

マルチモーダルセンサ情報を利用した遠隔指導システムにおける 効率的な通信品質制御のための指導状況モデルの構築

瀬戸淳也[†] 坂根裕[‡] 竹林洋一[†]

[†] 静岡大学大学院 情報学研究科

[‡] デジタルセンセーション株式会社

1 はじめに

ウェブに関する技術の高度化とともに、さまざまな教育機関においてインターネットを介した遠隔講義の機会が増加している [1]. 筆者らは講義形式の授業ではなく、絵画や書道、音楽、スポーツなどの演習を遠隔指導できる基盤システム構築を目指している. 一般的な遠隔授業システムでは、ビデオ会議システムを基盤として採用しており、教室内のカメラやマイクからの出力をミキシングし、最終的にはひと組の画像と音声として相互に交換する. 一方演習では、指導者は全受講者の作業を把握しつつ、適切なタイミングでアドバイスを与えることが機能的に求められる.

例えば絵画の演習では、全受講者の様子や描いている絵をデータとして指導者側に転送する必要があり、従来のビデオ会議システムでは要件を満たさない. また、データ転送量が増大することから、国を超えた遠隔授業などを考えると、ネットワークへの負荷が問題となる.

そこで筆者らは、指導者が注目している情報のデータは高品質で送信し、指導上必要のないものは低品質で送信することで、ネットワーク負荷の問題を軽減する方式を考えた. 本稿では、交換するデータの品質制御のために必要な、「指導状況」を表現するモデルの設計と品質制御の予備実験結果について述べる.

2 通信制御を行う遠隔指導システム

指導中の教室内ではさまざまな情報が飛び交っている. 例として演習中の教室内のイメージを図 1 に示す. 効率よく通信制御を行うためには、これらの情報の中から必要な物を必要なタイミングで選出し、通信制御を行う必要がある.

図 1 の遠隔指導の例では、指導者は指導のために、対象者の画像と音声が必要である. また、生徒は指導

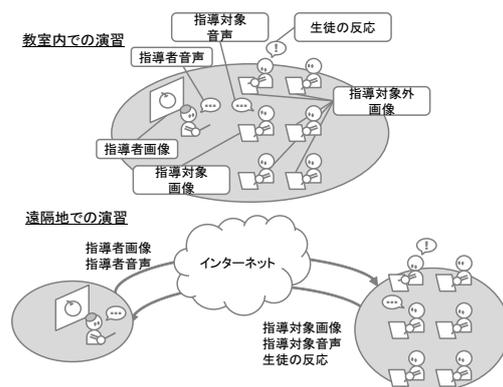


図 1: 教室内と遠隔地からの演習のイメージ

の様子を表す画像と音声が必要である. このとき、指導対象外の生徒の画像は重要度が低い.

このように、状況に応じて必要な情報は異なっているため、指導者と生徒がどのような状況にあるのかを把握することにより、通信制御を行うことが可能となる. そこで、「どのような指導をするのか」「指導者は何をしているか」「生徒はどんな反応なのか」「生徒同士は何をしているか」という 4 つの観点から指導中の状況を把握するためのモデルを構築した.

3 通信制御のための指導状況モデル

3.1 指導状況モデルの構築

まず、「どのような指導をするのか」という講義形態について考える. これは主となる情報の流れから、大学の講義のように指導者が情報を送る「講演」、指導者が教室を巡回し、生徒が情報を送る「演習」、そして、質疑応答のように互いが情報を送りあう「討論」の 3 つの形式が考えられる.

次に「指導者は何をしているか」という指導者の状態について考える. 考えられるのは、指導をしている、または生徒の作業を見守っているという 2 つであるが、指導している場合には特定の生徒に対する指導と全員への指導の 2 つが存在する. そのため、「巡回」、「個別指導」、「全体指導」の 3 つの状態とする.

Coaching situation model for efficient communication quality control in distant coaching system using multimodal sensory data

Junya SETO[†], Yutaka SAKANE[†] and Yoichi TAKEBAYASHI[‡]

[†] Graduate School of Informatics, Shizuoka University

[‡] Digital Sensation Co., Ltd.

