

木簡解読支援のための情報処理システムの実現

戸根 康隆[†] 未代 誠仁^{††} 中川 正樹^{††} 馬場 基^{†††} 渡邊晃宏^{†††}

東京農工大学 工学部[†] 東京農工大学^{††} 奈良文化財研究所^{†††}

1. はじめに

日本では『木簡』と呼ばれる墨を用いて木片に文字情報を記した歴史的文書が奈良時代に広く使われていた。これまでに国内各地の遺跡から出土した木簡の総数は約 32 万点に上り、そのうち 17 万点以上が奈良の奈良平城京跡から出土したものである。

しかし、出土した木簡の多くには割れや掠れ・にじみなどの破損・汚損が見られ、文字情報を完全に留めたものは珍しい。このため、木簡の解読は専門家にとっても困難であり、これまでに解読が行われた木簡は数万点にとどまる。

我々は、コンピュータによる文字認識、画像処理、検索などを用いた木簡解読支援システム（以下、本システム）を提案、実装した。しかし、これらの機能を専門家が有効に活用するためには、専門家に使いやすいユーザインタフェースを提供する必要がある。特に、直感的な操作性を犠牲にすることなく、操作に必要な手間を減らすことが重要であると考えた。また、解読の最中に複数の仮定が並立する場合、ユーザがこれらの仮定を比較検証するための仕組みが必要である。

本報告では、上記の点を踏まえた本システムの改良について述べる。

2. 木簡解読支援システム

本システムは文字認識、画像処理、地名検索の 3 つの機能、およびこれらをユーザに提供する GUI から構成されている。また、GUI には機能ごとにウィンドウを持つマルチウィンドウインタフェース方式を採用している。（図 1）

本システム上において、ユーザは木簡画像から任意に選択した領域に対して画像処理・文字認識処理を行うことができる。また、地名検索機能を利用して解読結果を検証することが可能である。

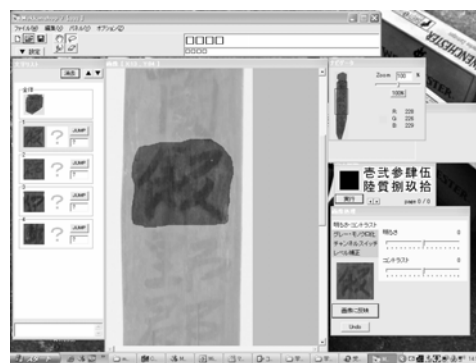


図 1：本システム実行画面

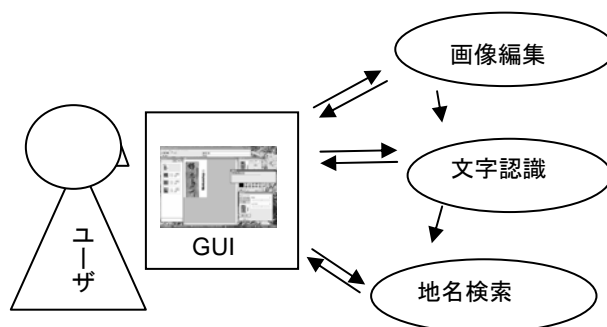


図 2：システムの構成

これらの機能を 1 つのアプリケーションとして提供することで、市販のアプリケーションを組み合わせるよりも円滑に解読作業を進めることができる。（図 2）

3. 木簡解読支援システムの改良

本システムは解読が円滑になるよう支援するという点から見て有効との意見を頂いた。しかし、本システムに対してこれまでユーザから得られた指摘・評価からは更なる支援の方法があると考えられる。具体的な改善案を以下に示す。

(1) タイリングパネルウィンドウ方式への転向

本システムではウィンドウが 7 つあり、各ウィンドウ同士が互いに重複するために十分な表示領域を得られない。各ウィンドウ間の優先順位が確定している場合はポップアップ形式を取ることも可能だが、解読作業においてユーザが必要とする情報には優劣をつけられない。よって、ユーザが

Implementation of Information Processing System to Support Reading Mokkan
Yasutaka Tone[†], Akihito Kitadai^{††}, Masaki Nakagawa^{††},
Hajime Baba^{†††}, Akihiro Watanabe^{†††}

