

局所特徴認識アルゴリズムによる車両の認識

吉田 達哉[†] 影澤 政隆[†] 塘中 哲也^{††} 池内 克史[†]

[†] 東京大学生産技術研究所
〒 153-8505 東京都目黒区駒場 4-6-1
^{††} 三菱重工業株式会社高砂研究所
〒 676-8686 兵庫県高砂市荒井町新浜 2-1-1

E-mail: [†]{yoshida,kagesawa,ki}@cvl.iis.u-tokyo.ac.jp, ^{††}tomonaka@wt.trdc.mhi.co.jp

あらまし 本稿で、我々はロバストな車両認識システムを紹介する。本システムの車両認識手法は binary features 法を利用している。binary features 法は、局所的な特徴に注目した物体認識手法である固有窓法を元にした手法であり、オクルージョンや照明条件の変動に対してロバストである。我々はこれまで、赤外線カメラによって取得した画像を用いて実験を行ってきたが、本稿では、広角のレンズを用い、屋外で撮影された可視光カメラの画像を利用して実験を行った。赤外線カメラより一般的な可視光カメラを使うことにより、さらに実用性の高いシステムとすることができる。実験では 95% 以上の高い認識率が得られ、システムの有効性が確認できた。

キーワード ITS, 車両認識, 画像処理, 固有窓法, Binary features 法

Vehicle Recognition with Local-Feature Based Algorithm

Tatsuya YOSHIDA[†], Masataka KAGESAWA[†], Tetsuya TOMONAKA^{††}, and Katsushi IKEUCHI[†]

[†] Institute of Industrial Science, University of Tokyo
Komaba 4-6-1, Meguro-ku, Tokyo, 153-8505 Japan

^{††} Takasago Research & Development Center, Mitsubishi Heavy Industries, Ltd., Shinhamma 2-1-1,
Arai-cho, Takasago, Hyogo, 676-8686 Japan

E-mail: [†]{yoshida,kagesawa,ki}@cvl.iis.u-tokyo.ac.jp, ^{††}tomonaka@wt.trdc.mhi.co.jp

Abstract In this paper, we describe a robust method for recognizing vehicles. Our system is based on local-feature configuration method, which is a generalization of eigen-window method. This method has following advances. It can detect even if part of vehicle is occluded, and it is robust against illumination changes. We have developed our system with infrared images, but in this paper, we apply our system on images of super wide-angle normal light images. In outdoor experiments, our system detects over 95% of vehicles.

Key words ITS, Vehicle recognition, Image processing, Eigen-window, Binary features method

1. はじめに

ITSにおいて、車両検出は基礎となる重要な技術であり、交通量計測、ETCなど幅広い応用分野で用いられている。現在車両検出には超音波センサやループディテクタなどが使われているが、近年では画像処理を利用したセンサも実用的なレベルに達しており、今後の発展、実用化が期待されている。従来のセンサでは車両を車両として認識することが困難であり、適用範囲が限られているが、画像認識技術を用いたセンサではそれが可能であり、またカメラという一般的なデバイスを用いることにより、他への応用、拡張範囲が広く、コスト面からも有効である。我々は、画像センサを利用した車両検出システムとして、局所特徴を利用した物体認識手法を元にした車両認識システムを提案して、赤外線カメラによって撮影された画像を用いて実験を行い、実用性を示してきた[1]。本稿では、このシステムを用いて、可視光画像での実験を行い、赤外線カメラと比較してより一般的なセンサである可視光カメラを用いたときの有効性を確かめる。

2. 認識アルゴリズム

我々のシステムにおける車両認識アルゴリズムは、物体認識の分野で発展してきた固有窓法[4]を元にした、Binary Features法[5]を使用している。固有窓法は固有空間法[3]を元にして改良を加えた手法である。固有窓法は、固有空間法において必要だった画像中における対象物体の切出し及び位置の補正を不要とし、オクルージョンにもロバストになっている。Binary Features法は固有窓法を2値のエッジ画像に適用した手法と考えることができ、照明変動に強く、計算が高速であるという特徴があり、屋外の画像をつかっての認識実験に適していると考えられる。以下で各手法について説明する。

