

修士学位論文

# 飾り文字認識のための構造解析手法 に関する研究

東北大学大学院 工学研究科 電気・通信工学専攻  
女川 俊一

# 目次

第1章	序論	2
1.1	研究の背景と目的	2
1.2	本論文の構成	4
第2章	マッチンググラフを用いた飾り文字認識	5
2.1	はじめに	5
2.2	飾り文字認識システムの概要	6
2.2.1	大局構造抽出	6
2.2.2	大分類	6
2.2.3	細分類	9
2.3	マッチンググラフを用いた構造解析手法	9
2.3.1	使用する特徴	9
2.3.2	マッチンググラフの定義	10
2.3.3	マッチンググラフを用いた構造解析手法の概要	11
2.4	マッチンググラフ作成の前処理	11
2.4.1	特徴抽出	11
2.4.2	ブリッジの導入	13
2.5	マッチンググラフの作成	14
2.5.1	マッチンググラフ作成の概要	14
2.5.2	初期部分グラフの作成	14
2.5.3	パスの延長	16
2.5.4	部分グラフの拡張	18
2.5.5	コスト計算	21
2.5.6	マッチンググラフ拡張の限定	24
2.6	認識実験	26
2.6.1	目的	26
2.6.2	予備実験	26
2.6.3	実験手法	27
2.6.4	実験条件	27

---

2.6.5	実験結果	29
2.7	まとめ	29
<b>第3章</b>	<b>適応的正規化の導入</b>	<b>31</b>
3.1	はじめに	31
3.2	適応的正規化	31
3.2.1	位置ベクトル作成	33
3.2.2	Global Affine Transformation(GAT)	33
3.2.3	Local Affine Transformation(LAT)	34
3.2.4	スムージング	35
3.2.5	適応的正規化の実行例	35
3.2.6	適応的正規化のコスト	35
3.3	認識実験	39
3.3.1	目的	39
3.3.2	予備実験	39
3.3.3	実験手法	39
3.3.4	実験条件	39
3.3.5	実験結果	39
3.4	まとめ	41
<b>第4章</b>	<b>複合化処理の導入</b>	<b>43</b>
4.1	はじめに	43
4.2	複合化方法	43
4.3	信頼度の妥当性	44
4.4	認識実験	45
4.4.1	目的	45
4.4.2	実験手法	45
4.4.3	実験条件	45
4.4.4	実験結果	47
4.5	まとめ	48
<b>第5章</b>	<b>結論</b>	<b>49</b>
5.1	本研究の成果	49
5.2	今後の課題	49
	<b>参考文献</b>	<b>51</b>

# 目 次

1.1	構造解析手法とパターン整合法	3
2.1	対応関係	5
2.2	飾り文字の例	6
2.3	飾り文字認識手法の構成	7
2.4	大局構造	8
2.5	大分類	8
2.6	対応付け	9
2.7	特異点・ストローク	10
2.8	マッチンググラフ	11
2.9	マッチンググラフによる構造解析手法の概要	12
2.10	特異点・ストロークの抽出	13
2.11	ブリッジ	14
2.12	マッチンググラフ作成の概要	15
2.13	初期部分グラフの作成	17
2.14	初期部分グラフの作成 (その 2)	18
2.15	パスの延長	19
2.16	部分グラフの拡張 (1)	21
2.17	部分グラフの拡張 (2)	22
2.18	変形コスト	23
2.19	整合コスト	24
2.20	部分グラフのコスト下限	26
2.21	重み定数の決定	27
2.22	実験手法	28
2.23	テストサンプル	28
2.24	正読例	30
2.25	誤読例	30
3.1	適応的正規化の導入	32

---

3.2	適応的正規化の概要	33
3.3	適応的正規化	34
3.4	適応的正規化の実行例 (その 1)	36
3.5	適応的正規化の実行例 (その 2)	37
3.6	適応的正規化の実行例 (その 3)	38
3.7	実験手法	40
3.8	テストサンプル	41
3.9	正読例	42
3.10	誤読例	42
4.1	複合化処理を導入した飾り文字認識システムの概要	45
4.2	信頼度と誤認識文字存在率	46
4.3	実験手法	46
4.4	テストサンプル	47
4.5	複合化処理を導入した結果	48

# 表 目 次

2.1	実験条件 . . . . .	27
2.2	実験結果 . . . . .	29
3.1	実験条件 . . . . .	40
3.2	適応的正規化を導入した際の認識率 . . . . .	41
4.1	実験条件 . . . . .	47
4.2	複合化処理を導入した際の認識率 . . . . .	47

# 第1章

## 序論

### 1.1 研究の背景と目的

近年，社会の急速な多様化，分散化に伴い，情報の複雑化，大量化が生じている．これに対処するために計算機が導入され，あらゆる分野で機械化が行われている．このような計算機の普及に伴い人間と機械との情報伝達が問題になっている．もし，我々が普段情報伝達に用いている，紙に書かれた，あるいは印刷された文字を文字認識システムにより自動的に計算機に入力できれば，人間と機械との情報伝達が容易なものとなる．

文字認識システムは多様な文書に対応できることが重要である．我々が普段目にする字体は一様なものではなく，多様なフォントが存在する．特に，雑誌等のタイトル・見出し，カタログ等の企業名・商品名に用いられる飾り文字は文字本来の目的である“情報伝達”に加え，デザイン性を高めることにより消費者の注意を引くといった特別な目的を持つため，既成のフォントにとらわれない多様な形状を有する．

そこで，本研究では飾り文字を対象とし，多様なフォントに対応できる飾り文字認識システムの開発を目指す．

既存のフォントに対する文字認識手法は，大きくパターン整合法と構造解析手法の2種類に分類できる．

パターン整合法は，パターン同士の重なり具合で評価し認識を行なう手法である．入力された未知の文字パターンとあらかじめ作成しておいた標準パターン（以後，辞書と呼ぶ）との重ね合わせを行ない，最も近いものを認識結果とする（図 1.1(b))．パターンの特徴量を用いた整合法として，文字画像から方向線分を抽出して辞書との近さをもとに認識を行う手法 [1][2] は日本語手書き文字認識において高い認識率を得ている．

構造解析手法は，入力文字画像から文字の線分などの構造情報を抽出し，線分の位置関係や接続関係などに着目して識別を行う手法である（図 1.1(a))．構造解析法として，山本による文字パターンの輪郭線を折線近似し文字自体を多角形近似することで線分を抽出し，

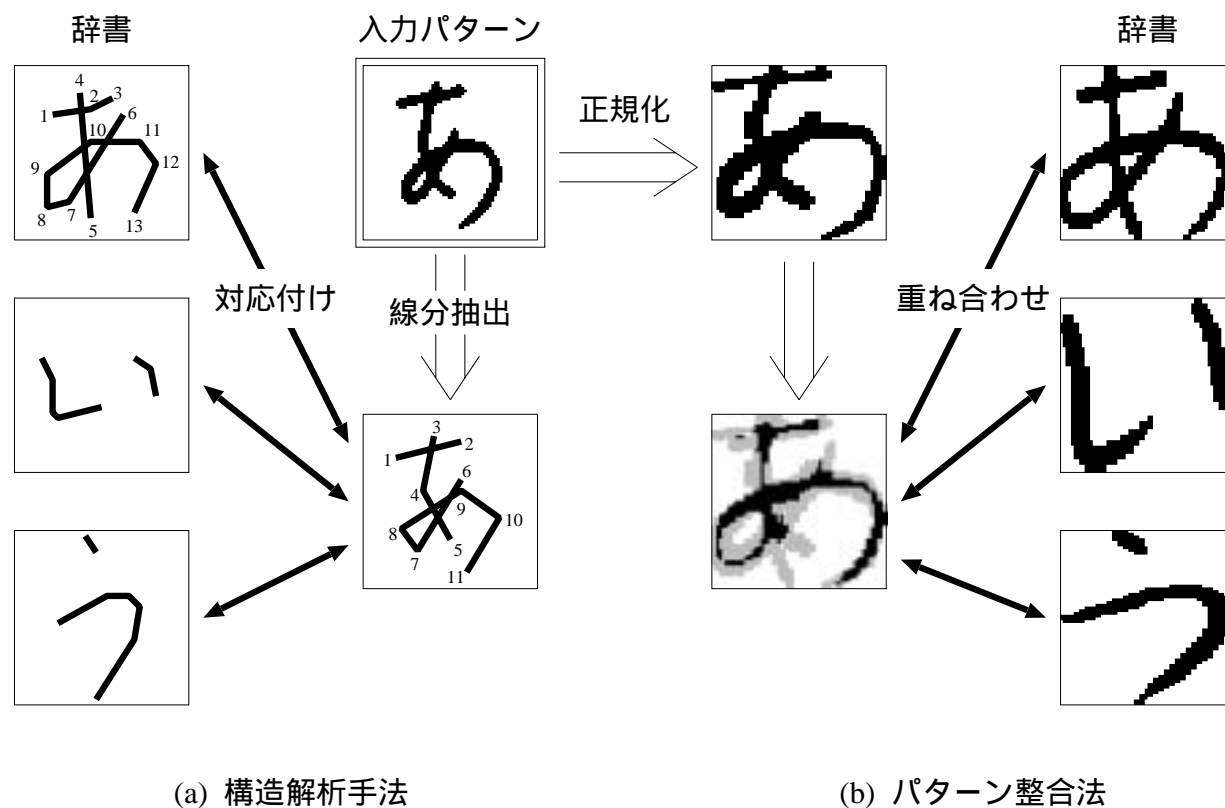


図 1.1: 構造解析手法とパターン整合法

線分間の対応付けを行う手法 [3] , Rocha らによる文字画像からストロークを抽出しグラフ表現に変換して形状を解析する手法 [4] 等が提案されている．前者は漢字のようなほぼ直線からなる文字に適した手法であり，後者は曲線のストロークも柔軟に扱うことのできるモデルになっている．

前述の手法ではいずれも印刷文字や手書き文字を対象とするため，特徴抽出は文字画像の黒画素の連結性に注目している．しかし，飾り文字の場合，黒画素の連結性に注目しただけではその文字の形状を表す特徴を抽出できるとは限らない．このような飾り文字に対し，複数のパラメータ (スケール) で画像にぼかしを与え尾根や谷等の幾何的な特徴 [5] を抽出し大局構造を求める手法 [6] が提案された．しかし，この手法により得られる大局構造は歪みが大きいため，その画像に適した認識手法が求められる．

飾り文字から得られる大局構造は同一字種ではほぼ同じトポロジとなる．このことは，構造解析手法による飾り文字認識の可能性を示唆している．そこで，構造解析手法として Rocha らの提案するマッチンググラフを用いた手法を元に，より大きな変形を有する文字構造に対応できるようなマッチンググラフの作成法を考案し，それを組み込んだ飾り文字認識手法を開発し認識性能の向上を目指す．



## 1.2 本論文の構成

本論文の構成は以下の通りである。

### 第1章 序論

研究の背景や目的，論文の構成を述べる。

### 第2章 マッチンググラフを用いた飾り文字認識

飾り文字を認識するための構造解析手法としてマッチンググラフを用いた手法を提案する。そして，その構造解析手法を組み込んだ飾り文字認識手法を提案し，提案手法により飾り文字の認識実験を行う。

### 第3章 適応的正規化の導入

第2章の結果を受け，マッチンググラフ作成のための前処理に適応的正規化を導入する。そして，認識実験を行いその有効性を確認する。

### 第4章 複合化処理の導入

第3章の結果を受け，第2章と第3章の認識結果の複合化を導入する。そして，その有効性を確認する。

### 第5章 結論

結論である。

## 第2章

# マッチンググラフを用いた飾り文字 認識

### 2.1 はじめに

本章では構造解析手法として、Rochaらの提案する形状分析モデル[4]に基づき、マッチンググラフを用いて飾り文字を認識する手法を提案する。本手法では図 2.1 に示すように未知入力と辞書との線分・曲線間の対応関係を調べ、その整合の度合をコストとして表し認識結果を出力する。認識対象となる飾り文字の例を図 2.2 に示す。

2.2 節では、飾り文字認識手法の構成について述べる。2.3 節では提案するマッチンググラフを用いた構造解析手法について述べ、マッチンググラフとその作成法について説明する。2.6 節では提案手法による飾り文字認識実験を行いその有効性を確認する。

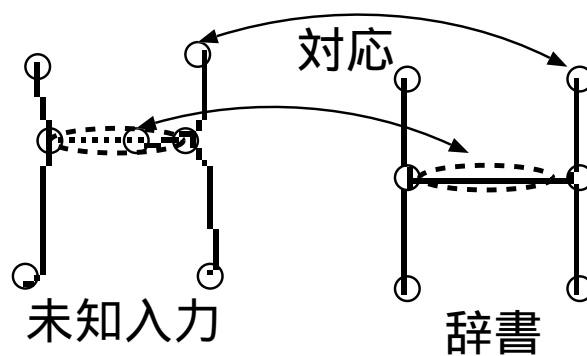


図 2.1: 対応関係

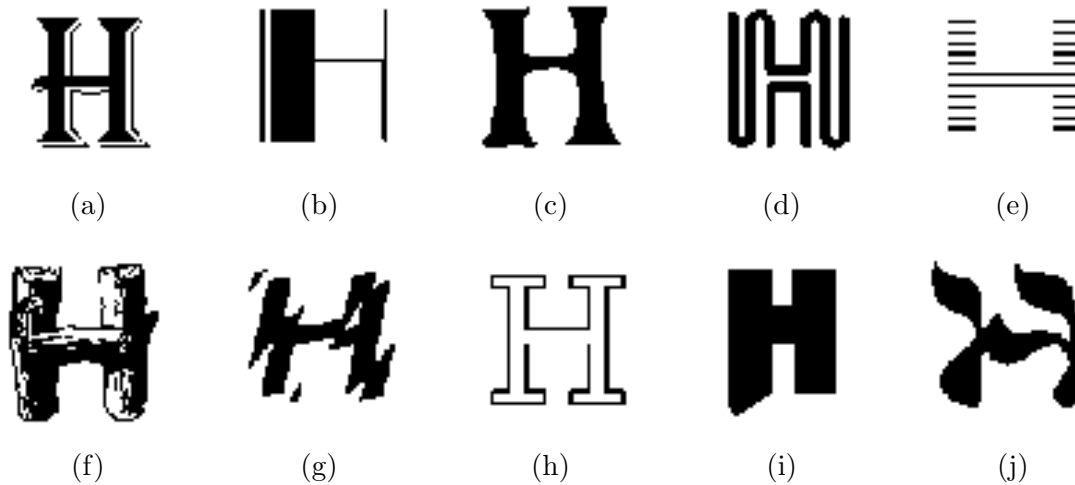


図 2.2: 飾り文字の例

## 2.2 飾り文字認識システムの概要

図 2.3 に提案する認識システムの概要を示す．最初にスケールスペース解析により飾り文字の大局構造を抽出する．次にパターン整合法により大分類を行う．最後に提案するマッチンググラフを用いた手法により細分類を行い，飾り文字認識を行う．

### 2.2.1 大局構造抽出

飾り文字に対して，文献 [6] で提案された手法により大局構造を抽出する．スケールスペース解析を行い濃淡画像の尾根点を抽出する．次に谷点を補間することにより文字の大局構造を抽出する．図 2.2 の画像から大局構造を抽出した結果を図 2.4 に示す．

### 2.2.2 大分類

次に 2.2.1 節で説明した手法により得られた各飾り文字の大局構造をパターン整合法で大分類を行う．その際のパターン整合法について説明する．

まず，大局構造を表す画像を  $64 \times 64$  に正規化する．次元圧縮と文字画像の変形の吸収のために，隣接する 4 画素を 1 つの小領域とし， $32 \times 32$  個の小領域を重複しないように設け，各小領域内の黒画素数をその小領域に対応する要素の値とする 1024 次元の特徴ベクトルを作成する．(図 2.5) 識別関数としてユークリッド距離を使用し，上位  $K$  個の候補を選出する．

