

進化的アルゴリズムを用いた顔検出器のための 学習用合成顔生成

松浦和博[†] 木村周平[†] 松村幸輝[†]

概要：顔検出器を作るには大量の顔写真が必要である。そこで本論文では、進化的アルゴリズムを用いて顔写真から人工的に「顔らしき合成画像」を作成することで学習用正解画像を増加し、少ない枚数の顔写真からでも精度の高い顔領域検出が行える検出器を作成することを目的とした。結果として、人間を撮影した顔写真に加えて、顔の各器官を再構成して作成した合成顔を学習に用いることで、少ない写真数で高い精度の顔検出器を作成できることを示した。

Generation of Synthetic Faces for Learning using Evolutionary Algorithm to Make Face Detector

Kazuhiro Matsuura[†] Shuuhei Kimura[†] Kouki Matsumura[†]

Abstract : In this paper, we aimed to make high quality face detector from a small quantity of face images by making up the "synthetic images like face" by using evolutionary algorithm because a large quantity of face images are necessary to make a face detector. In the result, we could show that we can make it.

[†] 鳥取大学大学院
Graduate School, The University of Tottori

1. はじめに

近年ロボットやセキュリティ分野、写真撮影など、さまざまな状況下で、カメラでの撮影画像中の顔領域を検出する技術が必要とされている。

顔領域の検出にはさまざまな手法があるが、近年では画像中の特徴を抽出し、その結果をあらかじめ学習済みの識別器に通す手法が中心となっている。顔領域検出に用いられる代表的な特徴抽出手法には、固有顔 (Eigenface)[1], Haar-Like[2], HoG (Histograms of Oriented Gradients)[3]などが挙げられる。また識別器は、単純ベイズ分類[4], バックプロパゲーションを用いたニューラルネットワーク[5], k近傍法[6], ブースティングなどが有名である。

いずれの手法にせよ、学習ベースの検出のためには学習用のデータとして、顔が写っている画像 (以下、正解画像) と、顔以外のものが写っている画像 (以下、非正解画像) がそれぞれ大量に必要となる。ところが人間の顔写真は個人情報にあたり[7], 一般的に撮影には本人の同意が必要となるうえ、その扱いも慎重を期すことになるため、大量に収集することが難しい。そこで本論文では、顔写真から人工的に「顔らしき合成画像 (以下、合成顔)」を作成することで学習用正解画像を増加し、少ない枚数の顔写真からでも精度の高い顔領域検出が行える検出器を作成することを目的とした。

2. 進化的アルゴリズム

進化的アルゴリズムは、生物の進化過程を模倣することで効率的に近似解を導くことを目的としたメタヒューリスティックアルゴリズムの一分野である。

本論文では、学習用の合成顔を作成するためにどの顔写真のパーツをどこに置くかという情報を進化的アルゴリズムによって制御している。以下では本論文で採用した進化的アルゴリズムの一種である遺伝的アルゴリズムについて説明する。

2.1 遺伝的アルゴリズム (GA)

遺伝的アルゴリズム (Genetic Algorithm 以下、GA) は、1975年に Holland によって提案された手法[8]である。

遺伝子と呼ばれる解候補および問題に対する適応度をもった「個体」を複数用意し、それら個体の集合 (おもに「集団」と呼ばれる) の中から適応度の高い個体を二つ選択する。選択された個体中遺伝子同士の一部を交換すること、または個体同士の位置関係をもとに新たな個体を生成することで、子個体を作る (この操作は「交叉」と呼ばれる)。個体の多様性を確保するために遺伝子の一部を変化させる「突然変異」と呼ばれる操作を行うこと

もある。こうしてできた子で新たな集団を形成し、再び適応度の高い個体を選択する。上記の操作を繰り返し、最終世代で最も高い適応度を持つ個体の遺伝子を解として出力する。

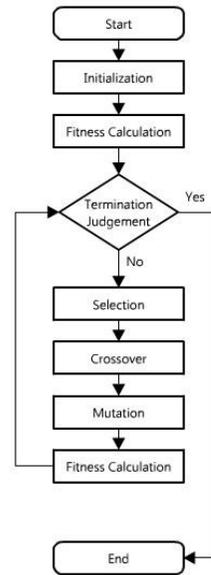


図 1 GA の基本的な流れ

Figure 1 The Basic Flow of GA.

