

機関番号：33919
 研究種目：基盤研究(c)
 研究期間：2008～2010
 課題番号：20500814
 研究課題名（和文） 学生の動きと教示情報との相関から講義中の学生の集中度を評価するシステムの開発
 研究課題名（英文） Development of a system that evaluates student's concentration from correlation of teacher's presentation and student's movement
 研究代表者
 田中 敏光 (TANAKA TOSHIMITSU)
 名城大学・理工学部・教授
 研究者番号：00262923

研究成果の概要（和文）：

本研究では、プロジェクタを使用した講義において、表示を切り替えたときに、学生がスクリーンを注視したり表示された文字を書き写したりすることに着目して、表示切り替え時刻における学生の動きや顔の向きの相関を調べることで、学生の講義に対する集中度を評価するシステムを開発した。また、この手法を、黒板を用いた一般の講義に適用するため、黒板に書かれた文字を画像処理により検出し、板書開始時刻を求めるサブシステムを作成した。

研究成果の概要（英文）：

This research developed a system that evaluates concentration of each student in a lecture room. In the lectures using presentation tools like Power Point, when the projected image is changed, usual students gaze screen or take notes. From this view point, the system tracks head of each student during the lecture, and then detects the time when the display is changed. If a student begins to move own head in a short delay from the change, the system judges that he or she is concentrated. Even if the head motion is not detected, when a student watches the screen continuously, the system gives the same judgment. In addition since classes using blackboard are popular in high schools, a subsystem was developed in order to apply the judgment system to the classes. The subsystem detects the start time of teacher's writing something on a blackboard.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	700,000	210,000	910,000
2009年度	800,000	240,000	1,040,000
2010年度	1,100,000	230,000	1,330,000
年度			
年度			
総計	2,600,000	680,000	3,280,000

研究分野：総合領域

科研費の分科・細目：科学教育・教育工学 教育工学

キーワード：授業学習支援システム，集中度評価，動き検出，頭部検出，顔検出，板書判定，画像処理

1. 研究開始当初の背景

教師の教育力の低下は多くの教育現場で問題となっている。また、大学や高校では、進入生の学力低下が深刻な問題であり、授業についていけない学生が増えている。しかし、

一般の授業、特に大学の講義は、多数の学生を相手に行われるので、授業中のよそ見や居眠りなどを教員がその場で全て把握することは困難である。このため、授業の内容や方法を改善するには、学生の受講状況を調べて

報告する補助手段が必要となる。

受講状況の調査では、アンケートを分析する手法が広く用いられている。また、一部では、他の教員や学外者による授業評価も行われている。しかし、このような報告者の主観にたよる方法では、講義の状況を総合的かつ客観的に評価することは難しい。

そこで、学生の動き（特に頭の動き）を調べることで客観的に講義を評価する研究が行われてきた。これまでの研究で、学生がほとんど動かないことから居眠りをしていると判定したり、顔が横を向く時間が長いことから授業を聞いていないと判断したりすることはできている。しかし、従来手法では、学生に動きがあっても、それが講義を聞いているために生じているのか、マンガを読んだり内職をしたりしているために生じたのかを区別することはできない。

2. 研究の目的

講義中の学生を撮影した動画像から、学生の動きを検出し、学生の授業に対する集中度を客観的に評価するシステムを開発する。学生の集中度は、教員が情報を与えるタイミングと学生の動きとの相関から判定する。

まず、大学の講義では一般的な形式である、Power Point等のプレゼンテーションツールを用い、資料をスクリーンに投影して提示する講義を対象として、集中度を判定する手法を開発する。このような講義では、教員が提示する情報が変わるタイミングは明確で、容易に検出できるため、学生の動きの検出と同期の判定方法を求めることが、研究の主目的となる。

次に、開発した手法を、高校までの授業では一般的な黒板を用いた講義に適用するため、教員の板書行動を検出し、板書を始める時刻を求める手法を開発する。

3. 研究の方法

(1) プロジェクタを用いた講義における集中度の評価

① 受講者の撮影方法

市販のハイビジョンビデオカメラで受講者を撮影する。ビデオカメラは、図1に示す様に、講義室前方の教材提示用スクリーンの左右に、それぞれ設置する。高さは、スクリーンのほぼ中央となるように調整する。このようにカメラを配置するのは、スクリーンを見ている受講者をなるべく正面から撮影することで、正面顔の検出を容易にするためである。また受講者を追跡する際、画像上で受講者が重なることを防ぐため、図1のようにスクリーン左のカメラで講義室の右側を、右側のカメラで講義室の左側を撮影する。

