

文字列をなぞることによる古地図の探索

渡邊 悠一¹ 寺沢 憲吾¹ 角 康之¹

概要：古地図は、歴史的背景を知るうえで重要な手掛かりの一つである。そのため、地図から文字領域を抽出することができれば地図同士を対応付けることができ、地図を比較することで新たな発見につながる。しかし、古地図は文字の向きがバラバラで様々な向きの文字列が書かれている。さらに縦書きや横書きの文字列が含まれているため、機械的に光学文字認識 (OCR) を適用するのは難しい。そこで本研究ではユーザが文字の向きに沿ってなぞることで文字を切り出しマッチングを行うシステムを提案する。手法としては、ユーザが地図中の文字列を文字の向きに沿ってなぞり、なぞった線上に存在するラベリングされた固まりを文字をみなすことで文字列を切り出す。マッチングは切り出した文字をスリット状に切り出し、その勾配分布特徴量を用いて行う。また本研究では、ユーザがなぞった文字をデータベースに追加するシステム構成とする。そうすることでなぞった文字は新たな検索対象となり、データベースの拡張につながる。なぞった文字というのはユーザの興味の対象であり、それらを登録していくことでデータベースは集合知としての価値を持つ。本稿では提案したシステムを用いて古地図から読み取れたことについての利用例をシナリオで示す。

Exploring Old Maps by Finger Tracing Characters

YUICHI WATANABE¹ KENGO TERASAWA¹ YASUYUKI SUMI¹

1. はじめに

古地図は、地理的情報だけではなく、統治者と市民との関係などの歴史的背景を知るうえで重要な手掛かりの一つである。そのため、複数の地図を探索することで、その年代や地域に応じた著者の意図や世界観の違いを明らかにすることができる。

近年、情報技術の発展により歴史資料がデジタルアーカイブ化がされてきた。そして、デジタル化された多くの古地図が Rumsay のサイト [1] や国立国会図書館のデジタルコレクション [2] のように一般公開されるようになった。古地図を研究者だけではなく、一般の人でも閲覧できるようになったことで古地図を用いて街歩きするなど日常生活の中で楽しむことができるようになってきた。しかし、デジタル化された古地図の多くは画像で保存されているため、そのままでは地図中の文字列を検索することはできない。このような問題に対処する方法としては、文書のように古地図に対しても光学文字認識 (OCR : Optical character

recognition) を適用する方法が考えられる。しかし、地図は図や文字、記号などが入り混じっており、文書のように文字認識させるのは難しい。さらに古地図は、文字の向きがバラバラで様々な向きの文字列が書かれている。また、日本語は英語と違い縦書きと横書きがあり、そのどちらもが古地図には書かれている。そのため、古地図に対して機械的に OCR を適用するのは難しい。

そこで本研究では、文字をなぞることを手掛かりとして、文字領域を抽出するシステムを提案する。複雑な文書を読むときに指で文字をなぞることは、人間にとっては簡単であると同時に、それはコンピュータが文字の向きや文字領域、ユーザの興味の対象を認識する上で重要な手掛かりとなる。

以下では、ユーザが文字をなぞることで文字領域を抽出し、それと類似した画像を持つ地図を探索するシステムの詳細について述べる。利用例としてユーザが地図中の文字をなぞることによって複数の古地図間の探索を可能とする例について示す。

¹ 公立はこだて未来大学

2. 関連研究

2.1 文書画像からの全文検索

デジタルアーカイブの多くは画像として保存されている。そのため、そのままでは必要な文字列を検索するのは難しい。このような問題に対応する方法として OCR を適用するなどし、テキスト化する手法が考えられるが、文書がくずし字など特有の文字で書かれていたりと様々な要因によってそれは難しい。こうした背景から文字認識をせずに画像のまま文字列検索を行うワードスポッティングという手法が Manmatha らによって提案されている [3]。そして、そのワードスポッティングに適した特徴量 [4] が提案されている。Manmatha らは、文字列を単語ごとに切り出すことでマッチングするワードスポッティングであったが、こうした単語切り出しを行わずにマッチングする手法なども研究されている [5] [6]。

