

# AR 技術に基づく反転講義のための講義映像作成支援システムの試作

大園 忠親 落合 裕也 片岡 瞳 新谷 虎松

授業と宿題の役割が反転した新たな授業形態である反転講義において、予習（宿題）用の教材である講義映像の編集における負担を軽減するための支援が必要である。反転講義における講義映像の作成は高負担であり、講義映像作成支援のための技術が求められている。講義の撮影とその記録映像の編集作業は煩雑であり多くの手間が必要である。例えば、撮影した映像に対して、生徒が講義内容を理解しやすくするための編集（例えば映像の不要部分の削除）は重要であるが、負担が大きい。本研究では、講義映像作成に関して聴覚的かつ視覚的な側面における支援に必要な技術の開発を行っている。本稿では、講義映像作成における視覚的な側面における支援に関連して、AR 技術を利用した即応的な映像編集技術について述べる。本システムにより、講義映像作成における負担軽減のみならず、より効果的な視覚的演出の作成が容易になることが期待される。

## 1 はじめに

授業と宿題の役割が反転した新たな授業形態である反転講義において、予習（宿題）用の教材である講義映像の編集における負担を軽減するための支援が必要である。反転講義における講義映像の作成は高負担であり、講義映像作成支援のための技術が求められている。講義の撮影とその記録映像の編集作業は煩雑であり多くの手間が必要である。例えば、撮影した映像に対して、生徒が講義内容を理解しやすくするための編集（例えば映像の不要部分の削除）は重要であるが、負担が大きい。

講義映像の作成において、聴覚的な観点からの支援および視覚的な観点からの支援の両方が必要である。講義映像作成支援に関連して、本研究の先行研究として、可聴性を阻害する話者の癖を調整し、音声明瞭性

の高い理想的な状態に講義映像を変換するための技術を開発してきた [1]。音声のみならず、講義映像を観やすくするための編集も重要である。例えば、資料の拡大表示などの強調効果の付与、講師の表示方法の変更、テロップの挿入などが挙げられる。

視覚的な観点からの講義映像作成支援のために必要な要素技術を検討すると、映像中の講師や資料の位置が検出可能であることが好ましい。講師や資料に対して、映像効果を適用するためである。ここで、AR (Augmented Reality) 技術を利用することで、映像中の立体的な座標を取得可能になる点に着目した。近年の AR 技術は、Apple 社の iOS 端末など、汎用的なモバイル端末で利用可能であり、特別な装置を必要としない点が優れている。本稿では、本システムにおける視覚的な観点からの講義映像作成支援について検討する。特に、視覚的な観点からの講義映像作成支援を可能にするための AR 技術の利用方法を検討する。

以降、本稿では 2 章で、開発済みの反転講義支援システムについて概説し、AR 技術に基づく講義映像作成支援の必要性について議論する。3 章では AR 技術に基づく講義映像作成支援システムの機能について述べ、最後に 4 章で本稿をまとめる。

---

Developing an AR-based Lecture Video Editing Support System for Flipped Teaching.

This is an unrefereed paper. Copyrights belong to the Author(s).

Tadachika Ozono, Hitomi Kataoka and Toramatsu Shintani, 名古屋工業大学大学院情報工学専攻, Dept. of Computer Science, Graduate School of Engineering, Nagoya Institute of Technology.

## 2 反転講義のための講義映像作成支援

本稿において、講義映像とは、講義を録画した動画とする。ここでの講義映像には、講師によるスライド形式の講義資料を用いた説明が録画されている。講義映像には映像および音声が含まれている。講義映像中の講義資料の映像をスライド映像と呼び、講義映像中の講師の映像を講師映像と呼ぶ。また、講義映像中の講師の音声を解説音声と呼ぶ。

ここでは、教室等における対面式講義を録画して反転講義用の講義映像として利用する場合、および反転講義用に新たに講義を録画する場合を想定する。

### 2.1 反転講義支援

反転講義を取り入れる教育機関やオンラインによる講義を提供するサービスの増加により講義映像の需要が高まっている。反転講義は、授業と宿題の役割を反転させ、生徒の学習意欲の向上や知識の定着を促すことを目的とし2000年頃から提案されている。反転講義では、生徒が宿題として講義前に自習を行い、教室ではグループ学習やディスカッションなど、発展した授業を行う。

