

劣化推定に基づいた辞書の自動選択による本棚画像中の文字認識

Recognition of Characters in Bookshelf Images using Automatic Dictionary Selection based on Estimated Degradation

正会員 澤木美奈子[†], 村瀬洋[†], 萩田紀博[†]
Minako Sawaki[†], Hiroshi Murase[†] and Norihiro Hagita[†]

Abstract We developed an adaptive dictionary-selection method for recognizing low-quality characters in scene images. First, the environmental conditions of the input image are estimated using an initial dictionary. Then, a relevant dictionary from several dictionaries reflecting different environmental conditions is automatically selected based on the estimation and used for recognition. Experiments on characters in images of bookshelves showed that this method achieves a higher recognition rate (89.8%) than that using one fixed dictionary (76.4%). Furthermore, recognition accuracy improved from 89.8% to 95.2% when contextual postprocessing was used.

キーワード：文字認識、情景内文字認識、テンプレート選択、本棚画像

1. ま え が き

室内に存在する様々な文字情報を認識する技術は、環境情報のデータベース化などを実現する上で重要な技術である。デジタルカメラのような簡易な入力手段も一般的になり、従来から認識対象とされてきた紙面上の文字以外にも様々な用途に文字認識を有効に利用することが可能となってきた。本論文では、書棚上の雑誌の背に印刷されている文字の認識について考察する。本や雑誌の名前を認識してコンピュータ上で扱える情報に変換することにより、例えば個人の書棚や図書館の蔵書管理、書籍販売の在庫管理や欠本の検索などを行うことができる。

これに類似した研究としては、情景内の文字認識に関する研究がある。情景内文字認識の多くの手法では、文字の単色性および複雑さに基づいて濃淡画像から文字部分を抽出し、その後に既存の文字認識手法により認識を行う。文字部分の抽出では、部分領域ごとに動的2値化を行う手法^{1)~4)}や、文字の単色性とエッジを用いての切り出し⁵⁾、空間周波数と輝度コントラストに基づく手法⁶⁾、などが報告されている。しかし、これまでの多くの研究は、文字領域を精度良く抽出することに重点がおかれ、その後段にくる文字認識部分についてはあまり考察されてこなかった。

ところが情景内の文字は、既存の文字認識アルゴリズムが対象とするような文書中の文字とはその性質が異なる。

例えば、文字部分と背景部分のコントラストが低い場合や、画像の部分によって文字と背景の条件が異なる場合がある上に、照明状況などの観測条件が大きく変化するために、観測された文字には、文字の変形やよごれやかすれという劣化がしばしば発生する。そのために、単純な類似度を用いた照合や、あるいは比較的劣化に強いといわれる補完類似度⁷⁾による照合を用いても、これらの文字を認識した場合、必ずしも高い認識率が得られない。

これらの大きな変形や劣化を吸収するためには、大きく分けて2つのアプローチがある。ひとつは、変形や劣化に不変な特徴量もしくは識別関数を用いることである。文字の輪郭方向に基づいた特徴量は変形にロバストであることが報告されている⁸⁾。また、参照パターンを変形させて入力パターンと照合する手法の検討もなされている^{9)~13)}。しかし、これらのいずれの手法も、変形にはロバストであるが、よごれやかすれという劣化の影響を受けやすい。もうひとつのアプローチとしては、参照パターンを変形や劣化に合わせて増やす、いわゆるマルチテンプレートによる手法が考えられる。例えば、学習パターンをあらかじめ種々に変形させ参照パターンに登録することにより、認識精度が向上することが報告されている¹⁴⁾。しかし、大きく変形したパターンをやみくもに参照パターンに追加すれば、カテゴリ間の分離が悪くなることが予想される。例えば、「9」が変形してループがつぶれたパターンを参照パターンとした場合、これは「1」や「7」のような別のカテゴリに類似しているため、かえって「9」「7」「1」の認識が困難になる場合も生じる。

本論文では、後者のアプローチを取るが、劣化した参照パターンを劣化の度合いに応じて複数の辞書に配分するこ

1998年8月17日, Proc. of 14th International Conference on Pattern Recognition で発表

1998年11月30日受付, 1999年10月5日最終受付, 2000年4月6日採録
† NTTコミュニケーション 科学基礎研究所

(〒243-0198 厚木市森の甲若宮3-1, TEL.046-240-3569)

