

研究報告書

「インタラクションの大規模結合による「学習場」の情報化」

研究タイプ: 通常型

研究期間: 平成26年10月～平成30年3月

研究者: 川嶋 宏彰

1. 研究のねらい

OCW (Open Course Ware) や MOOCs (Massive Open Online Courses) をはじめとした映像視聴型の e-learning 環境では、受講者が一方向的に講義映像を見るものが多く、実際の講義と比べ、双方向の対話的要素が不足している。受講者の視聴時におけるつまずきや集中状態は考慮されず、視聴後のテスト等で知識の評価を行うことが多いが、評価項目が部分的になるだけでなく、コースの受講自体を持続することが難しいといった問題がある。また、選択肢となる映像やコースが多岐にわたる場合は、しばしばどの映像を視聴するかという選択の困難が生じる。実際の講義では、講師が学習者の状態を考慮して対話的なやり取りを行うことで、学ぶ側の学習意欲が生起された状況、いわば「学習場」を維持し、講義外でも対話的アドバイスによって受講コースを推薦するなど、興味・関心が引き出される学びを支援している。

そこで本研究では、映像講義を選択・視聴する際の学習者の興味や集中度、理解度を推定するとともに、推定された状態に基づいて学習者へ可視化、さらには対話的フィードバックを行う学習支援エージェントを開発する。これにより、学習者自身が自らの状態(どこでつまづいているか、どのようなトピックに興味があるか)を自ら振り返ることができるようになり、学習意欲や学習効果が高まると期待できる。本研究では映像視聴型 e-learning におけるフェーズを、(A)映像視聴時(視聴フェーズ)、(B)コース・映像の選択時(選択フェーズ)、(C)視聴後の問題・課題解答時(フォローアップフェーズ)の3つに分け、それぞれの状況を想定した、人の状態推定とフィードバック手法を開発する。具体的には、(A)では、映像視聴時や視線等に基づき、視聴時のつまずき個所の同定や注意・集中状態を推定するとともに、推定状態を可視化・フィードバックする。(B)の選択時の状況では、選択肢が表示されたカタログ閲覧時の視線を解析し、推定された興味に基づいて対話的働きかけを行う。(C)では、対話的な課題解答時における表情、姿勢を解析することで、理解度を推定するとともに、適切なヒント提示や質問を生成する。

本研究では、これらを現在研究の進むラーニングアナリティクスとつなげる方式、さらには、学習支援エージェント間の大規模な情報共有により、分散的に学習場形成のためのコツを発見する方式および道筋を明らかにする。

2. 研究成果

(1) 概要

研究のねらいで述べたような3フェーズ、すなわち(A)視聴フェーズ、(B)選択フェーズ、(C)フォローアップフェーズのそれぞれにおいて、人へ適切なフィードバックを返す手がかりとして、人の心的状態を考慮する方法を検討した。これは Intelligent Tutoring System の研究に

代表されるように、人の認知・情報処理過程を考慮した働きかけが学習者の意欲や知識構築に対して有効であると期待できるからである。本研究では、上述の3フェーズにおいて、人の行動の背後にある注意・集中度、興味、理解度といった心的状態を推定する方法、およびその基盤技術についてそれぞれ研究を行い、以下の成果が得られた。

A. ビデオ講義映像視聴時の注視行動モデルと視聴状態推定手法の開発

講義映像を集中して視聴する典型的パターンが分かれば、そこからのずれによって集中の有無を知ることができる。しかし実際の視聴スタイルは、スライドに基づく講義であれば、講師が音声や指示棒で説明する箇所を追従する場合や、講師の説明に追従せずにスライドの構造を確認しながら先読みを行う場合など、複数のモードに大きく分けられる。このような複数モード混在型の視聴行動モデルと、それに基づく視聴者の集中状態推定手法を開発した。

B. カタログ閲覧時における興味推定アルゴリズムの開発

コースや映像カタログなど複数の選択肢が提示される状況において、閲覧者がどのトピックに関心があるか、さらにはどのような目的意識をもって選択しようとしているのかを知ることは、適切なコース・映像の推薦や、対話的選択支援を行う上で重要な手がかりとなる。そこで本研究では、閲覧者の注視行動を、各選択肢の属性情報を利用して統計的に解析することで、閲覧者の興味(評価基準)をリアルタイムに推定する手法を開発した。

