

研究論文

アイトラッキングカメラを用いた授業研究の基礎技術

谷 口 正 明

Fundamental Techniques for Lesson Study by Using the Eye-tracking Camera

Masa-aki Taniguchi

1. はじめに

授業研究^[1]では、授業内の教師及び生徒の授業中の言葉のやり取りを記録（逐語記録）し、教師の言葉かけ、教師の意思決定について分析・考察することが行われてきた。これを精度よく行うには、できるだけ多くの情報を記録することが大切である。

平山・後藤は「より精緻な授業記録は、子どもの変容状況を継続的総合的に把握し、授業過程を検討するために必要である」と主張してきた。そして、より詳細な授業記録を残すために、21世紀初頭から複数台のビデオカメラを用いて授業を撮影し、映像記録を画面合成器で合成するマルチ授業映像システムの開発を行ってきた。そして、7年ほど前から、眼鏡型の軽量なアイトラッキングカメラ（Tobii 社製 Tobii Pro Glasses 2）を用いて授業者の注視点記録し、これをマルチ授業映像システムの一部として取り入れてきた。^[2, 3] また、アイトラッキングカメラの映像を基にして、被験者の注視点の分布をヒートマップ化して、視野の広さ等^[4, 5]を研究してきた。しかし、眼鏡型アイトラッキングカメラの背景は、時々刻々と変化するため、授業中に教師が教室の何をどの程度見ているかは明らかにすることはできなかった。

私たちは、最近、商品マーケティングの分野で使われている Real-World Mapping (RWM) という、動画上の注視点を静止画に対応付けする手法を授業研究に応用することができるようになった。また、Areas of Interest (AOI) と言って、RWM で用いられる背景画像の中に「興味のある領域」を指定し、注視点がこの領域内に入っているかどうかを判定することができるようになった。これにより、教師の一人称視点映像の背景を固定するとともに、授業中に何をどの程度の時間見ているかを明らかにすることができるようになった。これを用いて、現在は一人称視点の映像分析を行っている。^[6, 7, 8]

この論文では、RWM と AOI に関する技術情報や分析方法をまとめておき、今後の分析を行なう際の基礎を記しておくこととする。なお、この論文の図をカラーで見ることができるサイトを用意した。^[9] これも参照されたい。

2. 授業動画の撮影方法

私たちは通常、授業中の生徒の状態を撮影するために、コンパクトビデオカメラ GoPro 社製 Hero7 Black と SONY 社製のウェアラブル 4K カメラを、教室の前の黒板に取り付けて教室全景の動画を撮影している。(図1) このようにすると、1台のカメラで教室の児童や生徒全員を撮影することができる。後述するように、この動画を Real-World Mapping を行う際の静止画を切り出すのに用いることもある。また、これらのカメラの解像度は 4K (3840 × 2160) である。これだけ詳細な画質であると、事後に動画を拡大して一人ひとりの授業中の行動や表情を追跡することができる。加えて、教室の後ろには、主に教師と黒板を録画するために、通常のビデオカメラで撮影している。これらとアイトラッキングカメラで撮影した一人称視点の動画をシンクロさせた上で合成し、マルチアングルのビデオを作成している。



図1. GoPro 社製 Hero7 Black と SONY 社製のウェアラブル 4K カメラ

私たちは、教師の一人称注視点を撮影するために、アイトラッキングカメラ (Tobii 社製 Tobii Pro Glasses 2) を用いている。このカメラは、45g の眼鏡型カメラと、データを保存する SD とバッテリーからなる 312g の記録ユニットからなっている。眼鏡型カメラには、背景を記録するカメラと、瞳孔を記録する複数のカメラが付いている。これにより、被験者が背景のどこを見ているかを記録することができる。これら2つのユニットはケーブルで結ばれてお



