

# カメラ目点：カメラを目に見立てたビデオチャットシステムの提案

清水 春翔\* 宮下 芳明\*

**概要.** アイコンタクトは信頼の形成や不安の低減など会話で重要な役割をもつ。だが、ビデオチャットでは互いにアイコンタクトを取ることが困難である。この問題に対処するために、視線補正やカメラ位置を移動させることによってアイコンタクトを可能にする手法が研究されている。しかし、既存研究では視線補正による違和感や、大掛かりなシステムの持ち運びの困難さが問題である。本研究では、ディスプレイに埋め込まれたパンチホールカメラか、小型 USB カメラと PC を使用するシステムを提案する。相手の目を検出し、カメラと相手の目の位置を常に重ねることができるシステムを作成した。カメラを瞳孔と見立てることによって、カメラを意識することなくビデオチャットでアイコンタクトを可能にする。

## 1 はじめに

遠隔地の人と会話をする方法としてビデオチャットがある。ビデオチャットはリアルタイムでオーディオとビデオを送受信し、相手の表情を見ながらコミュニケーションをとることができる。対面とビデオチャットでの会話において、大きな違いはアイコンタクトの有無である。アイコンタクトは非言語コミュニケーションのひとつで、信頼の形成や不安の軽減に大きな意味を持つ。非言語コミュニケーションは個人個人の関係において、重要な役割を果たしていると考えられている[1]。

しかし、一般的なビデオチャットでは、相手の目の映像と自分を撮影しているカメラレンズの位置が異なることからアイコンタクトをとることが困難である。ビデオ通話で視線を合わせるためにはユーザは直接カメラを見る必要がある。ゆえに、画面上の相手の映像を見ることができない。同様に、相手の目の映像を見ても相手に視線を向けることができない。そこで、現在は視線を合わせやすいようにディスプレイに取り付けられる小型ウェブカメラがある。ただし、こうした製品の使用例をみると相手の顔や目と重ねてウェブカメラを貼り付けている人はほとんどいない。また、映像に表示される相手が動くことにより、カメラと相手の位置がずれてしまうことも問題となる。この問題を解決するため、既存研究として、視線補正を用いる方法と、ハードウェアを用いた方法がある。



図 1. パンチホールカメラを目と見立てた様子

視線補正を用いる方法は、たとえ横を見ていても正面を向いているように映像を加工することである。ビデオコミュニケーションの違和感改善など様々な応用を見据えて古くから研究が行われてきた。しかし学習データの影響を受け、顔の同一性が保たれないという問題があった。塚本らは、目領域の入力画像への合成により、未学習の特徴を持つ利用者への個人学習を伴わずに視線変更を可能[2]にした。他にも、NVIDIA が提供する Maxine [3]がある。AI が視線を推定し、視線がカメラに向いているように補正することでアイコンタクトをしているように見せる技術である。この技術では常に相手と視線を合わせた映像にすることができる。しかし、会話において常に視線が合ってしまうのは不自然であると考えられる。たとえば考え事をする時、気まずい時には、あえて視線を相手から外し、強く伝えたいことがある時には意識的に目を見るなど、アイコンタクトには適切なタイミングや提示時間がある[4]。このことから、視線補正を用いる方法は完全に問題を解決できていないといえる。

Copyright is held by the author(s). This paper is nonrefereed and non-archival. Hence it may later appear in any journals, conferences, symposia, etc.

\* 明治大学



図 2. A: 上部の内蔵ウェブカメラ, B: 右耳近辺に設置する小型ウェブカメラ, C: 提案システム (小型ウェブカメラ使用) の比較

ハードウェアを用いる方法では、相手がどこを向いているのかわからない状況においてコミュニケーションを達成すべく、協同描画メディア ClearBoard[5]のコンセプトが提案されている。さらに、描画の保存再利用ができる ClearBoard-2[6]も作成されている。ClearBoard-2の使用で、相手がスクリーン上のどこを見ているのかが容易に分かる「gaze awareness」を実現した。さらに、吉富らはプロジェクタとカメラを用いて視線の調整が可能なテレビ電話システム[7]を提案した。他には、Tahsinらの設置型システム[8]がある。カメラが半透明スクリーンの後ろに取り付けられており、半透明スクリーンの後ろから映像を取得している。話し相手の目の位置に応じてレールに取り付けられたカメラを移動させ、視線を維持している。この方法はたしかに有用だが、半透明スクリーンに加え移動制御可能なカメラなど専用装置が必要であり、装置が大掛かりになってしまう。なので、持ち運びに適していない。

