

車載カメラ映像からの学習サンプルの自動収集による 標識検出器の高精度化に関する検討

出口 大輔[†] 道満 恵介[†] 井手 一郎[†] 村瀬 洋[†]

[†] 名古屋大学大学院情報科学研究科 〒464-8601 愛知県名古屋市千種区不老町

E-mail: [†]{ddeguchi, ide, murase}@is.nagoya-u.ac.jp, ^{††}kdoman@murase.m.is.nagoya-u.ac.jp

あらまし 本稿では、車載カメラ映像から多様な標識画像を自動収集し、高精度な標識検出器を構築する手法を提案する。遠く離れた位置の標識を精度良く検出するためには、遠くの標識を撮影した大量の学習サンプルを用いて標識検出器を構築する必要がある。しかしながら、遠く離れた標識のサイズは小さく、それらを自動もしくは半自動で大量に収集することは困難である。一方、ある時点では遠くの標識であったとしても、自車の走行に伴って遠くの標識も徐々に大きく撮影されるという特徴がある。このような大きく撮影された標識は自動検出可能であり、また、標識の時間方向への追跡も比較的容易である。そこで本稿では、まず近くで大きく撮影された標識を検出し、時間を遡りながら標識を追跡することによって、遠く離れた位置の低品質の標識を切り出して収集する。そして、自動収集した標識画像を学習サンプルとして利用し、標識検出器を構築する。提案手法をさまざまな環境で撮影した車載カメラ映像に適用した結果、F 値が約 0.95 の標識検出器が構築可能であることを確認した。

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

A study on an improvement of a traffic sign detector by automatic gathering of training data from in-vehicle camera images

Daisuke DEGUCHI[†], Keisuke DOMAN[†], Ichiro IDE[†], and Hiroshi MURASE[†]

[†] Graduate School of Information Science, Nagoya University Furo-cho, Chikusa-ku, Aichi, 464-8601 Japan

E-mail: [†]{ddeguchi, ide, murase}@is.nagoya-u.ac.jp, ^{††}kdoman@murase.m.is.nagoya-u.ac.jp

Abstract This paper proposes a method for constructing an accurate traffic sign detector by automatic gathering of training data from in-vehicle camera images. To detect distant traffic signs accurately, it is necessary to prepare numerous training samples of distant traffic sign images, and to construct a detector by using them. However, it is difficult to obtain those traffic sign images automatically or even semi-automatically, due to its size. On the other hand, the distance between a distant traffic sign and a vehicle shortens gradually when driving a vehicle, and it is captured as a large traffic sign eventually. It is easy to segment this large traffic sign, and also possible to track it back to the past relatively easily. Therefore, this paper gathers distant traffic sign images by detecting a large traffic sign and tracking it back to the past. Then, a traffic sign detector is constructed by using the obtained traffic sign images. We conducted experiments using in-vehicle camera images taken at various environments. Experimental results showed that the proposed method could construct an accurate traffic sign detector with a F-measure of approximately 0.95.

Key words Traffic sign detection, automatic gathering, training images, cascaded classifiers, in-vehicle camera

1. はじめに

近年、我々の運転する自動車にはさまざまなセンサが搭載されるようになってきている。その中でも車載カメラは急速に普及してきており、駐車支援を目的としたアラウンドビューモ

ニター (NISSAN)、安全走行支援を目的としたナイトビュー (TOYOTA) やインテリジェント・ナイトビジョンシステム (HONDA)、などで利用されている。ここ数年では、車載カメラを交通環境の認識・理解に利用する研究が広く行われるようになりつつあり、車載カメラを利用した高精度位置推定や道路

