

読書時の視点移動における時間特性

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

E-mail: {yukio, morita}@cv.kde.yamaguchi-u.ac.jp

英文の読書時におけるヒトの視点移動では、単語内の最初の視点が3~5文字目である時に、視点が次の単語へ移動するまでの時間が最小になる最適停留位置効果が現れる。この原因は、文字の識別率が視点から離れるに従って減少し、視野の右側の減少率より左側の減少率が大きいためであることが報告されている。本稿ではこの報告結果にもとづいて、最適停留位置効果が現れる視点移動をコンピュータ上に実現する。次に、中心窩の半径を6通りに変更させ、視点が次の単語へ移動するまでの時間を計測する。中心窩の半径が60ピクセルの時にその時間が最小になることを示す。さらにヒトの中心窩の角度が2°であることから、1文字が0.3cmの英文において視点が次の単語へ移動するまでの時間が最小になる読書時の文書と眼の距離は34cmであることを示す。

キーワード: 中心窩視覚、読書、視点移動、最適停留位置効果

Time Characteristic of Eye Movement in Reading

Yukio Ishihara[†] Satoru Morita^{††}

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

^{††}Faculty of Engineering, Yamaguchi University

2557 Tokiwadai, Ube, 755-8611 Japan

E-mail: {yukio, morita}@cv.kde.yamaguchi-u.ac.jp

In reading English literature, it is reported that the time required to move eyes to next word is the minimum when the first viewpoint is between third letter and fifth letter in a word, because the decrease rate of the rate matching a letter in the left side of visual field is bigger than the one in the right side of visual field as the distance between the letter and the center of visual field becomes to be long. The optimal viewing position is the position where the time required to move eyes is minimum. In this paper, we simulate the eye movement with the reported human characteristic in matching a letter in the right and left side for the center of the visual field, and show that the reported human optimal viewing position is appeared in the simulated eye movement. Next, we investigate the time required to move eyes to the word for the six kinds of the radius of the fovea. We show that the time is the minimum when the radius of the fovea is 60 pixels. Furthermore, we show that the optimal distance between the eye and the English literature is 34 cm in reading the English literature written by letters of 0.3 cm, because the angle of the human fovea is 2 degree.

Keyword: Foveated vision, Reading, Eye movement, Optimal viewing position effect

1. まえがき

ヒトが文書を読んでいる時の視点は、主に停留 (fixation) とサッカド (saccade) の繰り返しからなり、それらは初期視覚情報処理や言語情報処理と大きく関係している。視点移動と初期視覚情報処理、言語情報処理の関係は様々に報告されている。

Buswellは、読書時の視点と声に出して読んでいる位置の関係进行调查し[1]、視点は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報告している。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[2]。O'Reganらは、5~11文字長の単語を音読させた場合、語のどの文字位置に最初に停留したかによって音読潜時に違いが見られ、それぞれの単語長で音読潜時が最小となる最適停留位置があることを見出した[3]。また、O'ReganとLévy-Schoenは、単語へ1回のみの停留をする確率、2回の停留をする確率、3回の停留をする確率を最初の停留位置をもとに調べた。1回のみの停留をする確率は、最初に最適停留位置に停留した時に高く、2回又は3回の停留をする確率は最適停留位置から離れるに従って高くなると報告している[4]。この他に、視点移動で使用されている視点付近の情報を、McConkieとRaynerは移動窓法を用いて調べ、Raynerは境界法を用いて調べていることが文献[5]に示されている。

以上の研究では、ヒトの英文の読書時における視点移動の特徴を心理学的局面にもとづいて調査し、文書の文字を様々に変更し視点移動への影響を調べている。しかし、ヒトの眼の機能に制限を加え、視点移動への影響を調べるのは困難である。本稿では、視点移動をコンピュータ上に実現し、中心窩視覚に様々な制限を加え視点移動への影響を調べる。まず、単語内の最初の視点が3~5文字目である時に視点が次の単語へ移動するまでの時間が最小になる最適停留位置効果を実現する。最適停留位置が単語の3~5文字目にある原因は、文字列中の文字の識別率が視野の中心から離れるに従って減少し、右視野の識別率の減少率は左視野のそれより小さいことであるとNazirらは報告している[6]。本稿では、この報告結果にもとづき最適停留効果が現れる視点移動を実現する。次に、中心窩視覚の中心窩の半径を30、40、50、60、120、150ピクセルの6通りに変更し、視点が次の単語へ移動するまでの時間を計測する。中心窩の半径が60ピクセルの時、視点が次の単語へ移動するまでの時間が最小になることを示す。中心窩に含まれる文字列の長さは、読書時の文書と眼の距離に比例し、文書と眼の距離を離すに従い、中心窩に含まれる文字列の長さも長くなる。つまり、中心窩の半径を変更し中心窩に含まれる文