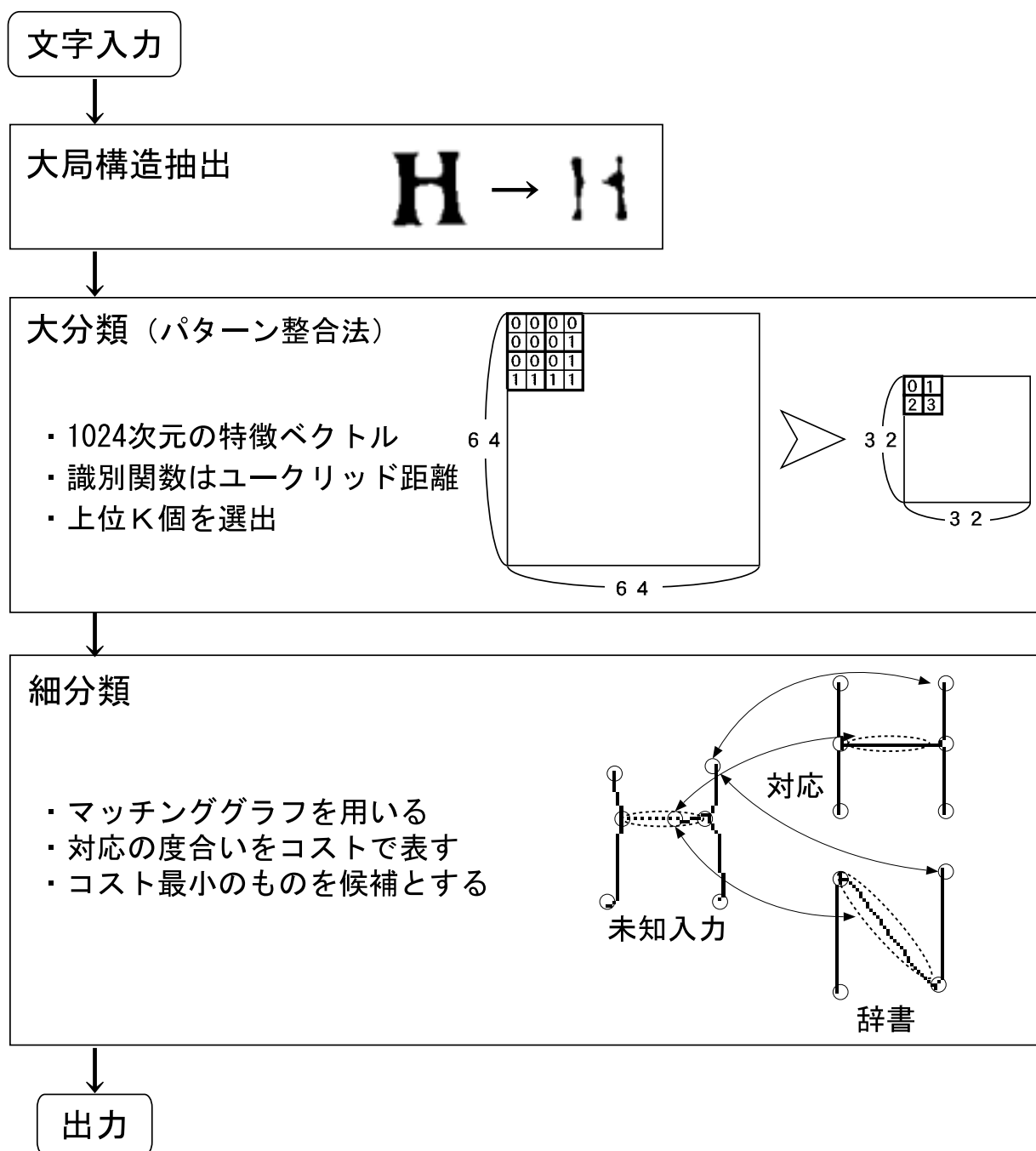


図 2.3: 飾り文字認識手法の構成

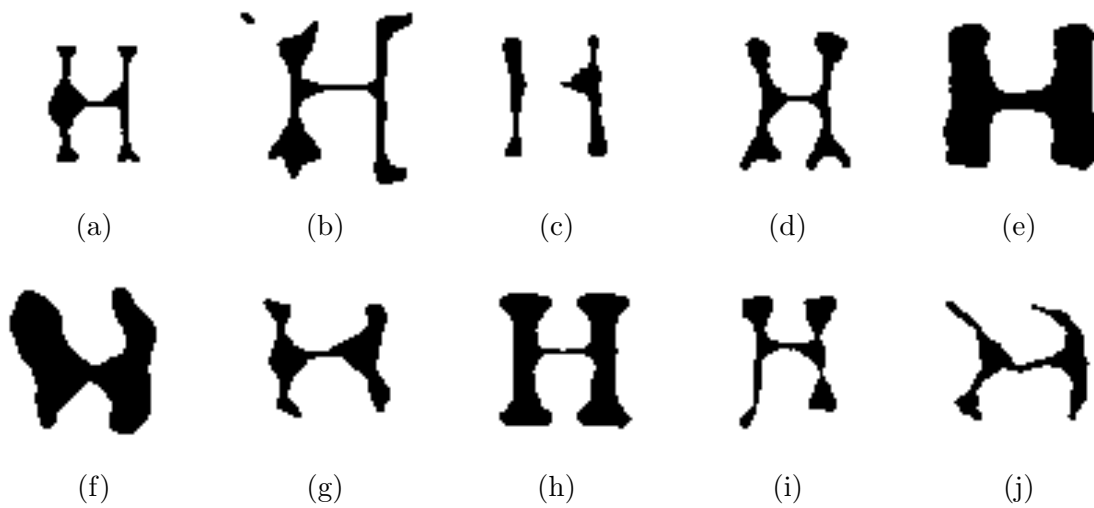


図 2.4: 大局構造

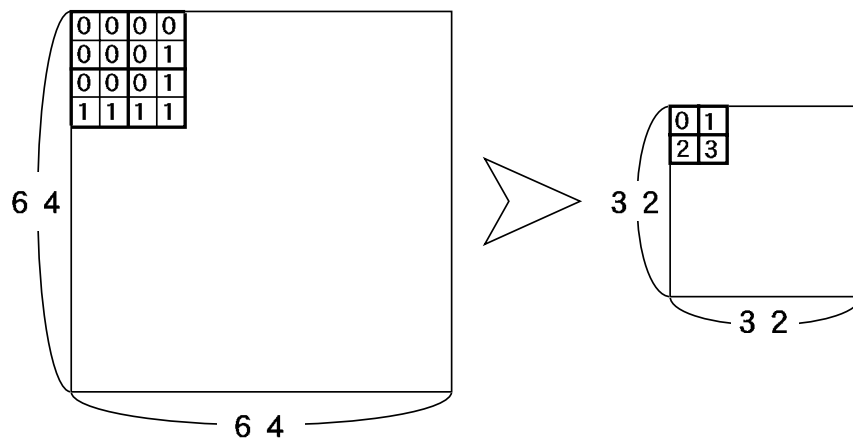


図 2.5: 大分類

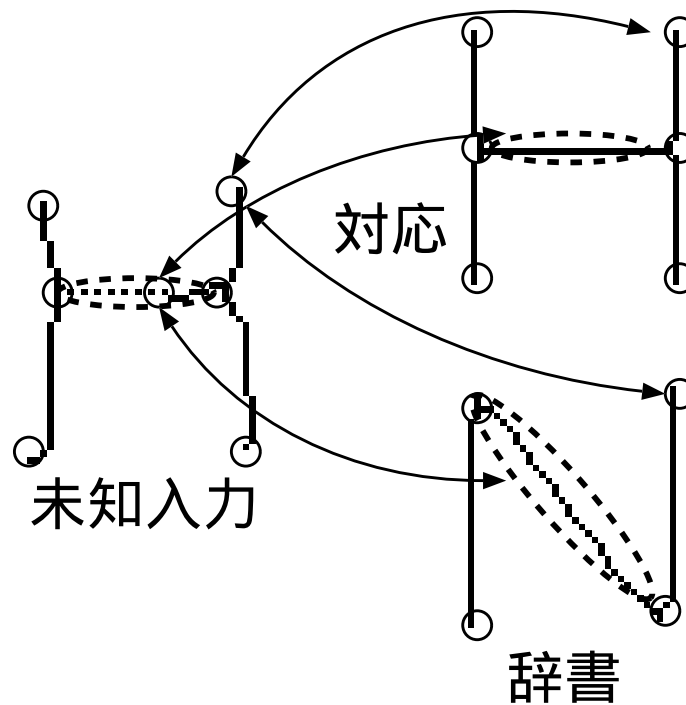


図 2.6: 対応付け

### 2.2.3 細分類

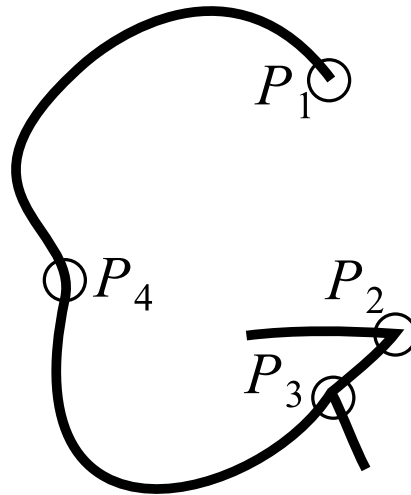
提案するマッチンググラフを用いた構造解析手法で細分類を行う。未知入力から得られた大局構造 (図 2.4) から線分と特異点を抽出し、マッチンググラフにより、辞書の線分と特異点との対応付けを行う (図 2.6)。対応付けをするために未知入力の線分を変形する必要があるが、その変形と整合の度合いがコストとなる。コストの最小のものを候補とする。マッチンググラフの詳細は 2.3 節で説明する。

## 2.3 マッチンググラフを用いた構造解析手法

2.3.1 節で使用する特徴について述べる。また、2.3.2 節ではマッチンググラフについて述べ、2.3.3 節ではマッチンググラフによる構造解析手法の概要について述べる。

### 2.3.1 使用する特徴

マッチンググラフに用いる特徴として、ストロークと特異点を使用する。大局構造を Hilditch 細線化法 [7] により細線化し、ストローク、特異点を抽出する。特異点としては、



特異点

図 2.7: 特異点・ストローク

- ストロークの端点 (図 2.7 点  $P_1$ )
- 鋭角の頂点 (図 2.7 点  $P_2$ )
- ストロークの交点 (図 2.7 点  $P_3$ )
- 凹形の頂点 (図 2.7 点  $P_4$ )

を抽出する．2つの特異点を結ぶ線分や曲線を改めてストロークと呼ぶ．

### 2.3.2 マッチンググラフの定義

本節ではマッチンググラフの定義について述べる．

まず，未知入力と辞書の線分・曲線のことをストロークと呼び，未知入力の連続する複数のストロークの組をパスと呼ぶ．標準文字パターンに対して，その構造を表すグラフを辞書として用意しておく．

マッチンググラフとは，その枝に未知入力のパスと辞書のストロークの対応関係を表すラベルを付け，その節点に未知入力の特異点と辞書の特異点の対応関係を表すラベルを付けたグラフである．マッチンググラフを作成することで，未知入力と辞書との，パス-ストローク間の対応関係，特異点の対応関係などを表すことができ，それをもとにマッチングの度合をコストとして計算することができる．図 2.8 にマッチンググラフの例を示す．図 2.8(a) の同色のパス-ストローク  $(f_1-g_1, \{b, f_2\}-g_2, f_3-g_3)$  を対応させたとする．また，未知

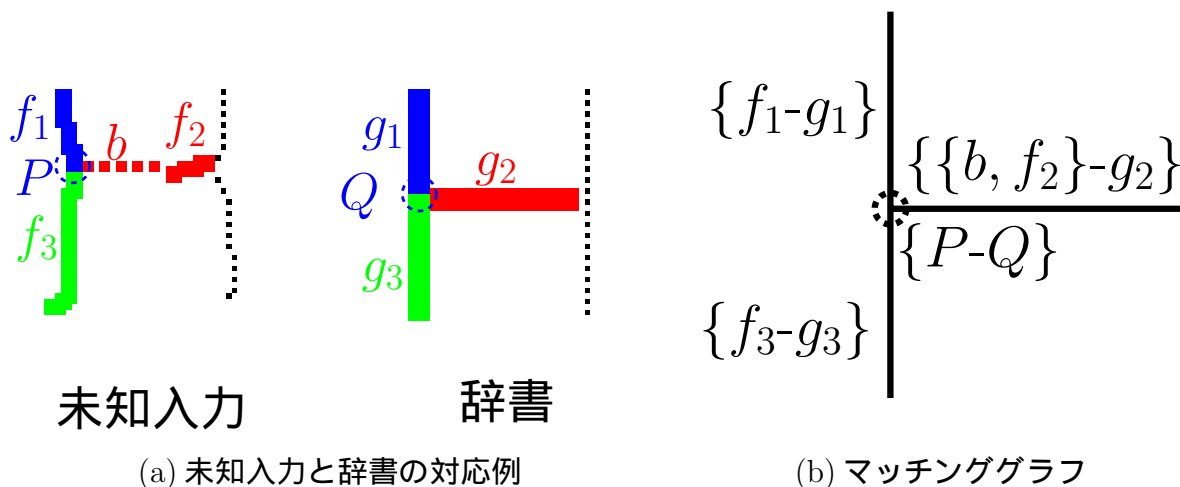


図 2.8: マッチンググラフ

入力の特異点  $P$  と辞書の特異点  $Q$  を対応させたとする．図 2.8(a) の対応関係を表すグラフ<sup>1</sup>は図 2.8(b) となる．

### 2.3.3 マッチンググラフを用いた構造解析手法の概要

図 2.9 にマッチンググラフによる構造解析手法の概要を示す．マッチンググラフを作成するための前処理として，特徴抽出，及び，仮のストロークの導入 (ブリッジ) を行う．

続いて，得られた特徴をもとにマッチンググラフを作成する．得られたグラフに対してコスト計算し，コスト最小のものを結果として出力する．

## 2.4 マッチンググラフ作成の前処理

本節では，マッチンググラフを用いて飾り文字の認識をするための前処理について述べる．前処理として特徴抽出，ブリッジの導入を行う．

### 2.4.1 特徴抽出

大局構造を Hilditch 細線化法 [7] により細線化し，ストローク，特異点を抽出する．特異点としては，2.3.1 節で述べた通り，

- ストロークの端点 (図 2.7 点  $P_1$ )

<sup>1</sup>表記が複雑になるので，本論文では以後図による説明を行う場合，図 2.8(a) のような未知入力と辞書との対応関係を直接示す図で説明を進める



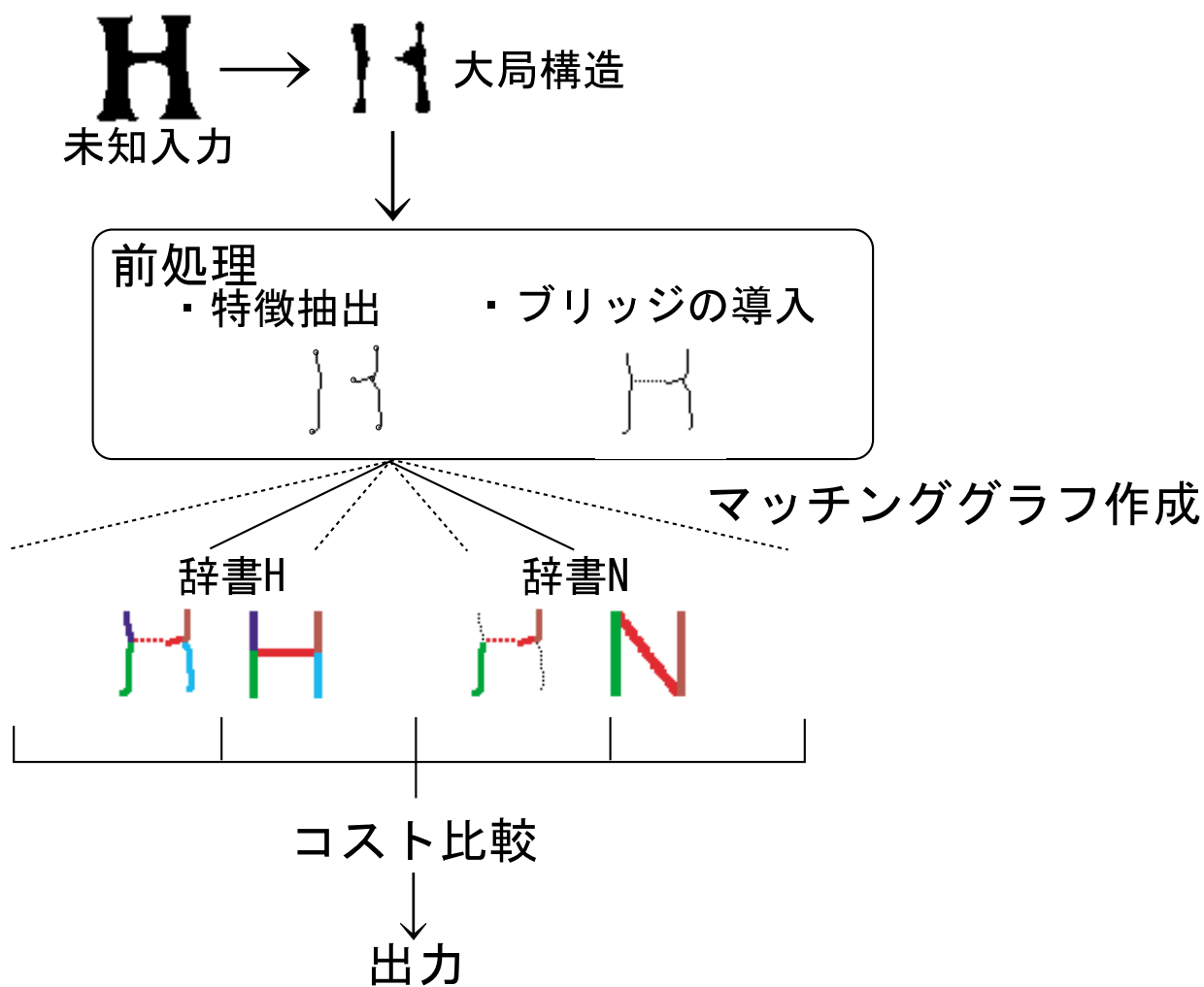
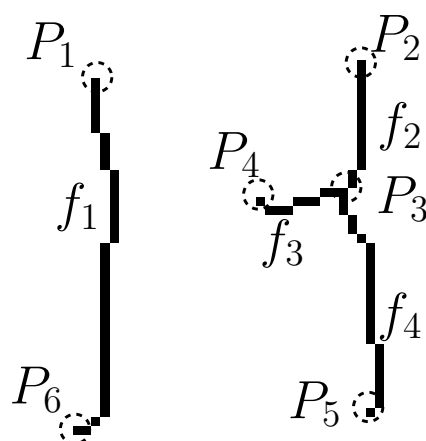


図 2.9: マッチンググラフによる構造解析手法の概要



特異点・ストロークの例

図 2.10: 特異点・ストロークの抽出

- 鋭角の頂点 (図 2.7 点  $P_2$ )
- ストロークの交点 (図 2.7 点  $P_3$ )
- 凹形の頂点 (図 2.7 点  $P_4$ )

を抽出し, 2つの特異点を結ぶ線分や曲線を改めてストロークと呼ぶ. なお, ストロークの凹形の頂点は端点から曲線ストロークをおって行って逆の曲率が最大となる点である. 具体例を図 2.10 に示す. 図 2.10 は図 2.4(c) の画像を細線化したものである. 図 2.10 における  $f_1, f_2, f_3, f_4$  は抽出されたストロークを,  $P_1, P_2, \dots, P_6$  は特異点を示す.

### 2.4.2 ブリッジの導入

ブリッジとはストロークの欠如に対応するために導入する仮のストロークである.

図 2.11(a) の実線は図 2.4 の飾り文字 “H” より, (b) の実線はある飾り文字 “A” より抽出された大局構造であるが, 一部のストロークが欠けてしまい, ストロークの対応が取れず誤認識の原因となる. これに対処するため仮のストローク (ブリッジ) を導入する (図 2.11, 点線). ブリッジは次の条件を満たす 2 点  $p, q$  が存在する場合に  $p, q$  間に導入する.

- 点  $p$  はストロークの端点であり, 点  $q$  は黒画素上の点である
- $d(p, z) < d(p, q)$  かつ  $d(q, z) < d(p, q)$  を満たす黒画素上の点  $z$  が存在しない

但し  $d(p, q)$  は 2 点  $p, q$  間のユークリッド距離である.

