

二人称映像を用いた疑似アイトラッカー

渡邊 康平^{1,a)} 満上 育久^{1,b)} 佐川 立昌^{2,c)}

概要: 本研究では、対面コミュニケーションにおいてお互いの人物が視野カメラのみを着用するだけで、アイトラッカーと同様に視野映像中の注視点を獲得できる手法を提案する。この手法ではまず、視野カメラを着用した人物がそのカメラで鏡に映った自身の顔を撮影した映像を用いて、頭部と視野カメラのキャリブレーションを行う。そして、自分の顔を映した相手の視野カメラ映像から自分の頭部方向と視線方向を推定し、それを自身の視野カメラ座標系に変換することで注視点を推定する。実際に対面コミュニケーションを行う2名の相互の視野カメラ映像に対して提案手法を適用してその計測性能を評価した。

キーワード: アイトラッカー, カメラキャリブレーション, OpenFace

1. はじめに

人と人が日常的に行う対面コミュニケーションにおいて、お互いのことを理解し合うためには、相手が話す言語情報だけではなく、非言語情報も重要となる。特に人の視線情報は、「目は口ほどにものを言う」ということわざにもある通り、その人物を理解するための有益な情報を含んでいる。このような観点から、人の視線情報は、消費者の購買予測 [1] や自動車運転者の状態推定 [2] など、様々な研究分野で広く利用されている。人の視線を計測するための装置としては、トビー・テクノロジー・ジャパン株式会社の Tobii Pro Glass3 [3] や株式会社ナックイメージテクノロジーの EMR-9[4] に代表される着用型のアイトラッカーが挙げられるが、一般に比較的高価なことからことから、多人数のコミュニケーション分析のために全員分用意するのは容易ではない。

そこで本研究では、双方の対面する二者の視野カメラの映像だけから、アイトラッカーと同等の視線情報（視野画像中の二次元座標として与えられる注視点）を取得する。提案手法では、まず、ウェアラブルカメラを装着した人物が、頭を動かしながら鏡の向こうの顔を撮影した映像を用いて、頭部に対するカメラの位置・姿勢を推定する。次に、OpenFace によって得られる視線情報は人の顔の大きさ・形によって大きな誤差をもつことから、事前の計測によってそのずれ方を学習しておき、補正を行う。また、性能評

価として、提案手法を用いて実際に対面コミュニケーションを行う二者の映像から、お互いの視野中の注視点を推定し、アイトラッカーで得られた真値と比較することで提案手法の有効性を確認した。

2. 関連研究

アイトラッカーには、ウェアラブルのものと同スクリーンベースのものが存在する。ウェアラブルは眼鏡やゴーグルに似た形状をしており人が着用して使用するもので、トビー・テクノロジー・ジャパン株式会社の Tobii Pro Glass3 [3] や株式会社ナックイメージテクノロジーの EMR-9[4] などがその代表例である。一方、スクリーンベースは、画面の縁などに据え置き画面を見つめる人物の眼球を計測するタイプのデバイスであり、Tobii Pro X3-120 [5] や富士通株式会社の EyeExpert [6] などが知られている。これらを含めアイトラッカーとして販売されているデバイスの多くは、近赤外線 LED で角膜に写る光の反射パターンを計測・分析を行う角膜反射法を用いており、比較的高精度に視線を計測することが可能だが、一般に高価であり、多人数のコミュニケーションの分析を行いたい場合にも気軽に多数個用意できるものではない。

より容易・安価に視線を計測する方法として、カメラを用いる方法が多数存在する [7], [8], [9], [10], [11]。最近では、顔認識のためのオープンソースライブラリである OpenFace [12] (図 1) を用いて顔向きや視線計測を行う事例も多く見られる。これらの方法は、上述のアイトラッカーに比べるとコストが低いが、多人数コミュニケーションにおける各人の視線を取得するために環境に設置したカ

