

SMART MOVIE: ARグラスとスマホを併用した 任意の場所で大画面かつ見やすい映像を視聴するシステム

武田 和大^{1,a)} 真鍋 宏幸^{1,b)}

概要: ARグラスで映像視聴する際、画面の大きさを自由に変更できるが、現実世界が透けて見づらくなる一方、スマホでは画面の大きさがデバイスの大きさに制限される。そのため両者を組み合わせた映像視聴システムが提案されたが、スマホの位置が固定されるという課題がある。そこで本研究では、ARグラスとスマホを併用してスマホを自由に移動できる映像視聴手法を提案する。提案手法の実装を行い、スマホを移動させて映像視聴できたが、遅延や2つの映像のズレが生じるなどの課題があった。



図 1 実装のイメージ図 (映像^{*1}は変更可能なものを使用)。

1. はじめに

Augmented Reality (AR) 技術が進歩しており、それを利用したコンテンツが増加している。AR を利用することで、現実世界に仮想オブジェクトを表示させたり、それらとのインタラクションを行うことができる。AR にはいくつかの実現形態があるが、光学シースルーなディスプレイを備えた頭部装着型デバイスである AR グラスを用いた場合には、ハンズフリーかつ 3 次元の入力を行うことができる。AR グラスは AR を体験するために設計されたデバイスであるが、映像視聴デバイスとしても利用することができる。AR グラスでは画面の大きさを任意に変更できるが、映像の背景に現実世界が透けてしまい、見づらくなってしまふ。一方、一般的な映像視聴デバイスとして用いられるスマホやモニタでは、画面の大きさがそれぞれのデバイスの大きさに制限されてしまふ。

この課題に対し花田ら [1] は、AR グラスとスマホを併用

して映像視聴する手法を提案したが、スマホの位置が固定されてしまふという課題があった。そこで本研究では、花田ら [1] の手法にあったスマホを移動できない問題の解決に取り組む。

2. 関連研究

複数のデバイスを組み合わせることで、それぞれのデバイス単体ではデメリットとなる面を補完し合うことができる。例えば、Jones ら [2] は、ディスプレイの周辺にプロジェクタで周辺情報を追加することで、映像への没入度を高める Illumiroom を提案した。この手法では、映像への没入度を高めることができるが、ディスプレイやプロジェクタが使用できる場所に制限されてしまふ。

AR グラスとその他のデバイスを併用した研究も行われている。MARVIS[3] では、AR グラスとタブレットを組み合わせてデータ分析を行った。机上での 6 つのユースケース例を実装し、AR グラスとタブレットで関連した別の内容を表示させたときに有用であったと報告されている。MultiFi[4] は、スマートウォッチの画面を AR グラスで拡張することで、ユーザへの高い負荷があるものの、それぞれ単体での作業時間より短縮できることを示した。これらの研究は、モバイルデバイスを入力デバイスや AR マーカーとして利用している。

AR グラスとスマホを組み合わせた研究もされている [5][6]。Normand ら [5] は、2 つのデバイスを組み合わせる手法が単体での入出力より好まれ、タスク完了時間も早くなることを示した。Hubenschmid ら [6] は、空間記憶や作業負荷などの面で最適なディスプレイサイズを求めた。これらは、入力デバイスとしてスマホのディスプレイを使

¹ 芝浦工業大学

^{a)} al20096@shibaura-it.ac.jp

^{b)} manabehiroyuki@acm.org

^{*1} <https://www.pexels.com/ja-jp/video/6084026/>

用している。

花田ら [1] は、スマホに中心部分の映像を表示し、AR グラスでその周辺映像を表示させることで映像視聴する手法を提案した。この手法は、映像の種類によっては没入度や臨場感を高めることを示した。しかし、映像の中心部分にスマホを固定する必要があった。

