

修 士 論 文

車両前方映像中の 道路標識の認識

平成 24 年度修了

三重大学大学院工学研究科
博士前期課程 情報工学専攻

苫米地 広貴

はじめに

現在、郊外の施設はすべてが自動車の保有を前提として出来上がっているため、自動車が必要不可欠になりつつあり、日本の自動車保有率は年々増加している。全国平均で1世帯当たり1台を超え、公共交通機関が整っている首都圏では1世帯当たり1台を下回っているが、地方では1世帯当たり1.5台を上回ることもある。人々の生活は豊かになっているが、自動車の全体量が増えるとともに交通事故件数も増加している [1]。しかし死者数は減少している。これは科学技術が発達し、安全運転支援システムと呼ばれる装置が開発されていることが例に挙げられる。このシステムは一部の自動車のみ搭載されているが、これからはより一般的になり普及していくと考えられる。

道路標識は常に変化する道路の通行方法等をドライバーに伝える役割を持ち、道路交通安全性・円滑性を確保する上で重要な存在である。しかし現在の道路標識システムでは、道路標識の認識はドライバーに委ねられており、標識の未確認・見間違いがしばしばある。この問題を解決するため、安全運転支援システムのように、前方の映像から自動で道路標識を認識することが出来れば、見落としにより起こる事故の危険性が減り、安全運転のサポートが出来るようになる。

これまでに標識認識のための手法は数多く提案されている。道路標識の構成色およびその比率を特徴量として認識を行うなど、標識の色特徴を用いるもの [2] と、テンプレートマッチングにより標識を認識するなど、形状特徴を用いるものである [3]。前者の利点は、道路標識が正面を向いていなかった場合であっても、標識に含まれる色の割合の変化が小さいため、認識が可能な点であり、欠点は、天候や時間帯によって色情報が変化し易く、認識に失敗する可能性があることである。後者の利点は、テンプレート画像に対する入力画像の全体的な明度の差に影響されにくいという点であり、欠点は、マッチングを行う際の計算量が多いことである。本研究では手書き数字認識の高精度化のために提案された若林ら [13] の濃度こう配特徴を道路標識に対して適用し、色や形状を限定することなく一般の道路標識を認識することを目指す。認識実験に用いる標識画像データセットとしては、Ruhr-Universitt Bochum の Real-time Computer Vision グループが提供する The German Traffic Sign Recognition Benchmark(GTSRB) データベースを使用する。

評価用の画像からこう配特徴を抽出し，SVM による認識実験を行ったところ 96.69%の認識率が得られた．低解像度の標識画像や，ノイズ，オクルージョン，ハイライト，ぶれを含む標識は誤認識をすることが多いため，これらの標識を認識出来るようにすることが今後の課題として挙げられる．

本論文では，第 1 章に研究背景と目的，第 2 章に従来研究，第 3 章に提案手法，第 4 章に実験方法と結果と考察，第 5 章にまとめと今後の課題をそれぞれ述べる．

目次

はじめに	i
第 1 章 緒言	1
1.1 研究背景	1
1.2 研究目的	2
第 2 章 標識認識に関する従来研究	3
2.1 色情報を用いる手法	3
2.2 形状情報を用いる手法	4
2.3 色情報と形状情報の両方を用いる手法	5
2.4 生成型学習法	5
第 3 章 提案手法	6
3.1 提案手法の概要	6
3.2 濃度こう配抽出	8
3.3 ヒストグラム平坦化	9
3.4 特徴抽出	9
3.5 次元削減	12
3.6 分類	14
第 4 章 認識実験	15
4.1 標識画像データセット	15
4.2 評価方法	19
4.3 実験	19
4.4 他手法との比較	19
4.5 考察	21
第 5 章 結言	32

5.1	まとめ	32
5.2	今後の課題	32
付録 A	付録	33
A.1	コンパイル方法	34
A.2	実行方法	34
A.3	修論研究発表資料	35
謝辞		36

第 1 章

緒言

1.1 研究背景

現在，郊外の施設は，すべてが自動車の保有を前提として出来上がっているため，自動車が必要不可欠になりつつあり，日本の自動車保有率は年々増加している．全国平均で 1 世帯当たり 1 台を超え，公共交通機関が整っている首都圏では 1 世帯当たり 1 台を下回っているが，地方では 1 世帯当たり 1.5 台を上回ることもある．人々の生活は豊かになっているが，自動車の全体量が増えるとともに，交通事故件数も増加している [1]．しかし死亡者数は減少している．これは科学技術が発達し，自動車運転の安全性が上がってきているからである．その代表としてあげられる装置にエアバッグ，ABS(Antilock Brake system)，スタビリティ・コントロール・システムがある．

また，飲酒運転の厳罰化や警察による厳しい取り締まりの成果もあげられる．飲酒運転による違反というものは昔からあったが，現在に比べると非常に罰則が緩かったため飲酒運転は絶えなかった．しかし，飲酒運転による交通事故は重大になることが多く，実際，死亡事故が増えてきたため，厳罰化が進んでいる．

さらに近年では，VICS と呼ばれる渋滞や交通規制などの道路交通情報をリアルタイムに送信し，カーナビゲーションなどの車載器に文字・図形で表示する画期的な情報通信システムが登場し，ルート案内しか出来なかったカーナビゲーションも音声案内により先の危険を知らせるまでに発達した．また，ステレオカメラで前方を監視し，衝突の危険性があれば自動ブレーキで車両を停止させるなどの安全運転支援システムが自動車に搭載され始め，今後の自動車の安全を担う存在として注目されている．

道路標識は常に変化する道路の通行方法等をドライバーに伝える役割を持ち，道路交通安全性・円滑性を確保する上で重要な存在である．しかし現在の道路標識システムでは，道路標識の認識はドライバーに委ねられており，標識の未確認・見間違いがしばしばある．安全運転支援システムのように，正確な道路標識認識が実現すれば，一時停止の見

落としや速度超過による事故を未然に防ぐことが出来る．

1.2 研究目的

本研究は，手書き文字認識のために開発された濃度こう配特徴ベクトルを基本として，道路標識の認識に適した認識システムを構築することを目的とする．

第 2 章

標識認識に関する従来研究

近年，車載カメラと画像認識技術を組み合わせた自動車の安全運転支援システム開発への期待が高まっている．これらのシステムの主な効果として，交通事故を未然に防ぐこと，ドライバーの判断負荷の軽減が挙げられる．安全運転支援システムは多くの種類があるが，本稿では車載カメラを用いた道路標識認識について検討する．正確な道路標識認識が実現すれば，一時停止の見落としや速度超過による事故を未然に防ぐことが出来る．

この章では，標識認識に関する研究について紹介する．標識認識の手法は大きく次の 3 つに分類することが出来る．

- 色情報を用いる手法
- 形状情報を用いる手法
- 色情報と形状情報の両方を用いる手法

以下では，上記 3 つのそれぞれに属する手法について簡単に述べる．

2.1 色情報を用いる手法

2.1.1 標識のカラー含有量を用いた手法

明珍ら [2] によって提案された，道路標識の構成色およびその比率を特徴量として認識を行う手法である．入力された標識画像のカラー含有率を求め，その割合から標識を認識する．この手法の利点は，道路標識が正面を向いていなかった場合であっても，標識に含まれる色の割合の変化が小さいため，認識が可能な点である．欠点は，天候や時間帯によって色情報が変化し易く，認識に失敗する可能性があることである．

2.2 形状情報を用いる手法

2.2.1 テンプレートマッチングを用いた手法

竜円ら [3] によって提案された、テンプレートマッチングを用いて認識を行う手法である。テンプレートと呼ばれる標識の一部の形状特徴を用いて、入力画像の部分画像とテンプレート画像の濃度値の差を計算し、差が最小の標識クラスを結果として出力する。この手法の利点は、テンプレート画像に対する入力画像の全体的な明度の差に影響されにくいという点である。欠点は、拡大・縮小・回転・照明変化を行い変化させたテンプレート画像が必要になるため、膨大な記憶容量が必要となること、マッチングを行う際の計算量が多いことである。

2.2.2 SIFT 特徴量を用いた標識認識

SIFT(Scale Invariant Feature Transform)[5] は、特徴点の検出と特徴量の記述を行うアルゴリズムである。検出した特徴点に対して、画像の回転・ノイズ・スケール変化・照明変化に頑健な特徴量を記述するため、イメージモザイク等の画像のマッチングや物体認識・検出に用いられる。高木ら [6] は SIFT 特徴量を道路標識を適応した。SIFT 特徴量を用いることで、道路標識の拡大・縮小・傾き・オクルージョンに対してロバストな認識が可能となることを確認した。

2.2.3 ピクトグラムの輪郭に基づいた認識手法

ピクトグラムとは、文字に代わって、情報や注意を示すために表示される視覚記号である。標識は限られた範囲で設置するため、基本的にピクトグラムを用いてドライバーに情報を伝達する。標識におけるピクトグラムは各標識で一意的に決まるため、形状特徴の一つと考えられ、標識認識にしばしば用いられる。

Paulo と Correia[7] は標識ピクトグラムの外面形状を利用する認識アルゴリズムを提案した。ピクトグラムの輪郭抽出は、動的輪郭モデル [8] によって輪郭形状を捉え、Curvature Scale Space (CSS:曲率尺度空間) [9] という特徴量を用いて行われる。CSS は局所的な特徴から大局的な特徴まで広く捉えることができ、雑音に強く、アフィン変換による形状の変形にも強い特色を持つ。しかし、輪郭形状の直線部分に歪みが発生した場合、CSS は直線部分に輪郭線を表現する上で必要のない冗長な特徴量を検出するという欠点が指摘されている [10]。

