

Title	パラメトリック固有空間法における階層的な辞書作成法
Author(s)	中村, 知彦
Citation	
Issue Date	1998-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1168
Rights	
Description	Supervisor:阿部 亨, 情報科学研究科, 修士

修 士 論 文

パラメトリック固有空間法における 階層的な辞書作成法

指導教官 阿部 亨 助教授

北陸先端科学技術大学院大学
情報科学研究科情報システム学専攻

中村 知彦

1998 年 2 月 13 日

目 次

1	序論	1
1.1	研究の背景及び目的	1
1.2	本論文の構成	2
2	パラメトリック固有空間法	3
2.1	はじめに	3
2.2	パラメトリック固有空間法における辞書作成法	3
2.3	パラメトリック固有空間法における認識法	6
2.4	パラメトリック固有空間法の応用例	7
2.4.1	距離画像を用いた固有空間法	7
2.4.2	局所固有空間手法による金属物体の安定認識	7
2.4.3	物体認識のための照明位置決定法	8
2.4.4	三次元物体のイメージスポッティング	8
2.4.5	動作識別	8
2.5	まとめ	8
3	パラメトリック固有空間法における階層的な辞書作成法	10
3.1	はじめに	10
3.2	従来の画像群分類法	11
3.2.1	階層的クラスター化法	11
3.2.2	非階層的クラスター化法	11
3.2.3	判別および最小自乗基準に基づく自動しきい値選定法	12
3.3	閉曲線上での判別および最小自乗基準で最適な画像群分類法	13

3.3.1	閉曲線上での画像群分類法	13
3.3.2	閉曲線上での判別および最小自乗基準法を用いた画像群分類法 . . .	16
3.4	閉曲線上での判別および最小自乗基準で最適な分類クラス数の推定法 . . .	19
3.5	閉曲線上での判別および最小自乗基準法を用いた画像群分類法の階層化 . .	23
3.6	まとめ	24
4	実験と考察	25
4.1	はじめに	25
4.2	パラメトリック固有空間法における辞書作成	25
4.2.1	使用データ	25
4.2.2	パラメトリック固有空間法における階層的な辞書作成法の流れ . . .	27
4.3	認識実験	33
4.3.1	提案手法を用いた認識実験	33
4.3.2	全物体に対する物体認識	36
4.3.3	各 object に対する姿勢推定	36
4.3.4	各 object に対する提案手法による二分木	37
4.3.5	UPGMA 法 vs. 提案手法による二分木を用いた認識実験	38
4.3.6	提案手法による判別および最小自乗基準で最適な木 vs. 提案手法で 作成された二分木を用いた認識実験	39
4.4	まとめ	40
5	結論	50

目 次

2.1	実験装置	4
2.2	実験に使用した物体	4
2.3	固有値寄与率	5
2.4	パラメトリック固有空間を用いた物体の表現	6
3.1	物体を一軸で回転させたときの見かけの画像の変化	13
3.2	パラメトリック固有空間を用いた物体の表現	14
3.3	提案手法による学習画像の分類法	15
3.4	固有空間上の学習画像の番号付け	16
3.5	クラス分離度 η_M^*	20
3.6	二次元上の対象のバイアスの除去法	21
3.7	クラス分離度 $\bar{\eta}_M^*$	22
3.8	判別および最小自乗基準で最適な分類クラス数の評価量 $Q(M)$	23
4.1	実験に使用した物体	26
4.2	パラメトリック固有空間の Object1, Object2, Object3 の表現	27
4.3	パラメトリック固有空間の Object1, Object2, Object3 の固有値寄与率	28
4.4	パラメトリック固有空間を用いた物体 1 の表現	28
4.5	パラメトリック固有空間を用いた物体 1 の表現	29
4.6	クラス分離度 η_M^*	29
4.7	クラス分離度 $\bar{\eta}_M^*$	30
4.8	$Q(M)$	30
4.9	第一段階 (クラス数 4)	31
4.10	第二段階	32

4.11	第三段階	32
4.12	第四段階	33
4.13	視点空間の分割	34
4.14	照合回数	35
4.15	物体の姿勢の推定誤差の度数分布	35
4.16	提案手法により作成された辞書の照合回数 (object1)	36
4.17	提案手法により作成された辞書の照合回数 (object2)	37
4.18	提案手法により作成された辞書の照合回数 (object3)	38
4.19	物体の姿勢の推定誤差の度数分布 (object2)	39
4.20	照合回数 (object1)	40
4.21	照合回数 (object2)	41
4.22	照合回数 (object3)	42
4.23	二分木 (object1)	43
4.24	二分木 (object3)	44
4.25	二分木 (object2)	45
4.26	物体の姿勢の推定誤差の度数分布 (object3)	46
4.27	物体 2 の各入力画像に対する認識結果出力までの照合回数	47
4.28	物体 2 の姿勢の推定誤差の度数分布	47
4.29	本手法による二分木 vs 本手法による判別および最小自乗基準に基づく最 適木の照合回数の比較	48
4.30	物体の姿勢の推定誤差の度数分布	49

第 1 章

序論

1.1 研究の背景及び目的

二次元画像から三次元物体を識別し、その物体の姿勢を検出する技術は、工業部品の分類や一般環境内での移動物体の監視など幅広い応用があり、これまで多数の研究がなされてきた。

実際の作業環境下で認識する物体の一連の画像を撮影し、それらの画像を学習した後、入力画像と学習画像の二次元照合により三次元物体認識を行うビジュアルラーニング手法の研究がなされている。この手法は、画像信号レベルを基にした二次元照合法であるため、撮影環境の影響を受け易いという欠点はあるものの、画像から物体の特徴を抽出する必要がないため、実環境で利用しやすい。

ビジュアルラーニング手法の中でもパラメトリック固有空間法 [4] は、画像符号化を基本とした二次元照合法であり、高次元な画像データを主成分分析により低次元の特徴ベクトルで表し低次元空間で照合を行う。そのため、三次元物体を二次元画像の集合体として記憶する際に記憶量を削減することができ、記憶されたデータと入力画像との照合に要する時間も短縮できる。

しかし、パラメトリック固有空間法では辞書として、視点や光源の位置により複雑に変化するすべての物体の学習画像を記憶する必要があり、膨大な記憶量を必要とする。また、認識時に、入力画像と辞書のすべての学習画像を照合する必要があり膨大な照合時間を必要とする。そのため実用化のためには、パラメトリック固有空間法における辞書の記憶量、照合の計算量をさらに削減する必要がある。

本研究では、パラメトリック固有空間法における辞書の記憶量、照合の計算量を削減するために、パラメトリック固有空間法における階層的な辞書作成法を提案する。具体的には、視点が連続的に変化する学習画像を間引くことにより辞書の記憶量を削減し、辞書を階層的に構成することにより照合の計算量を減少させる。

本手法ではパラメトリック固有空間法により物体の画像群が固有空間中の閉曲線上にのみ投影されることに注目し、閉曲線上で判別および最小自乗基準法を用いて、固有空間上の学習画像を再帰的に分類を行う。この分類を基に辞書の木を根から葉の順に作成することにより、探索時に coarse-to-fine 戦略に基づいて認識率の低下を抑えながら効率的に照合を行うことができる。また、閉曲線上で判別および最小自乗基準法を用いて分類された学習画像は、視点が隣接しかつ固有空間上で類似している画像でのみ構成される。そのため、分類された学習画像を一つの代表画像で表すことができ、多数の学習画像を少数の代表画像で表し辞書の記憶容量を削減できる。

1.2 本論文の構成

本論文の構成は次の通りである。

第二章では、従来から提案されてきた二次元照合による三次元物体認識手法であるパラメトリック固有空間法について説明し、実用化に際しての問題点について述べる。さらに、パラメトリック固有空間法が実際にどのような問題に対して応用が検討されているかについて述べる。

第三章では、パラメトリック固有空間法における辞書の記憶量、照合の計算量を削減するために、パラメトリック固有空間法における階層的な辞書作成法を提案する。まず、データの間引き、辞書の階層化を行うためにはデータをクラスタリングを行う必要がある。従来の一般的なデータの分類法について述べ、これらをパラメトリック固有空間法に応用したときの問題点を明らかにする。次に、固有空間上の学習画像を間引くために、閉曲線上での判別および最小自乗基準で最適な画像群分類法を提案する。そして、閉曲線上での分類に判別および最小自乗基準で最適な分類クラス数の推定法、辞書の階層化法について述べる。

第四章では、提案手法によるパラメトリック固有空間法における階層的な辞書作成法を用いて認識実験を行い、この結果についての考察を行う。

第五章では、本研究の結論を述べる。

第 2 章

パラメトリック固有空間法

2.1 はじめに

実際の作業環境下で認識する物体の一連の画像を撮影し、それらの画像を学習した後、入力画像と学習画像の二次元照合により三次元物体認識を行うビジュアルラーニング手法の一つとしてパラメトリック固有空間法が提案されている。

本章では、パラメトリック固有空間法について説明し、実用化に際しての問題点について述べる。さらに、パラメトリック固有空間法が実際にどのような問題に対して応用が検討されているかについて述べる。

2.2 パラメトリック固有空間法における辞書作成法

図 2.1 に示した装置を用い、撮影された図 2.2 の R 枚の画像を学習画像とし、これを元にパラメトリック固有空間法における辞書を作成する手法について説明する。

まず、得られた二次元画像をスキャンし N 次元列ベクトル $\hat{x}_i (i = 1, \dots, R)$ とする。

画像のセンサ感度の影響を除去するため、画像データを明るさで正規化を行う。正規化後の画像ベクトルを x_i とすると、ここではベクトル x_i の大きさが 1 になるように正規化する。

$$x_i = \frac{\hat{x}_i}{\|\hat{x}_i\|} \quad (2.1)$$

この画像集合を用い行列 X を構成する。

$$X = [x_1 - c, x_2 - c, \dots, x_R - c] \quad (2.2)$$

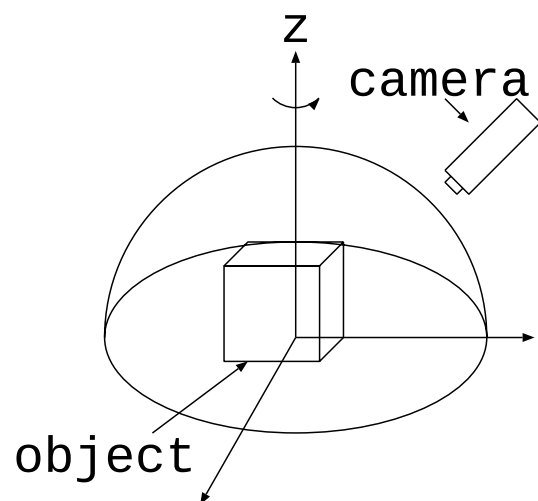


図 2.1: 実験装置



図 2.2: 実験に使用した物体

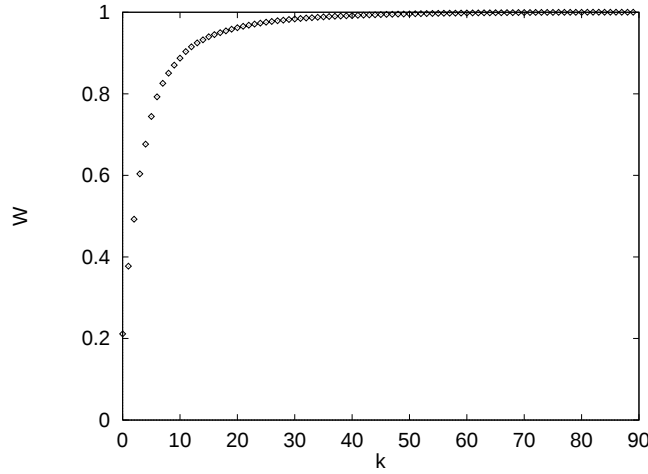


図 2.3: 固有値寄与率

ただし c は平均画像である。

次に、行列 X より、画像集合の共分散行列 $Q(N \times N)$ を構成する。

$$Q = XX^T \quad (2.3)$$

この共分散行列から、固有値 $\lambda_i (\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N)$ と固有ベクトル $e_i (i = 1, \dots, N)$ を導く。

$$\lambda_i e_i = Q e_i \quad (2.4)$$

固有ベクトルの計算手法としてさまざまな方法が考えられるが、次元 N が大きく固有ベクトルの計算が困難であるときは、特異値分解などを用いて求める手法などが提案されている [4]。それぞれの画像はこの固有値と固有ベクトルを用いて再構成することが可能であり、そのときの各固有ベクトルの重みは対応する固有値であることが知られている。ここで十分小さい固有値の項を無視することにより、学習画像の次元を落すことが可能である。有効な次数の決定には、次式の固有値寄与率 W_k と適当なしきい値 T_s を用いて決定することとする (図 2.3 は固有値寄与率 W_k の例である)。

$$W_k = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^N \lambda_i} \geq T_s \quad (2.5)$$

ここで、次元 N より十分小さい k 個の固有ベクトルにより構成された行列 $E = [e_1, e_2, \dots, e_k]$ を用い、次式より次元 N の画像ベクトル x_i を次元 k の固有空間ベクトル g_i へ投影するこ

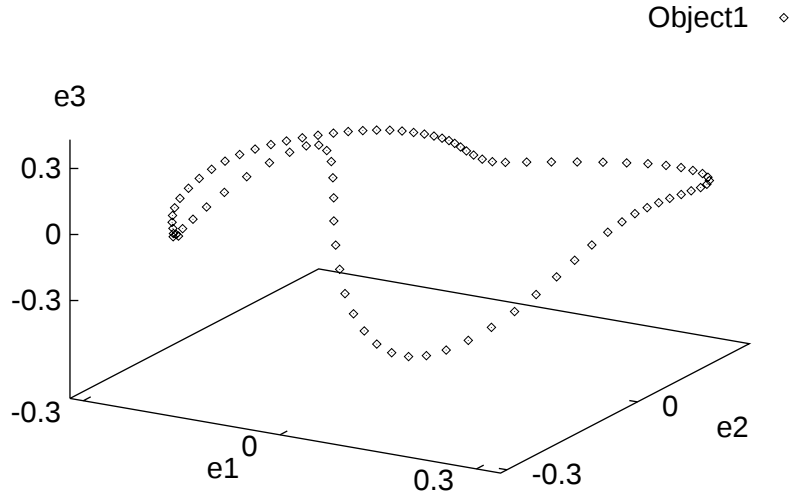


図 2.4: パラメトリック固有空間を用いた物体の表現

とが可能となる。

$$g_i = E^T(x_i - c) \quad (2.6)$$

パラメトリック固有空間法では、学習画像から固有空間上に投影されたすべての点 $g_i (i = 1, 2, \dots, R)$ を辞書として用いる。

例えば図 2.1 に示した装置を用い、視点を $0 - 360$ 度に変化させ撮影した学習画像を用いると、図 2.4 に示すように、辞書は固有空間内で一つの閉曲線を構成することになる。この閉曲線は対象物体や視点や光源の位置により複雑に変化するため、パラメトリック固有空間法ではそれぞれの対象や状況 (視点や光源の位置) に応じて異なる閉曲線を辞書として用意しておく必要がある。このため辞書の記憶容量が膨大となる。

2.3 パラメトリック固有空間法における認識法

前節で述べた辞書を用い、認識を行うためには、まず、入力画像を明るさで正規化する (そのベクトルを y とする)。次にこのベクトルを次式により全学習画像の固有空間上に投

影し、一つの点 z で表す。

$$z = E^T(y - c) \quad (2.7)$$

入力画像と学習画像の対応付けは、固有空間内の各点間のノルムで評価する。入力画像から固有空間上に投影された点 z に対応する画像 r を、辞書の点 $g_i (i = 1, \dots, R)$ の中より次式を用いて検索し、認識結果とする。例えば物体の辞書が図 2.4 の場合、入力画像から投影された点と図のすべての点を照合し、距離を最小にする点を認識結果とする。

$$d(r) = \min_i \|z - g_i\| \quad (2.8)$$

このようにパラメトリック固有空間法では、認識時に、入力画像と全辞書のすべての学習画像を照合する必要がある膨大な照合時間を必要とする。

2.4 パラメトリック固有空間法の応用例

2.4.1 距離画像を用いた固有空間法

照明条件に依存しにくい物体認識を実現するために、二次元照合による三次元物体の照合において、濃淡画像の代わりに距離画像を用いて物体の姿勢検出を行う手法が提案されている [1]。距離画像の二次元照合に固有空間を応用するにあたり、アクティブレンジファインダによる距離画像計測で生じるオクルージョンや、テクスチャによる反射光量の変化から生じる画素欠損の扱いが問題となる。欠損画素は距離情報を持たないため、固有ベクトルとの内積より固有空間へ投影することができない。そこで、距離画像を用いた固有空間照合を実現するために、画素欠損に依存しない固有空間照合を用いた姿勢検出手法が提案されている [2]。