さて、本研究の目的は、相手（一人）とビデオチャットする際に、視線を合わせたいときのみ合わせることができ、可搬性があるなど、現実的に実装・普及が容易なことである。

本研究では、パンチホールカメラや小型ウェブカメラを人の目に例えることによって、あまり違和感を与えずにアイコンタクトを可能にできると考えた。図1にパンチホールカメラを目と見立てた様子を示す。パンチホールカメラは、カメラの周りをディスプレイで覆ってしまう方式である。ノートパソコンやモニターではまだ多く採用されていないが、スマートフォンには多く採用されてきている。パンチホールカメラを使う方がより人の目に近いが、提案システムを実行できるスペックのパンチホールカメラ付属のPCがないため、小型USBカメラをパンチホールカメラに見立てた。

## 2 提案システム

本稿の提案システムはカメラとビデオチャットで表示される相手の目の位置が常に等しくなるようにウィンドウを移動させる。提案システムは任意のビ

デオチャットと仮想カメラを組み合わせる。

OBS Studio でビデオチャットの相手映像をキャプチャし、仮想カメラとして出力した映像を利用する。仮想カメラから WebGazer.js[9]のライブラリを用いて、目の座標を取得する。取得した目の座標は、Node.js をサーバーサイドで利用して socket 通信を行い、目の座標が取得されるたびにウィンドウの移動を行うプログラム側に送信される。

まず初めに、ユーザは OBS Studio の仮想カメラを起動し、ビデオチャットに表示される相手映像を仮想カメラの映像として出力する。その後、システムを実行するとウィンドウが生成され、ウィンドウには仮想カメラの映像が表示される。右目の座標を用いて、カメラと右目の位置が重なるようにウィンドウが移動する。仮想カメラを使用することによって、任意のビデオチャットと組み合わせる提案システムを使うことが可能になっている。これにより提案システム専用のビデオチャットを使う必要がない。

## 3 システムの比較と考察

最後に、システムを構築したうえで、他のシステムと比較を行った。図2は、内蔵ウェブカメラ、ディスプレイに取り付け可能な小型ウェブカメラ、小型ウェブカメラと提案システムをビデオチャットで使用した際の比較画像である。図2(AB)では相手の目とカメラの位置が異なることが見てとれる。このことからアイコンタクトが困難であることがわかる。一方で、図2(C)の提案手法では、相手の目を見たときに、カメラと視線が合っていることが確認できる。

本研究では、視線補正や大掛かりな装置を使用せず、既存のビデオチャットを使用し、アイコンタクトを可能にするシステムを提案した。今後は、既存のビデオチャットのみを使用した場合と提案システムを利用した場合の評価実験を行っていきたい。

なお、本手法は前述の通りパンチホールカメラが最も適していると考えられる。パンチホールカメラは画面角に設置されることが多いが、このシステムの有効性を受けて、画面中央に設置された製品が開発されることを望む。

## 参考文献

- [1] 中村克樹. 非言語コミュニケーションの意義. 学術の動向, pp. 28-31, 2004.
- [2] 塚本想也, 西保匠, 菅野裕介. 個人学習を伴わない視線変更モデル適応のための目領域画像合成. 第84回全国大会講演論文集, pp. 319-320, 2022.
- [3] NVIDIA Maxine.  
<https://developer.nvidia.com/maxine>.  
(2023/10/29 確認)
- [4] 福原省三. アイ・コンタクトと印象の評価が受け手の対人感情に及ぼす効果. 心理学研究, pp. 177-183, 1990.
- [5] 小林稔, 石井裕. ClearBoard : シームレスな協同描画空間のデザイン. 情報処理学会研究報告ヒューマンコンピュータインタラクション(HCI), pp. 133-140, 1992.
- [6] 小林稔, 石井裕. ClearBoard-2 における協同作業空間と会話空間のシームレスな融合. 情報処理学会研究報告グループウェアとネットワークサービス(GN), pp. 43-50, 1993.
- [7] 吉富千紗, 中嶋信生. テレビ電話におけるアイコンタクト効果の検討. 情報処理学会研究報告オーディオビジュアル複合情報処理(AVM), pp. 53-56, 2006.
- [8] M. T. Tausif, R. Weaver, and S. W. Lee. Towards Enabling Eye Contact and Perspective Control in Video Conference. *Adjunct Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, pp. 96-98, 2020.
- [9] A. Papoutsaki, et al. Webgazer : scalable webcam eye tracking using user interactions. *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, pp. 3839-3845, 2016.
- [10] Avasam obs-virtual-cam.  
<https://github.com/Avasam/obs-virtual-cam/releases>.  
(2023/10/29 確認)