

かな文字特徴抽出による「奥の細道」筆写本の筆者認識支援

宇田 紀之†, 浜 森太郎††, 木村 文隆†††, 三宅 康二†††

† 三重大学・情報処理センター †† 三重大学・人文学部・文化学科

††† 三重大学・工学部・情報工学科

〒514-0008 三重県津市上浜町1515

(和文アブストラクト)

本研究は、芭蕉文学研究用データベースに収納された「奥の細道」の筆写本3部を研究対象にし、文字画像データに計算機による統計解析を適応して、筆写本の筆者認識を試みた。文字の特徴量抽出には64次元加重方向指数ヒストグラム法を用い、筆者を特徴づける個性字種の識別には、K次近傍識別則を用いた。実験の結果、1)「を(遠)」において安定して見られる構造的特性、(2)「れ(禮)」、「な(奈)」における独自の文字変形パターンなど、従来、気づかれていなかった筆記特徴を発見した。検出した個性字種について、芭蕉自筆文書の文字と照合したところ、「を」、「な」に類似の筆記特性のあることが確認された。文字画像データベースにおける筆者認識の問題設定、及び、統計解析の効果と問題点について報告する。

Supporting for writer recognition on the copy book of

"Oku no hosomichi" using a feature extraction of kana characters

Uda Noriyuki† Hama Shintarou†† Kimura Fumitaka††† Miyake Yasuji†††

† Information Processing Center, Mie University

†† Faculty of humanities and social sciences, Mie University

††† Faculty of Engineering, Mie University

〒514-0008 1515, Kamihama, Tsu, Mie, Japan

(Abstract)

We have done a writer recognition experiment about three copybooks of "Oku no hosomichi (Travel reports of Touhoku Hokuriku area of Japan, in the Edo periods)", applying the statistical analysis on the kana character pattern. The data of our experiment was stored on the database for the literature research about a haiku master, Matsuo Bashou.

The weight direction code histogram method on 64 elements was used for feature extraction from character pattern. The distinct drawing individuality of writer was found on the following kana character "wo", "re", "na". Which was unknown among researchers of Bashou. This paper discussed about the problem setting and the effectivity of statistical analysis on the writer recognition.

1. はじめに

文字を書いた人物に関してなんらかの判断を下すことを筆者認識という。歴史学、国文学などの分野では、筆者認識は、資料の信頼性を決める最も基本的で重要な研究課題のひとつである。近年、文化遺産の保存や電子図書館サービス充実の要請から、文化的・歴史的文書資料のデジタル化が進んでおり、計算機による筆者認識支援への期待が高まっている。

計算機を使った筆者認識の研究は、急速な発展を遂げ^{1),2),3),4)}、一定の制約条件のもとでは95%以上の認識率を達成し、筆者認識の技術を応用した小切手署名照合システム⁵⁾などは、すでに実用段階にある。しかしながら、国内に限ってみれば、文化的・歴史的文書資料の筆者認識に、計算機を使った統計的手法が利用されることは稀である⁶⁾。計算機がこの分野で活発に利用されない理由のひとつは、(1)毛筆草書体の柔軟な文字変形、(2)文字品質の劣化とばらつき、(3)文字サンプル数の不足、など歴史的資料に固有する問題である。もうひとつは、筆者認識の問題設定に関わる問題である。たとえば、筆者認識の問題設定を正当化していた関連資料の信頼性に疑問がある場合、筆者照会システムの精度がどれほど高くても、その認識結果は意味を失う。専門的知識と経験を必要とする文化的・歴史的な文書資料の筆者認識では、関連情報を体系的に整理して、問題設定の妥当性自体を検証する研究用データベースの構築の必要性が言われる理由である⁷⁾。国内では、筆者認識に対応できる研究用データベースが、ほとんど整備されていないのが実状のようである。

本研究は、芭蕉文学研究用に開発されたデータベースを用いて、「奥の細道」筆写本の筆者認識実験を試みた。筆者の個性を表出する個性字種の抽出、文字の変形パターンの発見、及び、実験結果に基づく筆者認識の問題設定の評価について説明する。

2. 研究対象

2.1 芭蕉文学研究用データベース

芭蕉文学研究データベースは、芭蕉文学関連資料の収集と整理し、文献学的、書誌学的研究の支援を目的として、三重大学近世文化研究会(代表:浜森太郎)が作成し、管理運営しているデータベースである。このデ

