

無名付箋の画像資料への適用方法について

津田 光弘^{†1}

この報告の目的は、OCR 技術などで識別が難しいデジタルアーカイブ画像の文字位置や行を、テキストのように選択可能にすることである。このために、タイトル等メタデータを伴わない不可視かつ暗示的な図形アノテーション・モデルを設計した。このモデルを Web 環境で利用するためデータを作成し HTML5 で実装を行い、最後に人文科学分野での応用を考察する。

How to apply title-less annotations (“pre-tation”) to any digital archived image

MITSUHIRO TSUDA^{†1}

The purpose of this report is to be able to select character positions and lines of any digital image that is difficult to identify automatically by OCR Technology etc., like a text. For this, I designed an invisible and suggestive annotation model without metadata such as a title. In order to use this model in the Web environment, I designed data sample and made the prototype by HTML5, and, finally consider applications in the humanities.

1. はじめに

本研究の目的は、文字画像の内容を手作業で効率的に選択可能にする方法を確立することである。

今日の Web を中心としたテキスト検索や関連サービスは、既に私たちの情報活動の重要な基盤となっている。しかし、画像資料についてはまだそのような恩恵が少ない。中でも古文書、和書、書簡などが大量にデジタルアーカイブされ公開されているが、画像の閲覧という段階で留まっており、テキスト資料との差が大きいとの指摘もある。[1] Web ブラウザで任意の箇所を選択し取得や記録を望んでも、テキストのように選択できない。Google Books[2]のようなテキストコピーやリンクの仕組みを提供しているサービスはまだ少ない。OCR 以前に画像中の行や文字を選択するためにはそれらの配置情報が必要で各種手法も研究されている[3]が、自動的に識別することが難しい資料も多くある。例えば、図 1(C)のように文字間の不均等、斜めの行、文字の間に別文脈の行があるようなレイアウトの画像資料もある。このような資料は現状では人間の認識に勝る手段は無く、適材適所でデジタルに繋げてゆけば良い。内容検索には至らずとも、選択ができれば次の段階へ進めると考え本報告で提案を行う。

2. 画像資料と付箋

人による手法として、図形を用いた付箋（アノテーション）が知られる。画像の座標を基準とし図形のメタデータに座標値を保存することで、文字などの場所が記録できる。テキストのように対象の範囲を指定できない画像では、こ

の仕組みは有効である。また、アノテーションにタイトルやメモなどのメタデータを追加すればそれらを基に検索できる。昔から付箋紙が便利に使われてきたが、今日のアノテーションも利点は多い。

しかし、画像がコンピュータで利用されて以来、このアノテーションは知られている手法にもかかわらず目にする機会は少ない。今日の Web 地図上のマーカーは同手法の例外的な成功例である。使われない原因はいろいろあるだろうが、PC でマウスを用いてアノテーションを付けるという作業行為自体の問題をここでは取り上げたい。

マウス操作による小さなアンカー（＝アノテーション）の配置は経験上身体・精神的疲労（具体的には腕や肩、首が凝る）が大きい。ノート PC のトラックパッドでは慣れにも依るが細かな操作は厳しく、キー操作は論外である。一方、ペンやタッチ操作の情報デバイスでは上記の身体的な疲労は大幅に軽減される。だが、それでも画像上に位置を決め、範囲を定義し、テキストを入力する一連の行為はページを選んでブックマークするほどには受け入れられるものではない。要するに人間の操作行動から考えて精密な

