

# 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

花田 晶彦<sup>1,a)</sup> 真鍋 宏幸<sup>1</sup>

**概要:** 本論文では、ARグラスとスマートフォン（以下、スマホ）を併用し、スマホのディスプレイを拡張させる手法を提案する。スマホで出力される映像の周囲に、その外側の映像をARグラスから出力することで、個々の単一デバイスを用いた時とは異なる映像体験の提供を目指した。先行研究では主に、据え置き型のディスプレイとプロジェクタを併用した映像拡張について取り組まれてきた。提案手法では、プロジェクタではなくARグラスを用いる。そのため、背景は視聴の都度変化し、さらにそれが透過する。また、明るい環境下での視聴が想定されるだけでなく、手で把持したスマホを完全に固定することができないという違いがある。これらの違いは映像体験に影響を与える可能性がある。そこで、それらの違いがユーザの映像体験に与える影響について、複数の映像を用いた調査を行った。その結果、提案手法は新たな視聴体験を提供すること、また映像拡張が有益な効果をもたらす映像がある一方で、そうではない映像種類があることが明らかとなった。

## 1. はじめに

Augmented reality (AR) を多くの人が体験できるようになった。例えば、スマートフォン（以下、スマホ）を使用したARコンテンツ、ポケモンGO<sup>\*1</sup>である。ポケモンGOではスマホの位置情報を利用し、ポケモンを仮想オブジェクトとして表示させている。ARを導入したサービスやビジネスも増えた。IKEA Place<sup>\*2</sup>では、実際に販売されている家具を仮想オブジェクトにし、スマホで簡単にレイアウトやインテリアを考えることができる。また、車の部品の組み立ての現場や、電子部品の制作の現場などでは、ARグラスであるHoloLensを導入し、作業の効率化を図っている<sup>\*3</sup>。

このように、すでにARが様々な分野で利用され始めている。HoloLensなど現在のARグラスは、重くかさばり、視野角が狭いなどの課題がある。しかし、それらは技術の進展に伴い徐々に解消されていこう。そして、一人一人がARグラスを所持し、日常的に、ARグラスを装着し続ける未来が来ることが期待される。そこで本論文では、

日常生活の中で使用されるARグラスの活用方法に着目することとした。

そのような未来において、現在のスマホの機能を、ARグラスが完全に補うことは考えにくい。そのため、ユーザはスマホとARグラスの二つのデバイスを上手く使い分けて日常生活を送ることになるだろう。しかし、スマホは、高画質であるものの、ディスプレイのサイズが小さく、一方ARグラスでは、背景が透けてしまうため、映像が見づらくなることもある。それぞれのデバイスの問題を解決するために、二つを併用する選択肢が考えられる。両者を併用することで、スマホの高画質な映像体験と、ARグラスでの大画面での映像体験の両立が期待できる。実際、すでに2つのデバイスを併用する例はある。Illumiroom[5]では、プロジェクタとテレビを併用し、テレビの周囲にプロジェクタからの映像を投影させた映像拡張に成功している。

提案手法では、IllumiRoomでのテレビの代わりにスマホを用いて高画質の映像を出力し、プロジェクタの代わりにARグラスを用いて、ユーザの臨場感や没入感を高める拡張された映像を映し出す。IllumiRoomでは、ディスプレイが固定されている暗室での使用が想定されていたが、提案手法では、プロジェクタではなくARグラスを用いる。そのため、背景が決まっておらず、明るい環境での使用が想定される。また、提案手法は、ユーザが把持するスマホを用いているため、ユーザの多少の揺れが、提案手法の映像拡張に、違和感を与える可能性がある。提案手法を用い

<sup>1</sup> 芝浦工業大学

Shibaura Institute of Technology

<sup>a)</sup> ma21118@shibaura-it.ac.jp

<sup>\*1</sup> <https://pokemongolive.com/> accessed in Oct. 2022.

<sup>\*2</sup> <https://apps.apple.com/us/app/ikea-place/id1279244498> accessed in Oct. 2022.

<sup>\*3</sup> <https://news.microsoft.com/ja-jp/2020/10/06/201006-toyota-motor-started-introducing-hololens-2-in-gr-garage-nationwide/> accessed in Oct. 2022.

映像視聴は、映像の種類によって、ユーザの映像体験に影響を与えると考えられる。そこで、本研究では複数の映像を用いた調査を行い、その結果について報告する。

## 2. 関連研究

ARを用いることで、現実のオブジェクトの一部を拡大させ、ユーザの視覚的なサポートを可能にしたり [17]、現実のオブジェクトと仮想のオブジェクトを重ね合わせ、ハプティックデバイスによる視界の妨げを無くしたり [3]、現実のオブジェクトを透過させ [6]、目に見えない領域のものを、把握できるようになる。このように AR を用いることで、これまでにないインタラクションを提供することができる。提案手法も AR の一種であるが、提案手法では AR グラスとスマホを用いて映像領域を拡張させる。そこで重要な関連研究として、ディスプレイを拡張する手法 [2], [5] や、2つのデバイスを用いた AR 技術 [4], [11] を挙げる事ができる。以下でそれら関連研究について述べる。

