

# 広視野アイカメラにおける 顕著性マップを用いた視線推定結果の補正手法

重野 孝明<sup>1</sup> 間下 以大<sup>1,2</sup> 清川 清<sup>1,2</sup> 竹村 治雄<sup>1,2</sup>

**概要:** 広視野アイカメラを用いたアピランススペースでの視線推定において、アイカメラと頭部の位置関係の変動は、視線推定精度を大きく悪化させる原因である。本研究では、広視野アイカメラを用いた視線推定における、顕著性マップを用いた推定結果補正手法を提案する。視野画像に対して顕著性マップを計算し、アピランススペースでの視線推定結果と対応付けることにより、アイカメラと頭部の位置関係とは独立な、視線位置候補を取得する。得られた視線位置候補と対応付けられた視線推定結果の組をもとに、ホモグラフィ変換を計算する。このホモグラフィ変換を視線推定結果に適用することで補正を行う手法である。提案手法に対して実験を行い、ホモグラフィ変換による視線推定結果の補正に有効性が認められないという結果が得られた。

**キーワード:** 視線推定, 顕著性マップ, ホモグラフィ

## 1. はじめに

アイカメラとは利用者の視界の映像とともに注視点を記録する機器であり、近年の光学機器の小型化により急速に進化している。このような背景によりアイマークレコーダは工学、心理学、医学、産業などの様々な分野で視線の解析の手段として幅広く利用されている [3]。例えば、自動車運転中に熟練者と初心者では目の動きが異なることはよく知られており [13]、この違いを運転初心者にフィードバックすることができれば運転技術の向上に役立てることができると考えられる。また、歩行者が視界にある情報群から実際に得ている情報を分析することにより、歩行者にとってよりわかりやすい案内板の配置などを行うことができると考えられる。

現在一般的なアイカメラを、上記の自動車運転などの作業時の視線解析に利用するにはいくつかの問題がある。問題の一つとして視野角の確保があげられる。人間の視界は水平方向に約 190°、垂直方向に約 130° であり [19]、運転者や歩行者は前方にあるものだけでなく、サイドミラーや高い位置にある電光掲示板、案内板と車間距離など、様々な方向に対して同時に注意を払っている。そのため、利用者が何を見ているか、あるいは何を見逃したかを正確に解析するためには人間と同程度の視野角を持ち、周辺視野も

撮影できるカメラが必要となる。

カメラレンズと利用者の眼球との視差も問題となる。運転者や歩行者は手元の機器から遠方の標識まで、様々な距離にあるものを見ている。利用者の視界を撮影するカメラと、利用者の実際の視点との間に視差が存在する場合、利用者が見ている対象のカメラ座標上での位置は、その物体と利用者またはカメラとの間の距離に影響を受ける。すなわち、事前にキャリブレーションを行った際に用いた物体距離と異なる距離にある物体に対しては視線推定精度が低下する。このため、様々な距離にある物体を見る必要のある作業時の正確な視線解析に利用することは難しい。

実際の環境でアイカメラを利用するためには、他に装着位置の変動に対してのロバスト性も重要である。アイカメラと利用者の眼球との位置関係が変動した場合において、視線推定精度が低下する場合、長時間にわたる視線解析に利用することは難しい。また、利用者がアイカメラを装着するたびにキャリブレーションを行う必要があるため、利用者に対する負担が大きくなる。装着位置の変動に対するロバスト性を得ることにより、正確な視線推定を長時間にわたり実現できるだけでなく、装置を装着しなおした場合でも、過去のキャリブレーション結果を再利用できることから、キャリブレーションの実施回数を大幅に削減し、利用者に対する負担を低減することができる。

報告者の所属する研究グループでは、過去に住谷らが広視野と無視差の 2 つの要求を満たす、双曲面ハーフミラー

<sup>1</sup> 大阪大学大学院情報科学研究科

<sup>2</sup> 大阪大学サイバーメディアセンター

を用いた視点一致型広視野アイカメラ [12] を提案した。しかし、視線推定手法に問題があり、視線方向によっては正しい推定が行えないという問題があった。これに対して森らは利用者の視線方向や光学系の歪みの影響を受けにくいアピアランススペースの視線推定手法 [7] を提案し、この問題を解決した。森らの手法はアピアランススペースの視線推定手法を用いているため、微小な装着位置の変動に対しても、視線推定精度が大幅に低下することがわかっている。

本研究では、アピアランススペースの視線推定における、アイカメラ装着位置の変化に対するロバスト性を得るため、従来手法によって得られた視線推定結果を補正する手法を提案する。まず、視線推定を行うと同時に顕著性マップを計算し、実環境における視点位置候補を得る。その上で、視線推定の結果と顕著度マップの対応から推定結果から視点候補位置へのホモグラフィ変換を取得し、その変換を視線推定の結果に適用する。これによって、視線推定におけるアイカメラ装着位置の変動の影響を低減することを目指す。さらに、提案手法による補正が、アイカメラ装着位置の変動に対して有効であるかについて評価実験を行う。

