

論文の内容の要旨

階層内・階層間の確率論的統合による

自由手書き住所認識の研究

氏名 浜村 倫行

本論文では、自由手書き文章認識について論じる。制約なしで記載された手書き文章の認識は、板書の認識、手書きメモの認識、罫線のないノートの認識、郵便物の手書き住所の認識などの幅広い応用が期待される重要な技術である。特に手書き住所の認識は郵便物の自動区分機として早くから実用化されている産業上重要な技術である。

文字の認識結果から一階層上の単語、更にはもう一階層上の文章を認識する処理の難しさの本質は、文字と単語、単語と文章の間に強い依存関係があり、上位階層を考慮することなしに下位階層を決めることができないことにあると考えられる。このような問題は、全ての階層の認識を統合し、**top-down** のアプローチで解くべきであると考えられる。しかし、自由手書き文章において、これまで全階層の統合は成されておらず、**top-down** アプローチでの認識は実現されていない。そこで本研究では、**top-down** アプローチでの自由手書き文章の認識を扱う。また、自由手書き文章認識はクラス識別問題であるから、エラー率最小とするためにはベイズ決定則に基づいた識別を行うのが最適である。本研究では、ベイズ決定則に基づいた自由手書き文章の認識を論じる。

まず、自由手書き文章認識の代表例である手書き住所認識について、一般的な処理の流れと従来研究を 2 章で説明する。本研究では処理の流れの中で前半に位置している「前処理」、「二値化」、「行抽出」、「単語仮説生成」については従来法を用いるものとし、後半の「住所データベース探索」、「単語認識」を研究対象とする。

次に、環境依存による文字変形を論ずる。3 章では、音声認識で発展している環境依

存 HMM を手書き単語認識に利用する場合の問題点を明らかにし、手書き単語認識に有効である Partial Tied-Mixture モデルを提案する。環境による変動が支配的である場合、先に環境をクラスタリングするという従来法は有効に働くと考えられるが、手書き文字の場合、字体（ブロック体／筆記体）、手書き変形などによる変動が環境に比べ大きいいため、従来法はうまく働かないと考えられる。Partial Tied-Mixture モデルは、各環境の GMM の mixture 単位で互いに共有することを許すモデルであり、環境のクラスタリングと GMM 推定を同時に行うことに相当している。そのため、手書き文字のような環境以外による変動の大きな場合でも有効に働くことが期待される。Partial Tied-Mixture のパラメータ推定を尤度最大化基準で行い、EM アルゴリズムを適用することでパラメータの更新式を導く。CEDAR データベースを用いた手書き単語認識実験により、Partial Tied-Mixture の有効性が示される。

次に、単語認識について議論する。単語認識の解析的手法と全体的手法はそれぞれにメリット・デメリットがあるため、両手法を併用するのがよいとされている。本研究でも認識精度向上のため両手法を併用する。文字・単語・文章を通した top-down アプローチをベイズ決定則に基づいて実現するためには、単語認識の評価値は尤度か事後確率でなければならない。全体的手法では主に隠れマルコフモデルが用いられ、尤度を出力することができるが、解析的手法では従来事後確率の計算法は存在していない。

そこで 4 章では、解析的手法における事後確率の近似計算式を導出する。従来提案されているベイズ決定則に基づいた方法は、文字の切り出しが一意に定まりかつ語彙内単語の長さが一定の場合にのみ適用できる、非常に範囲の限定されたものであった。この方法を一般的な場合に単純に適用すると、単語長が長いほど不利になるという問題が生じる。そこで本研究では、まず一般の場合に計算すべき事後確率を示す。この事後確率はそのままでは計算困難であるため、近似・式変形により各文字仮説の尤度や事後確率などから計算できる形を導出する。更に、ベイズの定理の分母に対する粗い近似を改善する別の近似変形を示す。分母を語彙内単語の和の形に分解する方法（正規化事後確率比）、注目する単語仮説以外にも何らかの文字列が書かれているとする方法（拡張事後確率比）、分母にも何らかの文字列が書かれているとする条件を付ける方法（探索向け拡張事後確率比）が導出される。スウェーデンの手書き都市名单語を用いた単語認識実験により、提案法が従来法に比べ優れており、更に近似精度改善に伴い認識率が向上することを示す。

更に、5 章では、解析的手法と全体的手法をベイズ決定則に基づいて統合する方法を提案する。従来は解析的単語認識の事後確率を計算する方法がなかったため、統合方法もヒューリスティックにならざるを得なかった。我々は、まず計算すべき統合後の事後確率が、解析的単語認識での全情報と全体的単語認識での特徴ベクトル列を両方を条件としたものであることを示す。この事後確率はそのままでは計算困難であるため、4 章の成果をもとに近似・式展開を行い、計算可能な形を導出する。更に、統合時に初めて

問題となる 2 つの近似誤差を示し、その補正方法を提案する。1 つは解析的単語認識の事後確率計算にて、文字切り出し失敗を考慮していないことに起因する近似誤差であり、考慮した計算式を導出する。もう 1 つは、全体的単語認識にて、明らかに相関の高い各特徴ベクトルを独立とすることに起因する近似誤差であり、べき乗するというシンプルな補正により近似精度を向上する。スウェーデンの手書き単語画像を用いた単語認識により、2 つの近似誤差補正を用いることで従来法を凌ぐ認識率を達成できることを示す。

最後に、文章認識について議論する。文章認識問題は、**tree** 構造の探索問題と見なすことができる。6 章では、**tree** 構造の探索問題と見なすことのできるパターン認識問題全般を議論の対象とする。一般にパターン認識で用いられているビーム探索では、与えられた制限時間に合わせてビーム幅の調整を行う必要がある。しかし、一般に制限時間が各データごとに異なる場合が少なくない。このような場合、制限時間があらかじめ分からないためビーム幅の調整を行うことができず、認識精度の低下を招く。そこで、上述の問題の起こらない最良優先探索を用いる。最良優先探索では異なる **depth** のノードを比較する必要があるが、従来は適切な評価値がなかった。我々はノード間の比較にベイズ決定則を採用したベイズ最良優先探索(**Bayesian Best-Firstsearch: BB search**)を提案する。ベイズ決定則を用いるため、事後確率の近似計算式を導出する。スウェーデンの手書き住所認識実験により、ビーム探索に対する優位性が示される。