

電子ドキュメント上での書き込みを支援する 手書きアノテーション認識モデル

浅井洋樹[†] 山名早人[‡]

[†] 早稲田大学大学院基幹理工学研究科 〒169-8555 東京都新宿区大久保 3-4-1
早稲田大学メディアネットワークセンター 〒169-8050 東京都新宿区戸塚町 1-104

[‡] 早稲田大学理工学術院 〒169-8555 東京都新宿区大久保 3-4-1

国立情報学研究所 〒101-8430 東京都千代田区一ツ橋 2-1-2

E-mail: {asai, yamana}@yama.info.waseda.ac.jp

あらまし ドキュメント上にペンで情報を付加する手書きアノテーションは、資料に関するメモの記録や理解を深めるために広く行われている。電子ドキュメント上での手書きアノテーションを実現するためには、自由に書き込めながらも複雑な操作で思考を妨げない手書きの特長を生かせる UI 設計や、動的に変化するコンテンツに対応するためのアノテーション認識が必要となる。本研究では、電子ドキュメント手書きアノテーションシステム設計支援、およびアノテーション範囲認識を目的とした手書きアノテーション認識モデルを提案する。本提案モデルでは下線や囲い込み、縦線といった一般的な範囲選択アノテーションの自動検出に加え、筆記者の意図する選択範囲推定を実現する。認識精度評価の結果、83~96%程度の精度でアノテーションを認識可能であることを確認した。

キーワード 電子ドキュメント, 手書き, アノテーション

1. はじめに

資料上に情報を書き込むアノテーション行為は、ドキュメントと人間の間における重要なインタラクションの一つである[1]。紙のドキュメントに対して手書きで情報を書き込む手書きアノテーションは、ドキュメントに対する理解や情報付加、編集、校正など様々な場面で行われており、我々がドキュメントを扱う際に欠くことのできないものとなっている。このように紙のドキュメントに対して手書きアノテーションが行われてきたが、ディスプレイ上に動的に表示される電子ドキュメントに対しても手書きによるアノテーションを実現することによって、電子ドキュメントの利便性が向上すると考えられる。

アノテーションの基本的な動作はドキュメント上の強調または情報付加をする箇所を選択する「範囲選択」とコメントや図などの情報を記入する「情報付加」の2点で構成される。紙のドキュメント上における手書きアノテーションでは、ペンによる下線や囲い込みといった範囲選択とテキストや図の書き込みといった情報付加を行なっている。一方で電子ドキュメント上でのアノテーションシステムでは、マウスやタッチ操作によるドキュメント範囲選択とキーボードによる情報付加を行うものが一般的であり、手書きアノテーションとは異なるドキュメント範囲選択操作やメニュー操作が加わるユーザインタフェースが提供される。またキーボードでのアノテーションは手書きアノテーションと比較して文字変換作業を伴う場面では非効率であるとの報告[2]もあり、電子ドキュメントに対

しても手書きアノテーションを実現できるインタフェースに関する研究が行われている[3][4][5][6]。

しかしこれらのシステムは、紙のドキュメントと同様な手書きアノテーションが可能であるものの、ドキュメント中の特定の箇所を強調するアノテーションや、アノテーション情報を付加したドキュメント中の要素を理解可能なのはユーザのみであり、コンピュータがその情報を処理できる状態ではない。このため、手書きアノテーションを支援するインタフェースの設計や、ドキュメントの閲覧性向上、検索システムの開発が制限されてしまう。そこで本研究ではこの課題を解決するため、以下の2点を実現する手書きアノテーション認識モデルを提案する。

- 手書きによる範囲選択と情報付加の識別
- ドキュメント上の選択範囲推定

学習データを用いて上記を実現する提案認識モデルのパラメータを推定し、人間のアノテーション傾向に基づいたアノテーション検出モデルを構築する。本研究で提案する手書きアノテーション認識モデルを利用することにより、ユーザの負担を増やすことなくドキュメント範囲選択を検出可能となるため、手書きアノテーションシステムの設計においてアノテーションを支援するシステムが実現可能となる。また、認識したアノテーション情報を用いた情報ナビゲーションシステムや検索システムを構築することが可能となる。

2. 関連研究

人間のアノテーション行為を調査した報告の他に、動的にレイアウトが変化する電子ドキュメント特有の問題に対して取り組む研究や、コンピュータ上で手書きアノテーションを認識する手法に関する研究が行われている。また最近ではこれらの技術を応用した電子ドキュメント上での手書きアノテーションを実現するシステムに関する研究が行われている。本節では本稿で提案するアノテーション認識モデルに特に関連する部分であるドキュメントへの手書きアノテーションを調査した研究や手書きアノテーションを認識する手法に関する研究について述べる。

2.1. ドキュメントへの手書きアノテーション

本節ではアノテーション行為自体についてその種類や目的などを調査している研究について説明し、本提案モデルで検出するアノテーションの種類について議論する。Marshall は大学生・大学院生が講義で使用するテキストを対象にアノテーションに関する調査を実施し、その結果を報告している[1]。調査報告によると、収集した 150 冊以上のテキストを調べた結果、アノテーションの種類について次の 2 つの特徴に分類している。

