

郵便・物流自動化システム用 住所認識技術の性能向上

Improvement of Address Recognition Technologies for Letter and Parcel Sorting Systems

浜村 倫行

■HAMAMURA Tomoyuki

郵便・物流自動化システムでは、郵便物などを区分するのに必要な情報をそれらから読み取る文字認識技術が不可欠である。東芝は1960年代にこの自動化システム事業を始めて以来、文字認識技術を絶えず向上させ、現在では郵便番号に加え、住所や番地まで高精度に認識できるようにしている。枠の中ではなく自由に記載される住所は、一つ一つの文字の境目が曖昧であることから多数の読み方ができるため、住所データを利用した候補の絞り込みが重要な課題となる。

そこで当社は、単語の認識と住所データの探索を、統計的に最適とされる事後確率に基づいて行う手法を開発した。単語と住所の認識実験により、従来の認識・探索手法に比べ大幅な性能向上を確認した。

In logistics and postal automation systems, character recognition technology to acquire information for sorting destination addresses from scanned images is essential for letter and parcel sorting systems.

Toshiba has been engaged in the research and development of core technologies in this field since the 1960s. The accuracy of recognition of full address data from scanned images of postal items and parcels, including the street, floor or house number, and postal code, has been enhanced through these efforts. However, in the recognition of freely handwritten addresses, where the boundaries of each character may be unclear, a candidate reduction technique to narrow down the candidate addresses from among a large number of readings of characters in conjunction with an address database plays a major role in the recognition algorithm.

In response to this situation, we have now developed handwritten word recognition and address database searching algorithms for letter and parcel sorting systems based on an *a posteriori* probability that is assumed to be statistically optimal. We have conducted evaluation experiments using the newly developed word and address recognition technologies and confirmed that these technologies can recognize handwritten addresses with higher accuracy than conventional methods.

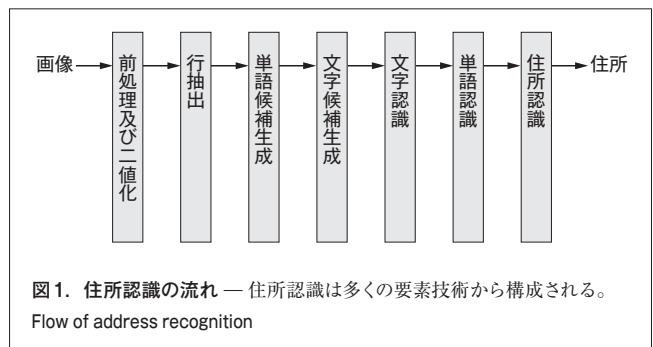
1 まえがき

東芝は、郵便・物流自動化システム事業を1965年に始めて以来、そのコア技術である文字認識技術を常に進化させ続けてきた。当初の認識対象は郵便番号枠の中に記載された数字だけであったが、自由な位置にある活字の数字や手書きの数字、両者の混在へと進化し、更に現在では住所や番地も認識できるようになり、それにより郵便物などを配達する順番に並べるまで自動化されている。

ここで適用されているのは、当社が開発し強みとなっている事後確率に基づいた住所認識技術である。ここでは、住所認識全体の流れと、当社独自技術である事後確率に基づく手書き単語認識及び手書き住所認識について述べる。

2 住所認識の流れ

住所認識は多くの要素技術の複合体である。まず、住所認識全体の流れと各要素技術の概要を述べる。ここでは欧米言語を対象として説明をするが、多くの要素技術は日本語住所認識でも共通である。



住所認識の流れを図1に示す。郵便書状や小包の画像が入力されると、ノイズ除去や、郵便物位置切出し、傾き補正などの各種前処理を施した後、文字部が1(黒画素)、背景が0(白画素)となるよう各ピクセルの輝度を二値に変換する二値化処理を行う。次に、同一行に属する黒画素をグループ化する行抽出処理を行う。郵便書状から行を抽出した結果の例を図2に示す。点線の長方形が抽出された行候補である。

続いて、行を単語に分割する単語候補生成処理を行う。単語間が十分に離れていない場合は、分割の仕方を1通りに定めることができないため、複数の分割候補が生成される。図2の



図2. 行抽出と単語候補生成のようす — 一点線が行で、実線が単語候補である。単語位置は一意に決められないため、複数の候補を生成する。
Example of line extraction and word hypothesis generation