Manmatha らは英語の手書き文書を対象としたのに対し、日本語の手書き文書を対象としたワードスポッティングが Terasawa らによって提案されている [7]。この手法は文書に対し、行の切り出しを行い、行をスリット状に切り出し、各スリットの特徴量を用いてワードスポッティングを行う。そして、この手法をもとに新たな特徴量を検討し、線幅の変化や微妙な位置ずれに頑健であるという特徴を持った勾配分布特徴量が提案されている [8]。この手法では、従来手法よりも高い精度であることが確認された。古地図も手書きで書かれているものが多いため、この手法が有効であると考えられるが、この手法は行切り分けを行うことが前提である。そのため、古地図に適用させるためには、文字領域や文字の向きを求める必要がある。

そこで本研究では、ユーザが地図中の文字をなぞることを手掛かりとして、文字の向きや文字領域を決定する手法を提案する。

2.2 文字領域抽出法

文書には、文字だけではなく図が含まれている場合があり、地図には図だけではなく記号なども含まれている。そのような文書や地図から機械的に文字領域を抽出する研究が多くなされている [9] [10]。近年では OCR 技術の発達により、文字領域の抽出に加え文字の認識もされるようになってきた [11] [12]。Pouderoux らによって提案された手法 [11] は地形図を対象とした手法で文字の連結性に着目した手法である。Chian らによって提案された手法 [12] は、ラスタ地図を対象としており、OCR ソフトの水平方向のテキスト認識を活用することに焦点を当てた手法である。この手法は、モルフォロジー演算を用いることで文字の向きを求めており、OCR ソフトで文字が逆さまであるかどうか判定を行っている。どちらの手法も認識は OCR を用

いているが、古文書や古地図などは特有の文字で書かれているものが多いため、OCR に頼っての認識は難しい。こうしたことから、ワードスポッティングが古地図には有効であると考えられる。

しかし、古地図は、文字の向きがバラバラで、様々な向きの文字列が書かれている。さらに縦書きや横書きが含まれている。寺沢らのワードスポッティング [8] を適用させるためには、これらの問題について対応しなければならない。

そこで本研究では、ユーザが文字の向きに沿ってなぞることで、文字の向きと文字領域を決定する。こうしたユーザがなぞった情報をもとに切り出した文字画像で寺沢らによって提案されたワードスポッティング [8] を用いてマッチングする。古地図に対しキーワードを用いて探索できるようになることで新たな発見につながると考えられる。また、本研究では、Simon らの研究 [13] で地図に注釈を付けることで情報量を高めるようにユーザがなぞった文字画像をデータベースに追加していくシステム構成とする。ユーザがなぞった文字は、ユーザの興味の対象であり、それらをデータベースに追加していくことでデータベースは集合知としての価値を持つ。

3. なぞることによる古地図の探索システムの提案

本章では、なぞりを手掛かりとした複数の古地図の探索手法となぞった文字をデータベースに登録していく流れについて述べる。図 1 はシステム概念図である。本研究では、「なぞること」を文字の向きに沿ってなぞることと定義する。本研究で使用した古地図は、国立国会図書館のデジタルコレクション [2] からダウンロードしたものである。

3.1 前処理

文字を切り出すための前処理として、地図画像をグレースケール化し、大津の手法 [14] を用いて二値化を行う。前処理は画像全体を処理するのではなく、システムで表示されている部分だけを対象とした。二値化した後に、連結成分に対してラベリングを行う。図 2b は図 2a に対して前処理を行ったものである。

