

リアルタイムに学習者の集中度を推定するシステムの作成 The System to Estimate Learner Concentration in Real Time

大政 悠暉[†]
Yuki Ohmasa

松本 慎平[‡]
Shimpei Matsumoto

加島 智子[†]
Tomoko Kashima

1. はじめに

近年、学習効果を計測するための指標として注目されているものに集中度がある。集中度計測のための取り組みとして接触式のセンサなどを使用する方式がある。しかし、センサを使用することにより、使用者の負担や機材を揃える手間が大きいといった問題がある、また、比較的使用者の負担が少ない手法として非接触式の WEB カメラを用いる方式があるが、法人向けのサービスが主であり個人で扱える集中度計測システムは実質的に存在しない。

そこで、本研究では従来のシステムよりも簡単に、オンライン授業などの様々な学習方法に対して学習効果を計測することのできるシステムを作成し、その有効性を検証する。このシステムを使用することで、学習方法のどのような要素が学習効果に影響しており、どのように改善していけばよいのかを分析するためのツールとして活用されることを期待している。さらに、学習者自身がシステムを活用することで、自分自身に最適な学習方法や休憩のタイミング、時間を確かめたり、オンライン授業等において、自身の集中度を確認することで、集中度の低下を予防したりすることによって学習の効率を向上させ、システムの使用者の学習効果を向上させることを目指す。

2. システム概要

本システムではオンライン会議や講義において、相手の話を傾聴しているときや PC 作業などを行っている際の集中度合いを測ることを目的としている。本システムは、表 1 に示すように大きく分けて指標検出部・時系列データ作成部・集中度推定部の 3 つに分割して考える。まず指標検出部では、WEB カメラを用いて集中度を取得する対象の顔画像をリアルタイムに取得し、その画像から対象の顔の向きや目の向き、閉目の有無、表情などの指標を取得する。時系列データ作成部では、指標検出部で取得できなかった空白データの補完を行い、過去一定区間の指標データから時系列を考慮したデータを作成する。そして集中度推定部では、時系列を考慮したデータを集中度推定モデルに代入して現在の集中度を算出する。

2.1 集中度の定義

本システムでは与えられた顔画像から、対象者がどの程度集中しているのかを判断する必要がある。例えば、集中度を判断するための集中度の定義として、集中力の向上を目的とした知覚刺激の先行研究 [1]が挙げられる。先行研究では、集中を「1 つの事物に継続的に注意を向けている」

状態と定義している。これは、普段はなるべく多くの刺激を処理するために向けられる注意が、1 つの物事を精密に処理するために継続的に向けられることがあるからである。そこで、本研究では上記の定義を集中度の定義とする。

表 1 システムのフロー

	詳細項目
Step 1 指標検出	1. 事前処理 2. 画像取得 3. 特徴点抽出 4. 表情検出 5. 頭部姿勢推定 6. 目線検出 7. 閉目検出 8. 描画
Step 2 時系列データ作成	1. 空白データ処理 2. 表情データ作成 3. 頭部姿勢データ作成 4. 視線データ作成 5. 瞬目データ作成 6. データリスト作成 7. データ出力
Step 3 集中度推定	1. 推定モデル代入 2. 出力

2.2 システムの詳細

集中度推定システムの構成は図 1 に示す。表 1 に示した 3 つの構成部分についての説明は以下の通りである。

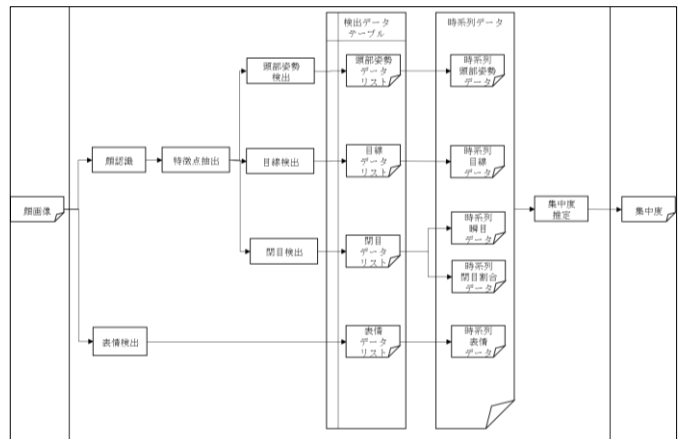


図 1 集中度推定システムの構成

2.2.1 指標検出部

指標検出部では、まず WEB カメラから使用者の肩より上の部分を連続的に取得する。その後、取得した画像に対して機械学習ライブラリである Dlib [2]を用いて特徴点抽

