

高解像度撮影における実時間視線推定の高精度化

佐竹 純二[†] 小林 亮博[†] 平山 高嗣^{††} 川嶋 宏彰^{††} 松山 隆司^{††}

[†] 情報通信研究機構 〒 619-0289 京都府相楽郡精華町光台 3-5

^{††} 京都大学 〒 606-8501 京都市左京区吉田本町

E-mail: [†]{j-satake,akihiro-k}@nict.go.jp, ^{††}{hirayama,kawashima,tm}@i.kyoto-u.ac.jp

あらまし 本研究では、大型ディスプレイを用いた情報提示システムにおいて、ユーザの顔や視線の方向を推定し、ユーザの興味や反応を認識することで、コンテンツをインタラクティブに制御することを目指している。カメラ画像のみを用いた視線推定法では、視線を虹彩中心と眼球中心を結ぶ直線として求める方法が一般的であるが、目領域の解像度が低いために誤差の影響が大きかった。そこで、高解像度カメラ (UXGA 30fps) を導入し、角膜による屈折を考慮した詳細な虹彩形状モデルを用いて、より高精度に視線を推定する方法を報告する。

キーワード 視線推定, 高解像度カメラ, 虹彩検出

Accuracy Improvement of Real-Time Gaze Estimation using High Resolution Camera

Junji SATAKE[†], Akihiro KOBAYASHI[†],

Takatsugu HIRAYAMA^{††}, Hiroaki KAWASHIMA^{††}, and Takashi MATSUYAMA^{††}

[†] National Institute of Information and Communication Technology

3-5 Hikaridai, Seika-cgo, Soraku-gun, Kyoto, 619-0289 Japan

^{††} Kyoto University Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-8501 Japan

E-mail: [†]{j-satake,akihiro-k}@nict.go.jp, ^{††}{hirayama,kawashima,tm}@i.kyoto-u.ac.jp

Abstract We develop an interactive information display that estimates face and gaze directions of a user in real-time, and interactively controls contents by recognizing the user's interest and reaction. The gaze estimation method that uses only the camera image generally estimates the gaze direction as a line that connects the iris center and the eyeball center. However, because the resolution of eye area was low, the influence of error was large. We report an accuracy improvement of real-time gaze estimation using high resolution camera and detailed iris model.

Key words gaze estimation, high resolution camera, iris detection

1. ま え が き

本研究では、ユーザの非言語情報（視線や顔きなど）からユーザの興味や反応を認識し、インタラクティブなやり取りを行う対話システムの開発を目指している [1], [2]。例えば、図 1 のような大型ディスプレイを用いた情報提示システムにおいて、ユーザの注意（顔向き、視線）や動作（顔き、首傾げ、指差しなど）からユーザの興味や反応を自動的に認識することで、ユーザの好みに合わせた旅行プランニングシステムやユーザの興味を先取りして説明するガイドシステムなどに応用できる。

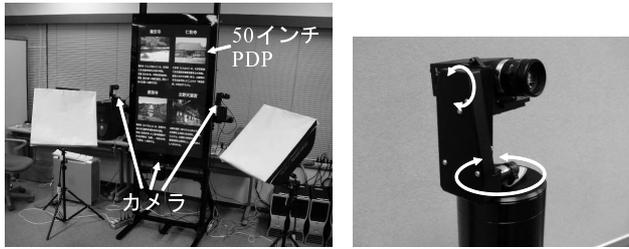
このような状況で、ユーザを拘束することなく視線を推定するには、顔画像のみを用い、視線を虹彩中心と眼球中心を結ぶ直線として求める方法が有効であると考えられる。しかし、顔

全体を捉えた画像では目領域の解像度が低いために誤差の影響が大きくなり、視線の推定精度が低いという問題がある。これに対し、赤外光源を利用する角膜反射法では、高解像度な目領域画像を用い、観測された瞳孔位置に対して角膜における屈折の補正も行っている。本稿では高解像度カメラを導入することで目領域の解像度の問題を解決し、顔画像のみから視線を実時間推定するシステムを構築する。また、角膜における屈折を考慮した詳細な虹彩モデルを用いることで、高精度に視線を推定する方法を提案する。

2. 高解像度撮影における視線推定の高精度化

2.1 従来の視線推定手法

従来の視線推定法の一つに、赤外光源を用いる角膜反射



(a) システム外観 (b) カメラ向き制御機構

図 1 情報提示システム

Fig. 1 Information display.