3.2 文字の切り出し

文字の切り出しは、ユーザがなぞった線上に存在するラベリングされた連結成分を文字とみなすことで行う。切り出しの手順は、ユーザが文字列をなぞった後、まずユーザがなぞった線の始点と終点をもとに文字列の向きを求める。次に、求めた向きをもとに画像を回転させ、再度連結成分に対してラベリングを行う。最後に、なぞった連結成分すべてを囲む外接矩形を求め、その矩形を切り出す。切り出し時、地図画像を回転する前にもなぞられた連結成分を囲む外接矩形を求める。それをもとにクエリのマッチング時

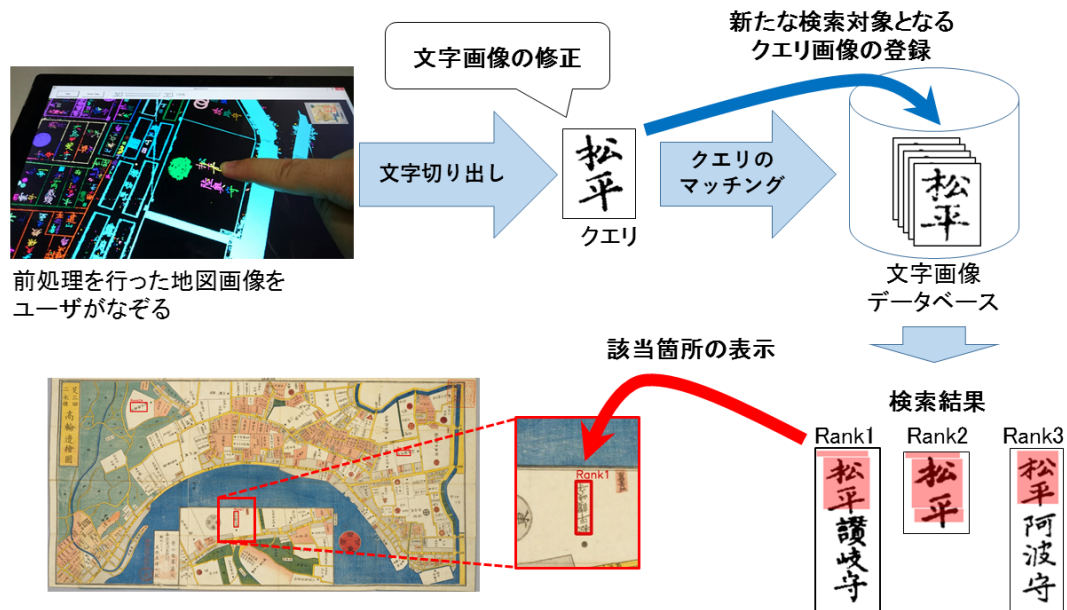


図 1: システム概念図

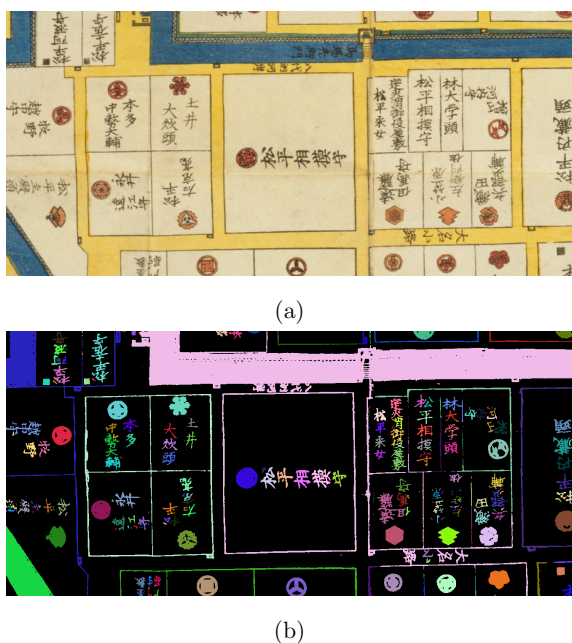


図 2: 二値化と連結成分に対してのラベリング (a) 入力した地図画像 (b) 出力された地図画像

になぞられ、データベースに s で登録されているかどうかを確認する。このなぞられているかの確認手法の詳細については、3.5 節で述べる。複数行の文字列がなぞられた場合は、それぞれの文字列に対し、同様の手法で文字列を切り出し、それらをつなげ一枚の画像とする。複数行の文字列は、それぞれの文字列が極端に大きさが変わらないと考えられるため、つなげる際に、最初になぞった文字画像に幅を合わせ、なぞった順番につなげる。図 3 は、ユーザがクエリとする文字をなぞっている図であり、図 4a はそのなぞった情報をもとに画像を回転させたものである。赤

枠で囲まれた文字列がユーザがなぞった文字列であり、図 4b は、外接矩形をもとに切り出した文字列である。

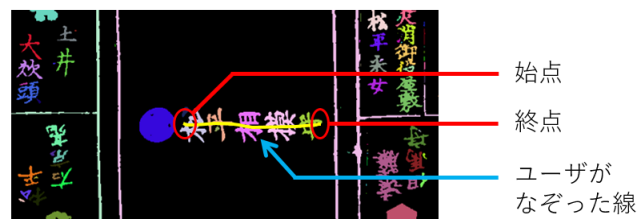


図 3: 文字領域と文字の向きの決定

3.3 文字の修正

文字の切り出しは、文字部分だけを切り出すのではなく、外接矩形で切り出しを行うため、区画線などの検索に適さない部分を文字と一緒に切り出してしまう。そこで、検索に適さない部分を修正できるようにするために消しゴムツールを実装した。図 5a のように区画線が含まれている場合、ユーザは消しゴムツールを用いて検索に適さない部分を修正する。図 5b は、図 5a を修正した例である。

