

# 視力に応じて表示を切り替える デジタルサイネージシステムのための 視点情報に基づく視力判別手法

濱本 昂<sup>1</sup> 寺田 努<sup>1</sup> 塚本昌彦<sup>1</sup>

**概要:** 近年、デジタルサイネージが公共の場所に広く普及しているが、一定時間ごとに表示を切り替えるような単純な表示パターンが多く用いられている。デジタルサイネージ閲覧者の特徴を踏まえて表示を変えることができれば、広告などのより効果的な情報提示を実現できる。特に、デジタルサイネージは視力の異なる様々な人々に閲覧されるため、ユーザの視力に応じてテキストサイズを大きくすることは効果的な情報提示につながると考えられる。そのため、筆者はユーザがデジタルサイネージ上に表示された広告を読む際、視力に応じて視点移動に差異が生じることに着目し、その差異からユーザの視力を判別できると考えた。本論文では、ユーザがデジタルサイネージを閲覧する際の視点情報から視力を判別し表示を変更することにより、効果的に情報提示をするデジタルサイネージシステムを提案する。本論文で提案した視力判別手法では、OpenFace で取得したユーザの広告閲覧時の視点情報から眼球運動の特性を考慮した特徴量を算出し、まず Random Forest を用いて、視点情報に基づくユーザの状態（読中、読む方向）と特徴（視力の良し悪し）で整理された 6 クラスに分類を行う。分類後、サイネージを読んでいない 2 クラスを除き、残りの縦または横方向に視点が動きながらサイネージを読んでいる 4 クラスのデータを視力が良い、悪いの 2 種類のクラスに統合する。さらに、2 つのクラスのうち視力が悪いクラスの占める割合が閾値を上回った場合、広告を閲覧するユーザを目が悪いと判別し、表示内容の変更を行う。評価実験では、提案システムによる視力判別精度を評価し、さらに本手法において目が悪いと判別する条件に用いる、視力が悪いクラスの占める割合の閾値の値を調査した。評価の結果、6 クラスの分類では、Leave-one-subject-out 交差検証を行い、被験者平均で F 値が約 0.48 であった。評価実験で得たデータに対し、最も適切な閾値を適応すると、視力判別精度は 87.5% であった。加えて、提案システムでの視力判別が可能な広告の特徴についても調査した結果、提案システムにより視力判別が可能な広告は、広告中で用いられている文字数が多く大小の文字が混在しているものであることが分かった。

## 1. はじめに

近年、デジタルサイネージと呼ばれる電子的な表示機器を使って情報を発信する媒体を目にする機会が増えている。デジタルサイネージはその多くが駅やデパートなど公共の場所に設置されており、広告や案内の表示機器として用いられる。デジタルサイネージは紙を用いた従来の看板と比べて初期の導入コストがかかるものの、簡単に動的な表示内容の切り替えができるという特徴を持つ。そのため、特定の目的に合わせた広告の提示を行ったり、ユーザの操作に応じてインタラクティブな情報提示を行ったり、動画を用いた広告でユーザへの訴求力を高めたりすることができ、従来の看板からデジタルサイネージへ移行しようとする

流れが加速している。

しかし、現在の多くの一般的なデジタルサイネージはどんなユーザに対しても画面遷移等決まったパターンでしか表示内容を変化させていない。これでは、デジタルサイネージの最大の利点である「表示内容の切り替えの容易さ」を最大限生かし切れていないとはいえない。一方で、近年のセンシング技術の発展に伴い、デジタルサイネージを見るユーザの動きや身体的特徴を計測することが可能になっている。そのため、デジタルサイネージを見るユーザの位置、体格、視力などの特徴からそれぞれに対して動的に表示を変えることができれば、より効果的な情報提示が可能となる。

特に、視覚情報を主とするデジタルサイネージでの情報提供において、細かい文字などを用いた表示は目の悪いユーザに対して最大の障壁となりうる。目の悪いユーザが

<sup>1</sup> 神戸大学大学院工学研究科  
Graduate School of Engineering, Kobe University

デジタルサイネージに近づいて情報を得ようとしても広告や案内では細かい文字を用いた表示が行われるため、思うような情報を得られないことに不便さを感じる。視覚情報を主とするサイネージにおいて視力に応じて表示内容を変更するのは効果的な情報提示に繋がると考えられるが、視力によるユーザの属性分けを行った研究は筆者の知る限り行われていない。

そこで本研究では、デジタルサイネージを閲覧するユーザの視力の良し悪しを判別し、目が悪いために文章を読めていないユーザに対して広告中のフォントの大きさなど表示内容を動的に変化させるシステムを提案する。デジタルサイネージは駅やデパートなど公共の場で用いられることが多いことから、本システム構築のためにはユーザに特別な装置の着用を必要とせず、サイネージを見るだけで視力判別可能な手法が必要である。そこで、筆者は広告閲覧時の視点情報に着目した。その理由は、視点情報はウェブカメラ等を用いて記録したユーザの顔部分の動画からも検出可能であるため、前述の条件を満たしつつ、デジタルサイネージ上に表示された広告を読む際に視力に応じて差異が生じると考えられるからである。

以上より本稿では、広告閲覧時の視点情報に基づいた視力判別手法と提案手法を用いた表示変更システムを提案し、その評価を行う。

本稿は以下のように構成される。2章で本研究に関連する研究について述べ、3章で提案手法について説明する。4章で評価実験とその考察を行い、最後に5章で本論文のまとめを述べる。

