

動的な入力画像学習における遺伝的アルゴリズムを用いた低コストな領域検出システムの設計

定村, 勇典 / SADAMURA, Yusuke

(出版者 / Publisher)

法政大学大学院情報科学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 情報科学研究科編 / 法政大学大学院紀要. 情報科学研究科編

(巻 / Volume)

11

(開始ページ / Start Page)

1

(終了ページ / End Page)

6

(発行年 / Year)

2016-03-24

(URL)

<https://doi.org/10.15002/00012880>

動的な入力画像学習における遺伝的アルゴリズムを用いた 低コストな領域検出システムの設計

Design of Low-cost Adaptive Image Detection System Using Genetic Algorithms

定村 勇典*

Yusuke Sadamura

法政大学大学院情報科学研究科情報科学専攻

〒184-8584 東京都小金井市梶野町 3-7-2

Email: 10k1115@cis.k.hosei.ac.jp

Abstract—In recent years, the object detection including the face detection becomes one of the most prosperous fields in the computer vision. Along with it, highly precise detection algorithms and high-speed circuits with low electric power consumption are demanded by various camera systems. The most of mainstream object detection algorithms spend more time on learning general pattern of the objects in advance. However, there are many situations that cannot be detected by the prior leaning. In order to solve this problem, it is needed to prepare various detectors for special situations. But, for low-cost devices, it is not effective to detect all situations. In this paper, we propose an efficient object detection system adapting to input pattern of special scenes. Our method uses the kernel-based matching with low-resolution image to learn and detect object. This method optimizes kernels using Genetic algorithms for approximating the distribution function more precisely, then detects objects by superposition of optimized kernels. However the algorithm requires long time to optimize the kernel. Therefore we estimate some parameters to improve the learning time and detection rate by simulation. The simulation results show that these parameters are effective for getting the kernel to learn the image.

I. 序論

顔検出をはじめとする物体検出はコンピュータビジョンで最も研究が盛んな分野の1つである。また近年、多類のカメラデバイスの普及と共に大きな進歩を遂げており、様々なカメラシステムにおいては省電力で高速かつ高精度な検出アルゴリズムが求められている。検出には事前学習により対象物の一般的なパターンを識別し、検出する手法が有用とされ用いられている。しかし実際の問題として、ユーザーの撮影環境によっては、事前学習による検出が困難で複雑な処理を必要とする場合も多く存在する。例えば顔検出においては、サングラスやマスク、照明や肌の色などにより顔に特有な部品のパターン認識ができない場合がある。そのため、考えられる限りの状況に対してあらかじめ機械学習による数多くの識別器を用意することが必要となるが、その万能な検出アルゴリズムを小型機器などの低コストなハードウェアで実装することは、速度や電力コストの関係上容易ではない。

そこで本研究では、事前学習だけでは対象物の検出が困難な状況を解決することに着目し、カメラと入出力装置を備えた端末機器を用いての、動的で少ない学習データ入力による物体検出システムを提案する。このシステ

ムはその都度新しい物体を少ない時間で学習することを目的としている。検出には単純なテンプレートマッチングの代わりにカーネル密度推定に基づいたカーネルマッチング法を行う。この手法はカーネルのサイズを可変とし分布関数を正確に近似することで、より精度の高い検出が行える [1]。我々のシステムにおいてこの検出処理を採用する。理由は以下である。

- サンプルパターン毎にカーネルサイズを最適化することで、少ないサンプルパターン数でも発生確率分布を比較的精度よく近似することができるため高い認識精度 (検出精度) が得られる。
- 単純なカーネルを用いることにより複雑な計算が不要となり汎用プロセッサによるソフトウェア処理を用いても高速な処理が可能で、FPGA(Field Programmable Gate Array) 等の書き換え可能集積回路を用いればハードウェア化が容易である。

しかし、このカーネルマッチング法ではカーネルサイズの候補の組み合わせが指数的に増大するため、全探索をしてカーネルサイズの最適解を求めることが困難である。そこで学習を行う処理として、遺伝的アルゴリズムを用い、準最適解を取得する。しかし、カーネルの取りうる値などを考慮せずに遺伝的アルゴリズムによる学習を行う場合は、領域候補として挙がる領域数が無くなったり増えすぎたりしてしまう。そのため遺伝的アルゴリズムが収束するまでには時間がかかる。我々はある程度の検出率を維持しつつ学習時間を早めるためのパラメータを設定し、シミュレータにより、そのアルゴリズムの有用性を調べる。その際、精度を測るための評価実験用にいくつかの画像データ用意し、短い時間であっても効率的に最適解となるカーネルを生成できるかを確認する。