ータベースの特徴は、1文字単位の文字データを要素として構築されているところで、文字データフィールドには、文字画像データの他に、出典コード、ポジションコード、漢字/かなの区別、品詞、活用形、見出し語(現代かな)、かな字母、語頭/語尾/語中の区別など最大14個属性が記述されている。文字画像データは、影印を含む出版物から、頁画像を一枚ずつスキャナ(解像度600dpi)入力し、ディスプレイを見ながら頁画像から行画像を一行ずつ、行画像から文字画像を一字ずつオペレータが手作業で切り出したものである。

2.2 「おくの細道」筆写本^{8),9),10)}

「おくの細道」は、元禄2年(1689年)の奥州北陸行脚の経験に基づいて執筆された松尾芭蕉の紀行文で、俳諧の弟子である曾良、素龍らの筆写と芭蕉の推敲を経て、定稿が固まってきたことが知られている。元禄初版本の原稿となった素龍筆の筆写本は、代々に継承され、柿衛本と西村本が現存する。通常、西村本を素龍本という。天理大学図書館が所蔵する天理本は、素龍本より前の段階で書かれたもので、墨による書き入れ(墨訂)と朱筆による補訂とが何箇所も加えられている。奥州北陸行脚に随行した曾良の筆写と考えられている。

平成7年(1996年)、大阪において発見された筆写本(新出本)は、「芭蕉直筆」として公表され、話題になったものである。冊子は、貼り紙・書き入れによるおびただしい数の修正が加えられている。修正と清書の前後関係から、新出本は天理本よりも以前に作成されたと考えられる。芭蕉直筆説に対しては、これを疑問視する見解もあり、現在も研究者にホットな話題を提供している。

これら筆写本の筆者認識には、資料の相互関係を検証する文献学的考察とともに、特定の文字の用字特性や筆跡の照会によって行われる。3つの筆写本は、元禄4年から元禄7年にかけての期間に作成されたものと推察され、同時期に芭蕉が記帳していた「三日月日記」や「幻住庵記」の文字を筆跡照会の参照資料とすることが多い。

2.3 かな文字の特性

統計的手法を有効にするためのサンプル数の確保と、筆者個性の把握の容易さの理由から、分析対象は「かな文字」に制限することにした。近世以前の文書では、

かな一文字に対して複数の漢字をあてはめる独自の文字表記が用いられる。かなの原形になる漢字のことを字母という。たとえば、「あ」の字母は、現代ひらがな「あ」の元にもなっている「安」の他に、「阿」「愛」などがある。

字種にもよるが、かなは字母の草書体と字母の文字構造を単純化した簡略書体(くずし文字)を併用して表記される場合が多い。字母のくずしにはいくつかのパターンがあるが、一人の筆者が、1字母に対して使用するくずし文字の種類はそれほど多くないことが経験的に知られている。

2.4 分析データ

「おくの細道」筆写本で使用されるかな文字は 118 字種ある。筆写本内でサンプル数が 50 個以上ある 28 字種のかな文字を分析の対象とした。文字の劣化が著しく、その形状からのみでは、ほとんど判読が不可能であると判断した文字については、分析対象から除外した。分析対象の字種とサンプル数を表1に示す。

筆写本間の文字サンプル数の差異は、修正による文字の増減もあるが、主に、使用するかな字母の切替えやかなの漢字化などが原因である。サンプルの字母以外の字母が使用される場合の主な代替字母を最終列に付記した。三日月日記の総かな文字数は 1252 字で筆写本の約1割の分量である。統計分析対象からは除外した。

3. 筆者認識

3.1 問題設定

筆者認識は、筆者候補を絞り込む筆者識別と、文字を照会して筆者候補に真偽の判定をつける筆者照会の2つを区別する。文学関連資料の筆者認識では、関連資料の記述内容による相互関係において筆者候補を識別し、筆跡照合による筆者照会がよく用いられる方法である。計算機の筆者認識は、文字画像データを画素単位で統計処理して抽出した構造的特徴量を比較し、筆者候補の可能性を確率と

して提示する。問題設定によって、計算機は、識別・照会の両方の道具として利用することができる。文学関連資料の筆者認識に、計算機が利用されるのは稀である。欠損やあいまい文字を含む文書資料の照会では、推

