

視線情報を用いた注視点の3次元位置推定

満上 育久[†] 浮田 宗伯[†] 木戸出正継[†]

[†] 奈良先端科学技術大学院大学 情報科学研究科 〒 630-0192 奈良県生駒市高山町 8916-5

E-mail: †{ikuhi-mi,ukita,kidode}@is.aist-nara.ac.jp

あらまし 本研究では、人間の視線運動を計測することでその人間が注視している対象点の3次元位置を推定するシステムを提案する。このシステムを構築する際、視線方向の計測精度が非常に重要になるが、人間の眼から注視点への正確な方向を知るには次のような問題を含む：1) 眼球計測装置の精度が不十分である、2) 人間は1点を非常に正確に注視しているつもりでも、視線は正確にその方向を向いていない。本研究では、まず1)についてはカメラの光学的な誤差要因を抑制することでキャリブレーション時に生じる誤差の軽減を図る。さらに、2)のように不可避な誤差を含む視線情報からでも注視点位置を高精度に推定することができるように、ある瞬間の視線情報のみではなく、人間が頭を移動させながら1点を注視することにより得られる複数視線情報を統合する手法を提案する。この手法は、各視線情報の信頼度を利用した確率的アルゴリズムとして実現されているため、高いロバスト性と精度を備えている。また、実験により提案手法の有効性を確認した。

キーワード 注視点計測、カメラキャリブレーション、3次元確率密度分布

Estimation of 3D Gaze Position Using View Lines

Ikuhisa MITSUGAMI[†], Norimichi UKITA[†], and Masatsugu KIDODE[†]

[†] Graduate School of Information Science, Nara Institute of Science and Technology

Takayama-cho 8916-5, Ikoma-shi, Nara, 630-0192 Japan

E-mail: †{ikuhi-mi,ukita,kidode}@is.aist-nara.ac.jp

Abstract We propose a new wearable system that can estimate the 3D position of a gazed point, by measuring binocular view lines. It is difficult to measure these lines with a certain accuracy by an eye tracker, because of errors caused by 1) difficulty in fulfilling the conditions of the device calibration and 2) the limitation that a human cannot gaze very accurately at a distant point. Concerning 1), the accuracy of the device calibration can be improved by considering the optical properties of a camera in the device. To solve 2), we propose a stochastic algorithm that estimates a gazed 3D position by integrating information of view lines from multiple head positions.

Key words gaze position measurement, camera calibration, spacial probability density

1. はじめに

人間と機械のインタラクションにおいて最も重要なことは、機械系が人間の注意対象・意図・判断等を正しく認識し動作することである。こうした意図抽出を目的として、従来カメラによるジェスチャー認識や、マイクによる音声入力などといったインタフェースに関する研究が多く行われてきている。しかし、これらの意図提示においては、コマンドパターンが複雑、ユーザにとって操作が複雑であるなどの問題があり、現在インタフェース研究において、よりユーザにやさしく使いやすい意図抽出手法に対するアプローチは多く行われている。それらのアプローチの中のひとつとして、人間の視線運動情報を利用する手法が挙げられる。この視線情報が特徴的なのは、以下のようである。

- 直接人間の興味対象の方向を示す。

- 視線を動かすことは無意識の動作なので、システム動作のために人間に余計な動作を強いる必要がない。

ユーザインタフェースの分野において、これらの特徴に注目した人間の意図抽出タスクの研究の例は多い[1][2][3]。その中でも、注視点の3次元位置推定タスクは、環境内のユーザの興味対象推定やロボットナビゲーションにおける目的地指定などの応用が可能となる点で非常に有用であり、本研究では人間の視線運動計測に基づいた注視点3次元位置推定を行うシステムを提案する。

人間は対象を注視する際、その奥行きを推定するために左右眼の視線方向と眼間距離情報に基づいた三角測量の原理を用いている。しかし、人間は注視した際にも視線を正確に注視点に向けることはできず、さらに視線は方向誤差 10° 程度の範囲でランダムに振動することが知られている。よって、たとえある瞬間の実際の視線方向を非常に正確に計測できたとしても、

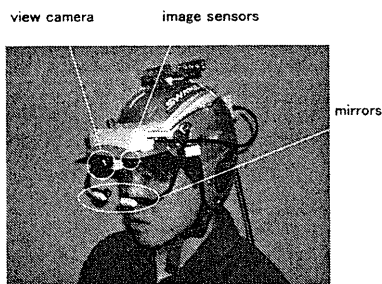


図1 両眼眼球運動計測装置 EMR-8
Fig.1 Binocular eyemark recorder EMR-8

人間の真の注視点の正確な位置推定は困難である。その上、両眼視による三角測量においてベースラインとなる眼間距離が一般に 65mm 程度と非常に小さく、注視点の位置推定をさらに困難にしている^(注1)。

