

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

高田真也¹ 西岡大¹ 齊藤義仰¹

Development of the Gazes Heat Map in Internet Broadcast using an Omnidirectional Camera

MASAYA TAKADA¹ DAI NISHIOKA¹ YOSHIA SAITO¹

1. はじめに

YouTube やニコニコ動画, U-Stream などに代表されるインターネット生放送サービスは, 近年コミュニケーションツールとして広く利用されるようになってきた. 平成 27 年度の通信利用動向調査[1]によると, インターネット利用者のうち, 半数以上がインターネット利用の目的に, 動画投稿・共有サイトの利用を挙げており, 多くのユーザが利用していると言える. また, これらの他にも Twitter や Facebook などの SNS でも動画のリアルタイム配信への対応が進んできており, 動画コンテンツは手軽に利用可能なメディアの 1 つとなってきた. これらのサービスでは, 放送者と視聴者がコメント等により, リアルタイムなコミュニケーションを取れることが特徴であり, 放送者はリアルタイムに視聴者のレスポンスを受け取ることができるため, 視聴者の反応や要望を放送内容にフィードバックすることが可能である.

動画配信共有サービスの大手である YouTube では, 近年普及しつつある全方位カメラを用いた動画への対応が進められている. 全方位カメラは複数の広角レンズにより周囲 360 度の映像を撮影することができるカメラのことで, 2015 年にオンデマンド配信, 2016 年にインターネット生放送サービスへの対応が開始された. 全方位カメラに対応した動画の再生方式は 360 度動画と呼ばれ, 動画プレイヤー上で操作することで視線を周囲 360 度全方向に向けることができる. 動画共有サービス最大手である YouTube の他にも, Twitter で利用することができる連携サービス「Periscope」や Facebook の Live360 などで 360 度動画のインターネット生放送サービスへの対応が進んでおり, 今後の爆発的な普及が期待される.

しかし, 従来の Web カメラやスマートフォンのカメラを使用したこれまでの動画配信の方法とは異なり, 全方位カメラを用いた放送では, 視聴者の視線を放送者が察知できないという問題が発生する. 図 1 に従来の Web カメラを用いたインターネット生放送と全方位カメラを用いたインターネット生放送の違いを示す. 従来の Web カメラを用い

た場合は単一のレンズが撮影方向を明示的に示しており, 視聴者は必然的にその撮影範囲内を視聴している. それに対し, 全方位カメラを用いた場合は, 魚眼レンズや複数のレンズが周囲 360 度の撮影を行っており, 撮影範囲が広すぎるため, 視聴者が撮影範囲の中でどこを見ているのかが分からない. 視聴者がどこを見ているのか, すなわち視聴者の興味がどちらに向いているのかを放送者が察知できない場合, コミュニケーションが円滑に行えない可能性がある. 本研究では, 全方位カメラを用いたインターネット生放送において, 放送者と視聴者の間で円滑なコミュニケーションが行えるように, 視聴者の視線を放送者に対して可視化する視線ヒートマップモデルを開発およびどうさの検証を実施した. 本稿では, 視線ヒートマップの実装と動作検証を行った結果, および動作検証結果から考察した改善点について報告する.



図 1: カメラによる視線認知の差異

2. コミュニケーションにおける視線の役割

これまでも, 行われた研究において, コミュニケーションにおけるノンバーバル情報の重要性に関して様々な提言がなされている[2][3][4]. 武川ら[5][6]によれば, コミュニケーションにおける視線はお互いの意図を伝える重要な役割を担っており, 何に対し興味, 関心を持っているのかを視線方向によって示しているとしている. つまりインタ

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

インターネット生放送において、放送者が視聴者の視線がどの方向に向いているのかを把握できないということは、視聴者の興味や関心、発言の意図などが感じ取れなくなる要因に成り得る。大野[7]の研究では、従来の遠隔ビデオコミュニケーションシステムに加え、視線分析を基にしたコミュニケーション相手の視線位置の可視化が試みられている。遠隔地間での協調作業において、対面状態と同等の作業効率を実現するためには、相手の関心対象をより明解に示すことが必要だと主張している。これらの研究から、コミュニケーションを行う上で、相手に関心を示しているものが何か、また発言における中心となる対象が何かを理解するために、相手の視線情報を把握することが有効であると言える。特に、相手が何に関心を示しているのかを示している。

つまり、全方位カメラを用いたインターネット生放送において、放送者が視聴者の視線を把握することができないということは、視聴者の意図を十分に理解できず、円滑なコミュニケーションが阻害される可能性が生じる。