¹ 広島市立大学

² 産業技術総合研究所

a) watanabe@sys.info.hiroshima-cu.ac.jp

b) mitsugami@hiroshima-cu.ac.jp

c) ryusuke.sagawa@aist.go.jp

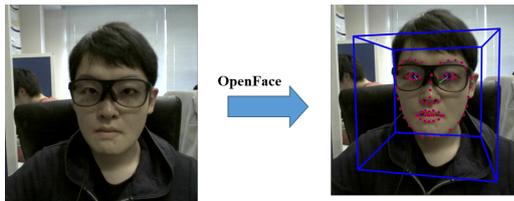


図 1 OpenFace による顔向き・視線方向の推定

メラで撮影した映像を用いると、顔や目の画質や解像度が不十分なため精度が低くなってしまいます。

Murakami らは、新たなアプローチとして人の眼球・頭部協調運動をモデル化することで、眼球の観測は用いず頭部運動の観測のみから注視方向を推定する手法を提案している [13], [14]。眼球の観測を用いないため環境カメラの映像からでも視線が推定できるが、人の自然な眼球・頭部協調運動を前提とした推定のため、一般にコミュニケーションの分析の対象となるような、人の意図・情動に影響を受けた目の動きを推定するには適さないと考えられる。

3. 疑似アイトラッカー

前章で述べたとおり、視線情報を取得するには通常、高精度な視線計測が行えるが高価なアイトラッカーを用いるか、安価だが視線計測精度が悪い環境カメラを用いるかを選択することになる。しかし、人と人のコミュニケーション分析という応用を想定すると、各人が視野を撮影するカメラさえ着用していれば、お互いのカメラに相手の顔が大きく映り込んでいるはずである。この視野カメラの映像を用いることで、高価なアイトラッカー無しでもそれと同等のデータを取得しようというのが、本研究のキーアイデアである。すなわち本研究では、対面する相手のカメラで撮影された二人称映像のみを用いることで、視線推定を実現させる疑似的なアイトラッカーを提案する。映像中から顔や視線を推定するには OpenFace を用いるが、上述のとおり画像ベースの方法のため一般に大きな誤差を持つため、それを補正する処理も施すことで、精度の向上を目指している。

視線の測定対象者は、頭部と視野カメラのキャリブレーションを事前に行う。この事前キャリブレーションで、測定者は固定された鏡の前で頭を上下・左右に振る。このとき、測定者のウェアラブルカメラ映像中の各瞬間では頭部の位置・姿勢が異なる鏡像群が得られる。測定者が頭を水平方向に振る状況で、測定者の頭部を固定したとき、これに対応する各瞬間の鏡像群の頭部の位置・姿勢は図 2 のように考えられる。測定者と鏡像群の位置関係を利用して測定者の頭部と視野カメラの位置・姿勢を算出する。また、この事前キャリブレーションの際に測定者は常に鏡像の視野カメラを注視することで、同じく図 2 の位置関係を活用して推定された視線情報に対して補正を行う。

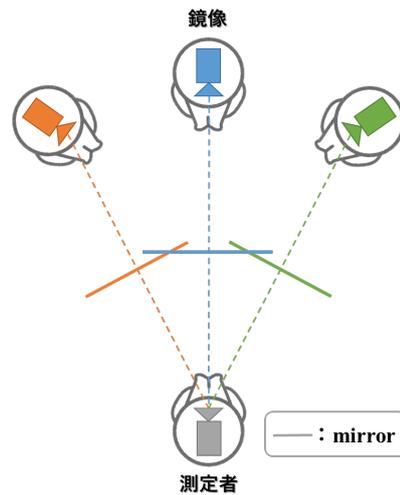


図 2 測定者の頭部固定時の測定者と各鏡像の位置関係

以上の事前キャリブレーションの後、対面する二人がお互いに視野カメラを装着した状態で、相手の視野カメラに映る測定対象者の視線方向と頭部方向を推定する。この視線情報に対して、事前に求めたパラメータで補正を行い、これを測定対象者の視野カメラ座標系に変換することで注視点を描画した。

3.1 事前キャリブレーション

本手法では、測定対象者は実験の事前に固定された鏡の前で頭を振るようなキャリブレーションを行う。このとき、測定対象者の頭部に取り付けられた視野カメラでは、顔の向きの異なる鏡像群が得られる。測定者の頭部を固定すると、この鏡像群の位置・姿勢は図 2 のように存在する。ここで、測定者の視野カメラで撮影される画像を水平方向に反転すると、鏡の影響を無視した、鏡像群が装着する視野カメラで測定者を撮影した画像に変換される。鏡像群から撮影したこのような画像群に対して OpenFace を適用すると、鏡像群のカメラ座標系で測定者の頭部の位置 t_i と姿勢 R_i が推定される。