### 2.1 ディスプレイの拡張

ディスプレイを拡張させるシンプルな方法は、ディスプレイを大型化させることである。しかし、サイズの大きなディスプレイは、ユーザの入力手段に大きな制約を強いる。制約の一つは、大型ディスプレイを用いた場合、ユーザはディスプレイ上のオブジェクトを直接操作することが難しいことである。Azam らは、大型ディスプレイの直接ユーザが操作できない領域に対して、フリスビーと呼ばれる手法を考案し、リモートでの操作を行う事で大型ディスプレイの操作を可能にしている [7]。我々の提案手法は、AR グラスを用いた映像拡張である。AR グラスを用いることで、実際に操作すると手が届かずユーザの操作が制約されるような、大型なディスプレイであっても、ハンドレイをはじめとしたジェスチャ入力、簡単に操作できるようになる。また、仮想のディスプレイであるため、ユーザは能動的にディスプレイサイズを変更することもできる。

人間の視覚特性として、中心視野は細部、色、質感を感じ [13]、周辺視野は、はるかに低解像度で、コントラスト [1] と動き [10] を感知する。この視覚特性を活かし、二つのデバイスを組み合わせることで、映像拡張を実現させた研究がある。IllumiRoom [5] では、映像の中心をテレビに表示し、その周囲の部屋の壁や床などの空間に対し、プロジェクタから映像を出力している。プロジェクタによって周囲に投影された映像が、見かけ上の動きを誘発することで没入感や臨場感を生み、新しい映像体験をユーザに提供した。投影されたプロジェクタの映像はテレビに比べ解像度が低い、前述した人間の視覚特性から、ユーザは違和感なく映像を視聴することができる。また、IllumiRoom に関連する事例として、テレビの周りに表示される映像を DNN を用いて推定する研究 [8], [15] もある。

視覚特性を利用した別な手法として、フォビエイテッドレンダリングがある [12]。フォビエイテッドレンダリングとは、VR の映像の全てを高解像度で表示するのではなく、ユーザが注視している領域だけを高解像度で描画し、そこから外側にむかって解像度を落とすように表示させる技術である。高解像度で描画する範囲が少なくなることで、ハードウェアへの負担が減少する。我々の研究も、周辺視野における人間の視覚特性を利用しており、スマホと AR グラスを組み合わせることで、広い視野角と高精細な中心視野の実現を可能にする。

人間の視覚特性を HMD に活用した研究の例として、HMD の中心視野に高解像度の映像、周辺視野に低解像度の映像を提示した研究 [18], [19] がある。高解像度の映像の周囲に低解像度の映像を組み合わせることで HMD の視野角を拡大させ、HMD の大きな課題の一つであった視野角の狭さによる制約を軽減し、映像に対する没入感を向上させた。これらの研究では主に VR 向けが想定されているのに対し、本研究では AR を対象としている。

AR グラスを用いて映像を視聴する先行研究については FoveAR [2] がある。FoveAR では、AR グラスに表示された映像の外側をプロジェクタで映すことで、広角の AR 体験を可能にしている。我々の研究は、現在発売されている HoloLens 2 の視野角 (対角 52°) が、FoveAR で使用されていた AR グラスである OST グラスの視野角 (対角 40°) と比べ広がったことを考慮しており、FoveAR におけるプロジェクタの役割を AR グラスにもたせ、FoveAR における AR グラスの役割をスマホにもたせている。また、提案手法では、プロジェクタではなく AR グラスで映像を表示させていることから、場所に固定されない使用ができる特徴がある。

### 2.2 スマホと AR グラスを併用した研究例

AR グラスでの入力は、スマホでの入力と比べ、物理フィードバックの欠如 [14] や腕の疲労などの課題がある。この課題に対し、AR グラスとスマホを併用する方法が試みられている。例えば、HoloLens 上の仮想オブジェクトを、スマホを用いて操作する研究 [9] では、HoloLens で採用されている空中ジェスチャよりもはるかに効率的な操作を可能としており、HoloLens とスマホの併用が有益であることが示されている。Unlu らの研究 [16] では、スマホ上に AR マーカを表示させ、それに対応する仮想オブジェクトを HoloLens 上で表示させることで、様々なアプリケーションを実装している。これらの研究 [9], [16] では、高解像度のディスプレイや、様々なセンサを搭載しているスマホの幅広い機能を活用して、AR アプリケーションを提案しているものの、あくまでスマホを AR グラスのコントローラとして利用している。一方、提案手法では、スマホを AR グラスのコントローラとしてではなく、映像を写す