3. 全方位カメラと360度動画

ここでは、本研究で対象とした全方位カメラおよび360度動画について述べる。

3.1 全方位カメラ

全方位カメラはカメラを中心とした周囲360度全方向の映像を撮影することができるカメラのことを指す。全方位カメラには、大きく分けて2つの撮影方式がある。1つ目は180度以上の広角を撮影できる魚眼レンズを用いた方式である。この方式では、1, 2個のレンズで360度動画の撮影が可能であるが、魚眼レンズを用いているため、レンズ中心から離れるほど歪みが発生してしまう。2つ目は複数のレンズを用いて、それぞれのレンズで撮影された映像を合成する方式である。Omniマウントと呼ばれる箱状のマウンターや専用の球体型マウンターに複数の小型カメラを搭載し、ソフトウェアで合成する方法と、複数のレンズを持った1つのカメラで撮影した映像を内部合成する方法がある。撮影された映像は付属しているソフトウェアなどを用いることで360度動画として視聴することが可能である。

現在、全方位カメラは様々な製品が登場している。代表的なものとしては、KodakのSP360アクションカメラ[8]やSamsungのGear360[9]などが挙げられる。本研究ではYouTubeのインターネット生放送サービスで、公式がサポートしているRICOHのTheta S[10]を用いた。Theta Sはdouble fish eyesと呼ばれる2つの魚眼映像が並んだ形式の映像が撮影可能である。1280×720(pixel)のHDサイズになっており、直径640ピクセルの魚眼映像1つ分で180度強の映像をカバーする。また、マイクも内蔵されているため、インターネット生放送での利用時はTheta Sを中心とした360度映像および音声を利用することが可能である。

3.2 360度動画

360度動画とは、動画プレイヤー上でドラッグ操作やフリック操作によって視聴方向を自由に変更することが可能な

動画のことである。スマートフォンやタブレット端末などでは、ジャイロセンサを用いることで端末の傾きを検知し、動画の視線を端末の動きに合わせて変化させることが可能である。OculusRift[11]やGoogle Cardboard[12]のようなHMDを利用する場合は顔を向けた方向をHMDでセンシングし、映像の切り替えを行う。映像ソースの入力はdouble fish eyes映像によって行われ、プレイヤー側で全天球映像に変換し、再生を行っている。

パノラマコンテンツは古くから認知されているコンテンツの1つであるが、ハードウェアやネットワークの発達により、普及が進んでいる。2012年に発表されたOculusRiftの発売から、HMDやパノラマコンテンツには関心が集まっていたものの、コンテンツ自体の少なさから爆発的な普及には至らなかった。しかし、2015年にYouTubeやFacebookなどのスマートフォンアプリで利用できるようになったことで、スマートフォンを持っていれば誰でも楽しめるコンテンツとなった。今後もスマートフォンをはじめとした様々なデジタルコンテンツへの対応が見込まれる。

360度動画に関する研究において、その活用法を模索する研究は近年多く進められている[13][14]。主に、カメラフレームの枠を超えた様々な表現の可能性や臨場感向上について論じられたものが多く、コミュニケーションにおける利用に関する研究は少ない。特に、360度動画の視点変更機能とコミュニケーションにおける視線の役割に着目した研究は今のところ報告されていない。

4. 提案手法

ここでは、視聴者の視線情報を放送者に可視化する視線ヒートマップの提案について述べる。また、視線ヒートマップを実装するにあたり対象としたユースケースについても述べる。

4.1 視線ヒートマップ

インターネット生放送における視聴者の視線情報を放送者に可視化するために、視線ヒートマップを提案する。視線ヒートマップは、視聴者が視線を向けている方向を集計し、全方位カメラを中心とした3D空間上の球体表面にヒートマップとして表示することで、集中した多数の視線および少数の視線の可視化を行う。図2に視線ヒートマップのモデル図を示す。

本手法において、放送者は放送用のインターフェースと共に視線ヒートマップを表示する。放送者は全方位カメラの撮影方向に対応付けられたヒートマップを見ることで視聴者の多くがどこに関心を示しているのか、多数から外れた視聴者がどこに関心を示しているのかを直観的に理解することができる。放送者は視線ヒートマップによって視聴者の視線を把握することで、視線というノンバーバル情報を加えて視聴者の反応に対応することが可能になる。