II. 先行研究

画像検出あるいは認識においては、幾何学的特徴を利用する方法と画像全体のパターンを用いる方法がある。

幾何学的特徴を利用する方法は、検出したい物体の画像において、その物体を構成する部品を抽出し、その位置関係や大きさなどの特徴を算出後、認識対象物との比較によって検出を行う方法 [4][5][6][7] である。例えば顔検出においては、固定の小領域に対して、あらかじめ定めた目・鼻・口などのブロック型の顔部品を抽出する空間フィルタを適用している。動画像での検出では、剛体運動のモデルや光学モデルなどの対象物の事前情報を組織的に利用して対象物モデルを作成し、各物体を識別

* Supervisor: Prof. Yamin Li

している。このような幾何学的特徴の機械学習による画像検出は、識別器を作成するために学習画像が数多く必要となることやその事前学習に時間がかかることなどの問題点がある。サングラス、マスクなどで特徴部位が隠れる状況は多種多様に存在し、この特殊な状況に対する学習データは少ない。そのため、その都度新しい物体として学習し識別させる場合は、柔軟に対応することが難しい。

一方、画像全体のパターンを用いて識別する方法は、このような物体に特有な部品の特徴を用いるのではなく、画像そのものをパターンとして扱う方法である。この方法は濃度値をそのまま特徴ベクトルにするマッチング処理などの方法と、濃度パターンを直交変換し、その変換係数を特徴ベクトルとして識別する方法に分類される。利用される直交変換として、離散フーリエ変換 (DFT)、離散コサイン変換 (DCT)、KL 変換 (主成分分析)、ウェーブレット変換などがある。これらは低次元で特徴を抽出したり、特徴となる成分の分布や形状からデータを高次元空間に投影し、似た物体の認識処理を行う際に多く用いられる。さらにこれらとニューラルネットワークなど他の認識手法を組み合わせるものもある。ニューラルネットワークによる画像の機械学習を行う際も、膨大な数の学習画像と学習時間を必要とする。そのため小さなカメラシステムにおいての新たな物体の学習は容易ではないと思われる。

画像のパターンでマッチングによる検出をする方法の中でも、テンプレートマッチングは画像信号に対して扱いやすい処理である。そのため古くから実用的な方法とされており、高速なステレオマッチング [8] を行う際に多く用いられている。テンプレートマッチングは特徴ベクトルとして濃度値データをそのまま用いて、あらかじめ登録しておいた小さな一部の画像領域と同じパターンが画像全体の中に存在するかどうかの比較を行い、統計的モデルを基に最も似通ったパターンを探す操作である。画像内にある対象物体の位置検出や、物体数のカウント、また探索範囲を限定して物体の追跡を行う際等に使われている。カメラから初めに得られる画像信号のほとんどは色濃度値であるため、直接生の色濃度値を特徴ベクトルとして扱う単純な処理であれば、ハードウェア化により高速な処理が望める。

テンプレートマッチングでは画像間の類似度を表す尺度に SSD (Sum of Squared Difference)、SAD (Sum of Absolute Difference)、NCC (正規相互相関、Normalized Cross-Correlation)、ZNCC (Zero-mean Normalized Cross-Correlation) 等がある。式 (1) の正規相互相関としてしられる NCC や ZNCC は照明変動に頑強であるが計算コスト高い。そのため、この部分の計算の簡略化や、ハードウェア化による高速化 [9] が図られている。またテンプレートマッチングでは大量の学習データとその計算時間が必要となるため、固有顔 [10] などによる、少ない画像でのマッチングが考えられている。しかし識別器となる画像の作成には、その学習データのために同じ照明条件でデジタル化された画像を多数用意し、特徴となる部分の位置合わせや解像度の再標本化をする必要がある。

$$S_{CNN}(x, y) = \frac{\sum \sum F(x+i, y+j)T(i, j)}{\sqrt{\sum \sum (F(x+i, y+j))^2} \sqrt{\sum \sum (T(i, j))^2}} \quad (1)$$

