

## 視野映像のセグメンテーションによる注視オブジェクトと 講義理解度の関係調査

### Investigation of Relationship between Gazed Objects and Lecture Comprehension

田丸 真彩<sup>1)</sup> 原田 翔太<sup>1)</sup> 方 昱<sup>2)</sup> 満上 育久<sup>3)</sup>  
Maaya Tamaru Shota Harada Yu Fang Ikuhisa Mitsugami

#### 1 はじめに

同一の講義を受講していても、受講生ごとに講義への理解度は異なる。これは、集中力・記憶力など受講生が元来持つ能力の違いが原因と考えられるが、それに加えて授業を受ける姿勢や態度、授業時間中の行動など「講義への取り組み方」の違いが理解度に影響を与えていると考えられる。実際、椋木ら [1] は、大学の講義における受講生の「振る舞い」と「理解度」には関係性があると報告している。また、「講義への取り組み方」のなかでも「注意の向け方」も理解度に関係があると報告されており、藤好ら [2] は英語のペーパーテスト受験時の注視行動とテストの点数には関係があると示し、視線移動情報とテストの正答数を用いることで、テストの点数を推定する手法を示した。黒田ら [3] は、視線計測のできる VR システムを構築し、講義動画内の講師の動作と受講生の視線を定量的に計測した。そして視線情報から受講生の理解度の高低を推定する手法を提案し、推定可能であることを実験的に示した。理解度を推定することができれば、授業の質の改善や向上に繋がり、受講生の学力向上に役立つ。そのためには、理解度とみている物体、注視オブジェクトにはどのような関係があるのかを調査する必要がある。

そこで本研究は、アイトラッカーを用いることで現実空間での講義中の受講生の視線情報を計測し、注視オブジェクトと理解度の関係を調査することを目的とする。黒田らの研究と同様に講義受講中の視線の定量計測を行うが、視野映像に対してはインスタンスセグメンテーションを施すことで、注視点の位置そのものではなく、注視点がどの物体に向いていたか(講師、スクリーンなど)に注目し、視線がどの物体を見ていたかを判定することで注視オブジェクトの時系列を獲得する。これらを理解度が高い受講生と低い受講生から取得し、両者にどのような違いがあるかを調査する。

#### 2 実験方法

##### 2.1 講義実施方法

本実験では講義動画をプロジェクタで壁面に投影したものをアイトラッカー着用の受講生に視聴させる形でデータを収集する。一斉授業を実施しデータ収集を行うことはデバイスの個数やセッティングの問題上難しく、少人数講義を複数回行う方法は、実験環境の一貫性を保証できないためである。この方法により、各受講者が視聴する内容が完全に同一な講義を複数回行うことができ

1) 広島市立大学

Hiroshima City University

2) ホンダ・リサーチ・インスティテュート・ジャパン

Honda Research Institute Japan

3) 立命館大学

Ritsumeikan University

るのでアイトラッカーを多数準備しなくても、多数の受講生の注視情報のデータを収集することができる。

##### 2.2 注視オブジェクトの認識法

本研究では、視野映像中の注視点座標を講義動画自体の注視点座標に変換し、講義動画そのものにインスタンスセグメンテーションを施す。そして、変換した注視点座標がどの領域に分類されるか判定することで注視オブジェクトを出力する。本実験は 2.1 節で説明した講義実施方法によりデータを収集しており、投影した映像を受講生が視聴しやすいよう部屋の照明を暗めにしていないため、投影されている講義動画がアイトラッカーの視野映像中で白飛びしやすくなる。したがって、視野映像に対するインスタンスセグメンテーションを行うと、講義動画内のオブジェクトを安定に検出できないという問題が生じるためである。