字列の長さを長くすることは、読書時の文書と眼の距離を離すことと同じである。視点が次の単語へ移動するまでの時間が最小になる中心窩の半径が60ピクセルであることから、英文の読書時においてその時間を最小にする文書と眼の距離を計算する。ヒトの中心窩の角度が2°であることから、1文字が0.3cmの英文において視点が次の単語へ移動するまでの時間が最小になる文書と眼の距離は34cmであることを示す。

2.では、中心窩視覚について述べる。3.では、視点の右側における文字の識別率の減少率は左側のそれに比べて低いという報告にもとづいて、最適停留位置効果が現れる英文の読書時の視点移動を実現する。4.では、中心窩の半径を6通りに変更し視点が次の単語へ移動するまでの時間を計測する。ヒトの中心窩の角度が2°であることから、1文字が0.3cmの英文において視点が次の単語へ移動するまでの時間が最小になる文書と眼の距離は34cmであることを示す。

2. 中心窩視覚

本章では、視点移動を実現する際に使用する中心窩視覚について説明する。

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[7]が広く使われている。

Wilsonはlog-polar mappingモデルに従い空間的不定の受容野配置において次に述べるモデルを提案し、ヒトのコントラスト感性機能を説明している[8]。受容野の中心は、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。

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

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

以上のモデルでサンプリングを行う際に、同一ピクセル内の複数回のサンプリングを防ぐために $R_1 - R_0 = 1$ となるように R_0 、 Ov を用いて Cm を次式で求める。

$$Cm = \frac{2}{(2R_0 + 1)(1 - Ov)} \quad (2)$$

$(R_0 > 0, 0 < Ov < 1)$

式(1)は、式(3)を用いて次式で表される。

$$R_n = R_0 \left(1 + \frac{1}{R_0}\right)^n \quad (3)$$

以上のモデルで、受容野の中心のRGB値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は次の通りである。生成される中心窩画像の中心(C_x, C_y)を基点として、画像の点(x, y)での角度 θ と半径 R を求める。

$$R = \sqrt{(x - C_x)^2 + (y - C_y)^2} \quad (4)$$

$$\theta_c = \arccos \frac{x - C_x}{R} \quad (0 \leq \theta_c \leq \pi) \quad (5)$$

$$\theta_s = \arcsin \frac{C_y - y}{R} \quad \left(-\frac{\pi}{2} \leq \theta_s \leq \frac{\pi}{2}\right) \quad (6)$$

θ は θ_c, θ_s より以下のように求める。

$$\theta = 2\pi + \theta_s \quad \text{if } \theta_s < 0, \theta_c < \frac{\pi}{2} \quad (7)$$

$$\theta = 2\pi - \theta_c \quad \text{if } \theta_s < 0, \theta_c \geq \frac{\pi}{2} \quad (8)$$

$$\theta = \theta_c \quad \text{if } \theta_s > 0 \quad (9)$$

$R\theta$ 画像を使用して、求めた角度 θ と半径 R に対応する $R\theta$ 画像のデータを画像の点(x, y)に書き込み、中心窩画像を生成する。

3. 読書時の視点移動の実現

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ほぼ同じであり、視点は行の先頭から行の終わりまで行に沿って移動する。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つの処理を用いる。

- (i) 単語を見る処理
- (ii) 次の単語を探す処理
- (iii) 次の行を探す処理

これらの処理の流れを視点移動を用いて説明し、その後(i) (ii) (iii)のそれぞれの処理を注目領域を用いて説明する。まず、処理の流れを視点移動を用いて説明する(図1)。視点が行の先頭にある時、(i)と(ii)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(ii)により前方に空白が検出される。空白が検出されると、(i)と(ii)の繰り返しにより、視点は移動してきた行を逆に移動し、行の先頭に戻る。この時、視点が行の終りに向