ここで、測定者の視野カメラ映像を n 枚の画像群に変換するとき、測定者の頭部(顔の中心部)を統一座標系として、対応する鏡像群のカメラ位置 T_i ($i = 0, 1, 2, \dots, n-1$) は、

$$T_i = -R_i^{-1}t_i \quad (1)$$

と表現される(図 3)。

このように、鏡像群の視野カメラで撮影した画像群及び測定者と鏡像群の位置関係を利用することで、頭部と視野カメラのキャリブレーションを行う方法を以下の 3.1.1 項及び 3.1.2 項で示す。

3.1.1 OpenFace の顔基幹点を用いたキャリブレーション

OpenFace では、出力として顔の位置・姿勢に加えて、図 1 のように、顔の基幹点(図中の赤丸群)が推定される。こ

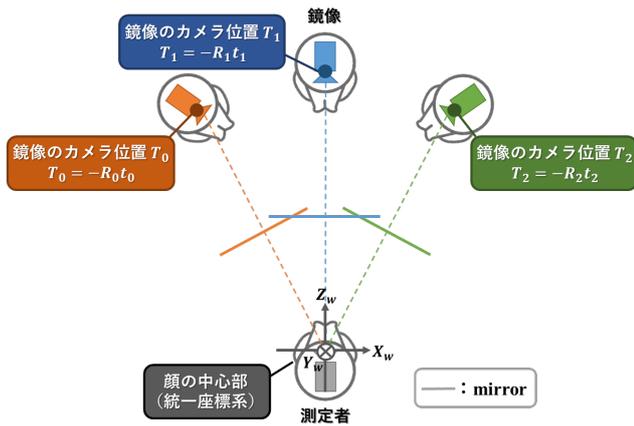


図 3 OpenFace の顔基幹点とウェアラブルカメラの位置関係

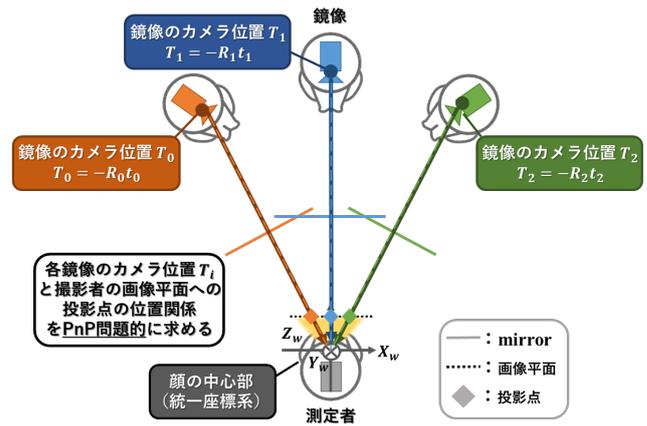


図 5 測定者の画像平面への各鏡像のカメラ位置投影点の位置関係

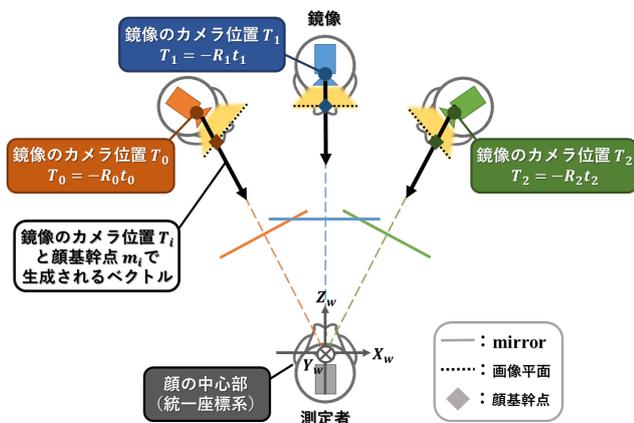


図 4 測定者と各鏡像及び顔基幹点の位置関係

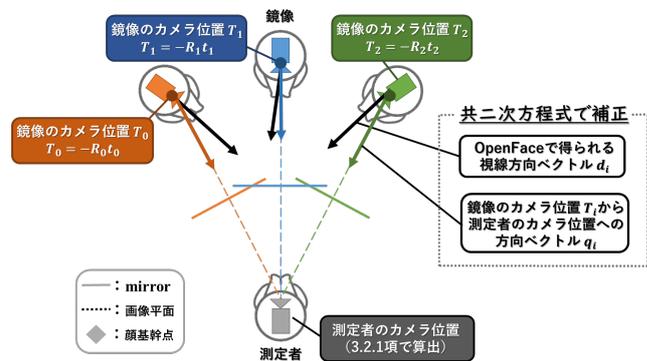


図 6 視線情報の補正に用いる情報

の基幹点のうち、鼻筋の最上点が装着する視野カメラのカメラ部とおよそ同じ位置であることから、本研究では、この基幹点を、鏡像群で撮影されるカメラ画像中における測定者のカメラ部と仮定した。