2.4.2 局所固有空間手法による金属物体の安定認識

工場内のピンピッキング作業に代表される重なり合っている複数の金属物体を各々認識することを目標とし、物体の部分的な見え方を用いた局所固有空間手法が提案されている [9]。また効率的な局所ウィンドウを選択するために、特徴点評価、類似度評価、信頼度評価という三つの評価指標を設け、複数の隠れを各々含んだ金属物体の安定な認識が実現している。

2.4.3 物体認識のための照明位置決定法

物体の見え方は照明の方向や照明の色により大きく変化する。物体認識にとって都合の良い照明とは、対象となる物体間で画像間の類似性が少なくなるような照明である。そこで、照明の変化による見え方の変化をパラメトリック固有空間で表現することにより、この問題を効率良く解く手法が提案されている [5]。

2.4.4 三次元物体のイメージスポッティング

二次元照合により任意の方向を向いた三次元物体認識を複雑な背景画像から切り出し、同時にその方向と大きさを検出する手法が提案されている [6]。これは、パラメトリック固有空間表現を多重解像度表現した画像に階層的に適用し物体領域を切り出している。

2.4.5 動作識別

固有空間法により動作識別を行う手法が提案されている。動画像を用いた認識として、歩行動画像から人体を切り出し、その画像集合から固有空間を構成し、この空間上の軌跡で歩行動作を表現し個人の識別を行っている [7]。

また、ロボットモデルにもとづく動画像解析で人体の姿勢データ列を求め、これらから固有空間を構成し、この空間上に動作を表現して動作の識別を行う手法も提案されている [8]。

2.5 まとめ

パラメトリック固有空間法 [4] は、画像符号化を基本とした二次元照合法であり、高次元な画像データを主成分分析により低次元の特徴ベクトルで表し低次元空間で照合を行う。そのため、三次元物体を二次元画像の集合体として記憶する際に記憶量を削減することができ、記憶されたデータと入力画像との照合に要する時間も短縮できる。このため、種々の応用が研究されている。

しかし、パラメトリック固有空間法では辞書として、視点や光源の位置により複雑に変化するすべての物体の学習画像を記憶する必要があり、膨大な記憶量を必要とする。また、認識時に、入力画像と全辞書のすべての学習画像を照合する必要があり膨大な照合時

間を必要とする。そのため実用化のためには、パラメトリック固有空間法における辞書の記憶量、照合の計算量をさらに削減する必要がある。

本研究では、この問題について検討する。

第 3 章

パラメトリック固有空間法における階層的な辞書作成法

3.1 はじめに

本章では、パラメトリック固有空間法における辞書の記憶量、照合の計算量を削減するために、パラメトリック固有空間法における階層的な辞書作成法を提案する。具体的には、視点が連続的に変化する学習画像を間引くことにより辞書の記憶量を削減し、辞書を階層的に構成することにより照合の計算量を減少させる。

データの間引き、辞書の階層化を行うためにはデータのクラスタリングを行う必要がある。まず、従来の一般的なデータの分類法について述べ、これらをパラメトリック固有空間法に応用したときの問題点を明らかにする。次に、固有空間上の学習画像を間引くために、閉曲線上での判別および最小自乗基準で最適な画像群分類法を提案する。そして、閉曲線上での分類に判別および最小自乗基準で最適な分類クラス数の推定法、辞書の階層化法について述べる。

3.2 従来の画像群分類法

3.2.1 階層的クラスター化法

代表的な階層的クラスター化法として、UPGMA クラスター化法 (平均連結クラスター化法) [10] が提案されている。

UPGMA クラスター化法は、融合されつつある二つのクラスターの間類似度の平均値に基づいて、クラスターを形成するクラスター化法である。

クラスター化法の最初の段階では、各対象は別々のクラスターの中にあると見なされる。各クラスター化法の段階で、二つの最も類似したクラスターを融合し、クラスターの数一つだけ減らして行く。最終段階では、対象をすべて含む一つのクラスターを形成し、木状の辞書が作成される。木の各ノードの代表点は、そのクラスターを構成している点の重心とする。

クラスター間の距離として、平均距離 (それぞれのクラスターに属する個体間の距離の平均) を用いる。

$$D_{fg} = \frac{1}{n_f \cdot n_g} \sum_{\alpha \in C_f} \sum_{\beta \in C_g} d_{\alpha\beta} \quad (3.1)$$

$$D_{fg}^{(i+1)} = \frac{n_h}{n_g} D_{fh}^{(i)} + \frac{n_l}{n_g} D_{fl}^{(i)} \quad (3.2)$$

ただし、 D_{fg} はクラスター C_h と C_l が融合されて C_g となったときの C_g と $C_f (f \neq h, l)$ との距離である。 $n_g = n_h + n_l$ で、 n_g, n_h, n_l はそれぞれ対応する添字を持つクラスターに属す個体の数。

対象の識別では、未知の対象と、各クラスターの代表 (重心) の間のユークリッド距離を計算し、最も類似したクラスターを探索し、これを繰り返して木の辞書を辿り、結果を検出する。

UPGMA クラスター化法では葉からクラスターを融合して木を作成するため、探索時に木の辞書を間違えた経路で辿る場合がある。

3.2.2 非階層的クラスター化法

代表的な非階層的クラスター化法として K-means 法が提案されている。

K-means 法の標準的なアルゴリズムは、 K を類 (クラスター) の数として、

- X において K 個の点を類の初期代表点とする。
- 全ての点を 1-MN 方式で識別することにより類別する。
- 各類の平均値 (重心) を求めて類の更新代表点とする。
- 手続き 2, 3 を収束するまで繰り返して類別を更新する。

である。

K-means 法ではクラスタリングの結果が初期値に依存するという欠点がある。

3.2.3 判別および最小自乗基準に基づく自動しきい値選定法

しきい値で対象を複数のクラスに分類する手法として、判別および最小自乗基準に基づく自動しきい値選定法が提案されている [3]。自動しきい値選定法では、分離されるクラスの濃度レベルでの分離度を最大とする判別基準の立場から、濃淡画像を二値化し対象領域と背景に分離するしきい値を選定する。この方法は同時に原濃淡画像の最小自乗近似の意味でも最適なしきい値選定法となっている。

自動しきい値選定法は判別および最小自乗基準で最適性が保証されており、対象を分離する何らかの特徴のヒストグラムが与えられた分類の問題に対して一般的に適用でき、簡単に標準的な自動しきい値選定法として広く応用できる。

自動しきい値選定法は対象の一つの特性値に基づいて分離する手法であるが、対象の特性値が多数の場合へも応用できる。判別および最小自乗基準に基づく自動しきい値選定法を用いて多次元の対象を分離する場合、各次元ごとに対象を分離する。そのため、対象を分類するクラス数の論理的な組合せは、対象の次元数が $N (\geq 2)$ の場合、 M_i^* を各 i 次元における対象の分類クラス数とすると、計 $M_T^* = \prod_{i=1}^N M_i^*$ 種類となる。

このとき、対象の次元数が少ない場合は対象を少数のクラスに分類できるが、対象の次元数が多い場合は対象を分類するクラス数が爆発してしまう可能性がある。パラメトリック固有空間法では物体の画像を 15 ~ 30 次元程度で近似するため、自動しきい値選定法をパラメトリック固有空間法にそのまま適用すると対象を分類するクラス数は爆発してしまう。

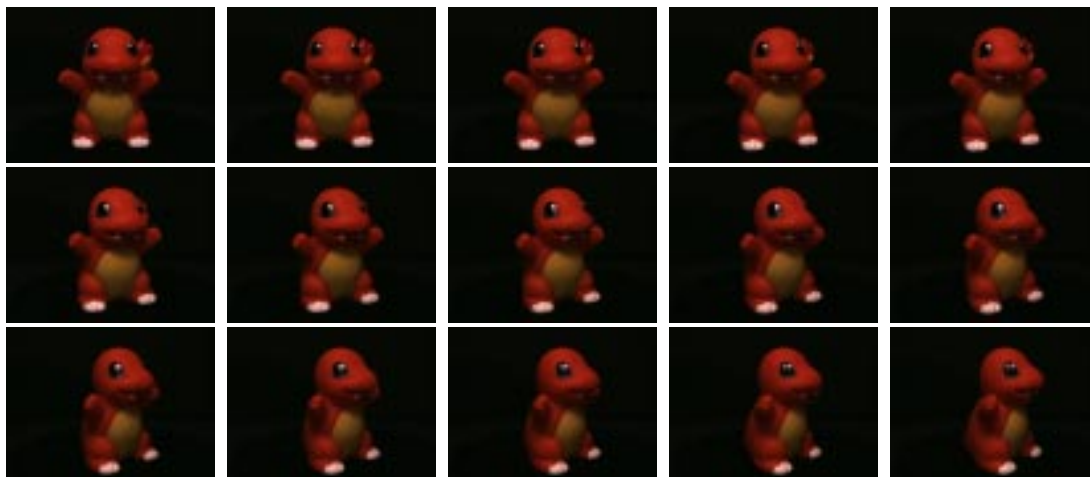


図 3.1: 物体を一軸で回転させたときの見かけの画像の変化

3.3 閉曲線上での判別および最小自乗基準で最適な画像群分類法

3.3.1 閉曲線上での画像群分類法

図 2.1 に示した装置を用い、視点を $0 - 360$ 度に変化させ 4 度ごとに撮影した図 3.1 のような人形の一連の学習画像を用いると、パラメトリック固有空間法における辞書は図 3.2 に示すように、物体の回転に伴い固有空間上で一つの閉曲線を構成することになる (ただし、表示のため三次元表示)。

パラメトリック固有空間法は、視点が隣接した画像は極めて相関が高いという性質を利用して、画像を圧縮する手法である。図 3.1 のような人形の一連の学習画像を用いると、パラメトリック固有空間法により、 $2304(80 \times 60 \times 3)$ 次元の画像データが 95.4% 程度の固有値寄与率で 19 次元で近似できる。また、2 枚の画像 x_i, x_j の類似度は固有空間上のそれらの点 g_i, g_j の類似度に近似でき、

$$\|x_i - x_j\|^2 \approx \|g_i - g_j\|^2 \quad (3.3)$$

図 3.1 のように視点が隣り合った 2 枚の画像は極めて相関が高いため、それらから固有空間上に投影された点の距離は近くなる。このため、視点が連続的に変化する画像は 3.2 に

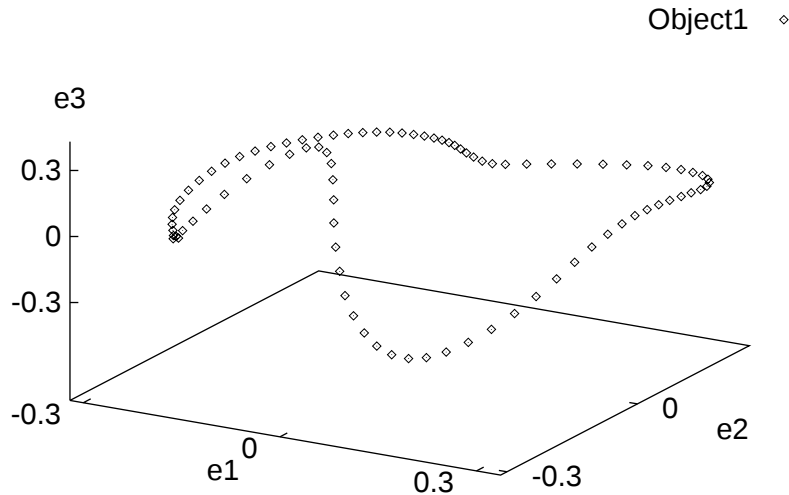


図 3.2: パラメトリック固有空間を用いた物体の表現

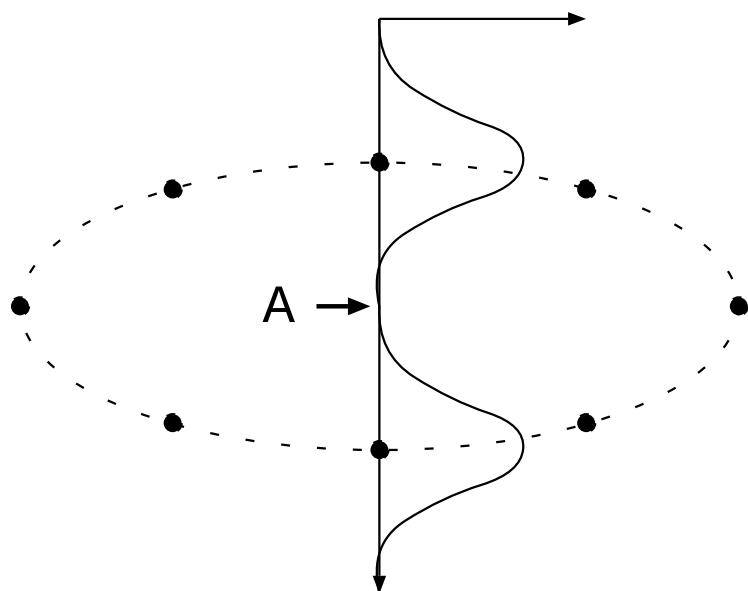
示すように、物体の回転に伴い固有空間上で滑らかな点列を描く [4]。図 3.3(a) のように認識対象となる物体の画像は閉曲線上の近辺に分布するはずである。

そのため、視点が連続的に変化する物体の学習画像から固有空間上に投影された点を分類する場合、閉曲線の内部の図 3.3(a) の A の箇所に境界がないといけない。

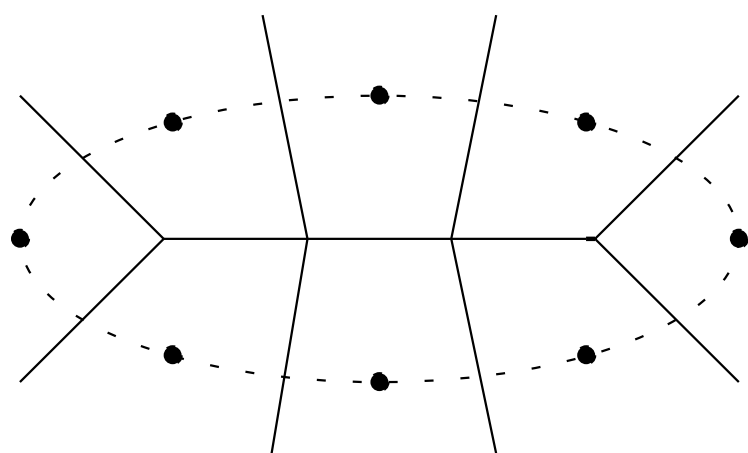
学習画像の分類に UPGMA クラスタ化法などを用いると、閉曲線の内部の図 3.3(a) の A の箇所に境界があるという保証はない。

本研究では、パラメトリック固有空間法により認識対象となる物体の画像が固有空間中の閉曲線上の近辺に分布するということに注目し、図 3.3(b) のように、学習画像を閉曲線上で分類する手法を提案する。学習画像を閉曲線上で分類することにより、分類された学習画像は視点が隣接した画像で構成され、学習画像に存在しない画像をも分類された画像群にある程度含ませることができる。さらに、クラス分離度が良くなるように分類することにより、分類された学習画像は視点が隣接しかつ固有空間上で類似している画像でのみ構成される。それにより、分類された学習画像を一つの代表画像で表すことができ、多数の学習画像を少数の代表画像で表し辞書の記憶容量を削減できる。