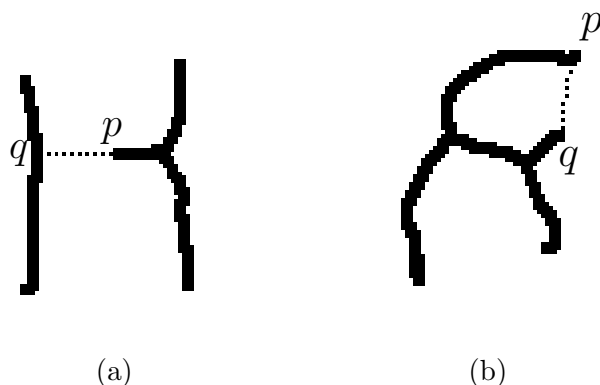


図 2.11: ブリッジ

## 2.5 マッチンググラフの作成

Rochaらのモデルではマッチンググラフを作成していく際にその時点でのコスト最小のグラフのみを考慮する．この手法では，比較的変形の小さい文字には対応できるが，飾り文字のような変形の大きい文字に対しては局所解に陥り易い．そこで，提案する手法では複数の候補を残し，最終的に最適なグラフ（コスト最小のグラフ）を選出する．

### 2.5.1 マッチンググラフ作成の概要

マッチンググラフ作成の概要を図 2.12 に示す．

マッチンググラフを作成するには，最初に初期部分マッチンググラフ（以下単に初期部分グラフ）を作成する．

次に作成した初期部分グラフを拡張した部分マッチンググラフ（以下単に部分グラフ）を作成する．これを再帰的に繰り返すことにより，文字全体を表すマッチンググラフを作成する．

最後に作成された全グラフのコストを比較し，コスト最小のグラフを解とする．

以上を実行するために，2.5.2 節で初期部分グラフを作成する手法，2.5.3 節で初期部分グラフのパスを延長し新たな部分グラフを作成する手法，2.5.4 節で部分グラフを拡張する手法を説明する．

### 2.5.2 初期部分グラフの作成

ある未知入力と辞書との組が与えられた際，最初に文字の一つの特異点とその周囲のストロークに対応させた初期部分グラフ  $M_1, M_2, \dots, M_N$  を作成する．本節では初期部分グラフの作成法について述べる．初期部分グラフの例を図 2.13 に示す．図 2.13(a) が与えら

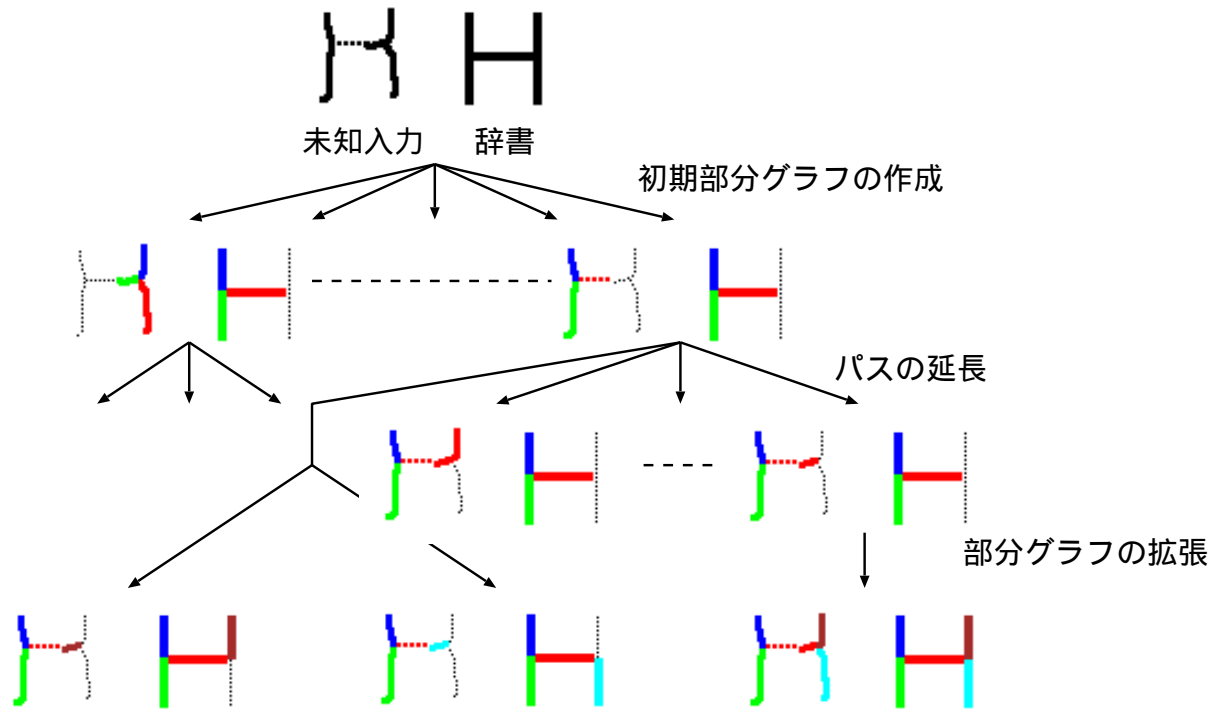


図 2.12: マッチンググラフ作成の概要

れた未知入力と辞書の組であり, (b), (c), ..., (g) が作成された初期部分グラフの例である.

まず辞書の一本のストローク  $g$  とその端点である特異点  $Q$  を選択する. 選択した辞書のストロークに対応させるべき未知入力のストロークあるいはブリッジとして, 任意の 1 つ  $f$  を選ぶ. 辞書のストローク  $g$  と特異点  $Q$  の位置関係 ( $Q$  は  $g$  の上端か下端か, 左端か右端か, など) により  $f$  の 2 つの特異点のうち一方を  $Q$  に対応させるものとして選ぶ. それを  $P$  とする. 次に未知入力の特異点  $P$  の周囲のストロークを  $f$  から時計周りに探索し, ストローク  $f_1, f_2, \dots, f_m$  を見つける. 同様に辞書の特異点  $Q$  に隣接するストロークを  $g$  から時計周りに探索し,  $g_1, g_2, \dots, g_n$  を見つける. この順番を保持する対応付けを全て求める ( $m = n$  なら一意). そして, それぞれの対応付けに関して初期部分グラフ  $M = (\{P-Q\}, \{f-g, f_{i_1}-g_{j_1}, \dots, f_{i_{n'}}-g_{j_{n'}}\})$  ( $n'$  は  $m, n$  の小さい方の値,  $s < t$  ならば  $i_s < i_t$  かつ  $j_s < j_t$ ) を作成する. これを辞書の各一組のストロークと特異点に対し, 未知入力のすべてのストロークに対して対応付けを行い, それぞれの対応付けに対する初期部分グラフを作成する.

未知入力の前処理において本来 1 点となるべきものが, 短いストロークになってしまうことがある. そこで, あるストロークが次の条件を満たす場合, そのストロークを削除し, そのストロークの両端の特異点を同一の特異点とみなして, 同様の対応付けを行った部分グラフも作成する.

- 両端の特異点が共に 3 本以上のストロークの交点

- 接しているストロークの中で最短

図 2.13 は辞書のストローク  $g_1$ , 特異点  $Q_3$  の組に対して初期部分グラフを作成した例であり, 同色のストロークは対応関係にあることを示し, 破線の円に囲まれた特異点是对应関係にあることを示している. 初期部分グラフ  $M_1, M_2, \dots, M_5$  はそれぞれ最初に未知入力ストローク  $f_1, f_2, \dots, f_5$  と辞書のストローク  $g_1$  を対応させた場合に作成した初期部分グラフであり, 初期部分グラフ  $M_6$  はブリッジ  $b$  と辞書のストローク  $g_1$  を対応させた場合に作成した初期部分グラフである.

また, 図 2.14 の例では辞書のストローク  $g$ , 特異点  $Q$  の組に対して未知入力ストローク  $f$  と特異点  $P$  を対応させた場合の初期部分グラフを作成した例であり, 同色のストロークは対応関係にあることを示している, 初期部分グラフとして  $M_1 = (\{P-Q\}, \{f-g, f_1-g_2, f'-g_3\})$   $M_2 = (\{P-Q\}, \{f-g, f_1-g_1, f'-g_3\})$   $M_3 = (\{P-Q\}, \{f-g, f_1-g_1, f'-g_2\})$  が作成されるが, 図 2.14 の例では  $f'$  は  $f, f_1, f_2, f_3$  より短く, かつ, その両端の点  $P, P'$  がそれぞれ 3 本のストロークの交点となっているので,  $f'$  を削除し, 点  $P, P'$  を同一の点とみなして, 同様の対応付けを行い, 部分グラフ  $M_4 = (\{P(P')-Q\}, \{f-g, f_1-g_1, f_2-g_2, f_3-g_3\})$  を作成する.

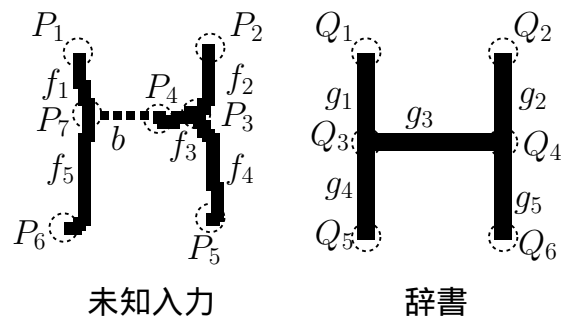
### 2.5.3 パスの延長

前節で作成された初期部分グラフは未知入力と辞書のストロークが 1 本対 1 本で対応 (一対一対応) している. しかし, 本来一本のストロークとして抽出されるべきものが, 大局構造の歪みや文字の飾りの効果などが原因で, 複数本のストロークとして抽出される場合がある. 従って, 未知入力の複数のストロークからなるパスと辞書のストロークとの対応 (多対一対応) 考慮する必要がある. そこで本節では, 一対一対応の部分グラフ  $M$  が与えられた場合, パスの延長を行うことで多対一対応の部分グラフ  $M'_1, M'_2, \dots, M'_N$  を作成するための手法について述べる. パスの延長の例を図 2.15 に示す. 図 2.15(a) の  $M$  が与えられた一対一対応の部分グラフであり, 図 2.15(b),(c),(d) の  $M'_1, M'_2, M'_3$  が  $M$  のパスを延長して作成した多対一対応の部分グラフである.

与えられた部分グラフが  $M = (\{P-Q\}, \{f_{i_0}-g_{j_0}, \dots, f_{i_n}-g_{j_n}\})$  の場合, まず, 辞書のストローク  $g_{j_0}, \dots, g_{j_n}$  に対応するパスとして,  $\alpha_{00} = \{f_{i_0}\}, \dots, \alpha_{n0} = \{f_{i_n}\}$  を作成する.

次にこれらのパスをそれぞれ延長する.  $\alpha_{k0}$  を延長する際, まず,  $\alpha_{k0}$  の終点を  $f_{i_k}$  の特異点のうち  $P$  の反対側のもの (終点) として定める. それを  $P_{k0}$  とおく. このパスの終点に隣接するストローク  $f'_1, \dots, f'_{n''}$  を探索する. これらのストロークをパス  $\alpha_{k0}$  にそれぞれ挿入し, パス  $\alpha_{k1} = \{f_{i_k}, f'_1\}, \dots, \alpha_{kn''} = \{f_{i_k}, f'_{n''}\}$  を作成する. これらのパスの終点は挿入されたストロークの終点である. このパスの延長を長さが対応する辞書のストローク  $g_{j_k}$  の 1.5 倍以上になるか, 終点に隣接するストロークがなくなるまで再帰的に繰り返す.

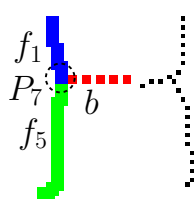
最後に作成されたパスの全組合せ (初期組合せは除く) に対して部分グラフを作成する.



未知入力

辞書

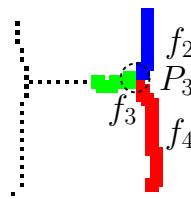
(a) 未知入力と辞書の組の例



未知入力

辞書

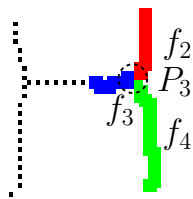
(b) 初期部分グラフ  $M_1$



未知入力

辞書

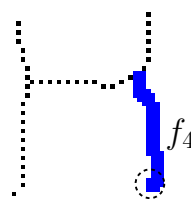
(c) 初期部分グラフ  $M_2$



未知入力

辞書

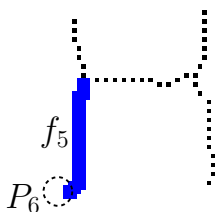
(d) 初期部分グラフ  $M_3$



未知入力

辞書

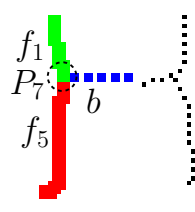
(e) 初期部分グラフ  $M_4$



未知入力

辞書

(f) 初期部分グラフ  $M_5$

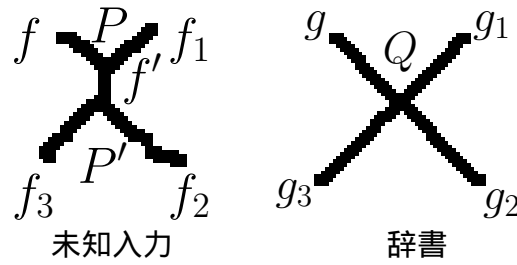


未知入力

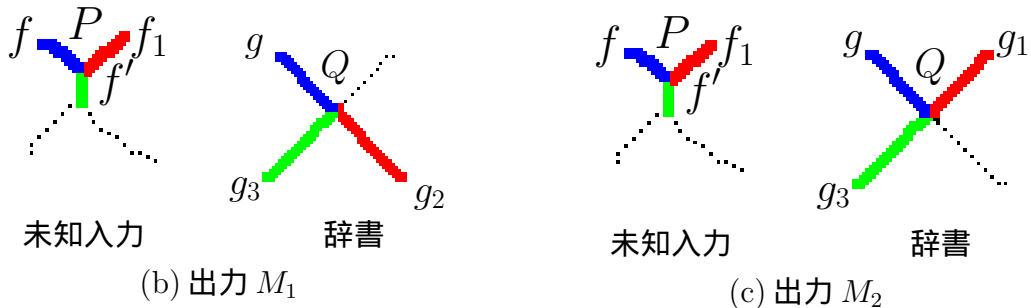
辞書

(g) 初期部分グラフ  $M_6$

図 2.13: 初期部分グラフの作成

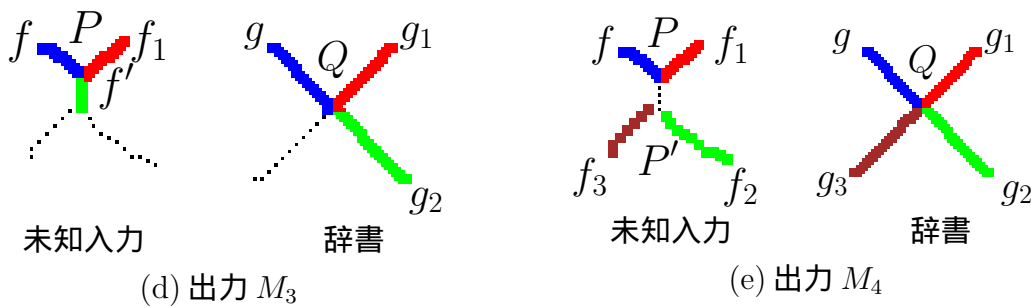


(a) 未知入力と辞書の組の例



(b) 出力  $M_1$

(c) 出力  $M_2$



(d) 出力  $M_3$

(e) 出力  $M_4$

図 2.14: 初期部分グラフの作成 (その 2)

図 2.15(a) の部分グラフ  $M$  の例では、まず、辞書のストローク  $g_1, g_3, g_4$  に対応する未知入力のパスとして、 $\alpha_{00} = \{f_1\}, \alpha_{10} = \{b\}, \alpha_{20} = \{f_5\}$  を作成する。

次に、パス  $\alpha_{10}$  を延長したパスとして、 $\alpha_{11} = \{b, f_3\}, \alpha_{12} = \{b, f_3, f_2\}, \alpha_{13} = \{b, f_3, f_4\}$  を作成する。但し、パス  $\alpha_{00}, \alpha_{20}$  についてはその終点の周囲に未知入力のストロークが存在しないので延長しない。

最後に作成されたパスを組み合わせ部分グラフを作成する。その結果  $M'_1 = (\{P_7-Q_3\}, \{\alpha_{00}-g_1, \alpha_{11}-g_3, \alpha_{20}-g_4\})$ ,  $M'_2 = (\{P_7-Q_3\}, \{\alpha_{00}-g_1, \alpha_{12}-g_3, \alpha_{20}-g_4\})$ ,  $M'_3 = (\{P_7-Q_3\}, \{\alpha_{00}-g_1, \alpha_{13}-g_3, \alpha_{20}-g_4\})$ (図 2.15(b),(c),(d)) を作成する。

### 2.5.4 部分グラフの拡張

本節では部分グラフを拡張する手法について説明する。

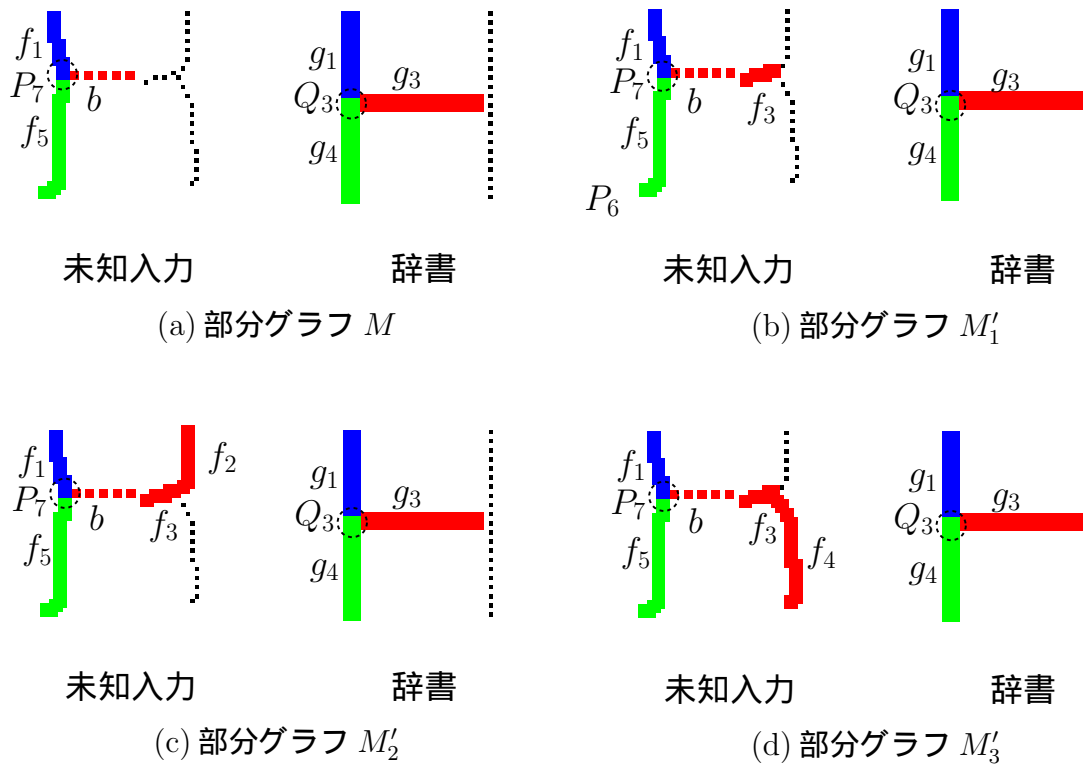


図 2.15: パスの延長



部分グラフ  $M$  が与えられた場合, 部分グラフ  $M$  に含まれる各パスの終点において, その特異点の対応付けが未だ行われていなければ, その特異点の周囲のストロークの対応関係を表す部分グラフを作成する. それらを  $M$  に結合することで部分グラフ  $M'_1, M'_2, \dots, M'_N$  を作成する. 図 2.16 は図 2.16(a) の部分グラフ  $M = (\{P_7-Q_3\}, \{\{f_1\}-g_1, \{b, f_3\}-g_3, \{f_5\}-g_4\})$  が与えられた場合の例であり, 図 2.17 は図 2.17(a) の部分グラフ  $M = (\{P_7-Q_3\}, \{\{f_1\}-g_1, \{b\}-g_3, \{f_5\}-g_4\})$  が与えられた場合の例である.