また、放送者の他にも視聴者が視線ヒートマップを表示することで、他の視聴者がどこに関心を向けているのかを知ることができる。他の視聴者の存在を視線の可視化によ

って感じることができ、視聴者間で行われるコミュニケーションも円滑にする効果が期待できる。

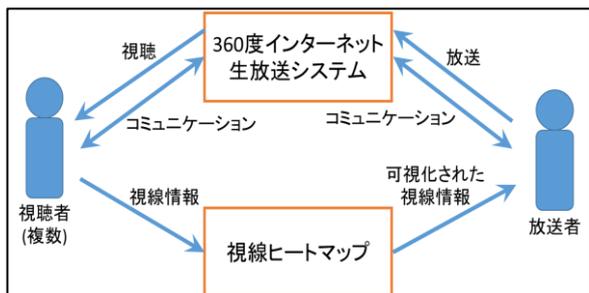


図2: 視線ヒートマップ

4.2 ユースケース

本研究において、全方位カメラを用いたインターネット生放送サービス利用のユースケースを、従来のインターネット生放送サービスの利用動向を基に定義した。津田ら[15]の調査では、インターネット生放送サービスの視聴および放送、両方の経験がある500人に対して放送内容に関するアンケートを行った。その結果、回答者が行った放送の内容について、視聴者参加放送、動物、エンタテインメント、音楽などが上位に上げられた。視聴者参加放送は、1人の放送者が多数の視聴者とのコミュニケーションを行う放送である。エンタテインメントではゲーム、音楽では楽器の演奏や歌唱などを放送している。今回、動物については、定点カメラを用いたペット等の観察を行うものであったため、コミュニケーションをメインとした放送ではないと判断し、対象から除外することとした。視聴者とのコミュニケーションをメインとした放送内容の中で上位に挙げられた視聴者参加放送、エンタテインメント、音楽を本研究では対象とする。放送者は1人のメイン放送者と数人のサブ放送者(サブ放送者は存在しない場合も有り得る)とし、放送者の中心に全方位カメラを設置するものとする。

津田らの調査結果から、個人で行われている放送において視聴者数が100人を超えるものがごく少数であることが分かっている。よって、本研究のユースケースにおける視聴者は数十人程度の不特定ユーザとした。放送内容は数人程度の放送者による、視聴者を巻き込んだトークやゲーム、音楽の演奏とした。これは、音楽やゲームをメインの放送内容とした放送であっても視聴者とのコミュニケーションによってプレイするゲームや演奏する曲の決定を行っており、ゲームや音楽のみで放送時間すべてを使用することはあまりなく、複数の内容を放送していることが多いからである。また、ネットワーク的な放送環境の観点から、室内での利用を想定し、周囲の状況を全方位カメラで把握可能な程度の大きさの部屋で行うものとした。

視聴者から放送者に対し行われるコミュニケーション手段は、多くのインターネット生放送システムで利用されていると同様に、テキストによるものを使用することとした。

5. 実装

ここでは、本研究で視聴者の視線を可視化するためにあたって用意した、全方位カメラを用いたインターネット生放送の放送環境及び、視線ヒートマップを実装した視線可視化アプリについて述べる。

5.1 システムの構成要素

本研究におけるインターネット生放送システムの構成要素は、放送者用のクライアントと視聴者用のクライアント、映像を仲介するWebRTCサーバ、視線情報を取得する視線サーバ、視線可視化アプリである。システムの構成図を図3に示す。ここでは、放送環境全般についての説明を行った後、視聴者用インターフェースと放送者用インターフェース、視聴者の視線情報を集積し、統計を取る視線サーバ、実際に視線を可視化する視線可視化アプリについての説明を行う。

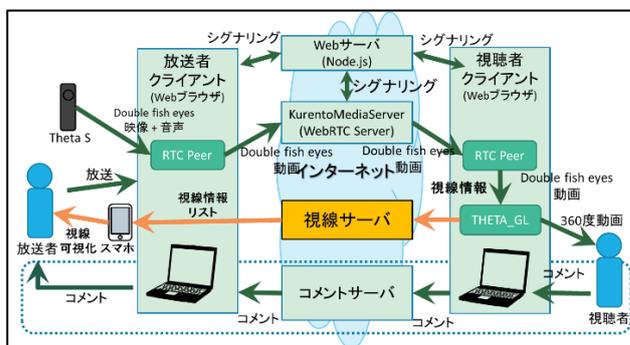


図3: 本研究のインターネット生放送システム構成図

5.2 全方位カメラを用いたインターネット生放送環境

既に全方位カメラを用いたインターネット生放送サービスはYouTubeなどで提供されている。研究開始当初、YouTubeで数回の放送を行い、その動作を確認したところ、放送者の行動や音声が見聴者のもとに届くまでに10秒前後の時間を要することが分かった。これはYouTubeにおいて、視聴者のPCスペックや回線状況に合わせて映像が配信できるように、YouTubeのサーバ側で低・中・高の3段階に映像をエンコードし直してから視聴者に映像や音声を送信することで発生する時間である。リアルタイムなコミュニケーションを行う上で、10秒前後のズレは話題の転換や次の発話が開始されてしまう可能性があり、視聴者も入力途中のコメントを差し控えるなど、円滑なコミュニケーションが行えない可能性があるかと判断した。そこで、本研究ではサーバサイドでのエンコード機能をカットしたアプリケーションによる処理時間の少ない、全方位カメラを用いたインターネット生放送システムを構築、使用することとした。コードの変更や機能追加などが容易に行える点も利点として挙げられる。

