

## 目次

はしがき	3
目次	4
I 本編	
木簡など出土文字資料の資源化のための機能的情報集約と知の結集	9
I-1 総論	11
I-1-1 研究の概要	11
I-1-2 研究の背景と目的	13
I-1-2-1 研究の背景	13
I-1-2-2 研究の目的	13
I-1-2-3 当初の研究計画	16
I-1-3 研究発表	20
I-1-4 研究の経過と成果	38
I-1-4-1 研究経過	38
I-1-4-2 研究成果	41
I-2 各論	50
I-2-1 木簡など出土文字資料の情報抽出の効率化	50
I-2-1-1 資料の情報を効率的に管理・活用するシステムの開発	50
I-2-1-1-1 アノテーションツールの開発・運用 — 台紙アノテーションツールの開発 (i)	50
I-2-1-2 情報そのものの拡充	52
I-2-1-2-1 奈文研保管木簡撮影フィルムデジタル化 (ii)	52
I-2-1-2-2 奈文研保管木簡のカラー写真撮影 (iii)	53
I-2-1-2-3 全国の出土木簡等の撮影 — 出土文字資料データの管理・活用 (iv)	53
I-2-2 木簡など出土文字資料の知の結集	57
I-2-2-1 画像検索機能の強化	57
I-2-2-1-1 木簡・くずし字解読システム—MOJIZO—、 及び同スマホタブレット版の開発 (v・vi)	57
I-2-2-1-2 形状による検索ソフトの開発 — 削屑断片接合処理技術の開発 (vii)	64
I-2-2-2 出土文字資料の字体の研究	66
I-2-2-2-1 木簡所用文字の標準字体の検討 (viii)	66
I-2-2-3 検索機能の強化	68

I-2-2-3-1	木簡字典の高次化 — 意味検索機能の追加 (ix) .....	68
I-2-2-3-2	文字認識とそのため画像処理 (x) .....	70
I-2-2-4	周辺データベース群の充実 .....	70
I-2-2-4-1	木簡研究論文データベース — 「木簡関係研究文献一覧」の作成 (x i) .....	70
I-2-2-4-2	古代地名検索システム (x ii) .....	73
I-2-2-4-3	正倉院文書字典 — 正倉院文書書状データベース (x iii) .....	78
I-2-3	出土文字資料統合データベースの構築 .....	83
I-2-3-1	統合データベース構築 .....	83
I-2-3-1-1	新木簡データベース「木簡庫」の開発 (x iv) .....	83
I-2-3-1-2	木簡庫のメタデータについて .....	85
I-2-3-1-3	木簡庫のトップページ .....	89
I-2-3-1-4	木簡庫の検索条件の設定 .....	92
I-2-3-1-5	木簡庫の検索結果ページ (一覧表示) .....	95
I-2-3-1-6	木簡庫の検索結果ページ (カード表示) .....	97
I-2-3-1-7	木簡庫の検索結果ページ (文字画像一覧表示) .....	100
I-2-3-1-8	カテゴリ検索 (意味検索) .....	101
I-2-3-1-9	「文字画像をさがす」「画像から」— MOJIZOへのリンク .....	101
I-2-4	研究成果の公開と活用 .....	103
I-2-4-1	木簡ワークショップの開催 (x x i) .....	103
I-2-4-2	木簡展示への活用 (x x ii) .....	104
	(附) 「平城宮跡出土木簡」の国宝指定 .....	107
I-3	まとめと展望 .....	109
II	論考編 .....	111
II-1	研究成果の展望 .....	112
II-1-1	渡辺晃宏「木簡字典15年の歩みと課題」 .....	113
II-1-2	渡辺晃宏「月城亥字出土木簡と日本古代木簡の比較」 .....	121
II-1-3	未代誠仁・高田祐一・井上幸・方国花・馬場基・渡辺晃宏・井上聡 「字形画像をキーとした情報検索による古文書デジタルアーカイブ活用への効果」 .....	139
II-1-4	渡辺晃宏「日本歴史学における木簡研究成果」 .....	148
II-1-5	未代誠仁・馬場基・渡辺晃宏・井上聡・久留島典子・中川正樹 「古文書字形デジタルアーカイブのための検索システムの試作」 .....	170
II-1-6	渡辺晃宏「古代木簡研究の意義と情報技術の価値」 .....	176

