

First Person Vision のための Inside-Out カメラの提案

清水 彰一^{†a)} 西尾 和晃^{†b)} 木村 誠^{††c)} 藤吉 弘亘^{†d)}

Proposal of Inside-Out Camera for First Person Vision

Shoichi SHIMIZU^{†a)}, Kazuaki NISHIO^{†b)}, Makoto KIMURA^{††c)},
and Hironobu FUJIYOSHI^{†d)}

あらまし 人の行動意図を認識しようとする First Person Vision では、人の状態とその人が何を見ているのかという情報が必要となる。そこで我々は、人の眼球と人の視界映像を同時に取得する Inside-Out カメラを提案し、そのカメラの構成を活かした注視点の推定法も提案する。Inside-Out カメラはハーフミラーを介して眼球を正面から、視界映像を眼球と同等の位置から撮影することができる。Inside-Out カメラでは、眼球を撮影した画像から得られる視線ベクトルと視界を撮影した画像から得られる注視点位置の関係を変換式で表すことが可能である。そのため、変換式のパラメータをあらかじめ推定することにより、視線ベクトルから注視点を推定する。評価実験では、指標を注視した際の両眼、両視界映像を撮影し、視界画像の指標位置を真値として視線ベクトルから推定された注視点との誤差を算出した。評価実験から視野角において約 1.5 度の平均誤差で注視点を推定可能であることを確認した。人間は 1 点を注視しているとき、注視箇所だけではなく視野角で約 2 度の範囲がはっきり見えていることが報告されていることから、この範囲を評価基準とすると、提案手法は、十分な精度をもつことが分かる。

キーワード First Person Vision, Inside-Out カメラ, 視線検出, ステレオ視

1. ま え が き

First Person Vision では、人の頭部にカメラを装着し、人の視界と同じような画像を得ると同時に他方のカメラで眼球の観測を行うことにより、人の行動意図を観測することを目的とし、2009 年に開催された CVPR のワークショップである Egocentric Vision においてカーネギーメロン大学の金出教授が提唱した試みである [1]。

従来、人の行動理解を行うためのセンサとして、ウェアラブルカメラが用いられている。久保田らは、頭部に装着したカメラから個人行動記録システムのために注目シーンの検出を提案している [2]。この注目シーンは、画像中の移動量を観測することにより検出され

る。また、青木らは、人物の動作特徴を抽出し、動作特徴を自然言語の概念に対応付けて、装着者の動作を自然言語テキストで表現する手法を提案している [3]。カメラは、人の移動量を安定して取得するために胸元に装着し、人物の移動量から人の移動方向認識を行い、手領域を検出することにより手の動きを認識している。このように、ウェアラブルカメラを用いて人の行動を認識する手法が提案されているが、人の行動意図までを観測・認識することができない。First Person Vision では、人の行動意図を観測することを目的とするため、人がどのように動いているかに加え、人が何を注目しているのかという情報も重要となる。そのため、First Person Vision では、視線を観測するカメラを使用し、人の行動意図の観測を行う。

視線を検出するためのカメラシステムは古くから製品化され、現在では人の視界を撮影するカメラを備えた製品が存在する [4]～[8]。また、視線を検出する手法も多く提案されている [9]～[12]。製品化されたカメラは、眼球を撮影するカメラと視界を撮影するカメラを有するが、カメラの設置構造上の問題により、眼球は正面から撮影されておらず、視界も実際の視界と異なる。

[†] 中部大学工学部情報工学科, 春日井市

Dept. of Computer Science, Chubu University, 1200 Matsumoto-cho, Kasugai-shi, 487-8501 Japan

^{††} 産業技術総合研究所デジタルヒューマン研究センター, 東京都
AIST Digital Human Research Center, 2-41-6 Aomi, Kotouku, Tokyo, 135-0064 Japan

a) E-mail: shiyou@vision.cs.chubu.ac.jp

b) E-mail: xeno@vision.cs.chubu.ac.jp

c) E-mail: makoto.kimura@aist.go.jp

d) E-mail: hf@cs.chubu.ac.jp

視線計測では、眼球を正面から撮影することにより、視線検出を容易にすることができる。また、視界は実際の視界と同等の位置から撮ることにより、オクルージョンの発生による見えの違いを軽減することが可能となる。

そこで我々は、ハーフミラーを介すことにより眼球を正面から、視界を眼球と同等の位置から撮影する Inside-Out カメラを提案する。また、Inside-Out カメラの構造上の特長を活かした注視点の算出法についても提案する。Inside-Out カメラは、合計4台のカメラがあり、左右の眼球を撮影する2台の Eye カメラ、視界を撮影する2台の Scene カメラにより構成される。注視点の算出には、あらかじめ Eye カメラの画像から推定した視線ベクトルから Scene カメラの画像に存在する注目点へ変換する変換式のパラメータを推定する。このパラメータを用いることにより注視点を視線ベクトルから算出することが可能となる。

2. 従来の視線計測装置

本章では、従来の視線計測装置として、頭部装着型と非装着型について述べ、その長所と短所を示す。

2.1 頭部装着型

頭部装着型の視線計測装置 [4]~[6] は、両眼と視界の撮影を行うことが可能である。眼球を撮影するカメラは2台、視界を撮影するカメラは1台で構成される。スポーツ選手の視線解析やヘッドマウントディスプレイを装着することでコンピュータゲーム等に利用されている。

2.2 非装着型

非装着型の視線計測装置 [7], [8] は、非接触であるため、自然な状態で被験者の視線を計測することが可能であるが、視界の映像を撮影することはできない。被験者への負担が少なく、子供と大人に関係なく視線計測を容易に行うことができる。

2.3 両タイプの比較

両タイプの長所と短所を表1にまとめる。頭部装着型は、頭部に装着することにより被験者が自由に移動できるため、スポーツ選手の視線解析等に利用されている。しかし、人体に装着する装置は被験者に負担がかかるため、計測される視線が日常のものとは異なる可能性がある。非装着型は、視界映像を撮影できず、被験者の行動は装置から眼球が撮影できる範囲に限られるが、被験者に「計測されている」という意識を与えることなく視線を計測できるため、自然なデータを得ることができるメリットがある。

表1 両タイプの比較
Table 1 Comparison of both types.