本研究では、放送映像の使用に関して、従来のFlash Playerを利用した動画配信の仕組みは使用しなかった。これは2016年にGoggle社がChromeブラウザにおけるFlash Playerサポート廃止を発表したこと[16]により、動画配信・共有サービス最大手であるYouTubeがFlashからHTML5

コンテンツへの乗り換えを発表しているためである[17]. YouTube の他, ニコニコ動画に関しても 2016 年 10 月 27 日より, 一部ユーザから順次 HTML5 版と題した HTML5 対応のプレイヤーの公開が進められている[18]. 動画配信・共有サービスで進められている HTML5 対応に合わせ, 本研究においても HTML5 で再生可能な WebRTC 技術を採用し, WebRTC の動画配信に対応した Kutrento Media Server[19]を利用した. WebRTC はリアルタイムコミュニケーションに必要な音声, 映像の配信やファイル共有などの機能を, プラグインを必要とせずに行うための API で, 特別なソフトウェアのインストールなどが不要なことから, ユーザが容易に利用できる.

また, 360 度動画の再生に関しては, Theta_GL[20]というライブラリを使用した. THETA_GL はインフォコム社が公開しているライブラリで, THETA S で撮影された double fish eyes 映像などを WebGL の 3D 空間における球体表面にマッピングすることで, 360 度動画として Web ブラウザ上で再生することができる. 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にも対応することができる.

Theta_GL では, WebGL の 3D 空間上に球体オブジェクトとカメラを設定している. 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おり, 球体の中心点にカメラ位置が設定されている. Theta_GL における 3D 空間上の球体オブジェクトとカメラの位置関係を図 4 に示す. カメラの方向は球体表面に緯度経度のように設定された (x, y) の座標によって表現されており, 球表面とカメラとの距離を示す z 座標を付加した (x, y, z) の情報として管理される.

本研究では 360 動画を再生しているプレイヤーから視聴者の視線の方向を (x, y, z) の座標として取得し, 視線方向を集計する「視線サーバ」に送信するようにした. 視線方向は視聴者のプレイヤーが読み込んだ Theta_GL 内でカメラ操作の変数として管理されている座標情報を取得している. 視線サーバで集計された視線の情報はヒートマップを作成するための点座標および半径, 強度の情報に変換される. また, 視線方向をインターネット生放送の映像や音声と対応付けるために, 視線方向の情報に対し, その情報を取得した時間の時刻情報を付与して送信することとした.

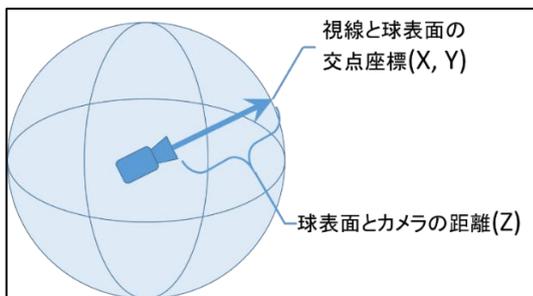


図 4: WebGL 上の 3D オブジェクトとカメラの関係

5.3 放送インターフェース

放送者および視聴者が使用するインターネット生放送インターフェースは Web サーバによって提供する. 放送者は

Web ブラウザでアクセスすることで図 5 に示すインターフェースを表示することができる. 放送開始のボタンを押下し, 使用するカメラをプルダウンメニューから選択するとカメラの映像が WebRTC サーバに送信され, 放送が開始される. 放送を停止する場合は放送停止ボタンを押下することで, WebRTC サーバへの映像送信が停止される. 放送開始とともに表示されるプレビューは視聴者が利用するプレイヤーと同じものになっており, 全方位カメラを中心とした自由に視聴方向を変更できる映像である. 視聴者から送信されたコメントはプレイヤー右側に時系列順にプレイヤー ID と共に表示される.

視聴者も放送者と同様に Web ブラウザでアクセスすることで視聴者用のインターフェースを表示することができる. 視聴者は視聴開始のボタンを押下することで WebRTC サーバから映像と音声を取得し, プレイヤによる再生が開始される. 再生が開始されると同時に, 視聴者の (x, y) で示される視線情報は後述する視線サーバへと送信される. また, 放送に対するレスポンスはプレイヤー下部に設置されたテキストフィールドより送信することが可能であり, すべてのユーザのコメントがプレイヤー右側に送信時間順で表示される.

