

移動する人物に対する視線推定 Calibration for Pedestrian's Gaze Estimation

井手 美優[†] 木谷 哲[†] 池西 俊仁[†] 吉田 道学[†]
Miyu Ide Tetsu Kitani Toshihito Ikenishi Michinori Yoshida

1. はじめに

人物の視線推定における視線誤差低減を目的とした校正について、移動している人物に対しては顔や目の識別精度が下がり、視線推定や校正の精度も悪化する問題がある。そこで、視線情報と相関の高い要素を確認し、その要素を用いて視線をモデル化することで、人物が移動するなどの環境下でも視線を校正可能な手法を検討し評価を行った。

2. 目的

人物の視線は行動意図理解のための重要情報の一つであり、様々な場面で活用される。視線推定では一般的に眼球等の動きにより視線方向や注視位置を推定するが、その注視推定位置と実際の注視位置との間には眼球形状の個人差や、視線取得センサと人物との位置関係等に起因する誤差が存在するため、その補正を目的として校正が行われている[1]。例えば、図 1 のように校正用の行列を用いて注視推定位置を正しい位置に射影変換する手法がある。しかし射影変換では、人物とセンサ間の位置関係が既知または変化が小さい場合が多く、位置変化を無視して常に一定の校正行列を用いると誤差が増大する。また、位置変化に対応した既存の校正手法として、動的な校正用行列を学習する手法や条件に応じた校正用行列を大量に用意する手法などがあるが、学習の場合は物理量の関係性が不明瞭であり条件の変更などの応用が困難なこと、学習データ増加の場合はデータ取得コストが非常に大きくなるのが課題となる。

そこで本稿では、人物とセンサ間の位置関係を変化させた場合の視線と注視点との相関の高い要素を確認し、その要素を利用して視線情報を回帰式によりモデル化することで、位置が変化した場合においても視線校正が可能な手法を提案する。

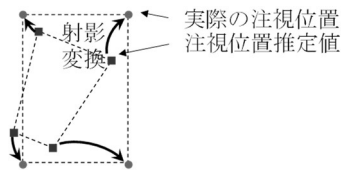
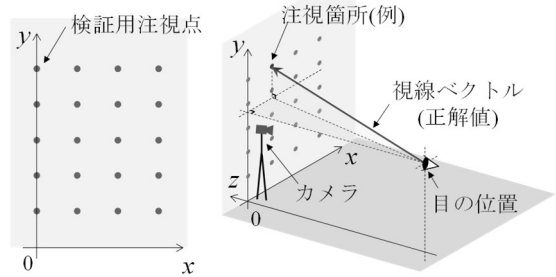


図 1 射影変換による視線校正例

3. 検証方法

図 2 に示す計測環境で実験を行い、提案手法を検証した。図 2(a)のように壁の上に注視用の点を設定し、図 2(b)のように、壁から離れた位置（最大約 4.5m）からその各点を見ているときの人物の様子をカメラで撮影、その画像情報をもとにオープンソースツール OpenFace[2]を用いて視線ベクトルを推定する。また壁と目までの距離を元に視線ベクトルから壁上の注視位置を推定する。視線ベクトルの正解値は目の位置から対象の注視点を結ぶベクトルとし、注視位置や視線ベクトルの正解値と推定値との位置誤差、角度誤差により精度評価を行う。



(a) 壁に設置する注視点 (b) 視線情報取得環境

図 2 検証環境

4. 視線校正方法の提案

4.1 視線校正方法の概要

提案手法の流れを図 3 に示す。従来手法では射影変換の際に用いる射影変換行列を、事前に 1 つまたは複数用いているが、本手法では視線と相関の高い要素を入力としたモデルにより求められた校正用注視位置推定値をもとに、決められた条件下に限らず各時刻でその条件に対応する射影変換行列を作成する。