## 2. 関連研究

### 2.1 デジタルサイネージを閲覧するユーザに応じて表示内容を変更する研究

近年のセンシング技術の発展に伴い、デジタルサイネージを閲覧するユーザの動きや身体的特徴などを計測することが可能になったことで、デジタルサイネージを閲覧するユーザに応じて表示内容を変更する研究は多数行われている。デジタルサイネージは公共の場に多く設置されているが、デジタルサイネージを閲覧するユーザの一部は閲覧することに対して抵抗を持っていたり、閲覧に対して不便さを感じている [1]。そこで、デジタルサイネージを閲覧するユーザに応じて表示内容を変更することにより、ユーザの興味を引くような情報提示やそれぞれのユーザに適した情報提示が可能となる。Vogel らは、ディスプレイとユーザの距離によって、それぞれのユーザに異なった映像を提示するシステムを提案した [2]。この研究ではディスプレイとユーザとの距離がディスプレイに表示されている内容に対するユーザの興味の指標になるという仮定のもと、ユーザの距離をセンシングすることで、いくつかに分けられた距離ごとに異なった映像・インタラクティブ要素を提供す

るシステムを実現している。鈴木らは、「一定距離以降から内容が認識可能になる」というモザイク画像の特徴を応用し、モザイク画像と通常解像度の画像とを重畳することによって、遠近2つの距離区分にいるそれぞれのデジタルサイネージ閲覧者に対して異なった画像を同時提供できる手法を開発した [3]。小玉らは、Kinect を用いて歩行者の位置に応じて画像を正対表示することで、より閲覧者の注意を引き付けるシステムを構築した。この研究では、システムを大学構内で稼働させた結果、生体表示を行わない場合と比較して、注視時間の増加や足を止める人が増加することを確認した [4]。スペインの児童支援団体 Fundacion ANAR は、見る角度によって絵柄が変化するレンチキュラーレンズを用いて、子供だけに特別なメッセージを表示させる看板を街に設置する Only for children というプロモーション活動を行った。井上らは、複数人の間の対人距離から恋人同士や友人同士といった人間関係を判別し、それに応じて適切な広告を表示するシステムを構築した [5]。この研究では、対人距離情報などから75%の精度で対人関係が判別できることが分かり、また、複数人の関係に基づく広告が、そうでない広告と比べて有効であることが認められた。このように、デジタルサイネージを閲覧するユーザの位置や身体的特徴に応じて表示内容を変更する研究は数多くなされており、デジタルサイネージ閲覧者に対して動的に表示を変更することは有効なアプローチであるといえる。しかしながら、ユーザの視力に応じて表示内容を変更する研究は筆者の知る限り行われていない。視覚情報を主とするサイネージにおいて視力に応じて表示内容を変更するのは効果的であると考えられるため、本研究ではユーザの視力に応じて表示内容を変更する。

### 2.2 ユーザの興味と視点情報の関係

人間の興味関心と目の動きの関係についての研究は数多くなされており、広告や Web ページ等の作成時に参考にされている。深澤らは、Web ページを閲覧しているユーザの視線と脈波の分析から、そのページにおけるユーザの興味箇所を特定する分析手法を提案した [6]。ユーザは Web ページ閲覧時、興味箇所では視線停留時間が長くなること、瞳孔径が大きくなること、脈波波高の振幅が小さくなることがわかった。青木らは、連続するストーリーの画像を見る際の視線によって観察者の興味を判断し、興味に応じてストーリーが変化していくインタラクティブデジタルシネマを開発し、それを用いてユーザの興味を持った画像と視点の関係性を調査した [7]。この研究により、ユーザは、特に「興味がある」と回答した画像の興味がある箇所をかなりの確率で見ていることが分かったが、「興味がある」箇所だけを見ていた事例は5%以下で、画面上の他の部分も含めて頻繁に視線を移動しながら観察していることがわかった。加藤木らは、テレビゲームにおけるキャラクタの動き

に対する視点情報の特性を分析し、興味対象に応じた映像分岐方法を提案し、視点情報から7体のキャラクタのうち興味のあるキャラクタを65%の正答率で推測することができた[8]。Ajankiらは、Web検索を行う際に、入力文字に続いて表示される予測候補を見る際のユーザの注視のパターンに基づき、80%の正答率でユーザが検索したい単語を推定可能なシステムを構築した[9]。

このように、ユーザの興味と目の動きには深い関係があり、デジタルサイネージに表示される広告という興味が重要な意味を持つコンテンツにおいて、視点情報を用いて情報提示を変えることは有効であるといえる。

### 2.3 視点情報とデジタルサイネージ

視点情報とデジタルサイネージを組み合わせにより良い情報提供を行おうとする研究は数多くなされている。人の視力と視線移動の関係を調査した研究に、Yangらのものがある[10]。この研究では、乳児眼振症候群（INS）患者にみられる特徴的な視線移動に注目し、INS患者に向けた視力測定の手法を提案した。彼らの視力測定手法はINS患者に特有の視線移動に基づいて研究を行ったものであるが、視力の悪いユーザにも視力の良いユーザに比べて特有の視線移動がみられた場合、視力判別出来る可能性がある。Zhangらは視線による制御機能を追加することで、よりインタラクティブなデジタルサイネージシステムを開発し、大学のロビーで一般ユーザに向けた実地調査を行った[11]。実地調査の結果、視線制御を加えたインタラクティブデジタルサイネージシステムは初心者ユーザにも簡単に視線制御を扱うことができ、一般的なデジタルサイネージと比べてよりユーザの興味を引くことが分かった。また、Yonezawaraはユーザの視線位置に応じ、従来の看板案内における補助的な情報をぬいぐるみの案内音声や動作を伴ってインタラクティブに提示するシステムを提案し、ロボットとのアイコンタクトはユーザにとって好ましい印象ロボットの視線の先により興味を引く共同注視が確認された[12]。このように、デジタルサイネージという情報提供媒体において視点情報を用いることは大きな効果を得られるといえる。

以上より本研究では、ユーザの視点情報に基づいてユーザの視力を判別し、表示内容を変化させるシステムを提案する。

## 3. 提案システム

本研究では、デジタルサイネージでユーザに効果的な提示を可能にするため、デジタルサイネージ上に表示される広告閲覧時のユーザの視点情報に基づき視力を判別し、表示内容を変化させるシステムを提案する。

