

Inside-Out カメラを用いた単眼による注視点推定手法の性能評価 Performance Evaluation of Gaze Estimation by Monocular using Inside-Out Camera

川田 健司†
Kenji Kawada

齊藤 剛史†
Takeshi Saitoh

1. はじめに

First Person Vision と呼ばれる 1 人称視点映像を用いた画像処理, それに基づく人間行動支援の枠組みが注目されている. そのためのデバイスとして Inside-Out カメラが提案されている. Inside-Out カメラは, ユーザの目の動きを捉える Eye カメラと視界を撮影する Scene カメラから構成される. 清水らは, 2 台の Eye カメラと 2 台の Scene カメラを搭載した頭部に固定するゴーグル型の Inside-Out カメラを開発し, 3 次元注視点の推定法を提案している[1]. しかしゴーグル型の Inside-Out カメラは, ユーザに負担がかかる. Devyver らは Eye カメラ 1 台と Scene カメラ 1 台の他に DaVinci プロセッサなどから構成され, メガネ等にカメラを搭載するグラス型の Inside-Out カメラを提案している[2]. 本研究でも Eye カメラ 1 台と Scene カメラ 1 台を搭載したグラス型の Inside-Out カメラを開発し, 単眼によるユーザの注視点推定手法を提案している. さらに特殊な器具を必要としない指先注視キャリブレーション手法を提案し, 注視点の推定を実現した[3]. また, 瞳孔中心検出に観測領域 ROI を自動的に設定することで処理コストを削減し, リアルタイム処理を実現した. しかし, 目と各カメラの光軸のずれから注視点推定精度が低下する問題および注視対象距離に対する検証が不十分であった.

そこで本稿では, カメラの位置を修正することで推定精度の向上を図る. さらに注視対象との距離変化による評価実験を通じて Inside-Out カメラを用いた単眼による注視点推定手法の性能を検証する.

2. Inside-Out カメラ

本研究では, 3D プリンタで製作した眼鏡フレームに脱着可能で小型・軽量な筐体に, カメラ 2 台を搭載する Inside-Out カメラを開発した. Eye カメラの周囲に赤外線 LED を取り付け, 暗瞳孔画像を取得する. 赤外光は不可視光線であるため, ユーザに視覚的刺激を与えることはない. Scene カメラは人の視界を撮影するため, 広角視野のカメラを利用する. また, ユーザへの負担を減らすため, 小型なカメラを用いる. 重量は 90g (コードを含む) である. 図 1 に開発した Inside-Out カメラを示す.

3. 注視点推定システム

本研究における注視点推定システムは, Eye カメラの画像内における瞳孔中心点と Scene カメラの画像内における注視点を線形近似によって対応付ける. キャリブレーション処理によって対応付けるパラメータを求め, このパラメータを用いて注視点を推定する.

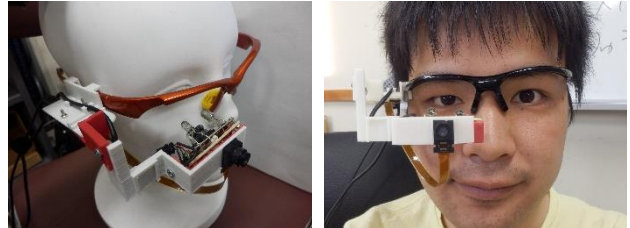


図 1 開発した Inside-Out カメラ

3.1 瞳孔中心検出

本システムでは Eye カメラの画像から瞳孔の中心位置を検出する. ここで従来手法[3]では画像全体から瞳孔中心位置を検出していた. この処理コストを軽減するため, システム起動時に観測領域 ROI を自動的に設定する. 具体的にはユーザの瞬き動作にフレーム間差分を適用して目領域を ROI として検出する. 次に ROI 内に対して組み合わせ分離度フィルタを適用し, 仮の瞳孔中心位置を検出する. その後, 仮の瞳孔中心位置から 8 方向の周辺領域に分離度フィルタを適用することで瞳孔のエッジ点を検出し, そのエッジ点を利用して近似楕円を推定する. 求めた近似楕円の中心を瞳孔の中心位置 (x_{eye}, y_{eye}) とする.