受講生の講義動画中での注視点座標を獲得するために、AR マーカを用いて視野映像中の注視点座標を講義動画自体の注視点座標に変換する。座標変換のイメージ図を図 1 に示す。アイトラッカーで計測される視線情報は、視野映像中における正規化 2 次元座標(視野映像の画面左上が (0,0)、右下が (1,1))として得られる。これを、視野映像の解像度 1920 pixel × 1080 pixel に合わせて縦横をスケールすることで、視野映像中での注視点座標を得る。視野映像内の講義動画の座標を獲得するために、あらかじめ講義動画周囲に AR マーカを貼布してデータ収集を行う。講義動画の投影位置は固定しているので、視野映像内の AR マーカを検出し、各マーカの左上の座標から講義動画の各頂点の座標を算出する。この

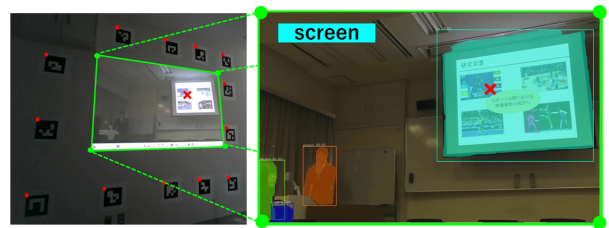


図 1 座標変換のイメージ図

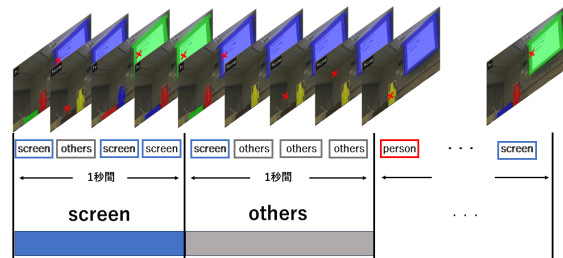


図 2 注視オブジェクトの可視化のイメージ図

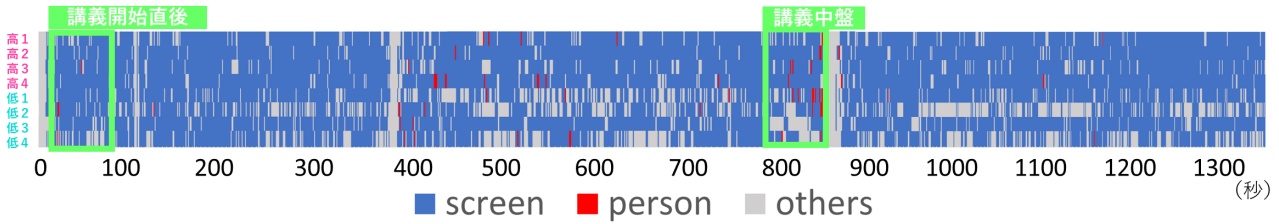


図3 講義動画全体の注視オブジェクト

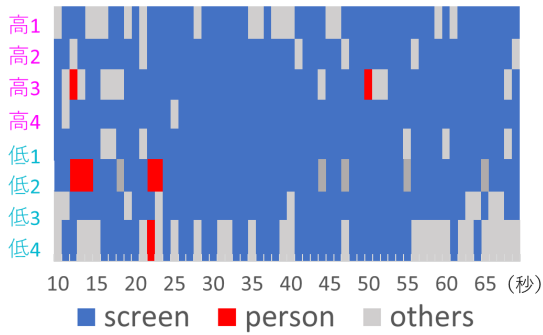


図4 講義開始直後

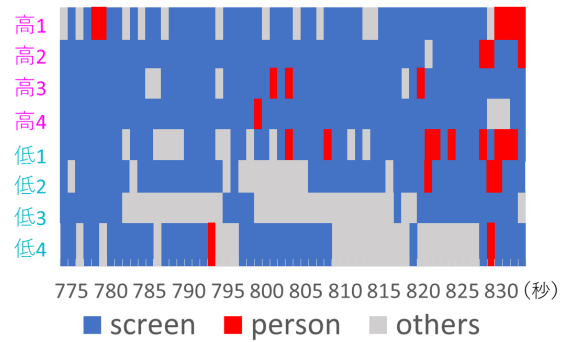


図5 講義中盤

処理で獲得した視野映像中の講義動画の各頂点の座標と、講義動画自体の各頂点の座標の平面射影変換行列を求め、それと視野映像中の注視座標の積を求めることで、視野映像中の注視点座標から講義動画内の注視点座標を求める。