### 3.1 想定環境

デジタルサイネージは駅やデパートなど公共の場で用い

られることが多く、大人や子供、視力の良い人や悪い人など様々なユーザが見ることが想定される。そのため、安全かつデジタルサイネージを見るだけで簡単に視力判別可能な手法が必要である。また、一般のデジタルサイネージはある特定の個人限定のものではなく、本システム構築のためにはユーザに特別な装置の着用を必要としない視力判別手法が必要である。

### 視線情報の取得に用いるセンサの検討

デジタルサイネージ閲覧時のユーザの視点情報を取得する方法として、装着型アイトラッカを用いる方法、非装着型アイトラッカを用いる方法の2つが考えられる。それぞれの特徴を説明した後、本研究ではどの方法を用いるべきかについて検討する。

装着型のアイトラッカはメガネ型のものが多く、内蔵されたカメラで眼球を撮影することにより眼球運動を測定する。また、カメラを用いず目の周りに電極を取り付けることにより眼球運動を測定するものもある。装着型は、頭部に装着するため場所を選ばずに眼球運動を測定することができるが最大の利点だが、デジタルサイネージ閲覧時のユーザの視点情報を取得するという想定環境においては、その度にユーザに対してメガネ型デバイスの着用を促すことはユーザの利便さを阻害するため、本研究で用いる視点情報取得方法としては適切といえない。

非装着型のアイトラッカは、内蔵されたカメラで顔全体を撮影することにより、ユーザの顔の向きと眼球運動などを測定し、ユーザの視点位置を特定する。非装着型はディスプレイ体型のものやカメラのみを用いるものなど、ユーザに特別なデバイスの着用等の負担を与えることなく視点位置の測定が可能である。そのため、装着型と比べて上記の想定環境で用いる視点情報取得方法としては適切であると考えられる。

本研究では非装着型のアイトラッカのうち、ケンブリッジ大学マルチコンプグループが開発したオープンソースの顔の動作解析ツールであるOpenFaceを用いてユーザの眼球運動を測定し、視点情報を取得する。OpenFaceはカメラにより取得した画像を用いて頭の向きや黒目の位置から視点を推定する。OpenFaceを用いる理由は、一般的なウェブカメラとPCのみで実行可能でありユーザに特別なデバイスの着用を必要としない点、ディスプレイからある程度離れても視点位置を測定可能でありユーザの行動範囲を制限しない点、複数人の視点情報を同時に取得できる点から、想定環境において最も適した視点情報取得方法であると考えられるからである。

### 提案システムの処理の流れ

以上を踏まえたうえで、提案システムの流れを図1に示す。まず、ユーザがデジタルサイネージに近づいて少し広告内容を見ようとする。その際の視点情報をカメラを用いて取得、得られた視点情報を分析した後、ユーザの視力の



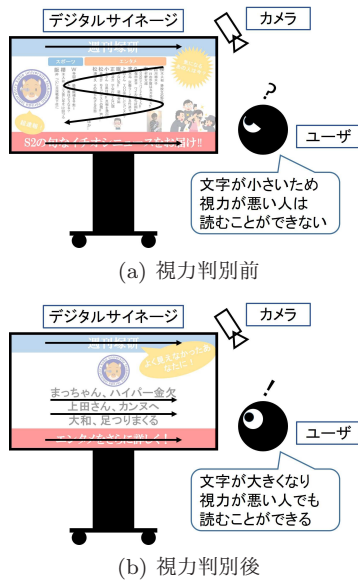


図 1 視力判別前後のシステムの流れ

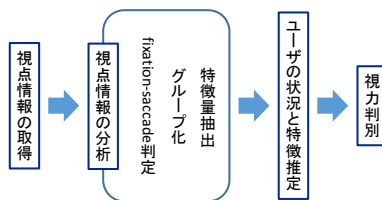


図 2 視力判別手法の流れ

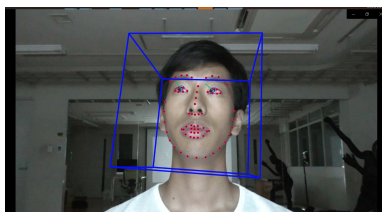


図 3 OpenFace による視点推定の様子

良し悪しを判別し、目の悪いユーザには大きなフォントの文字を使った広告など、より適した表示内容に変更する。

また、図 2 に本研究で用いる視力判別手法の流れを示す。まず、デジタルサイネージ上に表示される広告閲覧時のユーザの視点情報を取得する。その後、視点情報を一定時間で分割し処理を施したのち、それぞれの視点情報を複数の特徴量で表す。このようにして得られた特徴量を用いて、ユーザの視力の良し悪しの判別を行う。この流れを踏まえたうえで、以下にシステムで用いる視点情報取得方法と視点情報分析手法、視力判別手法を述べる。

### 3.2 視点情報の取得

提案システムでは、図 3 のように OpenFace を用いてユーザの眼球運動を測定し、視点情報を取得する。OpenFace はカメラにより取得した顔の画像における頭の向きや黒目の位置から視点を推定する顔の分析ツールで、視点情報は 2 次元座標 (X,Y) で得られる。

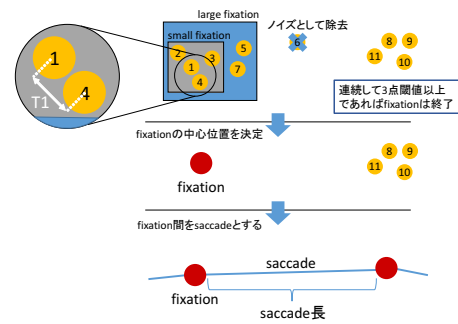


図 4 fixation 判定

### 3.3 データの前処理：特徴量抽出

提案する視力判別手法では、取得した視点情報をそのまま用いてユーザが文章を読んでいるか（読中）判別を行うのではなく、前処理によって眼球運動の特性を考慮した特徴量を算出することでより正確に読中判別を行うことができると考えた。