法 [3], [4] があり, 据え置き型の計測装置も市販されている. この方法は近赤外の点光源で目を照射し, 角膜で反射されたブルキニ工像をもとに角膜の曲率中心を求め, それと瞳孔中心を結ぶ直線から視線方向を推定する. 誤差 1 度程度の高精度な計測が可能ではあるが, 赤外照明装置が必要であることや, 光源距離やカメラ画角などから頭部位置が制限されるという問題がある.

カメラ画像のみから視線を推定する方法としては, 顔全体画像から頭部姿勢を求め, 視線を虹彩中心と眼球中心を結ぶ直線として求める方法 [5] ~ [7] があり, 本研究でもこの方法を用いている. 装置構成の手軽さが利点であるが, 目領域の解像度が低い (虹彩直径 10 ~ 30pixel 程度) ために誤差の影響が大きく, 推定精度は 5 度程度である. なお, これらの手法では, 虹彩中心を楕円の中心または黒色領域の重心として求めている.

頭部姿勢を用いず, 高解像度な目領域画像を用いて虹彩の楕円形状から視線方向を推定する方法も提案されているが, 推定精度 1 度を実現するためには虹彩直径 70pixel 程度を必要とする [8]. また, ノイズにロバストな虹彩検出法として LmedS を用いた手法 [9] が提案されており, 候補点の 50% が正しければ楕円推定が可能である. ハフ変換を用いた虹彩検出法も存在するが, 変形を考慮した楕円検出は実時間処理に至っていない.

2.2 試作システムの概要と高解像度撮影における視線推定の問題点

本研究では図 1(a) のような情報提示システムを街角に設置し, 訪れたユーザの興味や反応を認識してインタラクティブな情報提示を行うことを目指している. 従来の視線推定にはユーザに計測機器を装着させたり, 顔の動きを制限するものもあるが, 本システムではカメラ画像のみを用いてユーザの視線を推定する. ディスプレイの周りに設置した複数のカメラに制御機構 (図 1(b)) を設けることで立ち位置の変化や身長個人差に対応し, それぞれのカメラで独立に単眼推定処理を行うことで顔向きの変化によるセルフオクルージョンや眼鏡の反射などに対応する. カメラは Point Grey Research 社製の Grasshopper (UXGA (1600 × 1200), 8bit グレー画像, 30fps, 1/1.8 インチ CCD, IEEE-1394b) を用い, レンズは FUJINON 社製の HF25HA-1B ($f=25\text{mm}$, 左右) と HF16HA-1B ($f=16\text{mm}$, 下) を用いた. 図 2 はその撮影画像の例であり, 左カメラにおける虹彩直径は約 60pixel である. なお, 虹彩を検出するには照度が不十分だったため, 現在は写真電気工業社製の照明



(a) 左カメラ ($f=25\text{mm}$) (b) 下カメラ ($f=16\text{mm}$)

図 2 撮影画像例 (UXGA)

Fig. 2 Examples of captured image.

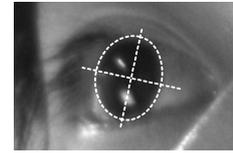


図 3 目領域の拡大画像

Fig. 3 Expansion image of eye part.

RIFA-F (50 × 50cm) を両脇に配置している.

撮影した顔画像から目領域を拡大したものを図 3 に示す. 図中の点線は虹彩の輪郭と, 視線方向から求めた軸である. これまでは目領域の解像度が低いためにあまり検討されていなかったが, 虹彩の輪郭が楕円ではなく, 角膜における屈折によって奥側がより膨らんで見えていることが確認できた. 虹彩の直径 (長軸) が 60pixel の時, この変形による虹彩中心のずれの影響は 2pixel 程度であるが, 顔とディスプレイの距離に比べて眼球は小さいため, 視線推定への影響は非常に大きい. 角膜反射法 [3] でも観測した瞳孔中心に対して屈折の補正を行っており, 高精度な視線推定を行うためにはこの屈折の問題を考慮する必要があると考えられる. 特に, 本システムではカメラを大型ディスプレイの周囲に設置しており, カメラと視線の角度が大きくなるため, この屈折に対する検討が必要となる. また, 高解像度画像を用いる際には処理速度も大きな問題となるため, 本稿では高速な推定方法を検討する.

2.3 屈折を考慮した虹彩の見え方のモデル化

文献 [3], [10], [11]などを参考に, 本研究では眼球構造を図 4 のようにモデル化した. 図中の数値はそれぞれの大きさ (単位は mm) を表している. また, カメラから見て手前側の屈折の影響は奥側に比べて小さいと考え, 手前側の虹彩輪郭の屈折は無視している. その他, 実際の眼球構造には以下のような特徴がある.