表1:かな文字のサンプル表

Table 1: Number table of kana character sample

かな	字母				代替字母
		新出本	天理本	素龍本	
あ	安	120	114	80	阿
い	以	59	62	70	
か	可	313	226	298	加
き	幾	81	81	80	
く	久	64	67	68	
こ	己	107	91	84	
さ	左	85	89	78	
し	之	227	239	224	
た	多	123	127	116	堂
つ	川	76	72	77	
て	天	374	375	371	
と	止	219	305	328	登
な	奈	125	128	131	
に	尔	201	420	446	耳、丹
の	乃	532	538	585	能
は	者	113	98	72	盤、ハ
ひ	比	97	84	96	
ふ	不	76	70	68	婦
へ	部	59	64	51	遍
ま	末	87	72	56	(漢字化)
も	毛	151	151	128	(漢字化)
ら	良	121	126	122	
り	利	176	186	203	里
る	留	182	210	206	流、ル
れ	礼	110	107	126	連
や	也	105	110	106	
よ	与	63	64	65	
を	遠	241	269	288	越
対象文字数		4287	4545	4823	
総文字数		10546	10596	10436	

論機能を働かせて情報処理する人間の照合能力が、計算機を上回る。ただし、人間の行う筆者認識は、判断基準があいまいで、先入観に支配されやすい欠点がある。一度に処理できる情報量にも限界がある。

筆者認識は、どの文字を識別・照合の参照文字にするかによって左右される。筆者の個性を特徴づけ、他の筆者の区別されやすい文字を個性文字、その字種を個性字種と呼ぶ。本研究では、個性文字、及び、個性字種を選び出すための識別手段として、計算機を利用する。筆者照会は、計算機が抽出した個性文字の文字パターンをディスプレイに表示し、人間の目で照会するようにした。処理の流れを以下に示す。

- (1) 天理本・素龍本・新出本のかな文字の特徴量を測定し、K次近傍識別則を適用して、筆写本

の筆者を特徴づける個性字種を識別する。

(2) 平均近傍の文字サンプルを個性文字として抽出して表示し、個性字種の構造特性を明らかにする。

(3) 個性字種の変形パターンを、文字形状のばらつき具合(分散)と、文字構造の簡略化の程度を示す輪郭画素比によって測定する。

(4) 三日月日記の同字種文字をすべて表示し、個性字種文字で発見した構造特性や変形パターンが存在するかどうかを照会する。

3.2 特徴量抽出

筆者認識では、筆跡照合をモデルにして線分(ストローク)に注目する構造解析法¹¹⁾、曲線処理を効果的に行う局所円弧パターン法¹²⁾などの方法がよく利用され、ほかに、スペクトル解析法¹³⁾、輪郭線の画素連結関係に注目する方向指数ヒストグラム法¹⁴⁾が用いられる。本研究では、かな文字を対象とするため、特に曲線処理に優れた加重方向指数ヒストグラム法を使用する。アルゴリズムを以下に示す(図1)。

- (1) 前処理した文字データに対して、文字部分の輪郭線を8連結で追跡し、各輪郭線方向を45度おきの4方向に量子化する。
- (2) 文字データを、縦7×横7の49個の小領域に分割し、それぞれの小領域内で4方向ごとのヒストグラムを求め、196次元(縦7×横7×4方向)の方向指数ヒストグラムを得る。
- (3) この方向指数ヒストグラムに重なりのある2次元のガウスフィルターをかけて、64次元に次元圧縮したものを加重方向指数ヒストグラム[h1,h2,h3,.....,h64]とする。

3.3 K次近傍識別則

加重方向指数ヒストグラムを用いて、筆者 w_i の字種 c_j の文字 $\{p_{ij}\}$ が、他の筆者の書く同字種の文字に対してどれだけ識別されやすいかを示す識別率 R_{ij} を、K次近傍識別則を適用して、以下の手順で求める。

- (1) 字種 c_j における全文字サンプル P_j と文字サンプル p_l とのユークリッド距離を求める。文字サンプル p_l と p_k のユークリッド距離 $d(p_l, p_k)$ は、以下の式による。

