

Title	局所固有空間法を用いたランドマーク認識法に関する研究
Author(s)	村上早苗
Citation	
Issue Date	1999-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1264
Rights	
Description	Supervisor:阿部 亨, 情報科学研究科, 修士

修 士 論 文

**局所固有空間法を用いた
ランドマーク認識法に関する研究**

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

村上早苗

1999 年 2 月 15 日

要 旨

本稿では, 局所固有空間法を用いて外界を撮影した画像中のランドマークの有無・存在位置を認識する手法を提案する. その際, 辞書に登録する学習画像群を判別分析を用いてクラスタリングし, 辞書を木状に構造化することで辞書探索の効率化を図る.

目次

1	はじめに	1
1.1	背景と目的	1
1.2	本稿の構成	2
2	固有空間手法	3
2.1	パラメトリック固有空間法	3
2.1.1	パラメトリック固有空間法による学習	4
2.1.2	パラメトリック固有空間法による認識	5
2.2	局所固有空間法	6
2.2.1	局所ウィンドウの学習	7
2.2.2	入力ウィンドウの対応付け	7
2.2.3	投票	7
3	学習ウィンドウの設定基準	9
3.1	学習ウィンドウの設定基準	9
3.2	実験	13
3.2.1	辞書作成	14
3.2.2	実験	15
4	階層的辞書構成	27
4.1	クラスタリングに基づいた学習ウィンドウの選択	27
4.1.1	クラスタリング	27
4.1.2	学習ウィンドウ設定箇所の選択	28
4.2	階層的クラスタリング	29
4.3	実験	31
4.3.1	学習ウィンドウ設定箇所の選択	31

4.3.2	実験	31
5	まとめ	35

目 次

2.1	パラメトリック固有空間法	4
2.2	局所固有空間法	6
3.1	入力ウィンドウ	9
3.2	輝度値の変化なし	10
3.3	多方向に輝度値の変化あり	11
3.4	1 方向のみ輝度値の変化あり	11
3.5	ずれに強い領域	12
3.6	ランドマーク	13
3.7	特徴ウィンドウ 3088 枚	13
3.8	分散の小さい局所ウィンドウ 2953 枚	13
3.9	重なり合う局所ウィンドウの削除 101 枚	14
3.10	学習ウィンドウ 58 枚	14
3.11	パラメータ 750mm で選択された学習ウィンドウ	14
3.12	距離 900mm で選択された学習ウィンドウ	15
3.13	距離 1500mm で選択された学習ウィンドウ	15
3.14	距離 750mm の入力画像, ランドマークの位置 (133,129)	19
3.15	距離 750mm の入力画像の 750mm の投票空間	19
3.16	距離 750mm の入力画像の 900mm の投票空間	20
3.17	距離 750mm の入力画像の 1500mm の投票空間	20
3.18	距離 900mm の入力画像, ランドマークの位置 (196,119)	21
3.19	距離 900mm の入力画像の 750mm の投票空間	21
3.20	距離 900mm の入力画像の 900mm の投票空間	22
3.21	距離 900mm の入力画像の 1500mm の投票空間	22
3.22	距離 1500mm の入力画像, ランドマークの位置 (180,204)	23
3.23	距離 1500mm の入力画像の 750mm の投票空間	23

3.24	距離 $1500mm$ の入力画像の $900mm$ の投票空間	24
3.25	距離 $1500mm$ の入力画像の $1500mm$ の投票空間	24
3.26	遮蔽のあるパラメータ $750mm$ の入力画像	25
3.27	遮蔽のあるパラメータ $900mm$ の入力画像	25
3.28	遮蔽のあるパラメータ $1500mm$ の入力画像	25
3.29	遮蔽のあるパラメータ $750mm$ の入力画像	26
3.30	遮蔽のあるパラメータ $900mm$ の入力画像	26
3.31	遮蔽のあるパラメータ $1500mm$ の入力画像	26
4.1	x 軸方向への学習サンプルの分割	29
4.2	y 軸方向への学習サンプルの分割	30
4.3	木構造	31
4.4	学習ウィンドウ設定箇所 ($750mm$)	31
4.5	学習ウィンドウ設定箇所 ($900mm$)	32
4.6	学習ウィンドウ設定箇所 ($1500mm$)	32

表 目 次

3.1	パラメータ $750mm$ の入力画像の投票結果	16
3.2	パラメータ $900mm$ の入力画像の投票結果	16
3.3	パラメータ $1500mm$ の入力画像の投票結果	16
3.4	パラメータ $750mm$ の推定結果	17
3.5	パラメータ $900mm$ の推定結果	17
3.6	パラメータ $1500mm$ の推定結果	18
3.7	遮蔽のある入力画像の推定結果	18
3.8	遮蔽のある入力画像の推定結果	18
4.1	パラメータ $750mm$ の推定結果	32
4.2	パラメータ $900mm$ の推定結果	33
4.3	パラメータ $1500mm$ の推定結果	33
4.4	遮蔽のある入力画像の推定結果	33
4.5	遮蔽のある入力画像の推定結果	34

第 1 章

はじめに

1.1 背景と目的

2次元画像から画像内に映し出されている対象を認識する手法は従来から多く研究されてきており, 1つの応用として, 室内環境内でのロボットの自律移動がある. 本研究では, ロボットの自律移動の際に目印となり得るような室内環境に現れる標識(誘導灯等)をランドマークとし, カメラで撮影した2次元画像からこのランドマークの有無・存在位置の認識を行なうことを目指す.

