

視線情報に基づいたテキスト幅の最適化

Text width optimization based on gaze information

富田 恭平*1 原 忠義*2 相澤 彰子*1*2
 Kyohei Tomita Tadayoshi Hara Akiko Aizawa

*1 東京大学大学院情報理工学系研究科
 The University of Tokyo

*2 国立情報学研究所
 National Institute of Informatics

In this research, we examine the possibility to facilitate human reading by optimizing text width of a document. Return sweep, which is fast gaze movement from the end of a line to the beginning of the next line in reading text, is basically reluctant movement evoked by the layout of the document. We suppose that the optimal text width can minimize the time for this extra processing, and examined whether gaze information in reading text can be used for specifying the text width.

1. はじめに

スマートフォンやタブレットなどの普及により、電子デバイス上で文書を読む機会が増えてきている。電子デバイスの画面サイズなどの環境は多様であり、そのような多様な環境に適したレイアウトを人間がすべて用意するのは大変である。そこで、近年はそれぞれの環境に適した文書のレイアウトを自動で生成する手法が数多く研究されている [Hurst 09]。

文書レイアウト最適化の先行研究は文書の見目の美しさに基づいて行われている物が多い。しかし、文書レイアウトは文章を読む際の読みやすさにも注目して構成すべきである。文書が読みやすいかどうかを知るための可能性の一つとして、読み手の視線の動きを観察することが考えられる。

我々は視線情報の中でも特に、行末から次の行の行頭に移る動作 (return sweep) に注目した。Return sweep はレイアウト上の制約から生じてしまう、文章の内容を読み取るためには余分な動作であり、その余分な動作に要する時間を最小化するような幅が最適なテキスト幅であると考えた。本稿では、文書を読む際の視線の動きを計測し、その情報が最適なテキスト幅を求める際に役立つかどうかを検討する。

2. 関連研究

英語テキストの最適な幅に関する研究は古くから存在している。例えば、Spencer らは 1968 年に一行の長さは 70 文字を超えない方が良いと提案している [Spencer 68]。一方、Jackson らは 1979 年に一行あたり 75 文字が適切であるとした [Jackson 79]。また、Dyson らは一行あたり 55 文字の場合に効率よくテキストが読めるという結果を報告している [Dyson 01]。この他にも一行の最適な長さについては多くの研究がなされているが、それらの主張は相互に矛盾しあっていることも少なくない。

視線情報を用いて最適なテキスト幅を検討した研究も存在している。例えば Khan らは、瞬きと return sweep のタイミングが重なるようなテキスト幅が良いのではないかと考え、実験の結果、用意した環境の中ではテキスト幅が 10.5 インチの時にはフォントは 10 ポイントの Verdana で、黒色が最良であると結論づけた [Khan 12]。

この研究は、既に存在しているレイアウトに対して視線情報を収集し、レイアウトの良さ、この場合はテキストの幅の適

図 1: Return sweep を構成する saccade と fixation の例

切さを評価している。一方、我々の手法は、読んだテキストの幅の評価にとどまらず、ユーザごとに最適な幅を提案することを目的としている。

3. Return sweep

3.1 Return sweep の定義

文字を読んでいる際の人間の目の動きは大きく分けて 2 つに分類される。1 つは 1 点を 200~300ms ほど注視する fixation と呼ばれる動作で、もう 1 つは saccade と呼ばれる fixation から次の fixation へと移る 30ms ほどの素早い動きである。Saccade はその特性により更にいくつか分類でき、文を読み進める動作 (forward-reading) や既に読んだ文を読み直すために前に戻る動作 (regression) などがある。 [Rayner 98]

Return sweep とは、テキストを読んでいる最中の、行末から次の行頭へと移る視線の動きのことである。図 1 に return sweep が 2 つの saccade から構成される例を示す。図中の赤い丸が fixation を表しており、fixation をつなぐ赤い線が saccade を表している。

