

視線を用いた自動辞書引きシステム

東中竜一郎 大野健彦

日本電信電話株式会社 NTT コミュニケーション科学基礎研究所
〒 243-0198 神奈川県厚木市森の里若宮 3-1
rh@atom.brl.ntt.co.jp, takehiko@brl.ntt.co.jp

概要

本稿では、電子文書の閲覧作業において、ユーザの視線情報から、ユーザの文書の理解度を自動的に測定し、自動辞書引きへの応用手法を提案する。具体的には電子文書中の単語領域に対する視線情報と、ユーザが辞書引きを必要とするかどうかを関連付ける関係式を決定木学習の手法を用いて作成し、辞書引き判定に用いる。被験者実験の結果、再現率が平均 17.5% (最大 36.4%, 最低 3.4%), 適合率が平均 29.9% (最大 58.1%, 最低 9.1%) でユーザの辞書引きを必要とする単語を推定できることが分かった。また、決定木の個人差は比較的大きいため、決定木は個人ごとに作成するのがよく、ユーザの辞書引き単語数が増加するにつれて、辞書引き単語の推定精度が向上することが示された。

キーワード: 視線, 視線測定システム, 自動辞書引き

1 はじめに

人とコンピュータの接する時間が増えるにつれ、コンピュータを利用して文書閲覧作業をおこなう機会も増加した。また、近年の文書検索技術は、従来閲覧の対象ではなかった文書の取得を可能にし、その結果、ユーザの閲覧する文書は量・多様性ともに増加している。このような背景において、文書閲覧の適切な支援手法に関する研究は、重要な研究課題である。

文書閲覧作業の支援手法の一つとして、文書内容の意味理解支援が挙げられる。意味理解支援とは、読み手にとって理解が困難な文書について、理解を改善する手法を指す。支援対象は主に、文書全体、段落、文、単語であり、特に単語は文書の最も基本的な構成要素であり、その知識の有無が文書理解を大きく左右することから、文書理解の観点において、単語理解を対象とした支援が最も重要である。

単語理解の支援としては、オンライン辞書が代表的であり、マウス操作によるポップアップ辞書は幅広く用いられている。しかし、閲覧文書数の増加と、それに伴う辞書引き単語数の増加を考えると、辞書を一度引く手間はわずかであっても、将来的に、その手間は膨大なものとなると思われる。そのため、ユーザが理解困難である単語について、自動的に辞書引きをおこなうシステムが実現できれば、その有用性は高いと考える。効果的な自動辞書引きを実現するためには、ユーザのモニタリングにより、辞書引き

が必要とされる単語を高い精度で推定する必要がある。また、そのモニタリング手法は、ユーザの本来の作業である文書閲覧作業をなるべく妨げないようにする必要があり、今回は、視線測定に着目した。

本稿では、文書閲覧作業におけるユーザの視線を用いた、ユーザが理解していないと思われる単語の自動推定手法と、自動辞書引きシステムへの応用について述べる。

以下、2章で関連研究を述べたのち、3章で本研究の課題について詳細に述べ、4章で本研究のアプローチについて説明する。5章で今回試作した視線による自動辞書引きシステムについて述べ、6章で提案手法の評価実験および実験結果について述べる。7章で考察を述べ、最後に結論と今後の課題を述べる。

2 関連研究

本研究は視線からユーザの意図を推定する研究の一種であるが、これまでそのような研究はあまり行われていない。なお、ここで言う、ユーザの意図とは、ユーザが実際に行いたいと思っている、または実際に行っている行為・表現などを指す。

Salvucci らは HMM (Hidden Markov Model) を利用して、視線からユーザの意図を自動的に推定する手法を提案した [7]。本手法はあらかじめユーザの作業を正規文法 (regular grammar) による書き換え規則集合として表現し (プロセスモデルと呼ぶ)、さらに各プロセスにおける視線の注視領域を定める。次に観察された視線データからどのプロセスモデルに

もっとも一致しているかを HMM によって算出する。Salvucci らは本手法を、視線キーボードを用いた単語入力システムに適用し、有効に機能することを確認している。本手法はプロセスモデルが定まっている場合には有効な手法であるが、本研究のように、ユーザの作業とは直接関係のない作業(本研究では辞書を引くという作業)を自動的におこなうことには適用できない。

