

修士論文

ベクトル量子化を用いた車両の認識とクラス
分類

Recognition and Classification of Vehicles
by Using Vector Quantization

2002年2月1日提出

指導教官

池内 克史 教授

東京大学大学院 工学系研究科 電子情報工学専攻

06402 吉田 達哉

内容梗概

近年注目されている研究分野に ITS がある。ITS は Intelligent Transportation Systems の略称で日本では高度道路交通システムと呼ばれている。ITS とは最新の科学技術を用いて、交通における諸問題を解決し知的で高度なサービスを提供することを目的とした新しい交通システムのことである。その中でセンシング技術は通信技術と並び必要度の高い基礎技術であり、高い性能が求められている。現在、画像センサに比べスポットセンサのほうがセンシング精度が高いといわれているが、様々な画像処理、認識技術と組み合わせることにより、画像センサも実用レベルに近づきつつある。画像センサはスポットセンサに比較して広範囲のセンシングが可能であり、状況認識、速度認識、クラス分類等の高度な機能を持ちうる有効なセンサであり、今後の発展が期待されている。

本稿では画像センサのうち、設置の容易性やコストの優位性を考慮し、最も望ましい単眼の可視光カメラを入力デバイスとし、高精度で車両を検出する車両認識システムを提案、実装する。可視光カメラは環境変動の影響を受けやすく、屋外の使用では画像の変動が大きい。またオクルージョンやぶれといった画像センサ特有の問題が起こる。提案するシステムはこれらの変動にロバストな認識手法を応用することにより問題の解決をはかる。実装したシステムの性能を屋外で撮影した実際の画像を用いて実験をすることにより確かめる。また、ニーズの高いクラス分類を認識システムの延長として提案し、認識システムと同様に実画像に対する実験を通して性能を評価する。

実験結果から、車両認識システムの高い認識精度とロバスト性が確認され、画像センサから車両のクラス分類を行うことのできる可能性が示された。

目次

内容梗概	i
1 序論	1
1.1 本研究の背景	2
1.2 本研究の目的	3
1.3 関連研究	3
1.4 本論文の構成	8
2 物体認識手法	9
2.1 物体認識手法と車両認識	10
2.2 画像認識アルゴリズム	11
2.2.1 固有空間法	11
2.2.2 固有窓法	14
2.2.3 Binary Feature 法	16
3 車両認識システム	19
3.1 車両認識アルゴリズム	20
3.1.1 画像処理ボードによる高速化	20
3.1.2 認識システムの実装	20
4 車両認識実験	26
4.1 実験背景	27
4.2 実験仕様	27
4.3 実験結果	31
5 クラス分類手法	36

5.1	課題	37
5.2	コンピュータグラフィックスによるモデル生成	37
5.3	クラス認識システム	38
6	クラス分類実験	40
6.1	実験仕様	41
6.2	結果	43
7	結論	51
7.1	実験結果の考察	52
7.1.1	車両認識実験	52
7.1.2	クラス分類実験	53
7.1.3	まとめ	53
7.2	今後の課題	54
	謝辞	56
	発表文献	57
	参考文献	58

目 次

1.1	ミネソタ大学による Blob Tracking Method の実験結果例	4
1.2	Model-Based Vehicle Recognition 実験の車両検出例	5
1.3	Model-Based Vehicle Recognition 実験の固定カメラでの実験例 . . .	6
1.4	Model-Based Vehicle Recognition 実験の車両上のカメラでの実験例 .	6
1.5	Silhouette Vision の実験例	7
2.1	固有空間法におけるモデル画像群の例	12
2.2	図 2.1 の画像群から求められた固有ベクトル画像	13
2.3	固有窓法の概要図	15
2.4	Binary Feature 法	17
3.1	IMAP-VISION ボードの構成図	21
3.2	車両認識システムの処理の流れ	22
3.3	投票操作の略図	24
3.4	エッジ画像での背景差分処理	25
4.1	テストコースで撮影された画像の例	27
4.2	画像を 9 分割した時の各位置における白い車両のモデル画像	29
4.3	画像を 9 分割した時の各位置における黒い車両のモデル画像	30
4.4	マスク画像の一例	30
4.5	モデル画像から抽出した特徴	32
4.6	抽出された特徴から生成したコード	32
4.7	車両認識実験の成功結果例	34
4.8	車両認識実験の失敗結果例	35
5.1	車両クラス分類システムの処理の流れ	38

6.1	歩道橋上から撮影された車両の画像例	42
6.2	実験に使用したセダンクラスのモデル画像	44
6.3	実験に使用したワゴンクラスのモデル画像	44
6.4	実験に使用したワンボックスクラスのモデル画像	45
6.5	実験に使用したハッチバッククラスのモデル画像	45
6.6	セダンクラスのモデル画像から抽出した特徴	46
6.7	ワゴンクラスのモデル画像から抽出した特徴	46
6.8	ワンボックスクラスのモデル画像から抽出した特徴	47
6.9	ハッチバッククラスのモデル画像から抽出した特徴	47
6.10	抽出された特徴から生成したコード	48
6.11	クラス分類実験の成功結果例	49
6.12	クラス分類実験の失敗結果例	50

表 目 次

4.1	認識実験の結果	33
6.1	クラス識別実験の結果	48

第 1 章

序論

1.1 本研究の背景

近年，発展した科学技術を利用して交通事故，渋滞といった道路交通問題を解決すること，自動走行，自動料金収受システムといった，快適な道路交通環境を提供することを目的とする，高度道路交通システム (ITS: Intelligent Transportation Systems) が注目を集めている．日本では国土交通省が中心となり，9 つの開発分野として，ナビゲーションの高度化，自動料金収受システム (ETC: Electronic Toll Collection System)，安全運転の支援，交通管理の最適化，道路管理の効率化，公共交通の支援，商用車の支援，歩行者等の支援，緊急車両の運転支援を挙げて，現在盛んに研究，開発が進められている．

ITS において中心となる技術の一つに，センシングがある．高度な情報やサービスの提供には，対象となる車両の位置や現在の交通状況などの情報を取得することが必要不可欠であり，センシングは情報通信技術と並んで，ITS における主要な基礎技術である．

センシングの中でも，得に車両検出は種々のサービスにおいて必要とされる重要な技術である．車両検出に使用されるセンサは，CCD カメラ，赤外線カメラなどの画像センサと，超音波センサ，ループディテクタ，レーザセンサなどのスポットセンサとに大別することができる．両センサにはそれぞれ異なった特徴がある．画像センサは広範囲をセンシングすることができ，画像認識，解析手法と組み合わせることで多くの情報を得ることができる．またデバイスのコストも低く設置も容易であるため，実用化と普及を考えた場合非常に有利である．しかし，画像センサは環境変動による影響を受けやすく，環境変化の大きい屋外での使用を前提としている車両検出では不安定なセンサであると言われている．また，オクルージョン（隠蔽）やぶれといった画像センサ特有の問題がある．一方，スポットセンサは対象となる地点において高い検出精度を持つため，高速道路などで実用化が進んでいる．しかし，広範囲での状況取得には不向きであり，車両と他の物体を区別することが困難であるため，例えば歩行者，自転車，看板や建物といった，検出対象である車両以外の障害物が多い市街地などでの実用化には問題がある．近年要求されている速度検出，車両のクラス分類，道路状況認識，事故検出などの高度な機能を実現することも，研究例 [23] はあるが同様の理由から向いていないと思われる．

1.2 本研究の目的

このような背景から，本研究では車両検出において今後の実用上有効であると考えられる画像センサに着目し，画像センサで問題となっているオクルージョンや環境変動といった問題に対する解決策として，それらにロバストなコンピュータビジョン分野の物体認識手法を用いることを考える．実際に画像センサを用いて高精度の車両検出を行うシステムを開発し，実装する．また，ETC や道路状況認識等の用途において，近年期待されている知的交通システムとして，センサから車両の種類を自動判別するシステムがある．そこで車両認識システムを発展させて，車種（クラス）を判別するシステムについても開発し実装する．どちらのシステムにおいても，屋外で撮影された実画像を使用して実験を行いシステムの有効性を検討する．

1.3 関連研究

画像センサは比較的歴史の長いデバイスであるため，画像センサを入力として車両を認識する手法は古くから研究されてきた．多くの手法が現在も研究されており，良好な結果を出しているものもあるが，それぞれ長所短所がありまだ研究段階であるといえる．ここでは，幾つかの代表的な研究を紹介する．

Blob Tracking Method

ミネソタ大学で研究されている Blob Tracking Method は，道路脇に備え付けられた 1 台のカメラから取得される画像を用いて，車両を検出する手法である [20]．車両のモデルは使用せず，幾つかの画像処理手法と拘束条件から車両を検出する．処理は以下のような流れになる．

- フレーム間差分処理により移動領域を検出する．
- 検出した差分画像を領域ごとに区切りラベル付けする．
- 各領域の周囲長，面積といったパラメータを求め，連続したフレーム間で領域 (blobs) ごとの関係付けをする．例えばフレーム i の領域 1 はフレーム $i+1$ の領域 2 であるというような関係グラフ (blob graph) を作成する．その際，領域は同時に分割と統合はしないという拘束条件を用いて可能性のあるグラフの候補を減らし，最適なグラフを求める．

- 得られたグラフから各領域を車両とみただて，長さ，幅，高さといったパラメータを求める．
- パラメータが車両と認められる範囲であるかを考慮し，複数の領域で車両が形成されているかどうかを求める．
- 領域 (blobs) レベルと車両レベルでトラッキングを行う．カルマンフィルタを用いてロバストなトラッキングを行う．

実験では入力画像シーケンス中において 85% の車両のトラッキングに成功している．実験結果の例を図 1.1 に示す．検出された車両を長方形で表している．失敗は大型車が 2 台と判定される，大型車で小型車が隠蔽されるといった状況で起こることが多い．



図 1.1: ミネソタ大学による Blob Tracking Method の実験結果例．

Model Based Vehicle Recognition

リーディング大学で研究されている車両認識システムは，変形可能な車両モデルを元に，単眼カメラの画像から車両をトラッキングする手法である [11][12][13][17]．自由度を持たせた三次元のワイヤフレームモデルを用意することにより，個々の車両における形状の変化を吸収し統一のモデルで扱うことができている．モデルと画像中の車両のマッチングは以下のようにして求める．まず，ワイヤフレームモデルの横方向を x ，縦方向（進行方向）を y ，垂直方向を z とし、モデルを入力画像内の道路上に配置したときの，各座標軸に投影されるモデルのフレームの累積を求

める ($m_x(u)$, $m_y(u)$, $m_z(u)$) . 画像中のグラディエントエッジについても同様の座標軸に投影して累積を求める ($s_x(u)$, $s_y(u)$, $s_z(u)$) . 両者の相関 ($c_x(u)$, $c_y(u)$, $c_z(u)$) を求めて相関の最大値に車両があると判断する . この操作は一次元のテンプレートマッチングと考えることができる . 各投影のヒストグラムを示した様子を図 1.2 に示す .

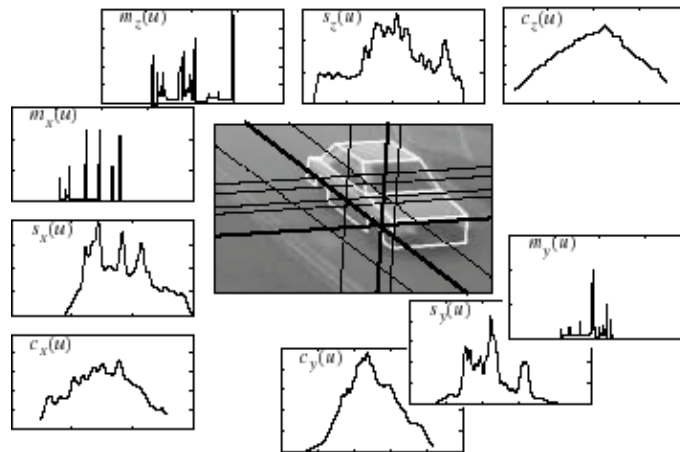


図 1.2: リーディング大学による Model-Based Vehicle Recognition 実験 . 一次元テンプレートマッチングによる車両の検出例 . 各軸のモデルフレームの投影 (m) , 画像エッジの投影 (s) と両者の相関 (c) が示されている .

このワイヤフレームモデルを使って , いくつかの状況で車両のトラッキング実験を行っている . 固定カメラで高速道路を撮影した画像を用いてトラッキングを行った実験では [13] , ROI (Region Of Interest: 関心領域) を設定し , ROI に対して上述したマッチング処理を行って領域に入ってきた車両の検出を行う . その後連続して検出した車両を時空間座標でトラッキングする . 車種ごとにモデルを持ち , より相関の高いモデルを求めることにより車種判別も行っている . 実験結果の画像例を図 1.3 に示す . 検出された車両上にワイヤフレームモデルが示されている . 車両にカメラを固定し , 移動する車両上で撮影した画像を使った実験では [17] , 白線などの入力画像内の路面の特徴をトラックすることにより自己位置を検出し , 上述の手法で画像内の車両をトラッキングし , 周囲の車両との衝突警報機能を付けたシステムを構築している . 実験画像例を図 1.4 に示す . 自己位置と画像内の車両位置を検出して , その位置関係を異なる視点からの画像としてワイヤフレームで示している .



図 1.3: リーディング大学による Model-Based Vehicle Recognition 実験 . 固定カメラで高速道路を撮影したときの車両検出結果例 .

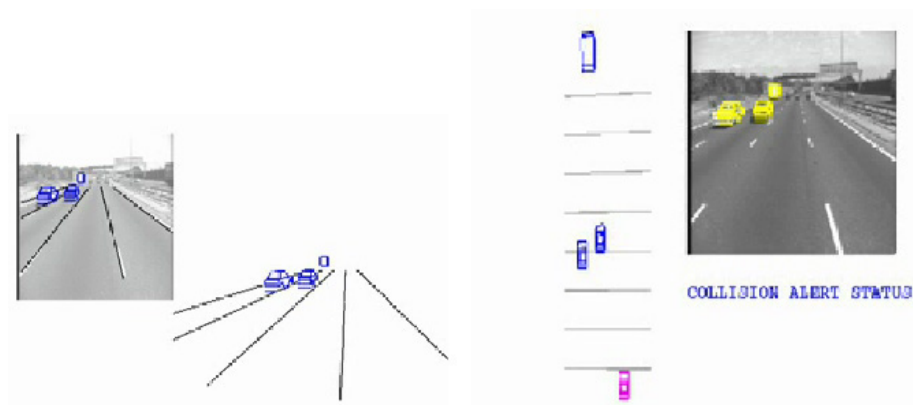


図 1.4: リーディング大学による Model-Based Vehicle Recognition 実験 . 車両に固定されたカメラで前方の車両を検出する実験例 .

