

文書に対する視点情報を利用した情報表示システムの提案

鈴木 若菜[†] 外山 託海^{††} 黄瀬 浩一[†] AndreasDengel^{††}

[†] 大阪府立大学大学院工学研究科

〒 599-8531 大阪府堺市中区学園町 1-1

^{††} ドイツ人工知能研究センター (DFKI)

Trippstadter Strasse 122 D-67663 Kaiserslautern, Germany

E-mail: [†]wakana@m.cs.osakafu-u.ac.jp, ^{††}Takumi.Toyama@dfki.uni-kl.de, ^{†††}kise@cs.osakafu-u.ac.jp,

^{†††}Andreas.Dengel@dfki.de

あらまし 本稿では、ユーザの読書をアシストするシステムを提案する。これは、読書中に本文のキーワードに関する情報が欲しい場合、ユーザが HMD を見ると、ディスプレイにその情報を表示するシステムである。本システムでは、装着型のアイトラッカとシースルー型のヘッドマウントディスプレイを組み合わせたデバイスを用いる。加えて、文書画像検索を用いてユーザが読んでいる文書を特定し、さらに、その文書のどこを読んでいるかをアイトラッカの注視点情報から特定する。ユーザは、ヘッドマウントディスプレイを見ると、直近に読んだキーワードを起点とする語彙集を参照し、キーワードに関する追加情報を得ることができる。したがって、本システムを利用することで、文書の本文中に無い情報を欲しいときにすぐ取得することができ、人と文書との関わりをより豊かにすることができる。本稿では、このシステムの基礎的な性能を評価し、文書画像検索と注視点情報を利用した、ユーザの利便性を重視したシステムの可能性について考察する。

キーワード 装着型アイトラッカ, シースルー型ヘッドマウントディスプレイ, 文書画像検索, Gaze Interaction

1. はじめに

日常生活において、新聞や書籍、資料など、文書に触れる機会は多々ある。文書を読んでいる最中に、知らない語句や意味の分からない語句があったとき、その媒体が印刷文書の場合、辞書を持ち出してきて引いたり、コンピュータに向かい、インターネットで検索したりすると考えられる。一方、電子文書の場合であれば、その語句をコピーし、検索ブラウザを立ち上げ、そこにコピーした語句をペーストして検索すると考えられる。しかし、情報が欲しいときにユーザが読んでいる場所の近くに知りたい情報がすぐに表示されれば、辞書を持ち出したり、コンピュータに向かう手間を省くことができる。また、ブラウザを立ち上げたり、該当語句をコピー・ペーストする必要がなく、より便利であると考えられる。

そこで、本研究では、印刷文書や電子文書といった媒体に依存せず、ユーザが欲しいときに欲しい情報をすぐに提示するシステムを開発することを目的とする。具体的には、文書中のユーザが読んでいるキーワードを特定し、ユーザが HMD (ヘッドマウントディスプレイ) を見ると、そのキーワードに関する情報を HMD に表示するシステムである。本研究の目的を達成するようなシステムを実現するためには、必要な要素が 3 つある。

第一に、文書に応じて固有の情報を提示するため、ユーザが読んでいる文書を特定する必要がある。これには、文書画像検



図 1 アイトラッカと HMD を組み合わせたデバイス

索が利用できる。文書画像検索は、カメラで撮影した文書画像とそれに該当する登録文書画像を検索する技術である。カメラベースの文書画像検索手法には、[1] や [2] などが挙げられる。しかし、これらの手法は、データベースが小規模であったり、処理時間が長いという問題点がある。そこで、本システムでは、それらの問題を解決する LLAH (Locally Likely Arrangement Hashing) [3] と呼ばれる手法を利用する。

第二に、ユーザが文書上のどのキーワードを見ているかを特定する必要がある。これが特定できれば、ユーザが欲しい情報を推定し、提示する情報を決定することができる。なぜならば、文書を読んでいて情報が欲しいとき、その語句をユーザは読んでいると考えられるためである。これには、ユーザの眼の

