

360度インターネット生放送における 視聴者 POV 提示手法に関する比較検討

高田真也¹ 齊藤義仰¹

概要：YouTube を始めとするインターネット生放送サービスで、全方位カメラに対応した 360 度インターネット生放送サービスが開始された。しかし、360 度インターネット生放送では従来の放送と異なり、カメラのレンズ方向から視聴者の視聴方向やその範囲を把握することができなくなった。コミュニケーション相手の視線が向いている方向は、相手の興味や関心を示す重要な情報である。相手の視線を把握できない場合、会話の中心が把握できなくなり、円滑なコミュニケーションが行えない可能性がある。この問題は 360 度インターネット生放送においても同様であると考え、視聴者の視聴方向(POV)放送者に提示する機能についての検討を行う。POV の提示の初期実装により判明した、可搬式全方位カメラを用いた 360 度インターネット生放送における制限項目を基に、3 種類の POV 提示機能を実装した。実装した POV 提示機能は、複数の振動モータを内側に配置したベルト型デバイス、任意の LED を点灯できる LED テープを円筒に巻き付けた LED 型デバイス、2 つのサーボモータを組み合わせて任意方向を向くことが出来るようにしたロボット型デバイスの 3 種類である。それぞれの POV 提示機能を 15 分ずつ用いる放送実験を実施し、放送者に対して聞き取り調査を行うことで、POV 提示手法の比較検討を行った。その結果、視覚に頼る POV の提示手法は放送者に好まれにくいことが分かった。また、放送者がロボット型デバイスに対して親しみを感じることが確認できた。

キーワード：インターネット生放送, 全方位カメラ, POV

1. はじめに

YouTube に代表されるインターネット生放送サービスは多くのユーザに利用されており、コメントを中心としたリアルタイムコミュニケーションを気軽に楽しむことが出来る。YouTube では 2015 年から全方位カメラに対応した 360 度インターネット生放送サービスの提供も開始しており、一般ユーザでも全方位カメラを用いた放送ができるようになった。360 度インターネット生放送とは、全方位カメラによって撮影される全天球映像を配信する放送形式で、視聴者は自身の興味に合わせて POV を 360 度全方向に変更することができる。POV とは Point Of View の略称で、視聴者が視聴している方向を示す。

しかし、360 度インターネット生放送には、放送者が視聴者の POV を確認する方法が存在しないため、視聴者の興味関心が把握できないという問題がある。従来のインターネット生放送では、単一のレンズを持った Web カメラが用いられており、Web カメラのレンズが限定的な矩形の撮影範囲を示していた。そのため、放送者は感覚的に視聴者の視聴している映像や撮影範囲を把握することができた。それに対し、360 度インターネット生放送で用いる全方位カメラは魚眼レンズや複数のレンズを搭載しているため、レンズ方向から視聴者の視聴している範囲がどこなのかを放送者が把握することができなくなった。これまでも、遠隔コミュニケーションツールにおける対話者の注視情報や視線情報が担う役割については多く論じられている[1][2]。その中で、対話者の注視情報は相手の興味や関心、話題の中心を示すということが明らかにされている。つまり、360 度インターネット生放送において、視聴者の POV は視聴者が全方位映像の中で何を見ているのかを示す情報で

あると同時に、視聴者の興味や関心を示す情報でもある。

我々はこれまでこの問題に対して、放送者への POV 提示や POV の分析アルゴリズムの検討に取り組んできた[3][4]。その結果、POV の提示によってコメントの伝達が正確に行えること、POV の分析によって 55.1%の精度で放送者に有用な POV を抽出することが可能となった。また、この分析手法を用いて抽出した POV を放送者に提示した際の効果について検証を実施した[5]。POV の提示により、放送者は視聴者の興味を容易に把握することが出来、コミュニケーションに消極的なユーザとの交流に役立つ手掛かりとして利用できる効果が確認できた。また、話題性に乏しいような POV であっても周囲の状況を把握するために利用できるという知見も確認できた。一方で POV の提示方法にはいくつかの課題が残っていたため、本研究では、ユースケースに適した POV 提示方法の検討を行った。