本稿では return sweep を下記のような saccade s_i と fixation f_j からなる列にて表現する。

$$\{s_0, f_1, s_1, \dots, f_n, s_n\} \quad (1)$$

ここで、 n は 0 以上の整数である。図 1 は、 $n = 1$ の場合の一例になっていると言える。

3.2 Return sweep とテキスト幅の関係

我々は、return sweep に要する時間はテキストの内容やフォントサイズの影響はあまり受けず、テキストの幅のみに依存するのではないかと考えた。実際にテキストの内容とフォントサイズがどの程度 return sweep に影響をあたえるかは 6.4 節と 6.5 節で検証するが、ここではテキスト幅以外の要素からの影響は無視できると考えて話を進める。

テキスト幅が長くなると、return sweep 中に余分な fixation が見られるようになり、1 回あたりの return sweep に要する時間は長くなる。逆に、テキスト幅が短くなると 1 回の return sweep に要する時間は短くなるものの、テキストの行数が増えるために return sweep の回数自体が増加する。つまり、return sweep 1 回あたりの所要時間と return sweep の回数はトレードオフの関係にある。

Return sweep は、レイアウト上の制約から生じる余分な動作であり、テキストの内容を理解するためには本来は必要のない動作である。そのような動作には時間をかけないほうが効率よくテキストを読むことができるので、本稿ではテキストを一読する際の return sweep に要する時間が最小になるようなテキスト幅を最適なテキスト幅とみなすことにする。また、この定義による最適なテキスト幅は読者によって異なりうることを 6.6 節にて実験で示す。

4. Return sweep の自動検出

Return sweep を用いた最適化を行うためには、集めた視線情報から return sweep を自動で検出する必要がある。本章ではその方法について議論する。本章で提案した手法の性能は 6.3 節にて検証する。

4.1 視線情報の前処理

我々の環境では、視線情報はスクリーン上の座標として 300Hz で観測されるので、そこから fixation と saccade を計算する必要がある。そのために、我々はまず視線情報に I-DT アルゴリズムを適用し fixation を得た [Salvucci 00]。その後、得られた fixation をつないだ直線を saccade であるとした。

4.2 ヒューリスティックアルゴリズム

視線情報を可視化して観察していると、ほとんどの return sweep には長い左向きの saccade が含まれるということ、その長い saccade の前後には短い左向きの saccade が連なっていることが多いということに気づく。これら 2 つの特徴を満たす saccade 列を検出するため、下記のヒューリスティックアルゴリズムを用いた。

1. Saccade を始点から終点に向かうベクトル $\mathbf{s}^{(k)}$ ($k = 1, \dots, N$) で表す。ただし、 N は全 saccade 数である。
2. $\mathbf{s}^{(k)}$ の x 要素 $s_x^{(k)}$ のみに着目して K 近傍法で saccade を 2 つのクラスに分ける。2 つのクラス加重心の x 座標を c^- と c^+ (ただし $c^- < c^+$) としたとき、 $s_x^{(k)} < \frac{2c^- + c^+}{3}$ を満たす $\mathbf{s}^{(k)}$ すべてに return sweep のフラグを立てる。
3. 2. でフラグを立てた saccade の前後をチェックし、左向きの saccade が連なっていればそれらにも return sweep のフラグを立てる。
4. 以上の過程で得られたフラグの立った連続した saccade 列がテキストの右端領域で始まり左端領域で終わっている場合、その saccade 列を return sweep とみなす。

5. テキスト幅の最適化

5.1 Return sweep の所要時間

テキスト領域が高さが h 、幅が w の長方形であるような文書中の各単語を一度ずつ注視しながら全文を一読する動作を考える。このとき、文書が $N+1$ 行で構成されるとすると、return sweep は合計 N 回行われることになる。k 番目の return sweep

に要した時間を t_k とする ($k = 1, \dots, N$)。このとき、文書を一読する際に return sweep に費やす合計時間は

$$\begin{aligned} \sum_{k=1}^N t_k &= N \cdot \frac{\sum_{k=1}^N t_k}{N} \\ &\propto h \cdot \frac{\sum_{k=1}^N t_k}{N} \\ &\propto \frac{1}{w} \cdot \bar{t}_{RS}(w) \end{aligned} \quad (2)$$