スクリーンとして使用し、AR グラスとの融合を目指している。

操作性の向上ではなく、映像を拡張させることを目的として、AR グラスと他のデバイスを併用する研究も行われている。例えば、スマートウォッチと AR グラスの組み合わせ [4] や、スマホと AR グラスの組み合わせ [11] がすでに取り組まれている。スマートウォッチと AR グラスを組み合わせることで、作業負荷は高いものの、タスク完了時間の点において、単一のウェアラブルデバイスによるインタラクションを上回ることが確認された [4]。また、文献 [11] では、スマホの映像を AR グラス上に拡張することで、大きなディスプレイでの出力が好ましい地図や、画面占有率の高いコンテンツの利用を可能にしている。その中で紹介されているコンテンツのひとつに、AR グラスとスマホを用いて Google マップを表示させた例がある。ディスプレイを拡張したことで、大きなサイズの地図を表示できるようになった。このコンテンツは、本研究が目指す視聴体験に類似している。我々は、Google マップのような情報提示のためのコンテンツではなく、映画などのように映像そのものがコンテンツとなるもの（画面全体を同時に視認できることが前提となっているコンテンツ）の視聴体験を想定している。そこで、先行研究では言及されていなかった臨場感や見やすさなどの、ユーザの視聴体験が映像の種類によって、変化するかを調査する点が異なる。

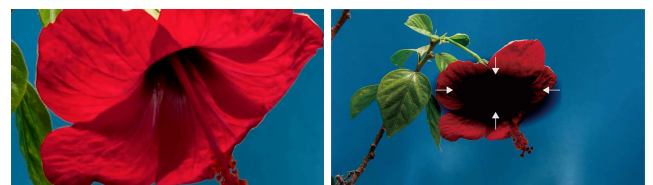
### 3. 提案手法

今回我々は、スマホの映像の周囲に AR グラスの映像を表示させる映像拡張手法（図 1）を提案する。スマホ単体での映像視聴は、高画質な視聴体験が可能であるものの、ディスプレイが小さいため、AR や VR などでの映像視聴と比べ、臨場感などの観点で劣ってしまう。一方で、AR グラス単体での視聴体験は、スマホと比べ、グラスから表示された映像の背景が透過してしまうことで、見づらい映像になってしまうことが考えられる。そこで、二つのデバイスを併用することで、高画質かつ、広いスクリーンでの映像体験の提供を目指す。

提案手法では、AR グラスを用いているため、AR グラス上に表示された映像は、第三者からは見えず、壁などの固定された広い空間を必要としない。そのため、複数のユーザが同じ部屋で同時に、それぞれ異なる種類の映像を体験できる。一方、据え置き型のディスプレイとプロジェクタを併用した映像拡張の研究 [5] では、プロジェクタが用いられており、映像を表示させる位置が壁に固定されている。そのため、映像拡張を実現するために必要な空間が広い。さらに、提案手法ではスマホを用いており、スマホの位置をユーザが能動的に変えることができる。そのため、ユーザの関心領域がスマホ内の映像ではなく、AR グラス上に表示された範囲の映像にある場合、スマホを移動させるこ



図 1: 提案手法



(a) スマホ側

(b) AR グラス側

図 2: それぞれのデバイスで表示させる映像

とで、高画質な関心領域の映像を視聴することができる。据え置き型のディスプレイとプロジェクタを併用した映像拡張の既存研究では、据え置き型のディスプレイであることからユーザが能動的に位置を変えることはできない。

また、提案手法と先述した既存手法 [5], [11] には、使用する環境や状況においていくつかの違いがある。一つ目は、提案手法では、スマホと AR グラスの両方の映像が重なり、図 1 のように同一の映像として映される点である。既存の研究 [2], [5] では、テレビや AR グラスで表示されるメインコンテンツに対し、没入感を高めるエフェクトなどのサブコンテンツを、メインコンテンツ周囲にプロジェクタで出力している。これは、スマホと AR グラスを用いて拡張された Google マップを表示した例 [11] でも同様であると我々は考えている。しかし、本研究では、スマホで表示される映像の外側の映像を AR グラスに出力している。つまり提案手法では、既存研究 [2], [5] と異なり、AR グラス上で表示される外側の映像の領域にも、メインのコンテンツの一部が表示されている。そのため、ユーザはスマホだけではなく AR グラスの映像にも注視する場合がある。二つ目は、AR グラスが明るい環境下でも使用できる点である。提案手法では、プロジェクタではなく AR グラスを用いるため、暗室での使用が想定されている据え置き型の映像拡張の既存研究 [5] と違い、部屋の明暗による制約を受けない。最後は、手でスマホを持っている点である。提案手法では、据え置き型のディスプレイではなく、スマホを使用している。スマホは手で把持した状態での使用が想



図 3: 使用した映像の種類

定されるため、ユーザの多少の揺れが、提案手法の映像拡張に、違和感を与える可能性がある。

提案手法には上記のような特徴があることから、視聴する映像の種類によって、映像視聴体験に違いが生じることが考えられる。そこで本研究では、提案手法での視聴方法と他の視聴方法の比較および、映像の種類による視聴体験の影響に関する調査を行うこととした。

#### 4. 実装

提案手法を実現するためには、AR グラスで表示されている映像の中で、どの位置にスマホが存在しているのかを常時トラッキングし、その位置に応じて2つのデバイスにそれぞれ表示させる映像を変更していく必要がある。しかし本研究ではまず、視聴する映像の種類の違いが、提案手法での映像視聴体験に違いを生むのかについて着目する。そのため、実験で用いたプロトタイプでは、スマホのトラッキングを行わず、スマホは特定の位置に手で固定された状況を想定することとした。