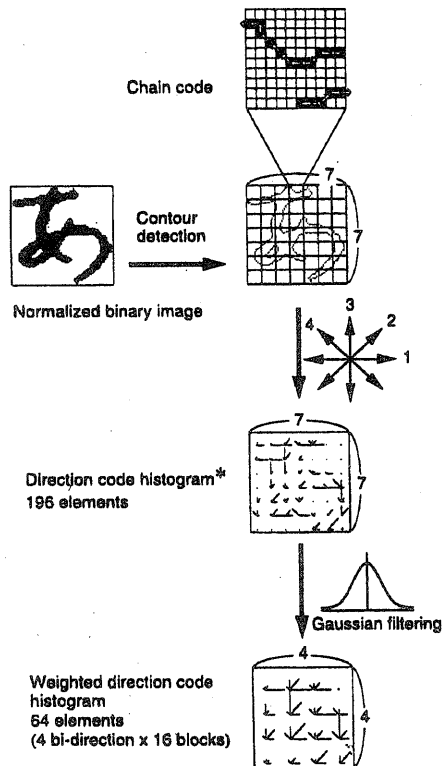


図1:特徴量抽出の処理の流れ
Fig 1:Flow chart of feature extraction

$$d(p_l, p_k) = \sqrt{\sum_{n=1}^{64} (h_{ln} - h_{kn})^2}$$

$$P_j = \{p_{l1}, p_{l2}, \dots, p_{lN_{ij}}\}$$

N_{ij} : 筆者 w_i の字種 c_j におけるサンプル数

$$p_l = [p_{l1}, p_{l2}, \dots, p_{l64}]^T \quad t: \text{転置}$$

- (2) ユークリッド距離の短い順に並び替え、上位 K 個をとる。そのなかに含まれる筆者 w_i の文字サンプルの数 k_i を求める。

- (3) 字種 c_j のなかで k_i が K の過半数を超えているサンプルの数 k_{ij} を求める。筆者 w_i の字種 c_j における識別率 R_{ij} は、筆者 w_i の字種 c_j のサンプル数 N_{ij} に対する k_{ij} の比で求める。

$$R_{ij} = \frac{\text{字種 } c_j \text{ で } k_i \text{ が } K \text{ の過半数を超えているサンプルの数}}{\text{筆者 } w_i \text{ の字種 } c_j \text{ のサンプル数}}$$

3.4 字種の変形パターン

筆者 w_i の字種 c_j の文字特徴を示す、分散 v_{ij} 、輪郭画素比 b_{ij} は、以下の式で求める。

字種 c_j の平均ベクトルを m_{ij} とする。

$$m_{ij} = \frac{1}{N_{ij}} \sum_{n=1}^{N_{ij}} h_{ijn}$$

$m_{ij} = [m_{ij1}, m_{ij2}, \dots, m_{ij64}]^t$ t :転置

文字形状のばらつきを示す分散 v_{ij} は、以下の式で求める。

$$v_{ij} = \frac{1}{N_{ij}} \sqrt{\sum_{n=1}^{N_{ij}} (h_{ijn} - m_{ij})^2}$$

輪郭画素比 b_{ij} は、文字輪郭線の画素数の2乗を文字の全画素数で割る式で求める。

4. 実験

4.1 方法

表1に示した「おくの細道」の筆写本にあるかな文字の画像データを用いて、筆者識別に関わる以下の2つの実験を行なった。

[実験1]文字画像データに加重方向指数ヒストグラム法を適用して64次元の特徴量を抽出し、K次近傍識別法によって字種 C_j ごとの識別率 R_{ij} と分散 v_{ij} を求める。

[実験2]字種クラスごとに輪郭画素比 b_{ij} を求め、ヒストグラムを作成する。

実験に先立って、以下の処理を行った。

(文字品質の平均化)

紙、墨など筆記条件や保存状態によって、筆写本の文字品質にはばらつきができており、これが識別率に影響することが考えられた。厚手の和紙に書かれた新出本の文字は、全体的に滲みやカスレにより輪郭線のディザが著しい。天理本の文字の輪郭線は、比較的滑らかであるものの、部分的にカスレやノイズの著しいところがある。素龍本は、輪郭線は滑らかで、ほとんど凹凸がないが、薄墨のためか本来あるべき連続線の一部が欠損して、判読困難な文字が含まれる。

