

# Reading-Life Log の実現に向けた万語計の提案

川市 仁史<sup>†</sup> Kai Kunze<sup>†</sup> 黄瀬 浩一<sup>†</sup>

<sup>†</sup> 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 Reading-Life Log とは、文書に対する「読む」という行動の記録である。これを解析することでユーザの趣向や考えを把握し、適した情報を提供することが可能になると考えられる。最もシンプルな Reading-Life Log の一つとして、ユーザの読んだ単語数を記録することが挙げられる。我々は、ユーザの読んだ単語数を記録するものを万語計と呼ぶ。万語計では、ユーザの眼の動きや読んでいる文書の情報を取得することにより、読んだ単語数を推定する。本稿では、4つの万語計について比較検討する。具体的には、(1) 単純に読書時間から推定する手法、(2) 読んだ文書の一行あたりの平均単語数と読んだ行数から推定する手法、(3) (2) の手法に視点情報を加味した手法、(4) 視点情報のみを用いた手法である。実験の結果、最も精度が高い手法は (3) であり、その誤差が 11.3 % であることがわかった。

キーワード Reading-life Log, 万語計, アイトラッカ, 文書画像検索

## 1. はじめに

近年、小型センサや大容量レコーダなどの技術進歩により、自分の行動をデジタルデータとして自動的に記録するライフログ [1] が多く研究されている。例えば、小型ウェアラブルカメラやモーションセンサを用いて、人の行動を記録するもの [2] [3] や、静電容量センサを用いて、食事や会話する際の喉の動きを記録するもの [4] がある。また、ばらばらに記録されたライフログを一つにまとめ、付加価値の高いサービスを目指す研究がされている [5]。このように、人の様々な行動を記録することで、過去を追体験することや、自分の行動を客観的に解析することが可能となり、自分の行動や考えから新たな発見が生まれる。

人の最も重要な行動の一つに、「読む」という行動がある。人は日常的に、読むことによって多くの情報を取得している。例えば、本を読んだり、新聞を読んだりすることにより、新しい情報や欲しい情報を取得している。また、読むことで取得する情報は、その人の年齢や趣味、学力などにより、人それぞれ異なっている。このことから、人が読んだテキストや、それをどのように読んだかを記録することができれば、その人の趣味や思考、学力を知ることができると考えられる。このように、読んだテキストやどのように読んだかなど、「読む」という行為を記録するライフログを、我々は Reading-Life Log と呼ぶ [6] [7]。

Reading-Life Log では、ユーザが読んだり学んだりした多くの情報を記録する。また、読んだ文書に対するユーザの理解度や興味の度合いを解析することもできる [8] [9]。Reading-Life Log の最終目標は、ユーザが読んだ全てのテキストと共に、それをどのように読んだか、どのように理解したかを記録することである。

このような Reading-Life Log を実現するためには、「読む」

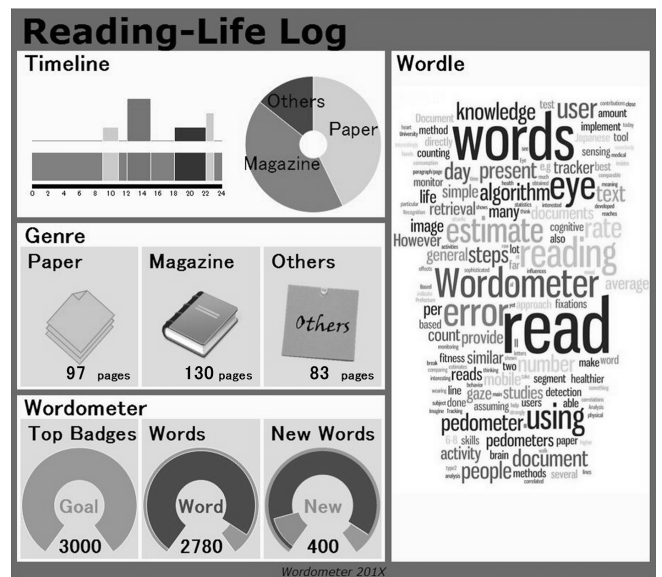


図 1 万語計による Reading-life Log

という行為に関する様々な情報を記録する必要がある。例えば、読んだ時間や冊数、どのように読んだか、読んでどう思ったかなどが挙げられる。その中でも、最もシンプルな情報として、ユーザが読んだ単語数が挙げられる。人が一日に読む単語の数と、その人の学力や一般常識には相関関係があることが知られている [10]。図 1 に示すようなインタフェースを通して、ユーザが一日で読んだ単語数がわかれば、その人の学力や一般常識をはかるための有効な手がかりになる。ユーザが読んだ単語数を推定する手法を我々はすでに提案し、これを万語計と呼ぶ [11]。

