

H-033

Particle Swarmに基づく人物追跡の一検討

A Study of Object Tracking Based on Particle Swarm

藤崎 陽一† 陶山 健仁†

Yoichi Fujisaki Kenji Suyama

1 はじめに

カクテルパーティー効果に代表される人間の優れた聴覚能力では、異種情報である視覚情報をも有機的に利用していると考えられている。従って生体並みの能力を有するセンシングシステムの構築を目指すならば異種センサ情報の利用が有効である。この技術はセンサフュージョンと呼ばれ盛んに研究されてきた [1]。

視聴覚センサフュージョンによる情報処理の1つとして移動話者追跡について考えよう。移動話者追跡を行なう場合、マイクロホンアレーを用いた移動音源追尾だけでなく、動画像情報を用いた話者位置の同定が有効である。本研究では、図1に示すようなシステムによる話者同定を目指し、その第一歩として単一カメラを用いた顔追跡について検討する。

顔追跡の一手法として、Particle Swarmに基づく手法 [2]がある。Particle Swarmに基づく手法は鳥や魚などの生物群を模したモデルであり、各パーティクルが追跡対象の特徴を記憶し、その特徴を用いて追跡を行なう。

本稿では、顔の追跡に適した特徴量の調査を行なう。実環境実験より、本稿で提案する特徴量が顔の追跡に適していることを示す。

2 Particle Swarmに基づく人物追跡

図2に、単一カメラで撮影したフレーム番号 t における入力画像 $I(x, y, t)$ から、パーティクル P_i の状態を算出する流れを示す。ここで、 x, y は入力画像の座標、 i はパーティクルの番号である。 $I(x, y, t)$ から特徴量抽出を行い、抽出された特徴量 f_m と P_i や群 S が持っている過去の情報に基づいてパーティクルの動作ルールを定め、 P_i や S の状態を定める。本稿では、特徴量として輝度値、画像勾配、移動量を用いた。さらに、マイクロホンアレーによる移動音源追尾の結果を模したデータを指標として加えた。

以降、文献 [2] をベースとした、パーティクルの構造や動作ルールについて述べる。

2.1 パーティクルの構造

P_i の状態を、 $P_i = (\mathbf{p}_i, \mathbf{v}_i, c_i, s_i, \mathbf{fl}_i)$ と表す。ここで、 \mathbf{p}_i は位置ベクトル、 \mathbf{v}_i は速度ベクトル、 c_i は追跡状況の評価で $0.0 \leq c_i \leq 1.0$ の値をとり、1.0に近いほど良い評価である。ま

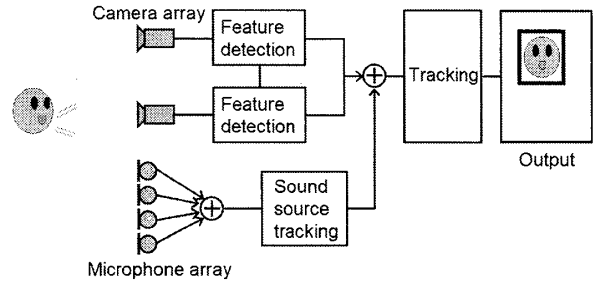


図1 目標とする話者追跡システム

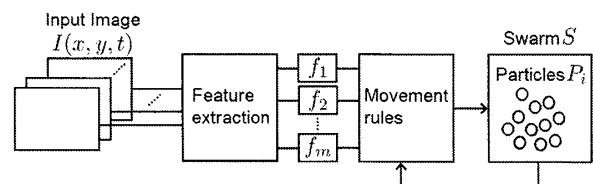


図2 Particle Swarmのモデル図

た、 s_i はパーティクルの速さ、 \mathbf{fl}_i は特徴リストで最近訪れたピクセルの輝度値の集合である。

2.2 群の情報

P_i の持つ局所的な情報だけでなく群の持つ大域的な情報を利用するため、群の重心 S_c 、速度 S_v 、予測重心 S_{pc} の3つの特徴を定める。 S_c 、 S_v は共に c_i による重み付き平均とし、 S_{pc} は S_c 、 S_v の加算とする。算出式は文献 [2] に従う。

2.3 パーティクルの動作ルール

P_i の動作は周囲の状況と群の動作の両方から決めるため、以下に述べる4つのルールで動作を定義する。

ルール1では P_i の周囲の状況から動作を決める。自身の位置を含めた $n \times n$ の N 近傍を解析し、最も追跡対象らしい方向に向かうための指標を算出する。解析対象を $p = (\mathbf{p}_k, c_k)$ ($k \in N$) とし、 \mathbf{p}_k を解析対象の位置ベクトル、 c_k を輝度値とする。指標を算出するための特徴量として、文献 [2] では色の類似度 α_k 、画像勾配 ΔI_{p_k} 、移動量 M_{p_k} の3つを用いている。3つの特徴量より、追跡対象らしさ β_k ($0.0 \leq \beta \leq 1.0$) は、

$$\beta_k = \alpha_k \cdot |\Delta I_{p_k}| \cdot (1.0 + M_{p_k}) \quad (1)$$

より算出する。 β_k を重みとし、 N 近傍の重み付き平均としてルール1の動作ベクトル \mathbf{V}_{i1} を次式より求める。

$$\mathbf{V}_{i1} = \frac{\sum_{k \in N} (\mathbf{p}_k - \mathbf{p}_i) \cdot \beta_k}{\sum_{k \in N} \beta_k} \quad (2)$$