実線の長方形が単語候補の例である。日本語の場合は、単語間が離れていないためこの処理は行わない。

更に、各単語候補（日本語の場合は各行）を文字に分割する。詳細は後述するが、単語分割と同様に文字分割も複数の候補が考えられ、複数の文字候補が生成される。

生成された各文字候補に対し、文字認識を行う。文字認識は、特徴抽出処理と認識処理に分けられる。特徴抽出処理では、文字候補画像から“特徴量”を抽出する。特徴量は、 d 個の数値の集まりで、 d 次元の特徴ベクトルとして表現される。画像を複数の領域に分け、領域内の輝度の平均値を用いたり、輝度勾配方向（輪郭方向）を数値化したりするなど、様々な特徴量が考えられ、性能の高い特徴量を採用している。図3は、 4×4 の領域に分けている例である。

こうして得られた特徴ベクトルを基に、認識処理を行う。認識処理にはSVM (Support Vector Machine) や、ニューラルネットワーク、部分空間法、疑似ベイズなど様々な一般パターン認識手法を用いることができ⁽¹⁾、性能の高い手法を採用している。文字認識結果は、各アルファベットにスコアが付与された形式で表すことができる。図3では文字認識結果をスコアの順にソートして表示している。

各文字候補の文字認識結果を用いて、住所データベースを利用しながら単語認識及び住所認識を行う。当社は、この単

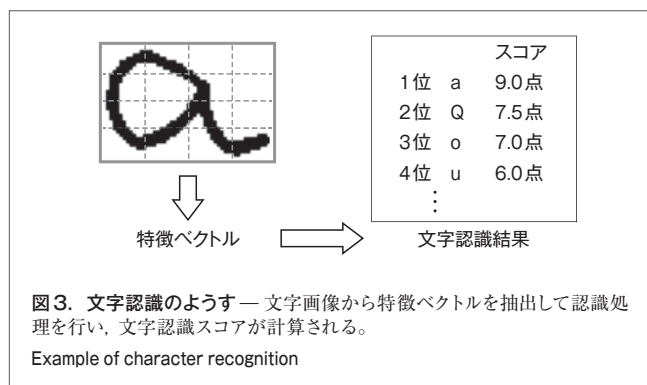


図3. 文字認識のようす — 文字画像から特徴ベクトルを抽出して認識処理を行い、文字認識スコアが計算される。
Example of character recognition

語と住所の認識を事後確率に基づく枠組みで行っており、高い認識精度を達成している。以下に、この二つの技術について詳細に述べる。

3 手書き単語の認識

3.1 手書き単語認識の難しさ

図4(a)は、フランスの住所に存在する単語である。“adiene”, “aclieru” など、様々な読み方ができるが、正解は“adresse”である。一つ一つの文字が認識できても、文字の境目がわからないと単語としては認識できないことがわかる。日本語でも、“明”と“日月”のように境目の違いから複数の読み方ができる例は多い。

しかし、おそらくフランス人はこの単語を住所として正しく読むことができる。それは“adiene”, “aclieru” という単語は住所に存在しないが“adresse”という単語は存在することを知っているためである。このように、単語を認識するためには単語の知識が重要であり、一般に“単語認識”とは単語辞書の中から正解単語を選び出す処理のことを指す。

3.2 単語認識の枠組み

文字認識や単語辞書を利用すれば境目を決めることができるが、境目を決めないと文字認識が実行できないため、両者は鶏と卵の関係にある。この問題を解決するため、図4(b)に示すようにあらかじめ多数の分割候補を生成することが一般に行われている。境目の候補は、輪郭形状の変化や黒画素の密度変化などを利用して決める。図4(b)の上二つの各文字候補画像の下には、簡単のためスコア1位の文字だけを文字認識結果の例として示している。