2.3 色情報と形状情報の両方を用いる手法

2.3.1 ファジィベースの認識手法

Fleyeh[11]により，標識の色情報と形状情報を用いたファジィベースのアプローチが提案された．画像中から標識を検出するために，まず色らしさの定義を，各ピクセルの Hue（色相），Saturation（彩度），Value（明度）に依存したファジィルールの集合に基づき決定する．その決定した色らしさと画像中の色を照らし合わせることで，特定の色を抽出できる．さらに，ファジィルールは矩形，三角形，楕円，八角系の形状を分類するために使用され，特定の色を持つ箇所と照らし合わせ，標識を認識する．この手法は，異なった環境状況において高いロバスト性があるが，オクルージョンのある標識は上手く扱うことが出来ない．

最後に学習法に関する手法を紹介する．

2.4 生成型学習法

計算機は学習したデータベースをもとに認識処理を行うため，様々な状況下に対応できる大量の学習用データ必要になる．しかし，道路標識認識においては，標識自体の変化やカメラと標識の位置関係，時間帯や天気等の違いから，撮影された標識画像は様々に変動するため，様々な状況下の全ての学習データを収集するのは難しく非現実的である．この問題を解消するために石田ら [12] は，生成パラメータの推定を通して実際の劣化に従い，自動的に学習データを生成する生成型学習法を提案した．この手法の利点としては，学習に用いる全ての学習画像を撮影によって収集する必要がないため，学習過程が効率化されていること，クラス数を増加させる場合でも，クラスに対する原画像の追加のみで対応出来ることが挙げられる．

第 3 章

提案手法

3.1 提案手法の概要

提案手法による標識認識の流れを図 3.4 に示す．本手法では，2 つの手法を検討した．

1 つ目の手法は，濃度こう配を抽出する際，上下左右から 2 つ分の小領域の特徴量を 0 で置き換え，擬似的に背景を除去する手法である．標識周辺の背景は常に変動するため，背景はノイズと考えられ，認識に悪影響を及ぼす．そのため，いかにノイズを除去し，特徴抽出するかがポイントになる．しかし，濃度こう配特徴は画像全体を特徴抽出の対象とするため，このノイズを含んでしまう．この問題を解消して認識性能を向上させるため本手法では，濃度こう配特徴作成の際，画像を小領域に分割すると同時に，周辺小領域を 0 で置換することで，擬似的に背景を除去し，特徴量の大きな変動を抑制する．

2 つ目の手法は，カラー画像の R（赤）要素，G（緑）要素，B（青）要素それぞれから濃度こう配を抽出し，その 3 つに主成分分析によって次元削減する手法である．文献 [13] の手書き文字認識では，入力画像から濃度こう配を抽出し形状を捉えているが，背景が白，文字が黒の 2 値画像のため，色情報は用いていない．しかし標識認識の場合は，各標識は形状と配色が決められているため，形状情報だけでなく色情報も用いることは認識率の向上に有効であると考えられる．

提案手法におけるこう配抽出方法の流れを図 3.1 に示す．

3.2 節で濃度こう配抽出について説明し，3.3 節から 3.5 節で図 3.4 の流れ図に沿って各処理について詳しく述べる．



図 3.1: 提案するこう配抽出手法の流れ

3.2 濃度こう配抽出

濃度こう配とは，各画素値の変化の度合いを示すものであり，2つの特徴量を持つ．この2つの特徴量は，こう配強度とこう配方向である．本手法では，Roberts オペレータを用いて，各画素における濃度こう配の強度と方向とを算出する．Roberts オペレータとは，こう配に基づいたエッジ検出法である． $u(45^\circ)$ 方向， $v(135^\circ)$ 方向それぞれの偏微分を計算することにより画像からエッジを検出することができる．

u 方向， v 方向それぞれのカーネルを図 3.2 に示す．画素 $p(x, y)$ におけるこう配の u 方向， v 方向の成分をそれぞれ $G_u(p)$ ， $G_v(p)$ とし，こう配強度を $|G(p)|$ ，こう配方向を $\theta(p)$ とする．ここで，こう配強度とこう配方向はそれぞれ式 3.1，3.2 で与えられる．図 3.2 のそれぞれのオペレータを適用することで，水平方向，垂直方向の成分である $G_u(p)$ ， $G_v(p)$ を取得し，式 3.1，3.2 により，こう配強度 $|G(p)|$ とこう配方向 $\theta(p)$ を求める．

以上により抽出した濃度こう配の画像例を図 3.3 に示す．これは，こう配強度を明度，こう配方向を色相によって表している．図 3.3 を見ると，標識内部の数字に強く特徴的なこう配分布があることが見てとれる．

$$|G(p)| = \sqrt{G_u(p)^2 + G_v(p)^2} \quad (3.1)$$

$$\theta(p) = \tan^{-1} \frac{G_v(p)}{G_u(p)} \quad (3.2)$$

1	0
0	-1

(a) 水平方向

0	1
-1	0

(b) 垂直方向

図 3.2: Roberts オペレータ

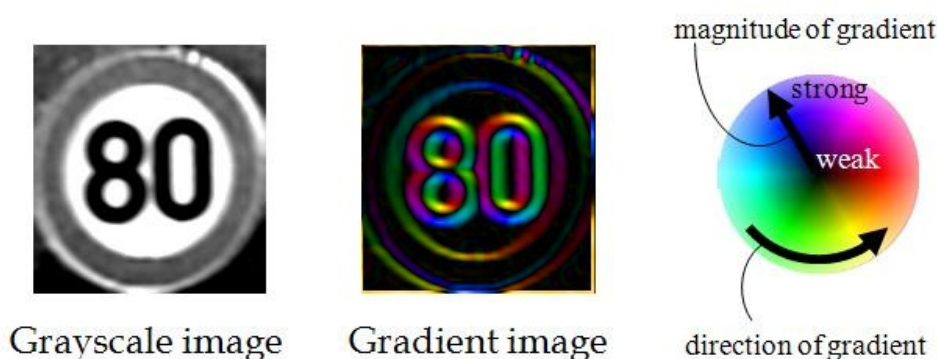


図 3.3: 濃度こう配の画像例

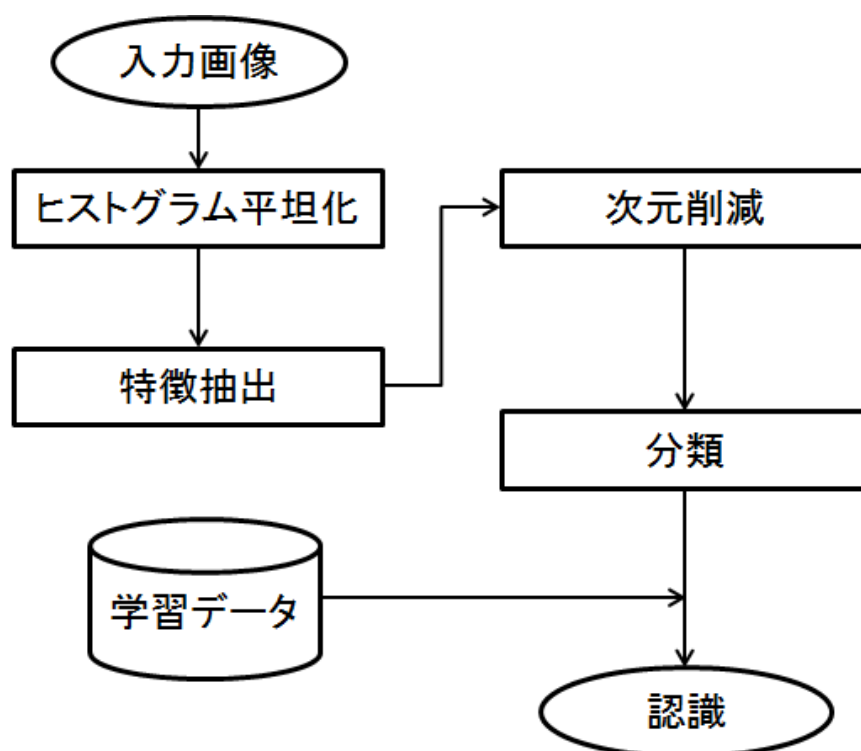


図 3.4: 提案手法の流れ

3.3 ヒストグラム平坦化

道路に設置される標識は太陽や街灯により照明変化が起きる．その結果，同種類の標識でもヒストグラムにおける画素の分布に偏りが生じてしまい認識に悪影響を及ぼす．この問題を解決するため，輝度値のヒストグラム平坦化 [15] の処理を行う．ヒストグラム平坦化し，コントラストを強調した結果を図 3.5 に示す．ヒストグラム平坦化によってコントラストが強調される理由は，次のように考えられる．一般の画像では，濃度の分布が密な部分（ヒストグラムの山）と疎な部分（ヒストグラムの谷）があり，ヒストグラムの平坦化によって，前者の部分が引き伸ばされ，後者の部分は圧縮される．一般に，濃度範囲が圧縮される部分に比べ，引き伸ばされる部分に属する画素の数が多く，全体としてはコントラストが強調される．

3.4 特徴抽出

本研究では，標識画像の見え方を濃度こう配ヒストグラム（HOG）を用いて表現する．濃度こう配ヒストグラムは，画像中の対象物の見かけを濃度こう配の局所方向ヒストグラ

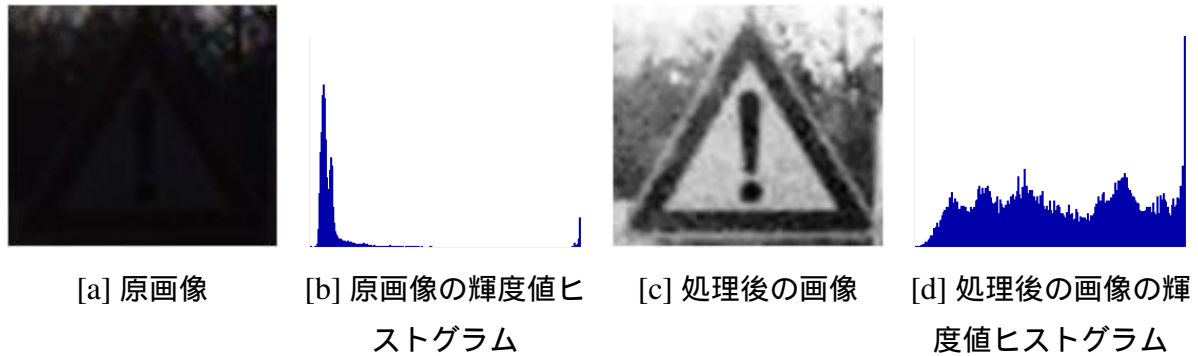


図 3.5: ヒストグラム平坦化

ムで記述する特徴量であり，手書き文字認識において高い分離性能を有する特徴量として知られている [13, 16]．本手法では，この濃度こう配ヒストグラムを標識画像から抽出する．