と表すことができる。ただし、式変形の途中でテキストの高さと行数は比例するという仮定、およびテキスト幅を変えても wh の値は大きくは変わらないであろうという仮定を用いた。

また、式 (2) では $\bar{t}_{RS}(w) = \frac{\sum_{k=1}^N t_k}{N}$ とした。このように置くと、 $\bar{t}_{RS}(w)$ は return sweep 1 回の平均所要時間を意味し、平均所要時間はテキスト幅と読者のみに依存することとなる。

この式 (2) を最小化することで、本稿で定義した最適な幅 w を得ることができるので、次の節では $\bar{t}_{RS}(w)$ の計算方法について議論する。

5.1.1 Return sweep 所要時間の推測

Return sweep を定量的に分析するために、本稿ではテキストを 5 つの縦セグメントに分割し、セグメント間の遷移列を分析の対象とした。テキストを等幅の 5 つのセグメントに分割し、誤差に対応するため左右のセグメントは幅を外方向に 1.5 倍にした。また、最左セグメントを Head (H)、最右セグメントを Tail (T)、中間のセグメントを左から I_1, I_2, I_3 、セグメントの外側の領域を Outside (O) とする。図 2 に、テキストを縦セグメントに分割した例を示す。

視線検出装置は主に縦方向に誤差が大きいという報告がなされており [Hyrskykari 06]、横書きの文書の場合、読者が読んでる行を特定することは難しい。この誤差の問題を避けるため、我々は縦方向に長いセグメントを用意することにした。

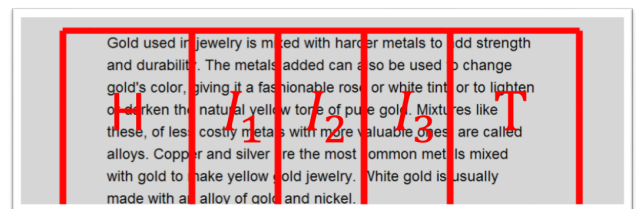


図 2: 縦セグメントへの分割例

次に、縦セグメントをもとにして各 return sweep に対して return sweep ベクトルを定義する。Return sweep ベクトルは 5 次元のベクトルであり、各要素が順に H, I_1, I_2, I_3, T セグメントに存在する fixation の個数を表している。

複数のテキスト幅 w_i ($i = 0, \dots, n$) に対する視線データ中の return sweep M 個すべてに対して return sweep を上の形式のベクトル $\mathbf{r}_{w_i}^{(k)}$ ($k = 1, \dots, M$) で表す。j 番目の要素の集合 $\{\mathbf{r}_{w_i}^{(k)}\}_j$ に対して線形回帰を適用することで、テキスト幅 w の時に j 番目の要素がとる値として期待される直線 $a_j w + b_j$ を得ることができる。Return sweep ベクトルの各要素は負の値には成り得ないことに気をつけると、任意のテキスト幅 w に対する return sweep ベクトルの j 番目の要素は次のように表すことができる。

$$\max(0, a_j w + b_j)$$

この結果から、任意のテキスト幅 w で期待される return sweep ベクトル $\mathbf{r}(w)$ を得ることができ、下記の式にて return sweep の平均所要時間を計算する。

$$\bar{t}_{RS}(w) = \|\mathbf{r}(w)\|_1 c_f + (\|\mathbf{r}(w)\|_1 + 1) c_s$$

ただし、 c_s, c_f はそれぞれ return sweep 内での saccade と fixation の平均所要時間である。なお、予備実験にて c_s, c_f の値はユーザによってばらつきがあるが同一ユーザであれば異なる文書に対しても近い値となることを確認しているため、 c_s, c_f は視線データからそれぞれのユーザごとに計算する必要がある。