これらの分割候補とその文字認識結果を、住所データベースから生成される単語辞書と照合し、全組合せの中でもっとも照合スコアの高い組合せを単語認識結果とするのが、一般

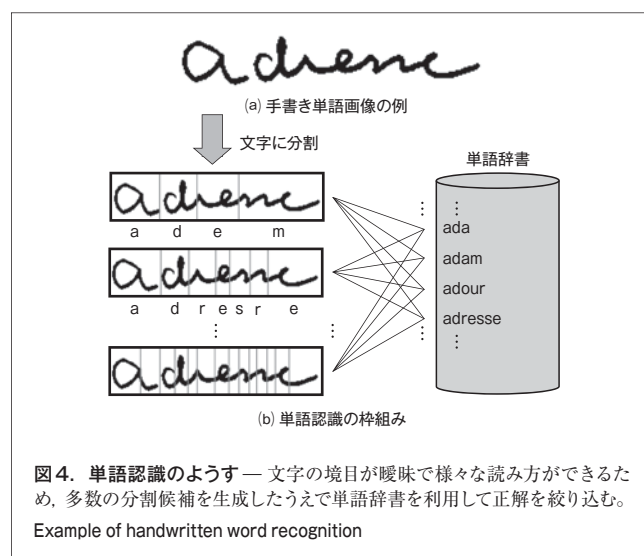


図4. 単語認識のようす — 文字の境目が曖昧で様々な読み方ができるため、多数の分割候補を生成したうえで単語辞書を利用して正解を絞り込む。
Example of handwritten word recognition

的な単語認識の枠組みである。

3.3 事後確率比法

照合スコアとして一般に用いられているのは、文字認識スコアの平均である。例えば図4(b)において、一番上の文字分割候補と単語辞書内の“adam”とを照合する場合、そのスコアは各文字候補画像の“a”, “d”, “a”, “m”の文字認識スコアを平均したものとなる。ここで、スコアとして事後確率を用いることができれば、統計的に最適であることが証明されている⁽¹⁾。

事後確率は一種の条件付き確率で、生成された各文字候補の文字認識結果全て (X) を条件としたとき、単語辞書内の第 i 単語 (w_i) が正解である確率 $P(w_i|X)$ を表す。図4(b)の例では、左側の全情報が X 、右側の“adam”などが w_i に相当する。

従来は $P(w_i|X)$ を計算することが難しかったが、当社はこの値を近似的に計算する事後確率比法を開発した。 $P(w_i|X)$ の近似式変形により、式(1)が導出される。

$$P(w_i|X) \approx P(w_i) \prod_j \frac{P(c_{ij}|x_j)}{P(c_{ij})} \quad (1)$$

ここで $P(w_i)$ は w_i の事前確率、 $P(c_{ij})$ は w_i の第 j 文字の事前確率、 $P(c_{ij}|x_j)$ は w_i の第 j 文字の事後確率である。式(1)の値が最大となる w_i を認識結果とする。

また、式(1)を基に、更に近似精度の高い式も導出している^{(2), (3)}。それらの認識性能比較実験の結果を図5に示す。従来の照合スコアに比べ大幅に性能向上していることがわかる。

4 手書き住所の認識

4.1 概要

郵便や物流の区分システムで最終的に必要な情報は、一つ

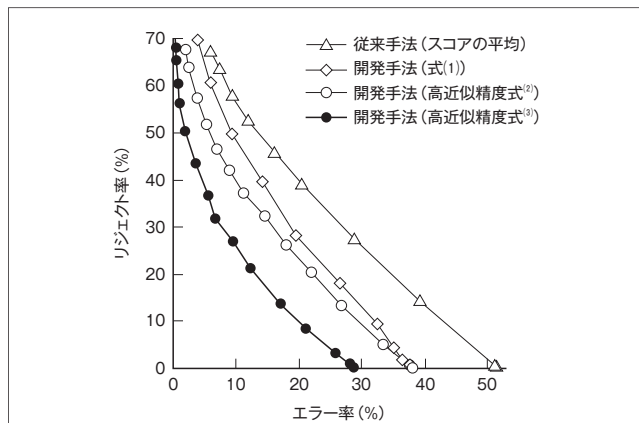


図5. 単語認識性能の比較 — 開発手法はリジェクト率及びエラー率が大幅に低減している。グラフの左下ほど高い認識性能を示している。
Comparison of performance of handwritten word recognition methods

一つの文字や単語の認識結果ではなく、住所である。住所は city, street, street No. のように階層になっているため、図6(a)に示すように木構造で表すことができる。すなわち、住所認識はこの木構造の末端ノード、図6(a)の例では street No. に属する八つのノードの中から正解を選び出す処理であり、木構造の探索問題とみなすことができる。

選び出すにあたり、全てのノードを対象に単語認識処理を行うのが理想的であるが、住所データベースの末端ノード数は数百万～数千万個あるため、膨大な処理時間が掛かる。そこで、制限時間内でいかに選び出すかが重要となる。一般には、上位階層から順に認識し、スコアの低いノードを枝刈りして絞り込むことで処理時間を削減する。