本研究では両眼の視線方向情報をベースとした注視点の位置推定の精度向上を目的とする。上記のように視線方向の誤差は不可避であるが、ベースラインを広げることで精度向上は可能である。当然眼間距離自体を広げることは不可能なので、ユーザが頭部を動かして複数の視点から1点を注視することで単眼カメラによるステレオ視と同様の効果を得ようというアプローチを提案する。この手法では、ある瞬間の両眼の視線方向だけではなく、複数視点からの多数の視線方向情報を取り扱う。それらすべてに不可避な方向検出誤差が含まれており、そこから精度よくひとつの解を求める必要がある。本研究では、各視線方向情報にガウス分布の方向誤差があると仮定し、ひとつの視線方向情報を1本の直線ではなく円錐状の確率密度分布で表現し、視線方向の信頼度とともに、各瞬間における両眼視による注視点の3次元位置推定結果をよる信頼度を考慮することで、誤差を含んだ視線群から確率的に最適解を推定値として採用する。

2. 両眼眼球計測装置の概要

2.1 装置の構成

視線方向情報を用いるので、その計測デバイスが必要となる。本研究では(株)ナックイメージテクノロジー社の眼球運動計測装置 EMR-8 を用いる^(注2)。EMR-8 はコントローラ部とヘッドユニット部から構成されている。ヘッドユニット部(図1)にはイメージセンサと反射鏡がついており、反射鏡が眼球周辺をイメージセンサに反射し観測させた上で、瞳孔/角膜反射方式によってイメージセンサ観測画像における眼球の位置を観測する。またヘッドユニットの中央には、視野カメラが装備されており、リアルタイムにユーザの視野の画像を観測している。コントローラ部は、これらの入力を受け取り、次節のキャリブレーションによって入力情報の対応関係を求める。この対応関係に基づいて視線方向を推定し、視野カメラ画像にその方向が重畳表示される。

2.2 装置のキャリブレーション

イメージセンサによって眼球を観測する際、眼球・イメージセンサ・視野カメラの空間的な相対位置関係は未知である。そ

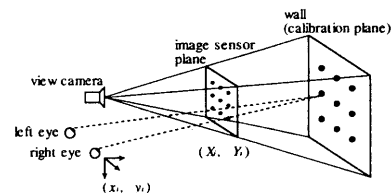


図2 装置のキャリブレーション
Fig.2 Device calibration

のためまず注視の際に、イメージセンサ画像における x, y 方向への眼球の変位量として定義される視線角度 (x_i, y_i) と視野カメラの画像座標系における注視点位置の対応関係を調べ対応式を得る。以後、得られた対応式をもとに眼球観測結果から視線方向を計算可能となる。ここでの事前の対応式作成のための作業をキャリブレーションと呼んでいる。この作業は(株)ナックイメージテクノロジー社によって提供されるオペレーションである。

具体的なキャリブレーション作業について述べる。まず実環境中に壁などの1平面を用意し、EMR-8 ヘッドユニットの視野カメラの光軸が壁面に垂直になるようにユーザ頭部位置を定める。このときのカメラと壁面間の距離 l_c をキャリブレーション距離と呼び、その壁面をキャリブレーション面と呼ぶこととする。視野カメラの画像座標系上で座標値が既知の9つの注視用点状ターゲットを (X_i, Y_i) ($i = 0, \dots, 8$) を実環境中のキャリブレーション面に投影し、投影された点をユーザが注視した際の眼球の視線角度 (x_i, y_i) を観測する(図2)。これらの座標値対から対応式を決定する。対応式は以下のような共二次方程式を採用している。

$$X_i = f(x_i, y_i) = a_0 + a_1 x_i + a_2 y_i + a_3 x_i^2 + a_4 x_i y_i + a_5 y_i^2 \quad (1)$$

$$Y_i = g(x_i, y_i) = b_0 + b_1 x_i + b_2 y_i + b_3 x_i^2 + b_4 x_i y_i + b_5 y_i^2 \quad (2)$$

この式を解くために必要な点は6つであるが、ここでは9点のデータを採用して最小2乗法で誤差最小となるような a_i, b_i を求める。

2.3 キャリブレーション時の誤差軽減

本研究の目的である視線方向情報による注視点位置計測を精度よく行うためには、この装置キャリブレーションを正確に行う必要がある。ここではキャリブレーションに際の誤差軽減処理を提案、実装し、その効果を確認する。

