

## 視線情報を用いた移動環境下における興味対象の検出

宮島 大和<sup>†</sup> 三上 弾<sup>†</sup>

<sup>†</sup>工学院大学情報学部  
東京都新宿区西新宿1丁目24-2

E-mail: †j119276@g.kogakuin.jp, ††mikami.dan@cc.kogakuin.ac.jp

**あらまし** 移動中に興味を感じたことが後に役に立つことは少なくない。しかし、移動中に興味を持ってもすぐ忘れてしまうこともしばしばある。人の視線には、興味や関心などの心的状態が現れることが知られており、視線情報の中でも注視の時間長及び回数分析により被験者の興味対象を推定できることが確認されている。そこで本研究では、注視時間と注視回数に基づいた興味推定を、メガネ式アイトラッカーを用いた移動環境下において実現し、興味対象を検出することを目指す。注視時間と注視回数の取得には、物体を注視し続ける「連続的な注視」と、以前に注視した対象物に対して再び注視する「断続的な注視」の2種類の注視検出が必要である。そこで本研究では、従来手法を活用した連続的な注視検出と、新規に提案する断続的な注視検出とを組み合わせた興味対象検出手法を提案する。評価実験の結果、約87%の興味対象を検出し、提案手法の有用性を示した。

**キーワード** 視線, アイトラッカー, 興味推定, 注視検出, Re-identification, YOLO

## Detecting objects of interest in a moving environment using gaze information

Yamato MIYAJIMA<sup>†</sup> and Dan MIKAMI<sup>††</sup>

<sup>†</sup> Kogakuin University  
1-24-2, Nishi-shinjuku, Shinjuku-ku, Tokyo

E-mail: †j119276@g.kogakuin.jp, ††mikami.dan@cc.kogakuin.ac.jp

**Abstract** What one finds interesting while on the move is often useful later on. However, it is often the case that people forget what they are interested in while on the move. It is well known that the mental states of interest and interest are expressed in people's gaze, and it has been confirmed that the subjects' objects of interest can be estimated by analyzing the duration and frequency of gazing among eye gaze information. In this study, we aim to realize interest estimation based on gazing time and gazing frequency in a moving environment using a spectacle eye tracker, and to detect objects of interest. To obtain gazing time and frequency, two types of gazing detection are required: continuous gazing, in which the subject gazes at an object continuously, and intermittent gazing, in which the subject gazes at an object that has been gazed at before and gazes at it again. In this study, we propose an object-of-interest detection method that combines continuous gazing detection using conventional methods and a newly proposed intermittent gazing detection method. The results of evaluation experiments show that the proposed method is useful, detecting about 87% of the objects of interest.

**Key words** gaze information, eye tracker, interest estimation, gaze detection, Re-identification, YOLO

### 1. はじめに

我々は通学や通勤を含めて日々多くの時間を移動に費やしている。移動は、周囲とのインタラクションであり、移動中に興味を感じたこと、関心を持ったことが後に役に立つことも少なくない。しかし、移動中に関心を持ってもすぐ忘れてしまうこともしばしばである。このような背景から、移動環境下で興味対象を自動検出することは大きな意味があると考えられる。

さて、興味や関心などの心的な状態がしばしば視線に現れることが知られている。そのため視線計測への大きなニーズが存在する。視線を計測する装置はアイトラッカーと呼ばれ、2種類のアプローチが知られている。第1は据え置き型であり、ディスプレイの上部や下部に配置される。据え置き型のアイトラッカーは、ディスプレイ内の視線の向いた位置を出力することが一般的である。また、視線計測対象者は、頭部姿勢変化を制限されてしまい、自然な状態での計測は難しいという側面もある。

第2は、メガネ型である。メガネ型のアイトラッカーは、名前の通りメガネのように頭部に装着するものである。メガネ型のアイトラッカーはワールドカメラと呼ばれる装着者視線の映像を取得するカメラを有し、ワールドカメラ映像内における視線の位置を出力する。これは、頭部に装着していることから、頭部姿勢変化に合わせてアイトラッカーも移動し、そのため、歩行しながら[1]など、自由に様々な状況で利用可能である。