図2. 眼鏡型アイトラッキングカメラ Tobii 社製 Tobii Pro Glasses 2

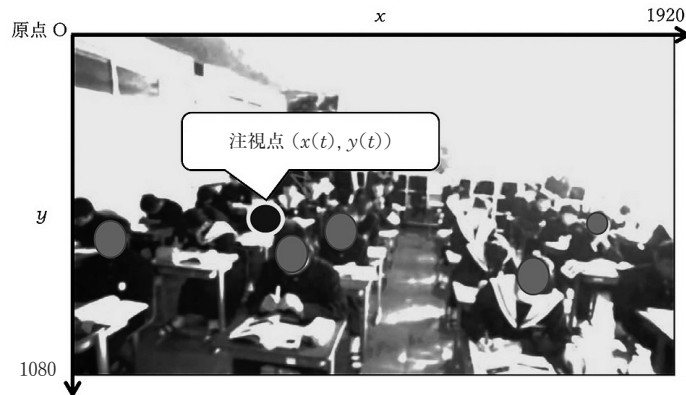


図 3. Tobii Pro Glasses 2 の座標。赤い丸は注視点を表す。

り、記録ユニットは被験者のポケットに入れるか、ベルトに吊り下げようになっている。

図 3 は、Tobii Pro Glasses 2 の画面と座標軸、注視点の座標を記したものである。Tobii Pro Glasses 2 の動画の解像度は、Full HD (1920 × 1080) である。座標の原点は画面の左上であり、 x -軸は横方向、 y -軸は縦方向である。赤い丸は注視点、つまり、ユーザが現在見ている箇所を表す。赤い円の半径は、見ている時間に比例する。注視点のデータは、時間 t に依存する $x(t)$ と $y(t)$ の組、2次元ベクトルデータ $(x(t), y(t))$ として記録される。

Tobii Pro Glasses 2 で撮影した動画データは、Tobii 社製 Tobii Pro Lab によって認識され、エクスポートすることによって取り出すことができる。エクスポートする際は、テキストファイル (TSV, Tab-Separated Values, タブで区切られた値)、あるいは、Excel ファイルの標準形式 (xlsx) で出力することができる。前者は無圧縮であるのに対し、後者は Zip 圧縮されているため、ファイルサイズの節約の意味では後者で出力するのが好ましい。また、データの分析を Excel で行うのであれば、後者を選択するのが良いであろう。

3.Real-World Mapping を用いた注視点の分析方法

Real-World Mapping (RWM, 実世界への写像) とは、動画上の注視点を静止画上にマッピング (対応付け) する方法である。これは、Tobii 社製の分析ソフト Tobii Pro Lab を用いて行うことができる。

Tobii Pro Lab では、自動的に RWM を行ってくれる (Assisted Mapping) が、手動で行うこと (Manual Mapping) もできる。Tobii Pro Lab の RWM の自動認識は、動画と静止画の角度などが違うと失敗することがある。(図 4 のようなケースでは、おそらく自動認識できない。) このような場合は手動で対応付けを修正することが必要である。

Tobii Pro Lab では、RWM を行った際も、Gaze Plot や Heat Map を作成することができる。



図4. RWM のイメージ。左の動画から右の静止画へ注視点を対応付けしている。

この一例を図5と図6に示す。Gaze Plot とは注視点のバブルチャートである。バブルの半径は注視時間に比例している。Heat Map とは、注視している頻度に応じて色分けをしている図である。

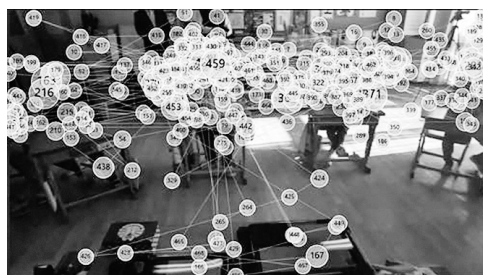


図5. Gaze Plot の例



図6. Heat Map の例

4. 対象とした授業

以下の授業の教師の一人称視点をこの論文の分析に用いた。

- 2020/2/5 2 限目 Y 先生の授業
- 特別支援学級 生徒数4名、プログラミング教育
- 冒頭4分弱（約232秒）の導入部分

この授業を分析に用いたのは、クラスの人数が4人と少ないため、被験者（教師）がどこを見ているかを判定しやすいこと、また、Excel を用いて手動の AOI を行いやすいためである。

この研究では、授業内容を研究するのではなく、RWM や AOI をどの程度正確に行うことができるかを評価することが目的である。このため、この授業冒頭の約4分間に限った。授業最初の4分間で、この教員は教卓付近に立ち、授業で行うこととそれぞれの目標を児童たちに伝えた。このため、大きく視野を変えることはなく、一人称動画の背景画像はほとんど RWM の静止画と同じであった。