本稿の構成を示す。2 節で顕著性マップを用いた視線推定、ならびに視線推定結果の補正に関する関連研究について述べる。3 節で住谷らの提案した広視野アイカメラ、および森らの提案したアピアランススペース視線推定手法の概念を紹介する。4 節では装着位置の変動の影響を低減する補正手法の設計と提案を行い、5 節では視線推定における提案手法の評価を行う。6 節では提案手法のまとめと今後解決すべき課題について述べる。

## 2. 関連研究

### 2.1 広視野アイカメラにおける視線推定手法

パッシブ型の視線推定手法に、アピアランススペースの視線推定手法がある [2], [6], [8], [10]。これらは入力画像の特徴を眼球の状態や虹彩の領域などの情報に変換せずに直接視線方向を推定する手法であり、眼球周辺の画像の変化に対して推定精度がロバストであるため、広視野アイカメラに適しているといえる。一般にアピアランススペースの視線推定手法は入力画像を学習データと照合することで視線方向を推定するため、利用者の顔に対するカメラの相対位置や向きが変化すると目の映り方が変化し、推定精度が低下するという問題がある。また、通常はカメラと肉眼の間に視差があるため、注視距離によって入力画像と視線方向のマッピングが変化してしまうという問題もある。

住谷らのアイカメラ [12] はヘッドマウント型であるため、利用者の眼球とカメラの相対位置は利用中に大きく変化しないと想定できる。これは視差の問題を解決しているといえる。しかし、住谷らの手法は虹彩を検出する簡易なモデルをもとにして視線の検出を行っているため、視線方向の変化に対して推定結果がロバストではないという問題

がある。森らは、住谷らのアイカメラにアピアランススペースによる視線推定手法を導入することにより、ヘッドマウント型の利点を残したまま、視線方向の変化に対してロバストな視線推定を行う手法を実現している [7]。

### 2.2 顕著性マップ

顕著性マップとは、画像および動画における人間の視覚的な注目度をあらわしたマップである。Itti ら [5] により提案され、色相、輝度、方向からなる特徴量を抽出することで計算される。Itti らの手法をはじめとし、運動、点滅、optical flow など、さまざまな特徴量による顕著性マップが考案されてきた [1], [4]。また、顕著性マップを利用した物体検出を行うこともできる [18]。

顕著性マップを視線推定に利用する場合、視線推定を行う環境やコンテキストにより、いずれの特徴量の重要性が増加するかは変化する。たとえば、自動車の運転中であれば、視界に存在する人間や、ほかの車両などに注意を払う必要があるため、運動特徴量をはじめ、歩行者、車両に対する物体検出結果の重要度が高くなる。そのため、視線推定を実際に行う場面に応じて適切な特徴量を抽出する顕著性マップを用いることが必要と考えられる。

### 2.3 顕著性マップを用いた視線推定手法

顕著性マップ利用した視線推定では、顕著性マップから視線位置の確率密度を計算する視線推定の手法がある [11], [16]。これらの手法は、カメラで直接撮影された動画のみを対象としているため、双曲面鏡を用いた広視野かつ、歪みのある画像に対してもそのまま適用できるかについては検討の余地がある。

## 3. 双曲面ミラーを用いたアイカメラと視線推定手法

2 節で述べたように、住谷らのアイカメラと森らの視線推定手法は、広視野かつ無視差であり、視線方向によらず安定した視線推定が可能である。しかし、装着位置の変動による入力画像の変化によって、推定精度が悪化してしまうという問題がある。カメラと眼球の位置関係はアイカメラを頭部に装着することによりある程度限定されるものの、再装着や頭部の運動によっても位置の変動が発生しないようにすることは現実的ではない。

本研究では、装着位置の変動による画像の変化に対するロバストな視線推定を実現する。そのために、住谷らのアイカメラと森らの手法をベースとする、顕著性マップによる視線位置候補の取得とホモグラフィ変換による視線推定結果の補正手法を提案する。本節では、提案手法のベースとして用いる、住谷らの広視野アイカメラと森らの視線推定手法の基本的な構成について述べる。