特徴抽出の流れは，まず入力画像中の対象領域に対してノイズ除去のために平滑化を行う．続いて，Roberts オペレータにより各画素における濃度こう配の強度と方向とを算出する．そして対象領域を複数の小領域に分割し，濃度こう配の局所方向ヒストグラムを生成する．最後に，生成された局所方向ヒストグラムに対して再標本化などの後処理を行い，次元数の削減を行う．

以下では，それぞれの手順について詳細を述べる．

1. 平滑化

対象画像 $I = I(p)$ の画素 $p = (p_x, p_y)$ に対して， 3×3 の移動平均フィルタを適用する．移動平均フィルタは，注目画素とその周辺の輝度値を用いて，輝度値を平均し，処理後画像の輝度値とする手法である．この処理により，画像中のノイズを除去する．

2. 濃度こう配抽出

Roberts オペレータを用いて，濃度こう配の強度と方向とを算出する．画素 p における濃度こう配の強度を $|G(p)|$ ，方向を $\theta(p)$ とすると，それぞれ以下の式で与えられる．なお， $G_u(p), G_v(p)$ は，それぞれ画素 p における濃度こう配の u, v 方向の成分である．

$$|G(p)| = \sqrt{G_u(p)^2 + G_v(p)^2} \quad (3.3)$$

$$\theta(p) = \tan^{-1} \frac{G_v(p)}{G_u(p)} \quad (3.4)$$

3. 小領域分割と局所方向ヒストグラムの生成

まず、それぞれの画素で算出された濃度こう配の強度と方向とを局所ヒストグラム化するために、対象領域を $n_x \times n_y$ の小領域に分割する．それと同時に、上下左右から n 分の小領域に対して 0 で置換する．さらに、量子化数 L により方向ごとの量子化を行う．続いて、 $n_x \times n_y$ 個の小領域それぞれの内部で、 L 次元の方向ごとに、濃度こう配の強度を積算していく．この処理によって、 $n_x \times n_y \times L$ の原特徴ベクトルを得る．

4. 再標本化による次元削減

抽出された特徴ベクトルの次元数を削減するために再標本化を行う．ヒストグラム化された濃度こう配強度を分割された任意の小領域 $(2i, 2j)$ ($i = 0, 1, \dots, \frac{n}{2} - 1, j = 0, 1, \dots, \frac{m}{2} - 1$) を中心とする 5×5 の領域について、方向毎にガウシアンフィルタを用いて平滑化し、領域数を $n_x/2, n_y/2$ に削減する．続いて、 L 段階に量子化された濃度こう配の方向に対し、ひとつおきに窓の中心を設定し、 $[1\ 4\ 6\ 4\ 1]$ の重みを用いて加重平均し、方向を $L/2$ に削減する．さらに同じように、この $L/2$ 段階に量子化された濃度こう配の方向に対し、ひとつおきに窓の中心を設定し、 $[1\ 2\ 1]$ の重みを用いて加重平均し、方向を $L/4$ に削減する．

なお、特徴抽出には様々なパラメータが存在するが、本研究では、小領域分割数 $n_x \times n_y$ は 13×13 、方向量子化数 L は 32、小領域を数値 0 で埋める数 n は 2 に設定した．さらに、R 要素、G 要素、B 要素の 3 種に対して特徴抽出を行い、それぞれの特徴量を連結させ、新たな特徴量とする．したがって、原特徴ベクトルは $8 \times 8 \times 7 \times 3$ の 1176 次元の特徴量となる．

5. 特徴ベクトルの正規化

画像間の様々なコントラストに起因する変動を最小限に抑えるため、抽出した特徴ベクトル z を以下の式に従って正規化する．

$$x = \frac{z}{|z|} \quad (3.5)$$

6. 変数変換

一般的に、カウントすることにより観測された正の値のみからなるヒストグラムは、図 3.6 のようにガンマ分布となることが多い．このとき、次式のような、べき変換によって観測値を変換すると、平均の周りの対称性と正規分布への近似がしばしば顕著に改善される．(図 3.6)

$$y = x^u \quad (3.6)$$

ここで、 x は変換される値で、 u はべき乗を示す．文献 [17] には $u = 0.4$ のときにガンマ分布が最も正規分布に近くなることが示されている．

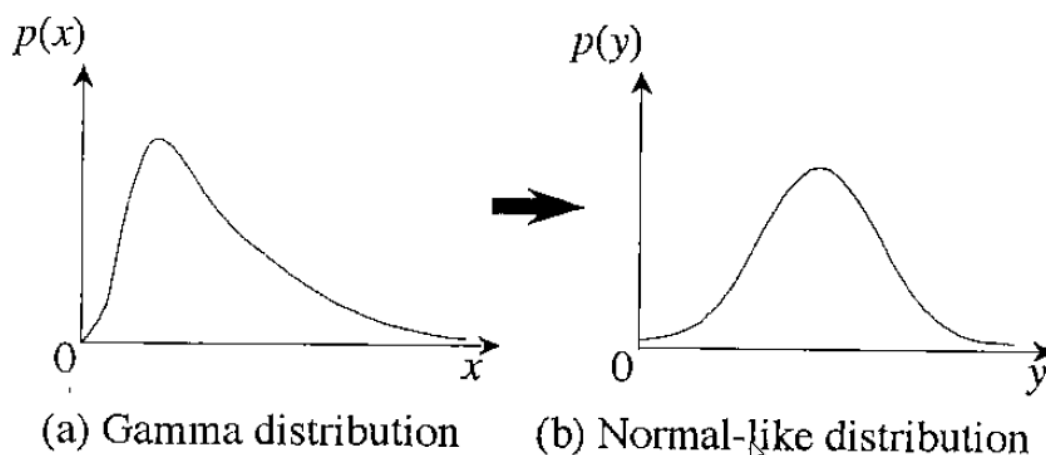


図 3.6: 変数変換

3.5 次元削減

3 種の濃度こう配特徴を抽出して生成される原特徴ベクトルはいずれも高次元となる。高次元の原特徴ベクトルには相関の高い要素の組が含まれ、冗長性を含んだ特徴ベクトルになる。また、高次元の特徴ベクトルを照合、認識に用いると誤確率や計算コストが増加する。そのため、特徴抽出により得られた特徴ベクトルに対し、主成分分析により次元削減を行う。これにより照合、認識に有効であると考えられる特徴が選択され、精度の向上が期待できる。

3.5.1 主成分分析 (Principal Component Analysis : PCA)

PCA は多次元空間上の特徴点を分散の大きい少数の低次元の直交部分空間に線形射影する手法である。多次元空間上の特徴点をより見やすくあるいは扱いやすくするために、固有空間を利用して少ない次元で表現する手法である。多次元の特徴量の低次元化により、照合、認識に有効な特徴の選択が出来る。例として、2 次元の特徴ベクトルで表現された 2 クラスのサンプルを 1 次元に削減する場合を図 3.7 に示す。図 3.7 において、全サンプル分布の分散が最大となる方向に主軸 y_1 をとる。サンプルを主軸 y_1 の投影により、1 次元データでも 2 クラスの分離が可能となる。

n 次元の原特徴ベクトルを PCA により n' 次元に削減する手順を以下に示す。

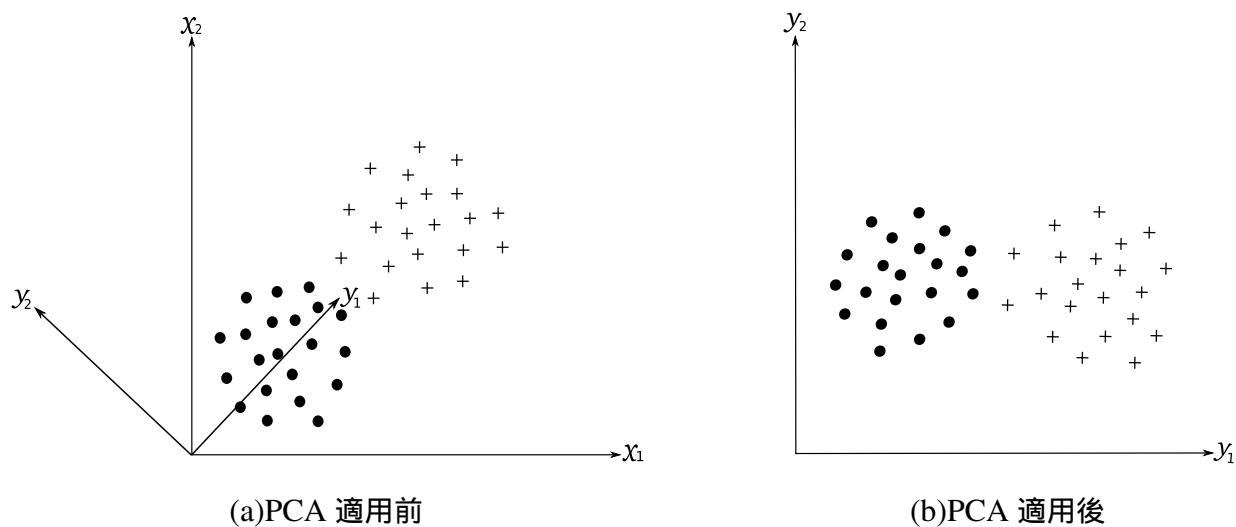


図 3.7: PCA による次元削減の例

学習データに含まれる画像 j ($j = 1, 2, \dots, m$) から抽出した特徴ベクトルを

$$\mathbf{x} = [x_1 \quad x_2 \quad \cdots \quad x_n] \quad (3.7)$$

と表す。

学習データの全平均ベクトル M ，共分散行列 Σ_t をそれぞれ以下の式により計算する。

$$M = \frac{1}{m} \sum_{j=1}^m \mathbf{x}_j \quad (3.8)$$

$$\Sigma_t = \sum_{j=1}^m (\mathbf{x}_j - M)(\mathbf{x}_j - M)^T \quad (3.9)$$

以下の式を満たす固有値行列 Λ ，固有ベクトル行列 Φ をそれぞれ求める。

$$\Sigma_t \Phi = \Phi \Lambda \quad (3.10)$$

固有値を大きい順にソートし，各固有値に対応する d 個の固有ベクトルを用いて，以下の式により原特徴ベクトルを d 次元まで削減する ($d \leq n$)。ここで \mathbf{x} は原特徴ベクトル， \mathbf{y} は次元削減後の特徴ベクトルである。

$$y_i = \Phi_i^T \mathbf{x} \quad (i = 1, 2, \dots, d) \quad (3.11)$$

$$\mathbf{y} = [y_1 \quad y_2 \quad \cdots \quad y_i \quad \cdots \quad y_d]^T \quad (3.12)$$

