

テンプレート切り替え型顔検出に関する検討 A Study on Face Detection Method with Various Templates

松尾 賢治
Kenji MATSUO

橋本 真幸
Masayuki HASHIMOTO

小池 淳
Atsushi KOIKE

1. はじめに

近年、セキュリティやエンターテイメント分野で顔認識技術が注目を浴びている。いずれのシステムにおいても画像内から高精度に顔を検出する技術が求められる。画像内から顔を検出する方法の一つとして、テンプレートマッチング[1]を利用した方法が挙げられる。これは予め多数のサンプル顔により作成された平均顔をテンプレートとし、画像内でテンプレートと類似している領域を探索し、類似度の高い領域を顔と判定する方法である。しかし、この方法の検出精度はテンプレートを作成する際に使用したサンプル顔に左右されるため、誰に対しても汎用的に使用できるテンプレートを作成することが困難であった。すなわち、テンプレートを作成する際に使用したサンプル顔の集合に含まれていない顔特徴を持つ人物に対しては、テンプレートとの類似度が低下し、顔が正しく検出できない問題があった。そこで本論文では、異なる複数のテンプレートを準備し、それらを複合的に、また切り替えて使用することによって、顔領域の検出精度を改善する方法を提案する。

2. テンプレートマッチング

テンプレートマッチングでは次の手順で画像内から顔を探索する。予め多数のサンプル顔から平均顔を作成し、それをテンプレート f とする。まず、検出の対象となる画像内を左から右、上から下のラスタースキャン順にテンプレートを適用し、各位置 (i, j) における矩形領域 $x(i, j)$ との類似度 S を測定する。次に、類似度がある設定されたしきい値 T を超えた場所の中で最も高い値を持つ位置の矩形領域を顔と判定する。ここで、本論文では類似度 S を次式(1)の相関値で定義した。 θ は $x(i, j)$ と f のなす角である。

$$S = \cos\theta = \frac{x(i, j)^T \cdot f}{|x(i, j)| |f|}, \quad (-1 \leq S \leq 1) \quad (1)$$

顔は各人物様々に異なり、その特徴は十人十色である。したがって、どのような顔特徴に対しても汎用的に使えるテンプレートを作成するためには、サンプル顔の数を増やして平均顔を作成する必要がある。しかしサンプル顔の数が増加するにつれて顔の特徴が平均化されてテンプレートが鈍ってしまい、誰に対してもある一定の類似度しか示さなくなる。この場合、平均顔から離れた特異的な顔特徴を持った人物に対しては、テンプレートが有効に作用しない。

本論文では、図1の手順で正規化された顔画像をテンプレート作成用のサンプル顔として使用する。鼻腔の midpoint と

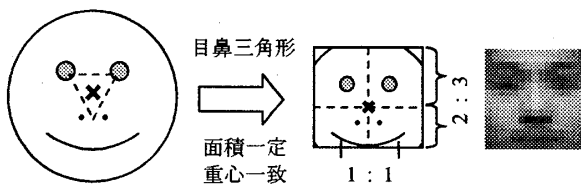


図1. テンプレート作成用サンプル顔の正規化手順

両目の3点を結ぶ三角形が人物に依存せずある一定の面積になるように画像を縮小し、さらに三角形の重心を切り出す矩形領域で縦方向に2:3かつ横方向に等分割する点に配置して、1辺15画素の正方領域として切り出す。ここで、テンプレートに使用するサンプル顔の数を変化させた場合の平均類似度の変化を図2に示す。テンプレート作成用のサンプル顔に300人、さらに類似度測定対象として100人の顔画像を用いた。図より、100人に対する類似度が低下していることが確認できる。

3. 提案方式

前節で述べた通り、サンプル顔の数が増加するにつれてテンプレートが鈍ってしまう。顔検出処理ではしきい値による判定を行うため、類似度がしきい値を下回り、顔であっても非顔と誤判定される問題があった。したがって、300人全てのサンプル顔を使って平均顔を作成するのではなく、無作為に100人ずつ3グループA, B, Cに分け、計3つの平均顔を作成する。100人ずつで平均顔を作成したことにより、各グループの平均顔に若干の偏りが生じる。図3に提案法の概念図を示す。これらの複数テンプレートを複合的に、また切り替えることにより、顔検出精度を改善する。実際に偏りが生じているか否かを平均顔同士の相関を測定し、表1にまとめた。ただし、グループA, B, Cそれぞれのテンプレートを f_A, f_B, f_C とし、300人全員のサンプル顔から作成したテンプレートを f_0 とした。それぞれ異なった顔特徴を持つことが確認できる。

