

# 主観的輪郭線に基づく段組を含む文章の読書時における視点移動 のシミュレーション

守田 了 石原 由紀夫

山口大学工学部

宇部市常盤台 2557

morita@kde.yamaguchi-u.ac.jp

段組を含む文章を読む際には、段組の構成を把握しつつ、行を読み進めなければならない。ヒトは、段組の境界を輪郭として捉えるが、実際にはその境界は多数の文字から構成されており、境界の輪郭線は存在しない。このように実際には存在しない境界の輪郭は主観的輪郭線と呼ばれている。主観的輪郭線が成立するメカニズムは明るさの対比や、体制化や奥行き手がかりが要因であることが特定の条件下で述べられているのが現状である。読書時の凝視時に文字を認識できるヒトの有効視野領域は英単語で2-4語程度である。そのため、周辺視の低い解像度の特徴を用いて、行や段の境界の主観的輪郭線を抽出できることが必要である。本稿では、低次視覚により得られる特徴をぼけ関数により群化することにより、行や段の境界となる主観的輪郭線が抽出できることを明らかにする。実際に段組を含む論文を読み進める視点移動をシミュレートし、視覚モデルの有効性を示す。

キーワード 主観的輪郭線, 群化, 中心窩視覚, 視点移動, 読書

## Simulating Eye Movement in Reading the Document Including Column Structure Based on Subjective Contour

Satoru MORITA and Yukio ISHIHARA

Faculty of Engineering, Yamaguchi University

2557 Tokiwadai, Ube, 755, Japan

Human read the document along lines while understanding the column structure to read the document including the column structure. Though human regard the column boundary as the contour, the boundary contour does not exist. We call it the subjective contour. It is discussed that the brightness contrast, Gestalt system and the depth are the factor of subjective contour on the special condition. In this paper, the letters are grouped on the document and the subjective contours are detected by blurring the features derived from the low level vision. We introduce the grouping mechanism in the eye movement model in reading the document including the column structure. We actually show the effectiveness by simulating the eye movement in reading the paper.

keyword: Subjective contour, Grouping, Foveated vision, Eye movement, Reading

# 1 はじめに

ヒトの読書時の視点移動は視覚心理学の分野で盛んに研究されている [1]. ヒトの視覚は網膜の中心に解像度が高く, 周辺に解像度が低い. 網膜の中心の視力の 50% 以下の視力を与える領域を周辺視野と呼び, 50% 以上の視力を与える領域を中心視野と呼ぶ. 読書時の凝視時に文字が認識できる有効視野領域は 2-4 語程度である. 周辺視がテキスト上の次の凝視位置を決定するのに重要であることが指摘されている [2][3][4]. 単語の長さ, 形態, 初頭文字, 終了文字の周辺情報を用いて効率のよい読みを達成維持できることが述べられている [5]. 改行運動時の初頭文字へのスムーズな移動には, サッカード運動の距離が長いだけに, 周辺視による方向定位が重要な手がかりになることが述べられている. 本稿では, 段組を含む文章を読む際の初頭文字と, 次の行の初頭文字を見つけるために周辺視を利用することにより, 読書時の視点移動をシミュレートする.

読書時の視点移動は大きく, サッカード, 凝視, 逆行, 行かえの 4 つからなる. 次の行の初頭文字に視点を移動する際には短期記憶を用いる [6]. 注目領域の単語が一定の解像度に達していない場合には, 読み進める方向とは逆方向に視点を移動する [7]. このような視点移動モデルを用いることにより, ヒトの視点移動に現れる最適停留位置効果などの視覚特性が現れることが明らかになっている [8][9][10]. 本稿でもこれらの知見に基づき, 視覚モデルを構築する.

段組を含む文章を読む際には, 段組の構成を把握しながら, 行を読み進める. ヒトは段組の境界を輪郭として捉えるが, 実際にはその境界は多数の文字から構成されており, 境界の輪郭線は存在しない. このように実際には存在しない境界の輪郭は主観的輪郭線と呼ばれている. 主観的輪郭線の知覚が成立するメカニズムは明るさの対比や, 体制化 [11] や興行き手がかり [12] が特定の条件下で要因であることが述べられているのが現状である. 本研究では, 低次視覚により得られる特徴をばけ関数により群化することにより, 行と段を境界づける主観的輪郭線が抽出できることを明らかにする.