C. 課題解答時の対話的働きかけに対する反応分析手法の開発

学習者が課題や演習を行う際には、対話的なフォローアップ、すなわち理解度に応じた足場架けとなる情報を提示し、学習者が知識を自ら整理・構築することが望ましい。対話的働きかけに対しては言語的応答が必ずしも得られないという問題に対し、本研究では学習者の非言語的反応(表情や姿勢変化)を記号化・分節化して分析するための時系列データクラスタリング手法を開発した。

(2) 詳細

研究テーマ A「ビデオ講義映像視聴時の注視行動モデルと視聴状態推定手法の開発」

講義映像視聴時の受講者の視線から、映像への集中の有無やつまずき箇所を、視聴時の確率的注視行動モデルに基づき推定する枠組みを提案した(図1)。ここで、集中度の低下、もしくはつまずきの生じた箇所の検出には異常検知と類似の考え方を取る。すなわち、あらかじめ同じ講義を受けた複数の受講者の視線データから、受講中の典型的な視線パターンを確率モデルとして学習しておく。受講者より観測された視線データにおいて、学習された視線パターンと大きく異なる視線パターンが見られた時

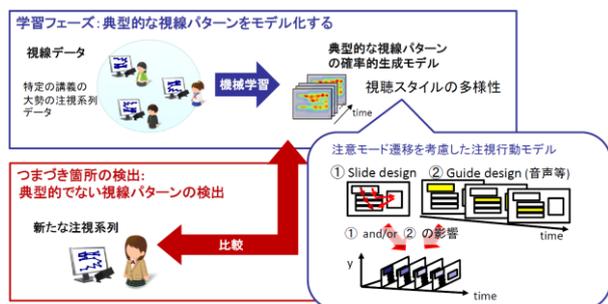


図1. 視聴時の確率的注視行動モデル

間およびスライド中の個所を、集中度の低下やつまづき箇所の候補とする(この基本的枠組みは、5.主な成果リスト(3)の4を参照のこと)。

本提案手法は、受講者の注意モードを導入する点に特徴がある。受講者の典型的な視線運動には、講師の音声で説明する内容や、動作・指示棒に対応する箇所を見るといった「ガイド追従モード」と、先読みなどのように、講師の話には追従せずに自分のペースでスライドを見る「ガイド非追従モード」が考えられる。この2つの注意モードをそれぞれモデル化し、視聴時の典型的な注視行動をこれらの混合モデルとして記述することで、複数受講者の視線データから、どのモードで視聴しているかを確率推論しながら機械学習を行うことを可能にした。

研究テーマB「カタログ閲覧時における興味推定手法の開発」

コースの受講や映像視聴は、学習者が時間的リソースをかけて行う活動であり、どのコース・映像を選択するかという意思決定は、選択者の興味・関心と十分に合致することが重要となる。そこで、図2のようにコース・映像カタログを閲覧・選択する状況において、シス

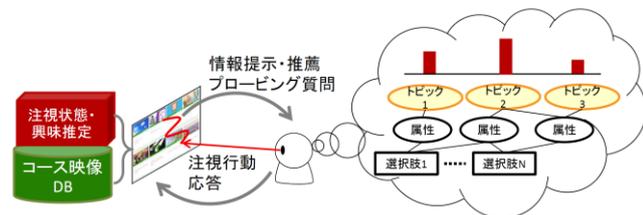


図2. 興味推定に基づく対話的選択支援

テムが対話的推薦や選択支援を行うために、閲覧者の注視行動から興味や評価基準(たとえばどのトピックに関心があるのか)を推定する以下の手法を開発した。

1. 注目属性検出のためのマルチスケール正確検定

カタログ閲覧者の興味を推定するには、その人が注目する属性を知ることが重要となる。しかし、閲覧時には選択肢の情報獲得フェーズと比較吟味フェーズが混在し(5. 主な成果リスト(1)の1を参照のこと)、さらに、選択肢が持つ複数属性のうち実際に着目している属性を分離することが難しい。そこで、各選択肢への注視を個別に解析するのではなく、閲覧者が見比べた選択肢群をまとめて解析し、有意に偏りを持った共通項として注目属性を検出する手法を考案した。これは、統計学における正確検定を様々な長さの解析時間窓に適用する手法であり、いつどの属性に、どれだけの長さ注目していたかをリアルタイムに検出することが可能となる(5. 主な成果リスト(3)の5)。

