

# 角膜表面反射と全天球画像を用いた全周型注視点推定

小川 太士<sup>1,a)</sup> 中澤 篤志<sup>1,b)</sup> 西田 豊明<sup>1,c)</sup>

**概要:** 本研究では、現実空間に対して注視点推定を行うべく、角膜表面反射画像と全天球画像を用いて、今の自分の位置から 360 度、全方向を対象に注視点推定を行う手法を提案する。注視点推定を行うために、全天球画像を複数方向の透視投影に変換し、角膜表面反射画像と各透視投影の位置合わせを行い、適切な透視投影と、座標変換関数を出力する。実験を通して、特徴物が写っている鮮明な角膜表面反射画像について、全周型注視点推定の有用性を確かめた一方で、全体的に角膜表面反射画像の鮮明さ、シーン画像の特徴量の多さに精度が大きく依存することも示唆された。本研究のように、注視点推定の範囲を拡大することができれば、3次元注視点マッピングなどへの応用可能性が期待される。

## Point of Gaze Estimation in All Directions Using Corneal Surface Reflection and Celestial Sphere Image

TAISHI OGAWA<sup>1,a)</sup> ATSUSHI NAKAZAWA<sup>1,b)</sup> TOYOAKI NISHIDA<sup>1,c)</sup>

### 1. はじめに

生体情報のなかでも、特に視線には多くの情報が含まれている。視線を向けること、さらには逸らすことにも、意図や心理が働いている場合が多い。例えば我々がある対象を注視したとき、その対象に興味がある、注意しているといった情報が読み取れる。このように、視線の滞在時間、および視線を向ける頻度は、人間の興味と直結していると考えられ、商業的、学術的に利用価値が高い。これらを記録し、検証する有用な手段として、ヒートマップがある。視線の滞在時間を可視化したヒートマップを生成する方法の1つとして、生成の媒体となるシーン画像、および注視点推定の技術を用いることが挙げられる。2次元平面画像に対しての注視点推定は、アイトラッキングなどの技術を用いて、現在多方面において開発、使用されているが、推定できる範囲はその画像上のみである。そこで我々は、空間上に対して注視点推定を行うべく、シーン画像として全天球画像を使用することで、今の自分の位置から 360 度、

すべての方向に対して注視点推定を行う手法を検討する。また、これを用いて、全天球画像に注視点のデータを集めることで、全天球画像上に注視点のヒートマップを生成することができる。これを注視点マップと呼ぶことにする。全天球画像を用いることにより、従来の2次元平面上の視線推定にはない、全周型の注視点推定が可能になる。今自分のいる位置から全ての方向に対して、注視点推定ができ、注視点マップが生成できるという点が今回の研究の大きな意義である。これを用いることにより、視線の解析の幅が広がり、利用価値がさらに高まることが期待される。また、近年研究が進んでいる角膜表面反射、その画像のみを用いて注視点推定を行うことから、有用性があると考えられる。

### 2. 関連研究

[虹彩研究] 本研究は角膜の表面反射を用いるため、角膜を扱う、とりわけ虹彩の研究から様々な知見を得ることができる。Bowyerら[2]は、目の虹彩バイオメトリクスの研究は、画像認識、虹彩分離、目のテクスチャ(表面の模様)の分析、テクスチャのマッチングの4つに大きく分類されることを述べ、この分野を新しく研究しようとしている人向けに、虹彩バイオメトリクスの現状とその重要性、応用

<sup>1</sup> 京都大学大学院情報学研究科  
Graduate school of Informatics, Kyoto University

a) ogawa@ii.ist.i.kyoto-u.ac.jp

b) nakazawa.atsushi@i.kyoto-u.ac.jp

c) nishida@i.kyoto-u.ac.jp

先を整理したうえで、虹彩バイオメトリクスに関する重要な文献をいくつか紹介している。Maら[5]は、空間フィルターを利用した新たな虹彩パターンによる個人認識手法を提案し、従来手法との比較実験を通して高い成果を残した。

**[角膜表面反射]** 角膜表面反射に関する研究として、Nishinoら[9]は、照度変化中における顔認識において、1人の被験者に対して、複数のサンプル画像が必要だった従来手法とは異なり、1枚の角膜表面反射画像のみを用いて、光の反射を逆投影することでシーンの光源環境を復元し、顔認識において高いスコアを残すことに成功した。また、Nishinoら[10]は角膜表面反射を利用して今見ているシーンを復元する、角膜イメージング法を提案した。そして、Nakazawaraら[6]は、角膜表面反射から眼球のモデルを構築し、画像の位置合わせ技術を組み合わせることで、角膜表面反射画像を用いて、シーン画像に対して注視点の推定を行うことができることを示した。Nitschkeら[11]は、幾何学的眼球モデルを構築することで、角膜表面反射からのシーン環境、および観測者とカメラ、シーンとの関係の復元方法について再検証し、角膜表面反射分野の包括的な説明と、その応用可能性について触れている。

**[SURF 特徴量]** 一般的に、2つの画像のマッチングを行う際には、2つの画像の共通点(対応点)を探す必要がある。そのためのアプローチとして、画像中から様々な特徴量を抜き出し、それらを比較して対応点を決定する手法がある。この指標となる特徴量の1つがSURFである。

SURF(Speeded Up Robust Features)は、Bayaらによって考案された特徴量である[1]。SURFは、Loweによって考案されたSIFT[4]の高速化版であり、画像認識において多く用いられている特徴量である。SURFは画像中のハイコントラストな点を検出し、検出した局所領域の勾配情報に対して主成分分析を適用し、特徴量ベクトルを形成する。SURFは特徴点について、オリエンテーションベクトルを持つため、向きと大きさが正規化でき、回転、スケール変化に対して不変である。本研究では、2画像間の位置合わせ精度を高めるため、SURFを利用することで、2画像間の変換行列を求めるために必要な対応点の数を少なくすると同時に、後述のRANSACを適用する対象となる点を抽出している。