部分グラフ  $M = (V, \{\alpha_0-g_{i_0}, \dots, \alpha_n-g_{i_n}\})$  が与えられた場合, それぞれのパス  $\alpha_j$  の終点  $P_j$  (対応する辞書の特異点  $Q_{i_j}$ ) で,  $P_i$  が未調査の場合, 2.5.2 節にて説明した初期部分グラフを作成する手法と同様にして, 特異点  $P_j, Q_{i_j}$  の周囲の対応関係を表す部分グラフ  $N_{j1}, N_{j2}, \dots, N_{jn}$  を作成する. 更に, 得られた  $N_{j1}, N_{j2}, \dots, N_{jn}$  に対し 2.5.3 節で説明したパスを延長する手法により, 多対一対応の部分グラフ  $N'_{j1}, N'_{j2}, \dots, N'_{jn'}$  を作成する. 適切な部分グラフを作成することができないパスも存在することを考慮して,  $N'_{j0} = (\{P_j-Q_{i_j}\}, \phi)$  も作成しておく ( $\phi$  は空集合). これを部分グラフ  $M$  に含まれる全てのパス  $\alpha_0, \dots, \alpha_n$  について実行する.

最後に作成された部分グラフの全組合せ (初期組合せは除く) に対して部分グラフ  $M$  と結合する.

本節で述べた部分グラフの拡張をすべての初期部分グラフに対し, すべての特異点を調べ終わるまで再帰的に行うことで, 文字全体の対応関係を表すマッチンググラフを作成し, 作成された全マッチンググラフの中からコスト最小のものを解として出力する. コストについては 2.5.5 節にて述べる.

図 2.16, 図 2.17 の例について説明する.

図 2.16(a) の部分グラフ  $M$  の例では,  $\alpha_1 = \{b, f_3\}$  の終点  $P_3$  で部分グラフ  $N'_{10} = (\{P_3-Q_4\}, \phi)$ ,  $N'_{11} = (\{P_3-Q_4\}, \{f_2-g_2, f_4-g_5\})$ , を得る.  $\alpha_0 = \{f_1\}$  の終点  $P_1$ ,  $\alpha_2 = \{f_5\}$  の終点  $P_6$  では隣接するストロークが存在しないため部分グラフとしては  $N'_{00} = (\{P_1-Q_1\}, \phi)$ ,  $N'_{20} = (\{P_6-Q_5\}, \phi)$  のみを得る. 部分グラフ  $M$  と  $N'_{00}, N'_{11}, N'_{20}$  を結合し, 図 2.16(b) の部分グラフ  $M'$  を作成する.

図 2.17(a) の部分グラフ  $M$  の例では,  $\alpha_1 = \{b\}$  の終点  $P_4$  で部分グラフ  $N'_{10} = (\{P_4-Q_4\}, \phi)$ ,  $N'_{11} = (\{P_4-Q_4\}, \{f_3-g_2\})$ ,  $N'_{12} = (\{P_4-Q_4\}, \{\{f_3, f_2\}-g_2\})$ ,  $N'_{13} = (\{P_4-Q_4\}, \{\{f_3, f_4\}-g_2\})$ ,  $N'_{14} = (\{P_4-Q_4\}, \{f_3-g_5\})$ ,  $N'_{15} = (\{P_4-Q_4\}, \{\{f_3, f_2\}-g_5\})$ ,  $N'_{16} = (\{P_4-Q_4\}, \{\{f_3, f_4\}-g_5\})$  を得る.  $\alpha_0 = \{f_1\}$  の終点  $P_1$ ,  $\alpha_2 = \{f_5\}$  の終点  $P_6$  では隣接するストロークが存在しないため部分グラフとしては  $N'_{00} = (\{P_1-Q_1\}, \phi)$ ,  $N'_{20} = (\{P_6-Q_5\}, \phi)$  のみを得る. 部分グラフ  $M$  と  $N'_{00}, N'_{20}, N'_{10}, \dots, N'_{16}$  それぞれを結合し, 図 2.17(b), ..., (g) の部分グラフ  $M'_1, \dots, M'_6$  を作成する.

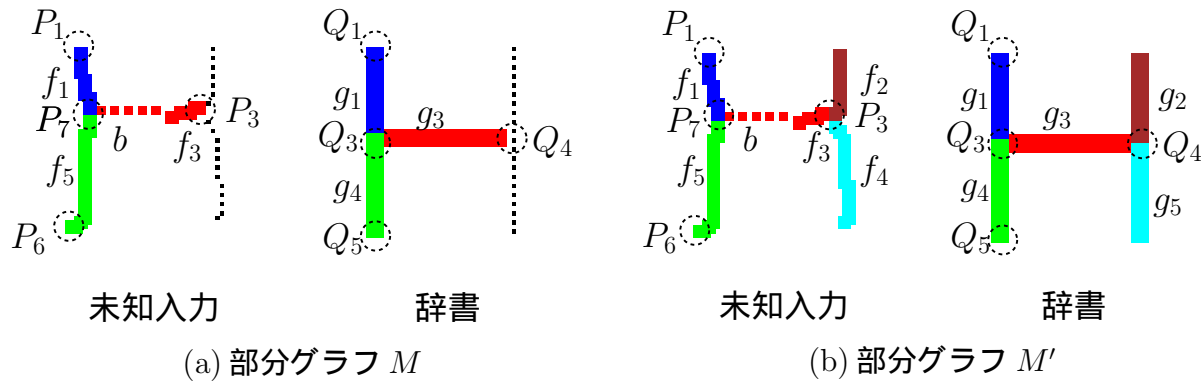


図 2.16: 部分グラフの拡張 (1)

### 2.5.5 コスト計算

マッチンググラフのコストは認識の基準になるように，次のように定義する．

$$\begin{aligned}
 cost = & w_1 \sum_i trans(\alpha_i, g_i) + w_2 \sum_i match(\alpha_i, g_i) \\
 & + w_3 \sum_i length(b_i) + w_4 \sum_i length(f'_i) \\
 & + w_5 \sum_i length(g'_i)
 \end{aligned}$$

上式の第1項の  $trans(\alpha_i, g_i)$  は未知入力のパス  $\alpha_i$  を辞書のストローク  $g_i$  に合わせて変形する際の変形コスト，第2項の  $match(\alpha_i, g_i)$  は未知入力のパス  $\alpha_i$  と辞書のストローク  $g_i$  との整合コストを与える．第3項は使用したブリッジ  $b_i$  の長さ，第4項は使用されなかった未知入力のストローク  $f'_i$  の長さ，第5項は使用されなかった辞書のストローク  $g'_i$  の長さである．また， $w_1 \sim w_5$  は重み定数で実験により定める．

変形コスト，及び整合コストについて詳しく説明する．

#### 変形コスト

未知入力のパスの対応の良さを知るために，対応する辞書のストロークに合うようにパスを変形し，変形の度合を示すコストを定める．

図 2.18(a) は対応する辞書のストロークが線分である未知入力の曲線を持つパス，折線パスにおける変形を示す．その変形コスト  $trans(\alpha_i, g_i)$  はパスの始点と終点を結んだ線分とパスの各点との距離の最大値  $d$  とする．図 2.18(b) は対応する辞書のストロークが曲線の場合のパスの変形を示す．その変形コストはパス上の特異点をはさむストロークが成す角を  $\theta$  とおいて， $d = \cos \frac{\theta}{2}$  と定める．特異点が複数存在する場合はそのような  $d$  の総和を変形コストとする．また，パスが一本の線分であり，対応する辞書のストロークが曲線である場合，曲線に歪みが生じて線分になることは無いという仮定をたて変形コストを  $\infty$  とする

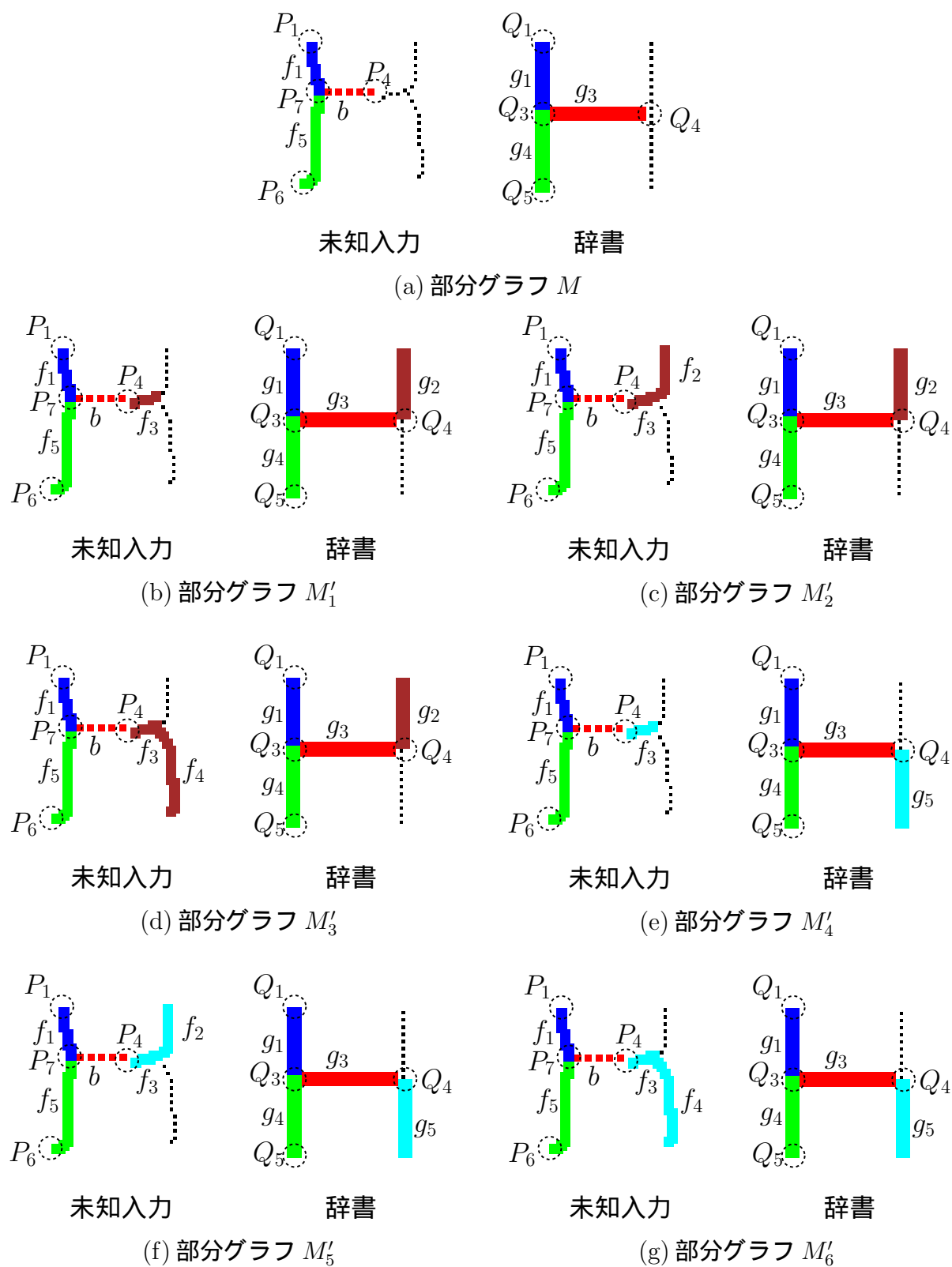
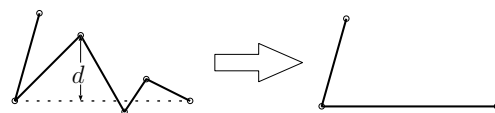
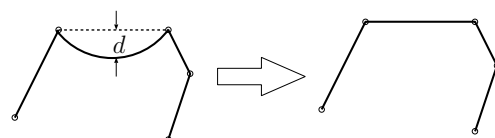
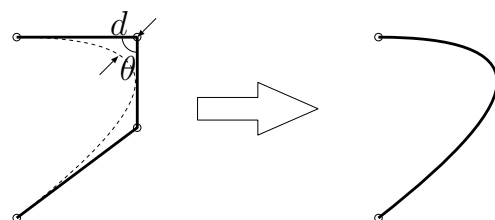


図 2.17: 部分グラフの拡張 (2)



(a) 線分への変形



(b) 曲線への変形

図 2.18: 変形コスト

### 整合コスト

整合コストは、辞書のストロークと未知入力のパスとの一対一対応に対して整合の度合を定めるものである。

- 辞書のストロークが線分の場合 (図 2.19(a))

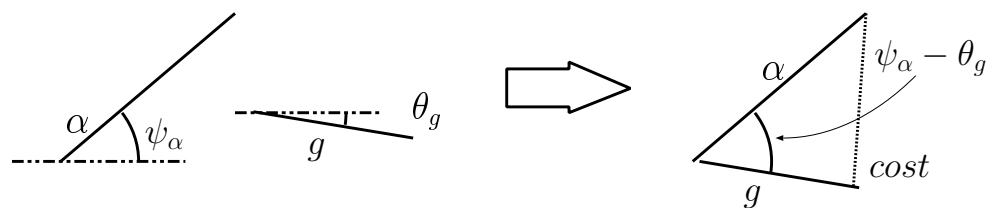
未知入力のパスの始点と終点を結ぶ線分の長さを  $l_s$ 、その始点における線分の水平方向に対する角度を  $\psi_\alpha$  とし、辞書のストロークの長さを  $l_g$ 、その始点におけるストロークの水平方向とに対する角度を  $\theta_g$  とする。このとき、線分間の整合コストを以下の式で定義する。

$$(l_s^2 + l_g^2 - 2l_sl_g \cos(\psi_\alpha - \theta_g))^{\frac{1}{2}}$$

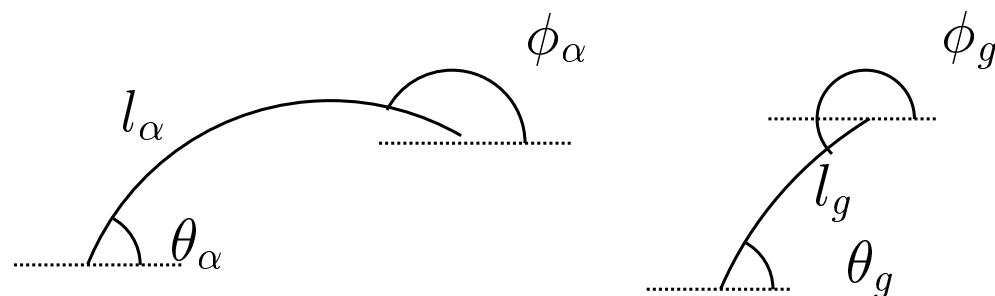
- 辞書のストロークが曲線の場合 (図 2.19(b))

未知入力のパスの長さを  $l_\alpha$ 、辞書のストロークの長さを  $l_g$  とする。また、未知入力と辞書の始点におけるストロークの接線の水平方向に対する角度をそれぞれ  $\theta_\alpha, \theta_g$  とし、未知入力と辞書の終点におけるストロークの接線の水平方向に対する角度をそれぞれ  $\phi_\alpha, \phi_g$  とする。このとき、曲線間の整合コストを以下の式で定義する。

$$\frac{1}{2}\varepsilon_{\text{arc}}(|\theta_\alpha - \theta_g|) + \frac{1}{2}\varepsilon_{\text{arc}}(|\phi_\alpha - \phi_g|) + \varepsilon_{\text{st}}(|l_\alpha - l_g|)$$



(a) 線分間の整合コスト



(b) 曲線間の整合コスト

図 2.19: 整合コスト

ただし,

$$\varepsilon_{arc}(\theta) = \begin{cases} 0 & \text{if } \theta < \theta_0 \\ \theta - 1 & \text{if } \theta \geq \theta_0 \end{cases}$$

$$\varepsilon_{st}(l) = \begin{cases} 0 & \text{if } l < l_g/4 \\ l - l_g/4 & \text{if } l \geq l_g/4 \end{cases}$$

$\theta_0$  は実験で定める.

