

視覚的顕著性に基づいたアイトラッカーの自動校正法

沖野 祐介† 中澤 篤志† 西田 豊明†

† 京都大学大学院情報学研究科

1 はじめに

視線情報はマーケティング・心理学・ユーザーインターフェース・コミュニケーション解析などの様々な分野の分析に利用されており、商用システム（アイトラッカー）として多くの製品が市販されている。特に、モバイル用途としてはヘッドマウント式のシステムが多く用いられている。これは、頭部に装着するヘッドセットに、眼球を計測するアイカメラと前方のシーンを計測するシーンカメラが取り付けられており、アイカメラで眼球の（アイカメラに対する）角度を計測し、これをシーンカメラ画像の座標系にマッピングすることで注視点を推定する仕組みである（Pupil Center Corneal Reflection (PCCR 法) 図1) [1]。本手法は広く用いられている完成度の高い方法であるが、眼球の角度（アイカメラに対する眼球の姿勢）をシーンカメラ画像にマッピングするために、事前のキャリブレーションが必要である。これは、環境中の数点を注視することで行うものであり、ユーザーにとっては煩雑な手順である。また、装置がズレてしまった場合でも再度のキャリブレーションが必要のため、特にウェアラブルグラスシステムなどで用いるには問題となる。

この問題を解決するため、明示的なキャリブレーションを行わず、注視点計測に必要なパラメータを得ることを考える。我々は、人の自然な注視行動はシーン画像の顕著性の高い領域に分布するという性質を利用し、人の眼球の動きとシーン画像の顕著性マップからキャリブレーションを行う事を考えた。本手法は、人の自然な注視行動を一定時間計測することによって自動的にパラメータ計算ができるため、明示的なキャリブレーションを必要とせずユーザビリティが高いと考えられる。

2 アイトラッカーのキャリブレーション

前述したように PCCR 法によるアイトラッカーのキャリブレーションとは、アイカメラから得られた眼球の姿勢とシーンカメラの画像座標を対応づける処理であり、様々な手法（マッピング関数）が存在する。我々は、カメラと眼球の相対距離は、両者とシーンとの距離に比

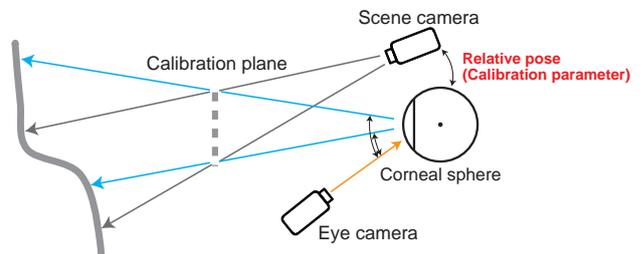


図1: Pupil Center Corneal Reflection (PCCR 法) の原理

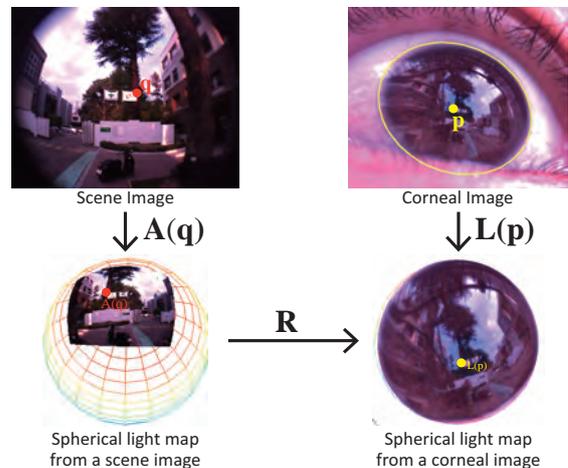


図2: シーン画像とアイカメラとの関係。Far scene assumption を導入することで、両者は鏡面座標系の回転 R で対応づけられる。

べて十分小さいことを仮定する (far scene assumption)。これにより、両者は同一の3次元光線環境を共有すると見なすことが出来る (図2)。この仮定を置くと、アイカメラ（眼球座標系）とシーンカメラの相対運動は、3次元の回転運動 (R) のみで表すことが可能である。すなわち、両画像間で3点以上の対応点が得られれば求める事が出来る。

3 視覚的顕著性

視覚的顕著性 (visual saliency) とは、画像中で注目しやすき点を、画像の特徴から推定しようとするアイデアである。視覚的顕著性の高さを輝度で表した画像を顕著性マップ (saliency map) と呼ぶ。すなわち、この値が高い点ほど、より注目されやすいことを表す。顕著性マップの推定は様々な手法が提案されているが、本

Calibration-free parameter estimations of eye tracking systems using visual saliency
 †Yusuke OKINO †Atsushi NAKAZAWA †Toyoaki NISHIDA
 †Graduate school of Informatics, Kyoto University.

