

中心窩視覚における読書時の視点移動の実現

石原由紀夫[†] 守田了^{††}

[†]山口大学大学院理工学研究科 ^{††}山口大学工学部

〒755-8611 山口県宇部市常盤台 2557

山口大学工学部

E-mail: {yukio, morita}@cs.csse.yamaguchi-u.ac.jp

人が文書を読んでいる時の視点は、単語の中心にくる傾向がある。これは、視点の移動において中心窩付近の文字の情報が使われているからである。文字の情報とは、文字の形、単語の長さなどである。そこで、本稿では中心窩付近の単語の長さを使用して、読書時の視点移動を実現する。さらに、中心窩付近の単語の長さを使用した場合と使用しなかった場合において視点移動を行ない、各単語のそれぞれの文字における視点の割合をそれぞれの場合において調べ比較する。そして、前者の方は視点が単語の中心にくる傾向があることを確認する。

中心窩視覚、読書、視点移動

Realizing Eye Movements using Foveated Vision in Reading

Yukio Ishihara[†] Satoru Morita^{††}

[†]Graduate School of Science and Engineering, Yamaguchi University

^{††}Faculty of Engineering, Yamaguchi University

Yamaguchi University

2557 Tokiwadai, Ube, 755-8611 Japan

E-mail: {yukio, morita}@cs.csse.yamaguchi-u.ac.jp

View point tend to be middle point of a word in reading, because letter information in the periphery is used for eye movements. Letter information is letter-shape, word-length. In this manuscript, we realize eye movements using word-length information in reading. In addition, we compare the eye movements using word-length information to the eye movement without using word-length information, and we confirm that view point tend to be middle point of a word for the former.

Foveated vision, Reading, Eye movement

1. はじめに

網膜の中心は、中心窩と呼ばれる。中心窩では密に情報を得て、その周辺では粗に情報を得ている。このため、私たちが物体を細かく識別できるのは、網膜内でも中心窩付近に限られている。そこで、眼球を動かし見ようとする物体を中心窩付近へもってくる必要がある。

中心窩周辺では密に情報を得てはいないが、そこで物体の移動が感知されると、瞬間にその方向に視線を向けるように眼球運動が行われる。このような眼球運動には、視線が物体を追いかける時に生ずる連続的な運動と、視線から離れている物体を捕らえるための跳躍的な運動とがある。前者を随従運動、後者を断続性運動（サッカード）と呼ぶ。この2つの眼球の運動によって、見ようとする物体を中心窩付近へもってくる。また、眼球運動の間の物体の一点を見る動作を凝視と呼ぶ。

一回の凝視で得られる観察対象の情報は、中心窩付近では密に得られ、その周辺では粗に得られる。そのため、観察対象の情報を密に得るためには、視線を移動して複数回観察を行う必要がある。この時の視線の移動は、行う作業によって異なると考えられる。作業が静的なものを対象としている場合には、視線の移動は断続性運動のみで行われる。本稿では、作業を読書として人の行なう断続的な視線移動をコンピュータ上に実現することを目標とし、人の行なう視線移動の心理学的局面を読書時の視線移動に取り入れることを考える。

人が文書を読んでいる時の視点移動には、中心窩付近の文字情報が使われていて、視点の右側10～11文字における文字の形の情報と視点の右側15文字における文字の長さの情報が使われているという報告がある[1]。また、読書時の視点と声に出して読んでいる位置の関係について調査[2]されていて、そこには視点の方が声に出して読んでいる位置より前方にあり、文の最初ではその距離が長く文の最後では距離が短いという報告がある。中心窩付近の文字情報を隠した場合と隠さなかった場合において読書時の視点移動と単に文字を探す時の視点移動を調べ、後者の方が前者に比べると中心窩付近の文字情報の変化による影響を受けなかったという報告もある[3]。そこで、本稿では単語の長さを使用した視点移動を実現する。

2章では、中心窩と視線の移動について述べる。3章では、人の行なう視線移動の心理学的局面を取り入れた読書時の視点の断続的移動について説明する。4章では、心理学的局面を取り入れた場合と取り入れてない場合において、実際にコンピュータ上で視点の制御を行い、それぞれを比較する。