3.2 キャリブレーション

本システムのキャリブレーション動作では, ユーザは指先を注視しながら一筆書きのような文字 (指先文字) を描く. このとき, Eye カメラの画像内から瞳孔の中心位置 (x_{eye}, y_{eye}) と, Scene カメラの画像内から指先の位置 (x_{scene}, y_{scene}) を検出する. 各フレームにおいて両方の検出に成功した場合, x, y の各座標系に対して RANSAC アルゴリズムにより 1 次式を推定し, 変換式のパラメータ $(a_x, b_x), (a_y, b_y)$ を算出する.

3.3 注視点推定

キャリブレーション処理で得られたパラメータを用いて注視点 (x_{gaze}, y_{gaze}) を推定する. 目の観測領域 ROI の左上座標を (x_{left}, y_{top}) とする. Eye カメラの画像中の ROI から瞳孔の中心位置 (x_{eye}, y_{eye}) を検出した後, 下式により注視点 (x_{gaze}, y_{gaze}) を算出する.

$$\begin{aligned} x_{gaze} &= a_x(x_{eye} + x_{left}) + b_x \\ y_{gaze} &= a_y(y_{eye} + y_{top}) + b_y \end{aligned}$$

図 2 に注視点推定の結果を示す. 図 2(a)は Eye カメラ画像から瞳孔中心位置を検出した結果, 図 2(b)は変換式により推定された注視点を強調表示した Scene カメラ画像である. ただし, 図 2(a)の矩形は ROI を示す.

†九州工業大学 Kyushu Institute of Technology

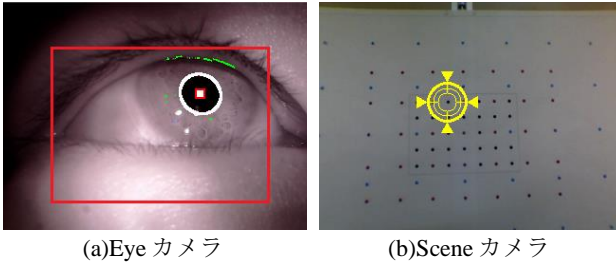


図 2 注視点推定

4. 評価実験

4.1 カメラ位置に対する実験

目と Eye カメラ, Scene カメラの光軸の位置がずれると注視点推定精度が低下する. このことを検証するため, Scene カメラの位置を目の光軸とずれている場合 (条件 1) と水平方向の光軸が一致した場合 (条件 2) において注視点推定実験を実施した. 実験では被験者男性 5 名 (S01~S05, S03 のみコンタクト装着) に対して Scene カメラから 50cm 離れた壁にある複数のランドマークを順に注視したときの推定された注視点と真値との角度誤差 E_{θ} [°] を算出した. Scene カメラの位置は条件 1 の場合, 右目に対して前に約 3cm, 右に約 4cm, 下に約 2cm, 条件 2 の場合, 左右のずれを限りなく 0 にし, 前に約 5cm, 下に約 2cm である. ランドマークは縦方向に各 5 点, 横方向に各 7 点の全 35 点を 5cm 間隔に描画した. キャリブレーションのための指先動作は 8 の字とした. 各被験者の角度誤差を表 1 に示す.

全被験者において条件 2 の方が条件 1 より精度が向上し, 平均で 0.76° 精度が向上した. 人間は注視している点から 2° 程度が明瞭に見えていることを考慮すると十分な精度である.