生徒にとって、講義内容を理解しやすく品質の高い講義映像が求められる。しかし、講師は、講義の様子を撮影し、撮影した映像の不要な部分を削除したり、生徒が講義内容を理解しやすくするために編集を行う必要があり、講義映像の作成コストは高い。そこで、講師の講義映像の品質向上、作成および編集の支援を行う必要がある。また、提供される講義映像の形式は、スライド形式、作業画面形式および擬似講義形式が考えられる。スライド形式は、講師がスライドなどの資料を用いながら講義映像の解説を行う形式である。作業画面形式は、黒板やパソコンの作業画面を映し、講師が作業を行いながら講義映像の解説を行う形式である。擬似講義形式は、実際に講師が映り黒板を使って擬似的な講義を行う形式である。いずれの講義形式でも共通である解説音声は生徒が講義内容を理解する上で重要である。特に、スライド形式および作業画面形式では、解説音声の質が講義内容の理解に大きく影響すると考えられる。そこで、本研究では、

解説音声の音声品質を対象とし、講義映像の品質向上を目指している。

また、オンラインで講義を提供しているJMOOC<sup>†1</sup>やKhan Academy<sup>†2</sup>などのサービスも登場している。講義映像がオンライン配信されることにより、多くの生徒が受講することが可能となった。講義映像を視聴する生徒が増えるとともに生徒個人の講義映像に対する要望や好みが多様化してくる。例えば、講義映像において、講義内容を理解するためにゆっくり聴講したい生徒や講義の概要だけ素早く聴講したい生徒がいる。講師は、生徒に合わせた講義映像を作成することは望ましいが、一人一人の生徒に合わせた講義映像を準備することは困難である。そこで、本研究では、生徒に対して、講義映像の視聴時に、個人の要望や好みで講義映像を再編集するインターフェースを提供する。講義映像を再編集するインターフェースを提供することで、生徒自身の好みに合わせた講義映像を視聴することができ、講義映像の内容の理解の向上に期待できる。他にも、講義映像の視聴に対するフィードバックは、講師が講義映像を作成する上で重要である。しかし、講義映像を視聴する際に、生徒がフィードバックをすることは、手間がかかる。そこで、本研究では、生徒の講義映像の視聴および編集の履歴を記録し、講師にフィードバックする機能を開発している。

### 2.2 講義映像作成支援システム

本研究の先行研究として開発した講義映像作成支援システムについて説明する。本システムでは、スライド形式の講義資料に対して解説音声を付与した講義映像の作成を支援することを目的とした。ここでは、解説音声を半自動的に改善するためのシステムを開発した。

図1は、開発済みの講義映像作成支援システムにおける講師が解説音声を編集するための画面である。本システムでは、主に解説音声の改善を扱っている。図1(A)には、スライド映像が表示されている。図1(B)は、解説音声の音声品質の編集値を設定する部分で

†1 <https://www.jmooc.jp>

†2 <https://www.khanacademy.org>



図 1 講師の講義映像の編集

ある。図 1(C) では、講義映像の再生制御をするための部分である。図 1(D) は、癖および視聴操作履歴を表示する部分である。ここでは、スライド映像のサムネイル画像が表示されている。特に、スライド映像とその解説音声の波形の対応づけが可視化されている。ここでは、講義映像の追加、削除、再生速度の変更および重要度を設定することができる。さらに、講義映像中の解説音声の音声品質の評価および視聴操作として、再生、停止、早送りおよび巻き戻しが発生した箇所が表示されている。

前述したように、本システムは、講義資料に対する解説音声の付与において、解説音声の改善を目的としているため、講義映像中には講師映像が含まれない。しかしながら、スライド映像のみでは、映像が単調になりがちであり、生徒の興味の維持が困難であると考えた。図 1 で示される講義映像作成支援システムは、講義映像における視覚的な演出効果のための編集支援が不十分であった。特に、講師映像の編集支援が不十分であった。

スライド映像のみの講義資料を改善するために、視覚的な演出が必要である。本研究では、構成論的に必要な機能を見いだすための探索的な研究として、様々な機能を試作することにした。例えば、スライド映像と講師映像の関係に関しては、スライド映像のみを表示すれば十分な場合や、講師によるポインティングのようにスライド映像と講師映像の両方が表示されなければならない場合がある。状況に応じて、スライド

映像と講師映像のを適切に組み合わせて表示することが好ましい。以降、本稿では、講義映像における視覚的な側面における編集支援を AR 技術を用いて実現する方法について述べる。

### 3 AR 技術に基づく講義映像作成支援

本システムは、アドホックな講義映像作成の支援を目的とし、撮影中における講師の動作や発話をトリガーとして、講義映像に映像効果を追加するシステムである。これにより、撮影前の綿密なシナリオの作成や撮影後の映像編集作業に要する負担の軽減を目的としている。本章では、講義映像作成支援における視覚的な側面への対応に関して、必要な機能に議論する。

