

眼球の表面反射と高速アクティブ光投影を用いた 非装着・事前校正不要な注視点推定

中澤 篤志^{†,††} ニチュケ クリスティアン[†] ラドコフ アレクサンダー^{†,†††} 竹村 治雄[†]

[†] 大阪大学サイバーメディアセンター 〒 560-0043 豊中市待兼山町 1-32

^{††} 科学技術振興機構さきがけ 〒 332-0012 埼玉県川口市本町 4-1-8

^{†††} (株) ソニー・コンピュータエンタテインメント 〒 107-0052 東京都港区赤坂 8-1-22

E-mail: †{nakazawa,christian.nitschke,takemura}@cmc.osaka-u.ac.jp, ††Alexander_Radkov@hq.scei.sony.co.jp

あらまし 注視点推定技術は心理学やライフサイエンス、マーケティングやユーザインタフェースの分野で広く使用されている。従来の注視点推定システムは、ユーザの装着装置が必要だったり、システムの使用前に校正（キャリブレーション）を行う必要があり使用が煩雑であるという問題点がある。またこの校正は、ユーザの前面に仮想的な平面を想定し、眼球の角度と平面上の位置を対応づけるという原理に基づいており、ユーザが校正時と異なる環境に移動したり複雑な奥行きを持つ環境では大きな誤差が生じる。本論文ではこれらの問題を解決するため、眼球表面上で反射する環境光の反射像を用いた新しい注視点検出手法を提案する。本手法は眼球に入射する環境の反射光から直接注視点を推定するため、従来法の問題点を解決することができる。まず我々は眼球の幾何モデルに基づき、注視点からの光が眼球表面反射画像中で反射する位置 (Gaze Reflection Point (GRP)) を求める手法を開発した。これにより、眼球表面反射画像中の GRP と環境画像中の対応点を求めることで注視点を推定することが可能だが、実際の眼球表面反射画像は輝度・解像度ともに低く、品質の良い画像が得にくいいため直接的な比較は困難である。これを解決するため、環境に高速度のパターン光を投影することで、眼球表面反射画像と環境中の点を高速かつ頑健に対応付ける手法を開発した。我々は高輝度 LED を格子状に配置しパターン光を高速にコントロールする高速度パターン光プロジェクタを開発し、高速かつ頑健な注視点検出を実現した。複数の被験者による注視点推定結果により、事前の校正や装置装着を必要とせず、従来法と同程度の精度で注視点を推定できることを確認した。

キーワード 注視点検出、表面反射、眼球、アクティブライティング

1. はじめに

環境中から人の注視点を検出する技術は、心理学やライフサイエンス、マーケティングやユーザインタフェース等様々な領域で利用されている。そのため、商用も含め数多くの注視点検出システムが提案されている [1]。従来システムは用途に応じて様々なものが提案されているが、いくつかの問題点も指摘されている。まず、多くのシステムはアイカメラ等のデバイスを装着しなければならず、また装着機器が不要なシステムであっても事前の校正処理が必要である。この校正は、ユーザの前面に仮想的な平面を想定し、アイカメラによって検出された眼球の角度と平面上の位置を対応づける原理に依っているため、ユーザが校正時と異なる環境に移動したり複雑な奥行きを持った環境の場合には大きな誤差が生じる。

これらの問題点に対し本論文では、新たな注視点検出手法として、眼球表面における環境の反射像を用いる手法を提案する (図 1)。すなわち、眼球表面上で反射する環境の像と、環境カメラで撮影した環境の画像とを直接対応付けすることで注視点検出を実現する。このようなアプローチを取ることにより、以下のような長所が得られる。

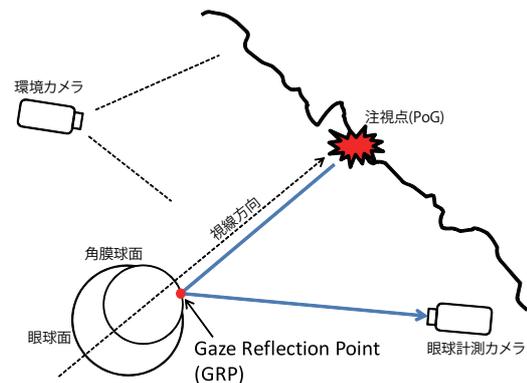


図 1 提案する眼球表面反射による注視点検出手法。眼球計測カメラから瞳の輪郭を抽出し眼球の姿勢を計測する。注視点 (PoG) からの光は眼球表面上の Gaze Reflection Point (GRP) で反射しカメラに入射する。提案する注視点検出法は、眼球表面反射像中の GRP と環境画像中の対応点を求めることで実現される。

キャリブレーション不要 本手法は環境画像中の点と眼球表面反射像中の点の直接対応づけで実現できるため、事前の校正が不要である。

装着機器不要 眼球の表面反射像はカメラにより遠隔から撮影できるため、機器の装着が不要である。

複雑な環境への対応 従来法はユーザの全面に仮想的な平面を仮定し注視点を検出するものであるため、キャリブレーションとは異なる環境では正しく動作しない場合がある。本手法は、環境画像と眼球表面反射画像を画像特徴レベルで比較するためこのような問題が生じない。

この推定方式を実現するために、以下の2手法を新たに開発した。

第1に、眼球表面上で視線方向から来る環境光を反射する点 (Gaze Reflection Point (GRP)) を眼球の幾何モデルを用いて求める手法を提案する。これにより、眼球表面反射画像での GRP における画像特徴と環境画像の画像特徴を比較することで、環境中の注視点を推定することができる。本論文では、透視投影に基づいた解法および弱透視投影に基づいた近似解法により GRP を導出する手法について述べる。