II-1-7	井上幸「古代木簡研究におけるデジタルデータの整理と集積」	181
II-1-8	馬場基「古代木簡研究における情報の活用と今後の課題」	185
II-1-9	未代誠仁・白井啓一郎・馬場基・渡辺晃宏・井上聡・久留島典子・中川正樹 「古文書デジタルアーカイブに対する横断的字形検索サービスの試作」	188
II-1-10	K. Shirai, Y. Endo, A. Kitadai, S. Inoue, N. Kurushima, H. Baba, A. Watanabe, M. Nakagawa ” Character Shape Restoration of Binarized Historical Documents by Smoothing via Geodesic Morphology ”	194
II-2	研究成果の活用と関連研究	199
II-2-1	Hung Tuan Nguyen, Nam Tuan Ly, Kha Cong Nguyen, Cuong Tuan Nguyen, 中川正樹 「変体仮名の認識」	199
II-2-2	小口雅史編「小口雅史編「日本古代史関係研究文献目録データベース」 (法政大学国際日本学研究所データサービス内)の構造について」	203
II-2-3	大山航「Merlin-Farberのハフ変換を用いた削屑木簡の再構成」	233
II-2-4	笹原宏之「上代における日本製漢字の作製と使用—その前史と展開—」	246
II-2-5	朱碧蘭「オフライン文字認識の精度向上及びオンライン文字認識機能の構築、 文字認識の研究」	272
II-2-6	岡谷陽平・白井啓一郎・宮尾秀俊・丸山稔 「木簡画像データベース用の単色・カラー画像の判別法 および木簡領域の検出法の検討」	278
II-2-7	白井啓一郎・杉本拓矢・藤森和博・宮尾秀俊・丸山稔 「最適な係数を用いた線形色変換による ノイズの低減された色主成分の抽出法」	291
III	記録編	
	木簡ワークショップの記録	307
III-1	第7回木簡ワークショップ —木簡の情報抽出と表現方法—	308
III-2	第8回木簡ワークショップ	336
III-3	第9回木簡ワークショップ	368
III-4	第10回木簡ワークショップ	399
III-5	第7回～第10回木簡ワークショップ当日配付資料	428
	あとがき	469

## 挿図目次

図1 「木簡庫」とこれまでのデータベース群 相互関係イメージ図……………	47	図32 項目検索（「内容」、「国郡郷里」）……………	94
図2 木簡破片注釈の流れ……………	51	図33 項目検索（「木簡の形」、「出典」・「年紀・人名」 ・「大きさ・樹種等」、「遺跡」）……………	94
図3 MokkaAnnotator II のユーザインタフェース…	51	図34 検索履歴の表示……………	95
図4 MOJIZO と検索の例……………	59	図35 検索結果テキスト一覧表示 （「長屋親王宮」で「すべて検索」）……………	96
図5 クライアント・サーバモデルによる MOJIZO の実装……………	61	図36 検索結果カード表示で「さらに見る」を 開いた状態の例……………	98
図6 MOJIZOkinによる画像処理の例……………	62	図37 別ウィンドウで画像をひらく……………	99
図7 利用者数の変化……………	63	図38 検索結果画像一覧表示（「長屋親王宮」で「テキス トから」検索）……………	100
図8 閾値処理による文字領域の取除……………	64	図39 平城宮跡の既指定木簡出土地と 新指定木簡出土地……………	108
図9 木簡破片画像の識別器……………	64	図40 課題の3レベル……………	200
図10 木簡破片画像の特徴抽出……………	65	図41 人文学オープンデータ共同利用センターが公開 している日本古典籍字形データセット……………	200
図11 特徴ベクトル……………	65	図42 人間でも読みにくい文字の組……………	200
図12 木簡破片の切り分け補助GUI……………	66	図43 DCRNの構成……………	201
図13 古代木簡使用文字字体一覧の一例……………	68	図44 データフォルダ配置……………	208
図14 意味検索の分類階層……………	69	図45-1 ウェブ公開用フォルダの配置……………	209
図15 タッチ操作に対応した新Mokkanshopの 画面……………	70	図45-2 IISで設定したサーバー上での ウェブ公開用フォルダの配置……………	209
図16 木簡庫における「研究文献情報」への リンクボタンとその拡大……………	71	図46 (a) 完全な形状の木簡と (b) 削屑木簡の例…	233
図17 研究文献一覧の表示……………	72	図47 再構成の (a) 成功例と (b) 失敗例……………	234
図18 「日本古代史関係研究文献目録データベース」と CiNiiへのリンク……………	72	図48 慶州月城塚字出土木簡1の「亘」……………	247
図19 古代地名検索システムのトップページ……………	73	図49 『日本上代金石文字典』の「亘」と 『豊後国風土記』の「亘」……………	247
図20 旧版古代地名検索システムの検索の流れ……………	77	図50 「纓」字をめぐる諸資料……………	249
図21 リニューアル版古代地名検索システムの 検索の流れ……………	77	図51 桂本『萬葉集』の「何怜」……………	250
図22 検索結果の一例（「なら」で検索）……………	78	図52 真福寺本『古事記』巻下と 兼永『古事記』巻下の「姦」……………	251
図23 正倉院文書書状データベースの トップページ……………	79	図52 真福寺本『古事記』の「刼」……………	252
図24 簡単検索と詳細検索の検索画面……………	79	図54 兼右本『古事記』の足偏の「游」……………	252
図25 検索結果一覧表示画面（「啓」で検索）と カード表示画面……………	80	図55 真福寺本『古事記』巻中の「申」……………	253
図26 拡大表示画面……………	81	図56 吉田本『日本書紀』巻1の「雲」「霽」……………	255
図27 語句検索の検索ツリー（開いた状態）……………	81	図57 乾元本『日本書紀』巻1の「雲」……………	255
図28 語句検索の検索結果（「書式」を選択）……………	81	図58 『豊後国風土記』の「霽」……………	255
図29 木簡庫のトップページ……………	89	図59 『篆隸万象名義書』巻23龍部の「霽」……………	255
図30 「木簡をさがす」の「すべて検索」、同「本文検索」、 同「カテゴリー検索」、「木簡画像をさがす」の「テ キストから」検索……………	90	図60 韓国羅州伏岩里遺跡出土木簡に 見える「亘」……………	256
図31 外国語版のトップページ（英語、中国語 （簡体字）、中国語（繁体字）・韓国語）……………	92	図61 法隆寺観音菩薩造像銘の「鵠」……………	257