5. Real-World Mapping を行う際の基礎実験

(1) Real-World Mapping の背景画像の違いによる自動認識率の変化

ここでは、RWM を行う際に背景画像の違いによって、Tobii Pro Lab を用いた自動認識率が変わるのかどうかを確認する。この実験で用いた背景画像（静止画）は、アイトラッキングカメラで撮影した動画から切り取って用意した静止画（図 7）と、黒板に設置した SONY 製の 4K ウェアラブルカメラで撮った動画から切り取った静止画（図 8）の 2 種類である。前者はアイトラッキングカメラで撮影した動画から切り取ったものであるため、当然ながら動画の背景動画とよく似ている。今回は子どもの数が 4 人と少ないため、一つの静止画に全員の子どもの顔が写った静止画を作ることができた。



図 7. 一人称動画から切り取った静止画



図 8. 黒板に取り付けたカメラの静止画

私たちは、それぞれ図 7 と図 8 を RWM の時の静止画として Tobii Pro Lab に読み取らせ、RWM の自動認識をさせた。RWM の自動認識はしばしば失敗する（動画とは違う点に対応付けをする）ため、手動で修正する必要がある。授業最初の約 4 分間で、自動認識できた点と手動認識せざるを得なかった点のデータ数を表 1 に記す。

表 1 によると、図 7 の方が図 8 よりも自動認識率が高い。また、図 8 では、自動認識で注視点が人間の半身分位ずれていたため、大規模な手修正（手動認識）を行った。このため、自動認識のデータが減り（約 2000 個減った）手動認識のデータ数が増えた。

表 1. 図 7 と図 8 を背景静止画とした時の自動認識、手動認識のデータ数

	図 7	図 8
自動認識	5760	3758
手動認識	1112	2688
合計	6872	6446

ところで、この表の「自動認識」は Assisted mapping gaze point X のデータの個数、「手動認識」は Manually mapped gaze point X のデータの個数である。「合計」は両者の和である。この表から、一人称動画から切り取った静止画（図7）の場合は、自動：手動 = 6:1 であった。一方、黒板に取り付けたウェアラブルカメラから切り取った静止画（図8）の場合は、自動：手動 = 1.4:1 であった。このように、後者の自動認識の成功率は大幅に減少した。

普通教室では、子どもの人数が多いため、クラスの全員の子どもを一つの写真の中に入れるためには、教師の一人称動画から切り取った複数の静止画から、一つの静止画を合成する必要がある場合がある。つまり、図7のような写真を動画からワンショットで取り出すことは難しい場合がある。一方、後者は、黒板の上部に取り付けたウェアラブルカメラの動画から切り取っているため、やや俯瞰した図となっている。そして、普通教室においても、隅の方の席に座っている子も見切れず、クラスの全員が入った写真となっていることが多い。ウェアラブルカメラから切り出した静止画は、自動認識率は下がり手作業が増えるが、それでも有用な場合はある。

図7、図8を背景の静止画として、RWMを行い注視点の Heat Map を作ったものが、それぞれ図9と図10である。これらを比較すると、全体的な傾向は変わらないことがわかる。



図9. 一人称動画の静止画上の注視点



図10. 黒板カメラの静止画上の注視点

(2) Excel の関数を用いた手動分析と Areas of Interest

Tobii Pro Lab には、Areas of Interest (AOI) という機能がある。これにより、被験者が画面上のある領域を見ていたかどうかを判定することができる。図5に AOI の一例を示してある。静止画に多角形（だ円、長方形を指定することも可能）の領域を手動で指定し、この領域内の注視点が入った場合には、被験者がこの領域にある物体を「注視している」と判定する。Excel の関数を用いてもこれと同等の分析を行うことはできるが、AOI を使うと作業を自動化することができる。

また、Tobii Pro Lab は、注視点のデータなどを Export することができる。これら注視点の座標データと持続時間のデータから、被験者が Excel を用いて誰（何）を見ているかを判断することができる。

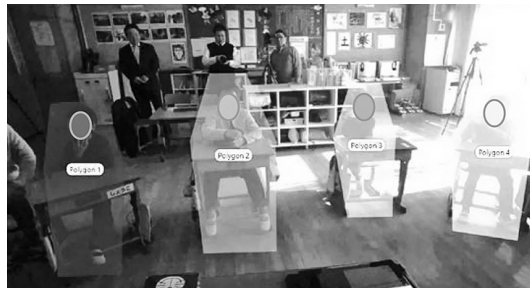


図 11. AOI のイメージ

この方法を正確に説明すると次の通りである。図 3 で説明したように注視点の座標 (x, y) は画像の左上を原点とする。そして、注視点が右側に行くにつれて x の値が大きくなり、下側に行くにつれて y の値が大きくなる。 x の最大値は 1920, y の最大値は 1080 である。Excel を用いて、例えば図 11 のどこを見ているかを判定する場合、 x の値が 110 と 410 の間で、 y の値が 180 と 750 の間であれば、一番左の人を見ているということとする。同様に、左から 2 人目、3 人目、4 人目の人を見ていると判定する条件も注視点の座標で判断する。これを表にまとめると、表 2 の通りとなる。