Silhouette Vision

Silhouette Vision[16] はオムロン株式会社が研究、開発した車両計測のための画像センサである。Silhouette Vision は屋外での環境変動にロバストなモデルとして車両側面の形状(サイドシルエット)に着目する。入力デバイスとしてステレオカメラを用いて、車両の特徴点を抽出した後、ステレオ画像処理により特徴点の空間位置を求める。特徴点抽出領域は車両の走行領域や車高等の事前知識を用いて画像内の一部に限定することができ、主にフロントガラス上下やルーフ前後の水平エッジが抽出される。特徴点の空間位置を、車両の側面に平行になるようあらかじめ設定した平面に投影することでサイドシルエットを得ることができる。得られたサイドシルエットをモデルとマッチングすることにより車両抽出が可能になる。専用のハードウェアに実装することにより 200ms 前後の処理速度で車両認識、追跡ができていいる。歩道橋上で撮影されたビデオ画像を用いて実験した結果、路面に影の映る晴れた日の昼、オクルージョンの起こりやすい渋滞時、コントラストの落ちる薄暮時にも 97%以上の通過台数計測性能が確認されている。実験画像例を図 1.5 に示す。左が入力画像で左が検出結果である。



図 1.5: オムロン(株)が研究している Silhouette Vision。左が入力画像，右は検出結果。

1.4 本論文の構成

本章では、本研究の背景、目的を述べ、関連研究の動向について簡単に紹介した。第 2 章では、屋外で撮影された画像から対象となる車両を認識する手法について、その元になるコンピュータビジョンの画像認識手法からシステムに使用する手法まで述べ、第 3 章では提案、実装するシステムの構成について述べる。

第 4 章では、第 3 章で紹介した車両認識システムの性能を検討するため、実際に屋外で取得した画像を用いて行った実験とその結果について示す。

第 5 章ではクラス分類手法を提案し、我々の車両認識システムへ導入する。車両認識とクラス分類を行う新しいシステムについて、原理と構成を述べる。クラス分類手法を導入したシステムに対しても、車両認識システムの場合と同様に屋外で撮影された実画像を用いて行った性能検証実験とその結果について、第 6 章で述べる。

最後に第 7 章で本論文の実験結果を元に手法やシステムの検討を行い、本研究の成果や今後の課題についてまとめる。

第 2 章

物体認識手法

2.1 物体認識手法と車両認識

コンピュータビジョンの分野では古くから物体認識の手法が研究されてきた。画像中から対象物体を識別することは、ロボット、モニタリング、ターゲットディテクションなど多くの分野でニーズがあり、積極的に様々な手法が研究、開発されてきた。なかでも対象をモデルとして保持し入力画像とのマッチングをとって対象を認識するモデルベース手法は、その有効性から長らく研究されている。

モデルベース手法としては、対象を三次元で保持して入力二次元画像とマッチングさせるものと、対象も二次元で保持し二次元同士でマッチングさせる手法とに大別できる。三次元のモデルを保持する手法は3DTM[8]や第1章で述べたワイヤフレームモデルによる車両検出法[12]などがある。3DTMは三次元のテンプレートマッチングのことであり、予め用意した対象物体の三次元ワイヤフレームモデルを用いて、入力画像中の対象物体（二次元）にマッチングさせる手法である。三次元のモデルを二次元画像にマッチングさせるため、二次元でモデルを保持する手法に比べて対象の三次元における位置及び姿勢を認識できるといったメリットがあるが、処理の負荷が大きいという問題がある。ただしこの問題も近年の計算機の性能向上に伴い、実時間でトラッキングを行うことに支障をきたすほどでは無くなりつつある。

一方の二次元でモデルを保持する手法としては固有空間法[10]などがある。二次元でモデルを保持することは、処理の速度及びメモリ資源の消費が抑えられるという利点がある。本研究のように対象を車両とした場合、三次元中の位置はかなりの拘束を持っており、二次元のモデルでも位置や姿勢を推定可能である。また対象の種類が多い場合、多くのモデルを保持しなければならないことと、車両認識では実時間性が重視され、処理の負荷が軽いことが望まれるため、二次元でモデルを保持する手法が有効であると考えられる。

固有空間法から派生した手法に Binary Feature 法[14]がある。Binary Feature 法は固有窓法[15]にベクトル量子化を適用して二値画像に対して処理を行う手法と考えることができる。Binary Feature 法はモデルを二値化した局所画像で保持するため、環境変化にロバストであり、処理が高速である。この特徴は屋外での使用と実時間処理が求められる車両認識システムに適していると考えられる。そこで、本研究ではこれらの手法を元にしベクトル量子化で二値画像を取り扱う手法を認識手法とする車両認識システムを提案する。

以下、本章では認識手法の理論について固有空間法から順に述べる。

2.2 画像認識アルゴリズム

認識システムに利用する Binary Feature 法は固有空間法から派生した固有窓法とベクトル量子化を組み合わせた手法と考えることができる。以下でこれらの手法について順に説明する。

2.2.1 固有空間法

概要

固有空間法は 3 次元物体の 2 次元画像中の「見え」を自動的に学習してモデルをつくり，作成したモデルから対象を認識する手法である。モデル学習に主成分分析の手法を応用し，低次元に圧縮したモデルをつくることができる。対象物体のモデル画像を画素数次元のベクトルと考えたとき，主成分分析でモデル画像ベクトルの固有ベクトルを求めることができる。このときモデル画像は求めた固有ベクトルの線形和と考えることができる。このとき寄与率の大きな固有ベクトルのみ用いることで，再現性をほとんど損なうことなく，モデル画像の次元を減少させることができる。モデル画像間の相関が高い場合，必要な次元数は少なく，高い圧縮率を得ることができる。また固有空間は画像間の類似度を表現しており，モデル画像と入力画像を固有空間上に投影してマッチングを行うことにより，認識に利用することができる。以下でその原理を紹介する。

主成分分析

固有空間法の基礎となる主成分分析 (PCA:Principal Component Analysis) 法 [4] をはじめに説明する。

画像を N 次元の列ベクトルとし (画像サイズ $n \times m = N$) 学習する M 個の画像を $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_M$ とするとき，平均画像 \mathbf{c} との差分により以下の練習行列 \mathbf{Z} を得る。

$$\mathbf{Z} = [\mathbf{z}_1 - \mathbf{c}, \mathbf{z}_2 - \mathbf{c}, \dots, \mathbf{z}_M - \mathbf{c}] \quad (2.1)$$

式 (2.1) から共分散行列 \mathbf{Q} が得られる。

$$\mathbf{Q} = \mathbf{Z}\mathbf{Z}^T \quad (2.2)$$

この \mathbf{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 $\mathbf{e}_i (i = 1, \dots, N)$ が求め

られる．

$$\lambda_i \mathbf{e}_i = \mathbf{Q} \mathbf{e}_i \quad (2.3)$$

閾値 T_ζ を用いて，大きな固有値に対応した，つまり寄与率の大きい k 個の固有ベクトルのみで画像を近似する．

$$W_k = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^N \lambda_i} \geq T_\zeta \quad (2.4)$$

以上より得られた行列 $\mathbf{E} = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k]$ を用いて，画像 \mathbf{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．

固有空間形成



図 2.1: 固有空間法におけるモデル画像群の例．

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．種々の光源環境及び姿勢で撮影された物体の画像は，サイズと明るさで正規化する．得られた画像群（図 2.1）

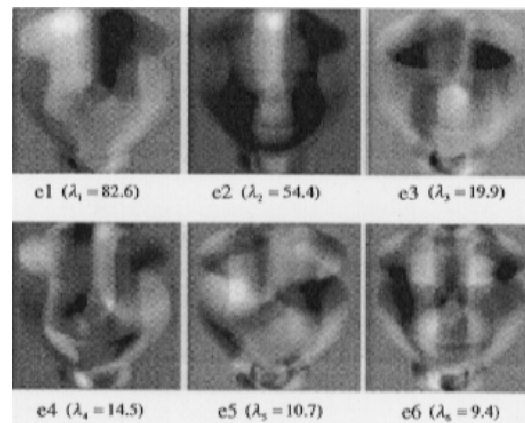


図 2.2: 図 2.1 の画像群から求められた固有ベクトル画像 .

に上述の手法を用いて固有ベクトル (図 2.2) を求め、各画像を固有空間上に投影する。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。

固有空間上での投影点間は、パラメータ変化において隣接した画像間では「見え」は大きく変化しないという仮定のもと、キュービックスプライン法 [6] で補間する。

複数の物体を認識対象とする場合、全物体の学習画像を用いて物体認識のための universal eigenspace を形成し、各物体ごとに姿勢決定のための object eigenspace を形成する。

認識手法

入力画像から物体領域を切出し、学習画像と同様に背景からのセグメンテーションをし、サイズ・明るさの正規化を行う。得られた正規化画像を universal eigenspace に投影し、雑音等を考慮した閾値以内の距離にある投影点に対応する物体と同一であると見なす。次いでその物体の object eigenspace に投影して、最近傍にある投影点に対応する姿勢及び照明環境であると見なす。以上の操作で入力画像に対応したモデル物体とその姿勢を求めることができる。

評価

固有空間法は物体の幾何形状モデルや反射特性を考える必要がないため、自動でモデルの取得ができる。また、効率よく圧縮されたモデルで物体認識を行うことが

できるため非常に有効な手法である。しかしデータベースの認識手法であるため、対象物体のスケールがモデルと大きく異なると対応できないほか、固有空間法には以下にのべるような幾つかの限界がある。

- セグメンテーションを行う必要があるため、使用できる環境が限られる。
- オクルージョンが生じないことが要求される。
- 背景ノイズ、照明環境、物体の移動・回転の影響を受けやすい。

これらの限界は屋外で撮影する車両認識では大きな問題となる。

2.2.2 固有窓法

固有窓法は局所領域に注目し、固有空間法を局所領域に適用する手法である。モデル画像において、画像全体を使うのではなく、検出しやすい特徴点の周辺の小領域に対して固有空間法を適用し、位置を考慮した投票方式で対象を認識する。これにより対象物体が部分的に隠蔽していたり（オクルージョン）、背景の変動等により対象物体がセグメンテーションできないといった固有空間法では困難な状況でも対象を認識できる。

原理

固有窓法においても、固有空間法と同様に主成分分析を使い画像ベクトルから固有ベクトルを計算して固有空間を求める。固有空間の作成にあたって、固有空間法のようにモデル画像全体に対してではなく、そのモデル画像に小さいウインドウを設け、その辞書ウインドウに対して主成分分析を行うことにより、固有空間を形成する。モデル画像中の各辞書ウインドウの位置は、投票により認識を行うさいに必要となるので記憶しておく。

モデル画像中の辞書ウインドウは、以下で説明するアルゴリズムにより検出容易性（detectability）、独自性（uniqueness）、安定性（reliability）の評価を行い、検出に有効なものを選択する。

検出容易性の評価は、対象物体中の検出、追跡しやすい特徴点を抽出することである。特頂点は以下の式のようなコーナー検出アルゴリズム [7] で求めることができる。

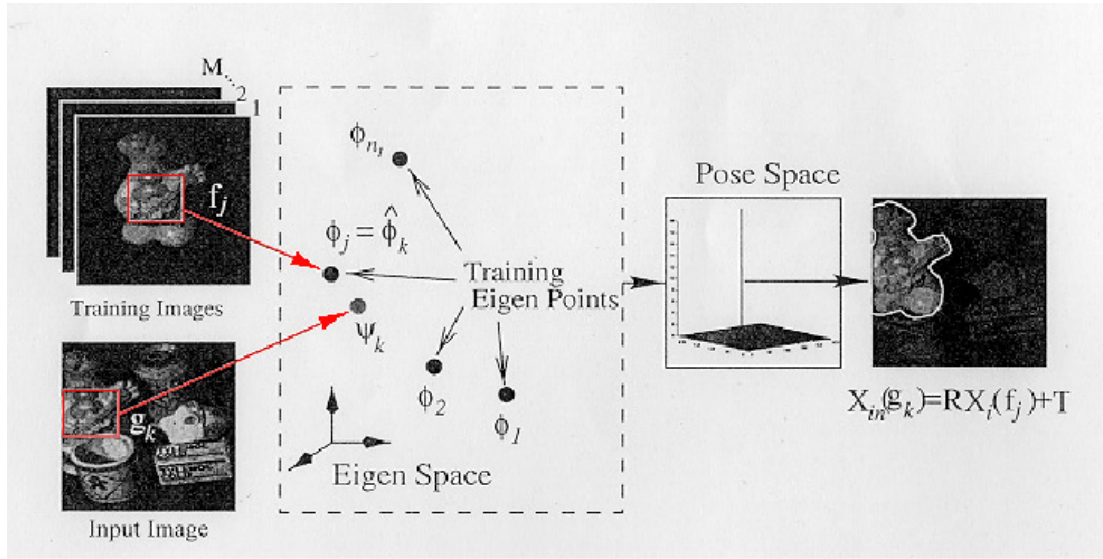


図 2.3: 固有窓法の概要図 .

$$G = \sum_{\chi \in \mathcal{R}} \left(\frac{\partial I}{\partial \chi} \right) \left(\frac{\partial I}{\partial \chi} \right)^T \quad (2.5)$$

$$\min(\lambda_1, \lambda_2) > \lambda \quad (2.6)$$

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

式 (2.5) で選ばれた特徴点を, 固有空間中での距離により類似度を評価をし, 紛らわしいものは排除する (独自性). つまり,

$$S_{l,m} = \|\phi_l - \phi_m\| \quad (2.7)$$

ただし, ϕ_l, ϕ_m : それぞれ辞書ウィンドウ f_l, f_m の固有空間上の投影点

式 (2.7) によりもとめた類似度が閾値 T_{sim} より低い場合, 対応する辞書ウィンドウ f_l, f_m は類似性が高く区別が難しい特徴であるとみなし, 削除される.

安定性の評価は, 物体が回転等により「みえ」が変化する際に, 特徴点の画像がどれだけ安定であるかを評価するもので, 大きく「みえ」の変化する特徴点は不安定と判断され削除される. 安定性は以下の式により表現できる.

$$R_j = \sum_{\delta \in \pm\Delta} \|\phi_j^{+\delta} - \phi_j\| \leq T_{rel} \quad (2.8)$$

ただし, $\phi_j, \phi_j^{+\delta}$: それぞれ辞書ウインドウ $f_j, f_j^{+\delta}$ の固有空間上の投影点, $f_j^{+\delta}$: 対象物体を角度 δ 回転させたときの辞書ウインドウ f_j T_{rel} : 閾値

以上の評価により求めた有効な辞書ウインドウを用いることにより, 安定した物体認識が可能になる.

