

車載カメラを用いた周囲環境認識

出口 大輔[†] 道満 恵介^{††} 井手 一郎^{†††} 村瀬 洋^{†††}

[†] 名古屋大学情報連携統括本部情報戦略室

^{††} 中京大学工学部情報工学科

^{†††} 名古屋大学情報科学研究科メディア科学専攻

あらまし 本発表では、車載カメラを用いた周囲環境認識の高度化を目的とし、車載カメラ映像から多様な学習サンプルを自動的に収集し、高精度な物体検出器を構築する技術に関して報告する。具体的には標識検出器を題材として用い、車載カメラ映像から標識検出器の構築に必要な標識画像（ポジティブサンプル）と背景画像（ネガティブサンプル）を同時に集める技術を紹介する。実際の車載カメラ映像を用いた実験の結果、高精度な標識検出器をわずかな手作業で構築できることを確認した。

キーワード 標識検出, 自動収集, 学習サンプル, カスケード型識別器, 車載カメラ

1. はじめに

画像からの物体検出はコンピュータビジョン・パターン認識の分野における最も重要な課題の一つであり、さまざまな手法の研究開発が広く行われている。また近年では、デジタルカメラにおける顔検出やサーベイランスシステムにおける歩行者検出といった様々なシステムで利用されるようになってきている。さらに、ITS (Intelligent Transport System) においても物体検出技術の重要性は高まってきており、安全かつ快適な運転環境を実現するための重要な技術となっている。しかしながら、我々が日々運転する道路環境には様々な物体（標識、信号、歩行者、他）が存在し、またそれらの見えは環境によって多様に変化する。そのため、これらの多様に変化するさまざまな物体を精度良く検出する技術が求められている。本発表では、特に交通環境の認識において重要な情報源となる道路標識に注目し、車載カメラ映像から道路標識を検出する技術を紹介する。

これまでに、さまざまな研究グループから標識を対象とした検出手法が提案されている [1]~[4]。Bahlmann らは、顔検出で広く用いられている AdaBoost 識別器 [5] を用い、車載カメラ画像から高速かつ高精度に標識を検出する手法を提案している [3]。この手法では、標識検出器の構築に大量の学習画像（標識画像と背景画像）の準備が必要であり、Bahlmann らは高精度な標識検出器を構築するための手作業でこれらの学習画像を大量に収集している。しかしながら、この作業に必要な人的コストは非常に高く、より簡便な手法が求められている。これに対し、道満らは道路標識の劣化モデルを考慮した生成型学習により、大幅なコスト削減を達成している [4]。しかしながら、図 1 に示すような実際の道路環境で観察される多様な見えの変化すべてを生成型学習で生成することは困難である。

上述の問題を解決するため、我々の研究グループでは



図 1 実際の道路環境で観察される道路標識の見えの変化。



図 2 本発表で対象とする道路標識。

遡及型追跡を用いて標識画像を自動収集する手法を提案した [6]。現在はこの手法を発展させ、標識画像と背景画像を車載カメラ映像から同時に収集する手法を実現している [7]。本発表では、文献 [7] で提案した手法と最先端の画像特徴量および識別器を組み合わせた場合の標識検出器の構築手法について紹介する。また、本発表で扱う標識の一覧を図 2 に示す。

2. 手 法

本発表では、走行中に撮影される車載カメラ映像から標識画像および背景画像を自動収集し、それらを学習サンプルとして利用することで高精度な標識検出器を構築する。

一般に、高精度な標識検出器を構築するためには、学習画像として用いる標識画像と背景画像に実際の道路環境で観察される多様な見えの変化が含まれている必要がある（幾何学的変形、照明変化、他）。従来 of 多くの手