2.3.1 視野カメラのレンズ歪み補正

一般にカメラによって撮影される画像は、カメラ内のレンズの影響で歪んでしまう。ここで利用している視野カメラも同様に歪みが起こるが、対応式を決定する際の (X_i, Y_i) にこの歪み誤差が含まれたままキャリブレーションが行われている。そこで、視野カメラについて Tsai のカメラキャリブレーション [5] を行い、リアルタイムで視野画像に対してレンズ歪み補正を行った。

2.3.2 壁面の垂直性の確認

実環境中、キャリブレーション面は視野カメラの光軸に対して垂直であるという仮定のもとに対応式を決定している。そこで、平面上に点が 3×3 の格子状に配置されたターゲットを用意し、それを壁面に配置し視野カメラで観測して、その見えを観測することで壁面の垂直性をリアルタイムに確認するようにした。

(注1): 人間は実際には、この両眼視線情報だけでなく、視野内のオブジェクトの形状変化や影の様子、色や大きさといった知識を無意識的に用いており、これらを統合して注視点位置を推定することができるため、比較的正確な位置推定が可能である。

(注2): EMR-8 はユーザ装着型のデバイスであるが、本研究の提案手法はユーザの眼の位置とその視線方向が分かればよいので、大野らの FreeGaze [2] などの非装着型の視線計測デバイスでも構わない。

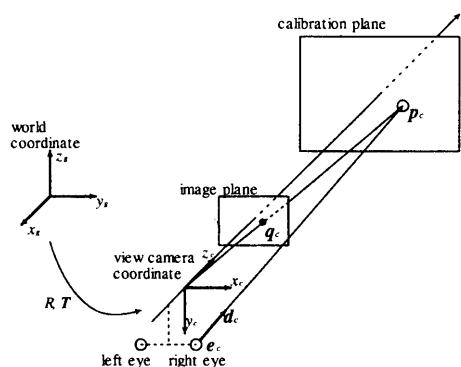


図3 視線直線
Fig.3 View line

2.3.3 軽減処理の効果

上記2つの誤差軽減処理の効果を確認する実験を行った。誤差軽減を一切行わない場合の視線方向の角度誤差は1.413度、上記2つの誤差軽減を行った場合は0.917度となり、誤差軽減の効果を確認できた。なお、角度誤差とは、視野カメラ画像の注視用点状ターゲットの位置から求められるターゲットの方向ベクトルと、眼球計測によって得られる視線方向ベクトルがなす角度である。

3. 視線直線と注視点位置推定

3.1 カメラ座標系における視線直線の表現

図3を用いてカメラ座標系における視線直線の決定方法について説明する。なお、本節におけるベクトル表記はいずれも視野カメラ座標系による3次元ベクトルである。図中の \$e_c\$ の位置にある眼球方向から得られる視線直線が、キャリブレーション面と交差する点を \$p_c\$ とする。キャリブレーション面は視野カメラ光軸に対して垂直に配置されており、点 \$p_c\$ は視野カメラの画像平面上の点 \$q_c\$ に投影される。2.2節で示したとおり、EMR-8のキャリブレーションの際には視線角度と視野カメラ画像2次元座標の対応関係を求めているので、観測視線角度によって得られるのは図の \$q_c\$ である。そこにカメラの中心からキャリブレーション面までの距離を与えることで \$p_c\$ が求められる。点 \$p_c\$ と眼球を結ぶ直線が視線直線なので、カメラ座標系における眼球の3次元位置 \$e_c\$ を実環境で計測し^(注3)、これら2点から視線直線を決定する。視線直線の単位長さ方向ベクトル \$d_c\$ は

$$d_c = \frac{p_c - e_c}{|p_c - e_c|} \quad (3)$$

となり、カメラ座標系における視線直線の一般式は、任意の実数 \$m\$ として \$e_c + md_c\$ と表現でき、以後これを \$(e_c, d_c)\$ と記述することとする。

3.2 世界座標系への変換

本研究では複数視点からの視線情報を扱う。そのため、3.1節のように各視線直線を各瞬間におけるのカメラ座標系で取り扱うのは不都合である。そこで、全視点共通の世界座標系を設定し、各視点の視野カメラ座標系の視線直線を世界座標系に変換する。この際、各視点での視野カメラの位置・方向といった外部パラメータが必要となるが、ここでは既知として議論を進める^(注4)。

(注3) : Appendix 参照。

(注4) : 本研究における視野カメラの外部パラメータ計測の実装法については、6.章で述べる。

ある視点での視野カメラの位置・方向が世界座標系において平行移動ベクトル \$T\$ と回転行列 \$R\$ で表されるとする(図3)。

$$T = \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}, R = \begin{pmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{pmatrix}$$

