

ライブプレゼンテーションに基づく 反転講義用動画作成のための AR パペットの開発

片岡 瞳^{†1,a)} 落合 裕也^{†1,b)} 大園 忠親^{†1,c)} 新谷 虎松^{†1,d)}

概要：本研究では、授業と宿題の役割を反転させた授業形態である反転講義のための動画作成支援システムを開発している。ここでは、反転講義用動画への AR パペットの導入を試みる。AR パペットは、生徒の理解促進や興味維持を目的とした、拡張現実空間内のエージェントである。動画撮影後の編集作業の負担を軽減するため、講師が撮影中に AR パペットを操作可能なシステムを検討する。また、講師と AR パペットによる対話形式の講義のための演出効果を実現する。本稿では、ライブプレゼンテーションに基づき動画作成中に操作可能な AR パペットの開発について述べる。

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。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(例えば映像の不要な部分の削除)は重要であるが、負担が大きい。

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などの視覚的な支援も重要である。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。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ため、視覚的な演出が必要であると考えた。そこで本研究では、反転講義の動画において、生徒の理解を助けるための機能として AR パペットを開発した。AR(Augmented Reality; 拡張現実感)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、特別な機器を必要としない点が優れている。AR パペットは、講師によって操作され、拡張現実空間内のオブジェクトとして存在するエージェン

トである。AR パペットを仮想の生徒として、講師と AR パペットとの対話形式の講義を演出することが可能であると考えた。対話形式にすることで、生徒の思考を促すことが期待できる。また、生徒にとって実際の講義時間における議論のための予行演習にもなる。また、AR パペットを講師役とする講義映像を作成することも可能である。これは、講師の映像を生徒が好む姿に変更可能という利点がある。

本稿では、ライブプレゼンテーションに基づく反転講義用動画作成において、講師による操作が可能な AR パペットの開発について述べる。本稿において、ライブプレゼンテーションに基づく反転講義用動画作成とは、講義映像撮影時に、講師がプレゼンテーションコンテンツの一部である AR パペットを、講義を継続しながら制御可能であることを指す。このような AR パペットの実現により、講師が講義の進行に応じて柔軟に AR パペットとの対話を取り入れることも可能である。ここでの AR パペットの役割とは、仮想の講師としてメインで講義を行うこと、人間の講師を補助しながら共に講義を行うこと、または仮想の生徒として人間の講師に疑問を投げかけることなどが想定される。また、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、同様に、講義映像中の講師の映像を講師映像、講義映像に AR パペットを追加した映像を AR 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。

以降の構成を以下に示す。2 章では、関連研究について

^{†1} 現在、名古屋工業大学

a) eye@toralab.org

b) ochiyu@toralab.org

c) ozono@toralab.org

d) tora@toralab.org

述べる。3章では、AR パペットの仕様について、見た目および操作方法の観点から議論する。4章では、AR パペットを用いた反転講義用動画作成システムについて、システム構成と機能を図や実行例を用いて説明する。5章では、開発した AR パペットの設計について議論する。6章では、考察について述べる。最後に、7章で本稿についてまとめる。

2. 関連研究

2.1 反転講義のための動画作成支援

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、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。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、講義映像中には講師映像が含まれない。本研究は、実際に講師が映り、AR パペットを用いてスライド形式の擬似講義を行う動画作成を支援する。

2.2 ライブプレゼンテーションにおける演出効果

プレゼンテーションやテレビ番組、広告といった様々な目的で、動画中にグラフィックスを組み合わせ、表現豊かなパフォーマンスを作成する演出が用いられる。グラフィックスを後処理として動画に追加する場合もあるが、天気予報やカメラアプリのビデオフィルターなど、リアルタイムでグラフィックスを操作する場合がある。このようなライブプレゼンテーションにおける演出効果を用いた動画作成支援の研究が行われている。

Saquib ら [2] は、プレゼンテーションの動画において、発表者のジェスチャや姿勢をトリガーとして、リアルタイムにインタラクティブな演出効果を追加するシステムを開発した。これにより後からプログラミングや煩雑な動画編集によるプレゼン映像とグラフィックスの合成を行う必要はなく、ライブパフォーマンス中に直感的に演出効果を制御可能である。Matulic ら [3] は、ライブプレゼンテーションにおいて、発表者のジェスチャーによって動作するアバターをスライドに表示するシステムを開発した。発表者がプレゼンテーションコンテンツの一部となり、データや図などの要素とインタラクションを行う試みである。ここで