本研究では、閉曲線上でクラス分離度が良くなるように学習画像を分類するのに、3.2.3



(a) 認識対象となる物体の画像は閉曲線上の近辺に分布



(b) 学習画像を閉曲線上で分類

図 3.3: 提案手法による学習画像の分類法

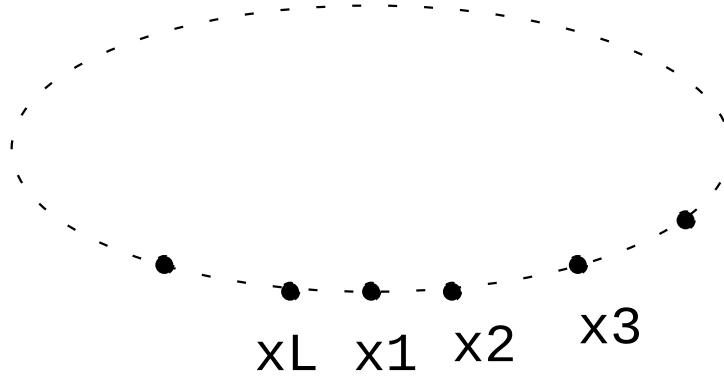


図 3.4: 固有空間上の学習画像の番号付け

で述べた手法を用いる。

3.3.2 閉曲線上での判別および最小自乗基準法を用いた画像群分類法

本節では、クラス分離度が良くなるように分類することにより、分類された学習画像を視点が隣接しかつ固有空間上で類似している画像でのみ構成する。

本研究では、閉曲線上でクラス分離度が良くなるように辞書の学習画像を分類する手法として、閉曲線上での判別および最小自乗基準法を用いた画像群分類法を提案する。

準備

パラメトリック固有空間法により投影された L 枚の学習画像を、視点の連続性に沿って図 3.4 のように番号付けを行い、 $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_L$ で表す。各学習画像は、他に同じ画像が存在しないため生起確率は全て $p = \frac{1}{L}$ となる。全学習画像の全平均、全分散はそれぞれ次式で与えられる。

$$\vec{\mu}_T = \sum_{i=1}^L \vec{x}_i p \quad (3.4)$$

$$\sigma_T^2 = \sum_{i=1}^L (\vec{x}_i - \vec{\mu}_T)^2 p \quad (3.5)$$

この全学習画像を、 $\vec{x}_L, \vec{x}_1$ 間の分類箇所と $M - 1$ 個の分類箇所 $k_i (i = 1, 2, \dots, M - 1)$

$$1 \leq k_1 < k_2 < \dots < k_{M-1} < L \quad (3.6)$$

により、 M 個のクラス

$$S_i = [k_{i-1} + 1, k_i] \quad (i = 1, \dots, M) \quad (3.7)$$

に分類するとする。ただし、 $k_0 = 0, k_M = L$ とする。このとき、各クラス i の生起確率 w_i 、各クラスの平均画像 $\vec{\mu}_i$ は、

$$w_i = \sum_{j \in S_i} p \quad (3.8)$$

$$\vec{\mu}_i = \sum_{j \in S_i} \frac{\vec{x}_j p}{w_i} \quad (3.9)$$

で与えられ、分類する箇所 $k_i (i = 1, 2, \dots, M)$ によらず常に次式が成り立つ。

$$\sum_{i=1}^M w_i \vec{\mu}_i = \vec{\mu}_T \quad (3.10)$$

また、各クラス i の分散は次式で与えられる。

$$\sigma_i^2 = \sum_{j \in S_i} \frac{(\vec{x}_j - \vec{\mu}_i)^2 p}{w_i} \quad (3.11)$$

判別基準

分類する箇所 $k_i (i = 1, 2, \dots, M-1)$ の良さを評価するために、判別分析で用いられる基準 (クラス分離度) η を導入する。

$$\eta(k_1, k_2, \dots, k_{M-1}) = \frac{\sigma_B^2(k_1, k_2, \dots, k_{M-1})}{\sigma_T^2} \quad (3.12)$$

ここで、

$$\sigma_W^2 = \sum_{i=1}^M w_i \sigma_i^2 \quad (3.13)$$

$$\begin{aligned} \sigma_B^2(k_1, \dots, k_M) &= \sum_{i=1}^M w_i (\vec{\mu}_i - \vec{\mu}_T)^2 \\ &= \sum_{i=1}^M w_i \vec{\mu}_i^2 - \mu_T^2 \end{aligned} \quad (3.14)$$

はそれぞれ、クラス内分散、クラス間分散である。クラス内分散、クラス間分散、および全分散の間には、次の基本的関係が $k_i (i = 1, 2, \dots, M-1)$ によらず常に成り立つ。

$$\sigma_W^2 + \sigma_B^2 = \sigma_T^2 \quad (3.15)$$

分類する箇所 $k_i (i = 1, 2, \dots, M-1)$ のクラス分離度の良さの評価として $\eta(k_1, k_2, \dots, k_{M-1})$ を最大、つまり σ_B^2 を最大とする $k_i^* (i = 1, 2, \dots, M-1)$ を判別基準で最適な分類箇所と

する。

$$\sigma_B^2(k_1^*, k_2^*, \dots, k_{M-1}^*) = \max_{1 \leq k_1 < k_2 < \dots < k_{M-1} < L} \sigma_B^2(k_1, k_2, \dots, k_{M-1}) \quad (3.16)$$

$\vec{x}_L, \vec{x}_1$ 間の分類箇所は固定していたので、 x_1 の位置がすべての L 枚の学習画像となる場合に対し計算し、判別基準で最適な分類箇所を決定する。

最小自乗基準

ここでは、得られる代表画像と元の学習画像の平均自乗誤差を考え、それを最小とする基準の下で最適分類箇所の組 $k_i (i = 1, 2, \dots, M - 1)$ を選ぶ場合について考える。

画像群を 2 つの代表画像 $\vec{a}_1, \vec{a}_2$ で近似する場合、平均自乗誤差は次式で与えられる。ただし、 p は画像の存在確率である。

$$e^2(\vec{a}_1, \vec{a}_2; k_1, k_2) = \sum_{i \in S_1} (\vec{x}_i - \vec{a}_1)^2 p + \sum_{i \in S_2} (\vec{x}_i - \vec{a}_2)^2 p \quad (3.17)$$

これをパラメータ $\vec{a}_1, \vec{a}_2, k_1, k_2$ に関して最小とすることにより、最小自乗しきい値選定法での最適しきい値が得られる。最適な $\vec{a}_i (i = 1, 2)$ は、

$$\forall j, \quad \frac{\partial e^2}{\partial a_{ij}} = 0 \quad (3.18)$$

より、

$$\vec{a}_i = \vec{\mu}_i \quad (3.19)$$

となる。つまり任意に固定された分類箇所 k_1, k_2 に対し、2 つの代表画像 $\vec{a}_i$ に関する最小値は、 $\vec{a}_i$ が各クラスの平均ベクトル μ_i のときに達成される。これらから、

$$\begin{aligned} \varepsilon^2(k_1, k_2) &= \min_{\vec{a}_1, \vec{a}_2} e^2(\vec{a}_1, \vec{a}_2; k_1, k_2) \\ &= \sum_{i \in S_1} (\vec{x}_i - \vec{\mu}_1)^2 p + \sum_{i \in S_2} (\vec{x}_i - \vec{\mu}_2)^2 p \\ &= \sigma_W^2 \\ &= \sigma_T^2 - \sigma_B^2 \end{aligned} \quad (3.20)$$

となる。つまり、ある分類箇所 k_1, k_2 で画像群を 2 つの代表画像で最良近似したときの最小自乗誤差 $\varepsilon^2(k_1, k_2)$ は、そのときのクラス内分散 σ_W^2 に等しく、 $\varepsilon^2(k_1, k_2)$ を最小とすることは η を最大とすることと同値である。

最小自乗近似の意味でも最適分類箇所の組 k_1, k_2 は、 σ_B^2 を最大とするものとして求められる。ここでは分類クラス数が 2 つの場合を考えたが分類クラス数が 2 つ以上の場合へもそのまま拡張できる。

分類クラス数が多い場合の対処

判別および最小自乗基準で L 枚の学習画像を M 個のクラスに分類する場合に最適な分類箇所 $k_i (i = 1, 2, \dots, M)$ を決定するためには、 σ_B^2 を最大とする M 個の分類箇所を探索すればよいわけだが、その組合せは $\binom{L-1}{M-1} \times L$ 通りあり、 M が大きくなるにつれ、組合せの数は指数的に増加し、計算量の点で総当り法はほとんど不可能になる。

そこで分類クラス数が多い場合は、動的計画法 (DP) により対処する。式 (3.12) の定数項 μ_T^2 を省略すると、最大化したい目的関数は

$$P_M(k_1, \dots, k_{M-1}) = \sum_{i=1}^M w_i \vec{\mu}_i^2 \quad (3.21)$$

と書ける。

$$P_M^*(L) = \max_{1 \leq k_j < L} P_M(k_1, \dots, k_{M-1}) \quad (3.22)$$

と定義すると、最適性の原理から、

$$P_M^*(L) = \max_{M-1 \leq k_{M-1} < L} [P_{M-1}^*(k_{M-1}) + w_M \vec{\mu}_M^2] \quad (3.23)$$

と定式化でき、動的計画法の解法に帰着される。この場合、計算の手間は $O(ML^2)$ になる。そして、 x_L, x_1 間の分類箇所は固定していたので、 x_1 の位置がすべての L 枚の学習画像となる場合に対し計算し、判別および最小自乗基準で最適な分類箇所を決定する。その結果、計算の手間は $O(ML^3)$ になる。

動的計画法により、最適分類箇所を決定する計算の手間が $O(\binom{L-1}{M-1} \times L)$ から $O(ML^3)$ に減少することが可能となり、分類クラス数が多い場合でも十分実用的になる。また、動的計画法による解法では、 M 種類のクラスに分類する場合、それ以下の任意の分類クラス数の最適分類箇所の結果が同時に得られる。

このように、学習画像が与えられると、分類クラス数 $M (M = 2, 3, \dots, L)$ それぞれの判別および最小自乗基準で最適な分類箇所 $k_1, \dots, k_M$ が決定できる。

3.4 閉曲線上での判別および最小自乗基準で最適な分類クラス数の推定法

本節では、 L 枚の学習画像を分類するクラス数 $M = 2, \dots, L$ から判別および最小自乗基準で最適な分類クラス数 M^* を推定する手法について述べる。

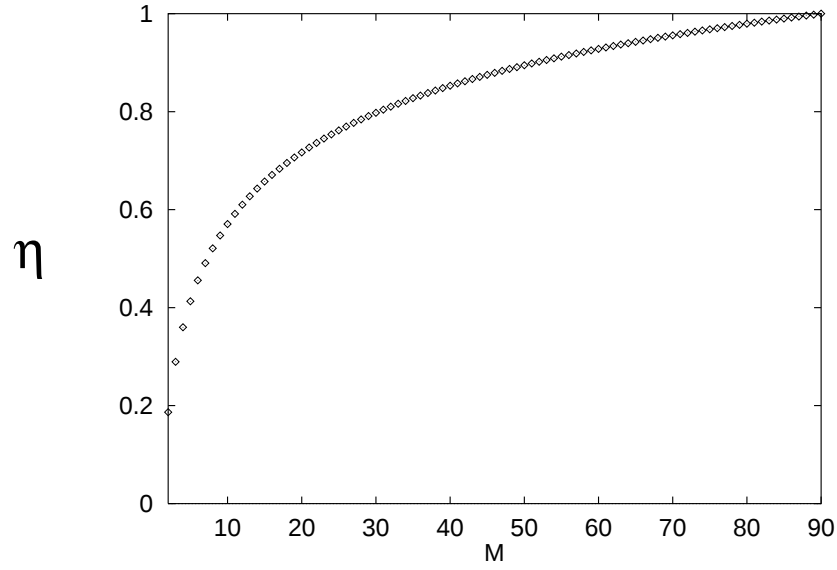


図 3.5: クラス分離度 η_M^*

分類クラス数 M を大きくするにつれ、 M 個の代表画像で最良近似したときの最小自乗誤差 σ_W^2 は常に単調に減少するので、 $\sigma_B^2 = \sigma_T^2 - \sigma_W^2$ 、従って、 $\eta_M^* = \sigma_B^2 / \sigma_T^2$ は、 M の単調増加関数となる (図 3.5 は図 3.1 の学習画像の各分類クラス数 M に対する最大の η_M^* を表す)。

$$\eta_2^* < \eta_3^* < \cdots < \eta_L^* = 1 \quad (3.24)$$

$M = L$ なら $\sigma_W^2 = 0$ であるから、 $\eta_L^* = 1$ となる。このため、比較のための意味ある測度を得るには、 η_M^* の持つこのようなバイアスを除去する必要がある。

大津 [3] は、 L 階調からなる濃淡画像を M 階調の濃淡画像で表現する際、しきい値選定法において、 η_M^* の持つ自然なバイアスを知るために、 L 階調が一次元上に一様に等間隔に分布する場合を用いている (画素値が一様に分布する場合なので生起確率 $p = 1/L$)。 L 階調を M 個に等間隔に分類するとき、全分散、クラス間分散は次式で与えられる。

$$\bar{\sigma}_T^2 = \frac{1}{12}(L^2 - 1) \quad (3.25)$$

$$\bar{\sigma}_B^2 = \bar{\sigma}_T^2 - \frac{1}{12} \left(\left(\frac{L}{M} \right)^2 - 1 \right) \quad (3.26)$$

$$\bar{\eta}_M^* = 1 - \frac{\left(\frac{L}{M} \right)^2 - 1}{L^2 - 1} \quad (3.27)$$

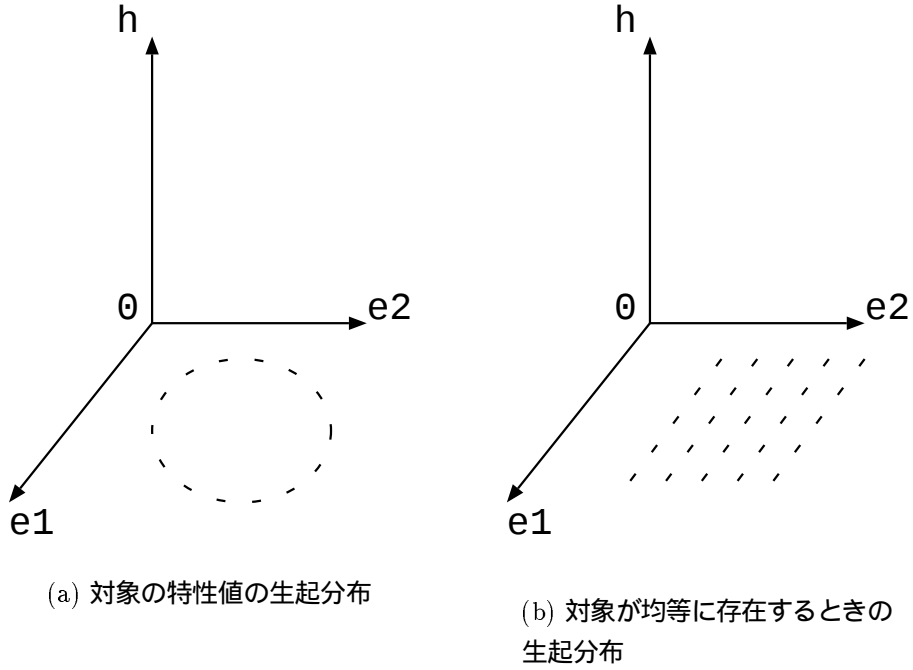


図 3.6: 二次元上の対象のバイアスの除去法

ここで、大津は実際的観点から $M \ll L$ を仮定し、分離度 $\bar{\eta}_M^*$ を次式で評価した。

$$\bar{\eta}_M^* = 1 - \frac{1}{M^2} \quad (3.28)$$

本手法では、この分離度 $\bar{\eta}_M^*$ を多次元に拡張し、それを用いて η_M^* の持つ自然なバイアスを除去する。このとき L 個の学習画像が k 次元の固有空間上に均等に存在する場合は図 3.6 のようになる。(ただし、図 3.6 では表示のため $k = 2$ の場合を示す。横軸が対象の特性値、縦軸がその特性値の生起確率を示す。(a) が対象の特性値の生起分布を示し、(b) が対象が均等に存在するときの生起分布を示す。) L 枚の学習画像をクラス数 M で等間隔に分類するとき、画素値が一様に分布する場合なので $p = 1/L$ 、

$$\sigma_T^2 = \frac{1}{12} (L^{\frac{2}{k}} - 1) k \quad (3.29)$$

$$\sigma_B^2 = \sigma_T^2 - \frac{1}{12} \left(\left(\frac{L}{M} \right)^{\frac{2}{k}} - 1 \right) k \quad (3.30)$$

この方法では k 次元での分離度 $\bar{\eta}_M^*$ は次式で表される。

$$\bar{\eta}_M^* = 1 - \frac{\left(\frac{L}{M} \right)^{\frac{2}{k}} - 1}{L^{\frac{2}{k}} - 1} \quad (3.31)$$

