

文書画像検索を用いたリーディングライフログの提案

川市 仁史[†] Kai Kunze[†] 黄瀬 浩一[†]

[†] 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

E-mail: †{kawaichi,kunze}@m.cs.osakafu-u.ac.jp, ††kise@cs.osakafu-u.ac.jp

あらまし リーディングライフログとは、文書に対する「読む」という行動をデジタルデータとして記録することであり、これを解析することでユーザの趣味、嗜好に合わせた情報の提供が可能になると考えられる。本研究ではアイトラッカと文書画像検索を組み合わせることで読んでいる文書と読んでいる位置を特定し、質の高いリーディングライフログの実現を目指す。本稿では、その第一歩として、視点情報における改行の判定法を提案する。また、改行判定を用いて読んだ単語数を推定する手法を提案する。実験の結果、改行判定の誤差は2.1行程度であった。また、単語数の推定誤差は、単純な推定法でも13.5%、SVRを用いた推定法で8.2%となり、提案手法の有効性が示された。

キーワード アイトラッカ、文書画像検索、リーディングライフログ、LLAH

1. はじめに

近年、ライフログ [1] が広く社会に認識され、多くの研究が行われている。ライフログとは、人の体験や行動を、音声や映像などのデジタルデータとして記録する技術、またその記録自体のことを指す。日頃の生活を記録することで、過去の体験を振り返ることができようになる。また、自分の行動を客観的に見ることで、自分の考えや行動パターンに新たな気づきを得ることができる。様々なライフログが提案、研究されており、ライフログとして人の行動の何を記録したら、どのように利用できるかが注目されている。

人の行動の中でも、眼に関する行動について考えると、人は日常的に多くの情報を眼から得て生活していることがわかる。特に、本や新聞、テレビや看板などの文字を読むことで情報を得ており、見たり読んだりする行動はとても重要な行動であると言える。このことから、人が読んだ情報を記録することができれば、その人物の趣味や嗜好を知ることができ、より高度なライフログが実現できると考えられる。

ところで、人がものを読むためにはそのものを見なければならぬ。人が見ている箇所を推定するデバイスとして、アイトラッカがある。アイトラッカを用いることで、ユーザの見ているものを特定することができ、また、どのようにものを見ているのかを解析することができる。

アイトラッカを用いて人の読んだ情報を記録するライフログとして、木村らのリーディングライフログ [2] がある。木村らの手法では、文字認識技術とユーザの視点を追跡するアイトラッカを組み合わせることで、看板など情景中の文字を記録する。この手法は、認識技術によるリーディングライフログの実現を目指している。

本研究では、アイトラッカと文書画像検索を組み合わせ、検索技術によるリーディングライフログを実現する。提案システ

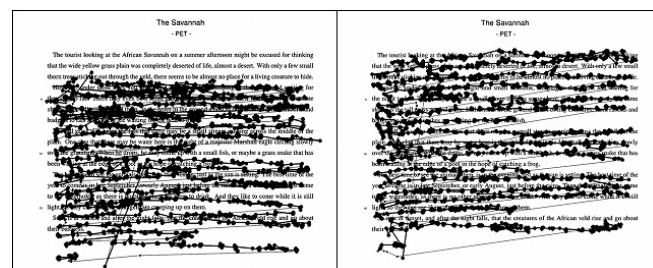


図1 被験者毎の視点情報のズレ

ムでは、ユーザの読んだすべての文書や単語を記録することを目的とする。文書画像検索とは、文書をカメラで撮影することにより、データベースからその文書に対応する文書画像を検索する処理である。文書画像検索を用いることで、ユーザの読んでいる文書を特定し、アイトラッカを用いることでその文書のどこを見ているのかを特定する。これにより、文書中の読んだ箇所や単語を記録することができ、ユーザがどのように文書を読んだのか解析することができる。また、ユーザの読んだ文書や単語を記録することにより、ユーザの趣味や文書に対する理解度などを特定できると考えられる。

