

# スマートフォンによる視線追跡手法を用いた インタラクティブな情報提示

田岡健太<sup>†</sup> 池松香<sup>‡</sup> 山中祥太<sup>‡</sup> 鳴海紘也<sup>†</sup> 坪内孝太<sup>‡</sup> 川原圭博<sup>†</sup>

<sup>†</sup> 東京大学 <sup>‡</sup> ヤフー株式会社

## 1 はじめに

本研究の目的は、スマートフォンにおける広告表示について、ユーザの本来の目的である記事の閲覧をできるだけ妨げない形で効果的に広告の情報を提示することである。そこで、スマートフォンで記事を開覧するユーザの視線を取得し、視線情報と連動した広告画像を提示する手法を提案する。具体的には、スマートフォンのインカメラでユーザの視線情報を取得し、視線が記事に向けられているときは文字情報のない簡素な画像を表示し、視線がターゲットとなる画像に向けたときだけ文字情報を含む詳細な情報を提示するシステムを実装した。また、クラウドソーシングを活用して本システムに関する393名の評価実験を行い、仮説検定により提案手法の有効性を検証した。以降の章では実装と評価実験の詳細について述べる。

## 2 視線追跡の取得

本研究では、視線情報の取得のために Apple 社が提供する ARKit [1] を使用した。ARKit を用いると、iPhone のインカメラの位置を原点として、撮影されたフレーム内の物体や人物を追跡できる。特に顔を検知した際には ARFaceAnchor という点が付与され、顔面のパーツの位置や角度を追跡できる。

本研究では、ARKit のトラッキングシステムを利用し、拡張現実空間における視線情報の座標を図1のようなスクリーン座標に変換し、画面上のユーザの注視箇所を推定した。

## 3 実験

### 3.1 提案手法

図2に、ユーザの視線と連動した広告画像を示す。十字の交点が注視点推定点である。推定された注視点が画像から離れているときは図2Aのように商品に関する写真だけの画像が表示され、注視点が画像に近づいたときは図2Bのように、商品名や値段を記した詳細な画像を表示する。

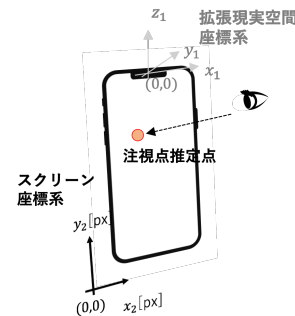
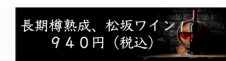


図1: スマートフォンのスクリーン座標系。

(A) 視線が記事に向いている場合

(B) 視線が広告に向いている場合



ノルウェーのノーベル賞委員会は現地10月10日(UTC+1)、2014年度ノーベル平和賞の受賞者を、パキスタンで女性や子供の教育を受ける権利を訴えたマララ・ユスフザイさん(17)と、インドの児童労働問題に取り組むカイラシュ・サトヤ

ノルウェーのノーベル賞委員会は現地10月10日(UTC+1)、2014年度ノーベル平和賞の受賞者を、パキスタンで女性や子供の教育を受ける権利を訴えたマララ・ユスフザイさん(17)と、インドの児童労働問題に取り組むカイラシュ・サトヤ

図2: 提案手法において、注視点(十字の交点)が(A) 広告に近いときと(B) 遠いときの広告画像(画像: アフロ)。

### 3.2 実験条件

Yahoo!クラウドソーシング<sup>1)</sup>にて iPhone X 以降の機種を所有するユーザを対象に393人の実験参加者を募集し、次のようなタスクを依頼した。

- 参加者はニュース記事に対するクイズという名目で記事を読んだ。
- ニュース記事が表示されている最中に、3通り(後述の提案・フラッシュ・静止)の手法で画面上部の固定された位置に広告画像を表示する。ただし各ユーザに対して2つの記事を表示し、それぞれに(提案・静止)または(フラッシュ・静止)の条件を割り当てた。
- 参加者は2つの記事を読んだあと、それぞれの記事に対する広告画像の想起度を測るテストと、不快感に関するアンケートに回答した。

想起度を測るテストでは、記事中に表示した画像とそれに類似した画像3つの計4つの中から、記事

Interactive Information Presentation Using Eye-Tracking Technologies Using a Smartphone

Kenta TAOKA<sup>†</sup>, Kaori IKEMATSU<sup>‡</sup>, Shota YAMANAKA<sup>‡</sup>, Koya Narumi<sup>†</sup>, Kota TSUBOUCHI<sup>‡</sup> and Yoshihiro KAWAHARA<sup>†</sup>