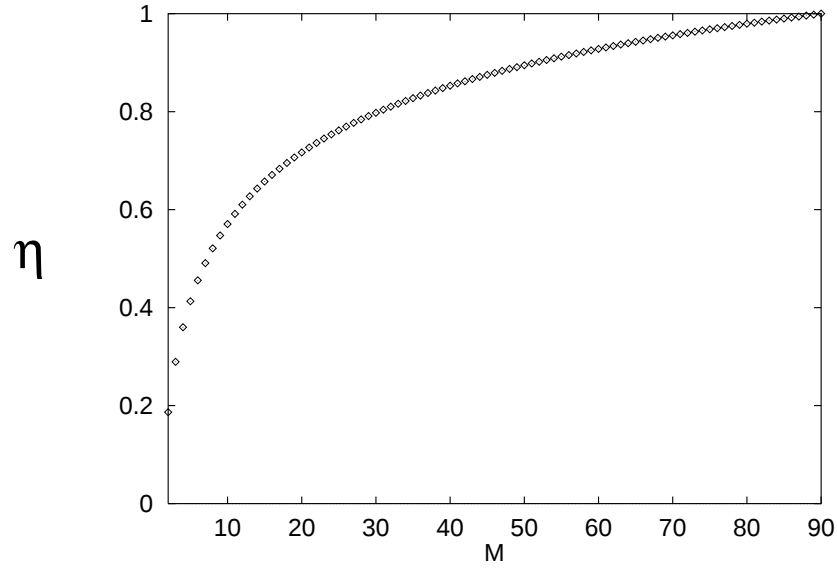


図 3.7: クラス分離度 $\bar{\eta}_M^*$

判別および最小自乗基準に基づく自動しきい値選定法と同様に $M \ll L$ を仮定すると、式 $\bar{\eta}_M^*$ の中から L の項が無くなる。しかし、固有空間上での閉曲線の構成画像により得られる η_M^* は、 L に依存する。また、学習画像の判別および最小自乗基準で最適な分類クラス数 M は $M = 2, 3, \dots, L$ のあらゆる分類クラス数を取り得るため、 $M \ll L$ という仮定は必ずしも成り立たない。

そこで、ここでは $M \ll L$ を仮定せず $\bar{\eta}_M^*$ として式 (3.26) を用いる。この $\bar{\eta}_M^*$ に対して次式が成り立つ (図 3.7 は、図 3.5 のバイアスを除去するための式 (3.26) による $\bar{\eta}$ 。ただし $L = 90, k = 19$)。

$$\bar{\eta}_2^* < \bar{\eta}_3^* < \dots < \bar{\eta}_L^* = 1 \quad (3.32)$$

本研究では、 η_M^* と各学習画像が均等に存在する場合の $\bar{\eta}_M^*$ を比較することにより、 η_M^* の持つ自然なバイアスを除去する。判別および最小自乗基準で最適な分類クラス数の評価量として次式を用いる。

$$Q(M) = \frac{\eta_M^*}{\bar{\eta}_M^*} \quad (3.33)$$

図 3.8 は、図 3.7 の η_M^* と図 3.7 の $\bar{\eta}_M^*$ を用いたときの式 (3.28) の $Q(M)$ を示す。これにより、

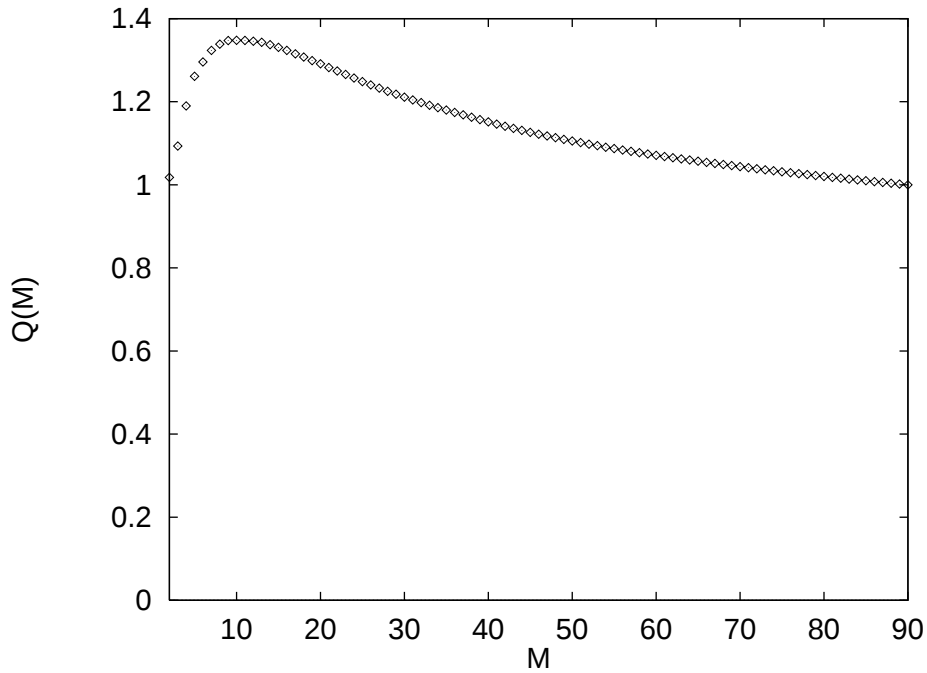


図 3.8: 判別および最小自乗基準で最適な分類クラス数の評価量 $Q(M)$

判別および最小自乗基準で最適な分類クラス数 M^* を選定する。

$$Q(M^*) = \max_{2 \leq M \leq L} Q(M) \quad (3.34)$$

図 3.8 では判別および最小自乗基準で最適な分類クラス数 M^* は $M^* = 10$ となる。

このように、学習画像が与えられると、閉曲線上での学習画像の判別および最小自乗基準で最適な分類クラス数を推定でき結果としてその分類箇所が決定できる。

3.5 閉曲線上での判別および最小自乗基準法を用いた画像群分類法の階層化

本節では、分類された学習画像集合に対しても、閉曲線上で再帰的に分類を行うことにより階層的な辞書を作成する。これにより照合の計算量を削減する。

パラメトリック固有空間法により認識対象となる物体の画像が固有空間中の閉曲線上の近辺に分布するということに注目し、閉曲線上での判別および最小自乗基準法を用いて、

学習画像から固有空間上に投影された点を再帰的に分類を行う。辞書の木を根から葉の順に作成することにより、探索時に coarse-to-fine 戦略に基づいて認識率の低下を抑えながら効率的に照合を行うことができる。

また、閉曲線上で判別基準法を用いて分類された学習画像は、視点が隣接しかつ固有空間上で類似している画像でのみ構成される。視点が少し移動したときに固有空間上の変化(距離)が小さい画像は、視点に対して安定であると考えられる。そのため、分類された学習画像集合を一つの代表画像で表すことができ、多数の学習画像を少数の代表画像で表し辞書の記憶容量を削減できる。

3.6 まとめ

パラメトリック固有空間法では辞書として、視点や光源の位置により複雑に変化するすべての物体の学習画像を記憶する必要があり、膨大な記憶量を必要とする。また、認識時に、入力画像と辞書のすべての学習画像を照合する必要があり計算に膨大な時間を必要とする。そのため実用化のためには、パラメトリック固有空間法における辞書の記憶量、照合の計算量をさらに削減する必要がある。

本研究では、パラメトリック固有空間法における辞書の記憶量、照合の計算量を削減するために、パラメトリック固有空間法における階層的な辞書作成法を提案する。具体的には、視点が連続的に変化する学習画像を間引くことにより辞書の記憶量を削減し、辞書を階層的に構成することにより照合の計算量を減少させる。

まず、データの間引き、辞書の階層化を行うためにはデータのクラスタリングを行う必要がある。従来一般的なデータの分類法について述べ、これらをパラメトリック固有空間法に応用したときの問題点を明らかにした。次に、固有空間上の学習画像を間引くために、閉曲線上での判別および最小自乗基準で最適な画像群分類法を提案した。そして、閉曲線上での分類に判別および最小自乗基準で最適な分類クラス数の推定法、辞書の階層化法について述べた。

第 4 章

実験と考察

4.1 はじめに

本論文で提案するパラメトリック固有空間法における階層的な辞書作成法を用いて認識実験を行い、この結果についての考察を行う。

まず、パラメトリック固有空間法における辞書作成について述べ、パラメトリック固有空間法における階層的な辞書作成法の流れを示す。実験の有効性を示すために、比較実験 (UPGMA 法 vs. 本手法による二分木、本手法による判別および最小自乗基準で最適な木 vs. 本手法による二分木) を行う。

4.2 パラメトリック固有空間法における辞書作成

4.2.1 使用データ

実験では図 4.1 に示す 3 種類の物体を用いた。まず、図 2.1 に示した装置を用い、object1,2,3 の物体を視点を $0 - 360$ 度に変化させ 4 度ごとに撮影した図 4.1 のような一連の学習画像 (80×60 画素) を作り、行列 X を作成する。学習行列は、共分散行列の主要な固有値と固有ベクトルを計算することにより、固有空間上の点に投影される。95 % 程度 ($W_k = 0.95$) の固有値寄与率を得るためには、固有空間として図に示すように、66 次元程度必要であった。つまり、 $(80 \times 60 \times 3) \times 3$ 次元が 66 次元に圧縮される。図 4.2 は各物体の学習画像から固有空間上に投影された点であり、合計 3 つの閉曲線を構成している (た



(a) object1



(b) object2



(c) object3

図 4.1: 実験に使用した物体

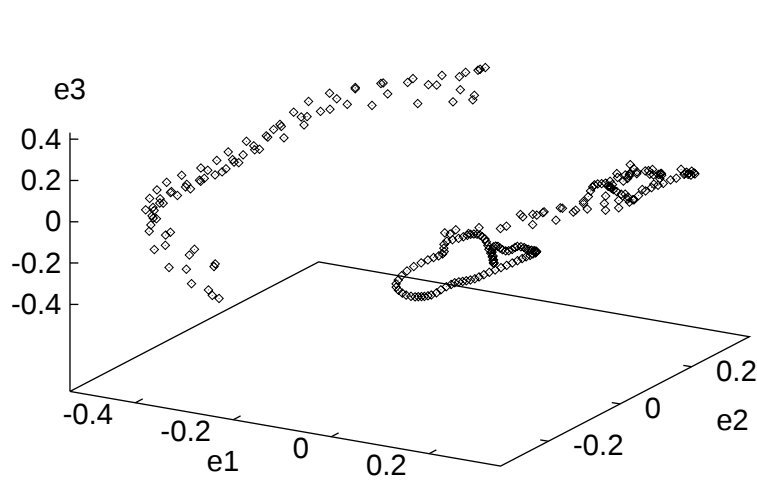


図 4.2: パラメトリック固有空間の Object1, Object2, Object3 の表現

だし表示のため三次元表示)。

4.2.2 パラメトリック固有空間法における階層的な辞書作成法の流れ

object1 を視点を $0 - 360$ 度に変化させ 4 度ごとに撮影した図 4.1(a) のような人形の一連の学習画像を、パラメトリック固有空間法で 2 次元 (固有寄与率 37.7%) に圧縮し、学習画像から固有空間上に投影された点を図 4.5 に示す。

この例を用い階層的な辞書作成の様子を説明する。まず、図 4.1(a) の学習画像から固有空間上に投影された図 4.5 の点集合を、閉曲線上で判別および最小自乗基準を用いて分類する。図 4.6 は、図 4.5 の各分類クラス数 M に対する最大の分離度 η_M^* を表す。 η_M^* のバイアスを除去するために、学習画像から投影された点集合が 2 次元上に均等に存在し (画像枚数 $L=90$)、それを均等な M クラスに分類する場合の分離度 $\bar{\eta}_M^*$ を用いる (図 4.7)。

図 4.6 の η_M^* 、図 4.7 の $\bar{\eta}_M^*$ から、式 (3.28) の $Q(M)$ を求めると図 4.8 のようになる。図 4.8 の $Q(M)$ より、第一段階では、学習画像集合を 4 つのクラスに分類する。そのとき図 4.5 の点集合は図 4.9 の代表画像 (重心) で近似できる。