計算コストを抑えた他のマッチング手法として、パターン認識において多く用いられるカーネル法の密度推定で画像検出をするカーネルマッチング法がある。これはサンプルパターンの分布から各カテゴリの確率密度関数の推定を行う手法である。未知パターン X がカテゴリ A に属する確率 $p(X|A)$ をカテゴリに属する N_A 個のサンプルパターン $S_1^A, S_2^A, \dots, S_{N_A}^A$ を用いて推定したものは式 (2) で近似できる。

$$\tilde{p}(X|A) = \frac{1}{N_A} \sum_{i=1}^{N_A} K(X - S_i^A) \quad (2)$$

この式はサンプルパターンに与えられた一定サイズの核関数 $K(\cdot)$ の重ね合わせにより確率密度関数を近似することを意味している。カーネルを用いた移動体の追跡 [11] などでは複雑な特徴量や一般的なガウス関数の核関数を用いることがあるが、本研究のシステムでは計算量を減らすために核関数としては簡単に扱える式 (3) の一様関数を用いる。

$$K(\mathbf{x}) = \begin{cases} \frac{1}{2d} & (|x_k| \leq d \ (k = 1, 2, \dots, n)) \\ 0 & (otherwise) \end{cases} \quad (3)$$

d はカーネルサイズを表すパラメータである。つまりカーネルマッチング法とは各カテゴリに属するサンプルパターンのカーネルを包含する数を用いて、未知パターンが各カテゴリに属する確率を求め、その確率から未知パターンのカテゴリを識別する方法である。サンプルパターンを一つの点とするならばカーネルはその 1 点と周りの領域を意味する。なお今回の実験ではカーネルはサンプル画像の各画素濃度値の幅となる。

またカーネルマッチング法で実用的な検出精度を得るためには、数多くのサンプルパターンが必要となる。我々はこの問題を解決するため、各サンプルパターン毎に各カーネルサイズを可変としたカーネルマッチング法を用いる。しかしこの手法では、各画素の数 n 毎にカーネルサイズ d があり、 d の候補数を L 種類とした場合、1 つのサンプルパターンに対し L^n 個のカーネルサイズ候補の組み合わせが存在する。このように指数関数的に組み合わせが増加するため、最適解を全探索で求めることが困難となる問題が生じる。

計算機科学における組み合わせ最適化問題に対しては、進化的計算手法が有用とされる。上記のような準最適な組み合わせを得る場合は遺伝的アルゴリズム (GA, Genetic Algorithms) [2] を適用することができる。GA は生物の遺伝子による進化過程を工学的に模倣した近似最適化手法である。カーネルサイズを遺伝子にコード化し、進化的に準最適なカーネルサイズが求められる。GA による探索を制御する上では、収束時間や多様性の面から探索と活用のバランスを考慮する必要がある [2][3]。

検出システムに限らず、多くのカメラシステムにおいては高速な演算が必要とされることが多い。そのため様々な特徴量の計算手法、階層的な部分領域のマッチングによる演算量の削減、並列演算器による高速化あるいは対象とする画像処理のハードウェア化実装など、多くの手法が研究されている。

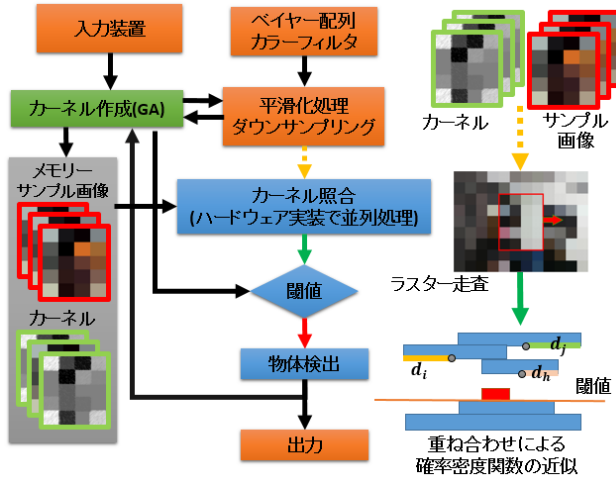


図 1. 提案システムの概要

III. 提案手法

A. 入力画像検出システムの概要

図 1 にこのシステムの概略を示す。まず対象物体の位置の入力を行い、そのサンプルパターンの濃度平均値で低解像度にモザイク処理するとともに、パラメータ設定のデータを保持する。次に GA で準最適化させたカーネルを用意する。GA では入力された領域を用い、次元数に応じて遺伝子を初期化後、評価関数を基に選択と交叉を繰り返すことで準最適解を求める。遺伝子の評価を行う際は、遺伝子の値からカーネルサイズを求め、入力が行われた際のモザイク画像にカーネルマッチング法を行う。同様にして複数のサンプルパターンを学習させる。学習後の検出処理には、得られたカーネルを使い、モザイク処理された入力画像をラスタ走査していく。対象画像とカーネルとの照合で重ね合わせによる確率密度関数の近似を行い、得られた値がある閾値を超えたら対象物と識別する。最後に得られた検出領域を出力する。これを継続的に実行させ、動的な学習と画像マッチングによる検出を行う。本研究では効率的な GA の作成や撮影環境による検出精度の調査のため、各パラメータを設定できるシミュレータを作成して評価実験を行った。