2次元画像内の対象を認識するには, 認識対象の3次元構造を利用する照合方法と2次元の見かけの画像のみを用いて照合する2次元照合手法の2つに大きく分類される. 現在, 2次元照合手法は実環境において良好な認識結果をもたらしており, そこで本研究では2次元照合手法の1つであるパラメトリック固有空間法 [1] を発展させた手法である, 局所固有空間法 [2] を用いてランドマークの有無・存在位置の認識を行なうことにする.

パラメトリック固有空間法はK-L展開により高次元の画像データを低次元の特徴ベクトルに圧縮し, 低次元の固有空間内で辞書の構成・認識を行なうため, 認識時の入力画像との照合に要する処理量を削減できる. また, パラメータ(カメラとランドマークの位置関係)を連続的に変化させた画像を固有空間上の多様体で表現することによって, 入力画像内のランドマークの識別と同時にパラメータも推定可能になる. しかし, 学習時にランドマーク全体を圧縮するため, 隠れのあるランドマークを認識することはできない. そこで, ランドマークの局所領域に固有空間法を適用させた局所固有空間法が提案された [2]. 局所領域に注目するため, 入力画像内にランドマーク全体が写ってなくても, 局所領域の照合によって認識可能となる.

一方, 2次元照合手法は視点の変化によってランドマークの見え方が変化し, これら全

ての画像をあらかじめ辞書として蓄えておく必要があり、認識時の照合に要する処理量が多くなる。また、局所固有空間法を用いることによって、用意する画像数は増加し、さらに認識速度の悪化を招く。

そこで本稿では、局所固有空間法を用いて外界を撮影した画像中のランドマークの有無・存在位置を認識する際の照合の効率化を目指し、認識に効果的な学習画像を選択する手法を提案する。さらに、辞書に登録する画像を判別分析を用いてクラスタリングし、辞書を木状に構造化することで照合回数を減らし、認識時間の削減を目指す。

1.2 本稿の構成

本稿ではまず2章において、2次元照合法として提案されたパラメトリック固有空間法について説明する。そしてその問題点を明らかにし、パラメトリック固有空間法の拡張として提案された局所固有空間法について述べる。3章では、辞書に登録する画像数を認識精度を低下させることなく削減するための画像の選択基準について考察する。4, 5章では更に学習画像を削減するのに効果的な選択方法を提案し、また、検索の効率を向上させるために辞書を木状に構成する。6章ではそれぞれの有効性を確認するための認識実験を行ない、その結果について考察する。

第 2 章

固有空間手法

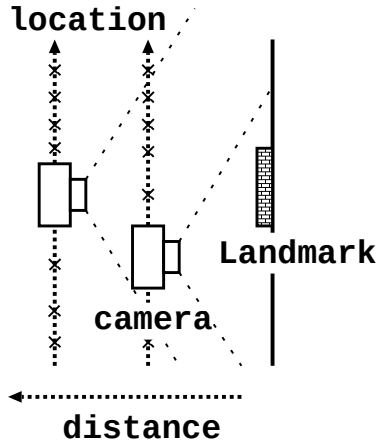
2.1 パラメトリック固有空間法

固有空間は物体認識における見かけ画像の表現や、画像処理一般における画像の表現に有用な空間であり、統計的特徴抽出手法として利用されてきた。これに対し、物体の統計的特徴抽出のみならず、物体の位置等の推定にもこの空間を用いて表現するパラメトリック固有空間法が提案された [1]。これは、異なる視点から固有空間法を発展させた手法で、その考え方は、パラメトリックに変化する画像を固有空間中の多様体で表現しようとするものである。

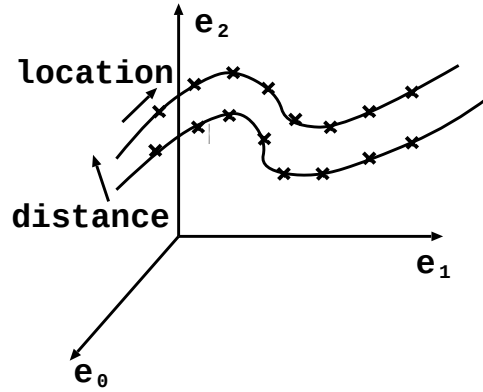
図 2.1(a) のようにカメラからの距離・位置を変化させてランドマークを撮影し、固有空間へ各画像を投影する。この場合図 2.1(b) に示すように、1 枚の画像は固有空間上で 1 つの点として表記される。この時、隣接した箇所で撮影した画像の相関は高いため、固有空間上でも隣接した箇所に投影される。つまり、カメラからの位置をパラメータとして連続的に変化させて撮影した画像は、固有空間上では多様体を構成し、パラメータ情報も固有空間上で表記できる。

パラメトリック固有空間法は学習段階と認識段階の 2 つのステップから成る。

学習段階では、ランドマークに対してパラメータ（カメラとランドマークとの位置関係）を変化させた画像を得、これを学習画像とし、固有空間内で辞書（多様体）を構成する。認識段階では、入力画像を固有空間に投影し、その時の点と固有空間上の多様体との位置関係から、ランドマークの認識・パラメータ推定を行なう。



