

自由歩行下での顔認識

細井 聖 川出 雅人

オムロン株式会社 新事業開発センタ ファジィ推進室
〒617 - 8510 京都府 長岡京市 下海印寺
Tel: (075) 953-3880 Fax: (075)952-0411
E-mail: {satosi-h, kawade}@zoo.ncl.omron.co.jp

1. はじめに

人と機械とのインタラクションの方法には、「意識的」なものと「無意識的」なものの二通りがある。前者の場合、人が機械に対して何らかのアクションを行う。それは、ボタンを押すことであったり、声をかけることであったり、顔を見せることであったりする。

このようなインタラクションを「意識的なインタラクション」と考えた場合、機械がインタラクションを開始するキー入力となるのは人の意識した行動である。反対に、後者は「無意識的なインタラクション」であり、機械がインタラクションを開始するキー入力は、人が意識していない行動でなければならない。たとえば、機械の近くを通りすぎた人に機械の方から反応し、何らかのインタラクションを働きかけるといったものは、「無意識的なインタラクション」と言える。

本稿ではこのような「無意識的なインタラクション」の一例として、機械の前を通りすぎる人が誰であるかを認識するシステムについて述べる。具体的には、CCDカメラの前を自由に歩行(以下、自由歩行)してくる人の顔を検出し、その人が誰であるかを認識する手法について報告する。

2. 自由歩行下での顔検出と顔認識

2.1 全体の流れ

全体の流れを図1に示す。タスクを2つ使用し、画像中に変化があったときに確実に秒間あたり指定された数のフレームを取得するタスクと、そのフレームをもとに時間をかけて顔検出・認識するタスクとに分かれる。

2.2 自由歩行

無意識の人の状態として、自由歩行の状態を考える。自由歩行下の人にはCCDカメラの方を見たり、CCDカメラのレンズに対して顔を近づけることはない。人はあくまでもCCDカメラの画角内を通り過ぎていくのみである。また、自由歩行下の人

正面にCCDカメラを設置することは不可能なので、図2のようにCCDカメラを出入り口の横に設置し、そこから得られるななめ顔の画像を対象とする。

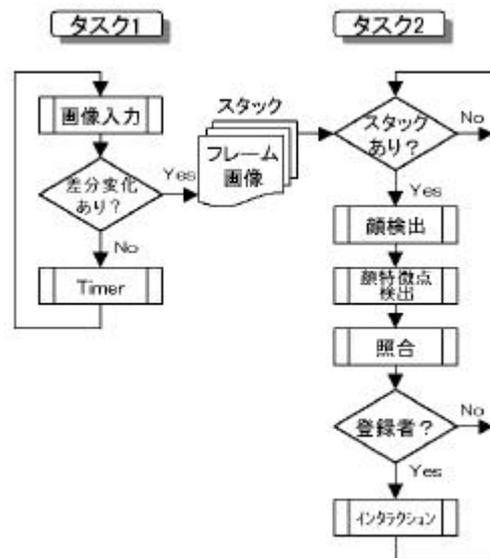


図1 全体の流れ

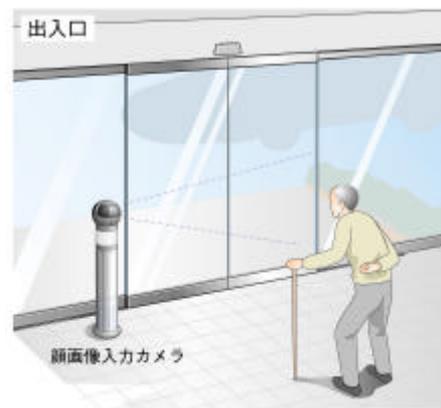


図2 カメラ設置例

2.3 入力画像のバッファリング

自由歩行下の人がCCDカメラの画角内に写っている時間は約1~2秒である。したがって、顔認識をするために必要な情報量を得るために、画像をバッファリングすることは有効である。バッファリングのタイミングとして連続した2枚のフレーム間の変動量に着目する。変動量があるしきい値以上である場合、カメラの画角内に人が写っているもの

とし、その時の画像を時系列画像データとしてバッファにため込む。自由歩行中に顔を認識するためには、4～6枚の画像が必要なため、Timer 制御により指定時間間隔で画像を取得する。

