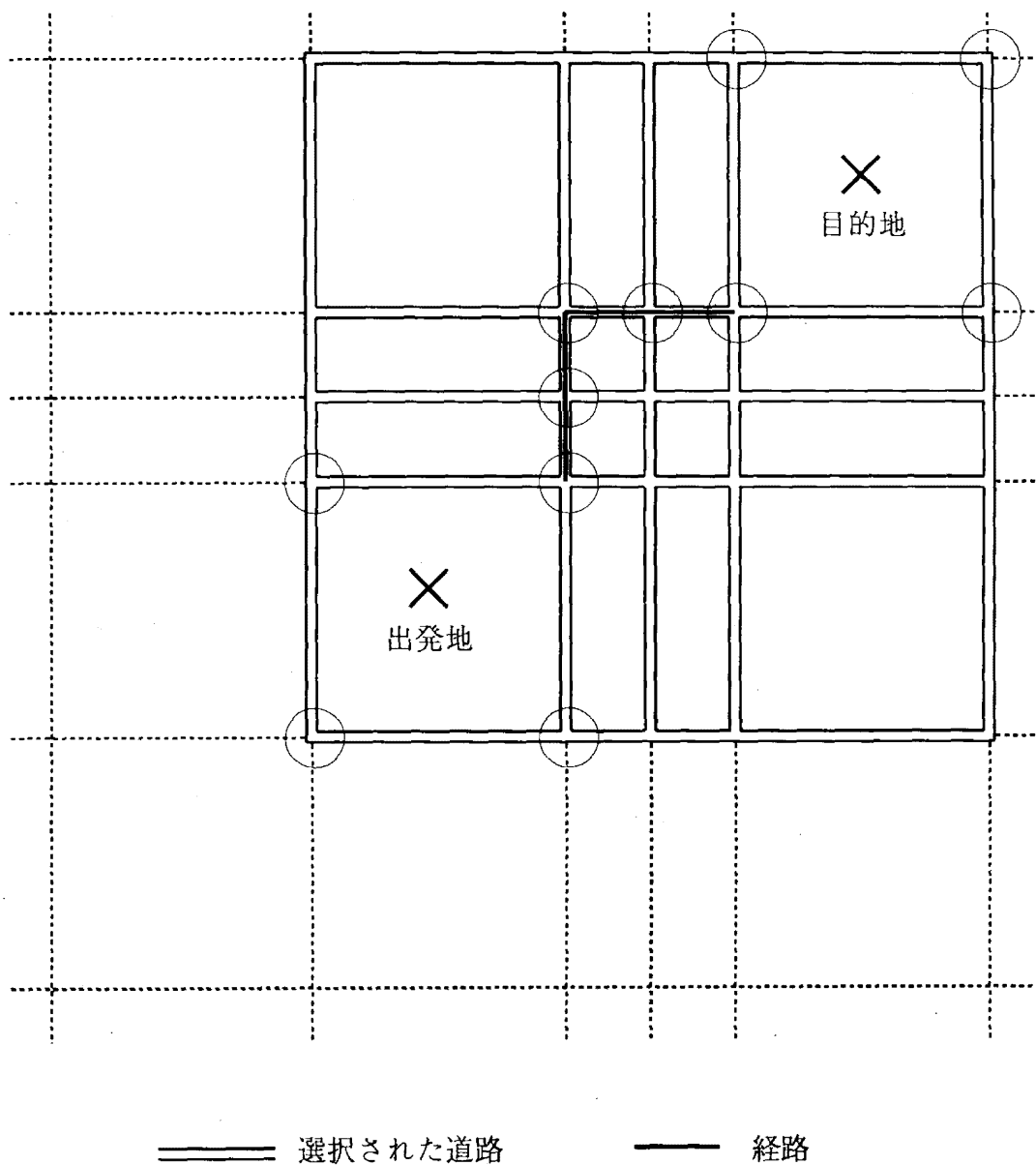


図 3.4: 上位層で選択される道路

図3.4は、上位層で略図を構成する道路を選択した例である。○はStep 1で求めた交差点で、2本線で描かれた道路が選択された道路である。

(3) 略図の構成

以上の手続きで得られた道路図を用いて、経路を表す略図を構成する。

【略図の構成アルゴリズム】

(入力) 上位層および下位層で選択された道路

(出力) 略図

(方法)

- Step 1. 入力された道路について道路網のネットワークを構成する。
- Step 2. 道路網のネットワークに対して、隣接する交差点間の方向をそれぞれ8方向に求める。
- Step 3. Step 2で求めた方向に矛盾しないように交差点を格子状に配置する。
- Step 4. Step 3で求めた交差点の位置にもとづいて、入力された道路で略図を構成する。
- Step 5. 下位層で信号機が存在する交差点を記号で示す。
- Step 6. 地名・目印となる施設を略図に加える。 ■

図3.5は、左下の○を出発地、右上の○を目的地と指示し、その間の経路を表す略図である。幅の広い道路が上位層に属する道路、幅の狭い道路が下位層に属する道路である。同一地域に対する実際の地図が図3.6である。

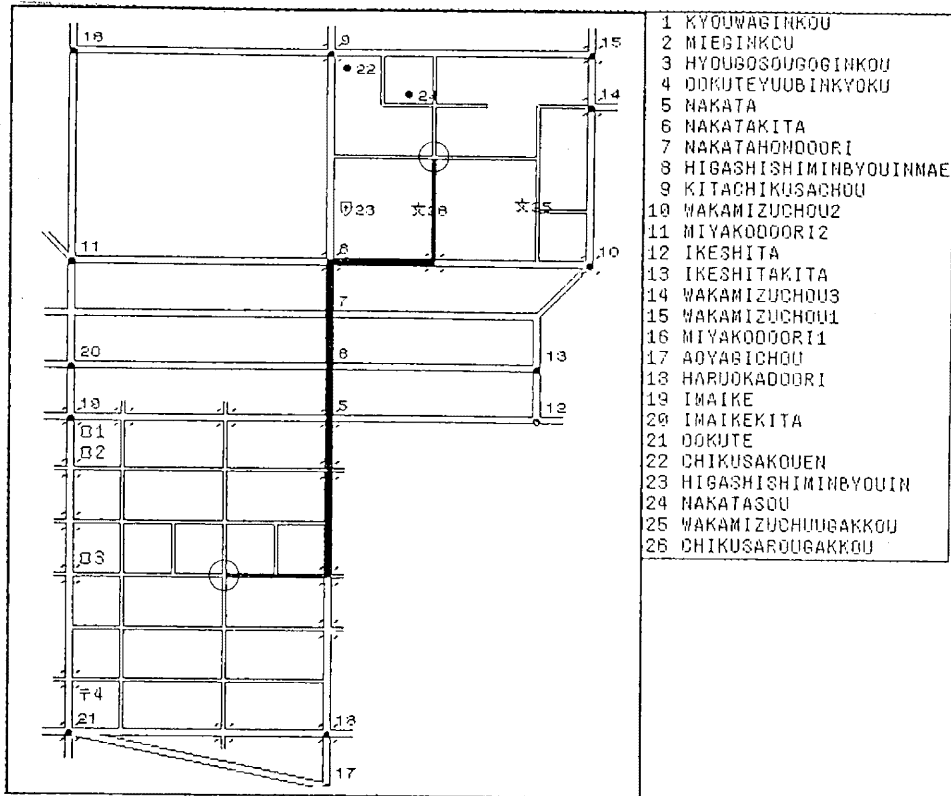


図 3.5: 経路を示す略図



図 3.6: 一般の地図

3.3.3 目的地を表す略図

目的地を表す略図を構成する場合には、経路に沿って道路を選択できない。したがって、目的地付近の主要な交差点をまず求め、その交差点から目的地への経路にもとづいて道路を抽出する。

【目的地を表す略図の構成アルゴリズム】

(入力) 目的地

(出力) 略図

(方法)

Step 1. 目的地に最も近い目標となる主要な交差点を求める。

Step 2. Step 1 で求めた交差点から目的地までの経路を求める。

Step 3. Step 2 で求めた経路にしたがって【経路を表す略図】の構成法を使い、略図を構成する。 ■

この手続きにおいて、求めた経路から略図を構成する上位層の道路を選択するとき、【上位層における道路選択アルゴリズム】を次のように変更する。

Step 1. 目的地が属するブロックに接続している交差点、それに隣接する交差点、経路上にある交差点を求める。

これにより、目的地を表す略図の方が経路を表す略図よりも広範囲に渡って大域的情報を提示する。これは、ユーザが経路を表す略図では、少なくとも出発地の場所を知っているのに対して、目的地を表す略図では、まったく知らない場合もあり、より大域的な情報を示すためである。

図 3.7 の例は、この変更したアルゴリズムで選択された上位層の道路である。また、図 3.8 は、○で表された目的地を表す略図である。

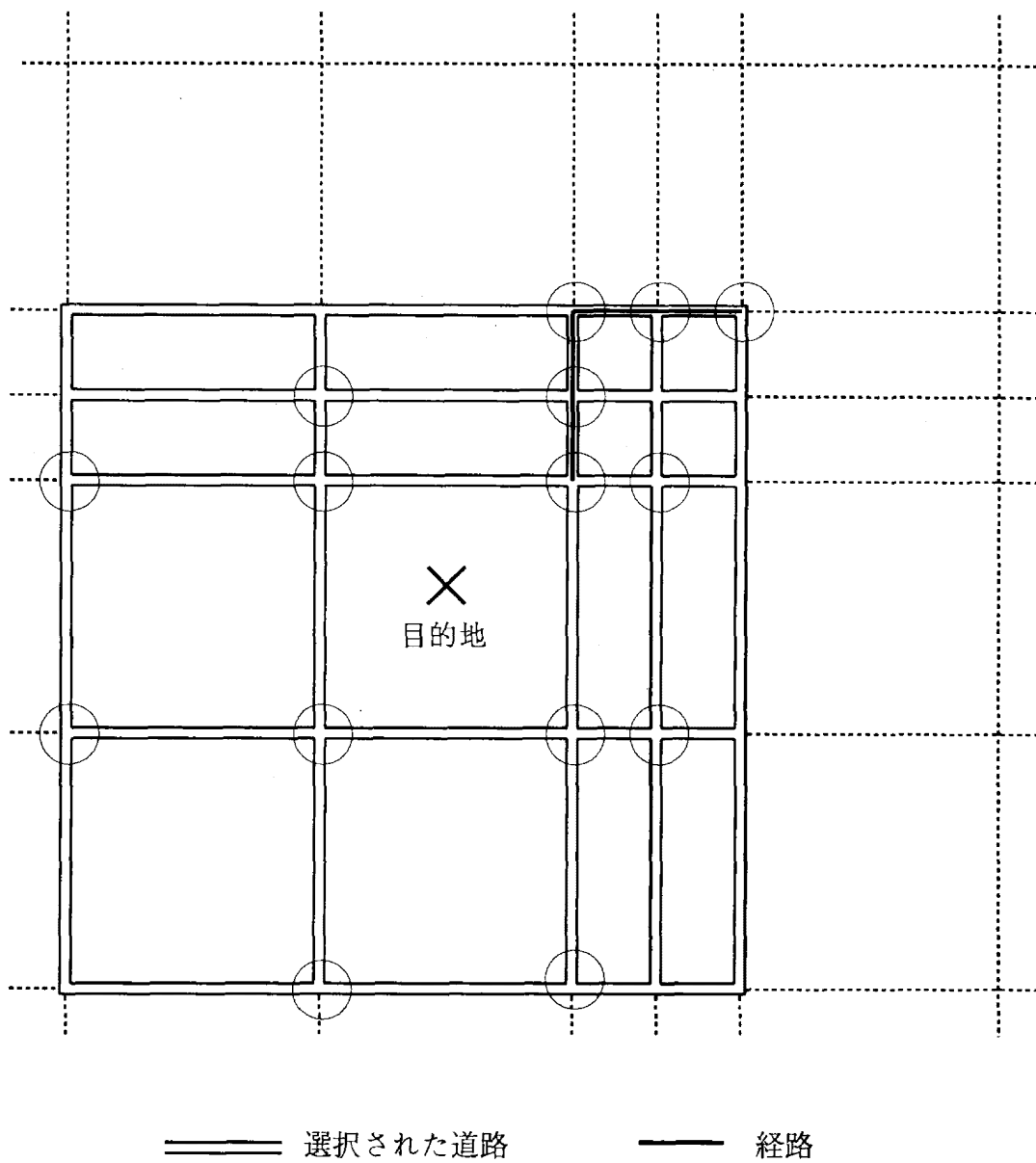


図 3.7: 上位層で選択される道路

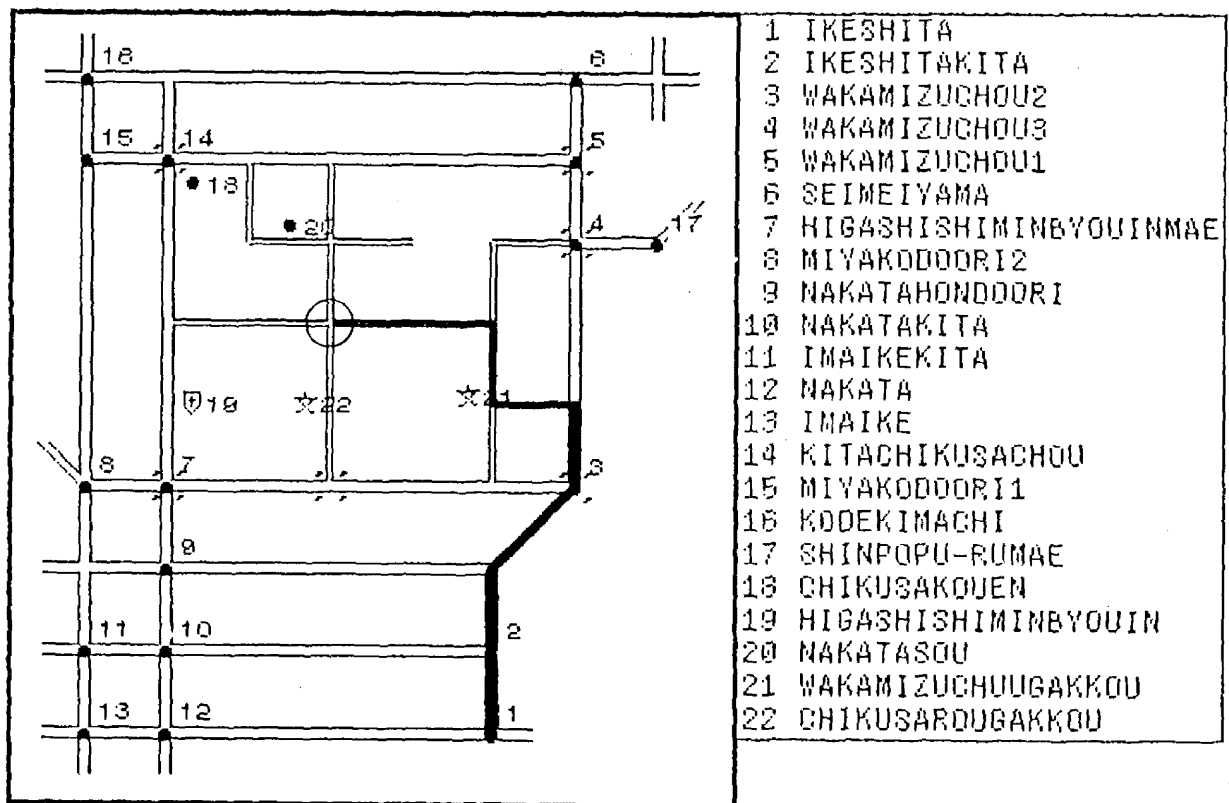


図 3.8: 目的地を示す略図

3.4 説明文の作成

道案内の経路情報を、音声や文章を使って説明する研究が行われているが [18, 19]、本節では、経路を表す略図を補完するために、経路に沿った道順を指示する説明文を作成する方法を報告する。説明文だけで道順を教えるのではなく、略図で表現された2次元平面の経路に時間的な進行状況を明示することにより、道案内の効果を向上させることが目的である [64]。したがって、まっすぐに進むときに会う目標物や交差点に関する情報など、略図から容易に判断される空間的情報は省略される。