テンプレートを1種類に限定せず複数のテンプレートに分割したことにより、検索の対象となる顔特徴に近いいずれかのテンプレートに整合する可能性が上がる。極端なケースを仮定し、サンプル顔が100人の男性、100人の女性、100人の外国人から構成されているとする。検索対象となる画像に男性が写っていた場合、300人全員をまとめて平均化して作成したテンプレート f_0 よりも、100人の男性画像から作成したテンプレートの方が高い類似度を示す。

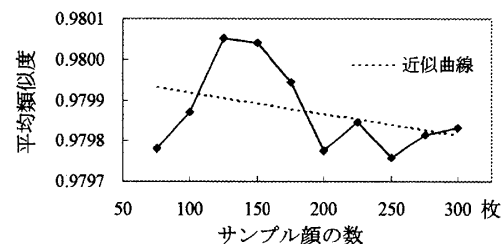


図2. サンプル顔と類似度の関係

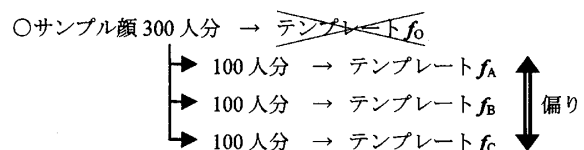


図3. 複数テンプレート切り替えの概念図

表1. 各テンプレート間の相関関係

	f_A	f_B	f_C	f_0
f_A	1.0000	0.9992	0.9903	0.9986
f_B	0.9992	1.0000	0.9925	0.9993
f_C	0.9903	0.9925	1.0000	0.9960
f_0	0.9986	0.9993	0.9960	1.0000

表2. 複合的に使用した場合の顔検出判定結果 (しきい値 $T=0.95$, 検出対象画像の総数 100 枚)

	閾値	目視	f_0 単一	f_A, f_B, f_C	差分
①	成功	成功	71	78	+7
②	失敗	成功	4	2	-2
③	成功	失敗	19	17	-2
④	失敗	失敗	6	3	-3

4. 計算機シミュレーション

提案方式の有効性を検証するために計算機シミュレーションを行った。テンプレートマッチングの前処理として、HSV 色空間を利用した肌色領域判定を行い、肌色を多く含む矩形領域のみを探索対象とした。また、大きさの変化にある程度適応できる顔検出を行うため、検出対象となる画像の解像度を3段階に縮小し、それぞれの縮小画像上でテンプレートマッチングによる顔検出を行った。以下、複数テンプレートを複合的に使用する場合と切り替えて使用する場合のそれぞれについて提案方式の有効性を検証する。

4. 1. 複合的に使用する場合

テンプレート作成用のサンプル顔に男性 300 人を用い、300 人全員のサンプル顔から作成したテンプレートを f_0 、無作為に 100 人ずつ 3 グループに分け、作成したテンプレートを f_A, f_B, f_C とした。また、解像度が 640x480 で背景が比較的複雑ではない男性の正面顔 100 枚を顔検出の対象画像とする。複数テンプレートを同時に適用し、最大の類似度を示したテンプレートとその検出領域を求める。この場合、3 種類のテンプレートを同時に適用することになるため、処理量が 3 倍になる。得られた結果に対し、次の 2 つの基準により顔検出の成否を判定した。

- ・しきい値による成否判定
- ・目視による成否判定

実際の顔検出システムでは顔の成否を前者の判定基準により自動的に判定するが、有効性を検証するために検出領域を目視で判断して、それが顔であるか否かを調べた。上述の観点から、顔検出結果には次の 4 ケースが考えられる。