動きを追跡することで、シーン映像中のどこを見ているかを特定するアイトラッカを用いることが考えられる。アイトラッカには、据置型と装着型があり、据置型のアイトラッカを利用したアプリケーションやサービスとして [4] [5] が挙げられる。しかし、あらゆる場面で文書を読むことを考慮し、本システムでは、読む環境を制限しない装着型のアイトラッカを利用することが適していると考えられる。

第三に、ユーザが文書を読んでいるすぐ近くに情報を表示するための表示デバイスが必要である。これには、目の前の風景を透過して見ることができるシースルー型の HMD が利用できると思われる。シースルー型の HMD であることから、ディスプレイが読書の妨げにはならず、さらに、印刷文書でも電子文書でも利用することができる。HMD は、図 1 に示すように、メガネのようにして装着するディスプレイ装置である。シースルー型の HMD を利用することで、読書の妨げにならず、また、印刷文書であっても電子文書であっても利用可能であると考えられる。そして、装着型であることから、目の前のディスプレイを視るのみの、手間いらずで情報を得ることができる。

本稿で提案するシステムでは、図 1 に示すように、装着型のアイトラッカとシースルー型の HMD を組み合わせたデバイスを使用する。そして、アイトラッカのシーンカメラで撮影した文書画像を検索質問として文書画像検索をする。本システムで利用する文書画像検索手法は、外観をキーとして検索する手法であるため、印刷文書でも電子文書でも検索可能である。そして、特定された文書とアイトラッカからの注視点情報からユーザが文書上のどこを見ているかを特定し、ユーザが直近に読んだ語句(キーワード)に関する補足説明や追加の情報をテキストおよび画像などを用いて HMD に表示する。本システムでは、ユーザが HMD に視線を移し、数秒間見続けると、キーワードに関する情報が目の前のディスプレイに表示され、それを参照することができる。また、本稿では、本システムの基礎的な性能を調査するために行った実験について述べ、システムの有益性について考察する。

2. 関連研究

この節では、本システムの関連研究について紹介する。

まず、文書画像検索に関連する研究について紹介する。HOT-PAPER [1] は、文書のテキストに基づく特徴量を用いる手法の 1 つである。この手法は、処理が高速であるという特徴を持ち、1 検索質問あたり約 300ms で検索可能である。しかし、報告されているデータベースのサイズは非常に小さく、5000 ページ以下である。加えて、検索精度も 60% 以下である。このようなスケーラビリティの悪さは、ユースケースが考慮されておらず、提供できるサービスに限りがあると考えられる。Mobile Retriever [2] も文書中のテキストに基づく特徴量を用いる手法である。この手法では 2 つの特徴量を用いており、複数の単語の shape code で定義される特徴量と、3 つの単語とそのオリエンテーションで構成される特徴量を使用している。この手法は、ある程度大規模なデータベースを用いて、高精度で検索することができる。しかし、単語の認識に文字認識技術の一つである

