

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 10 日現在

機関番号：17601

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25330407

研究課題名(和文) 講義中の受講者の振る舞いと理解度の関係解析

研究課題名(英文) Analysis of the relation between behavior of the students and their understanding level in class

研究代表者

椋木 雅之 (MUKUNOKI, Masayuki)

宮崎大学・工学部・教授

研究者番号：20283640

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：本研究では、講義中の受講者の振る舞いから、その受講者が講義内容をどの程度理解できているかを推定することを目指す。そのために、「ビデオカメラ等で観測した受講者の振る舞いを計算機により自動的に分類する技術」、「受講者に対して講義中に実施する講義の理解度アンケートや講義内容に関する小テストの結果から、受講者の理解度を客観的に推定する技術」、「求めた客観的理解度を Ground Truth とみなして、分類した受講者の振る舞いと関連を求めることで、受講者の振る舞いから客観的理解度を推定する技術」の3つの技術を開発する。

研究成果の概要(英文)：In this study, we develop a method for estimating the students' understanding level from their behavior in a classroom. In order to achieve that purpose, we develop following three technologies, that is, 1. the technology to classify a student's behavior from their videos automatically, 2. the technology to estimate the objective understanding level of the students from questionnaires and quizzes, and 3. the technology to clarify the relation between students' behavior and their understanding level.

研究分野：画像認識

キーワード：振る舞い認識 理解度推定 教育評価 知的学習支援 e-learning 学習ログ解析 学習様態獲得

1. 研究開始当初の背景

近年、我が国では、高等教育における FD (Faculty Development) が広がりを見せている。FD の一環として講義評価は多くの大学で実施されており、FD の中でも重要な位置づけにあると考えられる。講義評価の例として、講義過程を記録したビデオを他者が視聴し、講師の教示法について評価を行うという取り組みがある。このような取り組みでは、講義評価の研究者と講師がビデオを見ながらディスカッションを行うことで、講義における教示法の良い点、問題点が明らかになり、講義改善につながる。しかし、その効果は、ディスカッションにおける参加者の経験や直感に依存しており、客観的、定量的に講義を評価することは難しい。

ここで、講義において、受講者は講師の教示に対して反応を返し、講師はその反応を見て教示法を決定することが多い。すなわち、受講者の反応は、講義を成立させるに当たって重要な役割を担っていると言える。このことから、我々は、講義とは本来的に講師と受講者の双方向コミュニケーションと捉えており、講義評価においては講師の教示に対する受講者の反応にも注目する必要があると考えている。このとき、大学講義における一般的な形式である一斉講義においては、受講者の反応は主に非言語的な反応、すなわち振る舞いとして外部から観測される。そのため受講者の振る舞いは、講師受講者間のコミュニケーションが成立したか否かを評価する際の客観的、定量的データとなると考えられる。しかし、多数の受講者が受講する講義において、受講者の振る舞いを定量化する作業は、人手で行うには非常にコストがかかるため、半期 15 回程度の講義に適用することも困難である。即ち、多数の受講者が受講している講義室において、個々の受講者の振る舞いを自動的に分類する技術が望まれている。

一方で、講義の目的の一つは知識の伝達と言えるため、講義を評価する指標として受講者の講義内容の理解度も挙げられる。受講者の理解度が高い講義は、知識の伝達が適切に行えた良い講義とみなすことができる。受講者の理解度は、受講者本人への理解度アンケートや、講義内容に関する小テストで測定できると考えられる。しかし、理解度アンケートでは受講者の個人差によるばらつき、小テストでは問題の難易度差のばらつきがあり、これらの結果をそのまま使っても、理解度を測る客観的な指標とはならないと考えられる。偏差値等の指標を用いれば、より客観的な指標となるが、1 科目の講義中で実施する数問の小テストでは、このような指標自体の信頼性が問題となる。即ち、講義において実施できる程度の比較的小規模な理解度アンケートや小テストに基づいて、客観的に利用できる理解度の指標が得られることが望まれている。