説明文の作成手続きにおいても、道路網の階層性に注目して上位層と下位層に分けて文章を構成する。

【説明文の作成アルゴリズム】

(入力) 経路上の交差点リスト

(出力) 説明文

(方法)

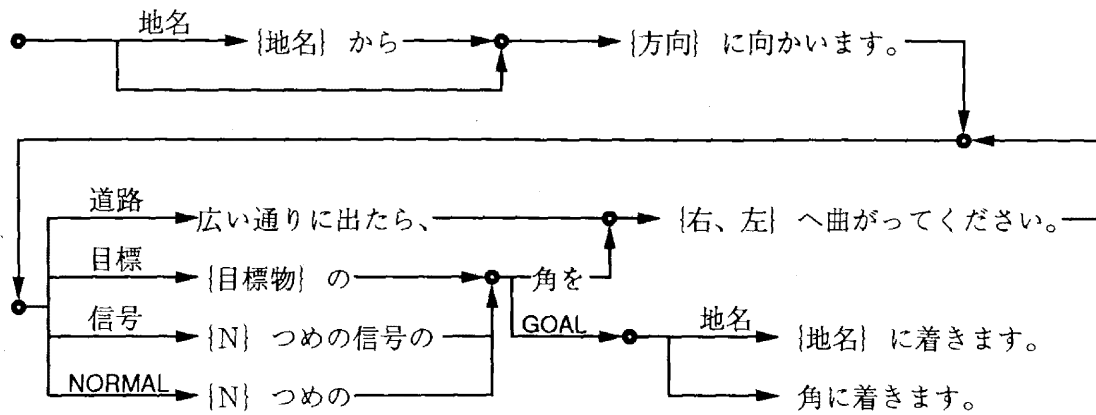
- Step 1. 経路上の交差点リストにもとづいて、それぞれの交差点で進むべき方向を調べる。
例 (SOUTH STRAIGHT LEFT STRAIGHT)
- Step 2. 経路上の特徴交差点を求め、特徴交差点における経路の情報を調べる。交差点の情報には、上位層の場合に進行方向、地名、道路名、交差点間の距離が、下位層の場合に進行方向、地名、目標物、右左折する交差点までの信号機や交差点の数、道路属性がある。図 3.9 に特徴交差点の情報リストの例を示す。リストは交差点番号、経路の方向、交差点の属性と属性値の組からなる。
- Step 3. 特徴交差点の情報リストより、図 3.10 のテンプレートにしたがって文章を生成する。このとき、特徴交差点が図 3.10 の矢印に示した属性を持つ場合には、その属性が付帯する矢印を選択する。また、{*} の中には、属性値が設定される。ただし、上位層の {距離} に対しては、1 km 以上の場合だけ、属性値が設定される。
- Step 4. Step 3 で構成した文章に対して、文と文の間に接続詞を挟み込み、一連のまとまった文章を作成する。次に、下位層と上位層の文章を接合する。このとき、同じ地名が続く場合には、後続の地名を代名詞に置き換える。 ■

((60 START (40 GOAL	(地名 仲田)(方向 北)(道路 仲田本通)(距離 0)) (地名 東市民病院)(距離 719))	上位層
((18 START (17 STRAIGHT (16 LEFT (23 GOAL	(方向 東)) (目標 (千種聾学校 東市民病院)) (信号 1)(目標 (千種聾学校 若水中学校)) (NORMAL 0 1)(目標 (千種聾学校 若水中学校)))	下位層

図 3.9: 特徴交差点情報リスト

図 3.11 は、図 3.5 の略図に対する説明文である。

下位層



上位層

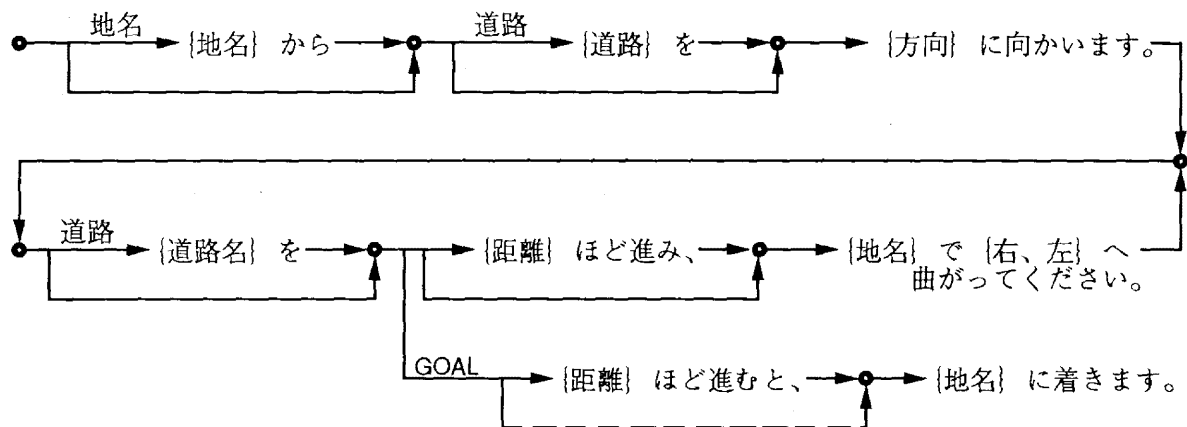


図 3.10: 文の生成規則

まず、出発地から東に向かいます。それから、広い通りに出たら、左に曲がります。すると、仲田にでます。

そこから、仲田本通を北に向かいます。東市民病院に着きます。

そこから、東に向かいます。そして、次の信号の千種聾学校、若水中学校の角を左に曲がって下さい。すると、次の千種聾学校、若水中学校の角が目的の場所です。

図 3.11: 説明文の例

3.5 実験結果と考察

本章で述べた各種のアルゴリズムにしたがって、道案内のための略図と説明文の作成システムを、ワークステーション XEROX 1100SIP で、Interlisp-D の下に実現した。地図の情報として、名古屋市千種区の地図 [21] (交差点数 4701、道路数 1732、地名・施設などの数 664) を使用した。

図 3.5、図 3.11 の例に示したように、略図と説明文を併用して提示することにより、経路の情報を、分かりやすく示すことができる。図 3.6 の実際の地図と比べると、不必要な情報が取り除かれて、提示したい情報はより簡潔に、かつ強調されて示されている。また、説明文を併用することによって、『3つ目の角を曲がる』という数の情報や『1 km ほど進む』という距離の情報を明示でき、略図で指示しにくい情報を簡単に扱うことができる。

略図の効果を調べるために次のような実験を行った。経路を示す略図および説明文を被験者の学生 20 名に与えて、その経路を実際の地図で記入させた。略図だけを参考にするグループ 10 名と、略図と説明文を参考にするグループ 10 名に対して実験した結果、説明文がない場合に正解者 7 名、説明文付きの場合に正解者 9 名であった。この実験で間違いの原因となったのは、実際の地図に信号機が示されていないために、略図で特徴的な交差点とした信号がある交差点の位置を間違えたことであった。この結果より、略図と説明文は道案内の情報として十分であると言える。

3.6 まとめ

本章では、道案内のための地図情報システムから検索・抽出された情報を、より分かりやすい形で表現するために、略図と説明文で提示する方法を述べた。略図とそれを補完する説明文は、実際の地図に比べて必要な情報を簡潔に表すことができる。少なくとも、本章で示した方法は道案内という応用目的に対して効果的な実現方法であり、実際の略図とほぼ同程度の機能性を有していると思われる。

本アプローチは、初めにある経路を求めて、その経路にもとづいて略図を構成している。このために、求められた経路により略図の形状、構造やそれに含まれる情報も変化する。本実験システムでは、経路は距離に一方通行などの交通規制や交差点での右左折のコストを考慮して検索するために、実用レベルの正確な情報提示が可能である。しかし、より分かりやすい略図を構成するためには、下位層における経路の検索方法を考慮する必要がある。例えば、なるべく明確な目標物が存在する経路を選択できることが重要で、そのための情報蓄積と検索方法を検討しなければならない。また、経路以外の道路も状況に応じて選別できる機能を実現して、略図に含む情報を精練させることも必要である。また、認知的科学的な視野から、情報提示の問題を検討することも必要である。

第 4 章

文字認識誤りモデルを用いた文字認識後処理機能の分析と評価

4.1 まえがき

文字認識の認識率を向上させるために、文法や語彙などの知識を用いた文字認識後処理が注目されている [34, 36, 43]。後処理は、文字認識部より出力される候補文字集合の中から、文法や語彙などの知識を用いて正しい文字列を選び出す方法である。

従来、文字認識後処理の性能評価としては、文字認識後処理後の文字認識率や訂正率(文字認識部が間違えた文字の中で訂正できた率)を用いることが一般的であった [33, 37]。しかし、文字認識後処理の入力データである文字認識部出力結果の認識率により、文字認識後処理後の認識率や訂正率は大きく変化する [46, 47]。そのため、これらで評価できることは文字認識後処理のある条件における特性だけであり、これだけでは十分な評価ができない。つまり、文字認識後処理の性能特性を分析するためには、文字認識部が出力する種々の認識率のデータに対する特性を分析する必要がある。しかし、認識対象画像の品質と処理結果の認識率との関係を正確に決定することはできないので、文字認識部から任意の認識率のデータを得ることは容易ではない。さらに、低認識率のデータを収集するためには、画像にノイズをかけたり、画像を回転させたりして故意に画像を劣化させる必要がある [48]。また、文字認識後処理の性能評価では、ターゲットとする文字認識部から出力される候補文字がテストデータとして使われる。しかし、候補文字集合の中に正解文字が含まれている確率や各候補文字に正解が存在する確率は、文字認識部の性能によって変化する。このように、文字認識系の条件によっても、文字認識後処理の結果は変動するので、文字認識後処理の性能を評価するためには、種々の文字認識部の出力データに対する特性も

分析する必要がある。したがって、文字認識後処理を評価するためには、文字認識部出力結果がどのような特性を持っているかを調べておく必要がある。

本章では、文字認識後処理を評価するための評価用データを容易に得るために、文字認識部が出力する候補文字データをシミュレートする文字認識誤りモデルを提案する。文字認識誤りモデルでは、文字認識部の出力において各候補文字に正解文字が存在する確率を、文字認識部出力の第1候補認識率から求める式を与える。このモデルにしたがって、評価用データを作成することができ、任意の条件における後処理の結果を比較することが可能となる。

以下、4.2節で文字認識誤りモデルを提案し、4.3節で文字認識後処理の分析評価法について述べる。4.4節で実際の評価例を示し、4.5節でまとめを行う。

4.2 文字認識誤りモデル

本節では、文字認識後処理を評価するための評価用データを容易に得るために、文字認識部が出力する候補文字データをシミュレートする文字認識誤りモデルを提案する。文字認識後処理で処理の効果に大きな影響を与えるものとして候補外文字がある。候補外文字があると、候補外文字自身だけでなく、その周囲の文字に対しても後処理により誤訂正されることがよくある。この影響による特性を評価できるように、文字認識誤りモデルを候補外文字を扱う候補外文字モデルと候補文字を扱う候補内文字モデルとから構成する。これらのモデルにより、文字認識部出力における各候補文字に正解文字が出現する確率(以降、正解存在確率と呼ぶ)と、候補文字の中に正解文字がない文字(以降、候補外文字と呼ぶ)の出現する確率(以降、候補外文字確率と呼ぶ)を推定する。そして、これらの確率にしたがって、既存のデータから文字認識後処理の評価用データを作成する。

4.2.1 候補外文字モデル

候補外文字モデルとして正解文字が候補文字集合に含まれない場合について考える。候補外文字出現の原因としては、認識対象画像にノイズがかかっている、認識対象画像が回転している、文字認識系の持っている文字の辞書と認識対象の文字が一致していない、文字切り出し誤りなどが考えられる。これらの原因に対して文字認識部が対策を行っているかどうかで、候補外文字の出現確率(候補外文字確率)は変化する。また、候補外文字確率と文字認識部の認識率(第1候補認識率)とは、相関があると考えられる。そこで、第 N 候補までに正解が入らない確率である候補外文字確率 P_x を次の式で表す候補外文字モデルを提案する。

$$P_x = (1 - P_1)^a \quad (4.1)$$

ただし、 P_1 は第1候補認識率を、 a は文字認識部の処理系によって決まる定数であり、文字認識系定数¹と呼ぶことにする。式(4.1)は、 P_1 が大きくなるにしたがって P_x が単調に減少する関数で、 P_1 が100%のとき P_x は0%となる。また、 P_1 が0%のとき P_x は100%となる。 P_1 が0%の場合は、文字認識部がまったく認識機能がない場合と考えられるので、 P_x が100%となるようにした。さらに式(4.1)で、 $(1 - P_1)$ は第1候補に正解が入らない場合であり、 P_x は第 N 候補までに正解が入らない場合(正解が第 $N + 1$ 候補以下に存在する)で

¹文字認識系定数 a は、候補文字の数 N によって変化する関数であるが、本論文では、候補文字の数 N も含めて文字認識の処理系と考えた。したがって、候補文字の数 N が違うものは違う処理系として扱う。

あるので、この2つの値に相関があることは妥当だと考えられる。

候補外文字モデルの妥当性を調べるために、実際のデータとの比較実験を行った。実験では、いろいろな認識率のデータを得るために、次のような操作を行って低品質の画像データを作成した。

- 文書画像にノイズを加える
- 文書を回転させる
- 文書画像の2値化レベルを操作する
- 文書を縮小コピーして文字につぶれを発生させる

このような画像データに対して、文字認識部を用いて処理を行い、処理結果の第1候補認識率 P_1 、候補外文字確率 P_x を調べた。実験は次の3つの文字認識部の処理系のデータとを比較した。

認識系1: 市街地距離による最小距離識別法 [49]

認識系2: ファジィクラスタリングを用いた構造化ニューラルネットワークによる文字認識法 [50]

認識系3: ハードクラスタリングを用いた構造化ニューラルネットワークによる文字認識法 [50]

図4.1に3つの認識系における文字認識結果の第1候補認識率 P_1 と候補外文字確率 P_x の関係を示す。ただし、実験では各認識系において1文字につき10個の候補文字を出力させた。

図4.1に示すグラフは、両対数グラフで描かれており、式(4.1)を表すグラフは、 $P_1 = 0\%$ 、 $P_x = 100\%$ を通る直線となる。このことから、図4.1に示された各処理系のグラフは、ほぼ式(4.1)の候補外文字モデルに一致する。図4.1から式(4.1)における文字認識系定数 a は、認識系1では1.71、認識系2では2.36、認識系3では2.74であった。この文字認識系定数 a は、式(4.1)から一意に決定できるので、それぞれの文字認識部の処理系に対して第1候補認識率 P_1 と候補外文字確率 P_x の値の関係を調べることで、容易に計算することができる。