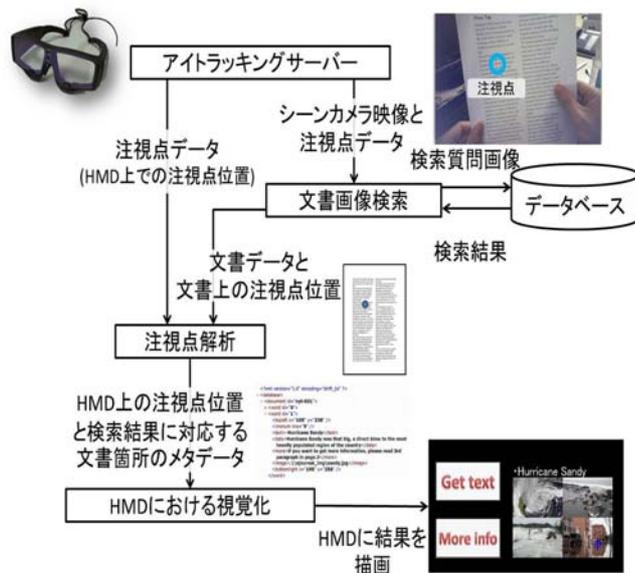


図 2 システムの流れ

OCR を用いているため、1 検索質問あたり 4 秒かかり、長い処理時間を要する。

次に、人の注視点を利用した研究や読む行為に関する研究について紹介する。人々が見たり読んだりしている場所を知ることができる技術として、アイトラッキング技術が挙げられる。人々の読書における注視点に関するテーマは、アイトラッキング分野において、長年、話題となっているものの 1 つである [6] [7]。この分野では、[4] などのように、人の注視点を利用したアプリケーションも開発されている。[4] では、コンピュータのディスプレイに表示した文書に対して、人の目の動きに基づいていくつかのインタラクティブなアクションやサービスをリアルタイムで読者に提供する。[5] は、アイトラッカを用いて、Web ページを閲覧している時のユーザの注視点を利用したシステムである。これらの人の視点を利用したインタラクティブなアプリケーションは、据置型のアイトラッカに依存している。また、[8] は、多重センサである body-worn センサを用いて、文書を読んでいるという行為を認識する手法を提案している。さらに、[9] は、HMD を用いたシステムで、注視点を用いて実世界のマーカーを選択し、システムをコントロールする。

3. 提案システム

本システムの流れを図 2 に示す。まず、アイトラッキングサーバからオンラインでシーンカメラの映像と注視点データを文書画像検索モジュールと注視点解析モジュールに送る。文書画像検索モジュールでは、ユーザが読んでいる文書に対応する文書画像を予め登録されているデータベースから検索する。さらに、検索結果の文書画像のどこを読んでいるかも特定する。注視点解析モジュールでは、ユーザが文書を見ているのか、HMD を見ているのかをチェックする。ユーザの注視点文書上のキーワード領域にあれば、XML ファイルに記述された対応するメタデータを HMD 視覚化モジュールに送る。そして、

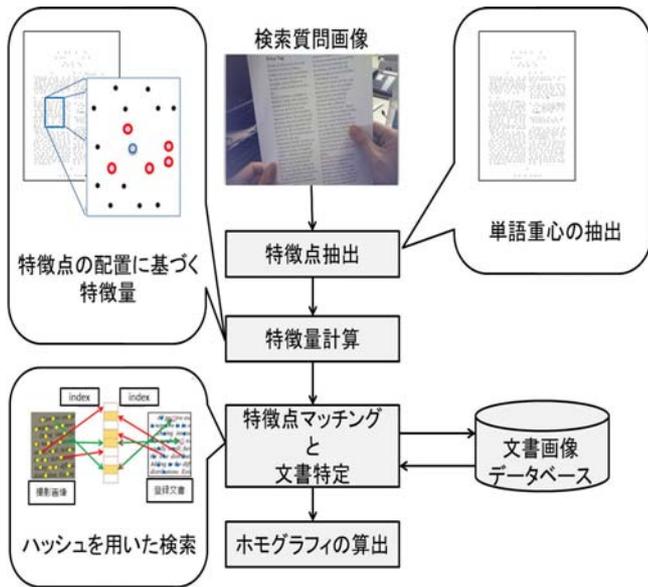


図3 文書画像検索の処理概要

ユーザがHMDを見れば、その情報が表示される。

3.1 アイトラッキング

本システムでは、SMI社の装着型アイトラッカETG(Eye Tracking Glass)を使用する。このデバイスには、ユーザの両眼を映すアイカメラが2つと、ユーザのしているシーンを撮影するシーンカメラが1つ内蔵されている。アイカメラでは、左右それぞれの眼球運動を検出する。眼球運動を検出するために、瞳の周りに6点の赤外光を照射し、眼で反射させる。このアイトラッカを使用するには、ユーザはまず、キャリブレーションをする。キャリブレーションでは、ユーザが実際のシーン中の1点または3点を見ることでユーザが実際に見ている点とシーンカメラ映像中の対応する点との位置合わせをする。キャリブレーションは、たいていの場合、長い時間を要さないが、一回うまく合わない場合があり、ユーザによっては正確にキャリブレーションできるまで何回かやり直しが必要な場合がある。本システムで使用するアイトラッカのサンプリングレートは両眼で30Hzであり、視野精度は 0.5° である。

