

全方位インターネット生放送における 視聴者の POV 情報再生ツールの開発

高田真也¹ 西岡大¹ 齊藤義仰¹

概要: 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サービスが開始された。しかし、全方位カメラを用いたインターネット生放送では従来のサービスと異なり、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を向けている方向は、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。そのため、我々はこれまでに視聴者の視聴方向の情報を可視化して放送者に提示する「視線ヒートマップ」を提案、実装した。視線ヒートマップは視聴者の POV 情報を球体状のヒートマップで表現することで放送者に対し可視化する。今後は視線ヒートマップの発展として POV 情報の分析を行い、放送の進行を支援する放送アシスタントシステムを開発していく。そのために必要となる POV 情報分析手法を検討するため、POV 情報再生ツールの提案と実装を実施する。

A Development of the POV Replay Tool in Internet Live Broadcast using an Omnidirectional Camera

MASAYA TAKADA¹ DAI NISHIOKA¹ YOSHIA SAITO¹

1. はじめに

情報技術の普及とともにインターネット生放送サービスが広く利用されるようになった。YouTube やニコニコ動画に代表されるインターネット生放送では、放送者と視聴者がコメントにより、リアルタイムなコミュニケーションを楽しめる。YouTube では 2016 年に全方位カメラに対応したインターネット生放送サービスの提供が開始された。このサービスでは視聴方向を 360 度全方向に向けられる。

しかし、全方位カメラを用いた放送では、視聴者の視聴方向を放送者が察知できないという問題が発生する。従来の放送では単一のレンズが撮影方向を示しており、視聴者は必然的にその撮影範囲を視聴している。それに対し、全方位カメラを用いた場合では、複数のレンズで撮影を行っており、レンズの方向から視聴者の視聴している範囲を把握することができない。コミュニケーションにおいて、視線情報が大きな役割を担うことは、すでに多くの研究で分かっている[1]。視聴者がどこを見ているのか、すなわち視聴者の興味はどこに向けられているのかを察知できない場合、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、視聴方向の情報を POV 情報と呼ぶ。POV は Point of view の意味で使用し、POV 情報とは視聴者が視聴用インタフェースを用いてどの方向の映像を見ているのかという情報を示す。そこで、全方位カメラを用いたイン

ターネット生放送における POV 情報をコミュニケーションにおける視線情報として捉え、放送者に対して POV 情報を提示することでコミュニケーションの円滑化を図る視線ヒートマップシステムの実装を行った[2]。実装したシステムでは視聴者の POV 情報を球体状のヒートマップとして放送者に提示することで、視聴者の POV 情報を放送者がリアルタイムに把握できるようにした。

しかし、実装した視線ヒートマップシステムでは POV 情報の可視化を行っているだけであり、ヒートマップの読み取りや意味づけを放送者が行う必要があった。実装システムの発展として、複数の POV 情報を分析して視聴者の放送に対する興味や興奮度を数値化する分析機能、および分析結果を基に放送者に進行のアドバイスを行うアドバイザ機能が必要になると考える。そこで、POV 情報の分析手法の検討が必要となる POV 情報の再生ツールの開発を行った。本稿では、POV 情報の再生ツールの概要と実装について報告する。2 章では、コミュニケーションにおける視線の役割について、関連研究とともに説明する。3 章では、視線ヒートマップについて、そのモデルや実装について述べる。4 章では、POV 情報再生ツールの概要と設計に関して説明する。5 章では実装について述べる。

2. コミュニケーションにおける視線の役割

これまでにも、行われた研究において、コミュニケーシ

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

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[3][4][5]. 武川ら[6][7]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[8]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, コミュニケーション相手と共に協調的な作業を行う際には, コミュニケーション相手の視線がどこに向けられているのかという情報は, その作業における作業効率や達成度に対して関係があると考えられる.

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, 視聴者の行ったコメント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. 視線ヒートマップ

本章では, 視聴者の POV 情報を放送者に可視化する視線ヒートマップについて述べる. また, 全方位インターネット生放送システムとともに, 視線ヒートマップの実装についても述べる.