3. 提案手法

本手法では大きく分けて三つのプロセスが存在する。

- A) 撮影済み顔写真のみにより学習した顔検出器の作成
- B) 撮影済み顔写真を用いた合成顔の生成
- C) 撮影済み顔写真および合成顔により学習した顔検出器の作成

以下では上記プロセスについて詳述する。

3.1 A) 撮影済み顔写真のみにより学習した顔検出器の作成

最初に、撮影済み顔写真のみを用いて顔検出器を作成する。本論文で使用している検出

器作成手法は、OpenCV[a][9]によって実装されている Haar-Like 特徴を用いた AdaBoost ベース識別[2]である。学習用の正解画像データは、Caltech101 データベース[10]に含まれる 5 人分 89 枚の顔写真から顔領域のみを 100px 平方へ縮小して切り出し、明るさおよびコントラストを調整したのち 24px 平方へ縮小したものを使用した。非正解画像データは同じく Caltech101 データベースの中から顔でないオブジェクトを写した 2980 枚の画像を加工なしで使用した。



図 2 学習用顔写真例

Figure 2 Example of The Pictures of Face for Learning.



図 3 切り出し・縮小後の学習用顔写真例

Figure 3 Example of Trimmed / Shrunk Pictures of Face for Learning.

ここで作成した検出器は、次のプロセスで GA の個体を評価するために使用する。

3.2 B) 撮影済み顔写真を用いた合成顔の生成

まず、合成のために A) で使用した 100px 平方の顔領域画像から「左目」「右目」「鼻」「口」の各器官を切り出す。これを最終的に 130px 平方の背景上に福笑いの要領で並べることで合成顔を生成する。ここで使用した背景とは、明るめの灰色 (RGB 値がすべて 210) で塗りつぶした PNG ファイルである。

個体は「パーツ番号」「パーツ中央の X 座標」「パーツ中央の Y 座標」の各要素で構成し

a) OpenCV は米国 Intel Corporation の米国及びその他の国における登録商標または商標である。

た。座標は左上を原点とし、右下に近づくにつれて値が大きくなるものとする。パーツ番号は、各器官あたり 89 枚ある画像のうちどれを使うかを表している。

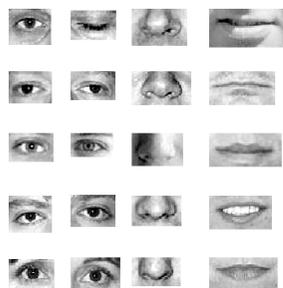


図 4 切り出したパーツ例

Figure 4 Example of Trimmed Parts.

3.2.1 初期集団の生成

集団サイズとして規定した数の個体をランダムに生成する。この時、パーツの位置情報の決定は各器官の位置関係を考慮している。なお、鼻の上に口が来る等、各器官同士の位置関係がおかしくならないよう、座標存在可能範囲に制限をかけている。

3.2.2 評価・選択

生成した各個体ごとに A) で作成した顔検出器を用いて「顔として判断できるかどうか」を評価する。そのため、本手法での評価値は正と偽の二値である。評価値が正である個体はエリート保存戦略[11]に基づき生存個体プールへ保存する。交叉のための親はこの生存個体プールの中からランダムに二つ選択する。

3.2.3 交叉

個体内の遺伝子は自然数の配列となるため、BLX- α [12]にて交叉を行う。本実験では α を 0.2 と設定した。一組の親個体から集団サイズの 10%の子個体を作り、親個体を選択し直してまた子個体を作る、という操作を 8 回行うことで集団サイズの 80%の個体を交叉によって生成する。

3.2.4 突然変異

残り 20%の個体は初期集団生成時と同様にランダムに生成するが、これは子個体遺伝子

の全部を変異させたことと同じであると捉えることができる。

3.2.5 評価・選択

交叉および突然変異にて生成した子個体群を再び A) で作成した顔検出器で評価し、正であれば生存個体プールへ保存する。

3.2.6 終了判定

生存プール内の個体数が指定数に達していた場合は繰り返し処理を終了する。そうでなければ選択処理に戻る。本実験では指定数を 300 とした。

3.2.7 書き出し

学習用データとして使用するため、生存プール内の全個体を 24px 平方の PNG ファイルとして書き出す。

3.3 C) 撮影済み顔写真および合成顔により学習した顔検出器の生成

A) で使用した顔写真および B) で作成した合成顔を使い、再び顔検出器を作成する。これを本手法の成果物とする。

4. 実験および結果

実装は OpenCV, OpenCvSharp[13], Microsoft Visual Studio 2005 Academic Edition[b]を用いて、C#にて行った。

テストは上記 Caltec101 データベースに含まれる 435 枚の写真を用いて行った。これは学習の際に使用した写真 89 枚を含んでいる。また各写真内に存在する顔は 1 つずつである。



図 5 合成・縮小後の合成顔例

Figure 5 Example of Compounded / Shrunk Synthetic Faces.

b) Microsoft, Visual Studio は、米国 Microsoft Corporation の米国及びその他の国における登録商標または商標である。

表 1 実験結果
Table 1 Result.