眼球の表面反射光は弱く解像度も高くないため、通常の撮影では画像特徴を用いてマッチングすることは困難である。そこで第2に、環境に対してパターン光を投影することで、眼球表面反射画像と環境画像をロバストに対応付けする手法を提案する。注視点の動的な動きに対応するため、非常に高速に時系列パターンを投影できる照明システム (High Framerate Programmable Illumination Projector (HF-PIP)) を新たに開発し、高速度カメラと同期して撮影するシステムを実現した。HF-PIP は格子状に設置された高輝度 LED、レンズ系、コントローラからなり、任意のパターン光を数百 Hz で投影することが可能である。この仕組みにより、眼球表面反射画像と環境の画像上の点を高速に対応付けし、注視点の検出を行う事が可能である。

以降2章では関連研究について述べ、3章では眼球の幾何モデルおよび GRP の導出について述べる。4章では新規開発した HF-PIP、暗瞳孔法による眼球姿勢の検出法およびシステム全体の構成について述べ、5章~7章で実験結果および考察を行う。

2. 関連研究

視線 (注視点) 検出においてよく知られている手法は、ユーザに頭部にアイカメラを装着し、頭部に対する眼球の角度を計測するものである [2] [3]。環境を計測する別のカメラがユーザの前方を観察し、注視点はこの画像中の位置として得られる。この方式はユーザの移動に対応できるが、機器の装着が必要であり、また環境カメラと眼球方向に対する校正処理が必要である。また計測精度は環境の形状や奥行きに左右される。別種のシステムでは、パン・チルト・ズーム (PTZ) ユニット等を装着したりモートカメラを用いて眼球画像を撮影するシステムもある [4] [5] [6]。この方式の場合機器装着の手間はないが、事前の校正は必要であり、また環境の形状に対応で

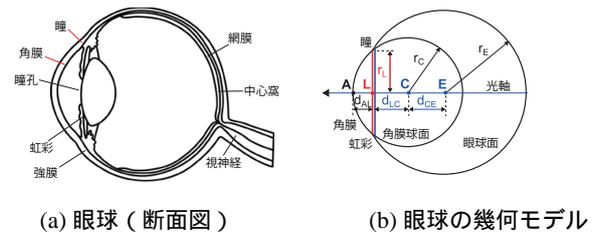


図2 眼球の幾何モデル

きない点に対しても同様の問題がある。

眼球の表面反射画像を用いて、人の観察環境を再現する試みも行われている [7]。しかしこれらの一連の研究では注視点計測は行っておらず、また画像は近接した位置から十分にセットアップされた環境で撮影されたものを用いている。

近年ではユーザーインタフェース等の用途を想定し、PCのマウスのクリックや画面のサリエンスマップと目領域の画像を学習することで注視点検出を実現するアプローチが提案されている [8] [9]。これらの手法はキャリブレーション等は不要であるが、想定された使用シナリオは限られており、また人の認知特性に依存しているため、提案手法の適用範囲とは異なる。

3. 眼球の表面反射を用いた注視点計測

図1に、提案する注視点計測の概念図を示す。環境中の注視点 (Point of Gaze (PoG)) からの光は眼球表面上で反射し、眼球計測カメラに入射する。この眼球 (角膜) 表面上の反射点を Gaze Reflection Point (GRP) と呼ぶ。提案する注視点計測は、眼球計測カメラでの GRP での環境反射像と、環境カメラの画像との対応を取ることで実現できる。

GRP を得るためには、眼球計測カメラに対する眼球の姿勢推定が必要である。以降ではまず眼球の幾何モデルおよび眼球の姿勢推定について述べ、次に2通りの GRP 導出法を述べる。

3.1 眼球の幾何モデル

図3.(a),(b) に実際の眼球および近似した幾何モデルを示す。ここから分かるように、眼球は単一の球面ではなく、角膜部分の球面 (角膜球面: Corneal sphere) と眼球本体 (眼球面: Eyeball sphere) の2つの異なる球面から構成した方が良く近似できる。この近似モデルにはいくつかのパラメータからなり、本研究では全ての実験において既定の値 ($r_L = 5.6mm, r_C = 7.7mm$) を用いる [10]。この幾何モデルは実際の眼球と完全に一致している訳ではないが、幾何的に単純であり解析的に扱いやすいという利点がある。また従来研究においても用いられており [11]、眼球の表面反射解析において十分な精度を持つと考えられる。

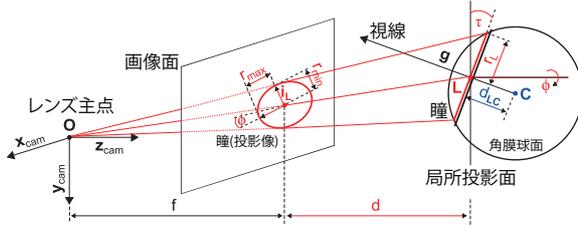


図3 瞳の投影像による眼球姿勢推定

3.2 眼球の姿勢推定

次に、眼球計測カメラ画像からの眼球の姿勢推定手法について述べる。弱透視投影モデルを仮定すると、瞳は中心 i_L 、半径 r_{max}, r_{min} 、回転角 ϕ の5パラメータで表される楕円として投影される(図3)。ここから、瞳中心の3次元位置 L からカメラ中心までの距離 d は $d = f \cdot r_L / r_{max}$ で得られる。ただし f はカメラの焦点距離である。

視線方向 g は以下で得られる。

$$g = \begin{bmatrix} \sin \tau \sin \phi & -\sin \tau \cos \phi & -\cos \tau \end{bmatrix}^T. \quad (1)$$