### 3.1 住谷らのアイカメラ

#### 3.1.1 全方位カメラ : Hyper Omni Vision

住谷らの提案したアイカメラ [12] は基本的な発想を全方位カメラ Hyper Omni Vision[15] から得ている。全方位カメラとは円錐や曲面、双曲面のミラーを用いて 360° の映像を 1 台のカメラで記録するシステムであり、Hyper Omni Vision は双曲面ミラーを用いた全方位カメラである。Hyper Omni Vision は単一視点であり、双曲面ミラーの反射により歪んだ画像を通常の透視投影画像に実時間で変換可能という特徴がある [14]。双曲面ミラーの内焦点は  $O_m(0, 0, +c)$ 、外焦点は  $O_c(0, 0, -c)$  である。この座標系ではミラーの形状は次式で表すことができる [15]。

$$\frac{X^2 + Y^2}{a^2} - \frac{Z^2}{b^2} = -1 (Z > 0) \quad (1)$$

$$c = \sqrt{a^2 + b^2} \quad (2)$$

式中の  $a, b, c$  はミラーのパラメタである。双曲面ミラーでは内焦点  $O_m$  に向かう光はすべて外焦点  $O_c$  へ反射されるため、カメラを外焦点  $O_c$  に設置すると内焦点  $O_m$  を視点とする光軸周り 360° の画像が取得できる。本座標系における空間中の任意の三次元座標  $P(X, Y, Z)$  に対応するカメラ座標系上の二次元座標  $u(x, y)$  は次式で表すことができる。ただし、 $f$  はカメラの焦点距離である。

$$x = X \times f \times \frac{b^2 - c^2}{(b^2 + c^2)Z - 2bc\sqrt{X^2 + Y^2 + Z^2}} \quad (3)$$

$$y = Y \times f \times \frac{b^2 - c^2}{(b^2 + c^2)Z - 2bc\sqrt{X^2 + Y^2 + Z^2}} \quad (4)$$

#### 3.1.2 広視野アイカメラの基本コンセプト

住谷らの提案したアイカメラの基本的なアイディアは、Hyper Omni Vision の光学系に対し、ミラーをハーフミラーに変えることおよびミラーの内焦点に利用者の眼球を配置することである (図 1)。前述のように、双曲面ハーフミラーの内焦点に向かうすべての光は外焦点に反射される。よって、小型のカメラを外焦点に設置することで、内焦点を視点とした広視野な映像が取得可能となる。これに加え、利用者の目を内焦点におくことで、彼らの視界と同一の映像をカメラで取得可能である。また、視線推定に用いる眼球画像を取得するためにハーフミラーに穴 (アイホール) が設けられている。住谷らの提案したアイカメラの主な利点を以下にまとめる。

- ・ 広視野

凸面ミラーを用いることで利用者自身の視角とほぼ同じ広視野な映像を取得可能

- ・ 無視差

双曲面の特性により利用者ともまったく同一の視点から映像を取得可能

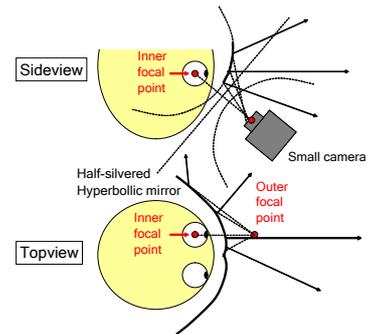


図 1 双曲面ハーフミラーを用いたアイカメラの基本構成 [12]  
Fig. 1 Basic component of eye mark recorder with hyperbolic half mirror[12]



図 2 入力画像の例  
Fig. 2 Example of input image

- ・ 視線推定が可能

目の動きを記録できるので、視線推定を行うことが可能

- ・ 構造がシンプル

1 台のカメラのみで利用者の視野と目の動きを記録できるため、他のカメラや同期機構が不要

### 3.2 森らの視線推定手法

森らは、住谷らのアイカメラの利点を活用したアピアランススペースでの視線推定手法を提案した [7]。アピアランススペースの視線推定とは眼球画像の特徴量と視線方向の関係を学習させ、推定に用いる手法である。

学習データの注視点とアイホール画像の特徴量の対応関係を用いて視線方向の推定を行う。アイホール画像とは入力画像 (カメラ画像) 中のアイホール周辺の矩形領域の画像である。入力画像の例を図 2 に示す。入力画像は中央のアイホール領域、その周辺の利用者の視野領域、一次反射ミラー領域などいくつかの領域がある。アイホール画像の例を図 4 に示す。

住谷らのアイカメラではカメラのレンズと利用者視点の間に視差が存在しないため、利用者の注視距離に関係なく、同じ視線ベクトル上にある点はカメラ画像上の同じ位置に映る。これによりカメラと眼球、注視対象の位置関係を考慮する必要なく、視線方向の推定を入力画像上の注視点推定と等価にみなすことができる。

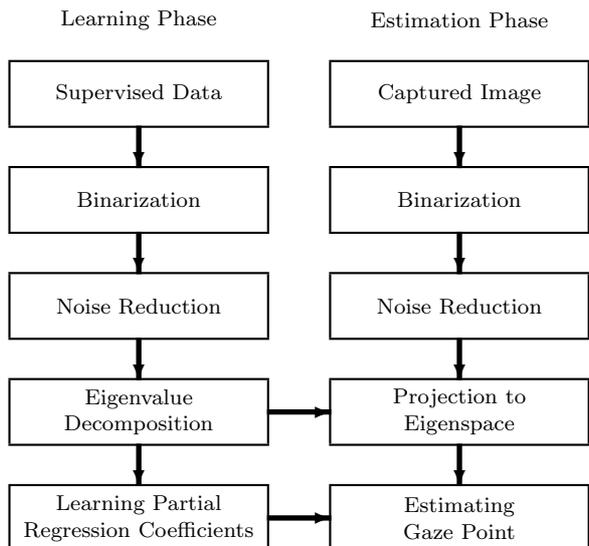


図3 推定手法のフローチャート [7]  
Fig. 3 Flowchart of Estimation[7]