† NTT Communication Science Laboratories

(3-1, Morino-ato-Wakamiya, Atsugi-shi, Kanagawa 243-0198, Japan)

とにより、カテゴリー間の分離を保つこととする。そして、あらかじめ用意された劣化の程度に対応した複数の辞書の中から、入力画像の劣化の程度に適した辞書を自動的に選択する。ここでは、局所的な領域では照明などの環境に関する状況は様である、と仮定して議論を進める。具体的には、ある文字列に着目し、その列に含まれる文字に共通する劣化の状態を推定し、しかるべき辞書を特定する。次に、この辞書を用いて認識する。実験により、この方法が、劣化の少ない文字データから作成した単一の辞書で認識するよりも、高い認識精度が得られることを示す。

2. 認識手法

2.1 本棚画像の性質

図1にデジタルカメラを用いて撮影した本棚の画像の例を示す。この例からもわかるように、これらの文字は、本の厚みの薄さ、文字線の欠損、本の傾きなどの様々な原因で文字形状そのものが変化する。さらに、同じ種類の雑誌であっても、背景色の微妙な違い、印刷濃度の違い、照明の違いなど、観測条件の違いによっても変化する。また、その後に行われる前処理の結果として生じる変化もある。「2値化」などを含む前処理は、本来、劣化を生じないように行う、もしくは劣化を除くために行うものであるが、実際には適切な前処理を行うことは困難であり、副作用が頻繁に発生する。本論文では、このような劣化要因により生じたよごれ・かすれを劣化と呼ぶこととする。情景内文字では、これらさまざまな劣化要因により、取り込んだ画像を単純に2値化しただけでは、画像としてかなり劣化したものとなる。

図2は、照明条件の異なった濃淡画像(256階調)に固定閾値(128)による2値化処理を施すことにより発生する文字の劣化の例である。照明条件や撮影条件の異なる情景内画像には、このように濃淡画像では非常に類似しているにもかかわらず、2値化した場合に文字線幅が大きく異なったり、周囲の雑音の影響の差が大きいというような劣化文字は頻繁に発生する。

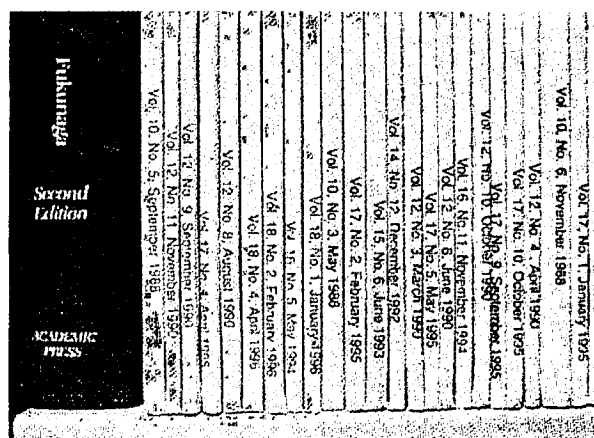


図1 本棚画像の例
Example of a bookshelf image.

このような画像に対して、前処理によって照明条件や撮影条件の影響を吸収する方法も考えられるが、照明条件や撮影条件が複雑に変化する情景内文字画像においては、多様な劣化要因が存在するために前処理では限界がある。

そこで本稿では、劣化した情景内文字を認識するため、まず良好な条件で撮影した文字パターンから作成した初期辞書を用いて認識し、その認識結果からどのような劣化状態かを推定し、その劣化状態に適した辞書に切り替えて再度認識する。これにより認識精度の向上が期待できる。

2.2 認識の流れ

ここで、本手法における本棚画像中の文字認識手法の手順について述べる(図3)。

(1) 前処理部

(a) 2値化 デジタルカメラで入力されたカラー画像を濃淡画像(図4(a))に変換した後に、適当な固定閾値で2値化する(図4(b))。

(b) 傾き抽出 本棚上の本や雑誌は、しばしば傾いて置かれているため、傾き抽出を行う。本の境界部分には、影により黒っぽい線が入る。そのため、ランレンクスに基づいてこの本の境界部分を検出し、その傾き角を求める。