各ウィンドウの管理を容易に行えるタイリングパネルウィンドウ形式が良いと考えられる（図3）。

(2) 抽出画像リストの座標に対応した表示形式

抽出画像のリストは抽出した順に時系列でリスト表示されていく。これではリストを見ただけでは木簡の全体的な文章構造を理解することは難しい。（図4）

よって、抽出画像リストは座標に即して表示する形式を取るほうが良いと考えられる。（図5）

(3) ユーザの試行錯誤の推移の保持

文字の読み取りが難しい木簡の解読作業において、ユーザは木簡上の文章に様々な仮定を立てる場合がある（以降「仮定の推移」とする）。このとき、各仮定を仮定パネルとして保持し、閲覧可能な状態に置くことで解読作業において自身の推理の過程を把握でき、解読作業は更に円滑に進むと考えられる。（図5）

(4) 仮定の推移の発生の自動判定

現在、新しい木簡解読支援システム（以下、新システム）における仮定の推移の発生判定はユーザ自身による手動判定と、システムの自動判定の2通りある。ここで、仮定の推移の発生に一定の法則を見出し、仮定パネルの生成を全てシステムが自動的に行えれば解読作業の大きな支援になると考えられる。

今回、新システムでは新規に抽出した領域と既存の領域の重複した面積を自動判定のパラメータとする。重複した面積が既存の抽出領域の面積と比較して「一定の割合」を超えた場合に仮定の推移が発生したとみなす。（図6）

上記の改良案をもとに我々は新システムを構築した。

4. おわりに

本稿では、木簡解読支援のための情報処理システムの改善について述べた。今後の課題は評価実験によって重複面積を自動判定する有効なしきい値の推定、及び地名検索の結果を活用した文字認識手法の実現などが挙げられる。

謝辞

本報告は、日本学術振興会科学研究費補助金基盤研究S: No. 15102001 の補助による。

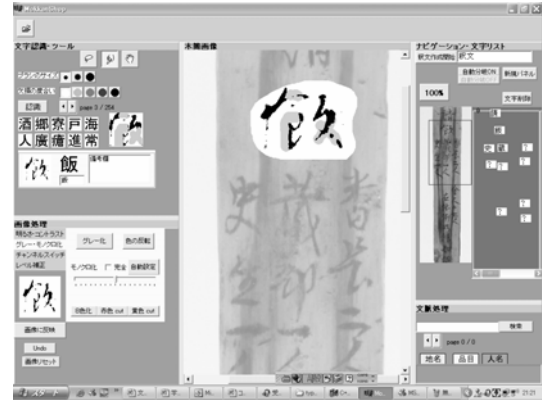


図3：新システムのGUI

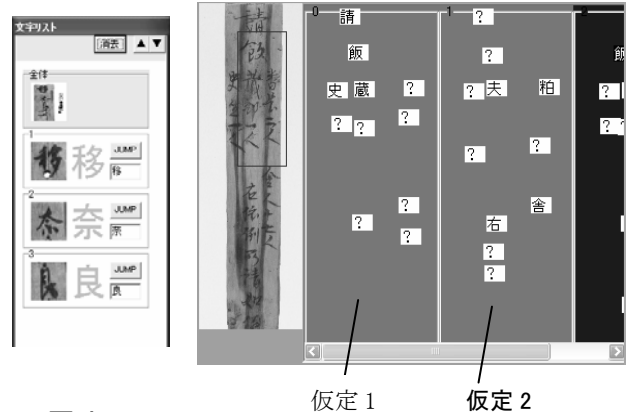


図4

図5

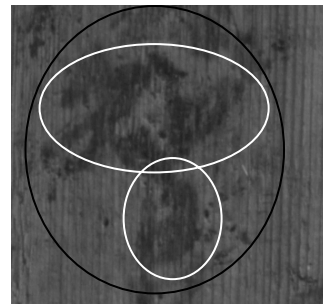


図6：「智」から「知、白」への仮定の推移

参考文献

- [1]A. Kitadai, et al.: Support System for Archeologists to Read Scripts on Mokkans, Proc. 8th ICDAR, Seoul, Korea, pp.1030-1034 (2005.9)
- [2]A. Kitadai, et al.: Context Processing to Read Text on Damaged Wooden Tablets, Proc. 10th IWFHR, La Baule, France, Vol.I, pp.581-586 (2006.10).