本稿では、より高精度でシンプルな万語計を実現するための

様々な方法を考案し、比較検証する。具体的には、以下の4つの万語計を考える。(1) 単純に読書時間から推定する手法、(2) 読んだ文書の一行あたりの平均単語数と読んだ行数から推定する手法、(3) (2)の手法に視点情報を加味した手法、(4) 視点情報のみを用いた手法である。これらの手法では、ユーザーの視点を記録するアイトラッカと、文書画像検索の一手法である LLAH (Locally Likely Arrangement Hashing) [12] を用いる。アイトラッカから得られるユーザーの視点情報と、LLAH から得られる読んだ文書の情報から単語数を推定する。実験の結果、単語数の推定誤差が最も小さかったのは(3)の手法で、その誤差は11.3%であることがわかった。

## 2. 関連研究

ここでは、文書と眼の動きに関する研究と Reading-Life log の研究事例について述べる。

文書と眼の動きに関する研究として、ユーザーの文書に対する理解度を推定する研究 [8] がある。この研究は、アイトラッカを用いて、ユーザーがどのように文書を読んでいるのかを記録し、その文書に対するユーザーの理解度を推定することを目的としている。理解度を推定するため、ユーザーが読んでいるのか、それともほかの行動をしているのかを推定する手法が提案されている。また、アイトラッカを用い、ユーザーの眼の動きから、どのジャンルの文書を読んでいるのかを推定する研究 [13] もある。この研究では、ユーザーの眼の動きから、漫画や雑誌、新聞などの文書のジャンルを推定することを目的としている。このように、ユーザーの眼の動きから情報を得る様々な研究がなされている。

また、眼の動きからユーザーに情報を与える研究も行われている。その一例として、text 2.0 [14] では、ディスプレイ体型のアイトラッカを用いて、電子文書に対するユーザーの視点を追跡する。そして、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。このシステムを用いれば、読んでいる単語に対応する挿絵を表示したり、効果音を出力したりすることが可能である。また、視点情報からユーザー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、読んだ箇所を知りこともできる。このように、眼の動きと文書の関係には多くの関心が向けられ、様々な研究がなされている。

Reading-Life Log の研究事例として、木村らの手法 [7] がある。木村らの手法では、モバイルアイトラッカと文字認識技術の一手法である OCR を用いている。アイトラッカと文字認識技術を組み合わせることで、情景内にあるユーザーが読んだ文字を認識し、記録することができる。情景画像とともに、ユーザーの見た文字を記録することで、見た情景と見た文字を結び付け、どこでどの文字を見たのかを知ることができる。この手法では、認識技術による Reading-Life Log の実現を目指している。一方で、我々はすでに検索技術である文書画像検索を用いた Reading-Life Log を提案している [6]。この手法では、モバイルアイトラッカと文書画像検索の一手法である LLAH を用いる。LLAH を用いてユーザーの読んだ文書を特定し、その文書画像上にユーザーの視点を交換することにより、ユーザーが文書を