## 2.5.6 マッチンググラフ拡張の限定

2.5 節で説明した手法により最適なマッチンググラフを見つけることができる. しかし, この手法では可能な部分グラフを全て作成してしまうため, 明らかに不要な部分グラフも作成してしまう. そこで, 本節では 2.5.4 節で説明した部分グラフの拡張を行う際にその拡張を限定する手法を提案する.

余分な部分グラフの拡張を限定するために部分グラフ  $M$  のコスト下限を次の式で定義

する .

$$\begin{aligned} \text{limit} &= w_1 \sum_i \text{trans}(\alpha_i, g_i) + w_2 \sum_i \text{match}(\alpha_i, g_i) \\ &+ w_3 \sum_i \text{length}(b_i) + w_4 \sum_i \text{length}(f_i'') \\ &+ w_5 \sum_i \text{length}(g_i'') \end{aligned}$$

上式第1項～第3項,  $w_1 \sim w_5$  は2.5.5節におけるコスト計算の式の定義と等しい. 第4項は部分グラフ  $M$  を拡張して作成されるどのグラフでも決して使用されない未知入力のストローク  $f_i''$  の長さ, 第5項は決して使用されない辞書のストローク  $g_i''$  の長さである.

コスト下限はその部分グラフ  $M$  を拡張して作成される部分グラフのコストの最小値以下の値をとる. 従って, そのグラフのコスト下限がその時点での全部分グラフのコストの最小値より大きい場合, その時点でのコストの最小値よりコストの小さなグラフを作成するのは不可能なので, 部分グラフの拡張を行わない.

上式第4項, 第5項の計算をするために必要な, 拡張後決して使用されないストロークの条件を以下説明する.

部分グラフ  $M$  に含まれるパス  $\alpha_i = \{f_1, f_2, \dots, f_n\}$ , それに対応する辞書のストローク  $g_i$  について考える. また, パス  $\alpha_i$  の始点を  $P_{si}$ , 終点を  $P_{ei}$  とし, それぞれに辞書の特異点  $b_{si}, b_{ei}$  が対応しているとする. この際, 決して使用されないストロークはグラフ  $M$  に含まれておらず,

1.  $f_j, f_{j+1}$  に挟まれる特異点に隣接している未知入力のストローク
2.  $P_{si}$  と隣接している未知入力のストローク
3.  $b_{si}$  と隣接している辞書のストローク
4.  $b_{ei}$  に隣接している辞書のストロークのすべてが部分グラフ  $M$  に含まれている場合,  $P_{ei}$  に隣接している未知入力のストローク
5.  $P_{ei}$  に隣接している未知入力のストロークのすべてが部分グラフ  $M$  に含まれている場合,  $b_{ei}$  に隣接している辞書のストローク

及び, それに連結しているストロークである.

上記の条件の理由を説明する. (1) については, そのストロークを使用すると未知入力と辞書とでトポロジの異なるマッチンググラフを作成してしまう. (2), (3) については,  $P_{si}, b_{si}$  はすでに部分グラフを作成するための調査済みであり拡張後調査されることはない. (4), (5) については,  $P_{ei}, b_{ei}$  の周りに関する部分グラフを作成する際, 一方のストロークが存在しないので部分グラフとして  $N = (\{a_{ei}-b_{ei}\}, \phi)$  しか作成されない.

図2.20(a)のグラフ  $M_1$  の例では未知入力のストローク  $f_4$  が, パス  $\{f_1, f_2\}$  に注目すると上記(1)より, 拡張後も決して使用されない. また, 辞書のストローク  $g_5$  が, パス  $\{f_6, f_7\}$

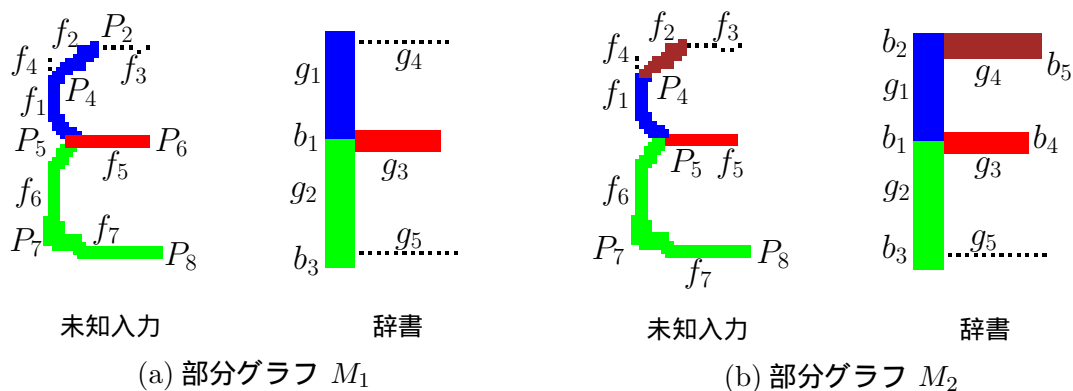


図 2.20: 部分グラフのコスト下限

の終点  $P_8$  及び辞書の特異点  $b_3$  に注目すると上記 (5) より, 使用されない. 図 2.20(b) のグラフ  $M_2$  の例では未知入力のストローク  $f_4$  が, パス  $\{f_2\}$  の始点  $P_4$  に注目すると上記 (2) より, 使用されない. また, 未知入力のストローク  $f_3$  が, パス  $\{f_1, f_2\}$  の終点  $P_2$  及び辞書の特異点  $b_5$  に注目すると上記 (4) より, 使用されない.

## 2.6 認識実験

### 2.6.1 目的

認識実験の目的は, パターン整合法により大分類を行った後にマッチンググラフを用いて細分類を行う飾り文字認識システムの有効性を確認するための実験である.

### 2.6.2 予備実験

コスト計算の際の重み定数を求めるための予備実験を行った. まず, 飾り文字と辞書とのマッチンググラフを作成し, 下のコスト計算の式

$$\begin{aligned}
 cost &= w_1 \sum_i trans(\alpha_i, g_i) + w_2 \sum_i match(\alpha_i, g_i) \\
 &\quad + w_3 \sum_i length(b_i) + w_4 \sum_i length(f'_i) \\
 &\quad + w_5 \sum_i length(g'_i) \\
 &= w_1 E_1 + w_2 E_2 + w_3 E_3 + w_4 E_4 + w_5 E_5
 \end{aligned}$$

の各項を要素とするベクトル  $E = (E_1, E_2, E_3, E_4, E_5)$  を作成する. このベクトルの作成をすべての飾り文字と辞書との組み合わせにおいて行い, 1つのクラスを作成する. また, その組み合わせと同数の零ベクトルを作成し, その零ベクトルからなるクラスを作成する. そ

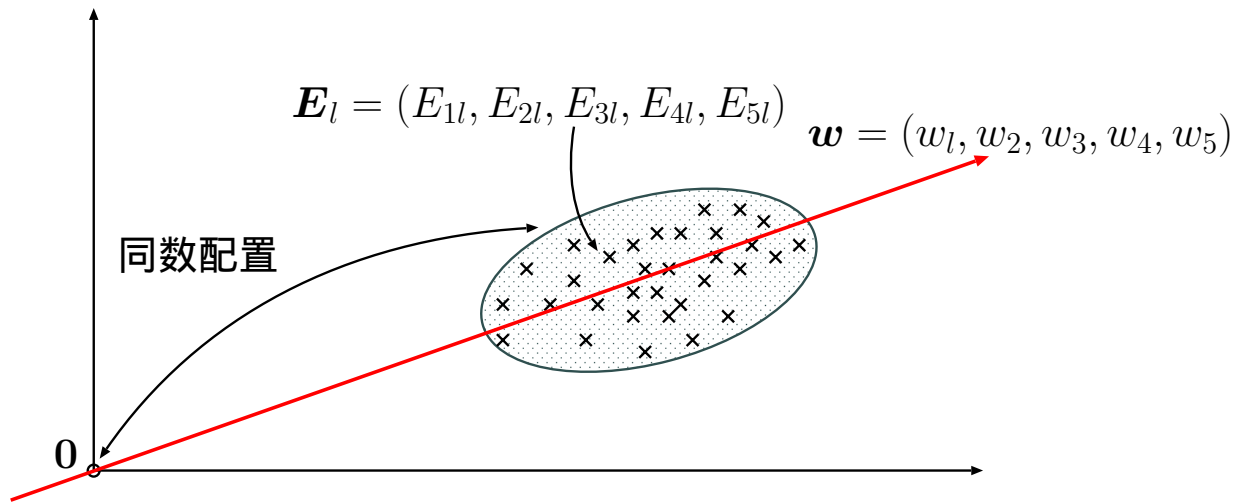


図 2.21: 重み定数の決定

表 2.1: 実験条件

認識対象	アルファベット大文字 26 文字
テストサンプル	飾り文字 10 字種
辞書作成用標準文字パターン	ゴシック体 10 字種

して, その2つのクラスを分割するために, Fisher 比最大となる軸  $w = (w_1, w_2, w_3, w_4, w_5)$  を求める. ここで求めた  $w_1, \dots, w_5$  が重み定数である (図 2.21).

予備実験の結果,  $w_1 = 0.2776551, w_2 = 0.5133225, w_3 = 0.3430888, w_4 = 0.5706035, w_5 = 0.4447392$  を得た.

### 2.6.3 実験手法

まず, 2.3 節にて提案した構造解析手法単独 (図 2.22 手法 A) での認識実験を行う. 次に, 本章で提案した, パターン整合法で大分類した後 2.3 節で提案した構造解析手法を用いて細分類を行う認識手法 (図 2.22 手法 B). により認識実験を行う.

### 2.6.4 実験条件

実験条件を表 2.1 に示す. また, 実験に使用した飾り文字の例を図 2.23 に示す.



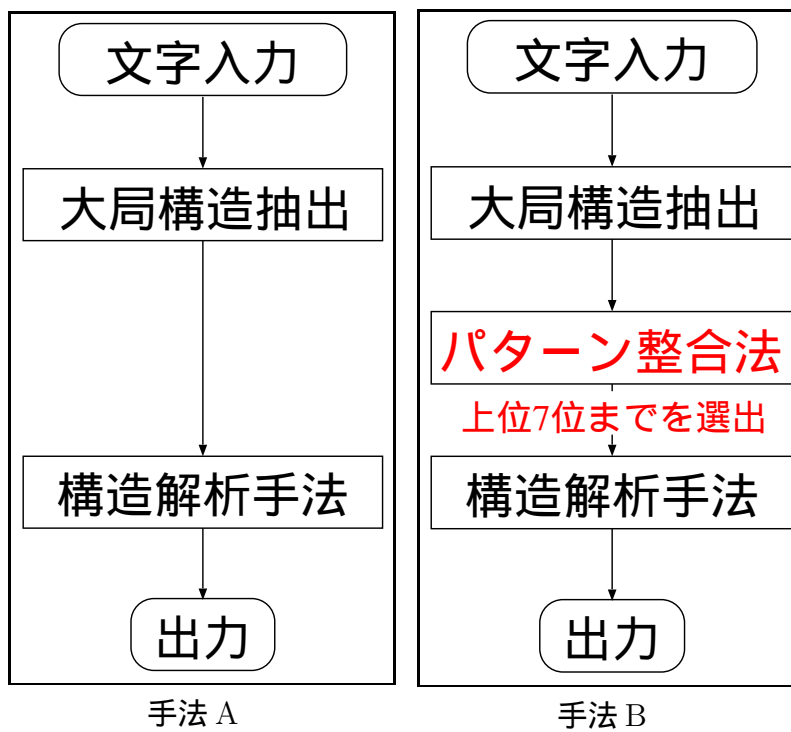


図 2.22: 実験手法

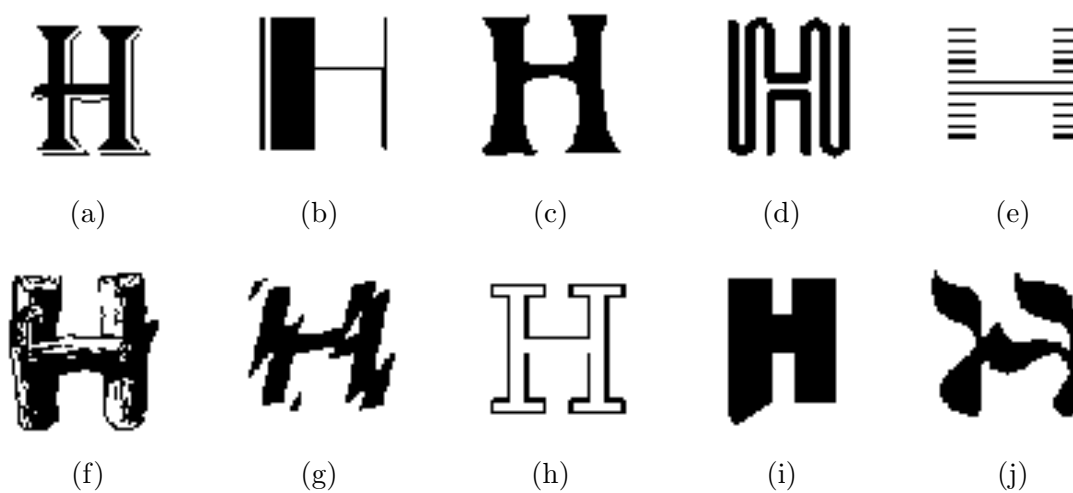


図 2.23: テストサンプル

表 2.2: 実験結果

手法 A	65.0%
手法 B	<b>69.6%</b>
パターン整合法だけの第 1 位認識率	57.7%
(パターン整合法の第 7 位累積認識率 90.8%)	

## 2.6.5 実験結果

実験の結果を表 2.2 に示す。

構造解析手法単独 (手法 A) では認識率は 65.0% であり、パターン整合法だけの場合より認識率が向上している。パターン整合法により大分類した後、2.3 で提案したマッチンググラフを用いた構造解析手法により細分類を行う手法 (手法 B) では 69.6% の認識率を得た。手法 A と比較して認識率が向上していることが確認できた。パターン整合法で大分類を行うことで、マッチンググラフが未知入力とは異なる辞書に対しても柔軟にマッチンググラフを作成してしまい誤認識を引き起こしてしまうのを抑制しているためと考えられる。

図 2.24 に提案手法で正しく認識されたマッチンググラフの例を示す。細破線は使用しなかったストローク、太破線はブリッジを表す。また、同色のストローク同士が対応関係にあることを表す。辞書 “H” に対するマッチンググラフにおいては未知入力、辞書のストロークは全て使用されている。一方、辞書 “N” に対するマッチンググラフは文字 “N” と同じトポロジを持つものの、未知入力に未使用のストロークが存在し、マッチンググラフのコストが大きくなっている。その結果この飾り文字は “H” と認識された。図 2.25 に提案手法で誤認識を起こしたマッチンググラフの例を示す。図 2.24 と同様に細破線は使用しなかったストロークを表し、同色のストローク同士が対応関係にあることを表す。辞書 “T” に対するマッチンググラフにおいても、辞書 “Y” に対するマッチンググラフにおいてもそれぞれの辞書と同じトポロジのグラフを得ることができた、しかし、大局構造の歪みの影響により辞書 “T” に対するマッチンググラフより辞書 “Y” に対するマッチンググラフの方が整合コストが小さくなり誤認識を引き起こした。このように、大局構造の歪みが誤認識を引き起こすものが見られた。

## 2.7 まとめ

本章ではマッチンググラフを用いた構造解析手法を提案した。また、この構造解析手法を組み込んだ飾り文字認識手法を提案した。提案したマッチンググラフを用いた構造解析手法、及び飾り文字認識手法の性能について調査するため認識実験を行った。その結果、パターン整合法で大分類を行った後にマッチンググラフを用いて細分類を行うことで認識精

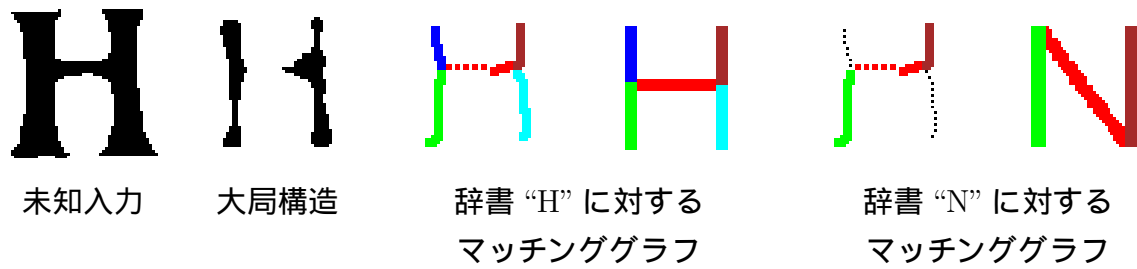


図 2.24: 正読例

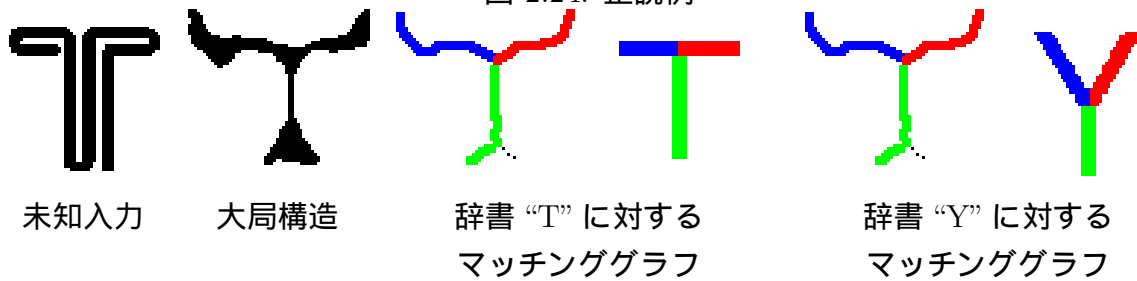


図 2.25: 誤読例

度を向上できることを確認した。また，大局構造の歪みが整合コスト・変形コストの値を大きくし誤認識の原因となっていことがわかった。

## 第3章

# 適応的正規化の導入

### 3.1 はじめに

第2章では、マッチンググラフを用いた構造解析手法と、それを組み込んだ飾り文字認識システムを提案した。その飾り文字認識システムにおいて、抽出された飾り文字の大局構造の歪みにより誤認識を起こすものが存在した。

