Inside-Out カメラによる 3 次元 Scan path の計測に関する研究

後藤祐斗　　指導教授：藤吉 弘亘

1. はじめに

システムがユーザの行動意図を理解して適切な支援または補助を行うために，生活の質の向上技術 (Quality of Life Technology:QoLT) の一環として First Person Vision(FPV) の研究が行われている [1]．FPV では，頭部にカメラを装着し，一人称の視点映像を用いて行動意図を理解することを目的としている．清水らは人の視界と視線が観測可能な Inside-Out カメラを用いた 3 次元注視点の推定手法 [2] を提案しているが，頭部運動を含む場合，正確な 3 次元位置推定ができないという問題がある．そこで本研究では，自己運動推定により頭部の運動を求め，3 次元 Scan path とシーンの 3 次元構造を復元するシステムを提案する．

2. Inside-Out カメラシステム

本手法に用いる Inside-Out カメラについて述べる．Inside-Out カメラは図 1 に示すように，ゴーグルの上部に人の眼球を撮影する Eye カメラを 2 台，下部に人の視界を撮影する Scene カメラで構成されたゴーグル型の撮影装置である．寸法は W160×H80×D100 で，重量は約 200g である．Inside-Out カメラは図 1(b) に示すように，Eye カメラが IR ミラーを介して人の眼球を正面から撮影し，Scene カメラがハーフミラーを介して人の視点と同一の映像を撮影するため，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．

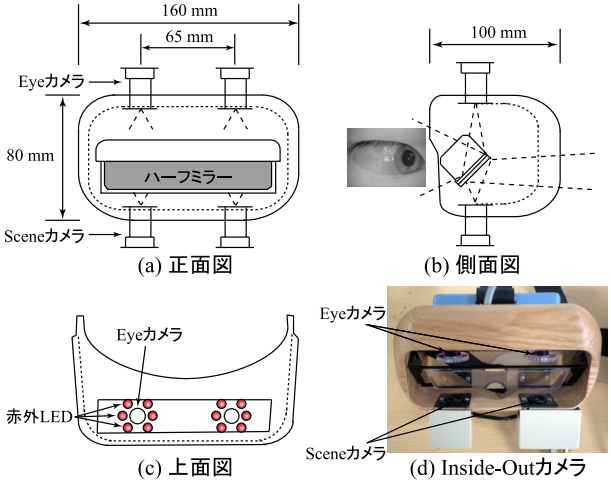

図 1：Inside-Out カメラの機構

3. 提案システム

本研究では，図 2 に示すように，動的に移動する人の 3 次元 Scan path と 3 次元のシーン構造の復元を目的とする．図 3 に，提案システムの流れを示す．Step 1 では，Scene カメラの映像から 3 次元空間を動的に移動する人の自己運動を推定する．Step 2 では，Scene カメラと Eye カメラの画像を用いて 3 次元注視点を推定する．また，2 台の Scene カメラ画像間でステレオマッチングにによりシーン構造を復元する．Step 3 では，自己運動時におけるカメラ座標から世界座標に変換することで，3 次元の特徴点群によるシーン構造と注視点の経路である 3 次元 Scan path を出力する．

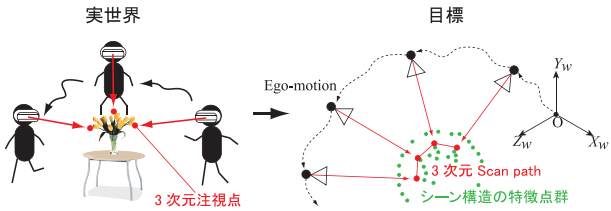

図 2：本研究の目標

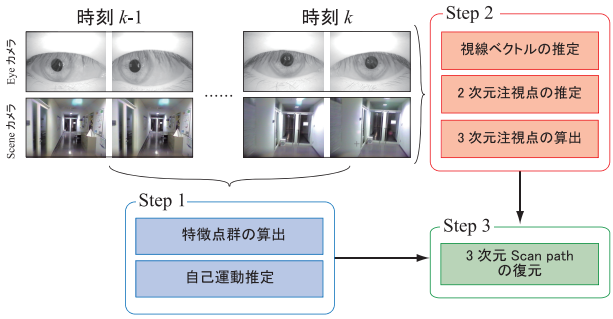

図 3：提案システムの流れ

3.1 自己運動推定

Scene カメラのシーケンス画像からの自己運動の推定には，Badino らが提案した手法 [3] を用いて KLT 法により追跡された特徴点群から，カルマンフィルタにより自己運動を推定する．カルマンフィルタは，システムモデルを逐次更新することで，各時刻における状態の最小二乗誤差を与える．