2. では主観的輪郭線を抽出する群化に基づく視覚モデルについて述べる. 3. では段組を含む文章の読書時における視点移動モデルを構築する. 4. では実際に段組を含む文章を読む視点移動をシミュレートし, 視覚モデルの有効性を述べる.

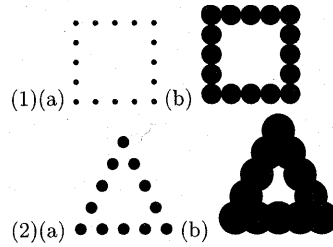


図 1: (a) 四角形の主観的輪郭線が現れる図形 (b) エッジをぼかすことにより現れる輪郭線 (c) 三角形の主観的輪郭線が現れる図形 (d) エッジをぼかすことにより現れる輪郭線

## 2 中心窩視覚と群化に基づく視覚モデル

### 2.1 中心窩視覚

ヒトは網膜の奥にある視細胞で光を感じることにより物体を見ている. 視細胞は杆体と錐体に分類される. 杆体は暗い場所で働き, 網膜の中心にある中心窩にはほとんど存在しない. 錐体は明るい場所で働き, 中心窩に多く存在する. 錐体の密度は中心窩の中心で  $150000$  個/ $mm^2$  であり, 周辺では  $10000$  個/ $mm^2$  に減少する [13]. このため, 明るい場所で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. 本論文では次のサンプリングモデルを用いて中心窩視覚を実現する. 解像度を単位面積あたりのサンプリング数の割合とし, 中心窩の中心から距離  $x$  の位置での解像度  $Reso(x)$  をシングモイド関数を用いて次式で表す.

$$Reso(x) = \frac{1 + \exp(-ab)}{1 + \exp(a(x-b))} (1.0 - 0.067) + 0.067 \quad (1)$$

定数  $a, b$  は, 解像度の減少率を決定するパラメータである. 解像度  $Reso(x)$  は中心窩の中心では 1.0 となる. 網膜上の錐体の密度は中心窩の中心での  $150000$  個/ $mm^2$  から周辺での  $10000$  個/ $mm^2$  へ減少するため,  $Reso(x)$  は周辺では  $10000/150000 \approx 0.067$  に減少する.

中心窩画像は原画像の各画素を解像度 (式 (1)) に応じた確率でサンプリングすることにより生成される. 解像度が 0.8 である画素がサンプリングされる確率は 80% である.

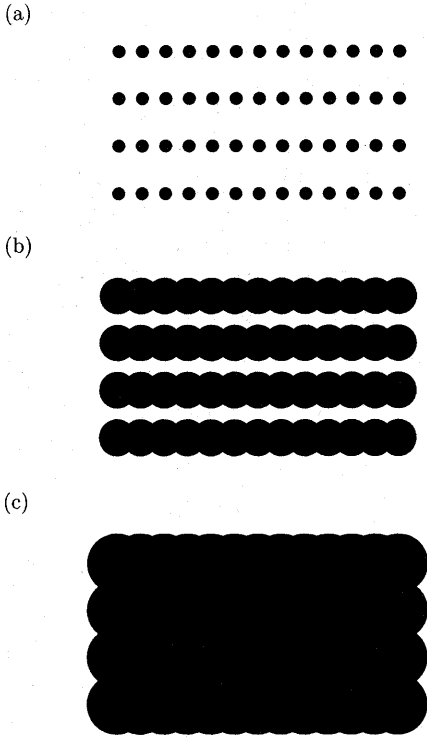


図 2: (a) 文章を読む際に現れる主観的輪郭線 (b) エッジをぼかすことにより現れる各行の輪郭線 (c) 文章の塊から得られる輪郭線

## 2.2 ぼけ関数と群化による主観的輪郭線の抽出

段組を含む論文を読む場合の視点移動をシミュレートするために、以下の2点に注意する。

- 行を認識できること。
- 段組を認識できること。

ヒトは点集合をより小さな集合にまとめて見る傾向にある。このように点集合をより小さな複数の部分集合に分ける操作を群化と呼ぶ。また、ヒトは実際には存在しない線、あるいはエッジを知覚することがある。これを主観的輪郭線と呼ぶ。図 1(1)(a),(2)(a) は主観的輪郭線が現れる図形の例である。それぞれの場合に四角と三角の実際には存在しない輪郭の知覚が成立する。また図 2(a) においては文章のように行と文章のかたまりの主観的輪郭の知覚が成立する。主観的輪郭線が現れる要因が数多く議論されている。