は、特に発表者とスライド上のコンテンツ両方に集中する必要がないという点に着目している。また、Trinh ら [4] は、仮想エージェントとの共同プレゼンテーションを実現するためのシステムを開発した。仮想的な共同発表者の存在が、発表者の不安を軽減し、自信を向上させたことを実証している。仮想エージェントが聴講者の注意を引くことで、発表者の緊張を軽減し、さらに発表の単調さを解消したといった結果が得られている。

本研究で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操作によって、リアルタイムに講義映像に AR パペットを用いた演出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する。

3. AR パペットの仕様

本章では、AR パペットの仕様を見たとおり操作方法の2点から検討する。まずは、制約条件について検討し、各制約条件を満たすための機能について議論する。

AR パペットが満たすべき制約条件は、(1) 必要な演出効果を実現可能であること、(2) 撮影中に講義を継続しながら講師自身で操作可能であること、(3) 簡便に利用可能であることの3点である。制約条件 (1) の必要な演出効果を実現可能であることとは、ここでは対話形式の講義に必要な演出効果が可能であることを意味する。多様な演出効果が可能であれば多様な講義映像を作成可能になるが、AR パペットの操作が煩雑になることが予想される。制約条件 (2) の撮影中に講義を継続しながら講師自身で操作可能であることとは、講義中の講師にとって認知的負荷が軽いということを意味する。対話形式の講義映像を作成する上で、撮影後の編集作業を可能な限り省略するために、講義中に AR パペットを講師自身で操作可能であることが有効であると考えられる。制約条件 (1) と (2) は、トレードオフの関係にある。演出効果の種類を増やしても、実際の講義映像撮影中には運用できないと予想される。制約条件 (3) の簡便に利用可能であることとは、汎用機器のみで実現可能であること、および AR パペットの操作者を別途必要としないことを意味する。制約条件 (2) を緩和するためには、講師以外が AR パペットを操作すれば良いが、講師と AR パペット操作者との事前の準備が負担になると考えられる。

制約条件 (1) を満たすために、ここでは AR パペットの演出効果として、(A) 挨拶、(B) 質問、そして (C) 了解の3種類を実現する。これは、講師と AR パペット間の挨拶に始まり、講師と AR パペット間の質疑応答により講義を進めることとしたからである。演出効果 (A)~(C) を実現するために、AR パペットは、おおよそ人間のような肢体を持ち、頭の位置と、腕の向きを制御可能にする。さらに、講義内容に AR パペットの見た目を合わせるため、ユーザ

がAR パペットを作成可能にすることも必要である。

制約条件 (2) を満たすために、半自律制御が可能なAR パペットを実現することとした。AR パペットの操作として、自律制御、半自律制御、および手動制御が考えられる。自律制御は、講師の認知的負荷の軽減に有効であると考えられるが、技術的に困難である。手動制御は、操り人形のように講師が全てを操作することを意味するが、認知的負荷が高い。ここでの半自律制御では、講師の簡単な操作により、演出を再生することを意味する。具体的には、(B) 質問は、ポインティング機能として実現する。すなわち、画面上のタップにより、AR パペットがその箇所をポインティングしつつ発話する。(A) 挨拶および (C) 了解は、ボタンのタッチによりAR パペットが動作付きで発話する。先にも述べたように、AR パペットは、AR 映像における演出効果と操作の容易さを両立する必要がある。AR パペットを詳細に操作することができれば多様な演出が可能になるが、操作が複雑になる。演出効果を3種類に限定した上で、半自律制御を可能とすることで、制約条件 (1) と (2) を両立すると期待される。

制約条件 (3) を満たすために、制約条件 (2) で示した半自律制御を実現する。また、汎用機器として Apple 社の iPhone もしくは iPad 上で実現することとした。iPhone もしくは iPad は、AR の機能を標準的に利用可能であり、また講義映像作成のための撮影機器としても優れている。

制約条件 (1)～(3) を満たすために、人のような四肢による3つの演出効果を、講義中に容易に操作可能な、拡張現実空間内のエージェントを実現する。また、本機能を iPhone もしくは iPad 上で実現することで、汎用機器による導入を可能にする。