文書画像検索を用いたリーディングライフログの実現において、読んだ文書はすべてデータベースに登録されている必要がある。しかし、これは現代の情報社会においては非現実的な条件ではない。一方、ユーザが読んだ単語を全て記録するためには、読んでいる箇所を正確に特定しなくてはならない。これはアイトラッカの精度に依存するものであり、アイトラッカが正確に人の視点を特定することができれば、容易に読んだ単語を記録することができる。

しかし、図1に示すように、アイトラッカにより記録される視点情報には、ユーザとアイトラッカの相性により文書上で数行程度のズレが生じることがある。そのため、既存のアイ

トラックの精度ではユーザの読んだ単語を正確に特定することは困難である。しかし、人の眼の動きと文書中の単語の分布には強い関連性があることは明らかである。そこで、イトラッカの精度によって生じるズレを補正し、単語の分布と視点情報との照合することを考える。視点情報と単語の分布とを照合する際、最も単純な照合方法として、一行分の単語の分布とその一行に対応する視点情報との照合が考えられる。これを実現するには、記録された視点情報に対して改行を特定することが必要である。

そこで本稿では、文書画像検索を用いたリーディングライフログの実現の第一歩として、提案手法により記録された視点情報と文書中の単語分布を照合するための改行の判定方法を提案する。加えて、改行判定結果の利用方法として、ユーザの読んだ単語の数を数える手法を提案する。これは、いわば読みに対する万歩計のようなもの（我々は万語計と呼ぶ）である。改行の判定結果から、ユーザの読んだ単語の数を推定することができれば、ユーザの読んだすべてのテキストを記録するリーディングライフログの機能の一部を実現できたと言える。

2. 関連研究

ここでは、イトラッカを用いた読書に関連する研究と、リーディングライフログの関連研究について述べる。

イトラッカと電子文書を用いた研究として、text2.0 [3] がある。text2.0 では、ディスプレイ一体型のイトラッカを用いて、電子文書に対するユーザの視点を追跡し、特定の単語に至ったとき、情報を出力するシステムである。例えば、読んでいる行に対応する挿絵を表示したり、効果音を出力したりできる。また、視点情報からユーザのわからなかった単語を推定し翻訳することや、どこまで読んだかわからなくなった時に、読んだ箇所を知ることができる。この手法では、電子書籍に対するサービスを実現するため、ディスプレイ一体型のイトラッカを用いている。そのため、ユーザはディスプレイを見ていることが前提であり、電子書籍以外の文書を読む場合には、文書に対するユーザの視点を記録することができない。

リーディングライフログに関する研究として、木村らの手法がある。木村らの手法では、ユーザ装着型のイトラッカと文字認識手法のOCRを用いている。イトラッカと文字認識を組み合わせることで、ユーザの見た看板や広告などの情景内の文字を認識し記録することができる。情景内の見た文字を記録することで、後で単語検索により、検索した単語の写った情景画像を出力することができ、どこでその文字を見たのかを知ることができる。この手法では、認識技術である文字認識を用いたリーディングライフログの実現を目指している。一方で、本研究ではリーディングライフログの実現に対して別のアプローチをし、検索技術である文書画像検索を用いたリーディングライフログの実現を目指す。

3. 文書と視点の関係

人の眼は、Fixation（注視）とSaccade（衝動性運動）を繰り返してものを見ている。Fixationとは、視点が狭い範囲に一

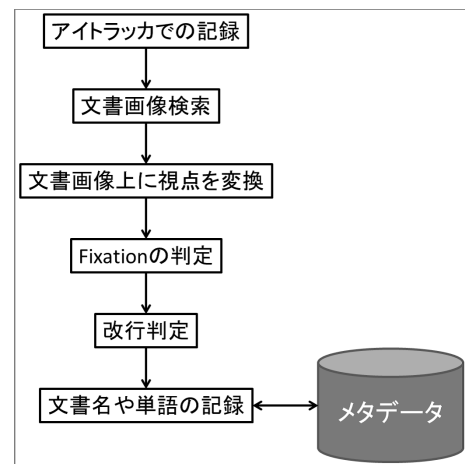


図2 システムの概要

定時間停留することである。また、SaccadeとはFixationとFixationの間に発生する素早い眼の運動である。人の眼は、物を見たり文書を読んだりする際、数百ミリ秒ほど注視をした後、視点を動かし他の場所を再び注視する運動を繰り返していることが知られている [4]。文書を読む際、これらの眼の動きは単語の分布や文書の難易度などに影響を受け、変化することが考えられる [5]。したがって、文書に対するユーザのFixationの分布などを調べることで、ユーザが読んだ単語や難しいと感じた単語の特定が可能であると考えられる。