視線情報の中でも、注視<sup>1</sup>の時間長および注視回数の分析により興味対象を推測可能であると確認され[2]、Webページのユーザビリティ評価[3]、写真閲覧時の潜在的嗜好の推定[4]など様々な分野で「注視」に注目した研究が行われている。しかし、その研究の多くが据え置き型を用いて行われている。これは、据え置き型では、ディスプレイ上の提示内容を制御可能であり、注視や注視対象を実験者が把握することが容易であるためと推測される。これに対して、メガネ式を用いた場合、ワールドカメラ映像は装着者の頭部姿勢変化によって大きく変化する。この動きは実験者が制御できるものではなく、その結果、注視の検出が極めて困難となる。

本研究では、[2]を参考に、注視時間と注視回数に基づいた興味推定を、移動環境、すなわち、メガネ式アイトラッカーで視線計測した状況下においても実現することを目指す。注視時間と注視回数を取得するためには、物体を注視し続ける「連続的な注視」と、以前に注視した対象物に対して再び注視する「断続的な注視」の2種類の注視検出が必要である。

メガネ式アイトラッカーにおける連続的な注視の検出の取り組みとしてSteilら[5]が挙げられる。メガネ式アイトラッカーでは、頭部姿勢変化や移動によりワールドカメラ映像における対象物の座標値が変化するが、Steilらは、2時刻における注視点周辺領域の画像間類似度を求めることで、同一対象を見ているのか否かを判定した。

断続的な注視を自動検出するためには、注視対象物を識別することが必要となる。自動車運転者の行動分析を行う目的で注視対象物を識別する取り組みが行われている。針ヶ谷らは、解析対象となる対象物を事前に学習した検出器を作成し、注視状態中のフレームに検出器を適用することにより注視対象物を識別した。[6]。しかし、解析対象が多い場合は検出器の作成が困難なことや、興味対象検出のように解析対象が不特定の場合など、注視対象が検出器の学習データに含まれていないときは注視対象物を識別することができないため、針ヶ谷らの手法を興味対象検出に使用することはできない。

以上を踏まえ本研究では、移動環境下においてメガネ式アイトラッカーでアイトラッキングした状況における興味推定というタスク設定に対し、Steilらの手法を活用した連続的な注視検出と、新規に提案する同一物体を際識別する断続的な注視検出とを組み合わせることによる解法を提案する。以降、2で関連研究、3で提案手法、4で実験、5で今後の課題を述べた後に、6でまとめる。

(注1)：ある対象に対して視線を向ける状態であり Fixation とも呼ばれる。対となる言葉にサッケードがある。これは急速な眼球運動を指す。

## 2. 関連研究：Re-identification

本研究の技術的貢献のメインは、メガネ式アイトラッカーにより取得した情報に基づく、断続的な注視の検出である。これは、同一物体を再識別する技術であり、Re-identificationの技術と深くかかわる。

複数のカメラ間あるいは、同じカメラの他の場面において同一物体を再識別する技術を Re-identification という[7],[8]。Re-identification は大きく2つの処理に分かれており、第1の処理では画像内から検出対象のオブジェクトをバウンディングボックスと呼ばれる外接矩形で検出する。第2の処理では、バウンディングボックス内から特徴量を抽出し、抽出した特徴量に類似するものが過去にあるかを判断する。類似するものがあれば同じクラスとして再識別し、なければ新規クラスとして保存する。

Re-identification は第1の処理で検出対象を事前に学習した検出器を用いることが一般的である。そのため、人を対象とした Person-Re-identification [7] や車などの乗り物を対象とした Vehicle-Re-identification [8] など検出対象が制限されることが多い。

Re-identification を断続的な注視の検出への適用を考えると、注視対象は人間や車など既存の検出モデルが利用可能なものだけでなく様々な対象が想定される。興味の対象となり得る全てのオブジェクトを検出可能に学習した検出器をあらかじめ作成することは現実的ではない。