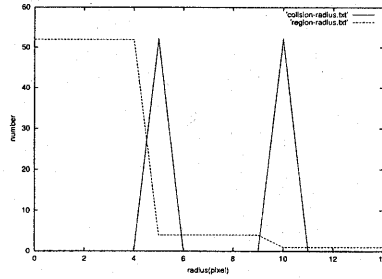


図 3: 各スケールにおける主観的輪郭線の本数のグラフとぼかされた特徴が他の領域と衝突するエッジ特徴の個数のグラフ

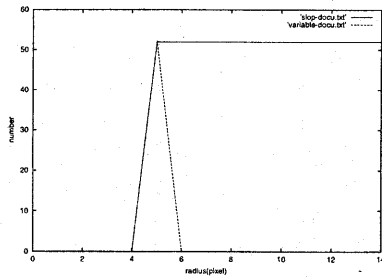


図 4: 各スケールにおける 1 回以上他の特徴と衝突するエッジ特徴の個数のグラフとその変化のグラフ

明るさの対比が要因で主観的輪郭線の知覚が成立したり、安定性、簡潔性を高めるように刺激布置を体制化することで主観的輪郭線の知覚が成立することが述べられている [11]。また奥行きが手がかりになって主観的輪郭線の知覚が成立することが述べられている [12]。主観的輪郭線の知覚が成立するメカニズムは特定の条件化で成立することが示されているのが現状である。

本研究では、低次視覚により得られる特徴をぼけ関数により群化することにより、実際には存在しない主観的輪郭線を抽出できることを明らかにする。

まず中心窩視覚モデルにぼけ関数を導入する。ぼけ関数は以下のように定義される。

$$F(x, y, \sigma) = \frac{1}{4\pi\sigma} \int_{-\infty}^{\infty} f(x, y) * g(x - \tau x, y - \tau y, \sigma) d\tau x d\tau y$$

但し  $g(x, y, \sigma)$  はガウス関数である。

主観的輪郭線はエッジや色などさまざまな特徴に対して現れる。ここでは読書を扱っているのでエッジ特徴に関する主観的輪郭線を抽出する。エッジ特徴は原画像を

$g(x, y)$  とすると,  $\frac{\partial^2 g(x, y)}{\partial x^2} = 0$  かつ  $\frac{\partial g(x, y)}{\partial x} \neq 0$  ならば  $f(x, y) = 1$ , それ以外を  $f(x, y) = 0$  とする. 網膜上の受容野から特徴画像  $f(x, y)$  が抽出される. 得られた特徴を画像平面でさまざまなスケール  $\sigma$  でぼかした画像  $F(x, y, \sigma)$  をしきい値に基づき 2 値化する. しきい値の計算は以下のように表される.

$$Fd(x, y, \sigma) = l[F(x, y, \sigma)]$$

但し,  $l[x]$  は  $x$  が  $k$  以上であれば  $l[x] = 1, k$  以下であれば  $l[x] = 0$  とするステップ関数である. 本稿では特徴画像に 1 と 0 の 2 値画像を用いているので,  $k = 0.5$  を用いる. 主観的輪郭線はエッジであるので,  $\frac{\partial^2 Fd(x, y, \sigma)}{\partial x^2} = 0$  かつ  $\frac{\partial Fd(x, y, \sigma)}{\partial x} \neq 0$  ならば  $Fe(x, y) = 1$ , それ以外を  $Fe(x, y, \sigma) = 0$  とする.  $Fe(x, y, \sigma) = 1$  となる輪郭線を  $\sigma$  における主観的輪郭線と呼ぶ. 本稿では 1 ピクセルの受容野から得られた特徴をぼかした場合に半径  $r$  ピクセルの円が生成される場合に, フィルタによるぼけの度合を表すスケールとして  $\sigma$  のかわりに  $r$  を用いる. 図 1(1)(a), (2)(a) はエッジをぼかすことにより, 四角と三角形の輪郭が現れる. スケールの異なるぼけ関数を用いることにより, 図 2(b) では各行の輪郭線が得られ, 図 2(c) では文章の塊の輪郭線が得られる. スケールの異なる主観的輪郭線集合から, 意味のある主観的輪郭線を抽出するために, スケールと主観的輪郭線の本数のグラフとぼかされた特徴が他の領域と衝突する個数のグラフを図 3 に示す. 図 3 から, 52 個あった領域は,  $r = 5$  において行のかたまりが 4 つでき,  $r = 10$  において段落のかたまりが 1 つできていることがわかる. 各特徴点の衝突回数は  $r = 5, r = 10$  の場合に衝突数が 52 であり, それ以外では 0 であることがわかる. 図 4 では, 各スケールにおいて 1 回以上他の特徴と衝突した特徴の個数のグラフとその変化のグラフを示している. このことからフィルタのスケールと主観的輪郭線の本数のグラフとフィルタの各スケールにおいてぼかされた特徴が他の領域と衝突する特徴の個数のグラフを求めることにより, 意味のある主観的輪郭線が抽出できることがわかる. 次に読書時の視点移動モデルにぼけフィルタによる主観的輪郭を抽出するしくみを組み込む.