筆写本ごとに、4近傍平均フィルターによる平滑化処理の回数を変えて、文字品質の平均化を行った。平滑

化処理の回数は、新出本9回、天理本5回、素龍本3回である。「あ」についてみると、平滑化処理を行わない場合の平均識別率は93.6%、平滑処理後の平均識別率は86.8%であった。

(識別次元の設定)

「あ」についてK次近傍識別の次元と識別率の関係を調べたところ、5次元100.0%、7次元98.2%、9次元98.2%、11次元91.6%であった。そこで、K次近傍識別の次元を11とした。

4.2 結果

字種ごとの筆写本の正識別率と分散を、3筆写本の平均識別率の大きい順に並び替えて表示したのが表2である。全体としては、新出本と素龍本の識別率が高く、天理本がやや低い。天理本の分散がやや高いところから、天理本ではバリエーションをもった文字遣いがなされ、一部が新出本や素龍本の文字に類似したものと見られる。

高識別率字種「を(遠)」、「れ(禮)」、「な(奈)」、「き(幾)」に共通するのは、字母の画数が多く、草書体のくずしに多様性があることである。逆に、低識別率字種「し」、「へ」、「つ」、「く」、「ら」は、1画のかな文字に単純化されて、形状にほとんどバリエーションがない字種である。平均識別率が90%を超えた4字種(「を」、「れ」、「に」、「な」)について、平均ベクトルとの距離が短い文字(平均近傍文字)上位7個を取り出して並べたのが図2である。高識別率の字種は、人間の目で確認できる形状の特徴をそなえており、筆者の個性を表わす個性字種として認知することができる。

5. 考察

人間の行う筆者認識は、書字動作を基にして形成された文字概念と現実の文字サンプルとの対応によって文字を認識する。筆者認識では、認識者が気づいた筆者の個性を強調する特異な形状の文字を特徴文字として認識し、文書全般にわたって、繰り返し現れる微妙であるが個人を特徴づける差異を看過する傾向がある。統計解析による筆者認識では、サンプル全体を対象として、平均的にみられる文字の構造特性、及び、変形のパターンを発見して、筆者識別を支援する。平均近傍文字の文字パターンを比較

表2: 実験1における正識別率を平均分散

Table.2: Correct identification rates and variance

かな	新出本		天理本		素龍本		識別率平均	分散平均
	識別率	分散	識別率	分散	識別率	分散		
そ	96.25	14.53	90.87	19.95	94.89	18.64	94.00	17.70
れ	90.00	19.53	85.85	23.68	100.00	17.18	91.95	20.13
にな	87.56	17.99	90.30	25.41	97.53	19.95	91.80	21.12
きな	91.85	17.36	88.89	22.44	93.64	20.90	91.46	20.23
きて	87.65	18.19	90.00	22.01	89.87	22.34	89.17	20.85
て	89.61	21.34	88.62	20.28	88.53	21.38	88.92	21.00
ひの	88.85	17.32	83.33	20.18	91.71	18.11	87.96	18.54
あ	87.45	20.63	84.21	23.82	90.69	22.04	87.45	22.16
さ	93.31	16.55	91.64	19.37	75.64	20.05	86.86	18.68
と	88.24	21.88	76.74	24.10	86.49	23.82	83.82	23.27
と	77.06	19.63	82.67	21.49	84.76	23.40	81.50	21.51
や	82.69	23.98	70.00	25.24	89.42	22.17	80.70	23.80
る	88.64	21.93	65.95	25.39	86.21	21.31	80.27	22.88
ふ	85.33	20.08	75.71	23.33	77.94	23.18	79.66	22.20
よ	82.54	17.28	78.33	19.92	75.81	19.71	78.89	18.97
た	81.30	22.27	74.80	25.01	74.78	23.88	76.96	23.72
り	81.01	21.11	71.57	23.76	77.07	20.86	76.55	21.91
も	76.19	22.29	73.43	23.41	79.89	24.66	76.44	23.45
い	66.10	28.12	70.49	27.57	85.71	23.11	74.10	26.27
は	93.64	16.86	75.79	17.66	52.78	20.65	74.07	18.39
ま	89.25	19.01	61.43	23.66	61.11	20.26	70.60	20.98
し	84.96	16.65	61.60	19.92	63.11	19.53	69.89	18.70
へ	79.63	25.24	59.02	34.33	60.78	31.94	66.48	30.51
か	90.45	22.76	40.24	25.15	76.85	25.21	65.85	24.37
こ	79.25	20.20	34.09	24.06	66.67	23.15	60.00	22.47
く	62.67	23.87	38.03	28.19	56.58	28.01	52.43	26.69
ら	65.62	22.34	33.33	27.80	46.27	23.51	48.41	24.55
ろ	42.86	22.32	28.00	23.06	38.52	21.50	36.46	22.29
全体	82.14	20.40	70.18	23.58	77.25	22.16	76.52	22.05