ここで τ は瞳の奥行き方向の傾きを表しており、 $\tau = \pm \arccos(r_{min}/r_{max})$ で求められる。またここから角膜球面の中心点 C は、 L から C までの既知の長さ $d_{LC}(=5.6\text{mm})$ と視線ベクトル g を用いて求めることができる。

3.3 眼球の表面反射と GRP の導出

眼球計測カメラによる眼球表面反射像の撮影は、反射光学系 (Catadioptric system) と同様であると見なせ、GRP は反射球面である角膜球面の幾何モデルと、眼球計測カメラから得られた眼球姿勢を用いて求めることができる。本論文では GRP の解法として、透視投影モデルに基づいた手法と弱透視投影モデルに基づいた手法を導出する。前者はより正確な推定モデルであるがカメラの内部パラメータの校正が必要であり、後者はカメラの校正は不必要だが、弱透視投影の性質上、瞳を平行投影する近似投影面での瞳の奥行き差と、カメラと近似投影面との距離の比に応じて誤差が大きくなる。しかし一般的な実験設定では、後者における誤差が問題になることは少ないと考えられ、計算量の差を考えると後者の近似モデルで十分である。

求めたい点は画像中の GRP の位置であり、これは GRP の3次元座標を求め、画像座標系に投影することで得られる。GRP の導出は3次元空間中の光線の軌跡を考える必要があるが、眼球への入射光と反射光を含む面を考えることで2次元平面上で解くことができる。以降では、3次元座標でのカメラ主点を O 、角膜球面の中心点 C 、注視点を P 、GRP を S 、および GRP における角膜球面の法線方向を n_S とし、これを前述した平面上で表した2次元点をそれぞれ O', C', P', S', n'_S と表す(図4)。こ

れらの間の変換を表す回転行列 R は以下で得られる。

$$R = \begin{bmatrix} r_x & r_y & r_z \end{bmatrix}^T \quad (2)$$

$$r_x = -\frac{C}{\|C\|}, \quad r_y = r_z \times r_x, \quad r_z = \frac{r_x \times g}{\|r_x \times g\|}. \quad (3)$$

3.3.1 透視投影モデルによる GRP の導出

眼球計測カメラのレンズ主点を原点とし、前述した反射系を含む2次元平面上の点を図4に示すように取ると、各点は以下のように表せる。

$$C' = \begin{bmatrix} -d_{OC} \\ 0 \end{bmatrix}, \quad g' = \begin{bmatrix} \cos \alpha \\ \sin \alpha \end{bmatrix}, \quad n_{S'} = \begin{bmatrix} \cos \beta \\ \sin \beta \end{bmatrix},$$

$$P' = C' + d_{CP} \cdot g', \quad S' = C' + r_C \cdot n_{S'}, \quad (4)$$

またこの平面上での視線方向の投影角 α は、以下で得られる。

$$\alpha = \arccos g'_x. \quad (5)$$

GRP となる角膜球面上の角度 $\beta \in [0, \alpha]$ は、PoG と眼球の距離によって異なる。GRP から PoG への直線の式は以下で得られる。

$$L(\beta) = S' + t_2 r'_2 \quad (6)$$

すなわち、求める角度 β は PoG と直線 L が最も近くなる角度 (β_0) を求めれば良い。

$$\beta_0 = \arg \min_{\beta} d(\beta). \quad (7)$$

ここで、直線 L と点 P との距離 d_{PL} を考える(図5(a))。また角膜表面での反射を考えると、角膜球面での入射光、反射光 r'_1, r'_2 も得られる。

$$d_{PL}(d_{CP}, \beta) = (r'_2 \times (P' - S')) / (r'_2 \cdot g'). \quad (8)$$

$$r'_1 = S' / \|S'\|. \quad (9)$$

$$r'_2 = 2(-r'_1 \cdot n_{S'}) n_{S'} + r'_1. \quad (10)$$

d_{PL} を d_{CP} で正規化することで、図5(a)における $d_{\alpha}(d_{CP}, \beta)$ が得られる。

$$d_{\alpha}(d_{CP}, \beta) = \tan^{-1}(d_{PL}(d_{CP}, \beta)/d_{CP}). \quad (11)$$

d_{CP} を十分遠方にとることにより β_0 は最小値に収束する。これは直線 L と P' は無限遠点で交差する状況である(図5(b))。すなわち、求める値 β_0 は、 β に対する以下の距離関数の根として得られる。

$$f(\beta) = \lim_{d_{CP} \rightarrow \infty} d_{\alpha}(d_{CP}, \beta) \quad (12)$$

ここから以下の解が得られる。

$$\beta_0 = \tan^{-1} \left(\frac{\tan \alpha (-2R^2 \cdot d_{OC} + R \cdot r_C + d_{OC})}{-2R \cdot d_{OC} + r_C}, R \right) \quad (13)$$

