

## AAMとパーティクルフィルタを用いた視線追跡

出水 祐樹<sup>†</sup> 佐野 睦夫<sup>‡</sup> 西口 敏司<sup>‡</sup>

<sup>†</sup> 大阪工業大学大学院情報科学研究科 〒573-0196 大阪府枚方市北山 1-79-1

<sup>‡</sup> 大阪工業大学情報科学部 〒573-0196 大阪府枚方市北山 1-79-1

E-mail: <sup>†</sup> m1m08a01@info.oit.ac.jp, <sup>‡</sup> {sano, nishigu}@is.oit.ac.jp

あらまし 近年、画像処理技術を利用して人間の行動を理解する研究が盛んに行われている。人間の行動を理解することは、人間とコンピュータとの自然なインタラクションを実現するのに必要不可欠である。特に視線は、人間の意図や注意を認識するのに大きな役割を果たしている。視線を検出することができれば、様々なインターフェースに適用できる。例えば、視線情報を利用して人間同士のコミュニケーションを促すエージェントなどが挙げられる。本稿では、AAM(Active Appearance Model)とパーティクルフィルタを用いたノイズにロバストな視線追跡方法について述べる。まず始めにAAMを用いて顔特徴点の抽出を行う。顔特徴点の情報をもとに顔向きを推定を行い、どの方向を見ているのかを検出する。次にAAMで求めた目領域に対して、パーティクルフィルタを用いて視線追跡を行った結果を示す。

キーワード AAM, パーティクルフィルタ, 視線追跡

## Gaze Tracking using AAM and Particle Filter

Yuki IZUMI<sup>†</sup> Mutsuo SANO<sup>‡</sup> and Satoshi NISHIGUCHI<sup>‡</sup>

<sup>†</sup> Graduate School of Information Science and Technology, Osaka Institute of Technology  
1-79-1 Kitayama, Oosaka-shi, Oosaka, 573-0196 Japan

<sup>‡</sup> Faculty of Information Science and Technology, Osaka Institute of Technology  
1-79-1 Kitayama, Oosaka-shi, Oosaka, 573-0196 Japan

E-mail: <sup>†</sup> m1m08a01@info.oit.ac.jp, <sup>‡</sup> {sano, nishigu}@is.oit.ac.jp

**Abstract** Recently, by using image processing technology, many methods to understand a human behavior have been proposed. It is essential to understand a human behavior to realize natural interaction between human and computer. Gaze-movement plays a major role to recognize human's intention and attention. Detecting the gaze-movement can make us develop various man-machine interfaces, such as communication agents. This paper proposes gaze tracking method using AAM and Particle Filter. At first, this method extracts the face feature points using AAM. Then it estimates a face direction based on information of the face feature points and exerts gaze tracking to eyes region of the estimated direction by Particle Filter.

**Key words** AAM, Particle Filter, Gaze Tracking

### 1. はじめに

近年、人間同士のコミュニケーションを助けるエージェントや人と人とのコミュニケーションに参加するロボットなどの開発が盛んに行われている。

人間同士のコミュニケーションにおいて、視線やうなずきなどの非言語情報は、大きな役割を果たしている。特に視線は、その人の興味ある対象に誘導される傾向があり、人間の意図や注意との関係が深い。話し手がどの方向を向くかによって話の注目度が変わった

り、聞き手が話者の方向を向いているかによって興味があるかどうか判断できる。このように人間がどこを注視しているのかを判定することができれば、人間とコンピュータ間の自然なインタラクションの手助けとなる。

本稿では、AAM[1]とパーティクルフィルタを用いた視線追跡を行う方法について述べる。

AAMは、表情の変化や顔向きにより変化する特徴点の座標値と輝度値を同時に表現することのできる統計モデルである。このモデルは、顔面の細かな形状変化

まで表現できる利点がある。

パーティクルフィルタ[2]は、時系列フィルタの一種であり、パーティクル(粒子)と呼ばれるそれぞれが重みを持った点を空間に散布し、その空間内全体の確率分布を近似する手法である。動きの変動やノイズがある場合でも頑健な追跡が実現できるとして、注目されている。