具体的には、まず境界部分の黒画素の連結(ラン)が雑音などにより分断される場合があるため、それをつなぐために2値画像を実空間での垂直方向(図4のy軸方向)に上下2画素ずつ膨張させる。その画像上の各x座標(x軸は実空間での水平方向とする)において、垂直方向に対し $\pm 2^\circ$ の間の 0.2° 刻みで複数回走査し、最大ランレンクスを与える角度をそのx座標における傾き角とする。膨張量および刻み角は、予備実験の結果を踏まえて決定した。また、最大ランレンクスが一定値以上の部分を本の境界線とする。

(c) 文字列抽出と傾き補正 次に、境界線で囲まれた範囲の中の文字列の幅を求める。文字列部分は、背景部分に比べエッジ部分が多いと仮定し、白黒反転回数により文字列部分を特定する。

具体的には、各x座標において、求められた傾き角の方向に走査し、画素値の白黒の反転回数の射影分布をとる。この回数が、一定値より多い範囲を文字列領域とする。求められた文字列領域の幅の中心のx座標における傾き角をその文字列領域の傾き角とし、その角度に応じて

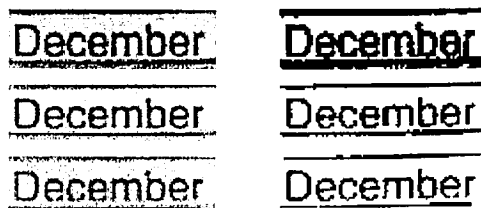


図2 照明条件の違いにより発生する文字劣化
(左)入力濃淡画像(256階調)、(右)2値画像
Character degradation caused by differences in lighting condition. (left) gray-scale image (256 levels), (right) binarized image.

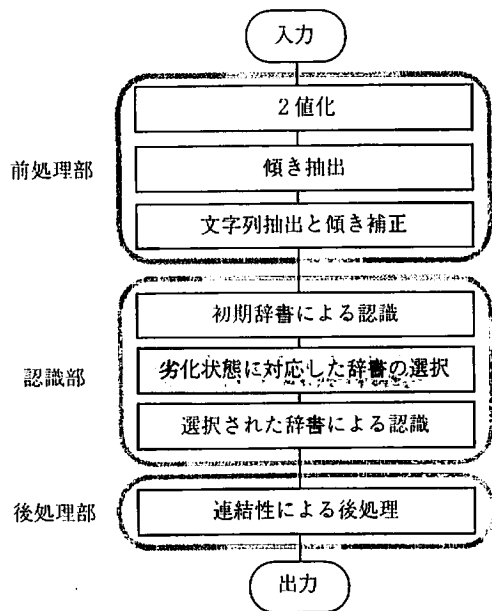


図3 認識の流れ
Flow of recognition procedure.

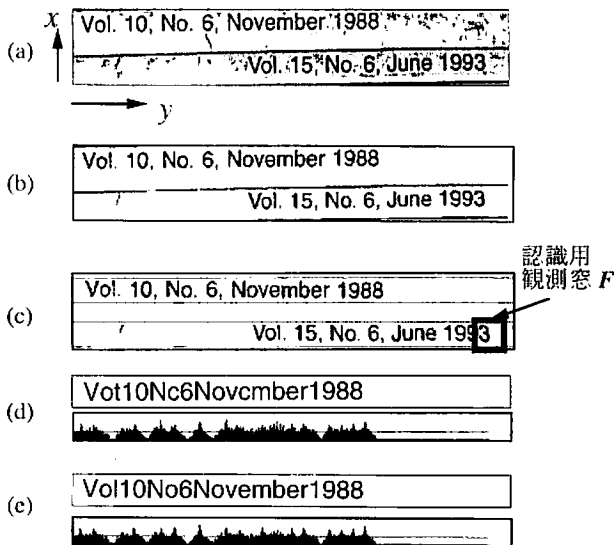


図4 認識過程
(a)入力濃淡画像、(b)2値画像、(c)傾き補正後の画像、(d)初期辞書による認識結果、(e)選択された辞書による認識結果
(d)、(e)について、上段：認識カテゴリー、下段：各ずらし位置での最大類似度分布、下段の直線：類似度閾値を表す。

Recognition process.