4. システム

4.1 システムの概要

図2にシステムの概要を示す。本論文で提案するリーディングライフログは、ユーザの読んだすべての文書や単語を記録することを目的とする。この目的を満たすためには、どの文書を読んだか、またその文書のどこを読んだかを特定することが必要である。そこで、それぞれを特定するために、文書画像検索とイトラッカを用いる。文書画像検索とは、文書をカメラで撮影することで、その文書に対応する文書画像を検索する処理である。提案システムでは、文書画像検索の一手法であるLLAH(Locally Likely Arrangement Hashing) [6]を用いる。LLAHは、射影歪みや紙面の非線形的な湾曲に対してロバストであることが知られている。一方、イトラッカは眼の動きからユーザの視線を追跡する装置である。提案システムでは、メガネ型のイトラッカを用いることで紙媒体の文書に対する、ユーザの視点の動きを追跡し記録する。これらの技術を組み合わせることにより、LLAHによって特定された文書画像上に、イトラッカが追跡したユーザの視点を記録することで、ユーザがどの文書のどこを読んだかを特定する。以下では、それぞれの処理について詳しく説明する。

4.2 LLAHによる文書画像検索

図3にLLAHによる文書画像検索の概要を示す。文書画像検索は、検索する文書を撮影し、対応する文書画像をデータベースから出力する。登録文書画像は、特徴点抽出によって点の集合に変換され、登録処理に入力される。登録処理により、

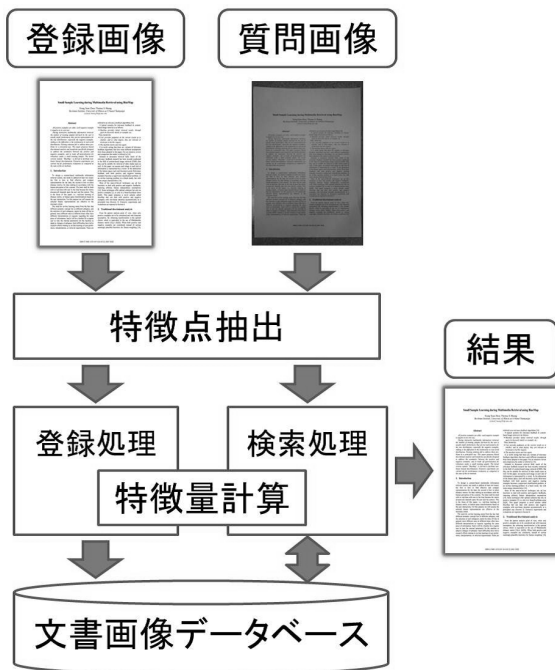


図 3 LLAH の処理の概要

各特徴点はそれぞれの特徴量に基づいてデータベースの登録される。検索処理では質問文書画像から特徴点を抽出した後、登録処理と同じ方法で特徴点から特徴量を計算される。そして、得られた特徴量を用いてデータベースにアクセスし、対応する文書画像を出力する。以下に、各処理について示す。

4.2.1 特徴点抽出と特徴量計算

LLAH では、特徴点の配置に基づいて文書画像を特定している。そのため、射影歪みやノイズが生じた場合でも同一の特徴点を抽出できるように単語領域の重心を特徴点としている。LLAH では、ある特徴点 s の特徴量に、点 s の近傍 4 点から以下の式で得られる値を用いている。

$$\frac{P(A, C, D)}{P(A, B, C)} \quad (1)$$

ここで、 $P(A, B, C)$ は頂点 A, B, C からなる三角形の面積である。また、2 つの三角形の面積比である式 (1) の値は、アフィン不変量であり、局所領域では射影変換にほぼ不変となる。特徴点 s の特徴量を計算する際、点 s を含む近傍 m 点から 4 点を選ぶ組み合わせを考える。そして、 ${}_m C_4$ 通りあるすべての組み合わせからアフィン不変量を求め、それらを一列に並べたものが一つの特徴量となっている。