カーネルマッチング法は、各濃度値がカーネルサイズ内に含まれた数をカウントし、閾値との比較で検出を行う。そのためこの部分に関しては単純な組み合わせ回路として表現でき、並列な専用ハードウェアとしての実装が容易である。これにより入力画像と各カーネルの照合は並列させて実行することが可能で、検出処理にかかる時間はサンプルパターン数に依存しないで検出が行える。ただし集積回路のサイズには制限があるため、数多くのサンプルパターンによる識別をする場合は、その回路規模を考慮しなければならない。FPGA などの再構成可能な集積回路であればカーネル数を増やしての実装も可能とされる。また動的再構成可能な進化ハードウェアであれば遺伝的アルゴリズムを FPGA により回路化したものも多く報告されており [12][13]、FPGA による並列 GA のフレームワークにこのシステムの実装が期待できる。

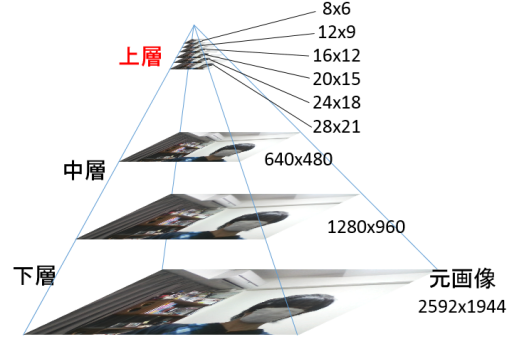


図 2. 画像のピラミッド構造化

B. 画像のピラミッド構造化による領域検出と物体の大きさへの対応

対象画像の画素数が大きくなると演算量もそれに比例して増大するため、高速化をするためには実質的な演算量を低減する必要がある。我々の研究では対象画像とカーネルの照合を行うために図 2 に示すような画像のピラミッド構造化 [7][14] をする。

まず平均濃度値を持つ画像を階層的に用意しておき、始めは低解像度の画像を用いて高い解像度は必要としない大まかな探索を行う。そして検出された部分領域の周辺を更に中層や下層の高い解像度で再探索し、詳細な認識や判定の識別を行う。全探索に近い探索ができ、演算量を大幅に低減することが可能となるため、画像のピラミッド構造化は多くの検出システムで用いられている。例えば画素数を 640x480 から縦横半分の 320x240 にすると演算量は 1/4 になり、また縦横 1/32 の 20x15 にすると演算量は 1/1024 に低減できる。

我々のシステムでは上層の低解像度画像を用いて高速に対象物の領域検出する。検出後の処理としては、特徴点を抽出や幾何学的特徴量を用いた教師付き分類器による、より詳細な識別処理が考えられる。

低解像度画像を平滑化処理で作成する場合、平均濃度値のためノイズによる影響や不均一な背景にも対応できる。また対象物のサイズ変化にも対応することができる。すなわち全ての入力画像とそのカーネルの画素数を固定することで、大きな物体は低解像度の対象画像で検出され、逆に小さい物体はより高い解像度で検出される。検出する画素数が全て同じ大きさになることで、カーネルがほぼ同じ回路規模で実現できハードウェア化が容易になる。

C. カーネルマッチング法による画像検出

テンプレートマッチングは実用的な対象画像検出手法であるが、距離計算が複雑であり、容易に高速化できない点やハードウェア規模が大きくなる点が課題である。

一方、カーネルマッチング法はカーネルと呼ばれる規則の要素を重ね合わせることで全体の規則を形成することができる。カーネル法はパターンの識別において多く用いられており、そのカーネル密度推定はサンプルパターンの分布から各カテゴリの確率密度関数の推定を行う手法である。判別する際にカーネル法とよく組み合わせて用いられるサポートベクターマシンは、カーネル法により拡張された高次元特徴空間の非線形データ構造を線形構造に変換し、それらのカテゴリを分類する。この

カーネル法を取り入れたデータ解析手法は、主成分分析や正規相関分析、クラスター分析やk平均法など多くのアルゴリズムが存在する。

しかし本研究では検出を主としており、データ解析などを用いず、識別には複雑な計算を必要としない。そのため低コストなハードウェア実装が容易で、ある程度の精度と高速処理が期待できる。よって本研究の少ない入力画像データによる画像マッチングの検出手法ではカーネルマッチング法を用いる。