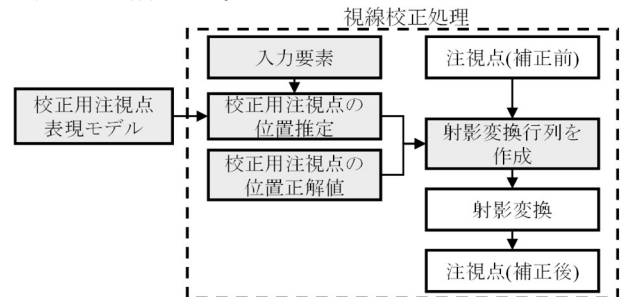


図 3 視線校正処理の提案手法

4.2 視線推定の検証

注視点とカメラから計測した目の位置の関連があると仮定し、3 章に示した方法で壁上の注視点を 1 秒間ずつ見ている際の注視位置推定値を求め、正解値とのずれを検証した。被験者は 20~40 代の男女 3 名 A, B, C とし、事前に実験内容を説明し実験参加への同意を得た。被験者は $(x, y, z) = (20, 0, -450)$ の位置に両足の中心が来るように立つ。カメラはおおよそ被験者の正面となる $(x, y, z) = (5, 156, 0)$ にカメラ①を、被験者から見て右側である $(x, y, z) = (125, 156, 0)$ にカメラ②を設置する。カメラ座標系から見た被験者の左目と右目の中心点の x (左右) 方向位置を x_{obs} , y (上下) 方向位置を y_{obs} とすると、各カメラから見た目の中心位置 (x_{obs}, y_{obs}) は表 1 の通りである。

各注視点を見ていた際の視線ベクトルと壁までの距離をもとに算出した 1 秒間の注視位置推定値を注視点毎にプロットした結果を図 4 に示す。図 4 中の×は壁に設置した実

際の注視対象点の位置、●は注視位置推定値を示し、異なる注視点を見ている時の推定結果を色分けして表示している。図 4 より、視線ベクトルから求まる注視位置の推定値は正解値に対してずれが生じ、そのずれ方は被験者により異なり、同じ被験者でもカメラ位置により変化している事が確認できる。これは、目の中心位置(x_{obs}, y_{obs})により変化する傾向があることを示していると言える。

目の中心位置(x_{obs}, y_{obs})と視線情報との相関を検証するため、この検証で得られた注視位置推定値との相関係数を算出する。射影変換による校正を行うため、注視点のうち四隅にある左下・左上・右上・右下の計 4 点を見ている時の注視位置推定値に着目し、 x_{obs} と y_{obs} の 2 つの要素について四隅の各点 x, y 座標との相関係数を算出し、 $-1 \sim 1$ までの値の絶対値をカラーマップで可視化した結果を図 5 に示す。この結果より、目の中心位置(x_{obs}, y_{obs})と四隅の点の注視位置推定値には相関があることを確認した。

表 1 カメラから見た目の中心位置座標

被験者	カメラ①		カメラ②	
	x_{obs} [cm]	y_{obs} [cm]	x_{obs} [cm]	y_{obs} [cm]
A	-15	-15	105	-15
B	-15	12	105	12
C	-15	4	105	4

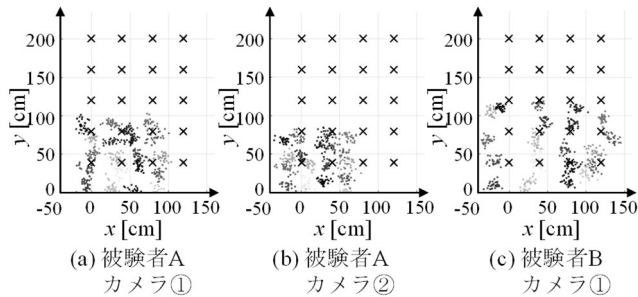


図 4 注視点推定値と正解値の比較

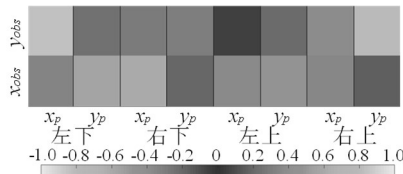


図 5 目の位置と注視点推定値の相関係数

4.3 視線情報表現モデルの定義

4.2 の結果をもとに、射影変換行列に算出する校正用の既知の 4 点(x_i, y_i) ($i = 1, 2, 3, 4$)を見ている際の

注視位置推定値(\bar{x}_i, \bar{y}_i)を、カメラから見た目の中心位置(x_{obs}, y_{obs})を説明変数とした一次式(1)で表される回帰モデルで表現することとした。ここで a, b, c は回帰パラメータである。

$$\begin{cases} \bar{x}_i = a_{1i}x_{obs} + b_{1i}y_{obs} + c_{1i} \\ \bar{y}_i = a_{2i}x_{obs} + b_{2i}y_{obs} + c_{2i} \end{cases} \quad (i = 1, 2, 3, 4) \quad (1)$$