このように候補外文字モデルでは、簡単な式で候補外文字確率を推定する。また、認識系による処理結果の違いをひとつのパラメータだけで吸収することができる。

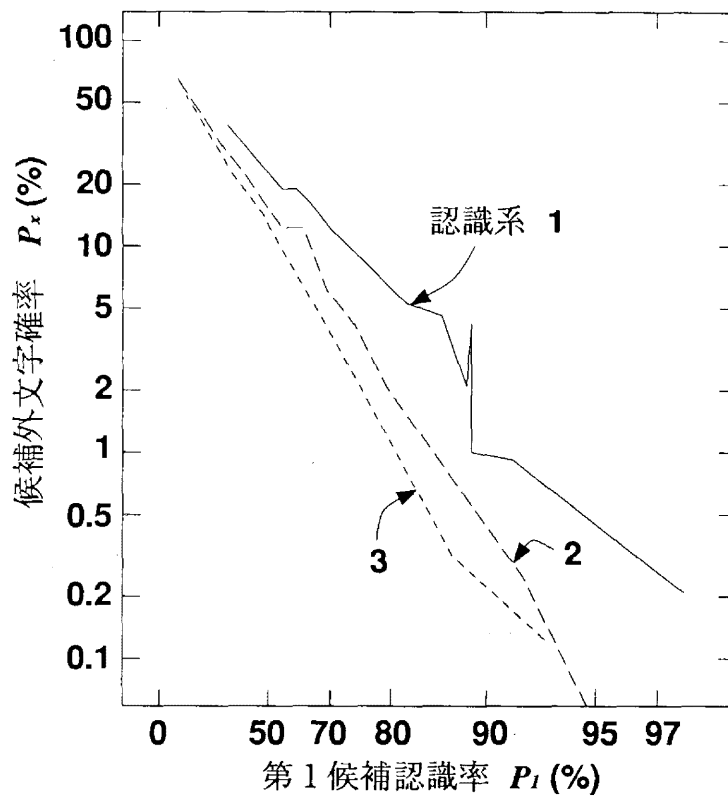


図 4.1: 実際のデータにおける候補外文字確率

4.2.2 候補内文字モデル

候補内文字モデルでは、正解文字が候補文字集合の中に含まれる場合について考える。文字認識部では処理結果として、1文字につき N 個の候補文字を出力する。これらの各候補文字が正解である確率を候補内文字モデルにより推定する。

第 i 候補文字の正解存在確率 P_i を次の式で表す候補内文字モデルを提案する。

$$P_i = \frac{P_1}{j^b} \quad (2 \leq i \leq N) \quad (4.2)$$

ただし、 P_1 は第1候補認識率、 N は候補文字の個数、 b は次の式(4.3)より求める。

$$\begin{aligned} \text{候補内文字確率} &= 1 - P_x \\ &= \sum_{j=1}^N \frac{P_1}{j^b} \end{aligned} \quad (4.3)$$

ただし、 P_x は候補外文字確率である。候補内文字モデルを用いれば、第 i 候補文字の正解存在確率 P_i は、第1候補認識率 P_1 から決定することができる。

式(4.3)は、第1候補認識率 P_1 が高くなれば b は大きくなり、第1候補認識率 P_1 が低くなれば b は小さくなる。これは、第1候補認識率が高ければ各候補文字の正解存在確率は第1候補認識率だけが大きく、その他の候補文字の正解存在確率は低いという急峻な分布になり、第1候補認識率が低ければ各候補文字の正解存在確率はなだらかな分布になることを示している。

候補内文字モデルの妥当性を調べるために、実際のデータとの比較実験を行った。実験では、候補外文字モデルの実験と同様に、3つの認識系に対していろいろな品質のデータを処理し、その結果をモデルを用いて計算した各確率と比較した。図4.2に、それぞれの認識系に対して、実験から得た実際の処理結果における各候補文字の正解存在率と、第1候補認識率から候補内文字モデルを用いて推定した各候補文字の正解存在確率とを比較したものの一部を示す。

図4.2から分かるように、実験から得た実際のデータ値と候補内文字モデルにより推定したデータ値は、よく一致している。また、他の認識対象文書を用いた実験でも、同様の結果となった。よって、候補内文字モデルを用いれば、各候補文字の正解存在確率を第1候補認識率から推論することができる。また、実際のデータ値を式(4.2)により最小自乗近似法で求めた定数 b の値と、式(4.3)から文字認識系定数 a と第1候補認識率 P_1 より求めた定数 b の値との比較を表4.1に示す。表4.1より、実験データから求めた定数 b と式(4.3)

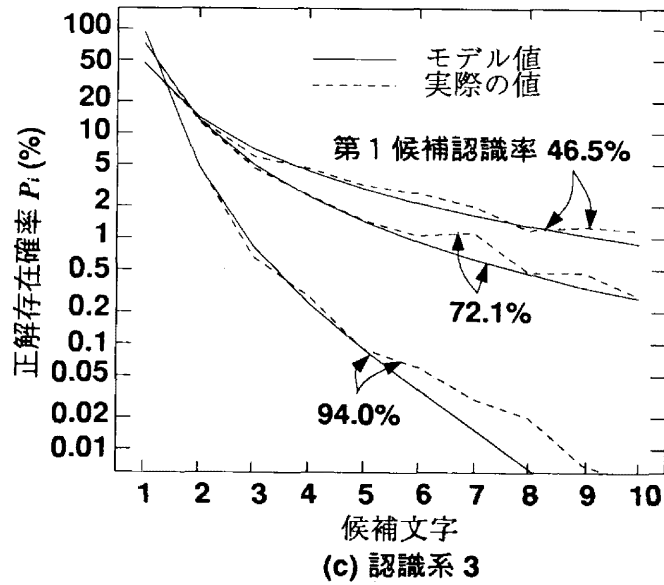
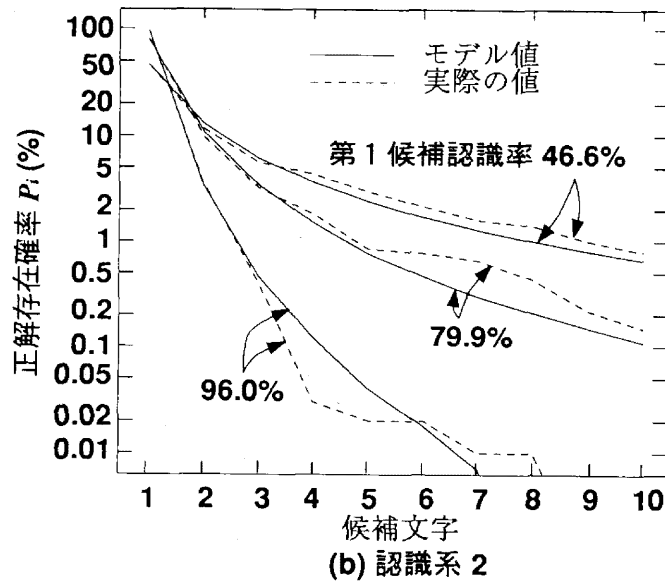
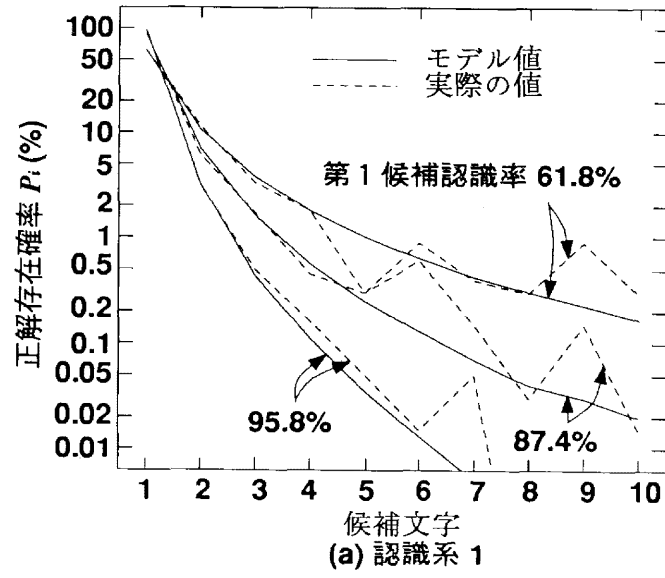


図 4.2: 候補内文字モデルと実際のデータの比較

表 4.1: 候補内モデルにおける定数 b の実際のデータとの比較

認識系	第1候補認識率 $P_1(\%)$	文字認識系定数 a	実験データから 求めた b	式(3)から 求めた b
1	95.8	1.71	4.92	4.91
1	87.4	1.71	3.76	3.64
1	61.8	1.71	2.64	2.55
2	96.0	2.36	4.78	4.85
2	79.9	2.36	2.87	2.87
2	46.6	2.36	1.86	1.84
3	94.0	2.74	4.45	4.30
3	72.1	2.74	2.48	2.43
3	46.5	2.74	1.74	1.72

から求めた定数 b はよく一致しており、文字認識定数 a が求めれば、式(4.3) から定数 b を求めることができる。このように、式(4.2) より第1候補認識率を決めれば、どの文字認識部の処理系に対しても各候補文字の正解存在確率を計算することができる。

4.3 文字認識後処理の評価法

本章では文字認識後処理の評価法の具体的な内容を述べる。

評価は、文字認識誤りモデルを用いて作成した評価用データに対して後処理を行い、その結果を分析することによって行う。評価項目は、文字認識部の認識率に対する依存特性、文字認識部の認識処理方法に対する依存特性を考える。以下、評価用データの作成法と各特性の評価について説明する。

4.3.1 評価用データの作成

評価用データの作成は次の手順で行う。

文字認識後処理の評価で、ある決まった文字認識部と組み合わせたときの性能を評価したい場合には、まず、文字認識系定数 a を求める。文字認識系定数 a は、候補外文字モデルの式 (4.1) から計算することができ、第1候補認識率 P_1 と候補外文字確率 P_x の値の関係を調べることで決定できる。次に作成したいデータの第1候補認識率 P_1 から、候補外文字モデルの式 (4.1) を用いて、候補外文字確率 P_x を求める。さらに、候補内文字モデルの式 (4.2)(4.3) を用いて、第 i 候補文字の正解存在確率 P_i を求める。

各候補文字の正解存在確率 P_i 、候補外文字確率 P_x の値が計算できたら、既存のデータから評価用のデータを作成する。ここで用いる既存のデータとは、実際に文字認識部で処理を行った出力結果であり、正解文字が分かっているデータである。

まず、各文字位置に対して、正解存在確率 P_i 、候補外文字確率 P_x にしたがって、正解文字がどの候補位置（第1候補～第 N 候補）あるいは候補外に入るかを定める。次に既存のデータにおいて、正解文字を決定された候補位置に配置し、他の候補文字を順にシフトして、評価用データを作成する。ただし、正解文字が候補外文字に入ると決定された場合は、正解文字を候補文字から取り除き、候補文字を順にシフトさせて、第 N 候補には空白文字を入れる。

図 4.3 に既存のデータから評価用データを作成した例を示す。図 4.3 では、文字認識系定数 $a = 1.71$ 、候補文字の個数 $N = 10$ の認識系に対して、第1候補認識率 $P_1 = 0.853$ のときの評価用データを作成した例である。まず、式 (4.1) から候補外文字確率 P_x は、

$$P_x = (1 - P_1)^a = (1 - 0.853)^{1.71} = 0.0377 \quad (4.4)$$

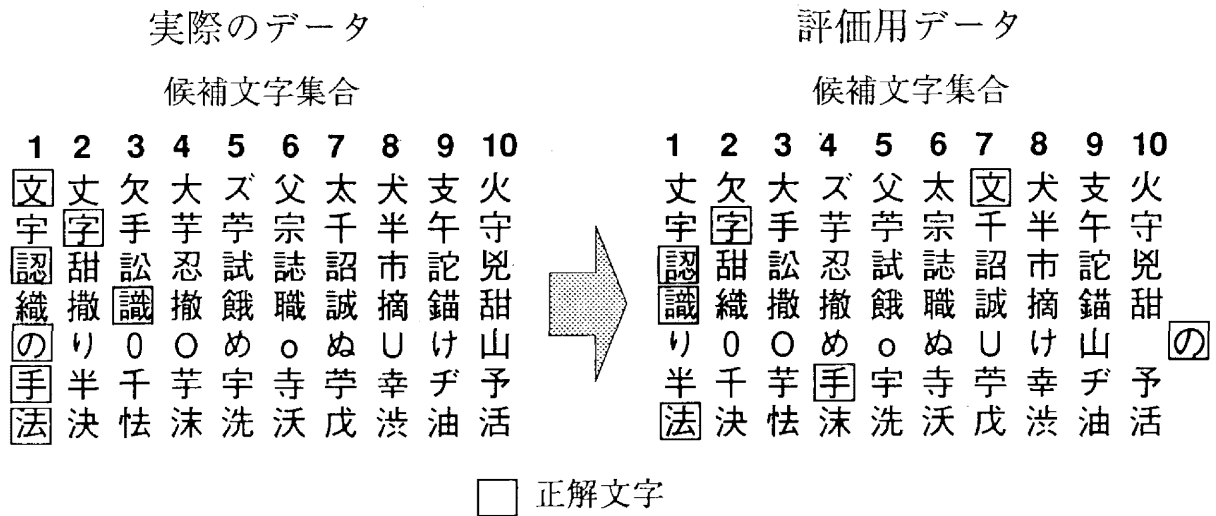


図 4.3: 評価用データの例

式(4.3)から定数 b は、

$$\sum_{j=1}^N \frac{P_j}{j^b} = 1 - P_x$$

$$b = 3.48 \quad (4.5)$$

となり、式(4.2)から各候補文字の正解存在確率 P_i を求めると、

$$\begin{aligned} P_2 &= 0.0766, & P_3 &= 0.0187, & P_4 &= 0.00685, \\ P_5 &= 0.00315, & P_6 &= 0.00167, & P_7 &= 0.000977, \\ P_8 &= 0.000614, & P_9 &= 0.000408, & P_{10} &= 0.000282 \end{aligned} \quad (4.6)$$

となる。次に、先に求めた正解存在確率 P_i 、候補外文字確率 P_x にしたがって、各文字位置の正解文字がどの候補位置あるいは候補外に入るかを定める。図4.3の例では、正解文字が1文字めは第7候補、2文字めは第2候補、3文字めは第1候補と決定された。評価用データの作成においては図4.3の例の場合、1文字めは、まず正解文字を第7候補に配置する。このとき、既存のデータでは図4.3から正解文字が第1候補にあるので、既存のデータにおける第2候補文字『丈』から第7候補文字『太』をそれぞれ1つずつ上位の候補にシフトする。以下同様にして評価用データを作成している。

4.3.2 文字認識部の認識率に対する評価

文字認識後処理の評価で、文字認識部出力の認識率と後処理後の認識率との関係を求めることができれば、文字認識後処理の文字認識部認識率に対する依存特性を明らかにすることができる。これにより、従来のある一定条件下での特性だけではなく、より詳細な特性を調べることができる。また、複数の文字認識後処理を比較する場合に、文字認識誤りモデルにおける文字認識部の文字認識系定数 a を同じ値として比較することにより、従来の文字認識システム全体の総合的な評価だけでなく、文字認識後処理についての性能を比較することが可能である。

文字認識部の認識率に対する評価は、まず、文字認識誤りモデルにおける文字認識部の文字認識系定数 a を一定の値にし、文字認識誤りモデルを用いていろいろの第1候補認識率をもつ評価用データを作成する。そして、これらの評価用データに対して文字認識後処理を行う。それぞれのデータに関する文字認識後処理後の認識率を分析することにより、文字認識後処理が高認識率向きに作られているか、あるいは低認識率においても安定した性能を持っているかが分かる。

また、現在ターゲットとしている文字認識部以外と組み合わせて文字認識システムを構成したときの文字認識後処理の性能を評価することも可能である。組み合せたい文字認識部の文字認識系定数 a を調べ、この文字認識系定数 a を用いて作成した評価用データを利用すれば、現在組み合せていない文字認識部であっても、その文字認識部とで構成したときの文字認識後処理の評価ができる。

4.3.3 文字認識部の認識処理方法に対する評価

候補外文字の存在は、文字認識後処理の性能を大きく左右する。同じ第1候補認識率であっても、文字認識部の認識処理によって候補外文字確率は変化する。候補外文字のできる原因としては、ノイズ、回転、つぶれ、かすれ、認識文字と認識系の持っている文字の辞書との違いなどがある。これらの原因に対して認識率が低下しないような対策を持っている認識処理方法では、候補外文字確率は小さくなる。このように文字認識部の認識処理方法によって、候補外文字確率が変化するので、文字認識後処理の性能を分析するためには、文字認識処理法を変えたときに、後処理後の結果がどう変化するかを調べる認識処理方法に対する特性を考える必要がある。