5.2 最適なテキスト幅の予測

文書を逐次的に一読する際の return sweep の所要時間は式 (2) で表され、これこそが最適なテキスト幅を見つける際に最小化したい値であった。5.1 の方法にて任意の幅に対する return sweep 1 回の平均所要時間 $\bar{t}_{RS}(w)$ を得ることができれば、最適なテキスト幅 \hat{w} は、次の計算式にて予測できる。

$$\hat{w} = \operatorname{argmin}_w \frac{\bar{t}_{RS}(w)}{w}$$

6. 実験

6.1 実験環境

まず、我々が視線データを収集する際の環境を説明する。被験者の視線の動きを得るために Tobii TX300 アイトラッカーを使用した。モニターはアイトラッカー付属の、23 インチで 1920 x 1080 の解像度のもので使用した。このアイトラッカーを用いると、被験者が注視しているモニター上の点を 300Hz のサンプリングレートで得ることができる。被験者には目とモニターの距離が 65cm 程度になるような位置に座ってもらい、頭部を固定する器具などは用いずに視線の計測を行った。

6.2 視線データの収集

仮説や手法を検証するため、次の視線データ 1~3 を集めた。
視線データ 1: Wikipedia からダークエネルギーとサブプライム問題に関する、それぞれ 408 語、426 語で 900 ピクセルの幅の英文テキストを 39 人の被験者に読んでもらい、精度に問題のない 22 人分のデータを分析対象とした。

視線データ 2: アメリカの中学生向けの英語教材から文字サイズ 18 ポイントで 279 語のテキストと、文字サイズ 12 ポイントで 370 語のテキストを用意した。テキスト幅はともに 900 ピクセルであり、9 人の被験者のうち精度に問題のない 8 人分のデータを分析対象とした。

視線データ 3: アメリカの中学生向けの英語教材から、文字サイズ 18 ポイントでテキスト幅が 500, 700, 900, 1100, 1300 ピクセルのテキストを用意した。長さはそれぞれ 142, 200, 279, 362, 377 語であった。9 人の被験者のうち、精度に問題のない 8 人分のデータを分析対象とした。なお、うち 2 人分については、一部のみ精度の悪い部分があったため、その部分は手動で取り除いて使用した。

視線データ 2 と 3 のテキストの長さは、フォントサイズやテキスト幅が異なっても行数がほぼ同じになるような意図で選んだ。また、全ての視線データを集める際には、(1) 視線検出装置のキャリブレーション、(2) 視線検出装置で視線の動きを記録しながらテキストを一度読んでもらう、(3) テキストの内容理解に関する問題を解いてもらう、という流れで実験を進めた。

6.3 Return sweep の検出精度

4.2 節にて提案した return sweep 自動検出のアルゴリズムの性能を確かめる実験を行う。視線データ 1 で説明したダークエネルギーに関するテキストを読んだ視線データの内、6 人分を無作為に選んで return sweep のアノテーションを行った。視線データには計 6,350 個の saccade が含まれており、そのうち 502 個が return sweep を構成する saccade であった。表 1 の「自動検出」の行に 4.2 節の手法で検出した return sweep の精度を示す。

また、上述のアノテーションとは別の人によってなされたアノテーションの結果を、自動検出の結果と同様に評価したものを表 1 の「人手」の行で示す。F1 値がほぼ同じ値に到達していることから、このヒューリスティックのアルゴリズムは上限近くまで return sweep を検出できていることがわかる。

	Precision	Recall	F1
自動検出	0.846	0.906	0.875
人手	0.916	0.781	0.843

表 1: Return sweep 検出の精度

6.4 テキストの内容が return sweep に与える影響

6.4.1 実験設定

テキストの内容が return sweep に影響をあたえるかどうかを調べるために、視線データ 1 から return sweep の所要時間を比較する。そのため、2 つのテキスト A と B を読む際の視線データから 4.2 節の手法にて return sweep $\{r_1^{(A)}, \dots, r_m^{(A)}\}$ と $\{r_1^{(B)}, \dots, r_n^{(B)}\}$ を抽出した。その後、テキスト A と B での return sweep の所要時間の集合 $\{t_1^{(A)}, \dots, t_m^{(A)}\}$ と $\{t_1^{(B)}, \dots, t_n^{(B)}\}$ を求め、この 2 つの集合に対してウェルチの t 検定を適用して平均値に有意差があるかを調べた。