認識は以下のように行う. まず, 各モデル画像ごとに画像サイズ分の投票空間を準備する. 次に入力画像全点に辞書ウインドウと同じサイズのウインドウを設け, そのウインドウを固有空間に投影する. 投影点での最近傍の辞書ウインドウを検出し, その辞書ウインドウが属するモデル画像の投票空間に対し, 入力画像中のウインドウ位置とモデル画像中の辞書ウインドウの位置との相対位置に投票を行う. モデル画像 T の投票空間を $V(T, X, Y)$ (X, Y は投票空間の座標) とし, 入力画像 J の点 (x, y) のウインドウを $w(J; x, y)$ とする. また, w を固有空間に投影した時の最近傍の辞書ウインドウが $w_i(T_i; x_i, y_i)$ であるとする. ここで T_i は w_i が属するモデル画像, (x_i, y_i) は T_i 中における w_i の位置である. このとき, 投票空間 $V(T_i, x - x_k, y - y_k)$ に投票する. 入力画像中にモデル物体が存在すると, モデル画像中の辞書ウインドウが入力画像中にモデル画像と同じ位置関係で見つかる. このとき見つかる辞書ウインドウは位置の差に投票すると同じ位置に投票される. よって投票空間中の物体が存在する位置にピークが生じる. 適当な閾値でピークを検出することにより物体の認識が可能になる.

評価

固有窓法では物体を背景とセグメンテーションする必要はなく, 背景が動的であるような, セグメンテーションが難しい環境でも使用可能である. また認識対象物体にオクルージョンが生じていても, 閾値以上のウインドウが検出できればよいため, 認識が可能である. 問題としては計算量が膨大になり, 実時間処理に向かないことがあげられる.

2.2.3 Binary Feature 法

Binary Feature 法は, 上述の固有窓法をベクトル量子化を用いて二値画像に適用した手法であると考えることができる. 二値化エッジ画像を用いることにより, よ

り安定した処理と計算の高速化を実現している。



図 2.4: Binary Feature 法 . 左が認識対象の原画像 , 右が抽出された特徴ウインドウ .

原理

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 . 次いで固有空間法の辞書ウインドウと同様に特徴点 (binary features) 検出を行いウインドウを設ける (図 2.4) . 二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 [9] . 特徴点選択における相関性の評価は以下の式のように表現できる .

$$r_i(x, y) = \min_{\substack{-d \leq d_x \leq d \\ -d \leq d_y \leq d}} \{ D_H[\Omega\{e_i(x', y'); x, y, b\}, \Omega\{e_i(x', y'); x + d_x, y + d_y, b\}] \} \quad (2.9)$$

$$(d_x, d_y) \neq (0, 0).$$

ただし , $\Omega\{f(x, y); x_0, y_0, b\}$: (x_0, y_0) を中心とした $(2b + 1) \times (2b + 1)$ の領域を抽出する関数 ,

$D_H(\mathbf{r}, \mathbf{s})$: 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 binary features は Lloyd のアルゴリズム [5] でコードに圧縮される . Lloyd のアルゴリズムとは以下のようなものである . binary features に対し初期クラスタ中心を与えて , 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 binary features をグループ分けする . グループごとに , そのグループに属する binary

features の重心を求め、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重心は各要素ごとに 1 か 0 の多数決を行うことで求められる。このようにして各 binary feature がどのコードに属するかを求めて、その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、入力画像をモデル画像と同様に二値エッジ画像とし、全ての点にウィンドウを設け、コード化する (最も近いコードを求める)。コード化とは、モデル作成時に求めたコードに対して、入力画像中のウィンドウがどれに最も近いかを求めることである。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 binary feature 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。

評価

上述の手法同様に、小さいウィンドウを用いているため、背景の変動やオクルージョン、画像中における物体位置の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。また二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ハイライト等の光源環境の変化にロバストである。

第 3 章

車両認識システム

3.1 車両認識アルゴリズム

本研究の車両認識には，屋外環境でのロバスト性，実時間処理に向けた負荷の軽さを考慮して，上記の手法のうち二値画像を扱う Binary Feature 法を元にし，ベクトル量子化を用いた認識手法を使いシステムを構築する．本手法では二値画像を扱うため浮動小数点演算が必要ない．そこ処理の高速化を目指すため，処理には画像処理ボード，IMAP-VISION を用いる．

3.1.1 画像処理ボードによる高速化

ここでは車両認識システムを実装して，認識実験に使用する画像処理ボードについて述べる．ボードは，NEC 社製並列画像処理ボード IMAP-VISION である．IMAP-VISION の構成図を図 3.1 に示す．IMAP-VISION は大きく分けると画像処理部，コントロール部，ビデオインタフェース部にわけて考えることができる．画像処理部は 256 個の並列プロセッサと，各プロセッサに付随の内部メモリ (1KB/processor)，及び 16MB の外部メモリからなる．標準的な処理対象として 256×240 画素の画像を考えており，そのとき，対象画像の一行を並列プロセッサにより同時処理することが可能となる．コントロール部はシーケンサ，プログラムメモリ，バスコントローラなどから成り，ホスト PC との通信，プログラムなどのデータ格納，並列プロセッサやバス等の制御を行う．ビデオインタフェース部は，D/A コンバータと A/D コンバータにより，アナログビデオ入力的外部メモリへの書き込み，外部メモリ上の処理結果をアナログビデオ信号として出力する機能を有する．ただし，本研究の実装では入力を画像ファイルシーケンスとしてホスト PC から IMAP 上の外部メモリへ直接書き込むことにしたため，このアナログビデオ信号入出力機能は使用していない．

本画像処理ボードは，並列プロセッサを用い画像処理に特化した機能を持つことにより，高速な画像処理が可能である．

3.1.2 認識システムの実装

車両認識システムの処理の流れを図 3.2 に示す．認識対象車両の画像をあらかじめ準備しておき，モデル作成を行う．作成されたモデルを用いて，入力画像に対して認識操作を行うことによって，入力画像中の対象車両を認識する．モデル作成，認識

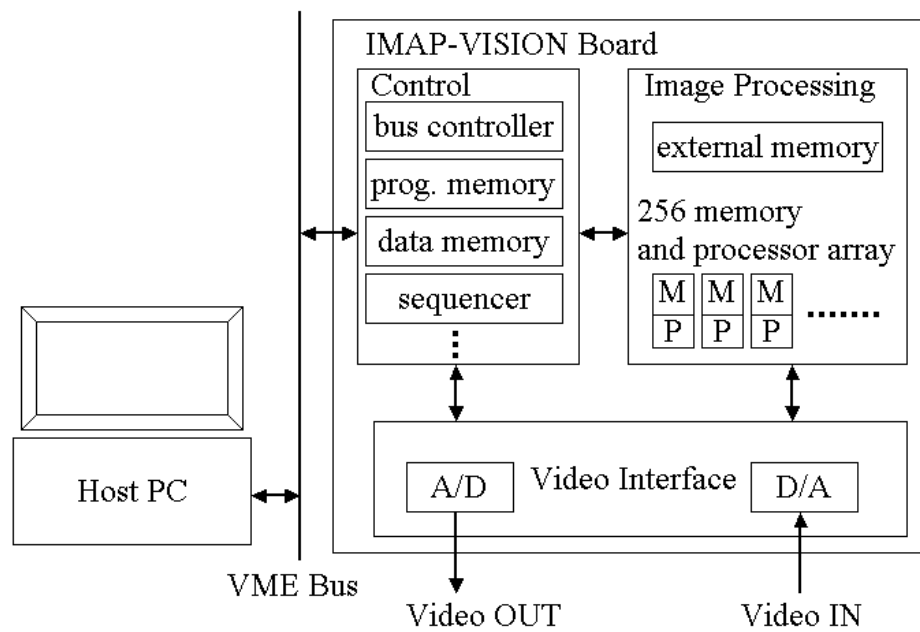


図 3.1: IMAP-VISION ボードの構成図 .

操作は固有空間法を元にしベクトル量子化を用いたアルゴリズムを用いている．それぞれの詳細を以下で述べる．

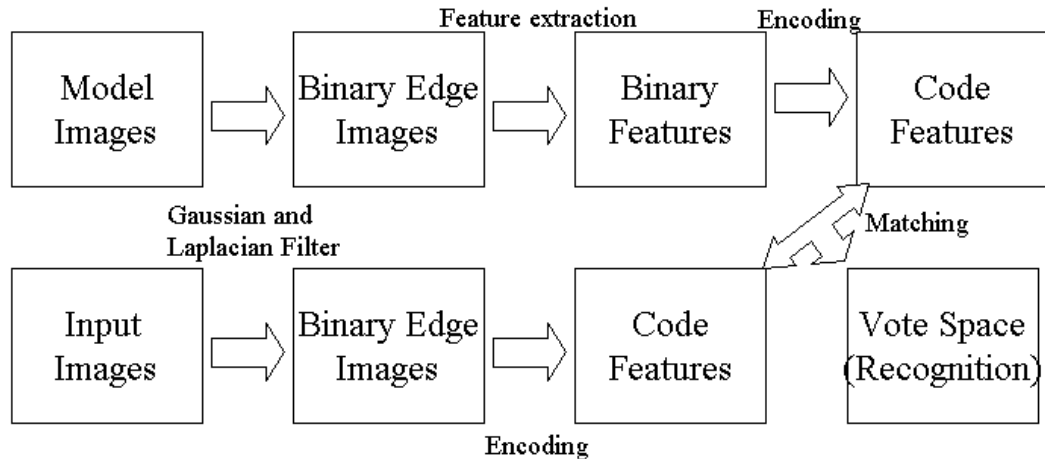


図 3.2: 車両認識システムの処理の流れ．

モデル作成

モデル作成では，入力画像と同じアングルで撮影された認識対象車両の画像を用いて，対象車両のモデルを作成する．モデル作成は Binary Feature 法を用いる．用意した認識対象車両の画像は Gaussian Filter 及びゼロクロス Laplacian Filter をかけて二値のエッジ画像にする．得られた二値化エッジ画像から特徴点を抽出する．特徴は Binary Feature 法と同様に，周囲との類似度を式 (2.9) により評価する．決められた数の特徴点を類似度の低い点から選択する．このとき特徴点画像のサイズ $N \times N$ を決めておき，特徴点は互いに重ならない点を選択する．得られた特徴を全モデル分をまとめて Lloyd のアルゴリズムでコード化する．各特徴は抽出元のモデル画像内における座標と対応するコードを保有し，モデルとして IMAP の外部メモリ上に保存される．また，作成されたコードの画像もモデルとして保存される．

モデル作成はホスト PC で行う．特徴点検出，コード化をホスト PC で行った後，IMAP の外部メモリに転送しモデルとして保持する．

車両認識

認識操作では，作成されたモデルを用いて入力画像中から認識対象車両を検出する．各モデルごとに投票処理を行って閾値を越えたものを検出することにより認識を行う．

まず，各モデル画像に対して入力画像サイズ分の投票空間を用意する．入力画像をモデル画像と同様に Gaussian Filter 及び Laplacian Filter で二値のエッジ画像にする．二値化した入力画像上の全ての点において，モデルの特徴と同じサイズのウィンドウを設ける．ウィンドウ内の画像とモデル作成時に作られた各コードとのハミング距離を計算し，最も距離の近いもの，つまり類似度の高いものをそのウィンドウのコードとする．各点のコードが得られたら，そのコードに対応する各特徴点を求め，その特徴点のモデル画像内の座標と，入力画像内のウィンドウの位置との差を求め，そのコードに対応する特徴点の所属するモデルの投票空間に投票する．入力画像中にモデル画像の特徴点とモデル内と同じ位置関係で見つかる点に投票される．モデルの特徴点とそのモデル内と同じ位置関係で入力画像中に多数見つかり，そこに認識対象車両が存在すると判断できるので，閾値でピークを検出して，対象車両を認識する．投票の様子を図 3.3 示した．本手法では投票の際ずれが生じることが予想されるので，安定化を計るため投票空間に 3×3 の平滑化フィルタをかける．フィルタの値は中心が 2 で残りは 1 である．

また，認識対象物体以外のエッジ成分は少ない方が認識に有利であるため，背景差分処理を行って余分なエッジを除去する．グレースケール画像に対する一般的な背景差分では，変動の多い屋外環境において不安定である．背景画像とは主に路面であり，認識の障害となるのは主に白線などである．安定な認識のためには白線から生じるエッジ成分を確実に除去したい．そこで本システムに導入する背景差分法として，背景，入力画像ともエッジ画像にしてから差分を行うことにする．つまり背景エッジ画像で 1 の部分が入力エッジ画像で 0 になる．2 値の入力エッジ画像 $I_i(x, y)$ ，2 値の背景エッジ画像を $I_b(x, y)$ ，差分後の画像を $I_s(x, y)$ とすると，

$$I_s(x, y) = \begin{cases} 1 & (I_i(x, y) = 1, I_b(x, y) = 0 \text{ のとき}) \\ 0 & (\text{それ以外のとき}) \end{cases} \quad (3.1)$$

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．背景差分の様子を図 3.4 に示す．エッジで背景差分を行うことにより，入力画像中の白線やその他背景のエッジ成分を完全に除去することができる．本手法では画像内で車両と背景のエッジ成分が重なると，車両のエッジ成分が消されてしまうという

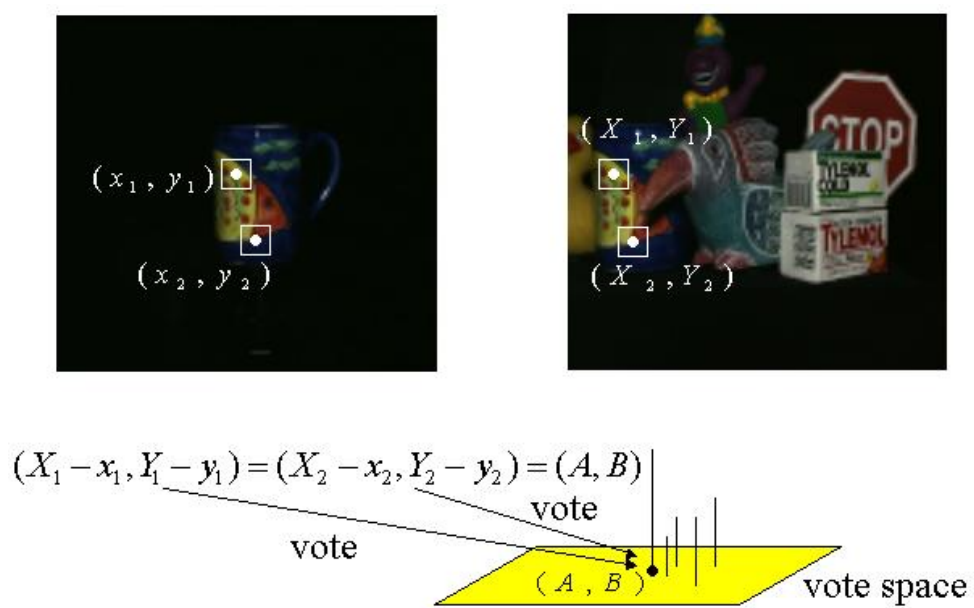


図 3.3: 投票操作．モデル画像と入力画像でマッチした特徴が複数個同じ位置関係で見つかりると同じ点に投票される．

問題が生じる．しかし認識手法は投票処理を行っておりオクルージョンに強いことから，重なって消されたエッジを他の物体によるオクルージョンとみなすと十分に認識可能であると考えられる．