今回使用したものは、スマホ端末の iPhone12 と AR グラスの HoloLens 2 である。はじめに、スマホと AR グラスで表示される映像をそれぞれ制作した。映像体験に影響を与える要素として、提示する映像の背景の明るさ、映像のスピード、映像のメインコンテンツの位置が考えられた。そこで、上記の要素による影響を調査するため、図 3 に示す 10 種類の映像を集めた。映像の明暗が影響を及ぼすのかを調べるため、背景が明るいバイクの映像 (図 3(1))、背景が緑色の遺跡の上空撮影の映像 (図 3(7))、そして背景が暗く黒色の花火の映像 (図 3(9)) を選定した。映像のスピードが提案手法での影響を及ぼすのかを調査するため、速度の速いジェットコースターの映像 (図 3(6))、速度の遅い開花のタイムラプスの映像 (図 3(2))、都市の早送り映像 (図 3(5)) を選んだ。スマホの内部に映像のメインコンテンツが収まっているか、そうでないかによって提案手法での視聴方法に違いが生じるのかを調べるために、観客の映像 (図 3(3))、花と鳥の映像 (図 3(4))、サッカーの映

像 (図 3(8))、相撲の映像 (図 3(10)) を選んだ。映像は、pexels<sup>\*4</sup>や istock<sup>\*5</sup>のサイトから収集し、編集した。なお、編集した映像は、全て映像改変が可能な映像である。使用した映像の画素数は、横 1920、縦 1080 である。スマホで表示される映像の画素数は一定 (2532 x 1170) であるのに対し、AR グラスで表示される映像の画素数は、ディスプレイサイズをユーザが能動的に操作できるため、可変である (最大で 2k)。

スマホには、映像の中心部を表示させ (図 2a)、AR グラスで映像の周辺部を表示させた (図 2b)。両者を組み合わせると、ユーザからは図 1 のように見える。図 1 では、AR グラスで表示された仮想ディスプレイの背後に、片手で保持したスマホが見える。また、スマホで表示されている領域と AR グラスで表示されている領域とは色味が異なることもわかる。AR グラス上で黒色に表示される映像領域では、後ろの背景が最もよく透過する。その特徴を踏まえ、図 2b のように、映像の中心部を黒く塗りつぶした映像を AR グラスに表示させた。図 2b 内の白の矢印は、映像の中心にいくに従って、黒色での塗り潰し度合いを上げている様子を示している。映像の中心部に向かって、塗り潰し度合いを徐々に上げることで、スマホと AR グラスから出力した映像との境界線がはっきり見えることを防ぎ、スマホと AR グラスの映像が自然に重なることを目指した。

次に、映像の再生を行う方法である、上記の方法で制作した映像を表示する web サーバを、Node.js を用いてローカルネット上に配置した。ユーザが AR グラス上の web ブラウザで表示されている待機ボタンを押し、その後、スマホ上のボタンを押す事で、AR グラス上およびスマホ上の映像が自動で同時に再生される。今回の実装では、ランダムな遅延がスマホの映像と AR グラスの映像間で発生した。同じ映像でも、再生するごとに、発生する遅延の時間は異なった。ユーザスタディでは、特に遅延が発生しやすい映像を視聴してもらう際は、被験者自身に、何度か再生

\*4 <https://www.pexels.com/> accessed in Oct. 2022.

\*5 <https://www.istockphoto.com/> accessed in Oct. 2022.



ボタンを繰り返し押ししてもらい、遅延による違和感が結果に与える影響を少なくするよう努めた。

## 5. ユーザスタディ

ユーザスタディを通して二つの調査を行った。調査1では、提案手法を用いて映像を見る時と、スマホのみで映像を見る時と、ARグラスのみで映像を見る時の3種類の映像視聴方法で、ユーザの使用感に違いが出るのかを調べる。調査2では、提案手法を用いた映像視聴方法が映像の種類によって、視聴体験に違いを生むのかについて調査する。

### 5.1 手順

実験には、20代の被験者8名(男性6名、女性2名)が協力した。眼鏡をかけていた人は2名、コンタクトを使用していた人は2名であった。ARグラスを日常的に使用している被験者はおらず、何度か使用したことがある人は3名だった。実験時間は40分から60分であり、被験者にはアマゾンギフト券1000円分を提供した。

被験者には蛍光灯がある室内で、椅子に座りスマホを持った状態で、被験者が負荷を感じない姿勢をとってもらい、10種類の映像を見てもらった。

被験者には、まずスマホとARグラスの位置合わせを行ってもらった。ARグラスが表示する仮想ディスプレイは、上の側面をつまむと移動することができ、ディスプレイの端をつまむことでサイズを調整できる。スマホを手に保持した状態でARグラスを着用してもらい、ユーザの手が届く範囲かつスマホと同じ奥行き位置に、ARグラスで表示される映像を移動するよう指示した。このとき、ARグラスで表示される映像がスマホの映像と違和感なく重ねられるよう、AR上のディスプレイのサイズを調整するよう伝えている。