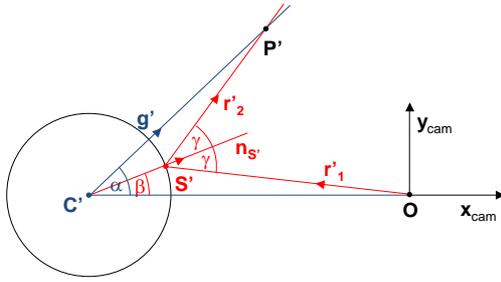


図4 眼球表面反射の関係．入射光・反射光は単一の面内を通過する．

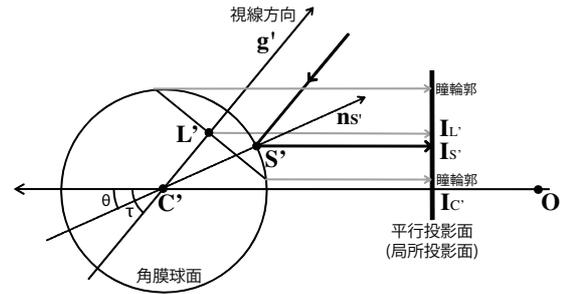
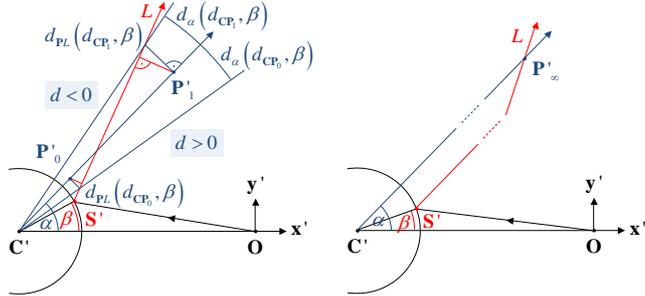


図6 弱透視投影モデルによる眼球（角膜）表面反射と GRP の関係．注視点からの光は視線方向と平行に入射し GRP で反射しカメラに入射する．



(a) Distance in visual angle (b) Case of unknown PoG

図5 注視点 P' に対する GRP S' (β) の推定．(a) 距離 f と角度の関係，(b) PoG が無限遠点にあるときの仮定．

ただし R は以下の 4 次方程式の解である．

$$\begin{aligned}
 P_4(x) &= a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 = 0 \quad (14) \\
 a_4 &= 4d_{OC}^2 (\tan^2 \alpha + 1), \\
 a_3 &= -4r_C \cdot d_{OC} (\tan^2 \alpha + 1), \\
 a_2 &= (r_C^2 - 4d_{OC}^2) (\tan^2 \alpha + 1), \\
 a_1 &= 2r_C \cdot d_{OC} (\tan^2 \alpha + 2) a, \\
 a_0 &= d_{OC}^2 \tan^2 \alpha - r_C^2.
 \end{aligned}$$

ここから 4 つの解 $x_{01}, x_{02}, x_{03}, x_{04}$ が解析的解法 [12] により得られ，対応する値 $\beta_{01}, \beta_{02}, \beta_{03}, \beta_{04}$ が得られる．最終的な解は以下の式で得られる．

$$\beta_0 = \begin{cases} \beta_{01} & \text{if } \beta_{01} > \beta_{02}, \\ \beta_{02} & \text{otherwise.} \end{cases} \quad (15)$$

カメラ画像における GRP の位置 $S'(\beta_0)$ は式 (4) で得られ，これを先に得られた回転行列 R およびカメラへの投影行列を用いて得ることができる．

3.3.2 弱透視投影モデルによる GRP の導出

眼球のカメラへの投影に弱透視投影モデルを導入し，PoG が視線方向と並行な直線上にあると仮定すると，GRP を大幅に少ない計算量で得ることが可能である．この仮定は，カメラと眼球との距離が角膜球面の大きさに比べて十分大きく，注視点の大きさが角膜球面の大きさに比べて十分大きいことを想定しているが，殆どの利用

ケースで成り立つと考えられる．

透視投影モデルによる導出法と同様に，この場合でも全ての反射は 1 つの平面内で起こると考えることができる．特に弱透視投影モデルの場合，この面は，瞳の投影像（楕円）の短軸を含み，画像面に対して垂直な平面である．図6に，この仮定を用いた場合の光線の経路を示す．まず，角膜表面での反射から以下の式が得られる．

$$C' \cdot ns' = g' \cdot ns'. \quad (16)$$

$$\begin{aligned}
 C' &= \begin{bmatrix} 0 & 1 \end{bmatrix}^T, \quad g' = \begin{bmatrix} \cos \tau & \sin \tau \end{bmatrix}^T, \\
 ns' &= \begin{bmatrix} \cos \theta & \sin \theta \end{bmatrix}^T.
 \end{aligned}$$

ここから，視線方向を表す τ を用いて GRP を表す角度 θ が得られる．また，弱透視投影の局所投影面上における瞳の中心と GRP の距離 $|I_{S'} - I_{L'}|$ を求めると，

$$\theta = \arctan((1 - \cos \tau) / \sin \tau), \quad (17)$$

$$|I_{S'} - I_{L'}| = r_c \cos(\tau - \theta) - d_{LC} \sin \tau, \quad (18)$$

となる．

次に，瞳の中心を物体中心と見なした弱透視投影を考えると，眼球計測画像での GRP の位置 $i_{S'}$ は以下で求められる．

$$i_{S'} = i_L + s \cdot v_{sm} |I_{S'} - I_{L'}|, \quad (19)$$

$$s = r_{\max} / (r_L \sin \tau). \quad (20)$$

ここで s は弱透視投影のスケール係数であり， i_L は瞳中心の座標， v_{sm} は瞳の投影像（楕円）の短軸を表す二次元ベクトルである．

4. システムの実装

提案手法を実装しシステムを構築した．システムの実装において問題となるのが，以下のような問題である．

1. 眼球の姿勢をどのように頑健に推定するか？
2. 環境中の注視点と眼球表面反射像の GRP をどのようにして頑健に対応付けるか？

第1の問題に対する解決法として、赤外光投影によりまず瞳孔を検出し、次に瞳の輪郭を検出する2ステップのアルゴリズムを用いた。瞳孔抽出は暗瞳孔法 [13] よりロバストに行えるため、瞳孔輪郭をまず検出しておくことで瞳の輪郭抽出も安定して実現できる。第2の問題に対する解決法として、環境に時系列パターン光を投影することで、環境の3次元点と眼球反射像の対応付けを取る手法を提案する。これにより、環境光の輝度が低い場合でも高速・ロバストに対応付けを行う事ができ、また時系列パターンを使うことで角膜のテクスチャの影響も除外できる。我々はパターン照明装置として、通常のデータプロジェクタに加え、新規に開発したLEDアレイによるプロジェクタも用いたシステムを実装した。

