

提示コンテンツのデザイン構造を用いた視線運動の意味理解

石川恵理奈[†] 米谷 竜[†] 川嶋 宏彰[†] 平山 高嗣^{††} 松山 隆司[†]

[†] 京都大学大学院 情報学研究科 〒606-8501 京都市左京区吉田本町

^{††} 名古屋大学大学院 情報科学研究科 〒464-8603 名古屋市千種区不老町

E-mail: [†]{ishikawa,yonetani}@vision.kuee.kyoto-u.ac.jp, ^{††}{kawashima,tm}@i.kyoto-u.ac.jp,

^{†††}hirayama@cmc.ss.is.nagoya-u.ac.jp

あらまし 本研究では、商品カタログのようなコンテンツを閲覧する際の視線運動パターン理解を目的として、コンテンツ中の画像や文章といった要素オブジェクトの構造的・意図的レイアウト（デザイン構造）を知識源として積極的に用いる新たな枠組みを提案する。提案する枠組みでは、従来用いられてきた「どのオブジェクトを見ているか」という情報だけではなく、注視オブジェクト集合の持つ類似関係、空間的近接・包含関係といった意味的關係によって視線運動パターンを記述する。本稿では、ユーザが画面に提示された複数商品を閲覧吟味しながら好みのものを選択するという状況を取り上げ、提案する枠組みをユーザの内的状態推定に適用する。

キーワード 視線運動, ユーザ状態推定, 意味的關係性, デザイン構造

Semantic Interpretation of Eye Movements Using Designed Structures of Displayed contents

Erina ISHIKAWA[†], Ryo YONETANI[†], Hiroaki KAWASHIMA[†], Takatsugu HIRAYAMA^{††}, and

Takashi MATSUYAMA[†]

[†] Graduate School of Informatics, Kyoto University

Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-8501 Japan

^{††} Graduate School of Information Science, Nagoya University

Furo-cho, Chikusa-ku, Nagoya, 464-8603 Japan

E-mail: [†]{ishikawa,yonetani}@vision.kuee.kyoto-u.ac.jp, ^{††}{kawashima,tm}@i.kyoto-u.ac.jp,

^{†††}hirayama@cmc.ss.is.nagoya-u.ac.jp

Abstract We propose a novel framework to understand eye movements and user states, which actively utilizes semantic relations and spatial layouts of the objects constituting the displayed contents, i.e., the designed structure of contents. The main contribution of this work is to interpret gaze patterns using various semantic relations (e.g., similarity relation, spatial proximity relations), which differs from the previous work annotating eye movements by the objects being looked at. In this paper, we assume the situation that catalog contents are displayed on a monitor and a user is browsing it, and apply the proposed framework to the estimation of user states.

Key words eye movements, user state estimation, semantic relation, designed structure

1. はじめに

何をどのように見ているかという視線運動は、人間の内的な状態をよく反映した振舞いとして古くから研究されてきた [1]. 神経科学、認知心理学などの分野では、視線は人間の高次認知プロセスを解明するうえで重要な情報を持つとされており、検証実験を通して様々な知見が得られている [2]. また特に近年では、視線運動から人間の会話参加態度や迷いなどの内的な状態

を推定したり [3]~[5], 提示コンテンツのユーザビリティを評価する [6] といった工学的応用も盛んに行われている. これらの研究を通して、人間の内的な状態を知るために視線運動を利用することの有効性が示されている.

視線運動は人間の内的状態だけではなく、コンテンツを構成する画像や文章といった要素オブジェクトの意味的性質やレイアウト構造などにも影響を受けて振る舞うため、その解析にあたっては、視線の持つダイナミクスのみならず、コンテンツ側

の情報も考慮する必要がある。本研究では、コンテンツの**デザイン構造**、具体的にはコンテンツ製作者によってデザインされたオブジェクトの意図的なレイアウト構造をモデル化し、視線運動解析において積極的に利用する枠組みを提案する。この枠組みでは、従来扱われてきた「どのオブジェクトを注視しているか」といった情報だけではなく、「オブジェクトの織りなすデザイン構造において、どのような意味的側面が注目されたか」に基づいて視線運動の解釈を行う。