表 2. 被験者が誰を見ているかを判断する条件

	左から 1 人目	左から 2 人目	左から 3 人目	左から 4 人目
x の値	(110, 410)	(660, 950)	(1210, 1450)	(1710, 1960)
y の値	(180, 750)	(110, 700)	(160, 650)	(170, 760)

この表で、「 x の値」が (110, 410) とは、 $110 < x < 410$ という意味であり、他も同様である。また、「左から 1 人目」とは、 $110 < x < 410$ かつ $180 < y < 750$ という条件なので、図形は長方形となる。したがって、図 11 に記してある AOI の領域の形とは少しだけ異なるが、条件を簡単にするためにこのようにした。これらの領域以外に被験者の注視点がある時には、「人以外」を見ている、と判定する。また、これらの領域の数値を変えることによって、「辛め」の判定や「甘め」の判定も行うことができる。

さて、これを利用すると、被験者が誰を見ているかという事を Excel で分析することができる。この結果と Tobii Pro Lab の AOI 機能を用いて分析した結果を表 3 に示しておく。「データ数 1」は Excel を用いて分析した注視点のデータ数である。「データ数 2」とは Tobii Pro Lab の AOI 機能を用いて分析した注視点のデータ数である。「時間 1」、「時間 2」とはそれぞれの対象を注視していた時間、「割合」とは 4 人の児童のいずれかを見ている時間を 100% とし、

それぞれの児童を見ていた時間の割合を百分率で表したものである。これによると、分析で用いた領域の形が違っても関わらず、ほぼ同様の割合であった。

表 3. Excel の関数を用いて行った注視点の分析と AOI を用いた分析

事象	データ数 1	時間 1 [sec]	割合 1	データ数 2	時間 2 [sec]	割合 2
左から 1 人目	273	5.5	14.10%	249	5	12.80%
左から 2 人目	1126	22.5	58.10%	1177	23.5	60.50%
左から 3 人目	405	8.1	20.90%	398	8	20.50%
左から 4 人目	134	2.7	6.90%	122	2.4	6.30%
合計	1938	38.8	100.00%	1946	38.9	100.00%
人以外	10161	203.2		10153	203.1	
総計	12099	242		12099	242	

(3) Excel を用いた手動分析と Gaze Plot

では、次に Excel を用いて作成したバブルチャート (図 12) と Tobii Pro Lab を用いて作成した Gaze Plot (図 13) を比較しよう。Gaze Plot は、注視点のバブルチャートであり、それぞれのバブルの半径は、注視の継続時間に比例している。両者は、バブルの半径の注視の継続時間への比例係数が違うため、バブルの半径が異なるが、同じ点にバブルがあるため、同等なチャートである。

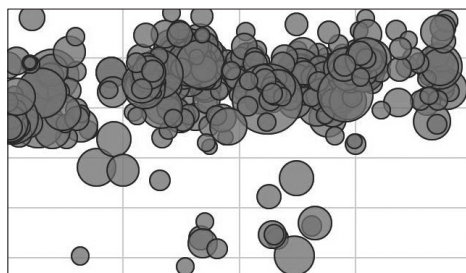


図 12. 手動分析のバブルチャート

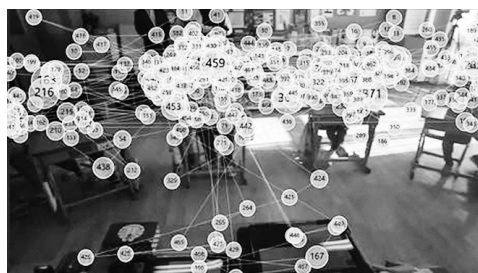


図 13. Tobii Pro Lab の Gaze Plot

(4) Excel を用いたより詳細な分析

Excel を用いると、注視点に関してより詳細な分析を行なうことができる。図 14 は、誰を見ているかを色分けしたバブルチャートである。これは、図 12 と同じデータから作成したが、誰を見ているかより詳しく表示することができる。(ただし、やや面倒である。)