- 角膜は透明で, 複雑な非球面構造をしている. 中央部よりも周辺部, 鼻側よりも耳側の方が曲率半径が大きく, 光学中心は鼻側に 0.2 ~ 0.6mm, 上側に 0.2mm 程ずれている.
- 強膜は乳白色で, 角膜と繋がっている.
- 虹彩は直径約 11mm で色素を持ち, 光を遮る絞りの役割をしている.
- 瞳孔は虹彩にある穴で, 瞳孔中心は角膜の幾何学的中心からわずかに鼻側にずれており, 縮瞳時にさらに鼻側にずれる.
- 水晶体は透明で, レンズの役割をする.
- 角膜と虹彩, 虹彩と水晶体の間には房水という透明な液体が入っている.

• 網膜上の中心窩はわずかに耳側にずれており、光軸と視線にはずれがある。

- 眼球の大きさは成人で直径約 23 ~ 24mm である。
- 眼球の旋回中心は角膜の後方約 13mm であるが、1 点を中心として回転する構造ではないため、決定は難しい。
- 眼球の形状や大きさには個人差があり、角膜や水晶体の歪み（乱視）や左右の視線が揃わないこと（斜視）もある。

以上のような特徴から、正確な眼球構造や屈折の計算式を求めるのは難しいと考えられる。そこで、本稿では虹彩の見え方を直接モデル化し、高速かつ高精度に虹彩中心を求める。なお、このモデルの曲率半径（厚み）や旋回中心は実際の顔画像を用いて設定する。また、両目・複数台カメラを用いて複数の視線情報を独立に求め、それらを統合することで誤差の影響を抑える。

光の屈折はスネルの法則を用いて計算可能であるが、本稿では虹彩の輪郭線を高速に計算するため、図 5 のような球面形状を持たせた虹彩モデルを提案する。モデルの曲率半径や旋回中心を正確に計算することは難しいが、本稿では実画像をもとに再推定誤差が最も小さくなるように設定した。以下、画像上の虹彩輪郭の計算方法について述べる。まず、楕円の傾きと短軸の長さは虹彩モデルの光軸とカメラとのなす角 α から求められる。また、画像上でのモデルの厚み l は以下の式で求められる。

$$l = c - \sqrt{c^2 - r^2} \sin \alpha$$

ここで、今回は実画像をもとに虹彩半径 $r = 5.5\text{mm}$ 、虹彩モデルの曲率半径 $c = 5.8\text{mm}$ 、虹彩中心と旋回中心の距離 $d = 13\text{mm}$ とした。虹彩検出処理の詳細については 3.4 で述べる。なお、光軸とカメラの角度 α が小さい場合 ($\alpha < \arccos \frac{r}{c}$, 約 18°) には接点が虹彩モデルの断面よりも後ろになり、画像上での虹彩の見え方は厚みのない、ただの楕円となる。

3. 視線推定処理の詳細

3.1 各座標系の定義

各座標系を図 6 のように定義する。座標系は全て右手系とし、ワールド座標系の原点をディスプレイの中央正面 1m の所に置く。ディスプレイの法線方向を Z_w とし、水平方向に X_w 、垂直方向に Y_w を定義する。各カメラ座標系はレンズ中心を原点、光軸方向を Z_c とする。また、ワールド座標系と画像座標系の関係は以下の式で与えられる。

$$\lambda \begin{bmatrix} u & v & 1 \end{bmatrix}^T = P \begin{bmatrix} X_w & Y_w & Z_w & 1 \end{bmatrix}^T$$

ただし、 P は射影行列であり、カメラキャリブレーションとレンズ歪みの補正は OpenCV の関数を用いている。

3.2 特徴点抽出

顔の特徴点の抽出には Active Appearance Model(AAM) [12], [13] を用いる。これは、AAM では顔部品の輪郭を抽出することができ、将来は眉や口の輪郭を用いた表情の認識や、発話タイミングの検出などにも利用できると思われるためである。特徴点抽出処理の詳細については本稿では省略する。

モデルの学習にはディスプレイ上の色々な場所を見ている複数枚の画像を用い、特徴点には両目、両眉、鼻、口の輪郭、計 45 点を用いた。作成した学習モデルの例を図 7 に示す。