<sup>†</sup>The University of Tokyo, <sup>‡</sup>Yahoo Japan Corporation

1) <https://crowdsourcing.yahoo.co.jp>

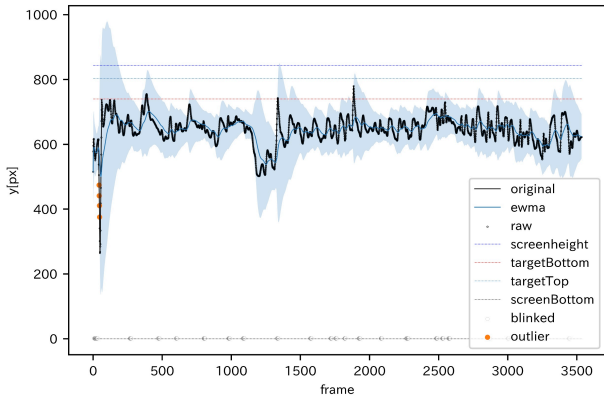


図 3: ある被験者の視線推定点の  $y$  座標の推移。

中に表示した画像を選択させ (3 点満点)、さらに広告画像に記載された商品の値段を記述させ (3 点満点)、計 6 点満点で計算した。なお、商品の値段についての自由記述問題は 3 桁の各位の数字に関するレーベンシュタイン距離を利用し 3 点満点となるよう計算した。また、不快感に関するアンケートでは、「記事で出された広告は不快に感じましたか?」という質問を 7 段階のリッカート尺度で選択させた。7 に近い方が不快度が高い。提案した情報提示手法に加え、比較対象として、視線に連動せずに 1 秒ごとに明滅する画像 (フラッシュ) と何も変化しない静止画像を用意し、提案手法と合わせた 3 通りの提示手法に対して後述の仮説検定を実施した。

### 3.3 仮説

検定に用いた仮説は以下の 3 つである：

- 帰無仮説 1：3 通りの提示手法に関して、ターゲットへの注視割合は等しい
- 帰無仮説 2：3 通りの提示手法に関して、広告の想起度は等しい
- 帰無仮説 3：3 通りの提示手法に関して、広告画像の不快感等は等しい

なお、注視割合はターゲットとなる画像の  $\pm 100$  px に視線が向いたフレームの割合を全体のフレーム数で割ったものとして計算した。

### 3.4 結果

視線データが上手く取れていない場合があるため、まずデータの選定を行った。図 3 は、ある被験者の時間フレームに対するスクリーン座標  $y$  [px] の観測データである。スクリーン最下端の  $y$  座標を 0 px としたとき、視線推定点の  $y$  座標が (スクリーンの最上端)  $+100$  px よりも大きい場合を外れ値とし、外れ値のフレーム数が全フレーム数の 10% 未満であること、さらに視線推定点の  $y$  座標の指数加重移動平均の平均値がスクリーン画面内にあることをデータ採用基準とした。

データの選定後、提案・フラッシュ・静止の 3 つの手法について、提案  $N = 89$ , フラッシュ  $N = 86$ , 静止  $N = 175$  のデータが集まった。3 つの提示手法に対する注視割合・想起度・不快感の平均値を表 1 に示す。

表 1: 提案・フラッシュ・静止の 3 つの提示手法における注視割合・想起度・不快感の平均値。

Factor	手法		
	提案	フラッシュ	静止
注視割合 [%]	29.00	28.41	29.03
想起度 (最大 6)	2.01	1.94	1.27
不快感 (最大 7)	3.71	4.10	3.38

3 つの factor に対する ANOVA を行ったところ以下の結果となった。

1. 注視割合： $p = 0.633 > 0.05$
2. 想起度： $p = 0.00246 < 0.01$
3. 不快感： $p = 0.0208 < 0.05$

帰無仮説 2 および帰無仮説 3 が棄却され、広告の想起度と不快感については 3 条件の間で有意差が生じることが分かった。さらに、仮説 2 (広告画像の想起度) に関して Tukey-Kramer 検定を行ったところ、提案と静止 ( $p = 0.008 < 0.01$ ) およびフラッシュと静止 ( $p = 0.020 < 0.05$ ) には有意差が見られた。一方で、仮説 3 (広告画像の不快感) に関しては、いずれのペアについても有意差は見られなかった。このことから、提案手法はフラッシュと同程度またはそれ以上に広告の想起度を向上させることが考えられる。

**おわりに** 本研究では、スマートフォンを用いて記事を読むユーザの閲覧を妨げず、かつ効果的な広告を提示するために、視線とインタラクティブに連動した画像の提示手法を提案した。また、393 人の参加者に対して検証実験を行い、提案・フラッシュ・静止の 3 条件に対して注視割合・想起度・不快感のデータを収集した。その結果、提案手法はフラッシュと同程度またはそれ以上に広告の想起度を向上させることが判明したものの、ある 2 条件に対する注視割合および不快感の有意差は得られなかった。そのため、今後は 1000 人規模の実験を行い、引き続き検証を行う予定である。

### 参考文献

- [1] Apple, ARKit, <https://developer.apple.com/documentation/arkit/>, Accessed on Jan 12, 2023.