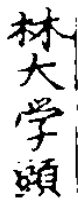
文字の切り出しは縦書きと仮定し行う。これは寺沢らによって提案された手法 [8] が文字列の向きでスリット状に切り出しているため、この向きを統一する必要があるためである。文字の切り出しを縦書きと仮定したため、横書きの文字列をなぞった場合、図 6a のような切り出しとなる。横書きの文字列を縦書きの文字列に修正できるようにするために回転ツールを実装した。回転ツールは、回転させる範囲を指定することで、その範囲の中心を軸に回転させる



(a) 文字領域

(b) 矩形をもとに切り出した文字列

図 4: なぞった文字の切り出し



(a) 区画線を含む文字画像



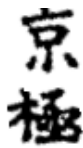
(b) 図 5a の修正例

図 5: 検索に適さない部分の修正

ことができる。ユーザは一文字ずつ文字に対して範囲を指定し、回転させることで横書きの文字列を縦書きに修正する。図 6b は、図 6a を修正した例である。



(a) 横書きの文字列



(b) 図 6a の修正例

図 6: 横書きから縦書きへの修正

3.4 クエリのマッチング

データベースは、地図から切り出した文字画像とそれらの地図中の位置座標を記述したファイルで構成されている。それらは、地図と紐づけされており、位置座標は文字切り出しの際に回転する前に求めた外接矩形の位置座標である。

マッチングは寺沢らによって提案されたワードスポッ

ティング [8] を用いて、切り出した文字画像をクエリとして行う (図 7)。マッチングでは、勾配分布特徴量を用いる。図 8 は勾配分布特徴量構成法の概念図であり、本研究では 5 つの小領域に分割して特徴量を記述した。また、文字の伸縮に対応するために DTW(Dynamic Time Warping) を適用し、マッチングを行う。

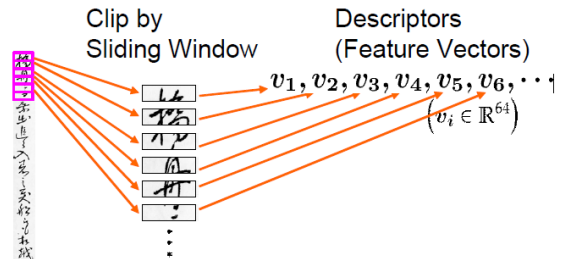


図 7: 特徴量記述 ([15] から引用)