3.1 視線ヒートマップの概要

インターネット生放送において, 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を提案した. 視線ヒートマップは, 視聴者が視聴している方向を集計し, 仮想球体の表面にヒートマップとして表示することで, 多数視聴者および少数視聴者の POV 情報の可視化を行う. 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を直観的に理解することができる. これにより, 放送者は多くの視聴者に関心を示しているものを把握した上で, コミュニケーションを行うことが可能になる. また, 多数の視聴者と異なる方向を視聴している視聴者の POV 情報を把握することで, ユニークな話題の採用や視線の誘導を行うことも可能になる.

3.2 全方位インターネット生放送システム

本研究では, コードの変更や機能追加などが容易に行えることから, 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, 使用した. 本研究では, 放送映像の使用

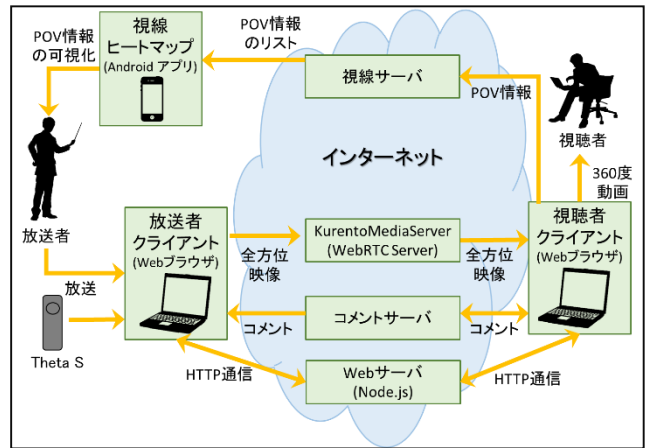


図1 全方位インターネット生放送システムの構成図

に関して, 従来の Flash Player ではなく WebRTC を利用した. 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[9]ためである. YouTube やニコニコ動画などのインターネット生放送サービスでも 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[10][11]. 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 Kurento Media Server[12]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[13]というライブラリを使用した. THETA_GL は RICOH 社が販売する全方位カメラ THETA S で撮影した全方位映像を, 360 度動画として Web ブラウザ上で再生することができる. 再生には Three.js を用いた WebGL の描画を用いている. 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, 球体の中心にカメラ位置が設定されている. 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)の座標によって管理される. 視聴者用インタフェースでは POV 情報を取得し, POV 情報を集約する「視線サーバ」に送信する. 視線サーバで集約された POV 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, POV 情報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, その情報を取得した時刻の情報を付加した.

3.3 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取得した POV 情報を集約し, ヒートマップ作成用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), 点が持つ熱量の情報が必要になる. 今回

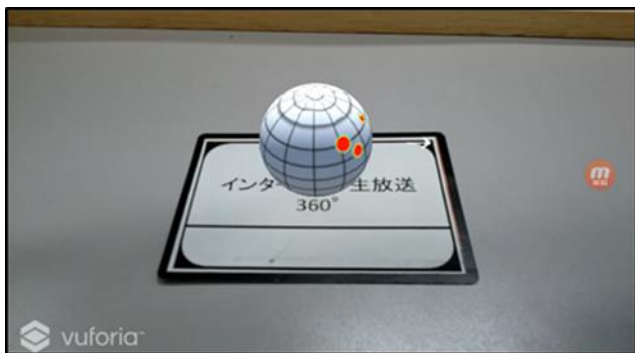


図2 アプリで可視化した視聴者のPOV情報

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、一定のものとした。視線サーバではこれらの情報を生成し、そのリストを保持する。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。

3.4 視線ヒートマップ

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AR技術を利用したAndroidアプリとして実装することとした。これは、Android端末は放送者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、生放送の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。アプリは、放送者がPCで放送開始、終了のみの操作を行い、放送中は手元で視聴者のPOV情報を確認するという利用方法を想定している。

アプリの実装はUnityによって行い、ヒートマップの作成はシェーダを用いた方式[14]を使用した。視線サーバに対するリクエストは毎秒1回の割合で行い、POV情報のリストを更新する。視線サーバが保持するPOV情報はWeb_GL上の座標を基にしているため、現実空間やUnityにおける3D空間の座標とは齟齬が発生する。そこで、ARマーカーを用いて基準方向をキャリブレーションすることでこの課題を解決することとした。AR機能の実装にはARエンジンVuforia[15]を用いた。視聴者のPOV情報をアプリによって可視化した様子を図2に示す。