画像撮影



固有空間表現

図 2.1: パラメトリック固有空間法

2.1.1 パラメトリック固有空間法による学習

あるランドマークに対してパラメータを変化させて M 枚の画像を得、これを学習画像とする。各学習画像をラスタースキャンし、画素値を要素とする画像ベクトル

$$\hat{\mathbf{x}} = [\hat{x}_1, \hat{x}_2, \dots, \hat{x}_N]^T \quad (2.1)$$

で各画像を表すことにする。ここで N は画素数である。

次にセンサ感度の影響を除くため、各画像のベクトル $\hat{\mathbf{x}}$ の大きさが 1 になるよう、明るさについて正規化を行なう。

$$\mathbf{x} = \frac{\hat{\mathbf{x}}}{\|\hat{\mathbf{x}}\|} \quad (2.2)$$

こうして得られた $\mathbf{x}$ をあるランドマークに対する画像集合として、

$$\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\} \quad (2.3)$$

と表す。そしてこの画像集合の各画像ベクトルを学習サンプルと呼ぶ。

この画像集合はパラメータを連続的に変化させた画像群なので、隣合う画像の相関は極めて高い。そこで、情報圧縮や類別のための統計的特徴抽出の代表的な手法である K-L 展開を用いて圧縮を行なう。これは、画像集合の情報を 2 乗誤差の意味で最適に圧縮でき、画像集合の共分散行列の最大 k 固有値に対応する直交固有ベクトルで張られる空間で、もとの画像を表現する手法である。

まず各学習サンプルから平均画像 $\mathbf{c}$ を引いた学習マトリクス

$$\mathbf{X} = [\mathbf{x}_1 - \mathbf{c}, \mathbf{x}_2 - \mathbf{c}, \dots, \mathbf{x}_M - \mathbf{c}] \quad (2.4)$$

を作成し, これから共分散行列

$$\mathbf{Q} = \mathbf{X} \mathbf{X}^T \quad (2.5)$$

を求める. この固有方程式

$$\lambda_i \mathbf{e}_i = \mathbf{Q} \mathbf{e}_i \quad (2.6)$$

を解き, k 個の大きい固有値 ($\lambda_1 \geq \dots \geq \lambda_k \geq \dots \geq \lambda_M$) に対応する固有ベクトル ($\mathbf{e}_1, \dots, \mathbf{e}_k$) を基底とする固有空間を得る.

一般に次元数の大きい行列の固有ベクトルの計算は困難であり, 特異値分解を利用する等の様々な解法が提案されている [5][6]. それらを用いることにより, 固有ベクトルの計算を可能にしている.

ここで, λ_k よりも小さい固有値の項を無視することにより, 次元を落す.

この次元数 k の決定法としてデータの全分散のうちのどれくらいの分散 (情報) が抽出されるかを示す指標である寄与率 η を計算し, 適当な閾値 λ を用いて次元 k を決定することにする.

$$\eta = \frac{\sum_i^k \lambda_i}{\sum_j^M \lambda_j} \geq \lambda \quad (2.7)$$

得られた k 次元の固有ベクトルによって構成されるマトリクス

$$\mathbf{E} = [\mathbf{e}_1, \dots, \mathbf{e}_k] \quad (2.8)$$

を用いて各学習サンプルを k 次元のランドマーク固有の空間へ以下の式で投影することができる.

$$\phi_i = \mathbf{E}^T (\mathbf{x}_i - \mathbf{c}) \quad (2.9)$$

こうして得られた ϕ_i ($i = 1, \dots, M$) を辞書として用いる. これは固有空間上で多様体として表現される.

2.1.2 パラメトリック固有空間法による認識

前節で求めた固有空間へ入力画像 $\mathbf{y}$ を以下の式により投影する.

$$\mathbf{z} = \mathbf{E}^T (\mathbf{y} - \mathbf{c}) \quad (2.10)$$

この z と各 ϕ_i のユークリッド距離 d を求め, 最小となる d を得る ϕ_i に対応する学習サンプル x^* が入力画像に対応する画像であるとし, その時のパラメータが推定すべきパラメータと判定する.

$$d(x^*) = \min_{i \in M} \| z - \phi_i \| \quad (2.11)$$

2.2 局所固有空間法

パラメトリック固有空間法は入力画像のランドマークの遮蔽・移動・背景により固有空間上の点の位置が大きく変化することがわかっている. そこで, ランドマークの特徴をよく表現している局所領域(局所ウィンドウ)のみに注目し, 各局所ウィンドウ毎にパラメトリック固有空間法を適用させた局所固有空間法が提案された [2]. 図 2.2の (a) のように各学習画像上に局所ウィンドウを設け, これを学習ウィンドウとし (図 2.2(b)), 学習ウィンドウを固有空間へ投影した図 2.2(c) を辞書として蓄えておき, これを用いて認識を行なう. つまり, ランドマークの特徴的な箇所だけに注目するため, ランドマーク全体が入力画像に映し出されていなくても, そのランドマークの特徴的な箇所が映し出されていれば認識可能であり, 入力画像内のランドマークの遮蔽に対処できる.