2.1 固有空間法

固有空間法は3次元物体の2次元画像における「見え」を学習して、モデルをつくる手法である。モデル学習に主成分分析の手法を用い、寄与率の大きいベクトルのみで張られる低次元の空間（固有空間）でモデルを表現できる。画像間の相関が大きい場合は、必要な次元は少なく、高い圧縮率を得ることができる。また固有空間はその空間内の距離で類似度を表現しているので、認識に利用できる。

固有空間法では次に説明する主成分分析法が用いられている。

画像を N 次元のベクトル（画像サイズを縦 n 、横 m としたとき $n \times m = N$ ）としたとき、学習する M 個の画像 $z_1, z_2, \dots, z_M$ とするとき、平均画像 c との差分により以下の練習行列 Z を得る。

$$Z = [z_1 - c, z_2 - c, \dots, z_M - c] \quad (1)$$

式(1)から共分散行列 Q が得られる。

$$Q = ZZ^T \quad (2)$$

この Q から以下の関係式を満たす固有値 λ_i と固有ベクトル $e_i (i = 1, \dots, N)$ が求められる。

$$\lambda_i e_i = Q e_i (\lambda_1 \geq \lambda_2 \geq \dots \lambda_N \geq 0) \quad (3)$$

閾値 th を用いて、大きな固有値に対応した k 個の固有ベクトルのみで画像を近似する。

$$W = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^N \lambda_i} \geq th (k \ll N) \quad (4)$$

以上より得られた行列 $E = [e_1, e_2, \dots, e_k]$ を用いて、画像 z_i （次元 N ）を固有空間上の点 ζ_i （次元 k ）に投影することができ、各モデル画像を固有空間上の点で表現できる。3次元物体の2次元画像中の「見え」は物体の形状、反射特性、空間中での姿勢、及び照明環境で決まる。このうち物体の形状、反射特性は不変であり、姿勢、照明環境はシーンごとに変化することに注目し、姿勢と照明環境をパラメータとして変化させ、認識対象物体のモデルを撮影し学習していく。同一物体の場合、画像間の相関は高く、非常に高い圧縮率を得ることができる。

2.2 固有窓法

固有窓法は、小さなウィンドウを用いてあらかじめモデルを作成し、投票により認識を行う手法である。

モデルの作成は以下のように行う。モデル画像中のウィンドウは、以下のアルゴリズムにより、検出に際し有効なコーナーが選択可能である。

$$G = \sum_{\chi \in \mathcal{R}} \left(\frac{\partial I}{\partial \chi} \right) \left(\frac{\partial I}{\partial \chi} \right)^T \quad (5)$$

$$\min(\lambda_1, \lambda_2) > \lambda \quad (6)$$

ただし、 $\chi = (x, y)^T \in \mathcal{R}$ 、 I ：画像の輝度値
 λ_1, λ_2 ： G の固有ベクトル、 λ ：閾値

さらに、固有空間中での距離により類似度評価をし、紛らわしいものは排除する。

$$S_{l,m} = \|\phi_l - \phi_m\| \quad (7)$$

ただし、 ϕ_l, ϕ_m ：それぞれウィンドウ z_l, z_m の固有空間上の投影点

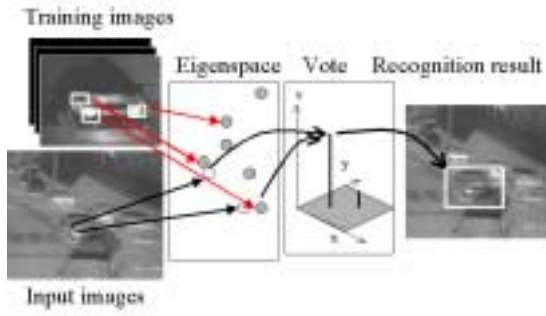


図1 固有窓法
Fig.1 Eigen window technique.