4. POV情報再生ツール

本章では、POV情報再生ツールについて述べる。まず、本研究におけるユースケースについて述べ、どのような放送を研究対象としているのかを説明する。その後、POV情報を分析する利点について説明する。最後にPOV情報再生ツールの概要と実装方法について述べる。

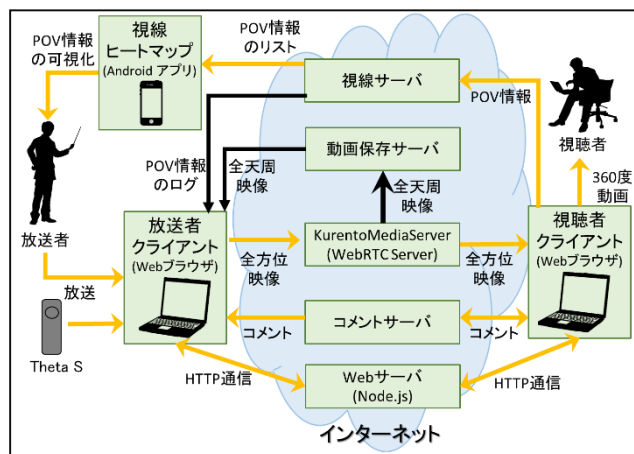


図3 実装システムの構成図

4.1 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[16]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。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。調査の結果から、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。放送内容は視聴者を巻き込んだトークやゲーム、音楽の演奏とした。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。

4.2 POV情報を分析する利点

全方位カメラを用いて作成されたコンテンツは、視聴可能範囲が広く、その全てを一度に視聴することができない。コンテンツ制作者の意図した被写体に気づきやすいコンテンツを作るため、視聴者の視聴体験を再現、分析する手法も存在する。VRCH[17]は360°VRのコンテンツ解析とユーザ行動のパターン解析を人工知能によって行うというものである。視聴中のオンデマンドの360度動画をリアルタイムに解析し、その結果を動画にオーバーレイ表示することが出来る。分析できるものは、人物や物体、話されている言語、雰囲気、シーン構成、ストーリーなどの様々な動画の構成要素である。このことから、データ分析の手法を用いて視聴者の視聴体験を分析することは、視聴者のニーズを把握することにつながり、放送の質向上が見込めるということが分かる。

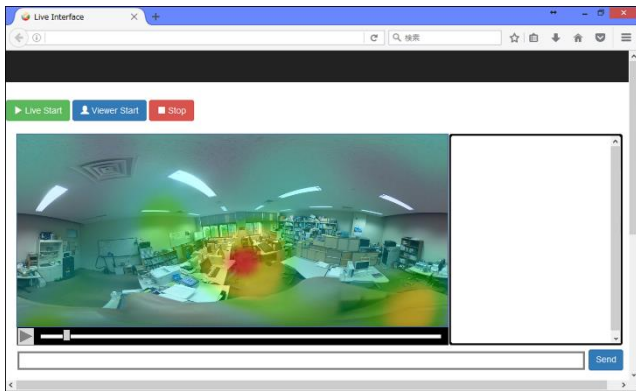


図 4 ヒートマップの再生イメージ

本研究において視聴者の POV 情報を分析する理由は 2 つある。まず 1 つ目はコミュニケーションにおける視線が担う役割が大きいという点である。2 章でも述べたように、コミュニケーションにおいて相手の視線情報を把握することで、相手の興味や関心、ニーズを把握することができる。しかし、インターネット生放送のような 1 対多の状況においてコミュニケーション相手全員の POV 情報を人間が把握することは困難である。そこで、POV 情報の分析をシステムにより行う必要がある。2 つ目は、視聴者からの反応のうち、常に得られる情報が POV 情報だからである。インターネット生放送サービスにおいて、視聴者が行えるアクションは限られている。多くのインターネット生放送サービスで共通しているものとしては、コメント機能、放送者からのアンケートに答える機能など、受動的なものしかない。それに対し、POV 情報は常に発信される情報であり、分析を行うことで、コメントなどを行わない視聴者であっても、そのニーズを把握することが可能になる。

以上のことから POV 情報を分析する手法を作成する必要があると考え、分析手法検討に必要なツールの設計と実装を行った。ツールを追加したシステムの構成を図 3 に示す。POV 情報の再生ツールや特定視聴者の視聴体験再生機能は、視線ヒートマップの実装および評価時に私用した全方位インターネット生放送システムをベースに作成した。ツールは放送終了後に使用するものとし、放送終了時に発行する ID を用いることで、放送時の映像やログなどを利用することができる。