2. 中心窩視覚

2.1. 中心窩

網膜の中心は、中心窩と呼ばれる。中心窩では密に情報を得て、その周辺では粗に情報を得ている。このため、私たちが物体を細かく識別できるのは、網膜内でも中心窩付近に限られている。マシンビジョンにおいては空間的不定のイメージサンプリングモデルとして log-polar mapping モデルが広く使われている。Wilson は空間的不定の receptive field 配置において一つのモデルを提案して、人のコントラスト感性機能がそれらからどのように現れているのかを説明している[4]。receptive field の中心は、中心窩の中心をもとに円状に位置づけられている。n 番目の円の離心率 R_n は次のようになる。

$$R_n = R_0 \left(1 + \frac{2(1-Ov)C_m}{2-(1-Ov)C_m} \right)^n$$

ここで、 R_0 は中心窩の半径であり、 C_m は中心窩の中心から receptive field の中心までの離心率に対する receptive field の直径の割合である。 Ov は、隣接する receptive field の部分的重複率であり、receptive field が他の receptive field に接触していれば $Ov = 0$ となり、receptive field が他の receptive field の中心に達していれば $Ov = 0.5$ となる。また、n 番目の円上での receptive field の半径は $\frac{C_m R_n}{2}$ であり、一つの円上の receptive field の数は $\frac{2\pi}{C_m(1-Ov)}$ である。

以上のモデルで、receptive field の中心の色をサンプリングして、横方向に θ 成分（角度線分）、縦方向に R 成分（半径成分）で表わされる画像を $R\theta$ 画像と呼ぶ。 xy を座標にもつ中心窩画像は $R\theta$ 画像をもとに生成される。 $R\theta$ 画像から中心窩画像への変換式は以下の通りである。

$$x = R \cos(\theta) + C_x$$

$$y = R \sin(\theta) + C_y$$

ここで、 C_x, C_y は中心窩画像上の中心の座標である。

本稿では、このモデルに従って中心窩画像を生成する。

2.2. サッカード

中心窩付近では、物体の情報を密に得ている。その周辺では、情報を密に得てはいないが、そこで物体の移動が感知されると、瞬間にその方向に視線を向けるように眼球運動が行われる。このような眼球運動には、視線が物体を追いかける時に生ずる連続的な運動と、視線から離れている物体を捕らえるための跳躍的な運動とがある。前者を随従運動、後者を断続性運動（サッカード）と呼ぶ。

サッカードを行う場合、次の視点の決定は、中心窩付近の情報から特徴を抽出することにより行う。抽出する特徴として“エッジ”を用いる。特徴抽出は、あらかじめ準備してあるエッジのテンプレートと $R\theta$ 画像を比較することにより行う。特徴が抽出された位置が次の視点決定に使用

される。

3. 読書時の視点移動

3.1. 読書

読書には、文書に沿って視点を動かしつつ、その情報を得るという作業が含まれている。この時の視点の動きにおいて、興味深い報告がある。McConkie, Kerr, Reddiz, Zolaらは、英語の文書を読んでいる時の視点は単語の中心にくる傾向があると報告している [1]。図1はその報告結果であり、横軸は単語の各文字の位置を示し、縦軸は各文字における視点の割合を示している。図1より、それぞれの長さの単語上における視点の割合は、単語の中心で高くなっていることが分かる。これは、次の単語が長い単語である時は大きく視点が移動して、次の単語が短い単語である時は視点が短く移動するということである。つまり、視点の移動には、中心窩周辺の文字情報が何らかの形で使われているということである。それでは、中心窩周辺のどんな情報がどのくらい使われているのであろうか。McConkie, Raynerは、次のような実験を行った。人の視点を追うシステムを用いて、現在の視点の右側数文字よりさらに右側の単語の長さの情報を隠した場合と隠さない場合の両方において文書の読みやすさとサッカードの距離を調べた。この実験より、視点の右側15文字分の間で単語の長さの情報が使われているという結果が得られた。

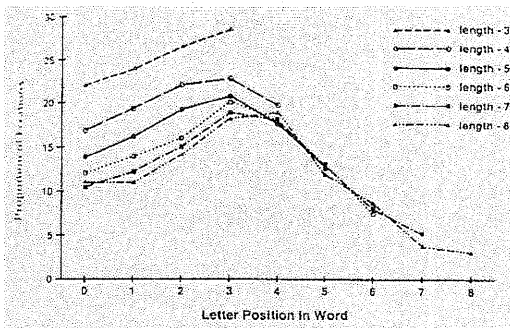


図1: 単語上の各文字における最初の視点の割合 (参考文献 [1] より転載)