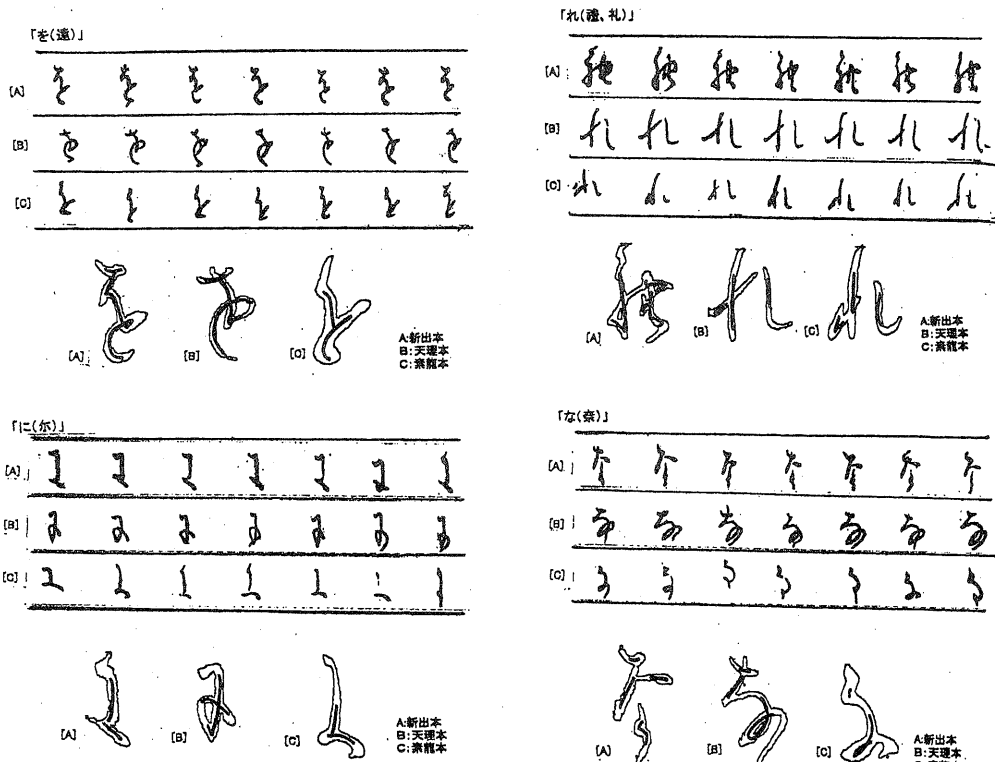


図2: 平均近傍文字(上位7個)

Fig.2 The nearest 7 characters neighboring by mean vector

して、高識別字種の特徴を分析してみることにする。

5.1 「を」の構造特性

分散値が低く、平均識別率が最も高い文字は、筆写本内において特徴的な字形の文字が、安定的に使用されていることを示している。

「を」について、筆写本ごとに平均近傍文字を比較してみると、新出本の「を」は、現代ひらがな「を」によく似た安定した構造であることがわかる。天理本「を」は、第3画が第2画の縦線に交差して内側に空白を含む弧型を形成する。この弧型が、天理本「を」の特徴となっている。素龍本「を」は、第1画の横線が第2画の縦線の中に一体化されて、存在が曖昧になる傾向がある。新出本、天理本、素龍本の「を」区別する筆記特性は、局所的で微妙な特徴ではあるが、それが安定して何度も繰り返されるため、筆者を識別する十分な文字特徴となる。これは、これまで、芭蕉研究者にも、ほとんど注目されることのなかった特徴である。

