3D Head Pose Tracking Using a Particle Filter with Nose Template Matching

Hitoshi Kubota Non-member (SUZUKI MOTOR Co., kubotah@hhq.suzuki.co.jp)
Masami Takeshi Non-member (SUZUKI MOTOR Co., takem@hhq.suzuki.co.jp)
Hideo Saito Member (Keio University, saito@ozawa.ics.keio.ac.jp)

Keywords: driver, head pose tracking, view direction tracking, particle filter, template matching, nose hole

In this paper, we propose a real-time tracking method for driver's head pose in real vehicle environment by using multiple NIR cameras, which are mounted in front of the driver's face. Fig. 1 show the system that we used in this research. Three NIR cameras are mounted on the automotive dashboard and NIR light sources are also mounted between the cameras. The cameras are synchronized by the same sync-signal.

The conventional researches for face pose estimation can be roughly categorize into two groups. First group is based on appearance of the face for face pose estimation. Another group is based on the positions of image features. In our method, we take the second strategy for avoiding that facial appearance might be affected by occlusion in the real environments. Particularly in our method, we take nose holes as the most important facial features, because the nose holes can be the most darkest area in the images captured by the cameras shooting the driver's face from the lower angle in our system [1].

In this method, we obtain the initial estimate of 3D position of the face based on the nose hole which is robustly detected from the input images. Based on the initial estimate of 3D position, we estimate the pose of the face by applying the particle filter [2], which has recently been paid a lot of attention as a robust method for tracking object. In the particle filter, we use matching scores based on the local feature points such as the edge of mouth and eyes of the facial parts that are previously registered in the system.

The 3D position and pose of the face are represented by 6 DOFs state vector that means translation and rotation. If we estimate the 6 dimensions vector using the particle filter, we need to use thousands particles that can not be handled in a real-time computation. Therefore, we need to consider the trade-off between the computational cost and the accuracy of the estimated vector. In our method, we initially estimate only the 3D position of the face by using robustly estimated nose-holes. According to the estimated 3D position, we can reduce the search space by the particle filter into 3 dimension means the pose of the face, so that we can robustly track the face pose with satisfied accuracy using only hundreds particles.

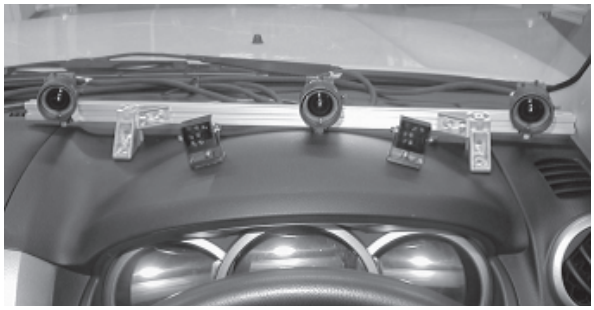


Fig. 1. Camera and lighting system



(a) (b) (c) (d)

Fig. 2. Result Images

Fig. 2 some example images of facial pose and position estimation in our system. The face pose vector is represented by the line from the nose, and 3D positions of the facial features are indicated by the cross dots in the images. The pose and position of the face can successfully be tracked not only in the case of the normal facial status as shown in (a) and (b), but also in the case of the different facial expression as shown in (c) and in the case of occluded mouth shown in (d).

References

(1) D.O. Gorodnichy: "On importance of nose for face tracking", Proc. FGR'02, pp.181-186 (2002-5)
(2) M. Isard and A. Blake: "Condensation - conditional density propagation for visual tracking", International Journal of Computer Vision, Vol.29, No.1, pp.5-28 (1998-8)

論　文

鼻孔のテンプレート探索とパーティクルフィルタによるドライバの頭部姿勢追跡

非会員　久保田　整*　　非会員　武士　正美*
正　員　斎藤　英雄**

3D Head Pose Tracking Using a Particle Filter with Nose Template Matching

Hitoshi Kubota*, Non-member, Masami Takeshi*, Non-member, Hideo Saito**, Member

In this paper, we propose a real-time tracking method for driver's head pose in real vehicle environment by using multiple NIR cameras. In order to achieve real-time performance and high-accuracy, 6-DOF of head pose is estimated by a particle filter with restricted state space. Firstly, the 3D position of nose hole is measured by template matching of stereo images. Because nose hole is the most dark area in the captured NIR images, we can robustly detect the position of the nose very precisely. From the 3D position of the nose, we can estimate the head pose as an initial estimate. Then, each hypothesis is updated by prior probability with a constraint of nose position. Thus it can reduce the number of particles with maintained accuracy. The experimental results indicate that this method is effective for head pose tracking.

キーワード：ドライバ，頭部姿勢追跡，視線追跡，パーティクルフィルタ，テンプレートマッチング，鼻孔
Keywords: driver, head pose tracking, view direction tracking, particle filter, template matching, nose hole

1. まえがき

近年，社会問題となっている交通事故による死亡者数を削減させるために，自動車の安全性能のさらなる向上が求められている。自動車の安全性能の向上策としては，従来は被害軽減を目的とした衝突安全技術が中心であったが，現在は事故自体を未然に防ぐことを目的とした予防安全技術の研究開発が盛んになっており，一部の機能は運転支援システムとして実用化されつつある。ドライバの運転モデルは『認知・判断・操作』の各過程の繰り返しで表されるが，運転支援システムは各過程におけるドライバのエラーを，情報提供，警告，制御等で支援することで事故を未然に防止する技術である(1)。