† 東京電機大学工学部電気電子工学科
School of Engineering, Tokyo Denki University

表1 パラメータの探索範囲

Parameter	Search range	interval
w_0	$0.1 \leq w_0 \leq 0.9$	0.2
w_1	1.0	-
w_2	$0.1 \leq w_2 \leq 0.9$	0.2
w_3	$0.1 \leq w_3 \leq 0.9$	0.2
w_4	$0.1 \leq w_4 \leq 0.9$	0.2
δ	$5 \leq \delta \leq 15$	5
μ	$10 \leq \mu(\text{Conventional}) \leq 100$	30
	$0.1 \leq \mu(\text{Proposed}) \leq 1.2$	0.4
γ	15	-
feature list size	3	-

表2 実験条件

Image size	640 × 480 [px]
Frame rate	30 [fps]
Time	20 [sec] or 10 [sec]
Number of particles	10 × 20 = 200
Search range	60 × 60 [px]

さらに、追跡状況の評価 c_i とパーティクルの速さ s_i を、

$$c_i = \max_{k \in N} \{\beta_k\} \quad (3)$$

$$s_i = 1.0 - c_i \cdot \mu \quad (4)$$

で算出する。

ルール2, 3, 4の算出式と P_i の p_i , v_i の更新式は文献 [2] に従う。

3 提案法

本稿では、ルール1における β_k の算出式の追加と P_i の速さ s_i の見直しを行った。

3.1 β_k の算出式の追加

ルール1において、 β_k の算出にマイクロホンアレーによる移動音源追尾の結果を模したデータ x_s の方向に追跡対象がいるとき β_k を強調するために、(1)式の後に次式を追加した。

$$\beta_k \leftarrow \beta_k \cdot e^{-|x_s - p_{k1}|/\gamma} \quad (5)$$

ここで、 p_{k1} は p_k の x 方向成分であり、 γ はパラメータである。

3.2 s_i の見直し

c_i が小さい、すなわち追跡対象が周囲に存在しない場合に、 P_i を加速させるため、 s_i を次式のように変更する。

$$s_i = \frac{\mu}{c_i} \quad (6)$$

4 実環境実験

提案法の有効性を確認するために実環境実験を行なった。従来法として s_i の算出に (4) 式を用いた手法 [2] と比較した。比較方法は、5つの動画パターンに対して、表1に示す7500通りのパラメータで追跡を行なった。そのとき、90%以上の正解率を示した場合を成功と見なし、成功率で比較する。ここ

表3 x_s による指標を用いなかった場合の成功率

Pattern No.	Proposed [%]	Conventional [%]
1	17.76	0.00
2	49.36	2.21
3	6.33	0.01
4	2.29	0.17
5	28.71	0.00

表4 x_s による指標を用いた場合の成功率

Pattern No.	Proposed [%]	Conventional [%]
1	23.69	0.00
2	30.37	3.12
3	26.51	0.00
4	24.53	0.05
5	30.80	0.00

で正解率は、全フレームのうち、鼻頭を中心とした(顔の横幅の2倍) × (顔の縦幅の2倍)の矩形範囲に群の重心が存在する割合と定義する。また、提案法で用いたマイクロホンアレーによる移動音源追尾の結果を模したデータ x_s は、あらかじめ測定した真値に平均0, 分散64の正規乱数を加算して作成した。その他の実験条件を表2に示す。

4.1 実験結果

各パターンに対する成功率を表3, 表4に示す。表3は x_s による指標を用いなかった場合、表4は x_s による指標を用いた場合である。これより、 x_s による指標の有無に関わらず提案法の方が良い結果であることがわかる。また、提案法において x_s による指標を用いなかった場合と用いた場合を比べたとき、パターン2以外は x_s による指標を用いた場合の方が良い結果となった。以上の結果より、パラメータ選択の柔軟性や追跡性能において提案法の有効性を確認できた。

5 むすび

本研究では、Particle Swarmに基づく顔の追跡に適した特徴量の調査を行い、新たな指標の組み込みについて検討した。提案手法では、マイクロホンアレーによる移動音源追尾の結果を模したデータの動作ルールへの組み込みと、パーティクルの速さの見直しを行った。実験結果より、提案手法は従来法と比べパラメータ選択の柔軟性や追跡性能の向上を示した。今後は、実際に移動音源追尾の結果を使用した場合の検討、ならびに、カラー画像への対応を検討する予定である。

参考文献

- [1] 上野元治, “視聴覚のセンサフュージョン技術により騒音の中から話し声, 聞き分け: リアルタイム信号処理システム試作,” 電子情報通信学会誌, Vol.76, No.4, p.457, 1993.
- [2] Luis Anón-Canalís, Mario Hernández-Tejera, Elena Sánchez-Nielsen, “Particle Swarms as Video Sequence Inhabitants For Object Tracking in Computer Vision,” IEEE. ISDA'2006, vol.2, pp.604-609, 2006.