

コミュニケーション中の動画像を用いてコミュニケーションの向上を目指したビデオ通話システムの実現

Realization of a Video Communication System with Improvement Using Video Images during Online

北野 雄大[†] 奥谷 五月[†] 荒木 英夫[‡]
 Kitano Yuuta Okutani iduki Araki Hideo

1. はじめに

2020年4月より2022年6月現在に至るまで、コロナ禍により人々が集まってコミュニケーションを取ることに様々な制限が加えられていた。そこで、実際に集まる代わりに遠隔コミュニケーションツールを用いた方法が普及した。コロナ禍以前にも、同等の技術は既に存在していたが、これらのツールの利用はすぐに会うことができない遠方とのコミュニケーションなどに限られていた。しかし、2022年6月現在までにオンライン会議やオンライン授業、オンライン研修、オンライン飲み会など様々な用途で活用された。

しかし、これらの遠隔コミュニケーションツールはビデオチャットや画面共有等の機能が主流で、画面を通じたコミュニケーションは会議等では有用である。しかし親睦を深める様なコミュニケーションでは伝わらないものも多いのが現状だ。

そこで、本研究では、WebRTCを用いたビデオ通話システムを利用してコミュニケーション中の動画像からの情報を抽出し、別の方法でフィードバックすることでコミュニケーションの幅が広がり、心理的距離を縮めることができると考えた。

本論文では、この提案システムの実現とその実験結果について報告する。

2. 既存システム

これまでも遠隔コミュニケーションを行うことを目的としたいくつかのシステムが構築されている。

その一つにハートトゥハートオンラインコミュニケーションシステムというものがある。このシステムでは、相手の存在感や生命感といった現象そのものを伝達することによってヒトの気配を伝えることが試みられている。webカメラで読みとった顔色の変化の低周波数成分を抽出し、オンラインコミュニケーション上の各々の話者の画面にヴィジュアル化した形でフィードバックすることで、参加者自身と他の参加者の顔色の変動を視覚的に認識しながらコミュニケーションできるユーザーインターフェースである。この研究のユーザーインターフェースの特徴は主に相手や自身の気配、存在感や生命感を伝達共有するために、ユーザーの血流の変動に伴う顔色の変化を、センサー等を利用せずにwebカメラで読み取ってフィードバックを行っている点である。このシステムは、心拍を読み取って相手にフィ

ードバックすることによってヒトの存在感や生命感、気配などを伝達共有が試みられている。[1]

しかし、このシステムでは、ヒトの動作・リアクション等が相手に伝わりづらいためであると考えられる。

また、オンラインコミュニケーションには人の位置関係が存在せず、また視線によるコミュニケーションも難しく、さらに音声についてもすべての人の声が混ざっているため、だれとコミュニケーションをとっているのかが判断できない。何より、利用者がコミュニケーションをとっている最中に会話のコンフリクトが発生するなど、これまでのコミュニケーションツールでは対面で行われるような情報や関係の取捨選択を行うことが難しい。しかし、快適なコミュニケーションを実現するためにはこのようなことが自然にできる必要があると考えた。

そこで本研究では、ヒトの顔の向きをコミュニケーション中の動画像から抽出して、相手にフィードバックすることによって、コミュニケーションの活性化を実現し、さらに新しいコミュニケーションの提供が可能なシステムを目指す。

3. 提案システム

現在提案するシステムは、大きくわけてビデオ通信システムと画像処理による頭部方向検出である。

3.1 ビデオ通信システムの詳細

ビデオ通信システムの構築にはWebRTCを利用した。WebRTC(Web Real-Time Communication)とはAPIを経由して、ウェブブラウザやモバイルアプリで高速なデータ通信をリアルタイムで行うためのツールで映像や音声などの容量の大きいデータをリアルタイムで送受信でき、P2Pの仕組みを持っているためビデオチャットやWeb会議などに利用されている。

本システムではWebRTCのP2P方式で実現している。

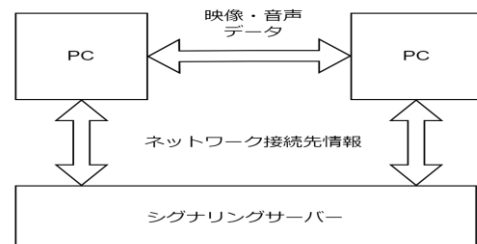


図1 本システムの通信の仕組み

[†] 大阪工業大学 大学院 情報科学研究科 Graduate School of Osaka Institute of Technology

[‡] 大阪工業大学 情報科学部 Osaka Institute of Technology

P2P 方式は直接通信したいクライアント同士が、図 1 内ではシグナルリングサーバーを利用して、相手のユーザー名や通信に必要な情報を取得し通信を開始する。通信そのものはシグナルリングサーバーを介せずに PC 同士が直接通信する。