現在我々が想定しているユースケースは、放送者が観光地に赴いて散策しながら放送を行うというものである。放送者は全方位カメラをリュックサックに固定し、放送用のノートパソコンを手を持って放送する。放送者が移動しながら放送を行う場合、手元を長時間注視することは安全面の問題がある。また、ノートパソコンで手が塞がるため、放送者の身体的な制限も発生する。これらの制限に対して適切な POV 提示手法を検討するため、複数の POV 提示手法を比較検討する実験を実施する。

2. ユースケース

360 度インターネット生放送は従来の放送形態と比較して撮影範囲が広いため、より多くの情報を視聴者に提供できるという特徴がある。そのため、空間全体の情報を提供する目的で利用されることが想定される。実際に、360 度

¹ 岩手県立大学大学院 ソフトウェア情報学研究科
Iwate Prefectural University,
Graduate School of Software and Information Science

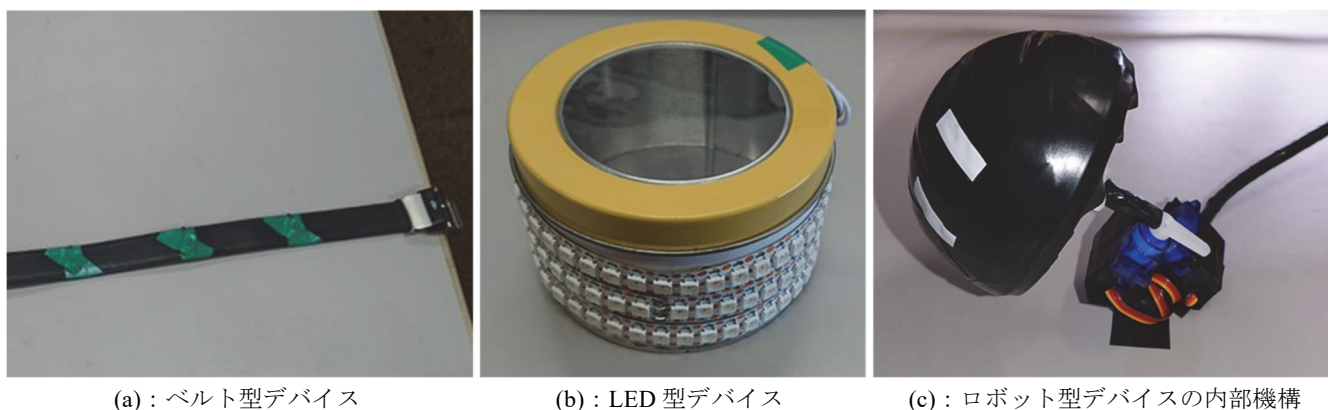


図1：各POV提示手法の外観

インターネット生放送は観光地の紹介や不動産物件の紹介などで利用されている[6][7]。一般的なユーザが行う具体的なユースケースとしては、放送者が観光地に赴いて散策しながら観光地の紹介を行う放送が考えられる。このような場合、放送者は全方位カメラをマウンタなどで身体に固定するか、スタビライザーを用いて手に持った状態での放送が考えられる。今回は、観光地において放送者1人が全方位カメラをリュックサックに固定し、放送用のノートパソコンを手に持って散策しながら放送を行うというユースケースを想定することとした。

本ユースケースにおいて、視聴者のモチベーションは観光地の情報をリアルタイムなコミュニケーションと共に得られることで、放送者に対して自身のニーズを発信することが期待される。放送者のモチベーションは観光地の散策をより良い体験にすることである。具体的には、視聴者からのコメントを確認しながら散策することで、同行者のような存在を疑似的に獲得することが出来る。視聴者からのコメントにより、放送者は自身だけでは気づくことが出来なかった周囲の事柄に気づき、経験することが出来るようになる。このユースケースにおいて、放送者は視聴者に対して見たいもののリクエストを募集することや、問いかけなどの双方向コミュニケーションを行うことが期待される。