さらに、カーネルマッチング法におけるカーネルサイズをサンプルパターン毎に可変とする拡張カーネルマッチング法 [1] を適用する。これは登録しておくサンプルパターンの画素毎にカーネルサイズを可変なものとしておく手法である。最適なカーネルサイズであれば、一定値で固定したカーネルマッチング法に比べ、少ないサンプルパターン数でも、発生確率分布を比較的精度良く近似させた検出を行うことができる。

カーネルマッチング法では、サンプルパターンの各画素濃度値に幅を持たせたカーネルをラスタ走査させていく。照合した対象画像の部分領域がカーネルの画素濃度値に含まれたか否かを調べていく。そして本システムでは式 (3) の一様関数を用いるため、包含されたカーネルの数がある閾値以上であれば検出領域の候補と判断する。この閾値による判断により画像中に複数の対象物が存在する場合であっても、同時に検出することができる。

D. 遺伝的アルゴリズムによるカーネルサイズの生成

個々のカーネルサイズを変更する拡張カーネルマッチング法では最適な組み合わせを求める必要があるが、高次元のパターン空間であるため解候補数が膨大になり、全探索することは計算上困難である。そのため最適解を得るために GA を用いる。進化的計算手法は GA 以外に進化戦略や進化的プログラミングがあるが、これらは適応度の評価回数が増える点や突然変異に複雑な処理を用いる点などの理由から使用せず、本研究では GA の中でも特に複雑な計算を必要としない単純 GA (SGA: Simple Genetic Algorithms) を用いる。

GA では取り扱う問題のパラメータが染色体に割り当てられ、染色体集団のうち各々の染色体には評価計算のための適応度が与えられる。その適応度の優れた染色体を次世代に残すように選択、交叉、突然変異を繰り返し、生き残った染色体が問題空間における最適解となる。終了条件は、適応度が最大値などの閾値に達したとき、または世代交代数があらかじめ決められた回数に到達したときなどが考えられる。GA の長所を活かすためには、解くべき問題に対して適切に GA を設計し、パラメータを適切に設定する必要がある。

図 3 に拡張カーネルマッチング法で用いるサンプルパターンと GA における染色体や遺伝子との関係を示す。まず GA でカーネルサイズの最適解を求めるために染色体集団を用意する。各染色体は各カーネル、染色体のもつ n 個の各遺伝子は n 次元パターン空間上の各次元のカーネルサイズにそれぞれ対応している。また各遺伝子の取りうる値は L 種類とし、あらかじめ L 個のカーネルサイズ値を用意する。8bit の色濃度値とする場合は 0 から 255 を L 個に分割した値を取りうることにする。この L を大きくすることでほぼ連続的にカーネルサイズを変化させることができ、より最適なカーネルサイズ

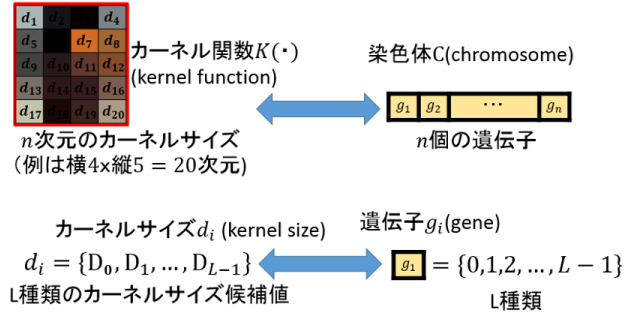


図 3. カーネルと遺伝子型の対応関係

を求めることが可能となる。また GA では染色体を進化させる上で、適応度関数を用意する必要がある。これは各々の生存確率である適応度を評価するものであり、本実験では式 (4) の適応度関数 $F(C)$ を用いる。

$$F(C) = \frac{1+a}{1+s} \quad (4)$$

a は対象物の領域の画像を含んだ回数、 s は候補として領域を検出した回数としたとき、この評価関数は検出された領域候補が正解領域であれば適応度を高くし、間違っていれば適応度を下げる評価関数となる。正解領域が入力された際の画像を学習用の対象画像としその各低解像度画像と染色体であるカーネルを照合させ領域候補を検出する。

本研究の検出システムでは 1 つのサンプルパターンにおけるカーネルの数である次元数を入力画像の縦横比で変動できるようにするため、GA の選択には適応度に比例して選択を行うルーレット方式を用いる。

IV. シミュレータの作成と効率的な学習を行うためのパラメータ設定

我々の検出システムをシミュレーションし、より適した検出アルゴリズムを検討するためにソフトウェア実装を行う。ウェブカメラを用い、GUI の操作で動的に画像領域の入力を受け取り学習させ、選択した検出方法でその候補領域を出力する。評価実験の際は同じ解像度の画像データと正解領域のリストを読み込み、検出率と実行時間も出力する。本研究で作成したシミュレータとその実行例を図 4 に示す。