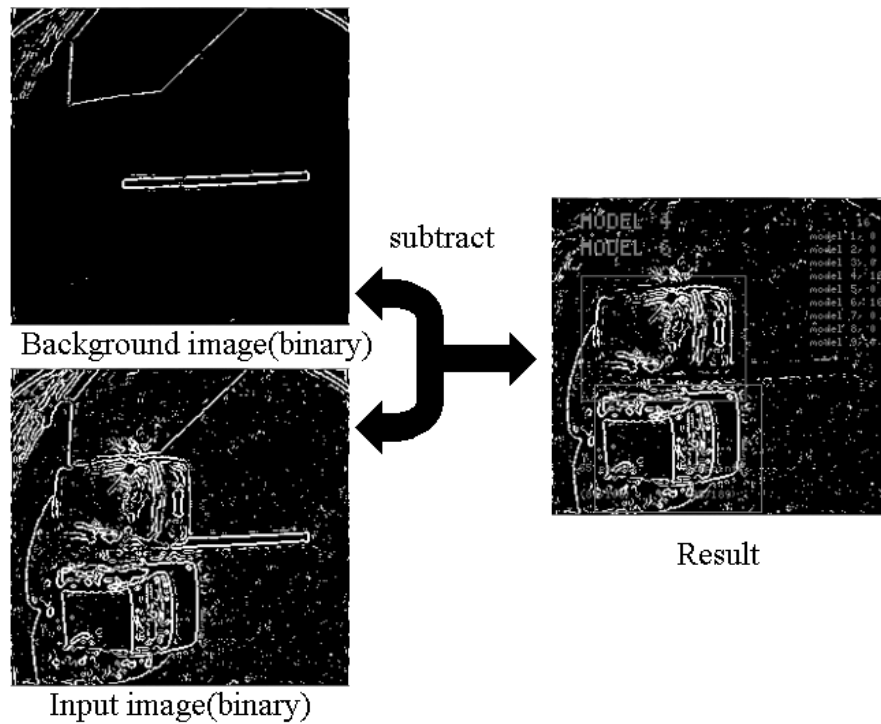


図 3.4: エッジ画像での背景差分処理．

認識処理は IMAP-VISION で行われる．画像の二値エッジ化，コード化から投票処理までを画像処理ボードの並列プロセッサで処理する．ただし入力画像は画像ファイルシーケンスとしてホスト PC に保存されており，ホスト PC から IMAP の外部メモリ直接書き込まれ，その画像に対して IMAP が処理を行う．画像の書き込みと IMAP の処理とは非同期で行われる．

第 4 章

車両認識実験

4.1 実験背景

第 3 章で述べた車両認識システムについて，その性能及び実用性を実際に屋外で撮影された画像を用いた実験を行うことにより検証する．

4.2 実験仕様

実験システムは第 3 章で紹介した通りである．ホスト PC の OS は FreeBSD ，メモリは 128MB ，CPU は Pentium3 の 533MHz である．ただしホスト PC のメモリや CPU の性能は基本的に認識処理の速度には影響しない．

実験に使用する画像は三菱重工株式会社のテストコースで，道路面に対して垂直に見下ろすかたちで設置された固定カメラを用いて撮影したものである．NTSC の CCD カメラであるが，一方向のみ広角に撮影できるシリンドリカルな魚眼レンズで撮影している．これは車両の進行方向は広い範囲を撮影し，車両が通過する間できるだけ長時間撮影したいが，車両の横方向に関しては広い視野はいと判断したためである．カメラは水平方向を広角に撮影できるようにレンズを装着して，水平方向が車両の進行方向になるように設置した．画像は 256 階調グレイスケール画像で画像サイズは 240×320 である．撮影した画像の例を図 4.1 に示す．ただし処理の際，IMAP-VISON ボードに合わせて画像の横方向の両端を切り 240×256 サイズの画像として扱う．

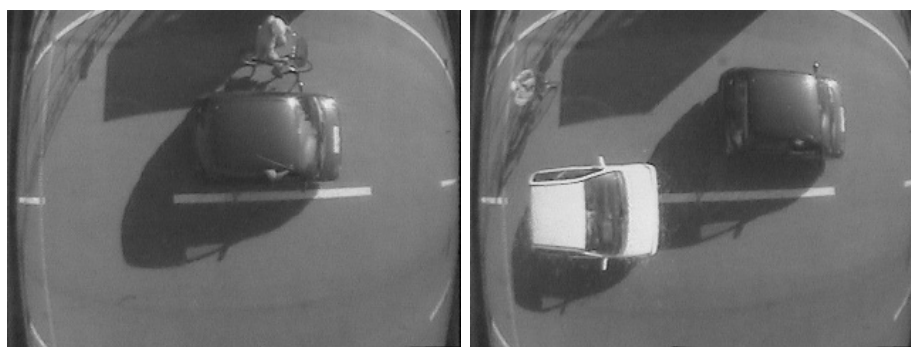


図 4.1: テストコースで撮影された画像の例．

次にシステムのパラメータを設定する．必要となるパラメータはモデル数，特徴点のウィンドウサイズ，特徴点数，コード数，投票の閾値である．

テストコースではワゴン型の黒い車両，白い車両及び自転車をいろいろな動き方で走らせて撮影した(図4.1)．本実験では認識対象を黒い車両，白い車両の両方に設定する．両車両は色の違いはあるが形状は似通っている．本研究の目的としては，個別に車を識別するのではなく，前提としてクラス分類への拡張を考えており，なるべく少ないモデルで同一車種と考えられる車両を識別したい．今回，テストコースで撮影された画像においては，先ほど述べたように撮影された車両は自転車を除くと同一車種と考えることができる．そこで共通のモデルを用意し認識することとする．白い車両と黒い車両，両方の画像を用意し，同数の特徴を抽出することによって共通のモデルを作成する．

しかし，画面水平方向に広角なレンズを使用しているため，画像内の位置によって車両の「みえ」が大きく変化してしまう．本手法ではイメージベースでモデルを作成しているため「みえ」が変化する状況では1つのモデルでは対処できない．この問題を解決する方法としては2つの手法が考えられる．一つは「みえ」ごとにモデルを持つことで，もう一つは画像に幾何学変換処理を施して透視投影図を再現し，位置による「みえ」の変化を補正する手法である．前者の手法では必要となるモデルの数が問題になり，後者では変換処理のコストと再現された画像の画質が問題となる．前者のモデルを複数持つ手法については，実際にモデルを複数持たせた予備実験を行い必要なモデル数を調べた．後者の変換手法については変換画像を作成し処理のコスト及び再現された画像の画質を検討した．予備実験の結果，複数モデルを持つ手法では必要なモデル数は9であることが解った．一方，幾何変換では計算コストはそれほどかからないが，レンズのパラメータが正確に得られなかったため正確な再現画像を求められないということ，カメラの中心から水平方向に離れた部分は原画像の解像度が低いため再現された画像が非常に荒くなるという問題があった．そこで本実験では9つのモデルを持つ手法で「みえ」の変化に対処することにした．

黒い車両，白い車両の実験に使用したモデル画像を図4.2，4.3に示す．それぞれ9つの画像が示されているが，その位置が画像を9分割したときの対応する位置を示しており，例えば図4.2の左上の画像は，画像を9分割した時の左上に対応する白い車両のモデルである．モデル画像は実験で使用する画像シーケンスから選んだ．対象の車両は切り出す必要があるため，マスク画像を用意して切り出した．マスク画像は手動で作成した．マスク画像の一つを図4.4に示す．

特徴点に設けるウインドウサイズは実験的に求め， 7×7 とした．また，一つのモ

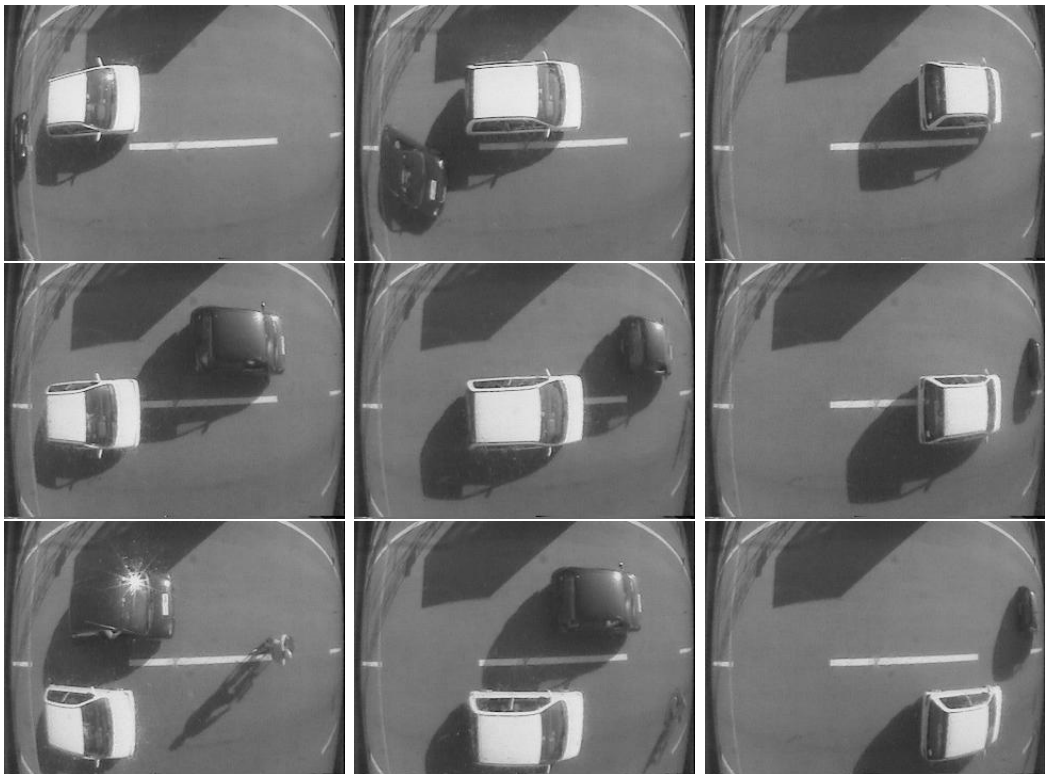


図 4.2: 画像を 9 分割した時の各位置における白い車両のモデル画像 .

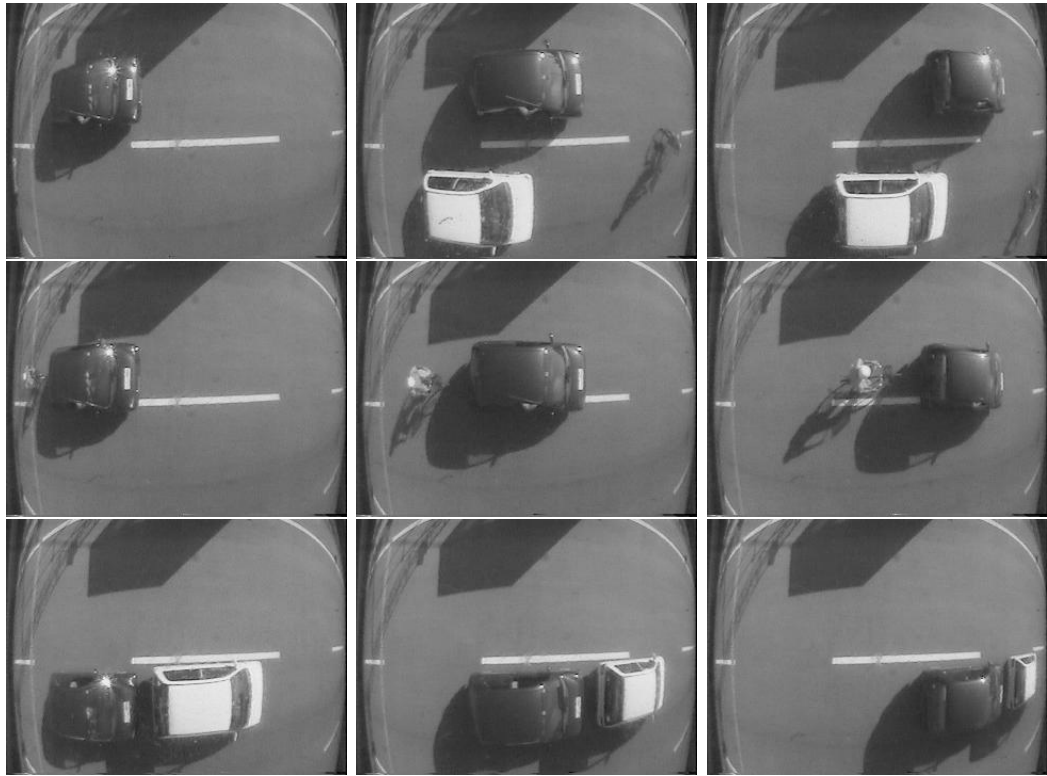


図 4.3: 画像を 9 分割した時の各位置における黒い車両のモデル画像 .



図 4.4: マスク画像の一例 . 画面左上用の白い車両のモデル .

