

手書き文字認識技術の Web アプリケーションへの応用

佐野 将史 織田 英人 中川 正樹
東京農工大学

1. はじめに

普及の兆しが見え始めたタブレット PC により、電子ペンで入力できる環境が徐々に増加している。今後、このような環境が一般的になれば Web アプリケーションでも電子ペンの利用が予想される。しかし、電子ペンによる入力は単なる点列データなので、意味のあるデータとして扱うには手書き文字認識技術が必須である。

2. 提案する Web アプリケーションの概要

現在、OCR など様々な手書き文字認識技術が開発され、実用化されているが、個々のプログラムに組み込む必要があったので、限られた範囲でしか利用されていなかった。本稿では、Web ブラウザ上で手書き入力した内容に手書き文字認識技術を適用することにより、Web ブラウザさえあれば誰でも簡単に手書き文字認識技術を利用できる Web アプリケーションを提案する。

3. 開発環境および実装

Web サーバは Apache HTTP Server で構築し、SQLite が標準で含まれる最新版の PHP を組み込んだ。基盤となる Web ページは XHTML で構築し、手書き入力を含む GUI は SVG で描画する。GUI の描画に必要な情報は、すべて Ajax[1] による非同期通信で取得する。利用者が描画した点列データは XSLT プロセッサで SVG から InkML に変換後、Web サーバ上に保存し、SQLite によるデータベースで管理する (図 1)。

4. 手書き文字認識技術の応用

枠あり単文字認識技術 [2] は、Ajax による非同期通信と組み合わせてキーボードに代わる文

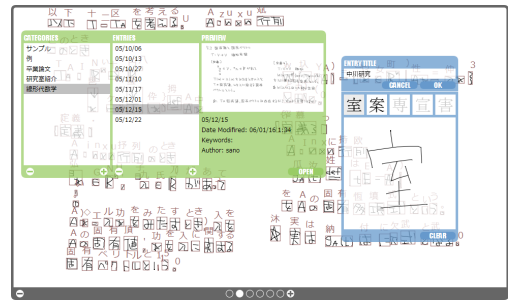


図 1. 動作画面

字入力手段、枠なし文字列認識技術 [3] は、描画内容の分割および認識結果の表示に応用した。

5. おわりに

SVG と Ajax による GUI を構築し、枠あり単文字認識技術および枠なし文字列認識技術を応用した Web アプリケーションを開発できた。

今後は、図や数式の分離技術を組み込むことにより認識精度を向上させるだけでなく、数式認識技術や文字列検索技術を応用することも検討している。

参考文献

- [1] Jesse James Garrett: "Ajax: A New Approach to Web Applications" (2005).
- [2] K. Akiyama and M. Nakagawa: "A Linear-Time Elastic Matching Algorithm for On-line Recognition of Handwritten Japanese Characters," The Transactions of IEICE, Vol.J81-D-2, No.4, pp.651-659 (1998).
- [3] M. Nakagawa, B. Zhu and M. Onuma: "A Model of On-line Handwritten Japanese Text Recognition Free from Line Direction and Writing Format Constraints," IEICE Trans. Inf&Syst., Vol.E88-D, No.8, pp.1815-1822 (2005).