このとき、世界座標系における眼球位置 \$e_g\$、視線方向ベクトルを \$d_g\$ とすると、

$$e_c = R e_g + T \quad (4)$$

$$d_c = R d_g \quad (5)$$

\$R\$ の逆行列 \$R^{-1}\$ を求めることで、\$e_g, d_g\$ が求まる。

$$e_g = R^{-1}(e_c - T) \quad (6)$$

$$d_g = R^{-1}d_c \quad (7)$$

3.3 両眼視線方向による位置推定

3.1, 3.2節の議論より、世界座標系における各視点からの両眼の視線直線が計算できる。本研究の提案手法は、各瞬間における左右眼の視線直線のみから得られる注視点の3次元位置推定結果を信頼度判定に利用している。ここでは、その各瞬間の両眼視線情報からの3次元位置推定の方法について述べる。

平面上の三角測量と異なり、両眼の視線は3次元空間中の2直線となるため、一般には交差せず、2直線の交点として解を得ることができない。そこで、両眼の視線直線の方程式について最小2乗法的に解を求める。具体的な計算法を以下に示す。

世界座標系における左右眼の視線直線が、\$(e_{gl}, d_{gl}), (e_{gr}, d_{gr})\$ であるとする。

$$e_{gl} = \begin{pmatrix} e_{xl} \\ e_{yl} \\ e_{zl} \end{pmatrix}, d_{gl} = \begin{pmatrix} d_{xl} \\ d_{yl} \\ d_{zl} \end{pmatrix}$$

$$e_{gr} = \begin{pmatrix} e_{xr} \\ e_{yr} \\ e_{zr} \end{pmatrix}, d_{gr} = \begin{pmatrix} d_{xr} \\ d_{yr} \\ d_{zr} \end{pmatrix}$$

左眼の視線直線に着目する。この視線直線の方程式は以下のようになる。

$$\frac{x - e_{xl}}{d_{xl}} = \frac{y - e_{yl}}{d_{yl}} = \frac{z - e_{zl}}{d_{zl}} \quad (8)$$

これから、以下の2式を導出する。

$$d_{zl}x - d_{xl}z = d_{zl}e_{xl} - d_{xl}e_{zl} \quad (9)$$

$$d_{zl}y - d_{yl}z = d_{zl}e_{yl} - d_{yl}e_{zl} \quad (10)$$

右眼に関しても同様の式が導出できる。これらを行列で表記すると、

$$\begin{pmatrix} d_{zl} & 0 & -d_{xl} \\ 0 & d_{zl} & -d_{yl} \\ d_{zr} & 0 & -d_{xr} \\ 0 & d_{zr} & -d_{yr} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d_{zl}e_{xl} - d_{xl}e_{zl} \\ d_{zl}e_{yl} - d_{yl}e_{zl} \\ d_{zr}e_{xr} - d_{xr}e_{zr} \\ d_{zr}e_{yr} - d_{yr}e_{zr} \end{pmatrix} \quad (11)$$

これを解くことにより、最小2乗法的に解がひとつつまり、この解をこの瞬間における注視点座標とする。

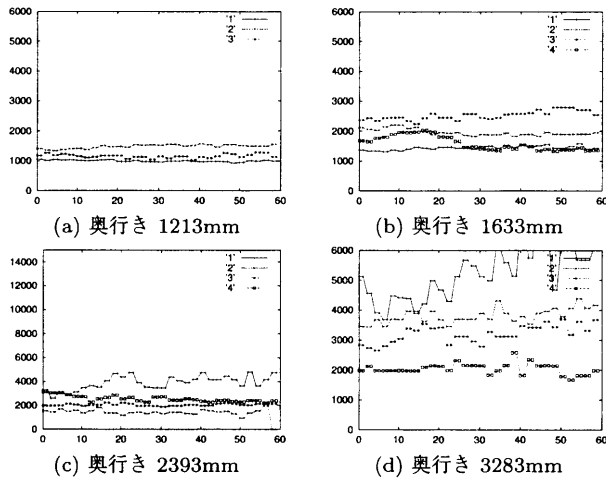


図4 各瞬間の両眼視線情報を用いた位置推定 (奥行きのみ)
Fig. 4 Estimation of a gazed position by instantaneous binocular lines

