

リアルタイム処理可能な汎用画像認識検出エンジン

1. 背景

近年のパーソナルコンピュータの処理速度の大幅な向上は、画像処理のような計算時間のかかる処理を、安価なパーソナルコンピュータでも高速に処理することを可能とした。本プロジェクトでは、特に画像認識、検出処理を高速に処理する手法を開発し、それによるリアルタイム処理可能な画像認識、検出処理の達成を目指す。

開発の概要は、開発者が以前に開発した Equivalent Retina Sampling(ERS)と呼ばれる手法にいくつかの改良を行い、c++言語を使って高速処理可能な形でインプリメントする。この結果、顔認識における肌色認識のような特定の対象に特化したヒューリスティックによらない対象検出をめざした。アルゴリズムの詳細は文献[1][2]に記述されている。

2. 目的

我々は文献[1][2]において、高精度でなおかつ高速に処理できる可能性のある対象検出手法を提案した。しかし、この手法はMATLABという非常に処理の遅いプログラムで記述されており、高速性の証明は不可能であった。そこで本アルゴリズムをC++言語でインプリメントすることで提案したアルゴリズムの高速性を証明するのが目的である。

3. 開発の内容

(開発したソフトウェアの動作環境、構成、機能等を図等を使用して記述)

近年、画像処理によって物体を検出する手法が多数研究されている。多くの手法は検出窓を動かしながら検出対象を発見するという手順を基礎としている。単純化するならば

$$(\text{画像が検出対象である度合い}) = P_{\text{subimage}}(\text{画像})$$

という関数を、様々なsub-image に適用することで対象を検出をする手法である。本プロジェクトでは、評価関数

$$\begin{aligned} &(\text{検出対象が位置 } (x, y) \text{ に大きさ } s \text{ で存在する度合い}) \\ &= P_{\text{entire}}(x, y, s, \text{入力画像全体}) \end{aligned}$$

を定式化し、この評価関数に基づく次の対象検出方法をインプリメントした。

$$\text{objectPosition} = \arg \max_{x, y, s} P_{\text{entire}}(x, y, s, \text{入力画像全体})$$

この定式化の結果、検出対象が入力された画像 I の位置 (x, y) に大きさ s で存在する確率 $P_{\text{entire}}(x, y, s, I)$ は x, y, s についての多項式と指数関数、対数関数の組合せで記述できるようになる。評価関数 $P_{\text{entire}}(x, y, s, I)$ は下図のように定式化する。まず、学習画像 $\{T1, T2, T3, \dots\}$ の特徴分布を混合ガウス分布で近似する。こ

の近似された結果を入力画像 I 上でサンプリングされた特徴と比較する。このとき適切な窓関数 Λ を使って適切でないサンプリング結果を無視する。

図1が、本プロジェクトで開発したソフトウェアの全体図である。

本プロジェクトでは主に2つのソフトウェアを開発する。1つは、認識、検出したい対象の特徴を学習アルゴリズムにより抽出するソフトウェア(図1上側の混合ガウス分布で特徴分布を近似する方法)。もう1つは抽出した特徴をつかってリアルタイムに対象検出するソフトウェアである(図1の下側)。