図6(a)で説明すると、まず最上位階層である city の認識を行い、例えば CHIBA であると確定する (絞り込む)。次に CHIBA

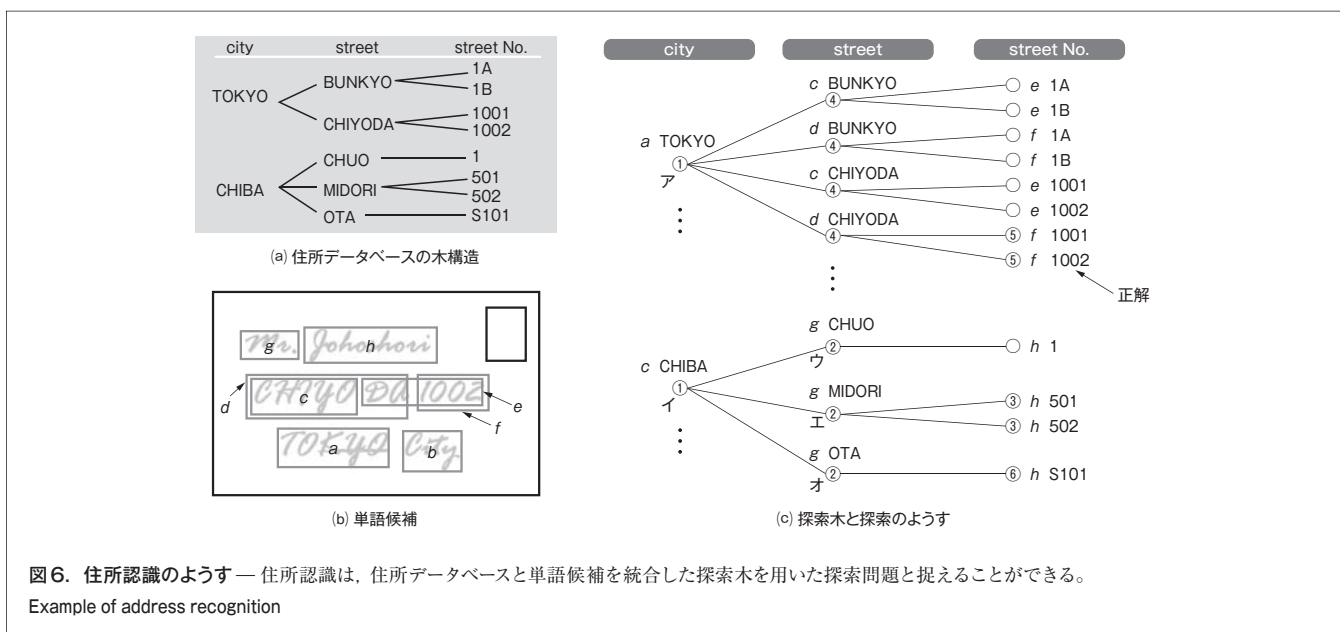


図6. 住所認識のようす — 住所認識は、住所データベースと単語候補を統合した探索木を用いた探索問題と捉えることができる。
Example of address recognition

の下に属するCHUO, MIDORI, OTAだけを単語辞書としてstreetの認識を行い、例えばMIDORIとOTAの2候補に絞り込む。更にその下に属するstreet No.を認識し…という順序になる。

4.2 実際の探索

実際の探索対象は住所データベースではなく、図6(c)に示すような単語候補と住所データベースを統合したものとなる。これは、住所データベースの各ノードに対応する単語候補は複数考えられるためである。日本語の場合は単語候補を用いることができないが、同様の探索木で表すことができる。

前述したような上位階層から順に絞り込んでいく方法は“ビーム探索”と呼ばれる。一般に、ビーム探索では、絞り込み度合いの調節が必要である。絞り込み過ぎると正解を枝刈りしてしまう可能性が高まり、逆に絞り込みが弱すぎると制限時間内に末端ノードに到達できなくなる。制限時間の中でできるだけ絞り込みを弱くする調節が必要である。

