

シーン内からの自動顔検出だけでなく、近年は高精度な顔部品検出や、それらを用いたアプリケーションが多く提案されている。本論文では、心理学的な実験における画像刺激で採用例の多い手法の一つである FUTON を具体的な応用例と設定し、従来は手入力で得ていた FUTON の顔特徴点列を高精度に自動抽出する手法を提案する。

1. はじめに

近年、顔認識において DNN (Deep Neural Network) を利用した応用例が目立っている。顔検出や顔認証における重要な技術として DNN は位置づけられており、これらの手法では DNN を用いることで、より有効な特徴量を選別することができるが分かっている[1]。一方で、顔部品の輪郭検出においては、画像や ROI を入力とするだけでなく、顔部品の輪郭情報を入力として与えるような機械学習[2 - 4]として枠組みを設計した方が、計算コスト、精度において有効であることが分かっている。すなわち、顔部品の輪郭の定義においては、ある程度の知見に基づく識別器が設計されることが望ましい。

我々は、心理学的な実験における画像刺激で採用例の多い手法の一つである FUTON[5]を具体的な応用例とし、従来は手入力で得ていた FUTON の顔特徴点列を自動抽出する手法を提案・評価している。本論文では、正面顔のみで構成された学習サンプルを用いていた先行事例での試みに加え、顔を正対した状態からやや斜めに向かせた例をサンプルに加えることの優位性について検証をする。

2. 輪郭情報に基づいた顔部品検出手法

シーン内の顔領域検出だけではなく、さらに詳細な顔部品の検出においては、形状回帰モデルに基づく様々な手法[2 - 4]が近年は提案されており、ロバストかつ非常に高速な処理となっている。これらは、何らかの顔検出器によって得られた顔領域に対して、各顔部品の点数等を予め決めておき、任意の局所画像特徴量を求める回帰を解くことで顔部品特徴点の座標を検出している。

本論文においては、手法[3, 6]で提案されている枠組みを用い、予め学習サンプルから得られる顔部品領域の特徴点群より平均的な顔形状を求めておき、各学習サンプルに対する最適な変位量を様々なパターンで学習させている。ここで、平均的な顔形状に対して1度の更新では最適解を得ることができないため、段階的に変位量を求めている。本手法においては、局所特徴量に HOG を用い、初期値から予め入力された特徴点への変位量を求める形状回帰モデルを用いて顔部品の特徴点検出器を生成している。

3. FUTON の顔特徴点列

顔部品特徴点の自動抽出手法においては、図 1 に示すように眉、目、鼻、口、顎を構成要素としている例が主となっている[2-4]。これらは一般的に、Point distribution model (PDM) [7]に基づく顔モデルであり、各部品においては、目や口は輪郭が閉じている例が多数であるが、眉については輪郭が閉じた領域として定義する例と、閉じていない曲線(実際には 1 ストロークで描ける直線の集合)のみで表現する例があり、同様

に鼻においても鼻梁線の扱いで複数の例が見られる。一方で、心理実験に用いる事を前提としている手法では、図 2 に示すように、頭部もしくは肌と頭髪の境目を部品として定義している例がある[5]。これは心理実験等に用いる事を前提とした場合、肌領域のモーフィング等の形状変化・操作が必要となることから予め頭部の顔部品を定義しているためであると考えられる。

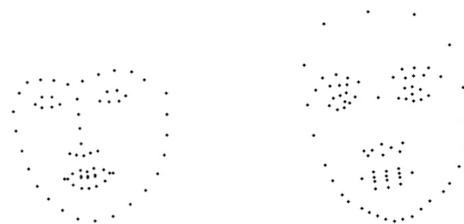


図 1 顔部品の例

図 2 FUTON の特徴点

表 1 FUTON の特徴点の詳細

部品	詳細	点数
目部品(左右)	瞳中心座標	1
	目輪郭	8
眉部品(左右)	眉輪郭	8
鼻部品	鼻根部	1
	鼻頂点	1
	小鼻	7
口部品	輪郭(外)	8
	輪郭(内)	6
顔輪郭	頭-顎	26

本手法では図 2 に示す FUTON の顔特徴点列[5]を自動抽出処理の対象とする。FUTON の顔特徴点は表 1 に示すような各顔部品で構成されており、全体で 83 点で構成されている(実際には画像の周辺部にも特徴点が配置されている)。FUTON の各特徴点は GUI の入力規則に従い、一定のルールに従って配置されている[5]。

4. 顔部品検出のための識別器生成

4.1 FUTON dataset

著者らは、FUTON の顔特徴点列および画像セットを多数有している。これらの顔画像は、基本的にカメラに正対した環境で正面を向いて撮影したものであり、図 3 に示すような、画

像サイズ 512x512 のおよそ中央へ縦方向に配置された頭部を含めた顔全体が収まっている。このような画像と、図 4 に示すような各顔画像の輪郭情報をマニュアルで入力した特徴点情報が紐付いたデータセットとなっている。すなわち、これらを学習サンプルとして用いる事で、識別器生成までの作業を効率的に進めることができる。

4.2 識別器の生成

本実験においては 20 代の男性 175 名、女性 196 名を学習サンプルとして識別器の生成をした。得られた識別器を用いた検出処理では、一つの顔領域に対する顔部品検出の平均速度は 2.84ms であり(CPU: Intel Core i7 2.7GHz, メモリ: 16GB)、他の手法[2-4]と同様にリアルタイムでの動作についても問題はない。