地図の作成に関する研究が盛んである。

安全運転支援システムを実現するためには、交通環境の認識・理解が必要不可欠な技術であり、高度な運転支援や安全・快適なサービスをドライバーへ提供するための重要な技術となる。特に、道路標識は交通環境を認識するための重要な情報であり、車載カメラ映像から道路標識を検出・認識する技術が注目を集めている。車載カメラ映像中の道路標識を認識するためには、道路標識検出が必要不可欠な前処理となる。これまでに、標識の色特徴を利用した手法 [1], [2] や、形状特徴を利用した手法 [3], [4], 顔検出の技術を応用した手法 [5], [6], など様々な方法が提案されている。Bahlmann ら [5] は、デジタルカメラの顔検出機能で広く利用されている高速な物体検出技術 [7] を標識検出に応用し、高速かつ高精度な標識検出器を構築可能であることを実験により示している。この手法では、Haar-like 特徴を特徴量として用い、AdaBoost アルゴリズムを用いて構築した強識別器をカスケード状に組み合わせることで、高速かつ高精度な標識検出器を実現する。しかしながら、安定して高い検出性能を発揮するためには、図 1 に示すような実際に起こりうる様々な変動パターンをバランスよく含んだ学習用画像を用意する必要がある。Bahlmann ら [5] は、手作業でこれらの学習画像を収集し、標識検出器の構築に利用している。しかしながら、実際に起こりうるすべての変動パターンを手作業で収集するには、数え切れないほど多くの標識画像が必要となるため、それらを手作業で集めることは不可能に近い。道満らは、生成型学習法を利用することでこの問題を解決する手法を提案している [6]。生成型学習は、対象をカメラで撮影する際に起こる種々の画像変化をモデル化し、さまざまな画像を生成して学習を行う手法である。しかしながら、現実にかかるすべての画像変化をモデル化することは困難であり、また、生成モデルのパラメータ決定が難しいといった問題がある。

そこで本稿では、車載カメラ映像から学習サンプルとなる標識画像を自動収集することにより、高精度な標識検出器を構築する手法を提案する。具体的には、図 2 に示す重要な円形の規制標識を対象とし、車載カメラ映像から多様な変動（照明変化、解像度変化など）を含む標識画像を大量に収集する。そして、自動収集した標識画像を学習画像として用い、高精度な標識検出器の構築を行う。これにより、従来大きな問題であった手作業による学習画像の収集コストを大幅に軽減する。

以降、2. で車載カメラ映像からの標識画像の自動収集による標識検出器の構築手法を説明し、3. で車載カメラ映像を用いた提案手法の評価を行う。4. で実験結果について考察を行った後、5. でまとめる。

2. 手 法

本稿では、走行中に撮影される車載カメラ映像から標識画像を自動収集し、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手法を提案する。高精度な標識検出器を構築するためには、解像度の低いものから高いものまで、多様な標識画像を切り出して収集する必要がある。しかしながら、図 3(a) のような遠方の低品質の標識（カメラと標識の距

図 1 実環境における標識の見えの変化



図 2 対象標識

離が遠い場合）は、半自動的であったとしても 1 枚の画像だけから正しく切り出して収集することは難しい。一方、図 3(c) のように近くで大きく撮影された標識は、認識や切り出しが容易である。また、一度大きく撮影された標識の位置が分かれば、時間を遡って標識を追跡することは比較的簡単である。そこで、まず近くで大きく撮影された標識を自動で見つけ、そこから時間を遡りながら標識を追跡することで低品質の標識を正しく切り出し、多様な学習サンプルを収集する。提案手法の処理の流れを図 4 に示す。提案手法は大きく分けて、(i) 標識画像の自動収集、(ii) 標識検出器の構築、の 2 つの段階からなる。これら 2 つの処理を車載カメラ映像に対して適用し、標識検出器の構築を行う。以降で具体的な処理手順を示す。

2.1 標識画像の自動収集

標識画像の自動収集手順は、(a) 大きく撮影された標識の検出、(b) 標識の逆方向追跡、の 2 つの処理に分かれる。大きく撮影された標識のみを検出する初期検出器は比較的容易に構築が可能である。一方、遠方で低解像度の標識の検出は困難であるが、もし正しい標識の位置が分かれば時間方向の追跡は容易である。本手法はこの 2 つの処理を組み合わせたものである。

これら 2 つの処理を走行中に得られる車載カメラ映像に適用し、標識画像を自動収集する。また、あらかじめ構築した標識の検出器 $\mathcal{H}_k$ を本処理の入力として用いる。

2.1.1 大きく撮影された標識の検出

ここでは、検出器 $\mathcal{H}_k$ を用い、車載カメラ映像中で大きく撮影された標識の検出を行う。まず、入力画像に対して検出窓の位置とスケールを変化させながら走査し、各検出窓で切り出した入力画像が標識であるかどうかを $\mathcal{H}_k$ を用いて判定する。そして、標識と判断された検出窓 $\mathbf{X}_i$ の集合 $\mathcal{M} = \{\mathbf{X}_i\}$ を求め