測定対象者の顔の中心部を原点とする鏡像群のカメラ位置は式 (1) で既に算出されている。続いて、鏡像群の視野カメラ撮影画像に写る測定者のカメラ部は、鏡像群のカメラ位置を始点とし、これと各鏡像の視野カメラ画像中のカメラ部への方向ベクトルで求まる直線上に存在することが分かる。このとき、各鏡像で算出される直線群の交点を最小二乗的に求め、これを測定者のカメラ位置とした。

3.1.2 PnP 問題解決によるキャリブレーション

3.1.1 項では、測定者のカメラ位置を推定した。本項では、これを利用して測定者のカメラ姿勢を推定する。

図 5 のように、鏡像群のカメラ位置 (式 (1)) と測定者のカメラ画像平面上へのその投影点の関係性を利用して、測定者のカメラ姿勢を PnP 問題で求める。本 PnP 問題を解く際には、OpenCV で公開されている関数 solvePnP Ransac を用いる。上記の関数を用いて、カメラの内部パラメータ、歪み係数と 3.1.1 項で求めた測定者のカメラ位置を初期解として与えることで、測定者のカメラ姿勢を推定した。

3.2 視線情報の補正

OpenFace を用いた視線推定は、眼球・瞳の大きさ等の個人差の影響によって、推定に誤差が生じることが知られている。

そこで、事前キャリブレーションでは、測定者のカメラ位置・姿勢の算出だけでなく、OpenFace で推定される視線情報の補正も同時に行う。具体的には、鏡の前で頭を上下・左右に振る際に測定者は鏡像のカメラ部分を常に注視する。これは、図 6 のように、鏡像のカメラから測定者のカメラへのベクトル $\mathbf{q} = (q_{xi}, q_{yi}, 1)$ と、測定者から鏡像へ向ける視線方向ベクトル $\mathbf{d} = (d_{xi}, d_{yi}, 1)$ は理想的には一致するはずであるが、後者は OpenFace による推定であるため誤差が生じる。この 2 つのベクトルが式 (2) のような共二次方程式で補正できるものとみなし、事前キャリブレーションの際に係数パラメータ $a_i, b_i (i = 0, 1, \dots, 5)$ を最小二乗的に算出する。以後に OpenFace で推定された視線はこの係数パラメータにより補正する。

$$\begin{cases} q_{xi} = a_0 + a_1 d_{xi} + a_2 d_{yi} + a_3 d_{xi}^2 + a_4 d_{xi} d_{yi} + a_5 d_{yi}^2 \\ q_{yi} = b_0 + b_1 d_{xi} + b_2 d_{yi} + b_3 d_{xi}^2 + b_4 d_{xi} d_{yi} + b_5 d_{yi}^2 \end{cases} \quad (2)$$

3.3 視野画像への視線情報マッピング

事前キャリブレーションの後、お互いに頭部に視野カメラを取り付けた二者が対面し、視線情報を推定する。この



図 7 ポスター発表を想定した実験結果

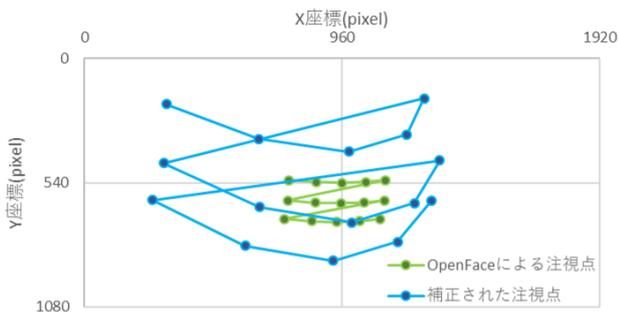


図 8 共二次方程式による補正の可視化

とき、視線の測定対象者を測定者、その視線を視野カメラで撮影する人物を撮影者とする。撮影者のカメラ映像を用いて測定者の視線を推定し、これに対して事前に求めた係数パラメータで補正を行う。この補正された視線情報を測定者の視野カメラ座標系に変換することで注視点を推定する。