3.6 分類

分類器には SVM(Support Vector Machine)[18] を用いる．SVM は，現在知られている多くの手法の中で最も認識性能が優れた学習モデルの一つであり，主に 2 クラスの分類を行う分類機械である．SVM が 2 クラス識別器として優れているモデルである理由に，学習サンプル中のクラス境界に最も近いサンプル（サポートベクター）との距離が最大となる境界面を求めるマージン最大化と，カーネル学習法による非線形の判別問題へ拡張が挙げられる．本研究では LIBSVM Ver.3.11[19] を用いた．

SVM は本来は線形分離不可能な問題には適用できない．線形分離不可能な問題へ対応するため，ソフトマージン法とカーネル関数法が考案されている．

ソフトマージン法とは，2 クラスの分布が線形分離不可能な場合に，学習サンプルがマージン内にあったり，決定境界の対立クラス側にある状況がある程度許容する方法である．しかし，対立するクラスに分類されてしまったり，正しく分類されているがマージン内に侵入してしまったサンプルにはペナルティを与える．ペナルティと，マージンの大きさの間のトレードオフを制御するパラメータとして， $cost$ という変数を用いる．

カーネル関数法とは，カーネル関数を用いてパターンを有限もしくは無限次元の特徴空間へ写像し，特徴空間上で線形分離を行う手法である．本研究では，多くの場合に最も高い分離性能を示す，RBF (Radial Basis Function) カーネルを用いた．RBF カーネルは，

$$K(x_1, x_2) = \exp\left(-\gamma |x_1 - x_2|^2\right), \gamma > 0 \quad (3.13)$$

で表される．カーネル関数を取り入れた一連の手法では，どのような写像が行われるか知ることなく，高次元空間中における 2 点間の距離や類似度を計算できるので，カーネルトリックと呼ばれる．

しかし， $cost$ と γ には適切な値を設定する必要がある．本研究では LIBSVM ライブラリに付属するパラメータサーチプログラムを用いて， $cost$ と γ の値を決定した．

LIBSVM ライブラリは 2 クラス分類を行うだけでなく，多クラス分類にも拡張ができる．多クラス分類とは，各データがあらかじめ用意されたクラスのどれに分類するかを言う．データ x が条件としてあった時のクラス分類の事後確率 $P(i|x)$ が最大となるようなクラスを選択するように通常設計する．本研究では SVM の基本的な構成のまま，多値の判別関数を直接構成する手法 (Multiclass-SVM) を利用している．

第 4 章

認識実験

4.1 標識画像データセット

本手法の有効性を評価するために、標識画像として Ruhr-Universitt Bochum の Real-time Computer Vision グループが提供する The German Traffic Sign Recognition Benchmark(GTSRB) データベース [20] の画像を用いる。このデータベースは IJCNN 2011 competition[21] で使用された。図 4.1 に道路標識の例を示す。GTSRB データベースの概要を以下に示す。

- ドイツの道路標識 (43 クラス) を含む (図 4.1 参照)。
- 同一の道路標識の画像 30 枚が 1 トラックになっている (図 4.2 参照)。
- 実地撮影時の各標識の出現頻度に合わせたトラック数を設定 (図 4.3 参照)。
- データベースに多様性を持たせるため、照明変化、ハイライト、ぶれ、オクルージョン、ノイズを含んだ画像が存在する (図 4.4)。
- エッジ検出器を用いる場合のため、標識の周辺には画像の 10%分背景が含まれる。エッジ検出器を使用しない場合は、標識の外接枠の座標データが存在するので、この座標データを用いて背景を除去する (図 4.5)。

本研究ではエッジ検出器を使用しないため、座標データを用いて画像中の背景を除去した画像を使用する (図 4.6 参照)。

4.1.1 画像サイズの正規化

GTSRB データベースの画像サイズは 15×15 ピクセルから 222×193 ピクセルと不均一である。そのため、特徴抽出の際、1 ブロックあたりに含まれる画素数が変動し認識に影響する。この影響を避けるため、本研究では、幅のサイズを 65 ピクセルになるように画像サイズの正規化を行う。



図 4.1: GTSRB データベースの全 43 クラスの画像例



図 4.2: 1トラックにおける標識画像例

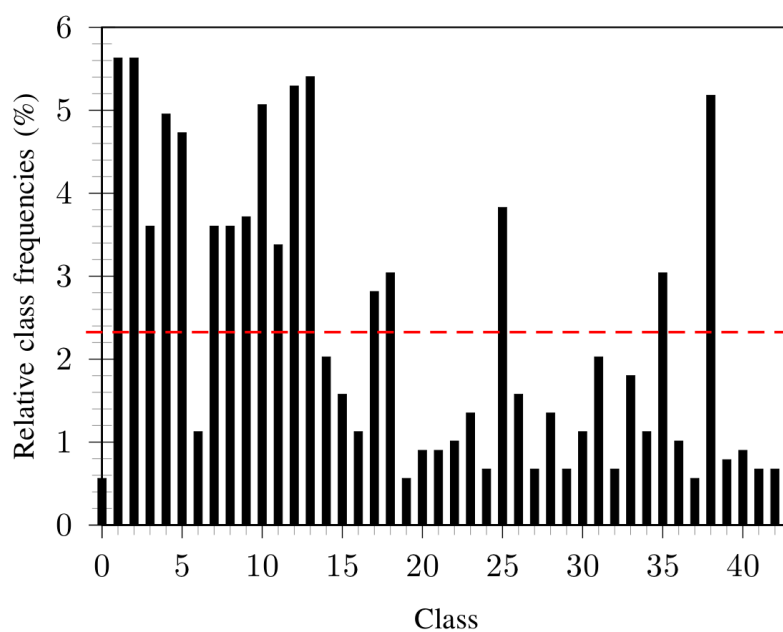


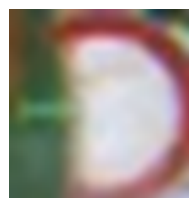
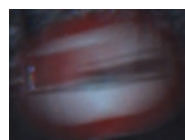
図 4.3: 各標識の出現頻度



[a] 照明変化



[b] ハイライト

[c] オクルー
ジョン

[d] ぶれ



[3] ノイズ

図 4.4: データベースの多様性



図 4.5: 標識の外接枠の座標データ



[a] オリジナル画像の例

[b] 背景除去画像の例

図 4.6: 外接枠の座標データを用いて背景を除去した画像の例

4.2 評価方法

実験は，SVM の学習用に 26,640 枚，評価用に 12,569 枚の計 39,209 枚を使用する．性能評価は，各手法の認識率を求めることで行う．入力画像に対し SVM が出力したクラスと，入力画像のクラスが一致すれば認識成功とし，認識率は，評価画像数に対する認識成功数の割合で定義する．よって，認識率を求める式は以下の通りになる．

$$\text{認識率 (\%)} = \frac{\text{認識成功数}}{\text{評価画像数}} \times 100$$

4.3 実験

提案手法の有効性を確かめるために 4 種類の特徴量を用いて認識実験を行った．便宜上，特徴量名とその内容を表 4.1 のように対応付けた．分類器には SVM を用いた．

特徴量名	内容
Gray 特徴量	グレースケール画像から特徴抽出
Gray_zero 特徴量	グレースケール画像から特徴抽出 + 周囲 2 ブロックの 0 置換
RGB 特徴量	カラー画像の RGB 要素から特徴抽出
RGB_zero 特徴量	カラー画像の RGB 要素から特徴抽出 + 周囲 2 ブロックの 0 置換

表 4.1: 特徴量名とその内容

4.3.1 実験結果

認識実験の結果を図 4.7 に示す．また，クラス分類結果を図 4.8，4.9，4.10，4.11 に示す．認識率の低い順に RGB が 95.08%，gray が 95.44%，gray_zero が 96.56%，RGB_zero が 96.69% となった．カラー画像の RGB 要素から特徴を抽出し，かつ上下左右から 2 つ分の小領域を 0 で置換した濃度こう配特徴量を用いた時，最良の認識率となり，提案手法の有効性を示すことが出来た．

4.4 他手法との比較

IJCNN 2011 competition 参加者上位 4 組の手法 [20] と結果を表 4.2 に示す．Human Performance は 36 人の被験者それぞれにランダムに選ばれた標識画像 350 枚を見せ，回答した時の認識率である．また，HOG1，HOG2，HOG3 は事前に提供された HOG 特徴