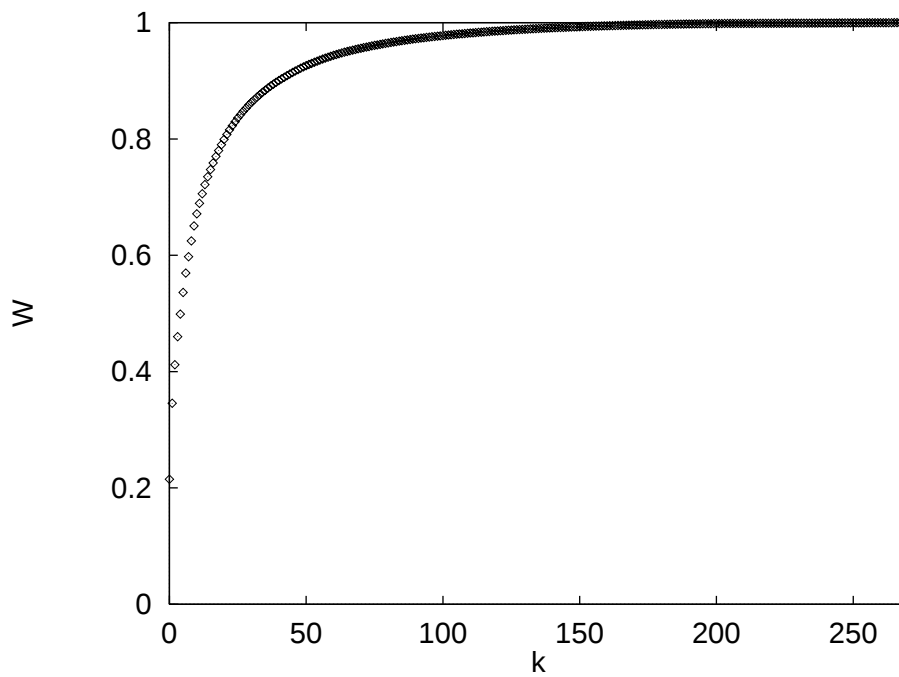


図 4.3: パラメトリック固有空間の Object1, Object2, Object3 の固有値寄与率

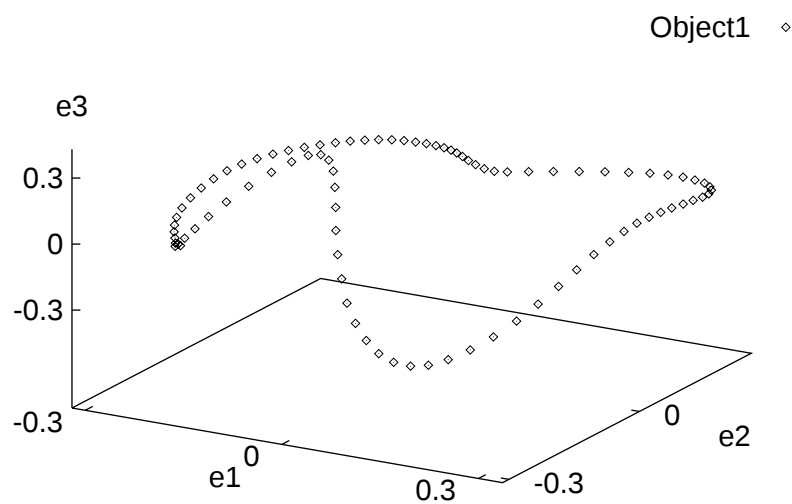


図 4.4: パラメトリック固有空間を用いた物体 1 の表現

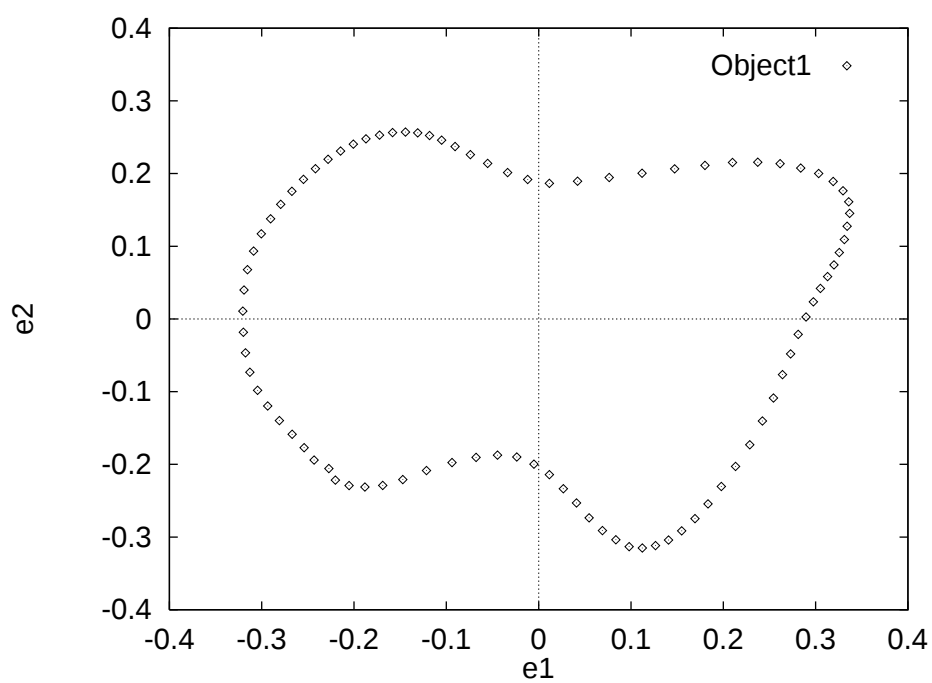


図 4.5: パラメトリック固有空間を用いた物体 1 の表現

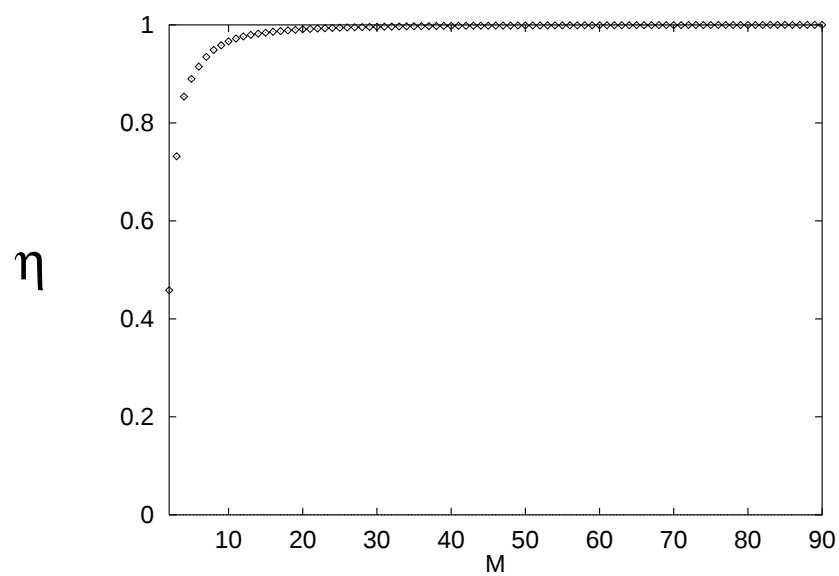


図 4.6: クラス分離度 η_M^*

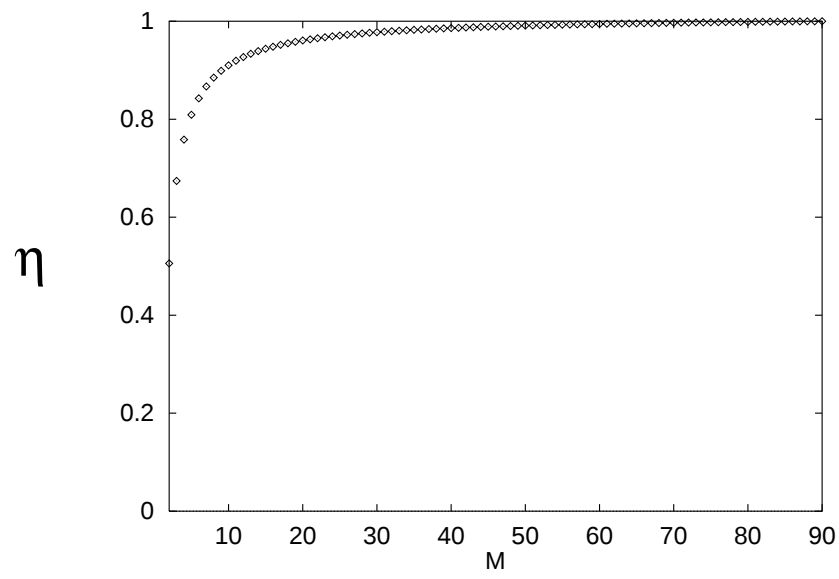


図 4.7: クラス分離度 $\bar{\eta}_M^*$

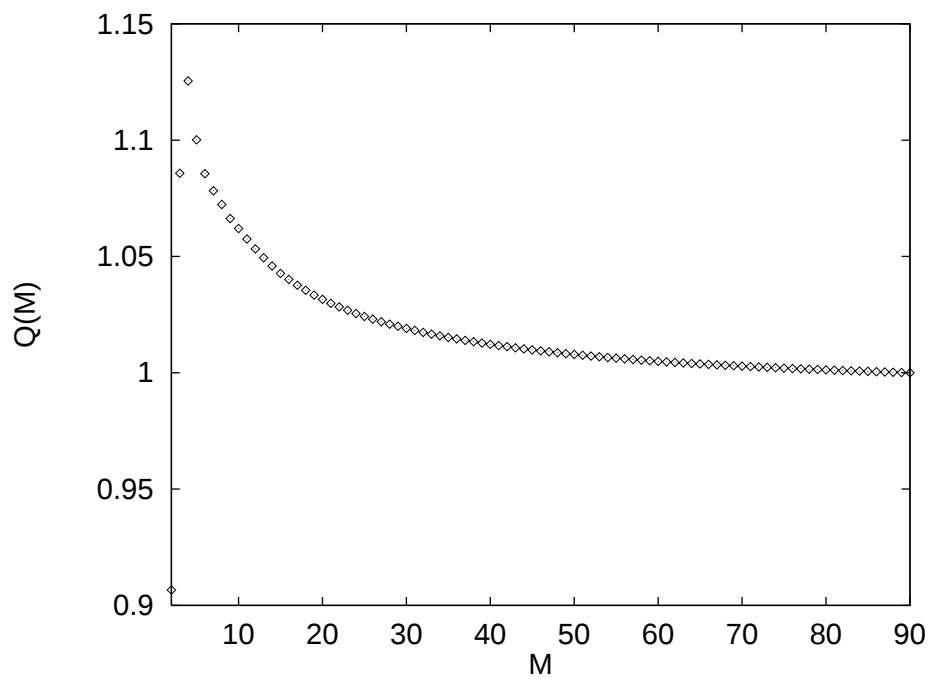


図 4.8: $Q(M)$

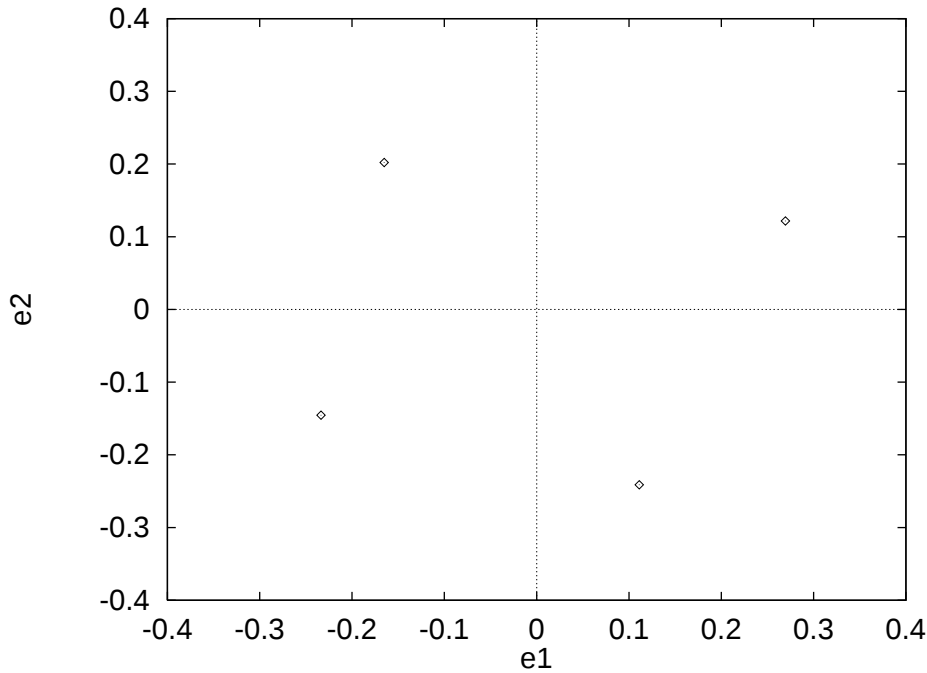


図 4.9: 第一段階 (クラス数 4)

分類された学習画像集合に対しても、閉曲線上で再帰的に分類を行うことにより階層的な辞書を作成する。ここでは、クラス内分散 σ_W^2 が $\sigma_W^2 < 0.001$ となったときそのクラスに属す学習画像の分類を終了した。同様に第二段階では図 4.10、第三段階では図 4.11、第四段階では図 4.12の代表画像で近似でき、この時点で全ての分類クラスの学習画像が $\sigma_W^2 < 0.001$ を満たし、分類を終了した。

図 4.13はそのときの視点空間の分割を示している。横軸が視点方向(0 – 360 度)、縦軸が変形過程数を表している。変形過程数の増加に伴い視点空間が階層的に分割されていく様子が明らかである。視点空間は、第一段階で 4 個のクラスに分けられ、第二段階で 9 個、第三段階で 18 個、第四段階で 29 個のクラスに階層的に分割されていく。結局、作成された辞書の木は、 $4 + 9 + 18 + 22 = 53$ のノードで構成され、90 個のノードから大きく間引くことが可能となり、辞書の記憶量を減少できる。

ただし、認識実験では、辞書の木を探索し、辞書の木の葉に辿った後に、その葉を構成しているすべての学習画像と照合を行った。これは、姿勢の推定誤差を調べ、認識時に辞書の木を間違えて辿っていないのか調べるためである。

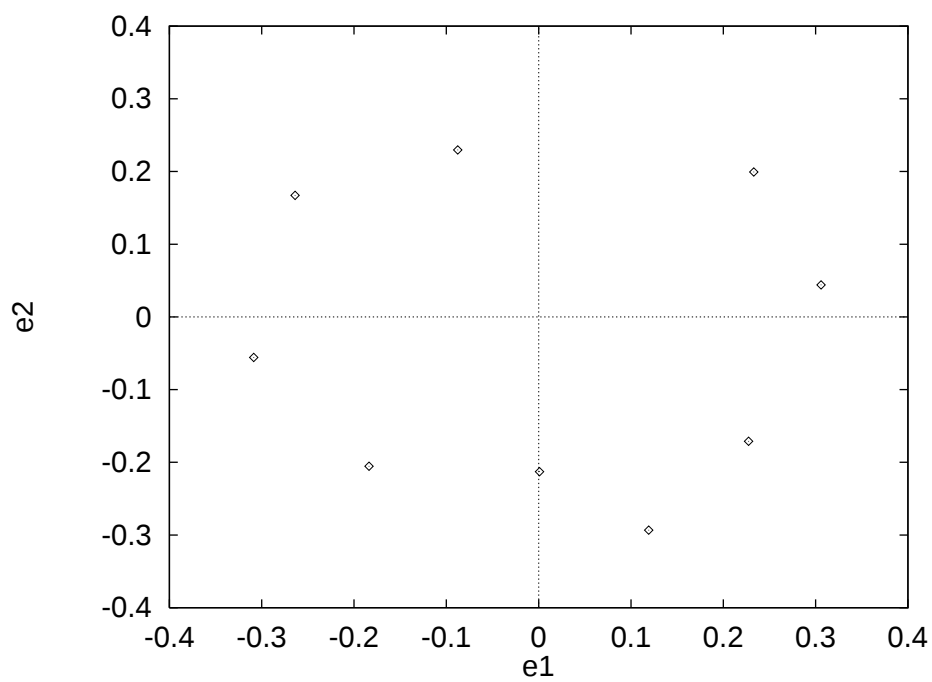


图 4.10: 第二段階

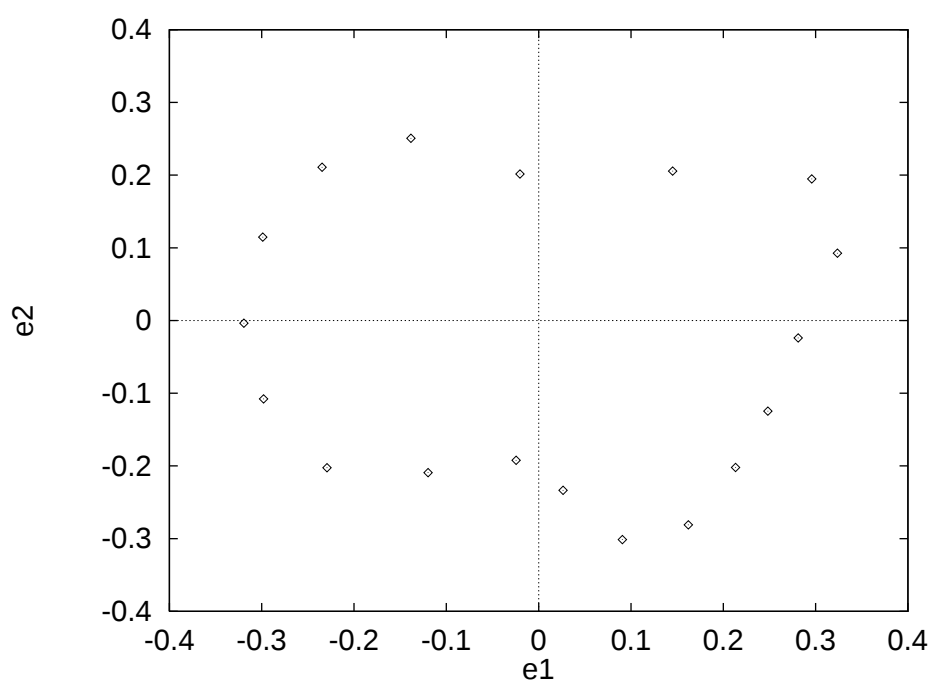


图 4.11: 第三段階

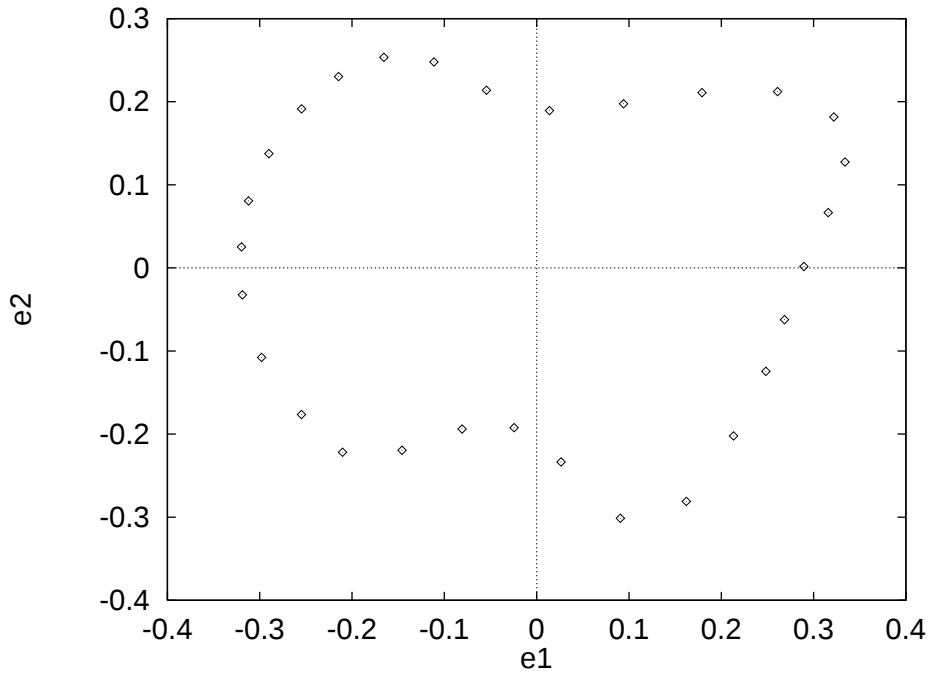


図 4.12: 第四段階

4.3 認識実験

4.3.1 提案手法を用いた認識実験

実験では、図 2.1 に示した装置を用い 2 度ごとに撮影された図 4.1(a) の 180 枚の画像を入力画像として、4 度ごとに撮影された 90 枚の学習画像で作成された階層的な辞書を用いて認識実験を行った。

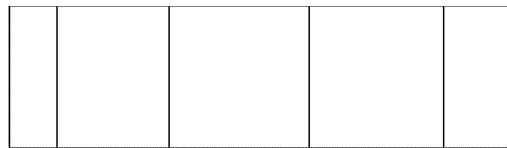
対象の認識では、未知の入力画像と、辞書の木の代表画像との間のユークリッド距離を計算し、最も類似した方を探索し、これを繰り返して木の辞書を辿り、認識結果とする。

図 4.14 は各入力画像に対する認識結果が出力されるまでの照合回数を示す。照合回数は平均で 12.9 回であった。階層的な辞書を用いて認識することにより、効率的に照合を行うことができた。

