

ビジュアル・エージェントのための
顔領域追跡手法

5C-6

塚本 明利

李 七雨

辻 三郎

(財) イメージ情報科学研究所

大阪大学

1. はじめに

本稿では動画像において顔領域を探索し追跡する手法について述べる。この手法は現在我々が実現を進めているコンピュータとの対話システム(ビジュアル・エージェント)への応用を目的としており、対象とする動画像は室内環境において撮影された未知の人物が任意の動作をしているものである。このため顔の大きさや向きは任意に変化しており、柔軟な処理が必要とされている⁽¹⁾。

顔を探索する処理として主成分分析による特徴ベクトルを利用するもの⁽²⁾等が提案されているが、処理の高速性や環境変化に対する柔軟性に問題がある。本手法はテンプレートマッチング⁽³⁾に基づいているため、特殊な特徴抽出処理は不要であり、相関演算をハードウェア化することにより高速に動作させることが可能である。また顔領域探索を柔軟に行なうため、平滑化およびMulti-Resolution処理を利用している。顔領域追跡においては顔の大きさや向きの変化した顔パターンを合成し、これもテンプレートとして利用することにより顔の向きの変化に対し柔軟な追跡を行なっている。

2. 顔領域探索処理

2.1 顔領域候補抽出

顔探索処理ではあらかじめ準備した顔画像を適当に平滑化及び縮小してテンプレートとし、入力画像とマッチングを行なうことにより候補領域を得る(図1)。テンプレートマッチングとはテンプレートを画像上に走査させながら相違度を求めてゆき、もっとも相違度が小さくなる位置を求めるものである。

画像Iのテンプレートと同じ大きさの領域AとテンプレートTとの相違度C(A)は式(1)により算出される。

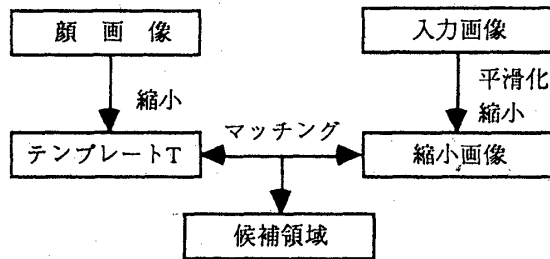


図1 顔領域候補抽出

$$C(A) = \sum_{i=0}^{T_w-1} \sum_{j=0}^{T_h-1} \left| (A(i,j) - \bar{A}) - (T(i,j) - \bar{T}) \right| / T_w T_h \quad (1)$$

ここでA(i,j)、T(i,j)はそれぞれ画像Iの領域A及びテンプレートTの座標(i,j)にある画素の輝度値、 \bar{A} 、 \bar{T} はそれぞれ画像Iの領域A及びテンプレートTの輝度平均値、 T_w 、 T_h はそれぞれテンプレートの幅及び高さである。式(1)は領域A、テンプレートTの輝度平均値の差を補正した1ピクセルあたりの輝度の差分である。

本手法はテンプレートに予め準備した顔画像を与えて入力画像とのマッチングを行なうものであるが、入力画像における顔の大きさが未知であるため、入力画像をいろいろな大きさに縮小し、別途縮小したテンプレートとのマッチングをとることにより大きさの違う顔の探索を実現している。

2.2 候補領域の絞り込み

次に得られた候補領域の絞り込みを行ない、顔領域を決定する。この絞り込みは顔候補領域とテンプレートとの明暗一致度に基づいて行なう。

画像Iの候補領域AとテンプレートTとの明暗一致度R(A)は(2)式を用いることによって算出される。

$$R(A) = (A_{\text{dark}} + A_{\text{blight}}) / (T_{\text{dark}} + T_{\text{blight}}) \quad (2)$$

ここで T_{dark} はテンプレート中の画素のうち輝度が閾値 th_L 以下である画素の数であり、 A_{dark} はそのような画素それぞれに対応する領域Aの大きさ $k \times k$ の部分領域のうち、実際に暗い(閾値 th_L 以下の輝度の画素がある)領域の数である。同様に T_{blight} 、 A_{blight} はそれぞれ閾値 th_H 以上の輝度を持つテンプレート内の画素数、候補領域A内の部分領域数である(kはテンプレートTと領域Aの大きさの比率)。

すべての候補領域に対する明暗一致度のうち、最大値を与える候補領域を結果として出力する。

2.3 処理結果

実際の画像について行なった処理結果を表1に示す。同表中正解率とは結果が正解であった割合、また正解受諾率とは候補領域内に正解が含まれていた割合である。テンプレートを作成するのに用いた顔画像は入力画像とは別のものであり、入力画像はそれぞれ異なる人物の正面顔がいろいろな大きさや背景下で写っている。表1よりテンプレートによって性能に違いはあるものの、人物の個人差や大きさ、背景の変化によらず精度良く顔探索を行なっている。

テンプレート	入力データ	正解率	正解受諾率
# 1	平滑化せず	0.500	0.66
	平滑化	0.611	0.77
# 2	平滑化せず	0.333	0.88
	平滑化	0.278	0.83

表1 顔領域探索結果 (データ数=18)