	Head-mounted	Contactless type
Subject movement	Unrestricted	Restricted
Burden on subject	Medium	Low
Accuracy	High	Medium
Visual-field capture	Yes	No

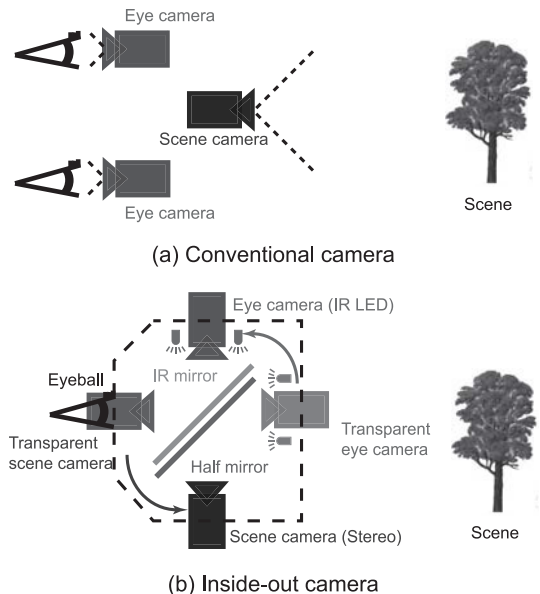


図1 従来の視線推定装置との比較
Fig. 1 Comparing conventional camera and inside-out camera.

2.4 従来の視線計測装置の問題点

従来の視線計測装置は、図1(a)に示すように、視界を撮影する Scene カメラが眼球とは異なる位置に設置されているため、実際の人の視界と撮影される映像に視差が存在する。複数カメラのステレオ視と同様、被写物体までの距離が短い近距離映像になると、人の視界とカメラ映像の視差は増大する。そのため、従来の視線計測装置は、注視対象をある程度離れた対象に設定する、あるいは注視対象を既知の平面上に限定するなど、注視点に関して何らかの拘束条件を設定することが一般的であった。

3. Inside-Out カメラ

本章では、提案する Inside-Out カメラについて述べる。

3.1 装置構成

これまでに試作した装置は、図2に示すように、

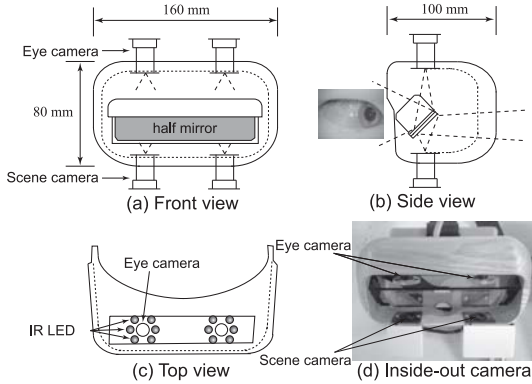


図 2 Inside-Out カメラの機構
Fig.2 Inside-out camera system.

ゴーグル上部に配置した眼球を撮影する Eye カメラ 2 台と、下部に配置した視界を撮影する Scene カメラ 2 台で構成されたゴーグル型の視線計測装置である。寸法は $W160 \times H80 \times D100\text{mm}$ であり、重量は約 200g である。材質は木を使用しており、カメラが発する熱を被験者に伝えにくいため、長時間の計測が可能である。Inside-Out カメラは図 1 (b) に示すように、透明なカメラが存在するような光学配置を実現している。その詳細は以降で述べる。

3.2 Eye カメラ

Eye カメラは、赤外ミラー、左右の眼球を撮影する 2 台の赤外カメラとカメラの周りに 6 個ずつ配置された赤外 LED で構成されている。LED はカメラの周囲に配置されており 750 nm～900 nm (中心波長 850 nm)、指向性 ± 40 度の近赤外光を発する。赤外カメラの画角は約 23 度、焦点距離は約 8 mm、重量は約 20 g である。赤外ミラーは、可視光の 95% を透過し、赤外光の 95% を反射するミラーである。赤外ミラーは赤外カメラの光軸と 45 度傾けて設置してあるため、上部に設置された赤外カメラは赤外ミラーを介して、人物の眼球を目の前から撮影した近赤外画像を $640 \times 480\text{pixel}$ の解像度で取得する。Eye カメラより取得した画像を図 3 に示す。赤外光は不可視であるため、視覚刺激を与えることなく眼球を撮影することが可能である。Eye カメラは、株式会社ニューオプトから購入することが可能である [13]。

3.3 Scene カメラ

Scene カメラは左右の視界を撮影する小型 CCD カメラ 2 台と 50% ハーフミラーで構成されている。CCD カメラの画角は約 80 度、焦点距離は約 4 mm、重量は