提案手法では、眼球運動の特性を利用して 6 つの特徴量を考案する。眼球運動は、blink, fixation, saccade の 3 つに分類される。blink とは、瞬きのことである。また、fixation は視点がある範囲内に一定時間停留することであり、saccade は fixation 間の素早い目の動きのことである。人は上記にあげた 3 つの眼球運動を繰り返すことでものを見たり文書を読んだりする。そこで、OpenFace から得られた視点の X,Y 座標情報から fixation-saccade を後述する手法で判定し、fixation と saccade の位置や方向、出現割合などを求め特徴量とする。

特徴量の抽出は、大まかに 3 つのプロセスに分類される。まず、得られた視点情報に fixation-saccade を判定する。その後、Kunze らが提案した視点情報の分析手法 [13] に基づき、得られた fixation と saccade をそれぞれ一定時間でグループにする。このようなグループごとに、特徴量を抽出する。以下にそれぞれについて詳しく述べる。

#### 3.3.1 fixation-saccade 判定

本手法では、fixation の判定手法に Buscher らの手法 [14] を用いる。図 4 にその手順を示す。図の黄丸は fixation-saccade 判定を施す前の視点位置であり、赤丸は fixation-saccade 判定を施した後の fixation の中心位置である。始めに視点が全て互いに距離  $T_1$  以内に連続して 4 点存在した場合、その 4 点を small-fixation とする。この時、 $T_1$  は視点の中心間の距離とする。そして、その次の視点と small-fixation の全ての視点との距離が閾値  $T_2$  以内であれば、その点を large-fixation とする。 $T_1$  と  $T_2$  は、 $T_1 < T_2$  の関係にある。閾値  $T_2$  以上である視点（図 4 の 6 番の点）はノイズとして除去する。Large-fixation に含まれない点が連続で 3 点現れると fixation は終了し、その最初の視点から新たに small-fixation 判定を始める。最終的には、fixation と判断された視点の重心を fixation の中心位置とする。そ

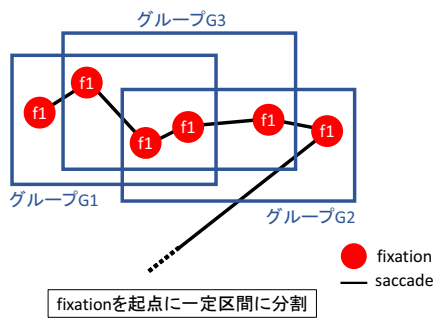


図5 視点情報のグループ化

表1 クラス分け

|                          | 目がいい (Good) | 目が悪い (Bad) |
|--------------------------|-------------|------------|
| 縦方向の文章を見ている (Vertical)   | GV          | BV         |
| 横方向の文章を見ている (Horizontal) | GH          | BH         |
| 文章を見ていない (Not See)       | GN          | BN         |

して fixation 間を saccade とし、その距離を saccade 長とする。このようにして視点情報に対し、fixation-saccade を判定する。

### 3.3.2 視点情報のグループ化

図5に視点情報のグループ化の概要を示す。連続する fixation と saccade に対して、グループ化を行う。得られた fixation の列を  $F = (f_1, f_2, f_3, \dots, f_N)$  とする。得られた  $F$  を一定時間で分割し、グループ化を行う。まず、最初の  $f_1$  を起点として、ウィンドウサイズを1秒間とし、1サンプルごとにウィンドウをスライドさせ、1秒間ごとのグループに分割する。上記にあげたグループから、それぞれ3.3.3項の特徴量を算出する。

### 3.3.3 特徴量の抽出

提案手法では、3.3.2項で分割した1秒間のグループごとに6次元の特徴ベクトルを抽出する。fixationに関するものには、fixation 持続時間の平均を用いる。saccadeに関するものは、saccade 長の平均長、saccade 長の最小値がある。ここで saccade 長とは図4に示す fixation 間のユークリッド距離である。また、saccade をベクトルとみなした際の合成ベクトルの x 成分、y 成分を用いる。これらに加えて、x 軸方向の視点情報に対し、一次元連続ウェーブレット変換を施して得た、視点移動の振幅の大きさの平均を用いる。

### 3.4 読中の視力判別

提案システムでは、視力が悪いユーザがデジタルサイネージを閲覧した際に、文字のフォントを大きくする等より適した表示に変更し、サイネージ上で効果的に情報を提示することを目的としている。そのため、視力が良いユーザを「視力が良い」と判別することよりも、サイネージを見ているが視力が悪いユーザを「視力が悪い」と判別することが重要である。眼球運動を考慮した6つの特徴量を

得た後、機械学習アルゴリズムの RandomForest を用いてこれらの特徴量からまずユーザの状況認識を行い、表1に示す6つのクラスに分類する。分類対象のクラスの選定について、日本語には縦方向に読む場合と横方向に読む場合があるため、読む方向を考慮して、縦方向の文章を見ている、横方向の文章を見ている、文章を見ていないの3種類に、視力の良し悪しを対応させた6クラスに分類することとした。

その後、GV, GH, BV, BH の4クラスは文章を読んでいる状態 (R), GN, BN の2クラスは文章を読んでいない状態 (NR) に分類し、文章を読んでいる状態 (R) のみを視力判別に用いる。文章を読んでいない際のユーザの視点移動には視力によってあまり違いがみられなかったため NR のデータはこの段階で破棄する。

提案システムは公共の場などに設置されることを想定しており、ユーザは通りすがりに短時間サイネージを眺める、といったことが考えられる。そのため、視力判別を行った後、サイネージの表示変更はユーザが広告を読み進める段階でリアルタイムに行うことが求められる。そこで、ユーザが広告の閲覧を開始した時点から一定時間ごと（2秒、5秒、10秒、15秒）に閾値を定め、R のラベル付けがされたデータのうち BV, BH の占める割合が閾値を上回っていた場合、広告を閲覧しているユーザは「目が悪い」と判別し表示を変更する。一方、BV と BH の合計が閾値を下回っていた場合、「目が悪い」とは判別せず、表示の変更は行わない。