3.4 基礎実験と結果

上記の位置推定法の挙動を確認するための基礎実験を行った。基礎実験の結果を図4に示す。視野カメラからの奥行き1213mm, 1633mm, 2393mm, 3283mmにおかれた注視用ターゲットを注視した際の推定結果を評価する。ここでは推定精度の差がもっとも現れやすい奥行きのみを比較、評価している。なお、各グラフの横軸は時間(1/30 sec)、縦軸は各時刻における両眼視線方向から推定されたターゲットの奥行き(mm)である。また、各奥行きに対して3ないし4回の注視試行を行っており、各グラフ中の1,...,4は各実験における試行の番号であり、グラフ間での番号の対応関係はない。

図4より、カメラからの奥行きが1.2m程度では比較的高い精度で推定できているのに対して、奥行きが増加するごとに精度は推定値は大きく揺らぎ、3m付近まで達すると推定値は真値とかげはなれていることが確認できる。この結果は[4]で指摘されているとおりの結果である。

4. 確率的アルゴリズムによる注視点位置推定

視線情報のように取り扱う入力に誤差が含まれていることが分かっている場合に、確率的なアプローチは非常に有効である。さらに、時系列画像におけるオブジェクトトラッキングの代表的手法である ConDensation アルゴリズム [6] のような多次元確率密度分布を反復的に更新する手法では、モデル設定が柔軟であり、視線の誤差モデルなどを反映させやすいと考えられる。本研究では、複数の視点からの左右眼視線方向情報から、人間の視線という特徴を踏まえた高精度の注視点推定を行うための確率的手法を提案し、その有効性を実験により確認する。

4.1 確率的アルゴリズムの詳細

提案アルゴリズムの処理は、3つのブロックに分けられる。以下、それぞれのブロックにおける処理について説明する。

4.1.1 前処理部

本アルゴリズムは、空間中に注視点座標の確率密度分布を形成し、順次更新していく手法であるが、視野カメラが観測している実空間全域に対応するには膨大なメモリ空間が必要となる。そこで、まず前処理として、注視点座標の解が存在するであろう解候補領域を推定し、そこにボクセル空間を設定する。解候補領域は、時系列的な視点・視線情報のうちの注視開始直後のある一定時間の情報から求めた初期解の座標を中心に形成する。

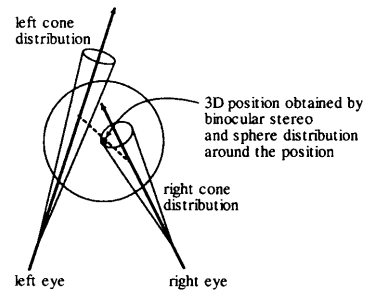


図5 確率密度分布の形成
Fig. 5 Generating a probability density distribution

複数視点から初期解を求める際には、式(11)を n 視点の視線直線のパラメータを取り扱うように拡張した以下の式を解く。

$$\begin{pmatrix} d_{z10} & 0 & -d_{x10} \\ \vdots & \vdots & \vdots \\ d_{z1n} & 0 & -d_{x1n} \\ 0 & d_{z10} & -d_{y10} \\ \vdots & \vdots & \vdots \\ 0 & d_{z1n} & -d_{y1n} \\ d_{zr0} & 0 & -d_{xr0} \\ \vdots & \vdots & \vdots \\ d_{zrn} & 0 & -d_{xrn} \\ 0 & d_{zr0} & -d_{yr0} \\ \vdots & \vdots & \vdots \\ 0 & d_{zrn} & -d_{yrn} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d_{z10}e_{x10} - d_{x10}e_{z10} \\ \vdots \\ d_{z1n}e_{x1n} - d_{x1n}e_{z1n} \\ d_{z10}e_{y10} - d_{y10}e_{z10} \\ \vdots \\ d_{z1n}e_{y1n} - d_{y1n}e_{z1n} \\ d_{zr0}e_{xr0} - d_{xr0}e_{zr0} \\ \vdots \\ d_{zrn}e_{xrn} - d_{xrn}e_{zrn} \\ d_{zr0}e_{yr0} - d_{yr0}e_{zr0} \\ \vdots \\ d_{zrn}e_{yrn} - d_{yrn}e_{zrn} \end{pmatrix} \quad (12)$$

これを解くことで、 n 組 $2n$ 本の視線直線に対する最小 2 乗法の解を得ることができる。

4.1.2 更新部

更新部では、前処理部で設定した解候補領域のボクセル空間に対して、1 対の視線情報が入力される度に、解の存在箇所に関する確率密度分布を形成し、ボクセル空間の各ボクセルに値を付加していく。

1 対の視線情報による確率密度分布の形成方法を図5に示す。まず、左右眼についてそれぞれ、視線直線を軸とし頂点を眼球位置とするような錐体を作成する。そしてその錐体の内部に軸周辺が密度が高く、錐体の表面に近づくほど密度が減少するような確率密度分布を設定する。具体的には、錐体の軸に垂直な断面が軸を中心とする 2 次元ガウス分布になるように設定する。

