

シーンの三次元形状と固視位置の整合に基づく視線検出器の自動校正

研究代表者

池田 聖

大阪大学 准教授

1 はじめに

視線計測は、様々な分野で応用されている。例えば、心理実験や市場調査などでは、視線情報は人の興味や関心、精神状態の解析などに用いられ、視線の入力インタフェースや中心窩レンダリングなどでは、計測した視線を実時間で利用する用途に用いられる。特に、カメラで撮影した眼の画像から使用者の注視箇所を計測する方式は、カメラの小型化、低価格化により広く普及し、バーチャルリアリティ (VR) 用もしくは拡張現実感 (AR) 用ヘッドマウントディスプレイ (HMD) では、視線計測器は標準搭載されつつある。

一般に、視線計測器は、使用者毎に校正する必要がある。視線検出器の校正原理を理解するには、人の視覚システムの知識が必須である。人は、対象物を視細胞が最も密集する中心窩で鮮明に捉える。中心窩と対象物を結ぶ線を視軸と呼ぶ。一方、眼球の回転中心と瞳孔中心を通る直線を光軸と呼ぶ。光軸と視軸は水平 4 から 5 deg, 垂直約 1.5 deg のオフセットがあり、オフセットには最大 3 度の個人差ある[1]。眼画像から瞳孔中心や光軸を求めることはできるが、中心窩の位置や視軸を画像から直接求めることはできない。

従来の視線計測器の校正は、視線計測を始める前、場合によってはその後にも、基準となるマーカを使用者に提示し、使用者にマーカを注視する動作が課される。様々な視線計測手法が存在するが、それらの共通点は、装置の取り付け位置由来と眼の個人差由来の偏差を含む何らかのパラメータを推定することである。注視しているときの視線方向とマーカとの位置関係からそれらのパラメータを決定することで、視線計測器の出力を正しい視線への変換関数が定まる。また、校正後も視線計測器の位置が変化すると、視線計測の精度が低下し再校正が必要となる場合もある。自動的に視線計測器を校正することができれば、視線計測後に一括解析する用途においては計測時間の短縮につながる。また、視線計測器を実時間利用する用途においては必要に応じて随時校正することが可能となる。

これまで、視線計測の自動校正に関する様々な手法が提案されてきた。自動校正手法は、学習ベース、顕著性ベース、パーシュートベース、シーンモデルベースに分けることができる。学習ベース手法は、頭部姿勢やシーン画像の入力と視線の出力の関係を大量に学習する。そのため、視線計測器は必要ないという利点はある。この手法は、挑戦的であるが精度は低い。顕著性ベース[3][4]、パーシュートベース[4]、シーンモデルベース手法は、そのユーザの視線方向に基づいた校正が行われるため、その問題がない。顕著性ベース手法は、多くの人の視線情報から学習された顕著性マップの極大点を視線方向とする手法であり、視線推定精度は顕著性マップの精度に依存する。現状の顕著性ベース手法の精度は十分高いとは言えず、学習するシーンに依存するという問題もある。パーシュートベース手法は、眼球の動きと視界内の同物体の動きの相関を利用する方法で、容易に任意の三次元シーンに拡張することができ。しかし、同じ動きをする物体が多い場合や、大きな物体のどの部分を注視しているのかを特定することが難しい。シーンモデル手法はシーンの形状が既知であることを前提としている。この前提は、VR アプリケーションと親和性が高い。しかし、多くの手法は、対象シーンとして平面スクリーンか、微分可能な曲面シーンに特化されていた。

そこで本研究では、次の特長を併せ持つ、視線計測器の自動校正手法の開発を目指した。

- ・ マーカを注視する能動的な校正動作は必要ない。
- ・ シーンを自由に見渡すだけの受動的な動作だけで良い。
- ・ マーカやマーカの代わりとなる特定の物体が必要ない。
- ・ シーンにはオクルージョンや様々な形の物体が含まれていても良い。
- ・ 計算量の高い顕著性マップ生成やオプティカルフロー推定、物体検出などの画像処理が必要ない。
- ・ 潜在的に眼球の個人差に由来する偏差を補正することができる。