まず始めに AAM を用いて顔特徴点の抽出をし、それらの情報をもとに顔向きを推定を行う。正面・15度・30度・45度のどの顔向きになっているのかを検出する。次にそれぞれの顔向きにおいて、AAM で求めた目領域画像にパーティクルフィルタを用いて視線追跡を行った結果を示す。

## 2. 顔向き推定

### 2.1 頭部姿勢推定の従来研究

従来、画像から人物の頭部姿勢推定の代表的なアプローチとして、特徴点ベースの手法とアピアランスベースの手法が知られている。特徴点ベースの手法では、目や鼻、口などの顔部品の特徴点の位置関係に基づいて、頭部姿勢の推定を行う。そのため、姿勢変動やオクルージョンなどにより特徴点を安定に抽出できない状況では、適用が困難となる。また、アピアランスベースの手法は、事前に姿勢既知の条件で撮影された大量の顔画像データを用いた学習のプロセスを要し、また、遮蔽や顔領域の切り出しの精度に敏感であるという欠点を有する。近年では、Active Appearance Model が注目されている。AAM は、形状とテクスチャの相関を考慮して顔形状を記述するモデルであり、顔面の細かな形状変化まで表現できる利点がある。

### 2.2 AAM の構築

まず1枚の学習画像中の物体の輪郭を構成する点の座標がすべて与えられた学習サンプル郡があるとする。各点の座標を並べたベクトルを形状ベクトル  $x$  とし、これを全学習画像について平均して得られる  $\bar{x}$  を mean shape(平均形状)と呼ぶ。形状ベクトル  $x$  は、主成分を列ベクトルとして並べた列直行列  $P_s$  を用いて表すと

$$x = \bar{x} + P_s b_s \quad (1)$$

となる。ここで、 $b_s$  は形状のパラメータである。

次に、すべての学習画像を共通の形状に変形すると、形状について正規化されたテクスチャベクトル  $g$  を得る。形状ベクトルと同様に、全学習画像のテクスチャ

ベクトルの平均を求めたものを mean texture(平均テクスチャ)  $\bar{g}$  とする。 $g$  についても、 $g$  からの偏差を主成分分析して得られる列直行列  $P_g$  を用いて表すと、

$$g = \bar{g} + P_g b_g \quad (2)$$

となる。ここで、 $b_g$  はテクスチャのパラメータである。

さらに、形状とテクスチャ間には相関があるので、それぞれの単位を正規化する対角行列  $W_s$  を用いて、式(1)と(2)より、

$$b = \begin{pmatrix} W_s b_s \\ b_g \end{pmatrix} = \begin{pmatrix} W_s P_s^T (x - \bar{x}) \\ P_g^T (g - \bar{g}) \end{pmatrix} \quad (3)$$

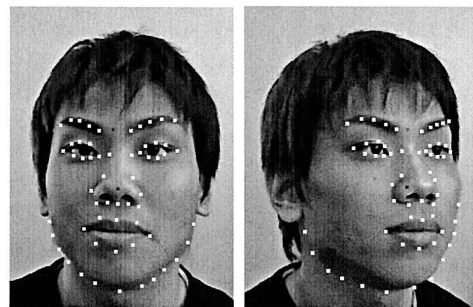
を作る。これを主成分分析すると、以下の形に近似できる。

$$b = Qc \quad (4)$$

ここで、 $Q$  は固有ベクトルであり、 $c$  は、物体の画像上の形状とテクスチャの両方を表し、アピアランスパラメータと呼ぶ。

構築した AAM を用いて、与えたアピアランスパラメータから式(4)で共通形状のテクスチャを求め、それをさらに式(1)で最終形状状態へ逆変形することで物体の画像を合成することができる。

本研究では、顔向きが左右様々に変動するため、水平方向の顔向きの変動に対応するモデルを作成する必要がある。そこで今回は学習画像を正面顔だけでなく、水平方向に顔向きが変動した画像を使用し、特徴点には、両目、両眉、鼻、口の輪郭、計58点を用いた。各学習画像の対する特徴点の配置例を図1に示す。図1(a)は正面顔に対して特徴点を配置しており、図1(b)は水平方向30度に対して特徴点を配置している。これらのモデルを使用して特徴点追跡を行い、顔向きを推定を行う。



(a)正面顔 (b)水平方向30度  
図1 各学習画像に対する特徴点の配置

### 2.3 顔向き の 推定 方法

顔向き の 推定 は AAM で 求め た 特徴 点 を 用い て 行 う . AAM で 求め た 両目 の 座標 と 鼻 の 座標 から , 目 と 鼻 の 重心 点 を 求め る . それら の 距離 に よっ て 顔 方 向 を 推定 す る .