第2章で提案した飾り文字認識システムにおいて、マッチンググラフの作成には辞書の知識を用いたがその前処理には一切辞書の知識を用いていない。図3.1に示すように、マッチンググラフ作成の前処理において、特徴抽出を行う前に辞書の知識を用いて歪みを除去することができれば認識精度の向上を実現できると考えられる。そこで、本章ではマッチンググラフ作成の前処理に、特徴抽出を行う前に Global Affine Transformation(GAT), Local Affine Transformation(LAT) を用いた適応的正規化 [8] を導入する。

3.2節において前処理に導入する適応的正規化について述べる。3.3節において飾り文字の認識実験を行い、導入した適応的正規化の有効性を確認する。

### 3.2 適応的正規化

本節では飾り文字から得られた大局構造から誤認識の原因となる歪みを除去するための適応的正規化 [8] について述べる。適応的正規化の概要を図3.2に示す。また、図3.3に図3.3(a)の飾り文字を図3.3(g)の辞書“T”に近づくよう適応的正規化を行った際の各処理における未知入力画像の変形の様子を示す。

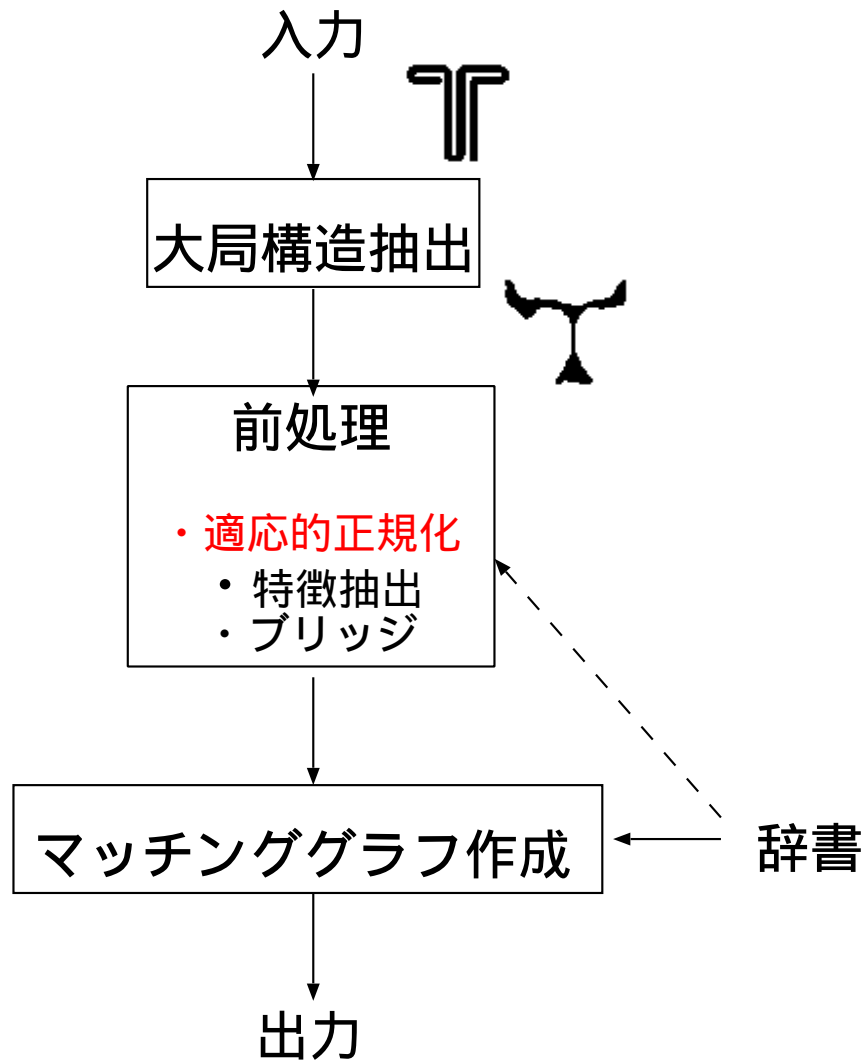


図 3.1: 適応的正規化の導入

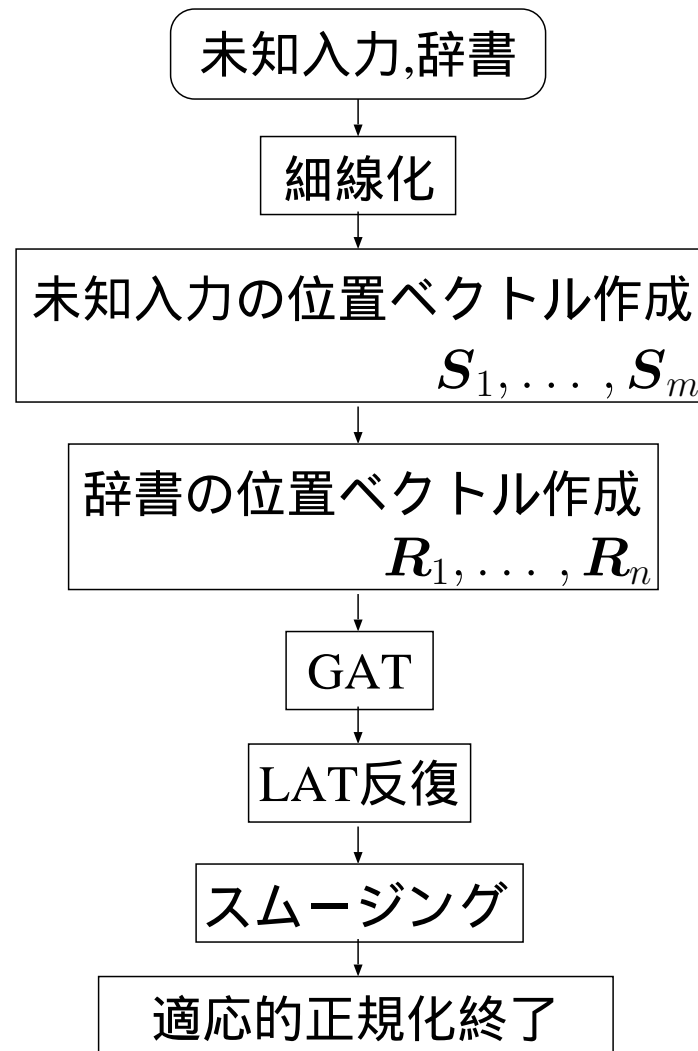


図 3.2: 適応的正規化の概要

### 3.2.1 位置ベクトル作成

最初に未知入力と辞書の細線化を行う (図 3.3(c)) . 次に未知入力の画像中の黒画素の座標を示す位置ベクトル  $S = \{s_1, s_2, \dots, s_m\}$  を作成し , 辞書の画像中の黒画素の座標を示す位置ベクトル  $R = \{r_1, r_2, \dots, r_n\}$  を作成する .

### 3.2.2 Global Affine Transformation(GAT)

前節で得られた位置ベクトルをもとに GAT (大局的アフィン変換) を実行する (図 3.3(d)) . 以下の式により , 未知入力の各位置ベクトル  $s_i (i = 1, \dots, m)$  を  $s_i^{\text{GAT}}$  に変換する .

$$s_i^{\text{GAT}} = As_i + b$$

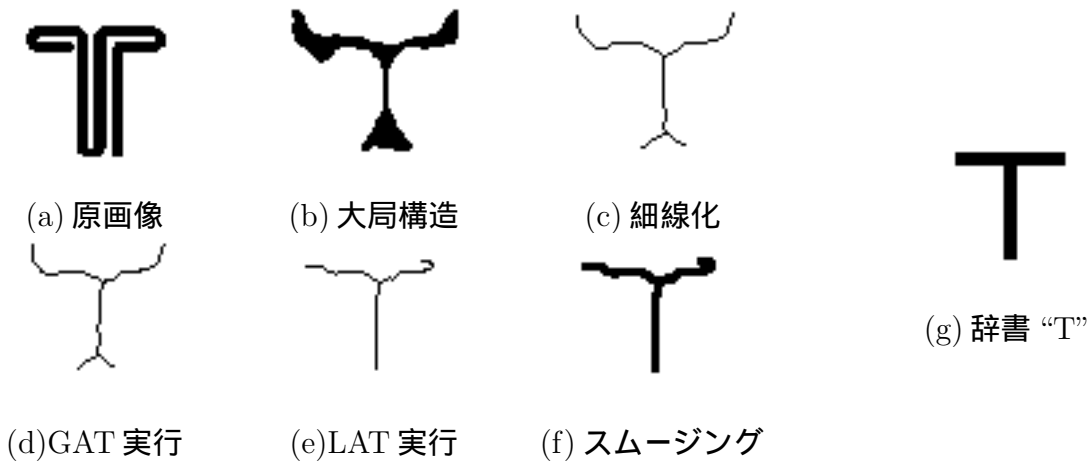


図 3.3: 適応的正規化

上式の  $A$  は  $2 \times 2$  の行列であり，回転，縮尺の変更を表し， $b$  は 2 次元のベクトルであり平行移動を表す．上式の行列  $A$ ，及びベクトル  $b$  は

$$\Phi = \frac{1}{2} \sum_i^m \sum_j^n \rho_{ij}(D) \|s_i^{GAT} - r_j\|^2 \quad (3.1)$$

を最小にするような行列，及びベクトルである．但し，式 3.1 の  $\rho_{ij}(D)$  は

$$\rho_{ij}(D) = \frac{\mu_{ij}(D)}{m} + \frac{\nu_{ji}(D)}{n} \quad (3.2)$$

であり，式 3.2 の  $\mu_{ij}(D)$ ， $\nu_{ji}(D)$  は

$$\mu_{ij}(D) = \exp \left[ -\frac{\|s_i - r_j\|^2 - \min_k \|s_i - r_k\|^2}{D} \right] \quad (3.3)$$

$$\nu_{ji}(D) = \exp \left[ -\frac{\|s_i - r_j\|^2 - \min_k \|s_k - r_j\|^2}{D} \right] \quad (3.4)$$

この  $\rho_{ij}$  は次節で述べる LAT でも用いる．GAT では， $D$  を次のように定める．

$$D = \frac{1}{2} \left[ \frac{1}{m} \sum_i^m \min_j \|s_i - r_j\|^2 + \frac{1}{n} \sum_j^n \min_i \|s_i - r_j\|^2 \right] \quad (3.5)$$

### 3.2.3 Local Affine Transformation(LAT)

前節で GAT により変換した位置ベクトルに対し，更に LAT(局所的アフィン変換) を反復し未知入力の変形を行う (図 3.3(e)) ．

前節により変換した  $s_i^{\text{GAT}}$  を  $s_i^{\text{LAT}(0)}$  とおく． $l$  回目の LAT により，得られる未知入力  
各位置ベクトル  $s_i^{\text{LAT}(l)}$  は以下の式により求められる．

$$s_i^{\text{LAT}(l)} = A_i^{(l)} s_i^{\text{LAT}(l-1)} + b_i^{(l)}$$

に変換する．上式の  $A_i^{(l)}$  は  $2 \times 2$  の行列であり， $b_i^{(l)}$  は 2 次元のベクトルであり， $A_i^{(l)}$ ，及  
び  $b_i^{(l)}$  は

$$\Psi_i = \frac{1}{2} \sum_j^m \omega_{ij}(\theta^{(l)2}) \sum_k^n \rho_{jk}(\sigma^{(l)2}) \|A_i^{(l)} s_j^{(l-1)} + b_i^{(l)} - r_k\|^2 \quad (3.6)$$

を最小にする行列，及びベクトルである．式 3.6 の  $\rho_{jk}()$  の定義は式 3.2 のものに等しい．  
 $\omega_{ij}$  は

$$\omega_{ij}(\theta^2) = \exp \left[ -\frac{\|s_i - s_j\|^2}{\theta^2} \right] \quad (3.7)$$

で与えられる．また，式 3.6 の  $\theta^{(l)}$ ， $\sigma^{(l)}$  は以下の式で与えられる．

$$\theta^{(l)} = \theta_0 / \sqrt{l} \quad (3.8)$$

$$\sigma^{(l)} = \gamma \times \theta^{(l)} \quad (3.9)$$

上式の  $\theta_0, \gamma$  は定数である．

### 3.2.4 スムージング

前節の LAT までを実行した細線の画像には細かい歪みが生じている．これを除去するた  
めスムージングの目的で黒画素の拡張処理を行う (図 3.3(f))．

スムージングまでを行ったものが適応的正規化の結果となる．

### 3.2.5 適応的正規化の実行例

図 3.4，図 3.5，図 3.6 に適応的正規化を実行した例を示す．それぞれの図において，(b)  
の画像に適応的正規化を実行した結果が (c), ..., (h) である．図より，大局構造の形状が辞  
書の形状に近づくように変形していることが確認できる．

### 3.2.6 適応的正規化のコスト

図 3.4，図 3.5，図 3.6 より，適応的正規化の効果で大局構造が辞書の形状に近づいてい  
るのがわかる．しかし，未知入力と辞書との字種が異なる際にも非常に形状が似通ってし



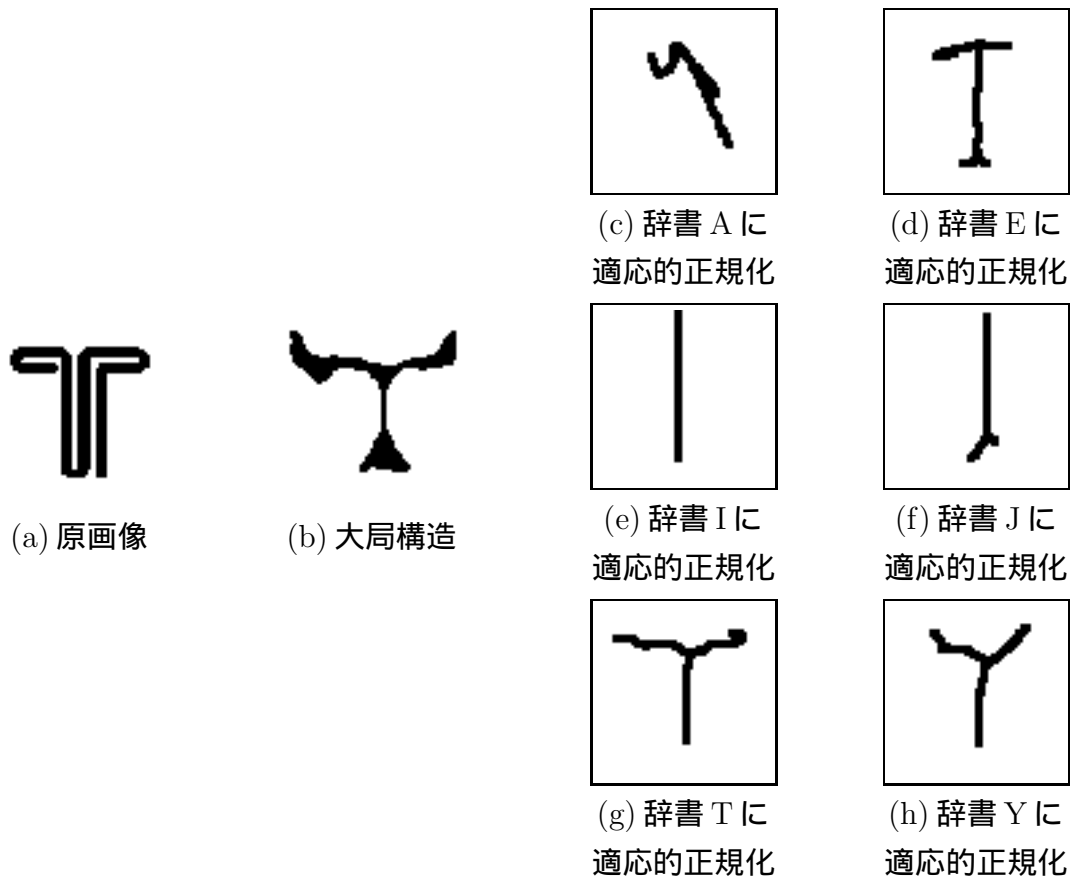


図 3.4: 適応的正規化の実行例 (その 1)

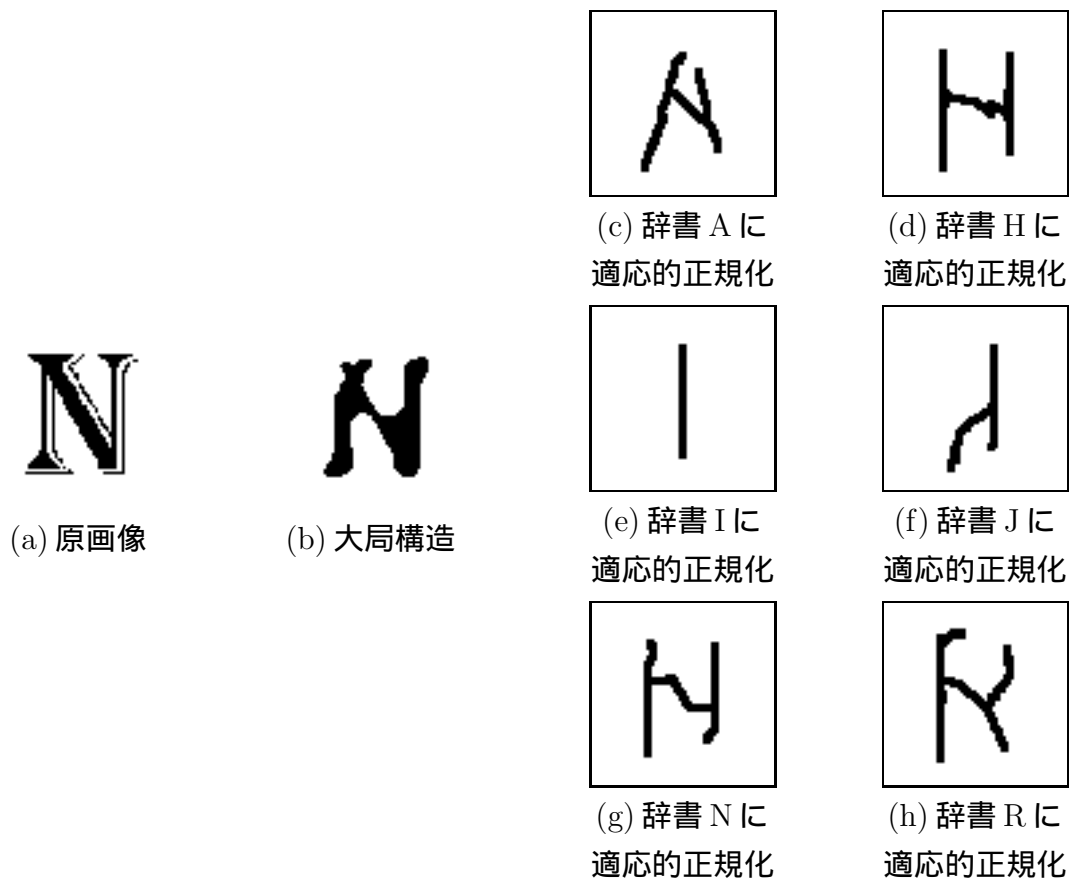


図 3.5: 適応的正規化の実行例 (その 2)