私たちは、以前の分析 [4, 5] で、RWM を用いずに、Heat Map あるいは Gaze Plot を作成していた。これは、この当時、まだ RWM という手法を知らず、これを行うことができなかったからである。

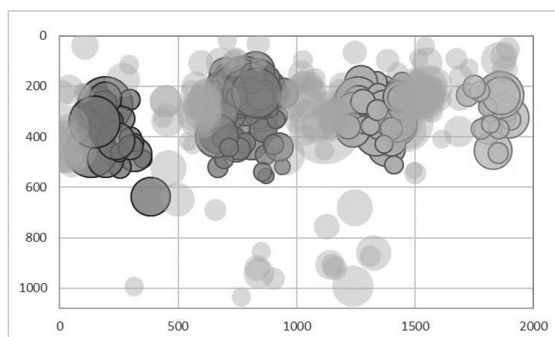


図 14. Excel で作ったバブルチャート

では、RWMを行わない Heat Map あるいは Gaze Plot は無意味なのだろうか。以下で考察してみよう。図 15 は、図 14 から人以外を視ていた注視点を除いたものである。

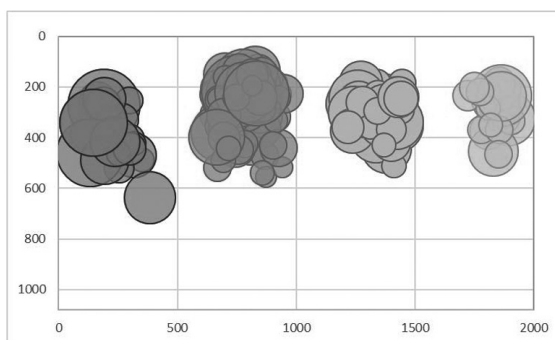


図 15. 図 14 から人以外を視ていた注視点を除いたもの

図 16 は RWMを行わなかった時の注視点のバブルチャートである。ただし、これまでと同様に、誰を見ていたかを色分けして示している。

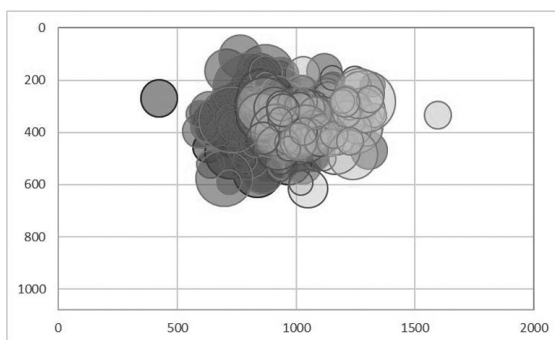


図 16. RWMを行わなかった場合の注視点のバブルチャート

図15と図16の比較により、RWMを行わなかったとしても、以下のことが言える。被験者は、正面の人を見る時、（おそらく体を正対させ）ほぼ真ん中の領域を見ていた。また、左側の人を見るにはやや左側の視野を使い、右側の人を見るにはやや右側の視野を使っていたことがわかる。このような訳で、RWMを行わない場合は、全体的に注視点の分布が中心寄りになるが、視野が広い人はやはりRWMを行わない場合も視野が広いのではないかと考えられる。ただし、この仮説は分析を行なった上記の例だけで成り立つので、取り扱いに注意が必要である。そして、分析事例を増やしていかなければはっきりとしたことは言えない。

6. まとめと議論

この論文では、アイトラッキングカメラ Tobii 社製 Tobii Pro Glasses 2 を用いた一人称視点の撮影方法とこれを用いたデータ分析の基礎技術についてまとめた。具体的には、Real-World Mapping (RWM) の方法の概要をまとめ、一人称視点動画から切り出した静止画を用いる場合と、教室の全景を映した動画を用いた場合の比較を行った。これによると、前者を用いた場合は 自動：手動 = 6：1 であり、後者の場合は、自動：手動 = 1.4：1 であった。すなわち、一人称視点動画から切り出した静止画を用いた方が、自動 RWM に成功しやすいという結果を得た。