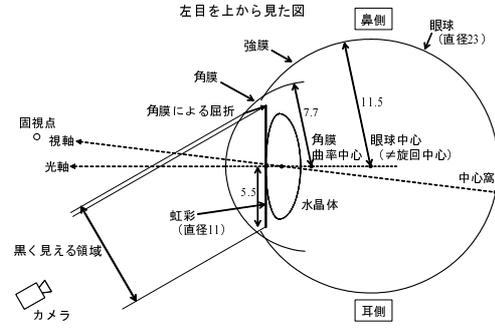
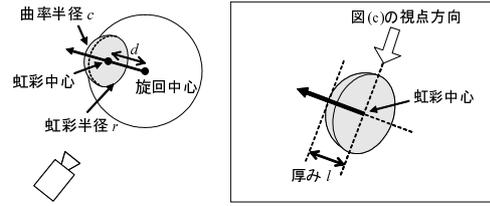


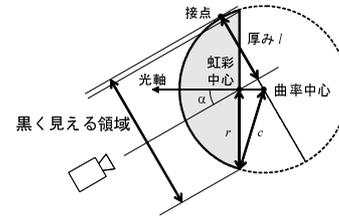
図 4 眼球構造モデル

Fig. 4 Structural model of eyeball.



(a) 虹彩モデル

(b) 画像上での虹彩の見え方



(c) 虹彩の輪郭線の算出

図 5 虹彩の輪郭形状の計算

Fig. 5 Calculation of iris outline.

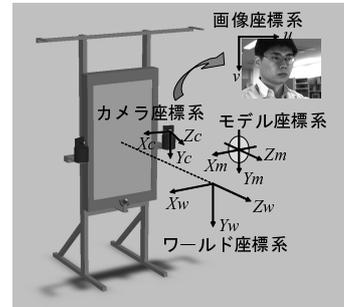


図 6 各座標系の定義

Fig. 6 Definition of each coordinate system.

3.3 頭部姿勢推定

頭部姿勢のパラメータ ξ は以下の 6 変数（回転 3 自由度，平行移動 3 自由度）で定義する。

$$\xi = \begin{bmatrix} \phi & \theta & \psi & t_x & t_y & t_z \end{bmatrix}^T$$

この時、 i 番目の特徴点のモデル座標 $(X_{m_i}, Y_{m_i}, Z_{m_i})$ からワールド座標 $(X_{w_i}, Y_{w_i}, Z_{w_i})$ への変換は以下の式で与えられる。

$$\begin{bmatrix} X_{w_i} \\ Y_{w_i} \\ Z_{w_i} \end{bmatrix} = R_z(\phi) R_y(\theta) R_x(\psi) \begin{bmatrix} X_{m_i} \\ Y_{m_i} \\ Z_{m_i} \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$



図 7 AAM の例

Fig. 7 Example of Active Appearance Model.

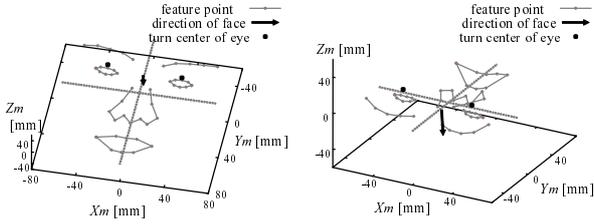


図 8 3次元顔形状モデル

Fig. 8 3D face shape model.

頭部姿勢の推定には最急降下法を用い、評価式 f は AAM で抽出した特徴点座標 (u_i, v_i) と、頭部姿勢 ξ から再投影した座標 $(u_i(\xi), v_i(\xi))$ との誤差で定義する。

$$\xi_{k+1} = \xi_k - a \frac{\partial f}{\partial \xi}(\xi_k)$$

$$f(\xi) = \sum_i w_i \sqrt{\{u_i(\xi) - u_i\}^2 + \{v_i(\xi) - v_i\}^2}$$

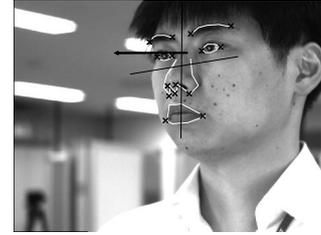
ただし、 a は学習係数である。また、 w_i は各特徴点の重みであり、AAM で抽出した特徴点 45 点の内、表情変動の影響を受けにくい点のみ（計 15 点、図 9(a) を参照）を用い、かつ、目の位置を重視するために以下のように設定した。

$$w_i = \begin{cases} 5 & (\text{目の両端, 計 4 点}) \\ 1 & (\text{眉の両端, 鼻腔周辺, 口の両端, 計 11 点}) \\ 0 & (\text{その他}) \end{cases}$$