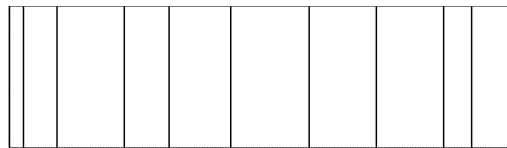
図 4.14 は各入力画像に対する認識結果の物体の姿勢の推定誤差の度数分布を示す。4 度ごとに撮影された 90 枚の画像を用いて辞書を作成し、2 度ごとに撮影された 180 枚の画像を入力画像として認識実験を行ったところ、姿勢推定誤差が最大でも 2 度であった。



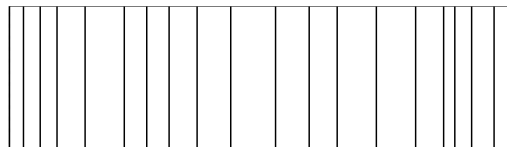
(a) 初期の視点空間



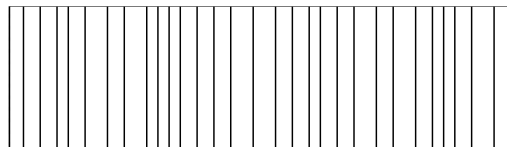
(b) 第一段階



(c) 第二段階



(d) 第三段階



(e) 第四段階

図 4.13: 視点空間の分割

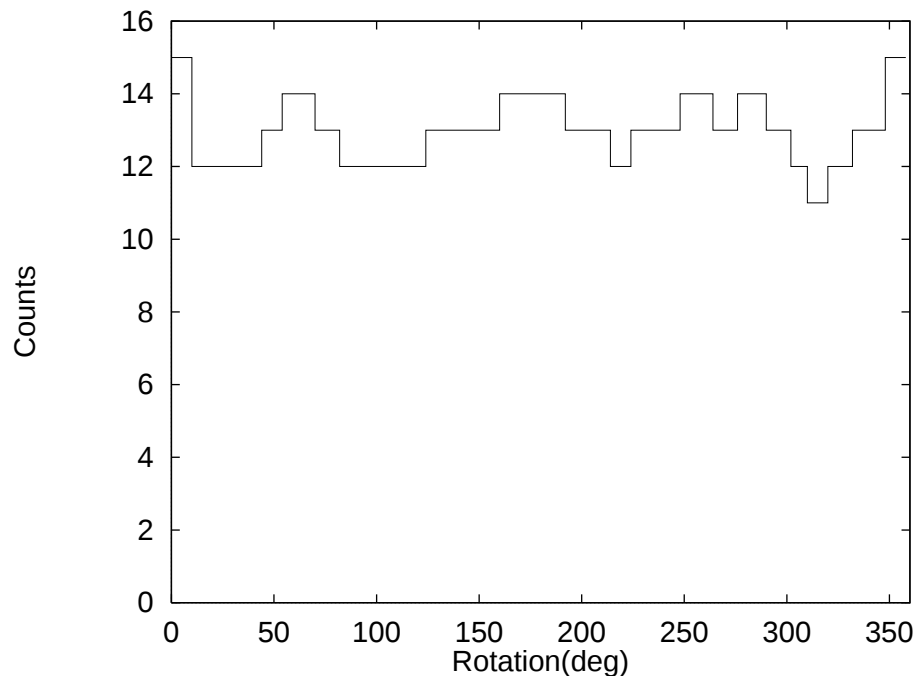


図 4.14: 照合回数

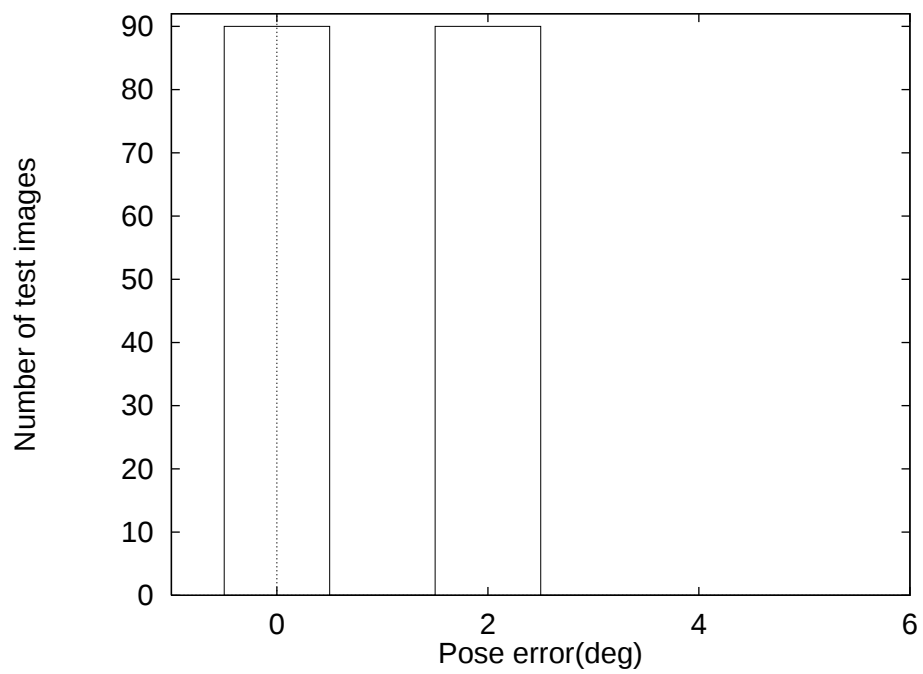


図 4.15: 物体の姿勢の推定誤差の度数分布

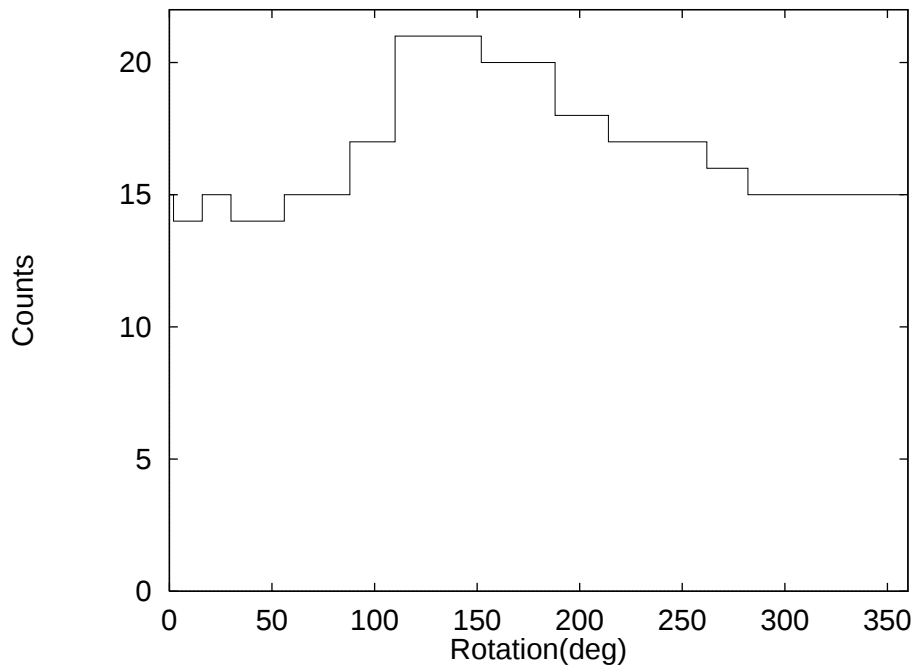


図 4.16: 提案手法により作成された辞書の照合回数 (object1)

この実験では、認識率の低下を抑えながら効率的に照合を行うことができた。

4.3.2 全物体に対する物体認識

各 object1,2,3 に対して提案手法による判別および最小自乗基準で最適な木の辞書を用いて物体認識を行った。object1,3 は姿勢推定誤差が最大でも 2 度であったが、object2 に対しては図 4.19 のように姿勢推定誤差が最大で 10 度の場合があった。しかし、物体を正しく認識することは可能であった。

4.3.3 各 object に対する姿勢推定

各 object1,2,3 に対し、それぞれの固有空間を作成し本手法により作成された辞書を用いて物体の姿勢推定を行った。各々、95 %程度 ($W_k = 0.95$) の固有値寄与率を得るためには、固有空間として object1 に対しては 19 次元、object2 に対しては 30 次元、object3 に対しては 37 次元程度必要であった。ここでは、クラス内分散 σ_W^2 が $\sigma_W^2 < 0.03$ となったときそのクラスに属す学習画像の分類を終了した。作成された辞書を用いて、各々の

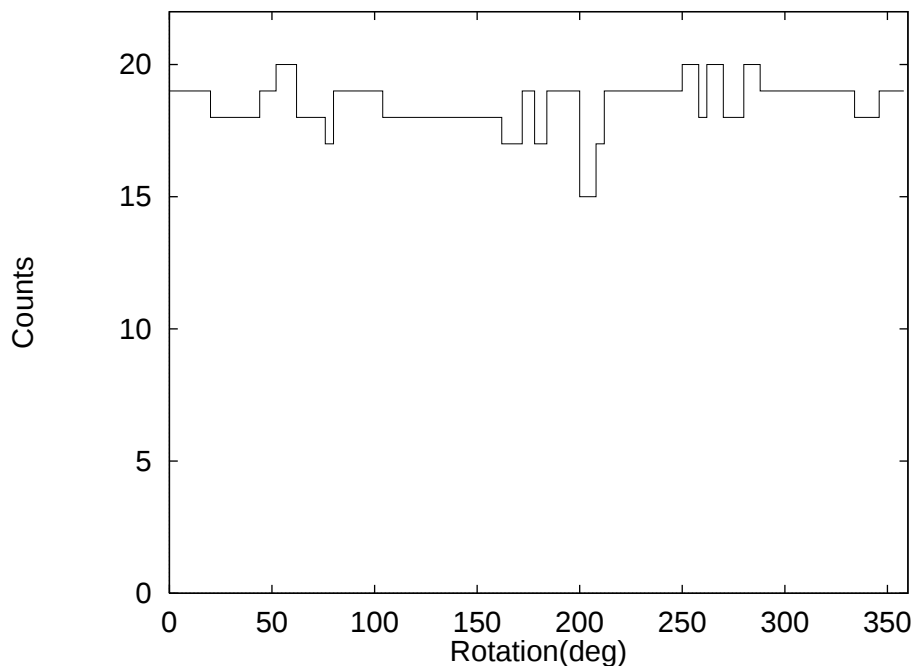


図 4.17: 提案手法により作成された辞書の照合回数 (object2)

object1,2,3 に対し姿勢推定を行った。物体 1,2,3 それぞれ姿勢推定誤差は最大で 2 度であり、照合回数はそれぞれ、object1 は平均で 17.2 回、最大で 22 回であった (図 4.20)。object3 は平均 26.2 回、最大で 28 回であった (図 4.21)。object2 は平均 18.6 回、最大で 20 回であった (図 4.22)。

このことから、本手法により作成された閉曲線上での判別および最小自乗基準で最適な木の辞書を用いたところ、物体の姿勢をほぼ正しく推定でき、照合回数を削減できた。

4.3.4 各 object に対する提案手法による二分木

提案手法により二分木の辞書を作成し、それを用いて物体の姿勢推定を行った。object1,2 はそれぞれ姿勢推定誤差は最大で 2 度であった。しかし、object3 は図 4.26のように姿勢推定誤差が大きい場合があった。

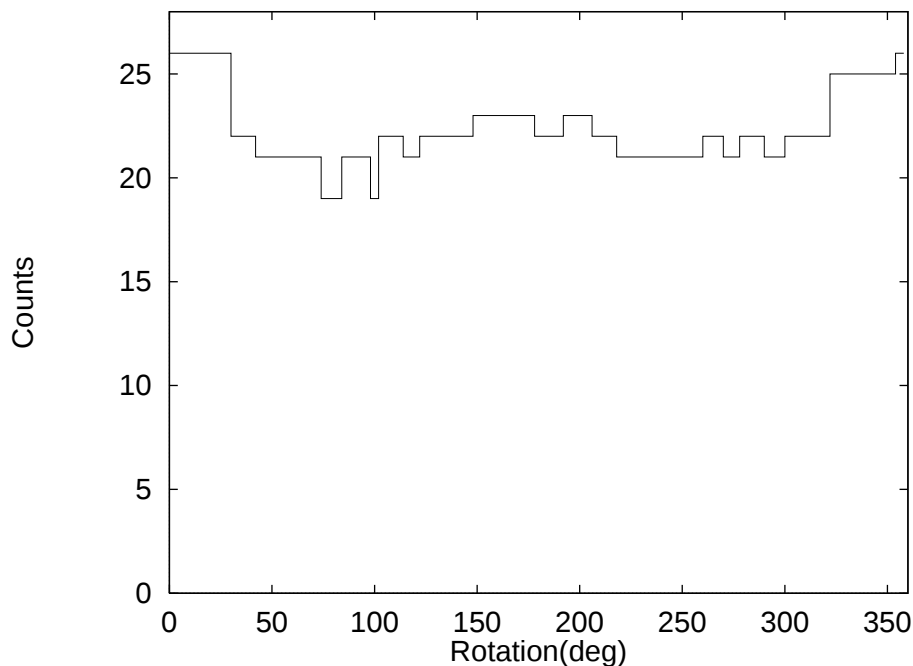


図 4.18: 提案手法により作成された辞書の照合回数 (object3)

4.3.5 UPGMA 法 vs. 提案手法による二分木を用いた認識実験

図 4.1(b) の人形を対象とし、4 度ごとに撮影された 90 枚の画像で、UPGMA 法、本手法でそれぞれ二分木の辞書を作成した。認識では、2 度ごとに撮影された 180 枚の画像を入力画像とし辞書探索を行った。図 4.27 より、本手法で作成された二分木では照合回数は最悪で 20 回、平均で 14.8 回であった。UPGMA 法で作成された二分木では照合回数は最悪で 22 回、平均で 15.3 回であった。

図 4.28 より、本手法で作成された二分木では最悪姿勢推定誤差は 2 度であった。UPGMA 法で作成された二分木では最悪姿勢推定誤差は 8 度であった。

このことから、UPGMA 法で作成された二分木よりも、本手法で作成された二分木は認識率の低下を抑えながら効率的に照合を行うことができた。この理由として、UPGMA 法で作成された二分木は辞書の木を葉から根の順に作成するため、探索時に木の辞書を間違えた経路で辿る場合がある。一方、本手法による二分木は辞書の木を根から葉の順に作成するため、探索時に coarse-to-fine 戦略に基づいて認識率の低下を抑えながら効率的に照合を行うことができたと考えられる。

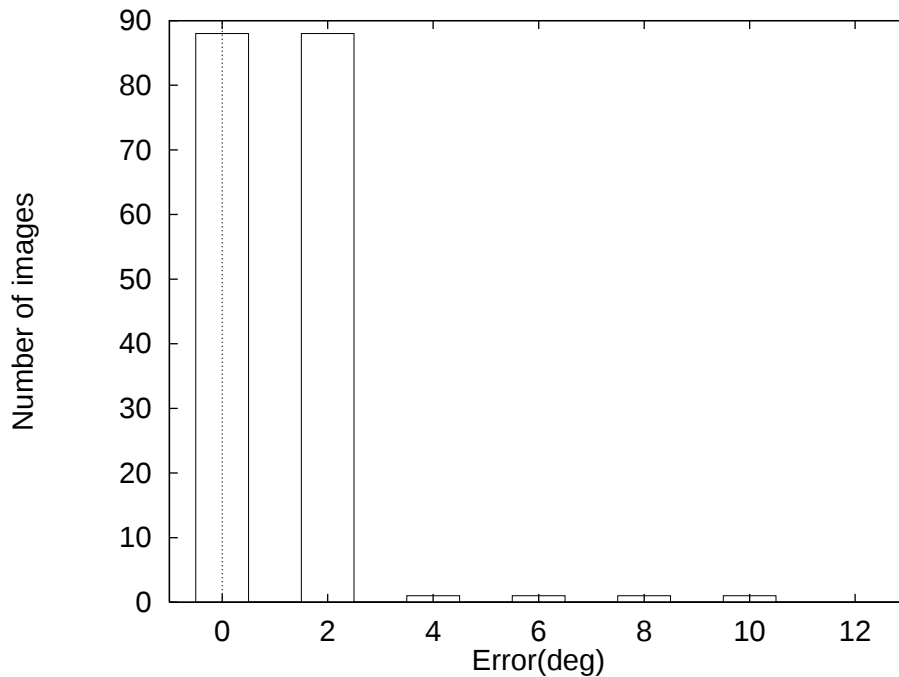


図 4.19: 物体の姿勢の推定誤差の度数分布 (object2)