式(7)によりもとめた類似度が閾値 T_{sim} より低い場合、対応するウィンドウ z_l, z_m は類似度が大きく区別が難しい特徴であるとみなし、削除される。求めた有効なウィンドウを用い、モデルとする。得られたモデル画像 T_i 中のウィンドウ w_i を $w_i(T_i, z_i; x_i, y_i)$ と表記する。 z_i はウィンドウ内の画像、 (x_i, y_i) は T_i 中のウィンドウの位置である。

認識システムでは、以下のような処理を行う。入力画像 J 内の全点にモデル画像のウィンドウと同じ大きさのウィンドウ w を設け、これについても固有空間に投影する。入力画像中のウィンドウ w を $w(J, z; x, y)$ と表記する。 z はウィンドウ内の画像、 (x, y) は入力画像 J 中のウィンドウの位置である。投票空間 V として、 $M \times R \times R$ の空間を用意する。 M はモデル画像の数、 $R \times R$ はウィンドウの相対位置を表す XY 座標である。このとき、 w の投影点での、最近傍のウィンドウ w_k を検出する。つまり、固有空間法と同様の手法で、 z ともっとも類似した画像 z_k を求める。 w_k が見つかったら、投票空間中の $V(T_k, x - x_k, y - y_k)$ に投票する。投票空間 $V(T, x, y)$ の値が r であるならば、入力画像 J 中の相対位置 (x, y) にモデル T の特徴（ウィンドウ）が同じ位置関係で r 個見つかったことになる。入力画像 J 中にモデルが存在する場合、複数の特徴がモデルと同様の位置関係で見つかるため、ピークが立つ。投票の多い点を閾値 th で検出して、対象物体を認識する。

固有窓法では、対象物体を背景とセグメンテーションする必要がなく、オクルージョンが生じていても、閾値 th 以上のウィンドウが検出できれば認識が可能である。また、対象物体の入力画像内の位置を、モデル画像と一致させる必要がない。

2.3 Binary Features 法

Binary Features 法は、上述の固有窓法を2値化エッジ画像において適用した手法であると考えることができる。2値化エッジ画像を用いることにより、

照明変化に対し安定した処理と、計算の高速化を実現している。

原理は以下の通りである。モデル画像に対しガウシアンフィルタで太線化、ラプラシアンフィルタでエッジ検出を行う。次いで固有窓法のウィンドウと同様に特徴点（binary features）検出を行いウィンドウを設ける。2値画像における特徴点としては、近傍との相関性が低い点を選択する。

$$r_i(x, y) = \min_{\substack{-d \leq d_x \leq d \\ -d \leq d_y \leq d}} \{ D_H[w(K, z; x, y), w(K, z_d; x + d_x, y + d_y)] \} \\ (d_x, d_y) \neq (0, 0) \quad (8)$$

ただし、 $w(K, z; x, y)$:画像 K における位置 (x, y) のウィンドウ。 z はウィンドウ内の画像。 $D_H(\mathbf{r}, \mathbf{s})$: バイナリベクトル $\mathbf{r}, \mathbf{s}$ 間のハミング距離。

選択された特徴はLloydのアルゴリズムでコードに圧縮される。特徴に対し初期クラスタ中心を与えて、各特徴が、ハミング距離が最も近いクラスタ中心に属するとし、全ての特徴をグループ分けする。グループごとに質量中心を求め、新たなクラスタ中心とする。これを適当な回数繰り返すことにより各特徴を代表するコード（＝クラスタ中心）が求まる。バイナリベクトルの質量中心は各要素ごとに1か0の多数決を行うことで求められる。各特徴ごとに対応するコードとモデル画像中での位置を記憶しておく。