開発法では頭部位置とシーンモデルが既知であることを前提とする。この前提は VR や AR においては、一般的な前提であり頭部の位置姿勢やシーンモデルは容易に取得することができる。この前提において、物体の表面に分布する固視中の注視点群の 3 次元位置を指標として校正パラメータを推定する。開発手法では、先行研究と同様に、物体上で注視点が一定時間密集する状態を固視と定義する。図 1 に示すように、固視中の注視点は物体表面上に密集して分布するが、視線計測誤差が大きいときには、固視中の注視点群の分散が大きくなることを利用して校正関数のパラメータを最適化する。

3 三次元固視位置を用いた自動較正法

開発した手法は、視線とシーンの3次元モデルの衝突点である3次元注視点から固視を検出し、固視中の注視点の再投影誤差から成る評価関数を最小化することで較正パラメータを推定する。ただし、HMD上で両眼に異なる画像が提示されていても、視線の計算では、両眼の中心である cyclopean eye を表す1台の仮想のシーンカメラ（ピンホールカメラ）を用いる。本手法で前提とする較正モデルは、一般的な三次元眼球モデルであり、 $\mathbf{g}, \tilde{\mathbf{g}}$ をそれぞれ光軸、視軸とすると、較正関数 f は次式で表わされる。

$$\tilde{\mathbf{g}} = f(\mathbf{g}; \boldsymbol{\theta}) = R_x(\beta)R_y(\alpha)\mathbf{g}$$

ただし、 $\boldsymbol{\theta} = (\alpha, \beta)$ は二つの較正パラメータの集合であり、視軸と光軸のオフセット角に対応する。 (α, β) はそれぞれ視線の水平、垂直方向の回転角度、 R_x, R_y はそれぞれ水平、垂直方向の回転行列を表す。

本手法は、固視中における注視点群の分散の和を評価関数として、評価関数を最小化する較正パラメータ $\boldsymbol{\theta}$ を求める問題と考えることができる。したがって、固視中の注視点群の検出、注視点群の分散の算出、評価関数の最適化が手法の肝になる。較正後の視線方向 $\tilde{\mathbf{g}}$ の延長線が、最初に物体表面上に衝突する点を3次元注視点とする。本手法では、既存の固視検出アルゴリズムにより各視線方向が同一固視かどうかラベル付けされる。 j 番目の固視には M_j 本の視線方向 $\mathbf{g}_i (i = 1, \dots, M_j)$ が割り当てられているとし、各集合を固視クラスタと呼ぶ。

固視クラスタの再投影誤差関数は、各固視クラスタに対して代表となる中心カメラ上で定義される。まず、 j 番目の固視クラスタに属する M_j 個のシーンカメラから、一つの固視に含まれる視線のうち、最も平均的な視線をもつシーンカメラを、中心カメラとして選出する。次に、各 \mathbf{g}_i に対応する3次元注視点を中心カメラに投影し二次元注視点 \mathbf{x}_i を得る。最後に、次式で表されるように、 N 個の固視クラスタの再投影誤差関数を同時に最小化することで、最適な較正パラメータ $\tilde{\boldsymbol{\theta}}$ を視軸方向として推定する。

$$\tilde{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \sum_{j=1}^N \frac{1}{M_j} \sum_{i=1}^{M_j} |\bar{\mathbf{x}} - \mathbf{x}_i|^2$$

実験

実験では、VR用のHMDに図2に示すような二種類のシーン提示し、18名の参加者が各シーン内で、約200sの間自由に見まわしながら歩いたときの視線データを取得した。固視検出アルゴリズムには、既存のI-VT法[6]に加え、二次元注視点に対するI-DT法[6]とI-VDT法[7]を用いた。I-VT法は、視線の角速度の閾値に基づいて、I-DT法は注視点の分散の閾値に基づいて固視を検出する。I-VDT法は、I-VT法とI-DT法を組み合わせた手法である。角速度閾値には80 deg/s、分散閾値には0.7 degを使用した。3種類の固視検出アルゴリズムをそれぞれ使用したときの提案手法と、光軸、視軸の精度を確認した。推定した較正パラメータの評価と、視軸に対応するパラメータの推定には、マーカを注視して得た視線データを用いた。図3は、18人の被験者に対する提案手法の絶対誤差を、図中のエラーバーは標準誤差を示す。I-DT法とI-VDT法の絶対誤差は、両シーンにおいて、光軸の絶対誤差がよりも有意($p < 0.05$)に低かった。一方、I-VT法は、スーパーマーケットシーンでは、光軸に対して有意差が確認されなかった。他にも以下の2点について確認した。固視検出に真値のパラメータにより較正された視線の時系列データを入力したのち、同じアルゴリズムにより較正パラメータを最適化しなおした場合、絶対誤差が1.2 degとなったことから、未較正の視線データに対する固視検出手法の精度が重要であることが分かった。また、本予算で購入した光学透過型

