

ビデオ講義における講師の顔の動きによる学習者への効果

村上 拓真[†]

兵庫県立大学 社会情報科学部[†]

川嶋 宏彰[‡]

兵庫県立大学 情報科学研究科[‡]

1 はじめに

ビデオ講義では、映像中の講師の振る舞いが学習効果に影響を与える可能性がある。たとえば [1] では、顔の向き（正面やスライド方向など）と体の向き（正面、横）の組み合わせでの学生の注意の配分と学習成績への影響を調査しており、体の向きに関係なく、講師がスライドを見ている方が学習者はよりスライドに注目し、よりよい成績になることを報告している。また [2] では、講師が学習者の視線を誘導をする場合としない場合（講師が常に前を向いている状態）が比較されており、視線を誘導することで有意に学習成績が向上するという知見が得られている。

しかし、これら従来研究の実験では、講師の体や顔向きは同じ方向を向き続けるなど、単純化された条件が用いられている。この理由として、実際の講師を撮影する際に様々な動きを許容すると、比較すべき映像において、顔の動かし方やその速さなどの条件を統制するのが難しくなることが挙げられる。一方で、講義内容に応じた講師の顔の動きのタイミングなどが、学習効果に影響を与える可能性がある。

そこで本研究では、映像生成技術を用いることで、講師の顔の動かし方に関して、および動き以外の声などの条件に関して統制のとれた講義動画を作成するとともに、ビデオ中の講師の顔向きやその動作タイミングが、学習者の注意の配分や学習成績に与える影響を調べる枠組みを提案する。

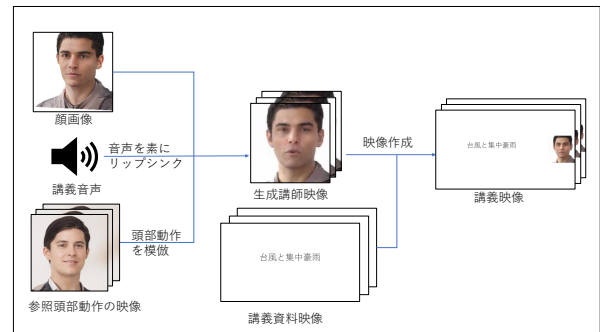


図1 講義映像生成の流れ

2 講義映像の生成

本研究では図1に示す流れで講義映像を生成する。はじめに、講義内容のスライドを作成し、音声合成^{*1}を用いて、講師の発話音声を生成する。これは、講師の声のトーンなど、顔向き以外の条件をそろえるためである。つぎに、どの実験参加者も知らないような講師の顔として、AI人物素材^{*2}の人物顔を用いる。この講師の顔を、Pose-Controllable Audio-Visual System[3][4]を用いて動かすことで、音声に合わせてリップシンクを行うとともに、指定した顔向きの動作を指定したタイミングで生成する。これにより、比較条件以外については、顔の動かし方やその速さ、声のトーンなどの統制が取れた講義映像を生成することができる。

顔の動きの指定には、正面およびスライド方向を交互に見るような参照用の動画が必要となる。以下で述べる実験では、参照用人物^{*3}の前向きから横向きの頭部動作を撮影し、この逆再生、および前向きと横向きの静止画を組み合わせることで参照用動画を作成した。

Effects of head movements in generated lecturer images on learners in video lectures

[†] Takuma Murakami, School of Social Information Science, University of Hyogo

[‡] Hiroaki Kawashima, Graduate School of Information Science, University of Hyogo

^{*1} VOICEVOX (<https://github.com/VOICEVOX>)

^{*2} <https://www.photo-ac.com/main/genface>

^{*3} 参照用動画ではどの人物の動作でもよい。

3 評価実験

講師の顔の動きが注意や学習効果に与える影響を確認するために、講義視聴中の視聴計測と受講後のテストを用いて実験を行った。

3.1 実験条件

講義映像: 台風と集中豪雨(約7分30秒), オーストラリアの歴史(約8分30秒), 都市工学(約9分30秒)の3科目の内容の講義映像を2節の方法で作成した。講師の顔の向きや動きとして, (C1) 正面のみ固定, (C2) スライド方向のみ固定, (C3) 正面とスライド方向を一定間隔で交互に変化, (C4) 重要語の説明時にスライド方向へ変化, の4条件を設けた。これを各科目について用意することで, 計12個の講義映像を用意した。重要語は受講後のテスト問題の答えに対応する単語とした。