1) Telegraphic

テキスト中では下線やハイライト、囲い込みといった書き込み、余白ではカッコ（本稿では縦線と表記）や記号、ページ全体の囲い込み、矢印といった書き込みがこれに該当する。

2) Explicit

テキスト中では行間のコメントや外国語単語の訳の書き込み、余白では短い文章やテキスト中では書ききれなかったコメントがこれに該当する。

一方 Wang らはアノテーションの役割について次の 2 種類に分類できると述べている[7]。

1) Non-Actionable

説明文や要約、強調、コメントといった情報を付加するアノテーション。

2) Actionable

挿入や削除、移動や置換といった編集操作を指示するアノテーション。

また我々の以前の研究[8]において、手書きノートにおいて特定の箇所を強調するアノテーションについて調査を実施した結果も、ここで挙げた報告と同様に下線や囲い込みといったアノテーションがおこなわれて

いることを確認している。これらの報告をもとに本研究で扱うアノテーションについて整理する。本研究ではアノテーションをドキュメントの範囲を選択する「範囲選択」とこれ以外の書き込みである「情報付加」の 2 ステップに分けて扱っている。範囲選択は Marshall の分類においてテキスト中でおこなわれる Telegraphic アノテーションがこれに該当するものであり、これを参考に提案認識モデルでは下線、囲い込み、縦線の 3 種類の範囲選択の認識に対応させる。

また Wang らの分類において Non-Actionable に該当する書き込みが本提案モデルの認識対象である。Actionable に該当する書き込みは、本研究の対象とする電子ドキュメント上においては編集操作を意味するものであり、アプリケーションにおけるメニュー操作や、本認識モデルへの入力とは別の段階でジェスチャとして認識すべきものであると考え、本研究での認識対象からは除外する。

2.2. 手書きアノテーションの認識

本節では前節で述べたような手書きアノテーションを認識する手法に関する研究について述べる。Wang らは電子ノートシステム上に書き込まれた手書きアノテーションを認識する研究[7]を行っている。この研究では、画像やテキストといった様々な種類のコンテンツが貼り付けられたノート上に書き込まれた手書きアノテーションを、図や下線、囲い込みや縦線など 7 種類に分類する手法を提案している。分類における特徴量としては Fonseca らの研究[9]で挙げている形状特徴量と、文字の書き込みと下線の位置関係といった意味特徴量を用いており、これらを機械学習で分類をおこなっている。また、その後の報告[10]ではコンテンツ間を接続する矢印や線といったアノテーションの認識について述べている。

これらの研究はアノテーションをおこなっている手書きストロークが検出できるが、本研究の対象としている電子ドキュメント上でのアノテーションにおいて、どの箇所を選択しているかを認識することはできない。本研究ではアノテーションの分類だけでなくドキュメント上の選択範囲も検出することで、電子ドキュメントのアノテーションシステムやナビゲーションなどの情報検索システムを構築する上で必要となる情報の認識を行う。

3. 手書きアノテーション認識モデル

本節では本研究で提案する手書きアノテーション認識モデルの詳細について述べる。本研究で提案するアノテーション認識モデルでは、手書きアノテーション支援インタフェースや情報閲覧・検索システムへの応用を想定したものである。入力データとして手書き

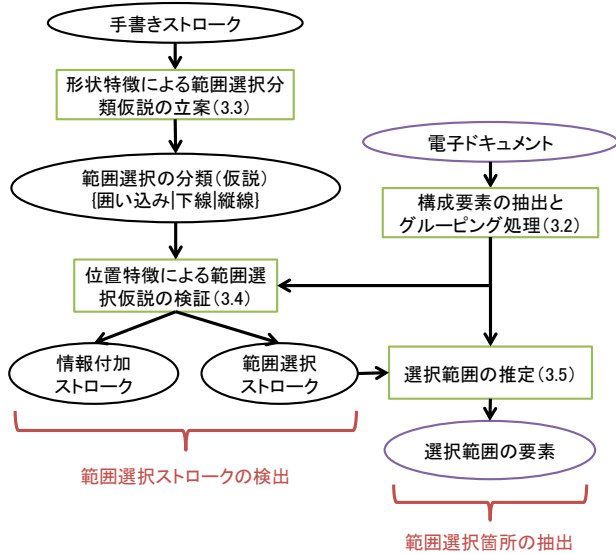


図 1 提案アノテーション認識モデルの概要

ストロークと電子ドキュメントデータを入力する．ここで入力として挙げた手書きストロークとはペンで筆記する際にペン先が画面に触れてから離れるまでの範囲を指す．本認識モデルは 1 ストローク分を入力データとして扱うため、一筆ごとにアノテーションの認識を処理することが可能である．認識モデルの出力として得られる情報は、入力手書きストロークが範囲選択ストロークであるかどうかの判定結果と電子ドキュメント上の選択範囲情報である．