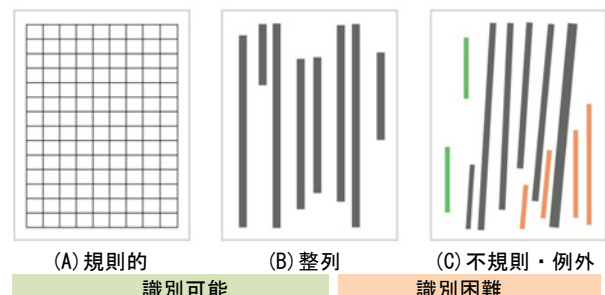


図 1 文字画像のレイアウトパターン

Figure 1 Layout pattern of text image

^{†1} イバレット
iPallet

要求をし過ぎるのだと経験から推察する。では、アノテーションを付けないアノテーションの仕組みがあれば良いのではないかということが提案の基である。

3. 無名付箋

3.1 ガイドとしての付箋

タッチデバイスへの拡張を想定し、最低限の操作でテキストのように対象位置を指定できる別種のアノテーションを設計した。

文字や列の位置のみを示すこのアノテーションを従来のタイトルやメモなどメタデータを記載したアノテーションと区別するため、無名付箋（あるいは、前付箋：プレテーション）と用語を仮定義する。この無名付箋の特徴は、

- 表示しない（表示してもガイドの役割）
- 自体を検索や情報の対象としない
- この座標情報を用いてアノテーションを作成する

3.2 無名付箋の作成、座標モデル

画像上の文字を選択するためには、文字の存在する位置をデータ化する必要がある。段落、文節、単語、文字などの内容に関するレベル、頁、行などの媒体に依存するレベルがあるが、容易な選択という操作を目標とするため、最低限、行の位置情報を扱うこととした。

図2でこの座標モデルを説明する。以下では、文字の配置は上から下、行の順序は右から左の日本語縦書き（tb-rl系）を想定する。

(a)は最も簡単な矩形を用いたモデルである。先の図1(B)のように文字がおおよそ正しく上から下に向けて並び、文字の大きさが多様である場合、外側を四角形で囲うこのモデルが簡単である。しかし、直筆の書簡など場合には行が多少斜めとなる場合もあり、(b)のように矩形を無理に適用すると範囲が大きくなって、現実のクリックやタッチ操作で隣接する行との干渉が生じる恐れがある。

そのため理想的には(c)のように線あるいはベクトルで行の開始点、終了点を定義することが望ましい。HTMLで実装する場合は(c)はDOM要素では作成できないが、HTML5ではSVG(Scalable Vector Graphics)のポリゴンで制作が可能である。更に、製作時の効率と文章構造を考慮すると、(d)のように「一筆書き」のポリラインが有効と判断した。

3.3 文字幅のデータ拡張

ユーザーインタフェースで文字を選択する際に、文字の大きさに相当する範囲を作成する必要がある。行のデータをポリラインで記述した場合に、文字の大きさを併せて定義する方法を検討した。

図3(a)は行データと共に文字の大きさを定義する。しかし、この方法は作成時の工程が増える。(b)は仮想的な文字の大きさの縦横1/2を最初のポリラインストロークで定義する方法。最後にも同様の定義を行う。(c)は(b)と同様であ

るが、個々の文字の高さは縦書きの場合必ずしも定義できないので、幅のみを定義したもの。これは行が折れている場合にも適す。(b)と(c)の違いは、(b)が文字の中心をデータ化する記述であり、(c)は文字の境界をデータ化する方法と考え得る。(d)は欧文など横書きlr-tb系のデータであるが、縦書きと類似構造で定義できる。