3. 提案手法

本研究では、AR グラスとスマホを併用し、任意の場所で大画面かつ高解像度な映像を視聴可能なシステムを提案する。映像の画面を AR グラスとスマホの 2 つに分割し、スマホで表示されていない範囲の映像を AR グラスでスマホの周辺に表示する点は、花田らの手法 [1] と同じであるが、スマホのトラッキングを行うことでユーザが自由にスマホを動かすことを可能にする点が異なる。そのため、映像の中心部分のみならず、映像の任意の場所をスマホで視聴することができる。これにより、映像全体を常に AR グラスで大画面で視聴しながら、ユーザが注視したい範囲をスマホの高解像度で視聴できる。

4. 実装

AR グラスとして Microsoft 社の HoloLens2、スマホには OPPO Find X3 Pro を用いて実装を行った。

スマホのトラッキングにはいくつかの方法が考えられるが、本実装では HoloLens2 のハンドトラッキングを利用した。スマホのトラッキング用に追加のデバイスが不要で、ライブラリを使えば簡単に実装することができる。スマホの持ち方を予め決めておき、ハンドトラッキングで得られる手の位置からスマホの位置を求める。具体的には、図 2 のようにスマホの四隅を両手の親指と人差し指で保持し、その 4 本の指で囲まれてできる長方形をスマホと推定している。スマホは回転させないようにし、常にスマホのディスプレイ面が AR グラスの映像と平行になるように保持する。推定したスマホの位置データをサーバに送信し、スマホ側ではその位置の映像のみを表示させる。スマホの位置推定の後、AR グラスでその範囲に対して黒色でマスク処理を行う。これは、AR グラスで黒色のオブジェクトを表示するとその部分は発光せず、何も表示していない状態と等しくなる性質を利用している。これにより、AR グラスの映像の内スマホと重なっている範囲に関しては、スマホの映像のみを見ることができるようになる。

HoloLens2 とスマホの映像を同時に再生させると、先に HoloLens2 の映像が開始され、映像開始時間に差が発生してしまうことがわかった。そのため、それぞれの動画を 10 回再生したときの開始時間の差の平均を求め、HoloLens2 ではその時間分待機させることで映像の同期を図る。

AR グラスとスマホのアプリケーションは共に Unity で作成し、Holographic Remoting Player を用いて、PC で処

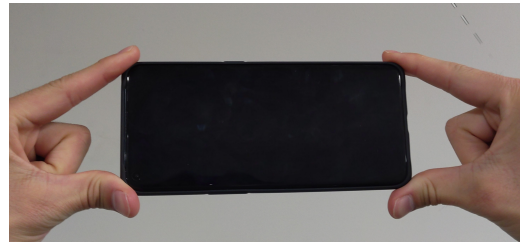


図 2 スマホを保持している様子。

理した映像を HoloLens2 に表示させている。スマホの位置情報の送信には Websocket を用いており、HoloLens2 からスマホに送信するため、Python と Cloudflare Tunnel を用いた Websocket サーバを作成した。なお、映像は改変可能なものを使用した。本手法で映像視聴している様子のイメージ図を図 1 に示す。

5. 実装の評価

今回の実装に対し、(1) AR グラスが提示する映像とスマホの映像の空間的なズレ、(2) AR グラスの映像とスマホの映像の時間的なズレ、(3) スマホを移動させたときの追従性、(4) 映像視聴したときの印象評価、の 4 点について評価を行う。(1) と (2) については、AR グラスとスマホの映像が空間的に連続していて、全体で 1 つの映像となっているのかを評価する。(3) は、スマホを移動させたときに、AR グラスとスマホの映像が更新されるまでの遅延時間や空間的なズレを評価する。

5.1 映像の空間的なズレ

縦横それぞれ 10px 毎に線を引いた 1080 × 1920px の映像を目の位置から約 0.5m 離れた所に AR グラスで表示し、スマホは映像の中心に固定するように約 5 分間保持した。スマホには 160px 毎に縦線が、90px 毎に横線が引かれた映像を表示させ、2 つのデバイスの境界線のずれを目視で 1px 単位で求めた。HoloLens2 の片目のディスプレイ解像度は 2K、視野角は対角に約 52 度、スマホのディスプレイ解像度は QHD+ である。手振れの影響もあるが、結果として水平方向には最大で約 15 から 20px、鉛直方向には約 10px のズレが生じた。