認識においては、入力画像を同様に2値エッジ画像とし、入力画像の全点に対しウィンドウを設け、コード化する（最も近いコードを求める）。固有窓法と同様に、入力画像中のウィンドウの位置と求めたコードに対応する特徴の位置の相対座標に投票し、閾値でピークを検出して対象物体を認識する。

上述の手法同様に、小さいウィンドウを用いているため、ノイズやオクルージョン、物体の画像内における平行移動にも対応できる。また固有窓法に比べ計算が大幅に単純化されており、特に浮動小数点演算を行う必要がないため、処理は早くなる。また2値化エッジ画像を用いているためグレースケールの画像を直接用いるよりも光源環境の変化にロバストである。

3. 車両認識システム

我々の車両認識システムは、画像からの物体認識を元にした認識システムである。モデル作成、車両認識処理に、前述のBinary Features法を使用してい

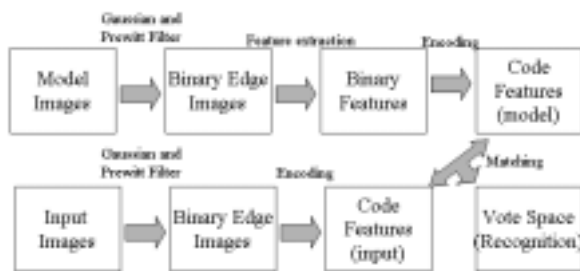


図2 車両認識システムの構成
Fig.2 Flow-chart of the vehicle recognition system.

る。システム全体の構成を図2に示す。これまで我々は、赤外線カメラの画像を入力画像に使用して実験を行い、システムの実用性を確かめてきた[1]。本稿では入力画像として、通常の可視光カメラの画像を使用し、実験を行う。可視光カメラで取得した画像のほうが、天候や時間変化等による変動が大きい、より一般的なデバイスであるため、拡張性、実用性が非常に高く、有効である。また、視野範囲を広げるため広角レンズを使用している。

3.1 モデル作成

モデル作成は以下の通りである。まず、モデルとなる車両の写っている画像を用意する。このとき車両を背景から切り出すための輪郭を用意する。モデルの写った画像にガウスフィルタ及びラプラシアンフィルタをかけて2値のエッジ画像とする。得られたエッジ画像から、輪郭を用いて車両部分を切出し、(8)式を用いて特徴抽出を行う。特徴は $n \times n$ サイズのウィンドウとして抽出する。得られた特徴をLloydのアルゴリズムによって m 個のコードに圧縮する。

3.2 認識

認識処理は以下のように行われる。入力画像に対して、モデル作成時と同様に、ガウスフィルタ及びラプラシアンフィルタをかけて2値のエッジ画像 $I_s(x, y)$ とする。得られたエッジ画像 $I_s(x, y)$ の全ての点において、その点を中心とした近傍 $n \times n$ のウィンドウを作り、そのウィンドウをコード化する。コード化とは、モデル作成で得られた m 個のコードとハミング距離を比較して最も近いものをそのウィンドウのコードとする処理のことである。

次に、各モデルにおいて投票空間を用意して、投票処理を行う。入力画像中の全点のコードを求めたら、そのコードと同じコードを持つモデル画像中の特徴を求め、入力画像中のその点の位置と、モデル画像中のその特徴の位置の差に投票を行う。これに

Models	9	Features	60
Codes(m)	20	Window size($n \times n$)	7×7
Threshold(T)	65	Iteration(Lloyd)	10

表1 実験におけるパラメータの設定

Table 1 Parameters of experiments

より、モデル画像における特徴のコードの位置関係と、入力画像中のコードの位置関係が等しい点に投票が重なり、投票が重なった点にモデル画像と同一の物体がある確率が高いということになる。投票空間の安定化のため、投票空間に 9×9 のフィルタをかける。フィルタの値は、中心が2で、他は1である。ある閾値 T 以上の投票があった点に対象車両があると見なし、認識する。