### 3.5 視力判別後のサイネージの表示変更

図1に、提案手法による視力判別前と後に表示する広告の例を示す。ユーザがデジタルサイネージに近づいて図1(a)に示すような広告を読んだ際に提案手法による視力判別を行った結果、視力が悪いと判別された場合、図1(b)に示すような文字のフォントが大きく、内容が簡潔な広告に表示を変更する。以上の表示変更により、視力の悪いユーザに適した情報提示を行うことが可能となる。

### 3.6 まとめ

本研究で提案する視力判別手法では OpenFace で取得したユーザの広告閲覧時の視点情報を特徴量とし、Random Forest を用いて学習を行い、GV, GH, GN, BV, BH, BN の6クラスに分類を行う。分類後、推定結果から有効なデータではないサイネージを読んでいない状態である GN, BN を除き、残りの縦または横方向に視点が動きながらサイネージを読んでいる4クラスのデータをクラス G (GV 及び GH)、クラス B (BH 及び BV) の2種類に統合し、2つのクラスのうちクラス B の占める割合が閾値を上回った場合、広告を閲覧するユーザを目が悪いと判別し、表示内容の変更を行う。



図 6 実験環境

表 2 被験者情報

| 被験者 | 視力   | 日常生活における視力矯正 | 視力分類 |
|-----|------|--------------|------|
| A   | 0.05 | 有り           | 悪    |
| B   | 0.1  | 有り           | 悪    |
| C   | 0.1  | 有り           | 悪    |
| D   | 0.3  | 無し           | 悪    |
| E   | 0.5  | 無し           | 良    |
| F   | 0.9  | 無し           | 良    |
| G   | 1.5  | 無し           | 良    |
| H   | 2.0  | 無し           | 良    |

## 4. 評価実験

### 4.1 実験内容

本章では 3 章で述べた提案システムによる視力判別精度を評価し、さらに本手法において目が悪いと判別する条件に用いる、クラス B の占める割合の閾値の適切な値を調べるための実験を行った。実験では、各被験者は視力測定を行った後、図 6 に示すように、縦 205 cm × 横 305 cm のスクリーンに単焦点プロジェクタにより投影された広告の内容を約 60 cm 離れた位置から読む。その際にユーザの顔付近の映像をビデオカメラで録画し、録画した映像に対して OpenFace を用いて視点情報を測定し、記録した。その後、広告閲覧開始時から終了時までの視点情報を用いて 6 クラスの推定精度を評価した。被験者は 20 代の男性 8 名で、全員裸眼で実験を行った。提案システムは公共の場での利用を考えているため、被験者本人のデータを学習データとせず、他者のデータから視力を推定する必要がある。そのため 6 クラスの推定では Leave-one-subject-out 交差検証法を用いて推定精度を評価した。実験で用いた広告は日本語の広告 7 枚、英語の広告 6 枚の計 13 枚である。表 2 に被験者情報を示す。なお、本論文では、評価環境での小さな文字が読める視力である 0.5 以上を目が良い、0.5 未満を目が悪いと分類した。

### 4.2 実験結果

#### 4.2.1 視点情報に基づくユーザの状態と特徴の 6 クラス分類結果

まず、実験により得られた視点情報に基づき表 1 で示し

表 3 視点情報に基づくユーザの状態と特徴の 6 クラス分類結果

|                   |    | Prediction |      |      |      |      |       | Recall |
|-------------------|----|------------|------|------|------|------|-------|--------|
|                   |    | GV         | GH   | GN   | BV   | BH   | BN    |        |
| Actual Activities | GV | 1871       | 43   | 180  | 251  | 17   | 762   | 0.60   |
|                   | GH | 67         | 4531 | 1029 | 49   | 1214 | 983   | 0.58   |
|                   | GN | 28         | 56   | 673  | 88   | 261  | 581   | 0.40   |
|                   | BV | 263        | 24   | 182  | 543  | 38   | 76    | 0.48   |
|                   | BH | 21         | 681  | 181  | 11   | 1826 | 311   | 0.60   |
|                   | BN | 99         | 81   | 314  | 194  | 86   | 261   | 0.25   |
| Precision         |    | 0.80       | 0.84 | 0.26 | 0.48 | 0.53 | 0.088 | 0.48   |

表 4 クラス G とクラス B の分類結果

|                   |   | Prediction |      | Recall |
|-------------------|---|------------|------|--------|
|                   |   | G          | B    |        |
| Actual Activities | G | 6512       | 1531 | 0.81   |
|                   | B | 989        | 2418 | 0.71   |
| Precision         |   | 0.87       | 0.61 | 0.75   |

たユーザの状態（読中、読む方向）と特徴（視力の良し悪し）の 6 クラス（GV, GH, GN, BV, BH, BN）に推定した結果を表 3 の Confusion Matrix に示す。ここでは被験者が広告を閲覧終了した後のデータをクラス分類した結果のみを記載する。Confusion Matrix とは、クラス分けが正しく認識された数を示す分類表であり、各セルは各クラスの出力回数を示している。表中の Actual Activities は正解データに基づく実際のクラス、Prediction は機械学習モデルの予測結果を表し、Precision は正解データのクラス分類が機械学習により分類されたクラスと一致したデータの割合（適合率）であり、Recall は機械学習により分類されたクラスが正解データのクラス分類と一致したデータの割合（再現率）である。色付けされているセルは同じクラス内で推定された数が最も多かったものである。Confusion Matrix において、対角線上の行と列に示されたクラスが一致しているセルが正しく認識を行えたときのものである。実験で行った分類では、F 値で 0.48 の認識率が得られた。表 3 より BN 以外のクラスは正しく推定が行われたデータが最も多いため、大まかには推定できていることがわかる。

#### 4.2.2 視力が良いクラスと視力が悪いクラスの分類結果

4.2.1 項の 6 クラス分類の後、有効なデータではない GN, BN のデータは破棄し、GV, GH, BV, BH のみについて、表 3 の結果をもとにクラス G とクラス B の 2 クラス分類を行った。分類結果を表 4 に示す。表 4 より、F 値で 0.75 の認識率が得られた。また、クラス G、クラス B とともに推定が正しく行われたデータが最も多くなった。