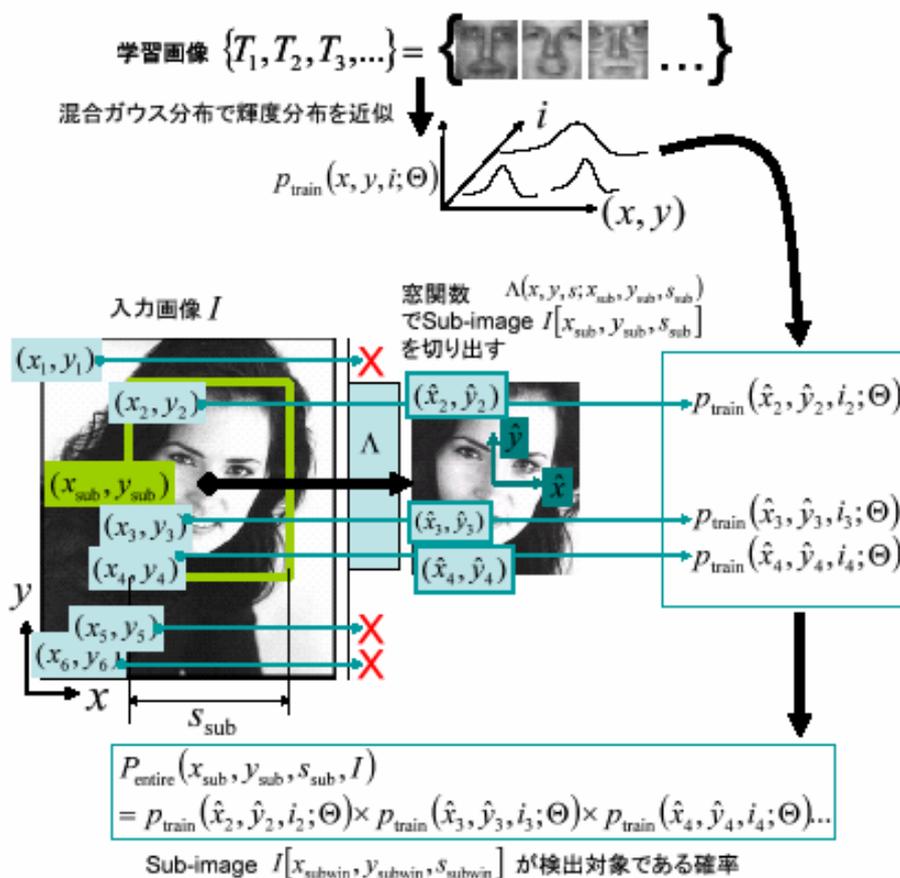


図1 提案システム全体図

4. 従来技術(または機能)との相違

(開発したソフトウェアの新規性、類似のソフトウェアと比較した場合の優位性等を記述)

本プロジェクトと同様にリアルタイムで対象検出を行える手法として Boosted Cascade[3]と呼ばれる手法がある。この手法と比べて、本プロジェクトの手法には次の利点がある。

1. 対象の大きさ(s)と角度()も高速に推定

Boosted Cascade は位置(x,y)のみしか推定できない

2. Boosted Cascade が見逃してしまう、凹凸の少ないパターンも検出可能

対象の大きさ、角度の推定が必要な場合には、従来手法よりも高速な処理を実現できると考えられる。

5. 期待される効果

(開発されたソフトウェアにより、波及的に活性化される技術分野・産業分野等の範囲やその効果を可能な限り具体的、定量的に記載)

画像検出したい対象が回転することが、困難さの原因になっている問題は多い。例えば農業の分野では、花粉の個数カウントを画像処理で行うことで、農作物の病気を予防する研究などが行われている。この問題では、検出器のとった画像を解析する時に、花粉が様々な方向に回転して写るために困難な問題となり、まだ実用化されていない。このような従来実用化困難な問題の解決が可能となる。

6. 普及(または活用)の見通し

(開発成果に関する利用者の具体的なイメージ[例えば、利用者数など]を可能な限り定量的に記載)
提案手法の特許が取得できたい、論文等で開発した手法の公開を行うことを検討している。

7. 開発者名(所属)

* 新妻 弘崇(産業技術総合研究所)

(参考)開発者URL

<http://www2s.biglobe.ne.jp/~niitsuma>

開発したプログラムのインターフェイスに当たる部分のソースコードを以下で公開

<http://www2s.biglobe.ne.jp/~niitsuma/STLLCV/>

(対象検出プログラムの主要部分は公開していません。)

参考文献

- [1] H.Niitsuma:“A Non-Parametric Trainable Object Detection Model Using a Concept of Retinotopic Sampling”, accepted to International Journal of Computational Intelligence and Applications (2003)
- [2] H.Niitsuma: “Trainable Object Tracking using Equivalent Retinotopical Sampling”, 7th International Conference on Knowledge-Based Intelligent Information & Engineering Systems, LNAI, Springer-Verlag(2003).
- [3] Paul Viola and Michael J. Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. IEEE CVPR, 2001

[記載要領]: 開発者の方々に問題のない範囲で下記要領にしたがって記載下さい。

書式: A4 (縦置き) で3 ~ 4枚程度、フォント(12ポイント、ゴシック)、

余白(上下左右とも25mm)、文字数等(37文字、38行)、

ページ数(1/1、1/2・・・とし、センタリング設定)。

ただし、テーマ名及びサブタイトルのみ14ポイント、センタリング設定。

テーマ名: 契約名を記載。

サブタイトル: 20字以内で記載(成果のイメージをつかんでもらうためのもの)。

1. ~ 6. までは提案書等をベースに記載。

7. 開発者名: 開発者名に加え、()書きで問合せ先となる所属等を記載。

なお、開発パートナー(会社)についても全て記載。

成果物を紹介する、または成果物に関連するURLがある場合には、URLを記載。

図表: 原則、図表は本文で引用した箇所に近い場所に添付(原稿末尾にまとめて添付することは避けてください)。

図表の前後には、空白行を1行設け、「表」のキャプションは表の上に、「図」のキャプションは図の下に記載。図番号・表番号はそれぞれの通し番号として記載。

ファイル名: 「契約番号(詳細)」をファイル名として保存。ex: 2004 情財第100号(詳細)

* 参考: 平成15年度成果報告集 <http://www.ipa.go.jp/about/jigyoseika/index.html>