運転支援システムにおける情報提供，警告，制御のタイミングは，ドライバの認知・判断ミスの有無に関わらず画一的に決められており，それらが不要な状況においても過剰に行われてしまい，結果としてシステムの信頼性を損なってしまうという問題が指摘されている。システムの信頼性を向上させるためには，ドライバ毎に異なるタイミングに即した警報提示や操作制御を行うことが有効であり，そのためには，ドライバの状態をモニタして認知・判断ミスの有無を推定し，その情報をもとにシステムを動作させることが必要とされる(2)(3)。

ドライバは運転操作時の認知・判断のもととなる情報のほとんどを視覚から得ていることから，認知・判断ミスの有無を推定するための手段として，ドライバの頭部姿勢および視線方向の3次元的な情報を検出することが有効と考えられる。そこで本研究では，ドライバの状態として頭部姿勢の3次元的な動きに着目し，自動車運転席前面のダッシュボート付近に設置した複数のカメラにより撮影されたステレオ画像から，運転中のドライバ頭部の3次元的な姿勢を推定する手法を提案する。

人間の顔画像から頭部姿勢（顔向き）を推定する手法は二つに大きく分けられる。第一は顔全体をパターンとしてとらえ，その見え方の違いにより直接姿勢を推定する手法(4)(5)である。第二は目，鼻，口といった局所的な特徴点を検出し，その位置から姿勢を推定する手法(6)(7)(8)である。本研究では，実際の運転走行環境のもとで外乱によるノイズやオクルージョンが多く含まれる状況においても安定して推定する必要があることから，顔全体の見え方ではなく，局

* スズキ (株) 開発部
〒 224-0046 横浜市都筑区桜並木 2-1
Development Department, SUZUKI MOTOR Corp.
2–1, Sakuranamiki, Tsuzuki–ku, Yokohama 224–0046
** 慶應義塾大学 理工学部
〒 223-8522 横浜市港北区日吉 3-14-1
Department of Information and Computer Science, Keio University
3–14–1, Hiyoshi, Kohoku–ku, Yokohama 223–8522

© 2008 The Institute of Electrical Engineers of Japan.

所的特徴の検出に基づいて姿勢を推定する手法を採用する。そして，局所的な特徴としては，まず，鼻孔周辺の画像特徴に着目し，この3次元位置を画像から推定する。鼻孔周辺の特徴に着目したのは，本研究におけるカメラの設置場所は実用上ダッシュボード付近に限定されるため，顔は斜め下方から撮影され，この場合に最も安定した特徴パターンが鼻孔となるからである[9]。

本手法では，このように安定して検出できる鼻孔パターンの検出により初期推定された頭部の3次元位置を元にして，ノイズやオクルージョンの発生する環境下においても，安定して人間の頭部姿勢を追跡できる手法として近年注目されているパーティクルフィルタ[10]を用いて，頭部姿勢を推定する。この推定の際の頭部姿勢は，目尻や口の両端といった局所的特徴に着目してあらかじめ登録されたテンプレートパターンによる評価値を用いる。

一般に，頭部の3次元的な位置と姿勢は並進成分と回転成分の6次元状態ベクトルで表されるが，パーティクルフィルタで6次元の状態ベクトルを精度よく推定するためには数千個オーダーのパーティクルを扱う必要があり，計算コストが増加してしまう問題があった。パーティクルフィルタを頭部の姿勢追跡に適用する場合は，精度と計算コストとのトレードオフの関係を解消させるための工夫が必要となるが，そのための手法としてパーティクルのサンプリング過程を工夫した手法が提案されている[11][12][13]。本研究では，車室内の環境下において安定している特徴点である鼻孔をテンプレートマッチングで初期探索し，その3次元座標値を拘束条件とすることにより状態ベクトルの探索空間の範囲を限定する。すなわち，パーティクルの仮説の予測時に用いる動作モデルに鼻孔位置の固定という条件を加える。これによりパーティクルのサンプリング空間の範囲を限定してサンプリング間隔を密にすることで，少ない個数のパーティクルでも高精度かつ高速に頭部姿勢を追跡することが可能となる。

さらに，物体追跡にパーティクルフィルタを応用するうえで問題となるのが初期位置探索問題である。パーティクルフィルタは事前確率により次フレームの探索エリアを推定するため，探索開始時や見失い時の再探索に対して対応できないという問題があった。本研究ではテンプレートマッチングによる鼻孔位置の初期探索を行い，その結果をパーティクルフィルタのサンプリング空間の限定条件としているため，探索開始時や見失い時の再探索に対してロバストな追跡が可能となっている。

2. システム構成

本研究では，ドライバの頭部を撮影するためにダッシュボード上に設置した3台の近赤外CCDカメラと近赤外線照明を使用する。Fig. 1に，実験に使用したシステムを車両に搭載したときの写真を示す。1台を顔の正面に，2台を左右10度の向きとなるように設置した。このとき，カメラの仰角は約15度とした。3台のカメラは同期信号を共通とし，同じタイミングで画像を撮影できるようにした。

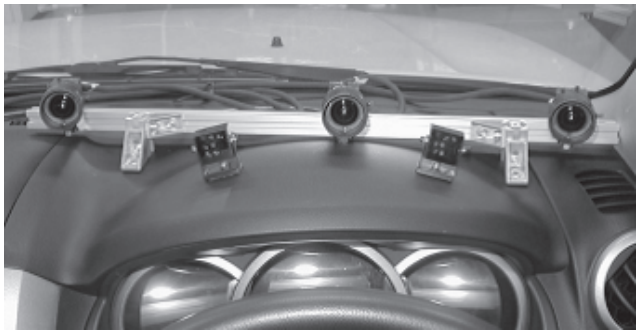


Fig. 1. Camera and lighting system.

Table 1. System specification.