デルから抽出する特徴数，抽出した特徴を圧縮するコード数，対象を検出する際の認識閾値も実験的に求め，それぞれ 60, 20, 65 とした．特徴数，コード数，認識閾値，画像サイズ，モデル数を決めると，コードの出現確率が一定であると仮定した時の誤検出率を求めることができる．コード数を n_c としたとき，全てのコードで出現確率が等しいと仮定すると，ある点での各コードの出現確率は $\frac{1}{n_c}$ である．特徴数を n_f ，認識閾値を T とすると，ある点において閾値 T 以上の投票が起きる確率 p_f は，以下の式で求めることができる．

$$p_f = \sum_{i=V}^{n_f} \binom{n_f}{i} \left(\frac{1}{n_c}\right)^i \left(1 - \frac{1}{n_c}\right)^{(n_f-i)} \quad (4.1)$$

モデル数を n_m ，画像サイズを (n_x, n_y) とすると，誤検出の確率 $p_{fa}(V)$ は以下の式で求めることができる．

$$p_{fa}(V) = 1 - (1 - p_f)^{n_m n_x n_y} \quad (4.2)$$

式 (4.1) を代入して

$$p_{fa}(V) = 1 - \left(1 - \left(\sum_{i=V}^{n_f} \binom{n_f}{i} \left(\frac{1}{n_c}\right)^i \left(1 - \frac{1}{n_c}\right)^{(n_f-i)}\right)\right)^{n_m n_x n_y} \quad (4.3)$$

となる．この式から特徴数，コード数とも増やすと誤検出の確率は下がることが解る．ただし抽出できる特徴数は限界があり，コード数は増えると投票数が下がり検出率が下がるため，どちらの数も実験的に適切な値を求めた．モデル画像 (図 4.2, 4.3) 中の対象車両を抜き出し，特徴を抽出した結果を図 4.5 に示す．また，抽出された特徴から求めたコードを図 4.6 に示す．

以上のように定めたパラメータで実験を行った．

4.3 実験結果

テストコースで撮影された 262 枚の画像用いて，車両認識システムの性能評価実験を行った．使用した 262 枚の画像中，白い車両が写っている画像が 143 枚，黒い車両が写っているものが 234 枚ある．入力画像と認識結果画像の例を図 4.7, 4.8 に示

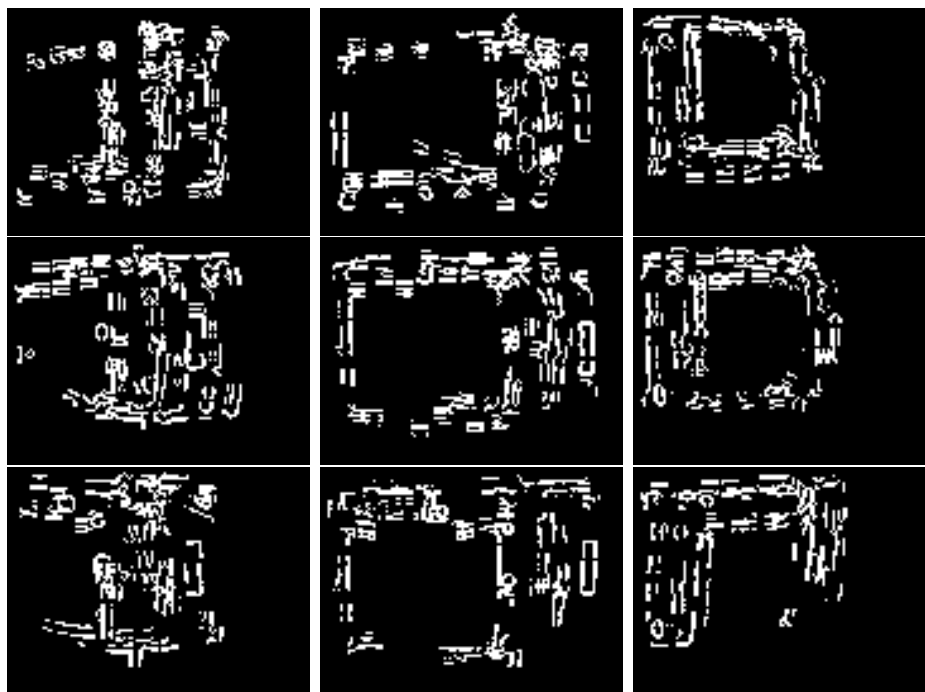


図 4.5: 画像を 9 分割した時の各位置における車両のモデル画像から抽出した特徴 .



図 4.6: 抽出された特徴から生成したコード .

す．図 4.7 は成功結果例，図 4.8 は失敗結果の画像例で，左が入力画像，右が識別結果の出力である．結果画像は処理後の二値エッジ画像で，車両を認識した場合，その位置に長方形で検出結果を表示している．図 4.8 の失敗例では車両が画像内にあるのに検出されない未検出の例を上にも，車両が存在しない位置に車両がいると判定する誤認識の例を下に示している．実験画像は図 4.7，4.8 に示したように，二台の車両間を自転車が行く，二台が近接して行くといった画像認識では困難な状況の画像も含まれている．また，実験画像は晴天の日中に撮影されたもので，明瞭な影が道路面や車両上にかかっており，これも認識の妨げとなっている．実験結果を表 4.1 に示す．表 4.1 は，画像中の各車両を位置も含めて正しく認識した数と，その車両が存在する入力画像全ての数を成功数 (Success) として，両者比率を成功率 (Ratio) として示し，車両の存在しないところに車両が存在すると判定した誤認識の数と，実験回数 (つまり画像枚数) を誤検出 (False alarms) の欄に示している．黒い車両，白い車両とも 95% をこえる高い認識率で認識することができた．

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

表 4.1: 認識実験の結果．

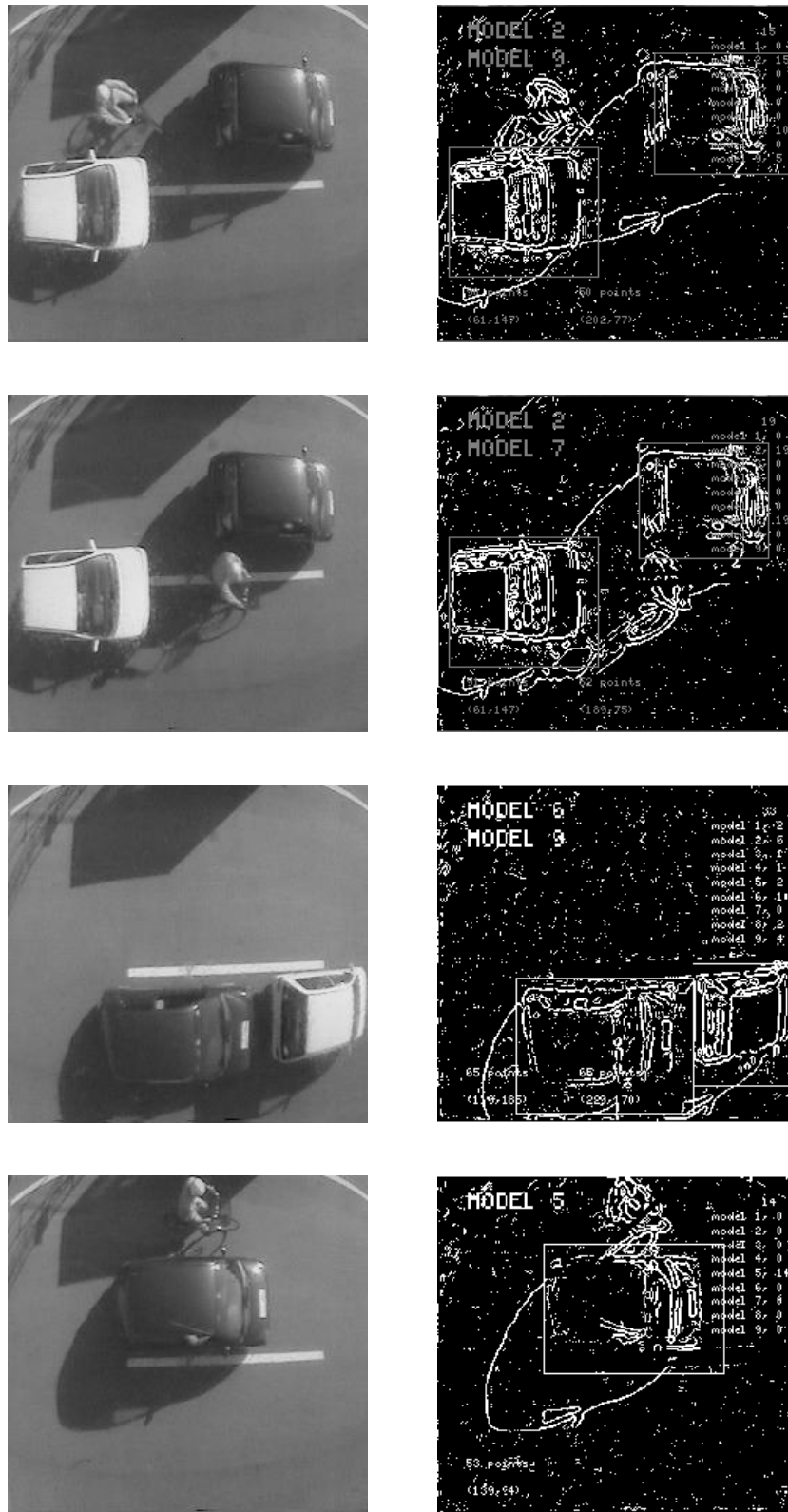


図 4.7: 車両認識実験における成功結果の画像例．左が入力原画像，右が認識結果．

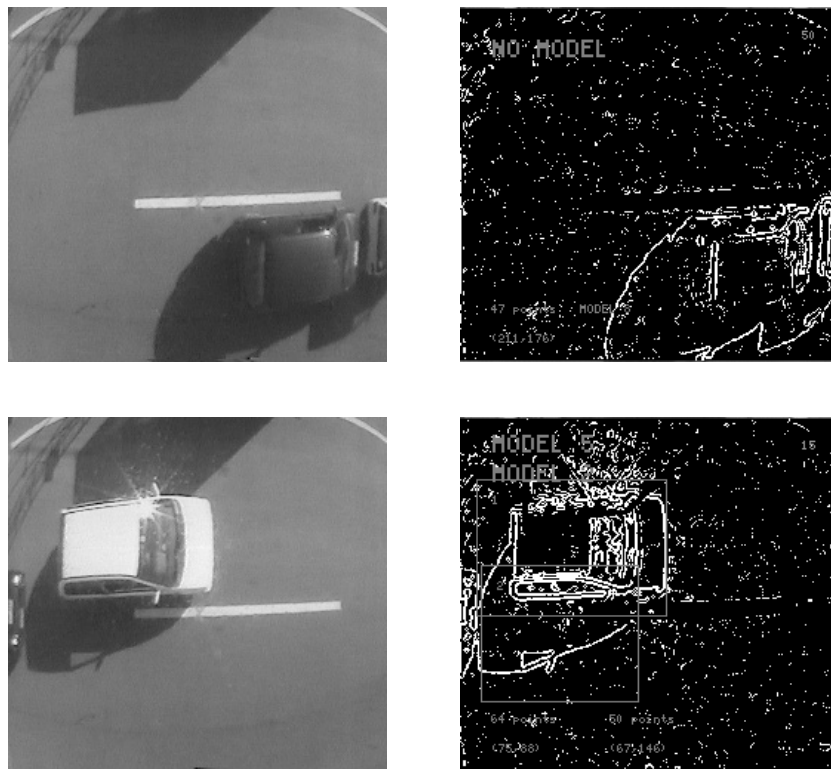


図 4.8: 車両認識実験における失敗結果の画像例。左が入力原画像，右が認識結果で，上は未検出，下は誤認識の例。

第 5 章

クラス分類手法

第 4 章で，提案した車両認識システムが，可視光カメラを入力デバイスとして使用した場合に問題となるオクルージョンや影の問題に影響されず，高い認識率で車両を認識することが可能であることを示した．本章では我々の車両認識システムのさらなる知的化を目指すため，車両認識システムのクラス分類（車種分類）への応用を考える．クラス分類は自動料金収受システム，道路管理システムといった ITS における多くの分野で必要とされている技術である．第 1 章で紹介した研究 [11][13][23] の他にも，画像センサを使った研究例 [22]，レーザセンサを使った研究例 [21] などがあり，必要性が高く注目されている研究課題である．赤外画像を利用して車両のクラス認識を行った先行研究 [18] を踏まえ，車両認識システムのクラス認識への応用を検討する．

5.1 課題

クラス分類のためには，現在のシステムをクラス分類に適した形にする必要がある．我々のシステムにクラス分類を適用する場合，単純な手法であるが車両の各モデルに対しクラスのラベルを付けて認識する手法が最も適していると思われる．ただし，我々の用いている認識手法はモデルベースの手法であるため，認識にはあらかじめ認識対象車両の画像を準備しておき，モデルとして保持しておく必要がある．クラス分類のためには分類対象の車両の画像を全て用意する必要がある．各クラスのモデルは，そのクラスに分類される車両の「みえ」の変動を吸収するくらいの数が必要となる．全ての車両に対してクラス分類を行うことを考えると，大量のモデル画像が必要となりモデル画像を収集する時点で相当の労力と，時間を浪費してしまう．この問題を解決する方法として我々はコンピュータグラフィックス（CG）で車両の画像を生成して，モデルとして使用することを考える．

5.2 コンピュータグラフィックスによるモデル生成

近年の CG 技術の発展により，CG のクオリティは大幅に向上しており，実写と変わらないほどのリアリティを持つ画像を生成することも可能になっている．また，CG で画像を生成する時には，光源環境，物体の姿勢，視点などを自由に変化させて画像を生成することができる．画像中の物体は，幾何形状，表面の色や反射パラメータ等を設定することにより任意のものを生成することができる．我々が必要として

いる車両のように人工的な物体は，比較的これらのパラメータを与えやすくリアルな画像を生成しやすい．また幾何データなどは市販されているものもあり手に入りやすい．以上のような利点を考えると，大量のモデル画像を収集するにはCGで生成するのが適していると考えられる．

CGの画像はリアリティが高くても，やはり実写画像との間には差が存在する．特に屋外環境における画像を生成する場合，その複雑な光源環境を再現することは難しい．しかし我々の車両認識システムは，画像を二値のエッジ画像として使用するため，細かな違いは十分に埋めることができると考えられる．そこで，実際にCGでモデル画像を生成してシステムに実装し，実画像内の車両を認識できるかどうか，実験を通して検証を行う．

5.3 クラス認識システム

我々が実験で使用するクラス認識システムは，第4章の実験で実装した車両認識システムと同様のシステムで，各モデルに対してどのクラスに属しているかをラベル付けしている．車両認識システム同様，IMAP-VISION ボードを画像処理に使用している．クラス認識システムの処理の流れを図5.1に示す．