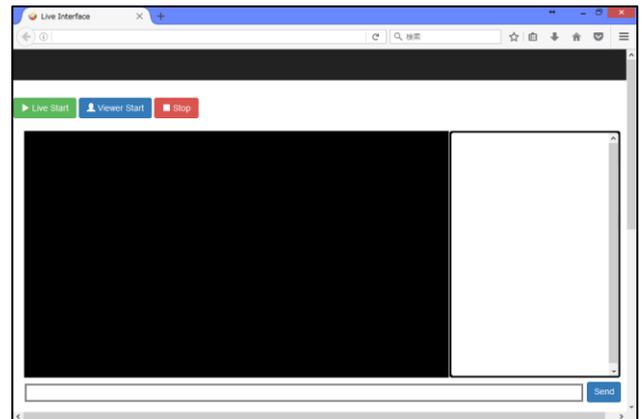


図 5: 放送者および視聴者用インターフェース

5.4 視線サーバ

視線サーバの役割は, 複数の視聴者クライアントから送信されてくる視線の方向情報を集計し, ヒートマップ作成の情報に変換することである. 3D 空間におけるヒートマップ作成には, 熱源となる点の座標 (x, y, z) および点が影響を及ぼす範囲(点からの半径), 点を持つ熱量の情報が必要になる. 今回の実装において, 点を持つ熱量と影響力に関しては変化をつけず, 定量的なものとした. 視線サーバでは, 送られてきた球中心から見た方向を示す (x, y) の情報から, 3D 空間上における視線方向と球表面の交点となる座標 (x, y, z) を算出し, そのリストを保持する. このとき, リアルタイム性を意識し, 一定時間が経過した点の情報についてはリストから自動的に削除されるようにした.

視線サーバは視線可視化アプリからのリクエストを受信すると, 自身が保持している視線方向のリストを可視化アプリに対して送信する.

5.5 視線可視化アプリ

視線ヒートマップを実装する上で、球体や半球、円筒状のディスプレイ、複数のプロジェクターを使用した投影など様々な手法が考えられる。しかし、これらの手法では使用する機材が特殊かつ高価であり、個人でインターネット生放送を行う場合には入手が困難である。また、機材を使用するにあたり全方位カメラで撮影している映像と視線ヒートマップの表示の向きを細かく調整する必要があり、実用が難しい。そこで、本研究では視線ヒートマップの可視化に AR 技術を利用した Android アプリを採用することにした。これは、放送を行うユーザが一般的に持ち得るデバイスであり、全方位カメラに合わせて AR マーカーを設置することで、撮影映像と視線ヒートマップの表示方向を容易に一致させることができるからである。視線可視化アプリは、放送者が PC で放送者用インターフェースを操作しながら、手で視聴者の視線をアプリにより視線ヒートマップとして表示するという利用方法を想定している。

視線可視化アプリでは、視線サーバから受信した視線方向のリストから視線ヒートマップを作成し、放送者に対する可視化を行う。アプリの実装は Unity によって行っており、ヒートマップの作成は Alan Zucconi によって提案されたシェーダを用いた方式[21]を使用した。この方式では 3D 空間上にヒートマップ用のデータリストを基に、中心点と半径、熱量のデータに合わせて透明な球体を配置する。シェーダが設定されたオブジェクトはこの透明な球体と重なった個所表面にその熱量に応じてヒートマップを表示する。作成したアプリでは、位置関係を分かりやすくするために、縦と横に線を引いた画像をマッピングした球体と、シェーダを適用した一回り大きい球体が重ねてある。

視線サーバに対するリクエストは毎秒送信され、視線方向のリストを更新していく。更新されたリストは即座に 3D 空間上に配置した球表面にヒートマップとして反映され、視聴者の視線をリアルタイムに可視化することができる。可視化アプリで表示する視線ヒートマップと全方位カメラの撮影した映像の方向を対応付けさせるため、可視化アプリの表示には既存の AR ライブラリを使用した。本研究では PTC ジャパンが提供する AR エンジン Vuforia[22]を用いて実装を行った。Vuforia では事前に登録した AR マーカーをアプリ起動時にカメラで認識することで 3D オブジェクトを重ね合わせて表示ができるようになる。認識に用いる AR マーカーを全方位カメラのレンズ方向に合わせて設置することで、アプリ内で表示したヒートマップと全方位カメラが撮影した映像との一致を図ることが可能である。

実際に視聴者の視線を視線可視化アプリによって表示した様子を図 6 に示す。図 6 において、設置された紙にはアプリにマーカーとして登録した図が印刷されており、その上部にヒートマップ表示用の球体が重ね合わせて表示されている。重ね合わされた球体の表面には、緑色のヒートマップが表示されている。また、球体右側に表示された矢印は全方位カメラとの設置方向調整のために表示しているも