システムモデル　時刻 $k-1$ での位置ベクトルを $\boldsymbol{p}_{k-1} = (X, Y, Z)^{\mathrm{T}}$，速度ベクトルを $\boldsymbol{v}_{k-1} = (\dot{X}, \dot{Y}, \dot{Z})^{\mathrm{T}}$ と表す．時刻 $\triangle t_k$ 後のカメラ位置と速度は，前時刻のカメラ位置と回転行列 $\boldsymbol{R}_k$，平行移動ベクトル $\boldsymbol{t}_k$ を用いて式 (2) で表す．

$$\begin{aligned} \boldsymbol{p}_k &= \boldsymbol{R}_k \boldsymbol{p}_{k-1} + \boldsymbol{t}_k + \triangle t_k \boldsymbol{R}_k \boldsymbol{v}_{k-1} \\ \boldsymbol{v}_k &= \boldsymbol{R}_k \boldsymbol{v}_{k-1} \end{aligned} \tag{1}$$

ここで，位置ベクトルと速度ベクトルを合わせて状態ベクトル $\boldsymbol{x}_k = (X, Y, Z, \dot{X}, \dot{Y}, \dot{Z})^{\mathrm{T}}$ とする．状態ベクトル $\boldsymbol{x}$ は，変換行列 $\boldsymbol{A}_k$ と入力ベクトル $\boldsymbol{B}_k$ から，線形システムモデル式 (2) で表現できる．$\boldsymbol{\rho}_k$ はガウスノイズである．

$$\begin{aligned} \boldsymbol{x}_k &= \boldsymbol{A}_k \boldsymbol{x}_{k-1} + \boldsymbol{B}_k + \boldsymbol{\rho}_k \\ \boldsymbol{A}_k &= \begin{bmatrix} \boldsymbol{R}_k & \triangle t_k \boldsymbol{R}_k \\ \boldsymbol{0}_{3\times 3} & \boldsymbol{R}_k \end{bmatrix} \\ \boldsymbol{B}_k &= (\boldsymbol{t}_k^{\mathrm{T}}, 0, 0, 0)^{\mathrm{T}} \end{aligned} \tag{2}$$

計測モデル　KLT 法により取得した特徴点はその座標 (u, v) とステレオマッチングにより得られるカメラ視差 d を合わせたものを $\boldsymbol{m} = (u, v, d)^{\mathrm{T}}$ とする．状態ベクトル $\boldsymbol{x}_k$ との関係式は非線形測定式 $\boldsymbol{h}$ により式 (3) で表現できる．B はカメラのベースライン，$\boldsymbol{\nu}$ はガウスノイズである．

$$\boldsymbol{h}(\boldsymbol{x}_k) = (u, v, d)^{\mathrm{T}} = \frac{f}{Z}(X, Y, B)^{\mathrm{T}} + \boldsymbol{\nu} \tag{3}$$

式 (3) により画像座標上の特徴点ベクトル $\boldsymbol{m} = (u, v, d)^{\mathrm{T}}$ から状態ベクトル $\boldsymbol{x}$ の推定が可能である．ここで自己運動 $\boldsymbol{R}$ と $\boldsymbol{t}$ は逆関数 $\boldsymbol{g}()$ を用いて，式 (4) の最小化により特徴点群 $\boldsymbol{m}_i = (u_i, v_i, d_i)^{\mathrm{T}}$ から推定する．

$$E = \arg\min_{\{\boldsymbol{R}, \boldsymbol{t}\}} \frac{\sum_{i=1}^{n} \omega_i^2 (\boldsymbol{m}_i' - \boldsymbol{h}(\boldsymbol{R}\boldsymbol{g}(\boldsymbol{m}_i) + \boldsymbol{t}))^2}{\sum_{i=1}^{n} \omega_i^2} \tag{4}$$

3.2 三次元 Scan path の推定

Inside-Out カメラから得られる Eye カメラと Scene カメラの画像を用いて 3 次元 Scan path の復元を行う．

3.3 視線ベクトルと 2 次元注視点の算出

視線ベクトルを推定するための角膜曲率中心を，プルキニエ像を用いて算出する．プルキニエ像から角膜曲率中心 $\boldsymbol{c} = (u_c, v_c)^{\mathrm{T}}$ を算出し，その角膜曲率中心と瞳孔中心 $\boldsymbol{p} = (u_p, v_p)^{\mathrm{T}}$ を結ぶ直線を視線ベクトル $\boldsymbol{v} = (u_v, v_v)^{\mathrm{T}}$ とする．各時刻 k における視線ベクトルは $\boldsymbol{v}_k = \boldsymbol{p}_k - \boldsymbol{c}_k$ で算出できる．