#### 3.1 機能

講義映像の視覚的な観点からの講義映像作成支援について検討する。ここでは、視覚的な観点からの講義映像作成支援として、(i) 講師映像と資料映像の合成を支援、(ii) 対話形式の講義映像の作成を支援、(iii) 講義映像の強調を支援の 3 点を検討することとした。

(i) の講師映像と資料映像の合成を支援とは、講師映像とスライド映像の表示方法を動的に制御することである。例えば、最初にスライド映像のみを表示し、講師によるポインティングが必要な場合は、主にスライド映像を表示しつつ、講師映像をスライド映像の観やすさを考慮しながらスライド映像上に重畳表示する。これにより、スライド映像のみの単調な講義映像よりも、変化に富む講義映像を作成することが可能になる。ここでは、システムを簡便に利用可能にするために、1 台のカメラ付きモバイル端末による撮影を可能にすることを目標とする。

(ii) の対話形式の講義映像の作成を支援とは、講師と生徒との対話形式の講義映像を作成することを支援することである。対話形式の講義映像に関しては、特に初等教育や導入教育における、馴染みやすさやわかりやすさの向上を期待している。ここでは、講師が、講師と生徒の役割を一人二役で演じることとする。これは、生徒役との打合せに必要な準備の負担を減らすことを狙っている。講師は、人形などの仮想の生徒との対話を行う。例えるならば、腹話術により対

話形式の講義をすることを支援する。本稿では、生徒役の対話エージェントのことを、パペットと呼ぶ。

(iii) の講義映像の強調を支援とは、講師映像およびスライド映像を強調するための映像効果を半自動的に挿入することである。例えば、講師の自己紹介時に講師の氏名を講師の頭上に表示することや、重要点の説明時にその概要をテロップとして講義映像に挿入することなどが考えられる。ここでは、撮影後の映像編集作業を減らすために、撮影中の講師の身体動作等をトリガーとして、自動的に強調効果を講義映像に挿入するための機能を実現する。

(i)~(iii) を実現するために、1台のカメラ付きモバイル端末による撮影から、講師の特定の身体動作/発話等をトリガーとして、視覚的な講義映像編集を行うことが有効であると考えた。よって、本研究では、講師の特定の身体動作、発話およびスライド操作をトリガーとして動作するイベント駆動型の講義映像編集機能を実現することとした。すなわち、特定の身体動作等をトリガーとして、講義映像の撮影中に映像上の演出を挿入する。ここでは、AR機能等を利用して、映像中の講師やスライドの認識および音声認識を行う。

AR技術に基づく講義映像作成支援機能として、具体的に(i)~(iii)に対応する機能として、それぞれ、講師映像編集機能、パペット機能、およびテロップ表示機能の3つの機能を実現する。講師映像編集機能は、講義資料と講師映像の表示方法を、自動的に切り替える機能である。パペット機能は、講義映像上に表示されるエージェントとの対話形式の講義を実現するための機能である。テロップ表示機能は、講義映像にテロップを挿入する機能である。以降、これらの機能を制御するためのトリガーについて説明し、その後、各機能について説明する。

### 3.2 トリガー

本システムでは、講師映像編集機能、パペット機能、そしてテロップ表示機能を動作させるためのトリガーとして、講師の特定の身体動作、発話およびスライド操作を用いる。

トリガーとして利用する身体動作は、ポーズとモーションに分類される。ポーズとは、特定の姿勢で静止



図2 (a) 無編集モード

することである。例えば、直立不動、手を挙げる、首をかしげるみなどである。モーションは、身体を用いた特定の動きをすることである。例えば、礼、手を振る、頷くなどである。事前に定義されたポーズやモーションだけでなく、好みのポーズやモーションを定義することも可能にする。ポーズやモーションの認識には、姿勢認識技術により骨格抽出を行い、掌や肘などの位置を表す情報である骨格点を利用する。

トリガーとして利用する発話は、主に事前に定められた特定のキーワードを用いる。例えば、「ここが重要です」と発話することで、資料を強調表現することなどが考えられる。また、Apple社のSiri等の既存の対話システムを利用することで、より複雑なトリガーを作成することも可能になる。

トリガーとして利用するスライド操作は、スライドの表示/非表示の切替え、およびスライドのページ移動を用いる。本システムは、スライドの表示状態や、表示中のスライドのページ番号を管理している。