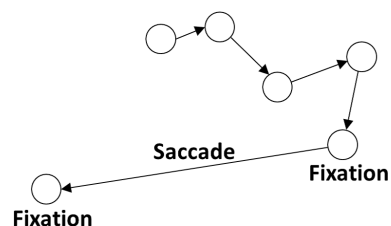


図2 Fixation と Saccade

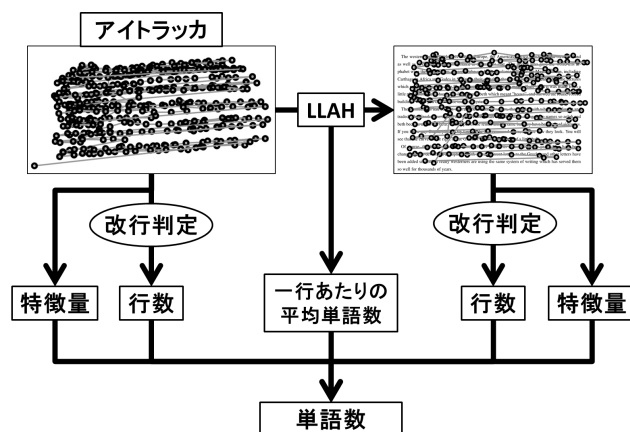


図3 万語計の概要

どのように読んだかの記録が可能になる。本研究では、アイトラッカと文書画像検索を用いた Reading-Life Log の実現を目指す。

## 3. 文書と視点の関係

文書を読むときの視点の動きについて述べる。図2に示すように、人は Fixation (注視) と Saccade (衝動性運動) を繰り返しながら物を見ている。視点が狭い範囲に一定時間留まることを Fixation という。また、Fixation と Fixation の間にする視点の素早い移動を Saccade という。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 Fixation と、Saccade を繰り返していることが知られている [15]。文書を読む時、Fixation や Saccade の動きは、その文書の難易度や単語の分布によって変化する [16]。したがって、ユーザーの視点を記録することで、読んだ文書の難易度や単語の分布、ユーザーの理解度の推定が可能であると考えられる。

## 4. 万語計

万語計は、ユーザーの視点と読んだ文書についての情報から、ユーザーが読んだ単語数を推定するものである。それぞれの情報を記録するため、ユーザーの視点を記録することができるデバイスであるアイトラッカと、文書画像検索の一手法である LLAH を用いる。アイトラッカを用いることで、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。また、LLAH を用いることでどの文書を読んでいるのかを特定することができる。図3に万語計の処理の流れを示す。万語計では、アイトラッカと LLAH



図 4 アイトラッカ

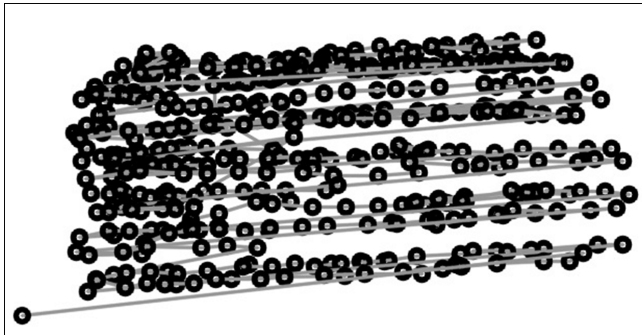


図 5 アイトラッカから得られる視点情報

を用いて、ユーザの視点情報を取得、解析することで、改行回数などの情報を得て、ユーザの読んだ単語数を推定する。以下では、万語計の概要について述べる。

#### 4.1 アイトラッカから得られる視点情報

図 4 に、本研究で用いるアイトラッカを示す。本研究では、SMI 社のモバイルアイトラッカ ETG (Eye Tracking Glasses) <sup>(注1)</sup> を用いてユーザの視点情報を記録する。アイトラッカは眼の動きを撮影するアイカメラと、ユーザの前方を撮影する視野カメラからなる。アイカメラによって、瞳孔の動きが撮影され、ユーザの眼の動きが検出される。そして、検出された眼の動きからユーザの視点特定され、視野カメラ映像上に出力される。

図 5 に、アイトラッカから得られる視点情報を示す。アイトラッカを用いることで、視野カメラ映像の各フレームに対する視点情報が得られる。得られる視点情報は、視野カメラ映像上の座標系で表現されている。また、Fixation や Saccade、Blink (瞬き) などの眼の動きについての情報も得られる。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情報であるため、頭の動きなどによるノイズを多く含む。具体的には、視点を動かすときに、頭も一緒に動いた場合、記録される視点情報には、実際の視点の動きに頭の動きも加味される。そのため、実際の視点の動きとは異なる視点情報が記録される。このように、アイトラッカから得られる視点情報は、視点の動きだけでなく、頭の動きなどによるノイズが含まれる。