3.2 文書画像検索

本システムでは、文書特定のために、文書画像検索手法の一つであるLLAH[3]を利用する。LLAHは、現実的な利用において生じる撮影方向の変化や隠れ、紙面の湾曲などの外乱にロバストである。加えて、リアルタイムで検索できるほどの高速性を持つという特徴がある。さらに、[10]では、[3]の手法を改良することによって、1000万ページのデータベースでリアルタイムでの検索が可能であることが報告されている。文書画像検索の処理概要を図3に示す。まず、カメラからの撮影画像を適応2値化し、2値画像を得る。得られた2値画像をガウシアンフィルタでぼかし、再度適応2値化すると、単語ごとに連結された画像が得られる。そして、連結成分の重心を計算して特徴点とする。次に、得られた特徴点の配置から特徴量を計算する。そして、検索処理では、抽出された特徴量と予め登録

されているデータベース中の文書画像の特徴量とのマッチングをとる。ハッシュ表を利用することで高速な処理を可能にしている。最後に、アイトラッカのシーンカメラからの映像と検索された文書画像との特徴点のマッチングにより、ホモグラフィを計算する。このホモグラフィをもとに、シーンカメラの映像上の注視点を検索結果の文書画像上に変換する。変換された注視点からその時ユーザが読んでいる語句を特定することが可能となる。

3.3 注視点解析

注視点解析モジュールでは、アイトラッキングサーバと文書画像検索モジュールの結果からの注視点データを受けとる。ここでは、ユーザがHMDを見ているかどうかを判定し、見ている場合、ユーザが直近に読んだキーワードに関する情報を送る。キーワードに関する情報は、XMLファイル形式のメタデータとして記述されている。

3.3.1 文書上の注視点

文書画像検索の結果を用いて、ユーザが文書のどこを見ているかを特定することができる。ここでは、特定された文書のどの語句を見ているかを判定するために、それぞれの語句の位置を含むメタデータを参照する。しかし、現在の装着型アイトラッカの精度では、キャリブレーションを行った状態であっても、文書上の単語全てを正確に特定するには限界がある。したがって、現在の精度で、読書の邪魔にならないサービス提供を考えると、キーワードの分布をある程度まばらにしなければならない。そこで、本システムでは、人名や地名といった固有名詞や専門用語などの語句をキーワードとして特定する。キーワードは、それぞれの文書に頻出するわけではないが、その文書を読むにあたっては重要であると考えられる。したがって、本システムではそれらのキーワードを語彙集として保持しておき、キーワードの中でユーザが直近に見たものを起点に情報を表示するようにしている。

3.3.2 HMD上の注視点

本システムでは、brother社のシースルー型のHMD(ヘッドマウントディスプレイ)AirScouterを使用する。これは、図1に示すように左右どちらかの眼の側に付けることが可能である。そして、本システムでは、必要な時にユーザがHMDを見れば、情報が表示されるようにしている。このHMDは、シースルー型であるため、HMDに焦点を合わせない限り、ユーザの読書の妨げになることは無いと考えられる。本システムでは、ユーザが文書を見ているのか、ディスプレイを見ているのかを、注視点がHMD上にあるか否かで判定している。したがって、HMDのキャリブレーションをする必要がある。キャリブレーションをすることで、シーンカメラの撮影画像中のどこにHMDがあるのかを特定することができ、その領域内に注視点があればHMDを見ていると判定する。HMDのキャリブレーションでは、HMDの四隅にそれぞれ順に点を表示する。図4に示すように、ユーザは、それぞれの点において、表示された点と対応するキャリブレーションウィンドウのシーン映像上の点を指定する。そして、対応点の関係からシーン映像とHMDとのホモグラフィを算出し、シーン映像上の注視点をHMD上