Camera	NIR B/W CCD Camera ×3
Light	IR LED
PC	Pentium4 3GHz WindowsXP
I/F	Matrox MeteorII-MC PCI
Image	512×480pixel 256GrayScale

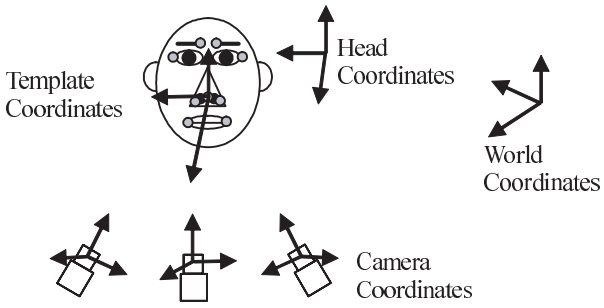


Fig. 2. Relation of coordinates.

Table 1にシステムの仕様を示す。CPUがPentium4 3.0GHz，OSがWindowsXPの汎用デスクトップコンピュータを使用した。画像入力ボードとしてMeteorII-MCを使用し，512×480pixelのグレースケール画像を3枚同時に取込む。

3台のカメラについて事前にキャリブレーションを行い，ワールド座標系と各カメラ座標系の対応関係を明らかにしておく。Fig. 2に各座標系間の関係を示す。頭部の姿勢を表すために，首の付け根を原点とする頭部座標系と鼻孔間中心を原点とするテンプレート座標系を用い，頭部座標系におけるテンプレート座標系から現在位置座標系への回転と並進（6DOFまたは4×4の変換行列）で表わすものとする。

なお，カメラが3台あるが，実際の頭部推定処理では，頭部の向きに応じて最も真正面に近い顔画像が撮影される2台のカメラが自動的に選択され，そのカメラにより撮影された2つの画像が頭部姿勢推定に利用される。つまり，カメラの台数を単純に増やすだけで，より広範囲な頭部姿勢の推定に拡張できるようになっている。

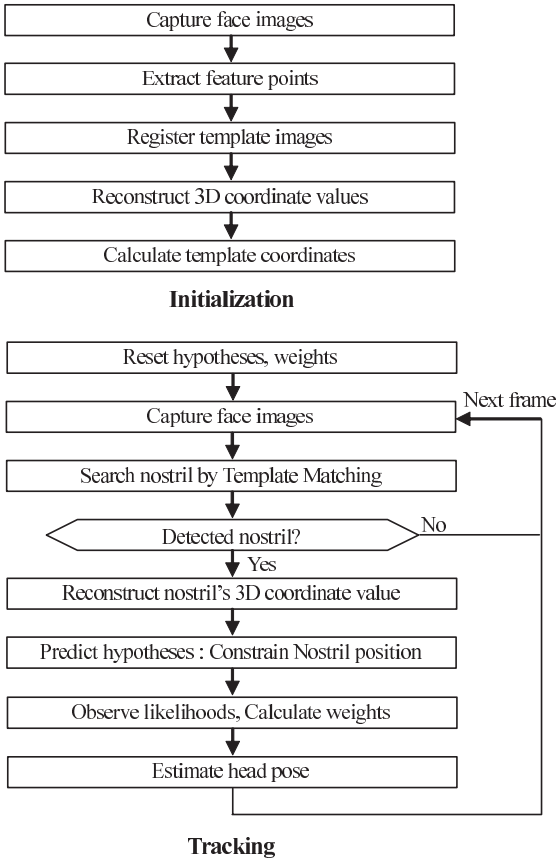


Fig. 3. Flowchart.

3. 頭部の姿勢推定手法

頭部姿勢の推定は，事前に実施しておく初期設定部と，実際に頭部姿勢を推定する追跡部から構成される。Fig. 3 に各部のフローチャートを示す。

〈3·1〉 初期設定部 初期設定部では基準となる顔のモデルを構築する。ドライバの正面顔画像から特徴点のテンプレートと 3 次元座標値の組を設定し顔のモデルとする。本研究では特徴点として，鼻孔間中心，左右の鼻孔端，口元，目尻，眉端の 9 箇所 (M=9) を用いた。また，テンプレートのサイズは 32×32pixel とした。正面顔画像 I_L,I_C,I_R を撮影した後，以下の処理を各特徴点 (m=1～M) について行う。

（1） I_C 上で特徴点の画像座標値 p_C を手動設定する。
（2） I_L,I_R のエピポーラ線上で対応する画像座標値 p_L,p_R を手動設定する。
（3） p_C,p_L,p_R を中心とするテンプレート T_L,T_C,T_R を切出す。
（4） スレテオ法により 3 次元座標値 P を求める。

さらに，各特徴点の 3 次元座標値 P_m をもとにテンプレート座標系を設定する。鼻孔間中心 P_{nose} を原点とし，顔の左右方向を X 軸，上下方向を Y 軸，前後方向を Z 軸とする。

〈3·2〉 追 跡 部 追跡部では，画像の撮影，テンプレートマッチングによる鼻孔間中心の探索に基づく位置推定およびパーティクルフィルタによる姿勢推定を繰り返し実行する。

まず，撮影された画像から鼻孔間中心をテンプレートマッチングで探索する。そして，鼻孔間中心が見つかった場合には，鼻孔間中心の 3 次元位置による拘束の元，パーティクルフィルタの各パーティクルが持つ仮説となる顔の姿勢を予測する。さらに各仮説の尤度は，各特徴点のテンプレートとの相関値により計算する。

なお，追跡の開始時におけるパーティクルフィルタの仮説は，初期設定部で求めたテンプレート座標系の姿勢，つまり顔が正面を向いているときの姿勢を基準として，回転の各成分についてガウスノイズ (σ=1.0) で拡散させることによって決定する。さらに，パーティクルフィルタの重みは総和が 1.0 となるように均等値で初期化する。