量である．それぞれの相違を表 4.3 に示す．この特徴量は，画像サイズを 40×40 に正規化し，グレースケールに変換した画像から特徴抽出してある．

表 4.2 と比較しても，提案手法はトップクラスの認識率であると言える．

チーム	手法	認識率
IDSIA	Committe of CNN+HOG3	98.98%
sermanet	Multi-Scale CNNs	98.97%
INI-RTCV	Human Performance	98.81%
VISICS	IKSVM+PHOG+HOG2	97.88%
VISICS	SRC+LDAs I/HOG1/HOG2	97.35%
noob	HOG+LDA+VQ	96.87%
INI-RTCV	LDA+HOG2	96.31%
INI-RTCV	LDA+HOG3	94.73%
INI-RTCV	LDA+HOG1	94.51%

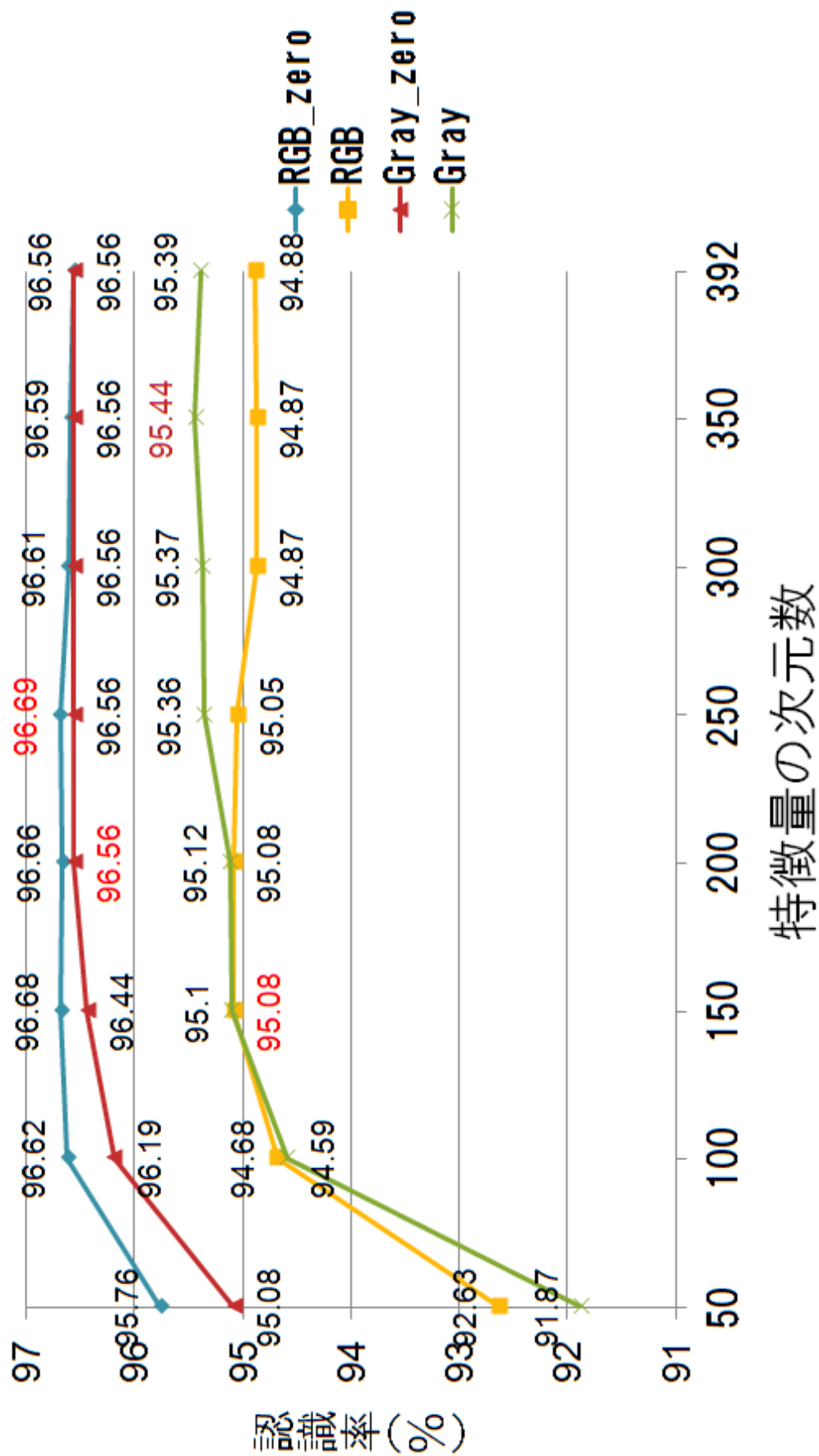
表 4.2: コンペティション参加者の手法と認識率

Name	Dimension	Cell	Block	Stride	Bins	Semicircle
HOG1	1568	5×5	10×10	5×5	8	true
HOG2	1568	5×5	10×10	5×5	8	false
HOG3	2916	4×4	8×8	4×4	9	true

表 4.3: HOG 特徴量のパラメータ

4.5 考察

4 種の特徴量を用いて実験を行った．図 4.8, 4.9, 4.10, 4.11 から読み取れることとして，誤認識はある程度決まった範囲に収まっていることが分かる．例えば，入力された画像が円形の標識であれば円形の標識に，三角形の標識であれば三角形の標識への誤認識が多い．ただ，図 4.9, 4.11 は小領域を 0 で置き換えているため，標識の形状は特に反映されていない．これは標識の形状は考慮せず，標識内部の数字やピクトグラムだけで認識を行った方が標識認識には適していると考えられる．また，RGB 特徴量は Gray 特徴量より認識率が劣るが，RGB_zero 特徴量は Gray_zero 特徴量より認識率は優れている．これは，RGB 要素から特徴抽出をすることで，特徴量の分離性が高くなるが，同時に冗長な特徴も増えてしまう．そこで周囲ブロックの 0 置換をすることで，冗長な特徴を削ぐことが出来，認識率が向上したと考えられる．特徴量の分離性を示すプロット結果を図 4.12, 4.13, 4.14, 4.15 に示す．その一方で，4.10, 4.11 を見る限り，青い標識を赤い標識に誤認識していることが分かる．標識の色には，例えば赤色でも青色の成分を少なからず含んでいたりする．そのため，RGB のこう配の差が出にくく，このような結果になったと考える．そのため，カラーを用いた濃度こう配を抽出する最は，お互いの色を含まないように画像を前処理しておくことが適当であるかもしれない．その前処理の簡単な例として，画素の RGB 値が (129,33,20) であれば一番値が大きいところを最大値にし，その他 2 つを最小値にする，すなわち (255,0,0) にする処理が挙げられる．さらに，全ての特徴量において誤認識している画像がある．それらの画像には，低解像度である，ノイズを含む，オクルージョンを含む，ハイライトを含む，ぶれている，といった特徴がある．それらの画像例を図 4.16 に示す．これらは一般物体認識において認識を困難にさせる要因であり，解決しなければならない問題である．それぞれの問題について，解決のための提案をする．低解像度については，画像を拡大して画素数を増やし，特徴量の小領域に属する画素数を増やすことが重要である．しかし，さらに重要なことは，拡大する際の補間法である．本手法では Imagemagick の convert コマンドにより画像の正規化を実現しており，補間法はバイリニア補間である．補間法は多種あり，低画素な標識にあった補間法を検討する必要がある．ノイズを含む，オクルージョンを含むものについては，局所特徴量の利用である程度対処可能であるが，ノイズが大きかったり，大部分隠れてしまった場合は非常に難しく，解決法は確立されていない．ハイライトを含むものについては，ハイライトを検出，除去し，本来の標識に復元する手法の適用が考えられる．ぶれについては，対象物やカメラがわずかに動くことによって起こるボケ程度ならば，Point-Spread Function(PSF, 点拡がり関数)を用いてあらかじめぼけた画像を学習しておくことで，ある程度対処は可能だが，図 4.16 のような極端なぶれを認識するのは不可能に近い．