以上のような認識を行う提案モデルの処理フローを図 1 に示す．本提案認識モデルでは、まず電子ドキュメントの情報からドキュメントを構成する文字や画像といった要素を抽出する．さらに文字データに対しては行単位でのグルーピングを行う．また、入力した手書きデータに対してはまずどの範囲選択ストロークに該当するかの分類仮説を立て、その後仮説と抽出したドキュメント要素位置情報から範囲選択ストロークの判定を実施する．また範囲選択ストロークとして認識されたストロークは統計データに基づいた範囲選択箇所の推定を行い、選択範囲のドキュメント要素を得る．以下の節で詳細を述べていく．

3.1. 検出対象アノテーション

本節では認識対象とする手書きアノテーションによる範囲選択ストロークの種類について述べる．2.1 節でも述べたようにドキュメント中の箇所を指定する手書きアノテーションには、強調や情報を付加する役割と編集操作を指示する役割が存在する．本研究では編集操作のジェスチャ認識を行うのではなく、ペンによって情報を付加するアノテーションを認識することが目的のため、[7]で挙げているアノテーションの種類

からこの目的を満たす下線、囲い込み、縦線の範囲選択ストロークに対応させる．認識対象となる 4 種類の範囲選択アノテーションを表 1 に示す．

表 1 検出対象範囲選択アノテーションの例

範囲選択分類	アノテーション例
下線	ような機械学習ではないため、最初から高い認る。 <u>オンライン手書き文字認識</u> [編集] オンライン手書き文字認識は、特殊なタブレットするもので、センサでペン先の動きやペンの上
囲い込み（単一行）	されている。 文字認識 [編集] 抽出後、個々の文字のイメー
囲い込み（複数行）	対象領域の特定 [編集] 対象領域を特定することで、手書き文字認識システムの認識率を向上させることが多く、例えば、ZIPコード（郵便番号）を対象とすれば、文字の縦幅はほぼ一定の数字で固定される。このため文字の特定の精度をかなり向上させることができる。 主なような技法がある。 • 使用される文字の種類を制限する。 • 特別な書式を採用し、特定の位置に特定の文字種が置かれるようにする。 文字抽出 [編集] 抽出された文字は、通常はそのまま使用されるが、抽出された文字は、抽出された文字から抽出された文字を抽出する。
縦線	ニューラルネットワーク [編集] ニューラルネットワークによる認識エンジンは、まず訓練用イメージ群で訓練済みネットワークを使って文字を認識する。ニューラルネットワークは、特徴を区別する属性を学んでおり、認識対象のイメージに同様の属性がないネットワークはセットアップが簡単だが、認識対象と傾向の異なる訓練データがないと、認識率は低くなる。 特徴抽出 [編集] 特徴抽出もニューラルネットワークと似たような方式で機能するが、プ

認識対象の範囲選択ストロークの考え方として、要素単位で選択する範囲選択ストロークと行単位で要素を選択する範囲選択ストロークで分類している．要素単位での範囲選択では下線と囲い込み（単一行）に対応し、行単位での範囲選択では囲い込み（複数行）と縦線に対応している．なお筆者によってはここで挙げた 4 種の範囲選択ストローク以外のパターンも存在する可能性があるため、本稿で提案する認識モデルでは開始・終了対応点を定義することで認識パターンを追加できるように設計している．

3.2. ドキュメントからの要素抽出とグルーピング

提案モデルが最初に行う処理として、ドキュメントに含まれる文字や画像などの要素情報を取得する．本節では電子ドキュメントから抽出する構成要素とそのグルーピング処理について説明する．抽出する情報は要素のサイズを示す矩形情報とその要素が表示されている座標情報である．なおこれらの情報はドキュメントの種類（HTML や PDF など）や要素の種類（文字・画像など）に依存しない情報のため、様々な電子ドキュメントに対して汎用的に利用できるように設計している．

ドキュメントの構成要素情報を取得した後に、文字要素に対しては行単位でのグルーピング処理を行う．これは行単位で選択を行う種類の範囲選択手法に対応す

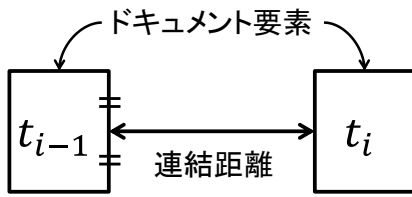


図 3 ドキュメント要素と連結距離

るためである．具体的には要素の出現シーケンス上で隣り合う要素同士の連結距離を求めることによってグルーピングを実施する．連結距離とはシーケンス上で隣り合う要素を囲む矩形において，その右辺の中点と左辺の中点を結ぶユークリッド距離を表すもので，その関係を図 3 に示す．

ドキュメント内に存在するすべての文字要素に対して連結距離を算出し，以下で算出される閾値 P を超えた場合を分割点としてグルーピングを行う．

$$P = \tilde{D} + 2 \times \bar{W}$$

ここで $\tilde{D}$ はページ内のドキュメント構成要素 $t \in T$ における連結距離の中央値， $\bar{W}$ は t の矩形幅における平均値をそれぞれ表している．