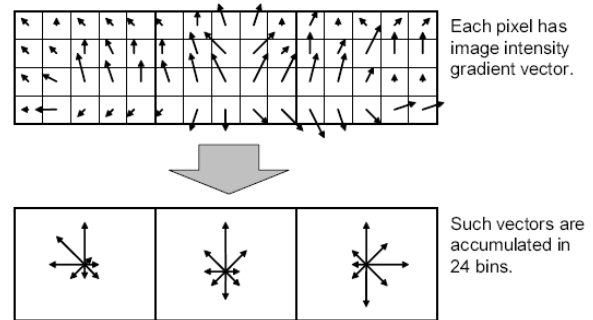
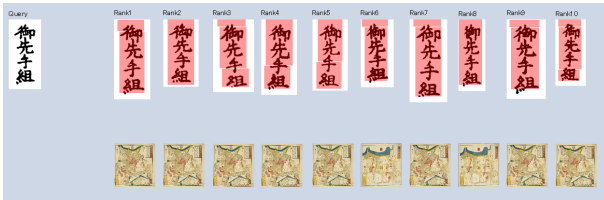


図 8: 勾配分布特徴量の構成法 ([8] から引用)

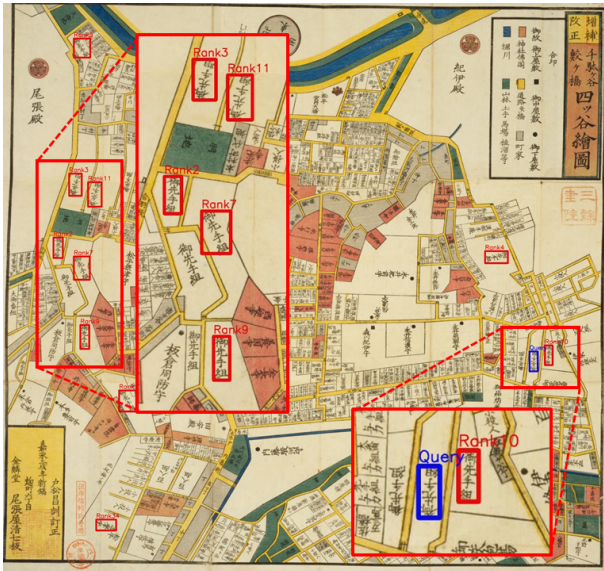
図 9a は、「御先手組」をクエリとした検索結果である。図中の左側にある文字列がクエリ (図 9b の青枠で囲まれた文字列から切り出した) であり、右側にある文字列は、左から順に類似度が高い文字列が並んでいる。検索結果の赤で塗られている部分は画像中でクエリと類似度が高い部分である。検索結果の下にあるのが、その文字が書かれている地図のサムネイルである。図 9b は、図 9a の結果を地図にハイライトしたものであり、図の赤枠で囲まれている文字列が類似度の高いものである。

3.5 新たな検索対象となるクエリ画像の登録

マッチング時にクエリとした文字がすでになぞられ保存されているかどうかを確認する。回転前に求めた文字列を囲む外接矩形の左上と右下の頂点の座標がデータベースに登録された位置座標にふくまれているかどうかでなぞられているかを確認する。その文字がまだなぞられていない場合は、ユーザが文字を登録する。文字が登録されていた場合は、ユーザは登録されている文字画像と今回なぞった文字画像を比較する。そして、今回なぞった文字の方がより



(a)



(b)

図 9: 結果: (a) ワードスポットティングの検索結果 (b) 検索結果をハイライトした地図

検索対象として適していると判断した場合に上書きする。なぞった文字を新たに登録していくことで、それらは新たな検索対象となり、データベースの拡張につながる。また、なぞった文字というのはユーザーの興味の対象であり、それらを登録していくことでデータベースは集合知としての価値を持つことにつながる。