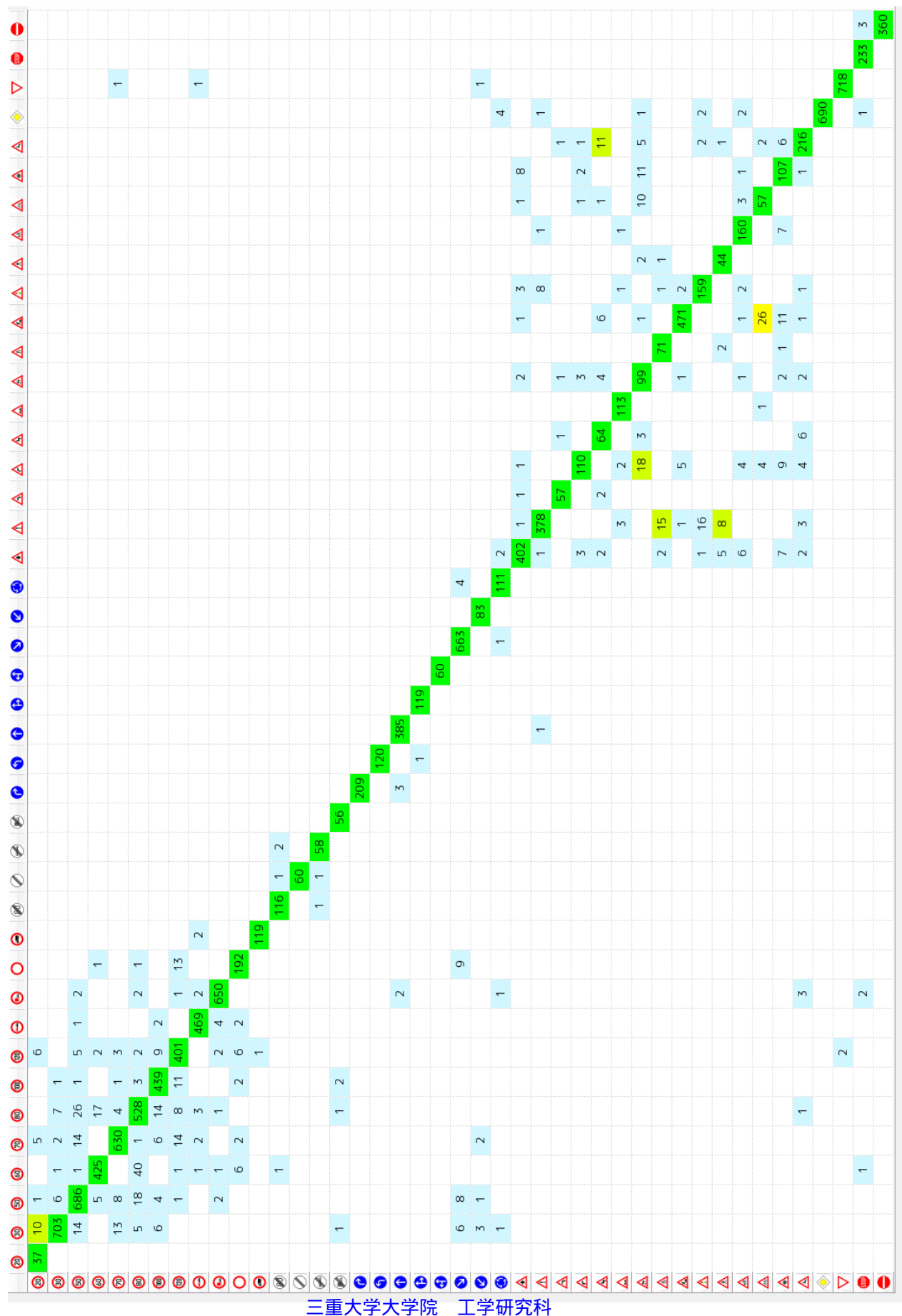


図 4.8: Gray 特徴量を用いた認識実験のクラス分類結果

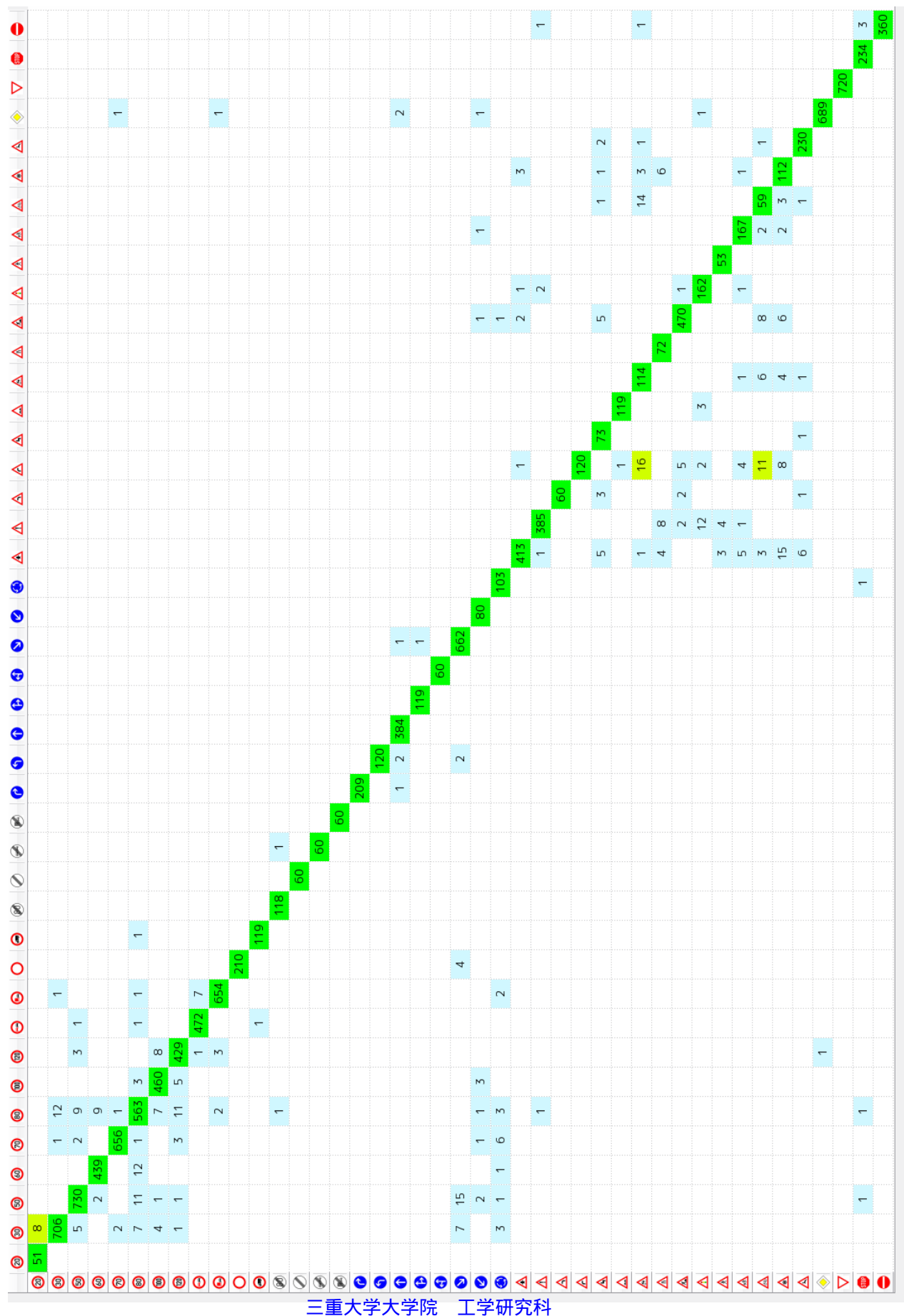


図 4.9: Gray_zero 特徴量を用いた認識実験のクラス分類結果

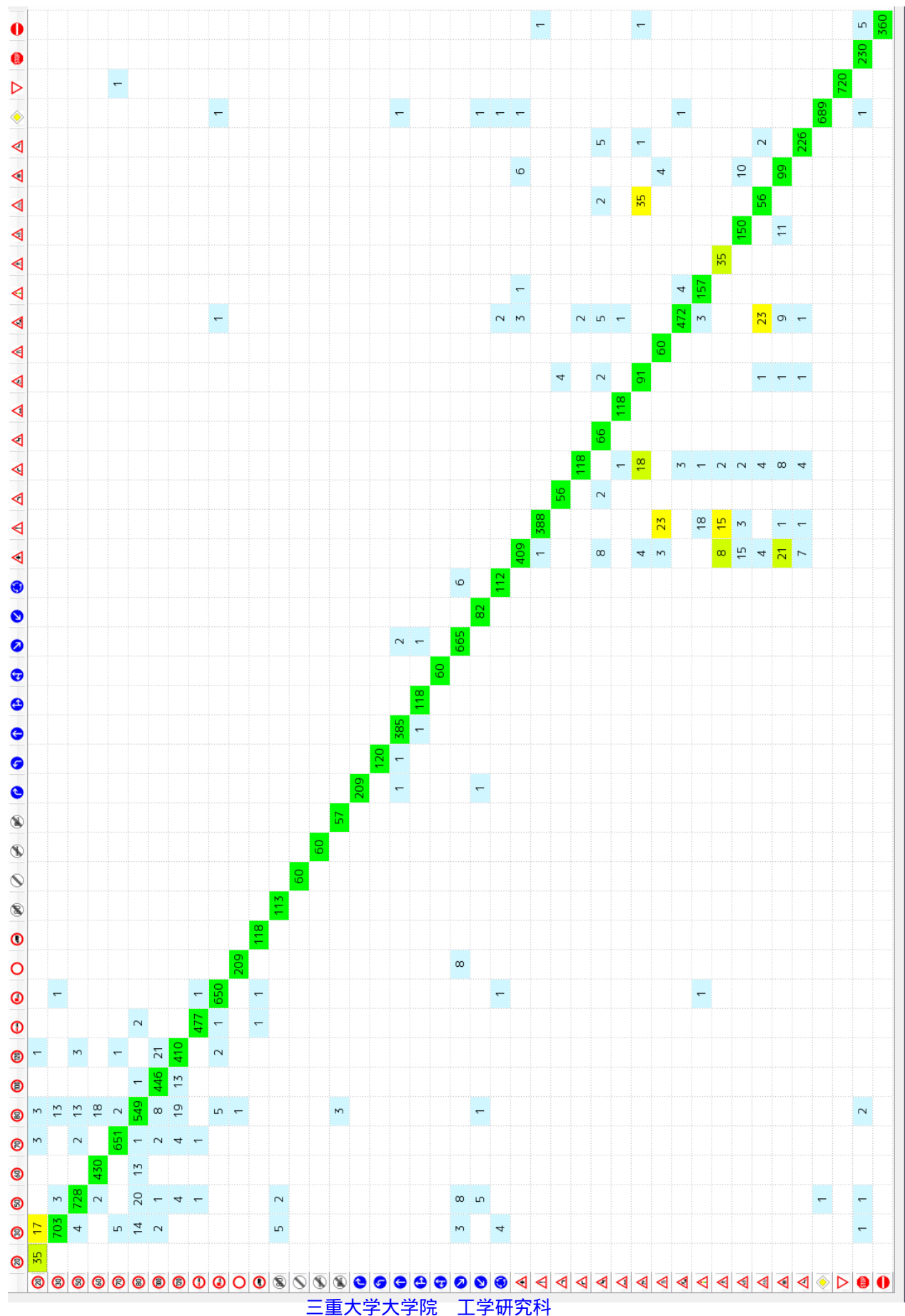


図 4.10: RGB 特徴量を用いた認識実験のクラス分類結果

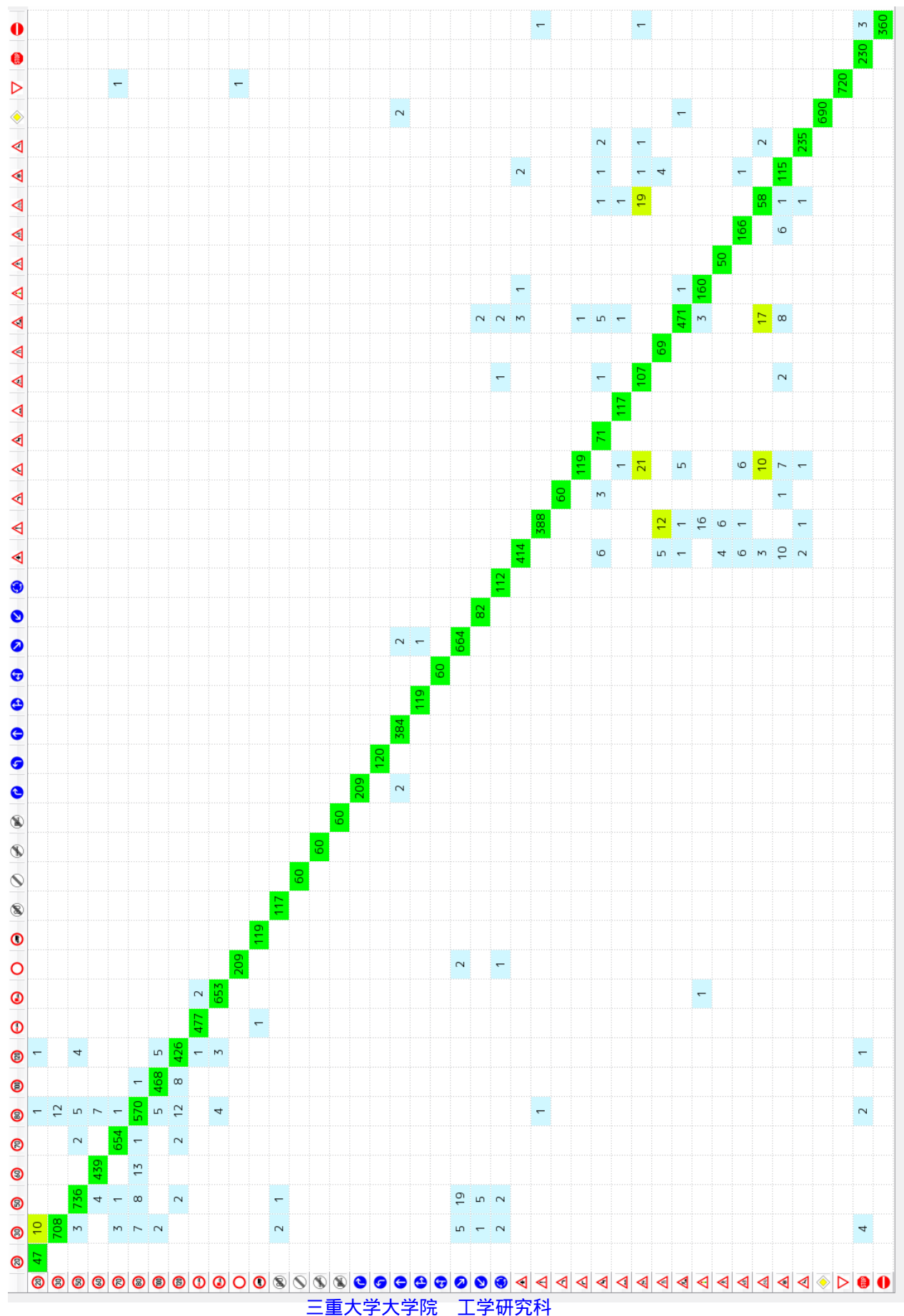
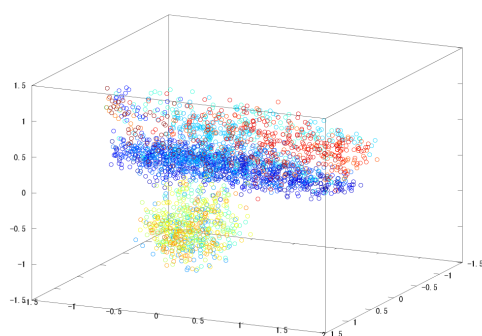
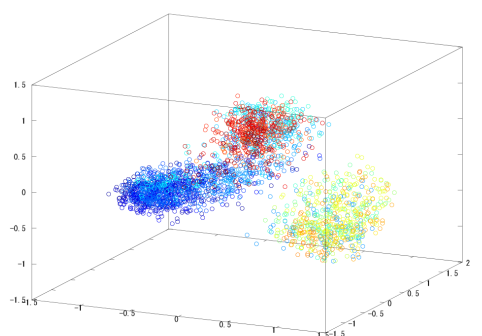
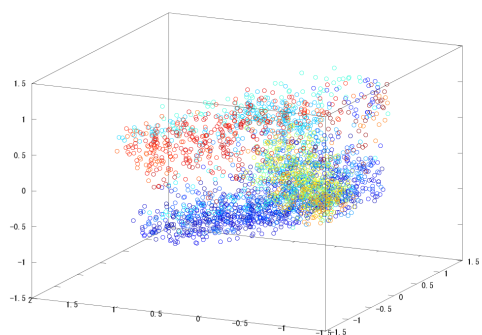