かって移動する時に各視点で抽出された特徴を保存した短期記憶を使用する[9]。視点は、短期記憶上の特徴を見ながら行の先頭へ移動するため、実際に視点は動かない。行の先頭に近づくと(ii)により前方に空白が検出される。空白が検出されると(iii)により次の行が検出され、視点が次の行に移動する。本稿では、英文の読書時の各行における視点移動のみを実現し、それに注目する。行の終わりから次の行の先頭への視点移動はヒトの一般的な視点移動とは異なるが、これには注目せず上のアルゴリズムに従うことにする。これらの3つの処理では注目領域を使用する。注目領域は英文の読書時において次の視点を決定するための、中心窩画像上の注目している領域である。ヒトが本を読む時に、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。本稿ではこのような視点移動を防ぐために、次の視点を注目領域の中から決定する。注目領域の長さは20文字分とし高さは文字の高さとする。

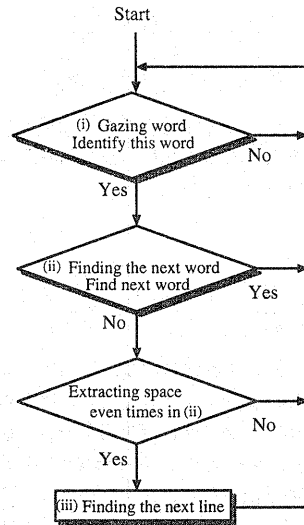


図1: 処理の流れ

次に、先に述べた(i) (ii) (iii)の処理を注目領域を用いてそれぞれ説明する。

(i)の処理は視点のある単語を注目領域で囲むことにより行う。注目領域内の単語が識別された場合は(ii)へ移り、それ以外の場合は注目領域内より視点を決定する。

まず、注目領域で単語を囲むために、現在の視点の中心窩画像より抽出されたエッジを用いて、現在の視点を中心に7文字分の長さをもつ領域を回転させ、領域内のエッジ密度を計算する。エッジ密度は領域内のピクセル数における抽出されたエッジのピクセル数の割合である。以前の注

目領域の移動方向とその逆の方向に重み付けを行い、それぞれの方向においてエッジ密度の高い方向を求める。次に、注目領域と同じ高さで文字の半分の長さをもつ領域をそれぞれのエッジ密度の高い方向へ移動させ、領域内のエッジ密度を計算する。エッジ密度の低い位置を単語の終りと判断する。エッジ密度の高い方向とその方向での単語の終りの位置より、単語を囲むように注目領域の角度と長さを調節する(図2)。単語間の空白を判断するエッジ密度のしきい値は次章で述べる。

文字の識別率は視点から離れるに従って減少し、その減少率は視点の左右で異なる。視点の右側における文字の識別率の減少率は左側のその半分であり、視点の右側の8文字目の識別率は0.1であると報告されている。本稿では、識別率を単位領域内のピクセル数に対する重み付けされたサンプリング数の割合とする。視点の左側のサンプリング数と視点の右側のサンプリング数は、視点の左側の8文字目の識別率が0.1であることを考慮し、それぞれ以下の式で重み付けを行う。

$$L_{weight} = \left(1 - \frac{r}{150}\right) \quad (10)$$

$$R_{weight} = \left(1 - \frac{1}{2} \cdot \frac{r}{150}\right) \quad (11)$$

r は、サンプリング点と中心窩の中心の距離である。注目領域内の単語の識別率 $Identify$ は、注目領域内のピクセル数 $pixel$ を用いて次式で表す。

$$Identify = \frac{1}{pixel} \left\{ \sum_{(x,y) \in L_{rgn}} L_{weight} + \sum_{(x,y) \in R_{rgn}} R_{weight} \right\} \quad (12)$$

