

中世英文学資料の認識および解析

— 意義・問題点および予備検討結果 —

山口 晃典[†] 内田 誠一[†] 千葉 淳一^{††} 飯田 周作^{††} 植竹 朋文^{††}
松下 知紀^{††}

[†]九州大学 大学院システム情報科学研究所 〒819-0385 福岡市西区元岡 744
^{††}専修大学 文学部・社会知性開発研究センター 〒214-0033 川崎市多摩区東三田 2-1-1
E-mail: †yamaguch@human.is.kyushu-u.ac.jp

あらまし 「Anglo-Saxon 語の継承と変容」プロジェクト(専修大学)では、1200-1600年代に作成された手稿写本や初期印刷本といった中世英文学資料のデジタルテキスト化を目指し、そのためのOCRの開発を行っている。本報告では、手稿写本に比べれば比較的扱いやすい初期印刷本(1500年前後の活字印刷本)について、それら文書の様々な性質を論じる。また「農夫ピアズ」初期印刷本の孤立文字データを用いて、小規模ながら実際に文字認識実験を行った結果について論ずる。

キーワード 中世英文学, 初期印刷本, OCR

Recognition and Analysis of English Historical Documents

— Purpose, Problems, and Preliminary Study —

Akinori YAMAGUCHI[†], Seiichi UCHIDA[†], Junichi CHIBA^{††}, Shusaku IIDA^{††}, Tomofumi
UETAKE^{††}, and Tomonori MATSUSHITA^{††}

[†] Kyushu University 744 Motoooka, Nishi-ku, Fukuoka-shi, 819-0385 Japan
^{††} Senshu University 2-1-1 Higashi-mita, Tama-ku, Kawasaki-shi, 214-0033 Japan
E-mail: †yamaguch@human.is.kyushu-u.ac.jp

Abstract One of the goals of the project “The Development of Anglo-Saxon Language and Linguistic Universals” (organized by Senshu University, Japan) is to develop an OCR system for medieval English manuscripts, that is, historical handwritten English documents. In this report, we will discuss characteristics of medieval English printed documents as a preliminary study towards the above goal. In addition, a preliminary recognition experiment was conducted on a small-scale character set from a medieval printed document, called “Pierce Plowman” (printed in 1550).

Key words Medieval English document, OCR

1. はじめに

様々な歴史的資産がデジタルアーカイブとして保存されている。歴史的な文書も例外ではない。単にスキャン処理によりデジタル画像化するだけでなく、その記述内容についても機械認識を介してデジタルテキスト化し、解析・利用する試みがなされている。国内に限っても、古文書翻刻支援システム開発プロジェクト[1], [2] や、木簡解読に関する検討[3], 毛筆文書認識[4] が挙げられる。古文書OCRに関する特集[5]も参考になる。海

外でも様々な検討が為されており、それらの一部は学術雑誌の特集号としてまとめられるほどになっている[6]。また比較的近代ではあるが、数学文書のretro-digitization[7]もその一種として挙げられる。

専修大学「Anglo-Saxon 語の継承と変容」は、中世英文学資料のデジタルテキスト化を主目的として、平成17年度に文部科学省私立大学学術高度化推進事業オープン・リサーチ・センター整備事業に選定された研究プロジェクトである[8]。具体的には、中世英文学資料について、それらの機械認識を介した