#### 4.2 文書画像上に変換された視点情報

図 6 に LLAH によって文書画像上に変換された視点情報を

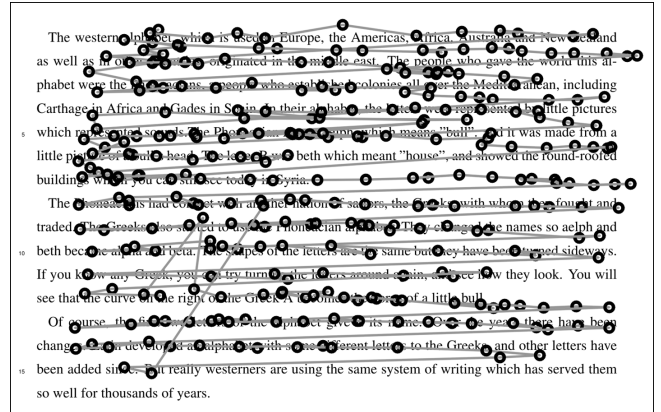


図 6 文書画像上に変換された視点

示す。LLAH を用いることで、視野カメラ映像上のユーザの視点を文書画像上の座標系に変換することができる。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。この対応関係から射影変換パラメータが計算でき、検索質問画像が登録文書画像のどこの範囲を撮影したものが特定される。この射影変換パラメータを利用することで、アイトラッカから得られる視点情報を文書画像上に変換することができる。文書画像上に視点情報を変換することにより、文書をどのように読んだのかを記録することが可能になる。また、文書画像上の座標系に変換することで、視点情報から頭の動きなどによるノイズが除去される。

#### 4.3 Fixation 判定

アイトラッカから得られる視点情報には、Fixation や Saccade などの情報が含まれる。Fixation と Saccade は、視野カメラ映像上の座標系で表される視点から判定されている。つまり、頭の動きなどが加味された視点情報に対する判定結果である。一方で、LLAH によって文書画像上に変換された視点情報は、頭の動きなどによるノイズが除去されている。そのため、視野カメラ映像上の座標系で判定された Fixation と Saccade を、文書画像上に変換された視点情報に対して、そのまま用いることができない。そこで、文書画像上に変換された視点に対して、Fixation を判定する。具体的には、判定する視点と、その視点以前の  $f - 1$  個数の視点を包含する矩形の面積を計算する。そして、矩形の面積が閾値以下であれば、その視点を Fixation とする。閾値以上のものは、すべて Saccade と判定する。この処理をすべての視点に行うことにより、文書に対する Fixation を特定する。

また、連続する視点が Fixation と判定された場合、それらの視点の平均座標を求め、一つの視点としてまとめる。Saccade と判定された視点情報は除去する。このような Fixation の結合と Saccade の除去により、図 7 のような視点情報が得られる。

#### 4.4 改行判定

改行判定法について説明する。改行判定を用いることで、ユーザの眼の動きから、文書の重要な情報の一つである行数を推定することができる。また、文書中のどれだけの単語を読んだのかを推定するための重要な要素となる。図 8 に改行時の視点の

(注1) : <http://www.eyetracking-glasses.com/>

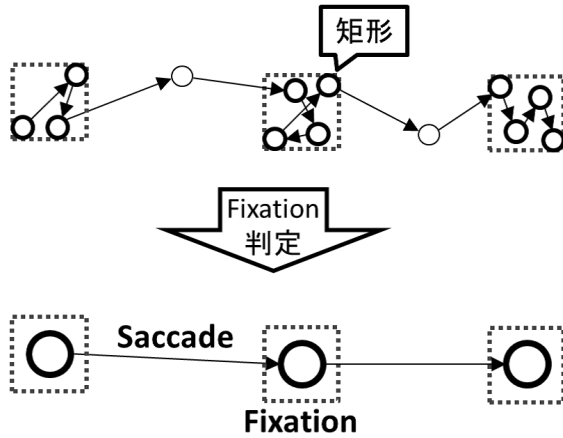


図 7 Fixation 判定

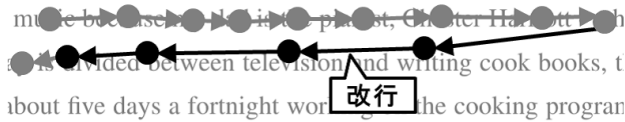
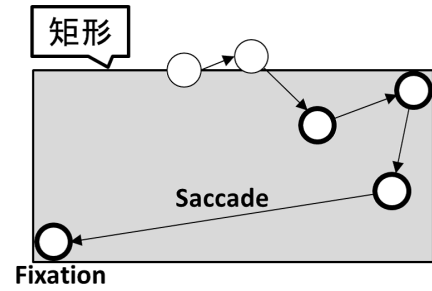
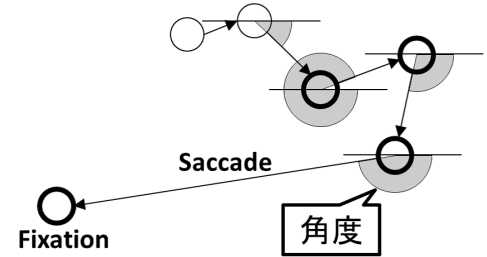