4.3 POV 情報再生ツールの概要

先に実装した視線ヒートマップは、全方位インターネット生放送におけるリアルタイムなコミュニケーションを支援する目的で、実装および評価を実施した。具体的な手法として、ヒートマップによってリアルタイムな POV 情報の分布を可視化することで放送者に直感的な理解を促した。その為、得られた POV 情報について、ログの書き出しは行うものの、ログデータの分析や利用をシステム上では行っていない。視線ヒートマップの評価実験の際に行った放送では、話題の転換時や話題が停滞した際に視聴方向の変化が多く、話題が盛り上がっている場合は視聴方向の変化が少なかった。このことから POV 情報のログを基に視

聴者のニーズや放送に対する興味を分析することが可能になると考えられる。そこで、POV 情報の分析手法を検討するために POV 情報再生ツールが必要となる。POV 情報再生ツールの機能は以下の 2 つである。1 つ目は、ヒートマップ表示の再生機能である。これは、どのタイミングでどこに視聴者の関心が集中していたのか、または関心が分散していたのかを分析するために必要となる。2 つ目は、視聴者の視聴体験の再生機能である。これは、特定の視聴者がどのように生放送を視聴していたのかを再現することにより、視聴方向変更操作に対する質的な調査を行うことが可能になるためである。

4.4 ヒートマップ表示の再生機能

YouTube が提供している全方位インターネット生放送サービスでは、全方位カメラを用いて作成した 360 度動画や全方位インターネット生放送時の映像の投稿サービスが提供されている。この投稿機能を利用した投稿者向けのアナリティクスサービスも提供されている[18]。当該サービスでは、視聴者がオンデマンド動画を見た際の視聴方向の動きを、動画と重畳表示されたヒートマップによって確認することができる。

本研究においても、サーバサイドで生放送の映像を保存し、ヒートマップを重畳表示する手法を用いて、視聴者の POV 情報を確認できる機能を実現することにした。ヒートマップの再生イメージを図 4 に示す。現在、全方位インターネット生放送システムで使用している映像は Theta S のライブビュー機能により取得したものであり、dual fish eyes というフォーマットである。これは、円状になった魚眼映像が 2 つ並んだもので、Theta S の表と裏に 1 つずつ搭載された広角カメラを用いて撮影された映像がそれぞれの円によって表現されている。また、動画中央部が現実空間における下部方向に指定されているため、dual fish eyes 映像にヒートマップを重畳表示しても直感的な理解は困難であると考えられる。

そこで、サーバサイドの動画保存機能とともに dual fish eyes 映像をエクイレクタングラー形式の映像に変換する機能を実装することにした。エクイレクタングラー形式とは、正距円筒図法とも呼ばれ、パノラマ映像を矩形に引き伸ばして表現する形式である。動画の保存機能については、WebRTC 形式で通信を行い、動画を保存する動画保存サーバを用意することにした。動画の変換には、WebGL 上でのシェーダを用いた手法[19]を利用することにした。全方位インターネット生放送システムと同様に Web 上で利用することができる。放送者は放送終了後に自身の放送の ID を入力することにより、自身の放送におけるヒートマップの変化を確認することができる。再生時には映像下部に表示したシークバーを動かすことで特定の時間における視聴者の POV 情報を確認することもできる。

4.5 特定視聴者の視聴体験再生機能

POV 情報の分析手法を評価する際に、様々な評価手法が考えられる。本研究では、放送直後に視聴者に自身の視聴

記録を見てもらい、興味や関心の度合いについてスライダーを用いて回答してもらおう手法を検討している。その際、視聴者が自身の視聴記録を正確に再確認できるようにする必要があるのである。

本機能は、WebGL 上でのカメラ方向を操作するメソッドを用いて、視聴者の視聴ログの再生を行う。視聴者の視聴ログは毎秒 5 回程度の割合で出力されており、映像の再生時間に合わせてログを読み込むことで、映像と当時の視聴方向の同期を図る。本機能を利用する際は視聴用インタフェースで視聴記録再生のボタンを押下し、自身の ID を選択することで視聴記録の再生が開始される。