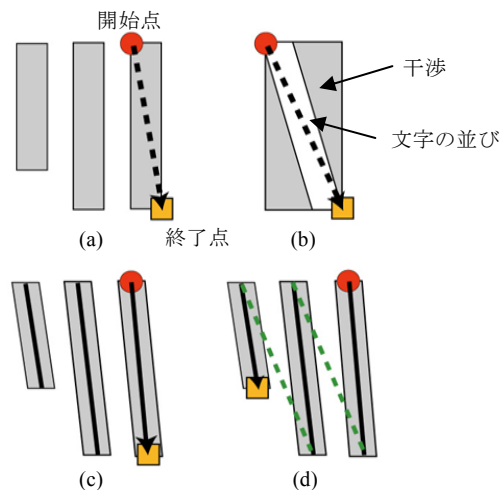


図2 行のモデル
 Figure 2 model of lines

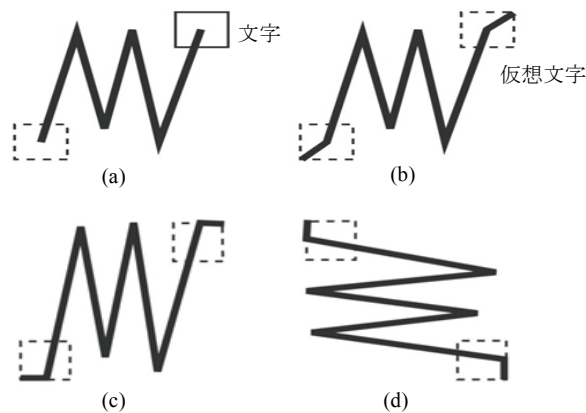


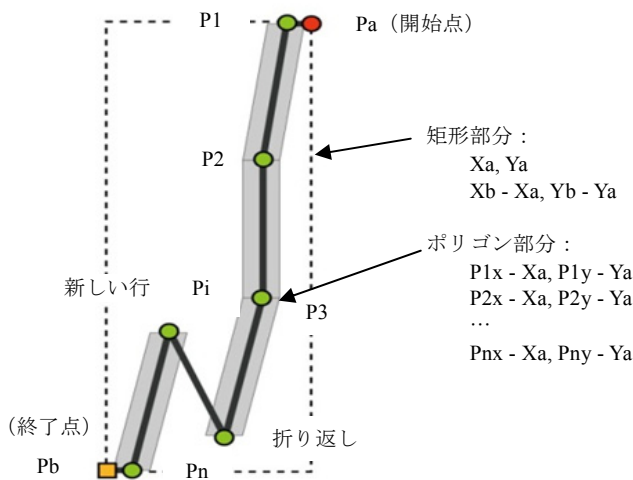
図3 仮想文字の幅
 Figure 3 width of virtual character

3.4 座標データ・モデル

以上の検討により、無名付箋の座標データモデルを定義した。

複数の行をグループとして考えた場合のtb-rl系の一般モデルを図4で説明する。座標系は左上に原点、右および下方向に増加。右上角の丸いポイントから右下の四角いポイントまでの矩形が範囲を表わす。それらポイントの内側のポリラインで行とその折り曲げあるいは文節、改行などの構造を定義する。

開始点以外はすべて開始点を基点とする相対座標である。



$Pa = [Xa, Ya], Pb = [Xb, Yb], Pi = [Pix, Piy]$
 とし、以下の各項目をスペースで連結した文字列

開始点座標 (座標原点は左上)
 点線範囲の矩形の幅と高さ

P1 点の原点からの相対位置
 P2 点の原点からの相対位置
 ...
 Pn 点の原点からの相対位置

図 4 データ構造 (座標データ規則)

Figure 4 data model

P1 の X 値がポリラインを中心とする選択の幅値 (負) に相当する. 座標は SVG に倣い各ポイントの XY 座標はコマ, 各ポイントの連結は半角スペースで行い, 文字列として扱う. 連結の順序は, 開始, 終了点から成る矩形座標データに続く形で, ポリラインデータを配置する. 折り返し点と文節あるいは折り曲げ点との区別はポリライン内の数値で判断する. 単位は基となる画像のピクセル値とするが, 比率を整数化して用いる方法もある. このような矩形とポリゴンを融合させた座標データモデルを用いることで, 今回対象とする文字画像のみならず, 一般的な画像の内容へも同一モデルが適用できる.