**[画像の位置合わせ技術]** 本研究では、画像の位置合わせ技術に、Fischlerらによって発表されたRANSAC[3]を用いている。RANSACはロバストなパラメータ推定であり、画像のマッチングをはじめとして、幅広い分野に用いられている。RANSACは、ランダムで取り出したデータからパラメータを求め、それを他のデータにも適用して誤差、あるいは一致度を求めることで、最も適するパラメータを探す手法である。

本研究は、角膜表面反射を用いて眼球姿勢を推定するとともに、角膜表面反射とシーン画像の位置合わせを行う際



図1 角膜表面反射画像および、角膜イメージングカメラ



図2 全地球画像

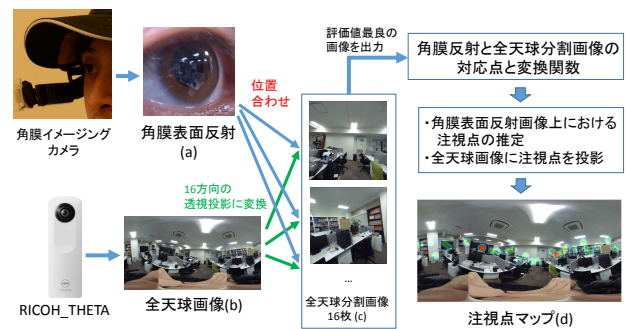


図3 注視点マップ生成の流れ

に、RANSACをSURF特徴量の検出点に対して適用し、画像の位置合わせを試みる。

### 3. 注視点マップの作成法

本研究の目的は、周囲360度を対象に注視点を推定し、注視点マップを生成することである。これを実現するため、注視点を推定する対象となる、複数フレームの角膜表面反射画像、および注視点マップを生成する媒体となる、1枚の全地球画像を用いる。図1は角膜表面反射画像のうちの1フレームと、それを撮影する角膜イメージングカメラであり、図2は全地球画像の一例である。全地球画像は、リコー社の製品、RICOH\_THETAを用いて撮影する。この画像は、上下90度、周囲360度の画像を1枚の画像上に収める全地球カメラである。本研究は、これに対して注視点推定を行うことを目標とする。注視点推定は図3の手順で行う。

#### (準備) 全地球画像の透視投影への変換

全地球画像はその特性上、我々が実際に見る風景と異なるため、角膜表面反射画像との位置合わせはそのままの状態では困難である。そこで、全地球画像(図

3(b) を一定角度の画角を持つ 16 方向の透視投影 (図 3(c)) に変換することにより, 位置合わせを行いやすくする. 以降, 変換後の画像を全天球分割画像と呼ぶ.

#### 眼球姿勢の推定

角膜表面反射画像から角膜の輪郭を取得する. このとき, 輪郭の形から眼球のモデルを導くことで, 眼球姿勢の推定を行う.

#### 角膜表面反射画像と全天球分割画像の位置合わせ

1point-RANSAC アルゴリズムを用いて, 角膜表面反射画像 (図 3(a)) と全天球分割画像の位置合わせを行う.

#### 全天球分割画像に対する注視点の推定

2つの画像の位置合わせの結果, および推定された眼球姿勢を用いることで, 全天球分割画像に対する注視点を導く.

#### 全天球画像に対する注視点の推定

導いた注視点を, もとの全天球画像の座標に変換する.

この過程を複数の角膜表面反射画像に対して行うことで, 全天球画像に対する注視点のデータが蓄積される. これを用いることで, 注視点マップ (図 3(d)) を生成することができる. 4章では, 各手順の詳細について説明する.

## 4. 注視点マップ作成システム

### 4.1 全天球画像の透視投影への変換

全天球画像 (図 2) は,  $x$  軸が水平方向の角度  $\phi$ ,  $y$  軸が鉛直方向の角度  $\theta$  に対応している.  $\phi$  は  $0 \leq \theta \leq 2\pi$ ,  $\theta$  は上下 90 度を映す全天球画像と対応するように,  $-\pi/2 \leq \theta \leq \pi/2$  とする. 全周型注視点推定を行うにあたって,  $x$  軸の全て, および  $y$  軸の  $-\pi/4$  から  $\pi/4$  を使用する.

生成しようとしている全天球分割画像を, 大きさ  $N * N$  ピクセルの 2次元平面画像とする. ここで画像の中央, 座標  $(N/2, N/2)$  を中心とし, そこから全天球画像の左右  $\pi/4$ , 上下  $\pi/4$  を範囲とする画素を分割画像に書き込んでいく.

全天球画像と全天球分割画像の関係を簡潔に示したものが, 図 4 である. いま, ある全天球分割画像  $\mathbf{I}_n$  の中心点を全天球画像における  $\theta = 0, \phi' = 0$  とし,  $\mathbf{I}_n$  の左上端を  $\mathbf{I}_n(0, 0)$  とすれば

$$\mathbf{I}_n(x, y) = (N(1 - \tan \phi')/2, N(1 - \tan \theta)/2) \quad (1)$$

$$(-\pi/4 \leq \phi', \theta \leq \pi/4)$$

が成り立つ.

全天球画像の鉛直方向の中央, つまり  $\theta = 0$  の部分に等間隔で置かれた 16 点それぞれを  $\mathbf{I}_n$  ( $1 \leq n \leq 16$ ) の中心点とし, 各  $\mathbf{I}_n$  について, 全天球画像の各画素に対して式 (2) を適用する. このようにして, 全天球分割画像 16 枚 (図 5) が生成される.



