

コンテンツ閲覧時における確率的注視行動モデル

下西 慶[†]川嶋 宏彰[†]松山 隆司[†][†]京都大学 大学院情報学研究科

1. はじめに

カタログコンテンツを閲覧し、比較吟味行動を行っているようなユーザに対して、閲覧行動中の視線運動からそのユーザの状態をダイナミックに推定することが可能になれば、そのユーザの状態に応じて戦略を変更しながらインタラクションを行い、ユーザの目的に適応した対象を推薦することが可能となる。視線運動からユーザの興味を探る研究はいくつか行われてきているが [1, 2], カタログコンテンツを閲覧するような場合の視線運動を扱うにあたっては、そのユーザの興味のみを考えることは十分ではない。たとえばコンテンツの中央に配置されている対象が閲覧されやすいといったように、その閲覧対象の配置に基づく外的な要因による影響を受ける [3, 4] ため、視線運動に与える外的な要因をも考慮して興味を推定する必要がある [5]。さらに、この配置に基づく影響は閲覧行動に対して、たとえばユーザの興味対象がコンテンツ内に無い場合に非常に強く現れたり、少数の興味対象の比較吟味のための閲覧行動においてはほとんど現れないというようにユーザの状況に応じて変化することが考えられる。そこで、ユーザの閲覧行動を複数の時区間に分割し、それぞれを解析することによって、内的、外的要因が与える影響の変化をとらえることを目指す。

2. 視線の確率的生成モデル

2.1 モデルの定式化

提案モデルにおいては、以下の対象を考慮する。

アイテム領域 コンテンツは N_{all} 個のアイテム集合 $\{I_1, \dots, I_{N_{all}}\}$ のうち N 個のアイテム提示領域によって重複なく $\{R_1, \dots, R_N\}$ に分割されるとする。領域 R_n に提示されているアイテムを $I(R_n)$ と表す。

属性値集合 アイテム集合に対して共通する複数の属性を考え、それらに対してとりうる属性値を全 Q 種の値 $\{V_1, \dots, V_Q\}$ と表す。

アスペクト K 個のアスペクト $\{Z_1, \dots, Z_k\}$ が考えられるとする。各アスペクトは属性値との関連度 $P(V_q|Z_k)$ によって特徴づけられる。

ここで、各ユーザの一連の注視行動をセッションと呼び、このセッション中においては興味アスペクト $\theta(s) = \{P(Z_k|s)\}$ は一定であるとする。図 1 に示す提案モデルにおける視線の生成過程は以下の通りである。各時刻において $\theta(s)$ に基づいてあるアスペクト z_t を決定する。次に条件付き分布 $P(V_q|Z_k)$ に基づいてある属性値 v_t が注目される。さらに、ユーザが属性値 V_q に着目している際に、あるアイテムが知識ベースを基にした $P(I_m|V_q)$ により選択、閲覧されるとする。この過程が各時刻独立に行われるとすると、セッション s の時刻 t において、アイテム I_m が閲覧される確率は、

$$P(i_t = I_m|s) = \sum_{q=1}^Q \sum_{k=1}^K P(i_t = I_m|V_q)P(V_q|Z_k)P(Z_k|s)$$

として与えられる。また、セッション s の長さを T_s 、時刻 t に閲覧されたアイテムを $I_{m_t}^s$ と表記する。このモデルを用いることにより、全セッションにおける注視アイテム系列集合 \mathcal{D} に対する尤度

$$P(\mathcal{D}) = \prod_{s=1}^S \prod_{t=1}^{T_s} \sum_{q=1}^Q \sum_{k=1}^K P(I_{m_t}^s|V_q)P(V_q|Z_k)P(Z_k|s)$$

を最大化するような最適化問題としてアスペクト $\{P(V_q|Z_k)\}$ や興味アスペクト $\{\theta(s)\}$ の学習が可能となる。

2.2 空間的な配置の影響の導入

本研究においては、モデルに空間的な配置による影響を取り入れることを目指す。そこで、アイテム自身ではなく、その提示領域が閲覧されるというモデルを考え、視線情報を注視アイテム系列そのものではなく、注視領域系列であると捉える。すなわち、アイテム $I(R_n)$ が閲覧された場合に、領域 R_n が閲覧されたものとする。