- ①しきい値判定で成功 \cap 目視判定で成功
- ②しきい値判定で失敗 \cap 目視判定で成功
- ③しきい値判定で成功 \cap 目視判定で失敗
- ④しきい値判定で失敗 \cap 目視判定で失敗

このうち可能な限り①のケースを増やし、②③④のケースを減らすことが重要である。全 100 人のうちそれぞれの発生数を測定した。結果を表 2 に示す。①について見ると、テンプレート f_0 を単一で使用した場合は 71 人の顔検出が成功したのに対し、複数のテンプレート f_A, f_B, f_C を複合的に使用した場合は 78 人の顔検出が成功し、7 人分だけ検出成功率が改善されたことが確認できる。また、複数のテンプレートを適用することによって逆に誤検出、つまり③のケースが増えることも考えられる。しかし、③の発生数も 19 から 17 まで減っており、誤検出率は増加しないことが確認できる。よって、提案方式は複数テンプレートを同時に使用して顔検出を行う場合に有効である。

表3. 切り替えて使用した場合の顔検出判定結果 (しきい値 $T=0.95$, 検出対象画像の総数 1000 枚)

	閾値	目視	f_D 単一	f_E 単一	f_F 単一	f_D, f_E, f_F 切替
①	成功	成功	775	753	766	833
②	失敗	成功	19	47	29	167
③	成功	失敗	91	112	108	
④	失敗	失敗	115	88	97	

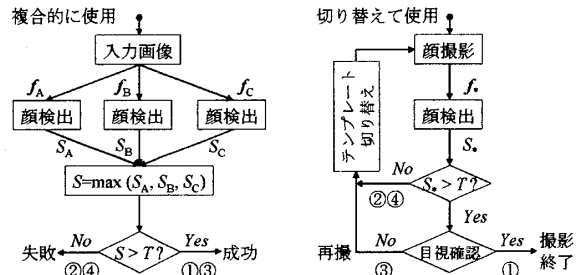


図4. 提案方式の処理アルゴリズム

4. 2. 切り替えて使用する場合

複数テンプレートを切り替えて使用する場合の有効性を検証する。これはカメラの前の人物がどの職業の顔特徴に似ているかを診断するエンターテイメントシステム[2]への適用を想定したものであり、全ての体験者に等しく均等なサービスを提供するために、顔検出率を可能な限り改善する必要がある。このシステムでは、まず顔画像を撮影と顔検出を行って、検出された顔領域を画面上に表示する。体験者は検出結果が正しいと判断すれば撮影終了し、次へ進む。正しくないと判断すれば再度顔を撮影する。このとき、再撮影のために別のテンプレートに切り替えることによって、初回撮影でテンプレートの顔特徴に一致していないことから検出失敗となった人物でも、2 回目、3 回目の撮影を行うことによって、正しく顔を検出できる可能性が増す。

新たにテンプレートを 3 種類 f_D, f_E, f_F 作成し、男女を含む被験者 1000 人を対象とした大規模な実験を行った。 f_D, f_E, f_F それぞれのテンプレートを単一で使用した場合と、切り替えて使用した場合の結果を表 3 にまとめた。表より、例えばテンプレート f_D を切り替えずに繰り返し使用した場合、①の成功ケースが 775 人で留まるのに対し、テンプレート 3 種類 f_D, f_E, f_F を切り替えて使用した場合は、①の成功ケースが 833 人に増加し、多種多様に異なる顔の検出精度を改善できることがわかる。この際の平均撮影回数は 1.45 回であった。また、図 4 にまとめた提案方式のアルゴリズムに示す通り、体験者の目視確認プロセスを加えたことにより、原理的にも②③④のケースの発生を低減できることが確認できる。

5. まとめ

テンプレートマッチングによる顔検出において、顔を精度良く検出できるテンプレートを作成すること自体が困難であった。複数のテンプレートを複合的に使用することにより顔の検出精度を改善する方法を提案した。

参考文献

[1] Brunelli R., "Face Recognition: Features versus Templates," IEEE Transaction on PAMI, Vol. 15, No. 10, Oct. 1993
 [2] 松尾他, "顔認識技術を応用した類似顔診断エンターテイメントシステムの開発," ITE2005 年次大会予稿集