森らの視線推定手法は学習フェーズと推定フェーズの2つからなる。図3は森らの提案手法の全体の流れを示している。学習フェーズでは大量の学習データから推定に必要な固有空間と回帰式を求める。学習データとは二値眼球画像とそれに対応した注視点の座標の組のことを指す。まず主成分分析を用いて眼球画像を固有画像に変換する。その後、重回帰分析を行うことで、固有画像と注視点の関係式の回帰係数を得る。固有画像とは入力画像に対しノイズ除去と情報圧縮を行った結果得られる画像である。推定フェーズでは入力画像を固有空間に投影し、学習フェーズで求めた回帰式に当てはめることで注視点の推定を実時間で行う。以降では視線推定で重要となる二値化、ノイズ除去、主成分分析、重回帰分析の説明を行う。

### 3.2.1 二値化、ノイズ除去

学習フェーズ、推定フェーズともに、最初にアイホール画像の二値化を行う。二値化の閾値は学習フェーズにおいて学習データを取得する前に、大津の手法 [9] を用いて決定する。大津の手法とは、各画素を閾値により輝度の高いクラス、低いクラスに分け、最もクラス間分離度の高い閾値を自動的に求める手法である。二値化後に膨張収縮処理によりノイズを減らす。図4,5にアイホール画像 (図4) から二値化、ノイズ除去を行った結果 (図5) を示す。なお、推定フェーズでは学習フェーズにおいて決定した閾値を定数として用いる。

### 3.2.2 主成分分析

固有画像は全学習データに対して主成分分析を行って得られた主成分空間に二値画像を投影することによって得られる。主成分分析を行うために、各二値画像を列ベクトル  $I_i$  に変換する。次に  $D$  を以下のように定義する。



図4 取得したアイホール画像  
Fig. 4 Eye hole image from camera



図5 二値化、ノイズ除去を行ったアイホール画像  
Fig. 5 Binarized and denoised eye hole image

$$D = [I_1 - \bar{I} \quad I_2 - \bar{I} \quad \dots \quad I_N - \bar{I}]^T \quad (5)$$

$\bar{I}$  は平均ベクトル、 $N$  は学習データ (二値画像) の数である。共分散行列を  $\sum_I = D^T D$  とする。  $\sum_I$  の固有値  $\lambda_0, \dots, \lambda_N$  は特異値分解によって算出することができる ( $\lambda_0 > \lambda_1 > \dots > \lambda_N$ )。  $v_1$  が第一主成分ベクトル、  $v_2$  が第二主成分ベクトル、  $v_n$  が第  $n$  主成分ベクトルである ( $N > n$ ) とすると、特徴ベクトル  $s$  は画像ベクトル  $I$  を次式のように  $n$  次元固有空間に投影することで得られる。

$$s = A^T (I - \bar{I}) \quad (6)$$

$$A = [v_1 v_2 \dots v_n] \quad (7)$$

$A$  は固有空間への射影行列である。

### 3.2.3 重回帰分析

偏回帰係数  $B$  は次式のように重回帰解析により求めることができる。

$$U = \begin{bmatrix} u_{01} & u_{02} & \dots & u_{0n} \\ u_{11} & u_{12} & \dots & u_{1n} \end{bmatrix} \quad (8)$$

$$B = \begin{bmatrix} b_{01} & b_{02} & \dots & b_{0n} \\ b_{11} & b_{12} & \dots & b_{1n} \end{bmatrix} \quad (9)$$

$$S = \begin{bmatrix} 1 & 1 & \dots & 1 \\ s_1 & s_2 & \dots & s_N \end{bmatrix} \quad (10)$$

$u = [u_{0i}, u_{1i}]^T$  は入力画像上での注視点であり、  $U$  は従属変数行列、  $S$  は独立変数行列である。よって回帰係数  $B$  は次式の  $F$  を最小化することにより算出できる。

$$F = (U - \hat{U})^2 = (U - BS)^2 \quad (11)$$

ここで、 $\hat{U}$  は注視点の推定値である。疑似逆行列を用いて  $B$  は以下のように表すことができる。

$$B = US^T[SS^T]^{-1} \quad (12)$$

視線推定の際には注視点  $\hat{u}$  は入力画像  $I$  を用いて次式で算出される。

$$\hat{u} = b_0 + B'A^T(I - \bar{I}) \quad (13)$$

ここで、 $b_i = [b_{0i} b_{1i}]^T$ ,  $B' = [b_1 \cdots b_n]$  である。

## 4. 顕著度マップを用いた視線推定結果の補正手法

### 4.1 視線推定結果の有効な補正手法を得るための課題

住谷らのアイカメラ [12] と森らの視線推定手法 [7] を組み合わせた視線推定を行うと、学習フェーズと推定フェーズにおける装着位置に変動があるとき、推定誤差が大きくなる。誤差が大きくなる原因と、アイカメラの要求仕様について述べる。

