

遺伝的プログラミングによる動画画像処理アルゴリズムの獲得

Automatic Construction of Video Image Processing Algorithm by Genetic Programming

小川 原也*¹
Genya Ogawa

喜瀬 勝之*¹
Katsuyuki Kise

長尾 智晴*²
Tomoharu Nagao

*¹ 富士重工業株式会社スバル技術研究所

Fuji Heavy Industry Ltd. Subaru Technical Research Center

*² 横浜国立大学大学院環境情報研究院

Faculty of Environment and Information Sciences Yokohama National University

We studied evolutionary dynamic image processing, using genetic programming. And we applied the pedestrian detection from video images in the scene with motion. The detection algorithm is automatically constructed by preparing the input and target images.

1. はじめに

近年、ITS(高度道路交通システム)の研究・開発が、盛んに行われているが、そのひとつとして、歩行者認識に代表される車載センサによる車外環境認識が挙げられる。車外の環境を認識することで、ドライバーが見落とした歩行者や障害物に対して、警告を発したり、衝突を回避し、事故を防ぎ、また、走行制御に認識情報を取り入れることで、より安全で快適な走行、さらには自動運転も可能になると考えられ、その意義は非常に大きい。

車載センサとして CCD カメラを用いた画像認識はこれまでも多くの研究が行われてきた。([Saneyoshi92][Motizuki04]など)画像センサは、画像自体が多くの情報を含むことから、非常に有用であるといえるが、屋外で自分自身が移動している状況で用いる場合、天候や、走行状態、走行環境など様々な条件に影響を受けるため、認識のロバスト性の確保が問題となっている。これまでは、それぞれの状況に対応したチューニングや、処理アルゴリズムの構築により、その適用範囲を拡大してきたが、すべての状況に対応する認識処理を構築することは開発コストの面も含めて大きな課題になっている。

このような状況に対して、我々は、種々の与えられた条件を学習することにより、自動的に認識処理を生成する手法の開発を行った。

画像処理アルゴリズムを自動構築する技術として、木構造状画像変換自動生成システム(ACTIT)がある。[Aoki99][Nagao02]これは、画像処理を画像フィルタを要素とした単位で捉え、その組み合わせを木構造で表現し、組み合わせを遺伝的プログラミングにより学習的に獲得するというものである。今回、このACTITを動画画像処理に拡張して(ACTIT-motion)、画像上の歩行者領域を抽出するシステムを開発した。

2. ACTIT-motion

歩行者領域を抽出する処理として、入力画像をFig.1のようなグレイ画像とする。そして処理の目標として Fig.2 に示すような画像上の歩行者領域を切り出した 2 値化画像とする。(処理は階調値画像で行う)画像処理は木構造の終端ノードに入力画像が入り、入力画像は終端ノードから1入力、2入力フィルタにより変換され、ルートから出力画像が出される構造とする。(Fig.3)ここで、画像フィルタは平均値フィルタ、ソベルフィルタ、2 値化フィルタといった基本的なフィルタを要素とする。

次に動画画像情報を扱うために終端ノードには現在の時刻の画像(画像 T)の他に過去の画像(画像 T-n)を入力することが出来るようにした。そして、遺伝的プログラミングにより学習で木構造の形状、終端ノードの入力画像の種類、中間ノードの画像フィルタの種類のを組み合わせを決定する。



Fig.1 入力画像 T



Fig.2 目標画像

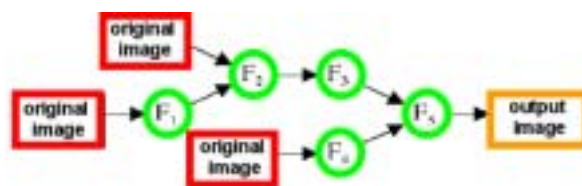


Fig.3 木構造処理

進化計算における適応度は出力された画像の目標とする画像に対する類似度として式(1)で定義する。

$$E = 1.0 - \frac{1}{N} \sum_{frame} \frac{\sum_{pixel} W \cdot |O - T|}{\sum_{pixel} W \cdot V} \quad (1)$$

E: 適応度 N: 学習セット数 V: 最大階調値
T: 目標画像 O: 出力画像 W: 重み画像

ここで、W は重み画像と呼ばれ、目標とする画像内での領域の重要度を表している。歩行者領域の抽出では、抽出したい領域を抽出するという目標と、抽出すべきでない領域を抽出しないという目標の重要度によって設定される。今回の実験では、それらの目標が、同程度の重要さを持つと考え、抽出、非抽出領域の面積比×重みが等しくなるように重み画像を設定している。

2. 実験

ACTIT-motion による歩行者抽出の実験を行った。入力画像は T~T-3 の 4 フレームとし、学習サンプルは 7 セット用いた。

連絡先: 小川原也、富士重工業株式会社スバル技術研究所、
〒181 8577 東京都三鷹市大沢 3-9-6 Tel:0422-33-7454 Mail: ogawag@tky.subaru-fhi.co.jp