求めた視線ベクトルとシーン画像上の 2 次元注視点 $\boldsymbol{g}_k = (u_g, v_g)^{\mathrm{T}}$ は，変換式 $\boldsymbol{g}_k = \boldsymbol{a}^{\mathrm{T}} \boldsymbol{v}_k + \boldsymbol{b}$ を用いて算出できる．線形方程式の傾きと切片はそれぞれ $\boldsymbol{a} = (a_u, a_v)^{\mathrm{T}}$，$\boldsymbol{b} = (b_u, b_v)^{\mathrm{T}}$ とする．

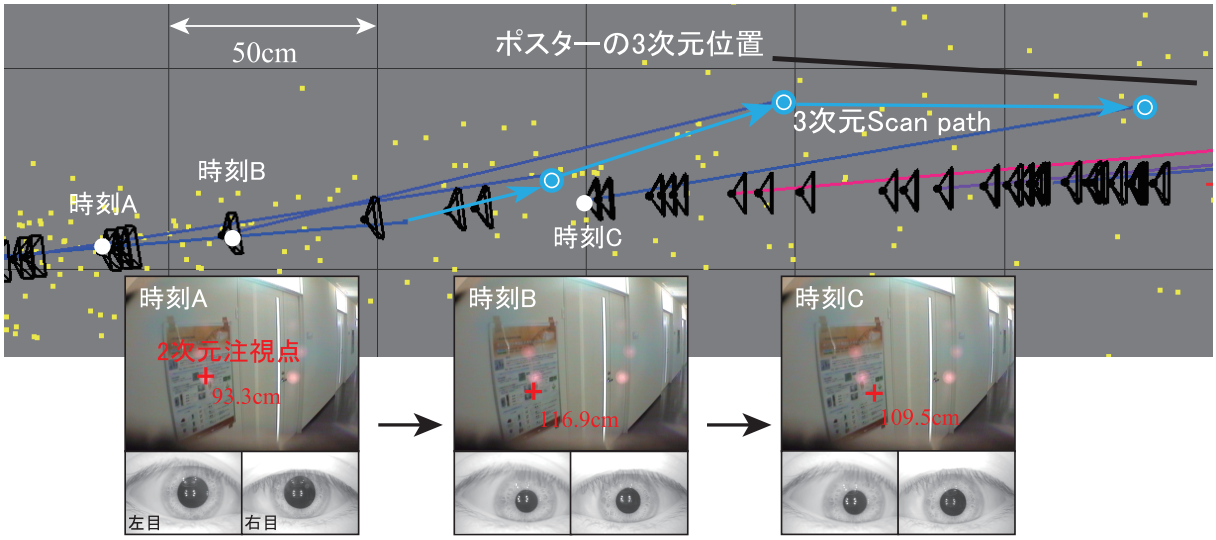

図 4：前進運動の 3 次元 Scan path の復元例

3.4 三次元注視点の推定

各 Scene カメラの 2 次元注視点からの画像上の点であるため，ステレオ視の問題として 3 次元の注視点を推定できる．しかし，算出した両眼の注視点は視線ベクトルに誤差を含むため，3 次元空間中で変わるとは限らない．そこで，最適補正を用いて，画像上の 2 次元注視点が一致するように位置の補正を行う．左の Scene 画像中の点 $\boldsymbol{m}_l$ と右の Scene 画像中の点 $\boldsymbol{m}_r$ が対応しているとき，それぞれの点を通過する視線の交差条件は，$\boldsymbol{m}_r^{\mathrm{T}}\boldsymbol{F}\boldsymbol{m}_l = 0$ を満たすように補正する (f_0 は任意の定数)．$\boldsymbol{F}$ は基礎行列を表す．最後に，カメラ座標における 3 次元注視点 $\boldsymbol{G}_k^c = (X^c, Y^c, Z^c)^{\mathrm{T}}$ は，ステレオマッチングを用いて算出する．

3.5 三次元 Scan path の復元

カメラ座標系における 3 次元注視点 $\boldsymbol{G}_k^c$ から 3 次元 Scan path の復元を行う．自己運動推定結果から時刻 k における回転行列 $\boldsymbol{R}_k$ と平行移動行列 $\boldsymbol{t}_k$ を用いて，式 (5) により世界座標系に変換する．$\boldsymbol{p}_{k-1}$ は時刻 $k-1$ におけるカメラの世界座標である．

$$\boldsymbol{G}_k^w = \boldsymbol{R}_k\boldsymbol{G}_k^c + \boldsymbol{t}_k + \boldsymbol{p}_{k-1} \tag{5}$$