認識処理方法に対する特性の評価は、まず、文字認識部の第1候補認識率を一定にし、文

字認識誤りモデルの文字認識系定数 a を変化させた評価用データを作成する。そして、これらの評価用データに対して文字認識後処理を行う。それぞれのデータに関する文字認識後処理後の認識率を分析することにより、文字認識後処理がどのくらいの性能の文字認識部と組み合わせが可能であるかを評価できる。

4.4 文字認識後処理の分析評価実験

前章の文字認識後処理の分析評価法の実用性を検証するために、実際の文字認識後処理法に対して本評価法を適用して、分析評価の実験を行った。

ここでは、候補文節の選択法として本論文 5.3.2 節で述べる最適文節選択法を用いた後処理法 [70] (後処理法 1)、探索木を用いた後処理法 [37] (後処理法 2) の 2 つの後処理法に関する分析評価を行った。後処理法 1 は、候補外文字に関する対策を行っており、低認識率においても効果があると考えられる後処理法である。実験では、文字認識部の出力する候補文字を 1 文字につき 10 候補とし、評価用のデータを作成するための既存のデータとして、パーソナルコンピュータ Panacom 上の OCR 清書プログラム [51] の出力を用いた。文字認識誤りモデルにおける上記文字認識部の文字認識係数 a は 1.71 である。なお、認識対象文書は、科学分野の書籍および一般の書籍のデータを使用した。認識対象文書は、ある特定の分野に偏ったものばかりを使用しない方が良い。これは、文字認識後処理で使用している単語辞書中に特定の分野についての語彙数が少ないと、その分野の文書に関する後処理結果が悪くなり、文字認識後処理の性能を正確に評価することができないからである。

4.4.1 文字認識部の認識率に対する評価

文字認識部の認識率を変化させたときの後処理の特性の評価を以下のように行った。

後処理法1に対して、本論文で提案する文字認識後処理の分析評価法により評価した結果を図4.4に示す。実験では5つの文書(1文書平均約700文字)を既存のデータとして、それぞれのデータをもとにして評価用のデータを作成した。図4.4において、第1候補認識率と後処理後の認識率との差が、文字認識後処理により誤りが訂正されて認識率が向上した部分である。今回評価した文字認識後処理法は、入力された候補文字集合の組み合わせの中で誤り訂正を行うもので、後処理後の認識率は、第10候補以内認識率より高くはならない。図4.4から、文字認識後処理によって向上する認識率は、第1候補認識率が低下するにつれて徐々に低下していることが分かる。

次に、後処理後の認識率の向上がよく分かるように、第10候補以内認識率を基準として、どのくらい訂正ができているかを修復率として定義して分析を行った。修復率の定義は、

$$\text{修復率} = \frac{\text{後処理後認識率} - \text{第1候補認識率}}{\text{第10候補以内認識率} - \text{第1候補認識率}} \quad (4.7)$$

である。図4.5に後処理法1と後処理法2に対して、分析評価した結果を示す。図4.5に示した修復率は、5つの文書に対する修復率の平均である。図から分かるように、第1候補認識率が高認識率の場合には、2つの後処理法で修復率はほぼ同じである。しかし、第1候補認識率が低認識率の場合には、後処理法1の方が後処理法2より高い修復率を示している。2つの後処理法を従来の評価法で評価した場合には、第1候補認識率を一定にして評価を行うので、第1候補認識率が高認識率の領域で評価を行えば、2つの後処理法の性能は、同等と判定された。一方、本分析評価法によれば、候補外文字に対する対策が施してある後処理法1の方が、第1候補認識率が低い場合において、後処理法2より高い性能を持っていることが分かる。また、後処理法2では第1候補認識率が20%以下の場合には、修復率が0%以下(すなわち後処理を行うことにより認識結果を逆に悪化させている)になることも分かる。

このように本論文で提案する評価法により、従来の方法では不可能であった文字認識部の認識率に対する特性の評価を行うことができる。

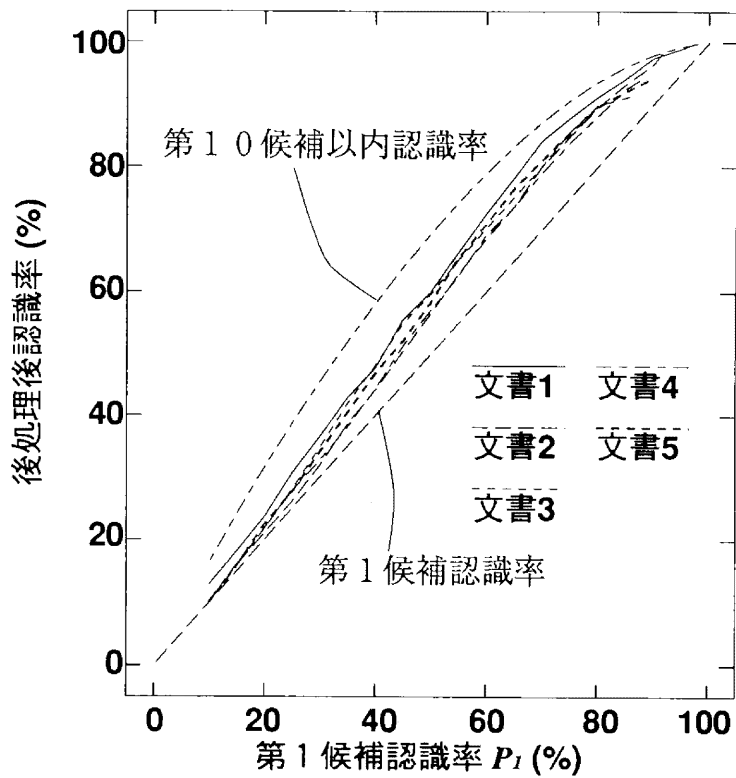


図 4.4: 文字認識部の認識率に対する特性

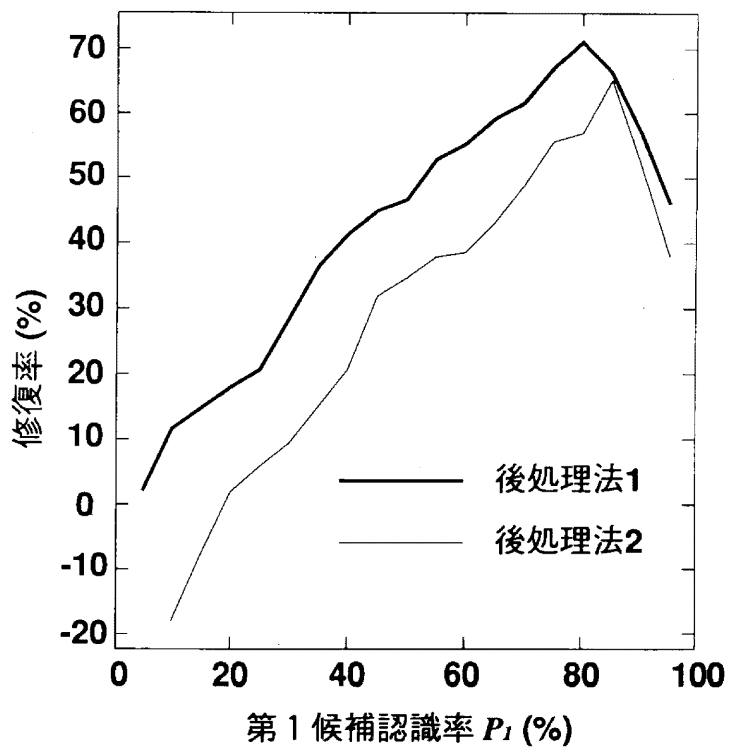


図 4.5: 文字認識部の認識率に対する評価

4.4.2 文字認識部の認識処理方法に対する評価

文字認識部の認識率に対する評価と同様に、後処理法1と後処理法2に対して、第1候補認識率を一定にして、文字認識誤りモデルにおける文字認識系定数 a を変化させたときの修復率の変化を図4.6に示す。図4.6は、第1候補認識率を90%と95%にして、5つの文書に対する修復率の平均を示したものである。図4.6から分かるように、第1候補認識率が90%、95%ともに後処理法1の方が修復率が少し高い。しかし、文字認識部の認識処理方法に対する傾向は、両後処理法ともほぼ同じ特性を示している。両後処理法とも、文字認識系定数 a が大きくなればなるほど修復率は上昇するが、文字認識系定数 a が一定の値以上であれば、ほぼ修復率は一定となる。第1候補認識率が95%のときには、文字認識系定数 a が1.6以上の処理系に対してほぼ一定の修復率を示し、第1候補認識率が90%のときには、文字認識系定数 a が2.0以上の処理系に対してほぼ一定の修復率を示している。よって、実験で用いた後処理法を用いた場合には、文字認識系定数 a が上記の値以上であれば、後処理法の最大の性能を出していると考えられる。したがって、第1候補認識率が95%での使用を考えるならば、文字認識部の文字認識系定数 a が1.6以上の処理系を、第1候補認識率が90%での使用を考えるならば、文字認識部の文字認識系定数 a が2.0以上の処理系を組み合わせることで文字認識システムを構成すれば後処理法の最大性能を発揮することができる。

このように本分析評価法により、実際に文字認識部と文字認識後処理を組み合わせることで文字認識システムを構成することなく、ある文字認識部と組み合わせたときの後処理の性能を評価することができる。これにより、文字認識部と後処理の組み合わせを設計することが可能となる。

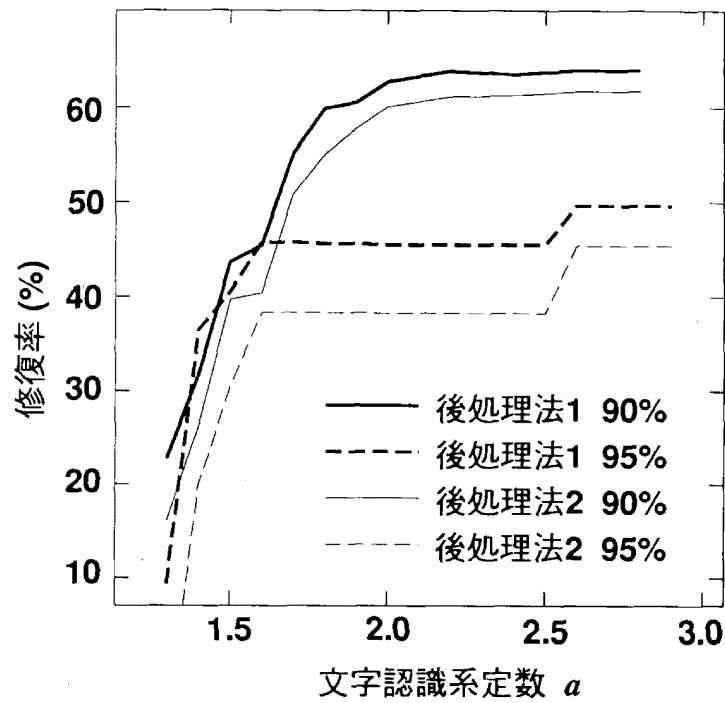


図 4.6: 文字認識部の認識処理に対する評価

4.5 まとめ

候補外文字モデルと候補内文字モデルを用いた文字認識誤りモデルを提案した。このモデルを用いることにより、任意の第1候補認識率を持つ評価用のデータを簡単に作成することが可能となった。この評価用データと実際に文字認識部から出力されたデータとを比較する実験により、文字認識誤りモデルが認識系によらず適用できることを確認した。

さらに、文字認識誤りモデルにより得られる任意の第1候補認識率の評価用データを用いて、文字認識後処理の文字認識部の認識率に対する特性、認識処理方法に対する特性を評価する方法を提案した。そして、実際の文字認識後処理法に対して本評価法を適用し、有効性を確認した。また、異なる文字認識部を持つ文字認識後処理に対しても、性能を比較することが可能であることを示した。本評価法を用いれば、文字認識後処理の適用認識率領域の判定、文字認識部を変更したときの性能の推測などの要求に十分応えることができると考えられる。

文字認識誤りモデルを用いれば、文字認識後処理における必要な候補文字の個数などの後処理の設計に応用することができると考えられる。さらに、文字認識後処理の入力データを出力する文字認識部についても候補外文字の出現確率の特性を評価するために応用できると考えられる。

第 5 章

キーワード情報を用いた文字認識後処理

5.1 まえがき

本章では、認識対象である文書の内容によって処理内容を変化させる文字認識後処理について述べる。

データベースの発達に伴い、高速で簡単なデータ入力の高まっている。この要求に答えるもののひとつとして文字認識が注目されている。しかし、日本語情報を扱う場合に、漢字は文字種が多く、類似文字も多いため、文書画像につぶれやノイズがあると、1文字ごとの認識処理だけでは、認識率の向上に限界がある。一方、人間が文書を読むときには、文法や語彙などの知識を用いることにより、少々のノイズがのった文書画像でも認識することが可能である。このような知識情報を用いた処理を文字認識に応用し、1文字ごとの処理では認識できなかった文字を訂正する文字認識後処理が最近注目されている [34, 36, 43]。

現在、後処理で用いられている知識には、一般的な文法や語彙などの知識がある。これらの知識情報を用いることにより、誤認識文字を 50 ~ 70% 訂正することができると報告されている [33, 37, 40]。しかし、手書き文字の認識など文字認識部での文字認識率が低い場合には、誤認識文字の訂正率は低下する。さらに、文字認識部での文字認識率がある値以下の場合には後処理を行うことにより、逆に、認識率を悪化させるといわれている。これは、訂正文字列の候補が多く、文法や語彙の知識だけでは正解を判断できないためである。一方、人間が文書を読む場合には、文書に書かれている内容を把握しながら文章を読むことにより、正しい文字列を判断することができる。よって、このような文書に書かれている内容を把握して、文字認識の後処理を行えば、文書の内容に沿ったより正確な文章に訂正することができると考えられる。また、文字認識後処理において認識率の向上を妨

げるものとして、文字認識部からの出力で候補文字の中に正解が含まれない文字(以降、候補外文字と呼ぶ)がある場合に、候補外文字の周辺に誤りが広がってしまうことがある。これに対処するために類似文字テーブルを用いて候補文字補完をする方法[33]が報告されている。しかし、漢字に対しては文字種類が多いため効果的な文字補完をすることは難しい。一方、人間の場合は不読文字をとばして、確実な文字から認識していると思われる。

本章では、文字認識後処理における以下の手法を提案する。ひとつは、文字認識部より出力された候補文字集合から、認識対象の文書に含まれるキーワードを自動的に抽出し、そのキーワードの知識を用いた後処理法であり、もうひとつは、候補外文字が周囲の文字に悪影響を及ぼせない文節の選択法である。これらにより、文字認識部での文字認識率が低い場合においても、後処理後の文字認識率を向上させることが期待できる。

以下、5.2節で文字認識後処理法の基本構成について述べ、5.3節でキーワードの知識と最適文節選択を用いた文字認識後処理法を提案する。5.4節で実験結果を示す。

5.2 文字認識後処理法の基本構成

文字認識後処理は、文字認識部が出力する文字に言語的知識を用いた処理を行うことによって、文字認識部の出力する誤った認識文字を訂正するものである。文字認識部から出力される文字は、一般にはいくつかの候補文字からなり、候補文字の中から最適な組み合わせを求めることにより誤り訂正を行う。

我々が提案する文字認識後処理の構成を図5.1に示す。文字認識部では、文字ごとの認識を行い、各文字に対して第1候補から第N候補までのN文字を候補文字として出力する。同時に、それぞれの候補文字に対して認識時の確からしさの度合である文字評価値も出力する。

後処理では、このN個の候補文字の中から文法・語彙・キーワードの知識を用いて、最も正しい組み合わせを出力する。後処理は以下の5つの部分からなっている。

- (1) 単語検索部 単語辞書を検索することにより、単語辞書内の単語に一致する候補文字の組み合わせを検索し、候補単語を出力する。
- (2) 文節検索部 文法辞書を用いて、文節を構成する単語の組み合わせを検索し、候補文節を出力する。
- (3) 文節評価値演算部 候補文節に含まれる文字の文字評価値、単語の長さ・頻度・品詞、文節の長さ、文書の内容を示すキーワード情報などから候補文節の確からしさを文節評価値として求める。
- (4) 文節選択部 文節評価値をもとにして、出力する文節を決定する。
- (5) キーワード抽出部 訂正文字列から文書の内容を示すキーワードを抽出する。