4. システムの利用例

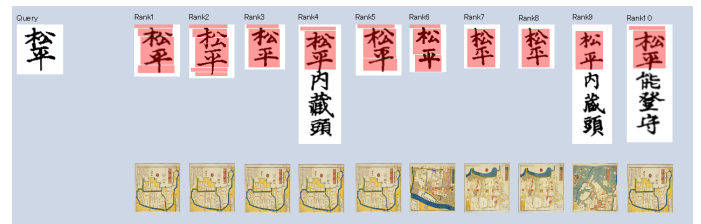
本章では、提案したシステムの利用例について述べる。提案システムを利用することで古地図から読み取れたことを、特定のキーワードをハイライトすることで明らかになった歴史的背景と、同じ地図上には多く書かれていない固有性の高いキーワードを手掛かりとした地図同士の比較の2つのシナリオで述べる。

利用例は約 200 年前に出版された江戸の切絵図 28 枚を用いて行った。これらの地図からは、農村地域や寺院、武家や一般市民の居住地といった地理的多様性など多くの歴史的背景を読み取ることができる。そのため、地図中の文字列をインタラクティブにクエリにし、複数の地図間で探索を行うことで歴史的背景のより深い理解につながる。

4.1 ハイライトすることで明らかになった歴史的背景

提案したシステムを使い、江戸切絵図を探索することで、ユーザが歴史的背景を知るのを容易にした例について示す(図 10)。ユーザは、江戸切絵図を閲覧しているときに、複数の地図上で特定のキーワード「松平」が数多くあることに気付いた。「松平」は代表的な大名(藩主)の姓であり、江戸には多くの居住地がある。

ユーザは、「松平」というキーワード(図 10b の青い枠で囲まれている)をクエリとして検索を行い、図 10a がその検索結果である。検索結果のサムネイルにより江戸の様々な地域に「松平」がいることがわかった。また、それらの多く(上位 10 件の内、6 つの検索結果)が江戸城周辺にいたことが明らかになった。図 10b は、検索結果の上位 30 位を地図にハイライトしたものであり、「松平」の多くが江戸城の周りにいたことを示している。事実、「松平」は江戸時代の藩主であり、幕府の親族でもあった。このように複数の検索結果をハイライトすることで歴史的背景を明らかにすることができ、この例では、幕府と「松平」に深いつながりがあったことを示した。



