

# 車載用歩行者抽出動画像処理のオンライン追加学習法

笠置 誠佑<sup>†</sup> 小川 原也<sup>†</sup> 喜瀬 勝之<sup>†</sup>

富士重工業株式会社<sup>†</sup>

## 1. はじめに

車両の安全性向上のため、実環境での歩行者の認識が必要とされている。これまで我々は、この歩行者を認識するアルゴリズムを、学習によって自動的に構築するシステムを開発した。このシステムでは、走行しながら得られる大量の情報を学習し、認識アルゴリズムを自動構築する事ができる。

本報告では、車載カメラで取得した画像から、歩行者の認識アルゴリズムを生成し、画像認識処理を実環境へ適応する手法を提案する。そして走行実験により、初期状態と比べて、歩行者の認識が高精度化されていくことを報告する。

## 2. 課題と手法の概要

これまでの歩行者認識手法<sup>(1)</sup>では、時間帯や場所、天候といった、車外環境における多様な条件に対応する必要があるため、十分な汎用性をもった認識アルゴリズムを開発することは難しい。

これらの手法に対し提案手法では、認識アルゴリズムの自動構築法<sup>(2)</sup>を基盤として、走行しながら多様な条件を学習する。この自動構築法は、画像の認識アルゴリズムを木構造フィルタ列（以下、木構造）で表現し、Fig.1 で表す入力画像を Fig.2 で表す教師画像にするための画像処理を、遺伝的プログラミングによって自動的に構築する。

本システムでは、この自動構築法を使って得られた木構造を複数使い、それらの出力を統合することで歩行者を認識する。そしてもし環境が大きく変化した場合は、この複数の木構造の中で、個々の木構造の追加・削除や、木構造の組み合わせを最適化することによって、より環境に適した歩行者の認識を行う。



Fig. 1 入力画像

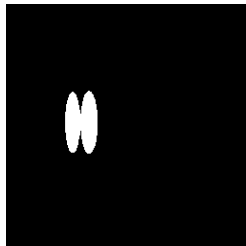


Fig. 2 教師画像

## 3. システムの流れ

以下に提案するシステムの全体図を Fig. 3 に示す。このシステムでは、図中の認識器によって、複数の木構造の実行と出力の統合とをリアルタイムで実行しリスクの認識を行っている。またその一方で環境の変化に対応するため、図中の学習器によって、木構造の追加削除を行い、認識器をバックグラウンドで更新している。

認識器では、木構造が出力した複数の画像を、1枚の画像に統合し、認識結果をモニタへ 10Hz でリアルタイムに出力している。一方学習器では、小川らが開発した教師作成部<sup>(3)</sup>が作る教師情報の入力をトリガとして学習が行われ、認識器の中にある木構造の追加・削除を行う。これら一連の、木構造の実行、統合、削除、追加を繰り返しながら、歩行者の認識と、多様な実環境への適応を行う。

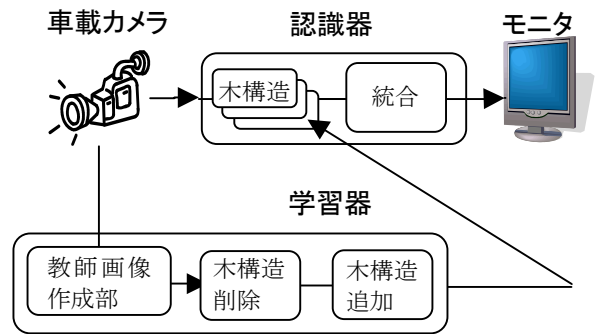


Fig. 3 オンライン学習の流れ

## 4. 木構造の統合

統合では、複数の木構造の出力画像を、式1に従い重ね合わせることで、1つの統合画像を出力している。式中の足切り閾値とは、統合画像の中で、少数の木構造しか出力しなかった領域を抑えるための閾値である。

$$O = \begin{cases} \frac{1}{N} \sum_{tree} P & \left( \frac{1}{N} \sum_{tree} P \geq Thr \right) \\ 0 & \left( \frac{1}{N} \sum_{tree} P < Thr \right) \end{cases} \quad (式1)$$

O : 統合画像      N : 木構造の数      P : 出力画像  
Thr : 足切り閾値

## 5. 学習と評価方法

Fig. 2 の教師画像と、木構造の出力画像を用いて、式2に従い木構造の評価を行う。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す。重要度は、抽出対象領域とそれ以外の領域との面積比によって設定される。今回は

Onboard online addition learning function for image processing of pedestrian detection

<sup>†</sup>Seisuke Kasaoki, <sup>†</sup>Genya Ogawa, <sup>†</sup>Katsuyuki Kise, Fuji Heavy Industries Ltd.