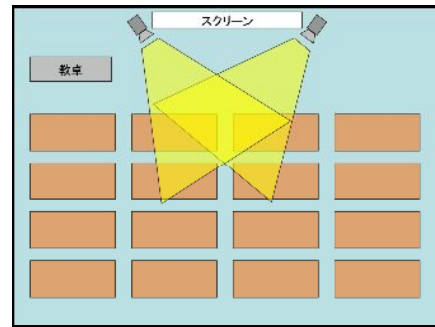


図1 受講者を撮影するカメラの配置

② 生徒領域の検出

頭部の位置と移動量の時系列から、画像上で、個々の受講者の頭部が動く範囲を推定し、矩形で近似する。これを生徒領域と名付ける。そして、それぞれの生徒領域ごとに、頭部の移動量を動作情報として抽出する。また、この中で正面顔の検出を行う。以下に、これらの手法の詳細を述べる。

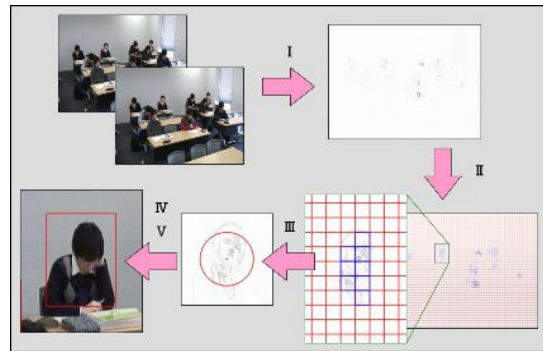


図2 生徒領域の検出手順

手順Ⅰ：撮影した動画像より一定フレームごとに画像を切り出し、RGBそれぞれで差の絶対値を求め、その平均値を濃淡画像として保存する。これを白黒反転したものを動作差分画像とする。例えば、図2左上に示した2枚の画像からは、右上の動作差分画像が生成される。

手順Ⅱ：動作差分画像を正方形格子で分割し、小領域ごとに画素値の合計を求め、値の高い領域を一定数選択する。選んだ小領域が4近傍で隣接していれば、それらをまとめてグループを作る。

手順Ⅲ：それぞれのグループについて、それを包括する初期円を決め、半径を縮小しながら中心を円内の画素の重心方向に移動する処理を繰り返すことで、円をⅡで求めた領域にフィッティングする。処理が終了したときの円の中心と半径から、頭部の位置と大きさを推定する。

手順Ⅳ：それぞれの時刻で求めた頭部位置のうち、時間軸上で連続しているものをグループにまとめる。

手順Ⅴ：Ⅳで求めたグループのそれぞれで、Ⅲで求めた頭部円の中心と半径の値を平均

する。そして、平均中心から、左右は平均半径の3倍、上下は平均半径の4倍の範囲を、生徒領域とする。

③ 動作情報の取得

講義室の照明条件は一定のため、生徒領域が正しく取得できれば、その中で生じる輝度変化は受講生の動きに起因するものと考えてよい。そこで、それぞれの生徒領域について、動画像1フレームごとに、前のフレームとの輝度値の差の絶対値を画素ごとの計算し、その合計をその生徒領域の動作情報とする。

④ 正面顔の検出

OpenCV(Intel Co.)の cvHaarDetectObjects 関数(前もって学習された分類器を用いてオブジェクトの検出を行う関数)を用いて、正面顔を検出する。本研究では、受講者がスクリーンの方向、つまりカメラの方向を向いているか否かが検出できればよいので、haarcascade_frontalface_alt.xml (OpenCVのサンプル)に含まれている、正面顔の特徴を学習した分類器を用いている。

⑤ スライド切替え時刻の取得

表示に用いる PC の画面を一定の時間間隔でキャプチャし、隣り合う時刻で画素値の差の絶対値を求めることで、差分画像を作る。そして、

1. 差分画像の画素値の合計が閾値を超える。
2. その後10秒以内に閾値を超える時刻が存在しない。
3. 値が変化する画素が、画像上に広く分布する。

の条件を満たす時刻に、表示が切り替わったと判定する。

上記の2は、スライドを連続して切り替えた時に、最後の時刻だけを検出するための条件である。また、3は、アニメーション等による部分変更を排除するための条件である。部分変更でも書き写しが必要となることはあるが、変化の範囲が小さいと安定して検出できないため、本手法では対象外としている。

⑦ 動画区間の検出

動画を流している間は、スライドの切り替えよりずっと小さい差分値が連続して発生する。そこで、動画検出用の小さな閾値を設定し、5フレーム連続して差分値がこの閾値を越えれば、その区間は動画であると判定する。時系列上に複数の動画区間が見つかった場合には、その間の時間差を調べ、一定の値より短ければ結合する。そして、その先頭の時刻を表示切り替え時刻として記録する。