4.2.2 登録・検索

LLAH では、文書画像の登録・検索にハッシュ表を用いる。登録処理では、各特徴点から得られる特徴量を基に、各点の文書 ID、点 ID、特徴量がハッシュ表に登録される。検索処理では、検索質問画像の各特徴点から得られた特徴量を基に、ハッシュ表からリストが得られる。そして、得られたリストの各項に対して特徴量が一致するかを調べ、一致していたら文書 ID に対応する文書に投票する。最後に、最大得票数を得た文書画像が検索結果として出力される。



図 4 アイトラッカ

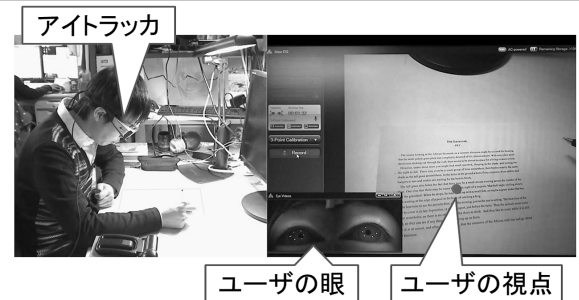


図 5 アイトラッカからの情報

4.3 アイトラッカ

アイトラッカはセンサによって眼球運動を捉え、ユーザーの視点を視野カメラ映像上に表示する装置である。本研究では、図 4 に示す SMI 社のアイトラッカ iViewX ETG (Eye Tracking Glasses) [7] を用いてユーザーの視点情報を記録する。図 5 にアイトラッカにより出力される視点情報の様子を示す。iViewX ETG は、メガネ型のアイトラッカであり、ユーザーの前方を映す視野カメラとユーザーの両眼を映す眼球カメラが内蔵されている。眼球運動を検出する際、iViewX ETG では眼の周りに円を描くように 6 点の赤外光を照射し、角膜で反射させる。眼球運動には、眼球の大きさなどによって個人差があるため、あらかじめキャリブレーションする必要がある。アイトラッカが視野カメラ映像上に表示する 3 点をユーザーに順番に見てもらい、それぞれの点での眼球運動を記録することで、眼球運動の個人差が補正される。検出された眼球運動は、視野カメラ映像上の視点情報に変換され、ユーザーがどこを見ているのかを特定することができる。

4.4 視点情報の変換

文書画像検索とアイトラッカを組み合わせることにより、ユーザーの読んだ文書の文書画像と視点情報が得られる。ユーザーが文書のどこを読んだのかを特定するためには、文書画像上での視点位置の特定が必要である。しかし、アイトラッカから得られる視点情報は、視野カメラ映像上の座標系で表されている。従って、視点情報の座標を文書画像上の座標系に変換する必要がある。そこで、LLAH で求める射影変換パラメータを利用し、視点情報を文書画像上に変換する。

