

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

岡 兼司

1 はじめに

実世界でのユーザの作業を効率的に支援可能なインタフェースを構築するためには、ユーザの行動を安定に計測することが重要となる。そこで我々は、コンピュータビジョンによりユーザの頭部の3次元的な姿勢を実時間推定するための手法について研究を進めている。

本研究では、突発的な動作への対応と十分な推定精度の両方を実現した実時間頭部姿勢推定システムを開発する。特に、パーティクルフィルタの頭部動作モデルにおける仮説の拡散を適応的に制御することにより、この問題の解決を目指す。加えて、任意のユーザが使用可能なシステムの構築に向けて、初期化部の完全自動化を実現する。

2 頭部姿勢推定システム

2.1 初期化部

初期化部では K 個の特徴点で構成される3次元モデルを自動的に生成する。このときに特別な拘束や事前の学習データを必要としないことが特長である。

まず、左カメラからの入力画像にオムロンで開発された OKAO ビジョンライブラリなどを適用し、 K 個の特徴点を検出する。それらの特徴点に対応する点を右カメラからの入力画像中で探索した後、各特徴点の3次元位置 M_k ($k = 1 \dots K$) を計算する。このとき同時に、左右のカメラに対する画像テンプレート $T_{L,k}$, $T_{R,k}$, および、追跡部で使用する頭部姿勢の初期値も獲得する。

2.2 追跡部の概要

初期化部で計算した3次元モデルと入力画像列から、パーティクルフィルタを利用して、3次元的な頭部姿勢(位置 (x_t, y_t, z_t) と回転 $(\phi_t, \theta_t, \psi_t)$) を推定する。

パーティクルフィルタでは、状態の確率密度関数を多

数の仮説の集合により表現する。また、それぞれの仮説は個々の重みを有している。それゆえ、この仮説群は任意の形状の確率密度関数を近似することが可能である。本手法では、 N 個の仮説 $s_t^{(i)}$ ($i = 1 \dots N$) を使用し、各仮説の重みを $\pi_t^{(i)}$ とする。

最初に、直前のフレーム $t - 1$ における仮説群 $\{(s_{t-1}^{(i)}; \pi_{t-1}^{(i)})\}$ と動作モデルをもとに、新たな仮説を生成する。まず、 $\pi_{t-1}^{(i)}$ に比例する確率により基準となる仮説 s'_{t-1} を選択する。そして、 s'_{t-1} を次式の動作モデルにしたがって s_t に移動する。

$$s_t = s'_{t-1} + \tau v_{t-1} + \omega \quad (1)$$

ここで、 τ は画像フレーム間隔、 v_{t-1} は直前のフレーム $t - 1$ で推定された頭部姿勢の速度である。また、拡散 ω は2.3節で説明する手法により適応的に制御される。この処理を N 回繰り返して、 N 個の新たな仮説 $\{s_t^{(i)}\}$ を得る。

次に、 $s_t^{(i)}$ に対応する重み $\pi_t^{(i)}$ をフレーム t での入力画像にもとづいて計算する。まず、 $s_t^{(i)}$ が表す姿勢にしたがって頭部モデルを移動し、移動したモデル上の各特徴点 M_k を、関数 \mathcal{P}_h ($h \in L, R$) により、画像平面上の点 $m_{h,t,k}^{(i)}$ に投影する。その後、 $m_{h,t,k}^{(i)}$ の周辺領域と対応する画像テンプレート $T_{h,k}$ とのマッチングスコアを、正規化関数 $\mathcal{N}_h(T_{h,k}, m_{h,t,k}^{(i)})$ により計算する。以上の処理を、全特徴点、全画像平面に対して適用したときのスコアの合計 $c_t^{(i)}$ を計算し、 $\pi_t^{(i)}$ を決定する。

$$c_t^{(i)} = \sum_{k=1}^K \sum_{h \in \{L, R\}} \mathcal{N}_h(T_{h,k}, \mathcal{P}_h(s_t^{(i)}, M_k)) \quad (2)$$

$$\pi_t^{(i)} \propto e^{-\frac{(2K - c_t^{(i)})^2}{2\sigma^2}} \quad (3)$$

ここで、 σ はガウス関数の標準偏差であり、経験的に3.0に固定している。また、 $\pi_t^{(i)}$ は、合計して1になるように正規化する。

以上により得られた仮説群 $\{(s_t^{(i)}; \pi_t^{(i)})\}$ から、重みが最大となる仮説の近傍に属する仮説群の加重平均により頭部姿勢を決定する。