L_{rgn} 、 R_{rgn} は、それぞれ注目領域内の視点の左側、右側の領域である。 $Identify$ が0.85以上であれば単語が識別され(ii)へ移る。0.85未満であれば単語が識別されず、注目領域内より次の視点を決定する。注目領域内の視点決定は、領域内のエッジの抽出点をサンプリングされた時の中心窩の中心からの距離に応じて式(13)に従い重み付けを行い、ランダムに視点を選ぶことにより行う。

$$W(i, j) = \left(1 - 1.01^{(Min_distance - Distance(i, j))}\right) \quad (13)$$

ここで、点 (i, j) はエッジが抽出されている点であり、 $W(i, j)$ 、 $Distance(i, j)$ 、 $Min_distance$ はそれぞれ、注目領域内の点 (i, j) の重み、注目領域内の点 (i, j) と中心窩の中心間の距離、 $Distance(i, j)$ の最小値である。

(ii) の処理は、(i) で計算されたエッジ密度の高い方向において、単語間の空白を探すことにより行う。文字と同じ高さで文字の半分の長さをもつ領域をエッジ密度の高い方向へ移動させながら領域内のエッジ密度を調べる。抽出された2つのエッジ密度の低い位置を次の単語の先頭と終

りと判断する。(i)と同様に次の単語を注目領域で囲みその領域内より視点を決定する(図3)。次の単語が検出されない場合は行の終わりとして判断し、(iii)へ移る。エッジ密度を調べる範囲の長さは12文字分とする。単語間の空白を判断するエッジ密度のしきい値は次章で述べる。

(iii) 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4)、エッジ密度の低い位置を行間の空白と判断する。エッジ密度を調べる領域の長さとは高さは5文字分の大きさとする。エッジ密度を調べる範囲は上下の2行分とし、行間の空白を判断するエッジ密度のしきい値は次章で述べる。

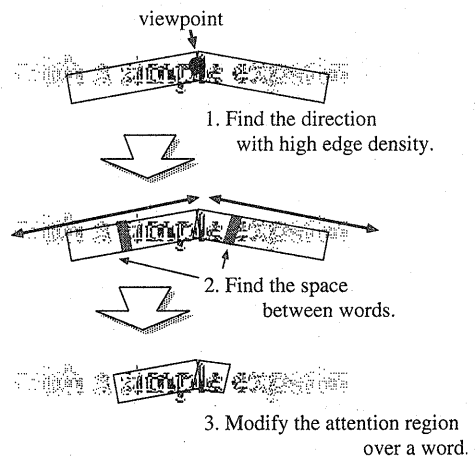


図2: 注目領域の変更

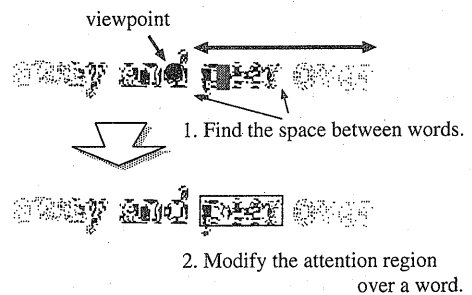


図3: 次の単語の検出

4. 中心窩の半径の単語識別時間への影響

実現した視点移動において、単語内の最初の視点が3~5文字目にある時に視点が次の単語へ移動するまでの時間が最小になる最適停留位置効果が現れることを確認する。次

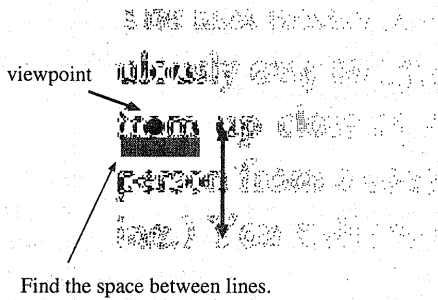


図 4: 行間の空白検出

に、6通りの中心窩の半径において視点が次の単語へ移動するまでの時間を計測し、その時間が最小となる中心窩の半径を調べる。

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) = 60 ピクセル、 C_m (中心窩の中心から受容野の中心までの離心率に対する受容野の直径の割合) = 0.33、 O_v (隣接する受容野の部分的重複率) = 0.95 とする。視野の半径は 300 ピクセルとする。また、前章で述べた視点を動かす処理、次の単語を探す処理、次の行を探す処理で用いられる空白を判断するエッジ密度のしきい値を経験的に次のように設定する。単語を見る処理と次の単語を探す処理においては、単語間の空白の有無を判断するエッジ密度を 0.95 % とし、次の行を探す処理においては、空白行の有無を判断するエッジ密度を 1 % とする。