本稿では具体的な状況設定として、カタログコンテンツを用いたユーザの選択行動シーンを取り上げ、その際のユーザの状態(情報取得している、選択している、など)の識別問題を通して提案する枠組みの妥当性を評価する。次節では、まず本研究で提案する枠組みの概要について述べる。コンテンツの構造、及び視線運動の表現方法の具体的内容については3節にて説明する。4節では提案する枠組みに基づくユーザ状態の識別手法について述べ、手法の有効性評価のための実験、およびその結果と考察を示す。

2. デザイン構造を用いた視線運動の意味理解

ディスプレイに複数商品の情報を含むカタログコンテンツが提示されており、1人のユーザがそれらの商品の中から好みのもの1つを選択する状況を想定する。ユーザは常にディスプレイに視線を向けており、視線データはディスプレイ上の注視点系列として、解析に十分な精度で獲得されるものとする。

このような状況において、ユーザの視線運動の持つ意味を理解するためには、以下の点を検討する必要がある。

- (1) 提示コンテンツのどのような側面に注目し、それをどのようにモデル化するか
 - (2) 視線運動のダイナミクスをどのように表現するか
- この(1)、(2)についてそれぞれ2.1節、2.2節で説明し、これら2つの表現を統合することによって得られる、新たな視線運動解析の枠組みの概要を2.3節で述べる。

2.1 コンテンツのデザイン構造

視線運動解析における提示コンテンツの表現は、分析対象となる視線運動の分類に基づき、大きく2つの側面から捉えられる。視線運動は、動いているものに目が引きつけられやすいなどといった外的な要因によって発生する外因性視線運動(exogenous eye movement)、及び人間の内的状態や意図に起因して発生するような内因性視線運動(endogenous eye movement)とに分類できる。

外因性視線運動を解析する際に用いられる提示コンテンツの表現としては、視覚的注意の引きつけられやすさ(顕著度)を計算するモデルである顕著性マップ(saliency map)が広く知られている[7]。Saliency mapを用いることで、視野中に存在する多数のオブジェクトについて、色やコントラスト、方向などの物理的特性から顕著度をボトムアップに計算できる。一方で、本研究で目的とするような内因性視線運動の解析のためには、より高次の意味的な構造が重要となる。

本研究では、コンテンツの意味的な構造として、コンテンツのデザイン構造、つまりコンテンツ制作者によってデザインさ

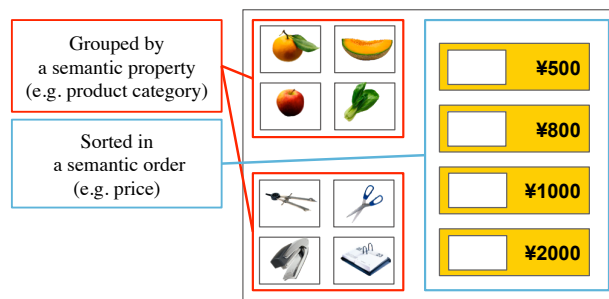


図1 コンテンツのデザイン構造の例

れたオブジェクトの意図的なレイアウト構造に着目する。例えば商品カタログなどのコンテンツにおいては、いくつかの商品がカテゴリごとに空間的に近接し配置されていたり、価格や製造年などの意味的な順序関係に応じて並べられていたりといったレイアウト構造が考えられる(図1)。このような構造は、コンテンツ製作者が商品についてどのような意味的特性(カテゴリや価格など)を強調したいか、という意図を反映したものであり、デザイン構造を扱ううえでは、オブジェクト間の意味的関係性(たとえば「同一カテゴリの商品である」「同一価格帯の商品である」)が重要となる。

そこで本研究では、このような意味的特性・関係性を陽に表現するモデルとして *content graph* を提案する。具体的には、各ノードにオブジェクトの意味的特性、エッジに意味的關係性をラベルとして持つような有向グラフを考える(詳細は3.1節)。デザイン構造を持つコンテンツにおいては、オブジェクトの空間的レイアウト(物理的側面)が *content graph*(意味的側面)を表現しているため、ディスプレイ上を移動する注視点、*content graph* 上の表現(たとえば部分グラフ)として解釈できることになる。すなわち、デザイン構造を視線解析にあたっての知識源として積極的に用いることで、視線運動が「どのような意味的側面が注目されているか」を指し示すものとして理解できるようになる。