図 4 2 画像間の関係

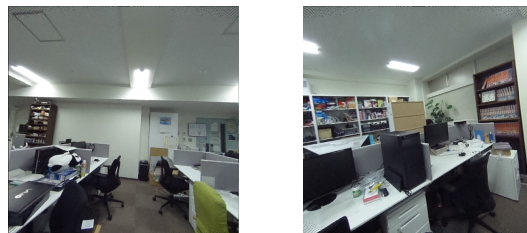


図 5 全天球画像から生成された全天球分割画像. これは周囲 360 度にわたり, 16 枚生成される.

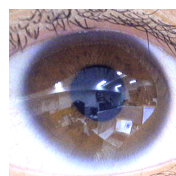


図 6 眼球領域

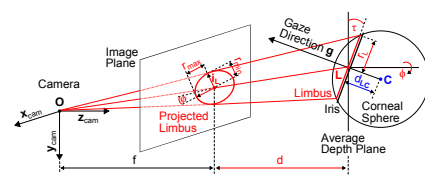


図 7 構築された眼球モデル [6]

### 4.2 眼球姿勢の推定

まず, 角膜表面反射画像から得られた眼球領域の画像 (図 6) から, カメラに対する眼球の 3次元位置を定める. 本節は従来法 [10],[6] と同様の手法を用いている.

角膜を球面とみなし, 瞳の画像上での中心位置を  $\mathbf{i}_L$ , 長軸, 短軸を  $r_{\max}, r_{\min}$ , 回転角  $\phi$  とすると, 角膜球面の 3次元位置は

$$d = r_L \cdot f / r_{\max} \quad (2)$$

ただし,  $\mathbf{L} = d\mathbf{K}_e^{-1}\mathbf{i}_L$  であり,  $\mathbf{K}_e$  はアイカメラの  $3 \times 3$  内部行列,  $f$  は焦点距離を示す. (図 7) また, Gaze Direction (角膜中心の表面と垂直な方向)  $\mathbf{g}$  は,

$$\mathbf{g} = [\sin \tau \sin \phi \quad -\sin \tau \cos \phi \quad -\cos \tau]^T \quad (3)$$

である. ただし,  $\tau = \pm \arccos(r_{\min}/r_{\max})$  のうち, 角膜輪部の傾きに一致する方である.

### 4.3 角膜表面反射画像と全天球分割画像の位置合わせ候補抽出

4.2 で, 眼球姿勢の推定を行ったので, 次は角膜表面反射画像と全天球分割画像の位置合わせを行う. なお, 本節は従来法 [6] と同様の手法を用いている.

角膜表面反射画像と, 全天球分割画像の位置合わせを行うには, 2つの画像を関係づけることができるような座標系を経由しなければならない. そこでいま, 全天球画像の撮影座標と, 眼球の位置が十分に近いことを仮定する. すると, 全天球分割画像と, 角膜表面反射画像は, 同一の

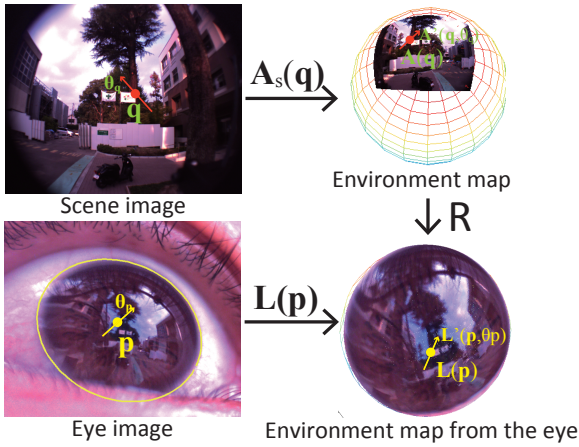


図 8 角膜表面反射画像と平面画像が共有する EM[7]

Environmental Map (EM) を共有する。(図 8)

角膜表面反射画像の座標  $\mathbf{p}$  における、角膜の表面に入射する光の式  $\mathbf{L}(\mathbf{p})$  を求めたい。すなわち、 $\mathbf{L}(\mathbf{p})$  は角膜表面の 3次元座標  $\mathbf{P}$  からシーン画像へのベクトルであるとする。

まず、アイカメラ中の点  $\mathbf{p}$  に入射する、正規化された逆投影ベクトル  $\mathbf{A}_e(\mathbf{p})$  は、

$$\mathbf{A}_e(\mathbf{p}) = \mathbf{K}_e^{-1} \mathbf{p} / \|\mathbf{K}_e^{-1} \mathbf{p}\| \quad (4)$$

である。 $\mathbf{K}_e$  はアイカメラの  $3 \times 3$  内部行列である。ここで、 $\mathbf{P}$  とアイカメラの距離を  $t_1$  とすれば、 $\mathbf{P} = t_1 \mathbf{A}_e(\mathbf{p})$  の関係が成り立つ。角膜球の半径を  $r_C$  とすれば、角膜球について、式  $\|\mathbf{P} - \mathbf{C}\|^2 = r_C^2$  が成り立つ。これを用いると、

$$t_1^2 \|\mathbf{A}_e(\mathbf{p})\|^2 - 2t_1 (\mathbf{A}_e(\mathbf{p}) \cdot \mathbf{C}) + \|\mathbf{C}\|^2 - r_C^2 = 0 \quad (5)$$

ただし、 $\|\mathbf{A}_e(\mathbf{p})\|^2 = 1$  である。これらより  $t_1$ ,  $\mathbf{P}$  を求めると、 $\mathbf{L}(\mathbf{p})$  は、

$$\mathbf{L}(\mathbf{p}) = 2(-\mathbf{A}_e(\mathbf{p}) \cdot \mathbf{N}(\mathbf{p}))\mathbf{N}(\mathbf{p}) + \mathbf{A}_e(\mathbf{p}) \quad (6)$$

$$\mathbf{N}(\mathbf{p}) = (\mathbf{P} - \mathbf{C}) / \|\mathbf{P} - \mathbf{C}\| \quad (7)$$