4. 屋外実験

3.章で述べた認識システムの性能評価をするために、実験を行った。また、背景差分を取り入れるシステムを試作し、性能の向上を試した。

4.1 評価実験

システムの性能を評価するため、以下のような実験を行った。

4.1.1 実験仕様

画像は広角レンズを使用した可視光カメラの画像で、カメラは上方から車を見下ろすように設置されている。この実験画像は、三菱重工業株式会社のテストコースで撮影されたものである。認識対象となるのは車体の色が白と黒の二種類の車両である。実験では対象となる2色の車を、共通のモデルで認識するものとする。

各パラメータ値は表1のように定める。広角のレンズを使っているため、画像中の写る個所によって車両の見えが異なる。そのためそれぞれに対応したモデルを用意しなければならない。そこで、先行して予備実験を行い、画面全体をカバーできるモデルの個数を求めた。この予備実験から、画像平面を縦横それぞれ3分割にして9個の領域を考え、それぞれに対応した9つのモデルを用意すれば十分であることがわかったため、9つのモデルを用意する。認識対象である2色の車両を単一のモデルで認識するため、モデル作成には、対応する位置にそれぞれの車両が写っている2枚の画像から作成する。実験で使った白い車両のモデル画像を図3に示す。黒い車両も同様にモデル画像を用意する。

特徴は、各色のモデル車両の画像から30ずつ抽出して、各モデルにつき60ずつ得る。得られた特徴の

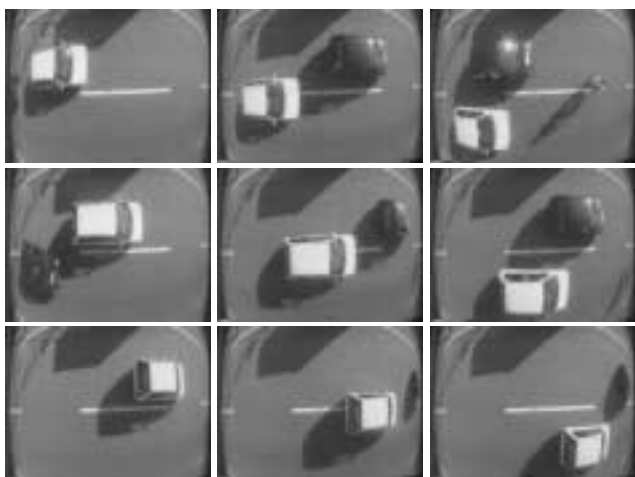


図 3 白い車両のモデル画像
Fig.3 Model images of a white vehicle.



図 4 抽出された特徴
Fig.4 Extracted features.



図 5 特徴から生成されたコード
Fig.5 Codes made from features.

例を図 4 に示す。

得られた 540 個の特徴をコード化の処理で 20 のコードに圧縮する。求めたコードを図 5 に示す。Lloyd のアルゴリズムの反復回数は 10 回である。

車両を検出する認識の閾値は 65 である。閾値は予備実験を行い適当な値を選んだ。これは、フィルタをかけられた投票空間に対しての値である。

4.1.2 実験結果

テストコースで撮影された 259 枚の画像を用いて認識実験を行った。実験画像例を図 6 に示す。実験に使用した画像中、白い車両が写っているものが 137 枚あり、黒い車両が写っているものが 227 枚ある。画像中には、2 台の車両が近接して走行しているもの、車両と自転車が並んで走行しているものなど、認識が困難であると考えられるものも含まれる。また、全ての画像で影がくっきりと写っており、認識の妨げ