〈3·2·1〉 鼻孔間中心のテンプレート探索 カメラ画像 I_L,I_C,I_R における鼻孔間中心の画像座標値をテンプレートマッチングにより探索し，ステレオ法により 3 次元座標値 P_{nose} を求める。以下に鼻孔間中心の探索の手順を示す。

（1） I_C 上でテンプレート T_L,T_C,T_R による探索を行う。相関値は正規化相互相関 (NCC) とする。
（2） はじめに多重解像度法（階数=2）で探索し，次に探索結果を中心とする小領域を元の解像度（階数=0）で再探索する。
（3） T_L,T_C,T_R の相関値の最大値で，成功/失敗をしきい値判定する。
（4） T_L,T_C,T_R のうち相関値が最大のものを画像座標値 p_C とする。
（5） I_L,I_R のエピポーラ線上で p_C に対応する座標値 p_L,p_R を探索し，相関値が高いほうの座標値を対応点とする。
（6） p_C と p_L または p_R から，スレテオ法により鼻孔間中心の 3 次元座標値 P_{nose} を求める。

鼻孔間中心の探索が成功か失敗かをテンプレート探索時の相関値によって判定する。相関値がしきい値以上なら成功としてパーティクルフィルタによる姿勢推定を行い，しきい値未満なら失敗として，鼻孔間中心がテンプレート探索により発見できるまで次のフレームに進む。本論文の実験では，このときの鼻孔間中心探索の成否を判定するための相関値のしきい値は 0.5 とした。

なお，〈3·2·3〉節で後述するように，鼻孔間中心のテンプレート探索に失敗し，再び探索に成功したフレームでは，探索に失敗する直前のフレームにおけるパーティクルフィルタの仮説をそのまま使用する。ドライバの頭部姿勢の追跡失敗のほとんどが，左右の大きな首振り動作を行った場合であり，追跡に失敗した直前の姿勢に再び戻ってくる可能性が高いためである。

テンプレートマッチングで顔の特徴点を探索するときに，頭部の移動による画像の拡大縮小や，回転による画像の変形のために，カメラ画像とテンプレート画像との間の相関が得られなくなってしまう問題がある。本研究では，画像

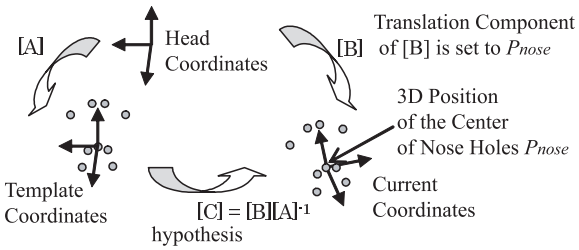


Fig. 4. Constraint hypothesis.

の拡大縮小に対しては多重解像度法による多段階ピラミッドを用いることで対応する。また，画像の変形に対しては異なる角度のカメラで撮った複数の画像のテンプレートを用いることにより対応する。運転中のドライバの姿勢変化は，初期位置を中心とした一定の範囲に限定されることから，上記の方法で対応可能である。

〈3･2･2〉 鼻孔間中心座標値の拘束方法 テンプレートマッチングによって探索できた鼻孔間中心の 3 次元座標値を固定とする拘束条件を，パーティクルフィルタの仮説予測時の動作モデルに加えることにより，サンプリング空間の探索範囲を限定することができる。

Fig. 4 に状態ベクトルの拘束方法を示す。状態ベクトルの仮説 [C] は，現在位置の座標系への変換行列 [B] とテンプレート座標系への変換行列 [A] の逆行列との積で表されることから，[B] の並進成分を鼻孔間中心の座標値 P_{nose} に固定し，回転成分のみを更新することにより，鼻孔間中心を固定した仮説が得られる。

〈3･2･3〉 パーティクルフィルタによる状態推定 パーティクルフィルタはベイズ推定の考え方に基づいた，高次元の非線形･非ガウス型の状態空間に適応可能な時系列フィルタである[10]。ベイズ推定は時刻 t における状態ベクトルを，時刻 t-1 から予測した事前確率と時刻 t における観測結果による事後確率から計算した確率密度分布により推定する手法であるが，状態ベクトルのすべての状態を計算することは現実には困難である。パーティクルフィルタは離散的な有限個のサンプル（パーティクル）のランダムサンプリングと重みから確率密度分布を近似計算することで現実的な利用を可能とした。

ある時刻 t における頭部姿勢を 6 次元の状態ベクトル X_t で表す。またパーティクルフィルタのサンプル数は N とする。はじめに，状態ベクトル X_t の一時刻前における仮説 $S_{t-1}^{(i)}$ から $S'^{(i)}_{t-1}$ を選択する。本研究では 1～N のすべての仮説をそのまま選択するものとし，重み $\pi_{t-1}^{(i)}$ による分別は行わない。

次に，動作モデルに基づいて $S'^{(i)}_{t-1}$ から $S_t^{(i)}$ を生成する。動作モデルとして (1) 式で表されるフレーム間における等角速度運動を仮定する。

$$S_t = S'_{t-1} + v_{t-1} + w \quad \cdots\cdots(1)$$

ここで，v_{t-1} は時刻 t-1 における回転成分と時刻 t-2 における回転成分との差分を示す。w は白色ノイズである。

Table 2. Selection of templates.