図 6 に文書画像上に変換された視点情報を示す。文書画像検索を行う際、質問画像中に検索したい文書以外の文書などが

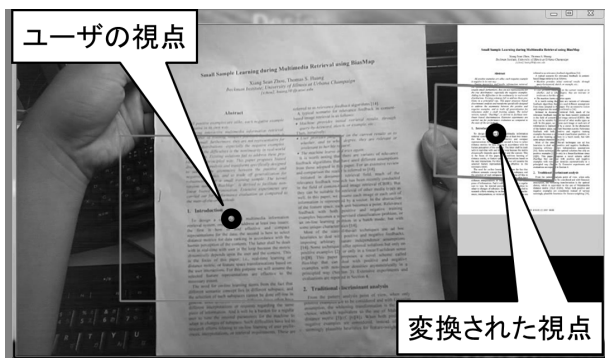


図 6 変換された視点

写っていると、検索精度が下がってしまうことがある。そのため、視野カメラ映像上に表示される視点情報を中心とした画像領域を切り出す。そして、切り出された画像領域を用いて、LLAH による文書画像検索を行う。LLAH では、検索過程で検索質問画像とそれに対応する登録文書画像の特徴点の対応関係が得られる。得られた対応関係から射影変換パラメータを推定することができ、検索質問画像の撮影範囲を特定することができる。また、特定された撮影範囲は、登録文書画像上に四角形で表示される。このように LLAH により得られた射影変換パラメータを用いて、視点情報を登録文書画像上に変換される。これにより、視点情報は文書画像上の座標に変換され、文書上のどこを読んでいるかを記録することができる。

4.5 Fixation の判定

文書画像上に変換された視点情報から、ユーザがどこを注視していたのかを特定するため、Fixation を判定する。記録された視点は、一定の狭い範囲に複数個存在することがある。これは、ユーザがそこにある単語を注視しているからであり、読んでいる可能性があると言える。そこで、記録した視点情報を解析し、Fixation を特定する。

判定する視点と、その視点以前の $f - 1$ 個数の視点を用いて Fixation を判定する。 f 個の視点を包含する矩形の面積が閾値以下であれば、その視点は注視していると判定する。この処理をすべての視点に対して行うことにより、文書中の注視点を特定する。

4.6 改行判定

図 7 に改行するときの視点の動きを示す。文書画像上に変換された視点から改行の視点の動きを判定し、一行分の視点情報を求める。改行するときの視点の動きは、読んでいた一行の行末から次の一行の行頭へ向かう動きであり、読んでいた時の視点の動きとは明らかに異なる。改行の眼の動きは、素早く大きな動きとなり、読んでいた時の視点の移動方向とは逆向きに移動する。このことを利用して、閾値処理により改行を判定する。

Fixation の判定法と同様に、判定する視点と、その視点以前の $l - 1$ 個数の視点から改行を判定する。 l 個の視点を包含する矩形の面積が閾値よりも十分に大きく、 $l - k$ 個目 ($k < l$) から l 個目までの視点が連続してそれまでの視点の向きと逆向きになっている時、 l 個目の視点を改行中の視点であると判定する。また、改行と判定された視点以降の一定数の視点も同じく改行

two children, Jimmy and Madeleine. We spend a lot of time just singing and dancing around the house. I grew up with music because my dad is the pianist, Chester Hott - who still plays by the way. My working day is divided between television and writing cook books, though TV takes most of my time. I spend about five days a fortnight working on **改行** as I appear in. I eat all sorts

図 7 改行

表 1 特徴量

視点情報	読書にかかった時間
	文書に対する Fixation の回数
	視点の総移動距離
	Saccade の総移動距離
	Saccade の平均移動距離

中の視点であると判定する。

4.7 万語計

改行判定の結果を用いて、ユーザの読んだ単語数を推定する万語計について述べる。改行判定により改行を正確に推定できれば、ユーザの読んだ単語の数を推定することが可能になり、リーディングライフログの一部の機能を実現することができる。

万語計の実現方法には様々なものが考えられる。ユーザが読んだすべての単語を記録することができれば正確な万語計を実現できる。しかし、一日や一週間など、一定期間にどれだけの単語を読んだのかを記録するのであれば多少の誤差は許容される。そこで本研究では、4.6 節で述べた改行判定を用いることで、万語計を実現する。

具体的には、二つの手法を提案する。一つは、次のような単純な単語推定法である。これは、ユーザの読んだページの一行あたりの平均単語数を N 、改行判定法から得られた改行回数を L としたとき、 $N(L + 1)$ として、ユーザが読んだ単語数を大まかに推定するものである。ここで、ページごとの N の値は文書画像検索によって得ることができる。