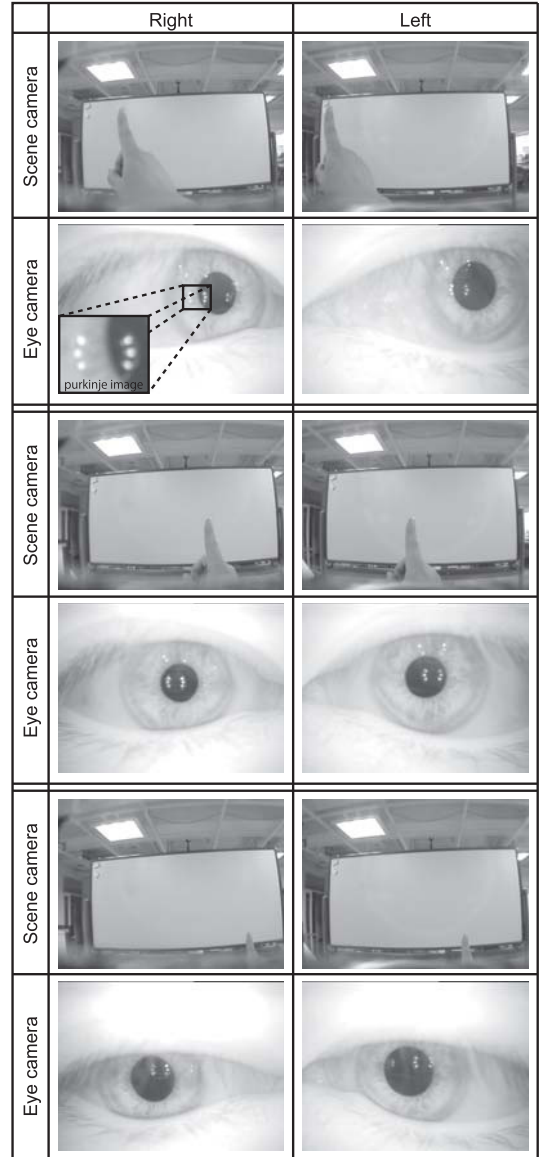


図 3 Inside-Out カメラにより得られる画像
Fig.3 Images captured with inside-out camera.

約 12 g である。50% ハーフミラーは可視光の 50% を反射し、残りを通過させる性質をもつミラーである。ハーフミラーも赤外ミラーと同様に CCD カメラの光軸と 45 度傾けて設置してあるため、ハーフミラーを介すことで、透明なカメラで光学的に人物視点とほぼ同一の位置から $640 \times 480\text{pixel}$ の解像度で映像を撮影することができる。また、ステレオカメラであるため、Tsai のモデル [14] や Zhang のモデル [15] を用い

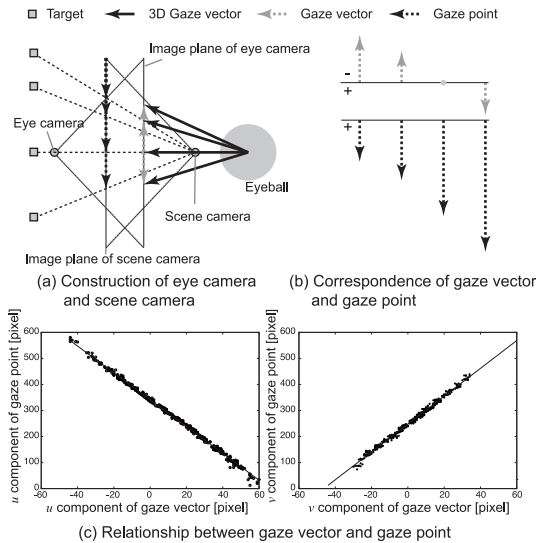


図 4 視線ベクトルと注視点の関係
Fig. 4 Relationship between gaze vector and gaze point.

てキャリブレーションを容易に行うことができ、視界内における注視点の三次元位置を推定することが可能である。カメラ間のベースラインは約 6.5 cm である。Scene カメラから得られる画像を図 3 に示す。

3.4 Eye カメラと Scene カメラの関係

Eye カメラと Scene カメラは、ハーフミラーを挟んで向かい合うように設置してある。したがって、図 4(a) に示すように各カメラの画像平面は平行に近い状態となっている。Eye カメラにより観測される視線ベクトルと Scene カメラで観測される物体の画像座標の関係には、物体に追従して視線が移動し、各カメラの画像平面が平行である場合、相関関係があることが分かる。図 4(b) の例では、視線ベクトルがマイナスへ変化するに従い、物体の座標位置が大きくなることが分かる。実際に指標となる物体を注視した際の視線ベクトルと指標の画像座標を各座標成分により表示した例を図 4(c) に示す。各カメラ間関係は図 4(c) のように線形になるため、線形変換により容易に扱うことが可能である。

4. Inside-Out カメラを用いた三次元注視点推定

一般に、注視点を推定するためには視線ベクトルが必要であり、眼球回転中心や角膜曲率中心、瞳孔中心より推定する手法が提案されている [9]~[12]。この視

線ベクトルは、注視する指標の位置に依存して動くため、視線ベクトルと指標の位置間に相関関係があることは明白である。提案する Inside-Out カメラでは、3.4 で述べたように、この相関関係を容易に扱うことが可能である。実際に、図 3 からは、注視点が左に移動するとプルキニエ像^(注1)と瞳孔中心から推定される視線ベクトルも左方向を向き、右に移動すると、右方向を向いている。図 4(c) からは、線形な関係があることが分かる。したがって、注視点と視線ベクトルには線形な関係があることから、視線ベクトルを推定することにより、Scene 画像における注視点を推定することが可能になる。以下に注視点推定の流れを示す。

- (1) 視線ベクトルの推定：瞳孔中心とプルキニエ像から視線ベクトルを推定
- (2) 注視点の算出：変換式により視線ベクトルを注視点に変換
- (3) 三次元注視点の算出：両眼の注視点を用いてステレオマッチングにより三次元注視点を算出

4.1 視線ベクトルの推定

大野らにより提案された視線検出法 [16] は、プルキニエ像から角膜曲率中心を算出し、その角膜曲率中心と瞳孔中心を結ぶ直線を視線ベクトルとする方法である。眼球回転中心と瞳孔中心を用いた手法より精度よく視線ベクトルが推定可能であると報告されている。そこで本手法においても、角膜曲率中心と瞳孔中心から視線ベクトルを推定する。