6.4.2 実験結果

22 人分の視線データのうち、有意水準を 0.05 とした場合に有意差があると認められたものは 1 人分のみであった。このことから、難しさが大きく変わらないテキストであれば、テキストの内容は return sweep には大きく影響を与えないと考えられる。

6.5 フォントサイズが return sweep に与える影響

6.5.1 実験設定

視線データ 2 に対して、6.4 節と同様の手法で、フォントサイズが異なるテキストを読む際の return sweep の平均所要時間に有意差があるかどうかをウェルチの t 検定にて調べた。

6.5.2 実験結果

8 人分の視線データのうち、有意水準を 0.05 とした場合に有意差があると認められたものは 1 人分のみであった。ただし、有意差は認められないものの、文字が小さいテキストを読む場合のほうが return sweep の平均所要時間が長くなっている視線データが 4 人分見られた。フォントサイズが極端でない場合は、フォントサイズは return sweep の平均所要時間に大きく影響を与えることはないと考えられるが、あまりにもフォントサイズに差がある場合は return sweep にも影響が出る可能性を考慮しなければならないと考えられる。

6.4 節と 6.5 節の結果は、我々がおいた、return sweep の平均時間はテキストの内容とフォントサイズには依存しない（テキスト幅と読者のみに依存する）という仮定を、内容の難易度やフォントサイズが極端でない場合において支持するものである。

6.6 最適なテキスト幅の個人差

まず、視線データ **3** から、最適なテキスト幅には個人差があることを示す。「テキストを一読した際に return sweep に要する時間が最小となるテキスト幅」が最適なテキスト幅であったため、8人分のデータから、テキストごとに return sweep の所要時間を計算する。異なる幅で使用したテキストの長さは異なるので、return sweep の所要時間はテキストの単語数にて正規化を行った。その結果、5種類の幅のテキストの内、500, 700, 900, 1100, 1300 ピクセルの幅が最適であった被験者がそれぞれ1人、1人、1人、2人、3人であった。この結果より、最適なテキスト幅は人によって異なることが示唆される。

6.7 提案手法の実装と評価

5.2 章で提案したテキスト幅の最適化を実装し、視線データ **3** を用いて実際に最適化を行った。ある被験者 u の、視線データを集めたテキスト中での最適な幅が $w_u \in \{500, 700, 900, 1100, 1300\}$ であったとする。我々の手法中で使用した目的関数 $f(w) = \bar{t}_{RS}(w)/w$ を使用して、500~1300 ピクセルの5種類の幅のうち最適であるものを推測し、それが w_u と一致するかどうかを確かめる。つまり、

$$\hat{w}_u = \underset{w \in \{500, 700, 900, 1100, 1300\}}{\operatorname{argmin}} f(w)$$

とした時に、 $w_u = \hat{w}_u$ が成立するかどうかを確かめる。その結果、8人中4人については w_u と \hat{w}_u が一致していた。

一方、残りの4人については、目的関数が単調減少の形になっているという共通点があった。これは、return sweep ベクトルの各要素を線形回帰している部分で、テキスト幅の増加に対する fixation 数の急速な増加が正しく捉えられていないからであると考えられる。

図3に、正規化された return sweep の合計所要時間と、目的関数を示す。上段が目的関数が正しく最適な幅を予測している例で、下段が失敗している例である。右側の目的関数は、式(2)より左側の合計所要時間に比例することが期待されている。実際、上段の例では両者の形は似ており、目的関数が正しく機能している様子がわかる。一方、下段の場合は、テキスト幅の増加に伴う return sweep 時間の増加が捉えきれず、単調減少となってしまっている。