(a) 遠くの標識

(b) 中間の標識

(c) 近くの標識

図3 カメラとの距離の違いによる標識の見えの変化

$\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \dots$

図4 処理の流れ

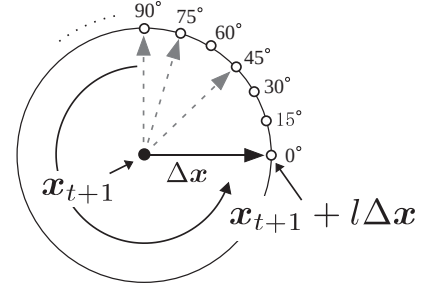


図5 エッジ検出の流れ

る．次に，各検出窓間の重なり率を基準に検出窓の集合 $\mathcal{M}$ を Mean Shift クラスタリング [8] し，複数の検出窓の集合 $\mathcal{M}_j$ ($j = 1, \dots, M$) に分割する．ここで得られる $\mathcal{M}_j$ には，検出に用いた $\mathcal{H}_k$ の性能により，標識以外のものも多数含まれる可能性がある．そこで，次に示す処理を適用することで，確実に標識と考えられる検出窓を抽出する．

上記の検出処理は，入力画像を走査しながら，検出器 $\mathcal{H}_k$ により標識かどうかの判定を行うため，検出対象の標識の周囲に標識と判断された検出窓が多数得られる．そこで，各 $\mathcal{M}_j$ に対して次式を評価することで，検出窓数の多いものを抽出する．

$$\|\mathcal{M}_j\| > \alpha \quad (1)$$

ここで， $\|\mathcal{M}\|$ は集合 $\mathcal{M}$ の要素数を表し， α は要素数に対するしきい値である．式 (1) を満たす集合 $\mathcal{M}_j$ から，集合内の検出窓の中心座標 $\mathbf{x}$ を求め，2.1.2 の初期値として用いる．

2.1.2 標識の逆方向追跡

本処理では，2.1.1 で得られた標識の検出窓を初期値とし，時間を逆方向に追跡することで低品質の標識を自動的に収集する．これは，時刻 $t+1$ における標識の中心位置 $\mathbf{x}_{t+1}$ を入力とし，時刻 t における標識の中心位置 $\mathbf{x}_t$ と半径 R_t を逐次的に求める処理に対応する．具体的には，(1) 標識強調画像の作成，(2) 標識の追跡，の2つの処理により $\mathbf{x}_t$ と R_t を求め，各時刻における標識を切り出す．

(1) 標識強調画像の作成

入力画像 $\mathbf{F}_t$ に対し，次式により強調画像 $\mathbf{G}_t$ を求める．

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

ここで， $r(\mathbf{x})$ ， $g(\mathbf{x})$ ， $b(\mathbf{x})$ は位置 $\mathbf{x}$ における $\mathbf{F}_t$ の赤色成分，緑色成分，青色成分を表す．次に，得られた画像 $\mathbf{G}_t$ に標

準偏差1のガウス平滑化を施し，画像 $\mathbf{G}'_t$ を得る．以降は，画像 $\mathbf{G}'_t$ に対して処理を行う．

(2) 標識の追跡

図5に示すように，時刻 $t+1$ の標識の中心位置 $\mathbf{x}_{t+1}$ を初期値として， $\Delta\mathbf{x}$ を変化させながら標識の輪郭に対応するエッジの抽出を行う．ここで， $\Delta\mathbf{x}$ は標識輪郭の探索方向（大きさ1に正規化されたベクトル）であり， l は標識の中心位置 $\mathbf{x}_{t+1}$ からの距離を表す．標識の輪郭に対応するエッジの抽出は， $\Delta\mathbf{x}$ 方向に l を増加させながら次式を評価することで行う．

$$\nabla \mathbf{G}'_t(\mathbf{x}_{t+1} + l\Delta\mathbf{x}) \cdot \Delta\mathbf{x} < 0 \quad (3)$$