3.3. 形状特徴による範囲選択ストロークの分類

ドキュメントの構成要素抽出とグルーピングをした後に，手書きストロークの処理を行う．本節では形状特徴を利用した範囲選択を行う手書きストロークの分類について述べる．本研究では入力された手書きストロークをアノテーションにおける範囲選択を表す**範囲選択ストローク**と，それ以外の書き込みである**情報付加ストローク**の 2 種類にわけて考えている．手書きストロークが入力された時点ではこの両者の手書きストロークが混在した状態で存在しているが，最初の処理である本ステップではこの両者の識別を行わない．ここではすべてのストロークを範囲選択ストローク仮定したうえで範囲選択ストロークの種類ごとに分類を行い，分類結果を仮説として出力する．情報付加ストロークとの識別については次のステップで実施する．

本ステップで分類する範囲選択ストロークは囲い込み，下線，縦線の 3 種類である．範囲選択ストロークの分類における特徴量として，手書きストロークの形状から得られる特徴量を利用する．本提案モデルは電子ドキュメント上での手書きアノテーションを支援するインタフェースでの応用を想定しているため，ユーザのアノテーション行為を妨げない処理速度が要求される．このため，より少ない特徴量でシンプルに判定できる手法を選択することが望ましい．そこで範

囲選択ストロークの分類特徴量としてそれぞれの範囲選択ストロークの特徴を顕著に表す次の 2 つの特徴量を採用する．

- 縦横比

$$\frac{H_{\text{stroke}}}{W_{\text{stroke}}}$$

- 密度

$$\frac{L_{\text{stroke}}}{W_{\text{stroke}} + H_{\text{stroke}}}$$

なお H_{stroke} ， W_{stroke} は手書きストロークの矩形における幅と高さ， L_{stroke} はストロークの長さを表している．また分類器として高速に判定可能で少ない特徴量において可読性に優れる決定木を用いる．なお決定木の構築には R の mvpart パッケージ¹による構築アルゴリズムを用いた．

3.4. 範囲選択ストロークの識別

本ステップでは 3.3 で得られた範囲選択ストローク分類の仮説と 3.2 で得られたドキュメント構成要素・グループ情報を用いて，入力された手書きストロークの範囲選択ストローク識別を行う．ここではストローク単位で範囲選択ストロークと，それ以外の情報付加ストロークの分類をおこなう．

本提案アノテーション認識モデルでは選択範囲をドキュメント上の開始要素と終了要素の検出によって推定する．開始要素，終了要素それぞれに対して手書きストロークとの位置関係を利用して判定を行なう．3.3 節で分類した範囲選択ストロークの分類に対してそれぞれアノテーション開始対応点・終了対応点を定義し，これとドキュメント要素との位置関係から算出できる相対座標を識別の特徴として用いる．本手法を採用することで，本研究で対象とした範囲選択スト

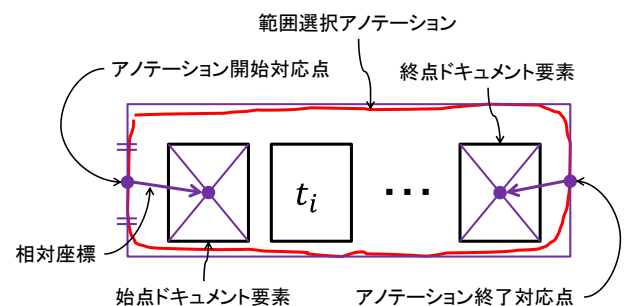


図 2 ドキュメント要素との相対座標

¹ mvpart: Multivariate partitioning, <http://cran.r-project.org/web/packages/mvpart/index.html>

ローク以外への対応がこれら 2 点の開始・終了対応点を定義するのみで対応可能となる柔軟なモデルを実現している．相対座標の説明に関する図を図 2，各範囲選択ストローク分類に対するアノテーション開始対応点・終了点の定義を表 2 にそれぞれ示す．

表 2 アノテーション開始・終了対応点の定義

範囲選択分類	アノテーション開始対応点(a_x, a_y)	アノテーション終了対応点(a_x, a_y)
下線	ストローク矩形の左辺中点	ストローク矩形の右辺中点
囲い込み(単一行)	ストローク矩形の左辺中点	ストローク矩形の右辺中点
囲い込み(複数行)	ストローク矩形の左上端	ストローク矩形の右下端
縦線	ストローク矩形の上辺中点	ストローク矩形の下辺中点

分類された仮説に従って，ドキュメント要素の重心に対するアノテーション開始・終了点との相対座標を算出する．ただしドキュメント要素の大きさの変化によって相対座標が変化することを防ぐため，ドキュメント要素幅と高さがそれぞれ 1 となるよう座標変換することで正規化を実施する．つまり算出する相対座標は以下の式で表せる．

$$(x_{rc}, y_{rc}) = \left(\frac{g_x - a_x}{W_{obj}}, \frac{g_y - a_y}{H_{obj}} \right)$$