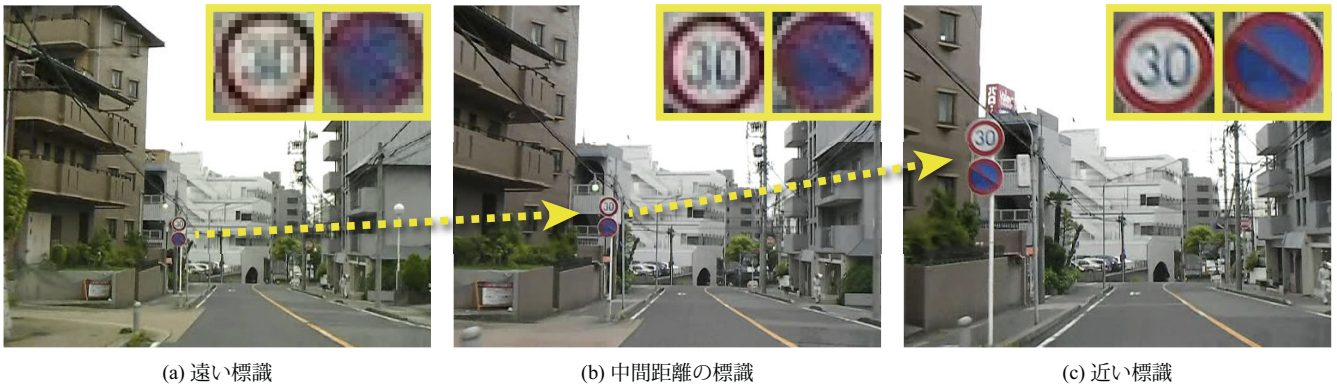


図 3 距離による道路標識の見えの変化.

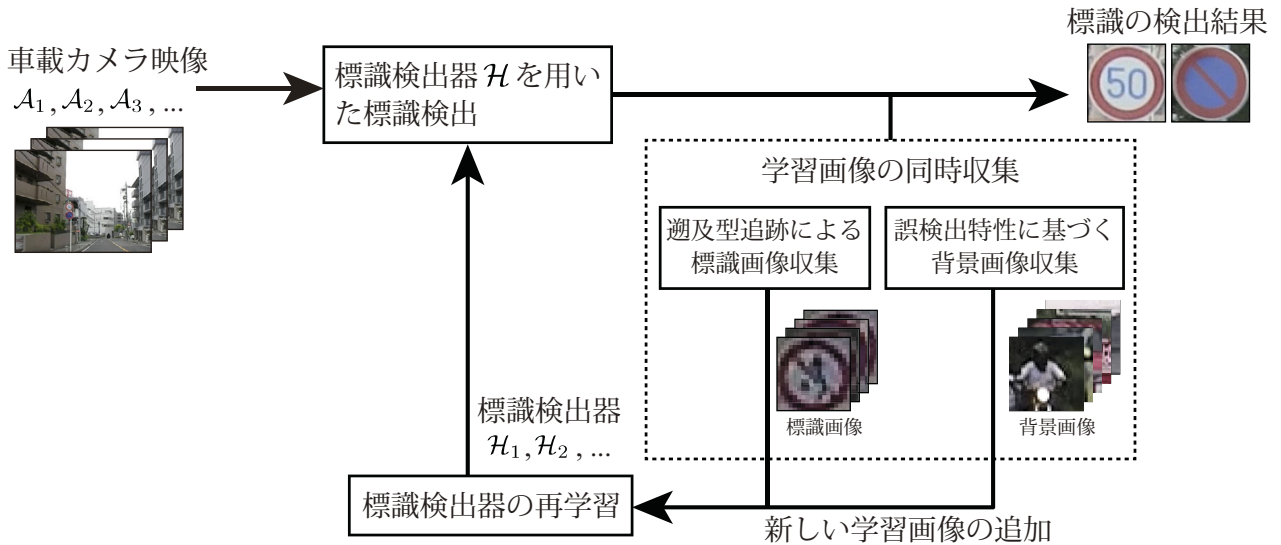


図 4 処理の流れ.

