

CNNと制約充足による手書き変体仮名認識

渡辺 悟[†] Marcelo Yukio Bressan Hossomi[‡] 鈴木 徹也[†] 相場 亮[†]

芝浦工業大学 大学院 理工学研究科 システム理工学専攻[†]

University of São Paulo, Institute of Mathematical and Computer Sciences[‡]

1 はじめに

古文書の翻刻は国文学の研究にとって基礎的な作業である。しかしそれには多くの知識と労力を必要とする。その理由の一つは古文書に使われている変体仮名にある。現代仮名とは異なり変体仮名においては、同じ読みをする仮名でも、字母やくずし方が異なるために、異なる形状の仮名文字が複数存在し、また異なる読みの仮名で極めて類似した形状を持つものも存在する。

我々は古文書の翻刻支援を目的として、制約充足による手書き変体仮名認識システムを提案してきた [1, 2, 3]。そのシステムでは、まず画像認識器が入力画像中の各文字の読みの候補を列挙する。次に制約解消器が、その読みの候補の組み合わせの中から最適な単語列を探索する。最後に、出力された複数の最適解の中から適切な読みを人間が選択する。

これまでの研究では、画像認識器が未完成だったため、各文字の読みの候補を手作業で作成し、制約解消器の評価実験を行った。そのため、システム全体の評価はできていなかった。

本研究では、画像認識器を作成し、制約解消器と組み合わせ、制約充足による手書き変体仮名認識システムの評価を行うことを目的とする。

以下、第2節では我々のこれまでの研究を簡単に説明する。第3節でCNNを利用した画像認識器、第4節で制約解消器の改善について説明する。第5節で実験結果とその評価を示し、最後に第6節でまとめと今後の課題を述べる。

2 制約充足による手書き変体仮名認識

2.1 制約充足問題

制約充足問題は変数とその領域、制約から構成される。制約とは、変数が同時にとることのできる値の集合を表す条件のことである。変数への値の割り当てを解と呼ぶ。

制約は優先度によって、必ず充足すべき必須制約と可能な限り充足すべき選好制約とに分類される。より優先度の高い制約を満たす解が最適解となる。

2.2 翻刻制約充足問題

翻刻制約充足問題 (R-CSP) とは翻刻支援を目的と

した制約充足問題である [1, 2]。変数は入力画像中の変体仮名、その領域は読みの候補となる。R-CSPの制約には、単語列制約（読みが単語列になるという制約）、等号制約、非等号制約、単語開始制約、単語終了制約がある。R-CSPは単語列制約を必須制約として必ず含む。その他の制約は必須制約もしくは選好制約である。

2.3 読みの割り当てグラフ

単語列制約を満たすために、単語辞書を参照して可能な読みを割り当てたグラフ (読みの割り当てグラフ) が解の探索に先だてて作成される。図1はある読みの割り当てグラフの一部である。各ノードは1つの単語を表しており、上部に変数名、下部にその変数へ割り当てられた文字が表記されている。最上流ノードから最下流ノードへのパスに沿って得られる変数への読みの割り当てが解の一つになる。

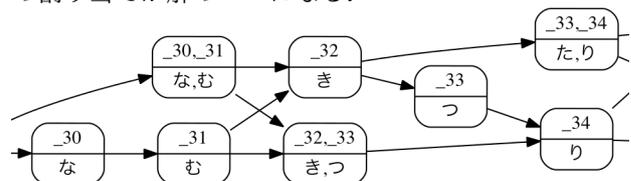


図1: 読みの割り当てグラフの一部

2.4 解の探索方法

コスト最小法を用いてヒューリスティックな解を探索する [3]。コスト最小法とは形態素解析に用いられる手法の一つであり [4]、単語の生起コストと接続コストを定義し、コストの総和が最小となるパスを解として選択する方法である。制約解消器の手順を以下に示す。

1. R-CSP から読みの割り当てグラフを作成する。
2. 読みの割り当てグラフにコストを付与する。
3. コストの総和が少ない順に上位 N 個の解を求める。
4. 解の中から最適解を求める。

3 CNNを利用した画像認識器

画像認識の手法として、画像認識の分野で大きな成果をあげているCNN(Convolutional Neural Network)[5]を利用する。CNNは多層パーセプトロンの一種で、ニューロンが平面上に並べられており、結合が局所的であるという特徴がある。

画像認識器は、正規化した変体仮名の画像をCNNによって字母に分類し、その字母に対応する現代仮名を読み候補として出力する。出力される候補は確率の高い上位 M 個である。また、その中で確率が X 以下の候補 (1 位は除く) は削除し、1 位の確率が Y 以上だった場合は 2 位以下の候補を削除する。

Recognizing Historical KANA Texts using CNN and Constraints

[†]Satoru WATANABE, [‡]Marcelo Yukio Bressan Hossomi,

[†]Tetsuya SUZUKI, [†]Akira AIBA

[†]Division of Systems Engineering and Science, Graduate School of Engineering and Science, Shibaura Institute of Technology

[‡]University of São Paulo, Institute of Mathematical and Computer Sciences

4 制約解消器の改善

4.1 読みの割り当てグラフ作成の高速化

読みの割り当てグラフを作成する時には単語辞書を参照する。これまでは単語辞書の実装に関係データベースを用いていた。今回ダブル配列 [6] で実装したトライ木を用いることにし、高速化を図った。