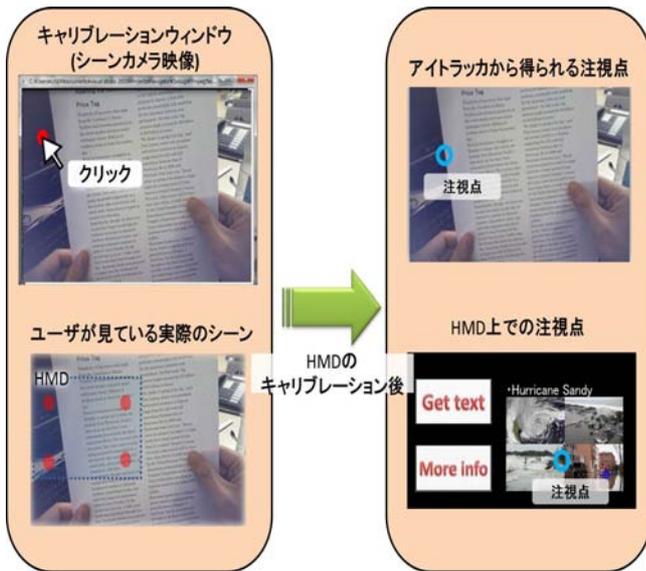


図 4 HMD のキャリブレーション

に変換する。これにより、HMD を見ている時の HMD 上での注視点の位置を求めることができる。

3.4 HMD での情報表示

ユーザが HMD を見ているとき、注視点解析モジュールは、XML ファイルに記述されたメタデータをディスプレイ視覚化モジュールに送る。視覚化モジュールは、ユーザが直近に見たキーワードを起点とするメタデータを受け取ると、HMD にその情報を表示する。そして、ユーザは読んだ語句に関する情報を得ることができる。HMD に情報が表示されている一例を図 5 に示す。ユーザの左右の見やすい方に HMD をつけることが可能であるため、HMD に表示する像も HMD をつける位置に応じて反転させることができる。

さらに、ユーザは HMD に表示されているボタンを注視することで、そのボタンを選択することができる。ここで、HMD のディスプレイ領域は狭いため、一度に表示できる情報量には限りがある。したがって、ボタンを利用し、情報を切り替えることで、ユーザにある程度の情報を提示することが可能となる。また、Midas touch problem と呼ばれる、ただ眺めているだけなのか、ボタンを選択しようとして見ているのかの区別をすることが困難である。この問題を回避するため、本システムでは、[11] で用いられている停留時間を用いて、ボタンを約 2 秒間注視することで選択したとして判定している。そして、ボタンが選択されると、視覚化モジュールは、選択されたボタンに応じて、表示する情報を切り替え、ユーザはさらに情報を得ることができる。

4. 実 験

システム全体の基礎的な性能評価をするための実験を行った。

4.1 文書画像検索の性能評価

まず、アイトラッカを用いた場合の文書画像検索の性能を評価するため、距離及び角度を変化させて実験を行った。この実験では、図 6 に示すように、1 カラムで書かれた A4 用紙に印



図 5 HMD の情報表示の一例

刷した文書との距離 d と角度 α を変化させたときの検索精度を計測した。加えて、それぞれの距離におけるガウシアンマスクサイズの値が 3 及び 7 の時の精度測定した。文書画像検索における登録文書画像数は 1000 ページのデータベースを用いた。使用したアイトラッカのシーンカメラの解像度は 1280×960 であった。精度は 30 秒間検索したときの正答した割合である。

シーンカメラのフレームレートは 25fps である。また、1 回の検索に要する時間は、シーンカメラの 1 フレームのキャプチャ速度より早い 40msec 以下である。図 7 と図 8 にアイトラッカのシーンカメラを用いて文書画像検索をした場合の精度を示す。この結果から、15cm から 30cm の距離で角度を垂直にして読む場合に精度が高いことがわかった。さらに、ガウシアンマスクサイズの値が 3 のときの方が、距離が 40cm でも高い精度で検索が可能であることが分かった。また、本システムでは、垂直から $-30^\circ \sim 30^\circ$ の間であればより高い精度でシステムが動作すると考えられる。これらの結果から、ガウシアンマスクサイズの値を 3 に設定し、文書を読む際に、ユーザが大きく頭部を動かさなければ、ある程度自由に文書を読んでも精度を保てると考えられる。