4.3.6 提案手法による判別および最小自乗基準で最適な木 vs. 提案手法で作成された二分木を用いた認識実験

装置 (図 2.1) を用い 4 度ごとに撮影された 90 枚の学習画像 (図 4.1(c)) で本手法により、二分木と判別および最小自乗基準で最適な木を辞書として作成した。この階層的な辞書を用いて認識実験を行った。実験では、2 度ごとに撮影された 180 枚の画像を入力画像とした。図 4.29 はそれぞれの辞書を用いたときの各入力画像に対する認識結果が出力されるまでの照合回数を示す。本手法により作成された二分木では照合回数が平均で 13.6 回であり、判別および最小自乗基準で最適な木では照合回数が平均で 26.2 回であり、二分木の方が照合回数は少なかった。

図 4.30 はそれぞれの辞書を用いたときの各入力画像に対する認識結果の物体の姿勢の推定誤差の度数分布を示す。この図から、入力画像によっては、本手法による二分木の辞書を用いても物体の姿勢の推定誤差が大きい場合がある。しかし、本手法による判別および最小自乗基準で最適な木を用いると、姿勢の推定誤差を抑えることができた。

常に 2 つのグループに分類されて作成された二分木の辞書よりも、各ノードごとに本

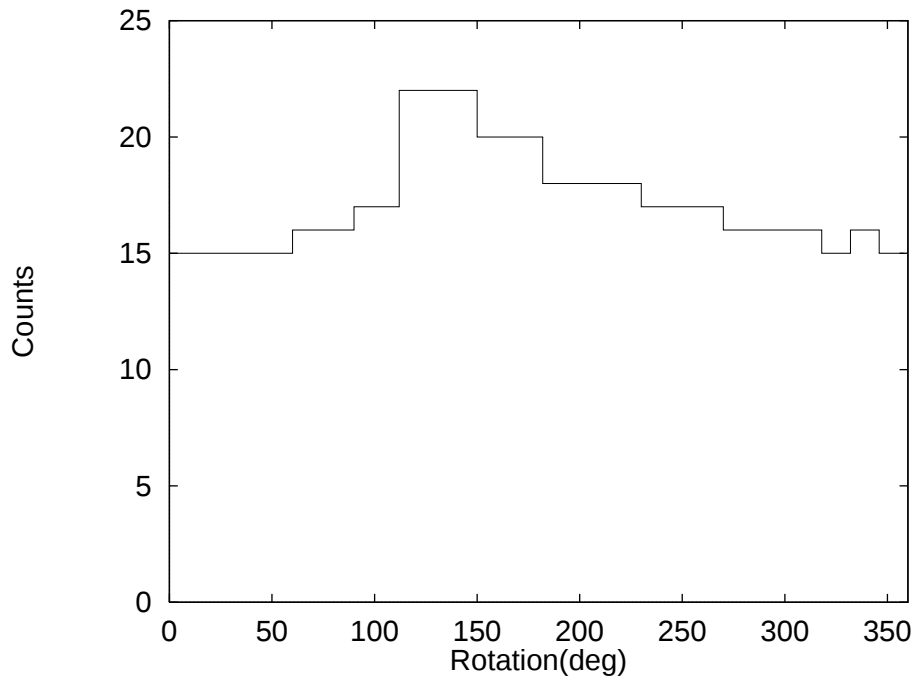


図 4.20: 照合回数 (object1)

手法により最適なクラス数の木の辞書の方が、照合回数は多くなるが良い認識率を得た。

4.4 まとめ

本研究によるパラメトリック固有空間法における階層的な辞書作成法を用いて認識実験を行い、この結果についての考察を行った。

まず、パラメトリック固有空間法における辞書作成について述べ、パラメトリック固有空間法における階層的な辞書作成法の流れを示した。実験の有効性を示すために、比較実験 (UPGMA 法 vs. 本手法による二分木、本手法による判別および最小自乗基準で最適な木 vs. 本手法による二分木) を行った。

パラメトリック固有空間法での記憶量、計算量を削減するのに本手法は有効であることを示した。

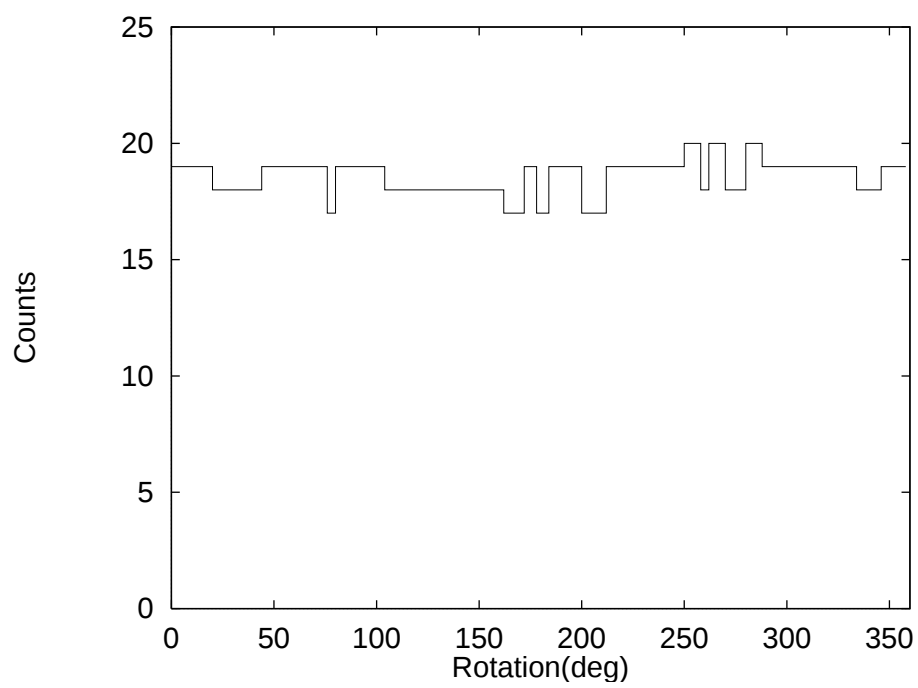


図 4.21: 照合回数 (object2)

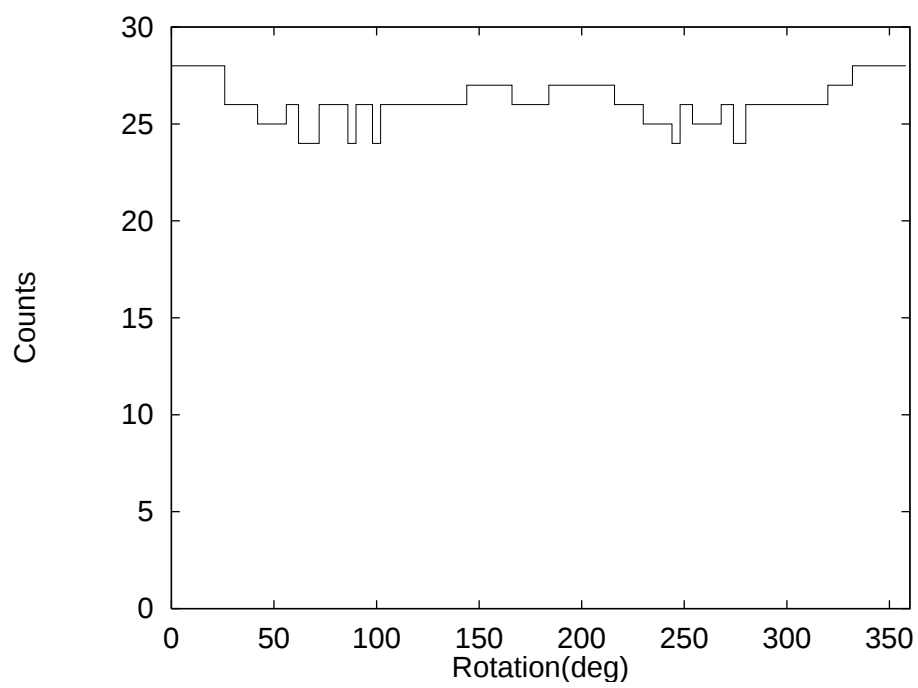


図 4.22: 照合回数 (object3)

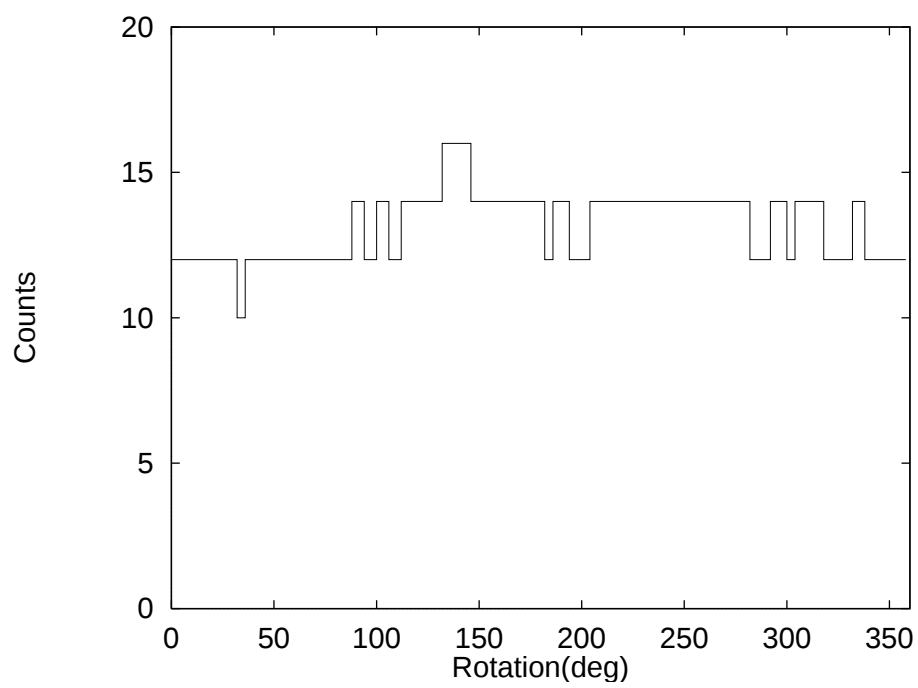


图 4.23: 二分木 (object1)

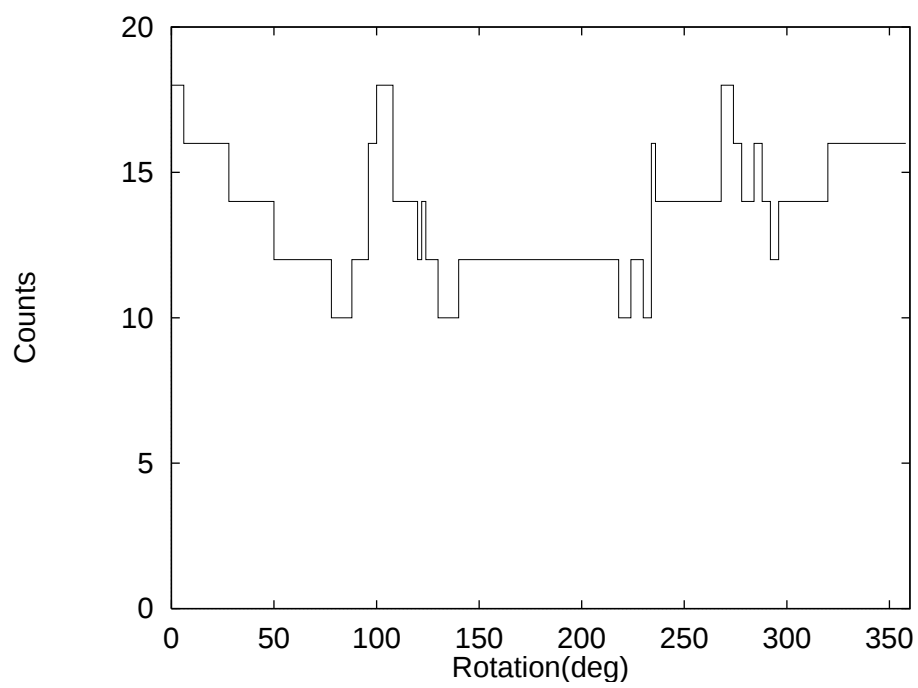


图 4.24: 二分木 (object3)

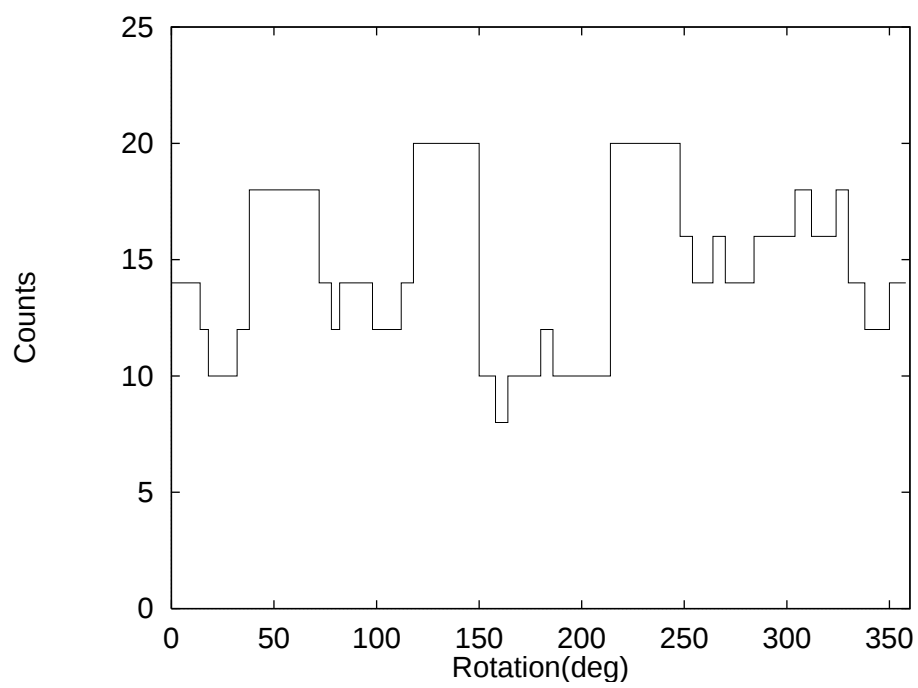


图 4.25: 二分木 (object2)

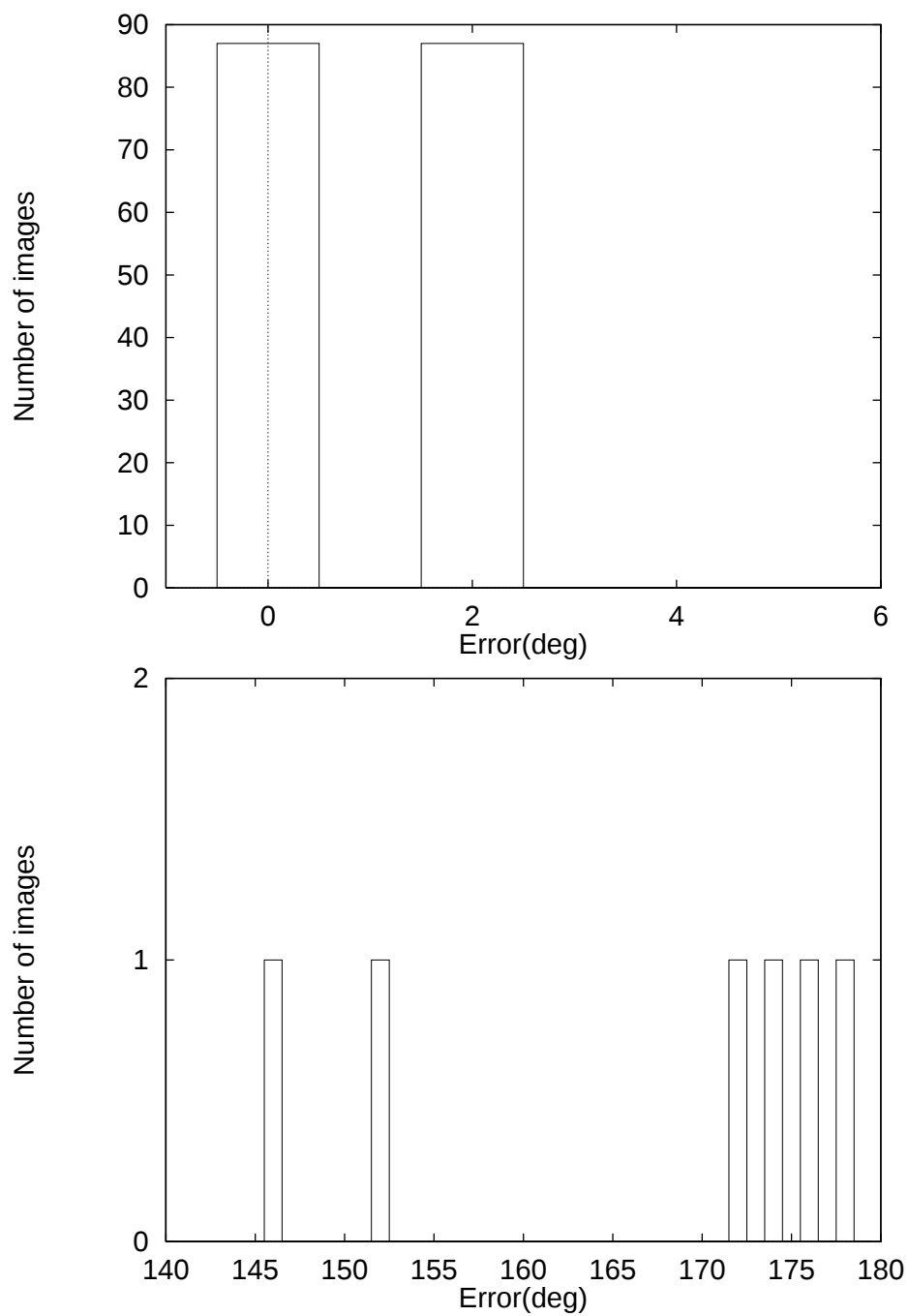


図 4.26: 物体の姿勢の推定誤差の度数分布 (object3)