検出手法としては正規相互相関を用いたテンプレートマッチング、カーネルサイズを固定したカーネルマッチング法、拡張カーネルマッチング法を選択できる。拡張カーネルマッチング法では、GA の最大世代数や染色体数や突然変異発生率などのパラメータを用意する。また正解領域の判定方法の選択でき、1 つのカーネルの解像度数やカーネルマッチング法のカーネルサイズとその検出精度を変動させられる。その他ヒストグラムの表示や画像に対して様々な濃度値変更処理が行える。

正解の判定は、検出された領域の中心座標が正解領域の範囲内であれば正解とする。複数の領域が検出されるカーネルマッチング法においては、その正解数が不正解数を上回ったならば正解とした場合の模擬実験は高い検出成功率を示した。またこの判定では領域サイズが綺麗に合わない検出候補も存在することになるため、より正確に領域をとらえているものを判定する場合に、検出領

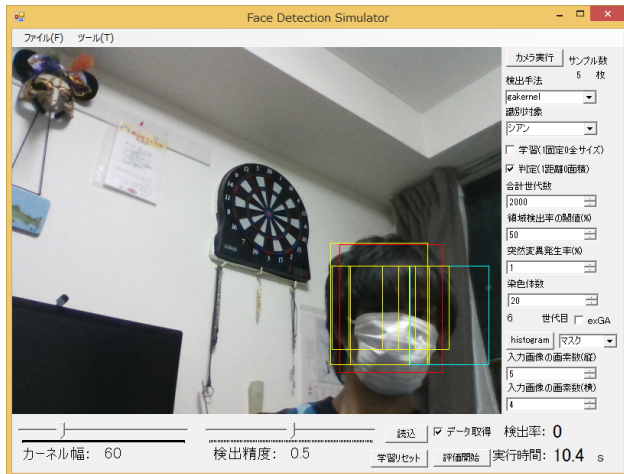


図 4. シミュレータ実行画面

域の面積と実際の正解領域の面積が半分以上重なりかつ正解領域の面積の2倍を超えないものを答えとする判定や領域を表す2点の近隣位置による判定も行う。

シミュレータでは、まずユーザーの入力で正解となる画像の位置とその範囲を入力する。その際、あらかじめ用意した画像データと正解領域のリストを読み込むことが可能である。次にピラミッド構造化に対応させ、正解判定をもとにGAを動作し最適解となるカーネルサイズを求める。その後、選択した手法で検出処理を行い、検出された領域を画面上に表示させる。また評価実験ではあらかじめ用意しておいたデータを読み込み、正解判定をもとに検出率とその実行時間を表示させる。

GAによる探索を制御する上では探索と活用のバランスが重要になるため、多様性や収束時間の面を考慮した初期集団や選択圧に関わる複数のパラメータを用意しGAにおける効率的な学習を行う必要がある[2][3]。例えば時間をかけた事前学習による拡張カーネルマッチング法の設定では、少ない世代交代数の場合にうまく選択圧がかからず、不適切に終了条件を迎えることが多く発生する。そこで本研究では初期集団や選択圧に関わるパラメータを用意し、シミュレーションにて多様性や収束時間の面を考慮した適切なGAの設計を行う。これによりGAの収束が適度に早まるようにする。

まずカーネルの取りうる値に上限値と下限値をパラメータ設定する。カーネルサイズが固定された従来のカーネルマッチング法のシミュレーションから、カーネルサイズの値が極端に0や255に近づく検出候補数が無くなったり増大しすぎたりすることが分かり、多くの遺伝子がこのような極端な値を持つ場合はその収束に時間を要してしまうことが考えられる。また拡張カーネルマッチング法で長い時間をかけ作成されたカーネルにも極端な値が入っていないことが多い。本研究では検出精度である閾値を0.5に固定して、検出を繰り返し評価実験を行い値を調べた結果、カーネルサイズの取りうる値の多くは64を中心とした前後の値であることが確認できた。そのため我々のシステムではカーネルには64前後の範囲内を候補数 L で分割した値を持たせ、この範囲内での最適解に近づけることとする。