## 3. 提案手法

### 3.1 提案手法概要

提案手法は、メガネ式アイトラッカーにより取得した一人称視点映像と視線情報を入力として、興味対象を検出し、興味対象画像を提示するシステムである。

図1に提案手法の概要を示す。本研究の提案手法は、大きく3つの手順によって構成されている。

(1) 連続的な注視の検出：撮影した一人称視点映像と視線情報を入力し、画像特徴に基づく注視状態の検出を行い、連続する注視を検出する。ここでは、単一の対象に対する連続的な注視が取得される。

(2) 断続的な注視の検出：手順1で検出した複数の連続的な注視の中から、同じ対象を注視している組み合わせを検出し、注視時間、注視回数を取得する。

(3) 興味対象の提示：注視時間、注視回数を基に興味対象を検出し提示する。

### 3.2 連続的な注視の検出

メガネ式アイトラッカー環境における連続的な注視は、Steilらのコンセプトを利用する[5]。そのコンセプトは、連続する2フレームにおいて視線位置周辺領域をパッチとして切り出し、類似度を求める。画像が類似している場合には同一対象を注視していると判断する。この処理を繰り返し、150msec. 以上の注視の継続を検出した場合に、連続的な注視としている。

本研究において上記コンセプトを実装するために、パッチ

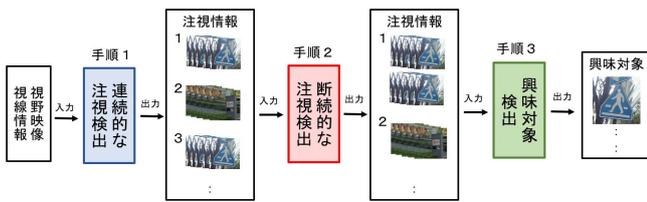


図1 提案手法概要図

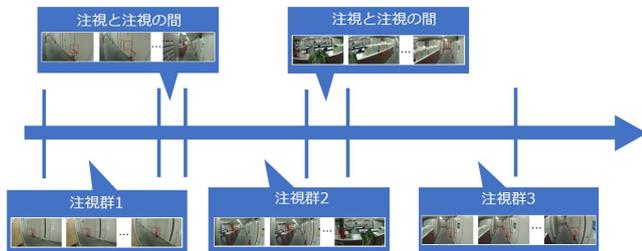


図2 連続的な注視の検出例

毎の画像特徴として、ImageNet [9] で学習された VGG-19 [10] を用いる。VGG-19 にパッチを入力し、クラスを出力する前のデータを特徴量として抽出する。

本ステップの出力例を図2に示す。時間軸上の、ある区間が注視群1、次の区間が注視群2と、複数の注視群に分かれて検出されている。各注視群は、連続的に同一オブジェクトに注視が集まっており、複数枚の類似した画像から構成されていることが分かる。

### 3.3 断続的な注視の検出

連続的な注視の検出により、複数の連続注視群に分割されている(図2参照)。断続的な注視の検出では、連続注視群間で同一のオブジェクトを検出することで、注視時間と注視回数を求めることを可能にする。

断続的な注視の検出には、Re-identificationのアプローチを用いる。従来のRe-identificationのアプローチでは、対象を検出する検出器の存在を前提としており、幅広い対象に注視が向く可能性があることから困難と考えられていた。しかしながら本研究では、前段階の処理により、注視群に分割されている。各注視群は、注視対象オブジェクトを含む複数の画像からなる。このことから、注視群の画像を用いることによりRe-identificationを実現することを狙う。

ここでひとつ課題がある。それは、注視先の座標はアイトラッカーで取得できるものの、その範囲、注視対象の大きさは不明であるということである。直接的な関係はないが、Steilらが、2つの時刻における注視点の類似度を評価する際には、視線位置の周辺を一定サイズのパッチで切り出している。これに対して我々は、注視対象までの距離により注視領域を変更する。同一の物体であっても遠方に存在する場合、見える大きさとしては小さくなる。近づくに従って、大きく見えてくる。ただし、一定範囲を超えて近くなると、より細かいパーツに注目が移るようになるため、一定以上大きくしないこととした。この様子を図4に示す。