ユーザが特定の作業をおこなっているか否かを視線から判定する手法として、Campbell らの Reading Detection がある [3]。本手法はユーザが文書読解中であるか、そうでないかを視線の幾つかの特徴的なパターンについて重み付けをおこない、そのスコアから判定している。Campbell らの研究は視線の特徴を利用する点において我々の研究と類似しているが、読解中であるかどうかの判定を経験則で判定しており、経験則の獲得には視線のパターンに熟知した専門家の知識を要するものと思われる。

高木は英文と和文の対訳コーパスを用いた翻訳支援環境を題材として、ユーザの「迷い」検出をおこなっている [8]。彼の試作した翻訳支援環境では英文と和文の対訳が矩形領域に並んでおり、ユーザの視線が各矩形領域間を遷移する状態を N-gram モデルを用いて分析し、ユーザの「迷い」と視線パターンの関係を求めた。N-gram で表される視線情報とユーザ意図である「迷い」を統計的に関連付けている点において本研究との類似点は多いが、扱う領域が例文単位と大きく、また検出された「迷い」を翻訳支援に用いることを目標としているものの、実際に支援をおこなうには至っていない。

3 課題

本研究においては、単語に対するユーザの理解度を視線からどのように推定するかが課題となる。算出方法の一つとして、設計者が経験則に基づき定める方法がある [3]。しかしながら、この手法には以下に述べる 2 点の問題点がある。

1. 精度の高い推定手法を設計するには、設計者の高い専門知識と時間・コストを必要とする。
2. 経験に基づき得られた推定手法が、必ずしもすべてのユーザにとって有効であるとは限らない。

例えば、視線停留時間が一定時間を越えた場合にその単語を理解していないと推定する手法の場合、ユーザが明確な目的を持たずにただ単語を見ていた場合も理解していないと判定してしまう。また、単語を理解していないと判定する最短停留時間の最適値はユーザによって異なる可能性もある。

以上の問題を解決するためには、それぞれのユーザ別に適切な推定基準を、設計者が特別な専門的知識を有することなく決定できる手法が必要となる。

4 アプローチ

本研究では、ユーザ毎の適切な推定基準を、機械学習により取得するアプローチを提案する。

具体的には、文書を複数の領域に分割し(今回は領域が単語単位となる)、それぞれの領域に関する視線情報を複数の特徴量で表す。同時にその領域に対するユーザの意図(今回は辞書引きが必要かどうか)を取得し、得られた特徴量とユーザの意図を関連付ける関係式を決定木を用いた機械学習の手法により求めることで、ユーザの意図を推定する。

4.1 単語領域に関する視線の特徴量

ユーザが文書を閲覧する際の視線の軌跡の例を図 1 に示す。図の丸い点は視線の停留点を示し、各停留点は停留順序に基づき直線で結ばれている。停留点から次の停留点へは視線のサッケード(跳躍運動)によって移動している。

本図を見ると、ユーザの視線は各単語領域に対し、領域への進入、領域内の移動、領域からの退出を繰り返しながら移動していることがわかる。そこで文書閲覧作業におけるユーザの停留点情報を、文書に含まれる各単語において求め、視線の特徴量とする。

具体的な特徴量は、表 1 に列挙する 14 種類である。表における「方向」とは視線のサッケード方向を示したものであり、円を 45 度単位で 8 方向に分割した扇形領域のいずれに属しているかを示す。扇形領域は -22.5 度から 22.5 度を方向 0 とし、反時計方向に 45 度単位で方向 7 までである。それぞれの領域について、何回該当するサッケードが発生したかを示す(図 2)。特徴量 1, 2, 3, 4 および 5 はそれぞれ方向 0 から方向 7 までの 8 属性を持ち、残りの 9 特徴量は各 1 個の属性を持つため、特徴量全体は 49 次元ベクトルである。ただし、瞳孔に関する特徴量 9 12 は、実験条件によって大きく変わることから、決定木作成には用いない。