FacePose	Image I_L	Image I_C	Image I_R
Left	T_C	T_R	T_R
Center	T_L	T_C	T_R
Right	T_L	T_L	T_C

このとき，前フレームで鼻孔間中心のテンプレート探索に失敗した場合は，仮説の予測が行われないことから S'_{t-1} は更新されない。追跡に失敗していた状態から回復する場合は，追跡に失敗する直前の仮説をそのまま使用する。ドライバの頭部姿勢の追跡失敗のほとんどが，左右の大きな首振り動作を行った場合であり，追跡に失敗した直前の姿勢に再び戻ってくる可能性が高いためである。

時刻 t における画像から尤度 $d_t^{(i)}$ の観測を行う。尤度として各特徴点の正規化相互相関値を合計した値を用いる。更新した仮説 $S_t^{(i)}$ に従って各テンプレートの 3 次元座標値を変換する。3 次元座標値を各カメラ画像上に投影して画像座標値を求め，テンプレートと同サイズの領域を切り出し，テンプレートとの正規化相互相関値を (2) 式により計算する。(3) 式によりカメラ画像 I_L,I_C,I_R におけるすべてのテンプレートの相関値 R を合計した値（-3M～3M）を求めて尤度 $d_t^{(i)}$ とする。

$$R = \frac{\sum_{y=0}^{H}\sum_{x=0}^{W} I(x,y)T(x,y)}{\sqrt{\sum_{y=0}^{H}\sum_{x=0}^{W} I(x,y)^2 \times \sum_{y=0}^{H}\sum_{x=0}^{W} T(x,y)^2}} \quad \cdots\cdots(2)$$

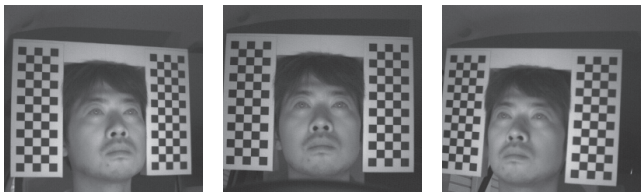
$$d_t^{(i)} = \sum_{j=1}^{3}\sum_{k=1}^{M} R_{j,k} \quad \cdots\cdots(3)$$

尤度計算の場合も，鼻孔間中心の探索時と同様に多重解像度法を用いる。さらに，顔の方向ベクトルとカメラの光軸ベクトルとが成す角度を求め，その値によって各カメラ画像 I_L,I_C,I_R で使用するテンプレート T_L,T_C,T_R を切り換える。これにより，顔の回転による見え方の違いに対応する。Table 2 にテンプレートの選択を示す。例えば，顔が左カメラの方向を向いている場合には，左画像 I_L 上での探索にはテンプレート T_C を使用し，中画像 I_C には T_R，右画像 I_R には T_R をそれぞれ使用する。このとき右画像 I_R では，いずれのテンプレートも高い相関が得られない確率が高く探索の必要性は低いと考えられるが，突発的な動きに対応するために T_R を用いた探索を含めて尤度を計算している。

次に，仮説 $S_t^{(i)}$ における尤度 $d_t^{(i)}$ の観測結果から重み $\pi_t^{(i)}$ を (4) 式のガウス関数により求める。標準偏差 σ は 1.0 とした。

$$\pi_t^{(i)} \propto \exp^{-\frac{(3M-d_t^{(i)})^2}{2\sigma^2}} \quad \cdots\cdots(4)$$

以上の処理を N 個の仮説に対して行い，仮説 $S_t^{(1)}$,…,



(cam L) (cam C) (cam R)

Fig. 5. Evaluation method.

$S_t^{(N)}$ に対する重み $\pi_t^{(1)},\cdots,\pi_t^{(N)}$ を求める。その後，すべての重みの合計が 1.0 になるように値を正規化する。最後に，仮説 $S_t^{(i)}$ と重み $\pi_t^{(i)}$ の加重平均を (5) 式により求めて状態ベクトル X_t の解とする。

$$E[X_t] \approx \sum_{i=1}^{N} \pi_t^{(i)} S_t^{(i)} \quad \cdots\cdots (5)$$

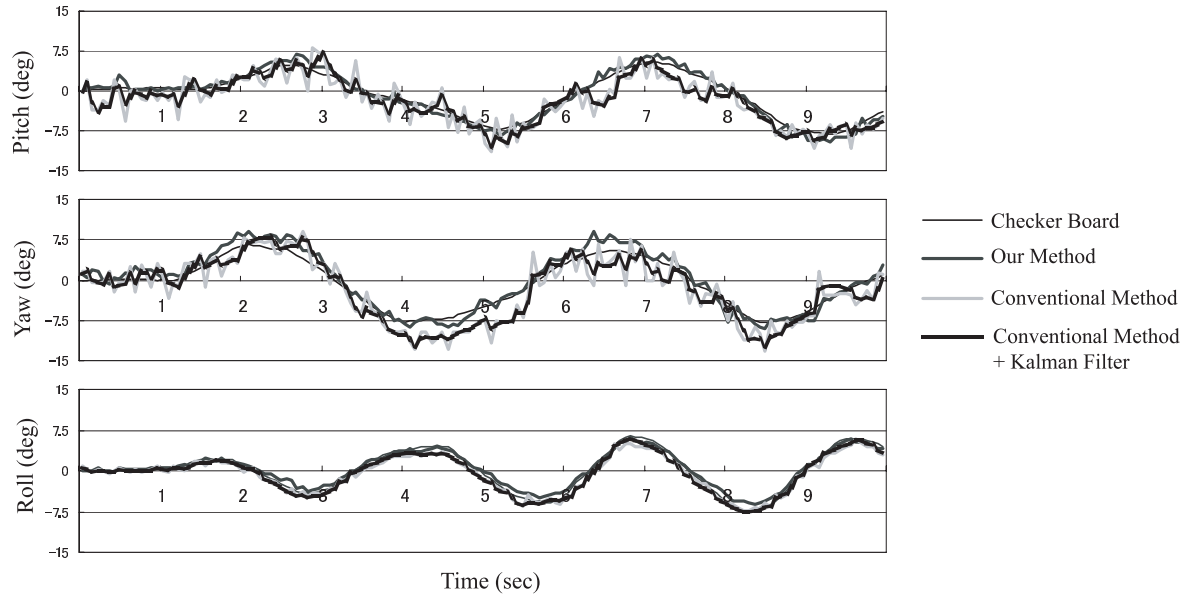


Fig. 6. Estimation Results.