さらに、10 人の被験者にアイトラッカを装着し、1 ページの印刷文書を通常文書を読む場合と同様に持ち、文書を読んでもらった。このときの文書画像検索精度を表 1 に示す。ほぼすべての被験者において、約 100% で検索することが可能であることがわかった。また、被験者には、Samsung SyncMaster の 24 インチのディスプレイに表示した PDF ファイルを印刷文書と同じ大きさになるように表示した文書を読んでもらった。表 1 に示した結果から、ディスプレイに表示されたデジタル文書も扱えることがわかった。これらの結果から、本システムで利用する文書画像検索手法では、アイトラッカのシーンカメラを使用して、手持ちの印刷文書と同様に、ディスプレイ表示した文書においても適用できると考えられる。

4.2 キーワード推定精度

次に、ユーザが読んだキーワードの検索精度について調査した。この実験では、13 人の被験者に A4 用紙に印刷された文書 1 ページを読んでもらった。使用した文書は、2012 年 10 月 30

表 1 ユーザごとと文書ごとの文書画像検索精度

被験者	A	B	C	D	E	F	G	H	I	J
距離 [cm]	35.0	40.0	40.0	30.0	30.0	35.0	30.0	35.0	40.0	35.0
角度 [°]	60	60	60	65	65	50	80	50	65	60
精度 [%](印刷文書)	100.0	100.0	100.0	100.0	100.0	97.31	100.0	100.0	99.80	100.0
精度 [%](ディスプレイ表示)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.87	100.0	100.0

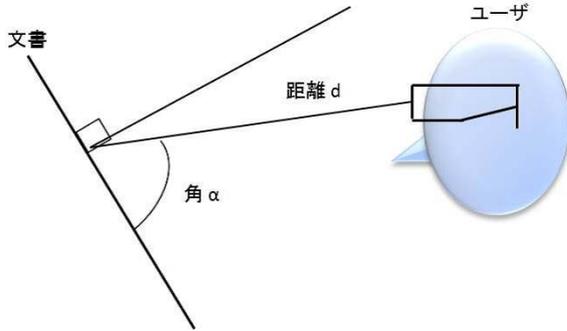


図 6 文書に対する距離 d と角 α

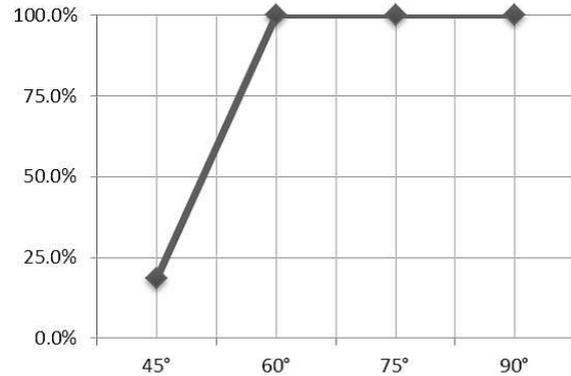


図 8 文書との角度と検索精度

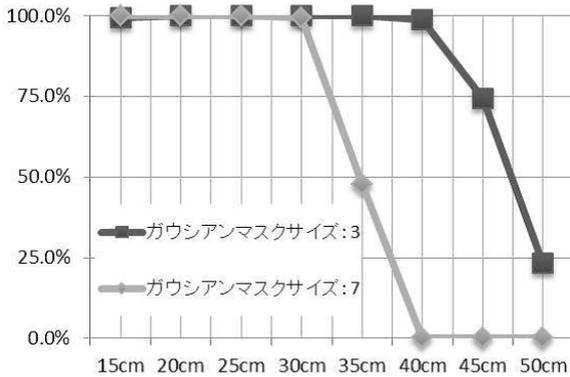


図 7 文書との距離と検索精度

日のニューヨークタイムズ紙の電子版で、見出しが“Awaiting the Storms Price Tag”の記事から引用した。そして、366語の中から7つのキーワード(Thomson Reuters, F.A.O Schwarz, Saks Fifth Avenue, Macy’s, Bloomingdale’s, Sears Holdings, and Kmart)を選び、被験者に文書を音読してもらった。そして、被験者がキーワードを読み上げたときに、そのキーワードが検出されているかを調べた。RecallとPrecisionの割合をヒストグラムで表した結果を図9に示す。RecallとPrecisionの値をそれぞれ R 、 P とすると、ユーザがキーワードを読みあげたときにそのキーワードが検出された場合を T_p 、異なるキーワードが検出された場合を F_p 、検索されなかった場合を T_n として、次式のようにして求めた。