アイトラッカーで計測した注視点座標はすべてが正確であるとは限らず、また、サッカードの影響を抑えるため、注視オブジェクトは一秒間で最も多く分類されたクラスとする。イメージ図を図2に示す。また、各クラスに色を設定することで、一秒ごとに出力された注視オブジェクトの可視化を行う。

### 3 実験結果

広島市立大学情報科学部の学生15名から2節で説明した実験方法によりデータを収集・分析した。なお、講義動画に対するインスタンスセグメンテーションには、PixelLib [5]、アイトラッカーはTobii Pro Glasses 3 [4]を使用した。受講生には講義動画視聴後、理解度を判定するために、講義内容に関する筆記試験(満点は22点)を課した。そのスコアの上位4名を理解度が高い群、下位4名を理解度の低い群とした。計8名の講義全体の注視オブジェクトを時系列を図3に示す。図3を観察すると、注視オブジェクトに特徴があらわれた区間がいくつか存在したので、講義開始直後と講義中盤の2つの区間を例に挙げて考察する。

講義開始直後では、理解度が高い群と低い群の受講生で注視オブジェクトに違いは表れず、どちらも講義スライドをよく注視していた。この区間の注視オブジェクト時系列を図4に示す。この区間は講義開始直後であるため、受講生たちは高い集中力が保たれており、講義に無関係なオブジェクトを注視しなかったと考えられる。

また、講義中盤では、理解度が高い群の受講生は講義スライドをよく注視していたが、理解度の低い群の受講生は講義に無関係なものを注視している時間が理解度が高い群よりも多いことが確認できる。この区間の注視オブジェクト時系列を図5に示す。この区間は講義スライ

ドに「まとめ」が表示され、理解度の低い群の受講生は「講義が終了する」と気が抜け、集中力が低下したと考えられる。

### 4 おわりに

本研究では、講義中の受講生の注視オブジェクトから理解度を推定することを目的として、アイトラッカーを用いて視線情報を計測し、インスタンスセグメンテーションと組み合わせることで講義への理解度の高低と注視オブジェクトの関係を調査する手法を提案した。そして、実際にこの手法を用いて15名の講義動画視聴時の視線情報を収集し、理解度が高い群の受講生と低い群の受講生の注視オブジェクトを時系列で観察した。その結果、2つの群で異なる傾向がみられる場面があることが確認できた。今後は、注視オブジェクトの判定方法に工夫を施し、より信頼性の高い注視オブジェクト時系列を用いて注視オブジェクトと理解度の関係性を考察し、またVRを用いることでより没入感のある環境で実験を行いデータを集める予定である。

### 謝辞

本研究は、JST、未来社会創造事業、JPMJMI22J1の支援を受けたものである。

### 参考文献

- [1] 椋木雅之, 美濃導彦, 京都大学学術情報メディアセンター, 「講義室での受講生の振る舞い観測と理解度推定の研究」, 人工知能学会全国大会, 2012年.
- [2] 藤好宏樹, 吉村和代, Kai Kunze, 黄瀬浩一, 「英文問題解答時の視点情報を用いた英語能力推定法」, 電子情報通信学会技術研究報告, Vol.115, No.22, pp.49-54, 2015年.
- [3] 黒田明日香, 満上育久, 「VR 講義動画視聴における理解度と注視分布の関係の分析」, 第18回情報科学技術フォーラム, 2019年.
- [4] Tobii Technology K.K., "Tobii Pro Glasses 3," <https://www.tobii.com/ja/products/eye-trackers/wearables/tobii-pro-glasses-3>
- [5] Ayools Olafenwa, "PixelLib," <https://github.com/ayoolaolafenwa/PixelLib>