図 8 改行時の視点の動き



(a) 視点を包含する矩形



(b) 視点の角度

図 9 改行判定

動きを示す。改行時の視点の動きは、改行前に文書を読んでいる時の動きとは明らかに異なる。改行する時、視点は読んでいた行の行末から次の行の行頭へ向かう動きをする。その動きは、素早く大きな動きとなり、それまで読んでいた時の視点の移動方向とは逆向きになる。これらの2つの特徴を利用して、視点の移動距離と角度から改行を判定する。

改行判定には、判定する視点とその視点以前の  $l-1$  個数の視点を用いる。図 9 に  $l$  個の視点を包含する矩形と、視点のなす角を示す。  $l$  個の視点を包含する矩形の面積の大きさと、判定する視点とその一つ前の視点とのなす角を求めて閾値処理により改行を判定する。矩形の面積が閾値より十分大きく、判定する視点が文書を読む時と逆向きになっているとき、改行と判定する。

#### 4.5 単語数推定法

表 1 に示す 4 つの万語計の単語数推定法について説明する。万語計を実現するために、様々な情報が利用可能である。ここでは、読むのにかかった時間と、アイトラッカから得られる視点情報、LLAH から得られる文書の情報の組み合わせについて考える。具体的には、(1) デバイスを用いず単純に読書時間から推定する手法、(2) アイトラッカと LLAH を用い、読んだ文書の一行あたりの平均単語数と読んだ行数から推定する手法、(3) (2) の手法に視点情報を加味した手法、(4) アイトラッカのみを用い、視点情報から単語数を推定する手法である。以下では、それぞれの手法について説明する。

手法 (1) では、ユーザが文書を読むのにかかった時間のみを用いて、学習により読んだ単語数を推定する。この手法では、アイトラッカから得られる視点情報や、LLAH から得られる文書の情報を用いない。学習には SVR (Support Vector

表 1 4 つの万語計

	推定法
手法 (1)	SVR (読書時間のみ)
手法 (2)	推定行数 $\times$ 行の平均単語数
手法 (3)	SVR (推定行数, 行の平均単語数, 特徴量)
手法 (4)	SVR (推定行数, 特徴量)

表 2 特徴量

視点情報	Fixation 回数
	視点の総移動距離
	Saccade の平均移動距離
	改行時の平均移動距離
	一行分の平均移動距離

Regression) を用いる。

手法 (2) では、改行回数と一行当たりの平均単語数を用いて読んだ単語数を推定する。この手法では LLAH によって文書画像上に変換された視点情報を用いる。改行判定から得られる改行回数を  $L$ 、ユーザの読んだ文書の一行当たりの平均単語数を  $N$  としたとき、 $N(L+1)$  として、ユーザの読んだ単語数を推定する。この手法では、改行判定に誤差が含まれると、推定単語数にも誤差が反映される。

手法 (3) では、手法 (2) で用いた情報に、視点情報から得られる特徴ベクトル  $\mathbf{x}_{doc}$  を加味して、 $f(N, L, \mathbf{x}_{doc})$  により読んだ単語数を推定する。関数  $f$  は、SVR による学習で定める。特徴ベクトル  $\mathbf{x}_{doc}$  は文書画像上に変換された視点情報から得られる 5 つの特徴量からなる。表 2 に 5 つの特徴量を示す。Fixation 回数とは、ユーザが文書を読む間にした Fixation の回数である。文書を読み始めてから、読み終わるまでに視点が移動した距離を、視点の総移動距離とし、総移動距離を Saccade