$$R = \frac{T_p}{T_p + T_n}$$

$$P = \frac{T_p}{T_p + F_p}$$

0~20%に入っている被験者は、アイトラッカのキャリブレーションがうまくいっていなかったことが原因であると考えられる。しかし、多くのユーザの場合、本システムはうまく機能していると考えられる。多くの場合、Precisionの値よりもRecallの値の方が高かった。これは、2つのキーワードが近接しており、実際に声に出して読んでいるキーワードと異なるものが検出される場合があることが原因であると考えられる。また、この結果から、被験者によって本システムの性能に差が出ることがわかった。4.1節の実験も含め、現在のアイトラッカのキャリブレーションの精度次第で高い推定精度が得られると考えられる。つまり、キャリブレーションが正確にでき、文書画像検索も正しい検索結果が得られれば、キーワードも正確に検出できると考えられる。

4.3 注視点を利用したインタラクション評価

HMD上での注視点を利用したインタラクションに関して、ユーザスタディを行った。この実験では、ユーザの注視点が生書からHMDに移ったことを検出できるか、また、ユーザがHMD上のボタンを選択した時に、システムが選択されたボタンに応じた動作をするかを調査した。4.2節の実験と同様、被験者は、A4用紙に印刷された文書を読んだ。文書中にはHMDを見るよう指示された文が挿入されており、そこに来たとき、その指示に従い、HMDを見てもらった。この時に、HMD上に情報が表示されたかどうかを確認した。そして、被験者にボタンを選択してもらい、システムが選択に応じた動作をしてい

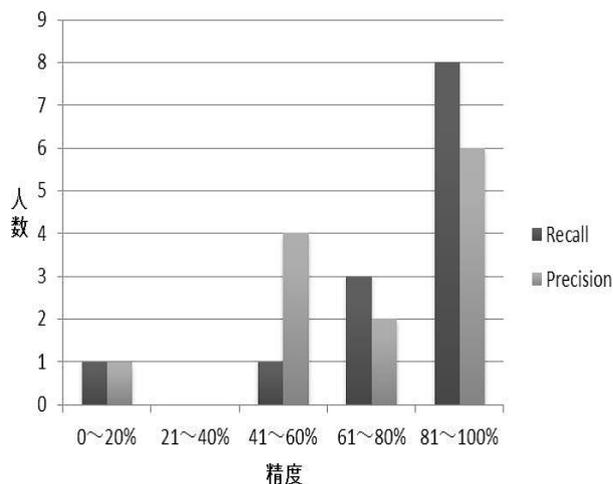


図9 単語推定精度

るか否かを確認した。

被験者がHMDを見た時のディスプレイ表示のRecallの値は、100%、Precisionの値は44%であった。これは、ユーザがディスプレイを見た時にはディスプレイに情報が表示されるが、一方で、ユーザがディスプレイを見ていなくてもHMDに情報が表示されることがあることを表している。そして、注視点を利用したボタン選択の精度は平均81%であった。この実験の場合も、被験者によって精度に大きな差が出た。思い通りにボタンを選択できる被験者もいれば、キャリブレーションが正確ではなかったため、正確な注視点が得られず、全く選択できなかった被験者もいた。このディスプレイを用いたインタラクションでは、アイトラッカのキャリブレーションの問題だけでなく、HMDのキャリブレーション精度も合わせて、両方の要素が精度に関与している。