4. 編集の実装と評価

座標を取得するための編集ツールは, Flash 製の付箋ツール[4]を用いた. 通常のアノテーション入力と異なり, メタデータは最低限マウス操作の結果として自動的に記録される座標値のみである. 規則的に入力する限り, 識別用のラベルも必要ない. 通常のアノテーションや行単位で入力す

る方法に較べ, 一筆書きの方式は制作時間が短く, 身体的負担も少ない.

5. 表示モデル

次に, 無名付箋の表示について検討する.

表示はアプリケーションの環境やデバイスに依存するため, 次の項目を条件とした.

- Web ブラウザ
- HTML5
- PC 及びタブレットデバイス

HTML5 の SVG 機能を用いて, ポリラインの各線分をポリゴン要素として作成する. 斜めの行であっても範囲は基本的には干渉しない.

JavaScript, CSS, jQuery といった標準あるいはデファクトスタンダードな技術で実装した例を図 6 に示す. 無名付箋は透明にするが, 説明用に着色している. 薄紫色の箇所が座標データモデルの数値を用いて作成した SVG のポリゴン要素のレンダリング結果である.

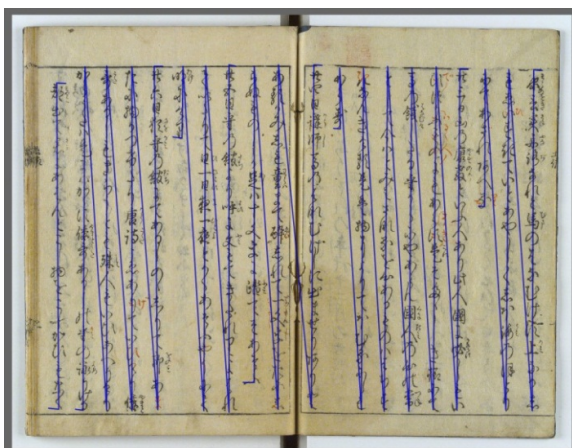


図 5 編集例

Figure 5 example of edit

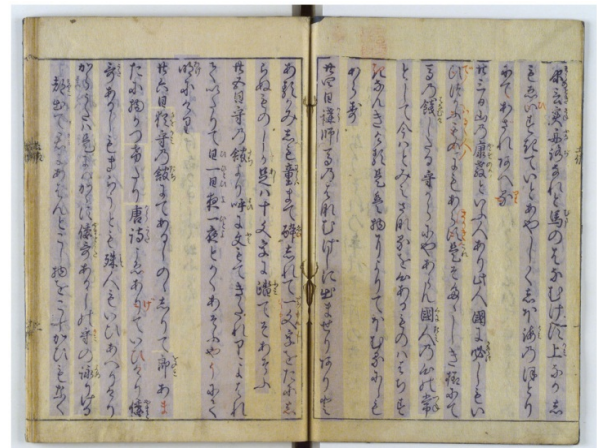


図 6 表示実装例

Figure 6 display of pre-tation on web browser

6. 選択モデル

無名付箋を利用して、画像文字をテキストのように選択する方法について説明する。無名付箋はあくまでガイドであってそれら自体の選択は行わない。無名付箋の座標情報を用いて選択の際に新しくオブジェクトを作成する。

既に作成した無名付箋のレイヤに、選択用のアノテーションのレイヤを DOM 要素で作成して重ねる。選択オブジェクトは無名付箋を再現するため、このレイヤでもその表示には SVG を用いる。

図 7 で、無名付箋を黒色の枠とする。これらは SVG のオブジェクトで、DOM および JavaScript 内では順序を持つ配列である。選択する際に、マウスボタンを押下した開始点（赤色丸印）からマウスボタンを放す終了点（黄色四角）までドラッグ操作によってマウスを移動させる（赤い矢印）と、両者のポイント座標と SVG オブジェクトの順序番号が得られる。この順序の増減に従って簡単なアルゴリズムで範囲にある SVG 要素を動的に作成・更新する。青色枠線の箇所がこうして得られた 3 つの SVG 要素のレンダリング結果である。順序値がマウスダウン時のものより大きければ後方への選択、小さければ図 7(b)のように前方と解釈する。