ここで g はドキュメント要素 t の重心座標， a は表 2 で定義したアノテーション開始・終了対応点の座標， W_{obj} と H_{obj} はドキュメント要素 t の幅と高さをそれぞれ表す．

人間の手書きアノテーションによるこれら相対座標の傾向から範囲選択ストロークの識別を行う．具体的には人間の手書きアノテーションから算出できる相対座標の分布を確率モデルに当てはめる．人間の操作によって表れる誤差は一般的に正規分布に従うことから，本認識対象の手書きアノテーションについてもズレの分布が正規分布に従うと仮定する．よって手書きアノテーションの相対座標のデータセットから正規分布の当てはめをおこなう．これによって得られた正規分布と入力された手書きストロークとの距離を定義することで，その距離をもとに範囲選択ストロークの識別を行う．本稿ではこの距離をアノテーション距離 (AD: Annotation Distance) と呼ぶことにする．AD はデータセットから得られた 2 変量正規分布と入力ストロークから得られた相対座標との距離を算出，すなわちマハラノビス距離をもとに次のように定義する．

$$AD(a, t) = \frac{1}{1 - \rho^2} \left(\frac{(x_{rc} - \mu_x)^2}{\sigma_x^2} + \frac{(y_{rc} - \mu_y)^2}{\sigma_y^2} - 2\rho \frac{(x_{rc} - \mu_x)(y_{rc} - \mu_y)}{\sigma_x \sigma_y} \right)$$

ここで $\rho, \sigma_x^2, \sigma_y^2, \mu_x, \mu_y$ は 2 変量正規分布モデルにおける相関係数，分散，平均をそれぞれ表している．

ページ内のすべてのドキュメント要素に対して入力手書きストロークにおけるアノテーション開始・終了対応点に対する AD を求め，得られた最小値がアノテーション開始・終了対応点ともに正規分布の信頼区間 99.9% に収まった場合は範囲選択ストローク，それ以外は情報付加ストロークとそれぞれ分類する．2 変量正規分布におけるマハラノビス距離の二乗値は自由度 2 のカイ二乗分布に従うことから，信頼区間 99.9% の境界を表す AD の閾値 AD_{thres} は以下のように算出できる．

$$AD_{thres} = \chi^2(2, 0.001) \sim 13.816$$

以上の手順により人間の手書きアノテーション傾向にもとづいた範囲選択ストロークの識別をおこなう．

3.5. 範囲選択箇所の推定

前節で範囲選択ストロークと識別された手書きストロークに対して，最後にドキュメント要素の選択範囲の推定を行う．前節で算出した AD が最小となるドキュメント要素 t_{target} をアノテーション開始・終了対応点，つまり以下の式で算出し，ドキュメント要素範囲選択箇所の始点・終点要素を求める．

$$t_{target} = \arg \min_t AD(a, t)$$

以上によりドキュメント範囲選択箇所が推定でき，コンピュータ側でアノテーション情報の認識を行うことができる．

4. 評価実験

本節では提案アノテーション認識モデルのパラメータ推定，およびモデルの認識精度の評価をおこなうために実施した，被験者からアノテーションデータを収集する評価実験について述べる．

4.1. 評価環境

まず実験の被験者や評価手順・環境について述べる．本収集実験の被験者は著者らの所属する大学の学生であり，被験者数は 10 名である．また実験を行うにあ

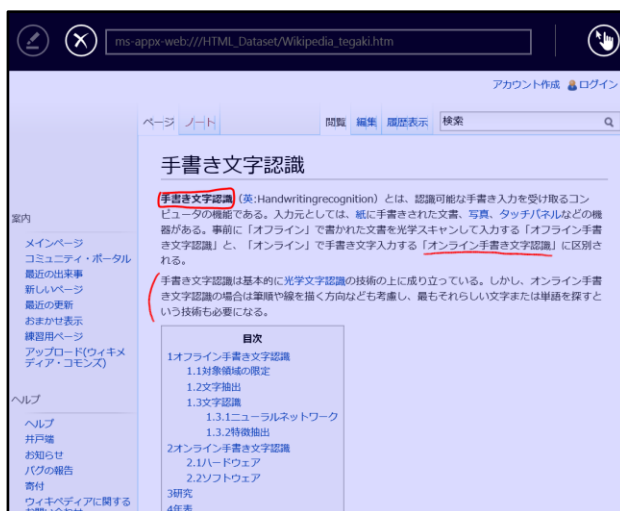


図 5 実験環境のスクリーンショット

たって手書きアノテーションデータを記録するアプリケーションを開発し、使用した。実験に使用したハードウェアは Sony VAIO Duo 11²である。ソフトウェアの開発環境は Windows8.1 および Windows ストアアプリケーションである。実験用アプリケーションのスクリーンショットを図 5 に示す。実験用アプリケーションは指定された URL が示す HTML ドキュメントに対して手書きアノテーションが可能となっている。搭載されている機能にはナビゲーションモードと書き込みモードの 2 つが存在する。ナビゲーションモードではドキュメントのリンク遷移やスクロールといったドキュメントに対する操作をタッチ操作でおこなうことができる。一方書き込みモードに移行するとドキュメントの位置が固定され、ペンによる書き込みのみを受け付ける状態となる。このモードでは手書きによる書き込み（アノテーション）と書き込みの消去（消しゴム）の操作のみ受け付ける状態となる。