今回の実験では、比較する3種類の映像視聴方法のうち、スマホ単体での映像視聴では、ARグラス越しに映像視聴を行わせており、かつ表示する映像にはARグラスで表示する領域を含む。つまり、スマホ単体の場合、映像は縮小されて全体が表示される。3種類の映像視聴方法および10種類の映像を、被験者ごとに、それぞれランダムな順番で見てもらい、記述式とマーク式のアンケートに答えてもらった。その後、被験者にインタビューを行った。

### 5.2 調査1

調査1では、3種類の映像視聴方法のうち、また使用したいと感じた映像視聴方法はどれなのか、また使用感にどのような差があるのかを調査した。3種類の視聴方法で特に差異を感じなかった場合は、ニュートラルという項目を選んでもらった。

図4より、花火(図3(9))を除いて、多くの映像で、スマホが選ばれた。花火(図3(9))では、提案手法を選んだ

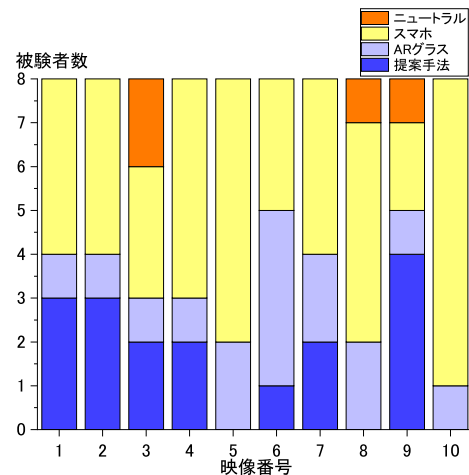


図4: どの視聴方法を再度使用したいと感じたか

被験者が4名であり、バイク(図3(1))、花(図3(2))でも3名の被験者が提案手法を選んだ。一方で、都市(図3(5))と相撲(図3(10))とサッカー(図3(8))では提案手法は選ばれなかった。

提案手法が選ばれた映像の特徴として、花火の映像は背景が黒く、バイクの映像はスマホ内部で映像が完結しており、花の映像は映像の速度が遅い特徴がある。一方で提案手法が選ばれなかった映像の特徴として、映像の中のメインコンテンツがスマホの画面内に収まりきっていないことが挙げられた。例えば、都市の映像(図3(5))は、ARグラス側にも、映像のメインのコンテンツが表示されており、相撲の映像は都市の映像と同様にARグラス側にも映像のメインコンテンツが表示され、さらに映像の速度が速いという特徴がある。上記の結果より、映像の特徴によって、提案手法への投票数に差があり、提案手法を用いるのが有益な映像とそうでない映像があるのが分かる。

次に、提案手法が他の視聴方法と比べ、使用感にどのような違いがあるのかについての結果である。被験者には、10種類の映像の視聴後に、3種類の視聴方法のうち、目の疲労度が最も高かったのは?、最も酔いやすいと感じたものは?、腕の疲れが最も大きかったのは?、映像に最も没入できたと感じたのは?、最も面白いと感じたのは?に対する回答(図5)を、3つの視聴方法からそれぞれ選んでもらった。先程と同様に、特に差異を感じなかった場合は、ニュートラルという項目を選んでもらった。

図5より、目の疲労度では4名が提案手法を選び、酔いやすさでは2名が提案手法を選び6名がニュートラルを選んだ。映像への没入のしやすさは、提案手法とARグラス単体が同数であった。どの方法で見るのが最も面白かったかに関しては、8名の被験者のうち、6名が提案手法を選び、2名がARグラス単体で見する方法を選んだ。ユーザが提案手法を選んだ理由として、「境目が綺麗にマッチしていて、技術として面白いと感じたため」「スマホの周りに視

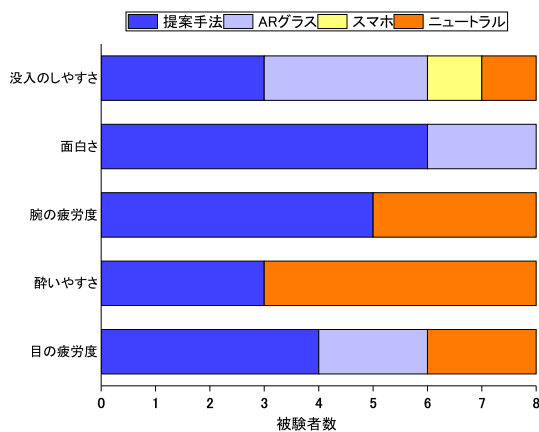


図 5: 3 種類の視聴方法の使用感の違い

野が広がる感じがしたから」「画面外に映像が広がるような見え方は、今まで経験したことがなかったので面白かった」という意見があった。このことから、提案手法が、今まで感じたことのない映像体験をユーザに提供することが分かった。酔いやすさ、目の疲労度は提案手法が多く選ばれたことから、ユーザに視覚的な負荷を与える可能性が他の手法より高いことが分かる。

調査 1 のアンケート結果より、映像の種類によって提案手法が好まれる映像と、そうでない映像があった。また、提案手法が、他の視聴方法に比べ、目の疲労度や酔いやすさなどが高いものの、最も面白く、ユーザに今まで体感したことのない視聴体験をもたらすことが示された。