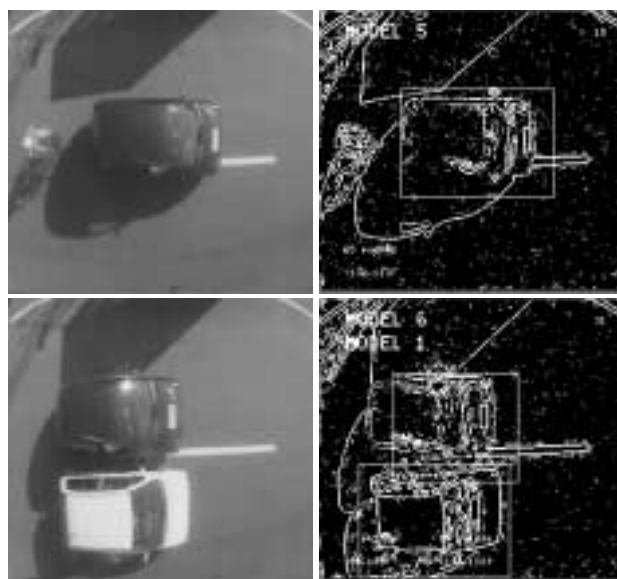


図 6 実験画像例 左:入力画像, 右:認識結果
Fig.6 Examples of experiments. Left images are input images, right images are recognition results.

	Success	Ratio
White vehicle	135/137	98.5 %
Black vehicle	129/227	56.8 %
False alarms	13/259	5.0 %

表 2 実験結果

Table 2 Results of experiments

になっている。実験結果を表 2 に示す。各車両の位置も含めた認識成功率と、車両がないところに車両があると判定した誤認識の全体に占める割合を示している。白い車両は高い認識率を得られたが、黒い車両は影の影響でエッジが消えてしまうことがあり、見落とされることが多かった。

4.2 背景差分の実験

認識率の向上を考え、認識処理時に以下のような背景差分処理を行うシステムを考え、実装して実験を行った。

4.2.1 背景差分処理

背景差分の概要を図 7 に示す。入力画像が与えられたとき、モデル作成時と同様にガウスフィルタ及びラプラシアンフィルタをかけて 2 値のエッジ画像とする。事前に背景画像を用意しておき、同様に背景も 2 値のエッジ画像とする。背景画像は入力画像と比較して、一定の間隔で更新される。得られたエッジ画像間で背景差分を行う。背景画像で 1 の部分が入力画像で 0 になる。つまり、2 値の入力画像 $I_i(x, y)$, 2 値の背景画像を $I_b(x, y)$, 差分後の画像を $I_s(x, y)$ とすると、

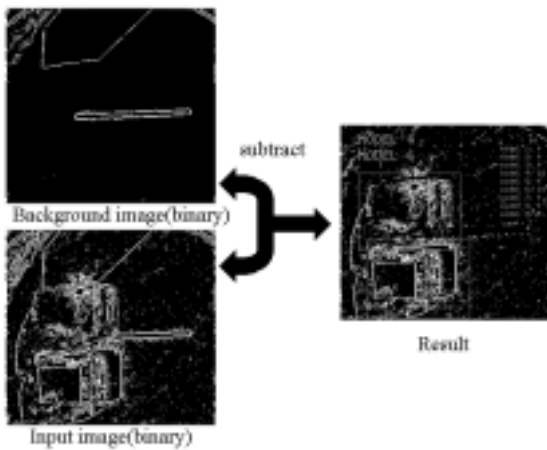


図 7 背景差分処理
Fig.7 Background subtraction.

	Success	Ratio
White vehicle	141/143	98.6 %
Black vehicle	226/234	96.6 %
False alarms	18/262	6.9 %

表 3 背景差分後の実験結果
Table 3 Results of experiments using background subtraction.

$$I_s(x, y) = \begin{cases} 1 & (I_i(x, y) = 1, I_b(x, y) = 0 \text{ のとき}) \\ 0 & (\text{それ以外のとき}) \end{cases} \quad (9)$$