なお、奥行き情報の取得には、単眼深度推定モデルである



(a) 入力画像 (b) 出力画像

図3 MiDaSによる深度推定

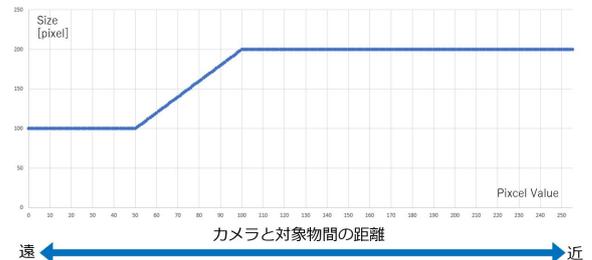


図4 注視領域の推移

MiDaSを用いる [11]。MiDaSは単眼カメラの画像から深度を推定することができる機械学習モデルである。図3にMiDaSによる深度推定の例を示す。図3(a)の画像を入力すると図3(b)のような深度マップを表す画像が出力される。深度マップの視線位置周辺の画素値に応じて注視領域の大きさを変更する。本研究では、一人称視点映像が1280×720pxの解像度のカメラを用いる。注視領域は正方形とし、正方形の1辺の長さをSize、深度マップの視線先の画素値をPixelValueとしたとき、図4に示すようにPixelValueが

100以上の時は

$$Size = 200$$

50未満の時は

$$Size = 100$$

50以上100未満の時は

$$Size = 200 * PixelValue * 0.01$$

とした。

図5に注視領域の変更例を示す。図5(a)の緑円が視線位置で、図5(b)(c)の青い領域が注視対象とする領域である。5(b)では広い領域を注視対象としているのに対して、5(c)では奥行きを考慮した小さい領域を注視対象としていることが分かる。

以上により、連続的な注視群として分割された画像において、奥行きを考慮した注視領域を得ることができる。ここから、Re-identificationの考え方に基づいたアプローチにより、対象ごとにクラスタリングする手法について述べる。なお、以下の処理は未だ検討中であり、本稿では3つの手法を並記する。

#### 3.3.1 手法1: YOLOによる物体認識

手法1では注視対象オブジェクトを含む画像が複数枚あることを利用し、現在の連続注視群を検出可能なアドホックなネットワークを学習することで、学習ベースのRe-identificationを実現することを狙う。連続注視群のフレームにおける、奥行きに合わせた注視領域を検出対象として、アノテーションし、これ



図5 注視領域の変化例

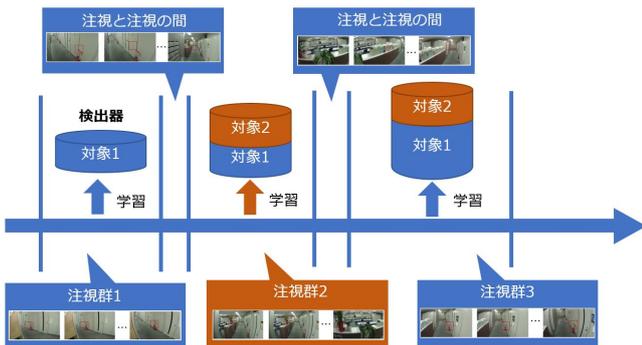


図6 手法1の流れ

らを教師データとして物体検出アルゴリズム YOLO [12] を用いて学習することにより、注視対象物の検出器を作成する。すなわち本手法1では、連続注視群のクラスに分類するネットワークモデルを1つ学習する。

手法1の処理の流れを図6に示す。まず、各連続注視群から作成した教師データを YOLOv5 を用いて学習し、注視対象物の検出器を作成する。注視時間が短い連続注視群は学習枚数が少ないこと、連続注視群のクラス数の増大に伴って学習時間が爆発的に増加することを考慮し、学習モデルには yolov5 の中でも演算負荷の小さいモデルである yolov5s を使用し、学習回数は 50epoch とした。