となる。ただし  $\mathbf{N}(\mathbf{p})$  は  $\mathbf{P}$  における角膜球の法線ベクトルである。

先ほど、全天球分割画像と、角膜表面反射画像は同一の Environmental Map (EM) を共有すると仮定した。このとき、2画像間の位置合わせ問題は、両画像から得られた EM どうしの変換を表す  $3 \times 3$  回転行列  $\mathbf{R}$  を求める問題として定式化できる。

$$\mathbf{L}(\mathbf{p}) = \mathbf{R} \mathbf{A}_s(\mathbf{q}), \quad \mathbf{A}_s(\mathbf{q}) = \frac{\mathbf{K}_s^{-1} [\mathbf{q}^T \ 1]^T}{\|\mathbf{K}_s^{-1} [\mathbf{q}^T \ 1]^T\|}, \quad (8)$$

ただし、 $\mathbf{K}_s$  は全天球カメラとシーン画像の  $3 \times 3$  対応行列である。

球面座標の変換  $\mathbf{R}$  を求めるには、2組の対応点が必要であ

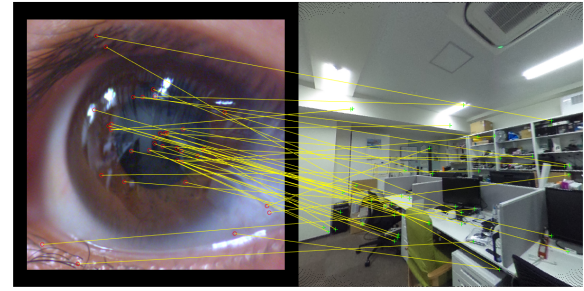


図 9 角膜表面反射画像と、ある全天球分割画像についての、SURF 初期対応点

る。しかし、角膜表面反射画像はノイズが多く、2組の正しい候補点を発見するのは困難である。そこで、今回は回転、スケールの変化に対して不変な画像特徴量 SURF を用いて、2画像の位置合わせの材料とする。SURF は回転に対して不変かつ、特徴量の方向を持っているため、これを用いることで、1組の対応点のみで  $\mathbf{R}$  を求めることができる。

今回は、SURF の特徴量のうち、位置  $\mathbf{x}$ , 特徴ベクトル  $\mathbf{F}(\mathbf{x})$ , 方向  $\theta^x$  を使用する。SURF の初期対応点検出(図 9)において得られた  $M$  ペアの特徴点と特徴量を、 $\{(\mathbf{p}_i, \mathbf{F}(\mathbf{p}_i), \theta_i^{\mathbf{p}}) | i = 1, \dots, M\}, \{(\mathbf{q}_i, \mathbf{F}(\mathbf{q}_i), \theta_i^{\mathbf{q}}) | i = 1, \dots, M\}$  とする。この中に、実際の対応と合致している組が少しはあると仮定し、それを見つけ出すための手続きを行う。

まず、各ペア  $i$  について、推定変換行列  $\mathbf{R}_i$  は、以下の式により表現できる。

$$\mathbf{R}_i = \begin{bmatrix} \hat{\mathbf{L}}_x & \hat{\mathbf{L}}_y & \hat{\mathbf{L}}_z \end{bmatrix} \begin{bmatrix} \hat{\mathbf{A}}_x & \hat{\mathbf{A}}_y & \hat{\mathbf{A}}_z \end{bmatrix}^{-1}, \quad (9)$$

ただし

$$\mathbf{L}_x = \mathbf{L}(\mathbf{p}_i)$$

$$\mathbf{L}_y = \mathbf{L}(\mathbf{p}_i) \times (\mathbf{L}(\mathbf{p}_i) \times \mathbf{L}'(\mathbf{p}_i, \theta_i^{\mathbf{p}}))$$

$$\mathbf{L}_z = \mathbf{L}(\mathbf{p}_i) \times \mathbf{L}'(\mathbf{p}_i, \theta_i^{\mathbf{p}})$$

$$\mathbf{L}'(\mathbf{p}, \theta^{\mathbf{p}}) = \mathbf{L}(\mathbf{p} + \mathbf{u}(\theta^{\mathbf{p}})) - \mathbf{L}(\mathbf{p})$$

$$\mathbf{A}_x = \mathbf{A}_s(\mathbf{q}_i)$$

$$\mathbf{A}_y = \mathbf{A}_s(\mathbf{q}_i) \times (\mathbf{A}_s(\mathbf{q}_i) \times \mathbf{A}'_s(\mathbf{q}_i, \theta_i^{\mathbf{q}}))$$

$$\mathbf{A}_z = \mathbf{A}_s(\mathbf{q}_i) \times \mathbf{A}'_s(\mathbf{q}_i, \theta_i^{\mathbf{q}})$$

$$\mathbf{A}'_s(\mathbf{q}, \theta^{\mathbf{q}}) = \mathbf{A}_s(\mathbf{q} + \mathbf{u}(\theta^{\mathbf{q}})) - \mathbf{A}_s(\mathbf{q})$$

$$\mathbf{u}(\theta) = \begin{bmatrix} \cos(\theta) & \sin(\theta) \end{bmatrix}^T$$

ただし、 $\hat{\mathbf{x}}$  は  $\mathbf{x}$  を正規化したものであり、 $\mathbf{L}'(\mathbf{y}_i, \theta)$  および  $\mathbf{A}'_s(\mathbf{y}_i, \theta)$  は、画像中の点  $\mathbf{y}$  における回転角  $\theta$  を投影する関数である。