⑧ 集中度の判定

図3は、教示情報の切り替わりタイミング時刻と受講者の動き、正面顔の検出結果を模式的に示したグラフである。PowerPoint等のプレゼンテーションツールを使った授業では、スライドを変えたときに教示情報の切り替わりになるため、切り替えがあった時刻を

1, その他を0とするパルス波形で表すことができる。また顔の向きについても、正面顔で判定しているため、正面顔が検出できた場合を1, できない場合を0で表すことができる。

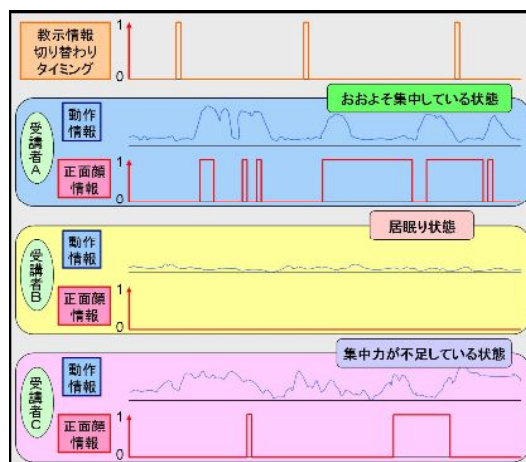


図3 受講者の動作の模式図

図中の受講者Aのように、切り替わりタイミングに少し遅れて動きが見られる場合には、揭示された情報に対して反応したと考えられ、この時刻に正面顔が検出できているので、集中していると判定して差し支えない。しかし受講者Cのように切り替わりタイミングで動いていても、正面顔がほとんど検出されなければ、他事をしている可能性が高いと考えられる。このように、顔の向きと動きを組み合わせると、集中度を判定する必要がある。

そこで、以下の基準で集中度を「高」「中」「低」「無」の4段階に判定する。

※ スライドが切り替わる時

- 切替え以前から正面顔を検出
 - 動作量が閾値以上 → 「高」
 - それ以外 → 「中」
- 切替え後10秒以内に正面顔を検出
 - 動作量が閾値以上 → 「中」
 - それ以外 → 「低」
- それ以外
 - 動作量が閾値以上 → 「低」
 - それ以外 → 「無」

※ 動画像が提示される時

- 正面顔を連続して3秒以上検出 → 「高」
- それ以外
 - 動作量が閾値以上 → 「中」
 - それ以外 → 「無」

(2) 板書行動の検出

① 黒板の撮影方法

webカメラを使って5秒間隔で画像を撮影する。横に長い黒板の場合は、図4に示すように、2台のカメラを平行に並べて撮影する。画像の大きさは、1600×1200画素とする。

講義の邪魔にならないように、カメラは最前列の席の机より低い位置に置かれるため、黒板を見上げるように撮影される。そこで、画像上の黒板の四隅の座標を用いて、撮影した画像を射影変換することで、正面から見た画像を得る。

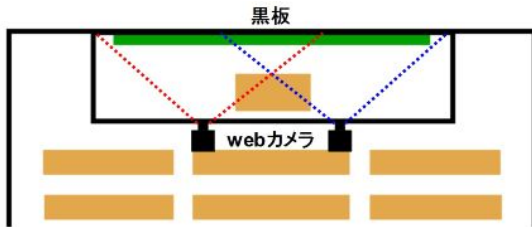
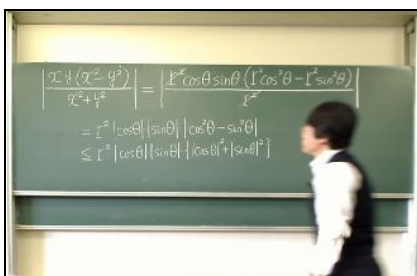


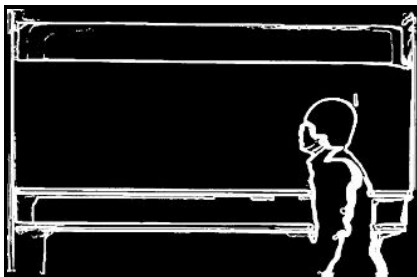
図4 黒板を撮影するカメラの配置



(a) 射影変換後の画像



(b) 平均値シフト処理の結果



(c) (b)の色エッジ



(d) 黒板領域

図5 黒板領域の抽出

② 黒板領域の抽出

教師が移動すると、図5(a)のようにぼけて写ることがある。この場合でも教師領域が正しく求められるように、前処理として、平均値シフト法により色の平滑化を行う。この画像から色エッジを求め、領域を分割する。次に、それぞれの領域の色の平均値を求め、黒板の色(濃緑色)に近い値を持つ領域を取り出す。これを黒板領域と呼ぶ。