そのうえで、ユーザは自らの興味に基づいてあるアイテムを閲覧する状態と、コンテンツ内の空間的な配置に基づいてある領域を閲覧する状態の 2 つの潜在状態を持っていることを仮定する。ここで、コンテンツ内で提示アイテムに関わらず領域 R_n が閲覧される確率が $P(R_n; \beta)$ としてパラメタ β によって与えられるとしたとき、実際にユーザが属性値 V_q に着目している際に領域 R_n が閲覧される確率を

$$P(R_n|V_q) := (1 - \alpha)P(I(R_n)|V_q) + \alpha P(R_n; \beta) \quad (1)$$

として、2 状態に対する混合比 α を用いて表す。この α は、視線運動に対して提示アイテムの空間的な配置が与える影響の度合いを示している。この拡張により、セッション s の時刻 t において、領域 R_n が閲覧される確率は、

$$P(r_t = R_n|s) = \sum_{q=1}^Q \sum_{k=1}^K P(r_t = R_n|V_q)P(V_q|Z_k)P(Z_k|s)$$

として与えられることになる。

3. 実験

提案モデルによる各セッション中のユーザの閲覧行動の解析を被験者実験を通して行った。この実験においては、図 2 のようにユーザに対して複数の料理画像を提示し、その中から教示に適した料理を選択してもらい、その際の視線を 1 分間計測した。用いた画像組は 5 組、教示は 3 種である。ここで得られた視線データ集合を用いて、ユーザの閲覧行動における空間的な配置が与える影響の時間的な変化 ($\alpha_{s,t}$) の解析を行った。

$\alpha_{s,t}$ の解析手法：コンテンツ内に提示されているアイテムの中で、ある属性値 V_q を持つアイテム数を N_q とすると、ユーザがその属性値に着目している際に、その属性値を持つアイテムを閲覧する確率を $P(I_m|V_q) = 1/N_q$

Probabilistic Gaze Behavior Modeling in Visual Content Browsing K. Shimonishi, H. Kawashima, T. Matsuyama (Kyoto Univ.)

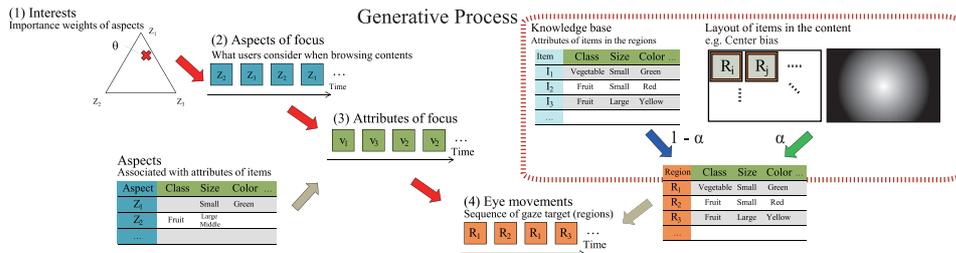


図 1: 提案モデルの概要

図 2: 実験環境

として等確率で与えた。また、全視線データを用いて、各領域の閲覧頻度分布を計算し、これを領域の閲覧パラメタ β とした。次に、式 (1) の α を一定と仮定し、 $\alpha = 0.5$ としてアスペクト、および各セッションの興味アスペクトを学習した。そして、この学習されたアスペクトおよび興味アスペクトを用いて、窓幅を $T_s/3$ 、ステップ幅を $T_s/12$ として固定幅窓を移動させながら各時区間における最適な $\alpha_{s,t}$ を計算した。この $\alpha_{s,t}$ はそのセッションの時区間における注視領域系列に対する尤度を最大化することで求めた。

図 3 に、推定された $\alpha_{s,t}$ を示す。この $\alpha_{s,t}$ は、各時区間における興味によるアイテムの閲覧と外的な要因による領域の閲覧の混合比を示している。すなわち、 $\alpha_{s,t}$ が高い時区間はそれだけ強く外的な要因による影響を受けていることを示している。また、これらのうち赤枠、青枠、緑枠の 3 つのセッションについて、図 4 に各時区間における各領域の閲覧頻度分布 (中段)、および領域の閲覧確率 β (上段)、各領域に提示されているアイテムに対して予測される興味 $\{P(I(R_n)|s)\}$ (下段) を示す。すなわち、 $\alpha_{s,t}$ が 1 に近いほど上段の分布に、0 に近いほど下段の確率分布に近い分布となっている。