頭部姿勢推定に用いる 3 次元顔形状モデル（図 8）は、各特徴点の座標、顔向きを表すベクトル、眼球旋回中心の座標の 3 つで構成され、学習用のステレオ動画をを用いて予め作成しておく。まず、ディスプレイ中央を注視している画像を用いて各特徴点の 3 次元座標をステレオ復元する。また、両目の両端計 4 点の重心とディスプレイ中央を結んだ直線を、形状モデルにおける顔向きとして学習する。次に、ディスプレイ上に表示した複数のマーカを順に注視したステレオ画像を用い、各マーカと虹彩中心の 3 次元座標から顔形状モデル中の眼球旋回中心を求める。なお、虹彩中心と旋回中心の距離 d は、注目座標の再推定誤差が最も小さくなった 13mm とした。

3.4 虹彩検出

高速に虹彩を検出するため、特徴点抽出と頭部姿勢推定の結果（図 9(a)）を利用する。まず、3.2 で求めた目領域（図 9(b)）に対して大津の二値化 [14] を行い、白色領域・黒色領域・目領域外の三つに分類する（図 9(c)）。次に、2.3 の虹彩モデルを用いて画像上での虹彩の輪郭形状を算出する。カメラからの奥行きは目頭と目尻の平均、視線方向は 3.3 で求めた顔向きに等しいと仮定して、虹彩の大きさや形状を決定する（図 9(d)）。内側の



(a) 特徴点抽出と頭部姿勢推定の結果



(b) 目領域 (左目) (c) 領域分類結果 (d) 探索形状

図 9 虹彩検出

Fig. 9 Iris detection.

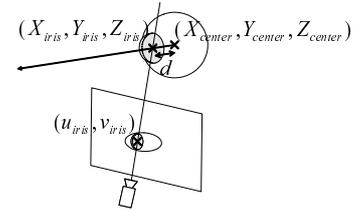


図 10 虹彩中心と視線の推定

Fig. 10 Estimation of iris center and gaze.

黒色領域が算出した虹彩形状であり、外側の白色領域は虹彩の断面を一回り大きくした楕円である（経験的に虹彩の 1.2 倍とした）。この領域分類画像 I_{src} と探索形状 I_{temp} を用いてテンプレートマッチングを行い、最も評価値 $E_{iris} (-1 \leq E_{iris} \leq 1)$ の高い位置を虹彩中心として求める。

$$I_{src} = \begin{cases} 1 & (\text{黒色領域}) \\ -1 & (\text{白色領域}) \\ 0 & (\text{目領域外}) \end{cases}, \quad I_{temp} = \begin{cases} 1 & (\text{黒色領域}) \\ -1 & (\text{白色領域}) \\ 0 & (\text{目領域外}) \end{cases}$$

$$E_{iris}(u, v) = \frac{\sum I_{src}(u + u', v + v') I_{temp}(u', v')}{\sum \{I_{temp}(u', v')\}^2}$$

3.5 視線推定

視線は眼球の旋回中心と虹彩中心を結ぶ直線と定義し、視線とディスプレイ平面の交点を注目場所（固視点）として求める。まず、旋回中心の 3 次元座標 $(X_{center}, Y_{center}, Z_{center})$ は、3.3 で求めた頭部姿勢から算出される。次に、旋回中心と虹彩中心との距離 d と、虹彩中心の画像上座標 (u_{iris}, v_{iris}) をもとに、虹彩中心の 3 次元座標 $(X_{iris}, Y_{iris}, Z_{iris})$ が以下の連立方程式から求められる（図 10 参照）。

$$\lambda \begin{bmatrix} u_{iris} \\ v_{iris} \\ 1 \end{bmatrix} = \mathbf{P} \begin{bmatrix} X_{iris} \\ Y_{iris} \\ Z_{iris} \\ 1 \end{bmatrix}, \quad \left\| \begin{bmatrix} X_{iris} \\ Y_{iris} \\ Z_{iris} \end{bmatrix} - \begin{bmatrix} X_{center} \\ Y_{center} \\ Z_{center} \end{bmatrix} \right\| = d$$

ただし、解が 2 つ求まるため、ディスプレイに近い方を正解の虹彩中心 $(X_{iris}, Y_{iris}, Z_{iris})$ とする。そして、ディスプレイ上