### 3 読書時の視点移動情報処理モデル

読書時の視点移動はサッカード, 停留, 逆行, 行変えの 4 つからなる. 注目領域は中心窩画像上の注目している領域である. ヒトが文書を読む時に視点は行に沿って進み, 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. 注目領域の初期値として注目領域の長さは視野の

半径とし高さは文字の高さとする. 進行方向と逆方向において注目領域の初期値を 180 度回転し, 密度の高い方向を探し, 視点から進行方向のエッジ数  $V_1(\theta)$  と進行方向の逆方向のエッジ数  $V_2(\theta)$  を求める. 値  $V_1(\theta)$  が最大となる方向  $\theta_1$  と値  $V_2(\theta)$  が最大となる方向  $\theta_2$  を求め, 方向  $\theta_1$  と  $\theta_2$  に注目領域を回転させる. 次に文字と同じ高さで文字の半分の長さをもつ領域を方向  $\theta_1$  と  $\theta_2$  へ移動させ, 領域内のエッジの数を計算する. エッジの数が 0 である位置を単語の終りと判断し, 単語を囲むように注目領域の長さを変更する (図 5). エッジの数を調べる範囲は視野内とする. この間の解像度が 60% を越えていなければ, 注目領域から次の視点を決定し, 60% を越えていれば次の単語から次の視点を決定する [7]. これにより文を読み進める方向とは逆方向への視点移動がシミュレートされる. 文字を理解するのに必要な解像度である 60% の値が変化することにより, 読み速度が変化する. 注目領域において, 次に視点が移動する確率が解像度の逆数に応じて重みづけられ, 注目視野領域上の点から確率的に視点が決定される.

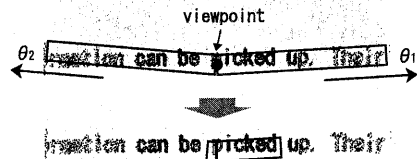


図 5: 注目領域の変更

短期記憶特徴点画像は次の視点移動の手がかりになる特徴点画像を, 視点位置で復元した中心窩特徴点画像を重ねて作成した画像である [15][6]. 行の終りに視点がある場合には, 視点は次の行の先頭に瞬時に移動する. 短期記憶に以前見た画像を蓄えておくことにより, 行の先頭に一気に移動できることが示されている [6]. ここで得られた短期記憶特徴点画像に対して, ぼかし処理を行うことにより, 行を形作る主観的輪郭線が抽出できる. この処理を視覚のモデルに導入することにより, 突然行の終りに視点が移動しても, 周辺視野を用いて主観的輪郭線を抽出し, 次の行の先頭を見つけることができる.

同様に文章の中心部分に視点を移動し, 受容野からの刺激をもとに, 特徴点の中心か画像を生成する. 複数回の視点移動から生成される短期記憶特徴点画像が生成される. この短期記憶特徴点画像にさまざまなスケールのフィルタをかけ, ぼけ関数のフィルタのスケールと領域数のグラフとエッジ特徴の衝突数のグラフを作成するこ