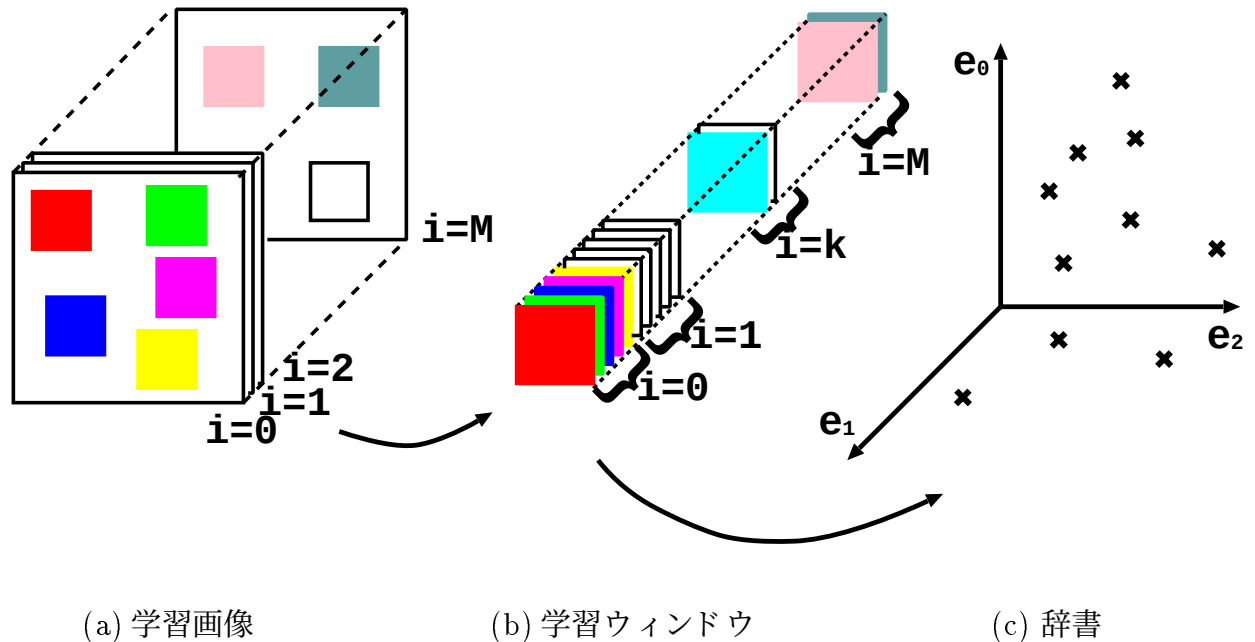


図 2.2: 局所固有空間法

2.2.1 局所ウィンドウの学習

学習サンプル $\mathbf{x}_i$ から n_i 枚の学習ウィンドウを選択する.

この時, 学習マトリクス $\mathbf{X}$ は次のようになる.

$$\mathbf{X} = [\mathbf{X}_1 | \mathbf{X}_2, \dots, \mathbf{X}_M] \quad (2.12)$$

$$\begin{aligned} &= [\mathbf{x}_1^1 - \mathbf{c}, \mathbf{x}_1^2 - \mathbf{c}, \dots, \mathbf{x}_1^{n_1} - \mathbf{c} | \\ &\quad \mathbf{x}_2^1 - \mathbf{c}, \mathbf{x}_2^2 - \mathbf{c}, \dots, \mathbf{x}_2^{n_2} - \mathbf{c} | \\ &\quad \dots \\ &\quad \mathbf{x}_M^1 - \mathbf{c}, \mathbf{x}_M^2 - \mathbf{c}, \dots, \mathbf{x}_M^{n_M} - \mathbf{c}] \end{aligned} \quad (2.13)$$

ここで $\mathbf{x}_i^j$ は i 番目の学習サンプルの j 番目の学習ウィンドウの画像ベクトル, $\mathbf{X}_i$ は i 番目の学習サンプルの画像ベクトルの集合マトリクスを示している. $\mathbf{c}$ は平均画像ベクトルである.

ここですべての学習ウィンドウ $\mathbf{x}_i^j$ は式 (2.9) と同様に共分散行列 $\mathbf{Q} = \mathbf{X} \mathbf{X}^T$ の固有値問題を解くことによって得られる固有ベクトルを用いて, 低次元の空間 (固有空間) ϕ_i^j へ投影される.

2.2.2 入力ウィンドウの対応付け

入力画像にも学習サンプルと同様に局所ウィンドウを設け, これを入力ウィンドウと呼ぶ. 入力マトリクス $\mathbf{Y}$ は,

$$\mathbf{Y} = [\mathbf{y}^1 - \mathbf{c}, \mathbf{y}^2 - \mathbf{c}, \dots, \mathbf{y}^{n_{in}} - \mathbf{c}]. \quad (2.14)$$

で求められる. このすべての入力ウィンドウ $\mathbf{y}^k$, ($k = 1, 2, \dots, n_{in}$) を固有空間へ投影し, 固有空間上での点 $\mathbf{z}^k$, ($k = 1, \dots, n_{in}$) と学習ウィンドウ ϕ_i^j とのユークリッド距離を求め, 各入力ウィンドウについてもっとも近接した点を検索し, その点に対応する学習ウィンドウを入力ウィンドウに最も類似していると判断する. ここで, 最も類似したウィンドウをその入力ウィンドウの対応ウィンドウと呼ぶ.