法では、このような学習画像を手作業で集めており、その人的コストは非常に高い。そこで本発表では、学習画像として用いる標識画像と背景画像を車載カメラ映像から自動的かつ同時に収集する手法を提案する。これにより、標識検出器の構築に必要な多様な学習画像を低コストに収集することを可能にする。

図 3(a) に示すように、一般に車から遠くの標識は解像度が低く、それらを精度良く検出して収集することはたとえ手作業であったとしても大変困難である。これに対し、図 3(c) から分かるように、車の走行に伴ってこのような解像度の低い小さな標識もいつかは大きな標識として観察される。このような大きな標識（車と標識の距離が近い場合）は、比較的容易に検出することが可能である。また、一度標識の位置が得られると、その位置を入力として標識を追跡することは容易である。これらの考えに基づき、提案手法ではまず大きな標識を検出し、そこから過去に遡って標識を追跡（遡及型追跡）することで、多様な標識画像を自動的に収集する。

一方、標識検出器により得られる検出結果を注意深く観察すると、背景の誤検出は画像内でランダムに出現し、また、時間方向に連続して検出されることもほとんど無

いことが分かる。そこでこの誤検出に関する特性を利用し、提案手法では検出結果の時間的かつ空間的な重なり率を計算することで誤検出らしさを定量化する。そして、この値に基づいて背景画像を自動的に収集する。

最終的に、これら 2 つの処理を同時並列に実行することで、標識画像と背景画像を自動的に収集する。具体的な処理の流れを図 4 に示す。以下の節で、それぞれの処理について簡単に説明する。

2.1 学習サンプルの自動収集

2.1.1 遡及型追跡による標識画像の収集

あらかじめ構築した標識検出器を用い、車載カメラ映像の各フレームから一定サイズ以上の標識を検出する。そして、大きな標識の位置を初期値とした遡及型追跡により、多様な大きさでかつ見えの変化を含む標識画像を収集する。具体的には、(1) 標識の検出、(2) 大きな標識の選択、(3) 標識の遡及型追跡、の 3 つの処理を適用することで、多様な標識を収集する。詳細に関しては、文献 [7] を参照されたい。

2.1.2 誤検出特性を利用した背景画像の収集

本処理では、車載カメラ画像に対して標識検出器を適

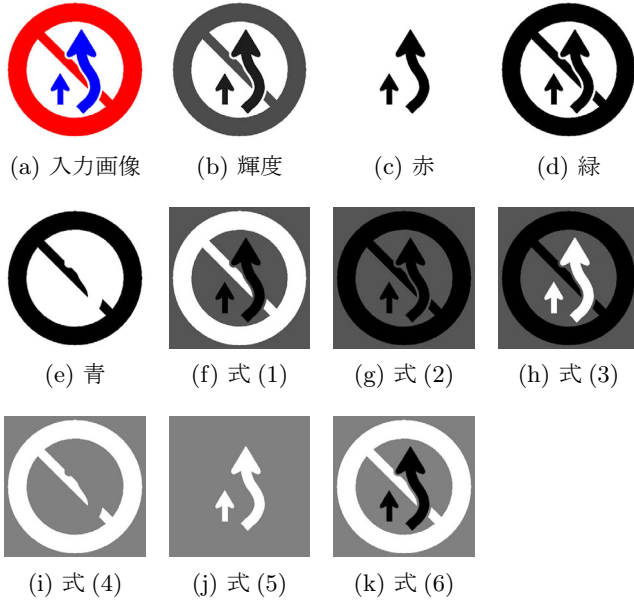


図5 MLBP 特徴量の計算に用いる色特徴画像の例.

用した結果として得られる誤検出を背景画像として収集する. 一般に, 標識に対して標識検出器を適用した場合, 標識は時間的にも空間的^(注1)にも連続して検出することが可能である. 一方, 背景は時間的にも空間的にも連続して検出されることはなく, 車載カメラ映像中に散発的に現れるという特性がある. そこで, これを誤検出に関する特性と捉え, 標識検出器により得られる検出結果の時間的かつ空間的な重なり率を計算することで, 背景らしさを評価する. この結果として得られる背景らしさの高いものを背景画像として収集する.