ポーズ、モーションおよび発話を複合的に利用したトリガーも考えられる。例えば、資料を指さしながら、「ここが重要です」と発話した場合には、講師が指差ししている手の付近を強調したりスライド映像を拡大すること等が考えられる。

### 3.3 講師映像編集機能

講師映像編集機能は、スライド映像と講師映像の配置および表示方法を編集するための機能である。例えば、スライド映像と講師映像が含まれる講義映像にお

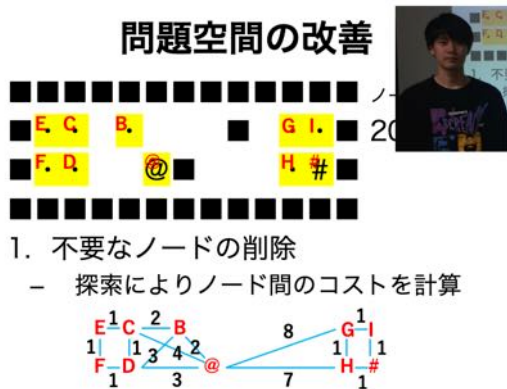


図3 (b) スライド強調モード

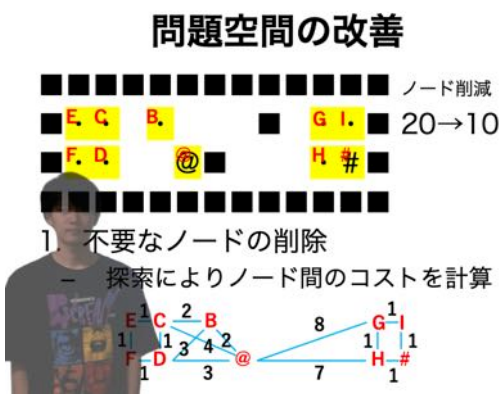


図4 (c) 講師シルエット映像モード

いて、スライド映像のみを表示し、講師映像を除去することが考えられる。講師映像編集機能では、(a) 無編集モード、(b) スライド強調モード、(c) 講師シルエット映像モードの3種類の映像効果がある。

(a) の無編集モードでは、本システムを用いて録画した、講義映像をそのまま表示するモードである。図2は、講師映像(画面左)とスライド映像(スライド「問題空間の改善」)をそのまま表示している様子を表しており、この映像を用いて (b) および (c) の映像を生成する。そのために、図2のような講義映像に対して、講師映像からの骨格点の検出および資料映像の検出を実行する。骨格点の検出には PoseNet [2] を用いている。

(b) のスライド強調モードは、図3のように、スラ

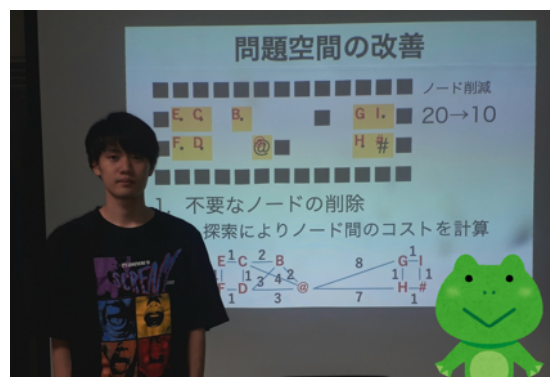


図5 パペット機能

イドを全面に表示するモードである。スライド資料の原本から得られた画像を用いることで、鮮明なスライド映像を表示することができる。図3の右上には、講師映像が表示されている。講師映像の表示/非表示の切り替えや、表示位置を制御することが可能である。姿勢認識を用いることで、1台のカメラで、このような講師映像の表示が可能になる。

(c) の講師シルエット映像モードでは、講師映像を半透明に表示することで、スライド資料の見やすさ(スライド資料映像領域の最大化)と、講師のジェスチャ(指さしなど)の表示を両立するモードである[3]。文献[4]では板書遮蔽問題(板書中の黒板が見えない問題)の解決方法として、講師映像を半透明化する手法の有効性が示されている。講師映像の透明度や表示する部位(全体、上半身のみ、手のみなど)をトリガーによって切り替えることができる。

### 3.4 パペット機能

パペット機能は、講義映像中におけるエージェント(パペット)を仮定の生徒として、パペットとの対話形式の講義を演出するための機能である。パペットとの対話形式の講義が、生徒の理解を促進することが期待される場合に利用する機能である。

図5は、パペット機能の実行例である。図5では、(a) 無編集モードに基づき、右下にパペット(ここではカエル)を表示している様子を表している。パペットの表示は、画面上に平面的に表示したり、AR空間内に立体的に表示することもできる。図5では、画面