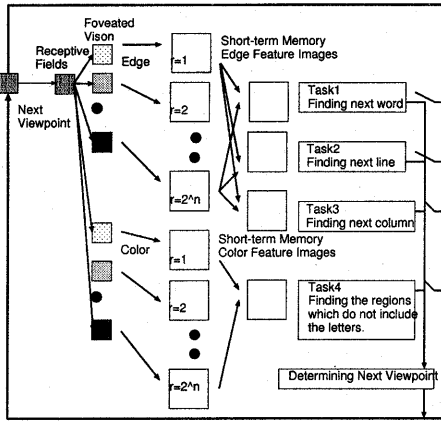


図 6: 段組を含む文章の読書時の視点移動モデル

とにより、段組の主観的輪郭を抽出する。段組をつねに把握しながら、文章を読み進め、文章が各段の最後まできたら、次の段の初頭文字に視点を移動する。段落は一般的規則から上の段落から、下の段落の左上に視点を移動する。最下部まできたら、右隣の最上部の段落の左上に視点を移動する。一番右下の段落を読み終るまで視点移動を続ける。

行の終りの空白を見つけた場合、次の行の先頭に視点を移動するタスクが実行される。段の終りの空白を見つけた場合、次の段の先頭に視点を移動するタスクが実行される。そうでない場合次の単語に視点を移動するタスクがおきる。並行して、色つきの図に視点を移動するタスクが実行される。図 6は提案する視点移動情報処理モデルを模式的に示している。

#### 4 段組を含む文章の読書時における視点移動のシミュレーション

実験では中心窩視覚の解像度の変化を決定するパラメータを  $a = 0.067$ ,  $b = 78.6$  とし、視野の半径を 1000 ピクセルとする。図 7は本実験で用いる中心窩視覚モデルの解像度の変化を示している。横軸は視点からの距離を文字単位で示す。縦軸は解像度を示す。図 8は実験に使用した論文 [14] である。図 9はフィルタのスケールに応じた領域の群化が示されている。図 9(a)(b)(c)(d)において用いたフィルタのスケールはそれぞれ  $r = 1, r = 2, r = 4, r = 8$  である。図 10では、エッジ特徴に各スケールでフィルタをかける。各スケールにおいて 1 回以上他

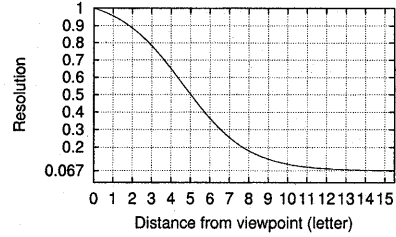


図 7: 中心窩視覚モデルの解像度の変化

の特徴と衝突するエッジ特徴の個数とその変化のグラフを示している。スケールと主観的輪郭線の本数のグラフとばかされた特徴が他の領域と衝突するエッジ個数のグラフを作成する。この結果から注目領域の高さを決定するばけ関数として  $r = 3$ , 段落を知るためのばけ関数として  $r = 7$  を選ぶ。  $r = 1$  では各エッジ点はそれぞれ別の領域になっているが、  $r = 4$  ではしだいに群化が進み、  $r = 8$  ではほぼ段組を構成する主観的輪郭線が生成されていることが分かる。図 11(a), (b), (c) はそれぞれ 657 視点目における中心窩視覚画像、657 視点目における中心窩視覚から得られたエッジ特徴点、657 視点目における短期記憶エッジ特徴イメージである。画像の大きさは  $846 * 1200$  ピクセルに対して、中心窩視覚の視野は 1000 ピクセルである。図 11(a) から視野の広さがわかり、図 11(b) から視野は広いにもかかわらずエッジ特徴から識別しうる単語は視野の中心の一行につき 2-4 単語程度であることがわかる。図 11(c) は次の行の初頭文字を見つけるための短期記憶特徴イメージである。この記憶イメージをたどることにより、視点は行の終りから次の行の先頭に瞬時に移動する。図 12(a),(b),(c) はそれぞれ論文読書時の 657 視点目における短期記憶イメージ、657 視点目における注目領域、657 視点目における視点移動である。図 12(b) より文字の大きさに応じた注目領域の高さが抽出できていることがわかる。図 12(c) より段落にしたがって論文が読み進められていることがわかる。図 6に示すように本モデルでは、文章の中に色つきの図を含んでいる場合、図を読むタスクと文章を読むタスクが並列して実行可能なため、図を読みながら文章を読む視点移動がシミュレートできる。