(a) 「松平」の検索結果

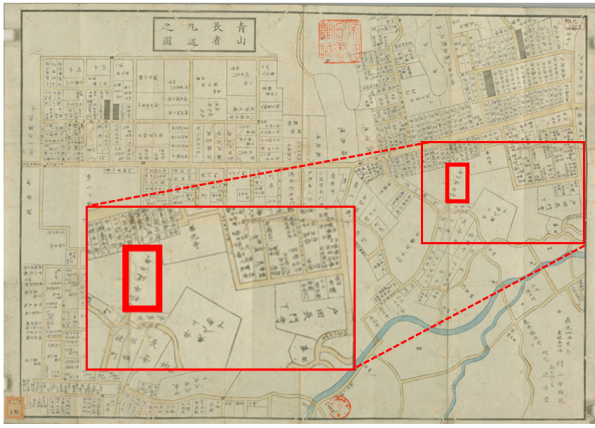


(b) 多くの「松平」が江戸城周りにいることを示す

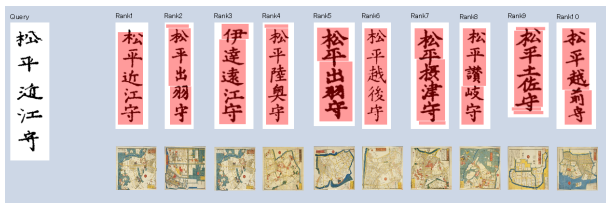
図 10: 重要なキーワードを検索することで多くの大名屋敷がヒット

4.2 固有性の高いクエリを手掛かりとした地図同士の比較

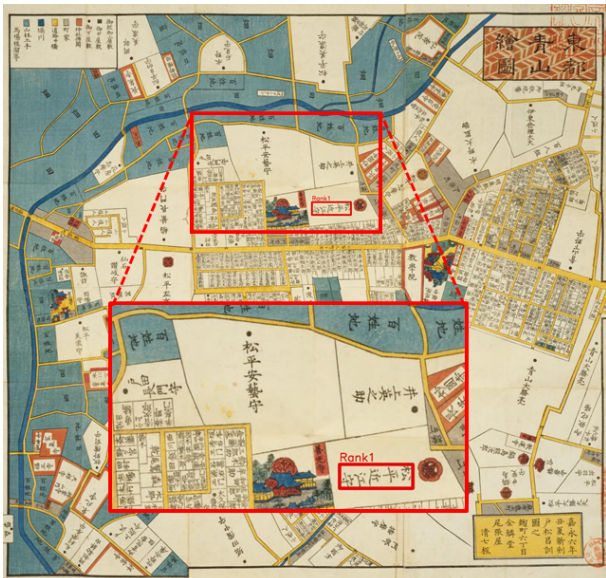
固有性の高いキーワードというのは同じ地図上に多くは



(a) クエリを切り出した地図 (青山長者丸辺図)



(b) 「松平近江守」の検索結果



(c) 類似度が最も高い文字列が書かれている地図 (青山渋谷絵図)

図 11: 同じ場所を描いた地図への横断 (a から b へ)

書かれない。つまり、そのような固有性の高いキーワードは、複数の地図間をつなぐ重要な手掛かりとなる可能性がある。江戸切絵図は、尾張屋や近江屋などいくつかの出版社から出版されている。同じ地域を描いているものもあるが、異なる出版社が出版した地図同士を対応させるのは難しい。

図 11a は 1881 年に近江屋によって出版された地図である。この例では、クエリとして「松平近江守」(図 11a の赤枠で囲まれたもの) というキーワードに注目した。このキーワードは、藩主の「松平」に官位を表す「近江守」を追加したものである。図 11b は「松平近江守」をクエリと

した検索結果である。その類似度が最も高い文字列はクエリとしたものと同じ文字列であり、別の出版である尾張屋が出版した図 11c の地図に描かれていた。それらは別々の出版社が出版した地図ではあるが、同じ場所を描いた地図であり、2つの地図は対応している。

地図同士を比較することで新しい発見が導かれる。例えば、クエリとした図 11a の地図には文字列しか書かれていないが、尾張屋の地図 (図 11c) には、家紋が描かれている。別の地図にアクセスし比較することで、クエリとしたキーワードが官位の高い大名であったことがわかる。このように固有性の高いキーワードをクエリとして用いることで、地図から地図への横断が可能となる。さらに、同じ地域を描いた別の地図へ横断することで、同じ地域を描いた異なる出版社の地図同士の対応が理解しやすくなる。対応関係を理解することで地図同士を比較することができる。

5. おわりに

5.1 まとめ

本研究では、ユーザが指でなぞった文字と類似した画像を持つ領域を表示することによって、古地図の探索を支援するシステムを提案した。なぞることは、コンピュータが文字の向きと文字領域を抽出する手掛かりとして有効である。切り出しは、ユーザがなぞった線上に存在する連結成分とみなすことで行う。そして、切り出した画像をスリット状にし、各スリットの勾配分布特徴量を用いてマッチングを行った。本研究では、ユーザがなぞった文字をデータベースに登録していくシステム構成とした。複数の検索結果をハイライトすることによって明らかとなった歴史的背景や特定の高いキーワードを手掛かりとした地図同士の比較での気づきの例について示した。