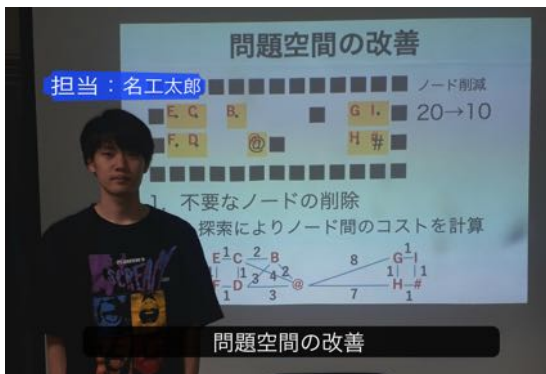


図 6 テロップ表示機能

上に重畳表示している。パペットの制御は、トリガーに基づき事前に設定されたシナリオ通りに再生するモードと、講師のジェスチャでパペットの挙動を制御するモードが考えられる。ジェスチャでパペットの挙動を制御する場合は、腹話術人形のようなものを検討している。また、ルールベースの反射的な演出も必要である。例えば、講師がパペットの方に顔を向けたときに、パペットも講師の方を向くような演出である。

AR 技術を利用したパペットの制御では、その演出効果の作成が手間である。演出効果作成の手間を軽減するために、AR 技術を用いた教示による演出効果作成手法を利用する [5]。

### 3.5 テロップ表示機能

テロップ機能は、講義映像に映像オブジェクト（総称してテロップと呼ぶ）を追加する機能である。テロップは、文字列および図形により表現される。テロップの表示方法は、パペットと同様に、講義映像に平面的に重畳表示するモードと、AR 空間内に立体的に表示するモードがある。テロップには、画面内もしくは AR 空間内の固定された位置に表示される静的テロップと、講師の身体動作等に合わせて移動する動的テロップがある。例えば、動的テロップの一つに、ポインター機能がある。ポインター機能は、掌の位置に円、直線、四角形などの図形を描画する機能である。掌の移動に伴って、図形も移動する。

図 6 は、テロップ機能の一例である。講師の頭上

に表示されている文字列「名工太郎」は、講師の氏名を AR 空間内に表示しており、講師の移動に追従する。画面の下に表示されている文字列「問題空間の改善」は、画面上に平面的に重畳表示されたスライドのタイトルを表示する演出である。例えば、スライドのページ移動をトリガーとして、スライド移動後にテロップを表示し、表示後一定時間後に消えるような演出が考えられる。

## 4 おわりに

講義映像作成における視覚的な側面における支援に関連して、AR 技術を利用した即応的な映像編集技術について述べた。具体的には、講師の身体動作や発話をトリガーとして利用可能な、講師映像編集機能、パペット機能そしてテロップ表示機能について説明した。授業と宿題の役割が反転した新たな授業形態である反転講義において、予習（宿題）用の教材である講義映像の編集における負担を軽減するための支援が必要であり、本システムにより講義映像の作成を支援することで反転講義が促進されることが期待される。

謝辞 本研究の一部は JSPS 科研費 JP16K00420, 19K12097, 19K12266. の助成を受けたものです。

## 参考文献

- [1] 松浦辰雄, 大園忠親, 新谷虎松: 反転授業における音声明瞭性を考慮した講義映像作成支援システムの開発, 第 16 回情報科学技術フォーラム, Vol. 16, No. 4, 2017, pp. 81-84 (FIT 論文賞受賞) .
- [2] Papandreou, G., Zhu, T., Chen, L. C., Gidaris, S., Tompson, J., Murphy, K., Personlab: Person pose estimation and instance segmentation with a bottom-up, part-based, geometric embedding model, Proceedings of the European Conference on Computer Vision (ECCV), 17p, 2018.
- [3] 丹羽 佑輔 片山 真也 白松 俊 大園 忠親 新谷 虎松: 発表者の実像を利用した動的なプレゼンテーション操作環境の実現日本ソフトウェア科学会第 30 回大会 (2013 年度) 講演論文集, 6p, 2013.
- [4] 奥本 隼, 山根恵和, 吉田光男, 岡部正幸, 梅村恭司: 講師のシルエットを透過表示した板書映像の生成とライブビューシステム, 日本教育工学会論文誌, Vol. 41, No. 2, pp. 177-186, 2017.
- [5] Hitomi Kataoka, Tadachika Ozono and Toramatsu Shintani: Realizing an Effect Editor for AR Pop-up Picture Books by Teaching Motions, IDDC2019, IEEE, pp. 989-994, 2019.