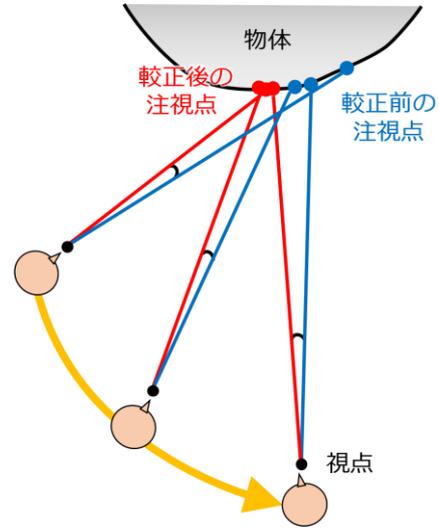


図1：注視点の分散による較正の原理

HMD の機能を精査し、開発手法が実環境にも適用できることを確認した。

まとめ

本研究課題により固視検出を前提とした視線計測器の自動較正手法を開発した。開発した手法は、シーンモデルを用いる手法の一種であり、平面ディスプレイを仮定する多くの従来法の拡張版である。本手法は、大きな頭部移動や、多数のオクルージョンを含む非平面のシーンに対応しており、バーチャルリアリティ用の HMD に搭載された視線検出器の較正に向いている。本手法は、用意に実環境に拡張可能であり、実環境および仮想環境中の任意のシーンに適用可能な視線自動較正手法である。その精度は、従来研究で報告されている任意のシーンに対応したどの手法の精度よりも高かった。開発した手法は、固視検出アルゴリズムによって精度が異なり、I-DT 法、I-VDT 法の絶対誤差に比べ、I-VT 法の絶対誤差が大きいことが確認された。未較正な視線データに対して頑健な固視検出手法について調査することが重要であるという課題も明らかになった。

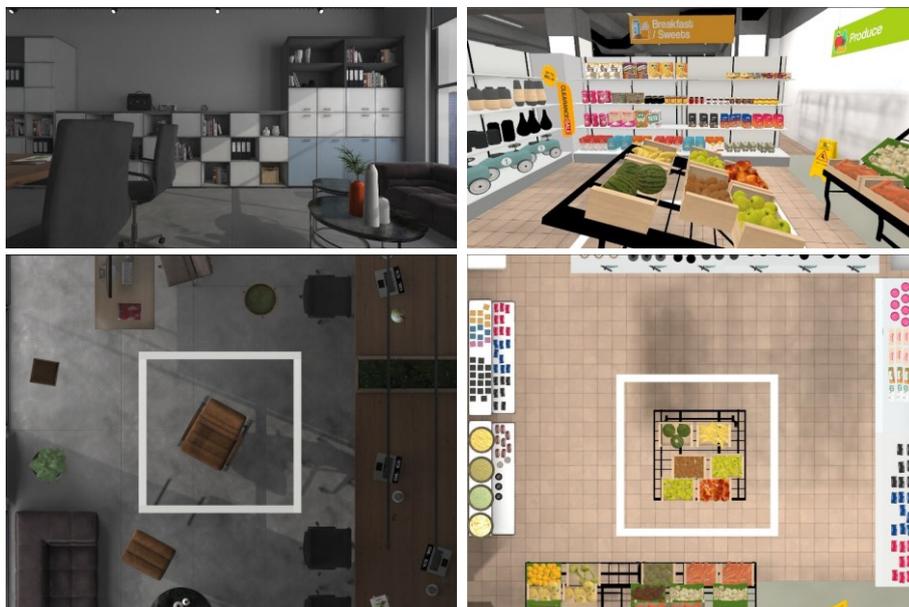


図 2 : VR シーン. (左)オフィス, (右)スーパー.