なお、コンテンツの意味的情報のうち、いくつかの側面は先行研究でも取り扱われている。例えば文献[3]では、各オブジェクト領域に対し、説明を行う会話エージェント、説明対象となっている商品などといったラベルを与え、注視対象のラベル系列に対し N-gram 解析を行うことで、ユーザの会話参加態度を推定している。また文献[4]では、翻訳タスクにおいて用例の和文・英文の領域へラベルを与え、対応する和文・英文領域間を交互に注視する視線パターンを「比較パターン」と呼び、ユーザの迷い状態と深く関連すると結論づけている。いずれの文献においても、基本的なアプローチはコンテンツ領域をいくつかのオブジェクト領域に分割し、各領域に意味的ラベルを与えるというものであり、コンテンツ中の意味的特性・関係性を陽に表現する本研究とは異なる。

2.2 視線運動の時区間構造木表現

静的なコンテンツを閲覧している際の視線運動は、素早い注視対象の切り替わりである**サックード**と、同一箇所を注視し続ける**停留**との2種類の眼球運動から構成される。このような視

線運動の分析にあたっては、「どこに目が向けられているか」という空間的特性と「どのような順序で注視を行っているか」という時間的特性とが重要になる。従来研究において用いられている分析手法は、この空間的特性に着目したものと時間的特性に着目したものとに大別できる。

空間的な分析手法としては、提示コンテンツ上の注視点（視線が停留した点）の分布へガウス関数を畳み込むことでヒートマップを作成する方法[8]や、K-means 法や mean shift 法などを用いて注視点をクラスタリングする方法などが挙げられ[9]、コンテンツ上での人間の注目箇所を可視化することができる。このような空間的分析手法は、最も一般的に用いられる視線分析手法であり、広告や Web ページのユーザビリティ評価などにも利用されている。しかしながら、一定時間内に発生した注視点の空間的分布を扱っており、この時間範囲内における注視点間の時間的順序関係といった視線運動の時間的情報は考慮されていない。さらに、この分析の時間範囲をどのように設定するかといった問題もある。

一方で、時間的な分析手法としては、スキャンパスと呼ばれる注視点の移動履歴を線で結んだ直線の集合を扱う方法がある[10]。スキャンパスには個人差によるばらつきが存在するため、実際の解析にあたっては提示コンテンツを複数オブジェクト領域に分割しておき、注視領域の遷移パターンが用いられることが多い[3], [4], [11]。このような時系列分析では、視線運動の時間的情報は表現されているものの、領域内における視線運動がうまく扱われないなど、分析結果がオブジェクト領域に依存するという問題がある。

そこで本研究では、視線のダイナミクスを表現するにあたり、「各時刻でどの領域を注視しているか」というフレームベースな表現ではなく、「いつ注視対象が切り替わったか（いつサックードが発生したか）」というイベントベースな表現を用いる。この注視対象が切り替わるタイミングに基づいて視線運動を時区間系列に分解し、各時区間における視線の振る舞いを分析する。これにより、あらかじめ決められたオブジェクト領域によらない視線ダイナミクスの解析が可能となる。

このとき、注視の切り替えには「同一商品内でのメディア間の切り替え」「異なる商品カテゴリ間での切り替え」など、切り替えの際のサックードの大きさに応じていくつかの意味が存在し、視線の振る舞いの分析にあたって重要な役割を果たす。そこで本研究では、サックードをスケールスペースを用いて検出することで、視線ダイナミクスを階層的な時区間系列として表現する。これにより、高スケールの時区間系列＝全体的に大きく見渡す際の特徴、低スケールの時区間系列＝コンテンツの端から細かく情報を探索する際の特徴など、スケールに応じてさまざまな「コンテンツの見方」を分析することが可能となる。なお実際の解析においては、時区間の包含関係を木構造（以下では時区間構造木と呼ぶ）を用いて表現する。

2.3 提案する枠組みの概要

2.1 節、2.2 節でそれぞれ導入した content graph と視線運動の時区間構造木表現を統合し、視線運動という物理的計測データを、コンテンツの意味的特性に基づいた表現へと変換す