ところで、改行判定に誤差が含まれると、推定単語数にも誤差が蓄積され、その結果、推定精度が低くなる。そこで、さらに視点情報から得られた特徴ベクトル \mathbf{x} を加味して、 $f(N, L, \mathbf{x})$ により単語数を推定する手法も考える。関数 f は学習データを用いて学習することで定める。具体的には、SVR (support vector regression) を用いて関数を学習する。この手法を、SVR を用いた単語推定法と呼ぶ。表 1 に視点情報から得られる特徴量を示す。特徴ベクトル \mathbf{x} は 5 次元ベクトルであり、文書に対する Fixation の回数、記録された視点の総移動距離、文書を読むのににかかった時間、Saccade と判定された視点の総移動距離、Saccade の平均移動距離の 5 つの特徴量から構成される。

5. 実験

提案システムの有効性を示すため、改行判定の精度を評価した。また、万語計の推定精度を評価した。

5.1 実験条件

被験者に複数枚の英語文書を読んでもらい、その際の視点情報を記録した。被験者は 10 人で、裸眼が 7 人、両目コンタクトが 2 人、片目コンタクトが 1 人であった。使用した文書は、ケンブ

表 2 改行判定の誤差

文書	行数	被験者									文書毎の 平均誤差
		a	b	c	d	e	f	g	h	i	
A	15	-1	2	0	1	3	-1	0	10	8	2.4
B	15	-1	2	0	2	2	2	1	11	2	2.3
C	12	3	10	1	3	0	2	1	13	6	4.3
D	19	-3	7	-2	-2	1	0	0	5	4	1.1
E	18	-1	1	1	1	-1	1	-1	10	4	1.6
F	19	-6	-2	-3	1	-1	-2	-3	10	2	-0.4
G	18	2	8	-1	0	2	1	-1	2	3	1.7
H	17	-1	5	1	2	2	2	0	5	5	2.3
I	16	1	2	0	5	2	4	0	7	1	2.4
J	21	-1	5	2	5	2	0	0	10	7	3.3
平均誤差		-0.8	4	-0.1	1.8	1.2	0.9	-0.3	8.3	4.2	2.1

リッジ大学 ESOL(English for Speakers of Other Languages) 試験のひとつである PET(Preliminary English Test) テストから得た 10 題である。各文書に対して、単語数に合わせた制限時間を設け、その時間内に読んでもらった。

実験の手順について説明する。被験者にアイトラッカを装着してもらいキャリブレーションした後、文書を一枚ずつ読んでもらう。読む際、文書は机に置いたままで自由な姿勢で読んでもらった。また、文書を読み終わる毎にキャリブレーションを繰り返して視点情報を記録した。そして、 k 人のデータに対して、 k -fold cross validation を行って評価した。

5.2 実験結果

10 人の被験者から記録した視点情報に対して改行を判定する。10 人の被験者のうち、1 人は記録中に居眠りをするのが頻繁にあり、視点情報に多くのノイズが含まれていたため、この人物の視点情報を使用しない。なお、この被験者は裸眼である。そのため、精度は 9-fold cross validation によって評価した。

表 2 に改行判定の誤差を示す。1 つの文書あたりの改行の推定誤差は、2.1 行であった。実験に用いた文書の平均行数は 17 行であることから、12.4% の誤差でユーザの読んだ行数が推定できることが分かった。改行判定の結果、被験者 h の改行判定の結果が最も悪くなった。原因として、被験者 h の視点情報は他の被験者のものに比べて瞬きなどのノイズが多かったことが挙げられる。また、被験者 h は他の被験者に比べて、改行直後に前の行の行末を見直す行為を頻繁にしていたことが挙げられる。そのため、改行でない箇所や読み返しを改行であると誤判定したことが考えられる。文書毎の改行判定の結果を見ると、推定改行数が実際の改行数より下回った文書が複数あった。これらの文書の多くが、単語数の極端に少ない行を含んでいた。その結果、単語数の極端に少ない行に対しては、提案手法の改行判定が有効でないことが分かった。提案手法では、連続する一定数の視点を包含する矩形の面積の大きさを改行判定の閾値としているため、また、記録された視点情報から、多くの被験者が単語数の極端に少ない行を注視する時間が極端に短かったため、改行を判定できなかったことが分かった。文書毎の誤差