#### 5 おわりに

受容野から得られたエッジ特徴にばけ関数をかけることによって得られた群化された領域の主観的輪郭線

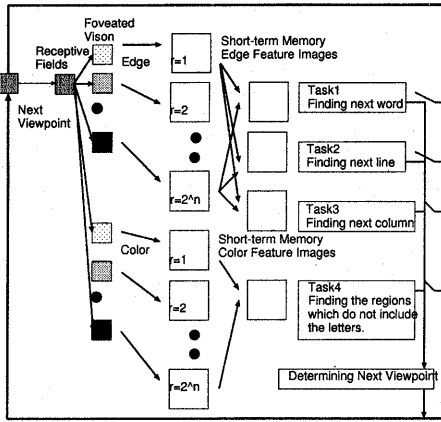


図 6: 段組を含む文章の読書時の視点移動モデル

とにより、段組の主観的輪郭を抽出する。段組をつねに把握しながら、文章を読み進め、文章が各段の最後まできたら、次の段の初頭文字に視点を移動する。段落は一般的規則から上の段落から、下の段落の左上に視点を移動する。最下部まできたら、右隣の最上部の段落の左上に視点を移動する。一番右下の段落を読み終るまで視点移動を続ける。

行の終りの空白を見つけた場合、次の行の先頭に視点を移動するタスクが実行される。段の終りの空白を見つけた場合、次の段の先頭に視点を移動するタスクが実行される。そうでない場合次の単語に視点を移動するタスクがおきる。並行して、色つきの図に視点を移動するタスクが実行される。図 6は提案する視点移動情報処理モデルを模式的に示している。

#### 4 段組を含む文章の読書時における視点移動のシミュレーション

実験では中心窩視覚の解像度の変化を決定するパラメータを  $a = 0.067$ ,  $b = 78.6$  とし、視野の半径を 1000 ピクセルとする。図 7は本実験で用いる中心窩視覚モデルの解像度の変化を示している。横軸は視点からの距離を文字単位で示す。縦軸は解像度を示す。図 8は実験に使用した論文 [14] である。図 9はフィルタのスケールに応じた領域の群化が示されている。図 9(a)(b)(c)(d)において用いたフィルタのスケールはそれぞれ  $r = 1, r = 2, r = 4, r = 8$  である。図 10では、エッジ特徴に各スケールでフィルタをかける。各スケールにおいて 1 回以上他

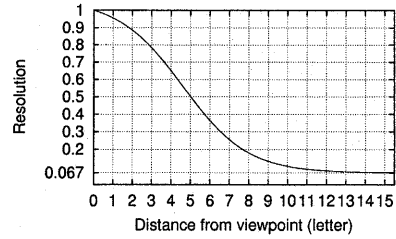


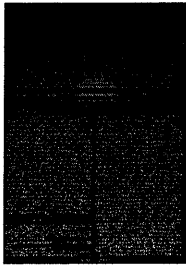
図 7: 中心窩視覚モデルの解像度の変化

の特徴と衝突するエッジ特徴の個数とその変化のグラフを示している。スケールと主観的輪郭線の本数のグラフとばかされた特徴が他の領域と衝突するエッジ個数のグラフを作成する。この結果から注目領域の高さを決定するばけ関数として  $r = 3$ , 段落を知るためのばけ関数として  $r = 7$  を選ぶ。  $r = 1$  では各エッジ点はそれぞれ別の領域になっているが、  $r = 4$  ではしだいに群化が進み、  $r = 8$  ではほぼ段組を構成する主観的輪郭線が生成されていることが分かる。図 11(a), (b), (c) はそれぞれ 657 視点目における中心窩視覚画像, 657 視点目における中心窩視覚から得られたエッジ特徴点, 657 視点目における短期記憶エッジ特徴イメージである。画像の大きさは  $846 * 1200$  ピクセルに対して、中心窩視覚の視野は 1000 ピクセルである。図 11(a) から視野の広さがわかり、図 11(b) から視野は広いにもかかわらずエッジ特徴から識別しうる単語は視野の中心の一行につき 2-4 単語程度であることがわかる。図 11(c) は次の行の初頭文字を見つけるための短期記憶特徴イメージである。この記憶イメージをたどることにより、視点は行の終りから次の行の先頭に瞬時に移動する。図 12(a),(b),(c) はそれぞれ論文読書時の 657 視点目における短期記憶イメージ, 657 視点目における注目領域, 657 視点目における視点移動である。図 12(b) より文字の大きさに応じた注目領域の高さが抽出できていることがわかる。図 12(c) より段落にしたがって論文が読み進められていることがわかる。図 6に示すように本モデルでは、文章の中に色つきの図を含んでいる場合、図を読むタスクと文章を読むタスクが並列して実行可能なため、図を読みながら文章を読む視点移動がシミュレートできる。