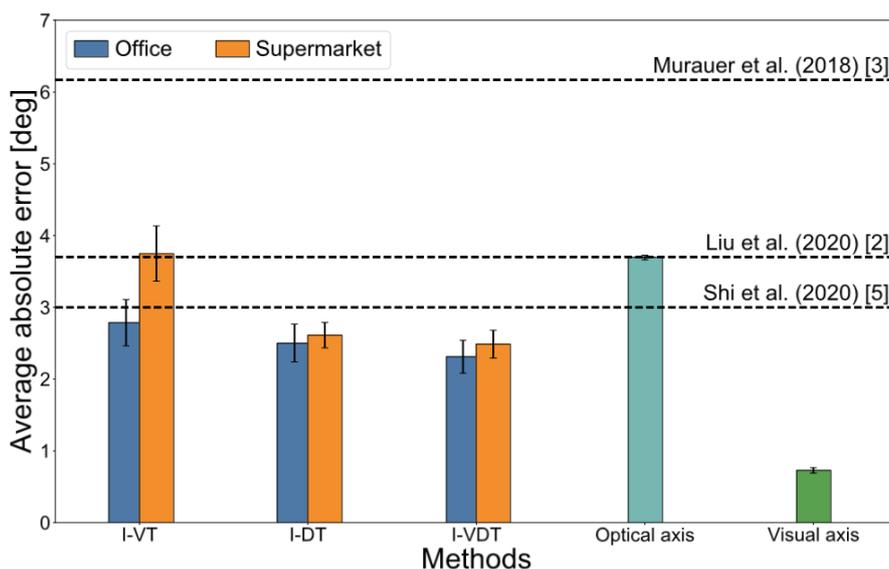


図 3 : 提案手法の絶対誤差

本研究では、固視検出を前提とした視線計測器の自動校正手法を提案した。提案手法は、シーンモデルを用いる手法の一種であり、平面ディスプレイを仮定する多くの従来法の拡張版である。本手法は、大きな頭部移動や、多数のオクルージョンを含む非平面のシーンに対応しており、VR 用や AR 用の HMD に搭載された視線検出器の校正に向いている。提案手法は、主に固視検出とパラメータ最適化から成り、最適化の指標は、同一固視中の注視点の再投影誤差を使用した。VR シーンを用いた実験では、提案手法は固視検出アルゴリズムと固視検出時の校正パラメータに依存することが分かった。一般的な三次元眼球モデルを想定した開発手法の校正パラメータの推定精度が約 2 deg で、任意シーンに対応するどの従来手法よりも高精度であった。また、固視検出アルゴリズムでは約 1 deg まで向上する可能性があることも確認した。

【参考文献】

- [1] R. H. S. Carpenter: “Movements of the Eyes,” 2nd Rev. Pion Limited, 1988.
- [2] D.W. Hansen and J. Qiang: “In the Eye of the Beholder: A Survey of Models for Eyes and Gaze,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, No. 32, Vol. 3, pp. 478–500, 2010.
- [3] M. Liu, Y. Li, and H. Liu: “3D Gaze Estimation for Head-mounted Eye Tracking System with Auto-calibration Method,” *IEEE Access*, Vol. 8, pp. 104207–104215, 2020.
- [4] P. Shi, M. Billeter, and E. Eisemann: “SalientGaze: Saliency-based Gaze Correction in Virtual Reality,” *Computers & Graphics*, Vol. 91, pp. 83–94 (2020).
- [5] M. Murauer, M. Haslgrubler, and A. Ferscha: “Natural Pursuits for Eye Tracker Calibration,” *Proceedings of the 5th International Workshop on Sensor-based Activity Recognition and Interaction*, pp. 1–10, 2018.
- [6] D. D. Salvucci, and J. H. Goldberg: “Identifying Fixations and Saccades in Eye-Tracking Protocols,” *Proceedings of the Symposium on Eye Tracking Research & Applications (ETRA)*, pp. 71–78, 2000.
- [7] O. V. Komogortsev, and A. Karpov: “Automated Classification and Scoring of Smooth Pursuit Eye Movements in the Presence of Fixations and Saccades,” *Behavior Research Methods*, Vol. 45, No. 1, pp. 203–215, 2013.

〈発表資料〉

題名	掲載誌・学会名等	発表年月
運動視差に基づく半透明像注視判定法	システム制御情報学会研究発表講演会	2022年5月
移動中の車両における窓型ディスプレイを用いた半透明提示像の注視判定	Ubiquitous Wearable Workshop	2021年12月
3次元注視点に基づく固視検出評価法	Ubiquitous Wearable Workshop	2021年12月