次に世代数を用いて閾値を前後させるパラメータを設定する。このパラメータ設定は多様性を無くさずに、全

体の収束を早めることを目的とする。SGAでは初期集団を乱数で生成するため、初めの世代の多くの遺伝子で検出候補数が無くなることや検出候補数が増えすぎることが生じてしまう。これでは正解領域を選ぶ確率が極端に少なくなり収束が遅くなりうる。収束を早めるために選択圧を強めた場合は、適応度が高いものが多く次世代の遺伝子に残されるため、本来の最適解とならずに収束を迎えることになってしまう[2]。特にこのシステムのGAにおいては交叉率が上がると最適なカーネルを作らずに終了することがある。そのためGAにおいて遺伝子の多様性を保つためには選択圧が強くなり過ぎないことが必要である。この状況を解決するには、初めの世代では候補数を減らしつつ、世代を増す毎に収束し過ぎない値にとどまることが望ましい。そこで各染色体の適応度の平均が低い場合(その際の候補数を多様性度合い v とする)や候補となる領域が1つもない場合には、各世代数 g ごとに検出領域判定となる閾値をパラメータ p を用いた $1/pg$ で加減算をする。これにより世代数の増加と共に閾値の変動値が減り、選択圧による多様性の消失を行わずに、適宜検出候補数を増減させることができる。

V. 実験結果

GAの評価のためにシーンを固定し、マスクと前髪で顔の部品が多く隠れる場合、屋内の逆光での顔画像検出、信号機の歩行者用押しボタン、玄関の靴、調味料の入れ物、冷蔵庫内のモノの検出を想定したテストデータを用意し、検出率と実行時間(Intel®Core 5-4300U™1.90GHzのプロセッサを使用した場合の1枚の未知画像を調べる際の平均時間)を調べ、撮影環境による検出率の良し悪しを調べた。これらのうちマスクと前髪で顔の部位が隠れる場合では、カーネル作成に使用するサンプルパターンを5枚、それとは異なる未知画像54枚(内5枚対象物無し)を用いた。カーネルは横4ピクセル×縦5ピクセルの20次元で、それぞれピラミッド構造化による6種類の低解像度画像324枚で検出を行った。検出方法にはテンプレートマッチング、値を固定したカーネルマッチング法、収束時間を考慮しない従来の拡張カーネルマッチング法、今回作成したGAによる収束時間を考慮させ改良した拡張カーネルマッチング法を用いた。テンプレートマッチングでは式(1)を用い、その値が最大値となる点が適度に設定した閾値を超えた場合に検出領域の候補とする。この閾値は物体が写っていない画像にて検出候補がなくなるように設定した。

カーネルマッチング法と拡張カーネルマッチング法ではカーネルと照合する際の閾値を全カーネル数の半分の割合に設定し、その値を超えたら検出領域の候補とする。正解の判定は、検出された領域の中心座標が正解領域の範囲内であれば正解領域とし、複数領域が検出されるカーネルマッチング法においては、その正解領域数が不正解領域数を上回ったならば正解とした。また v は低解像度画像の数の2倍に設定し、各染色体の平均検出候補数が0となった際は次世代の閾値を $1/pg$ 減らし、 v を超えた際には $1/pg$ 増やした。閾値を固定し最適解を求めた場合と、閾値を変動させて最適解を求めた場合のGAの実行時間やその正解率を比較する。この実験で用いたGAのパラメータを表Iに示す。この実験での検出率及び実行速度を表IIに示す。

表 I
用いた GA のパラメータ

個体数	20
世代数	1000
交叉確率	0.6
突然変異率	0.01
カーネルサイズの候補数 L	16
領域候補とする閾値	0.5
カーネルサイズ上限値	96
カーネルサイズ下限値	32
収束の度合い p	2
多様性の度合い v	12

表 II
実行結果

検出手法	検出率	検出時間 (学習時間)
テンプレートマッチング	42.6%	5.8ms
カーネルマッチング法 (KM 法)	1/16 検出点なし	4.7ms
	2/16 13.0%	4.7ms
	3/16 61.1%	4.7ms
	4/16 51.2%	4.7ms
	5/16 35.2%	4.7ms
	6/16 22.2%	4.7ms
	7/16 7.0%以下	4.7ms
従来の拡張 KM 法	81.4%	4.7ms(53s)
提案手法による拡張 KM 法	77.8%	4.7ms(6s)

評価実験により、少ない画像による学習であっても適度にパラメータ設定をした GA により、少ない時間で学習や検出を行うことができ、テンプレートマッチングの結果に対しても高い検出性能が得られた。照明が一定の向きで当たっている際などの検出には高い精度が得られたが、しかし評価したシーンの内、色の検出や輝度値での評価が難しいとされる局所的な明度変化や隠蔽のある画像が学習画像や検出対象となった場合には、検出性能が下がりカーネル作成に多少時間を要することがあるため、その点の改善を行わない限り全ての撮影環境下で優れているとは言えない。また今後収束させたい時間と精度のバランスをみて手動でパラメータ設定する必要がある点や検出候補として挙げられた領域が重なっている場合を考慮したクラスタリングによる分類で候補領域を1点とする必要がある。また学習に用いる画像や GA の乱数により学習時間が変動するため、精度を上げつつ学習時間を平均的にするように抑える工夫も必要となる。