4.1 システム構成

提案システムは、眼球計測カメラ、環境カメラおよびプロジェクタからなる(図7)。また、眼球計測カメラの周囲には赤外LEDが取り付けられている。全システムはマイクロコントローラで制御されており、同期動作する。

眼球計測カメラは、目領域を撮影し赤外LEDを用いたアクティブライティングにより、瞳孔と瞳を検出しカメラに対する眼球の姿勢を計測する。プロジェクタは環境に時系列パターンを投影する。

システムの動作シーケンスを図8に示す。最初の2フレームで瞳孔、瞳の検出を行い、その後環境へのパターン投影が行われる。このシーケンスを繰り返すことで動的なトラッキングを行うことができる。

4.2 赤外光照明による瞳孔・瞳の輪郭検出

瞳孔と瞳の輪郭検出は赤外光を用いて実現する。まず、暗瞳孔法を用いて瞳孔領域を抽出する。これは瞳孔と他の領域との光の反射・吸収特性の違いを利用した手法である [13]。眼球計測カメラに取り付けられた赤外LEDは、第1フレームではOFFとなり、第2フレームではONとなる。眼球領域の中で、瞳孔は外部からの光を吸収するため両フレームとも黒く写るが、他の領域は光を反射するため第2フレームでは明るく写る(図9-(a),(b))。ここから、第1フレームと第2フレームの撮影画像の差分を取ることで、値の小さな領域を瞳孔領域と見なすことができる。

次に、この領域の輪郭に対してRANSACによる楕円フィッティングを適用し、楕円 $B(x, y, c_x, c_y, a, b, \phi)$ を得る。ただし (c_x, c_y) は中心位置、 a, b は短径および長径、 ϕ は回転角である。

瞳輪郭の検出は、瞳孔輪郭の推定結果を用いて2ステップの最小化を行い推定する。まず、瞳孔の中心および回転角 (c_x, c_y, ϕ) を用い、径のパラメータ (a, b) のみを変化させ以下の関数を最小化し、瞳の径の初期パラメータ a_0, b_0 を得る。

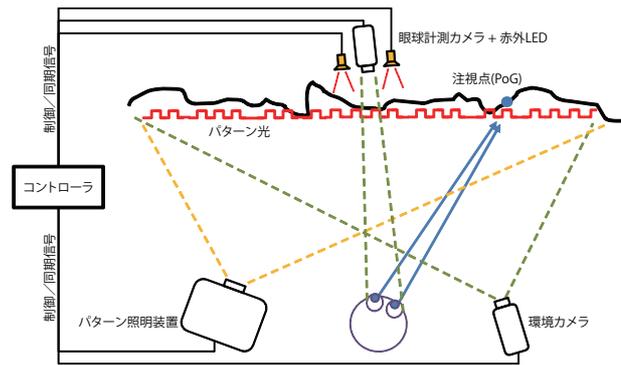


図7 システム構成。本システムは眼球計測カメラ、赤外線LED、環境カメラとプロジェクタからなる。全装置はマイクロコントローラによって同期動作する。

Frame #	1	2	3	4	5	...
赤外LED	OFF	ON	OFF	OFF	OFF	...
HF-PIP	OFF	OFF	ALL ON	CODE#1	CODE#2	...

図8 高速度パターンプロジェクタ(HF-PIP)による実装の場合の動作シーケンス。フレーム1,2で眼球計測カメラに付けられた赤外LEDが点滅し、眼球の姿勢推定を行う。フレーム3はHF-PIPのLEDが全点灯し輝度補正用のリファレンス画像を取得する。その後HF-PIPの時系列パターンを投影する。

$$eval(c_x, c_y, a, b, \phi) = \sum_{x,y} E_x(x, y) \cdot \text{sgn}(c_x - x) \cdot B(x, y, c_x, c_y, a, b, \phi) \quad (21)$$

$$[a_0, b_0] = \arg \max_{a,b} eval(c_x, c_y, a, b, \phi) \quad (22)$$

ここで $E_x(x, y)$ は入力画像の x 微分であり、また $a_0 > a, b_0 > b$ である。画像の x 微分のみを用いるのは瞼の影響を避けるためである。また、 $\text{sgn}(c_x - x)$ は、瞳の中心座標から外側方向に黒目から白目への変化があることを評価するためである。次に、得られた径の初期パラメータを用いて、楕円の全パラメータを同じ評価関数で収束させ、瞳輪郭のパラメータ $(c'_x, c'_y, a', b', \phi')$ を得る。

4.3 パターン照明装置

パターン光による画像間の対応点探索には複数枚の画像が必要なため、眼球のように動的に動く対象に対して適用するためには、高速な投影・撮影系が必要である。Narashimhanらは市販のDLPプロジェクタを用いて高速なパターン光投影を実現し、高速アクティブビジョンシステムを可能にした [14]。しかし、投影領域の広さは限られており、高速な投影を実現する反面、投影輝度は大幅に低下している。これに対し我々は高輝度の時系列パターン光を高速で投影するため、高輝度LEDにより構成する高速度パターン光プロジェクタ(High-framerate

Programmable Illumination Projector (HF-PIP)) を設計・開発し、高速な注視点検出手法を実現した。

本論文では、特性の異なるプロジェクタを用いることで推定精度がどの程度変化するかを評価するため、一般のデータプロジェクタおよび HF-PIP による実装の 2 種を実装した。