#### 4.2.3 各被験者の時間ごとの視力が悪いクラスと推定された割合

さらに、各被験者が広告を閲覧し始めてから 2 秒後、5 秒後、10 秒後、15 秒後、広告閲覧終了後までのそれぞれの視点情報について、表 5 に各被験者の時間ごとのクラス B と推定された割合を示す。全体的な傾向としては、被験者 A, C, D, E, F はクラス B と推定された割合が高く、被



表 5 各被験者の時間ごとのクラス B と推定された割合

| 被験者 | 視力分類 | クラス B と推定された割合 |      |       |       |       |
|-----|------|----------------|------|-------|-------|-------|
|     |      | 2s 後           | 5s 後 | 10s 後 | 15s 後 | 閲覧終了後 |
| A   | 悪    | 0.92           | 0.89 | 0.88  | 0.91  | 0.90  |
| B   | 悪    | 0.88           | 0.83 | 0.68  | 0.67  | 0.72  |
| C   | 悪    | 0.32           | 0.42 | 0.77  | 0.75  | 0.82  |
| D   | 悪    | 0.41           | 0.45 | 0.52  | 0.61  | 0.62  |
| E   | 良    | 0.71           | 0.72 | 0.75  | 0.69  | 0.66  |
| F   | 良    | 0.49           | 0.51 | 0.57  | 0.42  | 0.38  |
| G   | 良    | 0.43           | 0.41 | 0.27  | 0.32  | 0.21  |
| H   | 良    | 0.11           | 0.20 | 0.33  | 0.33  | 0.24  |

験者 B, G, H はクラス B と推定された割合が低かった。上記の結果をもとに、各時間ごとに視力判別をするためのクラス B と推定された割合の閾値を決める。図 7 に各時間ごとの目が良い被験者と目が悪い被験者のクラス B と推定された割合の分布を示す。目が悪い被験者は広告閲覧終了後までの視点情報を用いた場合には 4 名ともクラス B と推定された割合が 0.6 を超えていた。また、15 秒経過時点と広告閲覧終了時点のクラス B と推定された割合の分布には大きな差がなかった。一方で、10 秒以下の経過時間の場合は分布に差が表れた。特に広告閲覧開始から 2 秒後や 5 秒後までの視点情報からクラス B と推定された割合は、目の悪い被験者、目の良い被験者ともに広く分布しており、広告閲覧開始から 2 秒間や 5 秒間のみの視点情報では、正確な視力判別が難しいといえる。しかしながら、視力の悪い被験者のうち被験者 A, E は広告閲覧開始から 2 秒後や 5 秒後であっても、クラス B と推定された割合が 0.8 を超えており、時間経過に関わらず、クラス B と推定された割合が 0.8 を超えたのはこの 2 名のみであった。そこで、ユーザが広告閲覧開始から 2 秒以内又は 5 秒以内にクラス B と推定された割合が 0.8 を超えた場合、そのユーザは目が悪いと判別し、表示を変化させることが有効であると考えられる。また、前述のように広告閲覧開始から 15 秒経過後と広告閲覧終了後までの視点情報を用いた場合、4 名ともクラス B と推定された割合が 0.6 を超えていた。この結果から、本システムではユーザが広告閲覧を開始した時点から 2 秒後、5 秒後までの視力判別では閾値 0.8、15 秒後、広告閲覧終了時点までの視力判別では閾値 0.6 と設定し、定めた閾値を超えた場合、広告閲覧者を「目が悪い」と判別し、表示変更を行うことが適切であるといえる。

### 4.3 考察

#### 4.3.1 システムにより視力判別可能な広告の特徴

広告には、イラストを主としたもの、文字を主としたもの、日本語が用いられたもの、英語が用いられたものなど、様々なものが存在する。そこで、評価実験の結果をもとに、どのような広告において提案システムでユーザの視力判別が可能なのか検討する。表 6 に、取得した視点情報のうちクラス B と推定された割合と視力判別の可能性を示す。表

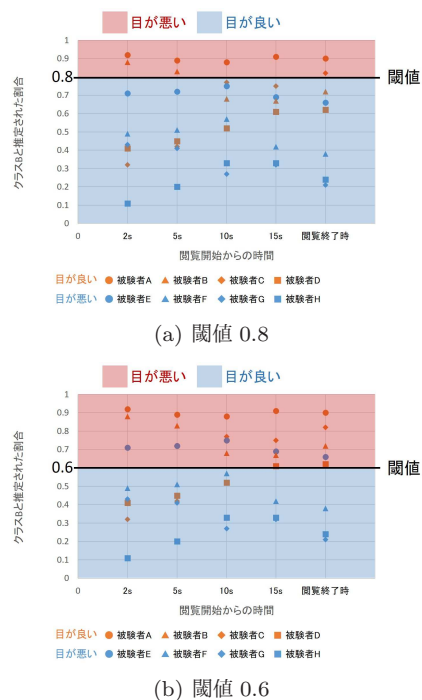


図 7 各被験者のクラス B と推定された割合と閾値の関係

表 6 広告ごとの視力判別可能性

| 広告 | クラス B と推定された割合 |            | 視力判別 |
|----|----------------|------------|------|
|    | 目が悪い被験者の平均     | 目が良い被験者の平均 |      |
| 1  | 0.53           | 0.48       | 難しい  |
| 2  | 0.61           | 0.36       | 可能   |
| 3  | 0.84           | 0.21       | 可能   |
| 4  | 0.76           | 0.32       | 可能   |
| 5  | 0.69           | 0.44       | 可能   |
| 6  | 0.81           | 0.30       | 可能   |
| 7  | 0.77           | 0.65       | 難しい  |
| 8  | 0.73           | 0.75       | 難しい  |
| 9  | 0.78           | 0.31       | 可能   |
| 10 | 0.76           | 0.29       | 可能   |
| 11 | 0.61           | 0.33       | 可能   |
| 12 | 0.80           | 0.29       | 可能   |
| 13 | 0.66           | 0.42       | 可能   |