#### 4.1.1 装着位置の変動にともなうアイホール画像の変化

眼球とカメラの位置関係が変動することにより、アイホール画像が変化し、視線推定精度が悪化することが森ら [7] により示されている。

#### 4.1.2 視線推定方法における広視野性の維持

森ら [7] によるアピアランススペースでの視線推定手法は、視線方向に対してロバストであることが示されている。広視野アイカメラにおける視線推定では、アピアランススペースでの視線推定における広視野性を保持している必要がある。

### 4.2 アプローチ

4.1 節での課題と要求仕様を満たすためには、眼球とアイカメラの位置関係の変動に影響しない手法により、視線位置候補を取得し、それをアピアランススペースにおける視線推定結果と対応付けることで、補正を行うアプローチを取ることができる。本稿では、眼球とアイカメラの位置関係の変動に影響しない視線位置候補の取得に顕著性マップを用い、得られた視線位置候補と視線推定結果の対応付けから、ホモグラフィ変換を用いることで補正を実現している。

#### 4.2.1 顕著性マップを用いた視線位置候補の検出

視界の画像から顕著性マップを計算することにより、アイホール画像とは独立に視線位置候補を得る。顕著性マップに用いる特徴量、ならびに物体検出は、利用するコンテキストに依存して変更する。

#### 4.2.2 視線推定結果と顕著性マップの関連付け

従来手法による視線推定結果のうち、視点が停留し、かつ顕著性の高い位置付近を示しているものを、顕著性マップと関連付けることにより、顕著性マップ上でもっともら

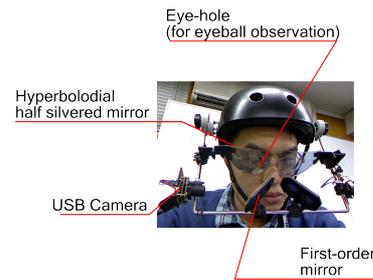


図 6 アイカメラ試作システム

Fig. 6 Prototype system of eye camera

表 1 試作システムのハードウェア構成

Table 1 Hardware Components of prototyping system

透明メタアクリル	
双曲面ハーフミラー	AL+SiO コーティング, 反射率 70%
	CPU : Intel(R)Core(TM)i7-3770
計算機	Memory : 8.0GB
	シキノハイテック, 52db
カメラ	1280×1024@30fps

しい注視点候補を取得することができる。これにより、広視野性を保持しつつ推定結果の補正を行うことができる。

#### 4.2.3 ホモグラフィ変換による視線推定結果の補正

視線推定結果と顕著性マップの組を複数取得することにより、視線推定の結果から、顕著性マップにて検出した視線位置へのホモグラフィを計算することができる。このホモグラフィ変換を視線推定結果に適用することで、視線推定結果の補正を行うことができる。

### 4.3 試作システムの実装

本研究の提案手法に沿って、試作システムを実装した。ベースとなった住谷ら、森らのアイカメラおよび視線推定手法から変更した部分を中心に実装の概要を説明する。

#### 4.3.1 アイカメラとハードウェア構成

試作したアイカメラは、双曲面ハーフミラー、ヘルメット、小型カメラ、計算機により成る (図 6)。また、本稿での試作システムでは、アイカメラの出力を、USB2.0 アイソクロナス転送によりそのまま計算機への入力とした。双曲面ハーフミラー、カメラ、計算機について、表 1 に示した機器を用いた。

#### 4.3.2 視線推定手法

森らの視線推定手法をベースとした提案手法による視線推定の工程を図 7 に示す。学習および推定に用いる予定の眼球画像を取得するたび、その画像に対して大津の手法により二値化閾値を決定し、得られた閾値に従って二値化を行い、学習および推定に用いる。これにより、常に最適な二値化を行うことができ、環境照度の影響を抑えることができる ([17])。二値化以降のノイズ除去、主成分分析、重回帰分析については森の手法に準ずる。