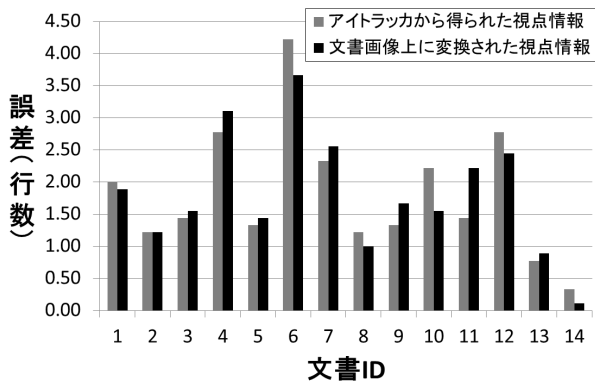


図 10 改行判定の誤差

の回数で割ったものを、Saccade の平均移動距離とする。各改行中に視点が移動した距離の平均を、改行時の平均移動距離とし、改行と改行の間にユーザの視点が移動した距離の平均を、一行当たりの視点の平均移動距離とする。

手法 (4) では、ユーザの視点情報のみを用いて、学習により読んだ単語数を推定する。この手法では、LLAH は用いず、アイトラッカから得られる視点情報のみを用いる。まず、アイトラッカから得られる視点情報に対し、Fixation の結合と Saccade の除去を施す。次に、改行を判定し、手法 (3) 同様に 5 つの特徴量からなる特徴ベクトル  $x_{raw}$  を抽出する。そして、得られた改行回数  $L$  と特徴ベクトル  $x_{raw}$  から、SVR による学習で関数  $f(L, x_{raw})$  を定め、読んだ単語数を推定する。

## 5. 実験

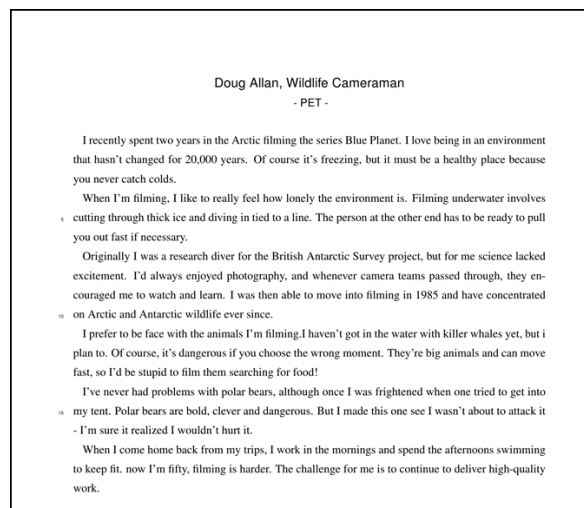
提案する 4 つの万語計を比較するため、それぞれの単語数推定精度を評価した。また、アイトラッカから得られる視点情報と文書画像上に変換された視点情報に対する、それぞれの改行判定精度を評価した

### 5.1 実験条件

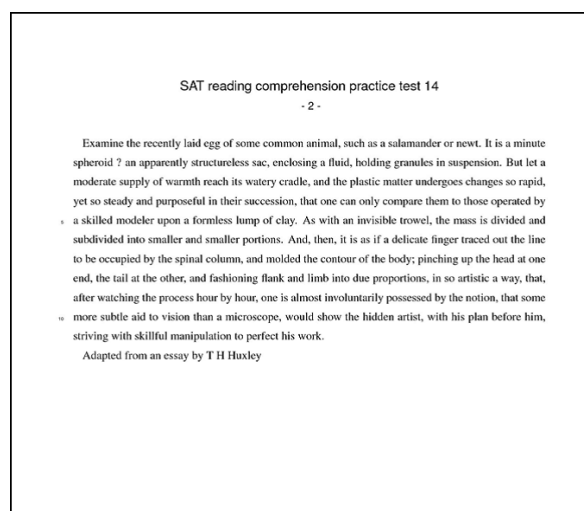
被験者 9 人が 14 枚の英文書を読み、その際の視点情報を記録した。実験手順について説明する。被験者はアイトラッカを装着し、文書を一枚ずつ読む。その際、文書は机の上に置いた状態で、自由な姿勢で読む。このようにして、文書一枚ずつの視点情報を記録した。SVR を用いる手法では、被験者 9 人のデータに対して、9-fold cross validation により、推定精度を評価した。