の注目座標 X_{gaze}, Y_{gaze} が以下の式から求められる．

$$\begin{bmatrix} X_{gaze} \\ Y_{gaze} \\ Z_{gaze} \end{bmatrix} - \begin{bmatrix} X_{center} \\ Y_{center} \\ Z_{center} \end{bmatrix} = \lambda \left(\begin{bmatrix} X_{iris} \\ Y_{iris} \\ Z_{iris} \end{bmatrix} - \begin{bmatrix} X_{center} \\ Y_{center} \\ Z_{center} \end{bmatrix} \right)$$

なお、現在のシステムではワールド座標系の原点をディスプレイの前方 1m に置いているため、 $Z_{gaze} = -1000$ である．

3.6 単眼視線推定結果の統合

3.5 で両目・各カメラについて独立に推定した視線情報を統合する．合計 6 つの視線情報（カメラ 3 台 × 左右の目）の内、顔の検出に失敗したものや虹彩検出の評価値が低いものを除き、残った視線情報の重み付き平均によってディスプレイ上の注目座標 $(\bar{X}_{gaze}, \bar{Y}_{gaze})$ を求める．なお、今回の実装ではこの重みに虹彩検出の評価値 E_{iris} を用い、虹彩検出成功の閾値を経験的に $E_{iris} > 0.40$ とした．

$$\bar{X}_{gaze} = \frac{\sum E_{iris} X_{gaze}}{\sum E_{iris}}, \quad \bar{Y}_{gaze} = \frac{\sum E_{iris} Y_{gaze}}{\sum E_{iris}}$$

4. 評価実験

4.1 視線推定精度の評価

ディスプレイ上に 20cm 間隔で 15 点のマーカ（図 11）を表示し、約 1m の距離からそれらを順に注視した様子を撮影した動画を用いて視線推定の性能評価を行った．

まず提案モデルの有効性を調べるため、学習用画像を用い、特徴点抽出や頭部姿勢推定など、虹彩検出以外の誤差を最小にして提案モデルと従来の円板モデルそれぞれで虹彩検出を行った（図 12）．なお、虹彩の大きさや形状は頭部姿勢をもとにそれぞれ算出している．図 12 は左カメラ同一フレームにおける左目の虹彩検出結果であり、マーカ⑦のように視線とカメラの向きが近い場合は同一の結果であるが、角度が大きくなると差異が生じる．特にマーカ⑧ではまぶたによる隠れや照明によるノイズによって虹彩抽出に大きな違いが生じた．また、マーカ⑨における虹彩中心の差は 2.2pixel であるが、同じ眼球旋回中心を用いた場合には円板モデルの視線推定誤差がディスプレイ上で 33mm 大きくなった．なお、視線とカメラの角度 α は注視したマーカ⑦、⑧、⑨でそれぞれ約 10° 、 20° 、 30° であり、 $\alpha > 30^\circ$ では奥の顔特徴点が隠れて頭部姿勢推定が困難になる．

次に、評価用動画（UXGA）に対して視線推定処理を行った．右・左・下の各カメラ単眼での視線推定結果と 3 台を統合した結果を表 1 に示す．公平な比較を行うためにそれぞれのモデルで学習した眼球旋回中心を用いたため、視線とカメラの角度が小さい場合でも推定結果に違いが生じている．特に、カメラから遠いマーカ（右カメラの①④や左カメラの③⑥⑨）で円板モデルの方が誤差が大きくなると予想していたが、実際には逆に円板モデルの精度が良くなってしまった．しかし、その他のマーカでは全体的に提案モデルの方が精度が良くなっており、円板モデルを用いて学習した旋回中心に誤差が含まれているためと考えられる．なお、同じ旋回中心を用いた場合にはカメラから遠いマーカでのみ違いが生じたが、誤差が大きくなる場合と小さくなる場合が混在した．これは虹彩検出のずれの影

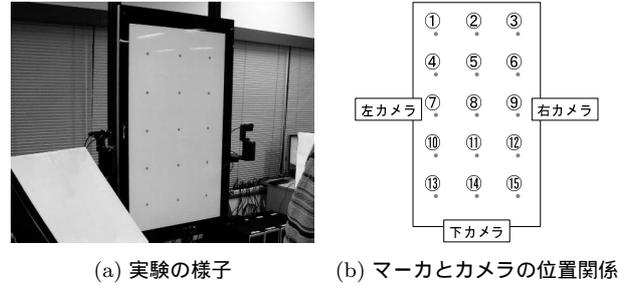


図 11 実験環境

Fig. 11 Experimental environment.