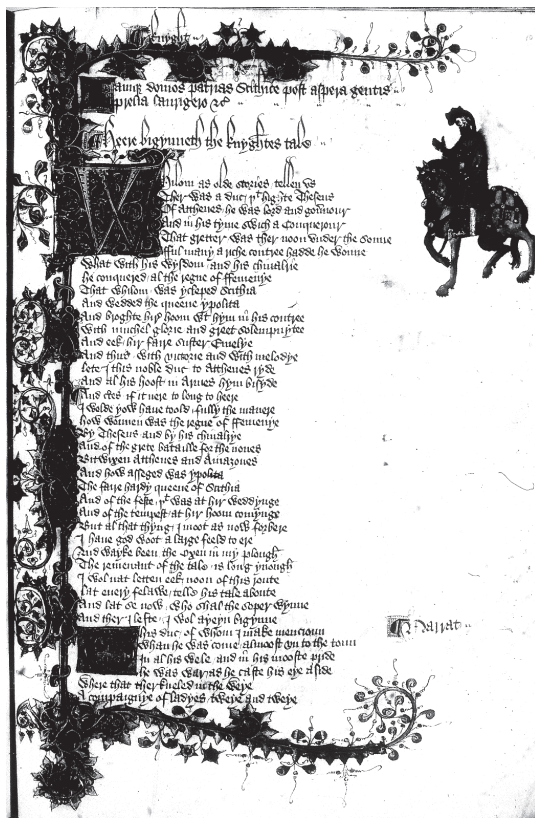


図1 手稿写本「エレスミア写本」

デジタルテキスト化を第一の目的としている。さらに第二の目的として、デジタルテキスト化した後に、各資料の様々なバージョン（後述）を相互比較することで、それらの差異解析および系統解析、ならびに制作年代および制作方言（制作地域）を推定することを目指している。最終的にはインターネットによる資料および解析結果の一般公開も考えている。

中世英文学資料には、人の手によってペンで書き写された「手稿写本」と、活字を用いて印刷した「初期印刷本」の2種類がある。手稿写本とは、活字印刷以前の時代に、一冊一冊手書きで作られた書籍である^(注1)。従って、意図的な編集だけでなく写本時のミスなどによって、写本のバージョン毎に内容が異なる。一方、初期印刷本とは、ゲーテンベルグの発明以来欧州中に広がっていった中世の活字印刷本である。これも版毎に多少の差異が認められる。このように中世英文学資料については、バージョン毎に細かな差異がある。こうした差異の伝播の様子を推定・解析することで、その作品全体に関する系統樹が推定できる。

手稿写本の例として「エレスミア写本」を図1に示す。エレスミア写本は丁寧な筆跡で書かれ、また保存状態も非常に良いため、多くの中世英文学者が解析を行っている。こうした手稿写本の機械認識の難しさについては、文献[9]にまとめられて

(注1): 手稿写本では、各章冒頭の装飾文字が、蔦模様を伴いながら赤/青/金色などで美しく彩色されていることが多い。また、各行末の余白部には、赤や青の細長い模様があしらわれることもある。さらに図1のようにページ周辺にも様々な絵や模様が描かれることもある。手稿写本では、これら絵模様も文字もすべて手書きで描かれたものであり、絵描きや写字生のコラボによる偉大な芸術作品として鑑賞しうものが多い。