### 5.2 実験結果

改行判定の精度は、アイトラッカから得られた視点情報のみを用いた場合と LLAH も利用した場合で差がなく、平均誤差はともに 10.6 % (1.82 行) であった。このことから、アイトラッカから得られた視点情報に対しても、文書画像上の座標系に変換された視点に対しても、2 行程度の誤差で改行判定が可能であることがわかった。図 10 にアイトラッカから得られる視点情報と文書画像上に変換された視点情報に対する文書ごとの改行判定の誤差を示す。改行判定の結果、最も結果が悪かったのは文書 6 で、誤差は、アイトラッカから得られる視点情報の場



(a) 文書 6



(b) 文書 14

図 11 単語数の少ない行を含む文書

合は 22.2 % (4.22 行)、文書画像上に変換された視点情報の場合は 19.3 % (3.67 行) となった。一方で、最も結果が良かったのは文書 14 で、誤差は、アイトラッカから得られる視点情報の場合は 3.0 % (0.33 行)、文書画像上に変換された視点情報の場合は 1.0 % (0.11 行) となった。この原因として、図 11 に示すように、文書 6 が文書 14 と比べて、文書中に単語数が極端に少ない行を多く含むことが挙げられる。単語数が極端に少ない行は、行末から次の行の行頭までの視点の移動距離が小さくなり、改行と判定できなかったことがわかった。このような単語数が極端に少ない行や、視点の移動距離が小さい読み返しの判定をする必要があると考えられる。

表 3 に 4 つの万語計の結果を被験者ごとに示す。4 つの万語計のうち、最も平均誤差が小さかったのは手法 (3) で、11.3 % (30.9 単語) だった。また、最も平均誤差が大きかったのは手法 (1) で、21.4 % (53.3 単語) だった。手法 (1) の精度が低かった原因としては、読むのにかかった時間が文書の単語数だけではなく、文書の難易度によって変動することが挙げられる。文書の単語数が少なくても、文書の難易度が高い場合は

表3 被験者ごとの単語数推定誤差 (%)

手法	被験者									平均 誤差
	A	B	C	D	E	F	G	H	I	
(1)	15.3	20.3	21.5	18.4	21.0	24.8	23.4	23.5	24.1	21.4
(2)	9.5	8.4	12.1	12.4	8.2	15.5	11.7	16.3	8.5	11.9
(3)	10.4	7.4	11.5	13.9	9.8	12.8	9.1	15.7	11.1	11.3
(4)	10.0	13.4	18.4	10.9	10.5	15.5	11.2	17.0	12.8	13.3

読むのに多くの時間がかかる。また、文書の単語数が多くて、文書の難易度が低い場合は読むのにかかる時間が少なくなる。そのため、読んだ時間だけを用いた手法では誤差が大きくなったと考えられる。アイトラッカと LLAH を組み合わせた手法 (2) と手法 (3) が単語数推定の誤差が小さい結果となった。このことより、ユーザの単語数を推定するのに、アイトラッカと LLAH を組み合わせた万語計が有効であることがわかった。LLAH を用いずアイトラッカから得られる視点情報のみを用いた手法 (4) の平均誤差は 13.3 % (34.2 単語) だった。このことより、LLAH のデータベースに登録されていない文書に対しても、単語数の推定が可能であることがわかった。

被験者毎に結果を見ると、最も推定精度が低かったのは被験者 H だった。この原因として、被験者 H の視点情報にノイズが多く含まれていたことが挙げられる。また、被験者 H の視点情報には、単語を読み返す視点の動きが多く含まれていた。そのため、単語の読み返しを改行と判定し、実際より多くの単語を読んだと判定されたと考えられる。

## 6. まとめ

本稿では、Reading-Life Log を実現する一手法として、ユーザが読んだ単語数を推定する万語計を提案した。また、二種類の視点情報に対する改行判定精度の検証と、4つの万語計についての比較検討をした。

実験の結果、改行判定の精度は、アイトラッカから得られた視点情報のみを用いた場合と、LLAH も利用した場合で差がなく、誤差は 10.6 % (1.82 行) だった。4つの万語計の比較検討では、最も単語数推定精度が高かった手法は、読んだ行数と一行当たりの平均単語数、視点情報から得られる特徴量を用いた手法で、誤差は 11.3 % だった。これらのことから、提案する万語計の有効性が明らかになった。

今後の課題として、単語数の少ない行や読み返しに対応した改行判定の検討が挙げられる。また、手持ち文書などの、より自然な読書姿勢への対応が必要であり、視点情報から得られる特徴量の選定や、推定法の改良などによる、万語計の精度向上を目指す。さらに、単語数だけでなく単語自体を記録することや、視点情報から読んだ文書のジャンルを特定する手法 [13] などと組み合わせることが挙げられる。