ので、矢印の方向が合うように紙の向きを調整することで、ヒートマップの表示が撮影した映像と対応する。



図 6：可視化アプリで表示した視聴者の視線

6. 動作検証

実装した視線可視化アプリについて、その動作を確認するために動作検証を行った。ここでは、実施した動作検証について述べる。

6.1 動作検証の概要

実装した視線可視化アプリを評価に用いることが可能であるかの検証、評価実験を想定した課題の洗い出しを行うため、動作検証を岩手県立大学コミュニケーション学講座研究室にて、2017年2月10日に実施した。協力者5名を、放送者1名、視聴者4名に分けて実装環境での放送を実施し、動作の確認および視線ヒートマップによる視線の可視化についてインタビューを行った。放送時間は30分、全方位カメラは放送用PC近くに三脚で固定して設置した。放送内容は事前に設置した4つの設置物に関する話題を視聴者から提供し、放送者のアドリブで雑談するものである。話題の提供は視聴者1人に対し1つをあらかじめ割り振ったものであり、お互いの話題は分からないようにした。また、実験時のシステムではコメント機能の実装を行っていなかったため、視聴者と放送者のコミュニケーションはスマートフォンアプリ「LINE」によって行うこととした。動作検証における放送者および視聴者の様子を図7に示す。

30分の放送後に放送者および視聴者全員に、視線方向に関するインタビューを行った。また、放送者に対しては視線方向に関する質問に加え、視聴者の視線の動きを視線可視化アプリによって把握することができたか、放送をする上で不便に感じたことは何かを質問した。

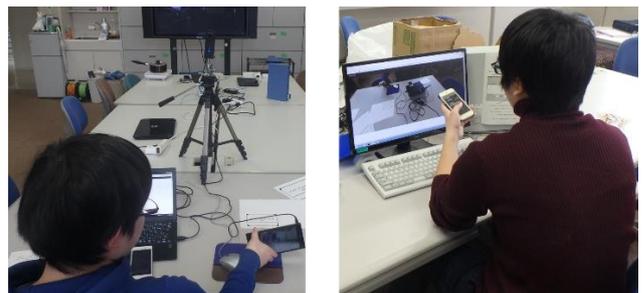


図 7：放送者(左)と視聴者(右)の様子

6.2 動作検証の結果

動作検証において、放送中に行われたコミュニケーションでは、問題は見られなかった。また、マーカーの向きや位置に関しても AR アプリの矢印表示をもとに容易に行えていた。しかし、三脚により全方位カメラを高い位置に固定したため、放送者や設置物を見るために視線の下方集中が見られた。

放送者からは「視聴者の視線をリアルタイムに把握することができ、視聴者の様子を把握しながら放送が行えた」という意見が得られた。一方で、視線ヒートマップの下部が見づらいつという点が挙げられた。これは AR アプリにおいてマーカーがカメラフレーム内に収まっていないとオブジェクトの表示が行えないため発生した。先に述べた視線の下方集中と併せ、現状では視聴者の視線の完全な把握は困難であると考えられる。また、視聴者も含めた実験参加者からは、PC の放送画面とスマートフォンを同時に操作することが負担になるとの意見が得られた。視線ヒートマップを表示した球体の下方や裏側を確認するためには、AR マーカーを回り込むような操作が必要になったため、スマートフォンを PC と同時に操作することが重度の負担となってしまったものである。このことから、放送を妨げずに視線ヒートマップの確認が行えるような、表示方法および操作に対する負担軽減を考察する必要がある。

7. 視線ヒートマップの表示方法検討

今回の動作検証で得られた意見の中で、視線ヒートマップの下部や裏側を見るために重度の負担が必要となる問題について、現状では放送および評価実験が困難であると判断した。放送画面を見ながらのスマートフォン同時操作について、スマートフォン端末自体を大きく動かすような操作は放送者に対する負担が大きいため、端末を移動させずに視線ヒートマップの表示を変化させる機能を実装することで解決を図る。

7.1 視線可視化の提示方法について

動作検証の結果から、下方に集中した視線情報を放送者に提示するために、視線ヒートマップの回転操作機能が必要であると判断した。当該機能はスマートフォンのタッチスクリーンを利用し、視線ヒートマップを表示した球体をスワイプ操作により表示の回転ができるようにするものである。ただしこの時、視線ヒートマップを表示した球体が回転することで AR マーカーとの同期が取れなくなり、全方位カメラが撮影する映像と視線ヒートマップにズレが発生してしまう。そのため、スワイプ操作後にユーザの指がタッチスクリーンから離れた時点で、球体の回転をリセットすることで、この問題を一時的なものとし、いつでも同期した状態に戻せるようにする。この機能により、視線ヒートマップの下方や裏側の表示をするためにスマートフォン端末を大きく動かす必要がなくなり、視線ヒートマップの確認に対する負担を軽減できる。