歩行者の認識なので、歩行者領域とそれ以外の領域との面積比×重みが等しくなるように重み画像を設定している。また得られる評価値は1を最高として0～1の範囲をとる。

$$E = 1.0 - \frac{1}{N} \sum_{set} \frac{\sum_{pixel} W \cdot |O - T|}{\sum_{pixel} W \cdot V} \quad (式2)$$

$E$  : 評価値     $N$  : 学習セット数     $V$  : 最大階調値  
 $T$  : 教師画像     $O$  : 出力画像     $W$  : 重み画像

## 6. 木構造の削除

認識器にある木構造を更新するためには、木構造の削除と追加が必要になる。この削除のため、持てる木構造の上限を  $N$  個とする認識器の中から、少なくとも1個以上で構成される木構造の集合を作り、その統合画像を評価する。そしてこの木構造の集合を作る作業と評価を、1から  $N$  個まで順に、全ての木構造の組み合わせで行い（組み合わせは  $2^N$  通り）、最も評価が高かった組み合わせを、新しい木構造として認識器に使用する。仮に  $N-3$  個の組み合わせが最高だったとすると、3個の木構造を削除することができ、かつ残った木構造は、お互いを補う出力ができる。

## 7. 木構造の追加

削除作業によって上限  $N$  個から3個の木構造が削除された後、新しい木構造を追加する。ただし削除された木構造がない場合、追加は行われない。追加作業では、削除後に残った木構造が作る統合画像の中で、誤認識した箇所にも重みを付けた後、追加用の木構造を自動構築法<sup>(2)</sup>によって生成する。この重み付けには前述した重み画像を使う。重み付け後に新しく作られた木構造は、削除後に残った木構造が誤認識した箇所を出力しにくくする。このため、生成した木構造を追加すれば、統合画像の出力を補正することが可能になる。なおこの追加作業は、認識器が持てる木構造の数が、上限に達するまで続けられる。

## 8. 実験結果

約45分の実走行での試験を行い、机上で充分学習させた学習初期の状態からオンライン学習を行った。学習画像での削除・追加結果を Fig. 4 と Table1 に、未学習画像の結果を Fig. 5 と Table2 に示す。なお図中の濃い赤色部分が認識した歩行者抽出領域である。

学習結果では、教師画像に対して、初期の状態よりも歩行者に対する認識が向上しているのが分かる。また未学習結果でも、学習初期のシーンより、学習後期の方が、認識率が向上している。

### ①学習結果（左から初期状態、削除後、追加後）

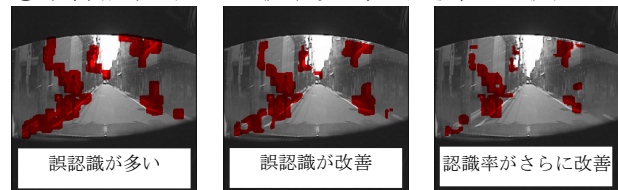


Fig. 4 学習結果

Table1 学習結果

	初期状態	削除後	追加後
評価値	0.80	0.83	0.89

### ②未学習結果（左から学習初期、学習後期）

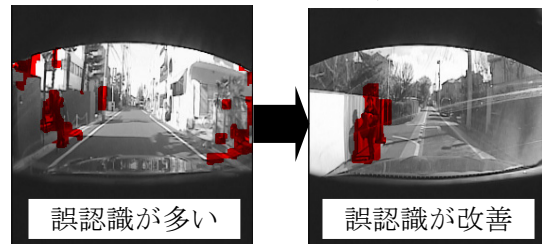


Fig. 5 未学習結果

Table2 未学習結果

	学習初期	学習後期
評価値	0.79	0.88

## 9. まとめ

提案アルゴリズムにより、オンラインで教師情報を取得し学習するほどに歩行者の認識率が向上し、実環境に合わせて画像認識処理を適応していることが分かる。今後は、木構造の学習方法と統合方法が課題となる。木構造の課題としては、ノイズがある教師情報に対しても学習できる枠組みが必要である。また統合方法の課題としては、個々の木構造の出力を平均化する現行の方法から、個々の木構造の特性を活かす新しい統合手法が必要である。

なお今回の研究にあたり多大なるご指導を頂いた横浜国立大学大学院 長尾智晴教授に、深く感謝申し上げます。

## 参考文献

- (1) 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およ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”, 電子情報通信学会論文誌 D-II, vol. j87, No. 5, 2004
- (2) 小川原也, 喜瀬勝之, 長尾智晴: “遺伝的プログラミングによる動画処理アルゴリズムの獲得”, 人工知能学会全国大会 2005, 3F4-02
- (3) 小川原也, 喜瀬勝之, 長尾智晴: “車載用歩行者抽出動画処理のオンライン学習による獲得”, 第69回情報処理学会全国大会 発表予定