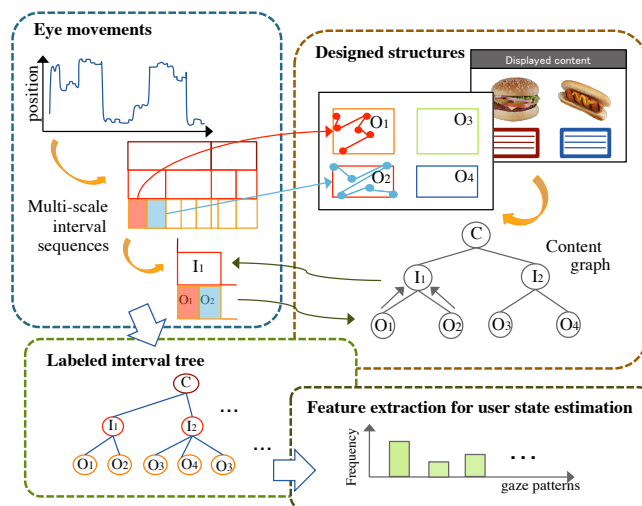


図2 提案する枠組みの概要

る(図2)。具体的には、時区間構造木の各時区間に対して、その時区間をもっとも特徴付けるような意味的特性のラベルを、content graph より獲得し与える。各々の時区間にそれぞれラベルが付与された時区間構造木を、ここでは前述の構造木と区別して「ラベル付き時区間構造木」と呼ぶ。ラベル付き時区間構造木では、各スケールにおける視線の振る舞いが、content graph におけるノードの時間的遷移として表現される。

このとき、出現するラベルの遷移パターンは、ユーザがどのような状態・目的でコンテンツを閲覧しているかに依存すると考えられる。例えば商品カタログ閲覧時の重要な視線パターンとして、ある対象 A と別の対象 B とを交互に注視するといった見返しパターンが挙げられるが[12]、この見返しパターンの持つ意味は、対象 A, B がどのような意味的関係を持つかによって異なる。見返しが同一商品を指す画像、文章領域で行われていれば、この視線パターンは商品を「吟味」としていると解釈できるが、異なる商品間で行われていれば、各々の商品を「比較」としていると捉えられる。ラベル付き時区間構造木はこのようなユーザの意図の違いを表現することが可能であり、木の構造を統計的に学習しておくことで、新たに観測された視線運動から意図を推定することも可能である。

3. 視線運動とデザイン構造の表現法

ここでは前節で導入した、content graph や時区間構造木の表現法、およびそれらの統合法について、本稿で用いる具体的方法を述べる。

3.1 Content graph

カタログコンテンツを構成する要素オブジェクトは、画像や文章、見出しなど様々なタイプのメディアが想定され、また、それらのオブジェクトが指し示す商品は、商品カテゴリや価格、製造元など多様な意味的特性を持っている。2.1 節で述べたように、content graph においては、オブジェクトの意味的特性や、それらの特性間の意味的関係がノードと有向エッジによって表現される。

本稿では商品の持つ様々な意味的特性のうち、商品カテゴリ

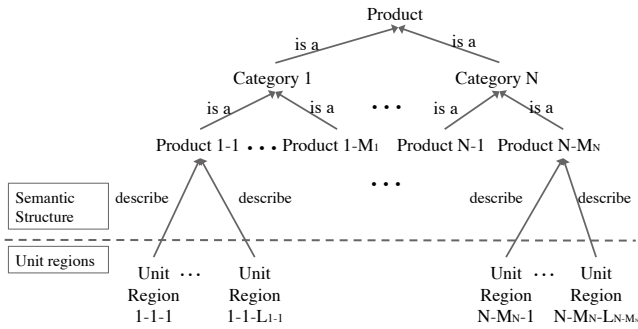


図3 content graph

に注目する。商品カテゴリは、カタログなどのコンテンツを扱う際には、重要な意味的特性の一つであると考えられる [13]。例えば、カタログに記載されている商品は、「りんご」は「果物」カテゴリに含まれるなどといったように、いくつかの商品カテゴリに分類できる。この時「果物」は「りんご」の意味的特性の一つであり、上位下位関係である *is-a* 関係を用いて「りんご *is-a* 果物」と表現できる。この関係は、content graph においては、「りんご」「果物」の各ノードを *is-a* というラベルが貼られた有向エッジで結ぶことによって表される。

また、提示コンテンツ内において分析の最小単位となるオブジェクト領域を、本稿では便宜上 unit region と呼び、各 unit region と、それが指し示す対象との関係を *describe* と呼ぶ。本稿では、この unit region は商品を説明する画像、文章の各領域に対応する。Content graph においては、各 unit region に対応するノードと、それが説明する商品に対応するノードとが *describe* というラベルが貼られた有向エッジによって結ばれる。このようにして、商品とそのカテゴリとの関係を content graph として表現すると図3のような木構造で表される。