放送者が移動しながら放送を行う際にPOVが把握できないと、放送者は視聴者が興味を持った対象に気づかず通り過ぎてしまう可能性がある。また、視聴者がコメントによりリクエストを行ったとしても、放送者はコメントの対象がどの方向に存在するのかをコメントだけから推測する必要がある。そこで、視聴者のPOVを放送者に対して適切に提示する手法を検討する必要がある。

3. 比較手法

適切なPOVの提示手法を検討するために、いくつかのPOV提示手法を作成し、比較することにした。我々が事前に行ったPOV提示効果の検討における知見を基にPOV提示手法の候補を決定することとした。放送者が移動しながら放送を行う際には2つの制限が発生する。まず1つ目は手元への注視行為がごく短時間に制限されることである。近年スマートフォンの普及に伴い、歩きスマホや自動車運転中のながらスマホなどが問題視されている。移動を伴う行動中に手元の情報に注意が集中すると事故につながる危険性がある。次に2つ目は身体的な制限である。放送者は背

面に背負ったリュックに全方位カメラを固定しており、手には放送用のノートパソコンを所持している。大きさにもよるが、片手または両手がノートパソコンによって占有されており、重量がある装置の使用や、複雑な操作を必要とする装置の使用は困難である。また、これらの制限によってPOVの提示によるコミュニケーション支援の効果が阻害される可能性を考慮し、コミュニケーション支援を重視した提示手法についても検討する必要がある。

1つ目の制限に対し、安全面に考慮した手法としてベルト型のデバイスを検討する。また、2つ目の制限に対し、軽量かつクリップによってノートパソコンに固定することが可能なLED型デバイスを検討する。また、これらに加えてコミュニケーションを重視した手法としてロボットエージェントを参考にしたロボット型デバイスを検討する。

ベルト型デバイスは、スマートフォンなどのバイブレーション機能に利用される振動モータをベルトの内側に複数配したものである。利点として、視覚に頼らずにPOV提示を行うため、デバイスへの注視が不要で、放送者は周囲に注意を払いながらPOVを確認することが出来る。一方で、水平方向のPOVについては提示できるが、垂直方向の情報については提示が困難であることが予想される。LED型デバイスは、個別に発行させることが可能なLEDテープを筒にらせん状に巻き付けたものである。筒の底部にはノートパソコンに固定するためのクリップがある。利点として、ベルト型では困難な垂直方向のPOV提示が可能になる点が挙げられる。一方で注視が必要になるため、POVの確認を行うためには一度立ち止まる必要がある。ロボット型デバイスは、サーボモータを2つ使用することで水平方向および垂直方向に稼働するドーム状の頭部を持った手乗りサイズのロボットである。LED型デバイス同様、ロボット型デバイスの底部には固定用のクリップが用意した。図1に各デバイスの外観を示す。aがベルト型デバイス、bがLED型デバイス、cがロボット型デバイスの内部機構である。今回の比較において、これらの比較はPOV提示手法に必要な要件を洗い出すための初期段階としての位置づけである。