3.2 画像処理による頭部方向検出の詳細

画像処理による視線検出・頭部方向検出の構築には Python とそのライブラリである OpenCV と Dlib を利用した。Dlib とは C++・Python の機械学習系のライブラリである。

頭部位置検出では顔認識用の 68 点学習済みデータ (shape_predictor_68_face_landmarks.dat) を利用している。その中の眉の内側・目の内側・鼻の頭と両側・口の外側・唇の下・あごの特徴点を利用して顔の向きを検出している。それをもとに頭部方向の検出を行う。

3.3 カメラの同時使用

Windows では通常 1 台のカメラに 1 台のソフトウェアを割り当てるようになっている。そこで ManyCam と OBS Studio という 2 つのソフトを使うことによってカメラ画像をスプリットして Python ソフトとビデオ通話を同時に行えるようにしている。

4. 実験と結果

4.1 実験方法

本システムが正常に動作しているかを確認するために 2 台の PC を用いて 1 対 1 の実験を行った。

今回はビデオチャット、頭部方向検出した時の顔の向きの値の表示の 2 つが正常に動作しているかを検証した。

4.2 実験結果

実験の結果を図 2 と図 3 に示す。図 2 が Python による頭部方向検出している様子である。図 3 がブラウザ上でコミ

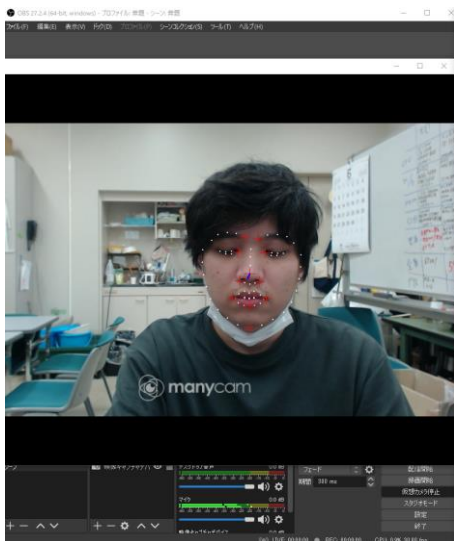


図 2 Python による認識処理

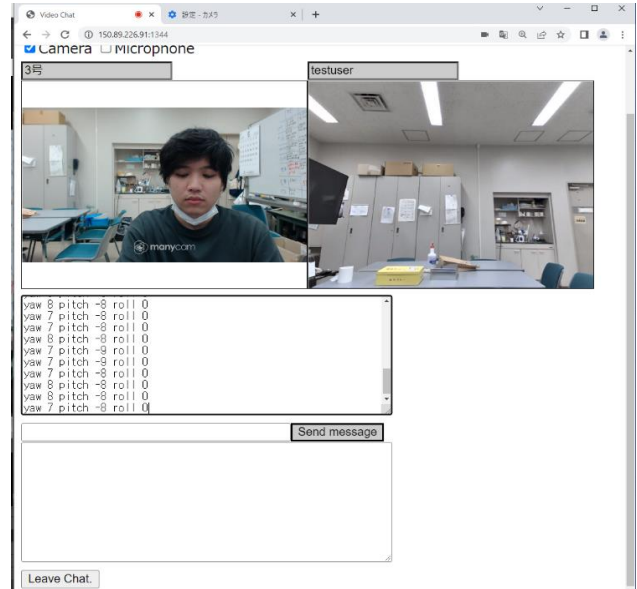


図 3 イメージ図

ュネーションを取っている様子である。

図 2 上にある点は、OpenCV と Dlib により顔の特徴点を抽出したものであり、これの特に頭部の向きを数値化したものをブラウザ経由で相手に送信している。

図 3 のブラウザ上では、画像処理した映像ではなく、通常のカメラ画像が転送されていることがわかる。また、ブラウザにビデオの双方向通信が表示されていることがわかる。ブラウザの左側の画像はユーザ (3 号)、右側の画像はユーザ (testuser1) である。画像の下のテキストボックスで頭部方向の値が表示されていることが分かる。今回は動作確認の為に、ユーザ (3 号) の頭部方向の値を送ることとした。その結果、顔の向きを変えると値が変化することが確認できた。

以上のことから提案システムが正常に動作したことを確認した。

5. おわりに

今回の実験では、2 つの確認項目が実行できたと確認され、本システムは正常に動作している。

今後の課題として、最大 4 人までのビデオチャットの実現を目指している。このシステム上で 4 人を升目上に並べた状態でビデオチャットを行い、誰がどのヒトを見ているかが詳細にわかるようにすることを目指す。

また、目線や顔の色等を画像処理によって抽出してそれをユーザーにフィードバックすることによって、コミュニケーションの活性化を測る予定である。

参考文献

- [1] 林 晃世, 平林 真実, 小林 昌廣, “ハートトゥハートオンラインコミュニケーション”, エンタテインメントコンピューティングシンポジウム (EC2021), (2021)