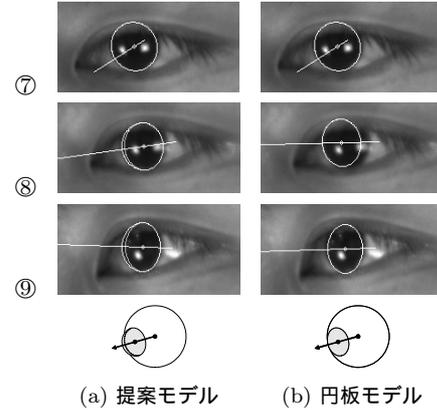


図 12 虹彩検出結果

Fig. 12 Results of iris detection.

響（15mm/pixel 程度）よりも顔特徴点抽出などの誤差が大きいためと考えられる．また、虹彩検出の評価値 E_{iris} は提案モデルの方が円板モデルよりも高くなる傾向があり、画像上でも実際の虹彩輪郭に近い形状であることが確認でき、より正確な虹彩輪郭形状を得られていると考えられる．

比較のため評価用動画を XGA に縮小して同様の視線推定を行った結果（表 1 の右欄）、統合後の平均誤差が 78.3mm となった．UXGA カメラを用いることで虹彩検出精度が上がり、XGA より約 25mm 視線の推定精度が向上したと考えられる．

UXGA カメラ 3 台統合による推定誤差 52.9mm は角度に直すと約 3 度であり、より高精度な推定を行うためには特徴点抽出の安定化や表情変化・顔形状の個人差などへの対応が必要と考えられる．なお、今回の実験では眼鏡を外しており、眼鏡ありの場合には屈折による位置ずれやフレームによる隠れ、反射などによって推定精度が低下し、単眼では推定が不可能なケースも発生する．しかし、本システムでは複数台のカメラを独立に用い、それらの推定結果を統合することによって眼鏡ありの場合でも平均推定誤差 100mm 程度に抑えられている．

4.2 高解像度撮影における処理時間

現在、各カメラはそれぞれ別 PC（Intel Core2 Duo 2.66GHz、メモリ 2GB）で処理し、その結果を 1 台の PC に集めて統合している．画像の取得から推定結果の送信までにかかる時間を表 2 に示す．なお、処理の高速化のために Intel Integrated Performance Primitives を用い、レンズの歪み補正には OpenCV の関数 `cvRemap`（最近傍）、顔の特徴点抽出には縮小画像を用いたピラミッド処理、左右の虹彩検出の並行処理などを行って

表 1 視線推定誤差 [mm] (各 50 フレームの平均, 右・左・下は単眼の推定結果)

Table 1 Error of gaze estimation.

マーカー	評価用動画画像 (UXGA)								縮小動画画像 (XGA)			
	提案モデル				円板モデル				提案モデル			
	右	左	下	統合	右	左	下	統合	右	左	下	統合
①	89.1	13.3	68.5	38.2	<u>64.1</u>	24.4	48.1	29.5	96.1	64.4	38.3	32.5
②	63.6	64.7	23.1	12.8	78.3	70.3	39.9	31.0	86.4	83.3	47.2	36.5
③	109.9	59.8	37.0	44.9	114.2	<u>25.7</u>	39.0	35.1	113.7	92.4	55.9	72.3
④	100.1	66.3	59.3	46.1	<u>95.8</u>	84.8	64.7	61.7	112.1	98.2	79.9	72.8
⑤	93.5	92.3	46.2	34.3	102.5	107.5	54.6	46.1	43.5	185.0	59.6	48.1
⑥	71.5	42.3	17.1	28.8	92.2	<u>32.2</u>	13.3	43.2	108.3	82.7	35.7	57.9
⑦	95.5	108.6	62.3	72.5	98.4	125.1	74.1	85.2	167.4	133.2	113.8	97.1
⑧	82.5	54.4	32.9	39.1	89.7	77.3	42.1	46.7	93.1	35.5	59.6	30.9
⑨	116.0	83.9	46.5	70.3	130.6	<u>78.0</u>	49.0	78.5	179.1	170.9	68.5	110.6
⑩	—	—	60.5	60.5	—	—	71.3	71.3	—	—	132.5	132.5
⑪	—	—	44.8	44.8	—	—	56.3	56.3	—	—	26.5	26.5
⑫	—	—	19.0	19.0	—	—	25.4	25.4	—	—	99.9	99.9
⑬	—	—	103.6	103.6	—	—	110.3	110.3	—	—	116.6	116.6
⑭	—	—	119.7	119.7	—	—	128.1	128.1	—	—	164.5	164.5
⑮	—	—	59.0	59.0	—	—	65.7	65.7	—	—	76.3	76.3
平均誤差	91.3	65.0	53.3	<u>52.9</u>	96.2	69.5	58.8	<u>60.9</u>	111.1	105.1	78.3	<u>78.3</u>