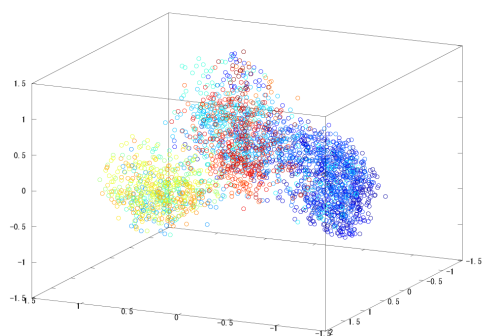
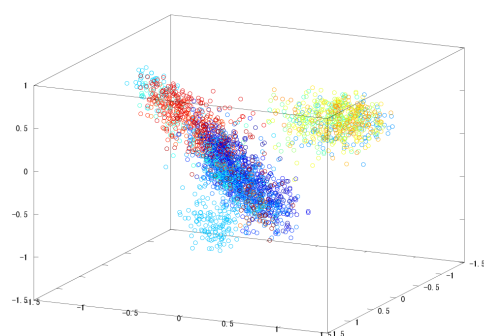
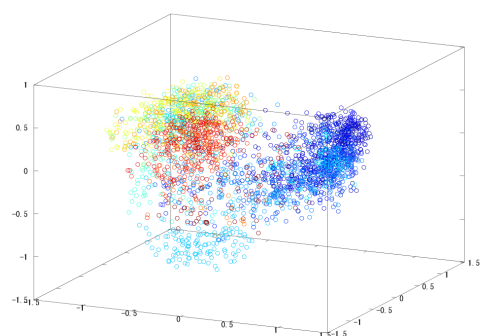
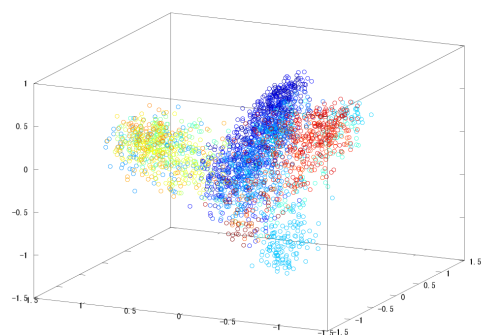
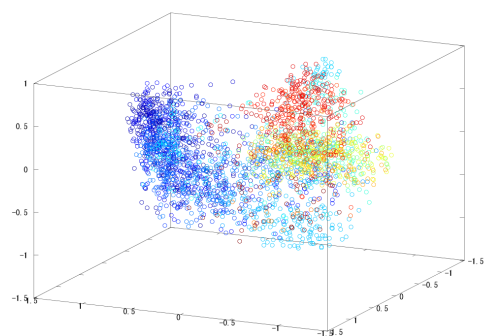


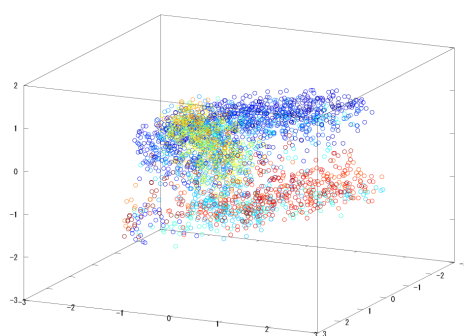
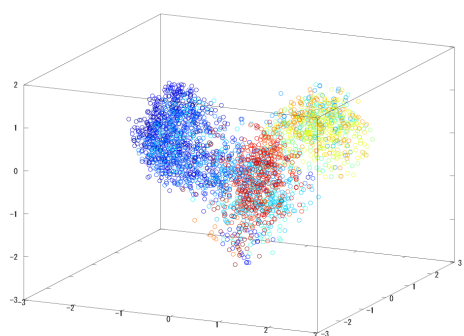
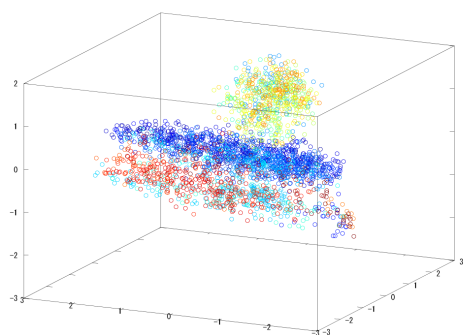
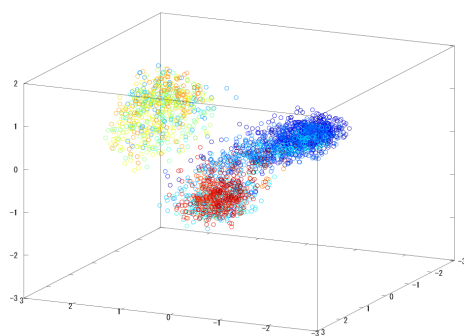
図 4.11: RGB_zero 特徴量を用いた認識実験のクラス分類結果