5.2 字母の変形パターン

「れ」、「に」、「な」の高い識別率は、平均近傍文字からも明らかのように、それぞれの筆写本における字母の変形パターンの差異に原因があることが考えられる。

「れ」は、字母「禮」「礼」の形状を留めた草書体系文字と現代ひらがな「れ」に類似した簡略体文字が併用される。各筆者本における「れ」の輪郭画素比ヒストグラムを図3に示した。新出本は、画数の多い字母「禮」の草書体系文字の使用が主であり、補足的に簡略字体「れ」が使用されていることがわかる。天理本の場合は、簡略字体「れ」とほぼ同様の割合で、草書体系文字も使用されていることが分かった(分散:23.68)。素龍本の場合は、扁と旁とが分離した簡略体系文字の使用がほとんどであると考えるとよい(分散:17.18)。同一字母における変形パターンの多様性も筆者識別の有力な手掛かりになると考えられる。

「に」は、字母「尔」から派生した変体がなく、現代ひらがなに対応する文字はない。画数が少なく、字画構成も単純なため個性の表れにくい文字とみられるが、筆写本に関しては、それぞれ独自の変形パターンがみられる。

「な」の字母「奈」は、くずしのバリエーションが豊富で、筆者の個性が表れやすい文字であり、特に、新出本と

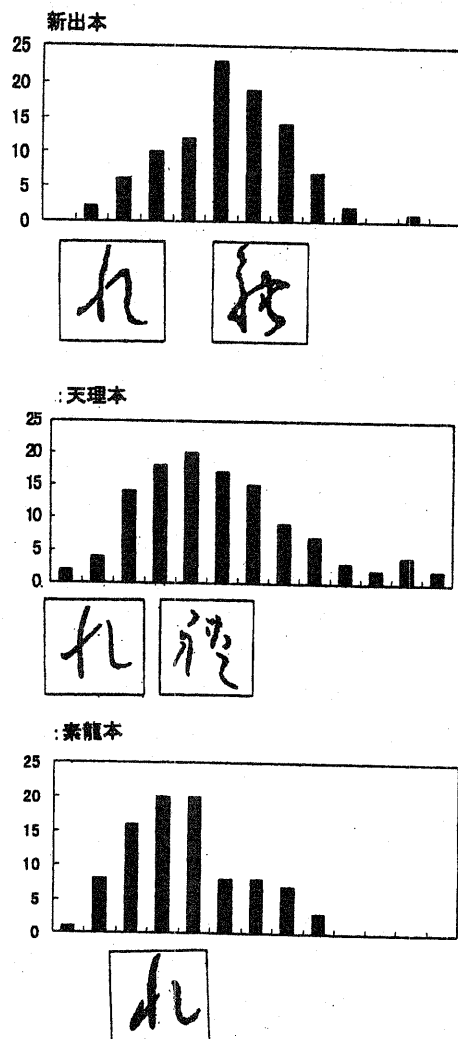


図3:「れ」の輪郭画素比のヒストグラム
横軸は輪郭画素比.02から0.48の間を14等分したもの
Fig 3: Histogram of rate outline pixel to character image

素龍本の変形パターンは安定的で識別率が高い(新出本:91.85 素龍本:93.64)。

5.3 筆者照会

芭蕉直筆とされる三日月日記から、筆者認識の個性文字と認識される「を」、「れ」、「な」の全文字を抽出し、表示したのが図4である。比較のため、矩形内に、新出本の平均最近傍文字を示した。三日月日記には、新出本には出現しないタイプの文字も含まれており、この比

較から結論を導くことは難しいが、三日月日記の「を」、
「れ」、「な」は、新出本の文字特性をよく共有しているこ
とがわかる。

6. まとめ

本研究は、芭蕉文学研究用データベースに収納され
た「奥の細道」の筆写本3部を研究対象にし、文字
画像データに計算機による統計解析を適応して、筆
写本の筆者認識を試みた。文字の特徴量抽出には6
4次元加重方向指数ヒストグラム法を用い、筆者を
特徴づける個性字種の識別には、K次近傍識別則を
用いた。本研究において発見された「を」の微妙な構造
特性、「れ」、「な」に見られる変形パターンなどの筆記
特性は、芭蕉研究者の間でもあまり言及されることな
かったものである。おそらく、計算機を使用しなければ、
気づかれることがなかった筆記特徴と思われる。とこ
ろで、これらの筆記特性が、筆者の個人的要因によるも
のであるのか、あるいは、筆写における何らかの制作
意図によって作為的に作り出されたものであるのかは、
筆者認識の根幹にかかわる興味深い問題である。芭蕉
研究者にフィードバックして今後の課題とすべきである。

