

協調的に学習する顔認識フレームワークの提案

田中 弘起[†] 岩崎 剛士[†] 高田 司郎[†]

近畿大学理工学部情報学科[†]

1. はじめに

人とロボットが協調して仕事を行うためには、ロボットは個人を同定する能力が必要であり、その一つの方法として、顔認識が挙げられる。顔認識の学習には、多量の顔画像が必要であるが、この多量の顔画像を手作業で用意するには、非常に手間がかかる。そこで、多量の顔画像を顔認識の訓練データとして自動作成することを目指す。

我々は、顔のトラッキング、顔の撮影、訓練データの作成、および顔の同定などを協調的に行うマルチエージェントシステムのフレームワークを提案している[1][2]。本稿では、顔のトラッキング、顔の撮影までを実現したので、報告する。

2. システム概要

図1 のシステムの概要を述べる。方向指定とズームができる2 台のカメラ（以下、ネットワークカメラ）を使用する。また、360 度の視野を持つ2 台の全方位カメラを使用する。実験環境は、10 メートル四方で9 名の学生の研究室とする。学生の出入りは不規則であり、他の研究室の教員や学生の出入りも想定している。

2.1. 顔トラッキングエージェント

2.1.1. 顔検出

ネットワークカメラから画像を取得して、OpenCV の物体検出モジュールを使用し顔検出を行う。OpenCV は、インテルが開発、公開しているオープンソースのコンピュータビジョン向けライブラリである。また、OpenCV には学習によって前もって獲得された分類器のカスケードが記述された xml ファイルが用意されている。これらの xml ファイルは、正面顔、右顔を検出する。しかし、左顔を検出する xml ファイルは用意されていないため、画像を反転させて右顔を検出する xml ファイルを用いている。

OpenCV の物体検出モジュールは処理が重いので、正面顔検出、右顔検出そして左顔検出を個々のコンピュータで動作させて負荷を分散している。

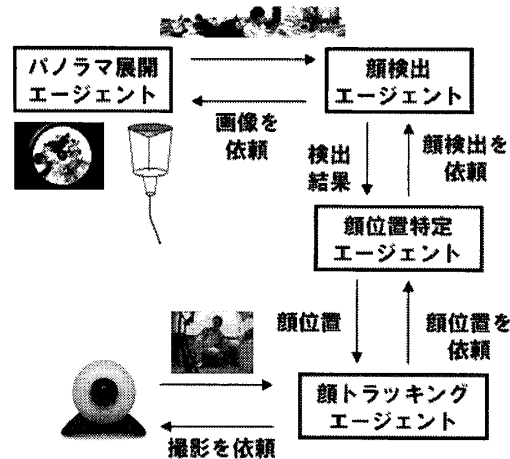


図1：システム構成

2.1.2. 顔のトラッキング

顔を検出した位置を注視するためにネットワークカメラを向ける。また、顔認識学習に最低限必要な大きさでなければ、ズームして撮影し再び顔検出を行う。

2.1.3. 顔領域切り出し

背景など顔以外の情報は、顔認識学習の妨げになるため顔領域を切り出す。これは、OpenCV の物体検出モジュールを用いて、検出した顔領域を切り出す。

顔の向きに応じて顔認識学習を行うために、切り出した顔領域を顔の向きで分類し保存する。正面顔検出、右顔検出、そして左顔検出を個々のコンピュータで動作させているため、どのコンピュータで顔を検出したか分れば顔の向きが分かる。

2.1.4. 顔の位置を依頼

一定時間、顔を検出しない場合、顔位置特定エージェントに顔の位置を依頼する。そして、依頼して獲得した顔の位置を3次元情報とネットワークカメラの位置を用いて、ベクトルの内積を計算しネットワークカメラの向く方向を求める。またノルムを計算して、獲得した顔の位置に近いネットワークカメラを求める。

2.2. パノラマ展開エージェント

全方位カメラから取得した全方位画像を全方位ライブラリである ODSSDK を使って、パノラマ展開する。

Learning Facial Recognition Framework
by Cooperative Agents.