まず、視線ベクトルを推定するために角膜曲率中心をプルキニエ像を用いて算出する。大野ら [16] の光源は一つであるため、プルキニエ像をカメラの光軸中心になるように補正を行うことで、角膜曲率中心を算出している。

一方、我々の提案する Inside-Out カメラは、カメラの周りに 6 個の光源があるため、その中心がカメラの光軸と仮定することができる。そのため、プルキニエ像群の中心を角膜曲率中心とすることが可能になる。プルキニエ像は、文献 [12], [16], [17] により提案されている手法により抽出可能である。このようにして求めたプルキニエ像群から画像上の角膜曲率中心 $C = [C_u, C_v]^T$ を推定することができる。

次に、瞳孔中心を推定する。瞳孔中心 $P = [P_u, P_v]^T$ の算出には、坂下らにより提案された手法 [18] を用いる。

(注1)：角膜表面における赤外光の反射光。

以上より推定した角膜曲率中心 C と瞳孔中心 P から視線ベクトルを算出する。視線ベクトル $V = [V_u, V_v]^T$ は、角膜曲率中心を基準とするため、以下の式により算出できる。

$$V = P - C \quad (1)$$

4.2 変換式を用いた注視点の算出

変換式を用いた注視点の算出では、4.1 により推定した視線ベクトルを用いて、Scene 画像上の注視点を推定する。

本手法では、変換式の変換パラメータを求めるオフライン処理と変換式を用いた注視点を推定するオンライン処理に分かれる。その処理の流れを図 5 に示す。

4.2.1 オフライン処理

オフライン処理では、視線ベクトルと注視点の関係を変換式で表現するためのパラメータを推定する。注視点 $L = [L_u, L_v]^T$ と視線ベクトル $V = [V_u, V_v]^T$ の各 u, v 成分の分布は、図 4(c) のようになる。図 4(c) から注視点と視線ベクトルの各成分は、比例していることが分かり、線形変換可能であることが分かる。線形変換の式は以下になる。

$$L_u = a_u V_u + b_u \quad (2)$$

$$L_v = a_v V_v + b_v \quad (3)$$

ここで、 $a = [a_u, a_v]^T$ は各成分における直線の方程式の傾きであり、 $b = [b_u, b_v]^T$ は切片となる。したがって、あらかじめ 2 点以上のキャリブレーションデータから、図 5(b) のように直線の方程式の傾きと切片を算出しておくことにより、視線ベクトル V を入力とした際の注視点 L を算出することが可能になる。

4.2.2 オンライン処理

オンライン処理では、オフライン処理で算出した変換パラメータ a, b を用いて、視線ベクトルから注視点を推定する。まず、視線ベクトルは、4.1 により算出する。算出した視線ベクトル V を各成分に分割し、式 (2), (3) により注視点を算出する。図 5(d) に視線ベクトルから注視点を算出する概念図を示す。

4.3 注視点の最適補正

4.2.2 により算出した両眼の注視点は、推定した視線ベクトルに誤差を含むため、Scene 画像上の注視点にもずれが生じる。そこで、文献 [19] にて提案されている最適補正を用いて、図 6 に示すように、Scene 画

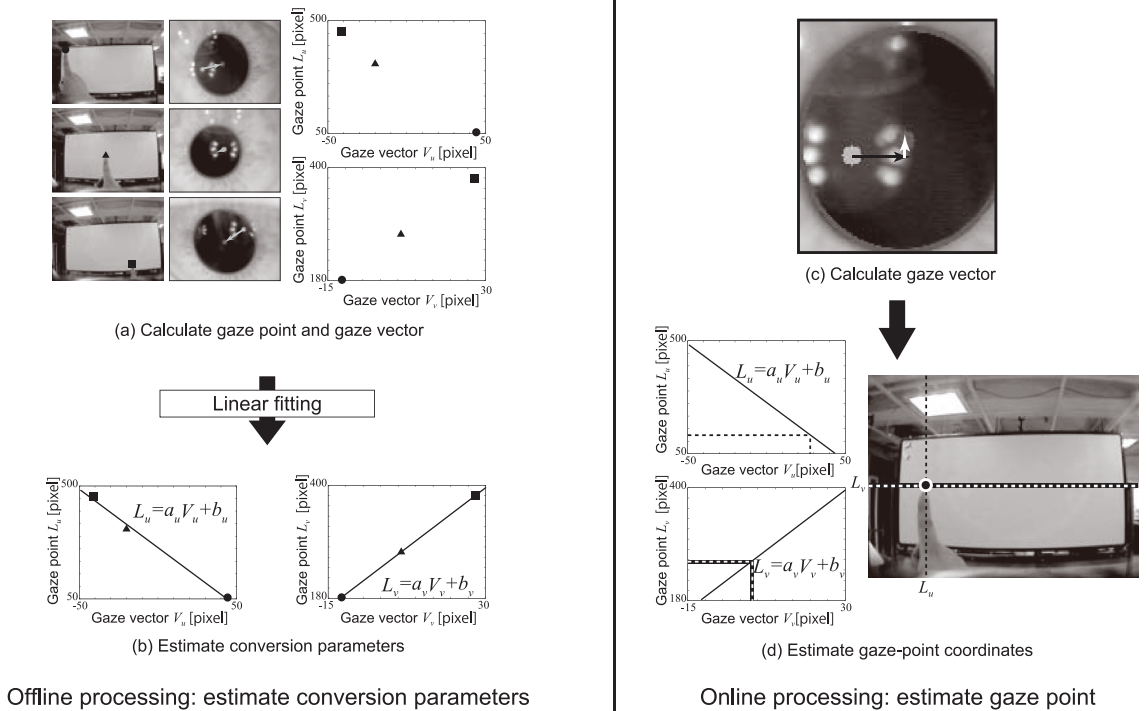


図 5 オフライン処理とオンライン処理
Fig. 5 Offline processing and online processing.

