

# AR空間における資料およびユーザの位置関係を考慮した ライブパフォーマンス支援環境の開発

片岡 瞳<sup>†</sup> 大園 忠親<sup>†</sup> 新谷 虎松<sup>†</sup>  
名古屋工業大学大学院情報工学専攻<sup>†</sup>

## 1 はじめに

講義やプレゼンテーションなどのパフォーマンスにおいて、対話の演出が効果的な場合がある。人間と仮想エージェントとのインタラクションを含むパフォーマンスについても多く研究されており、そのようなパフォーマンスを低コストで実現するという需要がある[1]。本研究では、人間と仮想エージェントとの共同のライブパフォーマンスを実現するため、拡張現実空間内の仮想エージェントであるARパペットを実現する。特に、講義のような資料を用いて視聴者に説明するようなパフォーマンスの動画作成を想定する。

ライブパフォーマンス支援システムにおいて、事前準備や動画編集の手間を軽減するという課題が存在する。そのため人間による手動制御を軽減し、ARパペットの半自律制御の実現を目指す。本研究では、AR技術を利用し、AR空間における資料およびユーザの位置関係を利用する。

## 2 ライブパフォーマンスにおける位置関係の考慮

AR技術を用いることで、映像中の立体的な座標を取得可能になる点に着目した。現実空間内の資料や人間の位置等を検出し、ARパペットとのインタラクションを実現する。そのような演出を実現するための機能として、特に(a)ポインティング機能、(b)資料検出機能、(c)画面レイアウト変更機能の3点について述べる。

(a)ポインティング機能は、ユーザのタッチ操作によって資料中の任意の点を指し示すための機能である。図1はポインティング機能の実行例である。ユーザは手元の操作用デバイスのタッチスクリーンを操作して

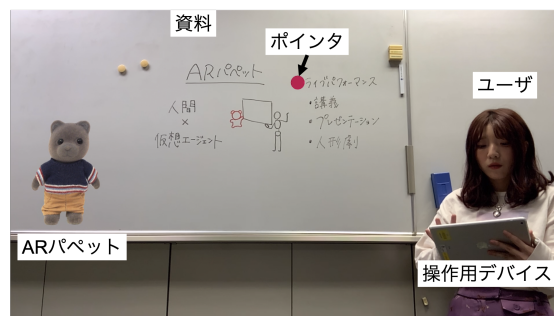


図 1: ポインティング機能の実行例

ポインティング箇所を入力している。映像中の資料の位置を考慮してポインタが表示され、ARパペットはポインタに視線を向けている。

(b)資料検出機能は、AR技術に基づく平面検出を用いて、資料の位置等を考慮したARパペットの自動設定を行う機能である。ユーザはあらかじめ資料の平面検出を行ってから撮影を開始する。この操作により、本システムは資料の位置・大きさ・向き等を認識する。ARパペットの位置・大きさ・向きは、検出した資料情報に応じて相対的に自動決定する。ユーザは撮影用デバイスのカメラで資料を映す動作のみでARパペットの配置を行うことが可能である。

(c)画面レイアウト変更機能は、資料、ARパペット、またはユーザの位置に応じて画面の一部分をクローズアップすることで、自動で画面レイアウトを変更するための機能である。AR技術に基づく平面や人間の座標検出により自動で位置関係を考慮した編集を行い、ユーザの編集作業の負担を軽減できると考えた。例えば資料の位置を元にトリミングを行うことで、資料のみをクローズアップするような演出を可能にする。

## 3 ライブパフォーマンス支援環境の実装

動画撮影環境には、資料、操作用デバイス、および撮影用デバイスが含まれる。映像は撮影用デバイスのカメラを用いて撮影する。操作用デバイスは、ARパペットを操作するためのデバイスであり、ユーザの手

Developing Live Performance Support System Considering Positional Relationship among User and Materials in AR Space

<sup>†</sup>Hitomi KATAOKA, <sup>†</sup>Tadachika OZONO and <sup>†</sup>Toramatsu SHINTANI

<sup>†</sup>Dept. of Computer Science, Graduate School of Engineering, Nagoya Institute of Technology.