4.2 単語領域に関するユーザ意図

ユーザの意図は、文書中の各単語に対してユーザが辞書引きを必要とするかどうかの 2 値で与える(辞書引き要求)。ユーザが辞書引きを必要とするときを 1 で表し、そうでないときを 0 で表す。辞書引きに関するユーザ意図は、実際にユーザが該当する単語に関する辞書引きをおこなうかどうかを観察することで得られる。

As the All-Star Series 2002 came to a thrilling conclusion on Sunday at Tokyo Dome in a 4-2 victory and a 4-3 series win for the visiting major leaguers, the race is just beginning for Hideki Matsui. The Yomiuri Giants free agent outfielder, who had a disappointing series, will now begin negotiating with major league teams. Unfortunately for the fans, who poured their hearts out for every one of the 28-year-old's plate

図 1: 視線の軌跡

特徴量 番号	説明
1.	対象となる単語に関してどの方向へ何度視線が移動したか
2.	対象となる単語に進入する視線がどの方向へ何度移動したか
3.	対象となる単語から退出する視線がどの方向へ何度移動したか
4.	対象となる単語の内部でどの方向へ何度視線が移動したか
5.	対象となる単語に関してどの方向へどのくらいの大きさを視線が移動したか
6.	対象となる単語のテキスト長
7.	対象となる単語の総停留時間
8.	対象となる単語に視線が関与した回数
9.	瞳孔の大きさの最小値
10.	瞳孔の大きさの最大値
11.	瞳孔の大きさの平均値
12.	瞳孔の大きさの標準偏差
13.	視線の移動距離の総和
14.	視線の移動距離の平均値

表 1: 各単語における視線の特徴量

4.3 文書閲覧時の単語理解度判定

以上で定めた視線の特徴ベクトルおよびユーザ意図の対応関係から、決定木学習 (C5.0[6]) を用いて辞書引き判定に用いる決定木を生成する。理解度の判定においては、文書閲覧中のユーザの視線から停留点