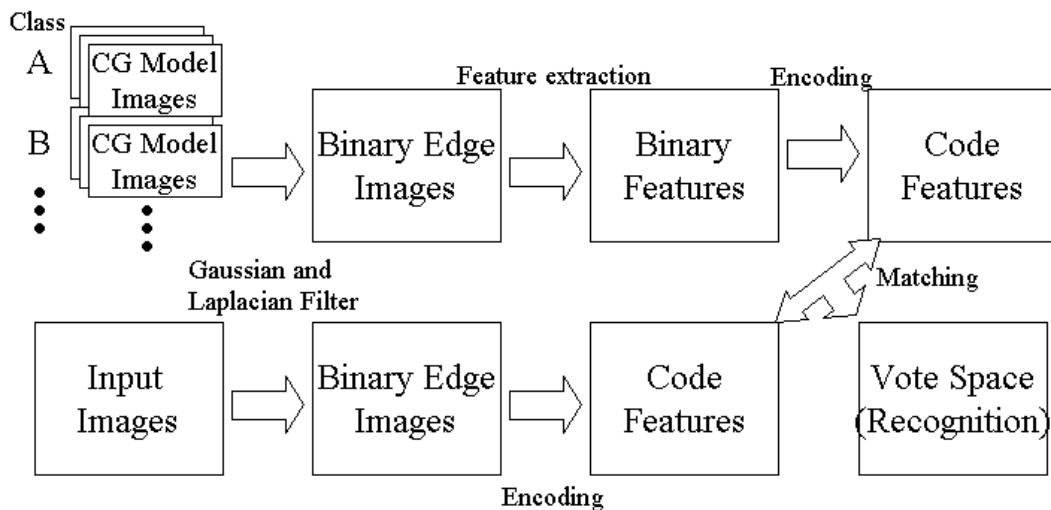


図 5.1: 車両クラス分類システムの処理の流れ．

認識を行う前にシステムにモデルを用意する．まず必要となるモデル画像を用意

する．必要となるクラスの必要数のモデル画像を上述したように CG で製作する．手間をかけずにモデルを収集するという主旨から，CG のモデル画像は CG モデリング・レンダリングソフトで，市販の形状データを用いて製作する．モデル画像に Gaussian Filter 及びゼロクロス Laplacian Filter をかけて二値のエッジ画像にする．得られた二値エッジ画像から，特徴点を抽出する．認識システムの時と同様に第 2 章の式 (2.9) で周囲との類似度を評価して決められた数の特徴点を周囲との類似性の低い点から選ぶ．特徴点からサイズ $N \times N$ の画像を特徴画像として使うので，特徴点は重ならないように選ぶ．得られた特徴は Lloyd のアルゴリズムでコード化し，IMAP の外部メモリにモデルとして保持される．各特徴は，抽出元のモデル画像，モデル画像内の位置，対応するコードを保有して IMAP の外部メモリにモデルとして保存される．

得られたモデルを利用して投票処理により対象を検出しクラス分類を行う．各モデル画像ごとに入力画像サイズの投票空間を用意する．入力画像をモデル画像のときと同様に Gaussian Filter 及び Laplacian Filter で二値のエッジ画像にする．二値化した入力画像上の全ての点において，モデルの特徴と同じサイズのウインドウを設ける．ウインドウ内の画像とモデル作成時に作られた各コードとのハミング距離を計算し，最も距離の近いもの，つまり類似度の高いものをそのウインドウのコードとする．各点のコードが得られたら，そのコードに対応する各特徴点を求め，その特徴点のモデル画像内の座標と，入力画像内のウインドウの位置との差を求め，そのコードに対応する特徴点の所属するモデルの投票空間に投票する．入力画像中にモデル画像の特徴点モデル内と同じ位置関係で見つかる点に投票される．モデルの特徴点とそのモデル内と同じ位置関係で入力画像中に多数見つかる点と，そこに認識対象車両が存在すると判断できるので，閾値でピークを検出して，対象車両を認識する．本手法では投票の際ずれが生じることが予想されるので，安定化を計るため投票空間に平滑化フィルタをかける．また，余分なエッジ成分を除去するため，第 2 章で説明した，式 (3.1) のエッジ画像間での背景差分処理を行う．

以上の様に車両認識システムと同様の処理で車両を認識し，検出されたモデルに対応するクラスで入力画像中の車両のクラスを判別する．

第 6 章

クラス分類実験

第 5 章で述べた車両クラス分類システムについて，その性能を検証するために屋外で撮影した画像を用いて実験を行う．

6.1 実験仕様

実験システムは第 5 章にあるように，第 3 章で紹介した車両認識システムと同じシステムにおいて，各モデル画像にたいしてクラスのラベルを付けたものである．また大量のモデル画像収集のため CG でモデル画像を生成してシステムに使用している．第 4 章の実験と同じく，ホスト PC の OS は FreeBSD，メモリは 128MB，CPU は Pentium3 の 533MHz である．

実験に使用する画像は東京大学の本郷キャンパス，農学部キャンパス間の歩道橋上からカメラを真下に向けて公道を走行している車両を対象に撮影したものである．第 4 章で使用した画像では同一の車両が繰り返し撮影されており，クラス分類には適していないため新たに実験用画像を撮影した．撮影時の天候は曇りで，第 4 章のときのような広角レンズは使用していない．画像はカラーで撮影したが，実験には 256 階調のグレースケールに変換して使用した．画像サイズは 240×256 である．撮影された画像の例を図 6.1 に示す．

まず撮影された画像データに写っている車両を分類し，四つのクラスとした．四つのクラスはセダン，ワゴン，ワンボックス及びハッチバック（軽自動車）である．車両のクラス分類法は様々なものがある．我々のシステムは画像センサを使っているため，本実験では形状による車両の分類法を用いた．各クラスに属す車両の画像例を図 6.1 に示している．第 4 章のときと同様に認識システムのパラメータを設定しなくてはならない．必要なパラメータはモデル数，特徴点のウインドウサイズ，特徴点数，コード数，及び投票空間にかけるウインドウのサイズである．各パラメータは第 4 章のときと同様に実験的に定めた．モデル画像は各クラスに対して 6 つとした．モデル画像は CG で作成している．作成には三次元モデリング・レンダリングソフトである Shade を用い，車両の幾何形状モデルは市販の Shade 用車両モデルデータ集を使用した．車両の形状や反射パラメータを人手で作成する必要はなく，データ集の車両形状やパラメータをそのまま用いるか多少の調整を行うことによって必要な車両の形状を得ることができ，任意の視点，光源環境下の画像を得ることができる．実験で使用したモデル画像を図 6.2，6.3，6.4，6.5 に示す．実写画像をモデルに使っていた第 4 章の車両認識実験のときは対象車両を切り出すためのマスク画像

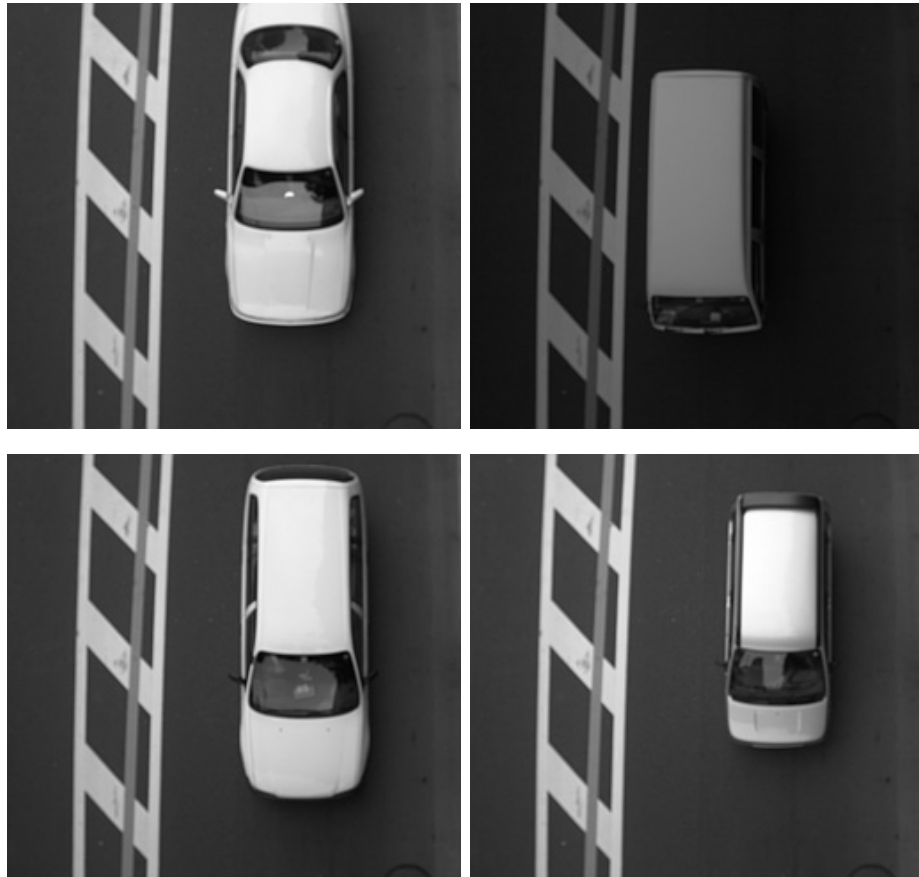


図 6.1: 歩道橋上から撮影された車両の画像例．それぞれ左上はセダン，左下はワゴン，右上はワンボックス，右下はハッチバックの例．

が必要であったが，CG モデル画像では背景を入れなければ容易に切り出しが行えるため，特にマスク画像は必要としない．

特徴点に設けるウインドウサイズは 11×11 とし，抽出する特徴数はモデルの周囲長に比例させて抽出し 42 個から 26 個とした．コード数は 20 とした．これらは上述のように実験的に求めた．投票処理の際に生じるずれを吸収するために投票空間にかけるフィルタは 15×15 の平滑化フィルタで，中心のみ 2 であとは 1 の要素を持ったフィルタである．ただし投票のときのずれはモデルのサイズによって変化すると考えられ，同じフィルタをかけるとサイズが小さいモデルが有利になるため，上述の様に特徴数をモデルのサイズで変化させこれに対処した．ウインドウサイズと投票空間にかけるフィルタのサイズが第 4 章の実験より大きくなっているのは，識別対象となる車両の画像内のサイズが大きくなったためである．また，今回の実験ではクラス分類に主眼を置いていること，実験画像には常に対象が一つ写っていることから，第 4 章の実験では必要であった認識の閾値は設定していない．投票の数が問題では無く，相対的に最も投票の大きい点を持つモデルが選択される．また実験の際，画像中に対象は一つとして実験を行った．図 6.2，6.3，6.4，6.5 のモデル画像から抽出した特徴を図 6.6，6.7，6.8，6.9 に，得られた特徴を圧縮して求めたコードを図 6.10 に示す．

以上の仕様で実験を行った．

6.2 結果

歩道橋上で撮影された 87 枚の画像を用いてクラス識別実験を行った．本実験ではクラス分類を主眼に置いているので画像シーケンスで同一の車両を何度も認識するのではなく，それぞれの画像には異なった車両が写っている．87 枚の画像中セダンが写っているものが 28 枚，ワゴンが写っているものが 20 枚，ワンボックスが写っているものが 22 枚，ハッチバックが写っているものが 17 枚ある．実験結果の画像例を図 6.11，6.12 に示す．図 6.11 は成功結果，図 6.12 は失敗結果の例で，左が入力画像，右が識別結果の出力である．結果画像は処理後の二値エッジ画像で検出した車両位置を長方形で示し，判定したクラスを左上に表示している．実験結果のコンフュージョンマトリクスを表 6.1 に示す．行はシステムにより判断された結果で，列は画像内の車両の実際のクラスを表しており，例えばセダンの行のワゴンの列はワゴンが写っている画像中でシステムがセダンであると判断した画像の枚数である．つま

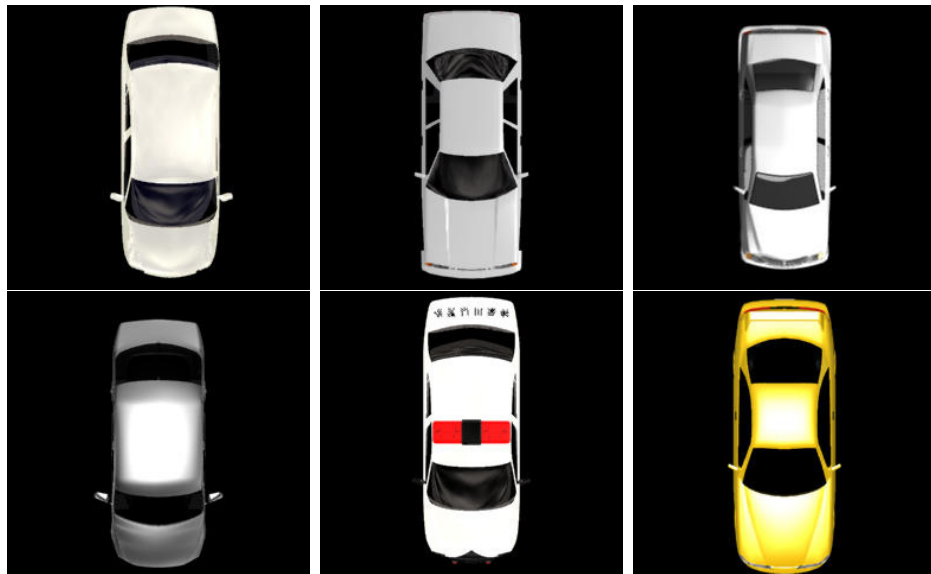


図 6.2: 実験に使用したセダンクラスのモデル画像 .

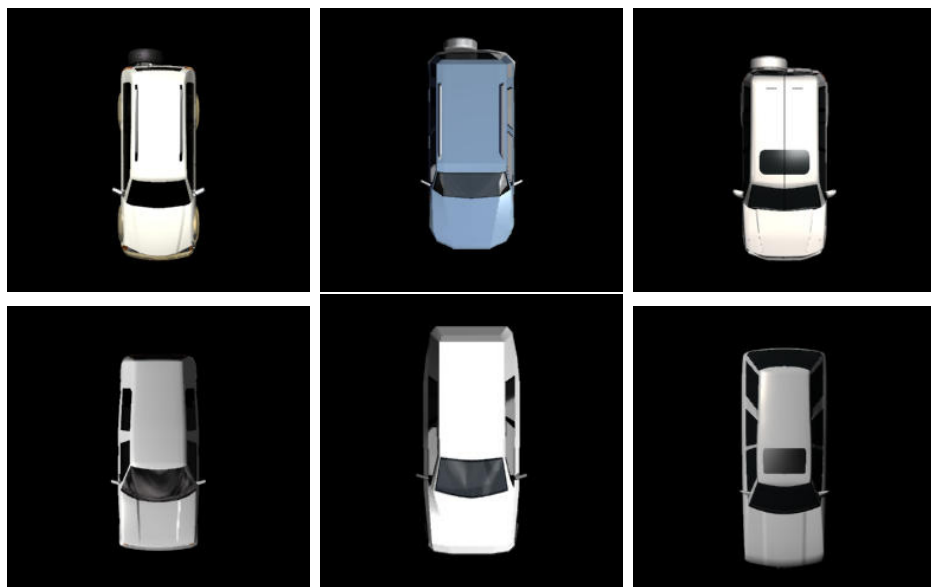


図 6.3: 実験に使用したワゴンクラスのモデル画像 .