4. 実験

実際にポスター発表の場面を想定した実験を行った。発表者と聴衆は、お互いに頭部に視野カメラを装着して約1メートル離れている。発表者の視野カメラで得られた聴衆の視線情報に対して補正と変換を行い、聴衆の視野カメラ上にマッピングした。

結果は以下の図7の通りとなった。①と②はそれぞれ聴衆の視線がポスター中から発表者に向けられる場面と発表者からポスター中に向けられる場面である。どちらの場面においても、本手法で推定された注視点が真値(Tobii Pro Glass3による計測値)と同様の動きをしていることが確認できた。

次に、本手法の事前キャリブレーションで求まる係数パラメータについて、この有用性について評価する。具体的には、OpenFaceで推定された格子点状の注視点群に対して係数パラメータを適応することで、補正の様子を可視化した。

結果は以下の図8の通りとなった。注視点群はよりスケールが大きく、非線形的に補正が行われていることが確認できた。

5. おわりに

本研究では、対面コミュニケーションにおいて、お互いの人物が視野カメラのみを着用するだけでアイトラッカーと同様に視野映像中の注視点を獲得できる手法を提案した。この手法では、視野カメラを着用した人物がそのカメラで鏡に映った自身の顔を撮影した映像を用いて頭部と視野カメラのキャリブレーションを行い、自分の顔を映した相手の視野カメラ映像から自分の頭部方向と視線方向を推定し、それを自身の視野カメラ座標系に変換することで注視点を推定する。今後の課題としては、提案手法の各ステップの処理を改良するのに加えて、時間軸方向の連続性を考慮して視線推定精度を向上させることを検討している。

参考文献

- [1] 金子雄太, 石橋健, 矢田勝俊, 「視線追跡データを用いた消費者の店舗内購買行動の分析」, PACIS2018 主催記念特別全国研究発表大会, 2018年.
- [2] 森みどり, 中易秀敏, 三好哲也, 「ドライビングシミュレータとアイトラッキングシステムを用いた運転者の眼球運動と車軌跡の同期解析」, 日本機械学会論文集(C編), Vol.79, No.803, 2013年.
- [3] Tobii Technology K.K., “Tobii Pro Glasses3,” <https://www.tobii.com/ja/product-listing/tobii-pro-glasses3/>
- [4] ナックイメージテクノロジー, 「モバイル型アイマークレコーダ EMR-9」, <https://www.eyemark.jp/product/emr-9/>
- [5] Tobii Technology K.K., “Tobii Pro X3-120,” <https://www.tobii.com/ja/product-listing/tobii-pro-x3-120/>
- [6] FUJITSU, 「小型視線検出システム EyeExpert」, <http://www.fujitsu.com/jp/group/fct/products/eyeexpert/>
- [7] 青山晃治, 山村毅, 大西昇, 杉江昇, 「一台のカメラによる顔と視線方向の推定」, 電情報通信学会技術研究報告, Vol.PRU95, No.233, pp.131-136, 1996年.
- [8] 佐橋正也, 青山晃治, 大西昇, 「耳と顔の特徴の位置関係に基づく顔画像からの視線方向の推定」, 電子情報通信学会技術研究報告, Vol.PRMU97, No.596, pp.113-120, 1998年.
- [9] 青木茂樹, 大西正輝, 小島篤博, 山田憲一, 三宅克実, 福永邦雄, 「運転者の目領域抽出による視線推定」, 電子情報通信学会技術研究報告, Vol.PRMU99, No.450, pp.15-22, 1999年.
- [10] 竹上健, 後藤敏行, 「角膜反射像と虹彩輪郭情報を併用した高精度視線検出法」, 電子情報通信学会技術研究報告, Vol.HIP97, No.509, 1998年.
- [11] 佐竹純二, 小林亮博, 平山高嗣, 川嶋 宏彰, 松山 隆司, 「高解像度撮影における実時間視線推定の高精度化」, 電子情報通信学会, 2008年.
- [12] Tadas Baltrušaitis, Peter Robinson, Louis-Philippe Morency, “OpenFace: an open source facial behavior analysis toolkit,” pp.1-10, 2016.
- [13] J. Murakami, I. Mitsugami, “Gaze from Head: Gaze Estimation without Observing Eye,” The 5th Asian Conference on Pattern Recognition, 2019.
- [14] J. Murakami, I. Mitsugami, “VR-based Eye and Head Motion Collection for Modeling Their Coordination,” IEEE 8th Global Conference on Consumer Electronics, 2019.