2.2.3 投票

各入力ウィンドウについて対応ウィンドウを検索し, 学習サンプル番号 i ($i = 1, \dots, M$) 毎に分類する. 各学習サンプル毎に投票空間を構成し, 入力画像中の入力ウィンドウの座標 (入力画像の右上を原点として入力ウィンドウ中心の座標) とその入力ウィンドウの対

応ウィンドウの学習サンプル中の座標（学習画像の右上を原点として学習ウィンドウ中心の座標）の差 j に投票する. これより, 入力画像内でランドマークがあると思われる位置 j^* に最も投票され, 最も投票数の多い番号に対応する学習サンプルの画像が入力画像に対応し, その時の入力画像内のランドマークの位置は j^* であると判断する.

第 3 章

学習ウィンドウの設定基準

局所固有空間法は、一般に、学習ウィンドウ数を増加により認識精度は上がるが、認識時の処理量も増加する。一方で、学習ウィンドウ数を削減すると認識時の処理時間は短くなるが精度は悪くなる。このようなトレードオフの問題に対処するには、認識に有効な箇所にものみ学習ウィンドウを設定する必要がある。そこで、そのための基準を設け、その有効性を検討する。

3.1 学習ウィンドウの設定基準

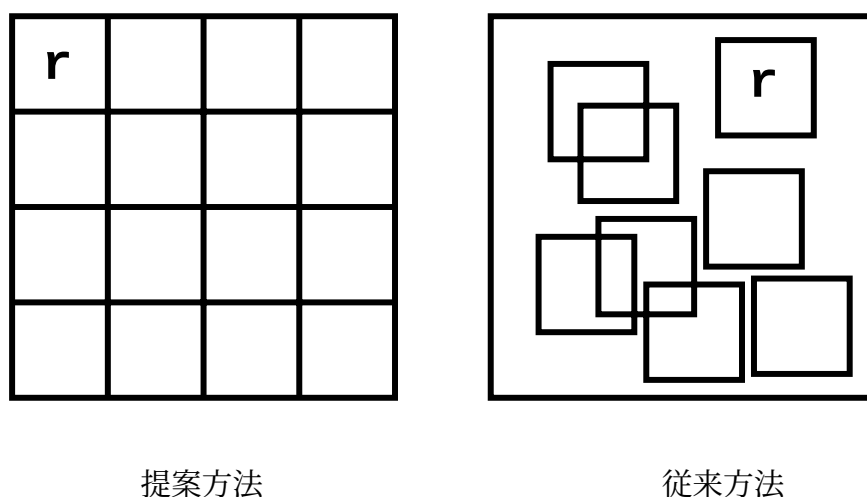


図 3.1: 入力ウィンドウ

入力ウィンドウに対する対応ウィンドウの検索時間の軽減を目指し、入力ウィンドウ・

学習ウィンドウの設定法について検討する.

まず, 入力画像に関しては, 画像が入力される毎に入力ウィンドウ R の設定箇所を選択するのではなく, 図 3.1 のように画一的に決定しておく.

学習ウィンドウに関しては, 対応ウィンドウの検索回数を減少させるために, 可能な限り, 辞書に登録する学習ウィンドウを少なくする. しかし, 学習ウィンドウの削減の仕方によっては, 正確にランドマークを認識できなくなる. そこで, 適切な学習ウィンドウの設定基準を以下のように定めた.

1. 学習画像中で特徴的箇所に局所ウィンドウを設定する.
2. 局所ウィンドウに含まれる特徴が, 局所ウィンドウ周辺であまり変化しない局所ウィンドウを選ぶ.
3. 似通っている局所ウィンドウ同士を削除する.

この 3 つの基準を満たした局所ウィンドウを学習ウィンドウとする.

1 つ目の基準は, 認識を容易にする箇所を検出するために設けた基準で, 多方向に輝度値が変化する箇所を特徴であると判断し, 選択することにする [4]. 学習画像内の任意に設

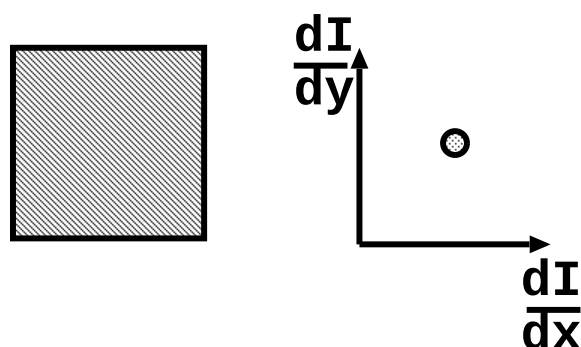


図 3.2: 輝度値の変化なし

定した局所ウィンドウ R 内の位置 $X = (x, y)$ の濃淡値 I から

$$G = \sum_{X \in R} \left(\frac{\partial I}{\partial X} \right) \left(\frac{\partial I}{\partial X} \right)^T \quad (3.1)$$

となる 2×2 マトリクス G を求める. このマトリクスが持つ 2 つの固有値 λ_1, λ_2 は, 次のような意味を持つ. 2 つの固有値ともに小さい場合は, 図 3.2 に示すように, 多方向への輝