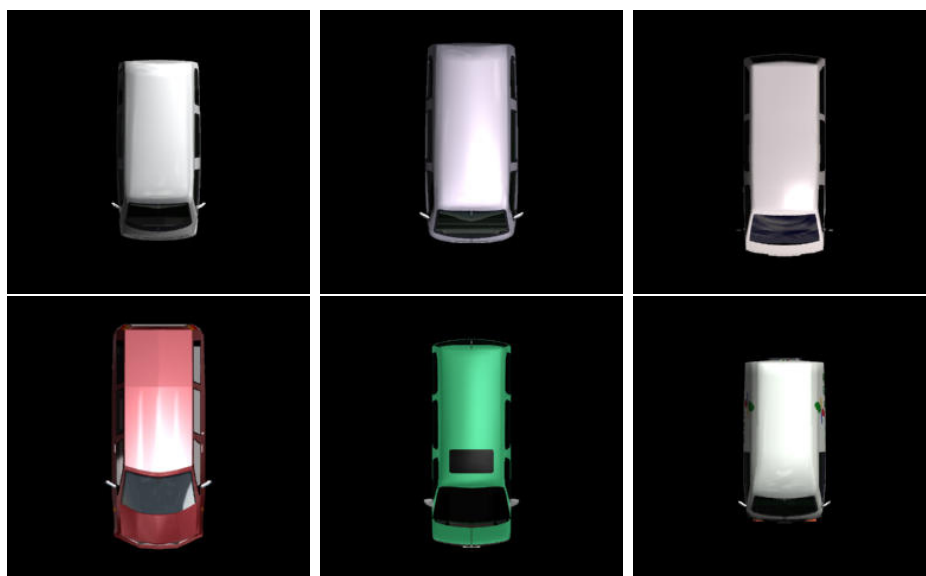


図 6.4: 実験に使用したワンボックスクラスのモデル画像 .



図 6.5: 実験に使用したハッチバッククラスのモデル画像 .

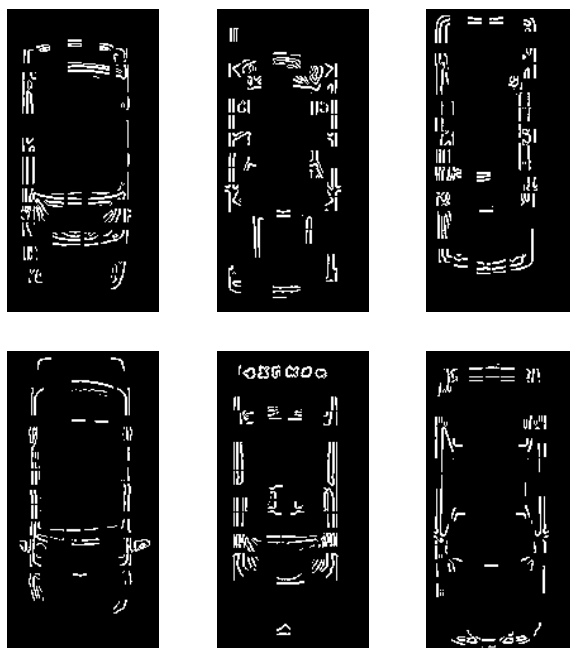


図 6.6: セダンクラスのモデル画像から抽出した特徴 .

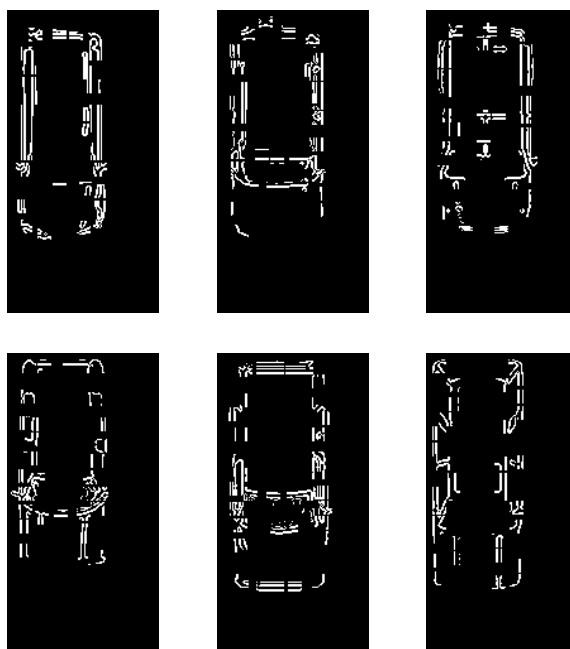


図 6.7: ワゴンクラスのモデル画像から抽出した特徴 .

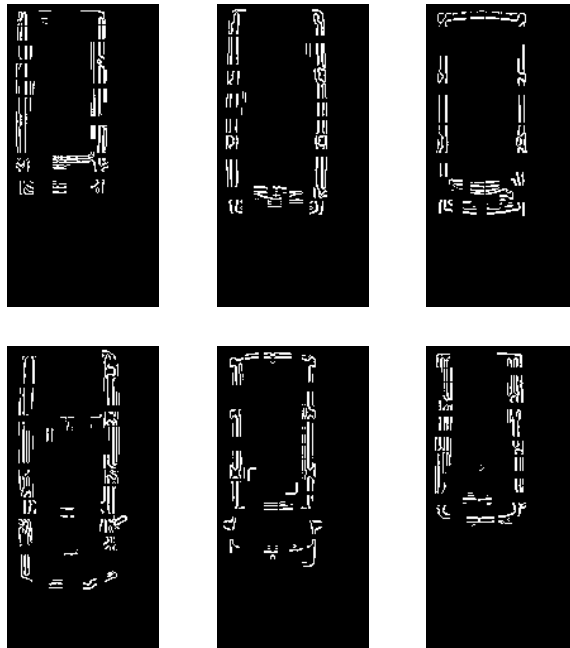


図 6.8: ワンボックスクラスのモデル画像から抽出した特徴 .

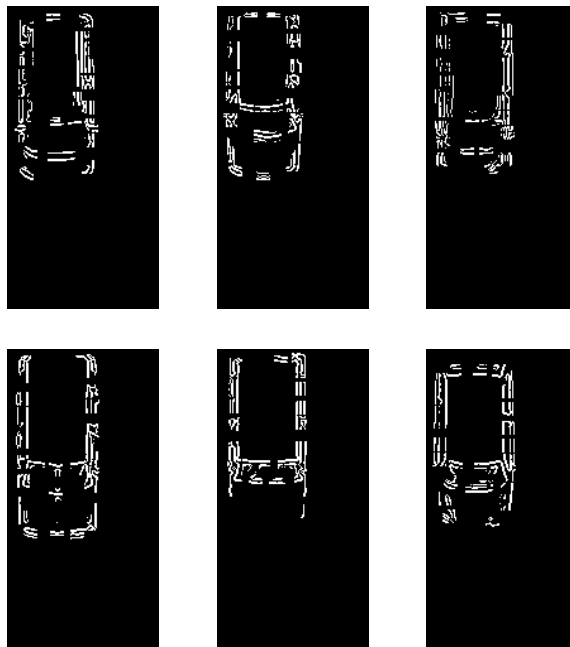


図 6.9: ハッチバッククラスのモデル画像から抽出した特徴 .

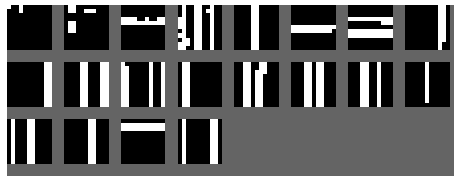


図 6.10: 抽出された特徴から生成したコード .

対角線上の数字が正解数であり，正解数の全画像数に対する割合を正解率としている．結果をみると全体で半数が成功している．ワゴン，ハッチバックは高い確率で成功しており，逆にワンボックスとセダンはあまり正解にならなかった．

Class as	True Class				total (as)
	sedan	wagon	one box	hatchback	
sedan	6	0	3	0	9
wagon	20	14	6	2	42
one box	1	4	10	1	16
hatchback	1	2	3	14	20
total (true)	28	20	22	17	87
accuracy	51%				

表 6.1: クラス識別実験の結果 .

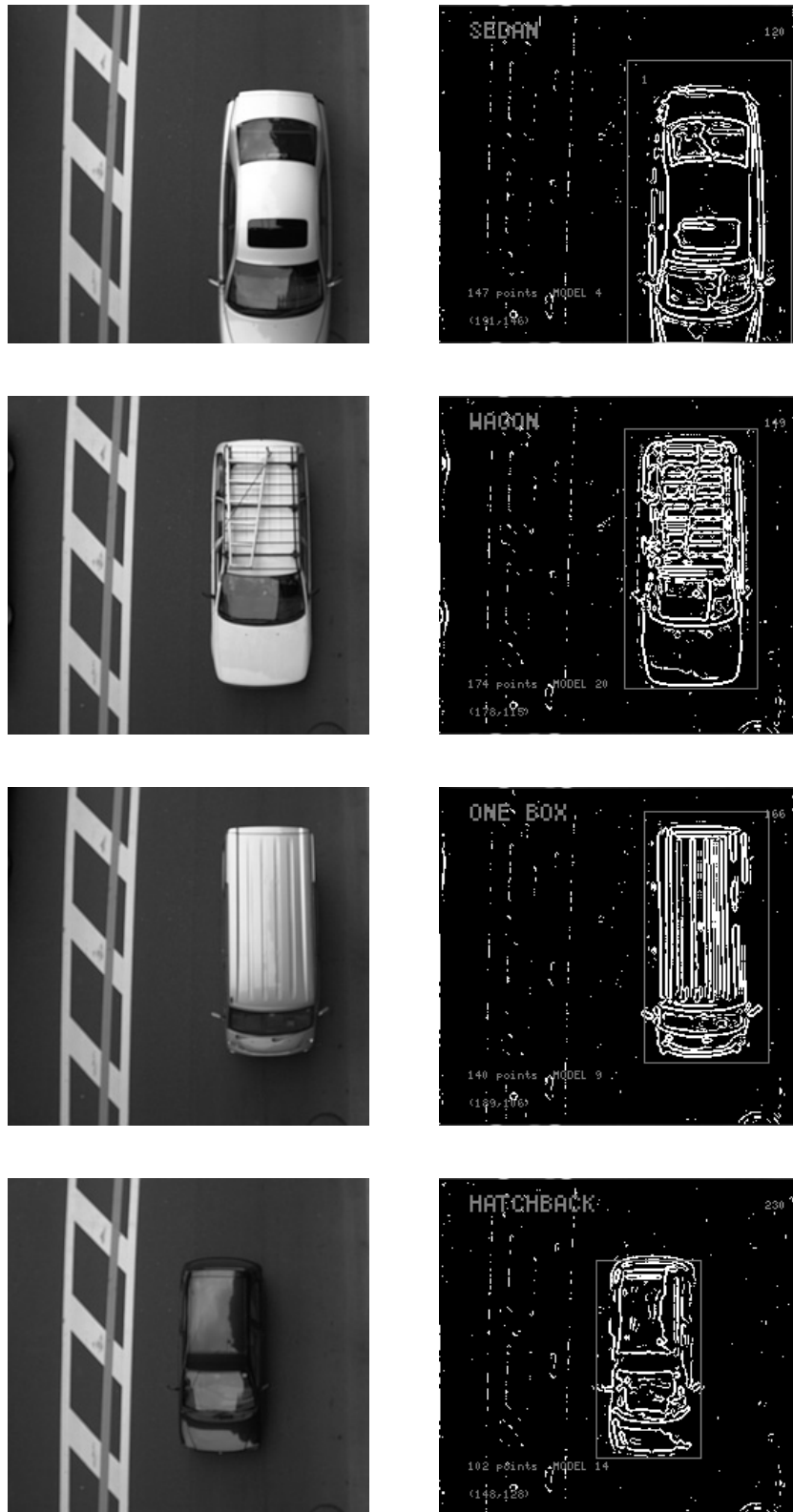


図 6.11: クラス分類実験の成功結果画像例 . 左が入力原画像 , 右が認識結果 .

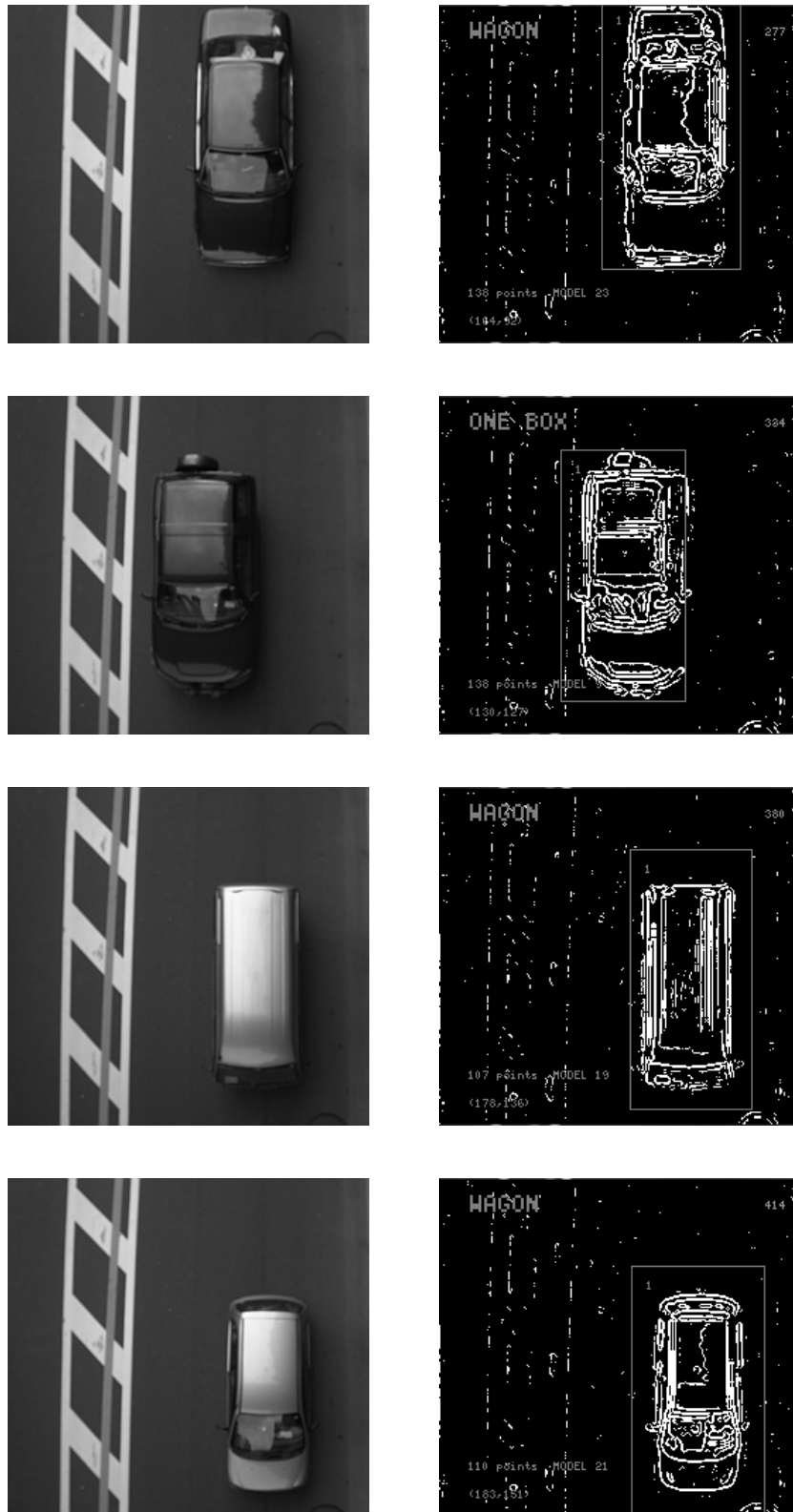


図 6.12: クラス分類実験の失敗結果画像例 . 左が入力原画像 , 右が認識結果 .