(a) input gray-scale image, (b) binarized image, (c) skew corrected image, (d) recognition results by using initial dictionary, (e) recognition results by using selected dictionary.

For (d) and (e) (upper) recognized categories, (lower) maximum similarity value at each position. The straight line indicates threshold values for recognition.

文字列画像を回転し、傾き補正を行う。また、大きさの正規化も同時に行う。図4(c)には、この処理で抽出された文字列を矩形領域として示してある。

(2) 認識部

(a) 初期辞書による認識 傾き補正後の文字列画像中の文字は、ずらし照合により認識する。各文字列上を認識用観測窓 F を垂直方向に走査し、各認識用観測窓と初期辞書内の参照パターンとの照合を行う。ここで、「初期辞書」とは、良好な条件で撮影された劣化の少ない文字パターンから作成した参照パターンからなる辞書を表す。また、照合には補完類似度⁷⁾を用い、位置ずれを考慮して垂直方向のみでなく水平方向への走査も行う。照合した各位置での最大類似度が、あらかじめ学習してある各文字カテゴリーの類似度閾値を越えた場合、そのカテゴリーを初期辞書による認識結果とする。認識結果同士が重畳している場合には、類似度の小さい方を認識結果から削除する。ここで用いるずらし照合は、文献7)に述べているものと同様である。図4(d)の上段に認識カテゴリー、下段に最大類似度分布を示す。

(b) 劣化状態に対応した辞書の選択 初期辞書による認識結果から入力画像の劣化状態(劣化指数)を推定する。その後、劣化指数に適合する辞書を選択する。辞書の作成方法および選択方法については3章で詳しく述べる。

(c) 選択された辞書による認識 選択された辞書を用いて、ずらし照合を再度文字列画像に適用して文字を認識する。図4(e)の上段に認識カテゴリー、下段に最大類似度分布を示す。

(3) 後処理部

ある限られた対象分野での文字認識においては文脈に基づく後処理が有効である。ここでは、連続する2文字のカテゴリーの連結性に着目した後処理を行う。具体的には、あらかじめ連結する可能性のある2カテゴリー(例えば、「e」の後のカテゴリーは{'b', 'c', 'm', 'p', 'r'}のいずれかである、等)をテーブルに登録しておく。認識結果を文頭からテーブルと比較し、2カテゴリーの連結関係で適合しないものが存在した場合、後方のカテゴリーを認識結果から削除し、近隣の認識候補の中から適合するものを探す。この結果得られた認識結果を出力する。

3. 参照パターンの作成と選択

3.1 辞書選択の基本方針

本手法は、ずらし照合に基づいて文字列を認識する方法を採る。その際に、劣化した入力パターンを正確に認識するために、あらかじめ様々な劣化に対応した辞書を複数用意しておき、劣化に対応した辞書を自動的に切り替えることにより認識する。その概念図を図5に示す。

印刷文字の劣化の要因としては、解像度/ぼけ/閾値処理、など様々な要因が知られているが¹⁵⁾、これら複数の要因を個別に分離して推定することは困難である。そのため本論文では、様々な要因からなるパターンの劣化を「劣化指数」という1次元の軸にあてはめて推定する。ここでは、入力パターンの劣化指数 I として、入力パターンと初期辞書中の最近傍参照パターンの黒画素数の比を用いる

(図6) 経験的に、今回対象としたような文字画像ではよごれとかすれはあまり同時に生じしない。通常、よごれが大きい場合には黒画素数が増加し、かすれが大きい場合には黒画素数が減少する。その結果、黒画素比により劣化の程度を近似的に表現しているとみなすことができる。ここで、劣化指数が1の場合、入力パターンは最も参照パターンに類似していることを意味し、小さい場合はかすれ、大きい場合はよごれに対応している。

学習段階では、事前学習用に用意されたパターンを用いて劣化指数に応じた辞書を作成する。

認識段階では、入力パターンを初期辞書で仮りに認識し、その認識結果をもとに劣化指数を推定する。この劣化指数に応じて適切な辞書を選択し、再度認識を行う。

ここで、いかに劣化指数を推定し、それに基づいていかに適切な辞書を選択するかが重要となる。本手法では、様々な劣化条件とみなせる範囲の複数文字の認識結果からそれぞれ得られる劣化指数を平均化することにより、初期辞書による誤認識によって生じる劣化指数の推定誤差を軽減する。本論文で取り上げる本欄画像においては、1冊の背の領域では、かすれ・よごれの劣化に関する状況は一樣とみなせると仮定し、1冊分の文字列中のそれぞれの文字から推定された劣化指数の平均をとり、文字列の劣化指数とした。