ここで， $\nabla \mathbf{G}'_t(\mathbf{x})$ は $\mathbf{x}$ における輝度の勾配を表し， $\cdot$ はベクトルの内積を表す．式 (3) を満たし，かつ， $\mathbf{x}_{t+1} + l\Delta\mathbf{x}$ における $\mathbf{G}'_t$ の画素値が最大となる点を探索する．この操作を $\Delta\mathbf{x}$ を反時計回りに 15° ずつ回転させながら適用し，標識の輪郭に対応する点の集合を得る（図5）．

次に，得られた点から任意の5点をランダムに選択し，文献 [9] の方法を用いて円を当てはめる．選択する5点を様々に変化させながら求めた円と輪郭点との距離を評価し，その値が最小となる円の中心を $\mathbf{x}_t$ ，半径を R_t とする．

$t \leftarrow t-1$ とすることで時間を遡りながら前述の操作を繰り返し適用し，各時刻における $\mathbf{x}_t$ と R_t を求めることで標識の切り出しを行う．

2.2 標識検出器の構築

本節では，前節までで自動収集した標識画像を学習画像と

Proposed method 1 (LRP) 本稿で説明した手法に従って、標識検出器 $\mathcal{H}_0 \sim \mathcal{H}_4$ を構築。

Proposed method 2 (Haar-Like) 2.2.1 の LRP 特徴量を Haar-Like 特徴量に変更し、標識検出器 $\mathcal{H}_0 \sim \mathcal{H}_4$ を構築。

Conventional method 手作業で抽出した 13 個の標識（初期標識検出器 $\mathcal{H}_0$ の構築時に利用）の切り出し位置を様々に変化させ、収集する標識画像の枚数（500～2,500 枚）を変化させて $\mathcal{H}_0 \sim \mathcal{H}_4$ を構築。

4. 考 察

表 1 から分かるように、提案手法により車載カメラ映像から学習サンプルを自動収集することにより、最大で F 値が 0.949 という高い精度の標識検出器の構築が可能であった。また表 1 は、 $\mathcal{H}_0$ を $\mathcal{A}_1$ の映像に適用して $\mathcal{H}_1$ を構築、 $\mathcal{H}_1$ を $\mathcal{A}_2$ の映像に適用して $\mathcal{H}_2$ を構築、というように学習サンプルを自動収集しながら標識検出器を逐次的に構築した場合の検出性能の変化を示している。これを見ると、順次使用する車載カメラ映像（異なる経路）を増やすことによって標識検出器の性能が向上することが確認できる。これは、カメラを搭載した車によりさまざまな環境を走行することで、逐次的に標識検出器の性能を改善できる事を意味している。これらから、提案手法はわずかな手作業のみで高精度な標識検出器の構築が可能であることを確認できる。

また、図 9(a) は、手入力で与えた 13 枚の標識画像（切り出し位置と回転の変化を含む）のみから学習サンプルを切り出した場合（Conventional method）の検出性能の変化を示している。具体的には、切り出し位置を変化させながら標識画像を 500～2,500 枚（500 枚きざみ）収集し、 $\mathcal{H}_0 \sim \mathcal{H}_4$ を構築している。図 9(a) のグラフから分かるように、手入力で与えた 13 枚の標識画像のみから標識検出器を構築した場合は、十分に標識の見えの変化を学習することができず、高い検出性能を得ることができていないことが分かる。特に、切り出し位置等を変化させて多数の学習サンプルを生成したとしても、F 値は最大で 0.79 程度である。これに対し、提案手法は手入力で与えた 13 枚の標識画像を利用して大量の学習サンプルを自動収集し、F 値で約 0.95 の検出器を構築可能であった。このように、提案手法はカメラを搭載した車でさまざまな環境を走行することで標識検出器の性能が改善できるため、従来の手作業で標識画像を収集する手法と比較して学習画像収集の手間を大幅に軽減できると考えられる。

一方、検出器の特徴量として Haar-Like 特徴量を利用した場合（Proposed method 2 (Haar-Like)）、高い再現率を得ることは可能であるが、図 9(b) から分かるように適合率は徐々に低下していくことが分かる。これは、2.1.1 で大きな標識を検出する際に多くの誤検出が発生し、自動収集した学習サンプルの中に標識以外のサンプルが多く含まれたためである。これに対し、本稿で利用した LRP 特徴量を用いた検出器の誤検出は少なく、結果として高い性能が得られたと考えられる。しかしながら、LRP 特徴量を用いたとしても僅かではあるが標識以