### 5.3 調査 2

調査 2 では、提案手法での視聴体験が映像の種類によって異なるのか調査を行った。視聴体験を決める視覚効果に大きく影響する要素として、臨場感と見やすさを選んだ。被験者には、調査 1 のアンケート回答後に、10 種類の臨場感および見やすさについての順位をつけてもらい、なぜその順位にしたのかを、記述式のアンケートで答えてもらった。

映像ごとの、臨場感に関する得票数を図 6 で表している。図 6 より、花火 (図 3(9)) は、1 位に 2 票、2 位、3 位に 1 票ずつ、バイク (図 3(1)) は、1 位に 2 票、2 位に 2 票、花 (図 3(2)) は、1 位に 2 票、2 位に 1 票入っていた。被験者からは、「上位の映像は、スマホと AR グラスの動画の直線や曲線がぴったり合っている気がした」「動きがスマホ内のみの方が臨場感があった」「背景が黒い動画はスマホと AR グラスが完全に一致する感じがし、違和感なく、スムーズに見える」という回答が得られた。花火のように、背景が黒く抽象度の高い映像や、バイクの映像のように、スマホ内で動きが完結していたり、花の映像のように映像スピードの遅いものが上位に選ばれることが分かった。下

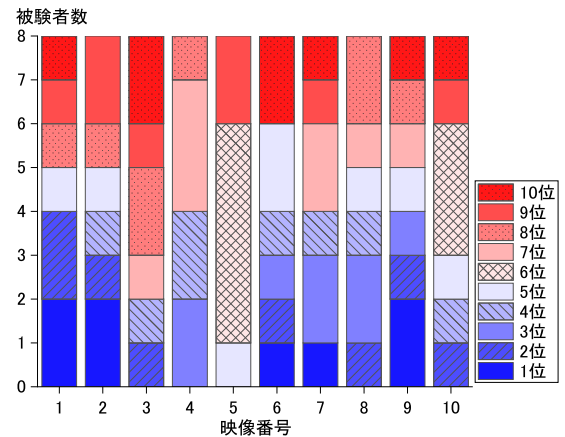


図 6: 映像ごとの臨場感についての順位

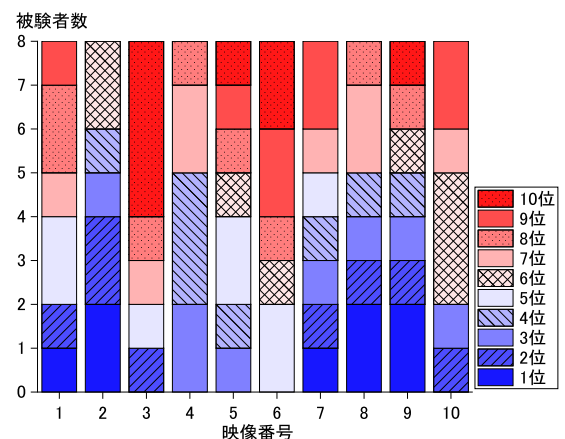


図 7: 映像ごとの見やすさについての順位

位に多く票が入った映像として、相撲 (図 3(10)) は、6 位に 3 票、9 位に 1 票、10 位に 1 票、都市 (図 3(5)) は、6 位に 5 票、9 位に 2 票、観客 (図 3(3)) は、8 位に 2 票、9 位に 1 票、10 位に 2 票であった。下位に票が多く入った映像は、映像の動きがスマホ内で完結していない映像が多く、提案手法を用いた映像視聴においては、スマホの内部で完結しているものが臨場感に作用することが分かった。

次に映像の見やすさについてである。映像の見やすさの得票数は、図 7 より、花 (図 3(2)) は、1 位に 2 票、2 位に 2 票、3 位に 1 票、サッカー (図 3(8)) と花火 (図 3(9)) は、1 位に 2 票、2 位に 1 票、3 位に 1 票入っていた。上位に票が多く入った映像は、調査 2 の臨場感の調査同様、サッカーや花の映像のように、スマホ内部で動きが完結している映像や、背景が黒く抽象度の高い花火の映像などが選ばれた。被験者に対して行ったアンケート結果でも、「テンポがゆっくりしていた」「視点が固定されていると、脳内で映像の補完が出来るため」といった回答が得られた。下位に多く票が入った映像として、都市 (図 3(5)) は、8 位に 1 票、9 位に 1 票、10 位に 1 票、観客 (図 3(3)) は、8 位に

1票, 10位に4票, ジェットコースター(図3(6))は, 8位に1票, 9位に2票, 10位に2票であった。被験者からは, 「観客の映像(図3(3))や都市の早送り映像(図3(5))は小さい点が動くので詳細が見えるスマホの方がいいと思う」といったフィードバックが得られた。下位の映像には, ジェットコースターのように映像の動きが速いものであったり, 都市や観客の映像のように, スマホ内部で動きが完結しない映像が選ばれた。被験者に対して行ったアンケートでは, 「下位の映像は遅延による違和感を感じた」「スマホとARグラスの画質の差があったことにより映像の境界を感じてしまった」といったフィードバックがあった。