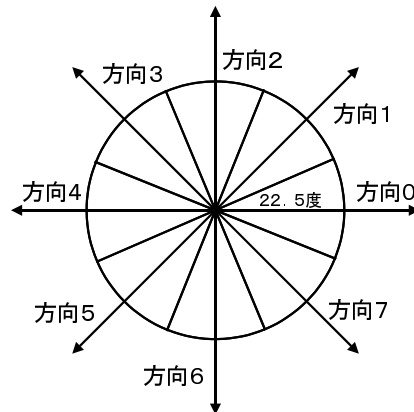


図 2: 視線のサックード方向

情報を取得し、決定木への入力とする。辞書引きを必要とすると判定した場合には、ただちに辞書引きをおこない、その結果を画面に表示する。

5 辞書引きを自動的におこなう文書ブラウザ

提案した単語理解度測定法および自動辞書引きシステムの動作を確認するために、英単語の自動辞書引きが可能な文書ブラウザを実装した。以下にその詳細を示す。

5.1 構成

試作した文書ブラウザは、ユーザの視線から決定木を生成する決定木構築段階および得られた決定木を利用して自動的に英単語の辞書引きをおこなう自動

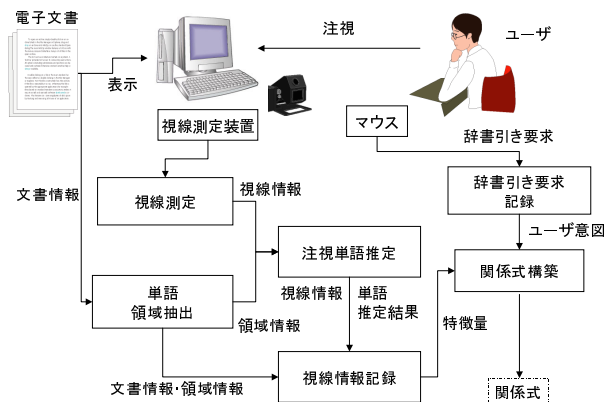


図 3: 決定木構成段階におけるシステムのフロー

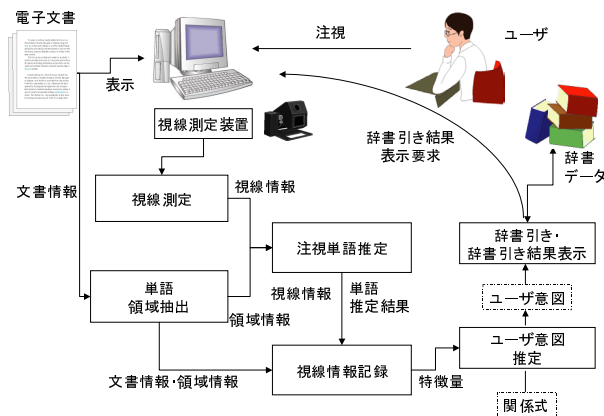


図 4: 自動辞書引き段階におけるシステムのフロー

辞書引き段階の 2 段階で動作する。

決定木構築段階 決定木構築段階はユーザの視線から決定木を生成する。動作フローを図 3 に示す。決定木構築段階では、まず閲覧用の文書から、文書に含まれる各単語の矩形領域座標を決定する単語領域抽出をおこない、次に、各単語について、WordNet[4] を用いて原型および品詞情報を取得しておく。ユーザが文書閲覧を開始すると、我々の開発した視線測定装置 FreeGaze[5] を用いて視線から停留点データを生成する。次に停留点データからユーザの注視している単語を決定し、該当する単語の特徴ベクトルを更新する。

ユーザが辞書引きをおこなうときは、該当する単語をマウスでクリックすれば良い。このとき、ユーザ意図としてクリックした単語の情報がファイルに記録される。

文書閲覧の終了後、得られた特徴ベクトルおよび辞書引きデータから、単語理解度判定用の決定木を生成する (図 6)。

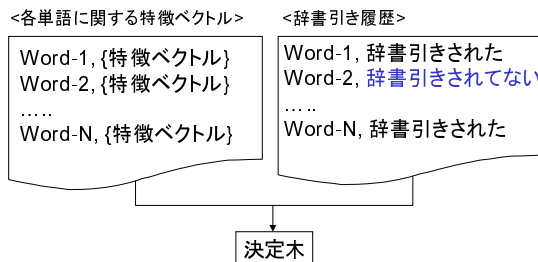


図 6: 決定木の作成

自動辞書引き段階 自動辞書引き段階のフローを図 4 に示す。自動辞書引き段階では、文書閲覧作業における視線の停留情報を取得して、文書中の各単語における特徴ベクトルを順次更新する。同時に、更新された特徴ベクトルを決定木で判定し、ユーザが単語を分からないと判定した場合には、画面左側には該当する単語の辞書引き結果を表示する。

5.2 実装システム

実装したシステムは Windows2000 上で動作し、Delphi, Visual C++ および Perl を用いて作成した。電子文書の表示には Internet Explorer 6.0 を Active X として使用した。英単語の品詞解析には WordNet を、辞書データには英辞郎 [2] を用いた。視線測定には FreeGaze を利用している。FreeGaze は 30Hz で視線測定可能であり、測定精度は視野角 0.8 度程度である。

ユーザの閲覧する画面のスクリーンショットを図 5 に示す。画面中央部には文書表示ウィンドウが表示されている。任意の HTML 文書を表示可能であるが、現在は測定した視線がユーザの注視している単語からずれることを防ぐため、行間をやや広く空けて表示している。辞書引き結果は画面左部に表示されている。ユーザの閲覧作業によって単語が次々と表示され、ユーザの希望する単語の辞書引き結果が消去されることを防ぐため、最新の単語は上部に、過去の単語は下部に、単語別のウィンドウでタイリング表示される。

視線測定風景を図 7 に示す。実際の利用においては、視線測定装置 FreeGaze が顔の動きを許容しないことから、実験途中で視線測定精度が低下することを防ぐため、顎載せ台を使用した。