3.2 学習

まず、事前学習用に用意されたパターンを、様々な劣化

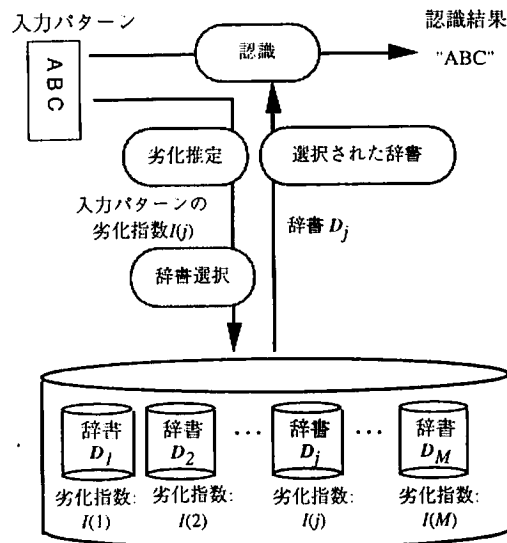


図5 劣化推定に基づく辞書の選択

Selection of dictionary based on degradation estimation.

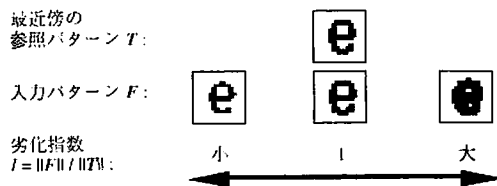


図6 劣化指数
Degradation index.

指数に対応するように劣化させる。実際の画像では、劣化は様々な要因で発生するが、ここでは擬似的に2値化閾値を変化させることにより、劣化指数を変化させる。実験的に、黒画素数比と2値化閾値はほぼ線形の関係をもつため、2値化閾値を変化させることにより、様々な劣化指数のパターンを生成することができる。

学習には、劣化させて作成したパターンから2種類の文字パターンセットを用意する。実際に辞書に用いるパターンセットを予備辞書、劣化指数とそれに対する適切な辞書との関係を求めるために学習段階で使用されるパターンセットを、辞書学習用パターンセットと呼ぶこととする。

具体的には、濃淡画像として入力した事前学習用の文字パターンを複数の閾値で2値化する。このようにして得られたパターンのうち、各劣化指数において、一部を予備辞書($P_i: i=1 \sim L$)とし、残りのパターンを辞書学習用パターンセット($T_j: j=1 \sim M$)とする。また、 $I(P_i), I(T_j)$ は、それぞれ P_i, T_j の劣化指数とする。ここで、閾値130で2値化した辞書を、比較的良好な参照パターンとみなし、初期辞書とする。初期辞書の劣化指数を基準値1.0とする。次に、劣化指数に対応した辞書を作成する。各劣化指数の辞書学習用パターンセット T_j を、すべての予備辞書で認識を行い、最大の認識率を与える予備辞書 P_{max} を選択する。次に、その辞書を辞書学習用パターンセットを100%認識できるようになるまで誤り訂正学習を行う。ここで誤り訂正学習とは以下の処理をいう。 T_j を P_{max} で認識し、誤読パターンを P_{max} に追加する。追加された辞書を用いて、再度 T_j を認識する。これを T_j が100%認識されるまで繰り返す。こうして作成された辞書を、辞書学習用パターンセットの劣化指数 $I(T_j)$ に対応した辞書 D_j とする。

なお、ここで入力パターンの劣化指数に適した辞書は、必ずしも同じ劣化指数で作成された参照パターンからなる辞書でない場合がある。この性質は使用する類似度の性質に依存し、本論文で用いる補完類似度では以下のような傾向をもつ。入力パターン F と参照パターン T としたとき、補完類似度は、 F と T が完全に一致した場合および F が T もしくは T が F に完全に包含された場合に精度良く認識できる。 F と T が完全に一致することはほぼ起こらないため、ある入力パターンの劣化指数に適した辞書としては、その劣化指数よりも小さい、もしくは大きい劣化指数の参照パターンからなる辞書となる傾向がある。