さらに、1 対の視線情報から 3.3 節の方法によって、1 対の視線情報のみから計算される解座標を求め、その座標を中心とした球体を作成する。この球体についても、内部に中心付近が密度が高いような確率密度分布を設定する。

以上の計算を行った上で、分布を更新する際には左右眼それぞれの円錐分布と球状分布の積分布をボクセル空間に付加する。両分布の積を取っているのは、その視線対の信頼度を考慮するためである。左右の視線が空間内で交差するときは、両円錐は球体の中心付近を通るため、積分布が大きな値を持つのにに対し、左右の視線が空間的にずれておじれの関係にあるときには、両円錐は球体の中心から外れたところを通るため、積分布の値は小さくなる。このようにすることで、左右の視線が正しく交差しない信頼度の低い視線対の影響を削減するという、視線信頼度を考慮した確率密度分布形成を行うことができる。

4.1.3 評価部

評価部は、更新部で形成された確率密度分布から、注視点座標の解を得る部分である。評価は、前処理が済んでいれば任意

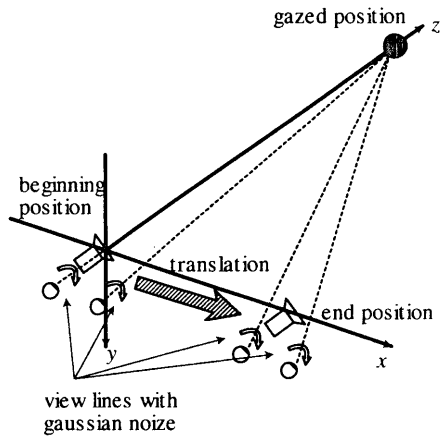


図6 シミュレーション実験環境
Fig.6 Experimental environment for simulations

のタイミングで行うことができる。解としては基本的に確率密度分布の最大値部分を採用する。

4.2 提案アルゴリズムの利点

上記のような確率的アルゴリズムを採用する利点について以下にまとめる。主に単純な最小2乗法による手法との比較を行っている。

(1) 視線の誤差モデルの考慮

人間の視線情報の誤差モデルを円錐状分布の幅や球状分布の大きさなどによって表現することで、視線の特性を踏まえた処理することができる。

(2) インパルス性のノイズに頑健である

人間の視線は、注視の際でも注視点から瞬間的に大きく外れることがある。最小2乗法による手法では、推定値がこれらの外れ値に引きずられることがあるが、本手法では少数の外れ値が混入しても推定値に影響を与えないことが期待される。

(3) 移動点注視への拡張可能性

本研究では空間中の固定点の注視しか取り扱っていないが、今後の拡張として運動している点を注視する際にも最小2乗法的な手法に比べて適用しやすいと考えられる。

5. 実験と考察

5.1 環境設定

4.章で述べた提案手法の有効性を確認する。ただし、頭部位置・方向を正確に計測できる環境のもとでの提案アルゴリズムの有効性を確認するために、頭部運動情報および視線情報を計算機シミュレーションで生成し、提案アルゴリズムを適用した。

実験環境は図6に示すように、ユーザが点を注視したまま頭部を回転することなくゆっくり20cm右方向に平行移動しながら1点を注視し続ける状況を想定している。視線情報については、実際の人間の視線運動を再現するために2.3.3節での結果をもとに、左右眼それぞれの視線の方向ベクトルに、標準偏差1.0度のガウシアンノイズの方向誤差として付加した。

5.2 結果と考察

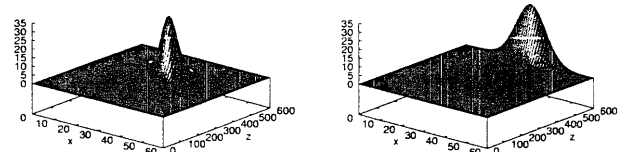
実験結果を表1にまとめる。性能比較のために、同じくベースライン拡張アプローチである式(12)の最小2乗法で求めた推定値を併記する。なお、提案アルゴリズム中のボクセルのサイズは10mmに設定したので、提案手法の推定値は10mm刻みの値を出力している。

表より、奥行き1000mmでは、開始点と終了点の視線に比較的大きな角度があるためにそれらの視線直線の方程式に関する最小2乗法が有効であり高精度に解を求まっているが、奥行