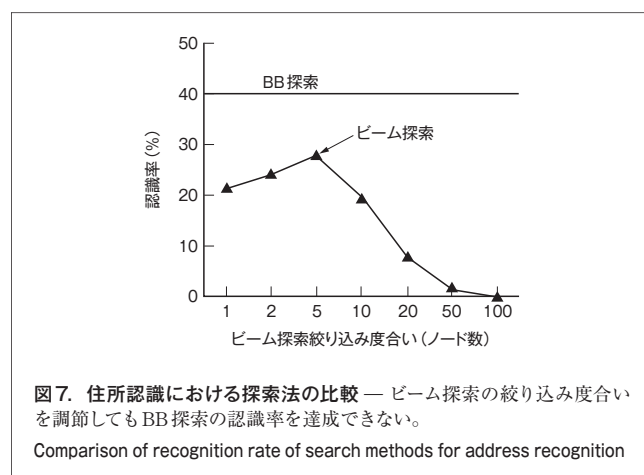
しかし、文字認識などの探索前の処理に掛かる時間が郵便物ごとにばらついている場合、探索に残される時間が変動するため、絞り込み度合いをあらかじめ調節しておくことができない。

そこで当社は、調節が不要な“最良優先探索”と呼ばれる探索法を採用している。最良優先探索とは、常にもっともスコアの高いノードを探索していく方法であり、ビーム探索とは異なりスコアの低いノードの枝刈りを行わない。

図6(c)に探索順序の例を示す。ノードの丸数字は制限時間内に処理できた探索順序である。まずcity階層(ア, イ)が処理される。次に、それらの中でイがスコア最大であったとすると、その子ノード(ウ, エ, オ)が処理される。ここまではビーム探索と変わらない順序である。ビーム探索では次にウ, エ, オの三つを比較するのに対し、最良優先探索では下位階層未探索のノード全て、つまりア, ウ, エ, オの四つを比較する。そして四つの中でエがスコア最大であったとすると、その子ノードが処理される。ビーム探索の場合は末端ノードに到達したら処理終了であるが、最良優先探索では制限時間内であれば次にスコアの高いノードの探索を進める。ビーム探索ではアを枝刈りした時点で正解到達が不可能であったところ、最良優先探索では制限時間内にアの子ノードに対しても処理が進み、⑤で正解に到達できている。このように、最良優先探索では制限時間の長さに応じ探索量が自動的に変わるため、ビーム探索のような絞り込み度合いの調節が不要である。

4.3 BB探索

最良優先探索を適用する際の課題は、スコアの計算方法である。単語認識結果として得られるスコアは、そのままでは用いることができない。なぜなら、最良優先探索では階層の異なるノード、例えばアとエを比較しなければならないからである。アを評価するには“a TOKYO”のスコア(例えば50点)だけ参照すればよいが、エを評価するには“g MIDORI”のスコア



コア(37点)だけでなくその親ノードである“c CHIBA”のスコア(83点)も考慮しなければならない。つまり、(50点)対(83点& 37点)という比較が必要となる。単純に83点と37点の平均(60点)を用いることも考えられる。

当社は、より良いスコアとして事後確率を近似的に計算し利用するベイズ最良優先探索(BB探索: Bayesian Best-First Search)を開発している⁽⁴⁾。図7に示すように、BB探索はビーム探索に比べ高い住所認識率を実現している。

5 あとがき

ここでは、郵便・物流自動化のコア技術である住所認識技術と、当社独自技術である事後確率に基づいた枠組みが認識技術の高精度化を支えていることを述べた。更なる精度向上を図り事業拡大に貢献していきたい。

文献

- (1) ビショップ, C.M. パターン認識と機械学習. 上, 下, 丸善出版, 2012, 349p., 433p.
- (2) 浜村倫行 他. 解析的単語認識における事後確率を用いた評価関数. 電子情報通信学会論文誌(D). **91**, 9, 2008, p.2325 - 2333.
- (3) Hamamura, T. et al. "An analytic word recognition algorithm using a posteriori probability." Proceedings of the 9th International Conference on Document Analysis and Recognition (ICDAR 2007). Curitiba, Brazil, 2007-09, International Association for Pattern Recognition (IAPR), 2007, p.669 - 673.
- (4) Hamamura, T. et al. "Bayesian best-first search for pattern recognition -application to address recognition-." Proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR 2009). Barcelona, Spain, 2009-07, IAPR, 2009, p.461 - 465.



浜村 倫行 HAMAMURA Tomoyuki, D.Eng.
電力システム社 電力・社会システム技術開発センター 自動化・画像応用システム開発部主務, 博士(情報理工学)。文字認識技術の研究・開発に従事。電子情報通信学会会員。
Power and Industrial Systems Research and Development Center