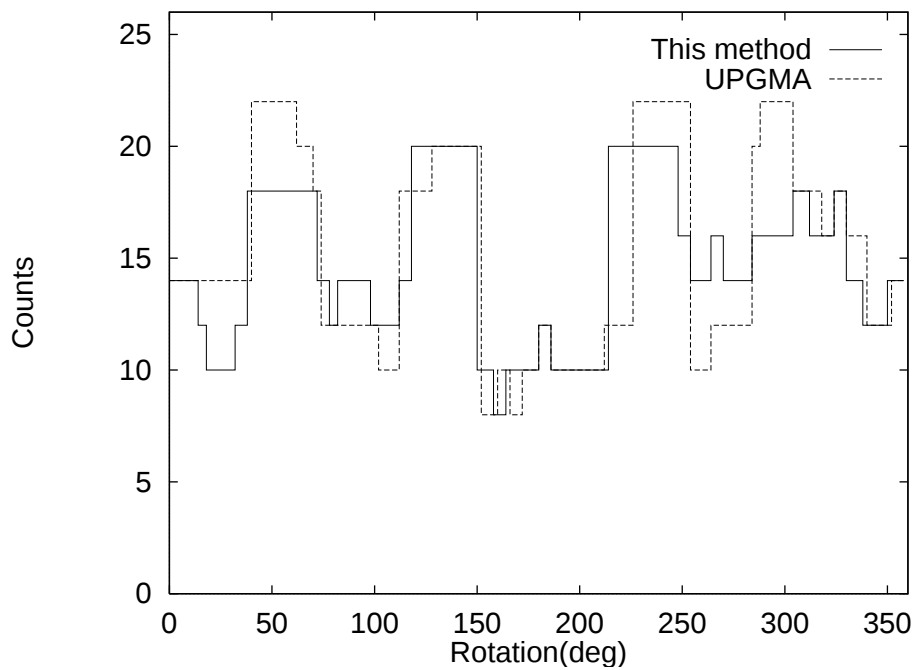


図 4.27: 物体 2 の各入力画像に対する認識結果出力までの照合回数

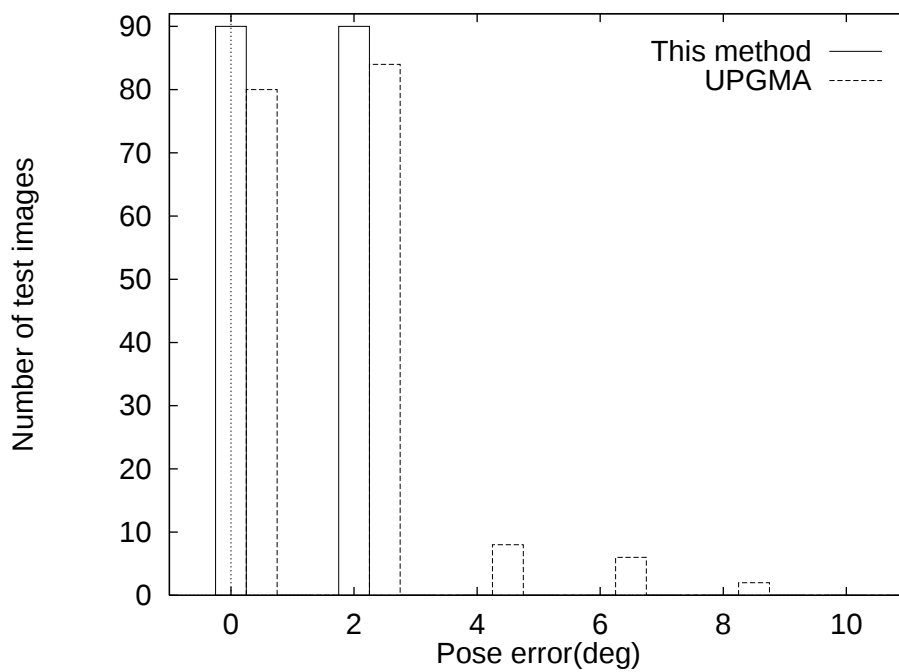


図 4.28: 物体 2 の姿勢の推定誤差の度数分布

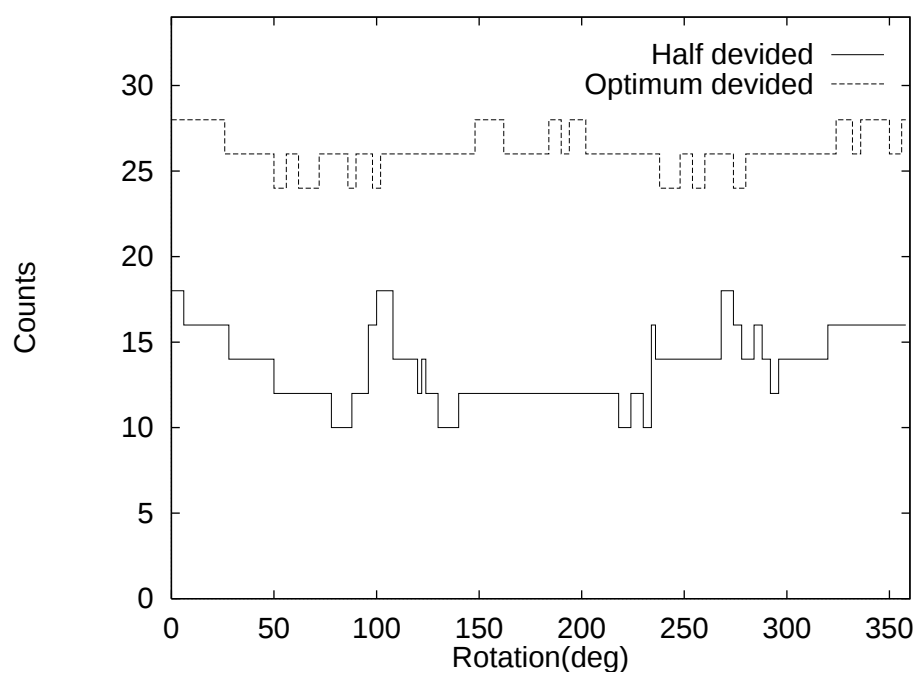


図 4.29: 本手法による二分木 vs 本手法による判別および最小自乗基準に基づく最適木の照合回数の比較

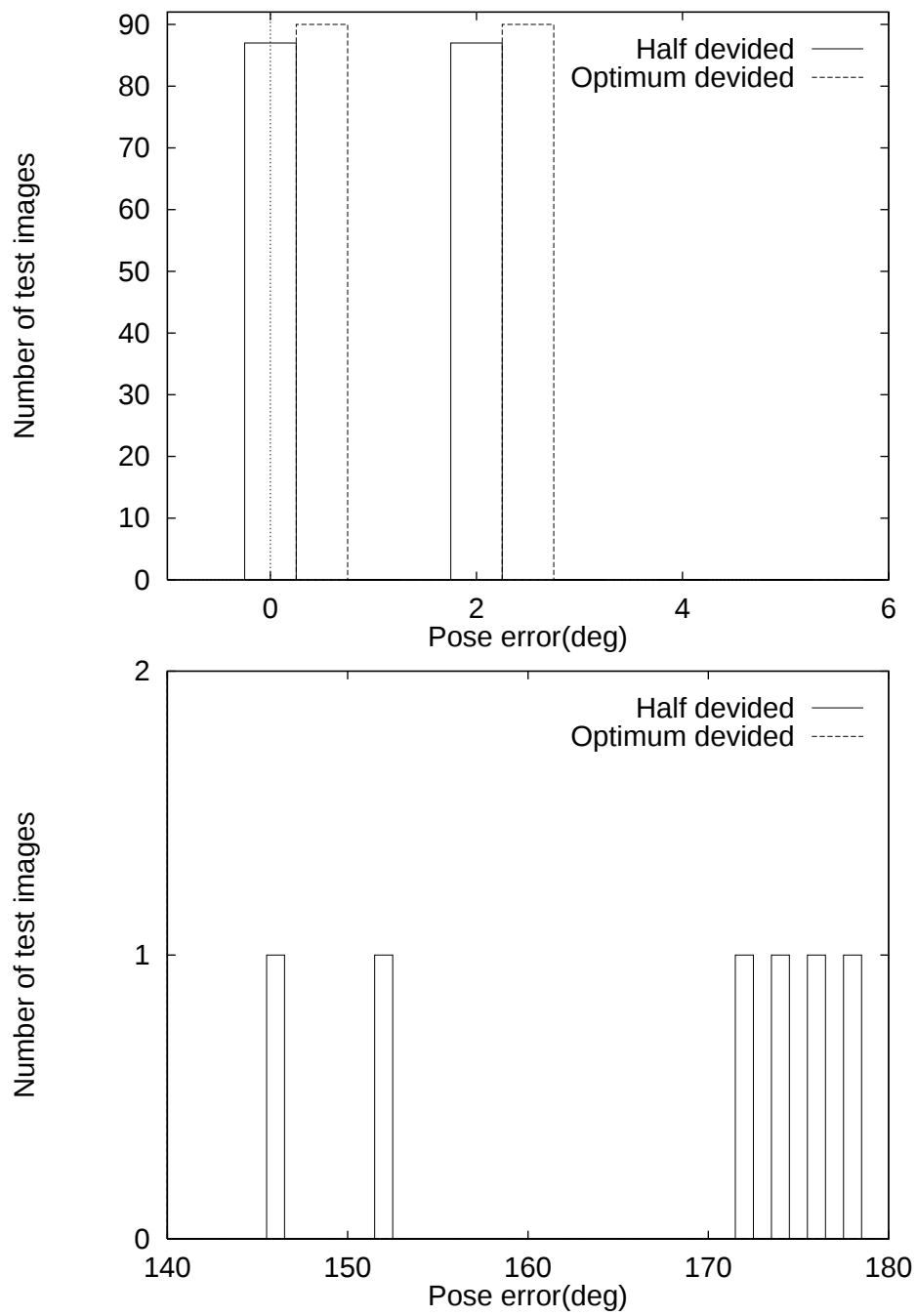


図 4.30: 物体の姿勢の推定誤差の度数分布

第 5 章

結論

パラメトリック固有空間法 [4] は、画像符号化を基本とした二次元照合法であり、高次元な画像データを主成分分析により低次元の特徴ベクトルで表し低次元空間で照合を行う。そのため、三次元物体を二次元画像の集合体として記憶する際に記憶量を削減することができ、記憶されたデータと入力画像との照合に要する時間も短縮できる。

しかし、パラメトリック固有空間法では辞書として、視点や光源の位置により複雑に変化するすべての物体の学習画像を記憶する必要があり、膨大な記憶量を必要とする。また、認識時に、入力画像と辞書のすべての学習画像を照合する必要があり膨大な照合時間を必要とする。そのため実用化のためには、パラメトリック固有空間法における辞書の記憶量、照合の計算量をさらに削減する必要がある。

そこで、本研究では、パラメトリック固有空間法における辞書の記憶量、照合の計算量を削減するために、パラメトリック固有空間法における階層的な辞書作成法を提案した。具体的には、視点が連続的に変化する学習画像を間引くことにより辞書の記憶量を削減し、辞書を階層的に構成することにより照合の計算量を減少させる。

本研究では、パラメトリック固有空間法における辞書は、物体の回転に伴い固有空間上で一つの閉曲線を構成するため、認識対象となる物体の画像は閉曲線上の近辺に分布するはずであるということに注目し、学習画像を閉曲線上で分類する手法を提案した。学習画像を閉曲線上で分類することにより、分類された学習画像は視点が隣接した画像で構成され、学習画像に存在しない画像をも分類された画像群にある程度含ませることができる。さらに、本研究では、閉曲線上でクラス分離度が良くなるように辞書の学習画像进行分类する手法として、閉曲線上での判別および最小自乗基準法で最適な画像群分類法を提案し

た。分類された学習画像は視点が隣接しかつ固有空間上で類似している画像でのみ構成され、分類された学習画像を一つの代表画像で表すことができ、多数の学習画像を少数の代表画像で表し辞書の記憶容量を削減できる。

本研究では、分類された学習画像に対しても、閉曲線上で再帰的に分類を行うことにより階層的な辞書を作成する。これにより照合の計算量を減少させる。この分類を基に辞書の本を根から葉の順に作成することにより、探索時に coarse-to-fine 戦略に基づいて認識率の低下を抑えながら効率的に照合を行うことができる。

提案した手法に基づいて作成されたパラメトリック固有空間法における階層的な辞書を用いて認識実験を行うことにより、パラメトリック固有空間法における辞書の記憶量、照合の計算量を削減できるかどうか検討した。クラス内分散がある程度小さくなったときそのクラスに属す学習画像の分類を終了したところ、提案した手法に基づいて作成された辞書は従来の辞書の画像数から大きく間引くことが可能となり、辞書の記憶量を減少させることが可能であった。また、認識率の低下を抑えながら効率的に照合を行うことが可能となった。このことより、本研究で提案したパラメトリック固有空間法における階層的な辞書が辞書の記憶量を削減し照合の計算量を減少するのに有効であると思われる。

しかしながら、固有空間を視点角度の位相を考えずに作成しており、学習画像の分類を視点角度の位相を考えて行っている。今後の課題として、固有空間を視点角度の位相を考えて作成する必要がある。

謝辞

本研究を行うにあたり、いろいろと御指導をしていただいた北陸先端科学技術大学院大学 阿部 亨 助教授に厚く御礼申し上げます。

また、さまざまな面で御教授いただきました北陸先端科学技術大学院大学 堀口 進 教授に深く感謝致します。

北陸先端科学技術大学院大学 下平 博 助教授には、サブテーマで熱心に御指導いただき深く感謝致します。

また、日頃よりお世話になった当研究室の皆様に心より感謝致します。

参考文献

- [1] 天野敏之, 日浦慎作, 井口征士, 距離画像の固有空間照合法に基づく物体の姿勢検出, 信学技報. PRU95-74, pp.145-150(1995)
- [2] 天野敏之, 日浦慎作, 山口証, 井口征士, 固有空間照合法に基づく距離画像からの物体の姿勢検出, 電子情報通信学会論文誌. Vol.J80-D-II No.5, pp.1136-1143(1997)
- [3] 大津展之, 判別および最小2乗基準に基づく自動しきい値選定法, 電子通信学会論文誌. Vol.J63-D No.4, pp.349-356(1980)
- [4] Murase.H and Nayar S.K, “Visual Learning and Recognition of 3-D Objects from Appearance”, International Journal of Computer Vision,14,5-24,1995.
- [5] H.Murase and Nayar S.K, ”Illumination planning for object recognition in structured environments”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.16,No.12,pp.1219-1227,1994.
- [6] H.Murase and Nayar S.K, “多重解像度と固有空間表現による三次元物体のイメージスポッティング”, 情報処理学会論文誌, Vol.36, No.10, pp.2234-2243,1995.
- [7] H.Murase and R.Sakai, “Moving Object Recognition in Eigenspace Representation : Gaint Analysis and Lip Reading”, Pattern Recognition Letters,17,pp.155-162,1996.
- [8] 大野宏, 山本正信, 固有空間法による動作の識別, 電子情報通信学会 技術研究報告, Vol.97,PRMU97-195,pp.71-76(1997).
- [9] 大場光太郎, 池内克史, 局所固有空間手法による金属物体の安定認識, 電子情報通信学会論文誌, Vol.J80-D-II,No.12,pp.3147-3154(1997).

[10] 西田英郎・佐藤嗣二 共訳, “実例クラスター分析”, 内田老鶴圃, 1992