4.2 注視距離に対する実験

提案システムを評価するために複数の注視距離において注視点推定実験を実施した. 実験では, 被験者男女 9 名 (S01~S09, S03 はコンタクト装着, S07 は眼鏡装着, S08 と S09 は女性) に対してランドマークを順に注視したときの推定された注視点と真値との角度誤差 E_{θ} [°] を算出した. Scene カメラからマーカまでの距離は 50cm, 100cm, 150cm の 3 通りとした. 視野角を揃えるため, マーカの間隔は距離ごとに 5cm, 10cm, 15cm とした. Scene カメラの位置は 4.1 の条件 2 であり, その他の実験条件は 4.1 と同様である. 各被験者, 各注視距離の角度誤差を表 2 に示す.

注視対象との距離が 50cm の場合, 被験者 9 名の平均角度誤差は 1.43° であった. 4.1 と同様に提案システムは注視距離 50cm の場合, 複数の被験者に対しても十分な精度を得られた. しかし, 注視対象が離れるとカメラの光軸のずれの影響で誤差は大きくなる. 本実験での光軸の影響による理論的な誤差は Scene カメラからキャリブレーション時の指先の距離が 50cm とすると, カメラ注視対象との距離 100cm で約 1.2°, 150cm で約 1.7° であり, 注視対象との距離が既知であれば変換関数に補正をかけることで十分な精度を得られる可能性がある.

提案システムはキャリブレーション時の指先の距離 (対応点取得距離) に依存していることは明白であるが, 距離に対して単調増加で誤差が大きくなることを考慮すると指先より 20cm 程度先までであれば十分な精度で推定できる.

表 1 Scene カメラ位置の精度比較

被験者	角度誤差 E_{θ} [°]	
	条件 1	条件 2
S01	2.32	1.24
S02	1.32	0.79
S03	2.42	1.67
S04	2.13	1.26
S05	1.32	0.73
Ave.	1.90	1.14

表 2 注視対象との距離変化に対する精度

被験者	角度誤差 E_{θ} [°]		
	50cm	100cm	150cm
S01	1.24	2.83	2.94
S02	0.79	1.58	1.57
S03	1.67	2.58	3.08
S04	1.26	1.96	2.23
S05	0.73	2.95	3.43
S06	1.10	2.82	6.35
S07	1.62	2.42	5.36
S08	2.19	4.15	5.79
S09	2.29	4.82	5.80
Ave.	1.43	2.82	4.06

今回の実験で注視点推定にかかる 1 フレームあたりの平均処理時間は, デスクトップ PC (CPU: Intel Core i7-4770, 3.4GHz) を使用し, 29.8ms であった.

5. おわりに

本稿では, 目とカメラとの光軸のずれを減らすためカメラの位置を修正し, 注視点の推定精度向上を図った. また, 提案システムの性能を検証するため注視対象との距離変化に対する評価実験を行った. その結果, カメラ位置に対する実験により, 角度誤差は光軸がずれている場合より約 0.8° 精度が向上した. また, 注視対象との距離変化による評価実験を通して, 人間は注視している点から 2° 程度が明瞭に見えていることを考慮すると, 提案システムはキャリブレーション時の指先の距離より 20cm 程度先までであれば十分な精度で推定できる.

謝辞

本研究の一部は, 日本学術振興会科学研究費補助金によるものである.

参考文献

- [1] 清水彰一, 西尾和晃, 木村誠, 藤吉弘亘, “First Person Vision のための Inside-Out カメラの提案”, 信学論, vol.J94-D, no.11, pp.1909-1918, 2011.
- [2] Michael S. Devyver, Akihiro Tsukada, and Takeo Kanade, “A Wearable Device for First Person Vision (FICCDAT workshop),” 3rd International Symposium on Quality of Life Technology, July, 2011.
- [3] Junki Iwagami and Takeshi Saitoh, “Easy Calibration for Gaze Estimation using Inside-Out Camera,” 20th Korea-Japan Joint Workshop on Frontiers of Computer Vision, 57, pp.292-297, 2014.2.