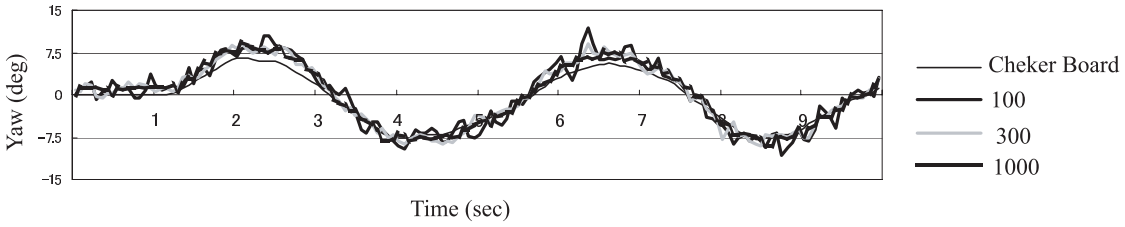


Fig. 7. Comparision of Particle Number.



(a) (b) (c) (d) (e) (f)

Fig. 8. Result Images.



(a) (b) (c) (d) (e) (f)

Fig. 9. Recovery from tracking failure.

4. 検証実験

〈4・1〉 姿勢の推定　構築したシステムにより頭部の姿勢推定実験を行った。この精度を検証するための基準値を求めるために，Fig. 5 に示すように頭部にチェッカーボードを装着して姿勢を計測した。チェッカーパターンの交点を抽出し，Zhang の手法 [14] に基づいた非線形最適化手法により推定した姿勢を基準値とした。

また，提案手法における鼻孔テンプレート探索とパーティクルフィルタによる頭部姿勢推定の有効性をを確認するために，これらを利用しない従来手法との比較を行った。ここでは，文献 [6] [15] で行われているように，頭部を撮影したステレオ画像内で複数個の特徴点をテンプレートマッチングにより探索し，ステレオ法により 3 次元座標値を求めて点群－点群間の変換行列として頭部姿勢を推定する手法を，提案手法の特徴である鼻孔テンプレート探索やパーティクルフィルタを使わない従来手法として位置付けて比較を行った。さらに，この従来手法の結果をカルマンフィルタでフィルタリングした結果も合わせて求めた。

上下 (Pitch)，左右 (Yaw)，前方軸回転 (Roll) の各動作を数秒周期で繰り返したときの推定結果を Fig. 6 に示す。このときのパーティクルの個数は 300 個とした。従来手法では，全体的に微小なノイズが見られるとともに，大きく値がずれている箇所がいくつか見られる。さらに，カルマンフィルタの結果では，微小なノイズは除去できているが，大きく値がずれる場合には対応できていない。それに対して，提案手法では微小なノイズ，大きな値のずれともに減少している。従来手法で見られる大きなずれのほとんどが周期のピーク付近で発生している。これは，ピーク付近で特徴点の画像の変形が最も大きくなることにより特徴点の座標値の計測誤差が大きくなってしまうことによるものと考えられる。提案手法では，周期のピーク付近においても基準値とほぼ一致した結果が得られており，安定した推定が可能となっている。

Fig. 7 にパーティクルの個数の違いを比較した結果を示す。パーティクルの個数を 100 個，300 個，1000 個とした場合の左右 (Yaw) 方向の姿勢変化の推定結果を比較した。300 個と 1000 個では大きな違いは見られず，どちらも基準値とほぼ一致している。100 個の場合は，大きなずれは無いが微小なノイズが一部見られる。このことから，本システムでは，従来千個オーダーの個数が必要だったパーティクルの数を数百個オーダーに減らしても安定した推定が実現できている。

Fig. 8 に姿勢推定時の画像例を示す。推定結果として顔の方向ベクトルを鼻孔間中心からの直線で描画し，各特徴点の 3 次元座標値を画像上に投影した結果を十字点で描画した。(a)～(c) の通常の姿勢変化に加えて，(d) の口を大きく開けた場合や，(e)(f) の口や眉を遮蔽した場合も推定できており，オクルージョンに対してロバストな推定が可能となっている。

〈4・2〉 再探索問題　物体追跡にパーティクルフィルタを応用するうえで問題となるのが再探索問題である。パーティクルフィルタは事前確率により次フレームの探索エリアを推定するため，探索開始時や見失い時の再探索に対してリサンプリングやスイッチングといった特別な処理を追加する必要があった。本研究ではテンプレートマッチングとステレオ法により鼻孔の 3 次元座標値を求め，その結果をパーティクルフィルタのサンプリング空間の限定条件としているため，探索開始時や見失い時の再探索に対して特別な処理を必要とせず，ロバストな追跡が可能となっている。

Fig. 9 に見逃し時の再探索の画像例を示す。顔の向きを右方向に変化させた後に再度正面に戻した場合に，(a),(b) の追跡している状態から (c) で追跡失敗となり，(d),(e) の後に (f) で再び追跡を回復できている。