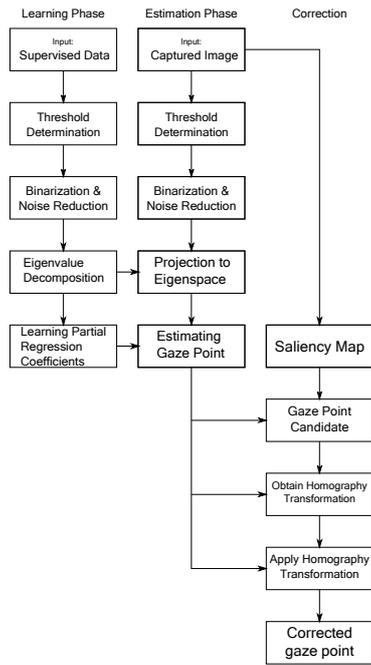


図 7 視線推定と推定結果補正手法  
Fig. 7 Eye gaze estimation and correction method

#### 4.3.3 推定結果の補正手法

##### 4.3.3.1 顕著性マップの計算

得られた視界画像に対し、顕著性マップを計算する。Ittiらの手法 [5] を用いる。

##### 4.3.3.2 推定結果と顕著性マップの対応付け

得られた視線推定の結果と顕著性マップを対応付ける。対応付けによって、視線推定結果と顕著性マップを条件付けた際にもっとも視線位置となる可能性の高い点を計算する。視線推定の誤差が正規分布に従うと仮定した場合、次式によって各点での視線確率分布が得られる。

$$p(\mathbf{x}|\mathbf{G}, \mathbf{x}_p) \propto \mathbf{G}(\mathbf{x}) \cdot \mathcal{D}$$

ここで、 $\mathbf{x}_p$  は視線推定結果、 $\mathbf{G}$  は得られた顕著性マップである。また、 $\mathcal{D} \sim \mathcal{N}(\mathbf{x}_p, \sigma^2)$  である。実装上は、右辺を最大化する点を視線候補点として得る。

##### 4.3.3.3 ホモグラフィ変換の計算と適用

ホモグラフィ変換を求めるために用いる視線推定結果と視線候補の組を選択する。このとき、選択される組の視線候補点を結んで得られる多角形の面積が、もっとも大きくなるものを選択する。抽出された点を用い、視線推定結果から、対応付けによって得られた視線候補へのホモグラフィ変換を計算する。ホモグラフィ行列を得た後、新たに視線推定を行った後、計算されたホモグラフィ変換を適用することで、補正が適用される。

## 5. 実験と考察

### 5.1 実験の目的

ホモグラフィによる推定結果の補正が有効であることを

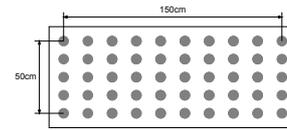


図 8 実験での注視点  
Fig. 8 Gaze point for experiment

確認するため、アイホール画像を人工的に平行移動させた場合、ならびにアイカメラを再装着した場合における補正の有効性を評価した。アイホール画像は、被験者から 60cm 離れた場所、図 8 に示す 50 箇所の注視点 (水平方向 100°, 鉛直方向 45° 相当) において、左右眼とも各 5 枚ずつ、計 500 枚を 1 セットとし、同じ被験者がアイカメラをセットごとに装着しなおし、2 セット取得した。ホモグラフィ変換の計算には、注視点のうち図 8 の四隅に属するものを 1 組ずつ、計 4 組を用いた。

### 5.2 評価に用いる視線推定の誤差

本実験では、アイカメラ搭載カメラ上での推定誤差を角度誤差  $\epsilon$  [degree] として評価する。 $\epsilon$  の算出方法を以下に示す。

カメラ上の注視点  $(x, y)$  に対応し、被験者視点を原点とする極座標系における視線方向  $(\theta, \phi)$  は次式で定まる。

$$\theta = \arctan \frac{-(y - y_c)}{(x - x_c)} \quad (14)$$

$$\beta = \arctan \frac{f}{\sqrt{(x - x_c)^2 + (y - y_c)^2}} \quad (15)$$

$$\phi = \arctan \frac{(b^2 + c^2) \sin(\beta) - 2bc}{(b^2 - c^2) \cos(\beta)} \quad (16)$$

ここで、 $f$  はカメラの焦点距離であり、 $(x_c, y_c)$  は全方位画像の中心のカメラ上での位置である。よって視線ベクトル  $\mathbf{V}$  は次式で表すことができる。

$$\mathbf{V} = \begin{bmatrix} \cos \theta \cos \phi \\ \sin \theta \cos \phi \\ \sin \phi \end{bmatrix} \quad (17)$$

正解の視線ベクトル  $\mathbf{V}$  に対し、推定結果の視線ベクトル  $\hat{\mathbf{V}}$  も同様にして計算可能である。角度誤差  $\epsilon$  は  $\mathbf{V}$  と  $\hat{\mathbf{V}}$  の成す角であり、次式で表される。

$$\epsilon = \arccos \frac{\hat{\mathbf{V}} \mathbf{V}}{|\hat{\mathbf{V}}| |\mathbf{V}|} \quad (18)$$