4. AR パペットを利用した反転講義のための動画作成システム

4.1 システム構成

図1は、本システムを利用した動画撮影環境におけるハードウェア設置の様子である。動画撮影環境には、講義用PC、講義スライド、操作用デバイス、および撮影用デバイスが含まれる。講師は講義用PCによって講義スライドを操作する。講義映像は撮影用デバイスのカメラを用いて撮影する。このとき撮影用デバイスは、講義映像にAR パペットを追加したAR映像を画面に表示する。本システムにおけるAR機能はApple社のARKitを用いて実装したため、撮影用デバイスはこれを利用可能なiOS端末である。また、講師は操作用デバイスのタッチ操作によってAR パペットを操作する。操作用デバイスには、撮影用デバイスの画面に表示されているものと同様のAR映像が転送されている。講師は、AR パペットを操作しながら作成されたAR映像を見ることでフィードバックを受け取る。本稿における実行環境は操作用デバイス、撮影用デバイス

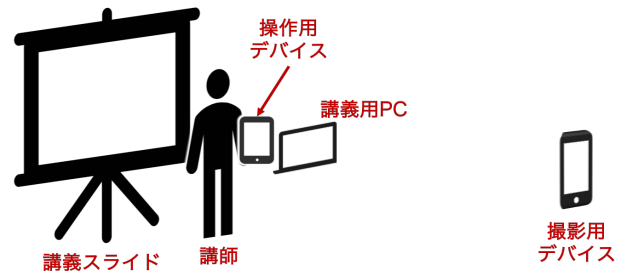


図1 動画撮影環境におけるハードウェア設置の様子

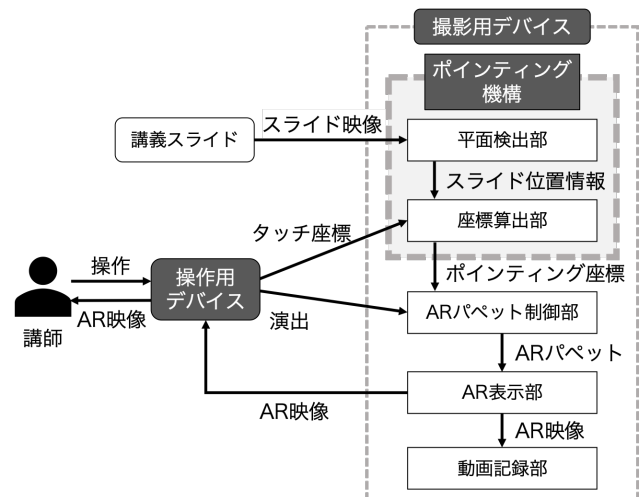


図2 システム構成図

ともに iPad を用いた。

図2に本システムの構成図を示す。本システムは、主に操作用デバイスと撮影用デバイスによって構成される。講師は、操作用デバイスのタッチスクリーンを用いて入力を行う。撮影用デバイス内のポインティング機構は、平面検出部と座標算出部を持つ。平面検出部は、カメラを利用しスライド映像を受け取り、平面を検出する。座標算出部は、スライド位置情報と操作用デバイス画面上のタッチ座標から、ポインティング座標を算出する。また、撮影用デバイス内のAR パペット制御部は、座標算出部から受け取ったポインティング座標を元にAR パペットのポーズを変更する。または、操作用デバイスから演出再生イベントを受け取った際は、対応する演出をAR パペットに適用する。AR 表示部は、講義映像にAR パペットを追加したAR映像を作成し、画面へ出力する。また、作成したAR映像は操作用デバイスへ転送し、操作用デバイスの画面に表示される。動画記録部はAR映像を保存する。

4.2 機能

スライド検出機能とAR パペットの自動設定機能について説明する。撮影用デバイスで本システムを起動すると、カメラ映像から平面検出を開始する。講師はあらかじめス

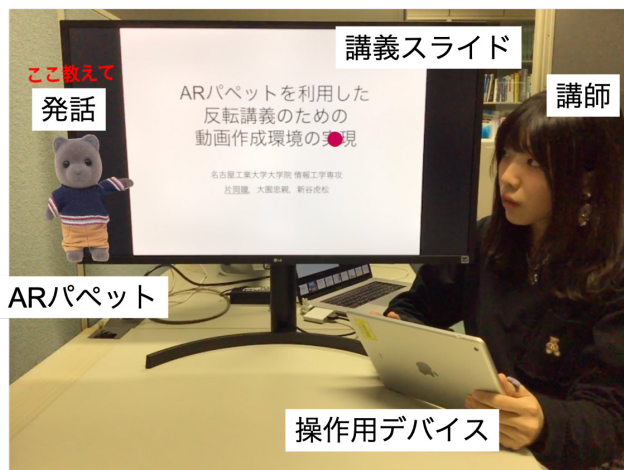


図 3 ポインティング機能の実行例

ライドを平面として検出を行ってから講義を開始する。この操作により、本システムはスライドの位置・大きさ・向き等を認識する。次に、検出したスライド情報に基づき、AR パペットの自動設定を行う。AR パペットの位置・大きさは、スライド情報に応じて相対的に自動決定する。講師は撮影用デバイスのカメラでスライドを映す動作のみで AR パペットの配置を行うことが可能である。