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 (A) (25240028) の補助による。

## 文 献

- [1] 相澤清晴, “ライフログ,” 映像情報メディア学会誌, vol.63, no.4, pp.445–448, 2009.
- [2] A.R. Doherty, N. Caprani, C.Ó. Conaire, V. Kalnikaite, C. Gurrin, A.F. Smeaton, Noel EO’ Connor, “Passively recognising human activities through lifelogging,” *Computers in Human Behavior*, vol.27, no.5, pp.1948–1958, 2011.
- [3] V. Loseu, H. Ghasemzadeh, S. Ostadabbas, N. Raveendranathan, J. Malan, and R. Jafari, “Applications of sensing platforms with wearable computers,” *Proc. 3rd International Conference on Pervasive Technologies Related to Assistive Environments* ACM, p.53 2010.
- [4] J. Cheng, B. Zhou, K. Kunze, C.C. Rheinländer, S. Wille, N. Wehn, J. Weppner, and P. Lukowicz, “Activity recognition and nutrition monitoring in every day situations with a textile capacitive neckband,” *Proc. 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication* ACM, pp.155–158 2013.
- [5] A. SHIMOJO, S. MATSUMOTO, and M. NAKAMURA, “Evaluating lifelog common data model and mashup api through practical application development,” *電子情報通信学会技術研究報告. SS, ソフトウェアサイエンス*, vol.111, no.107, pp.29–34, 2011.
- [6] 川市仁史, K. Kunze, 黄瀬浩一, “文書画像検索を用いたリーディングライフログの提案,” *電子情報通信学会技術研究報告*, vol.112, no.495, pp.255–260, Mar. 2013.
- [7] T. Kimura, R. Huang, S. Uchida, M. Iwamura, S. Omachi, and K. Kise, “The reading-life log - technologies to recognize texts that we read,” *Proc. 12th International Conference on Document Analysis and Recognition (ICDAR 2013)*, pp.91–95, Aug. 2013.
- [8] 吉村和代, 川市仁史, K. Kunze, 黄瀬浩一, “アイトラッカで取得した視点情報と文書理解度の関係,” *電子情報通信学会技術研究報告*, vol.112, no.495, pp.261–266, Mar. 2013.
- [9] K. Kunze, H. Kawaichi, K. Yoshimura, and K. Kise, “Towards inferring language expertise using eye tracking,” *CHI’13 Extended Abstracts on Human Factors in Computing Systems*, pp.217–222, Apr. 2013.
- [10] P.T. Terenzini, L. Springer, E.T. Pascarella, and A. Nora, “Influences affecting the development of students’ critical thinking skills,” *Research in higher education*, vol.36, no.1, pp.23–39, 1995.
- [11] K. Kunze, H. Kawaichi, K. Kise, and K. Yoshimura, “The wordometer - estimating the number of words read using document image retrieval and mobile eye tracking,” *Proc. 12th International Conference on Document Analysis and Recognition (ICDAR 2013)*, pp.25–29, Aug. 2013.
- [12] K. Takeda, K. Kise, and M. Iwamura, “Real-time document image retrieval on a smartphone,” *Proc. 10th IAPR International Workshop on Document Analysis Systems (DAS2012)*, pp.225–229, Mar. 2012.
- [13] K. Kunze, Y. Utsumi, Y. Shiga, K. Kise, and A. Bulling, “I know what you are reading: recognition of document types using mobile eye tracking,” *Proc. 17th annual international symposium on International symposium on wearable computers*, pp.113–116, Sept. 2013.
- [14] R. Biedert, G. Buscher, and A. Dengel, “The eyebook - using eye tracking to enhance the reading experience,” *Inforatik Spektrum*, vol.33, no.3, pp.272–281, 2010.
- [15] G. Buscher and A. Dengel, “Gaze-based filtering of relevant document segments,” *International World Wide Web Conference (WWW)*, pp.20–24, 2009.
- [16] G. Buscher, A. Dengel, L. vanElst, and F. Mittag, “Generating and using gaze-based document annotations,” *CHI’08 Extended Abstracts on Human Factors in Computing Systems* ACM, pp.3045–3050 2008.