本稿では、以上のように読書時の視点の移動には、視点の右側の単語の長さが関係しているという心理学的局面を取り入れて、読書時の視点制御を実現する。

3.2. 読書時の視点移動

読書時の視点移動は、文書の理解度、年齢(経験)により変化し、同じ人が同じ文書を読んだとしてもその時々で異

なる。しかし、読書時の一般的な視点移動は同じであり、次のようなものである。視点は行の先頭から行の終わりまで行に沿って移動し、行の終わりまで視点が移動すると、次の行の先頭へ視点が移動する。ふたたび、その行に沿って視点が移動する。読書時の一般的な視点移動はこの繰り返しである。本稿では、以上の視点移動を実現するために、次の3つの処理を考える。

1. 文書に沿って視点を動かす処理
2. 行の終わりを探す処理
3. 次の行を探す処理

これらの3つの処理を用いて視点移動を行なう。

3つの処理の流れについて説明する。3つの処理の流れは図2のようになる。1、2が基本的なループ処理で、2で行の終わりが偶数回検出されると、3が実行される。

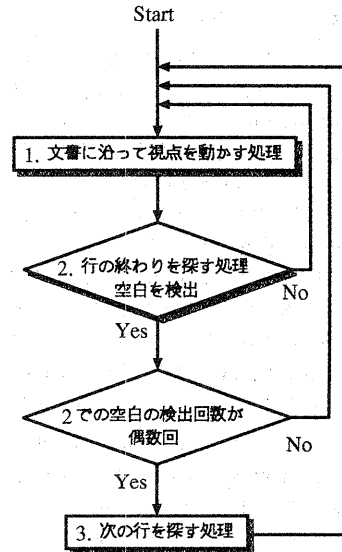


図2: 処理の流れ

次に以上の処理の流れを視点移動をもとに説明する。まず、視点が行の先頭にある時、1と2の繰り返しにより、視点が行の終わりまで進む。視点が行の終わりに近づくと2により前方に空白が検出される。空白が検出されると、1と2の繰り返しにより、視点は今移動してきた行を逆に移動し、行の先頭に戻る。行の先頭に近づくと2により前方に空白が検出されると3により次の行が検出されて、視点が次の行に移動する。

次にこれらの3つの処理をそれぞれ説明する前に注目領域を定義する。

人が本を読む時に、視点は行に沿って行上を進み、行の途中で視点が前の行や次の行に移動することはない。ここで、視点の移動にはある種の制限があると考えることができる。本稿では、視点移動を制限するものとして注目領域を考える。注目領域は次の視点を決定するための領域である。言い換えれば、次の視点は注目領域の中から決定される。次にこの注目領域を用いて先に述べた3つの処理をそれぞれ説明する。

[1. 文書に沿って視点を動かす処理]

この処理は注目領域を移動させることにより行なう。まず、現在の視点の周りにおいてエッジの密度が最大の方向へ注目領域を移動することを考える。現在の視点の周りでエッジの密度が高い方向を探すために、現在の視点の中心窩画像(図3(a))より抽出されたエッジ(図3(b))を用いて、図4のように注目領域をそれぞれ方向に回転させ、領域内のエッジの密度を調べる。その密度が最大の方向へ注目領域を移動する。

しかしこのままでは、視点は行に沿って進まず振動してしまう。なぜなら、注目領域は行上にあるため、エッジの密度の最大の方向と対称な方向においてもエッジの密度は大きく、注目領域が振動するからである。そこで、以前に注目領域が移動した方向を重視し、エッジの密度と以前の注目領域の移動方向の両方を考慮して、次の注目領域の移動方向を決定する。このようにすることにより、行上での視点の振動はなくなり、エッジの密度が高くかつ以前の移動方向とほぼ同じ方向へ視点を移動させることができる。また、現在の視点の周りのエッジの密度を調べて注目領域を移動させているので、行のゆがみにも柔軟に対応できる。

次に、前節で述べたように、視点の移動には視点の進行方向の語長が関係しているため、視点の進行方向の語長に応じて注目領域を変化させることを考える。本稿では、注目領域において単語間の空白を抽出し、注目領域が単語を覆うようにする(図5)。

以上のように注目領域を移動しその領域内から次の視点を選ぶ、という繰り返して視点を行に沿って移動させる。注目領域内での視点の決定は、領域内のエッジ抽出点において解像度の高低に応じ重みづけを行ない、解像度が低い点を優先して選ぶことにより行なう。

[2. 行の終わりを探す処理]