6 単語理解度推定法の有効性評価

提案した単語理解度推定手法の有効性を確認するために、試作した文書ブラウザを利用した単語理解度の推定実験をおこなった。

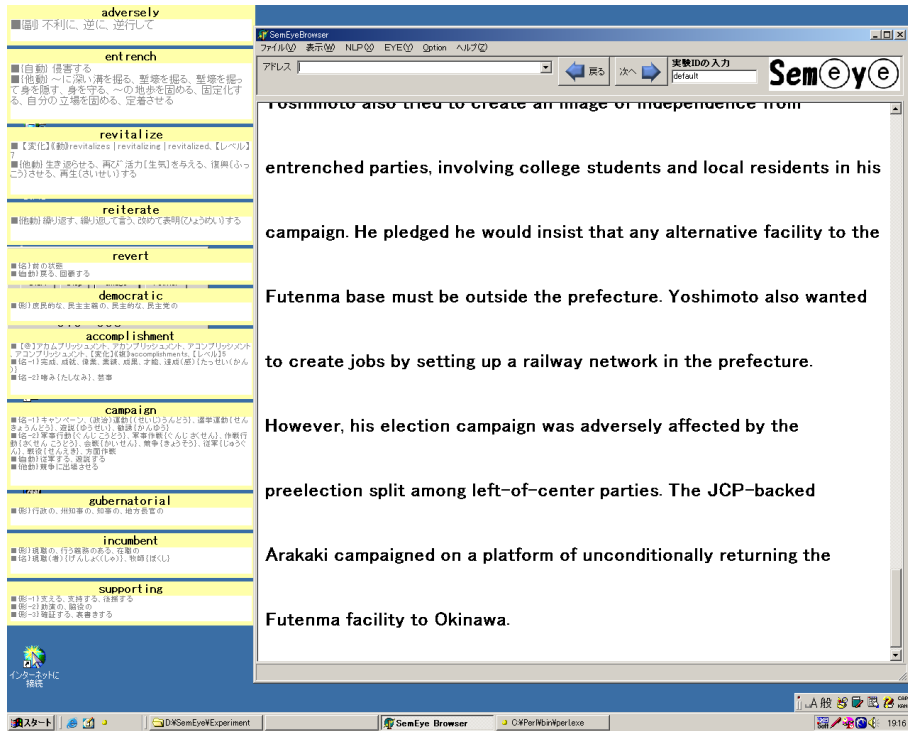


図 5: 試作した文書ブラウザのスクリーンショット



図 7: 視線測定風景

6.1 単語理解度の推定実験

実験手順 まず、視線測定装置の個人キャリブレーションをおこなない、個人パラメータを記録した。次に、十分な視線測定精度が得られていることを確認するため、ユーザが注視する単語と測定された単語が一致することを確認するツールを用い、単語を注視したときに確かに視線が単語領域内に含まれていることの確認をおこなった。

その後、被験者に文書ブラウザを用いて英文を提示し、あらかじめキャリブレーションによって得られた個人パラメータを読み込み、視線測定を開始した。

このとき被験者に、文書の内容を理解しながら文書閲覧をおこなうよう指示を与えた。途中、理解できない単語があった場合には、マウスを利用して単語をクリックすると、画面左側に単語の意味が表示されるようになっている。被験者には「文書の意味が理解できるように文書を読んでください」と指示を与え、単語のクリック回数は制限なしとした。

被験者が文書を末尾まで読み終えたとき、被験者は実験者に合図をおこない、実験者は次の文書を提示した。以上の作業を、被験者が題材の文書をすべて読み終えるまでおこなった。

被験者 被験者は著者を含む研究者および研究補助者 5 名である。本実験の参加に際して謝礼の支払いはおこなっていない。

題材 被験者に提示する題材には Daily Yomiuri Online[1] よりニュース 5 文書を用いた。文書内の単語数とトピックは表 2 のとおりである。

6.2 実験結果

推定実験において各被験者が辞書引きを行った単語数および文書全体の単語数に占める割合(辞書引き率)を表 3 に示す。本結果を見ると、辞書引き率には個人差が大きいことがわかる。

次に、実験によって得られた視線データから文書中の各単語における特徴ベクトルを生成し、辞書引

単語数	トピック
565	稲嶺知事再選
199	しし座流星群
236	日米安保
405	松井のメジャー挑戦
300	英語スピーチコンテスト
1705	

表 2: 実験に用いた文書のトピックおよび単語数

	辞書引き単語数	辞書引き率
被験者 1	27	1.6%
被験者 2	29	1.7%
被験者 3	19	1.1%
被験者 4	121	7.1%
被験者 5	135	7.9%