ま ず , 正 面 を 向い て い る と き の 両目 の 重心 点 間 距離 を  $D$  と す る . 任 意 の 角 度  $\theta$  を 向い て い る 時 の 両目 の 重心 点 間 距離 を  $D'$  と す る と , 顔 の 向き  $\theta$  は , 以 下 の 式 で 表 す こ と が 可 能 .

$$\theta = \cos^{-1}\left(\frac{D'}{D}\right) \quad (5)$$

ま た , 顔 が 左 右 どち ら を 向い て い る か は , 鼻 の 重心 点 と それ ぞ れ の 目 の 重心 点 の 距離  $d_1, d_2$  を 用い て 計 測 す る .

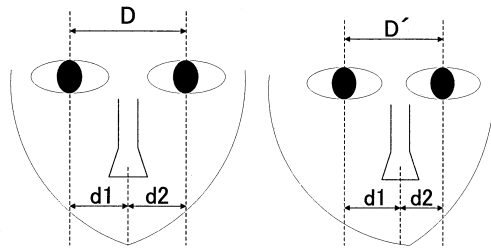


図 2 顔向き推定

な お 本 研 究 で は ,  $\theta$  が 7.5 度 未 満 の 時 は 正 面 , 7.5 度 以 上 22.5 度 未 満 の と き は 15 度 の 顔 向 き , 22.5 度 以 上 37.5 度 未 満 の と き は 30 度 の 顔 向 き , 37.5 度 以 上 は 45 度 の 傾 き と し た .

## 3. 視線追跡

### 3.1 視線認識の従来研究

視線推定の方法として、皮膚表面電極によって眼球付近の電位変化を測定する EOG 法や、サーチコイルが埋め込まれたコンタクトレンズを装着し、周辺に設置したアンテナを利用して眼球運動を測定するサーチコイル法などがある。これらの手法は特別な装置を使用し、高精度な測定が可能ではあるが、ユーザに負担がかかってしまう。

ユーザに負担をかけずにカメラ画像から視線を推定する手法は、目の画像パターン推定する手法と特徴抽出を行う手法の大きく 2 つに分けることができる。

画像パターンから視線を認識する手法としては、ニューラルネットワークを用いた手法[3]や、部分空間法を用いた手法[4]、SVD を用いた手法[5]などが提案されている。これらの手法は目の部品を正確に抽出せずに、低解像度画像を用いて視線推定を行っており、環境の変化や個人特徴に依存したモデルを用いている。応用としては環境の変化や個人特徴に非依存なモデルを構築することが望まれるが、環境や個人特徴に関する情報を含んだデータを大量に集めても、得られるモデルの視線方向に関する信頼性は高くない。

目の部分的な特徴抽出を行う手法としては、楕円形状から視線方向を推定する one-circle 法[6]がある。この手法は、楕円形状によって、楕円パラメータを用いた推定精度が低下するという欠点がある。ノイズにロバストな虹彩検出法として、LmedS を用いた方法[7]がある。目の画像から抽出された虹彩のエッジが 50% 以上の場合、正しくフィッティングできることが証明されている。しかし、瞬きの前後など上下のまぶたの間隔が狭いときは、黒目の輪郭エッジが少なく、正確な楕円フィッティングが困難となる。

そこで今回の瞳の追跡には、パーティクルフィルタを用いる。

### 3.2 処理の概要

提案する視線追跡の処理フローを図 3 に示す。

一連の処理の流れは、まず、ユーザがカメラに対して正面を向いたり、また水平方向に  $\pm 15$  度、 $\pm 30$  度、 $\pm 45$  度に顔を動かしている動画像を取得する。次に、あらかじめ顔の水平方向の変動に対応したアピラランスモデルを使用して、特徴点の追跡を行い、それらの特徴点をもとに目領域を求める。求めた目領域を拡大し、パーティクルフィルタを用いて視線の追跡を行う。