次に、検出器を注視群に対して適用する。視線位置周辺に物体が検知された場合、断続的な注視として検出し、検出した対象と同じクラスに注視群を教師データとして追加し再度学習を行う。物体が検出されなかった場合は新たなクラスとして再度学習を行う。この処理を繰り返すことにより、同一のオブジェクトを注視している連続注視群の組み合わせを検出する。

### 3.3.2 手法2: YOLO による物体認識 (専用検出器)

手法1では連続注視群のクラス数の増大に伴って学習時間が爆発的に増加し、長時間の映像に対して適用が困難となる可能性がある。そのため、手法2では 3.3.1 の手順において新規注視対象を検出した場合、専用の検出器を作成することにより実行時間の短縮を狙う。

手法2の処理の流れを図7に示す。新規対象が検出された場合に、図7に示す通り、新しい専用の検出器を作成する。これ

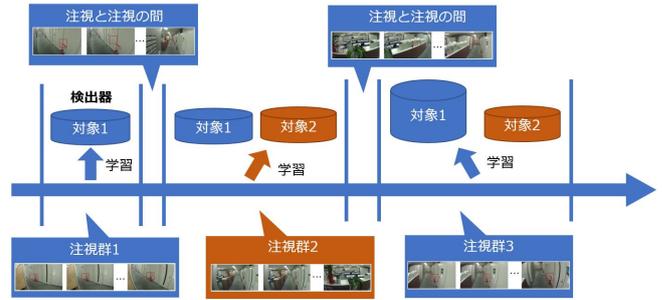


図7 手法2の流れ

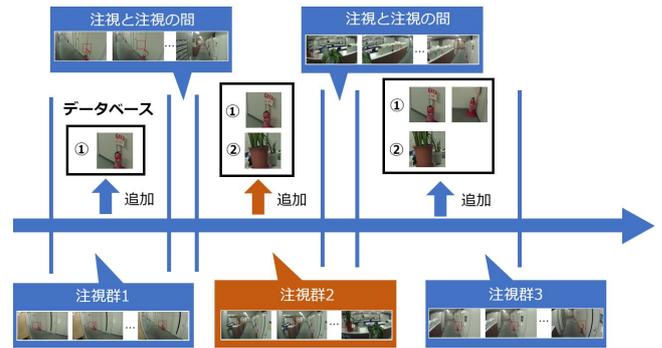


図8 手法3の流れ

により、図6に記載の、既存の検出器に学習データを追加し、再学習を行い、1つの検出器を作成する手法1に対して、大きな学習時間削減効果が期待できる。

### 3.3.3 手法3: 画像類似度

手法3は、パッチ間のシンプルなマッチングにより、同一対象か否かを判断する。この処理の流れを図8に示す。連続注視群フレームから奥行に応じて定められた注視領域をパッチとして切り出しデータベースに対して検索する。パッチと類似する画像がデータベースに存在すれば断続的な注視として検出し、データベースの検出したクラスにパッチを追加する。なければ新規クラスとして追加する。この処理を繰り返すことにより、同一のオブジェクトを注視している連続注視群の組み合わせを検出する。

なお、連続注視群に含まれる、誤検出、ぶれ、その他のノイズなど様々な、外れ値となるような画像の影響を低減するため、上記処理は、連続注視群から選択した代表画像を対象に行うものとする。

### 3.4 興味対象の提示

最後に、検出された注視回数、注視時間を基に興味対象を検出し、興味対象画像を提示する。

一般に複数回注視した対象や、長時間注視した対象は興味を持って見ていることが考えられる。そのため本研究では、2回以上の注視、1.2sec 以上の連続した注視、もしくはその両方を満たす注視が発生した際の注視対象を興味対象とする [13]。

興味対象を分かりやすく提示するために、注視群を代表するような画像を1枚選択し、代表画像を興味対象画像として提示する。図9に興味対象の提示例を示す。図9に示すように、標識、信号、看板など移動中に興味を持ったと考えられる対象を