なお、公立小学校4年生の通常学級の国語の授業に対して行った分析[7]では、教師が「児童の方」を注視していたデータ約4万個のうち、自動 RWM と手動 RWM の比は約 1:2 であった。この授業では、アイトラッキングカメラの動画から複数の静止画を切り出し、これらを合成して一枚の静止画を作り、RWM を行った。また、公立小学校4年生の通常学級の道徳の授業に対して行った分析[8]では、教師が「児童の方」を注視していたデータ約2.5万個のうち、自動 RWM と手動 RWM の比は約 1:2.4 であった。これは、児童数が多かったため見やすさを優先して、黒板に取り付けた SONY のウェアラブル 4K カメラから静止画を切り取り、これを RWM の背景静止画として用いた。このように、アイトラッキングカメラの動画から合成した静止画よりも、ウェアラブル 4K カメラから切り出した静止画の方が、教室の全景を捉えることができるため、（手作業は増えるが）有利である。

次に、Excel の関数を用いた手動分析の結果と Tobii 社の分析ソフト Tobii Pro Lab を用いた Areas of Interest (AOI) の分析結果の比較を行った。両者はほぼ同じという結果を得た。また、Excel を用いた手動分析と Gaze Plot を比較した。そして両者ともほぼ同じという結果を得た。これらにより、手軽に分析を行ないたいときには、Tobii Pro Lab を使うと良い。しかし、より詳細な分析を行ないたいときには、Excel を用いた手動分析を行なえばよいであろう。Tobii Pro Lab は、RWM, AOI, 注視点のデータのエクスポート, Heat Map や Gaze Plot の作成など、非常に多岐にわたる機能を持った分析ソフトである。ただ、中身で用いら

れているアルゴリズムやソースコードは公開されていない。このソフトを使うに当たって、この論文で行ったような簡単な整合性チェックを行うことは意味がある。

さらに、Excel を用いて、より詳細な分析を行なった。RWM を行った注視点の座標データと手動 AOI を組み合わせることにより、誰を見ているかを色で示したバブルチャートを作成した。また、RWM を行っていないバブルチャートも作成した。後者から、正面の人を見る場合は、視野のほぼ真ん中で見ていたこと、また、左側の人を見るにはやや左側の視野を使い、右側の人を見るにはやや右側の視野を使っていたことがわかった。また、RWM で視野が広いと思われる人は、RWM を行わないバブルチャートでも視野が広いと判定されるのではないかという仮説を提起した。これを実証するためには、より分析事例を増やしていかなければならないだろう。

謝辞

名城大学教職センターの教員、事務の方に感謝する。また、共同研究者の平山先生、後藤先生に感謝する。

参考文献

- [1] 河野義章編著「授業研究法入門」図書文化、2009 年。
- [2] 後藤明史，平山勉，「教師視点の映像記録を活用した授業の多元的記録・分析・構成方法の研究」，日本教育工学会研究報告集，JET14-1，pp.81-86，2014。
- [3] 後藤明史，平山勉，「教師視点の映像記録を活用した授業の多元的記録・分析・構成方法の研究 II」，中部教育学会第 63 回大会提示資料
- [4] 後藤 明史，平山 勉，谷口 正明，「教師視点の映像記録を活用した授業の多元的記録・分析・構成方法の研究 (3)」，日本教育工学会研究報告集 17 (1)，531-534，2017-03-04，日本教育工学会
- [5] Masa-aki TANIGUCHI, Tsutomu HIRAYAMA, and Akifumi GOTO (2017), "Analysis of Teachers' View Using an Eye-tracking Camera System —Difference between an Expert Teacher and a Novice Teacher—" Paper for International Conference for Media in Education, 2017.
<https://drive.google.com/file/d/0B5leUJsit5JGT2JvWmZMYTE1Z0E/view>
- [6] 平山勉，谷口正明，中山真樹，平山幸代，後藤明史，竹内英人，「教師視点映像記録を活用した授業者の授業改善 — 道徳とプログラミングの事例分析に基づいて —」，名城大学教職センター紀要 第 19 巻 125-138 2022 年
- [7] 谷口正明，後藤明史，平山勉，中山真樹，平山幸代，「教師の一人称視点をを用いた授業分析 — 熟練教師と初任教師の注視点に着目して —」(投稿中)
- [8] 平山勉，谷口正明，中山真樹，平山幸代，後藤明史，竹内英人，「教師視点映像記録を活用した授業者の授業改善」，名城大学教職センター紀要 第 20 巻 2023 年 (投稿中)
- [9] この論文の図をカラーで見ることができるサイトを用意した。URL は以下の通りである。
https://drive.google.com/file/d/1U4GYIDKP9g4d-VjEG_RTOK3_LI9ST7N9/view?usp=sharing