である。入力画像 $I_i(x, y)$ のかわりに、得られた差分画像 $I_s(x, y)$ を認識処理に使う。

4.2.2 実験結果

先の実験と同様の環境で、背景差分を用いた認識実験を行った。ただし背景の分の投票が減ったため、認識の閾値のパラメータを 65 から 45 に変更した。実験には、テストコースで撮影された 262 枚の画像を使った。実験画像例を図 8 に示す。実験に使用した画像中、白い車両が写っているものが 143 枚あり、黒い車両が写っているものが 234 枚ある。実験結果を表 3 に示す。各値の意味は、先の実験と同様である。先の実験では、黒い車両において、影の影響で輪郭が消えエッジが検出されにくいという問題が起こったが、背景差分をして閾値を下げることで、白い車両と変わらない高い認識率が得られた。

5. ま と め

屋外での可視光カメラからの画像を用いた実験で、我々のシステムの有効性を確認することができた。背景差分処理を行った実験で得られた認識率は 95% 以上であり、非常に有効な結果であるといえる。我々のシステムでは、従来の超音波センサ等では困難である、車両と他の物体の区別や、車両の位置の特定が可能であり、DSRC 等での応用が考えられる。ま

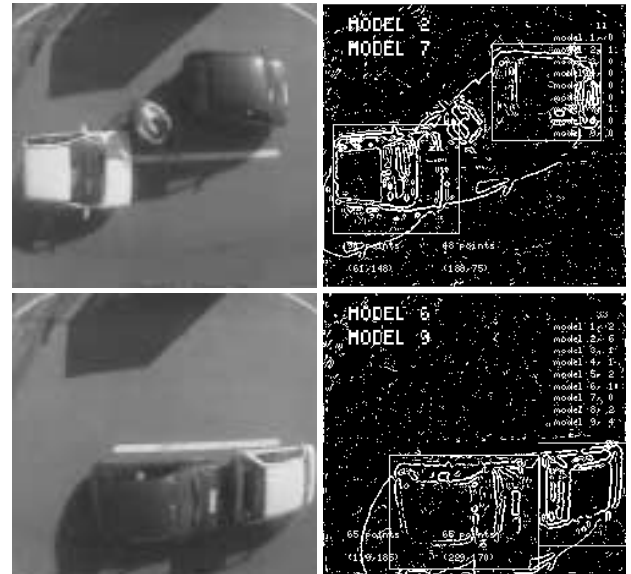


図 8 背景差分後の実験画像例 左:入力画像, 右:認識結果

Fig.8 Examples of experiments using background subtraction. Left images are input images, right images are recognition results.

た、我々のシステムは画像センサを用いており、監視カメラとの併用や、他機能の拡張が容易であり、実用性の高いシステムである。今後は、我々が取り組んでいる車種判別手法 [2] と組み合わせて、車種判別を行うシステムを検討していく予定である。

文 献

- [1] M. Kagesawa, S.Ueno, K. Ikeuchi and H. Kashiwagi, "Vehicle Recognition in Infra-red Images Using Parallel Vision Board," 6th World Congress on Intelligent Transport Systems '99, Toronto, U.S.A., Nov. 1999.
- [2] M. Kagesawa, A. Nakamura, K. Ikeuchi and H. Saito, "Vehicle Type Classification in Infra-red Image Using Parallel Vision Board," 7th World Congress on Intelligent Transport Systems 2000, Torino, Italy, Nov. 2000.
- [3] H. Murese and S. K. Nayar, "Visual Learning and Recognition of 3-D Object from Appearance," International Journal of Computer Vision, vol.14, no.1, pp.5-24, Jan. 1995.
- [4] K. Ohba and K. Ikeuchi, "Detectability, Uniqueness, and Reliability of Eigen-Windows for Stable Verifications of Partially Occluded Objects," IEEE Pattern Recognition and Machine Intelligence, vol.19, no.9, pp.1043-1048, Sept. 1997.
- [5] J. Krumm, "Object Recognition with Vector Quantized Binary Features," Proc. of Computer Vision and Pattern Recognition '97, pp.179-185, San Juan, Puerto Rico, June 1997.