

講義動画視聴時における同一コンテンツ学習者の頭部動作の同期性に基づいた集中度の推定

川村 亮介¹ 中谷 勇太² 豊田 雄志³ 新沼 厚一郎³

(株)富士通研究所¹ 富士通株式会社² Fujitsu Laboratories of America, Inc³

1 はじめに

学習施設でも採用されている動画学習では、対面で授業と異なり学習者の集中状態によらず授業が進むため、受講をやめてしまうことが大きな問題となっている。解決策の1つとして、集中度に応じて自動的にサポートを行うシステムが望まれており、そのための集中度の推定が重要になっている。設置の簡易さからウェブカメラを用い、学習者個々人の表情や姿勢の変化から集中度を推定する手法が提案されている[2,3]が、実際の学習時のシナリオでの評価はされていない。

本稿では、簡易な装置による実際の学習シナリオでの集中度推定を目的とし、まず、実際の学習時の様子を記録し、実際の学習時には表情や姿勢の変化に基づいた集中度の推定が難しいことを確認した。次に、解決手段として学習時に頻出する頭部動作の同期性に着目し、集中度推定手法の検討と評価を行ったので、これらの報告を行う。以下、実験の詳細および頭部動作の同期性に基づいた手法の評価について述べる。

2 データの収集実験

実際に動画学習を用いている塾において、動画学習時の様子を記録する実験を行った。

2.1 実験の設定

被験者はPCのディスプレイを用いて動画学習を行い、その様子はディスプレイ上部に設置されたウェブカメラを用いて撮影された(図1参照)。講義動画は、プロの講師によって作成・提供された(教科は現代文)。被験者は講義動画の視聴及び紙の資料を用いて、普段と同様に学習を行った。なお、被験者の姿勢、およびメモを取るなどの動作は制限をせず、動画の一時停止も自由としたが、早送り、巻き戻しは禁止とした。

被験者は実験を行った塾の周囲に住む大学生16名(女性:8名、男性:8名)とした。被験者は1本15分の講義を、1日4本視聴し、2日間で計8本(#1-#8)を視聴した。視聴する時間帯は、

Detecting loss of attention in video based lectures using synchrony of head movements with same video watchers

1 Fujitsu Laboratories Ltd.

2 Fujitsu Limited.

3 Fujitsu Laboratories of America, Inc



図1 実際の学習環境における実験時の様子

午前と午後の両方において視聴するように振り合分けた。実験前の時間の行動については制限をせず、各講義の終了後には確認テストと、アンケートを行った。アンケートでは、講義視聴中の集中度について、0-4(0:集中していない状態、4:集中している状態)の5段階のリッカート尺度での回答を課した。本稿ではその結果から、4・3と回答した場合は”集中している”、1・2と回答した場合は”集中していない”とした。なお、0と回答したものはなかった。

2.2 実験の解析

記録した被験者の映像の確認と解析を行った。まず、各#1-#8の動画においての”集中していない”の割合を表1に示す。表1の通り、全体として集中していると回答した視聴者が多いことがわかる。

表1 各動画での”集中していない”の割合

動画 No.	#1	#2	#3	#4	#5	#6	#7	#8
“集中していない”の割合(%)	25	31	50	56	31	19	25	36

次に授業の経験も豊富なプロの講師と共に映像の確認と学習者の集中に関して注目している点についての議論を行った。その中では、次のような知見が得られた。

- プロの講師であっても、表情や姿勢の変化から集中度を推定することは難しく、今回得られたデータにおいて、表情や姿勢の変化を利用した既存手法による集中度の推定は難しいと考えられる
 - 学習者個々の動きではなく、他者と比較した動きの違いに注目している
- また、観察の中でよそ見や居眠りといった姿勢