[†] 近畿大学工学部 Kindai University Faculty of Engineering

[‡] 広島工業大学情報学部 Hiroshima Institute of Technology Faculty of Applied Information Science

出を行う。次に、取得画像と特徴点の座標を用いて顔部分の範囲を抽出し、顔画像サイズの正規化を行う。そして、使用者の顔の特徴点を取得し、PAZ [3]を用いて表情検出を行い使用者の怒り、嫌悪、恐れ、喜び、悲しみ、驚き、無表情それぞれのパラメータを取得する。

続いて、図 2 に示す取得した特徴点を用いて yaw, roll, pitch の 3 軸の頭部姿勢を推定する。そして、取得した使用者の画像と特徴点の目の部分を用いて、使用者の視線と閉目の有無を調べる。

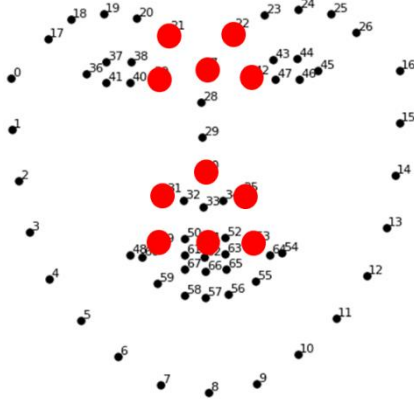


図 2 頭部姿勢推定に使用する特徴点の位置

視線の検出について、目の領域を判定するために、図 3 に示す特徴点 $p1$ から $p6$ を使用する。まずは式 1, 2 のように $p1$ から $p6$ までの x 座標の集合を X , y 座標の集合を Y とする。次に、式 3, 4 に示すように、それぞれの最大値と最小値を調べることによって *Minimum Point* と *Maximum Point* の 2 点を決定する。この 2 点で囲まれた四角形の領域を目の領域と判定する。その後、目以外の領域が黒目の推定に影響を与えないようにするために、カラー画像に対して $p1$ から $p6$ を用いて目の領域以外をマスクしたのちに、大津の 2 値化 [4]を用いて二値画像を取得する。続いて、先ほどの処理を行って取得した図 4(a)のような二値画像に対してモルフォロジー変換を行い、図 4(b)のように目の領域に被っている眉毛や陰の影を削除する。そして、その画像に対して図 4(c)の輪郭抽出を行い、最も大きな輪郭を黒目と推定する。最後に、図 5 のように目の領域の中心と黒目の中心の差から目線を検出し、検出した値を正規化したものを目線の値とする。