次節では、これらの処理の中で我々の後処理法に特徴的なキーワード情報の利用と最適文節選択法について述べる。

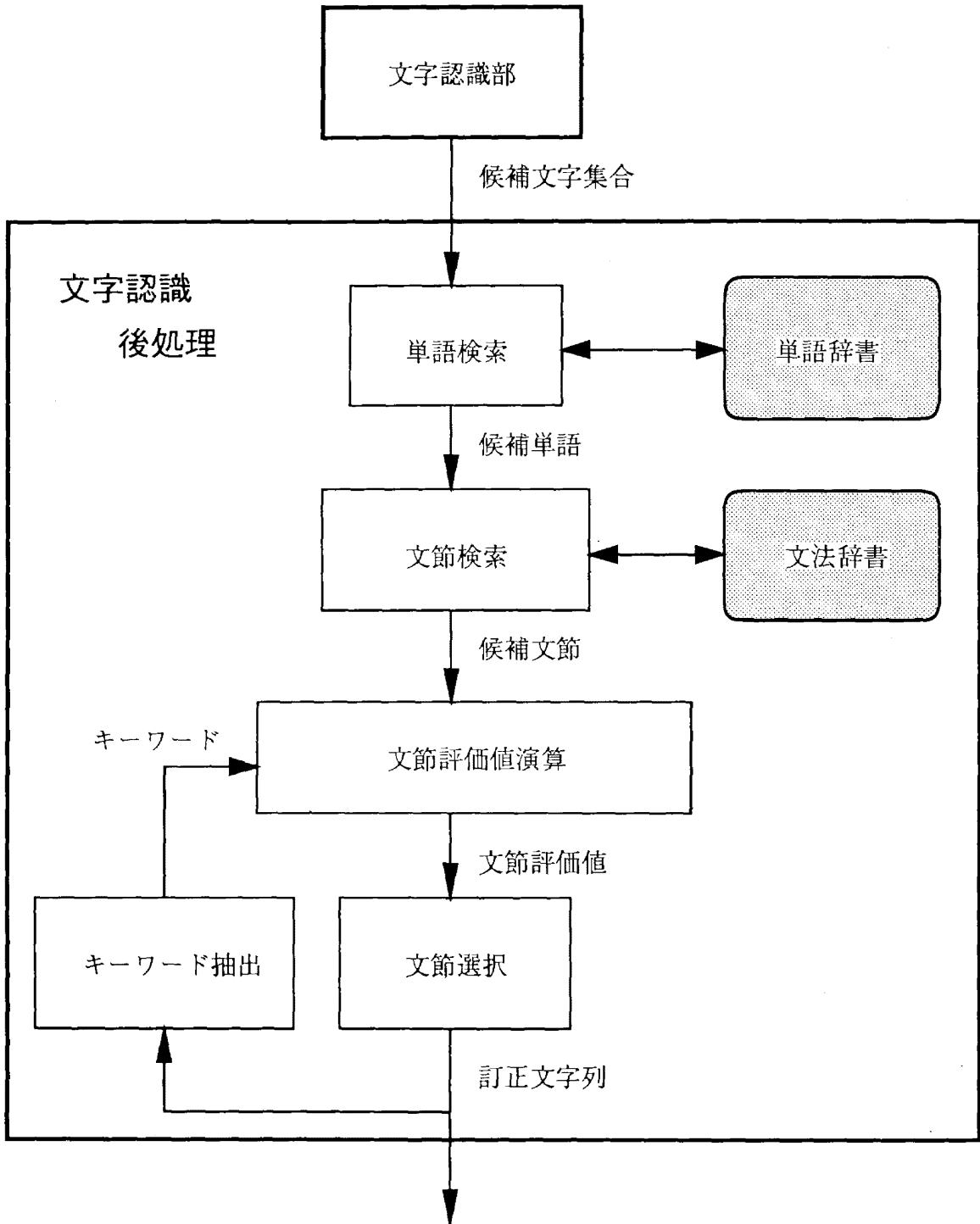


図 5.1: 文字認識後処理の構成

5.3 文字認識後処理法

5.3.1 キーワード情報の利用

文書の内容を表す情報として文書中のキーワードを抽出し、このキーワード情報の知識を用いて後処理を行う。キーワード情報を用いた後処理は、次の手順で処理される。

1. 文字認識部から出力された候補文字より、キーワード情報を用いずに後処理を行い、訂正文字列を得る。
2. 後処理結果の訂正文字列をもとにキーワードを抽出する。
3. 抽出された文書の内容を示すキーワード情報を用いて、再度、文節評価値を計算し直して後処理を行ない、訂正文字列を求める。

(1) Zipf の法則

キーワード情報の抽出は、Zipf の法則 [52] を応用して処理を行っている。Zipf の法則は、任意の文書中の単語の頻度分布に関する経験則で、第 1 および第 2 法則からなる。

(1) 第 1 法則 文書中に高頻度で生じる語についてその生起頻度 f の大きい順 r で並べたとき

$$f \cdot r = c \quad (5.1)$$

が成立し、生起頻度 f と順位 r とは反比例の関係にある。ただし、 c は文書に応じた定数である。

(2) 第 2 法則 文書中に低頻度で生じる語については、

$$\frac{I_1}{I_f} = \frac{f(f+1)}{2} \quad (5.2)$$

が成り立つ。ただし、 I_x は文書中に生起頻度 x で生じる単語の数である。第 2 法則は、低頻度の単語で同じ生起頻度の単語が複数ある場合に対応しており、このときは、生起頻度 x である単語の数 I_x が式 (5.2) を満たす。例えば、生起頻度 2 の単語の数 I_2 は式 (5.2) から $I_2 = \frac{1}{3}I_1$ となり、生起頻度 1 の単語の数の $\frac{1}{3}$ である。Zipf の法則はこのような単語の生起頻度に関する経験則であり、図 5.2 に Zipf の法則の概略図を示す。図 5.2 の実線がある文書における単語の頻度分布である。図の左側の領域では第 1 法則が成り立ち、右側の領域では第 2 法則が成り立っている。一般的に、第 1 法則と第 2 法則との境界領域 (図 5.2 の円で示

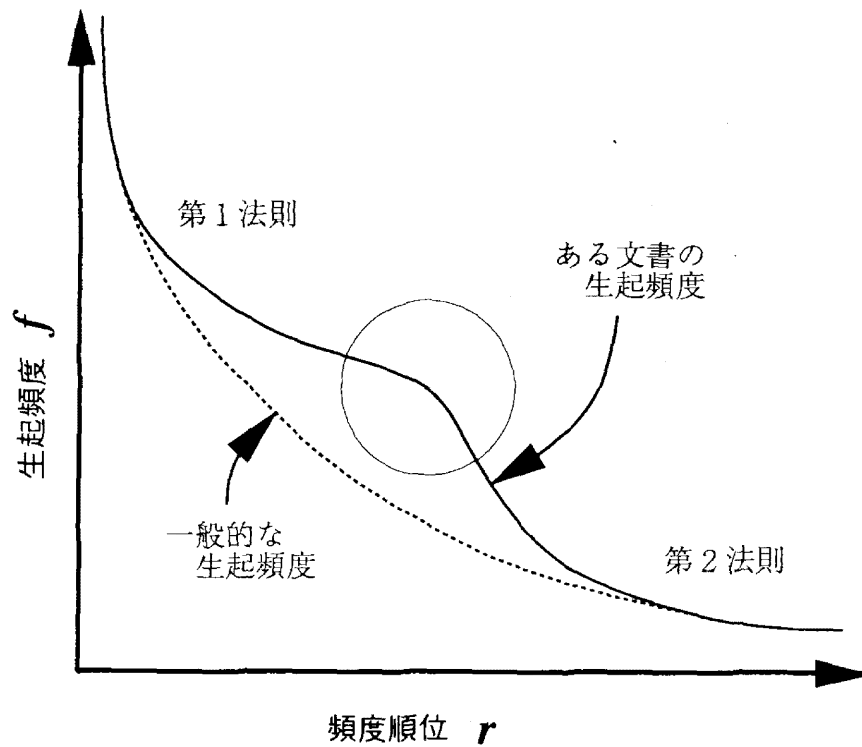


図 5.2: Zipf の法則

した領域) に含まれる単語が文書の内容を表すキーワードであると考えられている。これは、一般的に高頻度の単語はどの文書にもよく出現し、文書の内容を示すキーワードには成りにくく、また、低頻度の単語もキーワードには成りにくいことを表している。

さらに、あらゆる文書を集めた一般的な単語の頻度分布では、キーワードのような特徴的な単語がないために、図 5.2 の点線で示してあるような膨らみをもたない滑らかな曲線になることが知られている。

(2) キーワード情報の抽出

後処理において行うキーワード抽出は、候補文字集合という非常にあいまいで誤りを含んだ情報からキーワードを抽出しなければならないということで、一般のキーワード抽出法 [53] と大きく異なる。それで、図 5.3 に示すように、まず、候補文字集合を 1 度後処理することにより、訂正文字列と訂正文字列に含まれる文節の文節評価値を求める。文節評価値 E_j は、候補文節 j に含まれる文字の文字評価値、単語の長さ・頻度・品詞、文節の長さなどから候補文節 j の確からしさを求めたものである。ここで、訂正文字列に含まれる単語はまだ誤りを含んでいるため、単語の確からしさとして文節評価値 E_j を用いる。これは、文節評価値 E_j の高い単語ほど誤りでない確率が高く、キーワード抽出するための指標としてより信頼できると考えられるからである。また、Zipf の法則より文書中の単語の出現頻度と単語の一般的な出現頻度との差からキーワードが抽出できると考えられる。

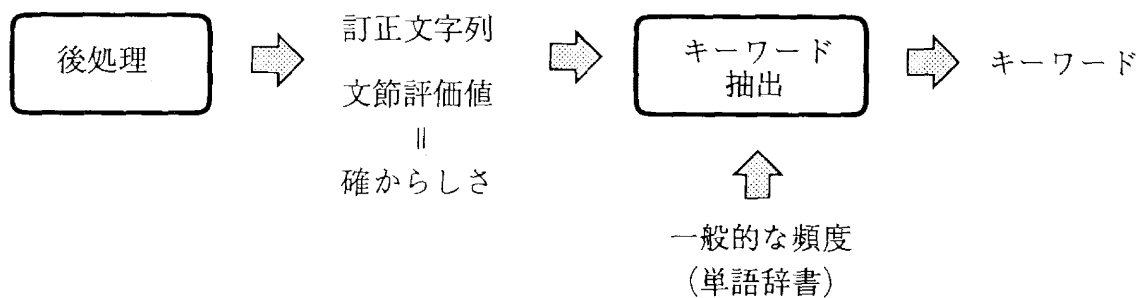


図 5.3: キーワードの抽出

このような考えから、次のようにして単語 w のキーワードへのなりやすさ K_w を求める。すなわち、文節評価値 E_j により重みづけられた文書中の単語の出現頻度と単語の一般的な出現頻度 F_w との差からキーワードへのなりやすさ K_w を求める。

$$K_w = \sum_{(\text{単語 } w \text{ を含む文節 } j)} \max((aE_j - F_w), 0) \quad (5.3)$$

ただし、 a は定数、 F_w は一般の文書に表れる単語の頻度統計から求まる値とする¹。定数 a は文節評価値 E_j と単語の一般的な出現頻度 F_w の大きさの整合をとるもので、キーワードにならないと思われる単語において $aE_j - F_w$ の値が 0 以下となるように、定数 a の値を実験により求める。

¹本章における実験では、松下電器製ワープロのかな漢字変換システムにおける累積変換回数¹の統計をもとにして、単語の出現頻度 F_w を決定した。

(3) キーワード情報を利用した後処理

キーワード抽出で得られたキーワード情報を用いて文節評価値を修正することにより、認識対象文書の内容を表すキーワードを把握した後処理を行う。文節評価値は、いくつかの要素から計算されるが、キーワード情報は認識する文章中の単語の頻度情報を反映しているものである。そこで、キーワード情報を単語の頻度に関する評価部分に組み込んで計算する。単語の頻度に関する評価値 E_{jh} は、単語の一般的な出現頻度 F_w にキーワードへのなりやすさ K_w を加えた形の式とする。これにより、認識対象の文書における単語の出現頻度に応じた評価ができる。さらに、単語の頻度が大きくなるほど、その値の相対的な差はあまり重要ではなくなるので、平方根をとる。よって、文節 j の単語の頻度に関する評価値 E_{jh} は、

$$E_{jh} = \frac{\sum_{(\text{文節 } j \text{ に含まれる単語 } w)} \sqrt{bK_w + F_w}}{(\text{文節 } j \text{ 中の単語の数})} \quad (5.4)$$

で計算する。ただし、 b は定数で、一般的な単語の頻度 F_w に対してキーワードへのなりやすさ K_w をどれだけ重視するかを調節するパラメータであり、文書のキーワード情報を重視する場合には、 b の値を大きく、キーワード情報をあまり使わない場合には、 b の値を小さくすることにより、キーワード情報の利用度を調節する。式 (5.4) の右辺において Σ を取っているのは、本手法では、接頭語や接尾語などを含んで構成されている複合語に対して、それぞれの語を別の単語として捉えているために、1文節中に複数の単語が存在することがあり、これらの単語に関する値の和を Σ で取っている。

式 (5.4) は、キーワードへのなりやすさ K_w が大きいほど、一般的な単語の出現頻度 F_w が大きいほど文節の確からしさを高くすることを意味する。このキーワード情報を反映した文節評価値を用いてもう1度訂正文字列を求めることにより、認識対象文書の内容を表すキーワードを把握した後処理ができる。

5.3.2 最適文節選択法

文節の選択は、候補文節の中から最も正解と思われる文節を選択する処理 [54, 55, 56] を行う。候補文節を選択するために重要なことは、文節の切れ目を正しく判断し、候補外文字の周辺に誤りが広がらないようにすることである。このことを実現するために、次の2つの手法を提案する。

1. 句読点を確実に抽出することにより文章を区切る。
2. 文章の中で最も確からしい文節から順に候補文節を選択していく。

これらの手法を用いることにより、候補外文字を含む文節が選択されるのは最後になり、候補外文字の周辺に誤りが広がることはなくなる。また、確実と思われる文節から選択されるので、これは、人間が不読文字の含まれる文書を読むときに、不読文字をとばして確実な文字から認識するのと同じ動作であると考えられる。

(1) 句読点の抽出

文節の選択を行うときに、処理をする単位を決定するために句読点を抽出する。これは、文節の区切りとして句読点が最も信頼できるからである。

ある文字が句読点であるかを決定するためには、第1候補文字だけから判断するのは適切でないので、ある文字のすべての候補文字から、その正解文字が句読点であるかどうかを決定する。句読点であるかどうかの判定には、同じ文字認識部で句読点を処理したときの候補文字集合の傾向を示す統計データを用い、その統計データと候補文字集合の類似度で判定する。句読点統計データは、それぞれの句読点に対して次の2つのデータからなる。

候補文字出現確率データ ある句読点 k が正解文字であるときに文字認識部が候補文字として出力する文字 c とその出現確率 G_{kc}

最多候補文字順位データ ある句読点 k が正解文字であるときに文字認識部が候補文字として出力する文字 c とその最多候補文字順位 R_{kc}

表 5.1 に正解文字が「。」のときの統計データの例を示す。表 5.1 から分かるように、必ずしも正解文字の出現確率 G_{kc} が最も高いわけではない。よって、句読点を抽出するには、すべての候補文字から判断する必要がある。統計データとの類似度は、文字認識部から出

表 5.1: 句読点の統計データ ($k = 『。』$)

候補文字 c	出現確率 G_{kc}	最多順位 R_{kc}
。	0.9	1
.	0.9	2
,	0.8	3
・	0.9	4
0	0.3	3
c	1.0	6
o	1.0	7
⋮	⋮	⋮