③ 文字の抽出

図6(a)と(b)に示すように、各時刻で、入力画像から黒板領域を切り出し、ソーベルフィルタと閾値処理で文字の輪郭を抽出する。次に、その時刻には黒板領域に含まれていないが、5秒前には含まれていた領域を求める。ここに書かれている文字がその時刻に教師に隠されていた文字となるので、5秒前の微分画像からその領域を切り出し、現在の微分画像に追加することで、隠された領域を補う。

$$\begin{aligned} \left| \frac{x^4(x^2-4^2)}{x^2+4^2} \right| &= \left| \frac{r^4 \cos \theta \sin \theta (r^2 \cos^2 \theta - r^2 \sin^2 \theta)}{r^2} \right| \\ &= r^2 |\cos \theta| |\sin \theta| |\cos^2 \theta - \sin^2 \theta| \\ &\leq r^2 |\cos \theta| |\sin \theta| |\cos \theta| + |\sin \theta|^3 \end{aligned}$$

(a) ある時刻のエッジ画像

$$\begin{aligned} \left| \frac{x^4(x^2-4^2)}{x^2+4^2} \right| &= \left| \frac{r^4 \cos \theta \sin \theta (r^2 \cos^2 \theta - r^2 \sin^2 \theta)}{r^2} \right| \\ &= r^2 |\cos \theta| |\sin \theta| - \sin^2 \theta \\ &\leq r^2 |\cos \theta| |\sin \theta| \end{aligned}$$

(b) 1時刻前(5秒前)のエッジ画像

$$-\sin^2 \theta$$

(c) (b)だけに含まれるエッジ

$$\begin{aligned} \left| \frac{x^4(x^2-4^2)}{x^2+4^2} \right| &= \left| \frac{r^4 \cos \theta \sin \theta (r^2 \cos^2 \theta - r^2 \sin^2 \theta)}{r^2} \right| \\ &= r^2 |\cos \theta| |\sin \theta| |\cos^2 \theta - \sin^2 \theta| \\ &\leq r^2 |\cos \theta| |\sin \theta| + |\cos \theta|^3 \end{aligned}$$

(d) 復元されたエッジ画像

図6 教師に隠された文字の復元

③ 板書の判定

5秒前の黒板画像との差分を求め、5秒間に増加した画素(書き加えられた文字のエッジを示す)を抽出する。これをラベリングして、領域数(文字数)を求める。そして、領域数が3以上、または、最も大きな領域の画素数が90以上150以下の場合に、その時刻に板書されたと判定する。

ただし、図5(a)に示したような、上下スライド式の黒板では、黒板の移動が板書と誤判定される恐れがある。そこで、差分文字領域数が50以上の場合を黒板の移動と判定して、取り除く。

④ 板書開始時刻の決定

連続した3時刻でいずれも板書と判定されたら、その最初の時刻を板書開始時刻として記録する。ただし、文字を書く途中で短い休みが入ることもあるので、2時刻(10秒)以下の短い時間を空けて板書と判定された場合には、前の板書動作と連続していると判定して、開始時刻を記録しない。

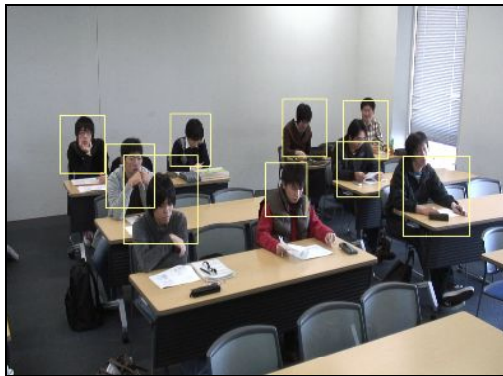
また、外光の変化などにより、板書と判定される場合があるので、15時刻以上連続して板書と判定されなかった後に、3時刻以下の短い板書が検出された場合には、判定を保留し、その後5時刻以内に板書の判定がなければ、板書を無効とする。