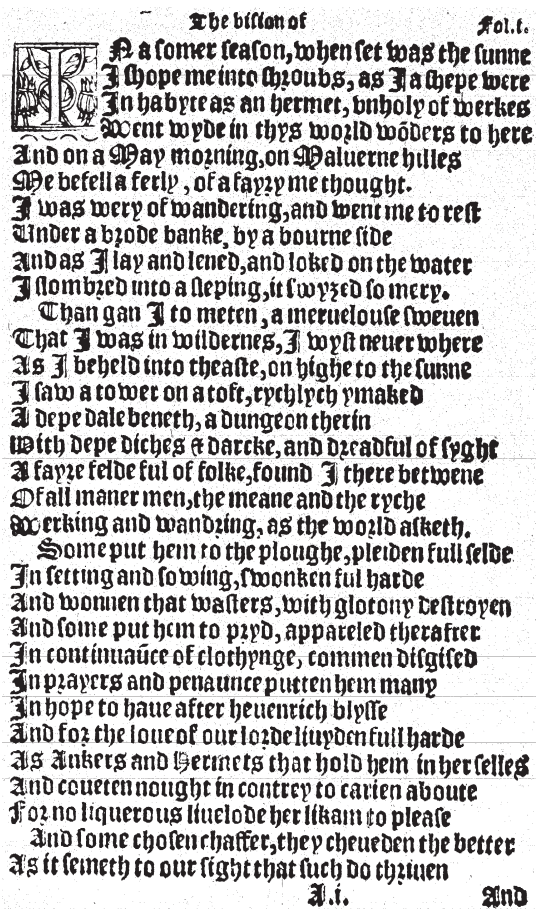


図2 初期印刷本「農夫ピアズ」

いる。

初期印刷本の例として「農夫ピアズ」を図2に示す。農夫ピアズについても、様々なバージョンがあり、Aテキスト(1362年頃)、Bテキスト(1377年頃)、Cテキスト(1399年頃)という3種類の改訂版が知られている。図2はBテキストに基づく1550年の活字印刷本である。農夫ピアズについては、A、B、Cそれぞれ20、17、31種類の手稿写本も知られている。

以下本報告では、本プロジェクトの課題を明確にするために、農夫ピアズのBテキスト初期印刷本(1550年)を中心的題材として、その特徴的性質を概観する。さらにそこから切り出した単文字パターンについて、予備的に行った認識実験の結果について述べる。

2. 初期印刷本の特徴

2.1 レイアウト

手稿写本の行切り出しや文字切り出しは困難であるが、初期印刷本では活字を用いているため、原理的には水平・垂直方向の空白を手がかりにこれら切り出しが可能である。しかし、後述のように合字やかすれ・つぶれ、紙面ノイズの影響があり、実際にはそう単純な問題ではない。

行切り出しの際には、各章の冒頭にある大型の装飾文字に配慮する必要がある。図2冒頭部にもその装飾文字の一例が見られる。また図3は農夫ピアズ中の装飾文字の一部である。

韻文(詩)では、レイアウトが独特なものとなる。脚韻や頭韻



図3 各章の冒頭にある飾り文字 (“A”, “H”, “I”, “T”, “W”, “Y”)

And hangen it aboute the cattes halfe, then here wee
 we hecher he rit oz rest, oz runne to pleye (moune ←
 And if hfm lift for to lake, than loke we might
 That the pouertie of the parish, should haue if thyne
 Perlös & her priekles : pleynd he to þ bishop (were ←
 That her parishes were poze, siche þ pestilēce : ime

図4 次行右端余白への回り込み印字

といった規則に基づく“まとまり”があり、複数のまとまりを1行内に入れることはない。このため、次のまとまりが来ると行末に余白を入れて改行する。行頭はすべて大文字で始まり、また行末が不ぞろいになる。(これは日本における詩のレイアウトでも同様である。) 図2の農夫ピアズもこの規則に従ってレイアウトされている。

このまとまりが1行幅以上になった場合、余白を避けるべく、次行右端余白に回り込んで印字されることがある。図4はその例である。こうした余白の節約のために、後述の上付き文字や省略記号も用いられる。それでも省略しきれない場合には、文章が改変されることもある(それがその後の版に伝搬していくことになる。)

2.2 印刷媒体

紙面に現れる裏写りや印刷時の汚れ、紙の質の劣化によるノイズは、OCRの様々なステップにおいて困難な問題を生じる。保存状態によりこれら劣化の程度は大きく異なるため、適応的なノイズ除去処理が必要になるであろう。なお、印刷媒体としては紙だけでなく羊皮紙も用いられる。

2.3 綴り字

初期印刷本だけでなく手稿写本にも言えることであるが、辞書を用いた文字認識候補の絞りこみを行う際には、いくつかの注意が必要である。例えば、中世の時点では綴り字が安定していない。また、単語中で“u”を意図的に“o”と書き、書写性と可読性を狙った記述をすることもある [9]。さらに植字工の不注意によるスペルミスなどを含んでいる可能性も高い。(そしてそれがその後の版に伝播していくことになる。)