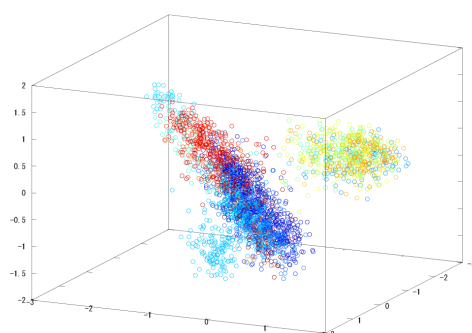
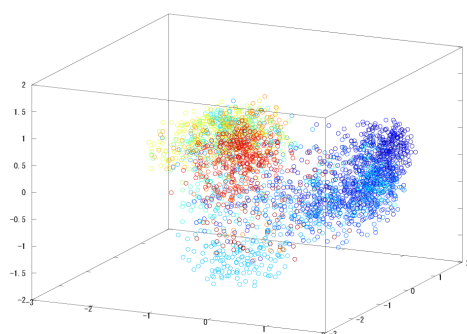
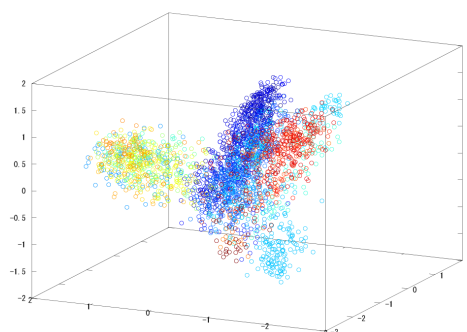
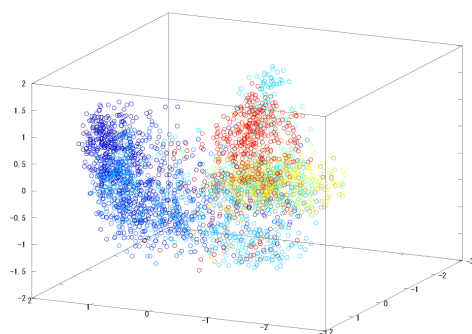
三重大学大学院 工学研究科
図 4.12: Gray 特徴量の分離性



三重大学大学院 工学研究科
図 4.13: Gray_zero 特徴量の分離性



三重大学大学院 工学研究科
図 4.14: RGB 特徴量の分離性



三重大学大学院 工学研究科
図 4.15: RGB_zero 特徴量の分離性



図 4.16: 誤認識する画像例

第 5 章

結言

5.1 まとめ

本研究では，標識認識に適した濃度こう配特徴ベクトル，SVM による標識画像の認識手法を提案し，その有効性を検証した．提案手法の特徴は，カラー画像から濃度こう配を抽出し，PCA による次元削減を行う点と，特徴量作成時，周辺小領域を 0 で置換し，標識内部の数字やピクトグラムで標識を認識する点である．

GTSRB データベースの標識画像 12,569 枚を用いた認識実験の結果，認識率は 96.69% であった．この結果は，従来のこう配抽出手法を用いたときの認識率 94.01 を上回り，提案手法の有効性が示された．

5.2 今後の課題

低解像度，ノイズ，オクルージョン，ハイライト，ぶれなどの，認識を困難にさせる要因を持つ標識を認識することが今後の課題である．

付録 A

付録

本研究に関するプログラムと画像データを以下のディレクトリに置く．

```
/net/xserve0/users/hiroki/research/main
```

ディレクトリ構造と各プログラムの概略を以下に示す．

```
main
|--65resize_Training_data/
|   # 画像サイズを幅 65 ピクセルに正規化した学習用データ
|
|--65resize_Test_data/
|   # 画像サイズを幅 65 ピクセルに正規化した評価用データ
|
|
|--feature_extraction/
|   |--pca_svm117600train
|   |   # 学習用データから特徴抽出するためのプログラム
|   |
|   |--pca_svm117600test
|   |   # 評価用データから特徴抽出するためのプログラム
|   |
|   |--pca_svm117600.sh
|   |   # 特徴抽出するためのシェルスクリプト
|
|--pca/
```

```
| #主成分分析を行う  
|  
|  
|--libsvm-3.11/  
  |--svm-train  
  | # 実験で用いる分類器を設計する .  
  |  
  |--svm-predict  
    # 分類器の性能評価を行う .
```

A.1 コンパイル方法

c++プログラムをコンパイルするときは、以下のように行う。

```
sh magickCompile [Input_File]
```

c プログラムをコンパイルするときは、以下のように行う。

```
gcc -o [Output_File] [Input_File] -lm
```

A.2 実行方法

特徴抽出は feature_extraction/において以下のコマンドで行うことができる。

```
sh pca_svm117600.sh
```

主成分分析は pca/において以下のコマンドで行うことができる

```
sh data_preparation_svm.sh
```

SVM による学習，評価は libsvm-3.11/において以下のコマンドで行うことができる。

学習

```
./svm-train -g [gamma] -c [cost] [Input_LearnData] [Output_ModelData]
```

gamma: ガンマの値

cost: コストの値

Input_LearnData: 学習用データから抽出した特徴量が記述されたファイル名

Output_ModelData:出力モデルファイル名

評価

`./svm-predict [Input_TestData] [Input_ModelData] [Output_File]`

Input_TestData:評価用データから抽出した特徴量が記述されたファイル名

Input_ModelData:./svm-train で出力されたモデルファイル名

Output_File:出力ファイル名

A.3 修論研究発表資料

平成 24 年度修論研究発表会で用いたプレゼンテーション資料を本論文の末尾に掲載する。

謝辞

本研究に関し，研究に対する基本姿勢や多くのアイデアを御教授下さった木村文隆教授，研究に対する指摘やプレゼンテーションの仕方などを御教授下さった大山航助教，進捗発表での数多くのアドバイスをいただいた若林哲史准教授，三重大大学三宅康二名誉教授に深く感謝します．また，日頃いろいろとお世話になった田中みゆき事務官，多くのアドバイスや精神的に助けていただいた研究室の先輩方，同期の皆様に感謝します．最後になりましたが，長きにわたる私の学生生活を支えてくれた両親に今一度の感謝を表して，本論文の結びといたします．

参考文献

- [1] 交通局交通企画課.“平成 24 年中の交通事故死者数について.” 警察庁.2013-01-07.
<http://www.e-stat.go.jp/SG1/estat/List.do?lid=000001105450>
- [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, Vol.104, No.740, pp.181 185, March 2005.
- [3] 竜円琢磨, 長坂保典, 鈴村宣夫. “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, Vol.99, No.609, pp.17 22, February 2000.
- [4] 澤和宏, 若林哲史, 鶴岡信治, 木村文隆, 三宅康二.“こう配特徴ベクトルと変動吸収分散行列による手書き漢字認識の高精度化.” 電子情報通信学会論文誌 D-II ,Vol.J82-D-II , No.11
- [5] David G. Lowe. “Distinctive image features from scale-invariant keypoints.” Journal of Computer Vision, 60, 2, pp.91-110,2004.
- [6] 高木雅成, 藤吉弘亘. “SIFT 特徴量を用いた交通道路標識認識.” 第 13 回画像センシングシンポジウム SSII07, LD2-06, June 2007.
- [7] C.F.Paulo and P.L.Correia. “Traffic Sign Recognition Based on Pictogram Contours.” Image Analysis for Multimedia Interactive Services, 2008, WIAMIS '08, Ninth International Workshop on, Klagenfurt, pp.67-70, May 2008.
- [8] M.Kass, A.Witkin, and D.Terzopoulos. “Snakes: Active contour models.” Int.J.Comput. Vis., vol.1,pp.321-331,1988.
- [9] F.Mokhtarian, M.Bober. “Curvature Scale Space Representation: Theory, Applications & MPEG-7 Standardization.” Kluwer Academic Publishers, Dordrecht, 2003.
- [10] Ivaro Enriquez de Luna, Carlos Miravet. “A decision support system for ship identification based on the curvature scale space representation.” SPIE Proceedings of SPIE Volume 5988, 59880K (Oct. 21, 2005).
- [11] H.Fleyeh, “Traffic Sign Recognition by Fuzzy Sets.” Intelligent Vehicles Symposium, 2008 IEEE, Eindhoven, pp.422-427, June 2008.
- [12] H.Ishida, T.Takahashi, I.Ide, Y.Mekada, and H.Murase. “Identification of degraded traf-

- fic sign symbols by a generative learning method.” Pattern Recognition, 2006, ICPR 2006, 18th International Conference on, Hong Kong, pp.531-534, August 2006.
- [13] 若林哲史, 鶴岡信治, 木村文隆, 三宅康二. “特徴量の次元数増加による手書き数字認識の高精度化.” 電子情報通信学会論文誌 D-II, Vol.J77-D-II, No.10, pp.2046-2053, 1994.
- [14] 田村秀行 (編): コンピュータ画像処理, pp.182-199. オーム社, 2002.
- [15] 田村秀行 (編): コンピュータ画像処理, pp.102-108. オーム社, 2002.
- [16] Tetsushi Wakabayashi, Shinji Tsuruoka, Fumitaka Kimura and Yasuji Miyake. “Increasing the feature size in handwritten numeral recognition to improve accuracy” Systems and Computers in Japan, Volume 26 Issue 8, pp.35-44, 1995.
- [17] 若林哲史, 鶴岡信治, 木村文隆, 三宅康二. “手書き文字認識における特徴量の次元数と変数変換に関する考察” 信学論 (D-II) , vol.J76-D-II , no.12 , pp.2495-2503 , Dec.1993 .
- [18] Vapnik, V. “The Nature of Statistical Learning Theory” Springer-Verlag, New York, 1995.
- [19] Chih-Chung Chang and Chih-Jen Lin, LIBSVM : “a library for support vector machines. ACM Transactions on Intelligent Systems and Technology” 2:27:1–27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [20] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel. “The German Traffic Sign Recognition Benchmark: A multi-class classification competition” in International Joint Conference on Neural Networks, 2011, accepted.
- [21] ruhr-universität bochum institut für neuroinformatik. “Results for IJCNN 2011 competition” German Traffic Sign Recognition Benchmark. <http://benchmark.ini.rub.de/section=gtsrb&subsection=results&subsubsection=ijcnn>