3.2 多重解像度解析によるサッケードの検出

視線計測装置により画面上の注視点が $\mathbf{x}_t \in \Omega$ として得られているとする ($\Omega \subset R^2$: ディスプレイ空間) この注視点の系列 $X = [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_T]$ からサッケード時刻を求め、視線運動の時区間構造木表現を獲得する方法について述べる。

注視点系列 X が与えられたとき、時刻 $t = 1, \dots, T$ について、注視点の移動速度の大きさは $v_t = \|\mathbf{x}_t - \mathbf{x}_{t-1}\|$ と与えられ、系列 $V = [v_1, v_2, \dots, v_T]$ が得られる。この系列 V に対し、多重解像度解析の一つである scale space filtering を施すと、次式のようにスケール σ_j において平滑化された信号が得られる。

$$F_j(t, \sigma_j) = V * N(0, \sigma_j) \quad (1)$$

ここで “*” は畳み込み演算子であり、 $N(0, \sigma_j)$ は平均 0、分散 σ_j^2 のガウス分布を表す。スケール $\{\sigma_0, \sigma_1, \dots, \sigma_J\}$ ($i < j$ のとき $\sigma_i < \sigma_j$) に対して波形集合 $\{F_0, F_1, \dots, F_J\}$ が得られ、各スケール $F_j(t, \sigma_j)$ における変曲点が、 $T_j = \{t | \partial_t F_j = 0\}$ として検出される。また、波形集合 $\{F_0, F_1, \dots, F_J\}$ ($j < J$) において、平滑化スケールを小さくしながら変曲点を追跡し、変曲点 T_j を σ_0 まで追跡したものを $T_j = [t_0^{(j)}, t_1^{(j)}, \dots, t_{(N^{(j)})}^{(j)}]$ とする。これにより、分節点集合 $\mathcal{T} = T_0, \dots, T_J$ を得る。この分節点集合より、 $I_i^{(j)} = [t_{i-1}^{(j)}, t_i^{(j)}]$ として時区間系列

$I^{(j)} = [I_0^{(j)}, I_1^{(j)}, \dots, I_{N^{(j)}-1}^{(j)}]$ を得る。この時区間系列の集合を $\mathcal{I} = \{I^{(0)}, I^{(1)}, \dots, I^{(J)}\}$ とする。

時区間構造木の獲得にあたっては、この時区間に対応するノード集合を用意する。そして、連続する上下のスケール間において、包含関係にある時区間に対応するノードをエッジで結ぶことによって、時区間構造木を構築する。

3.3 Content graph と時区間構造木の統合

前節までの処理によって、コンテンツのデザイン構造を有向グラフで表現した content graph と、視線運動のダイナミクスを階層的な時区間構造として表現した時区間構造木が獲得されている。2.3 節で述べたように、視線運動の解析にあたっては、時区間構造木の各時区間においてユーザの注目する意味的特性を content graph より獲得することによって、両表現を統合した表現、ラベル付き時区間構造木を構築する。以下では構築にあたっての具体的手続きを述べる。

まず各時区間に含まれる注視点系列の空間的分布から、その時区間において注視されている unit region (注視 unit region) の集合を求める。ここでは、unit region の持つ空間的領域を $\{\Omega_1, \Omega_2, \dots, \Omega_M\}$ ($\Omega_i \subset \Omega$) とする。ラベルの付与にあたっては、ある時区間 I_i^j に含まれる注視点をこの unit region の各領域へ投票する。 $j = 0$ のとき、つまり時区間構造木においてリーフにあたる時区間においては、投票数が最大である unit region をラベルとして与える。

$j > 0$ のときにおいては、content graph 内で各注視 unit region から出る有向エッジをたどり、全ての注視 unit region から共通にたどりつくノードのうち、unit region ノードから最も近いノードのラベルをその時区間において注目されている意味的特性とする。例えば、図3のような content graph においては、注視 unit region が全て商品 I を示していれば、商品 I というラベルが与えられ、注視 unit region が商品 I_i と I_j を示しており、これらの商品が同一の商品カテゴリ C に含まれるならば、カテゴリ C というラベルが与えられる。