4. 関連研究

これまでに遠隔コミュニケーションなどの環境において方向を提示する手法はいくつもの検討が行われている。ここでは、3つの関連研究について述べる。1つ目はナビゲ

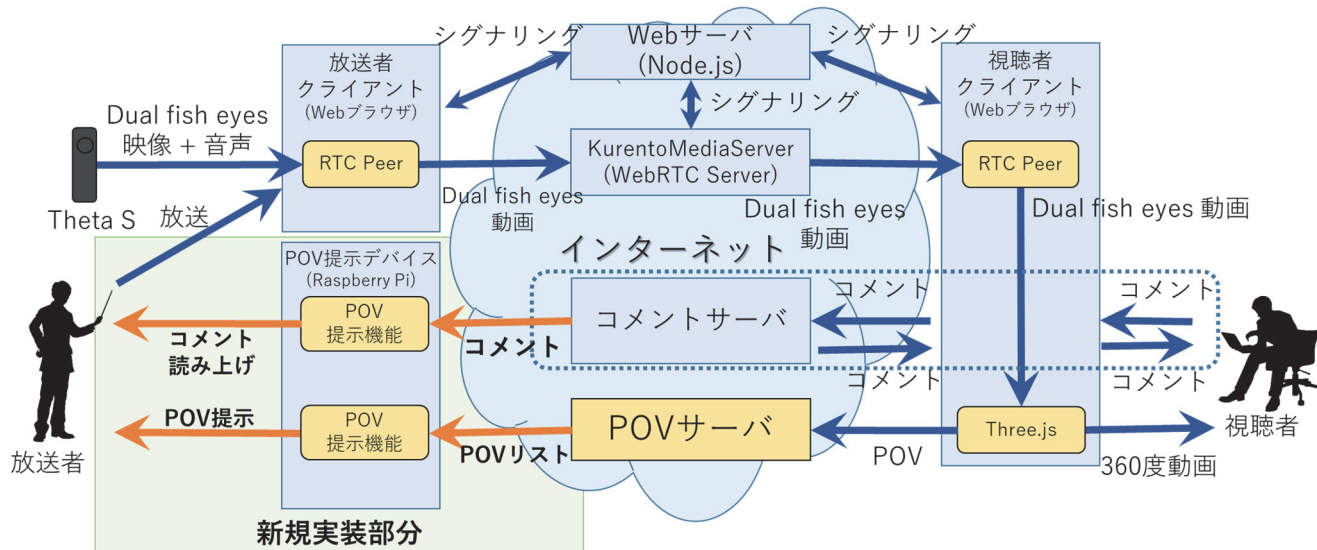


図2：システム構成図

ーションの分野において視覚障がい者向けの方向提示手法について検討を行った研究[8]である。2つ目は遠隔コミュニケーションにおけるLEDを用いた注視方向の提示手法に関する研究[9]である。3つ目は遠隔コミュニケーションにおけるロボットを用いた物理的な方向提示手法を検討したGestureMan[10]である。

Dragan Ahmetovicら[8]の研究では、視覚障がい者に対して水平方向の回転角を提示する手法として、音声を用いた4つの方向提示手法を検討している。1つ目は指定の方向を向いた際に高音が鳴動する方式で、2つ目は指定の方向に近づくと間隔が短くなる連続音が鳴動する方式である。3つ目は1つ目と2つ目の手法の組み合わせによるものであり、4つ目は方角に音階をスケールし、指定した方向が左右どちらの方向にあるのかを音階の提示によって表現する手法である。これらの方向提示手法は高い精度での方向提示を可能にしており、実用的な方向提示の可能性を示唆している。一方で、当該手法は水平方向にのみ対応しており、他の音声と同時に利用されることは想定できていない。音声ガイドなどと同時に提示するには方向の提示を他の情報に置き換えて提示する必要がある。本研究では方向の提示に用いる情報を音声から振動に置き換え、放送者を中心とした水平方向のリング状配置で提示を行う。

Omni-gazeは遠隔コミュニケーションにおいて、コミュニケーション参加者の視線方向をLEDによって提示する手法である。遠隔コミュニケーションには全方位カメラが用いられており、全方位カメラを覆うようにLEDパネルが配置されている。コミュニケーション参加者の視線方向を分析し、対応する方向のLEDを点灯させることで、立体的な視線方向の提示を行っている。立体的な配置のLEDを用いることによって現実空間に対応付けられた方向の提示が可能になっている。本研究ではLEDテープを円筒に巻き付けて用いることで同様の効果を期待する。