注視点真値	提案手法		最小2乗法	
	推定値	誤差	推定値	誤差
(0, 0, 1000)	(0, 0, 1010)	10	(5, 1, 951)	49
(0, 0, 2000)	(0, 0, 2010)	10	(18, 5, 1653)	348
(0, 0, 3000)	(0, 0, 3010)	10	(33, 11, 2031)	967
(0, 0, 4000)	(0, 0, 3930)	70	(47, 16, 2156)	1844
(0, 0, 5000)	(0, 0, 5000)	0	(58, 21, 2128)	2873

(単位は mm)

表1 シミュレーション実験結果
Table 1 Experimental results



(a) 注視点座標 (0, 0, 2000) (b) 注視点座標 (0, 0, 4000)

図7 確率密度分布の y = 0 断面

Fig.7 y = 0 section of probability density distribution

きが大きくなるにつれてその角度が小さくなっていき視線直線群が平行に近づいていくために最小2乗法による解は非常に不安定になり、推定結果も奥行き増加とともに真値から大きくずれてしまっている。

それに対して提案手法を用いた推定値は、奥行きが大きくなっても誤差は非常に小さく抑えられていることが分かる。真値を中心とするガウシアンノイズを加えているので、多数の視線による分布を重ね合わせることでノイズが平均化されているという側面があるが、最小2乗法と比較して、安定して高精度の解が得られていることより提案アルゴリズムの有効性が確認できたと言える。

次に、提案アルゴリズムで生成された3次元確率密度分布の y = 0 における断面を図7に示す。先にも述べたようにシミュレーションにおける視線誤差が多数の視線情報を用いていることによってキャンセルされる効果のため、いずれの試行においても図7(a)同様の確率密度分布が得られ、いずれの分布もピークが明確でかつ非常に高精度の推定を行っている。また、注視点(0, 0, 4000)の試行のみ推定誤差が大きいの、提案アルゴリズムの前処理において計算された初期値が真値大きくずれており、真値が解候補領域に含まれず、解候補領域内で推定値を探索したためであることが図7(b)より分かる。

6. おわりに

本研究では、人間が注視している点の空間的位置を推定するシステムを目指し、複数視点からの視線情報を利用することによって注視点の3次元座標を高精度に求める方法について提案した。不可避な誤差を含む視線情報を取り扱うため、その誤差モデルを踏まえた推定が可能な確率的アルゴリズムを提案、実装し、計算機によるシミュレーション環境での実験に適用し、その効果を確認した。

今後、実環境中で動作するシステムの構築を行うが、その際に問題となる点を以下に挙げる。

(1) 眼球計測の誤差軽減措置

システム構築上最も問題となるのは実環境におけるユーザ視線計測の誤差である。2.3節でその誤差要因の軽減を図っているが、依然誤差は大きく、注視点位置推定の精度向上の妨げと

なっている。今後は、装置のキャリブレーションも含めて、さらなる誤差軽減措置が必要である。

(2) 頭部位置・方向の計測手法の実装

実環境使用できるシステムを構築し、提案アルゴリズムの有効性を確認するためにはユーザ頭部の位置・方向を計測する手法が必要となる。視野カメラの時系列画像からカメラ外部パラメータを導出できる因子分解法 [7] [8] による手法や、環境側に設置されたステレオカメラによるユーザ頭部トラッキングを行う手法等を検討している。

(3) アルゴリズムの改善及び高速化

現在の実装では、ボクセル空間全体を繰り返し走査を行っている。分布更新の際の走査領域を限定するなどといった高速化が課題である。また、5.2節の図7(b)で議論した解候補領域設定の失敗によるエラーを回避するため、解候補領域を動的に変化させるといった改良が必要となる。

(4) 確率密度分布形状による推定値信頼度評価法

確率密度分布による推定を行っているため、その分布の形状から推定値の信頼度を評価できる手法についても検討し、実装したいと考えている。

(5) アプリケーションにおけるユーザインタフェース

ユーザの興味対象認識やロボットナビゲーションなど注視点の3次元位置計測システムを用いたアプリケーションを構築する際の、注視期間の判定や、周辺デバイスについて検討する必要がある。

さらに今後の課題としては、以下の項目に関しても検討する必要がある。

(1) 注視における個人差学習

注視実験の経験から、注視の際の視線方向の正確さ及びずれ方の特徴に個人差があるだろうという仮説を立て、その個人差をキャリブレーション時に学習し後の注視点推定に反映することで精度向上を図ることを検討している。

(2) 移動注視点の位置推定

4.2節で述べたように、環境中の固定点だけでなく運動している点でも注視によってその位置を推定しトラッキングできるようアルゴリズムを拡張することを検討している。

謝 辞

本研究は、科学技術振興事業団 (JST) の戦略的基礎研究推進事業 (CREST) 「日常生活を拡張する着用指向情報パートナーの開発」による。

Appendix.