以上の手続きにより、ラベル付き時区間構造木の各ノードには、unit region もしくは content graph における高次の意味的特性のラベルが与えられており、視線の持つダイナミクスとコンテンツの意味的特性との両面を考慮した分析が可能となる。例えば、任意のスケールにおけるラベルの時間遷移に着目することによって各時区間において注目されている特性が時間とともにどのように変化するか、ということを調べることができる(時間軸方向の解析)。また、任意の時区間において各スケールにどのようなラベルが現れるかに基づいて、ある意味的特性に着目しているとき、内部でどのような視線運動が起こっているかを分析することができる(スケール軸方向の解析)。

4. 実験

4.1 ユーザ状態の推定

カタログコンテンツに対する選択行動中の人間の状態としては、まず「どのような商品があるか確認し、商品の情報を獲得している状態」、次に「選択候補間の比較・吟味を行うことで

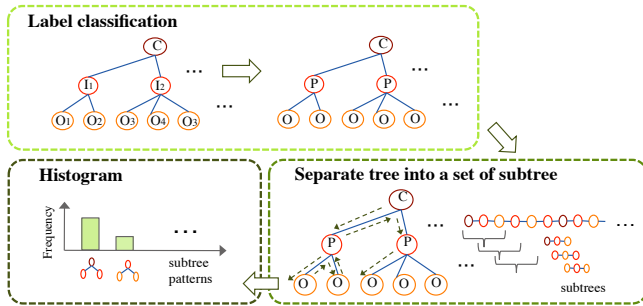


図 4 視線特徴の抽出

実際に選択する対象を決定している」状態、そして選択決定が完了した後、「自由にコンテンツを閲覧している状態」といった3状態を想定し、ある時点においてこのうち1つの状態をとると想定する。それぞれの状態を順に、「情報獲得中」、「選択決定中」、「自由閲覧中」とよび、本稿ではラベル付き時区間構造木の教師付き学習に基づき、観測される視線運動からこの3つの状態を識別することを目指す。

ラベル付き時区間構造木からの特徴抽出

統計的学習の導入にあたって、ラベル付き時区間構造木から視線特徴を抽出する(図4)。ここでは、ラベル付き時区間構造木より、親ノードを1つ、子ノードを2つ持つ構造の部分木を抽出し用いる。具体的には、部分木の親ノードのラベルを p 、子ノードのラベルを c_1, c_2 として、以下のように表されるラベルパターンの発現頻度 x を特徴とする。

$$x(L_i, L_j, L_k) = P(c_1 = L_i, p = L_j, c_2 = L_k) \quad (2)$$

本稿では、ラベル付き時区間構造木の各ノードのラベル L として、カテゴリに注目している、商品に注目している、商品内のメディアに注目している、という3種類を考える(図4(上))。

4.2 実験設定

本研究で用いるコンテンツ提示、及び計測環境を示す。コンテンツ提示には25.5インチディスプレイ^(注1)を用い、視線計測には角膜反射法に基づくアイトラッカー^(注2)を用いた。被験者はディスプレイ前に着席しており、ディスプレイからの距離は1000mmである。

前述したように本稿では、人間の内的状態として**情報獲得中**、**選択決定中**、**自由閲覧中**という3種類の状態を想定し、視線運動からこれらを識別することでモデルの妥当性を評価する。これらの各状態における視線データを獲得するため、被験者に対し以下のような教示を与えた。ここでは、各教示後に計測される視線データが上述の3状態に対応すると想定しており、統計的学習における教師信号としてはこの3教示が用いられることになる。

教示1 (30sec): ディスプレイに提示されるカタログを閲覧し、どのような商品が含まれるか確認する。

教示2 (時間無制限): カタログに掲載されている商品の中から親しい知人に対し、プレゼントを選ぶ。また、選択が決定すると同時に手元のボタンを押下する。

教示3 (60sec): 自由にカタログを閲覧する。

商品を選ぶ対象となる「親しい知人」は、事前に被験者に作成させた一覧の中から、実験者が試行ごとにランダムに指定した。また、教示1の段階であらかじめ商品を選択決定してしまうことを防ぐため、親しい知人の指定は教示2を与える際に行った。

4.3 結果、考察

8名の被験者の協力を受け、実験を行った。提示するコンテンツとしては、4つのカテゴリにそれぞれ4つずつ、計16個の商品を含む商品カタログを用意した。各被験者について、上述の3つの教示を1セットとして、コンテンツを変えながら10セットを行った。獲得された視線データの例を図5に示す。