図 5 に本手法を用いた特徴点の抽出結果の一部を示す。図中の矩形は顔領域検出の結果であり、この領域を基準として平均的な顔形状が初期値として配置される。図 5 の結果からも確認できるように、検出結果については眼鏡の有無については問題はない。

しかしながら、この識別器生成に用いた学習サンプルは、ほぼ全てがカメラに対して正対した姿勢の顔画像であったため、顔向きとカメラ光軸が異なる場合では検出性能に問題があることが分かっている[8]。そこで本論文では上記の学習サンプルに加え、プロフィール写真に使われる程度に斜めを向いた顔写真と、FUTON 特徴点入力結果を新たに準備した。これらの斜めを向いた写真は右向きが 22 例で左向きが 30 例であり、計 52 のデータを新たに加えて識別器を生成した。

4.3 検出性能の評価

4.2 節で得られた識別器を用いて、学習サンプルに含まれない男性 25 名、女性 114 名のデータセットに対して顔部品検出性能の確認をした。本論文では、予めマニュアルで入力された特徴点の座標を真値とし、本手法で得られた特徴点との位置の比較をした。

本実験においては、表 2 に示すように部品全体の真値に対する誤差を検証した。本手法で機械学習および検証に用いた画像セットは、512x512 ピクセルの画像サイズにおおよそ同じような大きさになるよう予め画像加工している。また、特徴点は顔の内部部品の誤差は全体的に小さくなるため、顔の輪郭(顎および頭部)と分け評価しており、それぞれ顔の外部部品のみでの誤差の平均値、内部部品のみでの誤差の平均値、また全体の誤差の平均値を算出した。正面顔のみで学習した識別器(以下 A とする)は顔の内部部品の方が誤差は少ないが、全体的に誤検出が多い結果となった。斜め顔を含めた識別器(以下 B とする)では顔の内部部品だけでなく、顔の外部部品の検出精度も向上している。特に外部部品については顎輪郭の検出精度が高まっているため、この結果のような誤差となった。

表 2 実験結果

	外部部品	内部部品	全体
識別器 A	35.73	28.25	30.59
識別器 B	10.15	6.52	7.66

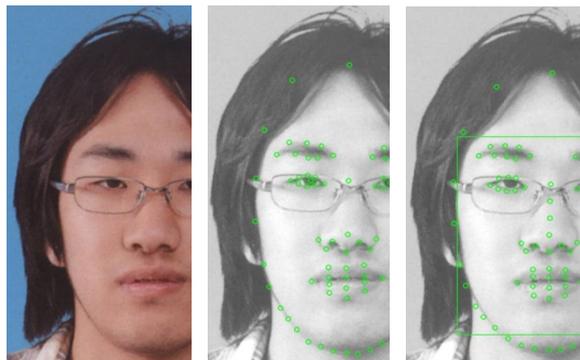


図 3 カメラで
取得した画像例

図 4 手入力
された特徴点

図 5 自動抽出
された特徴点

5. おわりに

本論文では、従来は手入力特徴点を得ていた FUTON の顔部品をより安定させて自動検出する手法を提案した。FUTON は心理実験等を目的としており、顔のモーフィングによる画像生成が主の目的であるため、本手法をこれらの応用事例へ展開することが今後の課題となっている。また、今回の実験では斜めを向いた顔に対する学習サンプルの検討はされているが、他の方向への向きや、その他の環境への考察も必要である。これらを含めたデータセットの拡充や吟味が今後にも必要となってくる。

参考文献

- [1] 山下隆義: 顔画像センシング技術の最新動向, システム制御情報学会, Vol.59, No.2, pp.44-46 (2015)
- [2] Xudong Cao, Yichen Wei, Fang Wen, Jian Sun: Face Alignment by Explicit Shape Regression, International Journal of Computer Vision, 107, 2, pp. 117-190 (2014)
- [3] Vahid Kazemi, Josephine Sullivan: One Millisecond Face Alignment with an Ensemble of Regression Trees, Computer Vision and Pattern Recognition (CVPR), pp.1867-1874 (2014)
- [4] Shaoqing Ren, Xudong Cao, Yichen Wei, Jian Sun: Face Alignment at 3000 FPS via Regressing Local Binary Features, Computer Vision and Pattern Recognition (CVPR), pp.1685-1692 (2014)
- [5] 向田茂, 蒲池みゆき, 尾田政臣, 加藤隆, 吉川左紀子, 赤松茂, 千原國宏: 操作性を考慮した顔画像合成システム:FUTON - 顔認知研究のツールとしての評価 -, 信学論(A), Vol.J85-A, No.10, pp.1126-1137 (2002)
- [6] Piotr Dollár, Peter Welinder and Pietro Perona: Cascaded Pose Regression, 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR2010), pp. 1078-1085 (2010)
- [7] Tim Cootes and Chris Taylor; Vijay Kumar: Active Shape Models-'Smart Snakes', In British Machine Vision Conference, pp.266-275 (1992)
- [8] 藤原孝幸, 橋本良平, 向田茂: 既存の顔特徴点列に対する形状回帰モデルに基づく顔部品検出手法, 動的画像処理実利用化ワークショップ 2018 (DIA2018)論文集, IS1-2 (2018)