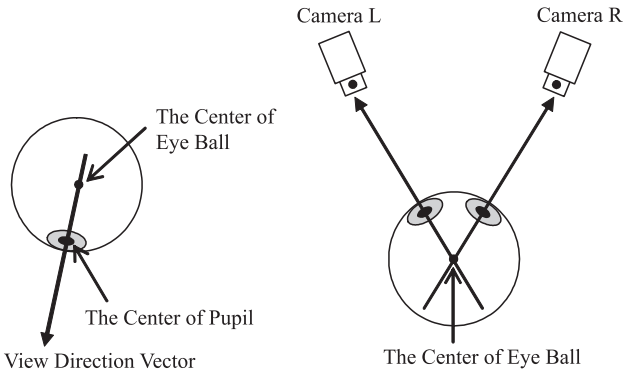
5. 視線の推定

頭部姿勢の推定結果を利用して視線方向を推定する手法の検討を行った。本研究では，照明として近赤外光を利用していることから，近赤外光を照射した場合に周囲より暗く写り抽出が容易な瞳孔中心を計測対象とした。瞳孔中心の 3 次元座標値の計測結果と頭部姿勢の推定結果から得られた眼球中心の 3 次元座標値により視線ベクトルの推定を行った [6] [16]。

Fig. 10 に視線推定のための眼球形状モデルと眼球中心位置の推定方法を示す。瞳孔中心と眼球中心を結ぶベクトルを視線ベクトルと仮定する。眼球中心は画像から直接的に検出することができないため，キャリブレーションによりあらかじめ推定しておく。頭部モデルを登録する際の画像撮影時に，カメラ L および R のレンズ中心を見ている状態での画像を合わせて撮影する。それぞれの画像における瞳孔中心の画像座標値を手動で決定した後，ステレオ計測により 3 次元座標値を計算する。さらに，カメラキャリブレーションで得られているレンズ中心の 3 次元座標値と瞳孔中心の 3 次元座標値から視線ベクトルを求め，それぞれのベクトルの交点を眼球中心の 3 次元座標値とする。

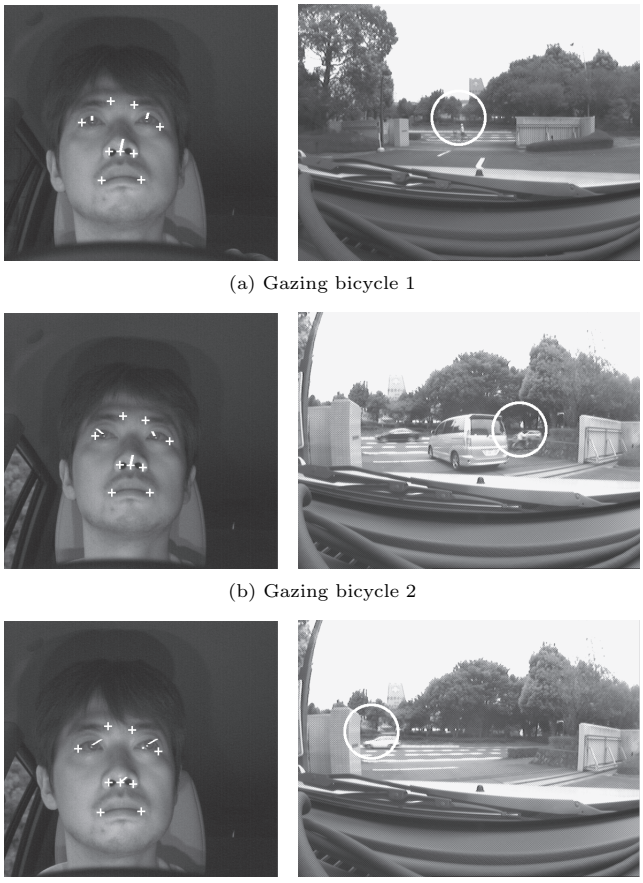
頭部姿勢の推定結果からカメラ画像上における目の領域を切り出し，二値化，ラベリング，重心計算により瞳孔中心の画像座標値を求める。さらに複数のカメラによるステレオ計測により 3 次元座標値を決定する。眼球中心の 3 次元座標値は頭部姿勢の推定結果を用いて初期位置を座標変換することにより求める。

Fig. 11 に視線の推定結果例を示す。車両前方を撮影するカメラを設置してカメラキャリブレーションを行い，姿勢推定のための 3 台のカメラとの座標系を合わせた。推定された視線ベクトルの方向の一定距離（10m）の 3 次元座標値をカメラ画像上に投影して白い円を描画した。走行している自転車を注視した場合 (a)(b)，自動車を注視した場合 (c) それぞれにおいて，視線ベクトルの方向に描いた円内に自転車 (a)(b)，自動車 (c) が存在しており，実際に注視したものを推定できている。この結果から，本手法により



(a) Eyeball Model (b) Calibration

Fig. 10. Eyeball Model and Calibration.



(a) Gazing bicycle 1

(b) Gazing bicycle 2

(c) Gazing car

Fig. 11. Estimation Results.

頭部を自由に動かしているドライバの視線方向を推定可能であることが確認できた。

6. ま と め

本研究では，運転中のドライバの顔画像から，パーティクルフィルタを用いて頭部姿勢を高速，高精度に推定するための手法を提案した。車室内において安定した特徴点である鼻孔間中心をテンプレートマッチングで探索した結果をパーティクルフィルタの仮説の拘束条件とすることにより，パーティクルのサンプリング空間の範囲を限定してサンプリング間隔を密にすることができ，数百オーダーの個数のパーティクルでも高精度かつ高速に頭部姿勢を追跡することを可能とした。

さらに，瞳孔中心の計測結果と頭部姿勢の推定結果から推定した眼球中心により，頭部を自由に動かしている状況においてもドライバの視線方向を推定可能であることが確認できた。

今後は，全プロセスの自動化をめざして，初期化時のテンプレートの登録プロセスを自動で行うための機能を追加する予定である。

(平成 19 年 10 月 10 日受付，平成 20 年 4 月 17 日再受付)