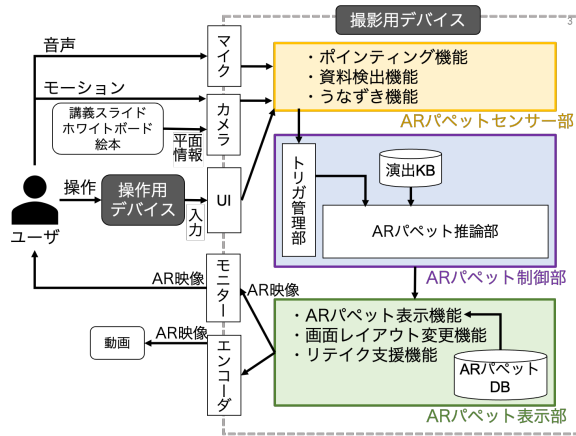


図 2: システム構成図

元にある。撮影用デバイスは、ユーザ、資料、および音声を含む映像を撮影し、また、AR パペットの表示や資料の検出など、AR 技術に基づく機能を担う。

図 2 に本システムの構成図を示す。撮影用デバイス内は、AR パペットセンサー部、AR パペット制御部、AR パペット表示部の 3 つに分けられる。AR パペットセンサー部は、マイク、カメラ映像、および操作作用デバイスからの入力を検知するための部分である。(a) ポインティング機能での操作作用デバイスからの入力や、(b) 資料検出機能でのカメラからの平面検出などを担う。AR パペット制御部は、AR パペットセンサー部からの入力に基づき AR パペットの行動を決定する部分である。トリガ管理部は AR パペットセンサー部が検出した情報を受け取り、それに応じた演出が格納された演出 KB をもとに、AR パペット推論部によって AR パペットの行動が推論される。AR パペット表示部は、AR パペット制御部が決定した AR パペットの行動を適用し、カメラ映像に AR パペットを追加した AR 映像として出力する部分である。AR 映像はモニターを通じユーザへのフィードバックにも利用される。また、(c) 画面レイアウト変更機能では、画面のクローズアップを適用する。AR パペット DB は、内容に応じて AR パペットの見た目を変更する場合を想定し、AR パペットのモデルが格納されている。

#### 4 評価実験

システムの性能を評価するために、ポインティング機能におけるタッチ操作から演出が実行されるまでの経過時間を計測した。ユーザはライブパフォーマンス中にリアルタイムに AR パペットを制御するため、操作から実行までの時間は重要である。ポインティング機能は操作作用デバイスの画面をタッチする操作によ

表 1: 平均経過時間

	平均経過時間 (s)
$\Delta t_1$	0.038
$\Delta t_2$	0.041

て動作する。操作作用デバイスでのタッチ開始時から撮影用デバイスでの AR パペットの演出開始時までの経過時間を  $\Delta t_1$ 、タッチ終了時から AR パペットの演出終了時までの経過時間を  $\Delta t_2$  とした。タッチの継続時間はおよそ 0.5~43.6 秒と様々で、100 回のタッチ操作を計測した。実行端末は、それぞれ操作作用デバイス iPad(6th generation) iOS 13.3、撮影用デバイス iPad Pro(2nd generation) iOS 14.0 である。

実験の結果を表 1 に示す。  $\Delta t_1$  および  $\Delta t_2$  の平均をそれぞれ表している。本システムは iPhone または iPad を利用するため、24~60fps での撮影が想定される。結果より、30fps 以下の動画ではほとんど影響がないと考えられる。また、ユーザはモニターを通して映像のフィードバックを受けながらポインタの位置合わせ等を行う。一般的なモニターは 60hz 程度であり、ユーザによっては遅延を感じる可能性がある。しかし、本研究の対象である講義動画等におけるポインティングという用途について実用的な性能であるといえる。

結果は  $\Delta t_1 < \Delta t_2$  であるが、  $\Delta t_1$  と  $\Delta t_2$  の差、すなわちタッチ開始時とタッチ終了時の差はごく小さい。したがって、ポインティング機能を実行中の処理による遅延はほとんど見られないと考えられる。また、タッチの継続時間による経過時間への影響は見られなかった。

#### 5 おわりに

本稿では人間と AR パペットとの対話の演出を含むライブパフォーマンス支援システムについて述べた。AR 空間における資料およびユーザの位置関係を考慮した演出として、ポインティング機能、資料検出機能、および画面レイアウト変更機能を実装した。評価実験では、本システムがライブパフォーマンスにおいて実用的な性能であることを示した。

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

#### 参考文献

[1] Y. J. Lin et al. "A Human-Computer Duet System for Music Performance" Proc. of the 28th ACM International Conference on Multimedia, pp.772-780(2020).