4.3.1 データプロジェクタによる実装

データプロジェクタを用いることで提案手法を容易に実装することが可能だが、一方でフレームレートや輝度、投影領域の広さに対して限界がある。本研究では、データプロジェクタが高解像度であることを利用し、PoG の推定精度を検証する目的のために導入した。使用する照明パターンとしては、3 次元形状計測で一般的に用いられる x, y 方向のグレイコードパターンを用いた。実験では両方向に 10 ビットのパターンを用いたため、1 推定に 20 フレーム必要である。すなわちシステムが 60Hz で動作する場合 3 fps で計測できる。

4.3.2 高速度パターン光プロジェクタによる実装

HF-PIP は、格子状に配置された高輝度 LED 群、レンズ、コントローラからなる (図 10)。各 LED は 0.05 ms (最速) で点滅をコントロール可能であり、LED 毎に異なる時系列の 2 値パターンを投影することができる。また各 LED の方向を個別に変えることができるため、広い領域をカバーするように照明光を配置できる。

各々の LED の時系列照明パターンにユニークな ID をもたせる事で、カメラの撮影画像列から ID を復元し光源となる LED を同定することができる (図 9(d)-(f))。一方、環境中の空間解像度は LED の数により決まるため、一般的にはデータプロジェクタよりも低い。この欠点を補うため、線形補間を組み合わせることで PoG の推定を行う。すなわち、眼球反射画像中において GRP から近い 3 つの LED の投影点を求め、それらと GRP との相対的な位置関係を求める。環境画像中からも各 LED の投影点を求め、既に得られた GRP との相対位置関係を用いて、環境画像中の注視点を推定する。

処理の流れは以下ようになる。まず、眼球反射画像列から時系列コード復元を行い、ピクセル毎の ID を求める。復元された ID を元にラベリングを行い、一定以上の面積を持つラベルを求め、その重心位置を各 LED に対する表面反射点とする。時系列コードとしては Hamming(7,4) コードを用いている [15]。これは、HF-PIP の各 LED の投影点は撮影画像の中では連続していないため、データプロジェクタの場合と異なりグレイコードではコード復元の頑健性がないためである。Hamming(7,4) コードは 7 bit のベースコード (ID = 1 ~ 64) と 3 bit のパリティビットからなるため、データ長は 10 bit (フレーム) である。眼球姿勢復元および輝度補正のために 3 フレーム必要なので、13 フレームで 1 回の注視点推定を行う事ができる。つまりカメラを 140 Hz で動作させた場合、10.8 fps での推定が可能である。また撮影画像中で瞳領域の動き

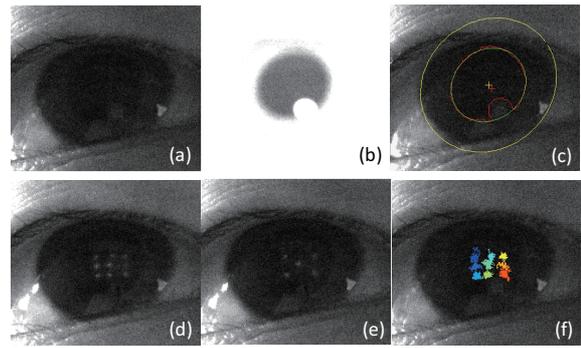


図 9 HF-PIP による実装での角膜表面反射光。(a) フレーム 1, (b) フレーム 2 (赤外光点灯), (c) 瞳孔および瞳輪郭の検出結果 (黄)。赤線は暗瞳孔法によって得られた瞳孔輪郭。黄十字マークは瞳輪郭の中心。赤十字マークは推定された GRP, (d) フレーム 3。(e) フレーム 13。(f) フレーム 3~13 を用いた表面反射光からのパターン復元結果 (9 LED)。異なる色は異なるパターン光 ID を示す。



図 10 高速度パターン光プロジェクタ (HF-PIP)。上段：光源 (LED)、下段左：パワーモジュール・コントローラ、下段右：実験環境と投影光 (9 LED)

を線形補完で追跡することで、中間フレーム (フレーム 4 ~ 13) における瞳の動きに対応している。

現在の HF-PIP は可視光 LED を使用しているが、赤外 LED に変えることも容易に可能である。これにより、ユーザが投影光に気づくことなく注視点推定を行う事ができる。

5. 実験結果

2 種類の実験を通して手法の評価を行った。実験 1 では、被験者が壁面上の点を注視する静的な状況での計測点の推定精度を評価した。この実験は、プロジェクタを用いたシステムと HF-PIP を用いたシステムの両者で実験を行った。これらの実装の間では、眼球の姿勢推定や

表 1 実験結果 (実験 1)

	誤差平均 [deg]	標準偏差 [deg]
被験者 1 (プロジェクタ)	2.37	-
被験者 2 (プロジェクタ)	2.91	-
平均 (プロジェクタ)	2.64	-
被験者 1 (HF-PIP)	3.11	0.953
被験者 2 (HF-PIP)	2.90	0.657
被験者 3 (HF-PIP)	4.60	0.896
被験者 4 (HF-PIP)	2.86	1.840
被験者 5 (HF-PIP)	4.69	0.730
平均 (HF-PIP)	3.63	-

GRP の推定手法の違いも評価するため、異なるアルゴリズムを適用した。プロジェクタによる実装では手動で瞳の輪郭抽出を行い、透視投影による GRP 推定手法を用いた。一方 HF-PIP による実装では、自動的な瞳輪郭抽出および弱透視投影による GRP 推定を用いた。実験 2 では、被験者が注視点を動かす状況での推定を行った。この実験はプロジェクタではフレームレートが不足するため、HF-PIP のみにより評価を行った。