3. 顔領域追跡処理

3.1 処理アルゴリズム

顔領域追跡処理では、顔探索処理によって得られた顔画像をテンプレートを用いて、以降のフレームにおける近傍の探索範囲においてマッチングを繰り返すことにより追跡を行なう。しかしながら対象としている画像は人物が任意の動作をしているため、人物の運動や顔の向きの変化によって顔のパターンが変動し、安定した追跡を行なうことは困難である。

本手法では顔探索処理によって得られた顔画像(カレントテンプレート)の他に、それを適宜拡大・縮小・回転させた画像もテンプレートとし、以降のフレームの画像とのマッチングに利用することにより顔の大きさや向きに追従した追跡を実現している(図2)。ここでは変形パラメータを長さ比0.9倍、1.1倍、および画像面上の±30°の回転としている。

入力画像 I_t に対しては以下の処理を行う。

(1) カレントテンプレートによる探索

カレントテンプレートを用いて顔位置近傍の探索範囲でマッチングを行なう。ここでは入力画像とテンプレートの縮小を行わずに(1)式の相違度を算出する。相違度の最小値が閾値 th 以下の場合にはその時のテンプレートの位置を顔の位置として出力する。

(2) 他のテンプレートによる探索

最小相違度が閾値を越えている場合には、テンプレートデータベースに登録されている他のテンプレートを用いて再度マッチングを行なう。そして相違度の最小値が閾値 th 以下の場合には、その時のテンプレートの位置を顔の位置として出力し、マッチしたテンプレートを次のカレントテンプレートとする。

(3) 顔位置の推定

全てのテンプレートでのマッチングの結果、最小相違度が閾値を越えている場合には、テンプレートデータベース内にはマッチする顔のパターンが存在しないと判断し、顔領域をそれまでの顔の位置の移動情報によって次のように推定する。

現在の時刻 $t(t \geq 2)$ における顔の推定位置は顔領域の左上座標を (x_t, y_t) とすると、 $x_t = x_{t-1} + \Delta x$ 、また $y_t = y_{t-1} + \Delta y$ とする。ここで $\Delta x = x_{t-1} - x_{t-2}$ 、 $\Delta y = y_{t-1} - y_{t-2}$ である。また顔の大きさは時刻 $t-1$ の時のものと同じとする。得られた推定領域の画像を新たにカレントテンプレートとする。

(4) 新規テンプレート作成

カレントテンプレートが更新された場合には、この拡大・縮小・回転画像も作成してテンプレートデータベースに追加する。このような処理を繰り返すことにより以降のフレームにおける顔領域を逐次求めてゆく。

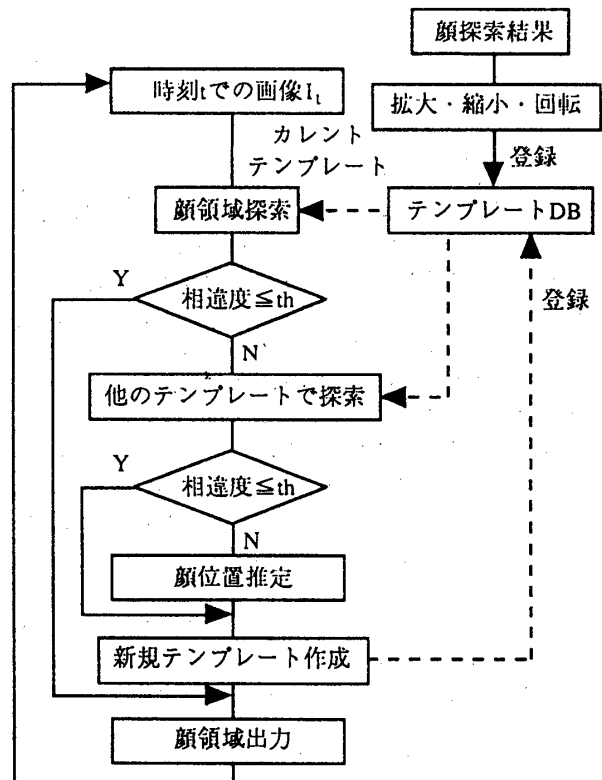


図2 顔領域追跡処理

3.2 処理結果

実際の画像に対し本処理を実行した結果、顔の向きや見かけの大きさの変化に追従した顔領域追跡を行なうことができた。しかし、カレントテンプレートがマッチしなくなった時に他のすべてのテンプレートとマッチングをとり直すことは非効率である。また新規テンプレートの登録によりテンプレートの枚数が非常に大きくなる。

効率良くマッチングを行なうためのテンプレートの選択法、およびテンプレートを大きさや顔の向きに分けて階層的に管理するなどの工夫が必要である。

4. まとめ

本稿ではビジュアルエージェントへの応用を考慮した顔画像探索/追跡の手法について述べた。現状では顔探索処理は顔追跡処理に先だつて動作するが、今後は顔探索処理を応用することにより顔検証を行ない、顔追跡処理の監視を行なうことにより安定性の向上を目指したい。

参考文献

- (1)小杉、「シーンの中の顔の探索と認識」、PRU91-104
- (2)M.A.Turk,A.Pentland,"Face Recognition Using Eigen-faces", CVPR'91
- (3)谷内田、「ロボットビジョン」、培風館、1990