では、文書 C で最も大きな誤差が生じた。文書 C は、文章は短いものであったが、単語数の極端に少ない行を多く含むわけではなかった。被験者 b 、及び被験者 h の文書 C に対する視点情報が多くのノイズや行を跨いだ読み返しを含んでいたことが、推定改行数の誤差を大きくしたと考えられる。

改行判定の結果から、ユーザの読んだ単語数を推定した結果について述べる。表 3 に単語数推定の結果を示す。単純な推定法の結果から求めたユーザが読んだ単語数推定の誤差は、平均で 47.95 個となった。実験に用いた文書の平均単語数は 279.3 個であることから、推定誤差は 17.2% であることが分かった。被験者 h の推定単語数の誤差が最も大きくなったのは、改行判定の誤差によるものである。表 4 に SVR を用いた単語数推定の結果を示す。SVR を用いた単語数推定の誤差は、平均で 40.5 個 (14.5%) となった。被験者毎の推定結果から、SVR の学習がうまくできていない被験者がいることが分かった。このことから、要素数を増やすことや、本実験で用いた 7 個の要素以外の視点情報を用いた学習を検討する必要がある。しかし、平均誤差は単純な推定法よりも良くなっていることから、ユーザの読んだ単語数の推定に SVR を用いることが有効であることが分かった。

表 5 にすべての文書に対する結果を示す。結果から、被験者 c の推定単語数の誤差が単純な推定法でも 0.3%、SVR を用いた推定法では 0.1% となり、ほぼ正解単語数に等しくなった。一方、被験者 h は単純な推定法で 47.0%、SVR を用いた推定法でも 21.8% となり最も悪い結果となったが、SVR を用いることで誤差を 1/2 に抑えることができた。被験者 a は SVR を用いることで誤差が大きくなった。原因としては、被験者 a が他の被験者に比べ Fixation 回数が少なかったことが挙げられる。総合的に見ると、10 文書 2793 単語に対して、推定誤差は、単純な推定法でも 13.5%、SVR を用いた推定法では 8.2% に留まることが分かった。この結果より、改行判定に基づく万語計は一定の有効性を持つことが分かった。

6. まとめ

本論文では、文書画像検索を用いたリーディングライフログを提案した。また、視点情報から改行を判定する方法、ならびにユーザの読んだ単語数を推定するアプリケーションである万語計を提案した。

改行判定の平均誤差は、2.1 行となった。実験に用いた文書の平均行数が 17 行であることから、提案する改行判定法が有効であることが分かった。また、単語数の推定精度としては、2793 単語に対して、単純な単語推定法で 13.5%、SVR を用いる単語推定法で 8.2% となり、手法の有効性が明らかになった。以上より、改行判定と文書画像検索を組み合わせることによって、万語計が実現可能であることがわかった。

今後の課題として、改行判定により特定できた一行分の視点情報と、対応する一行の単語の分布との照合、ならびに、視点情報と単語の分布を照合することにより、アイトラッカによる視点情報のズレを修正することが挙げられる。

表 3 単純な単語数推定の結果

文書	単語数	被験者									文書毎の 平均誤差
		a	b	c	d	e	f	g	h	i	
A	255	238.0	289.0	255.0	334.7	306.0	318.8	255.0	366.6	391.0	32.9
B	203	189.4	230.1	203.0	230.1	230.1	230.1	216.5	351.9	230.1	49.1
C	210	262.5	385.0	227.5	262.5	210.0	245.0	227.5	437.5	315.0	61.8
D	302	219.6	356.9	247.1	233.4	274.5	260.8	260.8	329.5	315.7	40.5
E	296	279.5	312.4	312.4	312.4	279.6	312.4	279.6	460.4	361.8	33.8
F	276	188.8	246.9	246.9	290.5	261.5	246.9	232.4	421.3	305.1	28.9
G	281	312.2	405.9	265.4	281.0	312.2	296.6	265.4	312.2	327.8	23.2
H	298	280.5	385.6	315.5	333.1	333.1	333.1	298.0	385.6	385.6	29.3
I	255	270.9	286.9	255.0	272.0	286.9	238.0	255.0	425.0	270.9	21.3
J	417	397.1	516.3	456.7	516.3	456.7	417.0	417.0	615.6	556.0	70.6
平均誤差		35.4	68.0	19.1	41.0	27.4	28.0	14.8	131.3	66.6	47.95