提示している。この提示画像を確認することにより、移動中に興味を持った対象をを振り返ることができる。

## 4. 実験

### 4.1 実験手順

提案手法の有用性を確認するために評価実験を行った。評価実験はメガネ式アイトラッカーを着用した被験者が街中を歩き計測したデータを用いた。視線の計測には図 10 の Pupil Labs 社の Pupil Core を使用した [14]。World Camera と Eye Camera により眼球運動と一人称視点映像を撮影する。一人称視点映像の解像度は 1280 × 720px, フレームレートは 60fps で撮影を行い、約 1 分 30 秒の動画を 12 本撮影した。World Camera 画像上での注視点情報の取得には、同社の Pupil Player を利用した。評価実験では、興味対象を正しく検出できているかを確認するため、図 11 のような計測した一人称視点映像に視点位置 (緑) を付与した映像を確認し、映像の中で興味があると判断した対象と、提案手法によって検出した対象 (図 9 参照) との一致率を調べた。また、各断続的な注視検出手法の検出精度、実行時間の違いについても調べた。

### 4.2 実験結果

はじめに、断続的な注視の検出結果を述べる。断続的な注視は全体で 178 回発生した。表 2 に各手法における断続的な注視の検出回数と検出率を示した。表 2 に示すように、手法 1 が最も多くの断続的な注視を検出できたことを確認した。

次に、興味対象の検出結果について述べる。興味対象は全部で 70 個観測された。提案手法の処理の都合上、同じ対象であるにも関わらず、別の興味対象と検出し、検出が重複することがある。図 12 に検出の重複例を示す。図 12(a) は興味対象 1, 図 12(b) と興味対象 2 と検出された例である。図 12(a)(b) はともに同じ興味対象であるにも関わらず、別の興味対象として検出されている。このような際、本稿では、検出対象の種類数と、提案手法によって検出された興味対象の総数をそれぞれを検出件数、総検出数と定義する。図 12 の例だと、検出件数は 1, 数総検出数は 2 となる。

表 2 に各提案手法における興味対象の検出件数と、興味対象の内検出できた割合を表す検出率を示した。表 2 から分かるように、手法 1 では検出率が 0.87 を示しており、おおよその興味対象を検出できていることがわかる。このことから提案手法の有用性を示したと考える。手法 1 が他の手法に比べて高い検出率となったのは、断続的な注視を多く検出したことにより、注視回数に基づいた興味対象の検出数が増加したことが要因だと考えられる。

表 3 に各手法における総検出数と、検出が興味対象と適合した割合を表す適合率を示した。表 2 に示すように、適合率は手法 1 が最も低い結果となった。手法 1 は断続的な注視を多く検出したが、進行方向や建物の壁などの興味対象とは考えられないが複数回見た対象を断続的な注視と判断し、興味対象と検出したため、適合率が他の手法に比べて低下した。

表 4 に各提案手法における映像と視線情報を入力してから興味対象検出までの実行時間の平均を示した。表 4 から分かるよ

表 1 各手法の断続的な注視検出率

|      | 手法 1 | 手法 2 | 手法 3 |
|------|------|------|------|
| 検出回数 | 145  | 68   | 75   |
| 検出率  | 0.81 | 0.38 | 0.42 |

表 2 各手法の興味対象検出率

|      | 手法 1 | 手法 2 | 手法 3 |
|------|------|------|------|
| 検出件数 | 61   | 52   | 52   |
| 検出率  | 0.87 | 0.74 | 0.74 |

表 3 各手法の適合率

|      | 手法 1 | 手法 2 | 手法 3 |
|------|------|------|------|
| 総検出数 | 89   | 84   | 85   |
| 適合数  | 67   | 65   | 71   |
| 適合率  | 0.75 | 0.77 | 0.84 |

表 4 各手法の実行時間の平均

|          | 手法 1  | 手法 2  | 手法 3 |
|----------|-------|-------|------|
| 実行時間 (秒) | 36843 | 15789 | 1213 |