進化戦略は MGG アルゴリズムを用い、個体数 100、生成する子個体は 50、交差率は 1.0、突然変異率は 1 個体で突然変異が起こる確率を 0.9 とした。また、木構造の大きさの制限として中間ノード数 40 以内に設定している。進化計算の終了条件は 5000 世代とした。世代における適応度の結果を Fig.4 に示す。

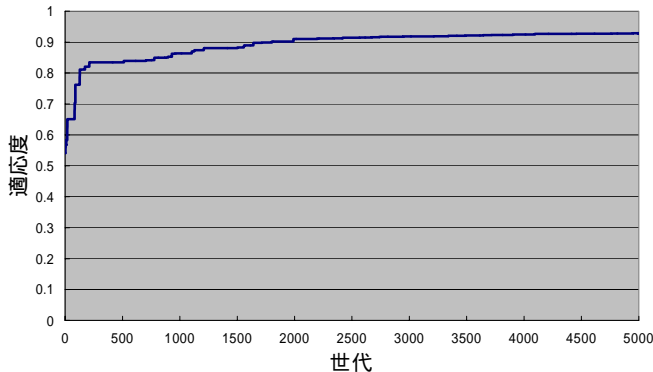


Fig.4 進化過程

終了世代でのエリート個体の学習セットでの出力結果を Fig. 5 に示す。



Fig.5 出力結果

4. 考察

上記計算を初期個体を変更して複数回行った。それぞれの試行でまったく同様の木構造は生成されなかったが、多くの木構造で Fig.6 のような基本構造が見受けられた。

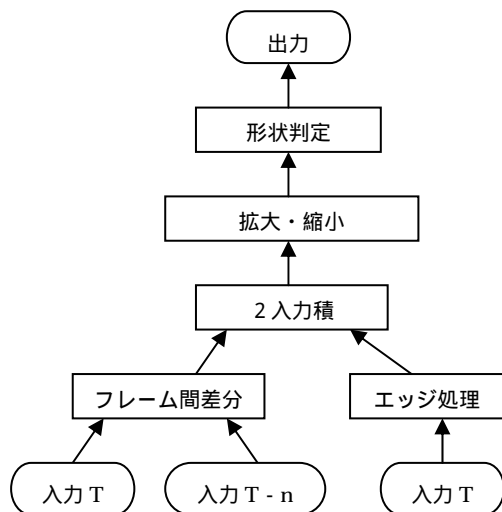


Fig.6 獲得した木構造の基本構造

このことより、今回与えた歩行者抽出目標では、フレーム間差分による動きの情報、エッジの情報が重要であるという知識の学習が出来ていると考えられる。

次に、複数回の試行でのそれぞれの木構造の出力については、歩行者領域はどの木構造でも抽出が行われていたが、歩

行者以外の抽出すべきでない領域も抽出されることが多かった。これは、歩行者に共通する特徴については、十分な学習が行われたが、歩行者以外に共通する特徴は十分学習できていないことを示していると考えられる。この要因としては今回与えた学習セット数が少なかったことが大きいと思われるが、歩行者以外というくくりは非常に広範囲で多様な対象を含むため、それらすべてを網羅して学習させることは、実際問題として難しい。

ここで、複数回の試行の結果に注目すると、どの木構造でも歩行者領域はほぼ抽出が行われているという傾向のほかに、歩行者領域以外を抽出してしまった領域はそれぞれの木構造で異なる場合が多いということがわかった。そこで、結果画像として、それら複数の木構造からの出力を重ね合わせ、画像上それぞれの画素で抽出されている木の本来の割合を表示した画像を作成した。歩行者領域は多くの木構造で抽出されているので高い輝度値になるが、歩行者以外の領域は輝度値が低くなる。6 本の木構造の出力を重ね合わせた画像を Fig.7 に示す。

上が学習を行った画像、下が未学習画像での結果である。画像は元の画像に結果の画像を赤色で重ねて、透過するように作成している。赤色が濃いところほど結果画像で輝度値が高くなっていることを表している。



Fig.7 入力画像と重ねあわせ後の出力結果

5. まとめ

車載カメラから得られる動画画像情報を処理し、歩行者領域を抽出するアルゴリズムを自動的に生成する技術の開発を開発した(特許出願中)。歩行者抽出以外の処理についても、目標を設定するだけで、アルゴリズムの自動構築は可能で、車載カメラの他、移動ロボットの環境認識や、監視カメラなど、動画画像処理を必要とする分野への適用も容易で、有効であると考えている。今後は学習セットの増加など、目標達成精度の向上を図る。

参考文献

- [Saneyoshi92] 実吉敬二、埴圭二、喜瀬勝之:“3 次元画像処理による道路形状と障害物の認識”、自動車技術会、vol46, No.4, 1992.
- [Mochizuki04] 望月大介、矢野良知、横山智訓、大熊繁:“車載カメラによる歩行者検出のための背景除去及びアクティブ探索を用いたテンプレートマッチングの高速化”、電気情報通信学会論文誌 D- , vol.j87, No.5, 2004.
- [Aoki99] 青木紳也、長尾智晴:“木構造状画像変換の自動構築法 ACTIT”、映像情報メディア学会誌 vol53.No6, 1999.
- [Nagao02] 長尾智晴:“進化的画像処理”、昭晃堂、2002.