力される候補文字集合から文字位置ごとにそれぞれの句読点 k との類似度 T_{ik} を計算する。各候補文字に対して、統計データでの出現確率が高いほど、候補文字順位が高いほど、統計データの最多候補文字順位との差が小さいほど、正解文字が句読点であるとの決定に影響が大きいので、より大きく評価するのが望ましい。そこで、文字位置 i の文字が句読点 k である類似度 T_{ik} は、第1位から第 n 位までの各候補文字について候補文字出現確率データ、候補文字の順位、最多候補文字順位との差をそれぞれ評価し、それらの積を求め、得られた値の第1位から第 n 位までの候補文字に関する和をとることにより求める。類似度 T_{ik} の計算を式 (5.5) に示す。

$$T_{ik} = \sum_{m=1}^n \{G_{kC_{im}}(n+1-m)(n-|m-R_{kC_{im}}|)\} \quad (5.5)$$

ただし、 C_{im} は、文字位置 i の第 m 候補文字である。文字位置ごとに各句読点 k の類似度 T_{ik} があるしきい値を越えるかどうかを調べ、類似度 T_{ik} がしきい値を越える場合は文字位置 i を句読点の位置として抽出する。

このように句読点の統計データをもとにすべての候補文字集合の情報を用いて類似度を求めるので、正解文字が第1候補に上がってなくても句読点を抽出することができる。

(2) 最適文節選択

句読点抽出によって得られた句読点の位置で文書を分割し、それぞれの句読点で挟まれる文字列内において、候補文節の中から最適な文節を選択する。以下の手順で候補文節の選択を行う。

1. それぞれの候補文節 j に対して、図 5.4 に示すように接続評価値 H を求める。接続評価値 H は、文節 j の文節評価値 E_j 、文節 j の前に接続する文節 p の文節評価値 E_p 、文節 p と文節 j の接続のしやすさ S_{pj} 、文節 j の後ろに接続する文節 q の文節評価値 E_q 、および文節 j と文節 q の接続のしやすさ S_{jq} から式 (5.6) により求める。

$$H = E_j + E_p S_{pj} + E_q S_{jq} \quad (5.6)$$

2. それぞれの候補文節に対して H を最大にする文節 p, q を求め、そのときの接続評価値 H をその候補文節 j の接続評価値 H_j とする。
3. 句読点で挟まれる文字列内の候補文節において、最も大きい接続評価値を持つものを選択する。
4. 選択された候補文節を出力する修正文字列と確定する。
5. 選択された候補文節と文字位置が重なる候補文節を取り除き、句読点で挟まれるすべての文字列が確定するまでこれらの処理を繰り返す。

ただし、ここで用いられる接続のしやすさ S_{pj} は 0 または 1 の値をとり、文節 p の最後の単語の品詞と文節 j の最初の品詞の関係から値が決まる。 S_{pj} が 0 となるのは 2 つの文節が接続できないときであり、 S_{pj} が 1 となるのは 2 つの文節が接続できるときである。例えば、文節 p の最後の単語が動詞連用形で文節 j の最初の単語が名詞の場合 S_{pj} は 0 であり、文節 p の最後の単語が動詞連体形で文節 j の最初の単語が名詞の場合 S_{pj} は 1 である。

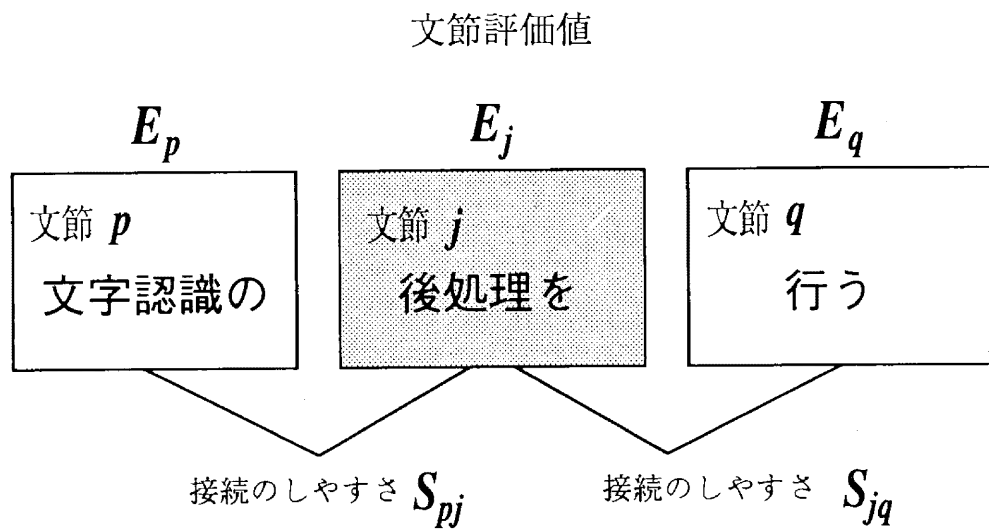


図 5.4: 接続評価値

5.3.3 実験結果と考察

文字認識後処理の実験結果を示す。後処理への入力データは、パーソナルコンピュータ Panacom 上の松下電器製 OCR 清書プログラムを用いて、第 10 候補までの候補文字を認識したものを使用した。計算機は、ワークステーション PanaStation 4/75(28.5MIPS,4.2MFlops) を使用した。単語辞書は、表記・品詞・頻度情報からなり、約 32,000 語が登録されたものを用いた。また、認識対象文書は、科学分野の書籍および一般の書籍から約 20 件の印刷文字データ(約 1 万 3 千文字)を使用した。認識対象文書の文字認識部での第 1 候補認識率は 77.5%~97.5%(平均 89.1%) で、低品質のデータから高品質のデータまでを用意し、種々の第 1 候補認識率をもつ文書に対して文字認識後処理の効果を調べた。

(1) 後処理法の基本効果

第 4 章で述べた文字認識の評価法を用いて評価を行う。後処理後処理結果の評価の基準として、以下に定義する修復率を尺度とする。本論文の後処理法では、文字認識部で候補文字になった文字の組み合わせのみで訂正文字列を構成するので、後処理後の最大認識率は、文字認識部での第 10 候補以内認識率である。よって、実験では第 10 候補以内認識率を基準にして、どのくらい訂正できているかを修復率として表す。修復率の定義は、

$$\text{修復率} = \frac{2 \text{ パス認識率} - \text{第 1 候補認識率}}{\text{第 10 候補以内認識率} - \text{第 1 候補認識率}} \quad (5.7)$$

である。ここで、2 パス認識率とは、キーワード情報を用いた後処理の結果の文字認識率である。第 1 候補認識率とは文字認識部での第 1 候補の認識率である。また、キーワード情報を利用せずに後処理をした 1 度めの後処理の結果の文字認識率を 1 パス認識率と呼ぶことにする。図 5.5 と表 5.2 に実験の結果を示す。図 5.5 は、各認識対象文書に対する文字認識部での第 1 候補認識率と文字認識後処理後の文字認識率との関係を示したものである。表 5.2 は、文字認識部の認識率と後処理後の認識率の詳細データを、文字認識部の認識率別に 5 文書について示したものである。

図 5.6 は、4.3 節で述べた文字認識誤りモデル [66, 67, 70](文字認識系定数=1.71) を用いて作成した評価用のデータに対して、後処理の効果を調べたものである。文字認識誤りモデルとは、文字認識部が文字を誤認識する仕方をモデル化したものである。このモデルにより、任意の第 1 候補認識率に対する各候補文字の正解存在確率を求めることができ、この正解存在確率にしたがって評価用のデータを作成することができる。図 5.6 により、文字

表 5.2: 後処理後の文字認識率

文書名	第1候補 認識率	1パス 認識率	2パス 認識率	第10候補 以内認識率	修復率	文字数	計算時間 (sec)	処理速度 (char/s)
1	97.5	99.8	99.8	99.8	100.0	473	6.8	69.6
2	94.8	98.3	98.5	99.9	72.2	848	13.1	64.8
3	88.9	93.8	94.0	95.8	73.9	836	9.7	86.2
4	85.6	90.3	91.2	94.7	61.5	548	4.7	116.6
5	77.7	85.9	86.5	92.6	59.0	631	9.5	66.8

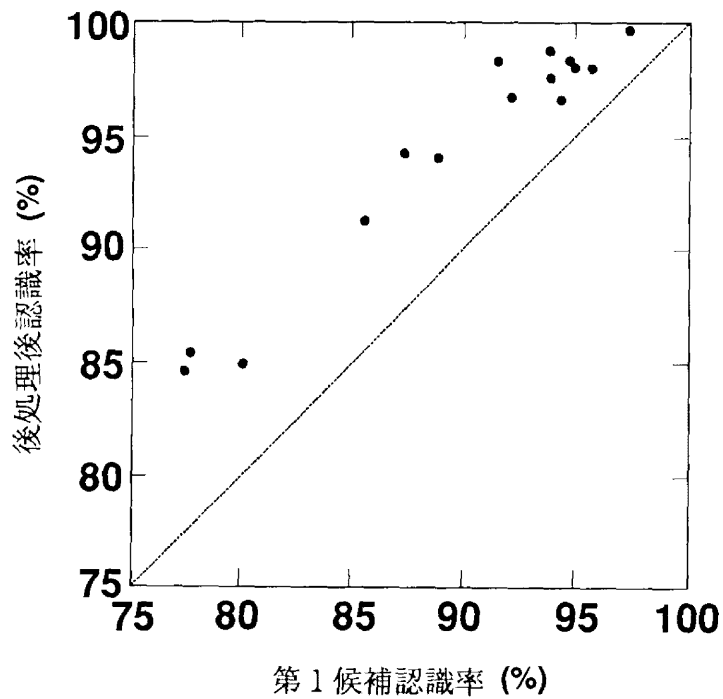


図 5.5: 第1候補認識率と後処理後の認識率

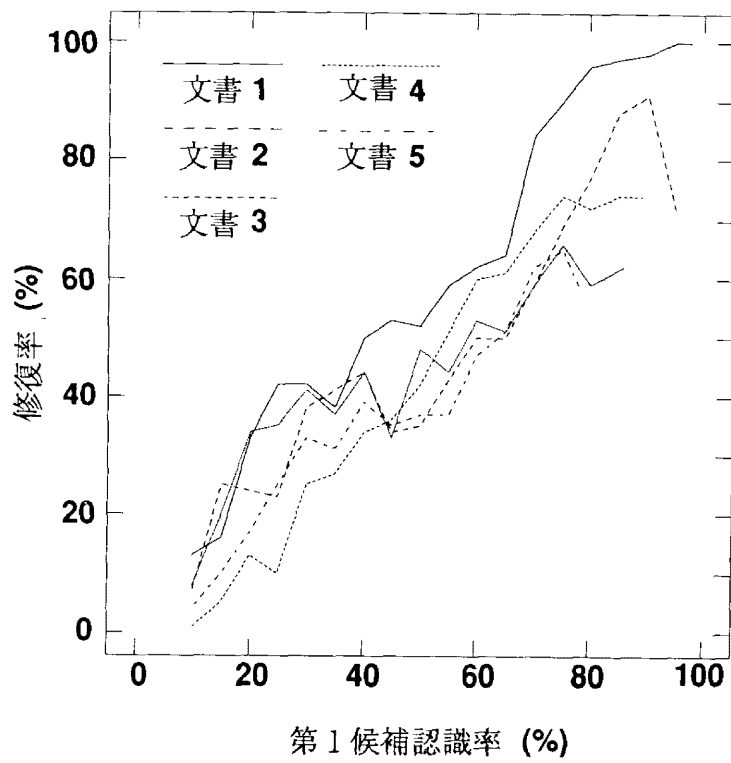


図 5.6: 文字認識部の認識率と後処理の修復率

認識部の第1候補認識率と後処理の効果との関係を知ることができる。つまり、文字認識部の文字認識率が非常に低い値のときでも、修復率は正の値となり、後処理を行うことによって文字認識率が低下しないことがわかる。また、文字認識部の文字認識率が高くなるほど修復率も高くなり、データによりバラツキはあるが、文字認識部の文字認識率が60%~70%以上であれば、修復率は50%を越える。図5.6において、第1候補認識率が大きくなるにつれて修復率が単調に増加していない部分がある。例えば、文書4における第1候補認識率45%の辺りで修復率のグラフは下に凸になっている。このようにグラフがデコボコになるのは、評価用データにおいて候補外文字となった文字が長い単語に属しているか短い単語に属しているかによって、修復率が変動するからである。

(2) キーワード情報利用の効果

抽出されたキーワードの例を表5.3に示す。表5.3は『ニューロ制御エアコン』に関する文書中の単語を、キーワードへのなりやすさ K_w の大きい順に並べたものである。この表より文書のキーワードとなる単語ほどキーワードへのなりやすさ K_w の値が大きく、キーワードが抽出できていることがわかる。一方、『もの』や『こと』などの単語は一般的な出現頻度 F_w が高いため、式(5.3)の演算の結果、キーワードへのなりやすさ K_w は0となり、キーワードとして重要でないとして出力されている。

表 5.3: キーワード抽出の例

キーワード	キーワードへのなりやすさ K_w
快適	1893
エアコン	1849
認識	1624
ニューラル	1535
⋮	⋮
もの	0
こと	0

また、キーワード情報を利用することにより、訂正が可能になった例を図5.7に示す。この例はキーワード情報を用いない1度めの後処理では誤って出力されたが、文書の他の部

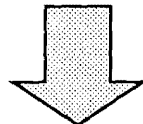
1 パス認識結果

現場学 省機能 文字 認識 前証言語

キーワード

文字, ニューロ, 学習, 認識, 言語

2 パス認識結果



現場 学習 機能 文字 認識 前記 言語

図 5.7: キーワード情報を利用した後処理

分から『文字』、『ニューロ』、『学習』、『認識』、『言語』がキーワードとして抽出され、2度目の後処理で『学習』、『文字』、『前記』と訂正することができている。このように、訂正文字列の候補が複数あり、従来法では正解が判断できない場合においても、キーワード情報を用いることにより、正解を得ることができた。

表5.2では、1パス認識率と2パス認識率の差はあまり大きくないが、これは、キーワード情報を利用して認識率が向上するのは、文字認識部において、認識結果の上位の候補文字の組み合わせに正解以外の単語が存在する場合だけだからである。しかし、キーワード情報を利用することにより、従来不可能であった文書の内容にもとづいた判断ができ、文字認識率が向上することが分かった。これにより、文書の内容に関する知識をキーワードとして抽出し、その知識を使って文字認識後処理を行うことの有効性が確認できた。

(3) 最適文節選択法の効果

句読点の抽出では、抽出された句読点が文節の絶対的な区切りとして用いられるので、確実に句読点として判断できるものだけを抽出する必要がある。文字認識部における句読点の第1候補認識率95.6%のデータによる実験で、統計データを用いた句読点の抽出は、句読点が正解文字であるものの96.8%を抽出することができた。抽出できなかった句読点は、すべて正解が候補文字に含まれていない候補外文字であった。また、句読点以外の文字を句読点として抽出することはなかった。一方、従来よく用いられている第1候補文字が句読点であり第1候補と第2候補との文字評価値の差があるしきい値以上のとき句読点と判定する方法[33]では、句読点の68.4%しか抽出することはできなかった。

ここで、本論文で提案した最適文節選択法の効果を、従来よく用いられている文節の探索木を作りながら前から探索していく方法と比較する。図5.8に文字認識誤りモデルを用いて実験した文字認識部の第1候補認識率と修復率との関係を示す。文字認識部の認識率が高い場合には、両手法で修復率の差はあまりない。しかし、文字認識部の認識率が低い場合には、本手法の文節選択法を用いた方が高い修復率を示し、従来手法では、第1候補認識率が約20%以下では修復率がマイナス、すなわち後処理をしたことにより認識率が低下している。これは、探索木で候補外文字以降の文節がうまく探索できず、候補外文字の周辺で誤訂正が起こるからである。一方、最適文節選択法では、候補外文字の周辺の文節に悪影響を与えないように処理していることから誤訂正は少なくなる。このため、候補外文字が多くなる認識率の低い領域で両手法の差が顕著になる。