4.2 画像認識結果の順位を利用したコスト補正

以下の式でノードの生起コストを補正する。

$$cost' = cost + \alpha |cost| \sum_{i=1}^{|w|} (rank(c_i) - 1) \quad (\alpha \geq 0)$$

w はそのノードに対応する単語、 $|w|$ は単語 w の長さ、 c_i は w の i 番目の文字、 $rank(c_i)$ は c_i の画像認識結果の順位、 α は補正係数である。M=1 ときは $rank$ が常に 1 になるので補正項は α に関わらず 0 になる。

4.3 必須制約を満たさないノードの削除

読みの割り当てグラフ作成時、単語列制約以外の必須制約を満たさないノードを削除する。必須制約を満たさない解は最後に取り除かれるので、グラフ作成時にノードを削除することで、無駄な探索を抑えられる。

5 実験

5.1 画像認識実験

CNN の実装には Pylearn2 [7] を使用した。CNN の訓練データとして、変体仮名の字典 [8] を使用した。認識対象の文字画像は伊勢物語 [9] の一部を使用した。文字画像は人手で複数の切り出し方を考慮して切り出した。複数の切り出し方とは、例えば、ある画像領域について 1 文字とみなす場合と 2 文字とみなす場合の 2 通りで切り出すということである。

M=4, X=0.001, Y=0.95 で実験を行った。候補に正解が含まれる割合は 0.7846 (102/130)、順位の分布は 1 位 72 個、2 位 14 個、3 位 10、4 位 6 個、1 位正解率は 0.5538、候補数の平均は 3.216 となった。正誤判定は正しく切り出した画像にのみ行った。

5.2 制約充足実験

単語の生起コストと接続コストは中古和文 UniDic [10] のものを使用する。

グラフ作成について、画像認識実験で作成した R-CSP に対して実験を行ったところ、関係データベースでは 323 秒かかっていたが、ダブル配列では 2.71 秒でグラフを作成することができた。

M と α の値を変化させたときの平均正解率を表 1 に示す。表 1 から、画像認識結果の順位を利用したコスト補正によって平均正解率が向上しているのが分かる。平均正解率は M=3, $\alpha=0.25$ のとき最大値 0.5554 となった。ただし、 $\alpha=0.01$ の場合、 $\alpha=0.0$ の場合との差が小さい。また、M \geq 2, $\alpha=1.0$ の場合、M=1 の場合との差が小さい。

	$\alpha=0.0$	$\alpha=0.01$	$\alpha=0.25$	$\alpha=1.0$
M=1	0.5231			
M=2	0.4654	0.4646	0.4900	0.5308
M=3	0.4838	0.4838	0.5554	0.5308
M=4	0.4592	0.4685	0.5508	0.5385

さい。また、M \geq 2, $\alpha=1.0$ の場合、M=1 の場合との差が小さい。

5.3 評価と考察

画像認識での 1 位正解率 0.5538 に対して、制約充足での最高平均正解率は 0.5554 とほとんど同じであるが、作成された R-CSP には正しく切り出されていない画像の認識結果も含まれているので、制約充足は有効であると考えられる。

パラメータ M と α について考察する。M を小さくすると候補に正解が含まれる割合が減り、M を大きくするとその割合が増えるが、候補が増えることにより平均正解率が下がる。M が大きくても、コスト補正によって平均正解率の向上が見込めるが、 α を大きくすると 1 位候補が選ばれすぎてしまい、M を大きくした意味が小さくなる。よって、M=3~4, $\alpha=0.25$ 程度が良いと考えられる。

6 おわりに

CNN を利用した画像認識器の作成と制約解消器の改善を行い、制約充足による手書き変体仮名認識システムの評価実験を行った。実験により、制約充足の有効性を確認した。

今後の課題として、画像認識器の精度向上、文字の自動切り出し、システムのインターフェース作成が挙げられる。

参考文献

- [1] 新井侑太, 鈴木徹也, 相場亮. 手書き変体仮名認識における制約充足問題の拡張. 第 75 回全国大会講演論文集, pp. 331-333, 2013.
- [2] Yuta ARAI, Tetsuya SUZUKI, and Akira AIBA. Recognizing historical kana texts using constraints. In *Theory and Practice of Computation*, Vol. 7, pp. 151-164, 2013.
- [3] 渡辺悟, 鈴木徹也, 相場亮. 手書き変体仮名認識における単語の接続関係を用いた解の絞り込み. 情報処理学会論文誌, Vol. 56, No. 3, pp. 951-959, Mar. 2015.
- [4] 奥村学. 自然言語処理の基礎. コロナ社, 2010.
- [5] Yann Lecun, Lon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. In *Proceedings of the IEEE*, pp. 2278-2324, 1998.
- [6] 森田和宏, 泓田正雄, 大野将樹, 青江順一. ダブル配列における動的更新の効率化アルゴリズム. 情報処理学会論文誌, Vol. 42, No. 9, pp. 2229-2238, Sep. 2001.
- [7] Ian J. Goodfellow, David Warde-Farley, Pascal Lamblin, Vincent Dumoulin, Mehdi Mirza, Razvan Pascanu, James Bergstra, Frédéric Bastien, and Yoshua Bengio. Pylearn2: a machine learning research library. *arXiv preprint arXiv:1308.4214*, 2013.
- [8] 笠間影印叢刊行会. 字典かな: 写本をよむ楽しみ. 笠間書院, 新装版, 2003.
- [9] 鈴木知太郎. 御所本伊勢物語 冷泉為和筆 宮内庁書陵部蔵 影印本. 笠間書院, 1994-4-30.
- [10] 国立国語研究所. 中古和文 unidic ver.1.3. <http://www2.ninjal.ac.jp/lrc/index.php?UniDic%2F%C3%E6%B8%C5%CF%C2%CA%B8UniDic>. 参照 2014-01-08.