2.2 標識検出器の構築

本発表では, 画像特徴量として Zhang らが提案した MLBP (Multi-block Local Binary Pattern) 特徴量 [10] を用い, 識別器として Gentle AdaBoost 識別器を Multi-exit カスケード状に接続したもの [8], [9] を利用する.

まず, 図5に示す10種類の色特徴画像 $f_1 \sim f_{10}$ を車載カメラ画像から求める. ここで, f_1 は輝度画像であり, $f_2 \sim f_4$ は RGB 画像の各成分である. また, $f_5 \sim f_7$ は正規化 RGB 画像の各成分であり, $f_8 \sim f_{10}$ は対数 RGB 値 [11] の各成分である. $f_5 \sim f_{10}$ の具体的な計算式は以下の通りである.

$$f_5(\mathbf{x}) = \frac{r(\mathbf{x})}{r(\mathbf{x}) + g(\mathbf{x}) + b(\mathbf{x})}, \quad (1)$$

$$f_6(\mathbf{x}) = \frac{g(\mathbf{x})}{r(\mathbf{x}) + g(\mathbf{x}) + b(\mathbf{x})}, \quad (2)$$

$$f_7(\mathbf{x}) = \frac{b(\mathbf{x})}{r(\mathbf{x}) + g(\mathbf{x}) + b(\mathbf{x})}, \quad (3)$$

$$f_8(\mathbf{x}) = \log \frac{r(\mathbf{x})}{g(\mathbf{x})}, \quad (4)$$

(注1): 検出窓の位置が多少変化しても標識として検出されるため.

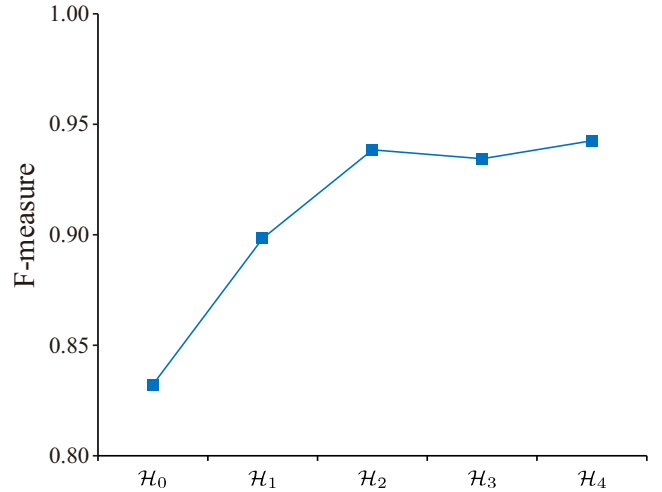


図6 標識検出器の反復的構築による性能の変化 ($\mathcal{H}_0 \sim \mathcal{H}_4$).

$$f_9(\mathbf{x}) = \log \frac{b(\mathbf{x})}{g(\mathbf{x})}, \quad (5)$$

$$f_{10}(\mathbf{x}) = \log \frac{r(\mathbf{x})}{b(\mathbf{x})}, \quad (6)$$

ここで, $r(\mathbf{x})$, $g(\mathbf{x})$, $b(\mathbf{x})$ は画素 \mathbf{x} の赤色成分, 緑色成分, 青色成分である.

これらの色特徴画像を 2.1.1 節と 2.1.2 節で収集した標識画像および背景画像に対して計算し, 得られる色特徴画像から MLBP 特徴量を求めることで識別器の学習に利用する.

また図4の処理の流れに示すように, 構築した識別器を 2.1 節の入力として用いることで, 新たな標識画像と背景画像の収集を行う. このように, 学習画像の収集と標識検出器の構築を繰り返し行うことにより, 標識検出器の精度が逐次的に向上することが期待される.