GestureManは遠隔コミュニケーションにおいて、コミュニケーション相手の周囲に存在する物への指示行為を遠隔操作者がロボットにより行うというものである。遠隔コミュニケーションにおける指示行為を再現するためにロボットのジェスチャを利用しているが、単純な腕部の稼働だけ

でなく、視線情報や身体各部分の方向などの身体性が情報伝達を補助していることを明らかにしている。視線情報やジェスチャなどの身体性は、周囲のコミュニケーション参加者の注意や関心を惹くため、参加者間でのコミュニケーションを誘発させる可能性がある。本研究では2つのサーボモータを組合せることで自由に視線方向を変化させられるロボット型デバイスを用いる。

5. 実装

本研究で用いたインターネット生放送システムは我々がこれまでにPOVの分析を行うため作成した放送システムに3種類のPOV提示機能を追加実装したものである。ここでは、既存システム部分と追加実装部分について述べる。

5.1 既存システム部分

既存のシステムはWebRTCおよびThree.jsを用いて実装したものである。HTML5に対応したWebRTCでは、映像や音声の配信やファイル共有など、リアルタイムコミュニケーションに必要な様々な機能が簡単に利用できる。映像の配信はWebRTCに対応したKurentoMediaServerを利用した。本研究ではRICOH社が販売する全方位カメラTHETA Sを用いて放送を実施しており、放送用のノートパソコンからTHETA Sの映像をKurentoMediaServerへと送信している。360度動画への加工にはThree.jsを使用しており、WebGL上の仮想空間上に球体オブジェクトを設置してその球表面の内側に映像を描画することで実現している。再生する映像のソースをWebRTCサーバから取得することでインターネット生放送として利用できる。図2にシステムの構成を示す。

本システムでは、WebGLの3D空間上にある球体オブジェクトの中心にカメラ位置が設定されている。球体の内側表面には360度動画の映像がマッピングされているため、球体の中心からカメラを通して映像を見ることで360度動画形式の映像になる。視聴者のPOVはカメラがどの方向を向いているのかを示しており、球表面の極座標系(ϕ , θ , r)によって管理されている。POVは100ミリ秒ごとにPOVサーバへと送信されており、POVリストのリクエストが発生すると、各視聴者のPOVを60秒分リスト化し送信する。

5.2 追加実装部分

追加実装部分では、音声ガイド機能と3種類のPOV提示機能を制御するための機能を実装した。音声ガイド機能と制御機能はRaspberry Pi環境を想定したプログラムとして実装した。

音声ガイド機能は、視聴者から投稿されたコメントを放送者に音声読み上げにより提示する機能である。従来のPOV提示機能は放送用パソコンの画面上で行っていたため、コメントの確認は同じ画面上に表示されるコメント欄を用いていた。しかし、今回比較するPOV提示手法では画面への注視を行わないため、コメントを読み上げることで放送者に視聴者の反応を提示することとした。音声の読み上げには骨伝導式のヘッドセットを利用しており、イヤホンなどの装着による周囲への注意が妨げられるということとを避けた。

各POV提示機能に用いる振動モータ、LEDテープ、サーボモータの制御はRaspberry PiのGPIOと接続し、制御することとした。使用した機材はRaspberry Pi 2 Model Bで、プログラムの作成にはPi4Jを用いた。1秒ごとにPOVサーバに対してPOVリストのリクエストを行い、POVの分析を実施する。この時、分析手法はこれまでに作成したPOV分析アルゴリズム[4]を用いて放送者に提示する必要があるPOV変化を抽出する。その後、提示の必要があるPOV変化がある場合はPOVのリクエストを5秒間停止し、各提示手法によりPOVの方向を放送者に提示する。また、POVの提示に気づかない可能性を考慮し、POVの提示を行う際には動作前にピープ音を再生することとした。