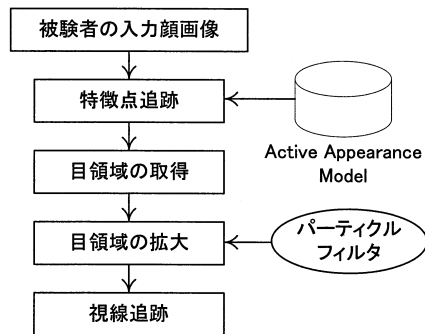


図 3 視線追跡の処理フロー

### 3.3 パーティクルフィルタの概要

パーティクルフィルタは、各時刻における状態ベクトルを逐次的に推定するための手法の一つとして注目されている[8]。ここで、一般的なパーティクルフィルタについて簡単に説明する。

パーティクルフィルタでは、時刻  $t$  における追跡対象の状態ベクトルを  $x_t$ 、観測ベクトルを  $y_t$  とし、観測値の集合を  $Y_t = \{y_1, \dots, y_t\}$  としたとき、追跡問題は  $x_t$  の事後確率分布  $p(x_t | Y_t)$  として推定できる。事後確率  $p(x_t | Y_t)$  は、ベイズの定理により次式のように事前分布と尤度の積に置き換えられる。

$$p(x_t | Y_t) = k_t p(y_t | x_t) p(x_t | Y_{t-1}) \quad (6)$$

ただし、 $k_t$  は正規化定数である。また  $p(y_t | x_t)$  は、ある状態  $x_t$  のときに、観測値  $y_t$  を得る確率（尤度）である。 $p(x_t | Y_{t-1})$  は、時刻  $t$  における  $x_t$  の事前確率であり、 $x_t$  のマルコフ性により、次式のように与えられる。

$$p(x_t | Y_{t-1}) = \int p(x_t | x_{t-1}) p(x_{t-1} | Y_{t-1}) dx_{t-1} \quad (7)$$

パーティクルフィルタのポイントは、2種類のサンプル集合  $S_{t|t-1}$ 、 $S_{t|t}$  をそれぞれ事前分布  $p(x_t | Y_{t-1})$ 、事後分布  $p(x_t | Y_t)$  に従って生成することである。これらの分布は式(1),(2)に従って推定される。パーティクルフィルタでは、これらの式を、次のような手順でサンプルに適用しながら、逐次的にサンプル集合を生成する。

1. 初期化：  $i = 1, \dots, N$  について  $s_{0|0}^{(i)} \sim p_0(x)$  を生成する。ただし、 $p_0(x)$  は、あらかじめ与えた初期分布である。 $t := 1$  として以下の手順を実行する。
2. 予測：各サンプルについて、時刻  $t$  における予測サンプル  $s_{t|t-1}^{(i)} \sim p(x_t | x_{t-1} = s_{t-1|t-1}^{(i)})$  を次の手順で生成する。
  - (a)  $i = 1, \dots, N$  について、 $l$  次元の乱数としてシステムノイズ  $v^{(i)} \sim q(v)$  を生成する。ただし、 $q(v)$  はあらかじめ設定されたシステムノイズ  $v$  の分布である。
  - (b) 時刻  $t-1$  の各サンプル  $s_{t-1|t-1}^{(i)}$  を遷移させて予測サンプルを生成する。

$$s_{t|t-1}^{(i)} = f_t(s_{t-1|t-1}^{(i)}, v_t^{(i)}) \quad (8)$$

3. 尤度設定：各サンプル  $s_{t|t-1}^{(i)}$  について、重み  $\pi_t^{(i)}$  を推定する。

$$\pi_t^{(i)} = \frac{p(y_t | x_t = s_{t|t-1}^{(i)})}{\sum_{i=1}^N p(y_t | x_t = s_{t|t-1}^{(i)})} \quad (9)$$

$p(y_t | x_t = s_{t|t-1}^{(i)})$  は、状態  $x_t$  が  $s_{t|t-1}^{(i)}$  であったときに、観測  $y_t$  を得る尤度である。