3.3 認識

入力パターンを初期辞書を用いて仮認識し、最近傍となる参照パターンを求める。その参照パターンと入力パターンの黒画素数の比を求めることにより、劣化指数を求める。入力パターン中に複数文字が記入されている場合には、それぞれの文字ごとに推定した劣化指数の平均値を入力パターン全体の劣化指数とする。これにより、初期辞書における誤認識による劣化指数の推定誤りの影響を少なく

することが期待できる。

実際に黒画素数を求める際には、ずらし照合を行うと複数の文字が認識用観測窓の中に入ってくるため、入力パターンの黒画素数が大きくなる。しかし、それは着目しているパターンの劣化とは無関係なものである。そこで、認識用観測窓中の画素のうち、最近傍参照パターンの外接矩形の範囲の黒画素数を求め、それと最近傍参照パターンの黒画素数との比を求める。

次に、求められた劣化指数に一番近い劣化指数 $I(T_j)$ を求め、対応する辞書 D_j を選択する。選択された辞書 D_j を用いて認識を行い、これを認識結果とする。

4. 実験

4.1 実験条件

本棚をデジタルカメラ(解像度 832×608 画素)で撮影したカラー画像を濃淡画像(画素値 0~255)に変換した画像を実験に用いた。本棚は、複数年にわたる技術誌をランダムに入れ替えながら、ほぼ固定の照明条件で撮影した。

カテゴリー数は、数字 0~9、英語の月の名前に含まれる文字(A, D, F, J, M-O, S, a-c, e, g-i, l-p, r-v, y)および V の計 37 カテゴリーを対象とした。予備辞書は 1 パターン/カテゴリー、辞書学習用パターンセットは 4 パターン/カテゴリーとした。テストパターンは、本棚画像 12 枚中に含まれる文字、合計 441 文字を用いた。

予備辞書および辞書学習用パターンセットには、手動で切り出した事前学習用の文字パターンを、2 値化閾値 80 から 150 での 10 きざみで 2 値化したものを用いた。辞書学習用パターンセットの例を図 7 にあげる。作成された辞書に順番に、辞書番号 1 から 8 までを付与した (D_j ; $j=1 \sim 8$)。これ以外の 2 値化閾値の範囲は、パターンの汚れやかすれが激しいため、用いないこととした。認識用観測窓のサイズは 24×24 画素とした。入力画像の 2 値化閾値は、経験的に 128 とした。

4.2 本棚画像の認識結果

本手法(後処理有/無)、単一辞書法(初期辞書、後処理無)、マルチテンプレート法(すべての予備辞書および辞書学習用パターンセットを合わせた辞書、後処理無)による認識結果を表 1 に挙げる。本手法を用いて辞書を切替えることにより、本棚画像中の文字認識では、初期辞書を用いた単一辞書法に比べ約 13% の認識率向上が得られた。また、事前学習用に用意されたパターンをすべて辞書に登録したマルチテンプレート法に比べた場合、本手法の方が認識に使用する参照パターン数が少ないにもかかわらず、高い認識率を示した。このことにより、やみくもに参照パターン数を増やすよりは、適切な参照パターンを選択することが有効に働く場合があることが示された。

さらに、本手法に後処理を追加したところ、95.2% を正読することができた。後処理を行った後の認識誤りの原因は、「8 が 3 に誤読される」というような数字の誤認識で

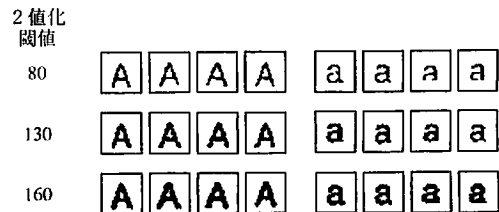


図 7 辞書学習パターンの例 (カテゴリー: 'A' と 'a')
Examples of dictionary learning patterns (category: 'A' and 'a').

表 1 本棚画像中の文字の認識率
Recognition rates of characters in bookshelf images.