まとめると、角膜表面反射画像中の点から全天球分割画像中の点に変換する関数  $\mathbf{W}$  は、

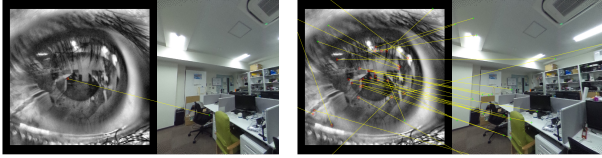


図 10 左: SURF 初期対応点のうちの 1 ペア, 右: そのペアで作った  $\mathbf{W}$  を対応点ペア全てに適用した様子

$$\mathbf{W}_i(\mathbf{p}) \equiv \frac{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{K}_s \mathbf{R}_i^{-1} \mathbf{L}(\mathbf{p})}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{K}_s \mathbf{R}_i^{-1} \mathbf{L}(\mathbf{p})} \quad (10)$$

となる. 4.4 節で,  $\mathbf{W}_i$  の中から最適な変換行列を得るためのアルゴリズムについて説明する.

#### 4.4 1point-RANSAC による評価

各行列  $\mathbf{W}_i$  の妥当性を評価するため, 画像の位置合わせにおいて広く用いられているアルゴリズム RANSAC (RANDOM SAMPLE CONSENSUS) を, 先ほどの各 SURF 初期対応点  $M$  ペアに対して適用する. これを以後, **1point-RANSAC** と呼ぶことにする.

角膜表面反射画像  $\mathbf{I}_{\text{eye}}$  と, 全天球分割画像  $\mathbf{I}_n$  ( $n = 1, \dots, 16$ ) における SURF 初期対応点のペア  $i$  ( $i = 1, \dots, M_n$ ) それぞれに対して, 以下を行う.

まず  $\mathbf{p}_i, \mathbf{q}_i$  の組から,  $\mathbf{W}_i$  を得る. ここで, 全てのペア  $1, \dots, M_n$  を評価に用いる. そのそれぞれの  $\mathbf{p}_j$  ( $j = 1, \dots, M_n$ ) に対して  $\mathbf{W}_i$  を適用した点 (図 10) を  $\mathbf{q}'_j$  としたとき,  $\|\mathbf{q}_j - \mathbf{q}'_j\| < \text{threshold}$  を満たすペア  $j$  の数を  $N_{\text{valid}}(i)$  とする. また,  $j$  の総数を  $N_{\text{all}}$  とする. このとき  $\mathbf{W}_i$  の評価値  $\mathbf{E}(\mathbf{W}_i)$  を

$$\mathbf{E}(\mathbf{W}_i) = \frac{N_{\text{valid}}(i)}{N_{\text{all}}} \quad (i = 1, \dots, M_n) \quad (11)$$

と定義する. ここで,  $\mathbf{I}_{\text{eye}}$  と,  $\mathbf{I}_n$  の対応に適するものとして

$$\mathbf{W}_{\mathbf{I}_n} = \arg \max_{\mathbf{W}_i} \mathbf{E}(\mathbf{W}_i) \quad (i = 1, \dots, M_n) \quad (12)$$

を満たす行列  $\mathbf{W}_{\mathbf{I}_n}$  を選択する.

この過程を, 各全天球分割画像  $\mathbf{I}_n$  ( $n = 1, \dots, 16$ ) に対して行い, 最も高い評価値を与えた  $\mathbf{W}_{\mathbf{I}_n}$  を結果とする. (図??) つまり, 角膜表面反射画像に最も近いと推測される画像  $\mathbf{I}_n^*$  および,  $\mathbf{I}_{\text{eye}}$  から  $\mathbf{I}_n^*$  への座標変換関数  $\mathbf{W}^*$  は

$$\mathbf{I}_n^* = \arg \max_{\mathbf{I}_n} \mathbf{E}(\mathbf{W}_{\mathbf{I}_n}) \quad (n = 1, \dots, 16) \quad (13)$$

$$\mathbf{W}^* = \mathbf{W}_{\mathbf{I}_n^*} \quad (14)$$

となる.

#### 4.5 注視点の推定

眼球姿勢, および 2 画像間の変換関数  $\mathbf{W}^*$  を取得したの

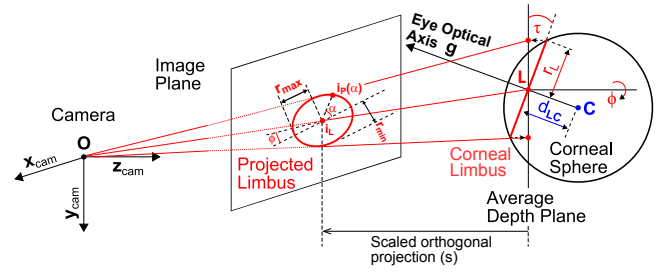


図 11 カメラと眼球の関係 [8]

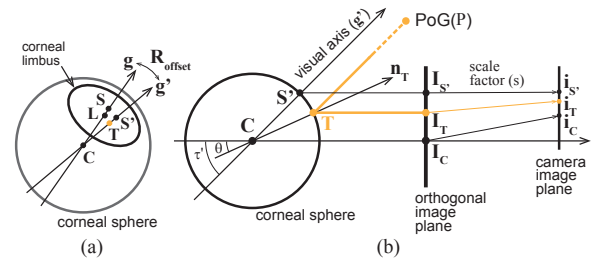


図 12 視線の方向, およびカメラと眼球の関係 [8]

で, 次は注視点の推定を行う. 本節は [8] と同様の手法を用いている.