実験参加者: 22人の実験参加者(大学生と大学院生)に各科目の映像を視聴してもらい, 視聴中の視線データと視聴後のテストの成績を取得した。実験参加者は各科目でいずれか一つの条件を視聴するものとし, 一人が同じ条件の映像を視聴しないよう割り付けた。

受講後のテスト: テストは10問前後用意し, 基本的に四択問題(一部は三択もしくは六択)とした。各問題について, 事前知っている内容か否かを確認する質問も設け, 5名以上が既知であると回答した問題はその正誤に関わらず点数に反映しないようにした。

視線の解析: 講師の顔の動きに対する学習者の視線の追従を調べるために, 画面を「講師」, 「スライド」, 「それ以外」の3つの範囲に分割し, 注視割合と視線移動回数を算出した。図2に受講中の視線推移の例を示す。注視割合では, 各範囲を注視した時間の割合を計算するとともに, 「それ以外」を除いた場合の割合も求めた。視線移動回数では, ある範囲から別の範囲へ移動した回数を算出した。また, 計測誤差を考慮して重要語の表示範囲を定めたうえで重要語の確認回数を算出した。

3.2 実験結果

学習成績: 実験から得られた学習者の正答数を用いてウェルチの t 検定を行った。これによ

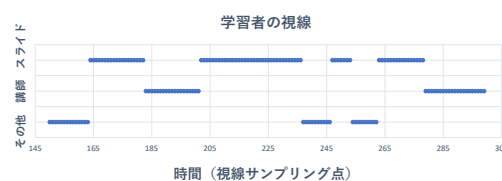


図2 学習者の視線の動きの例

て講師の顔の動き方によって成績に有意差が現れるのかを調べた。各科目において, 講師の条件による組み合わせ6種類で有意水準5%の両側検定を行った場合, 都市工学の(C1)と(C2)を比較した結果のみに有意差が見られた(ただしここでは多重性は考慮していない)。

学習時の視線: 注視割合は, (C4)の条件で, 各科目において講師の顔に注目している割合の平均が他の条件よりもやや多かった。ただし有意な差とはいえなかった。視線移動回数や重要語の確認回数についても, 条件による有意差は見られなかったものの, 視線移動回数の1分平均については, 本研究の(C1)と条件が同じである[2]と近い結果となった。

本研究では, 実験条件を統制するために, 生成講義映像を用いる実験の枠組みを提案するとともに, 実際の実験までを行った。一方で, 実験参加者の人数がまだ十分ではなく, 外れ値も多い。より多くの実験参加者により, 条件間の差の有無を検証することを今後の課題とする。

謝辞 本研究の一部は科研費 JP19H04226 の補助を受けて行った。

参考文献

- [1] Zhongling Pi, Ke Xu, Caixia Liu, and Jiumin Yang. Instructor presence in video lectures: Eye gaze matters, but not body orientation. *Computers Education*, Vol. 144, p. 103713, 2020.
- [2] Hongyan Wang, Zhongling Pi, and Weiping Hu. The instructor's gaze guidance in video lectures improves learning. *Journal of Computer Assisted Learning*, Vol. 35, No. 1, pp. 42–50, 2019.
- [3] Yasheng Sun, Hang Zhou, Ziwei Liu, and Hideki Koike. Speech2talking-face: Inferring and driving a face with synchronized audio-visual representation. In *IJCAI*, Vol. 2, p. 4, 2021.
- [4] Hang Zhou, Yasheng Sun, Wayne Wu, Chen Change Loy, Xiaogang Wang, and Ziwei Liu. Pose-controllable talking face generation by implicitly modularized audio-visual representation. In *CVPR*, 2021.