5.2 今後の展望

本研究での文字切り出しは、ユーザがなぞったラベリングされた連結成分を文字とみなすことで切り出す。そのため、連結成分をうまく抽出することができなければ、文字切り出しを行うことはできない。したがって、本研究での文字の切り出しは二値化の精度に依存する。ほかの古地図でも同様の手法で切り出すことができるかを確認する。

本研究では、登録されていない文字を新たにデータベースに登録するシステム構成とした。提案したシステムのデータベースは、ユーザの助けを借り、興味の対象を集めていき、徐々に成長していく。今後は、複数のユーザが地図を探索しての気づきを共有できるようにシステムを拡張し、データベースの成長に応じてどのようにユーザの満足度が変化するかを評価する。

参考文献

- [1] Cartography Associates: David Rumsey Map Collection, <http://www.davidrumsey.com/view/georeferencer> (参照 2017-05-10)
- [2] 国立国会図書館 : 国立国会図書館デジタルコレクション, <http://dl.ndl.go.jp/info:ndljp/pid/1286656?tocOpened=1> (参照 2017-05-10)
- [3] Manmatha, R., Han, C. and Riseman, E.M.: Word Spotting: A New Approach to Indexing Handwriting, Proc. CVPR1996, pp. 631-677, 1996.
- [4] Rath, T.M. and Manmatha, R.: Features for Word Spotting in Historical Manuscripts, Proc. ICDAR2003, vol. 1, pp. 218-222, 2003.
- [5] Gatos, B. and Pratikakis, I.: Segmentation-free Word Spotting in Historical Printed Documents, Proc. ICDAR2009, pp. 271-275, 2009.
- [6] Gatos, B., Konidakis, T., Ntzios, K., Pratikakis, I. and Perantonis, J.: A Segmentation-free Approach for Keyword Search in Historical Typewritten Documents, Proc. ICDAR2005, vol. 1, pp. 54-58, 2005.
- [7] Terasawa, K., Nagasaki, T. and Kawashima, T.: Eigenspace Method for Text Retrieval in Historical Document Images, Proc. ICDAR2005, vol. 1, pp. 437-441, 2005.
- [8] 寺沢憲吾, 長崎健, 川嶋稔夫 : 勾配分布特徴量による高精度手書き文字検索, 画像の認識・理解シンポジウム (MIRU2006), pp. 1325-1330, 2006.
- [9] Fletcher, L.A. and Kasturi, R.: A Robust Algorithm for Text String Separation from Mixed Text/Graphics Images, Proc. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 10, pp. 910-918, 1988.
- [10] Myers, G., Mulgaonkar, P., Chen, C.H. and Decurtins, J.: Verification-Based Approach for Automated Text and Feature Extraction from Raster-Scanned Maps, Proc. Graphics Recognition: Methods and Applications, pp. 190-203, 1995.
- [11] Pouderoux, J., Gonzato, J., Pereira, A. and Guitton, P.: Toponym Recognition in Scanned Color Topographic Maps, Proc. ICDAR2007, vol. 1, pp. 531-535, 2007.
- [12] Chiang, Y. and Knoblock, C.A.: An Approach for Recognizing Text Labels in Raster Maps", Proc. ICPR2010, pp. 3199-3202, 2010.
- [13] Simon, R., Sadiek, C., Korb, J., Baldauf, M. and Haslhofer, B.: Tag Clouds and Old Maps: Annotations as Linked Spatiotemporal Data in the Cultural Heritage Domain, Workshop on Linked Spatiotemporal Data 2010, Held in Conjunction with the 6th International Conference on Geographic Information Science (GIScience 2010), 2010.
- [14] Otsu, N.: A Threshold Selection Method from Gray-Level Histograms, Proc. IEEE Trans on Systems, Man, and Cybernetics, vol. 9, pp. 62-66, 1979.
- [15] 寺沢憲吾, 川嶋稔夫 : 文書画像からの全文検索のオンラインサービス, じんもんこん 2011 論文集, vol. 8, pp. 329-334, 2011.