被験者にはこの実験システムと紙に印刷された HTML ドキュメントを与えている。印刷された HTML ドキュメントには複数箇所がハイライトされており、被験者はハイライトされた箇所と同じ箇所を実験システム上に表示されているドキュメントから探し出し、手書きによる範囲選択アノテーションを行う。なお被験者には一回のアノテーションには一筆（1 ストローク）で完結させるよう指示した。実験では被験者に対して指定された 10 箇所をアノテーションし、これを 3 回繰り返す。アノテーションの種類は下線、囲い込み（単一行）、囲い込み（複数行）、縦線の 4 種類であり、被験者一人あたり合計で $10 \times 3 \times 4 = 120$ ストロークの

アノテーションを収集した。このようにすべての被験者から収集したアノテーションデータを用いてモデルのパラメータ推定および、認識精度の評価をおこなう。なお被験者がアノテーションに失敗した筆記データはあらかじめ取り除いている。

4.2. モデルパラメータの決定

本節では提案アノテーション認識モデルのパラメータ推定について述べる。前節で収集したデータを用いて正規分布フィッティングをおこない、提案モデルのパラメータを求めた。算出したパラメータを表 3 に示す。また、被験者から収集したアノテーションデータから算出したアノテーション開始・終了対応点の相対座標と正規分布フィッティングの結果をプロットしたグラフを図 6～図 13 にそれぞれ示す。点のプロットは実際の被験者から取得した相対座標データ、実線・破線の楕円はそれぞれ正規分布フィッティング時の 99.9%・95%信頼区間を表している。ここで示したパラメータはドキュメント中の文字サイズなど内容に依存するものではない、つまりドキュメントごとに依存しないように提案モデルを設計しているため、汎用的に本パラメータを用いることが可能と考えられる。

表 3 収集データより算出したパラメータ

		ρ	σ_x^2	σ_y^2	μ_x	μ_y
下線	開始	-0.102	0.050	0.015	0.454	-0.424
	終了	-0.016	0.080	0.015	-0.610	-0.424
囲込単行	開始	-0.022	0.093	0.008	0.727	0.040
	終了	-0.099	0.138	0.008	-0.747	0.040
囲込複数行	開始	0.102	0.494	0.066	1.839	0.926
	終了	-0.034	0.702	0.042	1.888	-0.781
縦線	開始	-0.030	0.773	0.031	1.704	0.599
	終了	0.231	0.465	0.054	1.563	-0.551

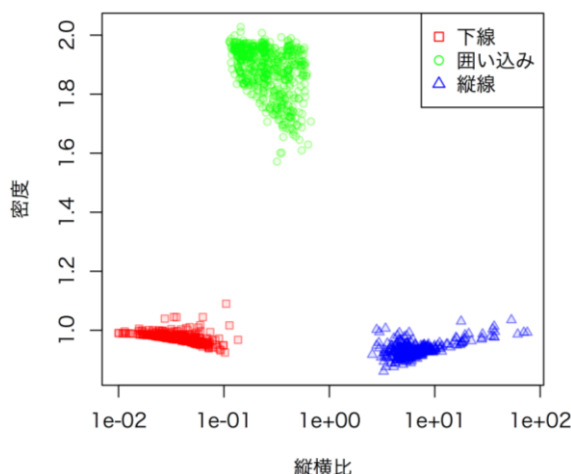


図 4 アノテーション分類の特徴量分布

² Sony VAIO Duo 11 (SVD1122AJ), <http://vcl.vaio.sony.co.jp/product/sv/svd1122aj.html>

4.3. 認識精度評価

パラメータの算出に加えて、被験者から収集したデータを用いて提案認識モデルの認識精度評価を実施した。まず 3.3 節で説明した形状特徴による範囲選択ストロークの分類における精度評価について述べる。被験者から収集したアノテーションデータに対して分類時に用いた特徴量を可視化した結果を図 4 に示す。この可視化結果を参照すると、縦横比と密度の 2 つの特徴量が分類する 3 クラスの範囲選択ストロークを分類する上で有効な特徴量であることが視覚的に確認できる。またこれらのデータに対して Leave-one-out 交差検定を実施したところ、100%の精度を達成することを確認した。

次に 3.5 節で説明した選択範囲の推定について検証をおこなった。本提案認識モデルを個人化せずとも堅牢な認識モデルであるかどうかを検証するために、被験者ごとのグループデータを分割し、交差検定をおこなった。交差検定によって算出した精度を表 4 に示す。