カメラ座標系における眼球位置の推定法

座標系として視野カメラ座標系を用いる。図8に模式図を示す。カメラからの奥行きが既知の値 z_t である注視用ターゲットと視野カメラの中心を結ぶ直線とキャリブレーション面 $z = z_p$ との交点 $r_c = (x_r, y_r, z_p)$ とする。 r_c は視野カメラの撮影画像から得ることができるので、ここから

$$t_c = (x_t, y_t, z_t) = \left(\frac{z_t}{z_p} x_r, \frac{z_t}{z_p} y_r, z_t \right) \quad (13)$$

となる。一方、この注視用ターゲットを注視した際の視線直線とキャリブレーション面との交点 $p_c = (x_p, y_p, z_p)$ も EMR-8 の眼球観測結果より得ることができる。視線直線は p_c と t_c を通るので、 (x, z) , (y, z) における視線直線の方程式はそれぞれ、

$$z - z_p = \frac{z_t - z_p}{\frac{z_t}{z_p} x_r - x_p} (x - x_p) \quad (14)$$

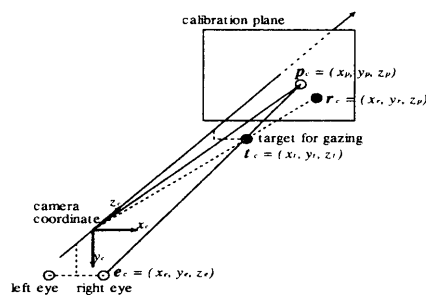


図8 眼球位置の推定法

Fig.8 Estimation of an eyeball's position

$$z - z_p = \frac{z_t - z_p}{\frac{z_t}{z_p} y_r - y_p} (y - y_p) \quad (15)$$

となり、整理して

$$z_p(z_t - z_p)x - (z_t x_r - z_p x_p)z = z_p(z_t - z_p)x_p - (z_t x_r - z_p x_p)z_p \quad (16)$$

$$z_p(z_t - z_p)y - (z_t y_r - z_p y_p)z = z_p(z_t - z_p)y_p - (z_t y_r - z_p y_p)z_p \quad (17)$$

この式を

$$Ax + Bz = C$$

$$Dy + Ez = F$$

と表し、カメラから注視点までの距離を変化させて複数作成することで、以下のような行列演算式を作成する。

$$\begin{pmatrix} A_1 & 0 & B_1 \\ \vdots & \vdots & \vdots \\ A_n & 0 & B_n \\ 0 & D_1 & E_1 \\ \vdots & \vdots & \vdots \\ 0 & D_n & E_n \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} C_1 \\ \vdots \\ C_n \\ F_1 \\ \vdots \\ F_n \end{pmatrix} \quad (18)$$

この式を解くと、 $e_c = (x_e, y_e, z_e)$ が最小2乗法的に求まる。これを左右眼についてそれぞれ行うことで、眼球位置を求める。

文 献

- [1] 吉川厚, 大野健彦: “視線を読む —ユーザにやさしい視線測定環境—”, NTT R&D Vol.48 No.4, 1999.
- [2] 武川直樹, 大野健彦, 深山篤, 澤木美奈子, 萩田紀博: “ビジョン技術におけるコミュニケーション理解環境—視線を用いたコミュニケーション支援—”, NTT R&D Vol.50 No.12, 2001.
- [3] A.Sugimoto, A.Nakayama and T.Matsuyama: “Detecting a Gazing Region by Visual Direction and Stereo Cameras”, Proc. of ICPR2002, Vol.3, pp.278–282, 2002.
- [4] 満上 育久, 浮田宗伯, 河野恭之, 木戸出正継: “両眼視線情報を利用した対象の位置推定とその応用”, 第16回人工知能学会全国大会, 2002.
- [5] R.Y.Tsai: “A efficient and accurate camera calibration technique for 3D machine vision”, CVPR, pp.364–374, 1986.
- [6] Isard, M. and Blake, A: “Condensation – conditional density propagation for visual tracking”, IJCV, 28(1):5-28, 1998.
- [7] Conrad J. Poelman and Takeo Kanade: “A Paraperspective Factorization Method for Shape and Motion Recovery”, CMU-CS-93-219, 1993.
- [8] 金出武雄, コンラッド・ポールマン, 森田俊彦: “因子分解法による物体形状とカメラ運動の復元”, 電子情報通信学会 D-II, Vol. J76-D-II, pp.1497-1505, 1993.