図 8 は本報告の最終的な例である。選択した範囲を着色

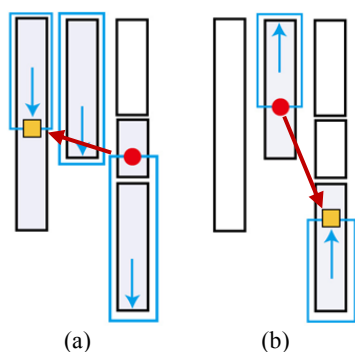


図 7 連続選択モデル

Figure 7 continuous selection model

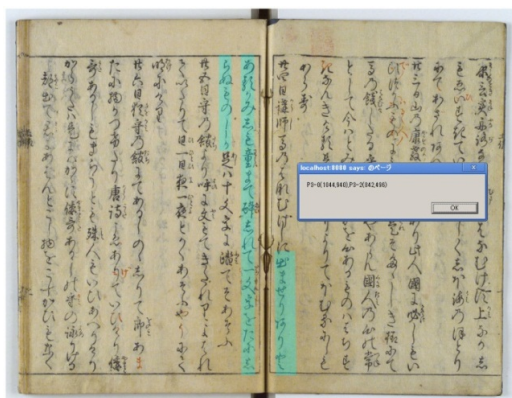


図 8 選択例

Figure 8 an example of selection

している。ブラウザ上でテキストを選ぶように画像中の文字位置を擬似的に選択できている。また、図の中では、アラートダイアログを用いてマウスボタンのアップ時に選択範囲の座標と無名付箋の ID を抽出し表示している。これは無名付箋で定義した内容を選択によって取り出せることを証明している。

7. 人文科学の場としての無名付箋

無名付箋を用いた選択は、その選択データを用いることで、あるいは蓄積することで各種の応用に結びつく。例えば、人文科学の分野を念頭において、

- 選択したオブジェクトを通常の図形アノテーションにしてリンクやメタデータを登録
 - C/S システム化し検索
 - ネットワークを用いた共有メモ、共有辞書
 - 座標データを文字認識用レイアウト分析に利用
- 単に画像であったものが、僅かなデータの付与で拡張できる可能性がある。資料の場の形成を期待する。

8. まとめ

文字などの内容選択のための指標化が困難な文字資料を対象に、タッチ操作等で簡単に文字箇所を連続的に選択する方法を設計し、データモデルを定義した。そして、HTML5 と JavaScript による実装でその実効性を確認した。また、選択の結果を用いた応用について考察した。

本報告で提案する無名付箋（プリテーション）は、人文系の文字画像にのみ限定したモデルではない。一般の画像においても、また、既に OCR で内容を取得できている資料にも広く用いることができる。

作業疲労の軽減を目的に可能な簡素化を行ったが、大量の資料を扱うプロジェクトでは手作業自体は現実的ではなく、画像処理による自動レイアウト認識を応用した省力化は常に考慮の必要がある。その一方で、本手法は利用者自身が作る必要もない道具として、今すぐにも実現可能な方法である。今後、汎用化と応用を進める予定である。

なお、本研究は国立歴史民俗博物館基盤研究「デジタル化された歴史研究情報の高度利用に関する研究」（平成 22 年度～平成 24 年度）の一部である。

謝辞 本検討にあたり、貴重なご意見をいただいた皆様に感謝の意を表する。

参考文献

- 1) 岡本隆明:デジタル画像資料を利用した文献研究に必要な環境について, 人文科学とコンピュータシンポジウム, pp.67-72, 2009.
- 2) <http://books.google.com/>
- 3) 猪村元, 田中譲:図形特徴に基づく手書き古文書画像の全文検索手法, 電子情報通信学会論文誌 D, vol.J92-D, no.11, pp.2022-2032, 2009.
- 4) <http://www.ipalletnexus.org/>