表 5 各手法の検出重複数

|      | 手法 1 | 手法 2 | 手法 3 |
|------|------|------|------|
| 2 重複 | 5    | 7    | 8    |
| 3 重複 | 0    | 0    | 4    |
| 4 重複 | 0    | 1    | 0    |
| 5 重複 | 0    | 0    | 1    |

うに、実行時間は手法 3 が一番短い結果となった。手法 3 は連続注視群毎に、データベースに画像を追加するだけなのに対し、手法 1, 2 は連続注視群毎に学習を行う必要があるため実行時間に大きな違いが出た。また、手法 2 は手法 1 に比べて実行時間が半分以下になった。このことから、個別の検出器を作成することにより、学習時間の短縮が行えたといえる。

最後に、表 5 に各手法における興味対象の検出重複数を示す。表 5 から分かるように、手法 1 が断続的な注視の検出率と同様に最も優れている結果となった。

## 5. 今後の課題

今回は被験者の注視領域を考慮し、深度情報に基づいたサイズでアノテーションを行った。しかし、物体検出器の作成に用いる教師データは対象物の外接矩形をアノテーションすることが望ましい。そのため今後は、視線位置周辺の構造的な意味情報を取得し、対象物の外接矩形でアノテーションすることにより、断続的な注視の検出精度の向上を目指していきたい。

また本研究では、画像類似度に基づく断続的な注視検出の際、画像特徴量抽出に ImageNet で学習された VGG-19 を用いた。しかしこの特徴量が今回のタスクにおいて適していたかは検証できていない。そのため、今後は他により適した特徴量抽出手法があるかを検証していく必要があると考えられる。