2.3 適応的拡散制御

本手法では、動作モデル(式(1))中の仮説の拡散 ω を適応的に制御することにより、ユーザの突発的な動きに対する追従性と注視時の精度の両方を向上することを目指す。

パーティクルフィルタでは、ユーザが突発的な動作を行った場合、追跡に失敗することが多い。これは、定義された動作モデルから大きく外れた上に、仮説が十分広範囲に拡散していないことが原因である。その一方で、あまり広範囲に仮説を拡散すると確率密度関数の近似精度が全体的に低下するために、結果として推定精度の劣化を引き起こすこととなる。

これに対して、本手法では仮説の拡散範囲を適応的に制御することにより問題の解決を図ることとする。本手法で用いる ω は平均が零ベクトルであるガウス雑音である。そして、共分散行列を、対角成分が $\beta_x^2, \beta_y^2, \beta_z^2, \beta_\phi^2, \beta_\theta^2, \beta_\psi^2$ である対角行列として定義する。ここで、 $\delta_t = (\beta_x, \beta_y, \beta_z, \beta_\phi, \beta_\theta, \beta_\psi)^T$ を「拡散制御ベクトル」とし、このベクトルの制御を通して仮説の拡散を制御するものとする。

本手法では、動作モデルによる姿勢予測の不確実性が姿勢変化の大きさに比例するものと仮定する。この仮定から、次式にしたがって、 δ_t は対応する速度成分の絶対値に関して直線的に増加するものと定義する。

$$\delta_t = \Gamma \hat{v}_{t-1} + \gamma \quad (4)$$

ここで、 \hat{v}_{t-1} は速度 v_{t-1} に対して各成分の絶対値を取った6次元ベクトルである。また、 Γ は 6×6 行列、 γ は6次元ベクトルである。このとき、 Γ と γ を実際の頭部姿勢推定結果をもとに逐次的に更新することにより、実際の頭部動作に即した拡散を実現する。

3 評価実験

本節では、適応的拡散制御の効果についての評価実験の概要と結果について述べる。

実験では、Intel Pentium4 3.0GHzとWindows XPが搭載されたPCを使用した。入力画像はデジタルカメラによるモノクロ画像(解像度 640×480)、画像テンプレートの大きさは 16×16 、パーティクルフィルタの仮説の総数は1000であった。このとき、本システムは毎秒30フレームで動作した。

入力画像列としては2種類用意した。画像列1ではユーザは比較的ゆっくり動作し、画像列2ではユーザは比較

画像	制御	x	y	z	roll	yaw	pitch
#1	なし	3.24	2.33	3.87	0.40	3.67	2.34
	あり	1.25	1.91	2.71	0.25	1.55	1.70
#2	なし	3.70	4.03	5.11	1.01	3.95	2.89
	あり	3.42	3.34	4.92	0.87	2.86	2.34

表 1: 適応的拡散制御「なし」の場合と「あり」の場合の平均二乗誤差の平方根 (x, y, z [mm] and roll, yaw, pitch [degree])



図 1: 推定結果画像の一例

的高速に動作した。いずれの画像列も20秒(600フレーム)分のデータを含む。これと同時に、磁気センサを頭部に装着して頭部姿勢を計測し、その計測値と本手法による推定値とを比較した。

表1は適応的拡散制御を適用した場合としない場合の平均二乗誤差の平方根である。この結果から、適応的拡散制御が精度の上で有効に作用していることがわかる。また、図1は適応的拡散制御を用いた場合の追跡結果の画像である。1行目は入力画像列2に含まれる画像であり、2行目は様々な状況下での追跡結果である。いずれの場合も安定した頭部姿勢推定を実現した。

4 おわりに

本研究では、複数のカメラからの入力画像をもとに実時間で頭部の3次元的な姿勢を推定するためのシステムを提案した。特に、パーティクルフィルタの頭部動作モデルにおける仮説の拡散を適応的に制御することにより、追従性と推定精度を同時に向上することに成功した。そして、本推定手法の性能の高さを評価実験を通して実証した。

謝辞

本研究の一部にはオムロン株式会社の顔検出・顔器官検出技術を利用した。