講義における受講者の振る舞いのデータ

と、その受講者の客観的理解度の組みが得られれば、どのような振る舞いが講義の善し悪しと関連しているかを解析することが可能となる。しかし、この関連性は自明ではない。原らは講義科目の最終回に授業アンケートを実施し、学生に自身の受講への取り組み姿勢を申告させ、試験の成績の関係を分析しているが、客観的なデータには基づいていないと言いが難い。より多くの受講者の振る舞い観測データに基づき、講義内容のトピック毎といった詳細な単位でデータを分析することで、データ主導による客観的な講義評価法の確立が望まれている。

2. 研究の目的

本研究では、以下の 3 つの技術を確立することを目指す。

(1) ビデオカメラ等で観測した受講者の振る舞いを計算機により自動的に分類・ラベル付けする技術

多数の受講者のいる講義室で、個々の受講者の振る舞いを認識する技術を開発する。近年、Kinect 等人物の動作を認識する技術が進歩してきているが、講義室のような広い空間に多人数がいる環境での振る舞い認識には、依然、技術的に課題が残っている。本研究では、このような環境で、実用性のある振る舞い認識を行う技術を実現することを目指す。

(2) 受講者に対して講義中に実施する講義の理解度アンケートや講義内容に関する小テストの結果から、受講者の理解度を客観的に推定する技術

我々は、比較的少数の理解度アンケート・小テストの結果に基づいて、項目反応理論を援用して、客観的理解度を推定する手法を開発している。より多くの講義についてデータを収集し、この手法を適用することで、手法の有用性を評価する。

(3) 求めた客観的理解度を Ground Truth とみなして、分類した受講者の振る舞いと関連性を求めることで、受講者の振る舞いから客観的理解度を推定する技術

振る舞いと理解度との関連は、自明ではないため、両者の組みからデータマイニングの手法を用いて、データ主導で関連性を求める。従来の講義評価で挙げられている教示法の善し悪しと、比較することで、求めた関連性の妥当性について考察する。受講者の振る舞いは外部から容易に観測できるため、振る舞いから理解度を推定できると、実用上有用と考えられる。

3. 研究の方法

(1) ビデオカメラ等で観測した受講者の振る舞いを計算機により自動的に分類・ラベル付けする技術の開発

講義室に設置したカメラから得られた受講者映像の振る舞いを分析することを目指す。振る舞いとは、人物の姿勢を時系列順に並べたものから意味のある区間を抽出したものである。また姿勢とは、ある瞬間での部位位置の集合であり、画像中の関節位置によって表現される。注目すべき振る舞いの定義については研究ごとに異なり、明確な指標が存在しない。そこで本稿では、振る舞いの構成要素である姿勢を獲得することを考える。

人物姿勢推定を妨げる要因として、人間の取りうる姿勢の多様性と、衣服や背景の変化による人物の外観の多様性が挙げられる。これらの要因によって、画像中での人物とその周辺は非常に大きな多様性をもち、姿勢推定の妨げとなる。多様性を制限した状況では、比較的高い性能を発揮する検出器を作成することが可能だと考えられる。ここで検出器は、画像中の関節位置の検出を行う。特定の状況に特化した検出器を複数作成し、推定対象となる画像に対してはそのいずれかに適切に振り分ける分類器を挟むことで、全体として高性能の検出器を構成することができる。この手法では、どのように多様性を制限するのか、どのように画像を分類するのか、の2点が大きな問題となる。本研究では、多数の受講者画像から得られた知見に基づいて3つのクラスを定義し、それぞれに適した検出器を作成することで推定精度の向上を図る。

(2) 受講者に対して講義中に実施する講義の理解度アンケートや講義内容に関する小テストの結果から、受講者の理解度を客観的に推定する技術の開発

受講者の振る舞いと理解度の関係を解析するためには、あらかじめ受講者の理解度を推定しておく必要がある。理解度を推定する手法としては、インタビューによる方法や、多数の問題を出題しその正解率をみる方法、講義の前後でテストを実施しその成績の変化をみる方法など、種々の方法が考えられるが、実際の講義で実施するには、講義の進行を妨げず、受講者に負担とならない方法が必要である。本研究では、受講者自身に理解度を申告してもらう理解度アンケートによる方法と、問題を出題しその正解率を用いる小テストによる方法を併用する。講義を数個のスパンに分け、スパン毎に理解度アンケートと小テストを実施する。スパンとして30分程度の比較的短い期間単位に講義を区切ることで、受講者の振る舞いと対応付けやすい単位での理解度推定を目指す。