## 6. おわりに

本研究では、移動環境において興味対象を検出する手法を提

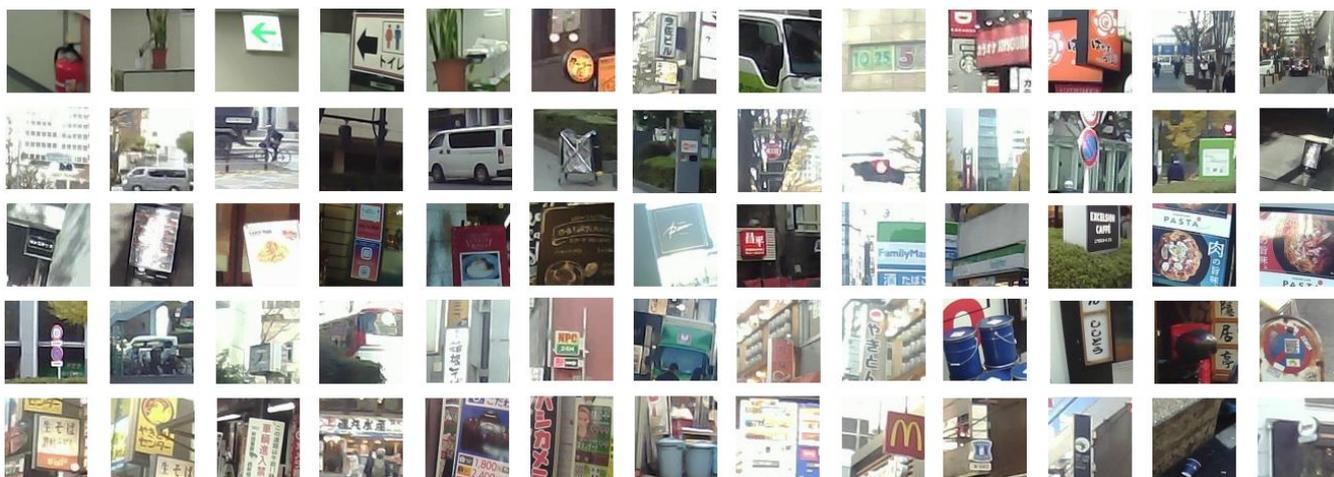


図9 興味対象の提示例

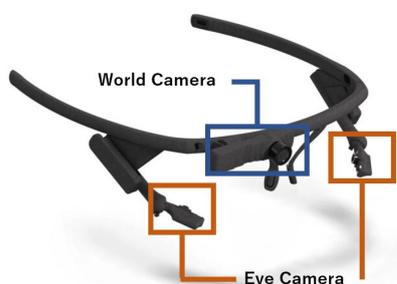


図10 実験に使用したアイトラッカー



図11 視線位置を付与した映像



(a) 興味対象 1 (b) 興味対象 2

図12 重複した検出の例

案した。従来の連続的な注視を検出する手法に加えて、断続的な注視を検出する手法を提案し、注視時間と注視回数から興味対象の検出を行った。実験の結果、学習ベースの Re-identification を用いた手法では、約 87% の興味対象を検出し、提案手法の有用性を示した。今後は視線位置周辺の構造的な意味情報を取得する手法や、適切な画像特徴の抽出手法を検討し、検出精度を向上させていきたい。

## 文 献

- [1] 岡本 康太郎, 内海 章, 山添 大丈, “視線計測を用いた商業施設における来店者行動の分析” 電子情報通信学会技術研究報告, pp.1-6, No.109, Vol.281, 2009.
- [2] 吉高淳夫, 西田 謙太郎, 平嶋 宗, “絵画鑑賞時の眼球停留の時間的な発生頻度に着目した注目状態の検出” 情報処理学会論文誌, pp.1267-1476, Vol.50, No.5, 2009.
- [3] 松延拓生, “ウェブサイトのユーザビリティ評価のための視線分析方法” ヒューマンインタフェース学会論文誌, pp.293 -302, Vol.16, No.4, 2014.
- [4] 小松 知世, 長竹 祐美子, 小館 亮之, 茂木 学, 深山 篤, 下村 道夫, “視線解析に基づく写真閲覧時の潜在的嗜好推定方法の提案” 電子情報通信学会論文誌 D, pp.1689 -1692, Vol.197-D, No.12, 2014.
- [5] Julian Steil, Michael Xuelin Huang, Andreas Bulling, “Fixation detection for head-mounted eye tracking based on visual similarity of gaze targets,” Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications, pp. 1-9, No.23, 2018.
- [6] 針ヶ谷 勇星, 白山 晋, “自動車運転者の行動分析のための注視情報抽出法” 人工知能学会インタラクティブ情報アクセスと可視化マイニング研究会, 第 24 回研究会研究発表予稿集, pp.50-57,2020.
- [7] Mang Ye, Jianbing Shen, Gaojie Lin, Tao Xiang, Ling Shao, Steven C. H. Hoi, “Deep Learning for Person Re-identification: A Survey and Outlook,” Pattern Analysis and Machine Intelligence, pp.2872-2893 Vol.44, No.6, 2022.
- [8] Xiangyu Zhu, Zhenbo Luo, Pei Fu, Xiang Ji, “VOC-ReID: Vehicle Re-Identification Based on Vehicle-Oriented-Camera,” Computer Vision and Pattern Recognition Workshops, pp.602-603, 2020.
- [9] Olga Russakovsky, Jia. Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, Li Fei-Fei, “ImageNet Large Scale Visual Recognition Challenge,” International Journal of Computer Vision, pp.211-252 Vol.115, 2015.
- [10] Karen Simonyan, and Andrew Zisserman, “Very deep convolutional networks for large-scale image recognition,” The 3rd International Conference on Learning Representations, 2015.
- [11] René Ranftl, Katrin Lasinger, David Hafner, Vladlen Koltun, Konrad Schindler, “Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-shot Cross-dataset Transfer,” Computer Vision and Pattern Recognition, pp.1623 - 1637, Vol.44, No.3, 2020.
- [12] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” Computer Vision and Pattern Recognition, pp.779-788, 2016.
- [13] 戸田 真人, 鷲見 和彦, 松山 隆司, “装着型視覚センサを用いた注目対象映像の獲得と理解” MIRU2005, IS4-150, pp.1388-1395
- [14] Pupil Labs, “Pupil core” <https://pupil-labs.com/products/core/>, (参照 Feb.20, 2023)