VI. 結論

本研究では事前学習では検出困難な状況に対して、小さなデバイスであっても動的な少ない入力画像による対象物学習と検出が行えるシステムを提案した。検出手法で用いる拡張カーネルマッチング法をより短い時間で効率的に学習させるため、GA のにパラメータを設定し、適した値を与えることで有効なカーネルが得られることを評価実験により確かめた。また複数の領域の検出も可能であり撮影環境下によってはその検出精度も高い値が得られた。しかし、照明環境の変化に対してはあまり良い性能とならないこともあるため、カーネルに用いる値をさらに検討していく必要がある。また今後の課題としては入力デバイスやカメラを用いて提案手法をハードウェア実装することや、対象物のクラスタリングによる

詳細な認識を他の手法と合わせた形での実装を行うことが挙げられる。

参考文献

- [1] K. Takashima, T. Nakamura, M. Yasunaga, and I. Yoshihara, "Face detection in images using genetic algorithms and its implementation onto an evolvable hardware systems," *Technical report of IEICE. PRMU*, vol. 100, no. 443, pp. 71–78, nov 2000. [Online]. Available: <http://ci.nii.ac.jp/naid/110003275206/en/>
- [2] 棟朝雅晴, 遺伝的アルゴリズム : その理論と先端的手法, pod 版 ed. 東京, Japan: 森北出版, 2015.7 2015.
- [3] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, 1st ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1989.
- [4] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1, June 2005, pp. 886–893 vol. 1.
- [5] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004. [Online]. Available: <http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94>
- [6] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vision*, vol. 57, no. 2, pp. 137–154, May 2004. [Online]. Available: <http://dx.doi.org/10.1023/B:VISI.0000013087.49260.fb>
- [7] C.-S. Chen, S.-H. Tsai, K.-Y. Peng, C.-T. Lin, and C.-C. Wen, "Image alignment using pyramid structure with harris corner detection," in *Robot, Vision and Signal Processing (RVSP), 2013 Second International Conference on*, Dec 2013, pp. 200–203.
- [8] 岩田健司, 中村良介, 田中良夫, 増田知記, 町田亮介, 小島功, and 関口智嗣, "異なるアーキテクチャのメニーコアプロセッサにおけるステレオマッチングプログラムの高速化と性能評価," *情報処理学会論文誌コンピューティングシステム (ACS)*, vol. 3, no. 3, pp. 178–188, sep 2010.
- [9] S. MURAMATSU, Y. OTSUKA, Y. KOBAYASHI, and E. SHIMIZU, "Strategy of high speed template matching and its optimization by using ga," *The transactions of the Institute of Electronics, Information and Communication Engineers. D-II*, vol. 83, no. 6, pp. 1487–1497, jun 2000. [Online]. Available: <http://ci.nii.ac.jp/naid/110003223934/en/>
- [10] O. N. A. Al-Allaf, "Review of face detection systems based artificial neural networks algorithms," *CoRR*, vol. abs/1404.1292, 2014. [Online]. Available: <http://arxiv.org/abs/1404.1292>
- [11] U. Devani, V. Nikam, and B. Meshram, "Super-fast parallel eigenface implementation on gpu for face recognition," in *Parallel, Distributed and Grid Computing (PDGC), 2014 International Conference on*, Dec 2014, pp. 130–136.
- [12] T. Tachibana, Y. Murata, N. Shibata, K. Yasumoto, and M. Ito, "Flexible implementation of genetic algorithms on fpgas," in *Proceedings of the 2006 ACM/SIGDA 14th International Symposium on Field Programmable Gate Arrays*, ser. FPGA '06. New York, NY, USA: ACM, 2006, pp. 236–236. [Online]. Available: <http://doi.acm.org/10.1145/1117201.1117264>
- [13] K. Zhang, "A competitive reconfiguration approach to autonomous fault handling using genetic algorithms," Ph.D. dissertation, Orlando, FL, USA, 2008.
- [14] S. Jin, D. Kim, T. T. Nguyen, D. Kim, M. Kim, and J. W. Jeon, "Design and implementation of a pipelined datapath for high-speed face detection using fpga," *Industrial Informatics, IEEE Transactions on*, vol. 8, no. 1, pp. 158–167, Feb 2012.