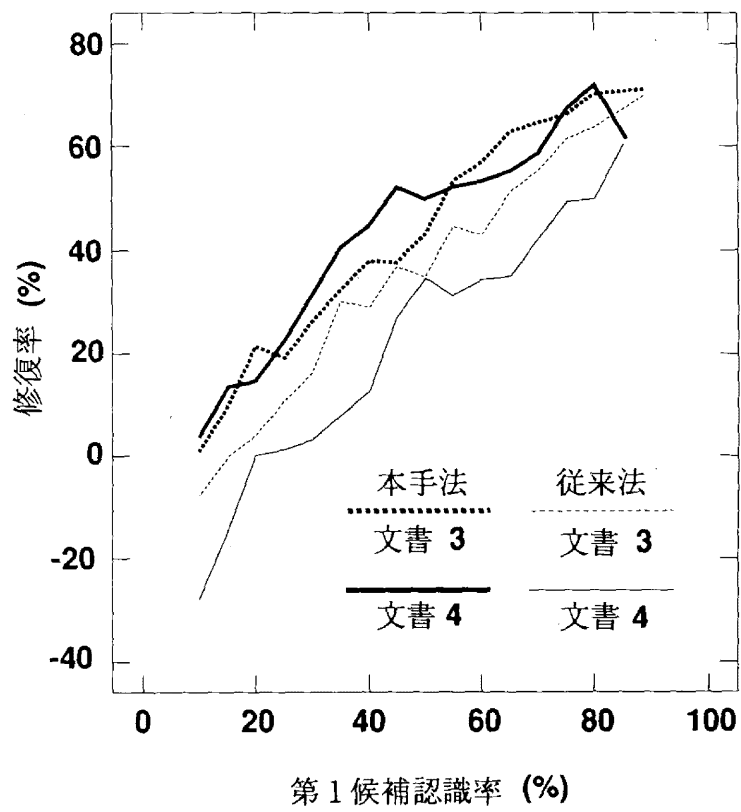


図 5.8: 本手法の文節選択の効果

(4) 後処理の例

図 5.11 に示す文書画像を文字認識部で認識し、その出力を後処理した例を示す。図 5.9 は、図 5.11 を文字認識部で認識した結果で、第 1 候補だけを示した。図 5.9 中で下線の引いてある文字は誤認識した文字である。図 5.10 は、図 5.9 を後処理した結果である。図 5.9 と同様に下線の引いてある文字は誤認識した文字である。この例では、文字認識部における第 1 候補認識率 94.47% が、後処理を行うことにより 2 パス認識率 99.17% に向上した。

全光演算方式ニューラルネットワーク

今回、実現した全光演算方式のニューラルネットワークは、新規に開発した「光ニューロン素子」と「2次元マルチレンズアレイ」および「学習メモリマスク」を組み合わせて構成したもので、光のみで並列的に演算処理を行い、文字や図形などの判断・認識をするものです。情報処理速度が速く（文字認識： ∞ 字/秒以上）コンピュータと接続する必要が全くないため、手軽に機器に組み込むことが可能です。

今後、情報処理やOA分野ではファックス、ワープロやコンピュータなどの入出力用端末装置に、またFAの分野では認識機能や判断機能を必要とする製造装置や検査装置などに「人工の目」として、広範な応用が期待できます。

量子化ニューロンを用いた機能別階層ネットワーク

手書き文字認識への適用

量子化ニューロンを用いた機能別階層ネットワークを英大文字の手書き認識に適用し、耐変動性に優れ、現場学習が可能なシステムを実現しました。システムが誤認識した場合には、現場で短時間に（1秒以下）追加学習させることができます。この現場学習機能を用いると、ユーザーの字癖等を現場で短時間に学習させ、ユーザーの文字に徐々に適応させることが出来ます。又、機能別階層ネットワークの構造は知識ベース処理の埋め込みに適しており、過去に得られた知見を有効に利用でき、他の認識判断処理や制御にも適用可能で、今後の応用発展が期待されます。

図 5.9: 文字認識部の出力（後処理前）

全光演算方式ニューラルネットワーク

今回、実現した全光演算方式のニューラルネットワークは、新規に開発した「光ニューロン素子」と「2次元マルチレンズアレイ」および「学習メモリマスク」を組み合わせて構成したもので、光のみで並列的に演算処理を行い、文字や図形などの判断・認識をするものです。情報処理速度が速く（文字認識：1000字/秒以上）、コンピュータと接続する必要が全くないため、手軽に機器に組み込むことが可能です。

今後、情報処理やOA分野ではファックス、ワープロやコンピュータなどの入出力用端末装置に、またFAの分野では認識機能や判断機能を必要とする製造装置や検査装置などに「人工の目」として、広範な応用が期待できます。

量子化ニューロンを用いた機能別階層ネットワーク

手書き文字雑音への適用

量子化ニューロンを用いた機能別階層ネットワークを英大文字の手書き認識に適用し、耐変動性に優れ、現場学習が可能なシステムを実現しました。システムが誤認識した場合には、現場で短時間に（1秒以下）追加学習させることができます。この現場学習機能を用いると、ユーザーの字癖等を現場で短時間に学習させ、ユーザーの文字に徐々に適応させることが出来ます。又、機能別階層ネットワークの構造は知識ベース処理の埋め込みに適しており、過去に得られた知見を有効に利用でき、他の認識判断処理や制御にも適用可能で、今後の応用発展が期待されます。

図 5.10: 後処理結果

全光演算方式ニューラルネットワーク

今回、実現した全光演算方式のニューラルネットワークは、新規に開発した「光ニューロン素子」と「2次元マルチレンズアレイ」および「学習メモリマスク」を組み合わせて構成したもので、光のみで並列的に演算処理を行い、文字や図形などの判断・認識をするものです。情報処理速度が速く（文字認識：1000字/秒以上）、コンピュータと接続する必要が全くないため、手軽に機器に組み込むことが可能です。

今後、情報処理やOA分野ではファックス、ワープロやコンピュータなどの入出力用端末装置に、またFAの分野では認識機能や判断機能を必要とする製造装置や検査装置などに「人工の目」として、広範な応用が期待できます。

量子化ニューロンを用いた機能別階層ネットワーク

手書き文字認識への適用

量子化ニューロンを用いた機能別階層ネットワークを英大文字の手書き認識に適用し、耐変動性に優れ、現場学習が可能なシステムを実現しました。システムが誤認識した場合には、現場で短時間に（1秒以下）追加学習させることができます。この現場学習機能を用いると、ユーザーの字癖等を現場で短時間に学習させ、ユーザーの文字に徐々に適応させることが出来ます。又、機能別階層ネットワークの構造は知識ベース処理の埋め込みに適しており、過去に得られた知見を有効に活用でき、他の認識判断処理や制御にも適用可能で、今後の応用発展が期待されます。

図 5.11: 後処理に用いた文書画像

5.3.4 後処理の問題点の分析

後処理を行った結果、最終的に誤った文字の分析を表5.4に示す。表5.4の表記で、候補外文字隣接は、候補外文字に隣接しているために誤った文字を、記号は、括弧などの記号が正解文字の場合に誤った文字を、その他は、単語辞書の未登録語、正解以外で正しいと思われる文節が存在することなどを表している。

記号に関する間違いについては、文法の知識で誤りを訂正できない文字(『と「,」と’)があり、これらの文字については、誤り訂正が困難である。候補外文字に隣接している文字に対しては、最適文節選択法により、候補外文字が含まれる文節以外の周囲の文字に悪影響が及ぶことはないが、同じ文節内の文字に対しては悪影響が及んでしまっている。よって、これらの文字を訂正するためには、候補外文字を含む単語を文書の内容などの知識から推論することにより、候補外文字の正解文字を推定しなければ訂正することはできない。

表 5.4: 後処理の誤り分析

文書名	文字数	候補外文字	候補外文字隣接	記号	その他
1	473	1	0	0	0
2	848	1	1	4	7
3	835	35	7	3	5
4	548	29	12	5	2
5	631	47	21	3	14
全体	12478	387	206	41	200

5.4 まとめ

文字認識後処理方法について述べ、実験により以下のことを明らかにした。

文書の内容に関する知識として、キーワード情報を取り上げ、文字認識部の出力のように誤りを含む曖昧な情報から、キーワード情報が自動的に抽出できることを確認した。そして、キーワード情報を利用した後処理の実験により、キーワード情報を用いることにより、後処理後の認識率が向上することを示した。さらに、認識率を向上させるためには、より高次の意味情報を自動的に獲得し、後処理に利用する方法を研究する必要がある。

統計データを用いることにより、句読点が第1候補でない場合においても句読点が抽出できることを示した。そして、最適文節選択法により、後処理後の認識率が向上し、文字認識部の認識率が低い場合においても最適文節選択法が有効であることが確認できた。

本後処理法は、文字認識部が低認識率においても有効であり、文字認識誤りモデルによる評価から、文字認識部の認識率が10%であっても、後処理を行うことにより認識率を上昇させることができる。

第 6 章

結論

6.1 本研究のまとめ

本論文は、知識情報処理における次の要素技術

- 知識情報処理を行うための知識のデータベースを人間が処理対象に対して持っている概念にもとづいた構造で表現する。
- 知識情報処理を行う場合に、知識情報処理を行う対象の状態を抽出し、その状態にしたがって処理を行う。

に対する研究成果をまとめたものである。

本研究では、知識情報処理の対象として、地図情報システムと文字認識後処理を選び、人間の直感的な思考の実現をめざした。

地図情報システムでは、市街地地図に含まれる道路網を道路のレベルにもとづいて階層化してデータを表現することを提案した。このようなデータ表現により、人間が行っている大局的な地図の認識と局所的な地図の認識を表すことができる。また、このデータ表現にもとづいた最適経路の探索方法を提案した。ここで探索される経路は人が地図を見て選択する経路に近く、探索のための計算量も少ないことを示した。さらに、地図情報システムで検索された情報をできるだけ簡単に分かりやすく提示するために、略図と説明文により地図情報を提示する方法を提案した。

文字認識後処理では、文字認識後処理の分析評価方法として、文字認識誤りモデルを用いる方法を提案し、文字認識後処理の評価を可能にした。また、文字認識における認識対象の文書の内容としてキーワードを取り上げ、キーワードを自動的に抽出し、その情報を

用いて文字認識後処理を行う手法を提案した。この手法によれば、文書の内容により後処理の内容が変わり、文書の文脈を理解しながら文字認識後処理を行う。さらに、文節の選択を確からしさの高いものから選択する手法を提案し、これらの手法を用いた文字認識後処理の有効性を示した。

本論文では、以上の研究成果を以下の5章に分けて述べた。

第1章では、研究の背景、従来の研究、本研究の方針について述べた。本研究では、柔軟な知識情報処理をめざして、研究の対象として、地図情報システムと文字認識後処理を取り上げた。地図情報システムでは知識表現の方法を、文字認識後処理では処理対象の状態にしたがって最適な処理を行う方法を目標とした。

第2章では、市街地地図の道路網を階層化してデータ表現する手法を提案した。そして、このデータ表現を用いて最適経路を検索する手法を提案した。これらの手法を用いた地図情報システムをワークステーション上に実現し、検索される経路が、人間が地図をみて選択する経路に似ており、検索時間も実用的であることを確認した。

第3章では、地図情報システムで検索された情報を略図と説明文を用いて提示する手法を提案した。略図により広範囲の大雑把な情報と狭い範囲の詳しい情報を同時に提示することができ、説明文により図では説明しにくい情報を表すことができることを示した。

第4章では、文字認識誤りモデルを提案することにより、文字認識部がどのような候補文字集合を出力するかをシミュレートできることを示した。この文字認識誤りモデルを用いて、文字認識後処理における性能を分析・評価する方法を提案した。実験により、文字認識部の認識率の変化や文字認識部の性能によって、文字認識後処理の性能がどう変化するかを分析・評価できることを確認した。

第5章では、文書の内容を把握するものとしてキーワードを取り上げ、文字認識後処理において、キーワードを自動的に抽出し、そのキーワード情報を用いて、文書の内容に即して文字認識後処理を行う方法を提案した。また、文節の選択を確からしさの順に選択する方法を提案した。これらの方法を用いた文字認識後処理により、後処理後の認識率が向上することを確認した。

6.2 今後の研究課題

より高次の知識情報処理を実現するためには、さらに次のような要素技術が必要である。

- ニューラルネットワークのようなパターン処理と知識情報処理を組み合わせることにより、パターン処理と記号処理のそれぞれの推論処理を統合し、柔軟な処理を可能にする。
- 知識情報処理における推論ルールを自動的に獲得し、推論対象にもとづいたルールで推論できることにより、柔軟な知識情報処理を可能とする。

次に、本論文で述べたそれぞれの研究に対する今後の課題について述べる。

本論文で述べた道路網のデータ表現は、地図情報の検索においては、非常に有効である。しかし、道路網データの更新が問題となる。更新される道路が上位の階層に含まれている場合には、その更新によって下位の階層のデータを再構成しなければならず、処理にかなりの手間を要する。さらに、道路の混雑など動的に変化する情報や各種の交通網（鉄道、バス、航空機など）の情報を階層の概念の中に取り込み、地図情報をより総合的に検索できる方法を研究する必要がある。地図情報の提示において、略図は非常にすぐれた方法である。本論文で述べた略図の構成法では、略図を道路情報のみをもとにして構成しているが、人が作成する略図では、略図に描かれている目標物の有効度（大きさ、知名度など）などによっても、略図の詳しさが変化していると考えられる。すなわち、有効度の高い目標物間の地図情報は大まかな情報のみで略図を構成する。このような目標物の有効度にも着目することにより、より分かりやすい略図の構成が可能となる。

本論文で、文字認識後処理において、文書内容の知識としてキーワードが有効であることを示した。しかし、文字認識後処理における問題の一つの候補外文字の訂正については、本研究では実現できていない。抽出できたキーワード情報を用いて、候補外文字の正解文字を推定することによって、文字認識後処理後の認識率が向上すると考えられる。さらに、文法や語彙的に不正確な部分を文字切り出し部にフィードバックすることにより、切り出し誤り位置を検出し、正しい切り出し位置を得ることができると考えられる。これにより、文字認識率も向上する。

謝辞

本研究を行うにあたり、数多くの方々にお世話になりました。お世話になった方々を以下に列記させて頂き、これらの方々には心より深く感謝致します。

本論文をまとめるにあたり、名古屋大学 杉江昇教授には励ましと懇切なるご指導を頂きました。

名古屋大学 稲垣康善教授、鳥脇純一郎教授からは、多くのご指摘とご指導を頂きました。

名古屋大学 渡邊豊英助教授には、励ましとご指導を頂きました。

中京大学 福村晃夫教授 (名古屋大学名誉教授)、名古屋大学 故吉田雄二教授には、熱心なご指導と数多くの助言を頂きました。

名古屋大学 杉野花津江助手には、暖かなご指導を頂きました。

また、名古屋大学工学部情報工学科 旧吉田研究室の皆様には、大変有益なご討論をして頂き、研究を進める上で多大なご協力を頂きました。心から感謝致します。

松下電器 中央研究所 新田恒治所長、電化本部 電化研究所 藤原義博所長、中央研究所電子機器基礎研究所 山下貞彦所長には、本研究を遂行する機会を与えて頂きました。

松下電器 中央研究所電子機器基礎研究所第2研究室 メ木泰治室長には、有益なご助言と本論文作成のための暖かいご配慮と激励を頂きました。

松下電器 コンピュータ事業部 横江亨課長には、OCR清書プログラムを提供して頂き、文字認識後処理の入力データとして役立てることができました。

松下電器 中央研究所電子基礎研究所 萱嶋一弘氏には、数多くの討論とご助言を頂きました。松下電器 中央研究所電子基礎研究所 山本浩司氏、小島良宏氏には、文字認識システムの作成で多大なお世話と数多くの討論を頂きました。

松下電器 中央研究所電子機器基礎研究所 丸野進氏、香田敏行氏、前川英嗣氏、今川太郎氏、伊藤哲氏、森家みち代氏、AV研究所 阪上茂生氏には、有益な討論とご協力を頂きました。また、日進ソフトウェア株式会社 杉本勲治氏、西川俊二氏には、文字認識システムの実現に際してご協力を頂きました。心からお礼を申し上げます。

最後に、学会・研究会などを通じて数多くの先輩諸氏の方々に暖かいご助言を頂いたことを感謝致します。

参考文献

地図情報システム関連

- [1] 坂内正夫：“地図データベース”，情報処理，Vol.27，No.10，pp.1153-1161 (1986).
- [2] 嶋田 茂，江尻正員：“日本語インタフェースを有する知識処理型マルチメディア地図情報処理システム GENTLE”，情報処理学会論文誌，Vol.27，No.12，pp.1162-1173 (1986).
- [3] 青木由直：“地域情報システム”，情報処理，Vol.27，No.10，pp.1162-1169 (1986).
- [4] 木戸出正継，恒川 尚：“画像情報処理におけるマルチメディアデータベース”，情報処理，Vol.28，No.6，pp.756-764 (1987).
- [5] 内藤重保，三枝博美，笠原 裕：“地理情報システムWINGによる都市空間管理”，第1回 AUTOCART JAPAN，1-1 (1985).
- [6] 伊理正夫監修：“計算幾何学と地理情報処理”，bit 別冊，共立出版 (1986).
- [7] 太田守重：“地理的情報の概念的データ構造”，第1回 AUTOCART JAPAN，1-3 (1985).
- [8] 嶋田 茂：“地図・図面情報処理におけるマルチメディアデータベース”，情報処理，Vol.28，No.6，pp.740-755 (1987).
- [9] Kobayashi,I.：“Cartographic Database”，Lecture Notes in Computer Science 80 (Chang, S.K. and Fu, K.S. eds.)，pp.322-350 (1980).
- [10] Sugie, M., Menzilcioglu, O. and Kung, H.T.：“CARGuide On-board Computer for Automobile Route Guidance”，Proc. of NCC 1984，pp.695-706 (1984).