獲得された計240試行の視線データ(8名×10コンテンツ×3状態)について、提案手法を適用してユーザ状態の識別を行った。ここではleave-one-out法を用いて、識別精度を評価する。提案手法による視線パターンの表現能力に関する基本的特性を調べるために、識別器としては線形識別器を用いた。本稿では、ベースライン手法として、まずコンテンツ上の空間的オブジェクト領域にラベルを付与し、注視対象のラベル系列に対するbi-gram解析を行った。このオブジェクト領域としては、同一カテゴリの領域、同一商品領域の2通りのスケールによる領域設定を行い、それぞれ個別に解析を行った。さらに、別のベースライン手法として、コンテンツ情報を用いず視線運動のダイナミクスのみを扱う解析も行った。具体的には、視線データの速さ成分を求め、与えられた2段階の閾値について、閾値以上の速さで視線が移動している時間の全体に対する割合を特徴量とし、k近傍法によって識別した。表1に識別精度を示す。

表1 3クラス識別精度[%]

提案手法	ベースライン手法		
	bi-gram(カテゴリ)	bi-gram(商品)	サッケード速度
59.7	47.1	42.9	48.8

本稿で提案したユーザ状態の推定手法では、特徴量として時区間木構造の部分木を用いた(式(2))。これは、異なる商品カテゴリ間、同一カテゴリ内で異なる商品間、同一商品内それぞれでの注視対象の切り替わりといった、複数スケールにおける視線運動を合わせて扱っていることに相当する。単一スケールにおけるbi-gram解析に比べ提案手法が高い識別精度を得たことから、選択行動時のユーザ状態を推定するにあたって、注視対象の切り替わりを階層的に捉えることが重要な意味を持つということが示された。またサッケードの速度をもとにした手法に比較しても高い識別精度が確認されたことから、視線運動のダイナミクスのみを扱うのではなく、コンテンツ情報を利用することで、視線運動の持つ意味をより詳細に表現できるといえる。

さらに、提案手法における、各々の状態間での識別しやすさを調べるため、各状態のペア間について2クラス識別を行った。

(注1) : MITSUBISHI Diamondcrysta RDT262WH, 25.5inch, W550.1mm/H343.8mm.

(注2) : Tobii 製視線計測システム X60EyeTracker, 頭部の移動可能範囲は幅400mm, 高さ220mm, 奥行き300mmである。

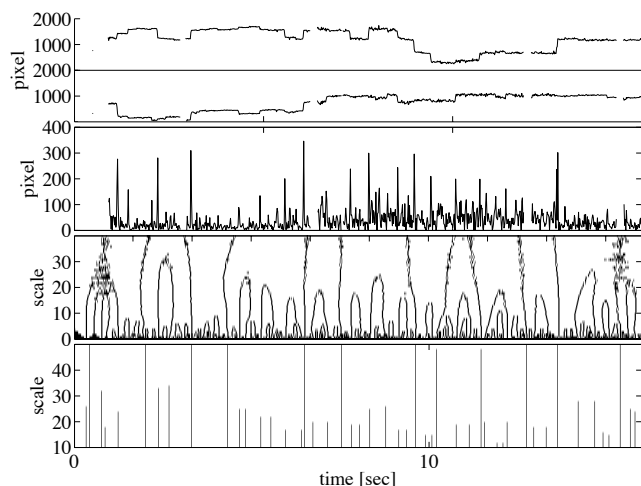


図 5 獲得された視線データの例。上から順に、x 軸方向の注視点座標、y 軸方向の注視点座標、注視点移動の速さ、平滑化波形集合から得られた零交差輪郭 (各スケールにおける変曲点)、変曲点追跡によって得られた分節点。

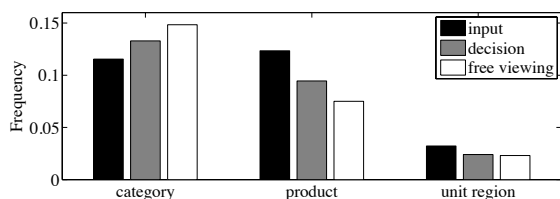


図 6 視線パターン出現頻度 (平均値)。左から順に、**カテゴリ**、**商品**、**商品内のメディア**のラベルが子ノードに与えられた部分木の出現頻度を示す (黒:情報獲得中、灰:選択決定中、白:自由閲覧中)。