ポインティング機能について説明する。生徒が講義内容を理解することを助けるために、AR パペットがスライド中の任意の点を指し示す機能が必要であると考えた。図 3 はポインティング機能の実行例である。図 3 右の講師は、手元の操作デバイスのタッチスクリーンを操作してポインティング箇所を入力している。操作デバイスの画面には、作成された AR 映像がリアルタイムに表示される。タッチ座標がスライド位置と重なった場合に、ポインティング機能が動作する。講師がポインティングを行なっている間、AR パペットはその点に向けて腕を回転する。図 3 の講義スライド上の点はポインティング箇所を示しており、AR パペットは点の方向を左腕で指し示している。また、ポインティング実行中にパペットの頭上には「ここ教えて」という発話が表示される。

次に、演出再生機能について説明する。演出再生機能は、あらかじめシステムに登録済みの AR パペットの演出を、任意のタイミングで再生するための機能である。図 4 は操作デバイスの画面である。画面下部には演出再生ボタンが表示されている。講師がボタンをタッチすると、対応する演出が AR パペットに適用され、再生を開始する。ここでもまた、講師は操作デバイスの画面を通して、演出が再生されている様子をリアルタイムに確認することができる。図 4 では、AR パペットを仮定の生徒として対話形式の講義を行うことを想定し、AR パペットが講師の解説を理解したという場面で再生するための「わかった」の演出を再生している。AR パペットの頭上には「わかった！」という発話が表示され、両手を挙げるモーションを実行し

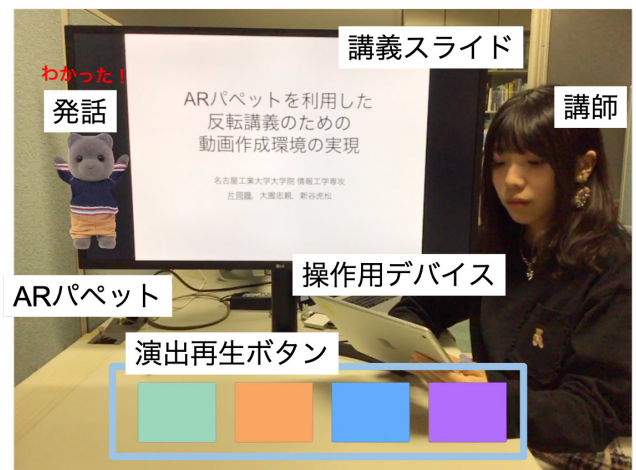


図 4 操作デバイスの画面。講師は演出再生ボタンをタップして AR パペットに「わかった」の演出を適用している。

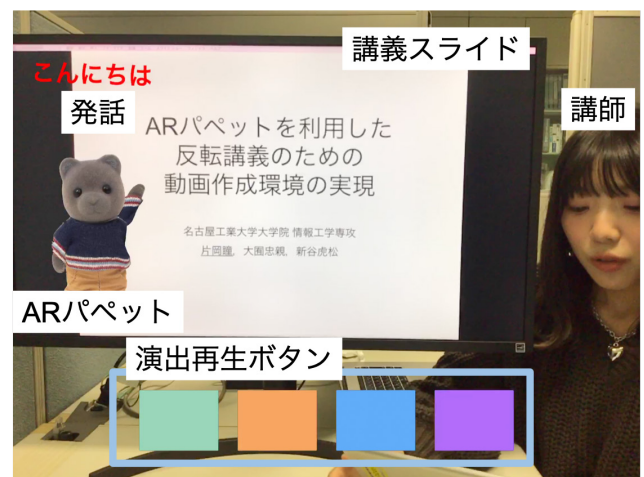


図 5 操作デバイスの画面。AR パペットに「あいさつ」の演出を適用している。

ている。図 5 も同様に、操作デバイスの画面である。講義の開始時に、AR パペットが講師に挨拶をするための「あいさつ」の演出を再生している。AR パペットの頭上には「こんにちは」という発話が表示され、右手を挙げて挨拶するモーションを実行している。