2.4 顔検出

バッファリングされた画像からななめ顔を順次検出する。顔検出手法として Gabor Wavelets Transformation + Graph Matching[1][2]を適用した。本手法では、顔画像の上いくつかの特徴点(図3)を設定し、その特徴点上の特徴量を Jet と呼んでいる。Jet は顔画像上の各特徴点において各種の方向性と周波数を持つ Gabor Kernel (図4)を使って Wavelets 変換処理を施し、その特徴点におけるパターンの周期性と方向性の特徴量として計算される。標準顔画像において、その特徴点間の空間的位置情報と、各 Jet を顔画像のグラフ表現として登録する。顔検出の際は、入力画像に対して、登録された標準顔画像のグラフ表現を平行移動・拡大/縮小・回転を行い、その類似度がしきい値を超える領域が顔であると判断する。これにより、4フレーム/秒の顔検出が可能である。

図3. 顔画像上の特徴点

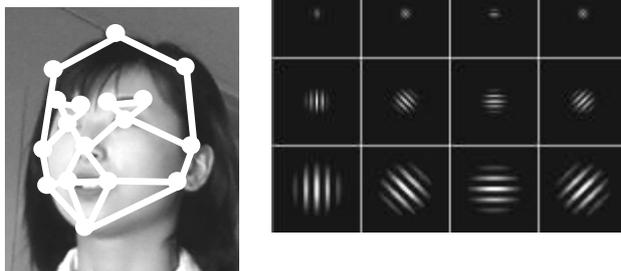


Figure 3. Facial feature points

2.5 顔認識

入力顔画像と登録画像の比較に使用する顔特徴点を精密に検出する必要がある。基本的アルゴリズムは顔検出と同様で、顔検出でラフに検出された顔領域に対してグラフ表現を当てはめ、周辺領域において平行移動・拡大/縮小・回転を行い、最も類似度の高い位置を求める。さらに各特徴点を局所的に平行移動させ、さらに類似度の高い点を求めていく。標準顔画像上の特徴点の数および各 Jet は顔検出の時よりも詳細なものを使用する。また特徴点を検出するために使用するグラフ表現には、性別・年齢層・表情・眼鏡の有無など各種の標準顔から計算された Jet が登録された Facial Bunch Graph (図5)を使用し、それらの中で最も類似度の高いグラフ表現の Jet を使うことで、各特徴点の検出率を上げている。

顔認識は登録された顔と入力された顔の特徴点間の Jet に関して、それらの類似度の総和が最も高い登録画像において、類似度の総和がしきい値を越えた場合、同一人物であると判断する。このような処理をバッファリングされたすべての顔に対して行い、類似度の平均値が最も高い登録者を認識結果と判定する。

自由歩行下で撮影した29人に対する顔検出・顔認識の実験結果を表1に示す。

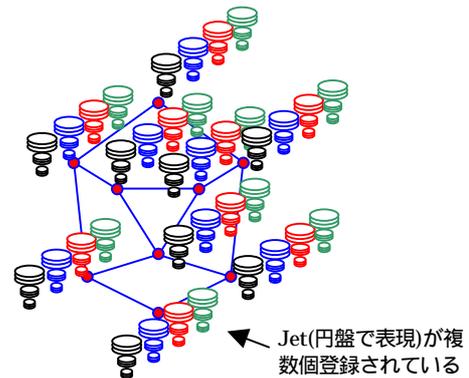


図5. Facial Bunch Graph

顔検出率	100.0%
顔認識率	96.6%

表1 29人による認識結果

3. おわりに

従来の顔認識が「意識的なインタラクション」を要求していたことに対し、本稿では「無意識的なインタラクション」の一例として自由歩行下の人を対象とした顔認識の手法について述べた。今後は、顔認識率の向上を図り、より安定したシステムを構築し、人間と機械の自然なインタラクションが可能なアプリケーションへと応用していく予定である。

参考文献

- 1) E.Elagin, J.Steffens, and H.Neven. Automatic Real-Time Pose Estimation System for Human Faces Based on Bunch Graph Matching Technology. Proceedings of the International Conference on Automatic Face and Gesture Recognition '98, 1998
- 2) J.Steffens, E.Elagin, and H.Neven. PersonSpotter - Fast and Robust System for Human Detection, Tracking and Recognition. Proceedings of the International Conference on Automatic Face and Gesture Recognition '98, 1998