表 4 SVR を用いた単語数推定の結果

文書	単語数	被験者									文書毎の 平均誤差
		a	b	c	d	e	f	g	h	i	
A	255	196.7	266.0	278.9	288.9	255.6	276.0	241.8	316.2	327.8	54.8
B	203	205.9	227.6	277.0	257.2	252.8	236.4	260.9	310.6	240.4	34.6
C	210	249.0	309.9	277.5	232.4	252.8	253.1	221.1	356.2	294.2	75.9
D	302	153.2	289.9	277.9	281.1	252.8	249.1	275.2	298.8	275.5	45.8
E	296	237.3	283.3	281.3	266.9	252.8	276.0	266.6	370.8	317.9	38.4
F	276	216.5	235.5	279.1	266.4	252.8	242.8	264.4	353.8	278.1	46.8
G	281	231.1	310.5	279.0	250.9	252.8	265.6	256.9	275.4	305.2	34.7
H	298	250.7	318.3	281.3	271.8	252.8	286.6	272.7	336.4	330.7	44.8
I	255	230.1	269.3	278.8	249.5	252.8	249.1	244.4	341.6	272.6	33.3
J	417	249.6	376.8	285.2	391.6	252.8	338.2	373.5	441.8	406.1	76.3
平均誤差		65.6	30.5	38.1	25.7	44.8	31.5	25.4	62.6	33.0	40.53

表 5 推定単語数の合計

	被験者									平均誤差
	a	b	c	d	e	f	g	h	i	
単純な単語推定	2638.8	3415.1	2784.6	3065.9	2950.5	2898.7	2707.2	4105.5	3459.0	376.1 (13.5%)
単純な単語推定の誤差 (5.5%)		622.1 (22.2%)	8.4 (0.3%)	272.9 (9.7%)	157.5 (5.6%)	105.7 (3.8%)	85.8 (3.1%)	1312.5 (47.0%)	666.0 (23.8%)	
SVR を用いた単語推定	2220.4	2887.5	2796.0	2756.8	2530.8	2672.9	2677.4	3401.5	3048.5	229.8 (8.2%)
SVR を用いた単語推定の誤差 (20.1%)		94.5 (3.4%)	3.0 (0.1%)	36.2 (1.3%)	262.2 (9.4%)	120.1 (4.3%)	115.6 (4.1%)	608.5 (21.8%)	255.5 (9.1%)	

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] <http://en.wikipedia.org/wiki/Lifelog>.
- [2] 木村崇志, 柿迫良輔, フォンヤオカイ, 内田誠一, 岩村雅一, 大町真一郎, 黄瀬浩一, “Reading-life log のプロトタイプ実装,” 画像の認識・理解シンポジウム (MIRU2012), p.5, aug 2012.
- [3] R. Biedert, G. Buscher, and A. Dengel, “The eyebook—using eye tracking to enhance the reading experience,” Informatik-Spektrum, vol.33, no.3, pp.272–281, 2010.
- [4] G. Buscher and A. Dengel, “Gaze-based filtering of relevant document segments,” International World Wide Web Conference (WWW), pp.●●–●●, 2009.
- [5] G. Buscher, A. Dengel, L. vanElst, and F. Mittag, “Generating and using gaze-based document annotations,” Proceedings and Extended Abstracts of the Conference on Human Factors in Computing Systems, pp.3045–3050, ACM, 2008.
- [6] 竹田一貴, 黄瀬浩一, 岩村雅一, “1,000 万ページのデータベ

スを対象とした長時間文書画像検索のためのメモリ削減と安定性向上,” 電子情報通信学会技術研究報告, vol.110, no.467, pp.●●–●●, mar 2011.

- [7] <http://www.eyetracking-glasses.com/>.