(3) 小テスト等で求めた客観的理解度をGround Truthとみなして、分類された受講者の振る舞いと関連を求めることで、受講者の振る舞いから客観的理解度を推定する技術の開発

これまでに、受講者の振る舞いと客観的理

解度の組みが得られるので、それらを用いて、関係性の解析を行う。この問題は、学習データに基づくパターン認識の問題として定式化できる。しかし、受講者の振る舞いと理解度の関係は単純ではなく、ある振る舞いが理解度と1対1に対応しているとは限らない。これに対して我々は、受講者は振る舞いと理解度の関係において、いくつかのタイプに分類することを目指す。

4. 研究成果

(1) ビデオカメラ等で観測した受講者の振る舞いを計算機により自動的に分類・ラベル付けする技術

手法

本手法は、検出器の学習と、それを用いたテストデータの姿勢推定の2つのステップから構成される。具体的には以下の手順となる。まず学習データ集合に対してクラス分類器による分類を行い、学習データ集合をクラスごとの小集合に分ける。次に小集合ごとに異なる検出器(以下、各クラス用部位検出器)を用意し、検出器に対応する小集合を学習させる。これにより、各クラスに対応した検出器を得ることができる。推定ステップでは、まず受講者画像をクラス分類器に入力する。次に分類結果に基づいて、対応する各クラス用部位検出器に受講者画像を入力し、推定姿勢データを得る。

受講者の部位の隠れに対応して、「頭が首や前腕部を隠す状態(クラス1)」「胴体の前で一方の前腕部が他方の前腕部を隠す(クラス2)」「その他(クラス3)」という3つのクラスを定義する。クラス1,2は姿勢推定が難しいものの、姿勢の多様性は小さい。クラス1,2を分離して扱うことで多様性を制限することが可能となり、これらに特化した性能の高い検出器を作成することができる。クラス3については、姿勢推定が難しいクラス1,2と分離することで性能の向上が見込まれる。

受講者画像に対して、まず顔検出を行う。顔が検出されなかった画像をクラス1とし、顔が検出された画像では、画像中の受講者の胴体周辺における横方向エッジの強さを求める。横方向エッジが十分強い場合にはクラス2とし、そうでない場合にはクラス3とする。

実験結果

提案手法と、クラス分類器を用いず単一の部位検出器で全てのクラスの学習データを学習する手法(以下、従来手法)を比較する。部位検出器にはYi Yangらの手法を用いる。Yi Yangらの手法では上半身が映った人物画像が入力されると、上半身の各部の座標値を求める。このうち頭頂部、首、両肩、両肘、両手首の座標値を対象とする。受講者姿勢は振る舞い分析に利用されるため、正解姿勢から大きく外れた姿勢が推定されることは望

ましくない。そこで、画像の大きさを縦 300 ピクセル、横 240 ピクセルとし、正解姿勢と推定姿勢の誤差が 20 ピクセル以内となる確率を評価尺度として用いる。人物画像には、京都大学で実際に行われた講義を撮影した映像中の受講者を 3 人分使用する。各画像にはあらかじめ正解クラスを手動で割り当てる。実験では第一に、クラス分類器の性能を測る。これは、各受講者について leave-one-out 交差検定を行うことで精度を求める。その後、各受講者について姿勢推定を行う。こちらも各受講者について leave-one-out 交差検定を行い、推定姿勢と正解姿勢の誤差が 20 ピクセル以内となる確率を求める。

クラス分類の結果を表 1 に示す。推定クラス 1 では精度が高くないものの、再現率が 100% となっている。また、正解クラス 2, 3 では、80% 以上の精度で正しい分類が行われていることが分かる。受講者は正面上方から撮影されており、顔が見えていても正しく顔検出が行われないことが多かったため、推定クラス 1 の精度が低くなったと考えられる。

誤差が 20 ピクセル以内となる確率を表 2 に示す。提案手法によって誤差が 20 ピクセル以内となる確率が大きくなるのが分かる。推定クラスを使用した場合、正解クラスを使用するより確率は小さいが、従来手法と比較して 30% ほど改善されている。