表 3: 各被験者の辞書引き単語数とその割合

き履歴と合わせて決定木学習をおこなった。導出に用いたデータの一例を表 4, 導出の結果得られた決定木の例を図 9 に示す。

決定木における duration は総停留時間 (特徴量 7), directions, vector_sizes, はそれぞれ表 1 の特徴量 1, 5 に対応し, 後続の数字は方向 0~7 (図 2) を示す。方向 7 への視線の移動距離が 476 を超え, 総停留時間が 2143ms を超える場合や, 方向 7 への視線の移動距離が 476 以下で, 総停留時間が 1623ms を超え, 方向 6 への視線の移動距離が 162 以下で, 方向 1 への移動距離が 18 以下の場合などに, その単語は辞書引きが必要であると推定される。なお, 移動距離の単位は pixel である。

次に, 導出された決定木を用いて, ユーザ理解度推定を行った。10 fold cross validation による推定結果を図 8 に示す。再現率は平均 17.5% (最大 36.4%, 最低 3.4%), 適合率は平均 29.9% (最大 58.1%, 最低 9.1%) であった。

7 考察

7.1 英単語の自動辞書引き

実験によって生成された決定木を文書ブラウザに組み込み、英文の文書閲覧をおこなったところ, 分からない単語の前後を見ていると実際に自動辞書引きがおこなわれた。停留時間のみを辞書引きの判定に用いるシステムと使用感を比較したところ, 不必要な辞書引きが削減されたように感じられた。しかし, 辞書引きが必要とされる場合に辞書引きがされないであったり, どのように単語を見ている場合に辞書

説明	データ
単語 id	id69
基本形	constellation
品詞	noun
特徴量 1	0,1,0,0,0,0,12
特徴量 2	0,0,0,0,0,0,2
特徴量 3	0,1,0,0,0,0,2
特徴量 4	0,0,0,0,0,0,8
特徴量 5	0,87,0,0,0,0,0,7739
特徴量 6	13
特徴量 7	3286
特徴量 8	13
特徴量 9	34.71
特徴量 10	38.16
特徴量 11	35.70
特徴量 12	1.07
特徴量 13	7826
特徴量 14	602
辞書引き	1 (必要)

表 4: 決定木作成に用いたデータの例

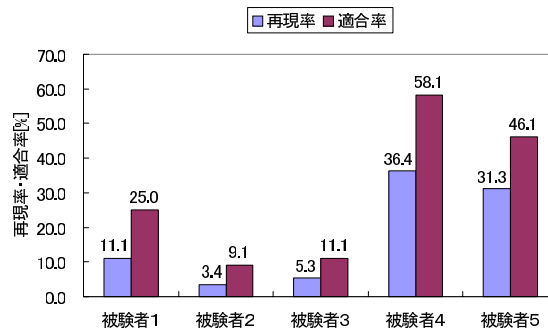


図 8: 被験者毎の再現率と適合率

引きがされるのか分からないであったりと, コントローラビリティの面で問題があるように思われた。

7.2 学習木の個人適合度

被験者間で決定木がどの程度共通であるかを調べるため, ある被験者のデータで学習した決定木を用い, 他の被験者のデータをテストデータとして再現率・適合率を求めた。その結果を図 10 および図 11 に示す。ただし, 図では同一被験者間の再現率・適合率は 0% としてある。

被験者によっては類似した決定木が構築されたものもあるが, そうでないものがほとんどであり, 再現率・適合率ともに決定木には大きな個人差が見られた。

また, すべての被験者のデータを用いてひとつの

```

vector_sizes_7 > 476:
...duration <= 2143: no (46/7)
: duration > 2143: yes (26/3)
vector_sizes_7 <= 476:
...duration <= 1623: no (2311/10)
duration > 1623:
...vector_sizes_6 > 18:
...vector_sizes_6 > 162: no (7)
: vector_sizes_6 <= 162:
: ...vector_sizes_1 <= 18: yes (12)
: vector_sizes_1 > 18: no (4)
vector_sizes_6 <= 18:
...duration > 2515:
...directions_0 <= 4: yes (10/1)
: directions_0 > 4: no (29/7)
duration <= 2515:
...vector_sizes_0 > 231: no (109/3)
vector_sizes_0 <= 231:
...directions_7 <= 0: no (12/2)
directions_7 > 0: yes (4)