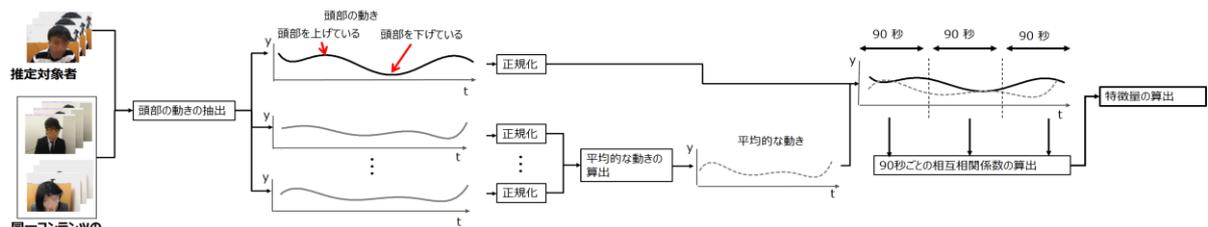


図 2 特徴量抽出処理のステップ

や表情の大きな変化はほとんど見られなかったが、ノートテイクのために、頭部を下に動かす動作が多く確認できた。そこで、各動画での、被験者が下を向いている、もしくは画像中から顔が検出されない割合を算出した。その結果を表 2 に示す。

表 2 各動画での被験者が頭部を下げていた割合

動画 No.	#1	#2	#3	#4	#5	#6	#7	#8
頭部を下げていた割合 (%)	23	37	39	40	37	46	42	42

表 2 を見ると、#1 を除いた動画では 4 割程度が下を向く、もしくは顔が画像から検出されないことがわかる。居眠りのように長期間下を向くような動作はほとんど確認できなかったため、ノートテイクのような短いスパンで下を向くという動作が原因と考えられる。また、このように顔が下を向いてしまう場合、表情を利用した集中度の推定は困難となる。

3 頭部動作の同期性に基づいた推定の検討

集中していない場合、他の学習者と比較して、頭部動作が起こらない、または遅れて動作をすることが、実験データの観察から確認された。そのため、あるタイミングでの他の学習者との頭部動作を比較することで、集中度の推定に有用だと考えられる。なお、集中度の推定ではないが、Fujii ら[1]も頭部動作の同期具合の可視化が、集団授業での教師の補助に有効であることを検証している。

3.1 アプローチ

本稿では頭部動作の同期性に着目し、それぞれの被験者の頭部動作と、その他の被験者の頭部動作の平均との相互相関係数を算出する。この流れを図 2 に示す。まず、鼻の頂点をはじめとした顔や頭部の特徴点の位置を取得し、被験者ごとに動画 1 本の視聴時での最大・最小値で正規化する。これらの時間的変化を頭部動作とする。さらに、同一コンテンツを視聴している他者の頭部動作の各時刻での平均を算出し、90 秒毎に推定対象の学習者の頭部動作との相関係数を算出する。90 秒毎に計算された相関係数を特徴量

とし Local Linearly Embedding で次元圧縮する。

3.2 評価

評価においては、「集中している」と「集中していない」の 2 つの状態を分類する分類モデルを SVM を用いて各講義ごとに構築する。15 人分の特徴量を訓練データ、残りの 1 人分の特徴量を評価データとする Leave one-person out cross validation を行い、平均の F-score で評価する。なお、データの不均衡を考慮しデータ拡張を行った。また、表情を用いた手法との比較を行うため、[3]を参考にし、Action Unit (AU) と Head Pose (yaw, pitch, roll) の 15 分間の平均を特徴量とした場合での分類も行った。推定結果の F-score を表 3 に示す。

表 3 推定結果の F-値

	Chance rate	AU + Head Pose	提案手法
F-値	0.35	0.53	0.78

表を見ると、提案手法が最も高くなっており、表情変化と比較しても、頭部の動きの同期性が集中度の推定に有効である可能性が示された。

4 終わりに

本稿では、簡易な機器を用いた実際の学習シナリオでの集中度の推定を目的とし、まず、実際の学習シナリオでの学習時の様子を記録し、観察・解析を行った。その結果から、同一コンテンツ視聴時の頭部動作の同期性に基づいた推定手法を検討、評価した。その結果、実際の学習環境で取得したデータで F-値において表情や頭部の向きを用いた場合を上回る精度を示すことを確認した。今後は、被験者数や講義の種類を増やした場合について検討していきたい。

参考文献

- [1] Katsuya Fujii, et al. 2018. Sync Class: Visualization System for In-Class Student Synchronization. In Proc. AH. ACM, 12:1-12:8.
- [2] Tarmo Robal, et al. 2018. Webcam-based Attention Tracking in Online Learning: A Feasibility Study. In Proc. IUI. ACM, 189-197.
- [3] Jacob Whitehill, et al. 2014. The faces of engagement: Automatic recognition of student engagement from facial expressions. IEEE Trans. Affective Computing 5, 1, 86-98.