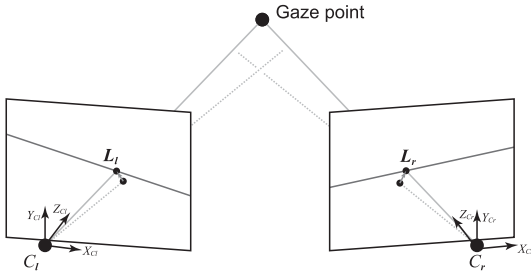


図6 注視点の最適補正
Fig. 6 Optimal correction of gaze points.

像上の注視点同士が対応するように位置の補正を行う。

左側 Scene 画像中の点 L_l と右側 Scene 画像中の点 L_r が対応しているとき、それぞれの点を通る視線同士が交差する条件は、各点を同次座標 $\mathbf{m}_l = [L_{u_l}/f_0, L_{v_l}/f_0, 1]^T$, $\mathbf{m}_r = [L_{u_r}/f_0, L_{v_r}/f_0, 1]^T$ (注2) で表現すると、式 (4) を満たす。

$$\mathbf{m}_r^T \mathbf{F} \mathbf{m}_l = 0 \quad (4)$$

ここで、 $\mathbf{F}$ は基礎行列を表し、左右の画像の幾何学的関係を示す。

最適補正では、視線同士が交差するように、Scene 画像上の注視点を移動させ、その移動量が最小になるように定める。左側 Scene カメラの最適補正後の Scene 画像座標 $\hat{\mathbf{m}}_l$ と、右側 Scene 画像座標 $\hat{\mathbf{m}}_r$ はそれぞれ式 (5), (6) を用いて算出する。

$$\hat{\mathbf{m}}_l = \mathbf{m}_l - \frac{(\mathbf{u}\xi)\mathbf{P}_k\mathbf{F}\mathbf{m}_r}{(\mathbf{u}\mathbf{V}_0[\xi]\mathbf{u})} \quad (5)$$

$$\hat{\mathbf{m}}_r = \mathbf{m}_r - \frac{(\mathbf{u}\xi)\mathbf{P}_k\mathbf{F}^T\mathbf{m}_l}{(\mathbf{u}\mathbf{V}_0[\xi]\mathbf{u})} \quad (6)$$

ここで、 $\mathbf{P}_k$ は第 3 成分を 0 にする対角行列である。

$$\mathbf{P}_k = \text{diag}(1, 1, 0) \quad (7)$$

$\mathbf{u}$ と ξ は、それぞれ基礎行列 $\mathbf{F}$ とデータ $\mathbf{m}$ を九次元ベクトルに書き直したものである。

$$\mathbf{u} = [F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{32}, F_{33}]^T \quad (8)$$

$$\xi = [L_{u_l}L_{u_r}, L_{u_l}L_{v_r}, f_0L_{u_l}, L_{v_l}L_{u_r}, L_{v_l}L_{v_r}, f_0L_{v_l}, f_0L_{u_r}, f_0L_{v_r}, f_0^2]^T \quad (9)$$

ここで、行列 $\mathbf{V}_0[\xi]$ は 9×9 の行列であり、文献 [19]

にて定義されている式である。

算出した $\hat{\mathbf{m}}$ を $\mathbf{m}$ に代入し、式 (5), (6) を再度適用し、収束するまで繰り返す。

4.4 三次元注視点の算出

両眼の注視点を用いて、三次元注視点を算出する。注視点は、2 台の Scene 画像上の点であるため、ステレオ視の問題として三次元の注視点を算出することができる。本論文では、ステレオマッチングにより、三次元注視点を算出する。ステレオマッチングの際に使用する類似度を算出するウィンドウのサイズは、人間は 1 点を注視しているとき、注視箇所だけではなく視野角で 2 度程度の範囲がはっきり見えていることが報告されている [20]。そこで、Scene 画像上に視野角 2 度を投影した 20 pixel をウィンドウサイズとして用いる。

5. 評価実験

本章では、提案手法を評価するために、実際に指標を注視したデータを採取し、指標位置の真値と推定した注視点との誤差を算出する。また、三次元注視点を推定した例も示す。

5.1 実験データの採取

評価実験を行うためのデータは、以下の条件のもとに採取した。

- 注視対象：図 7 に示すように、九つの指標から構成するボードを使用
- ボード上の注視順番：九つの指標を左上から順番に注視
- 指標の大きさ：被験者が裸眼で指標を視認できる大きさに適応的に変更し、指標の中心を注視
- ボードの奥行範囲：1 m から 5 m
- ボードの奥行間隔：50 cm
- 一つの奥行に対する撮影回数：一つの奥行に対してボードを 6 箇所配置するため、54 回撮影
- 1 被験者に対するデータ数 (組)：54 (組) $\times$ 9 (奥行の幅) = 486 組

ここで 1 組は、1 回の撮影時に得られる 2 枚の Scene 画像と 2 枚の眼球画像の計 4 枚の画像を表す。

- 被験者：5 人
- 被験者の状態：頭部の固定なし。ただし、可能な限り同一姿勢を保つ

5.2 誤差推定結果

提案手法では、あらかじめ変換式のパラメータを算

(注2)： f_0 は数値計算を安定化させるための任意のスケール定数。

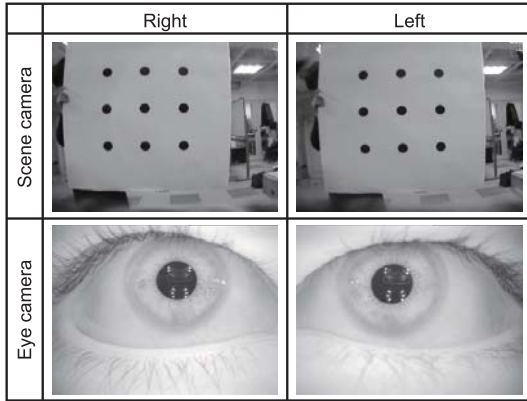


図 7 実験データの一部

Fig. 7 A part of the experimental data.