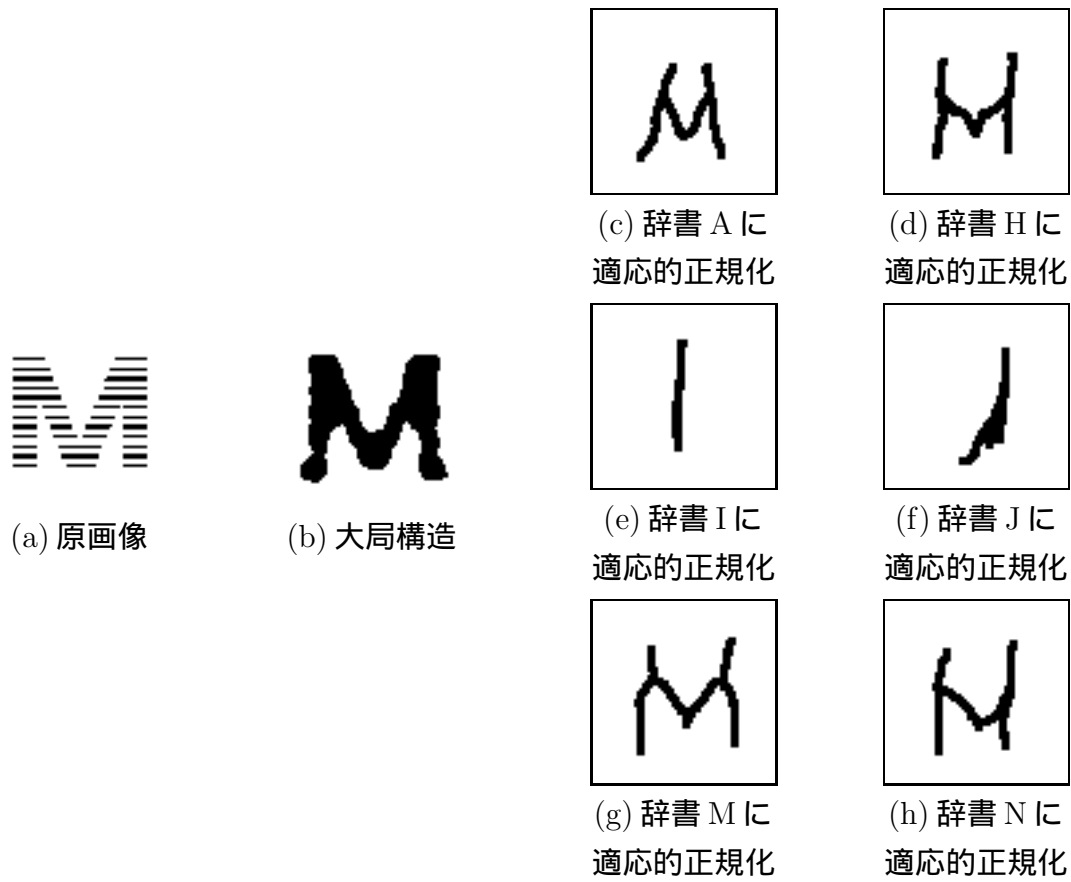


図 3.6: 適応的正規化の実行例 (その 3)

まっていることもわかる．そこで，マッチンググラフを作成する際，適応的正規化によって変形した度合いに応じてコストをを加える．

未知入力の各位置ベクトル  $s_i$  が適応的正規化により  $s_i^*$  に変形した場合，適応的正規化のコスト  $D$  を

$$D = \sum_i \|s_i^* - s_i\|$$

で定義する．これは，未知入力の各黒画素の移動量をユークリッド距離で表したとき，その総和となっている．

### 3.3 認識実験

#### 3.3.1 目的

本節で行う認識実験の目的は適応的正規化の有効性を確認することであり，前章の実験の手法 B と比較を行うことでその有効性を示す．

#### 3.3.2 予備実験

前章と同様の予備実験を行い，重み定数  $w_1 = 0.3603173$ ,  $w_2 = 0.4479289$ ,  $w_3 = 0.4083128$ ,  $w_4 = 0.5075907$ ,  $w_5 = 0.4950728$  を得た．但し， $w_6$  に関してはヒューリスティックな手法を用い， $w_6 = 1$  を得た．

#### 3.3.3 実験手法

手法 C として，パターン整合法により大分類を行い，前章で提案したマッチンググラフを用いた構造解析手法により細分類を行う．但し，マッチンググラフ作成の前処理として適応的正規化を導入する (図 3.7)．

#### 3.3.4 実験条件

実験条件を表 3.1 に示す．また使用した飾り文字の例を図 3.8 に示す．

#### 3.3.5 実験結果

実験結果を表 3.2 に示す．なお，比較のため本章で提案した適応的正規化を導入していない場合 (前章，手法 B) の認識率も同表に入れた．

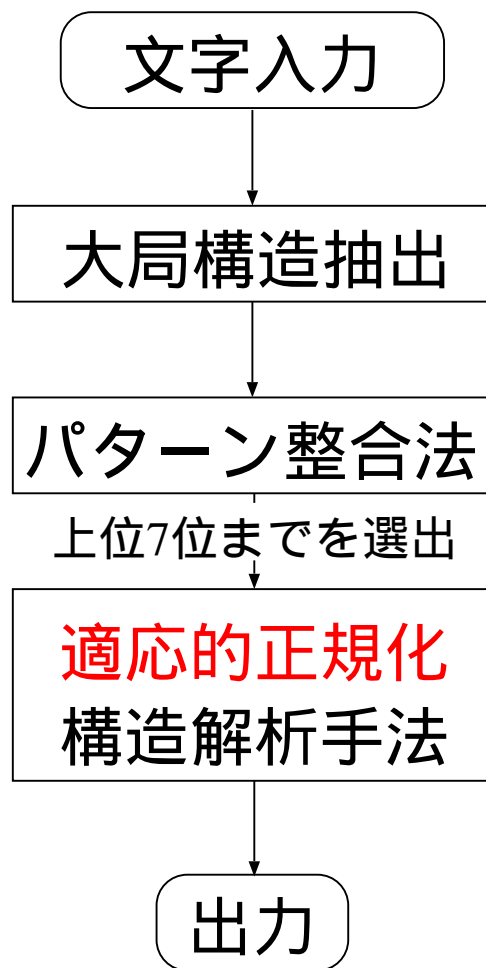


図 3.7: 実験手法

表 3.1: 実験条件

認識対象	アルファベット大文字 26 文字
テストサンプル	飾り文字 10 字体
辞書作成用標準文字パターン	ゴシック体 10 字体

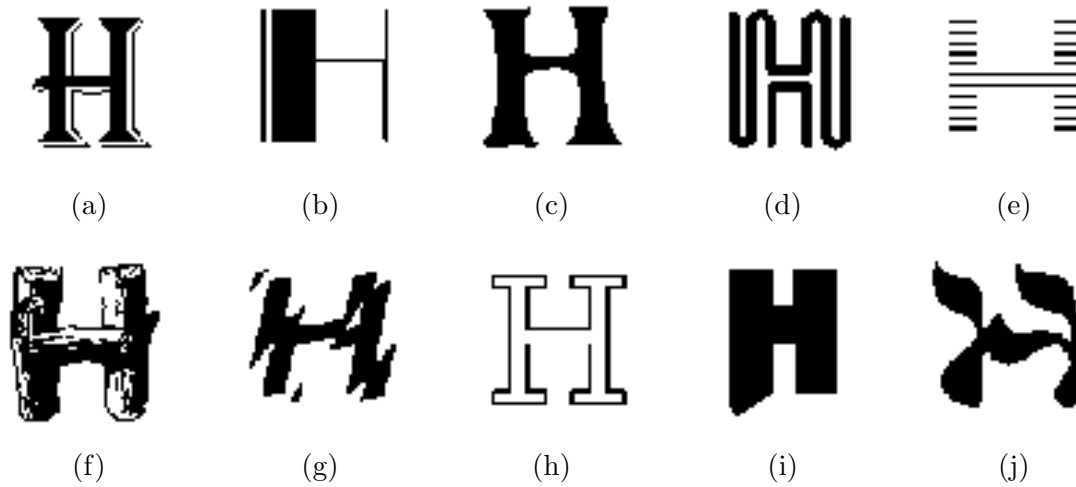


図 3.8: テストサンプル

表 3.2: 適応的正規化を導入した際の認識率

適応的正規化未導入 (手法 B)	69.6%
適応的正規化導入 (手法 C)	70.0%

(パターン整合法の第 4 位累積認識率 90.8%)

認識実験の結果，適応的正規化未導入 (手法 B) の際には誤認識を起こしていたが，適応的正規化を導入 (手法 C) することにより正しく認識した飾り文字は 8.1%，反対に手法 B では正しく認識していたが，手法 C では誤認識を起こしていた飾り文字は 7.7% あった．すなわち，適応的正規化未導入の場合と比較し，0.4% に認識率の向上が確認できた．このことは適応的正規化の有効性を示している．

図 3.9 に正しく認識できた飾り文字の例を示す．適応的正規化を導入しなかった場合，“Y” と誤認識を起こしていたが，適応的正規化を導入することにより，大局構造の歪みを除去することができ正しく認識できた．

図 3.10 に誤認識を起こした飾り文字の例を示す．適応的正規化を導入しなかった場合，正しく認識することができていたが，適応的正規化を導入することにより，形状が “N” よりも “R” に近づいてしまい，その結果誤認識を起こしていた．

### 3.4 まとめ

本章ではマッチンググラフを作成する際，大局構造の歪みがコストを増大させ誤認識を引き起こすことを防ぐため，マッチンググラフ作成の前処理において，特徴抽出を行う前に適応的正規化を導入した．この正規化の有効性を確認するため認識実験を行った．その

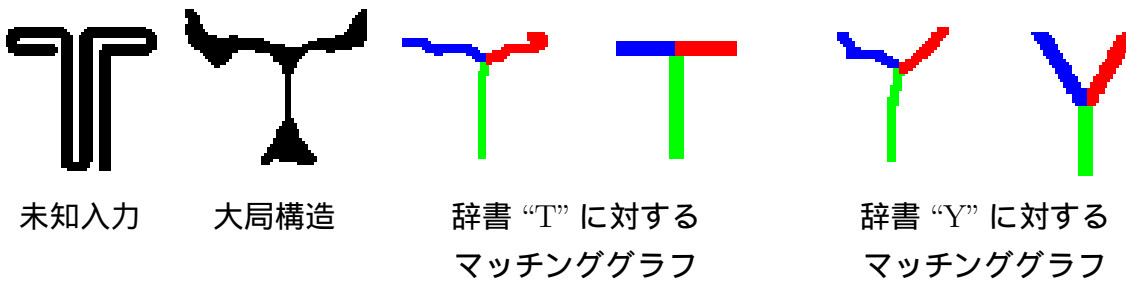


図 3.9: 正読例

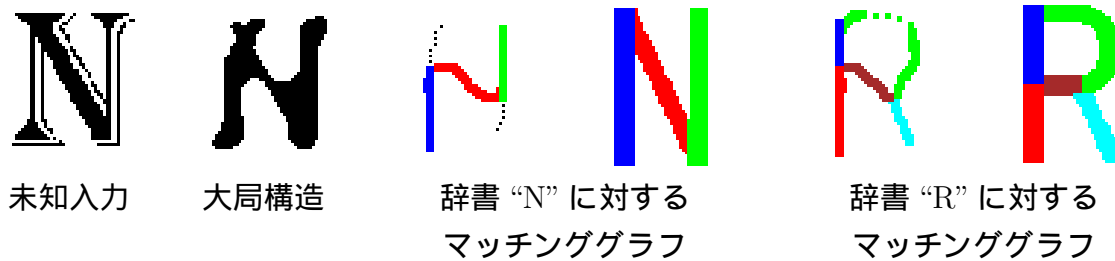


図 3.10: 誤読例

結果，認識率の向上が確認でき適応的正規化の有効性が確認できた。

しかし，適応的正規化を導入することで，異なる字種の辞書との対応関係を表すマッチンググラフのコストまでが小さくなってしまい，適応的正規化を導入していなければ認識可能であった飾り文字も誤認識を起こしてしまう例が見られた．複数の文字認識アルゴリズムの複合化により認識精度の向上が可能であることは既に報告されている [9] が，適応的正規化を導入しなかった際の認識結果と適応的正規化を導入した際の認識結果を複合することで認識精度の向上が望めると考えられる。

## 第4章

# 複合化処理の導入

### 4.1 はじめに

第3章において導入した適応的正規化により、第2章において提案した飾り文字認識手法の精度を向上させることができた。しかし、適応的正規化を導入することで、未導入の際には正しく認識できていた飾り文字が誤認識を起こしている例も確認できた。このことから、適応的正規化を未導入の際の結果と導入した際の結果を複合することにより高精度な認識を行える可能性がある。そこで、本章では、適応的正規化を未導入の認識結果と適応的正規化導入の認識結果を複合する処理を加えた飾り文字認識システムを構築し、その性能を実験を通して確認する。

4.2節では複合化を導入した飾り文字認識システムについて提案し、その複合方法及び複合の基準となる信頼度の定義について述べる。4.3節では複合化の基準として定義した信頼度の妥当性について調査する。4.4節において、認識実験を行うことで、提案した複合化処理を導入した飾り文字認識システムの有効性を確認する。

### 4.2 複合化方法

複数の文字認識手法の複合化手法を考えた場合、大きく分けて二種類の方法がある。一つは、各文字の認識結果に対して、複数の文字認識手法の内いずれか一つの結果を選択する方法である。各文字認識手法に対して各入力文字毎に認識結果の信頼性を示す評価値を与え、最も評価値が高かった手法の結果をその入力文字の結果とする。この方法の場合、文字認識手法の数だけ評価値を計算すれば、後はその大小で認識結果を比較するだけである。よって出力候補数に制限されることなく評価値の与え方を決定できる。また、信頼性の評価方法によっては、ある程度複合化処理に有する時間を抑えることが可能である。し



かし，評価をどのように取るかだけでなく，認識手法ごとに重み付けや異なる評価値の与え方などを行って，評価が一定になるようにする必要がある．

もう一つは，複数の文字認識手法を全て合わせて新たな認識結果を得るものである．一つの入力文字に関して，各文字認識手法で予め決められた数だけ候補文字が距離値を持って出力される．その出力間で共通する認識結果の距離値の例えば和をとり，その値で改めて小さなものから順に並べ換えて最終的な認識結果とするわけである．この方法を用いると，距離値が小さいものから順に上位候補とするような文字認識手法であれば，容易に他の手法と組み合わせることができる．しかし，文字認識手法の出力候補数に制限があると，認識結果で存在したり存在しなかったりする文字が現れる．そういう文字に対しては，存在しなかった文字認識手法においては最下位候補の距離値を使う，あるいは入力文字と全認識対象文字とで評価を行うということをする必要がある [9]．

本研究では，前処理の方法は異なるものの，基本的な文字認識手法は共通したものをを用いている．このことから，認識結果の信頼性を示す評価値の与えかたは適応的正規化を導入したものと導入しなかったものとの全く同じ計算式を使うことにほぼ問題はないと考えられる．そこで，複合化方法として，複数の文字認識手法の内いずれか一つの結果を選択する方法を採用することにした．また，認識結果の信頼性の評価方法であるが，認識結果の信頼度を

$$\text{信頼度} = \frac{\text{二位候補のコスト}}{\text{一位候補のコスト}}$$

と定義し，適応的正規化未導入の結果，適応的正規化導入の結果それぞれから信頼度を計算し，信頼度の値の大きい方の結果を最終的な結果とする．上式で求められた信頼度は，一位候補が正解である場合に大きく，そうでない場合は小さくなる傾向があることが確認されており，実際に複数の認識結果の複合 [10][11]，文字認識結果の確信度 [12] などに用いられている．今回用いる複合化処理を導入した飾り文字認識システムの概要を図 4.1 に示す．

### 4.3 信頼度の妥当性

前節において，一位候補と二位候補のコスト比を信頼度と定義し，適応的正規化未導入の場合と適応的正規化導入の場合で信頼度の大きい方の認識結果を最終的な結果とすることにした．この手法は全く異なる文字認識手法の複合化においても有用であった [9]．更に，本研究の場合には基本的に認識手法は同じであるので，この信頼度により二者を比較することに問題はないと思われる．しかし，万全を期すため，ここで予備実験により確認しておく．

実験の方法として，信頼度と全誤認識文字数に対するその信頼度以上の結果中に含まれる誤認識文字数の割合の関係を求めた．実験結果を図 4.2 に示す．この結果から分かるように，両者の傾向は類似しており，どちらも信頼度が大きくなるにつれて急激に誤認識文