このとき、ある(x_{obs}, y_{obs})に対し、注視位置推定値(\bar{x}_i, \bar{y}_i)を既知の 4 点(x_i, y_i)に補正するための校正用射影変換行列を求めることができる。その射影変換行列を、その x_{obs} と y_{obs} に対して可変な校正用行列 $H(x_{obs}, y_{obs})$ とすることで、未知の位置を見ている際に校正に使用可能な、カメラから見た目の中心位置(x_{obs}, y_{obs})に対応した射影変換行列を求める

ことができる。式(1)で定義される線形回帰モデルの回帰パラメータは、既知の 4 点(x_i, y_i)を見ている際の注視位置推定値(\bar{x}_i, \bar{y}_i)のデータの組み合わせをデータセットとして複数条件で用意し、数値解析ソフトウェア MATLAB 上の `fitlm` 関数により算出した。本稿で用いたデータセットにおける(x_{obs}, y_{obs})の条件は、4.1 の検証と同様のカメラ①、②と目の高さが異なる被験者 A, B, C に対し、立ち位置を(20, 0, -450)、(75, 0, -450)、(130, 0, -450)の 3 通りに増やし、計 $2 \times 3 \times 3 = 18$ 通りとした。

4.4 視線情報表現モデルを用いた校正の評価

4.2 で作成した視線情報表現モデルをもとに、4.2 におけるそれぞれの(x_{obs}, y_{obs})の条件ごとに校正用射影変換行列を算出し、校正を行った。図 4 と同様に、結果の例を図 6 に示す。図 4 と図 6 を比較すると、校正前は条件により異なるずれ方、分布をしていたにもかかわらず校正によりどの条件でもおおむね正しい分布に補正できていることを確認できる。また、各条件での視線ベクトルの x 方向、 y 方向それぞれの角度の正解値との誤差の RMSE と標準偏差を、全データに対して提案手法による校正前・校正後のそれぞれで算出し、表 2 に示す。この結果より、提案手法を用いた校正により x 方向では視線推定誤差の RSME を約 46.8%、 y 方向では約 68.8%低減でき、本提案手法の異なる条件下における視線校正に対して有効であることを確認した。

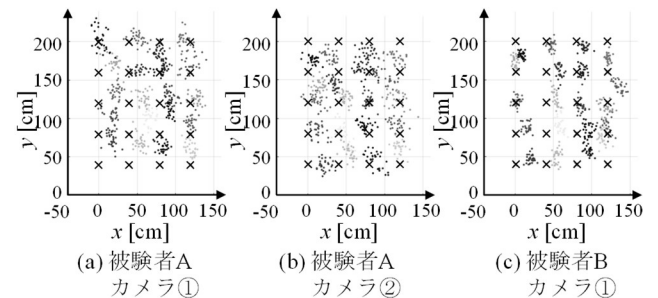


図 6 校正後の注視点推定値と正解値の比較

表 2 校正前後の視線ベクトル角度誤差

評価指標	x 方向誤差 [deg]		y 方向誤差 [deg]	
	校正前	校正後	校正前	校正後
RMSE	3.65	1.94	10.7	2.37
標準偏差	3.34	1.84	3.48	2.37

5. おわりに

本稿では人物が移動しカメラと対象者との位置関係が変化する場合にも適用可能な視線校正手法を提案し、カメラから 4.5 m 離れていても平均角度誤差約 2.1 deg で精度よく校正を行えることを確認した。今後は人の顔向き等、視線推定に有用な他の要素も利用し更なる高精度化を目指す。

参考文献

- [1] Y. Cheng, H. Wang, Y. Bao and F. Lu, "Appearance-Based Gaze Estimation With Deep Learning: A Review and Benchmark," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 46, no. 12 (2024).
- [2] T. Baltrušaitis, A. Zadeh, Y. Chong Lim, and L. Morency, "OpenFace 2.0: Facial Behavior Analysis Toolkit", IEEE International Conference on Automatic Face and Gesture Recognition, (2018).

† 三菱電機株式会社 Mitsubishi Electric