学習用 正解画像	学習用 非正解画像	正解顔数	検出数 (検出率)	誤検出数 (誤検出率)
顔写真 89 枚	2980 枚	435	337 (77.5%)	23 (5.29%)
顔写真 89 枚 + 合成顔 300 枚	2980 枚	435	376 (86.4%)	52 (12.0%)

顔写真のみを用いて検出器を作成した場合に比べ、合成顔を加えたものは検出率が上がっている。同様に、誤検出率も上昇している結果となった。

5. 考察

合成顔を加えたことで学習のための正解画像が増え、AdaBoost 識別の精度が上昇した結果、より顔を検出しやすくなったと考えられる。

誤検出率が上昇している要因としては、学習データにさまざまな合成顔を加えたため多様性が向上し、検出器内での「顔」の定義範囲が広がったと予想する。また、正解画像の枚数は 89 枚から 389 枚へ増えているが非正解画像はどちらも 2980 枚のため、非正解画像に対する正解画像の比率が変化した（正解画像：2.99%，非正解画像：16.4%）ことが影響している可能性がある。

6. おわりに

人間を撮影した顔写真に加えて、顔の各器官を再構成して作成した合成顔を学習に用いることで、少ない写真数で高い精度の顔検出器を作成できることを示した。

本論文では合成顔を生成する際、どのパーツを使うか、それをどこに置くか、という部分のみを制御した。今後はパーツの角度や大きさ、明るさなども制御することによって、より多様性のある合成顔の生成が可能と考える。また、本論文中で作成した顔検出器を使って新たに合成顔を作成し、それを用いて顔検出器を再作成する、という流れを繰り返すことで、徐々に精度を向上できる可能性があることについても検証する必要があるだろう。

参考文献

- 1) M.Turk and A.Pentland: Eigenfaces for recognition, Journal of Cognitive Neuroscience, Vol.3, No.1, pp. 71-86 (1991).
- 2) Paul Viola and Michael Jones: Rapid Object Detection using a Boosted Cascade of Simple Features, IEEE CVPR 2001, pp. 511-518 (2001).
- 3) Navneet Dalal and Bill Triggs: Histograms of Oriented Gradients for Human Detection, IEEE CVPR 2005, pp. 886-893 (2005).
- 4) Pedro Domingos and Michael Pazzani: Beyond Independence: Conditions for The Optimality of The Simple Bayesian Classifier, Machine Learning, pp. 105-112 (1996).
- 5) David E Rumelhart, Geoffrey E Hinton and Ronald J Williams: Learning representations by back-propagating errors, Nature, Vol.323, pp. 533-536 (1986).
- 6) E. Fix and J.L. Hodges: Discriminatory analysis - nonparametric discrimination: Consistency properties. Technical Report 4, USAF School of Aviation Medicine, Randolph Field, Texas (1951).
- 7) 個人情報保護に関する法律案, 第一章, 第二条, 第一項 (2003).
- 8) John H Holland: Adaptation in Natural and Artificial Systems, University of Michigan Press, Ann Arbor (1975).
- 9) Intel Corporation, Willow Garage, <http://opencv.willowgarage.com/wiki/>
- 10) Li Fei-Fei, Rob Fergus and Pietro Perona: Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories, Workshop on Generative-Model Based Vision, IEEE CVPR 2004, Workshop on Generative-Model Based Vision (2004).
- 11) K. A. De Jong: An analysis of the behavior of a class of genetic systems, Dissertation Abstracts International, Vol.36, pp.5140B (1975).
- 12) L. J. Eshleman and J. D. Schaer: Real-Coded Genetic Algorithms and Interval-Schemata, Foundations of Genetic Algorithms, Vol.2, pp. 187-202, (1993).
- 13) schimatk, <http://code.google.com/p/opencvsharp/>