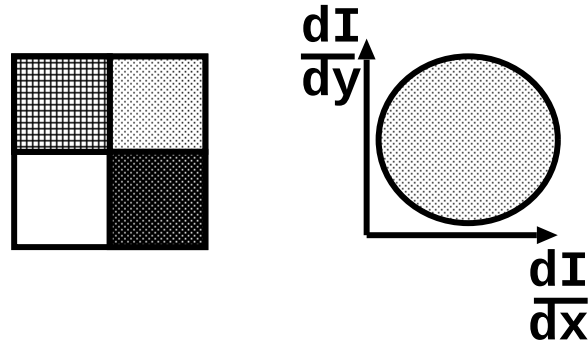


図 3.3: 多方向に輝度値の変化あり

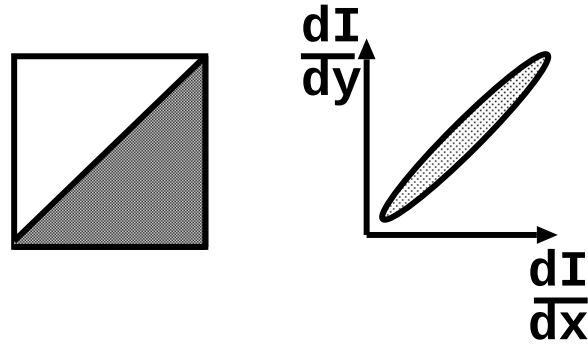


図 3.4: 1 方向のみ輝度値の変化あり

度値の変化が小さく、局所ウィンドウ内の輝度値はほぼ一定であることを表し、逆に両方とも大きい場合は、図 3.3に示すように、多方向への輝度値の変化が大きいことを表している。一方、1つの固有値のみが大きい場合は、図 3.4のように、1 方向への輝度値の変化のみ大きいことを示している。そこで、適当な閾値 λ に対して

$$\min(\lambda_1, \lambda_2) > \lambda \quad (3.2)$$

が成り立つとき、多方向に輝度値が変化しており、特徴を R は含んでいると判断し、選択することにする。ここで、選択された局所ウィンドウを特徴ウィンドウと呼ぶ。

2. は対応ウィンドウを求める際、全く違う学習ウィンドウを対応ウィンドウとして選択することがないように設定した基準である。今、ある特徴ウィンドウ R に対し、局所ウィンドウサイズを WS とすると、図 3.5に示すように、 R の周辺領域（ R を中心に $2WS \times 2WS$ 四方）に任意に局所ウィンドウを設置し、その局所ウィンドウを固有空間上に投影する。

周辺領域にそれぞれ設定した局所ウィンドウの固有空間上の点のばらつきが小さい時（分散が小さい時），この R は周辺領域内で安定しているので，この R を選択する．この時，選択された特徴ウィンドウ R の周辺領域に基準 1 で選択された特徴ウィンドウが存在するならば，その特徴ウィンドウは今選択された特徴ウィンドウ R と類似した特徴を含んでいると考えられるので，削除する．

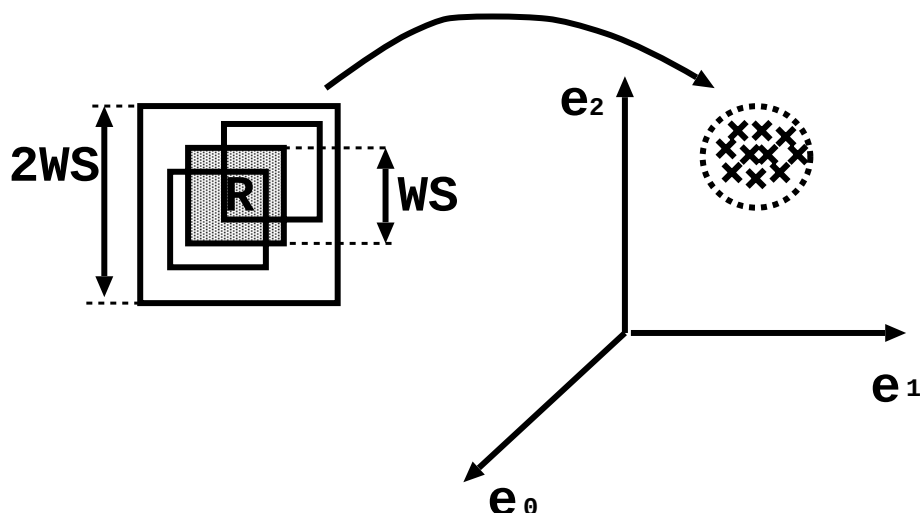


図 3.5: ずれに強い領域

3. は，誤認識を防ぐために設定した基準である．設定基準 1 及び 2 で選択された局所ウィンドウ内に含まれる特徴が固有空間中で一意に定まる保証はない．そこで，任意の 2 つの固有空間上の点のユークリッド距離を求め，距離が短い局所ウィンドウ同士は類似していると判断し，このような局所ウィンドウ同士を取り除く．

この 3 つの基準を満たす局所ウィンドウを学習ウィンドウとして辞書に登録することにする．

この学習ウィンドウ設定基準に基づいて学習サンプルに学習ウィンドウを設定した例を示す．学習サンプルとして図 3.6 のランドマークを用いる．画像サイズは 291×102 ，局所ウィンドウ（学習ウィンドウ）のサイズは 13×13 である．結果は図 3.7 から図 3.10 に示す．

図 3.7 において，2 方向に輝度値が変化している領域に局所ウィンドウを設定した．ここで選択された局所ウィンドウ数は 3088 枚である．次に図 3.8 において，周辺領域で分散が小さくなる局所ウィンドウを求めた．ここで選択された局所ウィンドウ数は 2953 枚である．図 3.9 において，重なり合う局所ウィンドウを取り除いた．ここで選択された局所ウィ