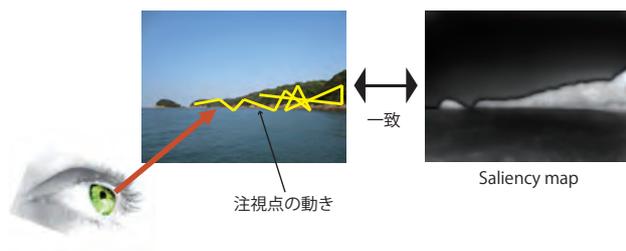


図3: 提案するキャリブレーションパラメータの推定法. 注視点の軌跡における視覚的顕著性の総和が大きくなることを利用して推定を行う.

研究では, Itti らによって提案された手法を用いた [2].

4 視覚的顕著性によるキャリブレーションパラメータ推定法

前章で述べた通り, 視覚的顕著性はシーン中の誘目性の高さを表しているため, 自然な注視行動を取った場合, 一定時間内の注視点の軌跡の視覚的顕著性の総和は高くなると考えられる (図3). これにより, 前述したパラメータ \mathbf{R} を求める. 推定に用いるフレームを $t = 1, \dots, T$, アイカメラ座標系における時刻 t での眼球光軸方向を $\mathbf{v}(t)$, シーンカメラ座標系における3次元光線から画像面への投影を \mathbf{L}^{-1} , 時刻 t におけるシーン画像の顕著性マップを \mathbf{S}_t とすると, 本問題は下記の最適化問題に帰着できる.

$$\mathbf{R}_{opt} = \arg \max_{\mathbf{R}} \sum_{t=1}^T \mathbf{S}_t(\mathbf{L}^{-1}(\mathbf{R}^{-1}\mathbf{v}(t))) \quad (1)$$

ここで, \mathbf{v}, \mathbf{L} は我々の既存手法 [3] を用いて求めている.

5 実験結果

提案手法の有効性を確認するため, まず屋内環境で仮想状況を生成し実験を行った. 屋内において白い紙を3カ所に張ったシーンを準備し, ユーザーに自由な注視行動を行わせた. この状況を角膜イメージングカメラで撮影し, 我々の従来手法を用いて角膜輪郭を検出, 眼球の光軸の動きを求めた. シーンと顕著性マップを図5(a)(b)に示す.

眼球光軸の軌跡およびサリエンス画像を3次元球面上にマッピングした結果を図5(c),(d)に示す. ここから, brute force 的に最適な回転行列 \mathbf{R} を求め位置合わせした結果を図5(e)に示す. 眼球光軸のマッピング点の中には外れ値も多いが, 正解点がサリエンス画像と良く一致しているのも確認できる.

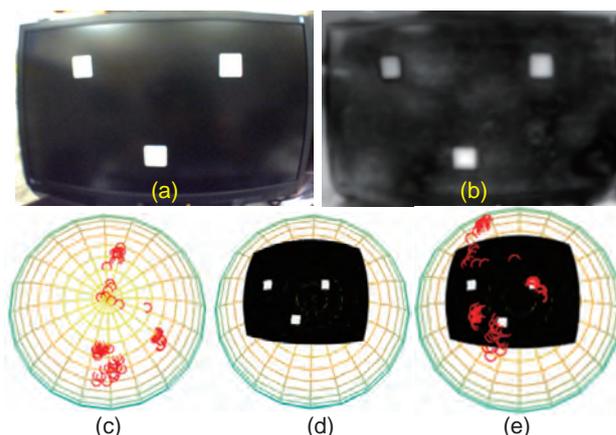


図4: (a) 実験シーンおよび (b) 視覚的顕著性. (c) 眼球光軸 (注視方向) および (d) シーン画像の視覚的顕著性の3次元球面座標系への投影結果, (e)(c) および (d) のマッチング結果.

6 おわりに

本研究では, アイトラッカーのアイカメラ・シーンカメラ間の幾何校正を, マーカー注視等の明示的なキャリブレーション手順を経ず, 自然な注視行動のみから行う手法を提案した. 人の注視点はシーンの顕著性マップと対応するという知識を用い, 一定時間内の注視点の軌跡が高い顕著性を示す部分に偏在するとした. 人工的に設定した屋内環境で実験を行い正しい解が得られることを確認したが, 今後実屋外環境での実験を進め有効性を確認する予定である.

謝辞 本研究は科学技術振興機構さきがけ, 文部科学省科学研究費助成事業研究課題番号 26270058, コミュニルタ科学技術振興財団, 科学技術振興機構・研究成果展開事業「センター・オブ・イノベーション (COI) プログラム」の支援を受けた.

参考文献

- [1] Hansen, D. W. and Ji, Q.: In the eye of the beholder: A survey of models for eyes and gaze, *IEEE Trans. on PAMI*, Vol. 32, No. 3, pp. 478–500 (2010).
- [2] Itti, L., Koch, C. and Niebur, E.: A Model of Saliency-Based Visual Attention for Rapid Scene Analysis, *IEEE Trans. on PAMI*, Vol. 20, No. 11, pp. 1254–1259 (1998).
- [3] Nakazawa, A., Nitschke, C. and Nishida, T.: Random resample consensus 法を用いた角膜表面反射とシーン画像の位置合わせ, 第17回画像の認識・理解シンポジウム (MIRU2014) (2014).