7.2 視線ヒートマップの回転表示における課題

先に述べたように、視線ヒートマップの回転を行う場合、回転中に表示されている箇所が全方位カメラで撮影している映像において、どの方向に対応するものなのかが咄嗟に分からないという障害が発生する。これは、視線ヒートマップを表示する上でベースとしている球体に直線のみが表示されていることが原因である。前述した通り、スワイプ操作の終了と共に回転をリセットすることで、この問題は一時的なものにはなるものの、視線ヒートマップの確認が直観的に行えないことは視線ヒートマップ本来の目的を達成するためには改善が必要である。これを解決するためにはいくつかの方法が考えられる。

7.3 回転表示における課題解決手法

その1つとして、全方位カメラの撮影した映像をベース球体に投影することで、重ね合された視線ヒートマップに対応する箇所を実際の映像で確認できるようにする方法がある。しかし、この方法ではリアルタイムの映像を投影するとスマートフォンの通信量が增大する可能性がある。また、周囲の映像を球体表面に投影する方法が最も適した表示方法であるかは不明である。

他にもマーカーに対して水平なリング状のオブジェクトを別途表示し、リングの各角度に対して集中度合いをヒートマップ表示し、集中する高さを色の濃淡によって示すという方法も考えられる。しかしこの方法では集中した視線の表示しか行えず、多数から外れた視聴者がどこに関心を示しているのかを放送者に対して提示することが難しい。

いくつかの表示方法がある中で、今回の実装では暫定的に、ベース球体がどの程度回転しているのかを放送者が理解できれば良いと仮定し、ベース球体の中心を原点とした3D空間におけるX軸Y軸Z軸の3軸方向をそれぞれ赤、緑、青の3色の矢印で表示する方法を使用する。3軸の表示を実装した様子を図8に示す。図8のベース球体中心から伸びた矢印はそれぞれX軸、Y軸、Z軸を示しており、回転表示を行った際もベース球体との相対方向は固定されたままになる。この3軸を表示する方法は3Dモデリングソフトなどでオブジェクトの回転を行う場合に回転度合いを表示するために使用される手法であり、オブジェクトがどの程度回転しているのかを示す場合には有効であると考えられる。これにより、放送者は放送画面のプレビューを確認することなく、視聴者の視線方向を把握することができる。

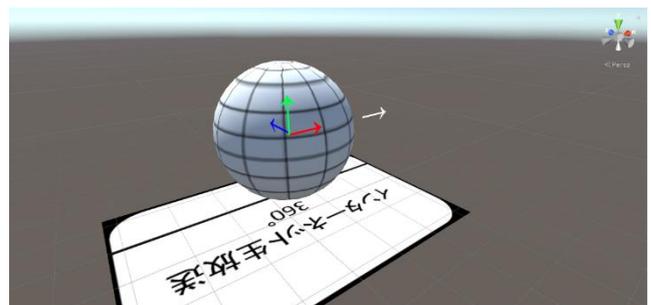


図 8 : 3 軸の実装を行った視線ヒートマップ

8. 評価

視線ヒートマップによってインターネット生放送におけるコミュニケーションの円滑化が図れるかを検証するために、評価実験を実施する。ここでは、評価の目的と方法を説明する。

8.1 評価の目的と方法

評価の目的は視線ヒートマップでコミュニケーションを円滑化することが可能であるかを検証することである。そのために、放送者と視聴者でのコミュニケーションが必要となるタスクを実験協力者に対して与え、タスクが完了するまでの時間を計測、視線ヒートマップの有無による対照実験を実施する。実験後に放送者と視聴者それぞれに対してインタビューを行うことで計測時間以外の観点からも分析が行えるようにする。

評価におけるタスクは動作検証で行った生放送の環境と同様に部屋に予め用意した設置物に関する質問を視聴者から放送者に対しコメントで行う。この時、放送者のコメントに対する回答が行われるまでの時間を計測することとする。これは、視聴者のコメントの意図や意味を放送者が理解するまでにかかる時間を計測する目的であり、放送者が視聴者の視線情報を持つ場合と持たない場合にどれだけの差が生まれるのかを比較することで、視線ヒートマップの有効性を比較できる。視聴者が行った質問の対象を、放送者がインターフェースで確認するまでの時間、質問に対する回答を行うまでの時間をそれぞれ計測する。これにより、視聴者の興味対象の把握と、コメントの意図を理解するためにかかる時間を分析することができるようにした。