図62	観智院本『類聚名義抄』佛中52ウの「嚙」…	258	図78	英単語“employees”の特徴点抽出…	274
図63	『播磨国風土記』と「御野国味蜂間郡 春部里戸籍」の「早」…	258	図79	木簡画像の種類…	279
図64	真福寺本『古事記』巻中、同巻上、乾元本 『日本書紀』巻1、吉田本『日本書紀』巻1、 「駿河国天平十年正税帳」の「鞆」…	259	図80	二つの尺度からなる特徴点 ( $f_1, f_2$ ) の分布…	280
図65	天文写本『倭(和)名類聚抄』 第1冊51ウ…	259	図81	非木簡領域の検出、及び検出領域の色の 置換が文字抽出用の色変換に与える影響…	282
図66	「御野国味蜂間郡春部里大宝二年戸籍」 (702)、「写後書所解」天平19年(747)、 「写経所解」天平20年(748)…	260	図82	非木簡領域の検出用に用いるそれぞれの 尺度の計算結果…	284
図67	『説文解字』の「櫃」、『篆隸書万象名義』の 「櫃」、『日本書紀』兼右本の「檣」、 『日本書紀』岩崎本の「檣」…	261	図83	グレースケール画像とカラー画像の 分類結果…	285
図68	『唐甫里先生文集』一卷と『新撰字鏡』 所引『小学篇』…	261	図84	低コントラスト画像として誤って分類 された画像の特徴点の座標…	285
図69	『日本上代金石文字典』、『上代木簡資料 集成』、狩谷掖斎、山田孝雄・香取秀真 『古京遺文』…	261	図85	非木簡領域の検出と領域色の置換の 成功例と失敗例…	287
図70	続々修3-2裏の「櫃」と続々修6-2の「檣」…	262	図86	尺度1の計算(式6)で用いる $Y_i - \bar{Y}$ の値の分布…	288
図71	西本願寺本『万葉集』巻7、細川家本 『出雲国風土記』(733)54ウ…	262	図87	尺度2の計算(式7)で用いる $\log(K \text{ box} * (Y_i - Y_i)^2)$ の値の分布…	288
図72	「擧」(異体字) 鋌(異体字) 十六隻 (異体字)…	263	図88	尺度3の値の分布…	288
図73	「鉈」の見える唐招提寺文書 「造唐招提寺用度帳」…	264	図89	異なる合成計数の組を用いた色の線形結合 (線形色変換)の結果…	292
図74	平城京左京三条二坊一・二・七・八坪 (長屋王邸跡)出土木簡の「鯛」…	265	図90	図88に示した画像の最適解…	299
図75	オンライン手書きパタンと オフライン手書きパタン…	272	図91	パラメータの設定と解の収束の様子…	299
図76	ペンや指などの入力デバイス…	273	図92	代表的な色変換方法と色成分分解方法 との比較…	300
図77	切り出し手法…	273	図93	線形判別分析(LDA)を用いた文字色と 背景色の分類法との比較…	301
			図94	強力な雑音除去能力や領域分割性能をもつ 標準手法との比較結果…	302
			図95	シミ検出：布地の上のコーヒーの シミの検出…	303

## 表目次

表1	MOJIZOに関連する活動履歴…	63	表5	検索対象にする項目としない項目…	214
表2	新統合木簡データベース「木簡庫」と 従来のデータベース群との関係…	84	表6	日本語表示用のNamazuテンプレート一覧…	216
表3	木簡データベース・木簡字典と木簡庫の メタデータ項目の比較…	88	表7	オンラインとオフライン認識手法の 長所・短所…	272
表4	データ項目一覧…	213	表8	カラー画像に分類された画像での 木簡領域と非木簡領域の検出精度…	286