```

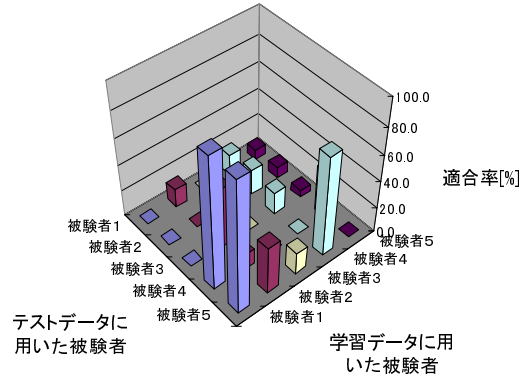


図 11: 被験者間の決定木の適合度 (適合率)

	5 文書	10 文書
被験者 1	11.1%	46.2%
被験者 2	3.4%	35.2%

表 5: 被験者 1 および被験者 2 における 5 文書と 10 文書での再現率

図 9: 導出された決定木の例

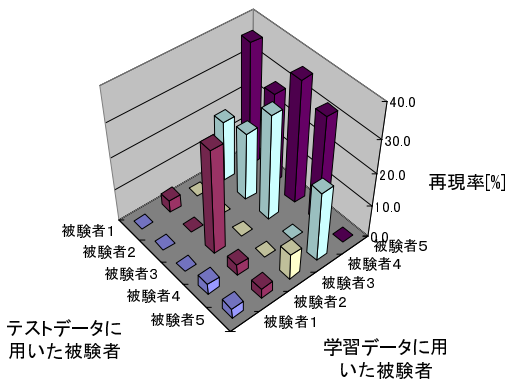


図 10: 被験者間の決定木の適合度 (再現率)

決定木を作成したときの再現率・適合率はそれぞれ 15.1% (49/324), 66.2% (49/72) であった。再現率の低さを考えると、辞書引き判定には複数の被験者による視線データから得られた決定木は有効であるとはいえない。

7.3 辞書引き単語数と学習結果の関係

被験者によっては、辞書引き単語数が余りに少ないため、決定木の学習が効果的におこなわれなかった可能性がある。そのため、被験者 1 と被験者 2 を対象に、やや難度の高い文書を 5 文書追加し、全体で 10 文書とし、決定木の学習をおこなった。被験者 1 と被験者 2 の辞書引き単語数は 27, 29 からそれぞれ 104, 89 に増加し、その結果、再現率と適合率は表 5

および表 6 のように変化した。辞書引き単語数が増加すると再現率が上昇していることがわかる。また、図 12 と図 13 に、被験者 1 と 2 における、辞書引き単語数と再現率・適合率の推移のグラフを示す。グラフでは、辞書引き単語数が 100 に近づくにしたがって、傾きがゆるくなっているため、決定木の学習に必要な辞書引き数が 100 程度であると予想される。

自動辞書引きという作業を考えたとき、適合率よりも再現率の方が重要である。なぜなら、必要とされるときに辞書を引かないことよりも、多少間違っても多めに辞書引きをする方が文書理解支援の観点から、リスクが少ないと考えられるからである。よって文書数の増加によって再現率が上昇したことは、学習結果が良くなったと考えることができる。

8 結論と今後の課題

本稿では、視線情報を利用した単語理解度の推定および自動辞書引きへの応用手法を提案した。具体的には電子文書中の単語領域に対する視線情報と、ユーザが辞書引きを必要とするかどうかを関連付ける関係式を決定木学習の手法を用いて作成する。

被験者実験の結果、個人毎に決定木を作成した場合、再現率が平均 17.5% (最大 36.4%, 最低 3.4%), 適合率が平均 29.9% (最大 58.1%, 最低 9.1%) でユーザの辞書引きを必要とする単語を推定できることが分かった。また、すべての被験者のデータを用いてひとつの決定木を作成したときの再現率・適合率は