5.2 映像の時間的なズレ

本システム使用時に映像が同期しているかを検証するため、それぞれのデバイスの映像開始時間の差を測定した。3 つの映像に対する開始時間の差を、30fps の HoloLens2 の内蔵カメラで 10 回測定し、その平均や標準偏差、誤差の最大値を求めた。結果は動画ごとに異なっており、ある動画は平均フレーム差が約 0.5 で標準偏差が約 1.75、最大 3 フレームの差があり、別の動画の平均フレーム差は約 2 かつ標準偏差は約 2、最大フレーム差は 5 であった。

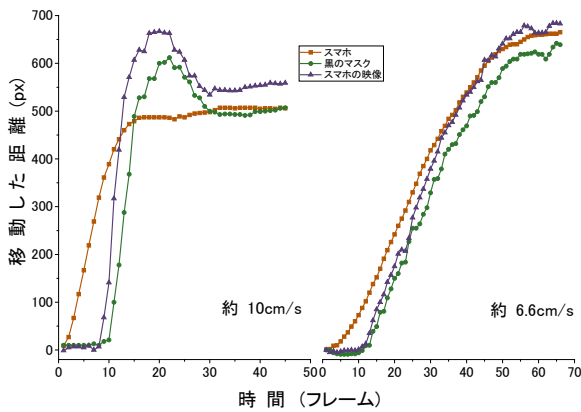


図3 トラッキングの追従性. (左) スマホを約 10cm/s で動かした時, (右) スマホを約 6.6cm/s で動かした時.

5.3 スマホの追従性

手でスマホを移動させた際, ハンドトラッキングの遅延や, サーバ経由によるスマホの映像更新の遅延などにより, スマホと AR グラスの映像にズレが生じる. ここではその遅延について評価を行った. 目の位置から約 0.5m 離れた所に, 10px 毎に縦線を引いた 1080 × 1920px の映像を用意し, 映像の中心部分にスマホを置いた状態を初期状態として, そこからスマホを水平右方向に動かした. 映像の同期評価のとき同様, その時の様子を 30fps の HoloLens2 の内蔵カメラで撮影し, スマホの位置, HoloLens2 で黒のマスク処理した位置, スマホに表示している映像の位置の3つについて, 毎フレーム目視で 1px 単位で求めた. このカメラはユーザの視点とは別の座標にあり, 実際にユーザが見ている映像とは異なるため, 初期状態を基準として評価する. 約 10cm/s と約 6.6cm/s の速度でスマホを動かしたときの結果を図3に示す.

図3の緑と紫のグラフの概形より, AR グラスからスマホに通信していることによる遅延は, ほとんどないことが分かる. また, 同じ移動距離の所を比べると, 先にスマホの映像が変化した後, AR での黒のマスクの位置が変化している. しかし, スマホを動かす速度が速いと, 両者の位置のズレが大きくなっている. また, 速度が異なる2つのグラフを比べると, スマホの移動速度が速いと, 実際のスマホの位置とハンドトラッキングにより求めたスマホの位置のずれが大きくなり, 実際の位置を超えてから実際の位置に戻るといった軌道をたどっている. スマホの移動速度に関わらず, スマホを動かしてから AR グラスで実際のスマホの位置推定までに最大約 10 フレームの遅延が発生しており, 速度が速いときにはオーバーシュートが約 15 フレームの間継続している.