この注視点の推定における目的は, 注視点 Point of Gaze(PoG) の角膜表面における位置, Gaze Reflection Point(GRP)  $\mathbf{T}$  の角膜表面反射画像上の座標  $\mathbf{i}_T$  を得ることである.

図 11 は, カメラが映す映像と眼球姿勢との関係, そして図 12 は別視点のもの, さらに角膜中心の表面と垂直な方向 (Eye Optical Axis)  $\mathbf{g}$  と, 実際の視線の方向 (Visual Axis)  $\mathbf{g}'$  の違い, およびその 2 つの関係について説明したものである. この視線の方向の違いには個人差があるが, その関係は  $\mathbf{g}' = \mathbf{R}_{\text{offset}} \mathbf{g}$  で表すことができる. ただし  $\mathbf{R}_{\text{offset}}$  は個人差パラメータである.

眼球は球であることを仮定しているが, 角膜表面反射画像は 2 次元画像である. そこで, 眼球の表面を平面に仮想的に投影した Orthogonal image plane (図??-(2)(b)) を考えると,

$$\begin{aligned} \mathbf{C} \cdot \mathbf{n}_T &= \begin{bmatrix} \cos \tau' & \sin \tau' \end{bmatrix} \cdot \mathbf{n}_T, \quad \mathbf{C} = \begin{bmatrix} 1 & 0 \end{bmatrix}, \\ \mathbf{n}_T &= \begin{bmatrix} \cos \theta & \sin \theta \end{bmatrix}, \quad \tau' = \arccos(g'_z), \\ \mathbf{g} &= \begin{bmatrix} g_x & g_y & g_z \end{bmatrix}^T, \quad \mathbf{g}' = \begin{bmatrix} g'_x & g'_y & g'_z \end{bmatrix}^T. \end{aligned}$$

また, 角膜の角度  $\theta$  は  $\tau'$  を使って表すことができ,

$$\theta = \arctan((1 - \cos \tau') / \sin \tau'). \quad (15)$$

である. ここで,  $|\mathbf{I}_T - \mathbf{I}_C| = r_C \sin \theta$  であり, Orthogonal image plane とカメラのイメージはスケール変化の関係にあることを考えると, GRP の座標  $\mathbf{i}_T$  は

$$\mathbf{i}_T = \mathbf{i}_L + s \left( -d_{LC} \begin{bmatrix} g_x & g_y \end{bmatrix}^T + r_C \sin \theta \begin{bmatrix} g'_x & g'_y \end{bmatrix}^T \right). \quad (16)$$

である。ただし、 $s$  は角膜の中心  $L$  を通る Orthogonal image plane と、角膜表面反射画像間のスケール倍率である。

#### 4.6 注視点マップの作成

ここまでの過程で、

- 角膜反射画像  $I_{eye}$  と全天球分割画像  $I_n$  の組、および  $I_{eye}$  から  $I_n$  への変換関数  $W^*$
- $I_{eye}$  における GRP  $i_T$

が得られた。これらを用いて、全天球画像に注視点を投影することが可能である。まず、 $i_T$  に  $W^*$  を適用し、 $I_n$  における注視点の座標を得る。この  $I_n$  はもともと、式 (2) によって全天球画像から生成されたものであるため、式 (2) から  $\phi', \theta$  が得られる。全天球画像における  $x$  軸の対応角度  $\phi$  と、先ほどの  $\phi'$  の関係は、各  $I_n$  を生成するとき 1 対 1 で決定されている ( $\phi = \phi' + n\pi/8 \pmod{2\pi}$ )。したがって、全天球画像上での GRP,  $u$  が得られる。

全部で  $F$  フレームある角膜表面反射画像のあるフレーム  $I_{eye}(k)$  ( $k = 1, \dots, F$ ) から  $u_k$  が得られるとする。このとき、各  $u_k$  について、 $\|u_k - x\| < \text{threshold}$  を満たす画素  $x$  それぞれに対して、評価  $(1 - \frac{\|u_k - x\|}{\text{threshold}})$  ポイントを与える。こうして全てのフレームを評価し、全天球画像の画素値行列  $A$  と同じ大きさの評価値の行列  $H$  をつくる。このとき、注視点マップ  $A + cH$  が得られる。ただし  $c$  は  $\max H(i, j)$  に応じて決まる係数である。注視点マップの例は、図 3(d) などに示している。

### 5. 実験

4.4 節において、角膜表面反射画像と全天球分割画像の位置合わせを、1point-RANSAC を用いて行う必要があることを述べた。つまり、注視点推定は大きく分けて次の 2 つのステップを通じて行われる。

- (1) 1point-RANSAC による評価で、角膜表面反射画像と全天球分割画像の対応点を 1 つ導く。このとき、対応点は正しくなければならない。
- (2) 対応点から得られた座標変換関数と、眼球モデルから導いた Gaze Reflection Point(4.5 節) を用いて、全天球画像上の注視点を推定する。

我々は実験で、(1) の成功率、および (1) が成功したものについて、(2) の注視点推定を行ったときの誤差を確認した。

#### 5.1 屋内

屋内の全天球画像と実際の注視点、および、それぞれの点に対応する角膜表面反射画像 (図 13) を用いて位置合わせの実験と注視点の推定を行った。屋内 1 は、画像全体に特徴物が多く存在するが、屋内 2 は、一部にこそホワイトボードや教壇が見られるものの、全体的には均質かつ規則的な空間であり、特徴物は少ない。

表 1 位置合わせ結果

	画像の数	位置合わせ成功数	成功率 (%)
屋内 1	28	28	100.00
屋内 2	10	9	90.00

表 2 注視点推定結果と実際の注視点の誤差

	平均 (度)	標準偏差	中央値 (度)
屋内 1 : 水平方向	4.826	2.818	4.410
屋内 1 : 鉛直方向	3.335	2.243	2.689
屋内 2 : 水平方向	2.877	3.459	1.377
屋内 2 : 鉛直方向	2.487	1.617	2.207



図 13 屋内 1 の角膜表面反射画像の一例



図 14 屋内 1 の推定注視点 (黄色枠). 緑丸は実際の注視点.