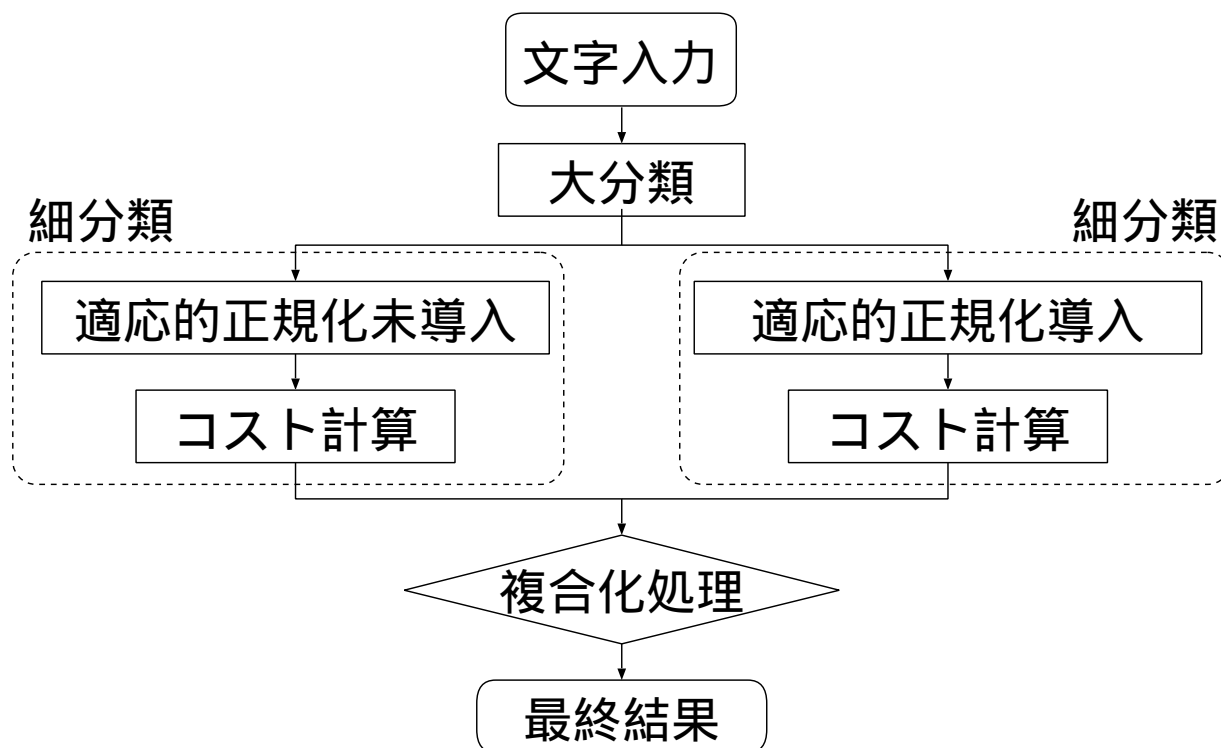


図 4.1: 複合化処理を導入した飾り文字認識システムの概要

字の存在比率が小さくなっている．このことから，前節で定義した信頼度は二者の比較に用いるのに妥当であることが確認できた．

## 4.4 認識実験

### 4.4.1 目的

本章の実験の目的は複合化処理の有効性を確認することであり，第2章の手法Bの認識結果と第3章の手法Cの認識結果と比較し，その有効性を示す．

### 4.4.2 実験手法

手法Dとして，マッチンググラフ作成の前処理に適応的正規化を導入しない場合(手法B)と適応的正規化を導入した場合(手法C)の認識結果を複合化する(図4.3)．

### 4.4.3 実験条件

実験条件を表4.1に示す．また，実験に使用した飾り文字の例を図4.4に示す．

## 誤認識文字の存在率

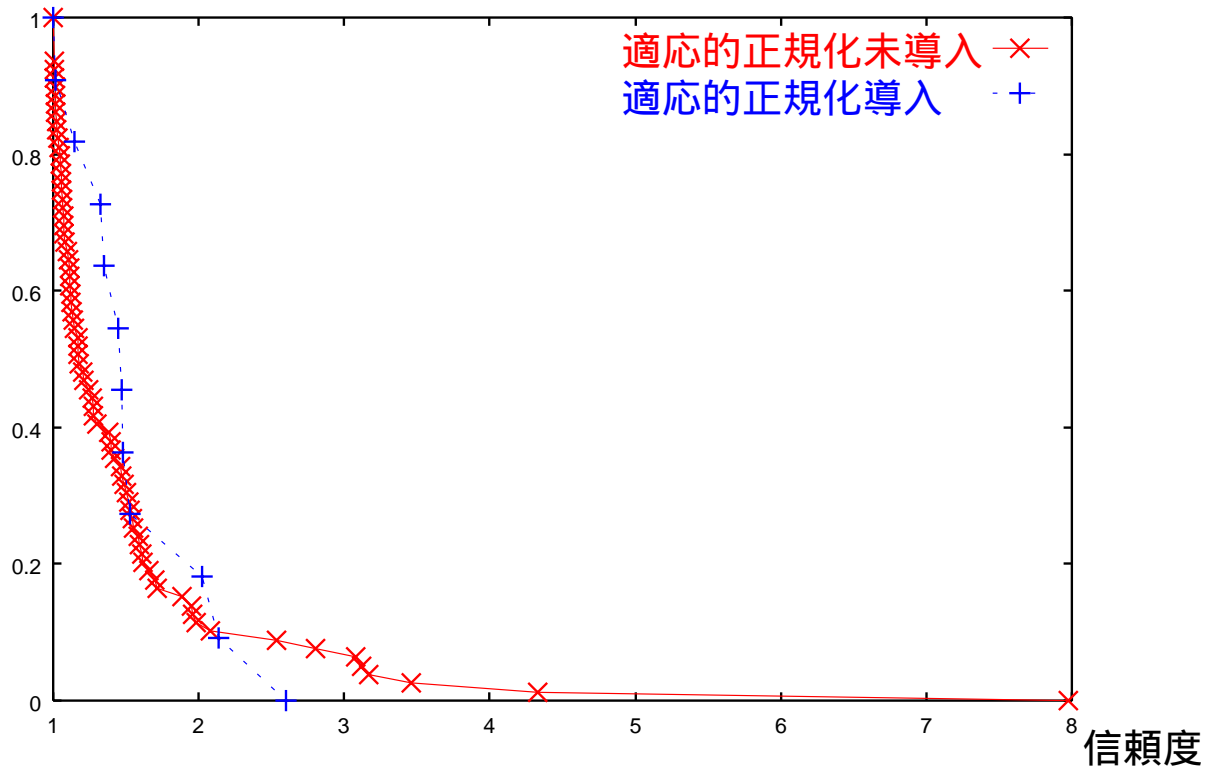


図 4.2: 信頼度と誤認識文字存在率

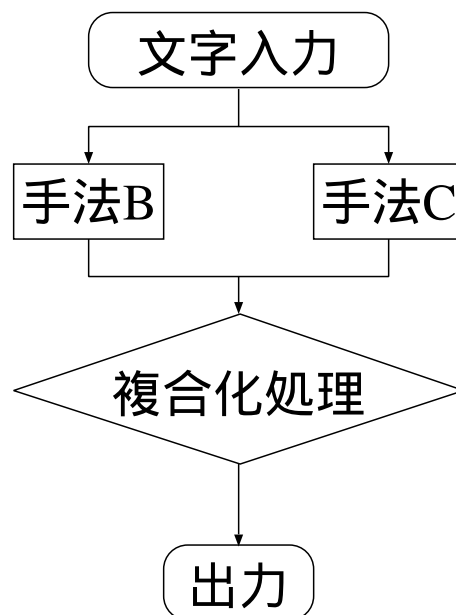


図 4.3: 実験手法

表 4.1: 実験条件

認識対象	アルファベット大文字 26 文字
テストサンプル	飾り文字 10 字体
辞書作成用標準文字パターン	ゴシック体 10 字体

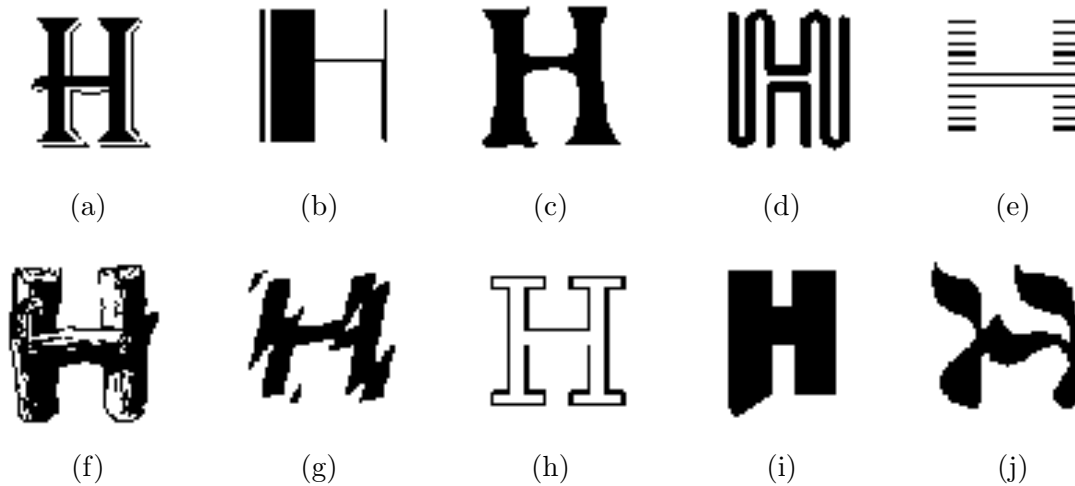


図 4.4: テストサンプル

#### 4.4.4 実験結果

実験結果を表 4.2 に示す．なお，比較のため，適応的正規化未導入の場合の認識率（手法 B）と適応的正規化導入（手法 C）の場合の認識率も同表に入れた．

実験結果により，手法 C と比較し 3.8% 認識率が向上し 73.8% の認識率を得たことが確認できた．実験結果の詳細を図 4.5 に示す．図の正誤正は，手法 B で正しく認識し，手法 C で誤認識，手法 D で正しく認識した文字の割合を示す．手法 B，手法 C の一方で正しく認識し，他方で誤認識を起こしていた文字（図 4.5 の誤正正，正誤正，誤正誤，正誤誤の和）は全体の 15.8% であった．複合化処理によって，そのうちの 75.3%（図 4.5 の誤正正と正誤正の和，全体の 11.9%）の文字が正しく認識した．このことから，本章で提案した複合化処理の有効性を確認することができた．

表 4.2: 複合化処理を導入した際の認識率

適応的正規化未導入（手法 B）	69.6%
適応的正規化導入（手法 C）	70.0%
複合化処理導入（手法 D）	73.8%

（パターン整合の第 4 位累積認識率 90.8%）

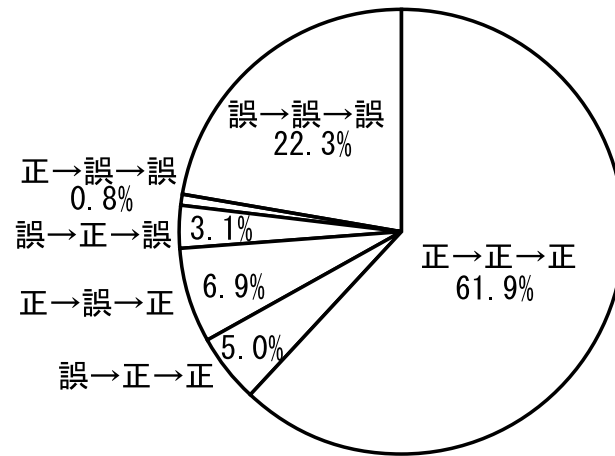


図 4.5: 複合化処理を導入した結果

## 4.5 まとめ

本章では適応的正規化を導入することで認識精度が向上するものの、反対に、適応的正規化の導入が原因で誤認識を引き起こしている文字が存在するという前章の結果を受け、適応的正規化を導入しなかった際の認識結果と適応的正規化を導入した際の認識結果を複合する手法を加えた飾り文字認識システムを提案した。複合化処理を行う際に用いる両認識結果の信頼性の評価方法として、一位候補と二位候補とのコスト比を信頼度として用いたが、その評価が信頼性を数値化する上で妥当なものであることを実験により確認した。また、提案した複合化処理を用いて実際に飾り文字認識実験を行い、その有効性を確認した。

# 第5章

## 結論

### 5.1 本研究の成果

本研究では、高精度な飾り文字認識手法の提案を目的とした。その研究の成果をまとめると以下の通りである。

第2章では飾り文字認識手法を提案した、飾り文字認識手法の有効性を確認するため、飾り文字を対象に認識実験を行い、69.6%の認識率を得た。

第3章では大局構造の歪みによる誤認識を防ぐため、マッチンググラフ作成の前処理において Global Affine Transformation, Local Affine Transformation を適用する適応的正規化を導入し歪みの除去を試みた。認識実験の結果、0.4%の認識率の向上が確認でき、70.0%の認識率を得、適応的正規化のある程度の有効性が確認できた。

第4章では複合化処理を加えた飾り文字認識手法を提案した。複合化処理では、一位候補と二位候補のコストの比を評価値として、二つの認識結果のうち一方を選択する方法をとったが、この評価値を用いるに当たり、その妥当性を実験により確認した。最後に、この複合化処理を導入した飾り文字認識システムの有効性を確認するため、実際に飾り文字認識実験を行った。その結果3.8%認識率を向上させることができ、最終的に73.8%の認識率を得ることができた。

### 5.2 今後の課題

今後の課題として以下のことが挙げられる。

本研究では飾り文字から抽出された大局構造がもとの文字のトポロジを保持していることから、一貫して構造解析手法により飾り文字認識を行ってきた。しかし、大局構造に大きな歪みが生じている場合に、作成されたマッチンググラフのコストが大きくなったり、あ

るいはその形状が類似するトポロジを有する文字に近づいてしまい誤認識を起こしている例があった．そこで，大局構造の歪みを除去できるように前処理の再検討，トポロジの類似している文字の高精度な識別が必要であると考えられる，

複合化処理を導入することで認識精度を向上させることができた．しかし，2つの認識手法で共通して誤っている文字は全体の22.3%であり，今回複合化処理により得られた認識率は73.8%(誤認識率26.2%)であった．更なる高精度化が必要と考えられる．

## 参考文献

- [1] 鈴木雅人, 大町真一郎, 加藤寧, 阿曾弘具, 根元義章. 混合マハラノビス関数による高精度な類似文字識別手法. 電子情報通信学会論文誌 (D-II), Vol. J80-D-II, No. 10, pp. 2752–2760, October 1997.
- [2] 若林哲史, 鄧巖, 鶴岡信治, 木村文隆, 三宅康二. 非線形正規化と特徴量の圧縮による手書き漢字認識の高精度化. 電子情報通信学会論文誌 (D-II), Vol. J79-D-II, No. 5, pp. 765–774, May 1996.
- [3] 山本和彦. 弛緩整合法による手書き教育漢字認識. 電子情報通信学会論文誌 (D), Vol. J65-D, No. 9, pp. 1167–1174, September 1982.
- [4] J. Rocha and T. Pavlidis. A shape analysis model with application to a character recognition system. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 16, No. 4, pp. 393–404, April 1994.
- [5] R. M. Haralick, L. T. Watson, and T. J. Laffey. The topographic primal sketch. *International Journal of Robotics Research*, Vol. 2, No. 1, pp. 50–72, January 1983.
- [6] 大町真一郎, 井上政樹, 阿曾弘具. マルチスケール画像を用いた飾り文字の構造抽出. 信学技報, PRMU98-30, pp. 1–8, 1998.
- [7] C.J.Hilditch. "Linear Skelton from Square Cupboards" In *Machine Intelligence 6*, pp. 403–420. B.Meltzer & D.Michie,Eds., Univ. Press, Edinburgh, 1969.
- [8] T. Wakahara and K. Odaka. Adaptive normalization of handwritten characters using global/local affine transformation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 20, No. 12, pp. 1332–1341, December 1998.
- [9] 松井俊弘, 山下郁生, 若原徹, 吉室誠. 文字認識アルゴリズムの複合化手法の検討. 信学技報, PRU92-33, pp. 65–72, 1992.



- 
- [10] 是角有二, 横塚志行. 複数の認識結果の複合による手書き漢字認識の高精度化. 電子情報通信学会総合大会, D-549, p. 275, 1995.
- [11] 山中清. 多フォント文字認識手法に関する基礎的研究. 東北大学大学院工学研究科 電気・通信工学専攻 修士学位論文, 1996.
- [12] 中山由美, 横塚志行. 文字認識結果の確信度に関する一検討. 電子情報通信学会総合大会, D-541, p. 267, 1995.

# 研究業績

女川俊一，大町真一郎，阿曾弘具，  
“非決定性形状分析モデルによる飾り文字認識，”  
画像の認識・理解シンポジウム (MIRU2000)，  
2000年7月．

女川俊一，大町真一郎，阿曾弘具，  
“マッチンググラフを用いた飾り文字認識，”  
電子情報通信学会技術研究報告，  
PRMU2000-137，2000年12月．

# 謝辞

本研究を進めるにあたり，東北大学大学院工学研究科 阿曾弘具教授には折りに触れて問題点を指摘して頂くなど，全般的な御指導を賜りました．心より感謝致します。

本論文をまとめるにあたり貴重な御意見を頂いた東北大学大学院工学研究科 川又政征教授，東北大学大学院情報科学研究科 堀口剛教授に深く感謝致します．

本研究全般に渡り親身に御指導、御助言を賜りました東北大学大学院工学研究科 大町真一郎助教授に心より感謝致します。

常日頃から研究に限らず貴重な御意見、御助言を戴く東北文化学園大学応用情報工学科 孫方博士に心より感謝致します。

また，後藤英昭氏（現在，東北大学情報処理教育センター助手）森大毅氏（現在，宇都宮大学工学部電気電子工学科助手）李信行氏（現在，（株）富士通研究所），佐藤俊治氏（現在，日本学術振興会特別研究員），菅谷至寛氏（現在，東北大学大学院工学研究科助手），李相憲氏（現在，東北大学大学院工学研究科博士後期課程在学中），岩村雅一氏（現在，東北大学大学院工学研究科博士後期課程在学中），加藤毅氏（現在，東北大学大学院工学研究科博士後期課程在学中），をはじめとする阿曾研究室の皆様には日々の研究におきまして，計算機環境を整えて頂き，また，有益な御意見を頂きました．また，研究以外にも多岐に渡ってお世話になりました．深く感謝致します．

最後に，温かなご支援を続けて頂いた家族に感謝致します．