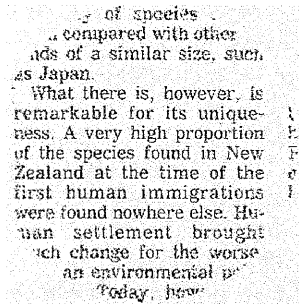
この処理は、視点の移動方向において空白を探すことにより行なう。

視点の移動方向において中心窩画像(図3(a))より抽出されたエッジ(図3(b))をもとにその密度を調べ(図6)、エッジの密度が少ない位置を行の終わりと判断する。

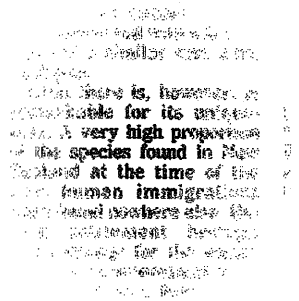
[3. 次の行を探す処理]

この処理は、視点の移動方向に対しその左右において空白を探すことにより行なう。

視点の移動方向に対しその左右において抽出されたエッジの密度を調べ(図7)、エッジの密度が少ない位置を行間の空白と判断する。



(a) 中心窩画像



(b) 中心窩画像より抽出されたエッジ

図3: 中心窩画像とエッジ

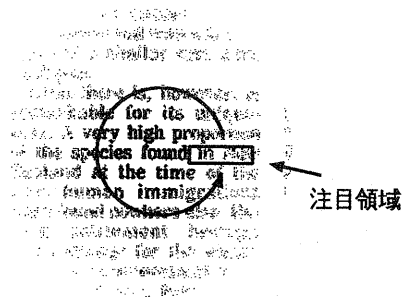


図4: エッジ密度の計算

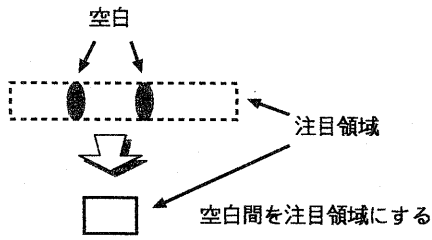


図 5: 注目領域の変更

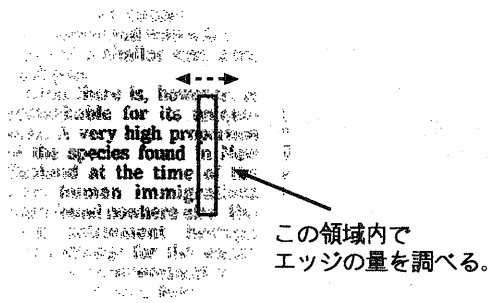


図 6: 行の終わりの空白検出

4. 実験

本章では、3章で述べた読書時の視点移動、つまり視点の進行方向にある単語の長さに応じて注目領域を変化させて行なう視点移動と、単語の長さに関わらず注目領域の長さを一定に行なう視点移動を比較する。そして、前者の方が人の行なう視点移動により似ていることを確かめる。両者を比較するための指標として、3章で述べた「視点は単語の中心にくる傾向がある」という報告をもとに、単語上での視点を使用する。単語上での最初の視点が単語の中心に移動する割合が大きければ、人の行なう視点移動により似ているとする。

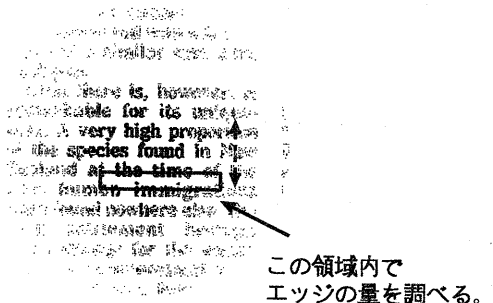


図 7: 行間の空白検出

中心窩画像の生成は Wilson のモデルを用いて行ない、 R_0 (中心窩の半径)=7、 C_m (中心窩の中心から receptive field の中心までの離心率に対する receptive field の直径の割合)=0.3、 O_v (隣接する receptive field の部分的重複率)=0.96 とする。また、視野の半径は200とし、注目領域の長さは単語間の空白の検出が可能な範囲を考慮し54とする。以上のようにパラメータ設定を行ない、

- 単語の長さに関わらず、注目領域を一定にする場合
- 単語間の長さに応じて、注目領域を変える場合

の2つの場合において視点移動を行なった。それぞれの場合において約1000単語を読ませ、それぞれの単語長において単語上での最初の視点を調べた。図8は、実験に用いた文書画像の一例であり、図9はその時の視点の軌跡を示している。図10、図11は先に述べた2つの場合における単語上での最初の視点の割合を示したものであり、横軸は単語の各文字の位置を示し、縦軸は各文字における視点の割合を示している。