実験環境を図 10 (下段右) に示す。プロジェクタによる実験は環境光を消した状況で、HF-PIP による実験は通常の屋内照明環境で実験を行った。被験者は壁面から約 2.8m 離れた位置に座り、壁面上の 20 点のマーカを注視する (実験 1)。眼球計測カメラは被験者の顔の少し下に位置し、顔とカメラとの距離は約 0.6 m である。環境カメラは被験者の後方に位置し、壁面のマーカが全て撮影できるように設置した。

実験機器として、NEC 社製 NP-U300XJD データプロジェクタ (1280 × 1024 pixels, 3000 lumen, 60 Hz)、Point Grey Dragonfly Express カメラ (640 × 480 pixels, 140Hz, B/W) を用いた。HF-PIP の光源としては、Philips 社製 Luxeon Rebel High-Power LED (3W, White) を用いている。パターン光投影、撮影および画像取得は PC (Intel Core i7-960 3.06 GHz, 8 GB RAM) によって行い、HF-PIP の画像処理は同 PC 上の Matlab R2010b および Image Processing Toolkit 環境で行っている。設計した HF-PIP は 42 個の LED まで同時に投影可能であるが、事前実験により LED を減らした状況でも同様の性能が得られる事が分かったため、本実験環境においては 9 つの LED を用いている。実験 1, 2 のマーカ配置および LED の投影位置を図 11, 12 に示す。

実験 1 においては、プロジェクタによる実装では 2 被験者、HF-PIP による実装では 5 被験者を対象に行った。各被験者は壁面上の各点を 20 フレーム (プロジェクタ, 1 計測)、300 フレーム (HF-PIP, 22-23 計測) の間注視する。また全ての計測は右目を用いて行った。表 1 および図 11 に推定精度 (誤差平均, 標準偏差) および実験結果をプロットした結果を示す。この座標系において、被験者の右目は $(x, y, z) = (0.9 \text{ m}, 0.3 \text{ m}, 2.8 \text{ m})$ に位置する。両実験結果のうち、被験者 1, 2 は同一人物である。

図 12 に実験 2 の結果を示す。この実験では被験者は、

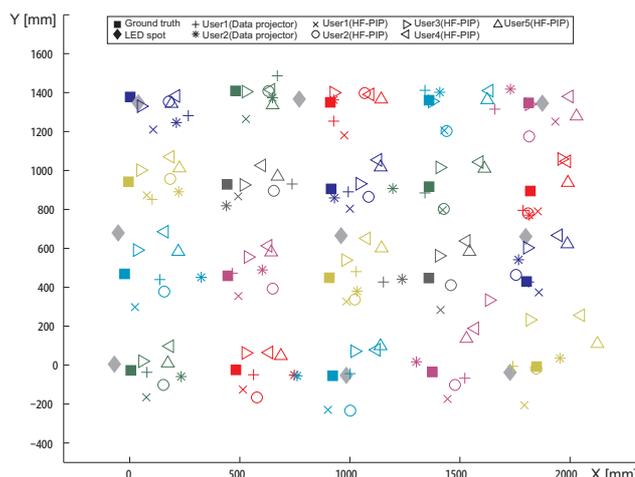


図 11 実験結果 (実験 1)。各シンボルは異なる被験者を表し、各色はマーカ位置を表している。

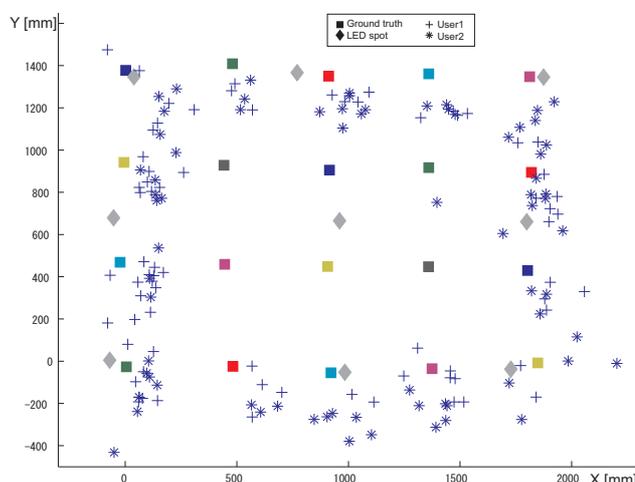


図 12 実験結果 (実験 2)。被験者は左上, 左下, 右下, 右上のマーカを通り、再び左上のマーカに戻るように視線を動かしている。異なるシンボルは異なる被験者を表す。

左上, 左下, 右下, 右上のマーカを通り、再度左上に戻るように矩形に注視点を動かしている。実験 2 は 2 被験者によって行い、各々 1000 フレーム (75-76 計測) の画像を取得し推定を行った。

6. 考 察

実験結果から、提案手法は従来手法と比較しても十分な推定精度があることを確認した。本手法はデバイスの装着や事前校正なしで実現されており、従来の注視点検出システムの問題点を克服しながら同様の結果を得たことは、注視点検出技術において大きな意味を持つものと考えられる。本手法では注視点推定のために複数枚の画像が必要だが、高速度な照明投光および画像取得により、注視点を動的に動かした状況でも安定して推定を行えることを確認した。

プロジェクタによる実装および HF-PIP による実装を

比べると精度に殆ど違いが無いことから、GRPの推定アルゴリズム(透視投影モデルと弱透視投影モデル)、プロジェクタの空間解像度、眼球輪郭抽出アルゴリズム(手動、自動)の違いは、本実験環境においては計測精度には殆ど影響しなかったと考えられる。ここから HF-PIP で実装した高速度推定、自動的な瞳輪郭抽出、弱透視投影による GRP 推定は有効であり、本システムが注視点検出システムとして実用に耐えるものであることが示されたと思われる。