3. 実験および考察

提案手法の有効性を確認するために, 実際の道路環境で撮影した車載カメラ映像を用いて実験を行った. 車載カメラとして SANYO Xacti DMX-HD2 を利用し, カメラをフロントガラス上に固定して走行することにより車載カメラ映像を撮影した. 撮影した車載カメラ映像の各フレームは 640×480 画素 (30 fps) である. 提案手法の性能を評価するため, 2,967 フレームからなる車載カメラ映像を5つの系列 ($\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_4$) に分割し, それらを提案手法の入力として用いた. これらの系列には $15 \times 15 \sim 75 \times 75$ 画素の標識が 4,973 個含まれていた. また, 標識の映っていない 640×480 画素の画像を10枚用意し, これらの画像から切り出した部分画像を提案手法における初期の背景画像として用いた.

3.1 実験手順

学習画像の収集と標識検出器の構築を繰り返し適用することによる性能変化を確認するため, 本実験では車載カメラ映像系列 $\mathcal{A}_0 \sim \mathcal{A}_4$ を順に入力することで標識検出

器 $\mathcal{H}_0 \sim \mathcal{H}_4$ を逐次的に構築した. 具体的な構築の手順は以下の通りである.

まず, 系列 \mathcal{A}_0 から手作業により 13 個の大きな標識を抽出した. そして, 切り出し位置と回転のパラメータを微小変化させることで 500 枚の標識画像を抽出し, これを初期の標識画像とした. また, 事前に用意した標識の映っていない 10 枚の画像から部分画像を切り出し, 標識検出器を構築するための初期背景画像とした. これらの標識画像および背景画像を用いることで, 初期標識検出器 \mathcal{H}_0 を構築した. 提案手法で必要となる手作業は, これらの 13 個の大きな標識の抽出と, 10 枚の標識の映っていない画像の準備のみである.

次に, 2.1 節で述べた処理に従い, 車載カメラ映像系列 \mathcal{A}_1 から検出器 \mathcal{H}_0 を用いて標識画像と背景画像を同時に自動収集した. そして, 得られた画像と検出器 \mathcal{H}_0 の構築に使用した画像に加え, 新しい検出器 \mathcal{H}_1 を構築した.

これらの処理を順次適用することにより, $\mathcal{H}_2, \mathcal{H}_3, \mathcal{H}_4$ の検出器を構築した.

3.2 実験結果と考察

3.1 節の手順に従って構築した検出器 $\mathcal{H}_0 \sim \mathcal{H}_4$ の性能を, 適合率と再現率の調和平均である F 値により評価した. ここで F 値は,

$$F \text{ 値} = \frac{2 \times \text{再現率} \times \text{適合率}}{\text{再現率} + \text{適合率}}. \quad (7)$$

により計算した. 検出器 $\mathcal{H}_0 \sim \mathcal{H}_4$ に対する F 値を図 6 に示す. また, 検出器 \mathcal{H}_4 による検出結果を図 7 に示す.

図 6 から分かるように, 提案手法を反復的に適用することにより, 検出器の性能が向上していくことが分る. また, 最終的に得られる検出器 \mathcal{H}_4 の性能は非常に高く, 図 7 からも標識のみを精度良く検出できていることが分かる. 一方, 今回の実験における手作業は, 13 枚の標識の入力, および 10 枚の標識の映っていない画像の準備のみである. これは, 非常に大変な手作業が必要であった従来手法の学習画像収集コストを, 提案手法により大幅に削減できることを示唆している. このことから, 低コストな標識検出器の構築に提案手法が貢献することが期待される.

4. む す び

本発表では車載カメラを用いた周囲環境認識の高度化を目的とし, 車載カメラ映像から多様な学習サンプルを自動的に収集し, 高精度な物体検出器を構築する技術に関して報告した. 特に, 交通環境の認識において重要な情報源となる道路標識に注目し, 低コストで高精度な標識検出器を構築可能な手法を紹介した. 具体的には, (1) 遡及型追跡による標識画像収集, (2) 誤検出特性に基づく背景画像収集, の 2 つの枠組みを同時に実行する手法を提案した. また, 最先端の画像特徴量および識別器を

用いて自動収集により達成される性能を評価した. 実際の車載カメラ映像を用いた実験の結果, 高精度な標識検出器を低コストで構築できることを確認した.