- [11] 加藤誠巳, 大西啓介: “階層化されたデジタル地図データベースに基づく都心部自動車用経路案内システム”, 情報処理学会データベース・システム研究会資料, 89-DBS-72-19 (1989).
- [12] 杉本克行, 加藤誠巳: “有向ネットワークにおいて閉路を含まない k 個の最短経路を求めるための手法”, 情報処理学会論文誌, Vol.26, No.2, pp.356-364 (1985).
- [13] 加藤誠巳: “都市交通における最適経路情報提供システム”, 情報処理, Vol.28, No.3, pp.307-314 (1987).
- [14] 嶋田 茂, 岩村一昭: “ニューロコンピューティングによる最適道路経路の探索方式”, 情報処理学会第 38 回全国大会講演論文集, 7F-2 (1989).
- [15] 今井 修, 坂下裕明: “都市計画窓口情報システム 背景画像を用いた地理情報システムの応用例について”, テレビジョン学会技術報告, Vol.13, No.63, pp.9-14 (1989).
- [16] 吉野祐介, 吉村 斉, 田中 稔, 市川忠男: “地図情報システムにおける基本操作機能”, 情報処理学会コンピュータビジョン研究会資料, 32-3 (1984).
- [17] 今村賢治, 大山 実, 橋田幸雄: “音声による道案内システム”, 昭和 62 年電子情報通信学会総合全国大会講演論文集, 1642 (1987).
- [18] 今村賢治, 大山 実, 橋田幸雄: “音声による道路案内システムにおける出力文の生成”, 電子情報通信学会論文誌 (D), Vol.J71-D, No.5, pp.926-929 (1988).
- [19] 加藤誠巳, 本間 靖: “文章合成による市街地歩行経路案内システム”, 情報処理学会第 37 回全国大会講演論文集, 4R-3 (1988).
- [20] Berge, C.: “Theory of Graphs and Its Applications”, New York: John Wiley and Sons (1962).
- [21] “名古屋市区分詳細図 5 千種区”, アルプス出版社.
- [22] 伯田 晃, 高橋友一, 小林幸雄: “概念情報と視覚情報に基づく画像同定の一考察”, 電子情報通信学会人工知能と知識処理研究会資料, AI88-30 (1988).

- [23] 林 浩一, 河合和久, 上原邦昭, 豊田順一: “画像データベースの意味検索の試み — 道路地図データベースの略地図による検索 —”, 情報処理学会知識工学と人工知能研究会資料, 44-2 (1986).

文字認識後処理関連

- [24] 坂井邦夫, 西村一夫: “文字認識技術の発展とOCRシステムの広がり”, 東芝レビュー, Vol.47, No.2, pp.84-87 (1992).
- [25] 黒沢由明, 麻田治男: “文字認識技術の将来展望”, 東芝レビュー, Vol.47, No.2, pp.111-114 (1992).
- [26] 木田博巳: “メディア変換技術の最新動向 (その2. 文字認識技術)”, ビジネス・コミュニケーション, Vol.28, No.8, pp.16-20 (1991).
- [27] 磯山秀幸: “住所文字列に対する文字認識後処理方式の検討”, 情報処理学会自然言語処理研究会資料, 82-3 (1991).
- [28] 鈴木 章, 宮原末治, 小橋史彦: “住所認識装置の選択後処理方式”, 電子情報通信学会画像工学研究会資料, IE88-133 (1988).
- [29] 大井勝則: “住所読み取りにおける丁目・街区認識方式”, 電子情報通信学会パターン認識・理解研究会資料, PRU92-40 (1992).
- [30] 阿部圭一, 秦野和郎, 福村晃夫: “辞書を利用する文字認識系の能力の評価”, 電子通信学会論文誌 (C), Vol.52-C, No.6, pp.305-312 (1969).
- [31] 梅田三千雄: “単語辞書を用いた文字認識における文字の確定能力”, 電子情報通信学会論文誌 (D-II), Vol.J72-D-II, No.1, pp.22-31 (1989).
- [32] 山田洋志, 大山 裕: “大語彙かな漢字変換方式の変換率評価”, 情報処理学会自然言語処理研究会資料, 87-5 (1992).
- [33] 杉村利明: “候補文字補完と言語処理による漢字認識の誤り訂正処理法”, 電子情報通信学会論文誌 (D-II), Vol.J72-D-II, No.7, pp.993-1000 (1989).

- [34] 杉村利明, 斎藤珠喜: “文字連接情報を用いた読取り不能文字の判定処理 —文字認識への応用—”, 電子通信学会論文誌 (D), Vol.**J68-D**, No.1, pp.64-71 (1985).
- [35] 村木一至, 浜田和彦: “OCRの認識誤り訂正に於けるテキスト適合性の評価”, 電子情報通信学会パターン認識・理解研究会資料, **PRU92-41** (1992).
- [36] 北村 正, 仲林 清, 大光明直孝, 中村 修: “単語知識を利用した手書き文字列処理方式”, NTT R&D, Vol.**39**, No.3, pp.429-436 (1990).
- [37] 高尾哲康, 西野文人: “日本語文書リーダ後処理の実現と評価”, 情報処理学会論文誌, Vol.**30**, No.11, pp.1394-1401 (1989).
- [38] Kise K., Shiraishi T., Takamatsu S. and Kusaka H.: “Improvement of Text Image Recognition based on Linguistic Constraints”, Proc. of MVA'92, pp.511-514 (1992).
- [39] 吉村賢治, 竹内美津乃, 津田建蔵, 首藤公昭: “コスト最小法を用いた日本語文の形態素解析”, 情報処理学会自然言語研究会資料, 60-1 (1987).
- [40] 五十嵐治一, 松田 透: “形態素解析を利用したOCR誤読文字の自動訂正”, Richoh Technical Report, No.**15**, APRIL (1986).
- [41] 木谷 強: “手書き文書の文字認識結果に対する後処理方式”, 情報処理学会自然言語処理研究会資料, 86-1 (1991).
- [42] 小黒雅己, 中村 修, 北村 正: “手書き単語組織別のための最良優先探索方式”, 電子情報通信学会人工知能と知識処理研究会資料, **AI90-62** (1990).
- [43] 小黒雅己, 中村 修, 北村 正: “手書き複合語文字列識別のための最適単語組探索方式”, 電子情報通信学会論文誌 (D-II), Vol.**J75-D-II**, No.1, pp.96-102 (1992).
- [44] 紺野章子, 本郷保夫: “日本語OCRの後処理に関する一手法”, 電子情報通信学会パターン認識・理解研究会資料, **PRU92-21** (1992).
- [45] 伊東伸泰, 丸山 宏: “OCR入力された日本語文の誤り検出と自動訂正”, 情報処理学会論文誌, Vol.**33**, No.5, pp.664-670 (1992).

- [46] 大町真一郎, 阿曾弘具, 木村正行: “低品質文字の認識に関する検討”, 1992年電子情報通信学会春季大会講演論文集, D-561 (1992).
- [47] 大槻恭士, 伊藤彰則, 牧野正三, 曾根敏夫: “言語情報を利用した文字認識における文字認識率と単語認識率の関係”, 情報処理学会第44回全国大会講演論文集, 5N-9 (1992).
- [48] 倉掛正治, 石井健一郎: “変形文字パターンを用いた手書き文字認識系の評価”, 電子情報通信学会論文誌(D-II), Vol.J74-D-II, No.9, pp.1208-1216 (1991).
- [49] 脇 康, 藤原啓恵, 竹之内磨理子, 横江 亨, 清水正博: “文書清書システム(3) —文字認識アルゴリズムとそのハードウェア—”, 昭和61年度電子通信学会総合全国大会講演論文集, 1512 (1986).
- [50] 小島良宏, 香田敏行, 丸野 進, 阪上茂生, 山本浩司, 木泰治: “構造化ニューラルネットワーク(NARA)によるマルチフォント漢字認識”, 電子情報通信学会ニューロコンピューティング研究会資料, NC91-31 (1991)
- [51] 清水正博, 竹之内磨理子, 脇 康, 藤原啓恵, 横江 亨: “文書清書システム(1) —システムとその概要—”, 昭和61年度電子通信学会総合全国大会講演論文集, 1510 (1986).
- [52] 伊藤哲郎: “情報検索 [ソフトウェア講座 19]”, 昭晃堂.
- [53] 木本晴夫: “キーワード自動抽出と重要度評価”, 情報処理学会自然言語処理研究会資料, 64-1 (1987).
- [54] 中村順一, 今永一弘, 吉田 將: “接続コスト最小法による日本語形態素解析の評価実験”, 電子情報通信学会言語理解とコミュニケーション研究会資料, NLC91-1 (1991).
- [55] 荒木健治, 枅内香次: “帰納的学習による形態素解析手法における適応能力の評価”, 電子情報通信学会言語理解とコミュニケーション研究会資料, NLC90-41 (1990).
- [56] 中村匡伸, 紺野章子, 本郷保夫: “DLM(Digraph Language Model)による文字認識後処理”, 電子情報通信学会パターン認識・理解研究会資料, PRU90-116 (1990).

本論文に関する著者の発表文献 ◎：論文，○：国際会議

道路網階層的表現関連

- ◎ [57] 丹羽寿男，吉田雄二，福村晃夫：“道路網の階層的表現にもとづく経路探索アルゴリズムと地図情報システムへの応用”，情報処理学会論文誌，Vol.31，No.5，pp.659-666 (1990).
- [58] 丹羽寿男，吉田雄二，福村晃夫：“市街地地図情報の蓄積と検索”，昭和61年電子通信学会総合全国大会講演論文集，1626 (1986).
- [59] 丹羽寿男，吉田雄二，福村晃夫：“市街地地図情報の蓄積と検索”，電子通信学会人工知能と知識処理研究会資料，AI86-27 (1986).
- [60] 丹羽寿男，吉田雄二，福村晃夫：“市街地地図情報データベースシステム”，昭和62年電子情報通信学会総合全国大会講演論文集，1412 (1987).
- [61] 丹羽寿男，吉田雄二，福村晃夫：“市街地地図情報データベースシステム”，昭和63年電子情報通信学会春季全国大会講演論文集，D-354 (1988).
- [62] 丹羽寿男，吉田雄二，福村晃夫：“Interlisp-Dによる地図情報データベースの実現”，情報処理学会記号処理研究会資料，88-SYM-46-2 (1988).

地図情報提示法関連

- [63] 丹羽寿男，吉田雄二，福村晃夫：“地図情報システムにおける略図生成機能の実現”，情報処理学会第38回全国大会講演論文集，4R-7 (1989).
- [64] 丹羽寿男，渡邊豊英，吉田雄二：“道案内地図情報システムにおける略図と文章の提示法”，情報処理学会情報システム研究会資料，90-IS-29-2 (1990).

文字認識後処理分析評価関連

- ◎ [65] 丹羽寿男, 萱嶋一弘, メ木泰治: “文字認識誤りモデルを用いた文字認識後処理の分析評価”, 電子情報通信学会論文誌 (D-II), (条件付採録).
- [66] 丹羽寿男, 萱嶋一弘, メ木泰治: “文字認識誤りモデルを用いた後処理の分析評価”, 1992 年電子情報通信学会春季全国大会講演論文集, D-562 (1992).
- [67] 丹羽寿男, 萱嶋一弘, メ木泰治: “文字認識装置の誤りモデル”, 1992 年電子情報通信学会秋季全国大会講演論文集, D-323 (1992).

文字認識後処理関連

- [68] Niwa H., Kayashima K. and Shimeki Y.: “Postprocessing for Character Recognition Using Keyword Information”, Proc. of MVA'92, pp.519-522 (1992).
- ◎ [69] 丹羽寿男, 萱嶋一弘, メ木泰治: “キーワード抽出と最適文節選択による文字認識後処理”, 電子情報通信学会論文誌 (D-II), (1993 年 6 月号掲載予定).
- [70] 丹羽寿男, 萱嶋一弘, メ木泰治: “文字認識後処理法と後処理による効果の分析 —キーワード情報の利用—”, 電子情報通信学会パターン認識・理解研究会資料, PRU91-135 (1992).
- [71] 丸野 進, 小島良宏, 山本浩司, 丹羽寿男, 萱嶋一弘, メ木泰治: “パターンと記号の統合化処理 C o P S による文書認識 (1)”, 1992 年電子情報通信学会秋季全国大会講演論文集, D-313 (1992).
- [72] 山本浩司, 小島良宏, 丹羽寿男, 丸野 進, メ木泰治: “C o P S による文書認識 (2) 神経回路網による文字切り出し処理 —フィードバック情報に基づくパラメータ制御—”, 1992 年電子情報通信学会秋季全国大会講演論文集, D-314 (1992).
- [73] 小島良宏, 山本浩司, 丹羽寿男, 香田敏行, 丸野 進, メ木泰治: “C o P S による文書認識 (3) —複数特徴量を統合した構造化ニューラルネットワークによる漢字認識—”, 1992 年電子情報通信学会秋季全国大会講演論文集, D-315 (1992).

- [74] 丹羽寿男, 山本浩司, 小島良宏, 萱嶋一弘, 木泰治: “C o P Sによる文書認識(4) 後処理における誤訂正単語のリジェクト”, 1992年電子情報通信学会秋季全国大会講演論文集, D-316 (1992).
- [75] 丹羽寿男, 伊藤 哲, 小島良宏, 山本浩司, 萱嶋一弘, 木泰治: “C o P Sによる文書認識(5) 後処理におけるキーワード情報を用いた候補外文字救出”, 1992年電子情報通信学会秋季全国大会講演論文集, D-317 (1992).
- [76] 伊藤 哲, 丹羽寿男, 萱嶋一弘, 木泰治: “複合語最適評価法によるキーワード抽出”, 1993年電子情報通信学会春季全国大会講演論文集, D-113 (1993).