表 4 提案モデルのアノテーション認識精度

		Top-1	Top-2	Complete
下線	開始	95.99%	100.00%	88.63% (STD: 0.0658)
	終了	92.64%	100.00%	
囲込単行	開始	92.64%	97.99%	83.28% (STD: 0.2107)
	終了	86.62%	97.99%	
囲込複数行	開始	97.98%	98.99%	96.97% (STD: 0.0407)
	終了	97.98%	98.99%	
縦線	開始	98.99%	98.99%	96.98% (STD: 0.0586)
	終了	96.98%	98.99%	

表中の Top-1, Top-2 は各範囲選択ストロークの開始・終了オブジェクト個々の推定精度をあらわしており、Top-1 は完全に一致した場合のみを許容した場合の推定精度、Top-2 はドキュメント要素オブジェクトシーケンス上での前後 1 要素の誤認識までを許容した場合の推定精度をそれぞれ示している。また Complete は範囲選択開始要素、終了要素ともに完全に認識成功した場合の認識精度をあらわしている。

本実験結果をみると、行単位で要素を選択する囲い込み（複数行）と縦線による範囲選択は 96%以上の精度で完全に認識をおこなえていることがわかる。一方要素単位で選択する下線や囲い込み（単一行）による範囲選択は行単位の範囲選択と比較すると 10%程度認識精度が低下することを確認した。またこれらの認識精度の値は他人のみの統計データを用いて推定したパラメータを利用しているため、パーソナライゼーション、つまりユーザの初期使用時にパラメータ学習を行わずにこれらの認識精度を達成できることを示している。さらに Top-1 と Top-2 の精度値を比較すると、最大 10%程度の差があることが確認できる。すなわち前

後 1 要素の誤差を修正できればさらに最大で 10%の認識精度向上が見込めることがわかる。

5. まとめと今後の課題

本研究では電子ドキュメント上における手書きアノテーション環境での支援を目的とした、アノテーション認識モデルを提案した。また代表的なアノテーション 4 パターンを本認識モデルに適用し、評価を行った結果、筆記者に依存せずに堅牢な検出をおこなえることが確認できた。

本研究での結果を踏まえ、認識精度向上のための検討課題としては、確率分布モデルの比較検討や、コンテンツ情報を元に分割点の重み付けモデルの導入である。これらを検討することでより認識精度の向上が見込まれる。また、今後は本認識モデルを利用したインタフェースを利用し、実際のアプリケーションでの有用性についても検証を行う必要がある。

謝 辞

本研究の一部は JSPS 科研費・挑戦的萌芽研究 (No.25540040) によるものである。

参 考 文 献

- [1] Marshall, C.C., Annotation: from paper books to the digital library, In Proc. DL 1997, pp.131-140 (1997).
- [2] Hamzah, M.D., Tano, S., Iwata, M., and Hashiyama, T. Effectiveness of Annotating by Hand for non-Alphabetical Languages, In Proc. CHI 2007, pp.841-850 (2006).
- [3] Yoon, D., Chen, N., and Guimbretière, F. TextTearing: opening white space for digital ink annotation, In Proc. UIST 2013, pp.107-112 (2013).
- [4] Plimmer, B., Chang, S.H.-H., Doshi, M., Laycock, L. and Seneviratne, N. iAnnotate: exploring multi-user ink annotation in web browsers, In Proc. AUIC 2010, pp.52-60 (2010).
- [5] Agrawala, M., and Shilman, M., DIZI: a digital ink zooming interface for document annotation, In Proc. INTERACT 2005, pp.69-79 (2005).
- [6] Kaneider, D., Seifried, T., and Haller, M., Automatic annotation placement for interactive maps, In Proc. ITS 2013, pp.61-70 (2013).
- [7] Wang, X., Shilman, M., and Raghupathy, S., Parsing Ink Annotations on Heterogeneous Documents, In Proc. SBM 2006, pp.43-50 (2006).
- [8] 浅井洋樹, 山名早人: オンライン手書きノートからの強調語抽出, 日本データベース学会論文誌, Vol.10, No.1, pp.67-72 (2011).
- [9] Fonseca, M.J., Pimentel, C., and Jorge, J.A., Cali: An online scribble recognizer for calligraphic interfaces, In AAAI Symp. Sketch Understanding 2002, pp.51-58 (2002).
- [10] Wang, X., and Raghupathy, S., Ink Annotations and their Anchoring in Heterogeneous Digital Documents, In Proc. ICDAR 2007, pp.163-167 (2007).