最初に、3. で実現した視点移動を約 3000 単語を用いて行い、単語内での最初の視点において、視点が次の単語へ移動するまでの時間を計測した。図 5 はその結果である。横軸は単語の各文字位置を示し、縦軸は視点が次の単語へ移動するまでの時間を示す。視点の右側における文字の識別率の減少率は左側のそれより小さく、単語の識別率が最大になる位置は単語の中心から前よりの位置になる。そのため各単語の 2~4 文字において、視点が次の単語へ移動するまでの時間が最小となっている。図 6 は、報告されているヒトの最適停留位置を示した図である。横軸は単語の各文字位置を示し、縦軸は視点が次の単語へ移動するまでの時間を示す。図 5 と図 6 を比較すると、両方において単語内での最初の視点が単語の前よりにある時に、視点が次の単語へ移動するまでの時間が最小となり、実現した視点移動において最適停留位置効果が現れることを確認した。

次に、中心窩の半径 R_0 が 30、40、50、60、120、150 ピクセルの 6 通りの場合において、約 3000 単語を用いて視点が次の単語へ移動するまでの時間を計測した。図 7 はその結果である。横軸はそれぞれの文字数の単語を示し、縦軸はそれぞれの単語において視点が次の単語へ移動するまでの

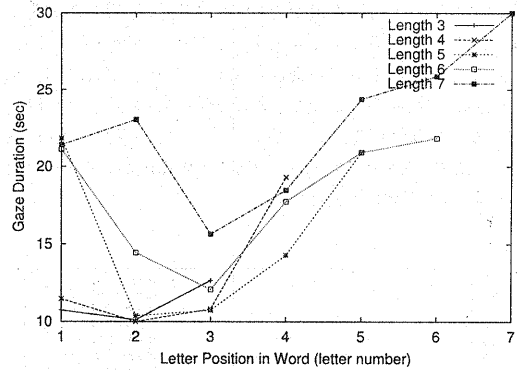


図 5: 最適停留位置

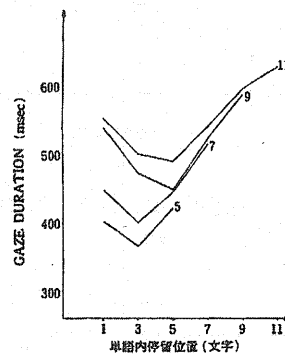


図 6: ヒトの最適停留位置 (参考文献 [4] より転載)

時間を示す。各単語において中心窩の半径が 60 ピクセルの時、視点が次の単語へ移動するまでの時間が最小になっている。

読書時の視点移動において 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 [5]。Ishihara らは、この特徴を取り入れ単語内での最初の視点が単語の中心にくる傾向が現れる視点移動を実現している [10]。本稿でも同様の特徴を取り入れ、単語内での最初の視点が単語の中心にくる傾向が現れる視点移動を実現した。

図 8 は、6 文字の単語において中心窩の半径を 30、40、50、60、120、150 ピクセルと変化させた時の、単語内の各文字が最初に視点となる割合を示している。中心窩の半径を 30、40、50、60、120、150 ピクセルと変更するに従い、単語内で最初に視点となる割合の高い位置が、単語の先頭から後ろに移動する。中心窩の半径が 60 ピクセルの時に、単語内で最初に視点となる割合の高い位置が単語の 2~4 文字目と

なり、最適停留位置と同じになるため、視点が次の単語へ移動するまでの時間が最小となる。実際に使用した文書画像では1文字が15ピクセルであり1文字を0.3cmとすると、中心窩の半径は4文字の長さと同じであり、1.2cmとなる。1.2cmの文字列が中心窩に含まれる時の文書と眼の距離は、ヒトの中心窩の角度が2°であることから、 $\frac{1.2\text{cm}}{\tan 2^\circ} \approx 34\text{cm}$ となる。文書と眼の距離が34cmの時、視点が次の単語へ移動するまでの時間が最小となる。