2.4 認識問題としての特殊性

中世文学資料研究のひとつの特殊性は、資料毎に多様な字体を持つという点である。この意味で手書き文字認識タスクとも似ているが、得られるデータ数が極めて限られている点で異なる。すなわち認識タスクが1冊子内で閉じており、例えば一般的な辞書データの作成は困難であろう。よって中世文学資料用のOCRにおいては、各資料毎に冒頭の数ページをラベル付けすることで、残りのページが認識できるような仕組みが効率、精度の面で有効であろう。



図5 「農夫ピアズ」の文字 “n”



図6 「農夫ピアズ」の文字 “A”



図7 「農夫ピアズ」における “r” と “s” の異体字



図8 合字活字の例

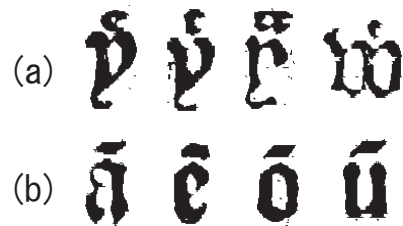


図9 省略のための特殊な表現

3. 文字形状

本節では「農夫ピアズ」初期印刷本(1550)を例として、中世英文学資料における文字パターンの性質について述べる。

3.1 字形変動

図5に農夫ピアズから抽出した“n”を示す。活字印刷本とはいえ、印刷時のインク塗布量や圧力、活字の磨耗や微妙な個体差があり、さらに紙面と2値化の影響もあり、様々な変動を起こしている。こうした変動については、4.3節で統計的に解析する。

3.2 カリグラフィ的文字

大文字にはしばしばカリグラフィ的な装飾が施される．図 6 に農夫ピアズから抽出した“A”を示す．カリグラフィを構成する非常に細い線は二値化により容易に掠れや他ストロークとの接合を起こす．

3.3 異体字

図 7 に示すように，文字“s”や“r”には異体字が存在する．こうした異体字の使い分けは，行の位置や読む時のリズム形成の規則に因ることもあるが[10]，文書によっては規則性がないものもある．

3.4 合字

活字間のスペースを不自然としないために，様々な合字活字が利用されている．ゲーテンベルク活字[11]やキャクストンによる初期印刷本で用いられたタイプ 2 活字[10]にも多くの種類のもが見られる．図 8 は合字活字の幾つかの例である．一見，接触していないもの（一つの連結成分になっていないもの）もあるが，垂直線で分離できないことから合字活字とわかる．

3.5 余白節約のための特殊文字

一定の行内に，適切な単語や文境界が来ないことがある．この場合，2.1 節で述べたように行末に余白を残す以外にも，前述のように次の行後半部に回り込む方法がある．さらに以下のように特殊文字を用いる方法が二つある．なお，これらの方法の使い分けは植字工に委ねられており，従ってその使用頻度などで植字工を区別できる可能性がある．

まず，図 9(a) のように，小さな上付き文字を用いる方法がある．図中の y 状の文字は thorn と呼ばれる“th”に相当する文字で，上に“e”，“t”，“u”を伴うことでそれぞれ“the”，“that”，“thou”と読まれる．また“w”は“t”を伴って“with”と読まれる．なお，これらの上文字付き文字は，それらの組で一つの活字として準備されている．

もう一つは，図 9(b) のように母音の上にアクセントをつけることで後続の“m”や“n”を略する方法である．これらアクセント付き文字も，一つの活字として準備されている．

4. 予備の実験

本節では中世英文学資料に対して予備的に行った実験について述べる．

4.1 実験試料

実験対象として農夫ピアズ，テキスト B の初期印刷本（全 234 ページ）のマイクロフィルムをスキャナにてデジタル画像化したものを利用した．このうち冒頭 4 ページから手作業で切り出した文字画像 44 カテゴリ，全 2918 文字を実験試料として用いた．1 カテゴリ当たりの文字数は様々であり，多いもので約 400 サンプル，少ないものでは 1 サンプルであった．4 ページ分と限られたデータであるため，本来カテゴリとしては存在するが，今回のデータセットには含まれないものも多数存在する．