得られた識別精度を表 2 に示す。2 クラス識別においては、**情報獲得中**と**自由閲覧中**とが他状態間に比べ高い精度で識別できている一方で、**情報獲得中**と**選択決定中**においては、比較的低い識別精度となっていることがわかる。**選択決定中**においては、商品を比較・吟味しながら商品情報を再取得するなどといった行動が取られる可能性がある。さらに、各状態において視線特徴がどのような傾向を示すか調べるために、それぞれの状態において得られる視線パターンの出現頻度の平均を求めた (図 6)。この結果についても 2 クラス識別の結果と同様に、**情報獲得中**と**自由閲覧中**の視線運動がより異なった傾向を示すことがわかる。また**情報獲得中**の方がより小さなスケールでの注視切り替わりが行われる傾向にあることから、**情報獲得中**にはよりコンテンツを細かく探索するといった注視戦略がとられることが示唆される。

表 2 2 クラス識別精度

状態	識別精度 [%]
情報獲得中 vs 選択決定中	66.2
選択決定中 vs 自由閲覧中	73.6
自由閲覧中 vs 情報獲得中	77.8

5. む す び

視線運動の意味理解のために、提示コンテンツを構成する要素オブジェクトの意図的なレイアウト構造であるデザイン構造を利用する枠組みを提案した。この枠組みにおいては、従来の「どのオブジェクトを見ているか」といった情報だけではなく、オブジェクトの意味的特性や特性間の関係性を用いて視線パターンを解釈する。本稿では、提案する枠組みを選択行動時のユーザ状態の推定へ適用し、実験を通して有効性を確認した。

本稿では特に商品カテゴリの上位下位関係に注目したが、今後はこれを一般化し、価格や製造年などによる順序関係を含む多様なオブジェクト間の意味的关系を考慮したモデルへの拡張を検討する。また提案する視線運動解析手法の応用先として、コンテンツのデザイン構造と視線運動から推定されたユーザの注目する意味的特性との対応をもとに、動的にレイアウトを生成するようなインタラクティブな情報提示システムへの応用に取り組む。

文 献

- [1] Alfred L. Yarbus. Eye movements and vision. *Plenum*, 1967.
- [2] 大野健彦. 視線から何がわかるか-視線測定に基づく高次認知処理の解明. 認知科学, Vol. 9, No. 4, pp. 565–576, 2002.
- [3] Yukiko Nakano and Ryo Ishii. Estimating user's engagement from eye-gaze behaviors in human-agent conversations. In *IUI*, pp. 139–148, 2010.
- [4] 高木啓伸. 視線の移動パターンに基づくユーザの迷いの検出 : 効果的な作業支援を目指して. *IPSJ Journal*, Vol. 41, No. 5, pp. 1317–1327, 2000.
- [5] Ryo Yonetani, Hiroaki Kawashima, and Takashi Matsuyama. Multi-mode saliency dynamics model for analyzing gaze and attention. In *ETRA*, pp. 115–122, 2012.
- [6] Laura Cowen. An eye movement analysis of web-page usability. Master's thesis, Lancaster University, 2001.
- [7] Laurent Itti, Christof Koch, and Ernst Niebur. A model of saliency-based visual attention for rapid scene analysis. *TPAMI*, Vol. 20, No. 11, pp. 1254–1259, 1998.
- [8] O Špakov. Visualization of Eye Gaze Data using Heat Maps. *Electrical Engineering*, Vol. 2, No. 2, 2007.
- [9] Anthony Santella and Doug DeCarlo. Robust clustering of eye movement recordings for quantification of visual interest. In *ETRA*, pp. 27–34, 2004.
- [10] David Noton and Lawrence Stark. Scanpaths in Eye Movements during Pattern Perception. *Science*, Vol. 171, No. 3968, pp. 308–311, 1971.
- [11] 大野健彦, 中谷桃子, 中根愛. マニュアルデザインにおける視線パターンと印象の関係. HCG シンポジウム予稿集, 第 787 巻, pp. 37–44, 2011.
- [12] J. Edward Russo and Larry D. Rosen. An eye fixation analysis of multialternative choice. *Memory & Cognition*, Vol. 3, No. 3, pp. 267–276, 1975.
- [13] Cai-Nicolas Ziegler, Georg Lausen, and Lars Schmidt-Thieme. Taxonomy-driven Computation of Product Recommendations. In *CIKM*, pp. 406–415, 2004.