調査2より, 提案手法を用いた際, 花(図3(2))の映像のように動きが遅いものや, バイク(図3(1))の映像のようにスマホ内で映像が完結しているものが, 臨場感および見やすさについて, 多く票を獲得していた。一方で, ジェットコースター(図3(6))のような映像の動きが速いものは, 他の映像に比べ見やすさが低く, 観客(図3(3))のようなスマホ内で完結しない映像については臨場感および見やすさがともに低かった。

また, 提案手法に対し, 被験者からは, 「慣れるまで位置を合わせたりするのが難しいが, 慣れたら面白いと感じた」「実際に使うと手と目が疲れそう」「スマホとARグラスの画質差や映像間の遅延が少し気になった」「スマホは画面の解像度など見やすかったが, ARグラスだと角度によって画面が見切れることがあった」といったフィードバックがあった。

## 6. 議論

実験で得られたユーザからのフィードバックから, 提案手法を実現するために必要な課題も判明した。具体的には, 二つのデバイスの画質の差異による映像間の境界の存在, スピードが速い映像での映像の遅延などによるスマホとARグラスの映像間の違和感などである。

### 6.1 ARグラスとスマホの映像間の境界の存在

今回の調査より, 花火(図3(9))などの背景が黒いコンテンツは, 境界の存在が目立たないことが分かった。このようにコンテンツによっては, 境界の問題を軽減できる場合がある。一方で, 二つのデバイスの画質や明るさの違い, 映像間のレイテンシー, 手や頭の揺れなどから, スマホの映像とARグラスの映像の境界の存在をユーザに認知させてしまうことがあった。

画質(色味の違い)による違和感を軽減させる方法として, スマホとARグラスの, それぞれのカラーフィルタを調整するソフトウェアを作成し, ユーザ目線で視聴した際の映像の色味を, 統一させることが挙げられる。明るさの違いによる違和感をなくしていくために, ARグラスとスマホの輝度を, ユーザが細かく操作し統一できるようにす

る必要がある。

今回の実装では, 境界を曖昧にするために, ARグラスの映像に黒色のグラデーションをかけている。このグラデーションの幅や強さなどを変更することで, 境界の存在をより目立たなくすることもできるかもしれない。

映像間のレイテンシーは, デバイスごと(スマホとARグラス)の, 再生開始のタイミングの違いにより発生していると考えられる。この問題は, 映像の始めのフレームをレンダリングするまで待機させることで, 改善される可能性があるため, 実際に上記の実装を行い, 調査していく必要がある。

今回の調査ではトラッキングの実装を行っていない。そのため, 手の揺れなどによって, 被験者にARグラスとスマホの映像間の境界の差を感じやすささせてしまった可能性がある。今後は, トラッキングを行った上で上記の結果に変化があるのか, 調べていく必要がある。

### 6.2 入力手法の改善

被験者の多くは, ARグラスのプッシュ操作などに苦戦していた。ARグラスの操作は, ユーザの慣れを必要とするものであり, 慣れていないユーザには負担を強いることとなった。関連研究で指摘されているように, スマホをARグラスのコントローラとして使用することで, 操作の負担を軽減することができる。

### 6.3 トラッキング

スマホを保持していることで生じる, 手の僅かな揺れによる映像間の違和感は, ユーザの充実した映像体験を損なう可能性がある。トラッキングを行うことで, 手の大きな動きに関しての違和感は軽減されるものの, 僅かな揺れに対してまで, トラッキングを行うと, ユーザの酔いを誘発してしまう可能性がある。また, ユーザの酔いやすさは映像のスピードによって変化する。例えば映像のスピードが遅い映像は, 精度の高いトラッキングを行ったとしても, スピードの速い映像に比べ, ユーザの酔いは抑えられる。さらに, 暗い映像ではそもそも境界を認識することが難しく, 精度の高いトラッキングを行う必要性が低い。そのため, 映像の種類によって, どの程度の精度でのトラッキングを行うべきか, 変える必要があると考える。また, トラッキングの実装で生じるデバイスへの負荷も考慮しなくてはならない。デバイスへの負荷の増加は, より映像のレイテンシーなどの問題を引き起こす可能性がある。

今回の実験の中では, 被験者自身に映像の位置合わせを行ってもらったが, 一部のユーザは, 位置合わせ自体が面白く感じていた。位置合わせをすることで, 映像を視聴する際の集中力が増したというフィードバックもあった。これらのフィードバックは予想外であったが, 位置合わせを行わせることで別の新たな視聴体験が提供できる可能性

を示唆している。ユーザ自身に位置合わせを行ってもらう場合と、トラッキングを用いて位置合わせを自動で行う場合の二つの場合を、ユーザ自身が選択できる仕組みが必要であると考えられる。

## 7. まとめ

本論文では、AR グラスとスマホを組み合わせた映像拡張手法を提案した。8名によるユーザスタディを行い、提案手法を用いる際、映像の種類によって、ユーザの映像体験に差異があることが分かった。また、アンケート結果より、提案手法での視聴体験が、AR グラスやスマホ単体の映像視聴よりも、新鮮で面白い映像体験をユーザにもたらしることが分かった。一方、二つのデバイスから出力される映像の境界の差がユーザに違和感を与えてしまうなど、解決しなければならない課題も明らかとなった。