表 1 $\mathcal{H}_0 \sim \mathcal{H}_4$ を評価映像に適用した結果

標識検出器	適合率	再現率	F 値
$\mathcal{H}_0$	0.969	0.485	0.647
$\mathcal{H}_1$	0.966	0.857	0.908
$\mathcal{H}_2$	0.964	0.926	0.945
$\mathcal{H}_3$	0.953	0.945	0.949
$\mathcal{H}_4$	0.941	0.954	0.948

図 7 対象外の標識を含む例

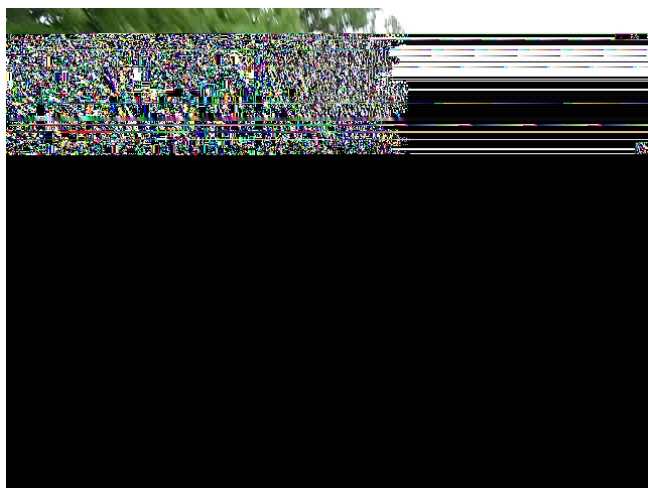


図 8 標識に部分的な隠れを含む例

外のサンプルが収集されることを確認した。今後、2.1.1 の処理における誤検出の抑制方法を検討し、学習サンプル収集の性能を改善する予定である。

図 10 と 11 は 2.1.2 の標識の追跡結果の一例を示している。図 11 から分かるように、提案手法は標識に部分的な隠れが存在する場合でも精度よく標識を切り出して収集できることが分かる。これは、標識の一部が隠れている場合でも、他の標識輪郭の抽出が可能なたためである。しかしながら、標識と背景の境界が不明瞭な場合や、標識の解像度が著しく低い場合に正しく標識を追跡することができなかった。今後、標識の境界抽出手法を改良し、追跡性能を改善する予定である。

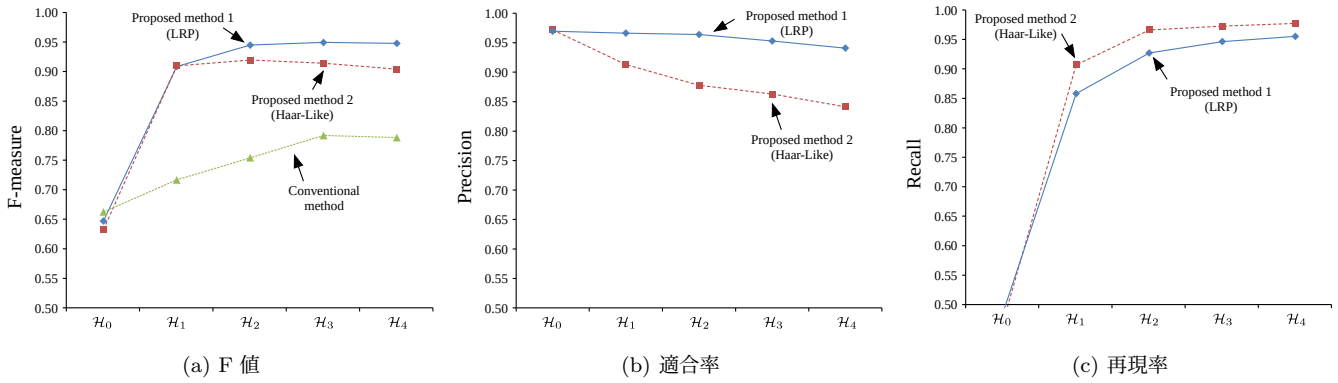


図9 提案手法と比較手法のそれぞれで構築した標識検出器 $\mathcal{H}_0 \sim \mathcal{H}_4$ の F 値, 適合率, 再現率.