いる．特に本手法では特徴点抽出や頭部姿勢推定結果を利用することで，高速に虹彩を検出する．以上の処理により，UXGA カメラを用いた場合でも 30fps での視線推定を実現した．ただし，画面表示は結果の確認用であるため省略している．また，以上の処理時間は前フレームの推定結果を利用する安定追跡時のものであり，顔を発見した最初のフレームでは約 100ms を要するため，初期探索部分にはまだ改良の必要がある．

5. むすび

本稿ではカメラ画像を用いた視線推定の高精度化について検討した．高精度な視線推定を実現するためには，角膜における光の屈折を考慮する必要があり，ビデオレートで処理するためには高速な推定方法が必要となる．球面形状を持つ虹彩モデルを用い，特徴点抽出や頭部姿勢推定結果を利用することで，高速かつ高精度に虹彩を検出する方法を提案した．そして，実際に UXGA カメラ 3 台を用いて精度 3 度，30fps で動作する視線推定システムを構築した．今後の課題としては，曲線部分に対する特徴点抽出の安定化，顔形状の個人差や表情変化への対応，初期探索部分の高速化などが挙げられる．

謝辞 本研究の一部は文部科学省科学研究費補助金特定領域研究 No.18049046 の補助を受けて行った．

文 献

- [1] 佐竹純二, 小林亮博, 川嶋宏彰, 平山高嗣, 水口充, 小嶋秀樹, 松山隆司, “インタラクティブな情報提示システムのための非装着・非拘束な視線推定”, 情処学 HCI 研報, no.99, pp.9-16, 2007.
- [2] 水口充, 浅野哲, 佐竹純二, 小林亮博, 平山高嗣, 川嶋宏彰, 小嶋秀樹, 松山隆司, “Mindprobing: システムの積極的な働きかけによる視線パタンからの興味推定”, 情処学 HCI 研報, no.99, pp.1-8, 2007.
- [3] 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.

表 2 1 フレーム当たりの処理時間

Table 2 Time of processing a frame.

処理内容	UXGA	XGA
前処理 画像取得	0.8ms	0.5ms
レンズ歪み補正	11ms	4.5ms
画像処理 顔特徴点抽出	15ms	10ms
頭部姿勢推定	3ms	3ms
虹彩・視線推定	1.3ms	1ms
後処理 画面表示	(10ms)	(7ms)
合計 (画面表示は省略)	31ms	19ms

- [4] C. Hennessey, B. Nouredin, and P. Lawrence, “A Single Camera Eye-Gaze Tracking System with Free Head Motion”, ETRA, pp.87-94, 2006.
- [5] R. Newman, Y. Matsumoto, S. Rougeaux, and A. Zelinsky, “Real-Time Stereo Tracking for Head Pose and Gaze Estimation”, FG, pp.122-128, 2000.
- [6] 松本吉央, 怡土順一, 竹村憲太郎, 小笠原司, “リアルタイム顔・視線計測システムの開発と知的インタフェースへの応用”, 情処学論 (CVIM), vol.47, no.SIG15, pp.10-21, 2006.
- [7] 山添大丈, 内海章, 安部伸治, “顔特徴点の追跡を用いた単眼姿勢推定”, MIRU, pp.319-324, 2007.
- [8] J.G. Wang, E. Sung, and R. Venkateswarlu, “Estimating the eye gaze from one eye”, CVIU, vol.98, pp.83-103, 2005.
- [9] 辻徳生, 柴田真吾, 長谷川勉, 倉爪亮, “視線計測のための LMedS を用いた虹彩検出法”, MIRU, pp.1-684-689, 2004.
- [10] 魚里博, “ヒトの目に特有な光学”, O plus E, vol.22, no.4, pp.418-430, 2000.
- [11] ぎもんしつもん目の辞典, <http://www.ocular.net/jiten/>
- [12] T.F. Cootes, G.J. Edwards, and C.J. Taylor, “Active appearance models”, PAMI, vol.23, no.6, pp.681-684, 2001.
- [13] M.B. Stegmann, B.K. Ersbøll, R. Larsen, “FAME - A Flexible Appearance Modeling Environment”, IEEE Trans. Medical Imaging, vol.22, no.10, pp.1319-1331, 2003.
- [14] 大津展之, “判別および最小 2 乗規準に基づく自動しきい値選定法”, 信学論 (D), vol.63, no.4, pp.349-356, 1980.