辞書選択法(本手法, 後処理有)	95.2%
辞書選択法(本手法, 後処理無)	89.8%
マルチテンプレート法(比較用辞書, 後処理無)	85.5%
単一辞書法(初期辞書, 後処理無)	76.4%

ある。参考までに、市販 OCR(本実験で用いた辞書学習用パターンセットによる学習は行っていない)で認識した場合の認識率は 10.7% であった。

5. 考察

本手法の有効性を確かめるため、図 4 の画像に対し、適切な辞書が選択されたかを確認した。ここで、図 4 の上の行を文字列例 1、下の行を文字列例 2 と呼ぶ。図 4 の画像を、用意された 8 つの辞書および初期辞書でそれぞれ認識した。その場合の誤読パターン数を表 2 にあげる。辞書を変更することにより、誤読パターン数は 0 から 3 まで変化した。劣化指数から選択された辞書番号は、文字列例 1 では 5、文字列例 2 では 6 であり、双方とも誤読パターン数は 0 であった。このように、本手法により、用意された辞書の中でエラーが少なくなるように辞書が選択されていることが確認された。

上記の点をパターン例を挙げて具体的に示す(図 8)。入力パターン(カテゴリー l)に対し、初期辞書による認識結果はカテゴリー l であった。この場合、入力パターンの変形が大きいため入力パターンは l に誤読された。選択された辞書には、入力パターンの変形に適した l の参照パターンが含まれており、入力パターンが l の参照パターンにほぼ含まれたため高い類似度値をとり、正読となった。

また、本手法の他の要因による劣化耐性を検討するために、擬似的に劣化させた入力画像の認識実験を行った。具体的には、入力画像に以下のフィルタを用いてぼかし処理を行った。

$$\begin{matrix} 0.25 & 0.5 & 0.25 \\ 0.5 & 1.0 & 0.5 \\ 0.25 & 0.5 & 0.25 \end{matrix}$$

この画像に対し、本手法を適用したところ、認識率は単一辞書(初期辞書)による認識率が 79.1% であるのに対し、辞書選択を行うことにより 85.0% と向上した。劣化(ぼか

表 2 辞書ごとの誤読パターン数(後処理前)
(文字列例 1: 図 4 の画像の上の行, 文字列例 2: 図 4 の画像の下の行)

Error pattern number for different dictionaries (results before postprocessing).
(Example 1: upper line of the image in Fig.4, Example 2: lower line of the image in Fig.4)

辞書番号	1	2	3	4	5	6	7	8	初期辞書
文字列例 1	2	1	1	0	0	1	2	0	3
文字列例 2	2	1	1	0	0	0	1	2	2

: 選択された辞書

入力パターン (r)	初期辞書	選択された辞書
1		

: 認識結果

図 8 辞書選択により正読になった例

Examples of correctly recognized pattern by the proposed method.

し処理)のない場合の認識率(89.8%)より若干の認識率の低下がみられたものの、この結果は、ぼけなどの劣化に対する本手法の適応可能性を示している。

情景画像中の文字を 2 値化した際に発生するよごれやかすれを低減させる手法として、適応的 2 値化を行う手法が考えられる。ここでは、入力画像に対し大津の適応的 2 値化¹⁰⁾を行った後に、単一の辞書を用いて認識する実験を行った。具体的には、入力画像中の本の境界によってはさまれた 1 冊の画像領域ごとに 2 値化閾値を決定した。その後単一辞書(初期辞書)で認識を行ったところ、認識率は、固定閾値の 76.4%より低い 69.6%であった。入力画像の性質に依存したより高度な適応的閾値設定を行うことにより性能は向上すると思われるが、最適な閾値の決定は困難である。

6. むすび

本論文では、低品質文字の認識を目的として、入力パターンに応じて辞書を選ぶことにより、認識率を向上させることができることを示した。

参照パターンによる仮認識の結果を用いて、辞書を選択することにより、きれいな参照パターンからなる辞書では、76.4%の認識率である本棚画像をデジタルカメラで撮った情景内の文字データに対し、89.8%の認識率を得ることができた。さらに、本の文字列に関する後処理を加えることにより、95.2%に認識率が向上した。

今後の課題としては、多様な要因による劣化への本手法

の耐性の評価、多種の劣化手法により生成された参照パターンの利用、マルチフォントへの拡張、認識対象の性質を積極的に利用した後処理の改良(言語処理など)が挙げられる。