## 参考文献

- [1] Adams, R. W.: *Peripheral vision and visual attention*, Iowa State University (1971).
- [2] Benko, H., Ofek, E., Zheng, F. and Wilson, A. D.: FoveAR: Combining an Optically See-Through Near-Eye Display with Projector-Based Spatial Augmented Reality, *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology*, UIST '15, New York, NY, USA, Association for Computing Machinery, p. 129–135 (online), DOI: 10.1145/2807442.2807493 (2015).
- [3] Cosco, F. I., Garre, C., Bruno, F., Muzzupappa, M. and Otaduy, M. A.: Augmented touch without visual obstruction, *2009 8th IEEE International Symposium on Mixed and Augmented Reality*, pp. 99–102 (online), DOI: 10.1109/ISMAR.2009.5336492 (2009).
- [4] Grubert, J., Heinisch, M., Quigley, A. and Schmalstieg, D.: MultiFi: Multi Fidelity Interaction with Displays On and Around the Body, *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, New York, NY, USA, Association for Computing Machinery, p. 3933–3942 (online), DOI: 10.1145/2702123.2702331 (2015).
- [5] Jones, B. R., Benko, H., Ofek, E. and Wilson, A. D.: IllumiRoom: Peripheral Projected Illusions for Interactive Experiences, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, New York, NY, USA, Association for Computing Machinery, p. 869–878 (online), DOI: 10.1145/2470654.2466112 (2013).
- [6] Kameda, Y., Takemasa, T. and Ohta, Y.: Outdoor see-through vision utilizing surveillance cameras, *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 151–160 (online), DOI: 10.1109/ISMAR.2004.45 (2004).
- [7] Khan, A., Fitzmaurice, G., Almeida, D., Burtnyk, N. and Kurtenbach, G.: A Remote Control Interface for Large Displays, *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology*, UIST '04, New York, NY, USA, Association for Computing Machinery, p. 127–136 (online), DOI: 10.1145/1029632.1029655 (2004).
- [8] Kimura, N. and Rekimoto, J.: ExtVision: Augmentation of Visual Experiences with Generation of Context Images for a Peripheral Vision Using Deep Neural Network, *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, New York, NY, USA, Association for Computing Machinery, p. 1–10 (online), DOI: 10.1145/3173574.3174001 (2018).
- [9] Lee, C.-J. and Chu, H.-K.: Dual-MR: Interaction with Mixed Reality Using Smartphones, *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, VRST '18, New York, NY, USA, Association for Computing Machinery, (online), DOI: 10.1145/3281505.3281618 (2018).
- [10] McKee, S. P. and Nakayama, K.: The detection of motion in the peripheral visual field, *Vision research*, Vol. 24, No. 1, pp. 25–32 (1984).
- [11] Normand, E. and McGuffin, M. J.: 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)*, pp. 123–133 (online), DOI: 10.1109/ISMAR.2018.00043 (2018).
- [12] Patney, A., Salvi, M., Kim, J., Kaplanyan, A., Wyman, C., Benty, N., Luebke, D. and Lefohn, A.: Towards Foveated Rendering for Gaze-Tracked Virtual Reality, *ACM Trans. Graph.*, Vol. 35, No. 6 (online), DOI: 10.1145/2980179.2980246 (2016).
- [13] Purves, D., Augustine, G., Fitzpatrick, D., Katz, L., LaMantia, A., McNamara, J. and Williams, S.: *Neuroscience 2nd edition*. sunderland (ma) sinauer associates, *Types of Eye Movements and Their Functions* (2001).
- [14] Robles-De-La-Torre, G.: The importance of the sense of touch in virtual and real environments, *Ieee Multimedia*, Vol. 13, No. 3, pp. 24–30 (2006).
- [15] Shirazi, M. A., Cho, H. and Woo, W.: Augmentation of Visual Experiences using Deep Learning for Peripheral Vision based Extended Display Application, *2019 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 1–2 (online), DOI: 10.1109/ICCE.2019.8662089 (2019).
- [16] Unlu, A. E. and Xiao, R.: PAIR: Phone as an Augmented Immersive Reality Controller, *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*, VRST '21, New York, NY, USA, Association for Computing Machinery, (online), DOI: 10.1145/3489849.3489878 (2021).
- [17] Villanueva, A., Zhu, Z., Liu, Z., Peppler, K., Redick, T. and Ramani, K.: Meta-AR-App: An Authoring Platform for Collaborative Augmented Reality in STEM Classrooms, *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, New York, NY, USA, Association for Computing Machinery, p. 1–14 (online), DOI: 10.1145/3313831.3376146 (2020).
- [18] Xiao, R. and Benko, H.: Augmenting the Field-of-View of Head-Mounted Displays with Sparse Peripheral Displays, *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, New York, NY, USA, Association for Computing Machinery, p. 1221–1232 (online), DOI: 10.1145/2858036.2858212 (2016).
- [19] Yamada, W. and Manabe, H.: Expanding the Field-of-View of Head-Mounted Displays with Peripheral Blurred Images, *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16 Adjunct, New York, NY, USA, ACM, pp. 141–142 (online), DOI: 10.1145/2984751.2985735 (2016).