第 7 章

結論

7.1 実験結果の考察

7.1.1 車両認識実験

第4章の車両認識実験では、コンピュータビジョンの物体認識手法を元にして提案し、画像処理ボードに実装した車両認識システムを用いて、実際に屋外で撮影された画像に対する性能を検証する実験を行い、高い認識性能を示すことができた。黒い車両と白い車両を同一のモデルで認識することができ、クラス分類にむけて少ないモデルで複数種類の車両を高い精度で認識できることが確認できた。

実験で使用した画像は、天気が晴れて影がはっきりと路面及び車両に映っている状態で撮影されたものである。我々の車両認識システムではロバストな検出アルゴリズムを用いているため、影を車両として認識することはほとんど起こらなかった。また、黒い車両は影ができると路面上の影と車体の色の区別が無くなってエッジが消えてしまう傾向があり、検出のさいに悪影響があるが、黒い車両も白い車両とほとんど変わらない精度で検出することができた。また、障害物として自転車が車両の間を走行したり、車両と並列に走行したりした画像シーケンスもあるが、自転車及び自転車が映す影による誤検出もほとんど起こらなかった。

今回の実験画像は真上から撮影されているためオクルージョンはほとんど起こっていないが、多少のオクルージョンでは認識性能に影響が無いことは赤外画像を使った先行研究 [19] で示されており、また本手法ではエッジ画像で背景差分を行っており、車両が白線など背景のエッジ成分と重なる際にエッジが消えるが、これをオクルージョンと考えると、やはり多少のオクルージョンは認識に影響ないといえる。以上のように、提案するシステムは可視光カメラで問題となる光源環境の変化や影に対して影響を受けずに、対象車両を高い精度で認識できた。

また、実験に使用した画像はシリンドリカルなレンズで撮影されているため同じ車両でも位置により「みえ」が変化していたが、画面を9分割にし各位置のモデルをもつことで画像全体の車両を認識することができ、広角レンズで広い範囲を対象にしても車両を認識できることが示された。

提案した車両認識システムが、広角レンズによる「みえ」の変化や類似した形状の車両における微妙な差違に影響されず高い精度で車両認識することが、実験を通して確認され、このことは、細かい個々の種類、つまり特定の車種、形状まで認識しようとした場合には問題になるが、クラス分類に向けては有効な手法であるといえる。

7.1.2 クラス分類実験

第6章のクラス分類実験では、まずCG画像で作成したモデルで対象車両を認識できることが確認できた。我々のシステムのように画像からモデルを作成する手法では認識対象のモデル画像が必要になる。室内の実験環境で必要な条件の画像を必要なだけ撮影できる場合と異なり、車両認識などの場合は設置されたカメラでまずモデル画像を必要なだけ集める必要がある。これは多くの場合非常に負荷の大きな仕事になるため、我々の提案したCGをモデル画像として使う手法は、その労力を軽減する非常に有効な手法となる。

クラス分類実験は正解率が約50%であった。実用性を考えると低い正解率であるが、他手法[11][13][18][21][22][23]と比較してみた際必ずしも劣った結果ではない。サイズで大型車と小型車を分類する研究[18][21]や2クラス間の分類を行う研究[13]では高い正解率の結果も報告されているが、車高、幅、長さといった情報で明確に区別しがたい場合のクラス分類研究では、高い正解率のものは報告されていない。

実験ではワゴンとハッチバックは正解率が高く、ワンボックスとセダンは正解率が低かった。後者の二つはモデル画像から抽出される特徴が輪郭に集中する傾向があるが、直接の原因であるかは解らない。ハッチバックが正解になる確率が高いのはサイズが小さく、投票のずれが小さいためであると考えられる。逆にセダンはサイズが大きく投票の際に生じるずれが大きいためと考えられる。またトランク部を除外した前の部分はワゴンと形状が非常に似通っているため、ワゴンと判定されてしまうことが多い原因になっていると思われる。実験した画像中には画面の端でトランク部が切れて見えなくなっており、ワゴンと判定されても仕方のない画像も数枚あった。

7.1.3 まとめ

最後に本研究の成果をまとめる。本研究では、画像認識手法とベクトル量子化を用いた車両認識システムを提案、実装した。この車両認識システムを用いて実際に屋外で撮影された画像に対して性能評価実験を行い、以下のような有効性を確かめることができた。

- 環境変動に弱い可視光カメラを入力デバイスとしながら、高い認識率で車両を認識した。
- 実験画像は可視光カメラで問題となる明瞭な影が映っているが影を車両として認識するといった悪影響はほとんど受けなかった。

- 障害物として自転車及び自転車の映す影が実験画像に映っているが，自転車を車両と認識することはほとんど無かった．
- 広角レンズで「みえ」が変化しているが必要なモデル画像を増やすことで対応できた．
- 同じ車種（クラス）の異なる車両を一つのモデルで認識できた．
- 黒い車両は影によって一部の境界のエッジが消えるが，白い車両と変らない精度で認識できた．

同一のモデルで複数の車両を認識できることが確認されたため，車両認識システムを拡張し，クラス分類システムを提案，実装し，同様に実験を通して性能評価を行った．クラス分類システムでは以下のような有効性が確認できた．

- CG で作成したモデル画像を使って，実画像中の車両を認識できた．このことは，実画像で多数のモデル画像を容易に収集することが難しい車両認識では非常に効率的な手法であるといえる．
- 50%以上の正解率でクラス識別を行うことができた．実用段階には至らないが「みえ」からクラス分類を行う有効性が確認された．

7.2 今後の課題

車両認識実験では高い認識率で可視光画像内の対象車両を検出することができた．拡張として考えたクラス分類システムではCGでモデル画像を作成して車両を認識することができたが，車両のクラス分類の精度は50%程度であった．実用的な段階で求められる精度にはまだ遠く，今後より高い精度でのクラス識別処理が求められる．サイズによる影響を受けるので，車両の領域を切り出し正規化して扱いたい画像センサがロバストでない屋外環境では正確に車両領域を切り出すことは困難であり，今後検討が必要である．

他には，クラス分類で用いられる手法を積極的に取り入れていくことで，より高い精度で車両のクラス識別ができると考えられる．我々のシステムでは車両を画像内から検出する段階でクラス識別を行うことを考えていたが，一般にクラス識別手法ではクラス間の差違が大きくあらわれる評価値を用いることで分類する．この手

法を我々のシステムに取り込むためには、前段階で車両検出を行い車両を画像中から切り出し正規化し、クラス分類の評価を行うことが必要であると考えられる。同様に今後の課題として取り組んでいきたい。

また実用化を考えた場合、実時間処理が求められる。画像処理ボードを用いているにも関わらず我々のシステムの 1 フレームの処理時間は数百 ms から数秒で、実時間処理はまだ不可能である。識別の正確性向上を優先する予定であるので、手法が複雑化することが想定されるが、近年のプロセッサの性能向上を考慮すると実時間処理も不可能ではないと思われ、処理の高速化で実時間処理を目指すことも今後の検討課題としたい。

謝 辞

実験で使用した画像の一部は三菱重工業株式会社のテストコースで撮影された画像を提供していただいたものです。画像の提供以外にも、お忙しい中何度も打ち合わせをしていただき、本研究に多大なご協力をいただいた三菱重工業株式会社の見持圭一さん、塘中哲也さん、杉本喜一さんに慎んで感謝いたします。

本研究を進めるにあたり、すばらしい研究環境を整えていただき、種々の業務や研究でご多忙な中、毎週のように打ち合わせをしていただき丁寧にご指導をしてくださった池内克史教授に心より感謝いたします。

本研究の指導や研究内容の相談に乗っていただくだけでなく、計算機環境の管理や、会議出席や発表の際には研究以外にもご指導、ご協力をしていただいた影澤政隆助手に深く感謝いたします。

打ち合わせで質問に答えていただいたり、毎回様々な助言や提案、指導をしていただいた、上条俊介講師、川崎洋さん、平原清隆さん、小野晋太郎さん、Shirmila Mohottalaさん、坂内研究室の松下康之さん、西田恒俊さんに感謝いたします。

研究でのアドバイスや協力だけでなく、同学年ということで研究生活全体でもお世話になった、橋本謙太郎さん、宮崎大輔さん、高橋徹さん、大石岳史さん、佐藤研究室の岡兼司さんに感謝いたします。

学会出席や事務手続きで大変お世話になった池内研究室の秘書のみなさま、研究に対するご指導の他にも研究生活を進めるにあたりご尽力していただいた佐藤洋一助教授、池内研究室の職員のみなさまに感謝いたします。

発表練習を開いていただくなど研究に対してアドバイスや協力をいただき、研究生活を送るにあたっていろいろと支えていただいた池内研究室の学生のみなさま、研究室こえてお世話していただいた佐藤研究室、坂内研究室のみなさまに感謝いたします。

最後に学生生活を支えてくれた家族へ感謝を送ります。

平成 14 年 2 月 1 日

発表文献

- [1] 吉田, 影澤, 塘中, 池内 “局所特徴認識アルゴリズムによる車両の認識” 電子情報通信学会研究報告 (PRMU研究会), 101, 302, pp. 9-14, Sep. 2001.
- [2] 吉田, 影澤, 池内 “局所特徴認識アルゴリズムによる車両の認識” 第19回 日本ロボット学会学術講演会, pp. 1119-1120, Sep. 2001.
- [3] T. Yoshida, M. Kagesawa, T. Tomonaka and K. Ikeuchi, “Vehicle Recognition with Local-Feature Based Algorithm Using Parallel Vision Board” *8th World Congress on Intelligent Transportation Systems*, Sydney, Sep. 2001.

参考文献

- [4] E. Oja, "Subspace Methods of Pattern Recognition," *Research Studies Press*, Hertfordshire, 1983.
- [5] R. M. Gray, "Vector Quantization," *IEEE ASSP Magazine*, pp. 4-20, April 1984.
- [6] W. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes in C*, Cambridge University Press, Cambridge, 1988.
- [7] C. Tomasi and T. Kanade, "Shape and Motopn from Image Streams: A Factorization Method-2. Point Features in 3D Motion," *Thechnical Report CMU-CS-91-105*, Carnegie Mellon Univ., Pittsburgh, Penn., Jan. 1991.
- [8] S. B. Kang and K. Ikeuchi, "Toward automatic robot instruction from perception—Recognizing a grasp from observation," *IEEE Trans. on Robotics and Automation*, 9, pp. 432-443, 1993.
- [9] J. Shi and C. Tomasi, "Good Features to Track," *Proc. of the IEEE Computer Vision and Pattern Recognition Conference '94*, pp. 593-600, June 1994.
- [10] H. Murase and S. K. Nayer, "Visual Learning and Recognition of 3-D Object from Appearance," *International Journal of Computer Vision*, 14, 1, pp. 5-24, Jan. 1995.
- [11] J. M. Ferryman, A. D. Worrall, G. D. Sullivan and K. D. Baker, "Visual Surveillance Using Deformable Models of Vehicles," *Proc. of International Symposium on Intelligent Robotic Systems*, Pisa, Italy, pp. 295-302, July 1995.

- [12] J. M. Ferriman, A. D. Worrall, G. D. Sullivan and K. D. Baker, "A Generic Deformable Model for Vehicle Recognition," *Proc. of 6th British Machine Vision Conference*, Birmingham, U.K., pp. 127-136, Sep. 1995.
- [13] G. D. Sullivan, K. D. Baker, A. D. Worrall, C. I. Attwood and P. Remagnino, "Model-based Vehicle Detection and Classification Using Orthographic Approximations," *Proc. of 7th British Machine Vision Conference*, 2, pp 695-704, Sep 1996.
- [14] J. Krumm, "Object Detection with Vector Quantized Binary Features," *Proc. of Computer Vision and Pattern Recognition '97*, San Juan, Puerto Rico, pp. 179-185, June 1997.
- [15] K. Ohba and K. Ikeuchi, "Detectability, Uniqueness, and Reliability of Eigen Windows for Stable Verification of Partially Occluded Objects," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19, 9, pp. 1043-1048, Sep. 1997.
- [16] A. SAito, M. Kimachi and S. Ogata, "Silhouette Vision: New Video Vehicle Detection Field Proven Robust and Accurate," *6th World Congress on Intelligent Transport Systems*, Toronto, Canada, Nov. 1999.
- [17] J. M. Ferriman, S. J. Maybank and A. D. Worrall, "Visual Surveillance for Moving Vehicles," *International Journal of Computer Vision*, 37, 2, pp. 187-197, June 2000.
- [18] M. Kagesawa, A. Nakamura, K. Ikeuchi and H. Kashiwagi, "Vehicle Type Classification in Infra-red Image Using Parallel Vision Board," *7th World Congress in Intelligent Transport Systems*, Torino, Italy, Nov. 2000.
- [19] M. Kagesawa, S. Ueno, K. Ikeuchi and H. Kashiwagi, "Recognizing Vehicles in Infrared Images Using IMAP Parallel Vision Board," *IEEE Trans. on Intelligent Transportation Systems*, 2, 1, pp. 10-17, March 2001.
- [20] O. Masoud, N. P. Papanikolopoulos and E. Kwon, "The Use of Computer Vision in Monitoring Weaving Sections," *IEEE Trans. on Intelligent Transportation Systems*, 2, 1, pp. 18-25, March 2001.

-
- [21] K. Dietmayer, J. Sparbert and D. Streller, "Model Based Classification and Object Tracking in Traffic Scenes from Range-Images," *IEEE Symposium Intelligent Vehicles 2001*, Tokyo, Japan, May 2001.
- [22] J. Till, "A Vision Based Automated Vehicle Classification system," *8th World Congress on Intelligent Transport Systems*, Sydney, Austraria, Sep. 2001.
- [23] S. M. Tabib, "Inductance-Pattern Recognition for Vehicle Re-Identification," *8th World Congress on Intelligent Transport Systems*, Sydney, Austraria, Sep. 2001.