各時刻 k の 3 次元注視点 $\boldsymbol{G}_k^w$ を結ぶことにより 3 次元 Scan path を復元できる．また，Scene カメラのステレオカメラにより抽出されたシーン構造を示す 3 次元特徴点の点群も 3 次元注視点と同様に復元する．

4. 実験

Inside-Out カメラにより撮影したシーケンス映像から，3 次元 Scan path の復元を行う．

4.1 実験概要

異なるシーンで撮影した 2 種類のシーケンスを用いて，提案システムの有効性を確認する．シーケンス 1 は，前進運動として廊下を前進し，壁に貼ってあるポスターを注視する動作である．シーケンス 2 は，周回運動として室内にある机の周りを歩き，机の上にあるオブジェクトを注視する動作である．

4.2 実験結果

図 4 に前進運動，図 5 に周回運動における 3 次元 Scan path と 3 次元シーンの復元結果を示す．シーン構造は黄色の点で，推定した 3 次元注視点を青い丸，3 次元 Scan path を青い矢印で示す．青の矢印は各フレームにおける 3 次元注視点を結ぶことで得られる 3 次元 Scan path を示している．前進運動では，被験者が廊下に沿って歩行しながらポスターを注視している様子が，復元した 3 次元 Scan path から確認できる．時刻 A のとき，被験者の 2 次元 Scan path はポスター上に計測されているが，3 次元 Scan path ではポスターの 3 次元位置より手前の空間を注視している．以上より，提案するシステムは，画像上の 2 次元 Scan path では取得できない注視情報を 3 次元 Scan path により取得可能であるといえる．

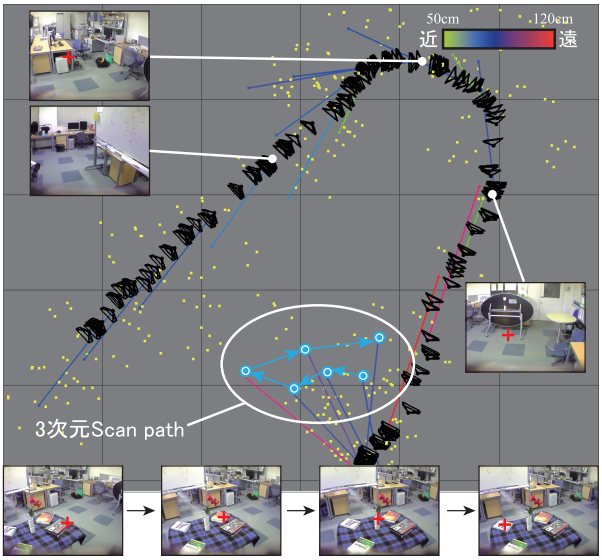

図 5：周回運動の 3 次元 Scan path の復元例

周回運動では，机の上にあるオブジェクトを注視するように 3 次元 Scan path が確認できる．3 次元のシーン構造と Scan path を取得することで，シーン特徴と注視点の 3 次元における位置関係から人の興味の有無や潜在的な意識が計測できるため，人の行動意図を理解するための情報として利用できると考えられる．

5. おわりに

頭部運動を含む人の運動を自己運動推定により計算し，3 次元 Scan path を取得する手法を提案した．2 次元 Scan path では取得できない注視情報を 3 次元 Scan path により取得可能であることを確認した．今後は，3 次元 Scan path の推定精度の評価について検討する．

参考文献

[1] T. Kanade and M. Hebert, “First-Person Vision”, Proceedings of the IEEE, Vol. 100, No. 8, pp. 2442-2453, 2012.

[2] 清水彰一，西尾和晃，木村誠，藤吉弘亘，“First Person Vision のための Inside-Out カメラの提案”，信学論 D，Vol. J94-D，No. 11，pp. 1909-1918，2011.

[3] H. Badino and T. Kanade, “A Head-Wearable Short-Baseline Stereo System for the Simultaneous Estimation of Structure and Motion, In IAPR Conference on MachineVision Applications (MVA)，2011.

研究業績

[1] Y. Goto and H. Fujiyoshi, “Recovering 3-D gaze scan path and scene structure from inside-out camera”, In *Proceedings of the 4th Augmented Human International Conference*, 2013.(発表予定)

（他 学会口頭発表 2 件，研究会技術報告 1 件）