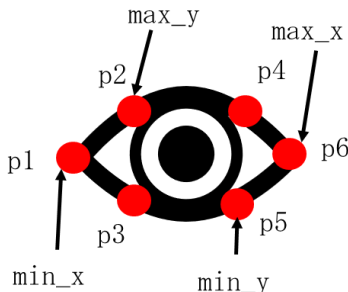


図 3 目の領域の判定

$$X = \{p1_x, p2_x, \dots, p6_x\} \quad (1)$$

$$Y = \{p1_y, p2_y, \dots, p6_y\} \quad (2)$$

$$\text{Minimum point} = (\min X, \min Y) \quad (3)$$

$$\text{Maximum point} = (\max X, \max Y) \quad (4)$$

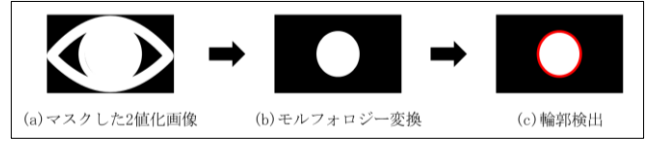


図 4 黒目の検出

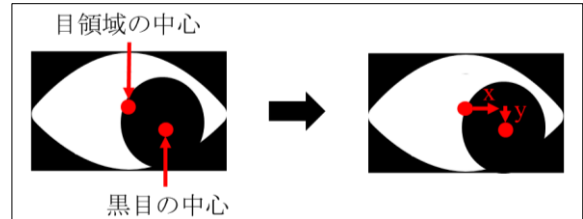


図 5 視線の算出

閉目の検出について、閉目の有無の検出には目のアスペクト比である *EAR* (Eye Aspect Ratio) [5]を使用する。*EAR* の算出には図 3 の $p1$ から $p6$ を使用し、式 5 に代入することにより算出する。本システムでは、*EAR* の閾値を 0.205 として、算出された *EAR* の値がそれを下回ると閉目していると判断する。

$$EAR = \frac{|p2 - p3| + |p4 - p5|}{2|p1 - p6|} \quad (5)$$

2.2.2 時系列データ作成部

まず、指標検出部で作成した過去 30 秒の各指標のデータを取得し、空白データの処理を行う。空白データの処理では、表 2 に示すとおり、顔が検出できなかった場合や、目を瞑っていて黒目が検出できず、目線が取得できなかった場合に、データの補完、検出の停止の処理を行っている。顔の検出の有無では過去 30 秒間の顔の検出の有無を遡り、検出割合が一定を下回る、もしくは一定期間顔を取得できていない場合は集中度の推定の停止を行う。目線の検出の有無では、一定期間以上連続して目線が検出できない場合に過去のデータを使用してデータの補完を行う。

表 2 空白データの処理

種類	条件	対応
顔の未検出	<ul style="list-style-type: none"> 直近 10 秒の半分以上のフレームが未検出 直近 5 秒間で検出されたフレームが 0 のいずれか 	集中度推定の停止
目の未検出	直近 15 秒のすべてのフレームで目の未検出	15~30 秒前のデータの平均値を取得

次に、時系列を考慮したデータを作成する。まず、表情に関するデータの作成として、過去 30 秒の指標リストより、各表情の評価値パラメータと各フレームで最も評価値の高かった表情を抽出して、過去 10 秒間の各表情の評価値の最大値、最小値、平均値と各フレームで最も評価値が高かった表情の個数を求める。次に、頭部姿勢に関するデ

ータの作成として、各指標のデータから頭部の yaw, roll, pitch の値を抽出して、変化の合計の平均値と分散値を求める。さらに、視線に関するデータの作成として、過去 30 秒間の両目それぞれの x 軸と y 軸の視線の向きから視線の平均値と分散値を取得する。また、表 3 の通りに過去 20 秒の両目それぞれの瞬目の有無と過去 30 秒間の閉目の有無から、過去 20 秒間の瞬目回数と過去 30 秒間の閉目の割合を取得する。最後に、データを作成した日時と経過時間を取得して、取得した各種データを過去 30 秒間の説明変数のデータのリストとして保存する。

表 3 瞬目データの加工

名称	値の取得方法	条件
瞬目	右の条件に沿ったものをカウント	直近 20 秒で目を開くまでの時間が 400ms 未満かつ前目を開いてから 500ms 以上経過
瞬目時間	瞬目と判断されたフレームの合計時間	直近 30 秒で目を開くまでの時間が 400ms 以上
閉目時間	30 秒 - 瞬目時間	なし
閉目割合	閉目時間 / (瞬目時間 + 閉目時間)	なし

2.2.3 集中度推定部

集中度推定部は、過去 30 秒間の説明変数のデータのリストを取得し、表情、顔の向き、視線、瞬目、閉目のデータを集中度推定モデルに代入して現在のフレームの集中度を取得する。また、取得した画像や指標、データ、集中度の GUI 表示、CSV ファイルへの出力を行う。

3. 予備調査

本研究では、顔の映像データから時系列を考慮した指標データを作成し、それを教師データとして集中度モデルを作成する。そして、そのモデルを使用して集中度推定システムを動作させる。そこで、時系列データ作成までのシステムの動作や、モデルの作成を実証するために予備調査を行い、モデル作成、精度の検証、システムの動作を行う。