4. 研究成果

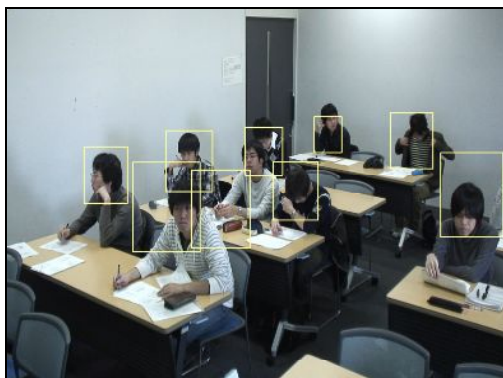
(1) プロジェクタを用いた講義における集中度の評価

① 生徒領域の検出実験

3分の動画像を入力して実験を行った。図7に示すように、左カメラの映像では9名中8名が、右カメラでは9名中7名が検出できた。認識率は83%である。この認識率は、講義の評価には十分な値といえる。



(a) 左カメラ



(b) 右カメラ

図7 生徒領域の検出

② 正面顔の検出実験

①で検出した受講者のうち、撮影条件の良い1名に対して正面顔の検出実験を行った。5.1と同じ3分の動画像を入力して処理した結

果、約90%を正しく認識できた。

③ 教示情報切り替わり時刻の取得実験

約85分の動画像を入力とし、スライドの切替え時刻と動画像の開始時刻を取得する実験を行った。処理結果を目視による判定と比較したところ、切り替え時刻では正検出42回、誤検出2回、未検出1回、動画像開始時刻では正検出8回、誤検出1回、未検出0回となり、ほぼ正しく検出することが出来た。

④ 集中度の判定実験

上記で得た結果を元に、集中度が高いと見られる受講者Aと低いと見られる受講者Bについて、人の評価とシステムの評価を比べたところ、表1に示すように、おおむね一致した。実験に使った動画像には書き写す動作が無いスライドも含まれているが、正しく検知することが出来た。

表1 目視判定との比較

	スライド切替え		動画区間	
	一致	不一致	一致	不一致
受講者A	9	2	3	0
受講者B	9	2	3	0

⑤ 結果の考察と今後の課題

実験結果から、ある程度以上の大きさで写っており、受講態度がはっきりしている受講者に対しては、目視による判定に対して70%程度の精度で判定できることが確認できた。しかし、一部が隠されていたり、小さすぎたりする対象に対しては、生徒領域の検出ができないことがある。このため、生徒領域の認識精度を向上する必要がある。

(2) 板書行動の検出

① 板書時刻の検出精度

3つの授業の黒板を撮影し、板書の開始時刻を求めた。この画像を目視により板書判定し、3章(2)節④項に示した開始時刻の判定条件を使って開始時刻を求めた。これを正解として、提案手法の適合率と再現率を計算した。結果を表2に示す。

表2 提案手法の認識率

		左カメラ	右カメラ
講義1	適合率	83.3%	90.9%
	再現率	58.8%	55.7%
講義2	適合率	64.3%	78.8%
	再現率	75.0%	70.8%
講義3	適合率	78.8%	78.8%
	再現率	68.4%	68.4%

講義により多少異なるが、おおむね70%程度の適合率と再現率が達成できている。

② 結果の考察と今後の課題

おおむね70%の検出精度は、悪い値ではないが、Power Pointを用いる講義におけるスラ

イドの切り替えタイミングの検出精度に比べると低いので、改善が必要である。

検出率が下がる原因を調べたところ、教師が写っていないカメラで板書と判定されたケースが多くみられた。現在は、左右のカメラの画像を独立に処理しているため、このような結果となってしまった。

照明の変化や教員の陰などにより差分が生じたことが誤判定の理由だが、教師がいない側では板書が生じないので、両方のカメラの教師領域の抽出結果を比較して、教師がいる側だけで判定すれば、誤判定を減らすことができると考えられる。現在、この点について改良を進めている。

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 8 件)

- ① 中島貴裕, 田中敏光, 佐川雄二, 教師の動きにロバストな板書の検出手法, 照明学会若手セミナー予稿集, 2011年3月5日, 名城大学.
- ② 中島貴裕, 田中敏光, 佐川雄二, 教師の動きに強い板書の検出手法, 電気関係学会東海支部連合大会, 2010年8月30日, 中部大学.
- ③ 澤田昌之, 田中敏光, 佐川雄二, 顔方向も考慮した講義受講者の中度評価, 電気関係学会東海支部連合大会, 2009年9月10日, 愛知工業大学.
- ④ 澤田昌之, 田中敏光, 佐川雄二, 顔方向を考慮した講義受講者の授業集中度評価, 電気関係学会東海支部連合大会, 2008年9月18日, 愛知県立大学.

[その他]

ホームページ等

<http://www-is.meijo-u.ac.jp/~tanaka/index.html>

6. 研究組織

(1)研究代表者

田中 敏光 (TANAKA TOSHIMITSU)

名城大学・理工学部・教授

研究者番号：00262923