今後の課題として, (i) 遡及型追跡の性能改善, (ii) 背景画像収集の高精度化, (iii) 大規模データによる評価, (iv) 様々な対象への適用, などが挙げられる.

謝辞 日頃より熱心に御討論頂く名古屋大学村瀬研究室諸氏に深く感謝する. 本研究の一部は, 文部科学省科学研究費補助金および JST 戦略的創造研究推進事業 CREST による. また, 本研究では画像処理に MIST ライブラリ (<http://mist.suenaga.m.is.nagoya-u.ac.jp/>) を使用した.

文 献

- [1] S. Maldonado-Bascón, S. Lafuente-Arroyo, P. Gil-Jiménez, H. Gómez-Moreno, and F. López-Ferreras, "Road-sign detection and recognition based on support vector machines," *IEEE Transactions on Intelligent Transportation Systems*, Vol.8, No.2, pp.264–278, 2007.
- [2] G. Loy and N. Barnes, "Fast shape-based road sign detection for a driver assistance system," *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol.1, pp.70–75, 2004.
- [3] C. Bahlmann, Y. Zhu, V. Ramesh, M. Pellkofer, and T. Koehler, "A system for traffic sign detection, tracking, and recognition using color, shape, and motion information," *Proceedings of IEEE Intelligent Vehicles Symposium*, pp.255–260, 2005.
- [4] 道満 恵介, 出口 大輔, 高橋 友和, 目加田 慶人, 井手 一郎, 村瀬 洋, "色変動を考慮した生成型学習法による道路標識検出器の構築," *電子情報通信学会論文誌, J93-D*, No.8, pp.1375–1385, 2010.
- [5] P. Viola and M. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, Vol.57, No.2, pp.137–154, 2004.
- [6] 出口 大輔, 道満 恵介, 井手 一郎, 村瀬 洋, "遡及型追跡に基づく標識画像の自動収集を用いた標識検出器の高精度化," *電子情報通信学会論文誌, J95-D*, No.1, pp.76–84, 2012.
- [7] D. Deguchi, M. Shirasuna, K. Doman, I. Ide, and H. Murase, "Intelligent Traffic Sign Detector: Adaptive Learning Based on Online Gathering of Training Samples," *Proceedings of 2011 IEEE Intelligent Vehicles Symposium (IV2011)*, pp.667–672, 2011.
- [8] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting," *Annals of Statistics*, Vol. 28, No. 2, pp.337–407, 2000.
- [9] M. Pham, V. D. Hoang, T. Cham, "Detection with Multi-exit Asymmetric Boosting," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1–8, 2008.
- [10] L. Zhang, R. Chu, S. Xiang, S. Liao, Stan Z. Li, "Face Detection Based on Multi-Block LBP Representation," *Proceedings of IEEE International Conference on Biometrics*, pp.11–18, 2007.
- [11] J. Berens, G. D. Finlayson, "Log-opponent chromaticity coding of colour space," *Proceedings of the 15th International Conference on Pattern Recognition (ICPR)*, pp.206–211, 2000.

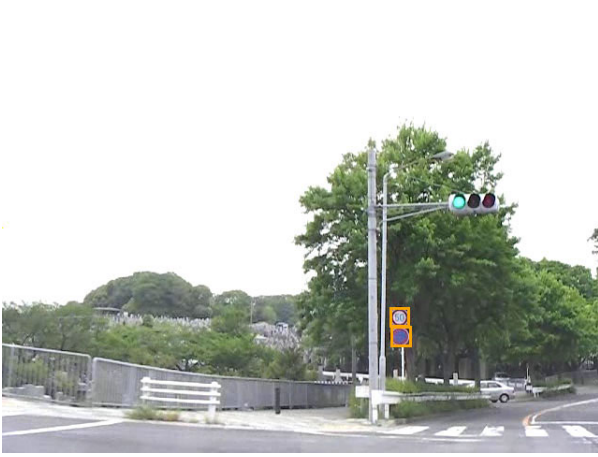


図 7 標識検出結果の一例.