図 8: 文書画像

図9より、3章で述べた視点の移動方法より、視点が文書の行に沿って行の終わりまで移動し、そして行の先頭まで戻り次の行へ移動しているのが分かる。図10、図11を比べると、注目領域を変化させた方は各単語における最初の視点が先頭文字に移動する割合が小さくなっている。これは、単語を覆うように注目領域を変化させたことにより、注目領域内の両側付近の文字上への視点移動が抑えられたからだと考えられる。したがって、注目領域を変化させた方が視点が単語の中心に移動する傾向があり、人の行なう視点移動により似ていることが分かる。

次に、図1と図11を比較する。2つの図を比較すると、6~8文字の長い単語において、図1では単語上での視点の割合の最大が3~4文字目にあるのに対し、図11では2~3文字目に最大がある。つまり、本稿で行った実験にお

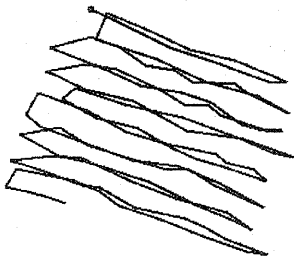


図 9: 視点の軌跡

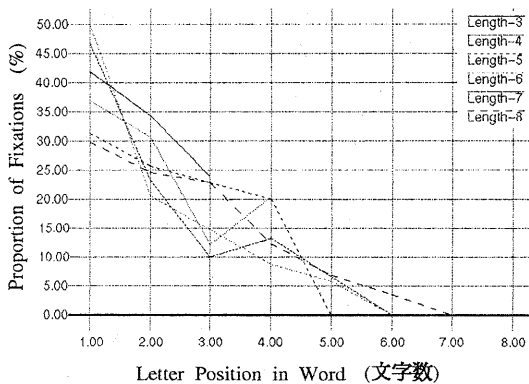


図 10: 単語上の各文字における最初の視点の割合 (注目領域を一定)

いては、長い単語において視点が単語の中心より前に移動する傾向が見られる。また、図 1 では長い単語において、単語の 6～8 文字目が視点となる割合がある程度大きい。

ここで、長い単語において視点が単語の中心より前に移動する傾向と注目領域との関係を考える。注目領域の最大長を短くした場合には、長い単語においては単語の先頭の文字に視点が移動し、後ろの文字には移動しなくなり、単語の中心より前での視点の割合が大きくなる。反対に、注目領域の最大長を長くした場合には、長い単語も注目領域で覆うことができ、視点は単語の中心に移動すると考えられる。したがって、本稿の実験において長い単語上で視点が単語の中心より前に移動した原因は、注目領域の最大の長さが短すぎて、長い単語の全体を覆うことができなかったからであり、注目領域の最大長を長くすることによって改善できる。

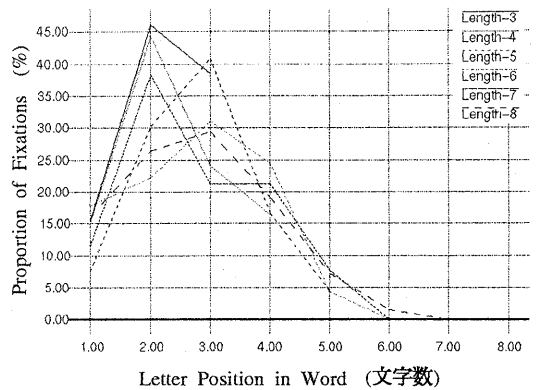


図 11: 単語上の各文字における最初の視点の割合 (注目領域を変化)

5. おわりに

人が文書を読む時の視点移動は、現在の視点の進行方向にある単語の長さに関係している。そこで、本稿ではこの心理学的局面を取り入れ、読書時の視点移動を実現した。さらに、現在の視点の進行方向にある単語の長さを視点移動に使用した場合と使用しなかった場合においてシミュレーションを行なった。そして、それぞれの視点移動を比較し、前者の方が人の行なう視点移動により似ていることを確かめた。

謝辞

第 2 著者は、本研究に関して議論いただいたユトレヒト大学 Prof. J. J. Koenderink に感謝します。

参考文献

- [1] K. Rayer, "Eye Movements in Reading," pp.7-20, Academic Press, 1983
- [2] R. A. Monty and J. W. Senders, "Eye Movements and Psychological Process", Lawrence Erlbaum Associates pub., pp.371-395, 1976.
- [3] R. A. Monty and J. W. Senders, "Eye Movements and Psychological Process", Lawrence Erlbaum Associates pub., pp.417-427, 1976.
- [4] S. W. Wilson: "On the retina-cortical mapping," Int. J. Man-Machine Stud. 18, pp.361-389, 1983.