出する必要があり、そのパラメータの精度は、キャリブレーションデータに依存することは明らかである。そこで、キャリブレーションに使用するデータ数を変化させた際の誤差についての検討を行う。評価実験は以下の手順で行う。

(1) キャリブレーションに使用するデータの組数を 2 組と決定

(2) 奥行が 1m のときのデータ (54 組のデータ) からランダムにデータの組数分選びキャリブレーションを実行

(3) キャリブレーションに使用したデータ以外のすべてのデータから注視点を推定し、推定した注視点と指標位置の真値のユークリッド距離を誤差として計算

(4) (2) から (3) の作業を 100 回施行

(5) 使用するデータの組数を増やし、組数が 9 になるまで (2) に戻り、繰り返す
ここで、真値として与える指標位置は、円検出によりサブピクセルレベルで得られている。以上の手順を各被験者に対して行い、誤差の評価を行う。

人間は 1 点を注視しているとき、注視箇所だけではなく視野角で 2 度程度の範囲がはっきり見えていることが報告されていることから [20]、真値から 2 度以内に領域を Scene 画像に投影した際の直径約 20 pixel を評価基準として評価を行う。

図 8 に横軸をキャリブレーションデータ組数、縦軸を全被験者における誤差の平均と 100 回の試行の最小平均誤差を示す。図 8 からデータ数を 4 組以上使用することにより、平均誤差を 15 pixel 以下とすることが

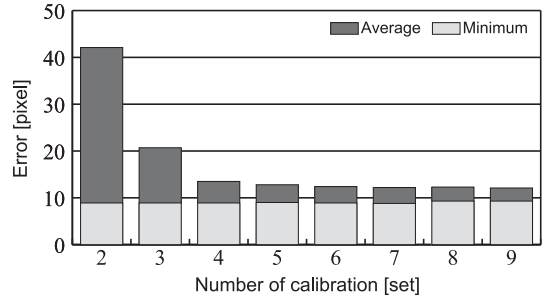


図 8 キャリブレーションに使用するデータ数に対する誤差 [pixel]

Fig. 8 Error to number of data used for calibration [pixel].

表 2 最大・最小誤差における二つの視線ベクトルの距離 [pixel]

Table 2 Distance of the two gaze vectors at minimum and maximum error [pixel].

	Left eye		Right eye	
	u	v	u	v
Minimum	51.92	27.66	49.78	27.06
Maximum	10.02	26.99	9.29	26.18

でき、誤差範囲である 20 pixel 以内に抑えることができることを確認した。平均誤差 15 pixel は、視野角では 1.5 度以内に真値があることを示す。また、平均誤差は減少傾向にあるため、組数を増やすことにより安定した注視点推定を可能にした。

次に、図 8 の最小平均誤差が約 10 pixel であることから、少ないキャリブレーションデータにおいても平均誤差が低く、精度の高い推定結果が得られることが分かる。これらのキャリブレーションデータを解析することにより、キャリブレーションに有効なデータ採取方法の手助けになると考えられる。

キャリブレーションデータに有効なデータとして、視線ベクトル方向や指標位置に偏りが少ないことが考えられる。そこで、データの分布について調査を行う。データの分布調査手順について以下に示す。

(a) キャリブレーションに使用するデータ数を 2 組と決定

(b) 上記実験における (2) から (4) までの処理を実行

(c) 最大・最小誤差における二つの視線ベクトル間の距離を算出

(d) (b) と (c) を 100 回繰り返すことにより、100 個の最大・最小誤差時の視線ベクトル間の距離を算出し、最大・最小誤差の平均値を算出

上記手順に従い最大・最小誤差における二つの視線ベクトルの距離を算出した結果を表 2 に示す。表 2 の u , v は、視線ベクトルの u 成分と v 成分を表す。

表 2 から v 成分に対する距離は最小と最大誤差ともに 26 pixel 以上の距離があるのに対して、 u 成分に対する距離は最小誤差の際に約 50 pixel, 最大誤差の際に約 10 pixel と約 40 pixel の差があることが分かる。