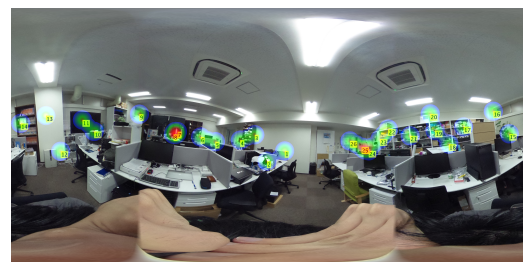


図 15 屋内 1 の全天球画像に生成された注視点マップ

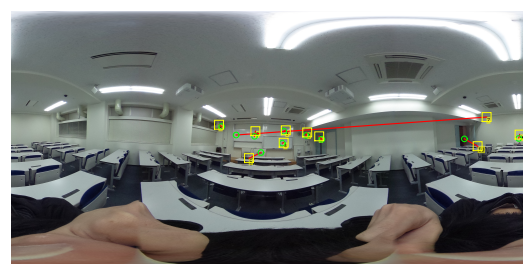


図 16 屋内 2 の推定注視点 (黄色枠). 緑丸は実際の注視点.

位置合わせを行った結果を表 1 に示した。また、それぞれの角膜表面反射画像に対して注視点推定を行った結果を、図 14, 図 16 に示した。図 16 を見ると、7 番の注視点推定の位置が大きくずれているが、これは位置合わせ結果が間違っていたものである。その他の推定結果は、実際のものとは少し離れているものもあるが、全体的には少しの誤差にとどまっているものが多いことがわかる。位置合わせが成功したものについての、注視点推定の角度誤差の計算結果を、表 2 に示した。

表 3 位置合わせ結果

	画像の数	位置合わせ成功数	成功率 (%)
屋外 1	30	28	93.33
屋外 2	30	24	80.00
屋外 3	10	7	70.00

表 4 注視点推定結果と実際の注視点の誤差

	平均 (度)	標準偏差	中央値 (度)
屋外 1 : 水平方向	4.687	9.782	2.306
屋外 1 : 鉛直方向	4.012	2.992	3.481
屋外 2 : 水平方向	9.598	19.395	2.644
屋外 2 : 鉛直方向	4.245	3.922	4.035
屋外 3 : 水平方向	2.252	1.700	3.550
屋外 3 : 鉛直方向	2.041	3.550	0.673

## 5.2 屋外

屋外の全天球画像と実際の注視点、および、それぞれの点に対応する角膜表面反射画像 (図 17) を用いて位置合わせの実験と注視点の推定を行った。10 点の実際の注視点を定め、屋外 1 と屋外 2 では被験者 3 名に、屋外 3 では被験者 1 名に、それぞれの点を注視してもらい、その角膜表面反射画像を撮影した。

位置合わせを行った結果を表 3 に示した。また、位置合わせが成功した角膜表面反射画像に対して注視点推定を行った結果を、図 18、図 19、図 20 に示した。位置合わせが成功したものについての、注視点推定の角度誤差の計算結果を、表 4 に示した。

図より、一部の点では位置合わせの結果が正しくても注視点推定が大きくずれているケースが発生していることが確認できる。その結果、注視点推定結果の誤差平均、標準偏差が大きくなってしまっているものがある。ここで誤差の中央値を見ると、誤差平均にかかわらず 0~4 度程度の値が出ていることから、大きくずれている点以外の、大半の点においては安定して注視点推定を行うことができるといえる。

## 5.3 考察

全天球分割画像と角膜表面反射画像の位置合わせが正しくなければ、正しい注視点を推定することができないことは既に述べた。つまり、注視点の推定で重要となる要素の 1 つに、1point-RANSAC の精度がある。位置合わせの結果の出力は、複数の候補の中から、4.4 節で扱った評価値が最大のものを選ぶ。今回の実験の際に、プログラム実行の様子を見ると、「正しい結果が出た位置合わせ」の中でも、他の候補点を大きく上回って出力した正解と、そうでないものに分かれていた。例えば、屋内 1 ではどの正解点の組も非常に良い評価値を出していたが、屋内 2 では、そうでない正解点の組が多く、「かろうじて」正解を出力することができたものがいくつか見られた。屋外の実験でも同様の現象が見られた。屋外 1 の全天球画像の中



図 17 屋外 1, 各被験者の角膜表面反射画像の一例

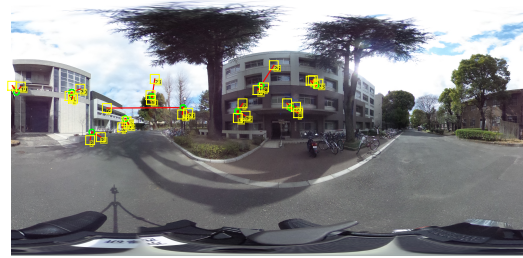


図 18 屋外 1 の推定注視点 (黄色枠). 緑丸は実際の注視点.



図 19 屋外 2 の推定注視点 (黄色枠). 緑丸は実際の注視点.



図 20 屋外 3 の推定注視点 (黄色枠). 緑丸は実際の注視点.