中では目の良い被験者と目の悪い被験者それぞれの平均のクラス B と推定された割合を示している。

本システムを用いる場合、広告 2, 3, 4, 5, 6, 9, 10, 11, 12, 13 は視力判別可能であると判断した。これらの広告は広告中で用いられる文字数が多く、大小の文字が混在している広告であるという共通点が存在する。例として、図 8 に被験者 A と G が広告 3 を閲覧した場合のそれぞれの視点移動を示す。目の悪い被験者 A は目の良い被験者 G に比べて、小さな文字部分をあまり閲覧しないために視点の滞留である fixation の回数が少なく、一度の saccade の距離も大きくなっていることが分かる。

一方、広告 1, 7, 8 は視力判別可能とはいえなかった。これらの広告はイラスト主体のものであるため広告中で用いられている文字数が少なく、大きい文字がそのほとんどを

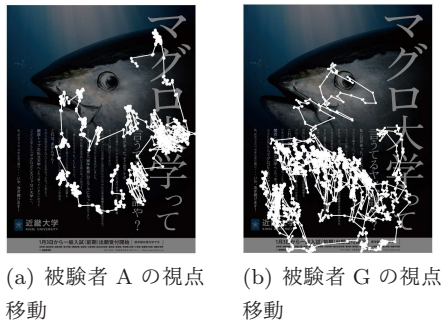


図 8 広告 1 を閲覧した際の被験者の視点移動

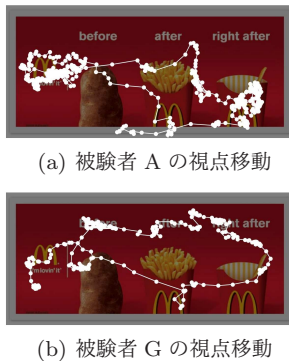


図 9 広告 6 を閲覧した際の被験者の視点移動

占めている。例として、図 9 に被験者 A と G が広告 8 を閲覧した場合のそれぞれの視点移動を示す。目の良い被験者 G は目の悪い被験者 A と比べて文字部分を中心に閲覧していることが分かるが、fixation の回数や saccade の距離に大きな違いは見られなかった。

また、被験者 2 名の言語能力は、共に日本語はネイティブレベル、英語は日常会話レベルであったが、日本語と英語の広告ともに上記の傾向がみられた。そのため、日本語の広告、英語の広告共に本システムで視力判別可能であると判断した。以上のことから、提案システムにより視力判別が可能な広告は、広告中で用いられている文字数が多く、大小の文字が混在しているものである。そのため、提案システムを用いて視力判別を行う場合には広告のデザインにも注意を払う必要がある。

また、デジタルサイネージには商業施設等で案内板として用いられるものもある。ユーザがデジタルサイネージに表示された案内を閲覧するとき、ユーザは注視を行うため提案手法は有効に働き、視力を判別することが出来ると考えられる。案内板は様々な視力の人々が情報を得るために活用するが、多くの案内板には小さな文字での表示が用いられているため、本システムによって視力の悪いユーザに表示を変更することは非常に有効であると考えられる。

#### 4.3.2 OpenFace のリアルタイムでの視線推定精度

本研究では、視線推定のために OpenFace を用いた。3 章で OpenFace はユーザに特別なデバイスの着用を必要としない点、ユーザの行動範囲を制限しない点、複数人の視

点情報を同時に取得できる点から、想定環境において最も適した視点情報取得方法であると述べた。また、OpenFace の特徴として一般的なウェブカメラと PC のみで実行可能であり、他の非装着型アイトラッカよりも導入コストが安価であることが挙げられ、実際に本システムを導入するには大きな利点となる。

しかしながら、OpenFace は他の非装着型アイトラッカと比べてリアルタイムで視線推定を行うと精度が低くなるという欠点が存在する。本稿で行った実験では OpenFace によるリアルタイムの視点推定ではなく、録画した動画に OpenFace を用いて視点推定を行った。そのため、他の非装着型アイトラッカに劣らない視点推定精度を得られたが、実際にシステムを運用する場合にはリアルタイムでの視点推定精度が低いことは大きな障壁となりうる。この問題は OpenFace の視点推定精度の向上もしくは OpenFace に代わる視点推定手法や非装着型アイトラッカの登場により解決され则认为る。

#### 4.4 議論

本論文で行った評価実験では、被験者はデジタルサイネージの代わりにスクリーンに投影した広告を閲覧し、実験環境において小さな文字を読むことが出来た視力 0.5 以上を目が良い、0.5 未満を視力が悪いと分類した。しかしながら、小さな文字を読むことが出来るか否か、という視力分類は周囲の明るさや被験者と広告の距離など様々な要因により変化するものである。そのため、実際にシステムを運用するにあたっては本論文の 0.5 を閾値とした分類にはとらわれず、その利用環境に合わせて視力分類を行ったうえで広告中の文字を読めていないなどの問題を抱えているユーザに表示を切り替えることが重要である。

さらに、本論文で行った評価実験では一つの広告を一人がみた場合についてのみ実験を行ったが、実際のデジタルサイネージは複数人が同時に画面を閲覧することが想定される。その場合には、複数のユーザそれぞれに対して視力判別を行い、もし目が悪い人が大半を占めているなら大きな文字の広告に変更するといったアプローチや、画面分割を用いて目の良い人、悪い人それぞれに適した情報提示に変更するといった対応が考えられる。

また、デジタルサイネージには商業施設等で案内板として用いられるものもある。案内板を閲覧する際、ユーザはそれぞれが求める情報を得るために広告を閲覧する時以上に注視を行うため、提案手法は有効に働き視力を判別することが出来ると考えられる。案内板は小さな文字も多く用いられているため、本システムによって視力の悪いユーザに表示を変更することは非常に有効であると考えられる。その際には、視点情報による興味推定などを利用し、正確にユーザが欲しい情報を拡大表示することが必要である。