ほとんどのカテゴリはアルファベット 1 文字に相当する．これら以外に，記号カテゴリが 1 個（“&”），続き字カテゴリが 3 個（“sh”，“sl”，“st”）存在した．また，異字体のために 1 アルファベットにつき 2 個カテゴリを用意したもの（“r”，“s”）が存

表 1 認識率 (%) .

	single	1/8	1/4	1/3	1/2	2/3	3/4	loo
16×16	80.5	95.5	96.2	96.3	97.2	97.0	97.3	97.3
32×32	84.0	96.6	97.1	97.2	97.7	97.6	97.7	97.8
64×64	82.1	96.2	96.8	96.7	97.0	97.1	97.2	97.1

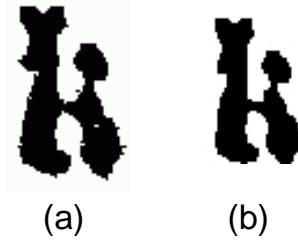


図 10 前処理．(a) 原画像．(b) 前処理後．

在した(図 7)．

サイズは文字ごとに様々であるが，縦横 70～100 画素に収まるものが多かった．

4.2 認識実験

4.2.1 実験概要

実験試料に対して認識実験を行った．特徴として文字の輪郭線の方向成分を用いた．各テストパターンについてすべての辞書パターンとの間に正規化相関を求め，最も相関の大きい辞書パターンのカテゴリを認識結果とした．

4.2.2 前処理

特徴抽出の前に画像に対して前処理を行った．まず実験試料の文字画像に対して膨張・収縮処理を行うことで文字の途切れや掠れを補正した．続いて収縮・膨張処理を行うことで雑音を除去した．

さらに以上の処理で得られた文字画像に対して大きさ正規化を施した．具体的には，画像中の黒画素の水平方向と垂直方向の長さを比較し，長い方の長さが全画像で一定になるように画像全体を拡大・縮小させた．このことにより，元の画像の形状を崩さないまま大きさだけを正規化できる．また，短い方の余白が均等になるように画像中央に配置した．図 10 に前処理前と前処理後の画像の例を示す．前処理の段階では，画像の大きさが縦横 64 画素になるように大きさ正規化を行った．

4.2.3 特徴抽出

前処理後の文字画像から以下の手順で方向特徴を抽出した．まず文字輪郭を抽出した後，輪郭線上の各点について，同一輪郭線上の前後の点を結んだ直線と水平方向との角度を求めた．その角度を 4 方向に量子化することで，各点における各方向成分を求めた．その後，各方向成分についてガウス関数によるぼけ変換を行った．図 11 はその抽出結果例である．

さらに画像を 2 通りに縮小して実験を行った．すなわちサイズをそのまま (64×64)，半分に縮小 (32×32)，1/4 に縮小 (16×16) したものの 3 種類を扱った．これら縮小処理により 4 画素もしくは 16 画素分の方向特徴が 1 画素に集積されることになる．なお縮小処理はぼけ変換の直前に行った．画像サイズを ($N \times N$) とすると，各パターンは結局 $4N^2$ 次元ベクトルで

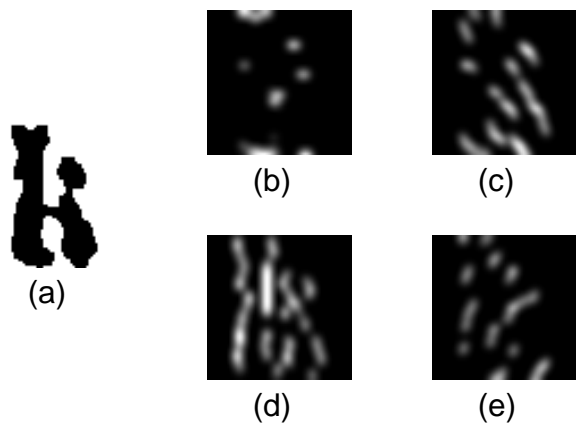


図 11 方向特徴 . (a) 対象画像 . (b) 水平方向成分 . (c) 右下がり方向成分 . (d) 垂直方向成分 . (e) 右上がり方向成分 .