5. パペットの設計

本章では、開発した AR パペットの設計について議論する。パペットの表示方法について、本研究では、パペットを画面上に平面的に表示する方法と、AR 空間内に立体的に表示する方法を検討している [5]。パペットを AR 空間内の 3D オブジェクトとして表示することで、カメラの方向によって様々な方向から AR パペットを撮影することや、現実の環境に応じた表現を実現することが可能になると考えた。現実の環境に応じた表現とは、例えば現実の床の上に AR パペットを立たせる、スライドの位置を考慮した移動、講師とのインタラクティブな振る舞いなどである。

前で述べたように、本システムでは講師のタッチスク



図 6 AR パペットの元画像

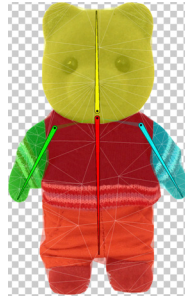


図 7 AR パペットの
パーツ分け・ボーン

リーンによる入力で AR パペットを操作する。スライドのポインティングは画面上のスライド映像をタッチすることで制御し、あらかじめ設定された演出を再生する場合にはボタンのタッチをトリガーとする。ここでは、作成中の動画に対し、画面をタッチする操作は直感的なインターフェースであると考えた。しかし、これは講義映像を撮影中に講師が操作用デバイスの画面を目視し手で操作を行うため、講師の動きが制限される問題があり、講義に影響を及ぼすケースが考えられる。この制限は、講師のジェスチャで AR パペットを制御する方法によって緩和できると考えた。PoseNet[6] によるリアルタイム姿勢推定を利用することで、動画作成のための機材を増やす必要がなく、撮影用デバイスのカメラ映像からジェスチャ制御を実現可能である。

本システムでは、あらかじめアニメーションエディタ上でパーツの位置やボーンの角度のパラメータを設定しシステムに登録することで、演出として利用している。AR パペットの演出効果作成において、パラメータを手入力で設定および調整することは手間がかかる。そこで、教示を用いて演出効果を作成する手法を利用することで、この手間を軽減できると考えた [7]。実物の人形を用いて動作の実演を行い、それを撮影することによって演出効果を作成する。ここでは、作成済みの演出効果を再生する機能として利用する場合と、リアルタイムに教示を行う場合が考えられる。

本システムで用いた AR パペットの作成方法について説明する。AR パペットは、2D 画像を利用したものと、3D モデルを利用したものの二通りを検討した。2D の AR パペットは、図 6 で示されるように、人形の写真から背景を切り抜いた画像を元に作成した。図 7 のように、頭、胴体、右腕、および左腕の 4 つのパーツに切り分け、それぞれのパーツにボーンを設定する処理を行う。図 7 は、AR パペットが 4 つのパーツに分けられ、それぞれ 1 つのボーンを持っていることを示している。