提案手法によって、誤差が 20 ピクセル以内となる確率は従来手法のおよそ 2 倍となった。村上らは、約 80% の精度で正解姿勢から講義への興味度を正しく推定している。本稿の推定姿勢を利用した場合には精度がより低くなることが予想されるものの、分類器を利用した姿勢推定の有効性が示された。

表 1：クラス分類結果(枚)

推定 クラス	正解クラス			計	精度
	=1	=2	=3		
=1	113	66	97	276	0.41
=2	0	79	4	83	0.95
=3	0	16	100	116	0.86
計	113	161	201	475	
再現率	1.0	0.49	0.50		

表 2：誤差が 20 ピクセル以内である確率

手法	使用クラス	確率
従来手法	不使用	0.35
提案手法	推定クラス	0.65
提案手法	正解クラス	0.69

(2) 受講者に対して講義中に実施する講義の理解度アンケートや講義内容に関する小テストの結果から、受講者の理解度を客観的に推定する技術

手法

理解度アンケートは、受講者の意識による主観的な理解度を測る指標である。アンケー

ト形式は五件法とし、選択肢は以下の五つとする。「1:非常によく理解できた」「2:よく理解できた」「3:だいたい理解できた」「4:まあまあ理解できた」「5:あまり理解できなかった」。

小テストは、出題した問題に対する正解率により理解度を測る指標である。小テストの形式は、直前のスパンでの講義内容に関する記述を提示し、その正誤を問う正誤問題を 5 問出題する。5 問中の正解した割合を、正解率とする。

実際に、異なる受講者が理解度アンケートに同じ回答をしても、受講者によって小テストの正解率が大きく異なる傾向があった。これは、主観的な理解度の解釈に個人差があることを示している。また、同じ受講者が理解度アンケートに対して同じ回答をした場合でも、小テストの正解率が大きく異なる場合があった。これは、小テストの問題に難易度差があったためと考えられる。小テスト間で難易度が等しくなければ、正解率の違いが小テストの難易度によるものなのか解答した受講者の理解度によるものなのかを判断できず、受講者の理解度を測る一定の評価基準になり得ない。

異なる難易度のテストを用いて解答者の能力を測る際に用いられる理論の一つに、項目反応理論がある。項目反応理論とは教育学分野における「試験」に関する理論であり、問題の正答率 p と解答者の能力 θ の関係が下式のロジスティック曲線で表現されるとするものである。

$$p(\theta) = c + (1-c) / (1 + \exp(-1.7 a (\theta - b)))$$

また、スパン j に対応する理解度アンケートにおける受講者 i の回答を x_{ij} 、このスパンの内容に対する受講者の能力（客観的理解度）を θ_{ij} として、両者の間に以下の関係が成り立つと仮定する。

$$x_{ij} = \theta_{ij} + y_i (x_{ij} - 1)$$

ここで、 x_i は受講者 i が理解度アンケートに「1:あまり理解できなかった」と回答した際の基準理解度であり、 y_i は受講者 i の理解度アンケート回答における各段階の基準理解度の間隔である。

本研究では、上式で得られる客観的理解度と小テストの成績との間に項目反応理論が成り立つと仮定し、各パラメータを求めることで、理解度アンケート回答の個人差、小テストの難易度差を補正する。

実験結果

実験では、京都大学で行われた 5 回の講義を通して、理解度アンケート・小テストを 13 回実施した。これにより 13 人の受講者から 73 組のデータが得られた。個人差、難易度差の補正結果の例を図 1 に示す。補正により、

ロジスティック曲線によくフィッティングしていることがわかる。

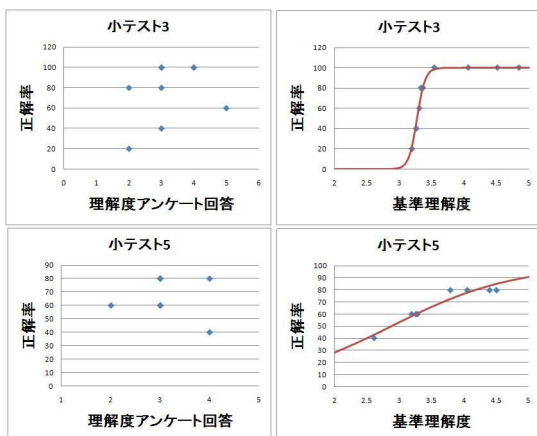


図 1: 理解度の補正結果

(3) 小テスト等で求めた客観的理解度を Ground Truth とみなして、分類された受講者の振る舞いとの関連を求めることで、受講者の振る舞いから客観的理解度を推定する技術