参考文献

- 1) 吉村ミツ、木村文隆、吉村功: 筆者識別に影響する要因の分析: 信学論(D), vol.66-D, No.1, pp.1-8, (1983)
- 2) 田口英郎: 書字による個人識別の技術, システム/制御/情報, Vol.35, No.7, pp.398-407, 1991
- 3) 吉村ミツ、吉村功: 類似パターンの頻度による筆者識別: 信学論 (D-II), vol.J72-D, no.12, pp.2051-2060, (1991)
- 4) 吉田公一、小柴良介、倉内秀文: 筆跡の測定と計算法—コンピ ュータによる筆跡測定, 計量国語学, Vol.18, No.4, pp.189-197, (1992)
- 5) 吉村ミツ、吉村功: 旅行小切手上の日本人署名の照合に関する 検討, 信学論(D-II), vol.J80, No.7, pp.1764-1773, (1997)
- 6) 吉村ミツ、吉村功: 筆者認識研究の現段階と今後の動向, 信学技報 PRMU96-48, 1996
- 7) 大谷徳蔵 (編)、芭蕉全図譜 / 松尾芭蕉 [著] 芭蕉全図譜刊行会, 岩波書店, p.599(1993)
- 8) 桜井武次郎: 影印おくのほそ道, 双文社出版, p.107, (1992)
- 9) 村松友次: 曾良本「おくのほそ道」の研究, 笠間書院, p.395 (1988)
- 10) 上野洋三、桜井武次郎: 芭蕉自筆「奥の細道」, 岩波書店, p.159(1997)
- 11) 井出正弘、山田新一、藤川英司: ストローク解析による筆者認 識, 情報処理学会全国大会論文集, Vol.45, No.2, pp.287-297 (1992)
- 12) 吉村ミツ、吉村功: 局所円弧パターン法を用いた筆者識別: 信学 論 (D-II), vol.J74-D, No.2, pp.230-238, (1991)
- 13) 尺長健、金子博、淀川英司: 2次元統計量の線分スペクトル分 解による手書き文字の筆者認識, 信学論(D), vol.J67, No.7, pp.776- 783, (1984)
- 14) 鶴岡信治、栗田昌徳、原田智夫、木村文隆、三宅康二: 加重方 向指数ヒストグラム法による手書き漢字およびひらがな文字認識, 信学論(D), vol.J70-D, No.7, pp.1390-1397, (1987)
- 15) 鳥脇淳一郎: 認識工学—パターン認識とその応用—, p.140, コロ ナ社 (1991)

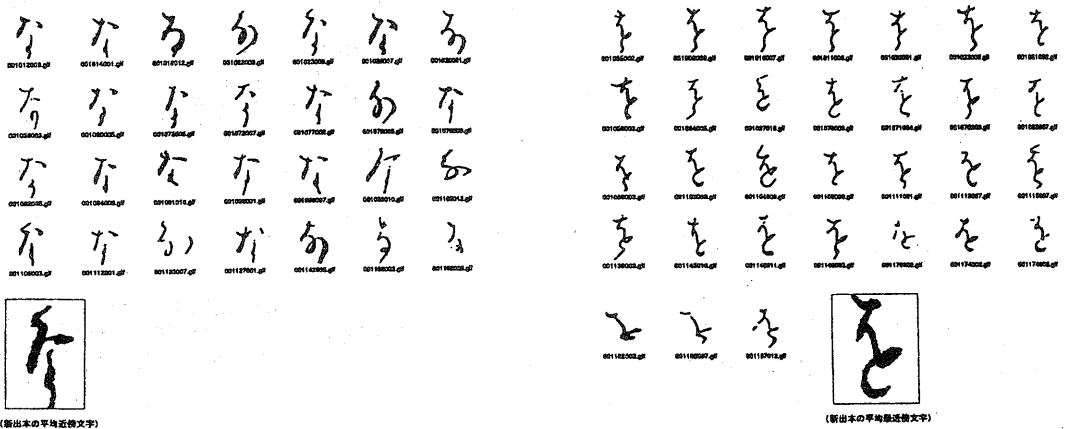


図4: 三日月日記における個性文字
Fig.4: The Characters express the personality
in the "Mikazuki Diary" which was written by Bashou.