AR パペットの利用方法について議論する。本システムを用いて動画作成を行う場合において、3 つの利用方法が

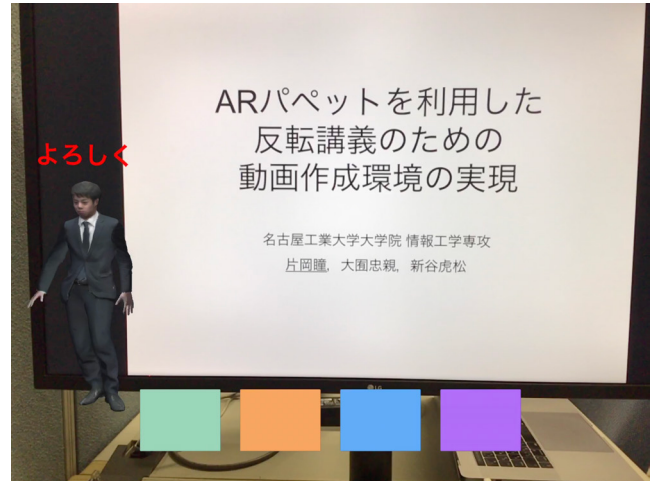


図 8 AR パペットを仮想の講師として動画作成を行なっている例。ここでは AR パペットとして、人間の 3D モデルを利用している。

考えられる。(a)AR パペットを仮想の生徒とし、講師との対話形式の講義を演出する方法、(b)AR パペットを仮想の講師とし、AR パペットによる講義を演出する方法、および (c) 複数の AR パペットによる講義を演出する方法である。(a) では、動画には講義スライド、講師、および AR パペットが含まれる。AR パペットが生徒の意見を代弁する演出によって、生徒の理解を促進することをねらいとする。AR パペットの操作は操作用デバイスを用いて行うため、講師が講義中に操作することも可能であり、または、講師とは別の操作者が操作することも可能である。(b) では、動画には講義スライドおよび AR パペットのみが含まれる。図 8 は、AR パペットを仮想の講師とした動画作成の例である。講師は動画に含まれないため、解説音声による講義と AR パペットの操作のみを行う。(c) では、例えば講師役と生徒役の AR パペットによる対話形式の講義などが想定される。

6. 考察

対話形式の講義に必要な演出効果の実現と、その操作方法について考察を述べる。本システムでは、講師と AR パペット間の挨拶と質疑応答という演出を実現するため、ポインティング機能、演出再生機能を実装した。いずれも操作デバイスの画面をタッチすることで動作する。ポインティング機能について、画面に表示されている講義スライドをタッチする操作はマッピングの観点において直感的であると考えられる。しかし、ポインティング中は画面を操作し続ける必要があるため、講師の動きが制限される。演出再生機能についてはボタン入力をトリガーとしており、これは非常に簡便な方法であるが、視認性や動きの制限の問題がある。ボタンと対応する演出を視覚的に確認しやすくするための解決策として、対応する演出のポーズを適用したパペットの画像を用いる方法が考えられる。

講師の動きが制限される問題に対する解決策として、5章で述べた通り、講師のジェスチャをトリガーとして演出を再生する方法が考えられる。この方法では、画面を目視する必要がなく、ビデオベースの姿勢推定技術により撮影機材を増やすことなく実現が可能である。しかし、正確さの問題や、状況によってはジェスチャを実行することで講義に影響を及ぼす可能性がある。また、講師はジェスチャを記憶することが望ましいため、ジェスチャと対応する演出のマッピングは直感的でなければならない。

また、AR パペットの見た目に關しての考察を述べる。動画を視聴する生徒の好みに合わせて生徒がAR パペットの見た目を変更したい場合や、講義内容に合わせて講師が動画作成時にAR パペットの見た目を変更したい場合が想定される。複数の3Dモデルから選択できるようにする機能、人形の画像などを取り込み2DのAR パペットを生成する機能などによって、見た目の変更が実現できると考えられる。

7. おわりに

ライブプレゼンテーションに基づく反転講義用動画作成において、講師の操作によって動作するAR パペットの開発について述べた。本システムは特別な機器を必要とせず、汎用的なモバイル端末のみで実行可能である。また、動画撮影中において、講師は操作デバイス（タブレット）のタッチスクリーンによる入力（ジェスチャ）でAR パペットを操作することが可能である。本システムは、撮影後の動画編集の手間を軽減することによって反転講義のための動画作成を支援する。

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, pp.81-84(2017)(FIT論文賞受賞).
- [2] N. Saquib, R. H. Kazi, L. Wei, and W. Li: “Interactive Body-Driven Graphics for Augmented Video Performance” Proc. CHI’ 19, ACM, Paper No. 622, 12p(2019).
- [3] F. Matulic, L. Engeln, C. Träger, and R. Dachsel: “Embodied interactions for novel immersive presentational experiences” Proc. CHI EA ’16, ACM, pp.1713-1720(2016).
- [4] H. Trinh, L. Ring, and T. Bickmore: “DynamicDuo: Co-presenting with Virtual Agents” Proc. CHI ’ 15, ACM, pp.1739-1748(2015).
- [5] 大園忠親, 落合裕也, 片岡瞳, 新谷虎松: 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, 日本ソフトウェア科学会 第36回大会, MACC, 51-L, 6p(2019).
- [6] G. Papandreou, T. Zhu, L. C. Chen, S. Gidaris, J. Tompson, and K. Murphy, Personlab: “Person pose estimation and instance segmentation with a bottom-up, part-based, geometric embedding model” Proc. ECCV, 17p(2018).

- [7] H. Kataoka, T. Ozono, and T. Shintani: “Realizing an Effect Editor for AR Pop-up Picture Books by Teaching Motions” IDDC2019, IEEE, pp. 989-994(2019).