4. フィルタ：  $s_{t|t-1} = \{s_{t|t-1}^{(1)}, \dots, s_{t|t-1}^{(N)}\}$  からそれぞれ  $s_{t|t-1}^{(i)}$  を重み  $\phi_t^{(i)}$  に比例する割合で  $N$  個復元抽出し、 $s_{t|t} = \{s_{t|t}^{(1)}, \dots, s_{t|t}^{(N)}\}$  とする。 $t := t+1$  として、2の予測へ。

### 3.4 追跡に用いる尤度推定

文献[8]のパーティクルフィルタの実装例では、対象モデルの輪郭と、入力画像のエッジがどれだけマッチしているかを評価している。エッジを用いた評価では、目尻などのエッジの多い箇所の尤度が高くなり、間違った場所に収束するという問題がある。そこで我々は、黒目の輝度を尤度としたパーティクルフィルタの字移送を提案する。

まず、切り出した目領域を拡大する。拡大した画像に対して、ランダムにパーティクル(粒子)をばら撒く。各パーティクルを中心とした  $M \times N$  矩形領域の範囲内で、輝度値が設定した閾値以下の場合の存在率を尤度  $\pi_i$  とした。

$$\pi_i = \frac{\sum_{q=i-M/2}^{i+M/2} \sum_{p=i-N/2}^{i+N/2} B_{qp}}{M \times N} \quad (10)$$

ただし、 $B_{qp}$  は次式のような輝度値が設定した閾値以下であるかどうかを判定する関数である。

$$B_{qp} = \begin{cases} 1 & \text{輝度値が設定した閾値以下である場合} \\ 0 & \text{その他} \end{cases} \quad (11)$$

例えば図4の(a)の場合、パーティクルに対して周りに輝度値が低い画素があるため尤度が高くなり、(b)は周りに輝度値が低い画素が少ないため、尤度は低くなる。



(a)尤度の高い粒子 (b)尤度の低い粒子

図4 パーティクルの尤度

尤度に従いパーティクルフィルタの重みを決定する。また、サンプル個数  $s$  によって正規化した重み  $w_i$  を次式によって求める。

$$w_i = \frac{\pi_i \times S}{\sum_{i=0}^S \pi_i} \quad (12)$$

## 4. 実験

### 4.1 実験環境

実験では、IEEE1394 カラーカメラを用いて、被験者とカメラとの距離が約 1 m の環境で行った。カメラの映像は 30[fps] にて撮影した 640×480[pixel] の映像であり、1 台の PC (CPU Intel Pentium4 3.00GHz, Memory 2.0GByte, OS WindowsXP) で処理を行った。AAM の学習画像は個人用と、複数人の平均のモデルを用意した。被験者には、正面に加えて、水平方向に ±15 度・±30 度・±45 度の方向 (図 5) を向いてもらい、顔向き推定を行った。また、それぞれの向きに対して視線追跡を行った。

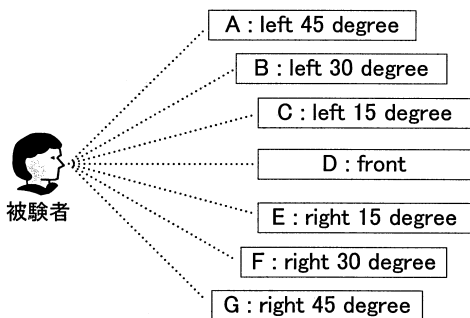


図 5 被験者の顔方向

### 4.2 顔向き推定実験

顔向き推定では、図 6 に示すように学習済みの人物や、平均モデルが未知の被験者に対して適合していた場合、高精度な顔特徴点の検出を行うことができ、顔向きの認識率も高い結果となった。