6. 比較実験

6.1 実験概要

実装した3種類のPOV提示機能を比較するため、放送実験を実施した。実施日は2020年12月20日で、実験協力者は放送者1名、視聴者5名の計6名である。放送者は岩手県盛岡市の高松の池周辺を散策するという内容で15分程度の放送を3セット実施する。高松の池は白鳥の飛来地としても有名な観光地であり、放送者は周囲の様子を基に視聴者との雑談を行いながら池周辺を散策する。各放送では3種類のPOV提示機能のうち1種類を用いる。視聴者は放送を視聴し、放送者の周囲にある風景や被写体に対してコメントを投稿する。放送者への注意事項として、放送用PCやPOV提示デバイスを注視する際には周囲に注意を払うように指示をした。また、視聴者からのコメントについては出来る限り応答を行うように指示をした。

3セットの放送終了後、放送者に対してPOV提示が発生した箇所についての聞き取り調査を実施する。聞き取り調査では、POV提示が発生したことに気づいたか、POVの方向が分かったか、そのPOVは放送に活用することが出来たかの3点について質問を行う。また、各POV提示手法について、方向の分かりやすさ、確認の手間がないか、どの程度参考にしたか、再び使いたいかの4点について5段階で評価を行ってもらい、そう感じた理由についても聞き取りを行った。聞き取り調査の最後には全体を通しての感想を聞いた。

表1：各POV提示に対する聞き取りの結果

使用した提示手法	POVの提示数	POV提示に気づいたか	POVの方向は分かったか	放送に利用できるか
ベルト型	24件	24件 (100%)	19件 (79.2%)	14件 (58.3%)
LED型	18件	17件 (94.4%)	8件 (44.4%)	9件 (50%)
ロボット型	21件	21件 (100%)	18件 (85.7%)	16件 (76.2%)

表2：POV提示手法に対する聞き取りの結果

使用した提示手法	方向の分かりやすさ	確認の手間がないか	どの程度参考にしたか	再び使いたいか
ベルト型	4	5	3	4
	振動を使っているため、確認の手間はないが、振動モータの間隔が広く、何を見ているのかわかりづらい時があった。他のものと比べて人目が気になりづらかった。			
LED型	2	3	3	2
	池の反射もあり、明るい環境では光っているLEDがどれなのかを判別することが難しかった。高さに関してはあまり変化が無かったため、確認できなくても良いと思う。			
ロボット型	4	4	3	4
	他のものより人目が気になるが、方向自体は分かりやすかった。コメントの読み上げがロボットの声のように思えて、親しみを感じた。			
全体	コメントを読み上げてくれるので、放送はしやすかった。用意されたPOVの提示手法はあまり負担に感じなかった。			

6.2 実験結果

放送実験を実施した結果、ベルト型デバイスを用いた放送では24回、LED型デバイスを用いた放送では18回、ロボット型デバイスを用いた放送では21回のPOV提示が発生した。表1に各POV提示への気づき、方向が分かるか、放送へ活用可能かの3点についての聞き取り結果を示す。POV提示に対する気づきについては、ほぼすべて気づいていることが分かった。POVの方向が分かるかについては、ベルト型、ロボット型では8割前後、LED型では4割程度の提示に対してその方向が理解できている。放送へ活用できるかについては、5割～7割程度のPOV提示が放送に活用可能であると判断された。表2に3種類のPOV提示手法についての聞き取り結果を示す。ベルト型とロボット型は4以上の評価が多く、放送者に使いやすいものとなったことが分かった。一方で、LED型は方向の分かりやすさについての評価が低く、他の2つの提示手法と比較して評価値が低かった。