	5 文書	10 文書
被験者 1	25.0%	59.3%
被験者 2	9.1%	57.4%

表 6: 被験者 1 および被験者 2 における 5 文書と 10 文書での適合率

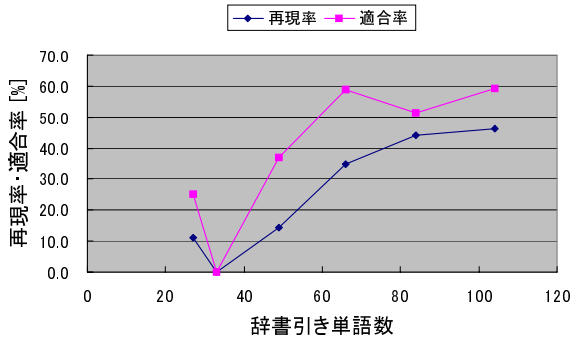


図 12: 被験者 1 における辞書引き単語数と再現率・適合率の推移

それぞれ 15.1% , 66.2%であった。決定木における個人間の差は大きく、すべての被験者の視線データから得られた決定木の再現率が低いこと、ならびに学習に用いる辞書引き単語数が増加するにつれ学習結果が良くなることから、辞書引き判定には、個人に特化した決定木を利用するのがよく、学習に用いる辞書引き単語数も、ある程度必要であると考えられる。今回の実験では、100 単語程度の辞書引き単語数は必要であるとの結果を得た。

よって、視線による自動辞書引きシステムは、個人個人のユーザにおいて、わからない単語の辞書引きをおこなわれるににつれて、決定木を更新し、再現率を向上させ、最適な自動辞書引きをおこなうシステムに発展させるのがよいだろう。

今回は自動辞書引き単語の推定法に関する評価実験をおこなったが、今後はコントローラビリティを考慮の上、自動辞書引きシステム自体の評価実験を行い、システムの使いやすさを検証する予定である。

謝辞

本研究をご支援いただく NTT コミュニケーション科学基礎研究所メディア情報研究部の村瀬洋部長、貴重なご意見を頂いたマルチモーダル対話研究グループ、環境理解研究グループの諸氏に感謝します。またお忙しい中、実験にご協力いただいた方々に深謝します。

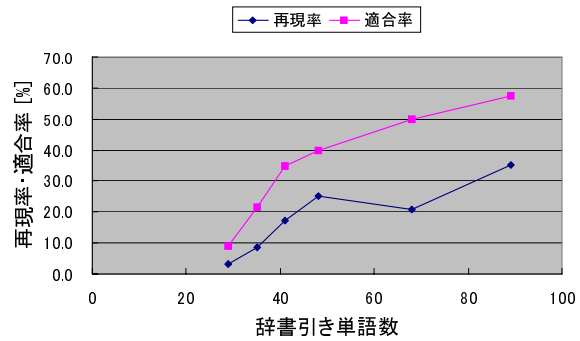


図 13: 被験者 2 における辞書引き単語数と再現率・適合率の推移

参考文献

- [1] Daily yomiuri on-line. <http://www.yomiuri.co.jp/index-e.htm>.
- [2] 英辞郎. <http://www.nifty.ne.jp/eijiro/index.htm>.
- [3] Christopher S. Campbell and Paul P. Maglio. A robust algorithm for reading detection. In *Proceedings of the ACM Conference on Perceptive User Interfaces (PUI '01)*, 2001.
- [4] Christiane Fellbaum. *WordNet An Electronic Lexical Database*. The MIT Press, 1998.
- [5] Takehiko Ohno, Naoki Mukawa, and Atsushi Yoshikawa. Freegaze: A gaze tracking system for everyday gaze interaction. In *Proceedings of Eye Tracking Research & Application (ETRA2002)*, pages 125–132, 2002.
- [6] J. R. Quinlan. *C4.5: Programs for Machine Learning*. The Morgan Kaufmann, 1992.
- [7] Dario D. Salvucci and John R. Anderson. Automated eye-movement protocol analysis. *Human-Computer Interaction*, 16:39–86, 2001.
- [8] 高木 啓伸. 視線の移動パターンに基づくユーザの迷いの検出-効果的な作業支援を目指して. *情報処理学会論文誌*, 41(5):1317–1327, 2000.