表現されることになる .

4.2.4 認識率

抽出した特徴を用いて認識実験を行った . 識別には正規化相関を用い , 各テストパターンについて最も相関の高い辞書パターンのカテゴリを認識結果とした . また , 交差確認法を用いて辞書パターン数を変化させそのときの認識率を調べた .

表 1 に認識率を示す . 同表の列方向は用いた画像のサイズ (単位 : 画素) である . 行方向については , “single” が各カテゴリ 1 辞書パターンとした場合 , “loo” が leave-one-out , “1/8” が全データの 1/8 を辞書パターンにして残りの 7/8 をテストパターンにした場合の交差確認法による結果である .

サイズ別に見ると半分に縮小したもの ($N = 32$) が最も認識率が高くなった . これは , 方向特徴を集積することで文字の太さの違いや位置ずれなどが多少吸収されたためと考えられる . 1/4 に縮小した場合 ($N = 16$) では認識率が低下しているが , これは縮小しすぎたために逆に他のカテゴリとの混同が起きたためと考えられる .

辞書パターンを全体の 1/8 しか用いなかった場合でも約 96% の認識率が得られた . これは一部の文字のだけを辞書パターンとして用いることで高い認識率が得られるということであり , 冒頭の数ページだけを辞書パターンとして登録することで資料全体を認識できる可能性があることを示している . ただし , 冒頭 4 ページには含まれないような文字カテゴリが多数存在するので , これらは別途収集する必要がある .

4.2.5 主な誤認識

辞書パターンを 1/2 としたときの主な誤認識の例を図 12 に示す . このように非常によく似たカテゴリの存在が誤認識を誘発している . さらに , “e”-“c” や “u”-“n” の例のように , 2 値化により文字の一部が欠損・変形し , 誤認識となったものもあった . 今回の交差確認法による評価においては単純にサンプルを分割して利用したので , こうした欠損パターンが辞書パターンに含まれることもあり , それらが最近傍パターンとして選択されることで多数の誤認識を起こしていた .

4.2.6 課題

認識実験の結果 , 一部の非常によく似たカテゴリの存在が誤

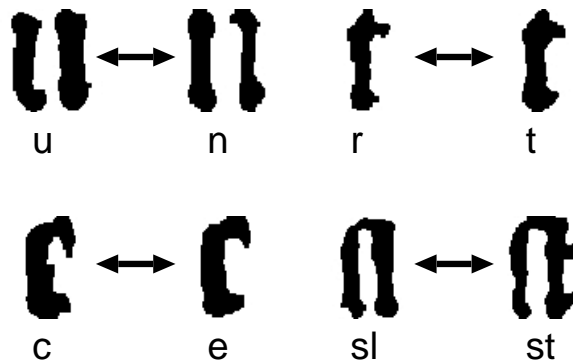


図 12 主な誤認識の例

認識を誘発していることがわかった . また , 2 値化の影響により形状が変形し誤認識となった例も存在した . これらに対処するのは非常に困難ではあるが , 以下に述べる形状変動のモデル化が有効である可能性もある . また , 前処理やノイズの影響を強くうけたパターンが辞書パターンに含まれないようにする処理が必要となる . さらに , 2 値化処理を施す前の濃淡画像データを用いることを検討してもよいと思われる .

4.3 形状変動の傾向

4.3.1 目的

前述のように , 活字印刷本とはいえ同一カテゴリの文字でも様々な形状のものが存在している . これらの変動要因としては , 印刷時のインク塗布量や圧力 , 活字の磨耗や個体差 , 紙面と 2 値化の影響があろう . 従ってその多くはノイズ的であり , 特定の傾向はない可能性もあるが , 細い部分の欠損や , 過剰インクによる屈曲点の丸まりなどの一定した傾向もあろう . もし一定した傾向がわかれば , その傾向をカテゴリ毎に収集することでカテゴリ間の分離に利用できると考えられる .