図 21 屋外 1 の全天球画像に生成された注視点マップ

心にある建物は距離が近く、窓なども多くあったため、その方向を見ていた画像については注視点推定は非常に安定していたが、それ以外の方向では位置合わせが失敗することもあった。このことから、画像中の特徴物の多さが、1point-RANSAC の精度に深く関わっていることがわかる。したがって、1point-RANSAC は、特徴物がよく写っている角膜表面反射画像に対しては、位置合わせの成功が良く望める一方で、何も特徴がないもの (例えば、単色の壁、床、空など) が大半を占める角膜表面反射画像に対しては、位置合わせを成功させにくいと考えられる。

見ている方向の明るさや、画像の明るさも位置合わせ精度に関わっていると考えられる。屋外1に比べて暗い場所が多い屋外2, 屋外3の実験では、屋外1に比べて位置合わせが失敗する点が多かった。これに加えて、角膜表面反射画像には多数のノイズが含まれている。角膜の光彩, まつげ, まぶたが角膜表面反射画像に映りこんでいることや、眼球への光の入射量や光源の種類(電灯, 太陽など)は位置合わせの精度に関わっていると言える。

注視点の推定においては、位置合わせが正しい場合でも、大半の点において少しの誤差が生じていることが確認できる。また、一部の点は大きく誤差が生じている。この誤差を生じさせている原因の1つとして、4.5節で述べた角膜中心の表面と垂直な方向(Eye Optical Axis)  $g$  と、実際の視線の方向(Visual Axis)  $g'$  の違いを表す個人差パラメータの誤差が考えられる。この個人差パラメータは正確に計測するのが難しいため、別のアプローチで調査する必要があると考える。また、仮定した眼球モデルと実際の眼球の違いや、角膜表面反射画像からシーン画像への変換関数  $W$  にも多少の誤差が生じていることが考えられる。その場合、位置合わせで検出できた対応点ペアと、その画像における注視点が離れているほど、誤差が大きくなるリスクが増大する。

## 6. 結論

本稿では、全天球画像を用いて全周型注視点推定を行った。角膜表面反射画像のみを用いて眼球モデルの構築を行い、全天球画像の分割による、平面画像への変換、角膜表面反射画像と、全天球分割画像の位置合わせによる、2画像の組み合わせの決定を通して、注視点を導いた。実験により、2画像の位置合わせにおいて高い成功率を出すことができる可能性、およびそこから推定された注視点の妥当性が示唆された。しかし一方で、環境の違いや角膜表面反射画像のノイズの多さにより、位置合わせの成功率の低下を招き、その結果注視点の推定の成功率が低下してしまうケースも見られた。したがって、2画像の位置合わせの側面では、鮮明な角膜表面反射部の安定した取得と、ノイズに対処する手法の検討が今後の課題として挙げられる。また、本稿では、周辺の環境が大きく変化しないという仮定のもとで注視点推定を行った。この仮定は確かに、現実的に様々な場面で成立しうが、例えば、混雑する車道のように、画像の一部環境が大きく変化する場面においては、有用性はまだ確認できていない。現実世界における様々な環境で、2画像の位置合わせが可能であるかどうかを検討することも今後の課題とする。注視点推定の側面から見ると、眼球姿勢や、注視点の軸(Visual Axis)には個人差により多少の誤差があることが示唆された。そこで、より精度の高い注視点推定のため、個人差の修正パラメータを導く手法を検討することを今後の課題とする。

全周型の注視点推定は、従来の平面に対しての注視点推定から一歩広がり、3次元的な概念が導入されている。注視点推定の範囲を拡大することができれば、将来的には角膜表面反射を通じて、3次元空間における注視点推定、および注視点マップの生成が可能になることが期待される。現実空間に即した注視点推定を行うことで、特に注視点推定や視線を扱う研究分野や、店内全体での注視点推定、マーケティングへの応用など、商業的分野においても、大きな利用価値が生まれることが期待されるから、角膜表面反射画像による注視点推定の応用可能性は十分に残されているといえる。

## 参考文献

- [1] Baya, H., Essa, A., Tuytelaars, T. and Van Gool, L.: Speeded-up robust features (SURF), *Computer vision and image understanding*, Vol. 110, No. 3, pp. 346–359 (2008).
- [2] Bowyer, K. W., Hollingsworth, K. and Flynn, P. J.: Image understanding for iris biometrics: A survey, *Computer vision and image understanding*, Vol. 110, No. 2, pp. 281–307 (2008).
- [3] Fischler, M. A. and Bolles, R. C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Communications of the ACM*, Vol. 24, No. 6, pp. 381–395 (1981).
- [4] Lowe, D. G.: Distinctive image features from scale-invariant keypoints, *International journal of computer vision*, Vol. 60, No. 2, pp. 91–110 (2004).
- [5] Ma, L., Tan, T., Wang, Y. and Zhang, D.: Personal identification based on iris texture analysis, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 25, No. 12, pp. 1519–1533 (2003).
- [6] Nakazawa, A. and Nitschke, C.: Point of gaze estimation through corneal surface reflection in an active illumination environment, *Computer Vision–ECCV 2012*, Springer, pp. 159–172 (2012).
- [7] 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).
- [8] Nakazawa, A., Nitschke, C. and Nishida, T.: Non-calibrated and real-time human view estimation using a mobile corneal imaging camera, *Multimedia & Expo Workshops (ICMEW), 2015 IEEE International Conference on*, IEEE, pp. 1–6 (2015).
- [9] Nishino, K., Belhumeur, P. N. and Nayar, S. K.: Using eye reflections for face recognition under varying illumination, *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, Vol. 1, IEEE, pp. 519–526 (2005).
- [10] Nishino, K. and Nayar, S. K.: Corneal imaging system: Environment from eyes, *International Journal of Computer Vision*, Vol. 70, No. 1, pp. 23–40 (2006).
- [11] Nitschke, C., Nakazawa, A. and Takemura, H.: Corneal imaging revisited: An overview of corneal reflection analysis and applications, *Information and Media Technologies*, Vol. 8, No. 2, pp. 389–406 (2013).