図10 標識追跡結果の一例 (右上はフレーム番号)

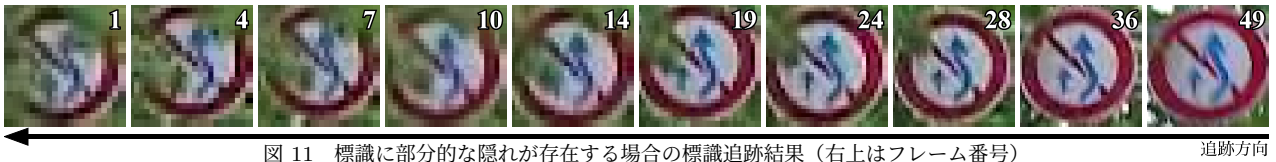


図11 標識に部分的な隠れが存在する場合の標識追跡結果 (右上はフレーム番号)

追跡方向

5. む す び

本稿では, 走行中に撮影される車載カメラ映像から多様な変動を含む標識画像を自動収集し, 高精度な標識検出器を構築する手法を提案した. 具体的には, まず近くで大きく撮影された標識を検出し, 検出した大きな標識を初期値として時間を遡りながら標識を追跡することで低品質の標識を切り出して収集した. 自動収集した標識画像を学習サンプルとして利用し, 標識検出器の構築を行った. また, 標識画像の自動収集と標識検出器の構築を繰り返し行うことで, 標識検出器の高精度化を図った. 提案手法を車載カメラ映像に対して適用した結果, 最大で F 値が 0.95 の検出精度を持つ標識検出器を構築可能であることを確認した.

今後の課題として, 自動収集した標識サンプルを用いたオンライン学習手法の開発, 大きな標識検出時の誤検出の抑制, 標識と背景の境界が不明瞭な場合における標識切り出し手法の改善, 多試料での評価, 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, JST 戦略的創造研究推進事業 CREST および文部科学省科学研究費補助金によった. 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] 内村圭一, 脇山慎也, 藤野麻衣子, “限定色表示を用いた円形道路標識の抽出”, 電子情報通信学会論文誌, Vol.J83-D-II, No.2, pp.855–858, February 2000.
- [2] 明珍甲太, 景山陽一, 西田眞, “カラー情景画像における円形道路標識の認識に関する検討”, 電子情報通信学会技術研究報告, PRMU2004-262, March 2005.
- [3] 竜円琢磨, 長坂保典, 鈴木宣夫, “円形状の道路標識の認識”, 電子情報通信学会技術研究報告, EID99-143, February 2000.
- [4] 荻木登, 三木成彦, “道路情景中の最高速度標識の検出および認識”, 電子情報通信学会論文誌, Vol.J77-D-II, No.7, pp.1393–1394, July 1994.
- [5] 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 2005, pp.255–260, June 2005.
- [6] K. Doman, D. Deguchi, T. Takahashi, Y. Mekada, I. Ide, and H. Murase, “Construction of cascaded traffic sign detector using generative learning”, Proceedings of International Conference on Innovative Computing Information and Control (ICICIC2009), ICICIC-2009-1362, December 2009.
- [7] P. Viola and M. Jones, “Robust real-time face detection,” International Journal of Computer Vision, Vol.57, No.2, pp.137–154, May 2004.
- [8] 岡田 和典, “ミーンシフトの原理と応用,” 電子情報通信学会技術研究報告, PRMU2007-308, March 2008.
- [9] I. D. Coope, “Circle fitting by linear and nonlinear least squares,” Journal of Optimization Theory and Applications Vol.76, No.2, pp.381–388, February 1993.
- [10] M. Hradis, A. Herout, and P. Zemcik, “Local rank patterns — novel features for rapid object detection,” Proceedings of the International Conference on Computer Vision and Graphics (ICCVG2008), Lecture Notes in Computer Science, Vol.5337, pp.239–248, November 2008.
- [11] R. E. Schapire and Y. Singer, “Improved boosting algorithms using confidence-rated predictions,” Machine Learning, Vol.37, No.3, pp.297–336, December 1999.
- [12] C. Huang, H. Ai, B. Wu, and S. Lao, “Boosting nested cascade detector for multi-view face detection,” Proceedings of the International Conference on Pattern Recognition (ICPR2004), Vol.2, pp.415–418, August 2004.