口頭ご指導頂く石井健一郎 CS 研企画担当部長に感謝します。また、傾き補正プログラムは、当研究所夏期実習生 Mr. Benoit Maire の協力によるものである。

(文 献)

- 1) J. Ohya, A. Shio, and S. Akamatsu: "Recognizing Character in Scene Image", IEEE Trans. Pattern Anal. & Mach. Intell., 16, 2, pp. 214-220 (1994)
- 2) 倉掛, 桑野, 新井, 小高: "認識技術を用いた映像中キーターゲットインデクシングの検討", 信学技報, PRU 95-237, pp. 15-20 (1996)
- 3) 安居院, 大竹, 中嶋: "書籍の背文字抽出のための領域分割", 信学全大, D-539, p. 7-251 (1991)
- 4) 岩元, 伊藤: "書棚画像からの書籍タイトル文字領域の抽出", 信学全大, D-343, p. 346 (1996)
- 5) 茂木, 有木: "ニュース映像中の文字認識に基づく記事の索引付け", 信学技報, PRU 95-240, pp. 33-40 (1996)
- 6) 劉, 山村, 大西, 杉江: "シーン内の文字列領域の抽出", 信学技報, PRU 95-222, pp. 47-53 (1996)
- 7) M. Sawaki and N. Hagita: "Text-line Extraction and Character Recognition of Document Headlines with Graphical Designs using Complementary Similarity Measure", IEEE Trans. Pattern Anal. & Mach. Intell., 20, 10, pp. 1103-1109 (1998)
- 8) T. Tsutsumida, T. Masui, T. Noumi and T. Wakahara: "Results of IPTP Character Recognition Contest and Studies on Multi-expert System for Handprinted Numeral Recognition", IEICE Trans. Inf. & Syst., E 79-D, 5, pp. 429-435 (1996)
- 9) 水上, 古賀, 島岡: "変位抽出を行う手書き文字認識システム", 信学論, J 80-D-11, 1, pp. 63-72 (1997)
- 10) A. K. Jain, Y. Zhong and S. Lakshmanan: "Object Matching Using Deformable Templates", IEEE Trans. Pattern Anal. & Mach. Intell., 18, 3, pp. 267-28 (1996)
- 11) E. Leven and R. Pieracchini: "Dynamic Planar Warping for Optical Character Recognition", Proc. of ICASSP, pp. 111-149-152 (1992)
- 12) C. W. Wu, R. M. Owens and M. J. Irwin: "Distortion Processing in Image Matching Problems", Proc. of ICASSP, pp. 2181-2184 (1990)
- 13) Roger K. Moore: "A Dynamic Algorithm for the Sistance Between Two Finite Areas", IEEE Trans. Pattern Anal. & Mach. Intell., PAMI-1, 1, pp. 86-88 (1979)
- 14) 村瀬, 木村, 吉村, 三宅: "パターン整合法における特性核の改良とその手書き平仮名文字認識への応用", 信学論, J 64-D, 3, pp. 276-283 (1981)
- 15) H. S. Baird: "Document Image Defect Models", in "Structures Document Image Analysis", ed. H. S. Baird, H. Bunke and K. Yamamoto, Springer-Verlag
- 16) N. Otsu: "A Threshold Selection Method from Gray-Level Histograms", IEEE Trans. Systems, Man and Cybernetics, SMC-9, 1, pp. 62-66 (1979)



さわき みさこ
澤木美奈子 1989年、慶應義塾大学理工学部電気工学科卒業。同年、日本電信電話(株)に入社。文字認識の研究に従事。正会員。



むらせ ひろし
村瀬 洋 1980年、名古屋大学工学部電子工学科修士課程修了。同年、日本電信電話公社(現 NTT)入社。以来、文字認識・図形認識・コンピュータビジョン・マルチメディア情報認識の研究に従事。現在、NTTコミュニケーション科学基礎研究所メディア情報研究部グループリーダー、工学博士。



あきた のりひろ
萩田 紀博 1978年、慶應義塾大学工学部電気工学科修士課程修了。同年、日本電信電話公社(現 NTT)入社。以来、文字・文書・画像認識の研究に従事。現在、NTTコミュニケーション科学基礎研究所メディア情報研究部長、工学博士。