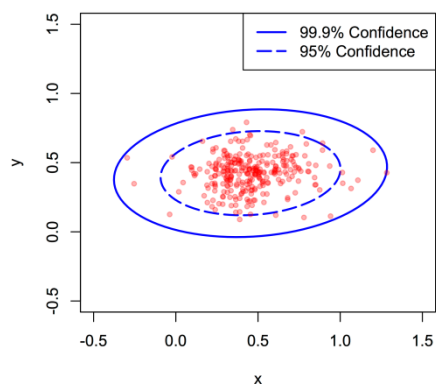


図 6 下線開始相対座標の分布

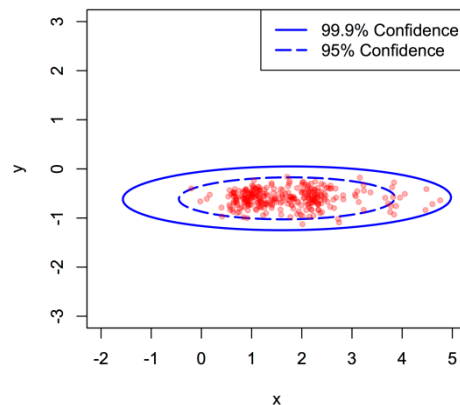


図 10 縦線開始相対座標の分布

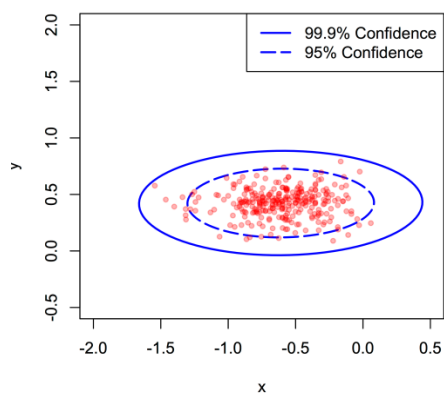


図 7 下線終了相対座標の分布

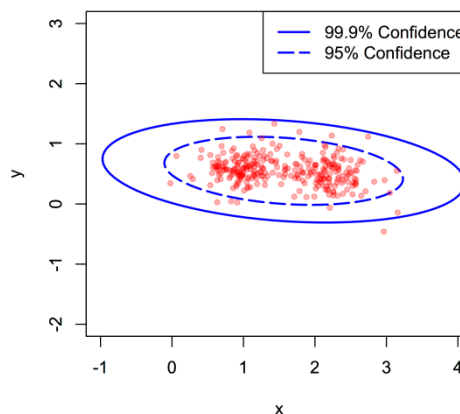


図 11 縦線終了相対座標の分布

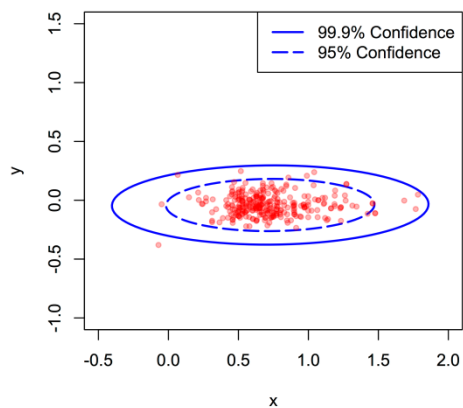


図 8 囲い込み（単一行）開始相対座標の分布

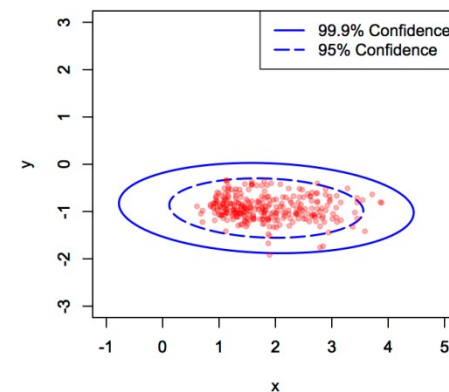


図 12 囲い込み（複数行）開始相対座標の分布

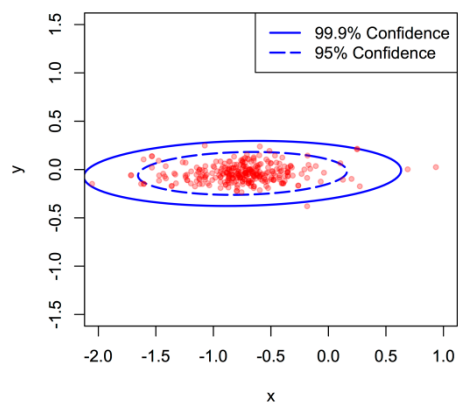


図 9 囲い込み（単一行）終了相対座標の分布

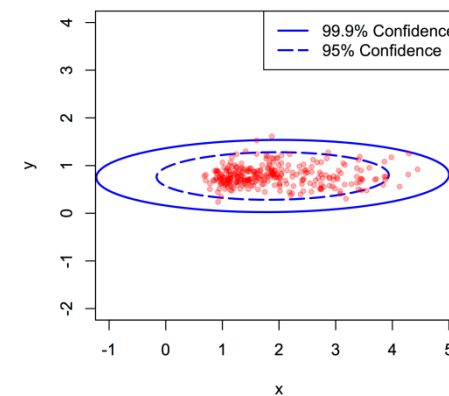


図 13 囲い込み（複数行）終了相対座標の分布