2. 選択時の評価基準候補の自動獲得と興味推定

選択時にどのような評価基準が用いられるかは、マーケティングの分野ではラダリングなどのインタビュー手法で獲得することがあるが、長時間かつコストをかけた調査が必要となる。そこで、上記1の手法を潜在意味解析に類似した機械学習と組み合わせ、カタログ閲覧を行う大勢の視線データから、閲覧時に「どの属性の組み合わせが注目されやすいか」を抽出することで、選択者の評価基準候補を自動獲得する手法を提案した。さらに、学習されたモデルに基づき、観測された視線運動の背後にある興味が評価基準に対する重みとして推定可能であることを示した。

研究テーマ C「課題解答時の対話的働きかけに対する反応動作分析手法の開発」

課題の解答時や演習時には、学習者の理解度に応じた対話的なフォローアップがしばしば行われる。その理解度を推定するうえで、学習者からの言語的応答は必ずしも得られるとは限らないため、人同士の対話では、表情や姿勢変化といった非言語的な反応を合わせて用いることで相手の理解度を推定している。しかし、非言語的反応の特徴は学習者によってそれぞれ異なり、表情変化のパターンも多様である。そこで、学習者に対して質問等の働きかけをシステムが行った際に観測される、人の表情変化や頭部動作の時系列データから、手動で定義されるよりも粒度の細かな動作特徴を抽出する時系列クラスタリング手法を新たに考案した(5.主な成果リスト(3)の 3)。これは、時系列データをスケールスペースと呼ばれる信号処理手法を利用して始めに細かく分節化し、各区間のパターンがそれぞれ異なる確率的時系列モデルに従うとして、多くの時系列モデルを推定しておく。その後、モデル間の距離定義に基づいて類似度行列を求め、この行列分解に基づいてクラスタリングを行う手法である。これにより、各個人に特有でかつ様々な時間スケールの反応動作特徴を、人手によるアノテーションを行うことなく記号化・分節化できるため、非言語的反応に基づく理解度推定の基盤技術として利用できる。

3. 今後の展開

今後はそれぞれの研究テーマで得られた知見を基に実フィールドでの評価を進める。

【研究テーマ A の展開について】

視聴状態推定の結果を要約し、学習管理システムや e ポートフォリオとの連携により学習者へフィードバックする。つまり、講義中のどの用語やトピックでつまづきが生じ、復習が必要かを可視化・推薦するシステムを構築し、学習効果を調査する。さらにこれを発展させ、映像コンテンツの最適化を試みる予定である。具体的には、講師の配置や発話の韻律情報、さらにはスライドのレイアウトや顕著性を変化させ、研究テーマ A の手法によりその効果を評価することを繰り返すことで、講義映像の最適化が可能と考えられる。既存の映像コンテンツを、注意・集中が適切に持続するよう自動再編集する技術を新たに確立する。

【研究テーマ B の展開について】

選択時の対話的やりとりが、選択後の行動にどのような効果を与えるかについて、特に学習コンテンツをどれだけ継続的に視聴するかという指標で調査を行う予定である。さらに長期的には、画面の前に座る状況だけでなく、学習者の日々の活動の中で、いつどの内容を視聴、予復習、もしくは演習するとよいかを推薦するシステムへ展開する。また、選択時の興味に基づく対話が選択後の行動へ与える影響の解析は、分野は異なるものの、購買後不協和 (post-purchase dissonance) とも密接に関係があり、学習コンテンツの選択のみならず、マーケティングやチャットボット型 e コマースをはじめ、広い分野への応用が考えられる。

【研究テーマ C の展開について】

教育工学における対話的学習支援システムでは、言語的な反応を基に次の足場架けとなる情報を学習者へ提示しているが、本研究で得られたマルチモーダルな反応分析を対話的学習支援システムへ組み込むことで、学習者からの言語的応答が不十分な場合

でも理解度を推定し、適切な学習支援につながることを示す。さらに、これを各種端末機器で利用できる環境を構築する。

4. 評価

(1) 自己評価

(研究者)

【研究目的の達成状況】

当初計画していた心的状態推定については、視線を用いた興味推定や視聴状態推定などの各手法を開発することができ、実験室規模での有効性評価を行った。これらの手法は、人への働きかけを制御するうえでの基盤技術となるが、一方で、評価のサイズは依然として十分ではなく、さらにフィードバックの効果、および学習者とのインタラクションを行う複数のノードを結んだ機械学習については、予備実験やアルゴリズムの検討にとどまり、十分な評価には至っていない。これについては、3. で述べたように、今後、実フィールドでの運用を行うシステムを構築し、数年のスパンで評価を行う計画である。