7. 考察

今回の比較検証において、放送者が最も評価値が高かったのはベルト型デバイスだった。原因としては、放送後の聞き取り調査から、POVの提示時に注視が必要なく、放送を中断せずにPOVの確認を行うことが出来るからだと考えられる。次に評価値の高かったロボット型デバイスについては、放送者がデバイスに対して親しみを感じたという意見を得ることができた。この結果は、ロボット型デバイスの狙いであるコミュニケーションの活性化につながるものであり、再び使用したいかという質問結果からも、放送者がロボット型デバイスに良い印象を抱いていることが考

えられる。

今回のユースケースにおいて、放送者が360度インターネット生放送を利用するモチベーションは、視聴者を観光の疑似的な同行者に見立て、自身が気付かないような周囲の事柄への気づきを得ることである。そのため、放送者は視聴者とのコミュニケーションを中断せずに、積極的に行うことが望ましい。今回の比較検証から、POV提示手法を実装する際には、長時間の注視を伴う視覚情報に頼った提示手法は避ける必要があるということが分かった。また、親しみを持てるロボットエージェントのような提示手法が望ましいということも分かった。

8. おわりに

本研究では、視聴者のPOVを放送者に提示する際のユースケースに適した提示手法について検討を行った。3つのPOV提示手法を実装し、比較することによって、POV提示に必要な要件を検討した。その結果、確認の手間、親しみという2つの評価観点を得ることが出来た。一方で、比較実験は1回しか実施できていないため、複数回の実験を実施し、より多くの意見から分析を進める必要がある。

参考文献

- [1] Roel Vertegaal, "The GAZE groupware system: mediating joint attention in multiparty communication and collaboration", CHI '99 Proceedings of the SIGCHI conference on Human Factors in Computing Systems, pp.294-301(1999).
- [2] David M. Grayson, Andrew F. Monk, "Are you looking at me? Eye contact and desktop video conferencing", ACM Transactions on Computer-Human Interaction (TOCHI) Volume 10 Issue 3, September 2003, pp.221-243(2003).
- [3] 高田真也, 西岡大, 齊藤義仰: 全方位カメラを用いたインターネット生放送における視線ヒートマップの開発, 分散協調とモバイルシンポジウム 2017 論文集, pp.1081-1088(2017).
- [4] 高田真也, 西岡大, 齊藤義仰: 360度インターネット生放送における視聴者POVの分析, 分散協調とモバイルシンポジウム 2019 論文集, pp.889-894(2019).
- [5] 高田真也, 齊藤義仰: 360度インターネット生放送における視聴者POV分析結果の提示, 研究報告グループウェアとネットワークサービス(GN) 2020-GN-110-6, pp.1-6(2020).
- [6] "渚の駅" たてやまでVR体験 | 館山市役所, <https://www.city.tateyama.chiba.jp/min-ato/page100266.html>, (2020年12月参照).
- [7] 不動産VR | 手間なく・早く・売れやすく | ハイクオリティのバースでVRをご提供!, <https://fudosan-vr.jp/>, (2020年12月参照)
- [8] D. Ahmetovic, F. Avanzini, A. Baratè, C. Bernareggi and G. Galimberti, L. Ludovico, S. Mascetti, G. Presti : Sonification of Rotation Instructions to Support Navigation of People with Visual Impairment, 2019 IEEE International Conference on Pervasive Computing and Communications (PerCom 2019), pp.332-341(2019).
- [9] Keisuke Shiro, Atsushi Okada, Takashi Miyaki, Jun Rekimoto : OmniGaze: A Display-covered Omnidirectional Camera for Conveying Remote User's Presence, HAI '18 Proceedings of the 6th International Conference on Human-Agent Interaction, pp.176-183(2018).
- [10] Hideaki Kuzuoka, Shin'ya Oyama, Keiichi Yamazaki, Kenji Suzuki, Mamoru Mitsuishi : GestureMan: a mobile robot that embodies a remote instructor's actions. , Proceedings of the 2000 ACM conference on Computer supported cooperative work, pp.155-162(2000).