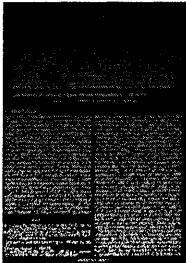
#### 5 おわりに

受容野から得られたエッジ特徴にばけ関数をかけることによって得られた群化された領域の主観的輪郭線を

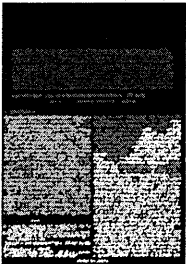
(a)



(b)



(c)



(d)

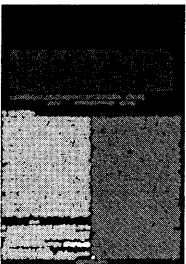


図 9: フィルタのスケールに応じた領域の群化 (a) フィルタの半径  $r = 1$  (b)  $r = 2$  (c)  $r = 4$  (d)  $r = 8$

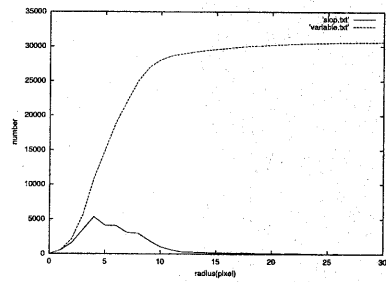


図 10: 各エッジ特徴に対してフィルタのスケールが小さい値から、初めて他の特徴と衝突するエッジ特徴の個数のグラフとその変化のグラフ

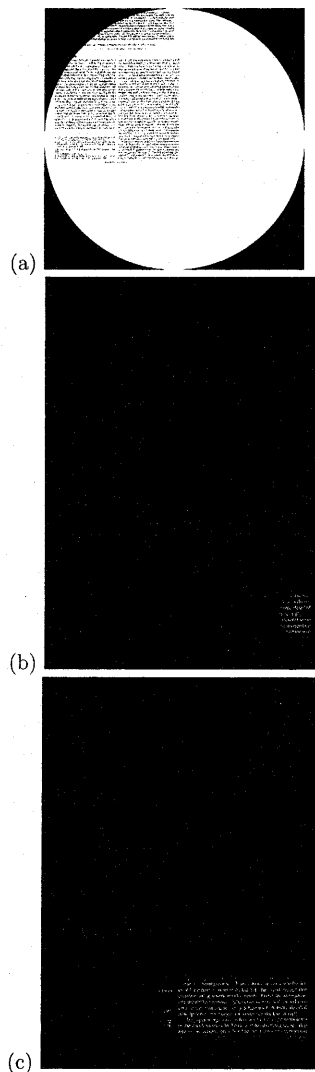


図 11: 論文読書時の視点移動 (a)657 視点目における中心か視覚画像 (b)657 視点目における中心か視覚から得られたエッジ特徴点 (c)657 視点目における短期記憶エッジ特徴イメージ

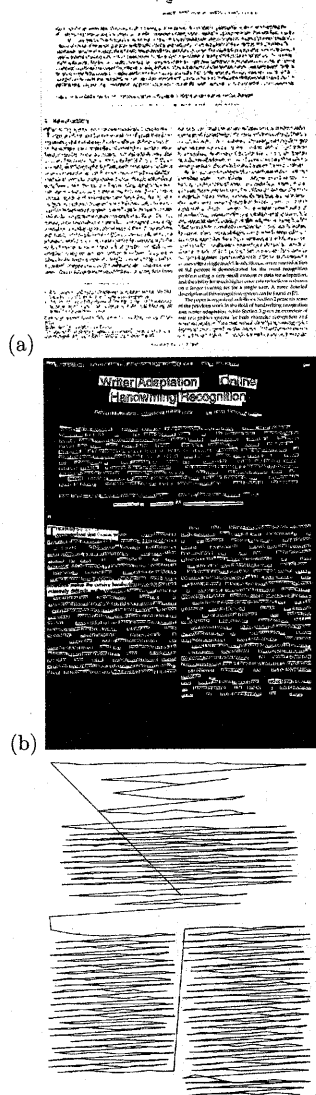


図 12: 論文読書時の視点移動 (a)657 視点目における短期記憶イメージ (b)657 視点目における注目領域 (c)657 視点目における視点移動