Hiroki TANAKA[†], Tsuyoshi IWASAKI[†] and Shiro TAKATA[†]

[†]Department of Informatics, School of Science and Engineering, Kinki University, Kowakae 3-4-1, Higashi Osaka 577-8502, JAPAN

2.3. 顔検出エージェント

パノラマ展開画は、鮮明でないため OpenCV の顔検出には適していない。そこで、以下の方法を用いて顔を検出する。

・ 肌色領域検出

YCC 表色系、HSV 表色系を個々に用いて肌色領域を検出した場合、顔以外の領域も検出される。そこで、YCC 表色系、HSV 表色系の両方で検出した領域を肌色領域とすることで、顔以外の領域が検出されることを減らしている。そして、肌色領域を白色、その他を黒色として 2 値画像を作っている。

・ ラベリング

膨張によって小さな孔や溝が除かれて、収縮により孤立点や突起を除く。そして NAIST のラベリングクラスを用いて、2 値画像の連続領域を抽出しラベルを付与して、各領域を顔候補領域とする。

2.4. 顔位置特定エージェント

顔領域候補の 3 次元情報を 2 台の全方位カメラの相対位置から求める。まず、顔領域候補の 2 次元情報から水平軸となす角度を求めて、垂直軸となす角度を多項回帰式を用いて求める。そして、水平軸となす角度を用いて 1 次関数の方程式を求める。これを 2 台の全方位カメラで行って、2 直線の交点を求めるためにガウス・ジョルダンの消去法を用いる。

交点と全方位カメラの位置を用いて全方位カメラからの距離が求められる。また、垂直軸となす角度と全方位カメラからの距離を用いて高さを求める。全方位カメラからの距離と高さを用いて顔領域候補の 3 次元情報を作成する。

3. 実験

実際に動かしてみて、どの程度の枚数の顔画像が撮影できるかを実験した(図 2)。この実験は、常にネットワークカメラが撮影できる範囲に、人がいるという状況で行った。

4. 実験結果

実験の結果、1 時間で 2000 枚程度の顔画像が撮影でき、その中で顔以外の画像は 40 ~ 60 枚程度だった。

実験によって顔トラッキングエージェントの顔検出は、精度が高いことが分かった。

しかし、顔検出エージェントの顔検出は色情報を用いているため精度が低い。そのため、顔位置特定エージェントが顔トラッキングエージェントに顔でない位置を間違えて伝えた場合、ネットワークカメラが顔でない位置を向いたまま、誤検出した位置を撮影し続けるという問題が生じている。



図2：実験の様子

5. 今後の課題

2 台の全方位カメラが複数の顔領域候補を検出した場合、顔位置を特定するために同一人物の顔領域候補ペアを選ばなければならない。しかし肌色領域だけでは、同一人物の顔領域候補ペアを選べない。そこで、同一人物の顔領域候補ペアを選ぶために、検討している方法を以下に示す。

- ・ 肌色領域の下に衣服があると仮定して、衣服の色で同一人物の顔領域候補であると判断する。
- ・ 顔の大きさと全方位カメラからの距離を求めて、同一人物の顔領域候補であると判断する。

また、訓練データの作成、顔の同定なども今後の課題である。

6. おわりに

ネットワークカメラと全方位カメラを協調的に動作させて、自動的に高精度の顔画像を撮影することができた。また、今後の課題について述べた。

参考文献

- [1] 高田司郎, 岩崎剛士, 田中弘起, “人顔認識の協調的学習フレームワークの提案”, JAWS-2008 予稿集
- [2] 岩崎剛士, 田中弘起, 高田司郎, “顔認識学習に用いる訓練データを自動作成するフレームワークの提案”, 平成20年度 関西支部 支部大会 公演論文集, pp. 219
- [3] 武岡さおり, 尾崎正弘, 川田博美, 足達義則, “学習者認識のための顔画像検出と顔の向き認識の基礎的実験”, 名古屋女子大学紀要49, 2003
- [4] 松橋聡, 藤本研司, 中村納, 南敏, “顔領域抽出に有効な修正HSV 表色系の提案”, 映像情報メディア学会, Vol. 49, No. 6, 1995
- [5] Christopher M. Bishop: “Pattern Recognition and Machine Learning”, 2006. 邦訳: 元田治他: “パターン認識と機械学習上下”