そこで , active shape model (ASM) [12] を用いて , 各カテゴリの固有変形情報を得る実験を行った . ASM とは統計的な変形モデルの一種である . 用意した学習パターンから得られる変形情報に対して主成分分析を行い , それによって得られる固有変形により基準パターンからの形状変動特性を表現する . なお , オリジナルの ASM では手動で変形情報を与えていたが , 本報告では後述のようにマッチング技術を利用することで , 基準パターンに対する各文字の相対的な変形を自動的に抽出し , それを用いて ASM を求めた .

4.3.2 前処理

各文字画像サンプルに対する前処理は前節と同様である . 前処理後のサンプルのうち , 途切れや断片化等により位相変化を起こしていないサンプルを手動で選別した . 前述のように多くの文字には掠れなどが生じているため , 十分なサンプルが得られないカテゴリも多かった . 以下の実験ではそれでも相当量のサンプルを確保できた 2 カテゴリ (“e” , “t”) を対象とした .

4.3.3 輪郭線マッチングによる変形抽出

各文字サンプルの変形をその最外郭輪郭線上の各点の変位として表現する . ただし変位は基準パターンの輪郭線に対する相対的変位とする . この相対的変位を求めるために , まず各サンプルについて最外郭輪郭を抽出した . 次に , この輪郭と基準パ

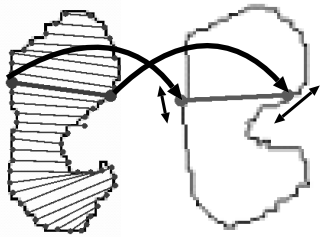


図 13 変形抽出のための輪郭マッチング

ターンの輪郭との間に最適マッチングを求め、輪郭線を構成する点間に対応関係を定めた。輪郭線は閉曲線であるために、1次元的なマッチングは適用できない。そこで図 13 のような手法を用いた [13]。すなわち、基準パターンの輪郭線上のある 1 点から時計回り・反時計回りに 1 個ずつ点を選んでいき、それら 2 点を組として輪郭線のスライス列を考える。各スライス毎に順次その両端を入力サンプルの輪郭上に対応付ける問題を考えると、これは動的計画法に基づくアルゴリズムで効率よく解け、輪郭線全体として最適なマッチングを求めることができる。

各カテゴリの基準パターンの選定法は幾つか考えられる。今回の実験では、以下の手法を用いた。すなわち、あるカテゴリの総サンプル数を N とすると、まず各サンプルについて、他のすべての $(N - 1)$ サンプルとそれぞれ輪郭線マッチングを行い、得られた $(N - 1)$ 個のコスト (変位量) の総和を求める。全 N サンプルのうち、この総和が最小となるものを、変形の意味で分布重心にあるとして、基準パターンとした。

4.3.4 主成分分析による固有変形抽出

基準パターンに対する各文字の相対的な変形を収集した。次いでそれらに対して主成分分析を行い、主成分ベクトルとしてカテゴリ内の固有変形情報を得た。

4.3.5 結果

2 カテゴリ (“e”, “t”) について得られた固有変形を図 14 に示す。主成分分析によって得られた固有変形のうち、第 1 主成分から第 5 主成分までを示している。対象とした文字が活字であるため、あまり大きな変形は見られない。一方で、印字条件の変動が抽出できていることがわかる。例えば “e” の第 1 主成分および “t” の第 1, 2 主成分がインク塗布量の違いによる変形、“e” の第 2, 3 主成分および “t” の第 3, 4 主成分が活字の凹領域のインク溜りによる変形であると考えられる。ただしそれより高次の主成分はノイズ的な変形のように見える。

4.3.6 課題

文字画像の最外郭輪郭について、基準パターンから見たカテゴリ内の固有変形を求めた。今回は十分なサンプル数を確保できた 2 カテゴリについて実験を行ったが、サンプル数が少ないカテゴリに対してはこの手法では固有変形情報を求めることはできない。また、今回用いたデータは 2 値化が施されているため、文字の一部が欠損・変形したものが多く、抽出された固有変形に少なからず影響を与えている。これを初期印刷本独特の変形と見るか単なるランダムノイズと見るかは今後の検討が必要であろう。2 値化を施さず、濃淡値レベルで直接変動を扱う