手法

講義をいくつかのスパンに分割し、各スパンで受講者の振る舞いと理解度のデータを収集する。1 スパンにおける各分類に属する振る舞いの頻度に基づき、多様な振る舞いのデータを定量的に収集する。受講者の振る舞いと理解度を関連づける手法には、大きく分けて「理解度の高い、あるいは低い受講者にどういった振る舞いが見られるかを分析する」「振る舞いにある特徴の見られる受講者の理解度を分析する」という二通りのアプローチが考えられる。ここで、講義において見られる受講者の振る舞いは様々であるが、それらと受講者の理解度との関係性が一意に定まるとは必ずしも言えない。また、一つの振る舞いと理解度との関係性の他に、複数の振る舞いの組み合わせと理解度との関係性があることが予想できる。そのため、受講者の振る舞いと理解度との関係性を明らかにするためには、「振る舞いにある特徴の見られる受講者の理解度を分析する」アプローチがより合理的であると考えられる。そこで本研究では、講義において観測可能な受講者の振る舞いに基づいて受講者をクラスタリングし、受講者の理解度と関連づけることで、講義における受講者の振る舞いと理解度との関係性を明らかにする。

具体的には、講義の各スパンにおける各受講者の振る舞い頻度を特徴ベクトルとし、k-means 法によりクラスタを生成する。k-means 法では、クラスタ数 k やクラスタ中心の初期値により、結果が変動する。そこで、種々の k および初期値によりクラスタリングを行い、信頼性の高い結果を採用する。信頼性の評価には、情報量基準 AIC を用いる。

得られたクラスタに対して、振る舞い頻度、基準理解度（受講者の意識による理解度、標

準化した理解度、難易度差を補正した正解率に基づき、各クラスタに属する受講者のタイプを目視により分析・意味付けする。

実験結果

京都大学で実施した講義 5 回分について、13 個のスパンを設定して、理解度アンケートと小テストを実施した。講義は、最大受講者 13 人が受講しており、計 73 組のデータを得た。このデータをクラスタリングした結果、以下の 5 つの典型的なタイプが得られた。

受動的タイプ：受動的受講行動の頻度が非常に高く、客観的理解度も母平均よりも高い。
PC タイプ：PC 端末閲覧・操作の頻度が非常に高く、客観的理解度は、受動的タイプには劣るものの母平均よりも高い。

逸脱タイプ：逸脱行動の頻度が非常に高く、客観的理解度は母平均を下回っている。

能動的タイプ：能動的受講行動の頻度が高く、客観的理解度は最も高い。

散漫タイプ：散漫行動の頻度が他のクラスタと比較して高く、受動的受講行動の頻度がやや低い。受動的タイプと比較して客観的理解度がやや低くなっているが、母平均とは等しい値である。

以上の 5 つのタイプから、能動的受講行動や受動的受講行動、PC 操作といった行動が客観的理解度と正の相関を持ち、逸脱行動が負の相関を持つという傾向が確認できた。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表](計 3 件)

相澤将吾, 椋木雅之, 美濃導彦, 三功浩嗣: 部位の重なりに応じた異なる検出器による受講者画像からの姿勢推定, 情報処理学会第 77 回全国大会, 2015/03/19, 京都府京都市。

永田裕太郎, 村上正行, 森村吉貴, 椋木雅之, 美濃導彦: MOOC における大規模学習履歴データからの受講者の学習様態獲得, 人工知能学会第 73 回先進的学習科学と工学研究会(SIG-ALST), 2015/3/5, 愛知県新城市。

川西 康友, 伍 洋, 椋木 雅之, 美濃 導彦, 勞 世竝: サーベイランスカメラ映像を対象とした人物照合の高精度化への取り組み, 電子情報通信学会 PRMU 研究会, 2013/12/13, 三重県津市。

6. 研究組織

(1) 研究代表者

椋木 雅之 (MUKUNOKI Masayuki)

宮崎大学・工学部・教授

研究者番号: 20283640