4.4 考 察

4.1~4.3の実験の結果から、本システムの精度としては、平均して十分に高い精度が得られたといえる。また、実験後、被験者にシステム全体に関していくつかの質問をした。総じて、肯定的な回答が得られた。例えば、「読書中に追加の情報を欲しいと思いますか。」という問いに対しては、77%以上の同意を得られた。この結果は、ユーザビリティの観点から見た場合、本システムのように文書を利用したインタラクティブなシステムの潜在的な利便性を表していると考えられる。しかし、数人の被験者はハードウェアに関して2つのデバイスを装着し、さらに、キャリブレーションをすることに対して、ストレスを感じるという意見もあった。これは、特に、キャリブレーションがうまくいかず何度もやり直した被験者が該当した。したがって、キャリブレーションに関しては、ユーザビリティを考慮したより便利なシステムに改良することが必要であるといえる。

5. ま と め

本稿では、装着型のアイトラッカとシースルー型のヘッドマウントディスプレイを組み合わせたデバイスを使用し、文書画

像検索を用いて、ユーザの読書をアシストするシステムについて述べた。さらに、このシステムの実現可能性を測るために、アイトラッカのシーンカメラを用いた文書画像検索の性能を評価し、文書を読んでいるときの単語推定法を提案した。本システムの基礎的な性能の評価実験により、本システムにおける今後の実現可能性と利便性について考察した。

今後の課題として、キャリブレーションの性能を向上させると共に、キーワードだけでなく任意の語句に関してユーザが読んでいるものを検出し、情報を表示できるようなシステムにすることが挙げられる。加えて、ユーザビリティの観点から、HMDにおけるユーザインタフェースをデザインする必要があると考えられる。

謝辞 本研究の一部は、JST CREST および日本学術振興会科学研究費補助金基盤研究 (B) (22300062) の補助による。

文 献

- [1] B. Erol, Emilio Antunez, J.J. Hull, "Hotpaper: multimedia interaction with paper using mobile phones," Proceedings of the 16th ACM international conference on Multimedia, pp.399-408, Jan. 2008.
- [2] X. Lui and D. Doermann, "Mobile retriever: Access to digital documents from their physical source," Int. J. Doc. Anal. Recognit., vol.11, pp.19-27, Sep. 2008.
- [3] T. Nakai, K. Kise, and M. Iwamura, "Use of affine invariants in locally likely arrangement hashing for camera-based document image retrieval," in Lecture Notes in Computer Science (7th International Workshop, vol.3872, pp.541-552, Feb. 2006.
- [4] R. Biedert, G. Buscher, S. Schwarz, Jörn Hees, A. Dengel, "Text 2.0," in Proc. of the 28th of the International Conference on Human Factors, pp.4003-4008, 2010.
- [5] Florian. Alt, A.S. Shirazi, A. Schmidt, Julian Mennen, "Increasing the user's attention on the web: using implicit interaction based on gaze behavior to tailor content," in Proceedings of the 7th Nordic Conference, pp.544-553, 2012.
- [6] K. Rayner, "Eye movements in reading and information processing: 20 years of research," Psychological Bulletin, pp.372-422, 1998.
- [7] G. Buscher, A. Dengel, R. Biedert, and L.V. Elst, "Attentive documents: eye tracking as implicit feedback for information retrieval and beyond," ACM Trans. Interact. Intell. Syst., vol.1, no.2, pp.9:1-9:30, Jan. 2012.
- [8] A. Bulling, J.A. Ward, and H. Gellersen, "Multimodal recognition of reading activity in transit using body-worn sensors," ACM Trans. Appl. Percept., vol.9, no.1, pp.2:1-2:21, Mar. 2012.
- [9] J.-Y. Lee, H.-M.P. Seok-Han Lee, S.-K. Lee, J.-S. Choi, and J.-S. Kwon, "Design and implementation of a wearable ar annotation system using gaze interaction," in Consumer Electronics (ICCE), 2010 Digest of Technical Papers International Conference on, pp.185-186, Jan. 2010.
- [10] K. Takeda, K. Kise, and M. Iwamura, "Real-time document image retrieval for a 10 million pages database with a memory efficient and stability improved "llah"," 2011 International Conference on Document Analysis and Recognition, pp.1054-1058, Sep. 2011.
- [11] T. Toyama, T. Kieninger, F. Shafait, and A. Dengel, "Gaze guided object recognition using a head-mounted eye tracker," in Proc. of the Symposium on Eye Tracking Research and Applications, pp.91-98, 2012.