文 献

(1) M. Nagai: “Prospects of Active Safty Technology”, *SAEJ*, Vol.57, No.12, pp.4-8 (2003-12) (in Japanese)
永井正夫：「アクティブセーフティ技術の展望」，自動車技術，Vol.57, No.12, pp.4-8 (2003-12)

(2) S. Kato and S. Tsugawa: “Cooperative Driving and Driving Assistance”, *SAEJ*, Vol.58, No.5, pp.74-79 (2004-5) (in Japanese)
加藤 晋・津川定之：「ドライバアダプティブ運転支援システム」，自動車技術，Vol.58, No.5, pp.74-79 (2004-5)

(3) Y. Umemura: “Driver Behavior Active Safety”, R&D Review of Toyota CRDL, Vol.39, No.2, pp.1-8 (2004-6)

(4) J. Xiao, S. Baker, I. Matthews, and T. Kanade: “Real-time combined 2D+3D Active Appearance Models”, Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR2004), Vol.2, pp.535-542 (2004-6)

(5) K.S. Huang and M.M. Trivedi: “Robust Real-Time Detection, Tracking, and Pose Estimation of Faces in Video Streams”, Proc. of the 17th International Conference on Pattern Recognition (ICPR2004), Vol.3, pp.965-968 (2004-8)

(6) Y. Matsumoto and A. Zelinsky: “An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement”, Proc. IEEE Fourth International Conference on Automatic Face and Gesture Recognition (FG2000), pp.499-505 (2000-3)

(7) S. Katahara and M. Aoki: “Motion Estimation of Driver’s Head from Nostrils Detection”, Technical report of IEICE. PRMU, Vol.101, No.125, pp.17-22 (2001-6) (in Japanese)
片原俊司・青木正喜：「運転者顔画像の鼻孔対検出と頭部の動きの推定」，信学技術研報，PRMU，パターン認識･メディア理解，Vol.101, No.125, pp.17-22（2001-6）

(8) K. Hayashi, M. Hashimoto, and K. Sumi: “Head Direction Estimation with Robust and Accurate Tracker to Facial Features”, *Trans. of the IEICE*, Vol.J84-D-II, No.8, pp.1762-1771 (2001-8)(in Japanese)
林健太郎，橋本 学・鷲見和彦：「頑健性と精緻性を備えた顔特徴点追跡による顔方向推定」，信学論，**J84-D-II**, 8, pp.1762-1771 (2001-8)

(9) D.O. Gorodnichy: “On importance of nose for face tracking”, Proc. 5th IEEE International Conference on Automatic Face and Gesture Recognition (FGR’02), pp.181-186 (2002-5)

(10) M. Isard and A. Blake: “Condensation - conditional density propagation for visual tracking”, International Journal of Computer Vision, Vol.29, No.1, pp.5-28 (1998-8)

(11) K. Oka, Y. Sato, Y. Nakanishi, and H. Koike: “Head Pose Estimation System Based on Particle Filtering with Adaptive Diffusion Control”, *Trans. of the IEICE*, Vol.J88-D-II, No.8, pp.1601-1613 (2005-8) (in Japanese)
岡 兼司・佐藤洋一・中西泰人・小池 英樹：「適応的拡散制御を伴うパーティクルフィルタを用いた頭部姿勢推定システム」，信学論，**J88-D-II**, 8, pp.1601-1613 (2005-8)

(12) J. MacCormick and A. Blake: “A Probabilistic Exclusion Principle for Tracking Multiple Objects”, International Journal of Computer Vision, Vol.39, No.1, pp.57-71 (2000-8)

(13) M.K. Pitt and N. Shephard: “Filtering via simulation: aux-

iliary particle filtering", Journal of the American Statistical Association, Vol.94, pp.590-599 (1999-6)
(14) Z. Zhang: "Flexible Camera Calibration by Viewing a Plane from Unknown Orientations", Proc. 7th IEEE International Conference on Computer Vision (ICCV'99), Vol.1, pp.666-673 (1999-9)
(15) B.K.H. Horn: "Closed-form solution of absolute orientation using unit quaternions", Journal of the Optical Society of America A, Vol.4, pp.629-642 (1987-4)
(16) A. Tomono and F. Kishino: "Gaze Point Detection Algorithm Based on Measuring 3D Positions of Face and Pupil", *Trans. of the IEICE*, Vol.J75-D-II, No.5, pp.861-872 (1992-5) (in Japanese)
伴野　明・岸野文郎：「顔と瞳孔の 3 次元位置計測に基づく注視点検出アルゴリズム」，信学論，**J75-D-II**, 5, pp.861-872 (1992-5)

久保田　　整



（非会員） 1965 年生。1989 年 3 月武蔵工業大学工学部電気科卒業。現在，スズキ (株) 勤務。主として コンピュータビジョン，画像処理などの研究に従事。自動車技術会，IEEE 会員。

武　士　正　美



（非会員） 1963 年生。1987 年 3 月豊橋技術科学大学エネルギー工学修士課程修了。現在，スズキ (株) 勤務。主として画像処理などの研究に従事。自動車技術会会員。

斎　藤　英　雄



（正員） 1965 年生。1987 年 3 月慶應義塾大学理工学部電気工学科卒業。1992 年 3 月同大学院理工学研究科博士課程電気工学専攻修了。1992 年同大学助手。同講師，同助教授を経て，2006 年同教授。1997 年～1999 年米国カーネギーメロン大学ロボット工学研究所訪問研究員。2000 年～2003 年科学技術振興事業団さきがけ研究 21「情報と知」領域研究員兼務。博士（工学）。主としてコンピュータビジョンに関する研究に従事。電子情報通信学会，情報処理学会，映像情報メディア学会，計測自動制御学会，日本 VR 学会，IEEE 会員。