3.1 被験者

近畿大学工学部の学生 4 人に対して実験を行った。実験内容は以下の通りである。

被験者は、まず WEB カメラを用いて顔の映像を録画した状態で、5 分程度のプレゼンテーションを 3 種類聴講する。その後、そのプレゼンテーションに関する 4 択問題をオンラインで解答する。問題は 10 点満点で、本実験では解答の点数を集中度とする。

3.2 モデル作成

モデルを作成するために、予備実験で取得した映像それぞれに対して、システムを動作させて時系列データを取得して、説明変数とする。また、その時系列データに対応する回答の点数を目的変数とする。これらを組み合わせることで教師データを作成する。そして、これらの教師データでモデルを訓練することによって予備実験におけるモデル

を作成する。今回の実験では、集中度を推定するためのアルゴリズムとしてランダムフォレストを採用する。

3.3 精度検証

作成したモデルに対して 5 分割の交差検証を行って精度を推定した結果、表 4 の結果が得られた。

表 4 予備実験の精度検証結果

項目	数値
正答率	0.80
適合率	0.83
再現率	0.83
F 値	0.83

表 4 より、約 80% 程度の精度で集中度を推定できることが確認された。ある程度の精度は確保できているが、被験者の人数が少ないため、人物や環境の変化に頑健でないと考えられる。

4. 学習モデルの作成

予備実験では、システムの動作やモデルが特定の条件下においてある程度の精度で集中度を推定できることを確認した。しかし、取得したデータ数が少ないため、より多くのデータを収集して学習モデルを作成するため、さらにデータの収集を行う実験を行う。また、このモデルが適切であるか分析を行う。今回は、PC を用いた講義における集中度の顔データと集中度データを取得するため Web カメラから個人の顔データを取得、集中度は講義内容を集中して聞いていれば回答ができ、聞いていなければ回答が難しいと考えられる小テストを実施し講義中の集中度合いを計測する。更に、10 名の受講者につき 1 名のスタッフが客観的な集中度合いの評価を行い、学習モデルのデータを取得する。

4.1 被験者

近畿大学工学部 1 年生、HTML 言語に関する講義受講者 107 人に対して実験を行う。被験者は各個人の PC にて講義で説明される資料を閲覧し、テキストを参考にしながら受講する。

4.2 実験内容

実験の実施は以下の手順で実施する。

- 事前調査：HTML の受講前の事前知識の確認（知識のある受講者はデータから除外）
- 講義の受講：約 30 分間の HTML に関する講義を聴講する。内容は 10 セクションで構成されている。
- 講義内容のテスト（講義内容の理解度）：説明された HTML に関する各セクションから問題を回答してもらう。

顔データの取得は、Web カメラで手順 2、3 の録画を行う。また、客観的な集中度を測定するため手順 2 において、約 10 名の受講者につき 1 名のスタッフがセクションごとに 3 段階で評価を行い集中度の評価を行う。

4.3 実験結果

本実験を実施した結果、アンケート調査で事前に 19 人を実験から除外し、69 件の有効なデータを収集することができた。



図 6 実験の様子

4.4 データ分析

実験で収集した HTML の問題の解答とスタッフが評価した集中度合いに対して分析を行い、傾向を観察する。

4.4.1 被験者ごとのデータの考察

図 7 は被験者ごとの集中度合いの点数の合計値とテストのスコアのグラフである。グラフを見ると、集中度合いに対して明らかにスコアが低いグループとスコアに対して集中度合いが低いグループがあることがわかる。これらのグループについて、被験者自身の能力や予習の有無により外れ値となっている可能性があるため、今回は除外する。外れ値を除外した後の相関係数は 0.629 であり、これらの指標には相関があることがわかる。