ンドウ数は 101 枚である。そして図 3.10において, 類似ウィンドウを削除し, これを学習ウィンドウとする。学習ウィンドウ数は 58 枚である。



図 3.6: ランドマーク

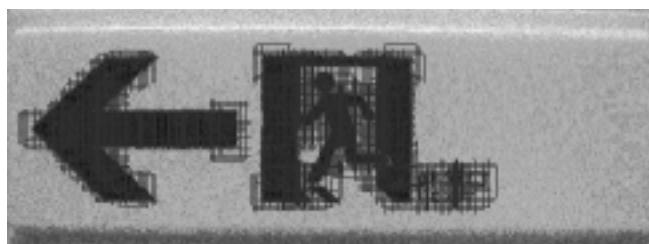


図 3.7: 特徴ウィンドウ 3088 枚

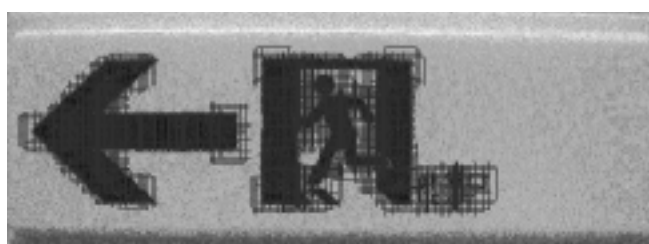


図 3.8: 分散の小さい局所ウィンドウ 2953 枚

3.2 実験

学習ウィンドウ設定基準によって選択された学習ウィンドウを辞書として蓄え, これを用いて入寮画像内のランドマークの位置推定を行なう。



図 3.9: 重なり合う局所ウィンドウの削除 101 枚

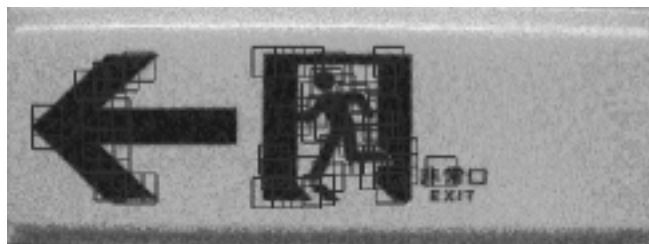


図 3.10: 学習ウィンドウ 58 枚

3.2.1 辞書作成

カメラとランドマークの距離を $750mm$, $900mm$, $1500mm$ としてランドマークを撮影し学習画像とする。学習画像サイズはそれぞれ 376×132 , 291×102 , 167×59 である。学習ウィンドウ設定基準に基づき学習ウィンドウを設定する。学習ウィンドウのサイズは 13×13 である。結果を図 3.11~3.13 に示す。選択された学習ウィンドウ数はそれぞれ 70 枚, 58 枚, 38 枚であり、選択された学習ウィンドウを固有空間へ投影し、それを辞書とする。

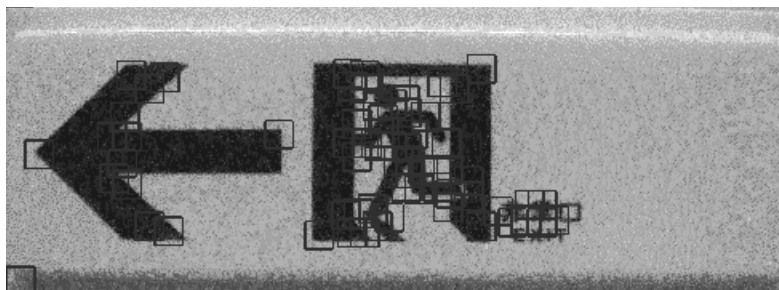


図 3.11: パラメータ $750mm$ で選択された学習ウィンドウ

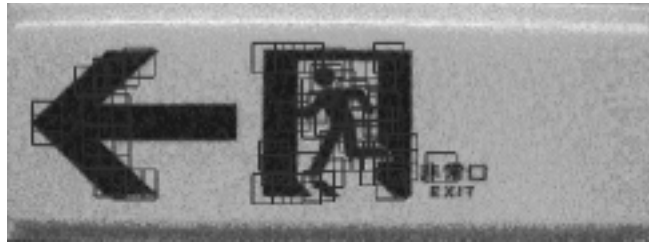


図 3.12: 距離 900mm で選択された学習ウィンドウ



図 3.13: 距離 1500mm で選択された学習ウィンドウ

3.2.2 実験

学習ウィンドウ設定基準に基づいて学習ウィンドウを登録した辞書を用いてランドマークの位置・パラメータ（カメラとランドマークの距離）推定を行なう。

入力画像として図 3.14を用いた時の投票結果を図 3.15～3.17, 表 3.1に示す。図 3.11の (130, 117) が最も投票数が多いので, ランドマークとカメラの距離は 750mm, 入力画像内のランドマークの位置は (130, 117) と判断する。ここでランドマークの位置は, 画像の左上を原点とした, ランドマークの左上の座標を指す。今, 位置推定結果のずれは解像度を下げていることが原因であり, ここでは, 正しく推定できているといえる。