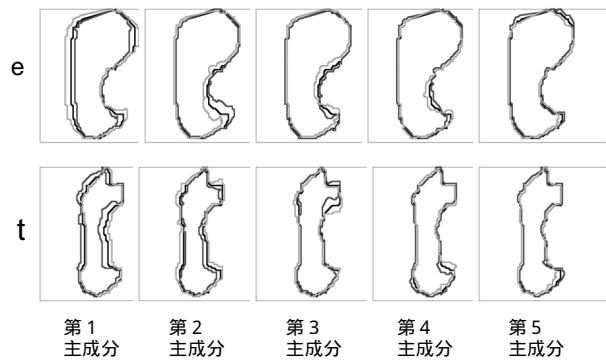


図 14 “e” と “t” の固有変形成分

方法も考えられる。

5. ま と め

本報告では中世英文学資料のデジタルテキスト化を目指し、まず初期印刷本について文章や文字形状等の様々な性質について論じた。例えば、デジタルテキスト化の際に問題となりうる、現在の文章では見られないような様々な特徴が存在することを述べた。また、様々な文字形状や特殊な文字など、単純に文字認識技術を適用するだけでは解決できないような問題が存在する。さらに、対象が歴史的に古いものであるため、紙面の劣化などが認識の際の障壁となりうる。

文字形状に関しては資料毎に多様な字体を持つため、各資料毎に冒頭数ページをラベル付けすることで残りのページを認識できるような仕組みが求められる。今回小規模ながら実際に文字認識実験を行った結果、96~98%程度の認識率が得られており、こうした冒頭データからの全体認識の可能性が示唆された。

文 献

- [1] 山田奨治, 他, “古文書翻刻支援システム開発プロジェクト報告: (1) —プロジェクト概要—,” 人文科学とコンピュータ研究会 45-1, 2000.
- [2] 山田奨治, 他, “古文書翻刻支援システム開発 (HCR) プロジェクト報告 (2),” 人文科学とコンピュータ研究会 50-2, 2001.
- [3] 蜂谷大翼, 末代誠仁, 齋藤 恵, 中川正樹, 馬場 基, 渡邊晃宏, “欠損を含む文字パターンを対象とした文字認識手法の試作,” 信学技報, PRMU2004-260, 2005.
- [4] 寺沢恵吾, 長崎 健, 川嶋稔夫, “文字切出しによらない毛筆手書き文字検索のための部分空間法,” 信学技報, PRMU2004-172, 2005.
- [5] 特集 挑戦 古文書 OCR, 人文学と情報処理, no. 18, 1998.
- [6] Special issue on the analysis of historical documents, Int. J. Document Analysis and Recognition, vol. 9, no. 2-4, 2007.
- [7] G. O. Michler, “Report on the retrodigitization project,” Archiv der Mathematik, vol. 77, pp. 116-128, 2001.
- [8] <http://anglo.fm.senshu-u.ac.jp/>
- [9] 松下知紀, 千葉淳一, 植竹朋文, “中世英語写本文字認識の問題点について,” 人文科学とコンピュータシンポジウム, 2007.
- [10] ロッテ・ヘリンガ著, 高宮利行訳, キヤクストン印刷の謎 — イングランドの印刷事始め, 雄松堂出版, 1991.
- [11] 凸版印刷印刷博物館編纂委員会, 印刷博物館誌, 凸版印刷, 2001.
- [12] T. F. Cootes, D. Cooper, C. J. Taylor, and J. Graham, “Active shape models - their training and application,” Comput. Vis. Image Und., vol. 61, no. 1, pp. 38-59, 1995.
- [13] 内田誠一, “DP マッチング概説 ~ 基本と様々な拡張 ~,” 信学技報, PRMU2006-166, 2006.