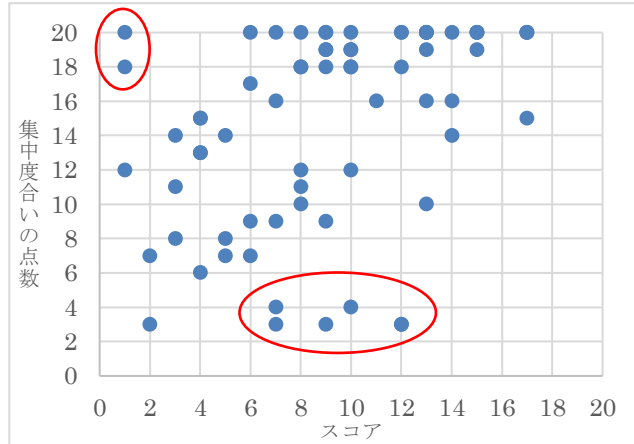


図 7 個人の問題のスコアと集中度合いの点数のグラフ

4.4.2 問題ごとのデータの考察

図 8 は問題ごとの正答率と、その問題が該当するセクションの集中度合いの被験者全員の合計値のグラフである。グラフを見ると、集中度合いの合計値に対して明らかに正答率が低い問題がある。この問題は筆記問題であり、該当セクション内であまり触れられず、かつ説明が始まってすぐの箇所まで被験者の準備ができていなかったため正答率が考えられる。したがって今回はこの問題を外れ値として除外する。外れ値を除外した後の相関係数は 0.768 であり、これらの指標には強い相関があることがわかる。

4.5 モデル作成

モデル作成及びモデルの評価の詳細は発表の際に示す。

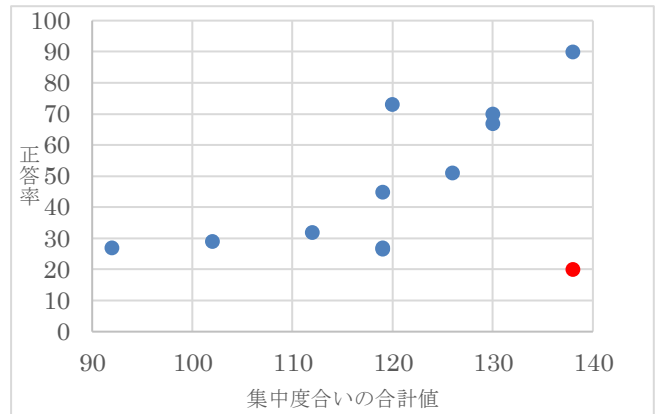


図 8 問題ごとの正答率と集中度合いの合計値のグラフ

5. おわりに

本研究では、従来手法よりも手軽に学習時の受講者の集中度度を推定するためのシステムの作成を行った。

予備調査の精度検証では、80%程度の精度を確認し、一定の有効性を実証した。また、環境の変化に対して頑強な学習モデルを作成するために実験を行い、データの傾向を分析した。

今後の課題として、今回は事前に知識がある可能性を考慮して、知識量について事前に調査を行ったが、授業の予習を行っていたり、能力が高く観察力のある人がいたりする場合は、本研究の実験で実施したテストにおいて集中度度を測ることが難しいことが明らかとなった。したがって、今後は知識が全くない内容にてモデル作成のためのデータ取得を検討する必要がある。また、より高精度で高速なシステムの作成を行ったり、システムの適用範囲を広げたりするために様々な状況下でのデータを収集する必要がある。さらに、本システムを活用した学習方法を考案し、既存の学習方法との違いを検証することで、より良い学習方法を提案することが可能となる。

謝辞

本研究は JSPS 科研費基盤研究(C)22K02815 の助成を受けたものです。

参考文献

- [1] 橋卓見, 岡部浩之, 佐藤未知, 福嶋政期, 梶本裕之: 注意領域への集中力向上を目的とした領域外における視覚刺激, 第 16 回日本バーチャルリアリティ学会大会論文集, (2011)
- [2] Davis E. King : Dlib C ++Library, Dlib (Online), available from <http://dlib.net/imaging.html>
- [3] Octavio Arriaga : Emotion classifier, PAZ-Documentation (Online) available from <https://oarriaga.github.io/paz/>
- [4] 大津展之: 判別および最小 2 乗規準に基づく自動しきい値選定法, 電子情報通信学会論文誌 D Vol.J63-D No.4 pp.349-356, (1980)
- [5] J. Cech and T. Soukupova, "Real-Time Eye Blink Detection Using Facial Landmarks," in Computer Vision Winter Work, 2016