### 5.3 学習画像を加工した場合の視線推定

#### 5.3.1 実験の内容

学習フェーズで扱ったアイホール画像を人工的に平行移動させた場合において、ホモグラフィ変換における補正を適用した場合の有効性について検証した。学習データ 1 セットに対して、そのセットに含まれるアイホール画像を

表 2 アイホール画像を上下に平行移動させた場合の視線推定誤差 (deg)

Table 2 Estimation error when eye-hole image shifted vertically(deg)

移動量	補正前	補正後
上方 4	9.70	20.0
上方 3	7.38	13.5
上方 2	5.72	10.0
上方 1	3.94	4.09
0	2.65	3.19
下方 1	2.40	5.06
下方 2	3.43	9.00
下方 3	5.10	17.6
下方 4	6.90	27.6

表 3 アイホール画像を左右に平行移動させた場合の視線推定誤差 (deg)

Table 3 Estimation error when eye-hole image shifted vertically(deg)

移動量	補正前	補正後
左方 4	8.88	35.2
左方 3	6.85	20.1
左方 2	5.18	11.4
左方 1	3.59	7.76
0	2.65	3.19
右方 1	2.57	3.11
右方 2	3.34	5.54
右方 3	4.62	7.73
右方 4	6.33	9.93

上下左右の各方向に 1 ピクセルずつ、4 ピクセルまでの範囲で平行移動させたテストデータを用い、視線推定、ならびにホモグラフィ変換による補正を行った。平均推定誤差を表 2,3 に示す。

### 5.3.2 考察

いずれの移動量においても、補正後の推定誤差が補正前の推定誤差より小さくなると認められなかった。補正により誤差が極端に大きくなった実験について精査したところ、補正前の推定結果は、実際の注視点に対して線形に分布しておらず、ある特定の領域に集中していた。さらに、ホモグラフィ変換の計算に用いる点を四隅の点についても補正前の推定結果が集中している領域に推定結果が含まれていたため、集中している推定結果が四隅の強い影響力を受け、すべて変換の影響を受けたためであるといえる。すなわち、ホモグラフィ変換の計算に用いる組の選択に問題があったため、補正が正確に行われなかったものと考えられる。

## 5.4 アイカメラの再装着を行った視線推定

### 5.4.1 実験の内容

学習フェーズと推定フェーズの間に、アイカメラの再装

表 4 アイカメラを再装着した場合の視線推定誤差 (deg)

Table 4 Estimation error with reattached eye mark recorder(deg)

再装着	補正前	補正後
あり	26.5	29.4
なし	2.65	3.186

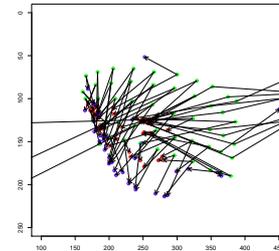


図 9 アイカメラを再装着した場合の推定結果と補正結果 (左眼)  
Fig. 9 Results of estimation and correction with reattached eye mark recorder(left eye)

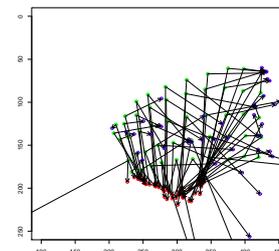


図 10 アイカメラを再装着した場合の推定結果と補正結果 (右眼)  
Fig. 10 Results of estimation and correction with reattached eye mark recorder(right eye)

着を行った場合の視線推定、ならびにホモグラフィ変換による補正を行った。比較対象として、学習フェーズと推定フェーズでまったく同じアイマーク画像のセットを用いた結果も示す。平均推定誤差を表 4 に、推定結果と補正結果を図 9. 12 に示す。図 9. 12 は、緑色の丸が実際に注視していた点、赤色の丸が補正前の視線推定結果、青色の丸が補正後の視線推定結果を表す。

### 5.4.2 考察

補正後の推定誤差が、補正前の推定誤差より小さくなることは認められなかった。再装着した場合における、アイホール画像の変動を精査したところ、下方方向に 16 ピクセル、左方向に 4 ピクセルの変動があり、さらに水平方向に対して約 9%の縮小が認められた。これらの変動は、5.3 節での変動量より大きく、したがって 5.3.2 と同様、ホモグラフィ変換を計算するために使う推定結果と注視点の組の選択が不適切であったためであると考えられる。

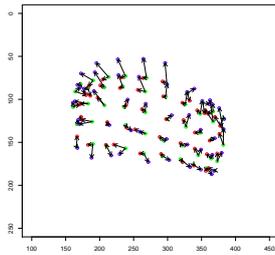


図 11 変動がない場合の推定結果と補正結果 (左眼)

Fig. 11 Results of estimation and correction without shifting(left eye)

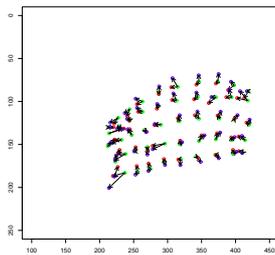


図 12 変動がない場合の推定結果と補正結果 (右眼)