特定視聴者の視聴体験を再生する機能は、2 次元の座標データとタイムスタンプによって構成される POV 情報ログを実際の映像とともに可視化することができ、視聴者に対するアンケートのほかにも、画像処理技術を利用することで、視聴者がどの設置物に興味を示していたのかを分析することも可能になる。将来的には、より詳細な調査を行うために画像処理技術の利用も必要になると考える。

5. 今後の展望

今後、実装したツールを用いて、視聴者の POV 情報に関する分析手法について検討を実施していく。まずは、視聴方向変更時の移動量について分析し、視聴者の放送に対する興味や集中の度合いを数値化する手法について検討する。また、POV 情報の分析手法の検討以外にも、実装したツールには様々な利点があると考えている。放送者は自身の放送において、視聴者の視聴方向がどのタイミングで集中しているのかを確認することができるようになる。この効果についても検討するため、今後の全方位インターネット生放送については、放送終了後に放送者に自身の放送を確認してもらい、簡単なアンケートに回答してもらう。

また、視線ヒートマップの発展として、視聴者の POV 情報を分析して放送者にアドバイスを行う放送アシスタントシステムの開発を行う。放送アシスタントシステムは POV 情報を分析した結果から、視聴者の放送に対する興味が低下した際に話題の転換を放送者にアドバイスする。さらに、放送に対する興味や集中度合いが高まった際には現在の話題に関する掘り下げを放送者にアドバイスする。近年、Deep Learning の登場により、Google Home[20]や、Clova WAVE[21]など、対話的にユーザの生活を支援するデバイスが多く登場している。対話的にアドバイスを行うことで、全方位インターネット生放送におけるアシスタントシステムも、自然な会話を行っているかのように放送の進行を支援することが可能になると考える。

6. まとめ

本研究では、インターネット生放送において、全方位カメラを用いることにより、放送者が視聴者の視聴方向を感じられなくなることを問題点として、視聴者の POV 情報をリアルタイムに可視化する視線ヒートマップを実装した。

今後はその発展系として視聴者の POV 情報を分析し、放送者の進行をサポートするアシスタントシステムを検討していく。そこで、POV 情報の分析手法を検討するため、視聴者の POV 情報ヒートマップを再生する機能と、特定の視聴者の POV 情報のログを再現する再生ツールの設計、実装を行った。今後は実装したツールを用いて POV 情報の分析手法について提案および評価を行っていく。

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] 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”, マルチメディア, 分散, 協調とモバイル シンポジウム 2017 論文集, (2017)
- [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] Mai Otsuki, Taiki Kawano, Keita Maruyama, Hideaki Kuzuoka, Yusuke Suzuki, “ThirdEye: Simple Add-on Display to Represent Remote Participant’s Gaze Direction in Video Communication”, Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp.5307-5312, 2017
- [6] 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
- [7] 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
- [8] 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
- [9] Intent to implement: HTML5 by Default, 2016, https://groups.google.com/a/chromium.org/forum/#!search/in/chromium-dev/HTML5%20by%20default/chromium-dev/0wWoRRhTA_E/_E3jf40AAJ, (2017 年 12 月参照)
- [10] 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, (2017 年 12 月参照)
- [11] 「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>, (2017 年 12 月参照)
- [12] Kurento, <http://www.kurento.org/>, (2017 年 12 月参照)
- [13] Theta_GL, 2016, https://github.com/mganeko/THETA_GL, (2017 年 12 月参照)
- [14] Arrays & shaders: heatmaps in Unity, 2016, <http://www.alanzucconi.com/2016/01/27/arrays-shaders-heatmaps-in-unity3d/>, (2017 年 12 月参照)
- [15] Vuforia | Augmented Reality, <https://www.vuforia.com/>, (2017 年 12 月参照)
- [16] 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
- [17] VRCHHEL, <https://vrchel.com/>, (2017 年 12 月参照)
- [18] 360° 動画のヒートマップレポート, <https://support.google.com/youtube/answer/7407544/>, (2017 年 12 月参照)
- [19] Stereoarts Homepage, 2015, <http://stereoarts.jp/>, (2018 年 5 月参照)
- [20] Google Home - スマート スピーカー&ホーム アシスタント - Google ストア, https://store.google.com/product/google_home?hl=ja, (2017 年 12 月参照)
- [21] Clova 公式サイト, <https://clova.line.me/>, (2017 年 12 月参照)