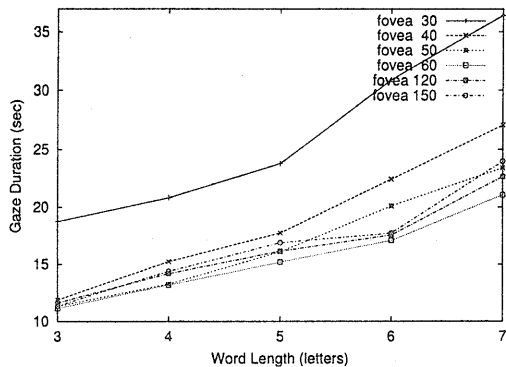


図7: それぞれの中心窩の半径において、各単語での視点が次の単語へ移動するまでの時間

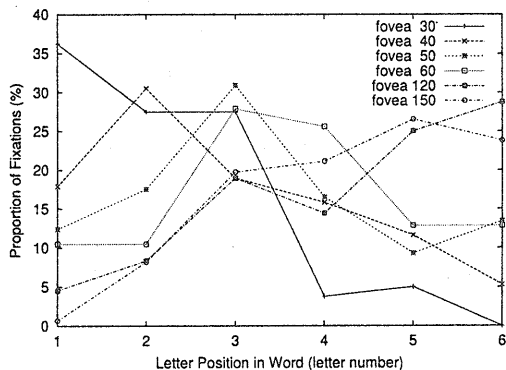


図8: 6文字の単語において、中心窩の半径を変更させた場合の単語上の各文字が最初に視点となる割合

実験において、3.で実現した視点移動に最適停留位置効果が現れることを確認した。次に、中心窩の半径が30、40、50、60、120、150ピクセルの6通りの場合において、視点が次の単語へ移動するまでの時間を計測した。中心窩の半径が60ピクセルの時に、視点が次の単語へ移動するまでの

時間が最小となることを示した。また、ヒトの中心窩の角度が2°であることから、1文字が0.3cmの英文において視点が次の単語へ移動するまでの時間が最小になる文書と眼の距離は34cmであることを示した。

5. むすび

ヒトの視点移動では文字が視野の中心から離れるにつれて識別率が減少する。視野の右側の減少率より左側の減少率を大きくすることにより、最適停留位置効果が現れる視点移動を実現した。次に、中心窩の半径を6通りに変化させ、視点が次の単語へ移動するまでの時間を計測した。中心窩の半径が60ピクセルの時に、視点が次の単語へ移動するまでの時間が最小となることを示した。また、ヒトの中心窩の角度が2°であることから、1文字が0.3cmの英文において視点が次の単語へ移動するまでの時間が最小になる文書と眼の距離は34cmであることを示した。

参考文献

- [1] R. A. Monty and J. W. Senders, "Eye Movements and Psychological Process," Lawrence Erlbaum Associates pub., pp. 371-395, 1976.
- [2] R. A. Monty and J. W. Senders, "Eye Movements and Psychological Process," Lawrence Erlbaum Associates pub., pp. 417-427, 1976.
- [3] O'Regan, J. K. and Lévy-Schoen, A., "Eye movement strategy and tactics in word recognition and reading. In : Attention and Performance XII: The Psychology of Reading," Lawrence Erlbaum Associates, pp. 363-383, 1987.
- [4] O'Regan, J. K., Lévy-Schoen, A., Pynte, J., and Bru-gallère, B., "Convenient fixation location within isolated words of different length and structure," Journal of Experimental Psychology : Human Perception and Performance, 10, pp. 250-257, 1984.
- [5] K. Rayner, "Eye Movements in Reading," Academic Press, pp. 7-20, 1983.
- [6] Nazir, T. A., O'Regan, J. k., and Jacobs, A. M. "On words and their letters," Bulletin of the Psychonomics Society, 29, pp. 171-174, 1991.
- [7] E. L. Schwartz, "Spatial mapping in the primate sensory projection: Analytic structure and relevance to perception," Biol. Cybernetics 25, pp. 181-194, 1977.
- [8] S. W. Wilson, "On the retina-cortical mapping," Int. J. Man-Machine Stud. 18, pp. 361-389, 1983.
- [9] 守田了, 石原由紀夫, "短期記憶とタスクモデルに基づく視点移動の実現," MIRU2000, I, pp. 206-214, 2000.
- [10] Yukio Ishihara, Satoru Morita, "Computation Model of Eye Movement in Reading Using Foveated Vision," BMCV2000, pp. 108-117, 2000.