8.2 評価における注意点

前回実施した動作検証において、放送者から得られた意見として、「放送者の部屋に用意した設置物は見慣れないものが多く、放送開始前に配置を把握してしまった」との意見から、実際にコメントで質問しないものもダミーとしていくつか追加で設置することとした。また、動作検証と異なる点として、コメント方法もインターフェース上で行えるように変更してある。

9. まとめ

本論文では、全方位カメラを用いたインターネット生放送において放送者が視聴者の視線を察知できないという問題点に対し、視線ヒートマップによる解決を図った。また、視線可視化アプリの動作検証の結果を元に、視線ヒートマップの表示方法の改善を提案した。表示方法の改善では一案としてベース球体を中心とした X 軸, Y 軸, Z 軸の 3 軸を付加表示することで回転表示時の視線ヒートマップ方向の同期を取る。

今後は今回の表示方法改善が有効であるかを評価し、視線可視化アプリが全方位カメラを用いたインターネット生放送において、放送者と視聴者の間で行われるコミュニケーションを円滑化することができるかを評価していく。ユースケースを元にした評価方法を検討し、複数回の実験と

その結果のフィードバックを行っていく。

参考文献

- [1] 総務省 : 通信利用動向調査 平成 27 年調査, 2016; http://www.soumu.go.jp/johotsusintokei/statistics/data/160722_1.pdf
- [2] 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
- [3] 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
- [4] Hiroshi Ishii, Minoru Kobayashi, “ClearBoard: a seamless medium for shared drawing and conversation with eye contact”, CHI '92 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp.525-532, 1992
- [5] 武川直樹, “コミュニケーションにおける視線の役割”, 電子情報通信学会誌 Vol.85(10), pp.756-760, 2002
- [6] Naoki Mukawa, Tsugumi Oka, Kumiko Arai, Masahide Yuasa, “What is connected by mutual gaze?: user's behavior in video-mediated communication”, CHI EA '05 CHI '05 Extended Abstracts on Human Factors in Computing Systems, pp.1677-1680, 2005
- [7] 大野健彦, “視線共有に基づく遠隔地間コミュニケーション”, 電子情報通信学会技術研究報告. HIP, ヒューマン情報処理 104(747), pp.55-60, 2005
- [8] PIXPRO - マスプロ電工 | MASPRO, <https://www.maspro.co.jp/products/pixpro/sp360/>
- [9] Gear 360 | ウェアラブル | by Galaxy, <http://www.samsung.com/jp/consumer/mobilephone/gear/gear/SM-C200NZWAXJP>
- [10] 製品紹介 | RICOH THETA, <https://theta360.com/ja/about/theta/s.html>
- [11] Oculus, <https://www.oculus.com/rift/>
- [12] Google Cardboard(商品情報)
- [13] 笠原俊一, 暦本純一, “JackIn : 一人称視点と体外離脱視点を融合した人間による人間のオーグメンテーションの枠組み”, 情報処理学会論文誌 56(4), pp.1248-1257, 2015
- [14] 片山康太郎, 柴田啓司, 堀田裕弘, “極座標方向への画像積分による全方位動画像中の車両検出の高速化”, 電子情報通信学会技術研究報告. IE, 画像工学 113(469), pp.125-128, 2014
- [15] 津田侑, 上原哲太郎, 森村吉貴, 森幹彦, 喜多一, “インターネット生放送におけるユーザの活動の分析”, システム制御情報学会論文誌 28(10), pp.407-418, 2015
- [16] Intent to implement: HTML5 by Default, 2016, [https://groups.google.com/a/chromium.org/forum/#!searchin/chromium-dev/HTML5\\$20by\\$20default/chromium-dev/0wWoRRhTA_E/_E3j40AAJ](https://groups.google.com/a/chromium.org/forum/#!searchin/chromium-dev/HTML5$20by$20default/chromium-dev/0wWoRRhTA_E/_E3j40AAJ)

- [17] YouTube Engineering and Developers Blog: YouTube now defaults to HTML5 <video>, 2016, https://youtube-eng.googleblog.com/2015/01/youtube-now-defaults-to-html5_27.html
- [18] 「動画視聴ページ HTML5 版 (β)」提供開始のお知らせ - ニコニコインフォ, 2016, <http://blog.nicovideo.jp/niconews/ni064261.html>
- [19] Kurento, <http://www.kurento.org/>
- [20] Theta_GL, 2016, https://github.com/mganeko/THETA_GL
- [21] Arrays & shaders: heatmaps in Unity, 2016, <http://www.alanzucconi.com/2016/01/27/arrays-shaders-heatmaps-in-unity3d/>
- [22] Vuforia | Augmented Reality, <https://www.vuforia.com/>