Fig. 12 Results of estimation and correction without shifting(right eye)

## 6. おわりに

本研究では、アイカメラと頭部の位置関係の変動により、アピランススペースでの視線推定精度が悪化する問題に対して、位置関係の変動とは独立な視線位置候補を利用した視線推定精度の補正手法を提案した。提案手法は、顕著性マップと視線推定結果を対応付けて視線位置候補を取得することで、視線推定結果から視線位置候補へのホモグラフィ変換を得る手法である。実験の結果、ホモグラフィ変換による補正が推定精度の改善に有効であるとは認められなかった。

今後の課題としては、ホモグラフィ変換などの線形変換以外の、非線形変換による補正手法の検討が考えられる。補正前の視線推定結果が注視点に対して線形に分布していないことから、有効性を検証する必要がある。正確な変換による補正が実現することで、再装着や頭部の運動などでアイカメラと頭部の位置関係が変動しても推定精度が悪化しない視線推定を行うことが可能である。

## 参考文献

[1] Avraham, T. and Lindenbaum, M.: Esaliency (extended saliency): Meaningful attention using stochastic image modeling, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 32, No. 4, pp. 693–708 (2010).

[2] Baluja, S. and Pomerleau, D.: Non-intrusive gaze tracking using artificial neural networks, Technical report, DTIC Document (1994).

[3] Duchowski, A.: *Eye tracking methodology: Theory and practice*, Vol. 373, Springer (2007).

[4] Harel, J., Koch, C. and Perona, P.: Graph-based visual saliency, *Advances in neural information processing systems*, pp. 545–552 (2006).

[5] Itti, L., Koch, C. and Niebur, E.: A model of saliency-based visual attention for rapid scene analysis, *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 20, No. 11, pp. 1254–1259 (1998).

[6] Morency, L.-P., Christoudias, C. M. and Darrell, T.: Recognizing gaze aversion gestures in embodied conversational discourse, *Proceedings of the 8th international conference on Multimodal interfaces*, ACM, pp. 287–294 (2006).

[7] Mori, H., Sumiya, E., Mashita, T., Kiyokawa, K. and Takemura, H.: A Wide-View Parallax-Free Eye-Mark Recorder with a Hyperboloidal Half-Silvered Mirror and Appearance-Based Gaze Estimation, *Visualization and Computer Graphics, IEEE Transactions on*, Vol. 17, No. 7, pp. 900–912 (2011).

[8] Ono, Y., Okabe, T. and Sato, Y.: Gaze estimation from low resolution images, *Advances in Image and Video Technology*, Springer, pp. 178–188 (2006).

[9] Otsu, N.: A threshold selection method from gray-level histograms, *Automatica*, Vol. 11, No. 285–296, pp. 23–27 (1975).

[10] Schiele, B. and Waibel, A.: Gaze tracking based on face-color, *Proceedings of the International Workshop on Automatic Face-and Gesture-Recognition*, Citeseer, pp. 344–349 (1995).

[11] Sugano, Y., Matsushita, Y. and Sato, Y.: Calibration-free gaze sensing using saliency maps, *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, IEEE, pp. 2667–2674 (2010).

[12] Sumiya, E., Mashita, T., Kiyokawa, K. and Takemura, H.: A wide-view parallax-free eye-mark recorder with a hyperboloidal half-silvered mirror, *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*, ACM, pp. 19–22 (2009).

[13] 萩原亨: 運転者の視点に関する研究 (1992).

[14] Yamazawa, K., Takemura, H. and Yokoya, N.: Telepresence system with an omnidirectional HD camera, *Proc. 5th Asian Conf. on Computer Vision (ACCV2002)*, Vol. 2, pp. 533–538 (2002).

[15] Yamazawa, K., Yagi, Y. and Yachida, M.: Omnidirectional imaging with hyperboloidal projection, *Intelligent Robots and Systems' 93, IROS'93. Proceedings of the 1993 IEEE/RJS International Conference on*, Vol. 2, IEEE, pp. 1029–1034 (1993).

[16] 宮里洸司, 木村昭悟, 高木茂, 大和淳司, 柏野邦夫: MCMC-based particle filter を用いた人間の映像注視行動の実時間推定, 画像の認識・理解シンポジウム (MIRU2009) 予稿集 (2009).

[17] 重野孝明: HDR カメラを用いた環境照度にロバストな広視野アイマークレコーダ (2013).

[18] 福地賢, 宮里洸司, 赤嶺一馬, 木村昭悟, 高木茂, 大和淳司, 柏野邦夫: グラフコストの逐次更新を用いた映像顕著領域の自動抽出, 電子情報通信学会論文誌 D, Vol. 93, No. 8, pp. 1523–1532 (2010).

[19] 野呂影勇: 図説エルゴノミクス入門, 培風館 (2003).