しかし、図 7 のように特徴点の検出、特に目領域に対しての特徴点検出がうまくいかなかった場合は、目の重心点間距離が変わり、顔向きの認識率は下がった。

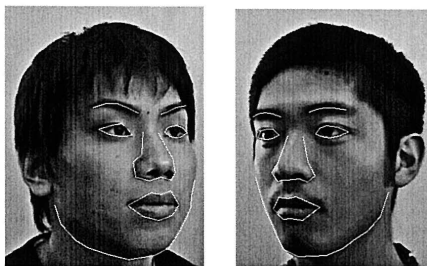


図 6 高精度な顔特徴点の検出

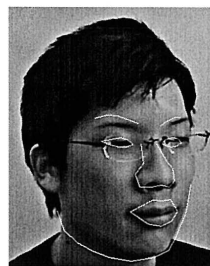


図 7 目の特徴点の誤検出

### 4.3 視線追跡実験

次に、視線追跡結果について述べる。視線追跡では、パーティクル数を 500 にし実験を行った。図 8 は、正面顔で色々な方向を向いてもらった追跡結果である。左右や上、右下など様々な方向に視線を動かし、追跡を行えている。図 9 は顔向きに変動のある場合の追跡結果である。顔向きが変動しても安定して追跡が行えている。また瞬きを含む場合でも図 10 に示すように追跡を行えている。

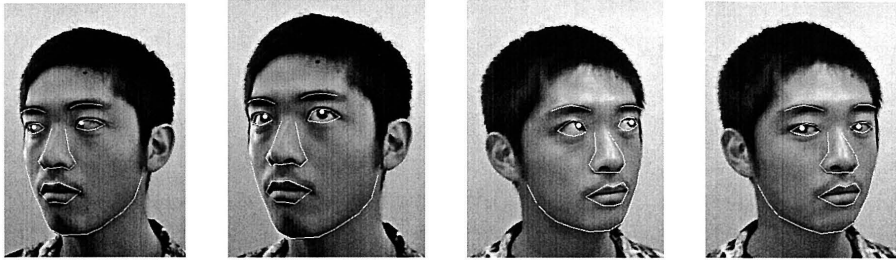
## 5. おわりに

本研究では、AAM とパーティクルフィルタを用いて、顔向きが変動した場合でも視線追跡を行う手法について提案した。実験結果より、顔向きが変わった場合でも安定して黒目の追跡を行うことができることを確認した。

今後の課題は、視線追跡だけではなく視線方向の推定を行い、様々なアプリケーションに適用させていきたいと考えている。



図 8 正面顔の視線追跡結果



(a)右 30 度の顔向き

(b)左 30 度の顔向き

図 9 顔向きに変動がある場合の視線追跡結果

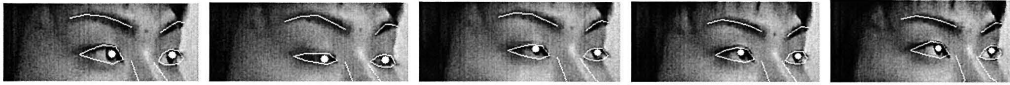


図 10 瞬きを含む追跡結果(左 30 度の顔向き)

## 文 献

- [1] T.F.Cootes, G. Edwards and C.J.Taylor.: "Active appearance models", Proc. ECCV, Vol. 2, pp. 484-498 (1998).
- [2] 樋口知之. 粒子フィルタ. 電気情報通信学会誌, Vol.8, No.12, 2005.
- [3] S. Baluja and D. Pomerleau, "Non-intrusive gaze tracking using artificial neural networks," CMU CS. Technical Report, CMU-CS-94-102, 1994.
- [4] 田中克己, 福井和広, "エージェントによるリアルタイムインタラクションを活用した. 視線入力インタフェース," 信学論 (D-II), vol.J85-D-II, no.9, pp.1464-1472, Sep. 2002.
- [5] 小野 泰弘, 岡部 孝弘, 佐藤 洋一, "目領域の切り出しの不定性を考慮した低解像度画像からの視線方向推定", 情報処理学会論文誌: コンピュータビジョンとイメージメディア, vol.47, no.SIG10 (CVIM15), pp.173-184, July 2006.
- [6] J.G. Wang, E. Sung, and R. Venkateswarlu, "Estimating the eye gaze from one eye", CVIU, vol.98, pp.83-103, 2005.
- [7] 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.
- [8] 加藤 丈和: パーティクルフィルタとその実装法, 情報処理学会研究報告, CVIM-157, pp.161-168 (2007)