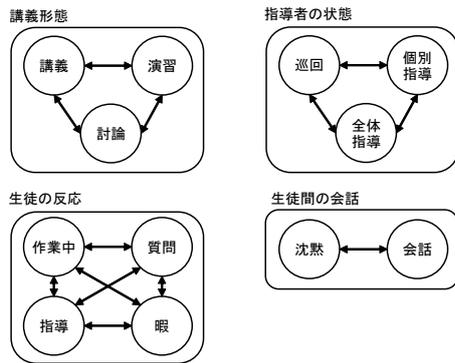


図 2: 指導状況モデルの全体図

表 1: 情報の重要度

	指導者画像	生徒画像	反応	会話
講演	高	無	中	無
演習巡回	無	中	中	低
演習指導	高	高(低)	低	低
討論	低	低	高	無

3つ目に「生徒はどんな反応なのか」という生徒の反応について考える。これは指導者の指導を受けている状態、黙々と作業を進めている状態、指導者の助言を求める挙手の状態の3つがある。しかし、挙手には何か質問事項があって行う場合と、作業を終えたことを示す場合で意味が変わってくる。そのため、「指導」「作業中」「質問」「暇」の4状態から構築する。

最後に「生徒同士は何をしているか」である。これには生徒同士の会話と、物の貸し借りなどの物理的なアクションが考えられる。しかしながら、遠隔からでは物理的なアクションには対応できないため、「会話中」と「沈黙」の2つの状態を構築する。

構築したモデルの概念図を図2に示す。授業は指導者の主導で進むため、通信すべき情報とその重要度は指導形態と指導者の状態から決定する。

モデルから推測される各状況における情報の重要度を表1に示す。表中の括弧は指導対象外の生徒を示す。また、音声については遠隔指導において常に一定の品質で通信することが求められるため、表中からは省略する。

「画像」は作業中の机上を写した画像を意味したもので、本モデルにおいて最も重要なセンサ情報である。また、「反応」は生徒から指導者への挙手等のリアクションを示した文字列の信号である。「会話」は生徒間のものであり、指導者との音声と区別するためテキストチャットを用いる。

表 2: 画像の品質と帯域使用量

	サイズ	fps	品質 [%]	帯域 [Kbps]
講演	大	2	100	413.93
演習巡回	中	1	75	985.6
演習指導	大(小)	2(0.5)	100	1125.21
討論	小	0.5	100	328.65

3.2 通信制御の予備実験

提案するモデルを用いた品質制御によって、どの程度のネットワーク帯域を使用するのかどうかを調査する実験を行った。実験では送信端末で各パラメータのJPEG画像を生成・送信し、受信端末で受信時の帯域使用量を計測した。結果から生徒数を20人とした場合の理論値を表2にまとめる。

画像のパラメータはWebで使用される画像サイズからアスペクト比の同じ代表的なものを抽出し、「大」512×288px、「中」384×216px、「小」128×72pxである。fpsは1秒間に送るフレームの数を示すものであり、割合はJPEGの品質を意味する。また、帯域は各パラメータで1秒間に使用する帯域の使用量である。

この結果から、ネットワークの揺らぎを考慮にいれる必要はあるが、音声やHTTP等のその他の通信、画像の中身による増加分を考慮しても、2Mbpsの回線があれば、指導状況モデルを用いた通信制御により、生徒20人との通信が可能であると考えられる。また、世界のネットワーク回線の平均速度は約3.3Mbpsであることから[2]、この規模であれば、多くの国の国内でシステムを用いることが可能であるといえる。

4 まとめ

遠隔指導システムの効率的な通信制御を実現するための指導状況モデルを提案した。実際に3つのパラメータを用いて画像品質制御を行った結果、2Mbpsの回線で20人の生徒との通信が理論的に可能であることを示し、遠隔指導に必要な情報の品質を適切に制御可能なことを明らかにした。今後、実証評価実験を進め、指導状況モデルの改良を進める。

参考文献

- [1] 櫻田武嗣, 萩原洋一: 全国18国立大学法人を結ぶHD対応遠隔講義システムの設計, 電子情報通信学会技術研究報告.SITE, 技術と社会・倫理, Vol.108, No.459, pp.91-95(2009).
- [2] Akamai: The State of the Internet 2ND QUARTER 2013 REPORT, Akamai(online), available from (http://www.akamai.com/dl/documents/akamai_soti_q213.pdf?WT.mc_id=soti_Q213) (accessed 2014-01-09).