【研究の進め方】

研究実施体制は学生を主体としたものであった。計測ツールの調査とデータ収集の補助のために、数か月程度研究補助者を雇用した年度はあるものの、プロトタイプ開発等に継続的にかかわることのできる研究補助者を長期的に雇用することが望ましかったといえる。

【研究成果の社会・経済への波及効果(今後の見込みを含む)】

教育の現場では一人一人の学習者にチューターや教員が対応し、理解が不十分なトピックの相談や、日々の学習履歴を見ながらアドバイスを行う、といったことが有効であるが、時間および人的リソースの不足から実施が困難な状況も多い。本研究成果は、学習履歴に基づくラーニングアナリティクスを一步進め、学習者のリアルタイムセンシングとフィードバックを実現するものであり、短期的には学習管理システムとの統合による、より詳細な学習履歴の蓄積と支援、長期的には人工知能型チュータリングまで、大きな需要が見込まれる。計測デバイスを簡略化する、もしくは新たなセンシング手法を開発することで、広く普及可能である。本研究は社会実装によりはじめて効果が検証できる部分も多く、今後社会への普及・展開へつなげたい。

(2) 研究総括評価(本研究課題について、研究期間中に実施された、年2回の領域会議での評価フィードバックを踏まえつつ、以下の通り、事後評価を行った。)

(研究総括)

教育は社会の重要な基盤サービスであり、本研究は情報通信技術の変革に伴って大きな変化点に立っている社会における重要な研究課題への挑戦である。学習者の基本的な情報の認知や様々な意思決定、学習者と教師や教材のインタラクションなど、教育学および情報科学として解明すべき問題が多種多様に存在する。本研究では、ビデオ教材と学習者のインタラクションの測定、学習者の種々の選択時の意思決定に関する興味の推定、教材と学習者の対話の分析など、MOOC などのオンライン教育を中心とした限定された教育環境に対する技術的な課題解決に取り組んで、興味ある成果を上げている。また、教育現場

における大規模な実験を行える状況にまでは至っていないものの、基本的な問題点の抽出と課題の解決の方向性を見出した点は評価できる。今後は、Society5.0時代の教育のあり方に立ち返り、新しい教育のあるべき姿から課題を整理し、ここで得られた成果や経験をもとに新しい教育情報科学を立ち上げる中心的なリーダーの一人となることを期待する。現在進めている Learning Analytics の研究グループへの参加や教育学や心理学の専門家との交流は重要であり、本さがけ領域を中心に得た人脈も活かしながら、研究者として新しい学問領域の立ち上げに果敢に取り組む基礎を築けたと考える。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

1. Erina Ishikawa Schaffer, Hiroaki Kawashima, and Takashi Matsuyama, "A probabilistic approach for eye-tracking based process tracing in catalog browsing", Journal of Eye Movement Research, 2016, 9 (7:4), 1-14.

(2) 特許出願

研究期間累積件数:0件

(3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

1. Singo Sawa, Hiroaki Kawashima, Kei Shimonishi, and Takashi Matsuyama, "Modulating Dynamic Models for Lip Motion Generation", The fourth International Conference on Human Agent Interaction, 2016.10.4-10.7.
2. Tomoki Nishide, Kei Shimonishi, Hiroaki Kawashima, and Takashi Matsuyama, "Voting-Based Backchannel Timing Prediction Using Audio-Visual Information", The fourth International Conference on Human Agent Interaction, 2016.10.4-10.7.
3. 川嶋宏彰, 松山隆司, "スケールスペースによる初期分節化を利用したモデルベース時系列クラスタリング", 第19回画像の認識理解シンポジウム (MIRU), PS1-71, 2016.
4. 植木康介, 川嶋宏彰, 松山隆司, "ビデオ講義における映像・音声情報を用いた受講生の視線予測モデル" 情報処理学会第78回全国大会, 6ZA-01, 2016.3.12. (学生奨励賞)
5. Kei Shimonishi, Hiroaki Kawashima, Erina Ishikawa Schaffer, and Takashi Matsuyama, "Tracing Temporal Changes of Selection Criteria from Gaze Information", ACM Intelligent User Interface (poster), pp.9-12, Sonoma, California, 2016.3.7.