7. おわりに

本稿では、テキストの最適な幅を求める際に、読み手の視線の動きに関する情報が役に立つかどうかを検討した。Return sweep と呼ばれる動作に注目して最適化を行うことで、8人の被験者のうち4人についてはテキスト幅の最適性を予測することができた。この結果から、テキストを読む際の視線情報、中でも特に return sweep に関する情報は、テキスト幅を決める際の有益な情報になり得ると言えるだろう。

また実験では、テキストの内容やフォントサイズは、return sweep の平均所要時間に大きな影響を与えないことが示唆された。さらに、その平均所要時間は読み手によっても異なることも示された。文章の諸条件や読み手の違いが return sweep に与える影響を調査した研究は我々の知る限り過去には無く、これらは本研究によって得られた新しい知見であるといえるのではないだろうか。

テキストを読む際の視線の動きは非常に多くの情報を含んでおり、我々はテキスト幅以外のレイアウトの最適化にも有益であると考えている。本手法の性能の向上と共に、テキスト幅とは異なるレイアウトの最適化を視線情報をもとに行うことは、今後取り組んでいきたい課題である。

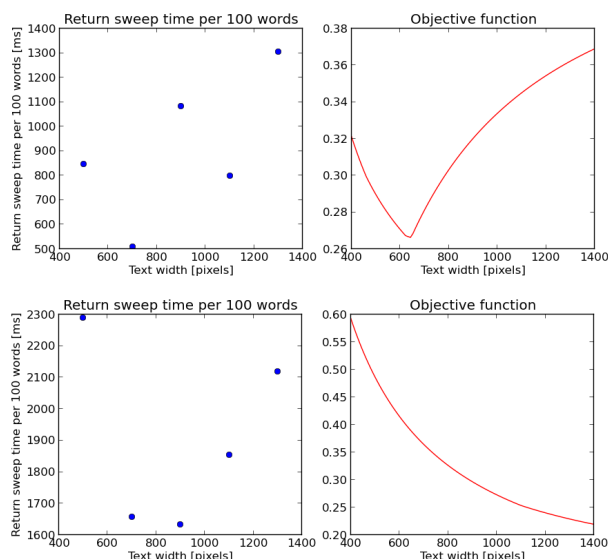


図3: 正規化した return sweep 所要時間 (左) と提案手法の目的関数 (右)。上段が最適化の結果うまく最適幅を予測できている被験者の例で、下段がそうでない被験者の例である。

参考文献

- [Dyson 01] Dyson, M. C. and Haselgrove, M.: The influence of reading speed and line length on the effectiveness of reading from screen, *International Journal of Human-Computer Studies*, Vol. 54, No. 4, pp. 585 – 612 (2001)
- [Hurst 09] Hurst, N., Li, W., and Marriott, K.: Review of automatic document formatting, in *ACM Symposium on Document Engineering*, pp. 99–108 (2009)
- [Hyrskykari 06] Hyrskykari, A.: Utilizing eye movements: Overcoming inaccuracy while tracking the focus of attention during reading, *Computers in Human Behavior*, Vol. 22, pp. 657–671 (2006)
- [Jackson 79] Jackson, M. D. and McClelland, J. L.: Processing determinants of reading speed, *Journal of Experimental Psychology: General*, Vol. 108, No. 2, pp. 151–181 (1979)
- [Khan 12] Khan, M. and Raza, M.: Article: Suitable Length of Text Line on the Bases of Eye Blink for Reducing Maximum Focus Losses, *International Journal of Computer Applications*, Vol. 37, No. 8, pp. 15–21 (2012)
- [Rayner 98] Rayner, K.: Eye movements in reading and information processing: 20 years of research, *Psychological Bulletin*, Vol. 124, pp. 372–422 (1998)
- [Salvucci 00] Salvucci, D. D. and Goldberg, J. H.: Identifying fixations and saccades in eye-tracking protocols, in *Proceedings of the 2000 symposium on Eye tracking research & applications*, ETRA '00, pp. 71–78, ACM (2000)
- [Spencer 68] Spencer, H.: *The visible word*, Iowa State University Press (1968)