図 3 赤枠のセッションおよび青枠のセッションは、ユーザの閲覧行動が興味による閲覧行動へと徐々に変化していき、ユーザがよく閲覧している領域の数が時間とともに減少していることが分かる。これは、ユーザが徐々に自らの興味に基づいて少数アイテムを見比べるような閲覧行動を取っているためであると考えられる。この一方で、図 3 緑枠のセッションにおいては、推定された $\alpha_{s,t}$ の値が時間を追うごとに上昇しており、さらにセッションの後半においてはどれか特定の対象を閲覧するということがあまりなくなっている。これは、セッションの前半で選択対象が決定してしまい、後半部においては領域による影響の方が大きくなってしまっているのではないかと考えられる。このように、セッションにおいて領域の与える時間変化のパターンも異なることが分かり、影響を一定として与える、または特定の傾向をあらかじめ与えておく、ということは困難であることが分かる。

4. まとめ

本稿では、ユーザのカタログコンテンツに対する注視行動のモデルを用いて、ユーザの閲覧行動におけるコンテンツの空間的な構造が視線に与える影響の時間変化についての検討を行った。その結果、各セッションにおいて空間的な配置の影響は一定ではなく、それぞれのセッションにおいて、影響の度合いは時間によって変化することが確認された。今後は外的・内的要因のバランスの時間的な変化も同時に推定することにより、たとえば既に選択が完了している、興味が定まっているような状況

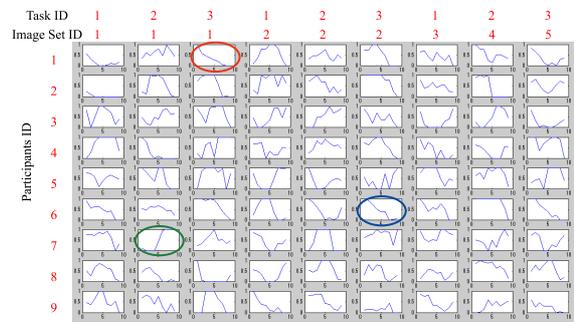


図 3: 各セッションの各時区間において推定された $\alpha_{s,t}$

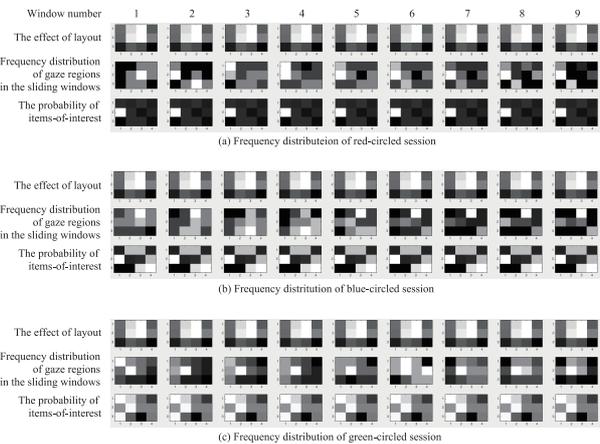


図 4: 3 つ (上から図 3 赤枠、青枠、緑枠) のセッションの各時区間における閲覧領域頻度分布。上段が学習された領域の影響、下段は推定される興味アイテムの分布。

であれば、その興味に基づいて新たな対象を提示してやる、といったようによりインタラクティブな推薦を可能にすることを旨とする。

謝辞

本研究の一部は、科学研究費補助金 26280075 および SCAT 助成の補助を受けて行った。

参考文献

- [1] 平山 高嗣, 角 康之, 河原 達也, 松山 隆司. 情報コンシェルジュ: Mind probing に基づくマルチモーダルインタラクションシステム. 信学技法 HCS, volume 111, number 190, pages 55–60, 2011.
- [2] Boris Brandherm, Helmut Prendinger, and Mitsuru Ishizuka. Interest estimation based on dynamic bayesian networks for visual attentive presentation agents. In *ICMI*, pages 346–349. ACM, 2007.
- [3] Ali Borji. Boosting bottom-up and top-down visual features for saliency estimation. In *CVPR*, pages 438–445, 2012.
- [4] Tilke Judd, Krista Ehinger, Frédo Durand, and Antonio Torralba. Learning to predict where humans look. In *ICCV*, pages 438–445, 2009.
- [5] 下西慶, 石川恵理奈, 米谷竜, 川嶋宏彰, 松山隆司. 視線運動解析による興味アスペクトの推定. ヒューマンインタフェース学会論文誌, volume 16, number 2, pages 103–114, 2014.