5.4 印象評価

3人に本手法で3つの映像を視聴してもらい, 上記の(1)から(3)についてのフィードバックをもらった.(1)は, 1

つの動画には全員がズレを感じていたが, ほかの2つの動画は人によって異なり, 2人はズレを感じ, 1人はほとんど感じなかった.(2)については, 1人は少し気になったものの, その他の2人はほとんど感じなかった.(3)は, スマホをゆっくり動かしたときにはさほど遅延が気にならない人が1人いたが, 全員が, 普段使用するときのようにスマホを動かすと遅延や空間的ズレをかなり感じていた.

6. 議論

6.1 ハンドトラッキングによる影響

5.4節で人によって空間的なズレの感じ方が異なった. 現在は図2のようにスマホを保持しているが, ハンドトラッキングでは指の腹の中心を取得しているため, 正確に指の腹をスマホの四隅に当てるのが難しい. また, スマホを保持する指の位置が変わるだけでもスマホに表示される映像が変化してしまい, 映像の空間的なズレが発生してしまう. そのため, スマホを保持する指の位置が異なったため, 人によって空間的なズレの感じ方に違いが生まれた可能性がある. そこで, 指の関節からスマホの位置を推定できるようにスマホの保持の仕方を変えることで, ユーザ毎の空間的なズレが軽減されることが期待される.

5.3節より, スマホの動かし始めてからハンドトラッキングでスマホの位置を更新するまでに約 10 フレームの遅延や, 速度が速いときに約 15 フレームの間オーバーシュートが起きているが, これは HoloLens2 のハンドトラッキング自体の問題である. そのため, ハンドトラッキングを利用している限り, スマホを動かしてから HoloLens2 で位置推定するまでの遅延は避けられない. また 5.4 節では, スマホを移動させた時の追従性について全員が, AR グラスで黒のマスク処理している部分やスマホの映像の遅延を感じていた. この追従性の遅延は, AR グラスとスマホの空間的なズレにも影響している. これらの問題について, HoloLens2 のハンドトラッキングのみならず, 深度カメラやスマホの IMU のデータを利用することで改善できる可能性がある.

6.2 映像の時間的なズレ

5.4節において, 時間的なズレをほとんど感じない人がいたが, 少し感じる人もいた. 図3では, AR グラスとスマホの通信による遅延はほとんどなかったが, 5.2節より, 映像再生毎に開始時間の差が変化するため, その差が大きかったことでズレを感じた可能性がある. 映像毎にも開始時間の差が変わるため, 今後はどのような映像に誤差が生じやすいのか, 誤差を減らせる実装方法はないかを調査していく.

7. まとめ

本研究では, AR グラスとスマホを併用し, 任意の場所

で大画面かつ見やすい映像を視聴可能なシステムを提案した。AR グラスのハンドトラッキングを用いて実装を行い、それによる映像の位置ずれや遅延などの評価を行った。映像の同期の誤差はあまりなかったが、ハンドトラッキングの遅延や実装上の制限によるものが大きく影響したため、その解決案などについて議論した。

参考文献

- [1] 花田 晶彦ほか: AR グラスとスマートフォンを併用した映像拡張手法における基礎検討, 情報処理学会研究報告 ヒューマンコンピュータインタラクション (HCI) 2023-HCI-201(34), pp. 1-8 (2023).
- [2] Jones, B. R. et al.: IllumiRoom: peripheral projected illusions for interactive experiences, *Proc. CHI '13*, pp. 869-878 (2013).
- [3] Langner, R. et al.: Marvis: Combining mobile devices and augmented reality for visual data analysis, *Proc. CHI '21*, pp. 1-17 (2021).
- [4] Grubert, J. et al.: Multifidelity: Multi fidelity interaction with displays on and around the body, *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 3933-3942 (2015).
- [5] Normand, E. et al.: Enlarging a smartphone with ar to create a handheld vesad (virtually extended screen-aligned display), *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, IEEE, pp. 123-133 (2018).
- [6] Hubenschmid, S. et al.: ARound the Smartphone: Investigating the Effects of Virtually-Extended Display Size on Spatial Memory, *Proc. CHI '23*, pp. 1-15 (2023).