## 5. まとめ

本論文では、ユーザがデジタルサイネージを閲覧する際の視点情報から視力を判別し表示を変更することにより、効果的に情報提示をするデジタルサイネージシステムを提案した。本論文で提案した視力判別手法では、OpenFaceで取得したユーザの広告閲覧時の視点情報から眼球運動の特性を考慮した特徴量を算出し、まず Random Forest を用いて、視点情報に基づくユーザの状態（読中、読む方向）と特徴（視力の良し悪し）で整理された6クラスに分類を行う。分類後、サイネージを読んでいない2クラスを除き、残りの縦または横方向に視点が動きながらサイネージを読んでいる4クラスのデータを視力が良い、悪いの2種類のクラスに統合する。さらに、2つのクラスのうちクラスBの占める割合が閾値を上回った場合、広告を閲覧するユーザを目が悪いと判別し、表示内容の変更を行う。

評価実験では、提案システムによる視力判別精度を評価し、さらに本手法において目が悪いと判別する条件に用いる、視力が悪いクラスの占める割合の閾値の値を調査した。評価の結果、6クラスの分類では、Leave-one-subject-out交差検証を行い、被験者平均でF値が約0.48であった。視力が悪いクラスの占める割合の適切な閾値に関しては、提案手法においては広告閲覧開始から2秒後、5秒後までの視力判別では閾値0.8、15秒後、広告閲覧終了時までの視力判別では閾値0.6を上回った場合、視力が悪いと判別することが有効であるとわかった。評価実験で得たデータに対し、最も適切な閾値を適応すると、視力判別精度は87.5%であった。加えて、提案システムでの視力判別が可能な広告の特徴についても調査した結果、提案システムにより視力判別が可能な広告は、広告中で用いられている文字数が多く大小の文字が混在しているものであることが分かった。そのため、提案システムを用いて視力判別を行う場合には広告のデザインにも注意を払う必要がある。

今後の課題として、提案手法で用いた特徴量の他に提案手法の視力判別の精度を上げるために有効な特徴量の検討を行う。また、リアルタイムの視点推定でも提案システムを用いて視力判別できるか検討を行う。さらにどのような広告でも目の悪いユーザが見やすい表示に切り替えられるアルゴリズムの実装を行う。

## 参考文献

- [1] C. Bauer, P. Dohmen, and C. Strauss: Interactive Digital Signage – an Innovative Service and Its Future Strategies, *Proc. of the 1st International Workshop on Frontiers in Service Transformations and Innovations (FSTI 2011)*, pp. 137–142 (Sep. 2011).
- [2] D. Vogel and R. Balakrishnan: Interactive Public Ambient Displays: Transitioning from Implicit to Explicit, Public to Personal, Interaction with Multiple User, *Proc.*

- of the 17th ACM User Interface Software and Technology Symposium (UIST 2004)*, pp. 137–146 (Oct. 2004).
- [3] 鈴木裕太, 小池英樹: モザイク画像を利用した遠近画像の同時提示手法, 第20回インタラクティブシステムとソフトウェアに関するワークショップ, Vol. 6, pp. 61–66 (Dec. 2017).
- [4] 小玉 駿, 須藤翔太, 渋谷 進: デジタルサイネージに向けた情報を正対表示させ歩行者を引き付けるシステム, 情報処理学会研究報告, Vol. 2014-HCI-157, No. 4, pp. 1–8 (Mar. 2014).
- [5] 井上智雄, 瓶子和幸: グループに適応する公共空間向け広告システム GAS, 情報処理学会論文誌, Vol. 49, No. 6, pp. 1962–1971 (June 2008).
- [6] 深澤 綾, 小俣昌樹, 今宮淳美: Web ページ閲覧時の視線停留時間と脈波波高に基づくユーザの興味箇所の特定, 電子情報通信学会技術研究報告, Vol. 107, pp. 1–6 (Jan. 2008).
- [7] 青木美奈, 勝本道哲: 視線はどこまで興味を反映するか—インタラクティブデジタルシネマにおける興味抽出のために—, 情報処理学会マルチメディア通信と分散処理ワークショップ論文集, pp. 318–322 (Nov. 2005).
- [8] 加藤木健太, 藤堂英樹, 柿本正憲: 視線追跡による好みキャラクタの判定, 情報処理学会第79回全国大会, Vol. 1, pp. 1023–1024 (Mar. 2017).
- [9] A. Ajanki, D. R. Hardoon, S. Kaski, K. Puolamäki, and J. S. Taylor: Can Eyes Reveal Interest? Implicit Queries from Gaze Patterns, *User Modeling and User-Adapted Interaction*, Vol. 19, pp. 307–339 (2009).
- [10] D. Yang, R. W. Hertle, V. M. Hill, and D. J. Stevens: Gaze-Dependent and Time-Restricted Visual Acuity Measures in Patients with Infantile Nystagmus Syndrome (INS), *Am J Ophthalmol* 2005, Vol. 139, pp. 716–718 (Nov. 2005).
- [11] Y. Zhang, J. Muller, M. K. Chong, A. Bulling, and H. Gellersen: GazeHorizon: Enabling Passers-by to Interact with Public Displays by Gaze, *Proc. of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2014)*, pp. 559–563 (Sep. 2014).
- [12] T. Yonezawa, H. Yamazoe, A. Utsumi, and S. Abe: Gaze-Communicative Behavior of Stuffed-Toy Robot with Joint Attention and Eye Contact Based on Ambient Gaze-Tracking, *Proc. of the 9th International Conference on Multimodal Interfaces (ICMI 2007)*, pp. 140–145 (Nov. 2007).
- [13] K. Kunze, Y. Utsumi, Y. Shiga, and K. Kise: I Know What You Are Reading - Recognition of Document Types Using Mobile Eye Tracking, *Proc. of the 12th International Symposium on Wearable Computers (ISWC 2013)*, pp. 113–116 (Sep. 2013).
- [14] G. Buscher and A. Dengel: Gaze-Based Filtering of Relevant Document Segments, *Workshop on Web Search Result Summarization and Presentation (WSSP 2009)*, Vol. 9, pp. 113–116 (Apr. 2009).