このことから、誤差を小さくするためには、視線ベクトル間の距離が大きくなるようなデータを採取することが重要であるといえる。

5.3 三次元注視点の推定

算出した注視点を用いて、三次元注視点を推定した例を図 9 に示す。図 9 は、各図の左側に実際に取得したシーンのある 1 組の画像を示し、右側に三次

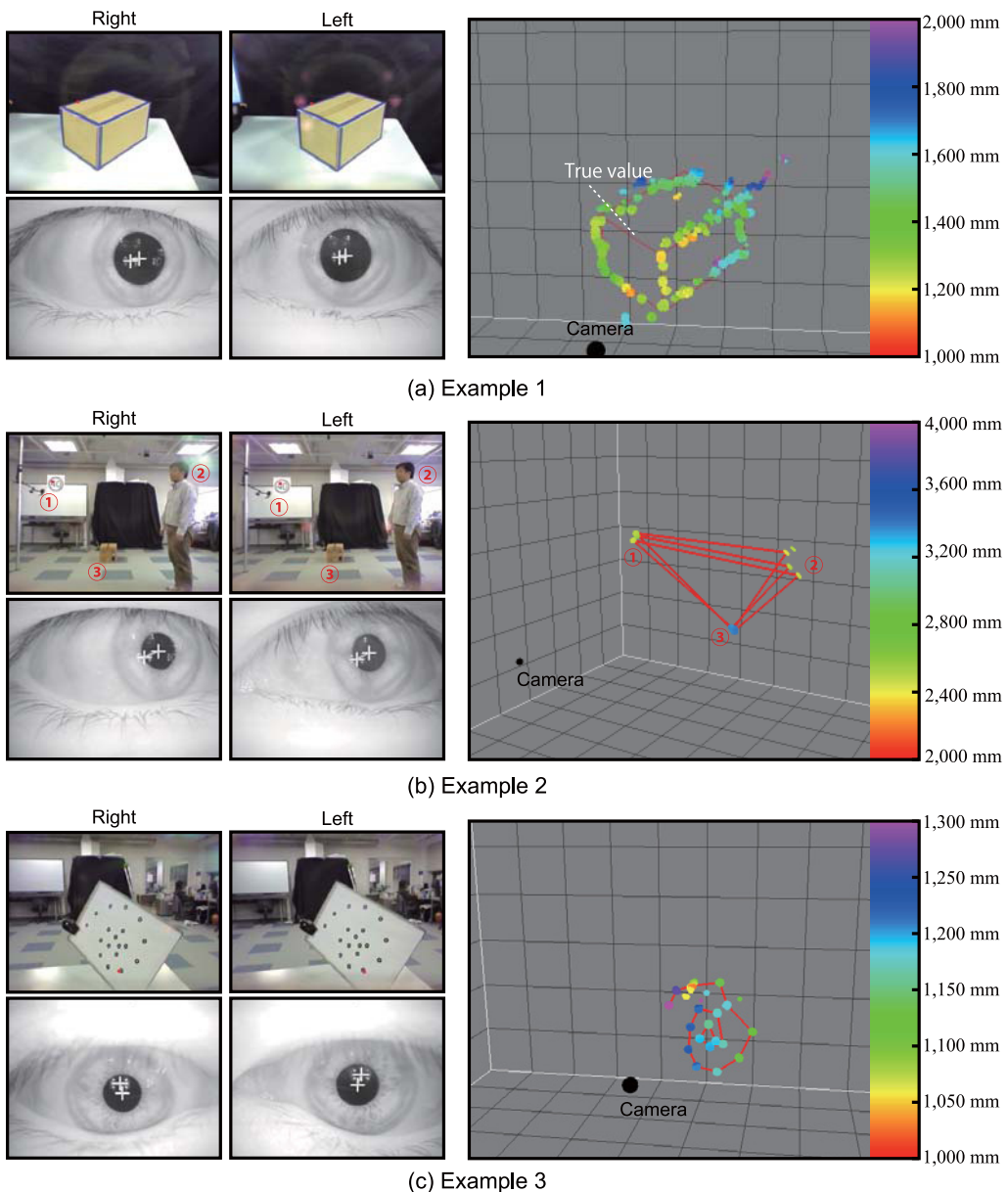


図 9 注視点の三次元位置推定
Fig. 9 Estimate 3D gaze point.

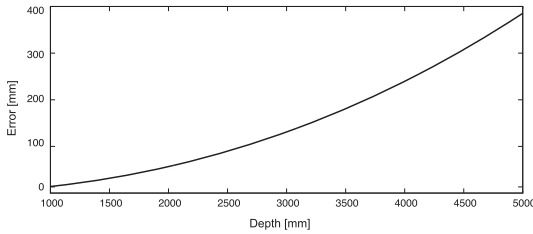


図 10 ステレオマッチングの精度
Fig. 10 Accuracy of stereo matching.

元注視点を算出した例を示す．三次元注視点の色は，Inside-Out カメラからの距離を表しており，右端のカラーバーの幅で色分けされている．

図 9(a) は，直方体の箱の縁をなぞるように視線を移動させた三次元注視点である．箱は約 1,000 mm 先に設置し，箱の幅，高さ，奥行はそれぞれ，約 350 mm × 約 350 mm × 約 450 mm である．箱の縁に沿って三次元注視点が得られていることが分かる．

図 9(b) は，①標識，②人間，③箱の順番に注視対象を変化させたときの三次元注視点である．各対象の奥行は，標識は約 3,500 mm，人間は約 2,500 mm，箱は約 4,000 mm である．三次元注視点を得られるため，注視対象の大まかな位置を把握することが可能となり，三次元 ScanPath を取得することができる．

図 9(c) は，約 1,000 mm 先に A0 の大きさの板上にらせん状に指標を配置し，それを注視した例である．これも図 9(b) と同様に三次元 ScanPath が得られていることが分かる．

三次元注視点は，ステレオマッチングにより求められるため，マッチング領域がユニークである場合に対応点の推定精度が高くなる傾向にある．また，三次元位置の推定精度は 2 台のカメラ間のベースラインに依存するため，Inside-Out カメラの精度を一意に知ることが可能である．対応点が 1 pixel ずれた際の Inside-Out カメラにおける誤差を図 10 に示す．一般的なステレオ視と同様に，対象までの距離に比例して誤差が増加することが分かる．

以上より，Inside-Out カメラでは三次元の注視点を推定可能であり，三次元空間中のどこを見ているかという情報を知ることが可能になる．そのため，人の行動意図を理解する際に，人がどこを見ているのかという情報を取得することが可能になる．

5.4 処理時間

提案手法の各処理において，CPU Xeon 2.27 GHz,

表 3 処理時間 [ms]
Table 3 Processing time [ms].