実験1の結果を見ると、推定結果のほとんどは注視対象とするマーカよりも右側に位置している事が分かる。また、その誤差傾向は人により違いがある事も確認できる。この理由として、眼球の幾何的前方向ベクトル(Optical axis)と視線ベクトル(Gaze axis)の間の差に起因すると推測できる。この値は3-5度の個人差のあるパラメータである[1]。一方で、本実験結果から複数の被験者の間で同様の誤差傾向があることが分かったため、多くの実験データサンプルから統計的な誤差傾向を解析し補正を行う事で、より精度の高い推定結果を得られる可能性がある。

実験を通して、下側に位置するマーカの推定結果は上部マーカよりも悪くなる傾向が見られた。これは、下部のマーカを注視する場合には瞼があまり開かれない状況となり、瞳領域が眼球計測カメラから見えにくくなるため、瞳の輪郭が捉えづらく眼球的姿勢推定が不安定になるためである。本注視点推定法は瞳の輪郭に基づく眼球姿勢推定結果に依っているため、これを安定的に得る方法は重要であることが理解できる。

今回開発した HF-PIP は高輝度 LED を格子状に配置し、それを環境に投光する手法を用いたが、他の環境照光方法を使うことも可能である。具体的には、LED を格子状に配置せずそれぞれの LED を別々に環境の異なる所に取り付け照光する、発光するデバイスを環境中の点に直接取り付ける等の実装が考えられる。このように提案システムは、状況に応じて様々なパターン投影方法を取ることができるため、適切な発光方法を用いることで注視点検出システムの適用可能性を広げる事が可能である。

7. 結 論

本論文では、眼球の表面反射像を用いた新しい注視点計測手法を提案した。従来法と比較して、本手法はアイカメラ等の装着デバイスを必要とせず、事前校正も不要である。また、本手法は複雑な形状を持つ環境においても、理論上、計測精度が落ちないという利点もある。技術的な新規性として、眼球表面上で注視点からの光を反射する点である Gaze Reflection Point (GRP) というアイデアを提案し、眼球の幾何モデルと姿勢推定結果を用いてこの点を解析的に求める手法を提案した。また環境と眼球表面反射像の点をロバストに対応づけるために環境へのアクティブ照光を用いるアイデアを提案し、高輝度

LED とコントローラからなる照光システムを開発し、同期した高速度撮影と組み合わせることで高速な注視点計測を実現した。本手法および提案デバイスは提案した注視点検出タスクのみならず、他の計測システムにも応用可能な汎用性を持つと考えられる。複数の被験者による二種の実験により、本システムが従来法と同様の精度を持ち、また動的な眼球の動きに対応可能であることを示した。

謝辞 本研究は、JST 戦略的創造研究推進事業さきがけの一環として行われたものである。

文 献

- [1] D. W. Hansen and Q. Ji: "In the eye of the beholder: A survey of models for eyes and gaze", *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, **32**, 3, pp. 478-500 (2010).
- [2] D. Li, J. Babcock and D. J. Parkhurst: "openEyes: A low-cost head-mounted eye-tracking solution", *Proc. Symp. on Eye Tracking Research & Applications (ETRA)*, pp. 95-100 (2006).
- [3] Tobii Technology AB: "Tobii Glasses Eye Tracker Product Description", <http://www.tobii.com>.
- [4] S. M. Kim, M. Sked and Q. Ji: "Non-intrusive eye gaze tracking under natural head movements", *Proc. Int. Conf. of IEEE Engineering in Medicine and Biology Society (IEMBS)*, pp. 2271-2274 (2004).
- [5] D. H. Yoo and M. J. Chung: "A novel non-intrusive eye gaze estimation using cross-ratio under large head motion", *Computer Vision and Image Understanding (CVIU)*, **98**, 1, pp. 25-51 (2005).
- [6] M. Reale, T. Hung and L. Yin: "Viewing direction estimation based on 3D eyeball construction for HRI", *Proc. Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 24-31 (2010).
- [7] K. Nishino and S. K. Nayar: "Corneal imaging system: Environment from eyes", *International Journal on Computer Vision (IJCV)*, **70**, 1, pp. 23-40 (2006).
- [8] Y. Sugano, Y. Matsushita, Y. Sato and H. Koike: "An incremental learning method for unconstrained gaze estimation", *Proc. European Conf. on Computer Vision (ECCV)*, pp. 656-667 (2008).
- [9] Y. Sugano, Y. Matsushita and Y. Sato: "Calibration-free gaze sensing using saliency maps", *Proc. Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 2667-2674 (2010).
- [10] R. S. Snell and M. A. Lemp: "Clinical Anatomy of the Eye", Blackwell Publishing, Malden, 2nd edition (1997).
- [11] C. Nitschke, A. Nakazawa and H. Takemura: "Display-camera calibration from eye reflections", *Proc. Int. Conf. on Computer Vision (ICCV)*, pp. 1226-1233 (2009).
- [12] E. W. Weisstein: "'Quartic Equation.'" From MathWorld-A Wolfram Web Resource", <http://mathworld.wolfram.com/QuarticEquation.html>.
- [13] R. Kothari and J. L. Mitchell: "Detection of eye locations in unconstrained visual images", *Proc. Int. Conf. on Image Processing (ICIP)*, pp. 519-522 (1996).
- [14] S. G. Narasimhan, S. J. Koppal, and S. Yamazaki: "Temporal dithering of illumination for fast active vision", *Proc. European Conference on Computer Vision*, pp. 830-844 (2008).
- [15] T. K. Moon: "Error Correction Coding: Mathematical Methods and Algorithms", Wiley-Interscience (2005).