2D Gaze point	Gaze vector	34.7
	Linear conversion	4.6×10^{-4}
	Optimal correlation	2.7×10^{-3}
3D gaze point		50.2
Total		84.9

メモリ 12 GB を使用して測定した処理時間を表 3 に示す．処理の多くは，視線ベクトルの推定と三次元注視点の推定に費やされていることが分かる．一方，視線ベクトルから Scene 画像上の注視点へ変換する処理は，処理が線形変換のため早く，最適補正は，反復回数が 1, 2 回で収束するため処理コストが低くなっている．処理時間を合計すると 84.9 ms であることから，1 秒間に約 11 FPS (Frame Per Second) で三次元注視点を計測することが可能である．

6. むすび

本論文では，First Person Vision のための Inside-Out カメラとそのカメラの構造を生かした注視点の推定法を提案した．Inside-Out カメラはハーフミラーを介して眼球を正面から，視界映像を眼球と同等の位置から撮影することを可能とした．注視点をあらかじめ算出した変換式のパラメータを用いて，視線ベクトルから算出手法を提案した．評価実験から視野角における約 1.5 度の平均誤差で注視点を推定可能であることを確認した．人間は 1 点を注視しているとき，注視箇所だけではなく視野角で約 2 度の範囲がはっきり見えていることが報告されていることから，この範囲を評価基準とすると，提案手法は，十分な精度をもつことが分かる．

文 献

- [1] T. Kanade, “First-person, inside-out vision,” IEEE Workshop on Egocentric Vision, CVPR, 2009.
- [2] 久保田敏司, 中村裕一, 大田友一, “個人行動記録システムにおける注目シーンの検出: 注目シーン検出の高精度化と環境カメラの利用,” 信学技報, PRMU2002-64, 2002.
- [3] 青木茂樹, 高橋昌史, 大西正輝, 小島篤博, 福永邦雄, “ウェアラブルカメラによる動作の認識とテキスト表現,” 信学技報, PRMU2003-49, 2003.
- [4] ナックイメージテクノロジー, “アイマークレコーダー EMR-9.”
- [5] Arrington Research, “アイトラッカー ViewPoint.”
- [6] ISCAN, “AA-ETL-500B.”
- [7] ナックイメージテクノロジー, “アイマークレコーダー EMR-AT VOXER.”
- [8] Tobii, “Tobii XL.”

- [9] R. Stiefelhagen, J. Yang, and A. Waibel, "Tracking eyes and monitoring eye gaze," Proc. Workshop on Perceptual User Interfaces, pp.98-100, 1997.
 - [10] 竹上 健, 後藤敏行, "角膜反射像と虹彩輪郭情報を併用した視線検出法," 信学論 (D-I), vol.J82-D-I, no.10, pp.1295-1303, Oct. 1999.
 - [11] 松田圭司, 永見武司, 山根 茂, "視線位置計測システムの開発," 信学技報, TL2000-2, 2000.
 - [12] 田中宏明, 疋田真一, 笠井 健, 竹田 仰, "二光源を用いた角膜曲率中心位置計測による視線検出法," 信学技報, MBE2009-128, 2009.
 - [13] ニューオプト, "眼振分析装置 ETS-500A (ET-60-L)."
 - [14] R.Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses," IEEE J. Robot. Autom., vol.3, no.4, pp.323-344, 1922.
 - [15] Z. Zhang, "A flexible new technique for camera calibration," IEEE Trans. Pattern Anal. Mach. Intell., vol.22, no.11, pp.1330-1334, 2000.
 - [16] 大野健彦, 武川直樹, 吉川 厚, "2点補正による簡易キャリブレーションを実現した視線測定システム," 情処学論, vol.44, no.4, pp.1136-1149, 2003.
 - [17] 大野健彦, 武川直樹, 吉川 厚, "眼球形状モデルに基づく視線測定法," 第8回画像センシングシンポジウム, pp.307-312, 2002.
 - [18] 坂下祐輔, 藤吉弘亘, 平田 豊, 高丸尚教, 深谷直樹, "高速楕円検出に基づく眼球回旋運動のリアルタイム計測システム," 電学論 (C), vol.127-C, pp.591-598, 2007.
 - [19] 金谷健一, 菅谷保之, 新妻弘崇, "2画像からの三角測量: Hartley-sturm vs. 最適補正," 情処学研報, CVIM, vol.2008, no.27, pp.335-342, 2008.
 - [20] 大野健彦, "視線を用いた高速なメニュー選択作業," 情処学論, vol.40, no.2, pp.602-612, 1999.
- (平成 22 年 11 月 22 日受付, 23 年 5 月 31 日再受付)



木村 誠

2002 慶應義塾大学大学院理工学研究科博士課程了。博士 (工学)。同年科学技術振興事業団 CREST 研究員。2006 産業技術総合研究所研究員。2010 年 4 月より (株)サムスン横浜研究所。コンピュータビジョン、画像処理に関する研究開発に従事。



藤吉 弘亘 (正員)

1997 中部大学大学院博士後期課程了。1997~2000 米カーネギーメロン大学ロボット工学研究所 Postdoctoral Fellow。2000 中部大学講師, 2004 同大准教授を経て, 2010 より同大教授。2006 米カーネギーメロン大学ロボット工学研究所客員研究員。工博。計算機視覚, 動画像処理, パターン認識・理解の研究に従事。2005 年度ロボカップ研究賞。2009 年度情報処理学会論文誌コンピュータビジョンとイメージメディア優秀論文賞, 2009 年度山下記念研究賞。情報処理学会, 電気学会, IEEE 各会員。



清水 彰一

2009 中部大学大学院後期課程了。博士 (工学)。2009 中部大学研究員。2010 年 10 月より三菱電機株式会社先端技術総合研究所。動画像処理・ITS の研究に従事。情報処理学会, 精密工学会各会員。



西尾 和晃

2008 中部大・工・情報工学卒。2010 同大学院修士課程了。2010 丸紅 OKI ネットソリューションズ (株) (旧沖電気ネットワークインテグレーション) 入社。映像監視システム, ビデオ会議システムの設計構築に従事。