同様に図 3.12, 図 3.13の投票結果を図 3.19～3.21, 表 3.2, 図 3.23～3.25, 表 3.3にそれぞれ示す。これより, 図 3.12はカメラとランドマークの距離は 900mm, 入力画像内のランドマークの位置は (182, 104), 図 3.13はカメラとランドマークの距離は 1500mm, 入力画像内のランドマークの位置は (182, 143) と推定される。

また, ランドマークとカメラの距離を一定にし, 入力画像内のランドマークの位置を様々なに変化させた。この時のランドマークの入力画像内の位置・パラメータ推定の結果を表 3.4～3.6に示す。

ほぼ正しく位置・パラメータ推定できたが, パラメータ 900mm や 1500mm において, パラメータの間違いや実際のランドマークの位置と推定されたランドマークの位置の間

表 3.1: パラメータ $750mm$ の入力画像の投票結果

投票空間	投票結果	
	位置	最大投票数
$750mm$	(130, 117)	13.4
$900mm$	(91, 104)	5.7
$1500mm$	(91, 104)	1.9

表 3.2: パラメータ $900mm$ の入力画像の投票結果

投票空間	投票結果	
	位置	最大投票数
$750mm$	(195, 52)	5.8
$900mm$	(182, 104)	7.7
$1500mm$	(195, 169)	3.8

にずれが生じた.

次に遮蔽のある画像のパラメータ・位置推定を行なう.

図 3.26~3.28 の入力画像に対する推定結果を表 3.26 に示す.

位置推定に関しては, パラメータ $750mm$, $900mm$ の入力画像に関しては, ほぼ正しく推定できていると言える. しかし, パラメータ $1500mm$ の入力画像に関しては実際のランドマークの位置と推定されたランドマークの位置の間にはずれが生じた. また, パラメータも誤認識を引き起こし, $750mm$ と推定された.

表 3.3: パラメータ $1500mm$ の入力画像の投票結果

投票空間	投票結果	
	位置	最大投票数
$750mm$	(169, 26)	3.8
$900mm$	(156, 117)	1.9
$1500mm$	(182, 143)	3.9

表 3.4: パラメータ $750mm$ の推定結果

ランドマークの位置	推定結果	
	位置	パラメータ
(10, 10)	(0, 0)	$750mm$
(30, 40)	(26, 39)	$750mm$
(60, 70)	(52, 52)	$750mm$
(100, 100)	(91, 78)	$750mm$

表 3.5: パラメータ $900mm$ の推定結果

ランドマークの位置	推定結果	
	位置	パラメータ
(10, 10)	(0, 0)	$900mm$
(30, 40)	(13, 0)	$750mm$
(60, 70)	(52, 52)	$900mm$
(100, 100)	(117, 91)	$750mm$

これより $750mm$, $900mm$ に関しては, ある程度の遮蔽に対処できることがわかる.

ランドマークがさらに隠されている図 3.29~3.31 を入力画像として推定を行ない, 結果を表 3.8 に示す. あまり良い結果を得ることはできなかった. これより, 特徴領域がほぼ遮蔽されていると, 正しい認識が行なえないことが確認された.

表 3.6: パラメータ $1500mm$ の推定結果

ランドマークの位置	推定結果	
	位置	パラメータ
(10, 10)	(13, -65)	$1500mm$
(30, 40)	(-13, 26)	$750mm$
(60, 70)	(53, 52)	$1500mm$
(100, 100)	(91, 78)	$1500mm$

表 3.7: 遮蔽のある入力画像の推定結果

入力画像		推定結果	
位置	パラメータ	位置	パラメータ
(193, 157)	$750mm$	(182, 143)	$750mm$
(165, 229)	$900mm$	(156, 221)	$900mm$
(166, 258)	$1500mm$	(130, 78)	$750mm$

表 3.8: 遮蔽のある入力画像の推定結果

入力画像		推定結果	
位置	パラメータ	位置	パラメータ
(68, 197)	$750mm$	(13, 13)	$750mm$
(135, 99)	$900mm$	(131, 166)	$1500mm$
(127, 287)	$1500mm$	(40, 114)	$750mm$



図 3.14: 距離 $750mm$ の入力画像, ランドマークの位置 (133,129)

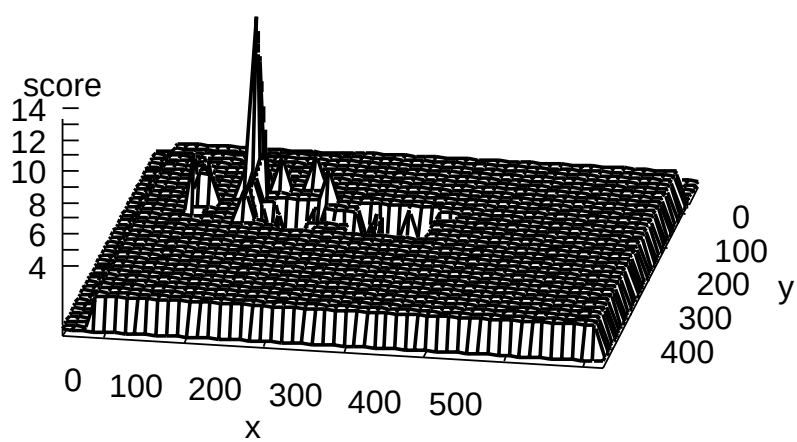


図 3.15: 距離 $750mm$ の入力画像の $750mm$ の投票空間

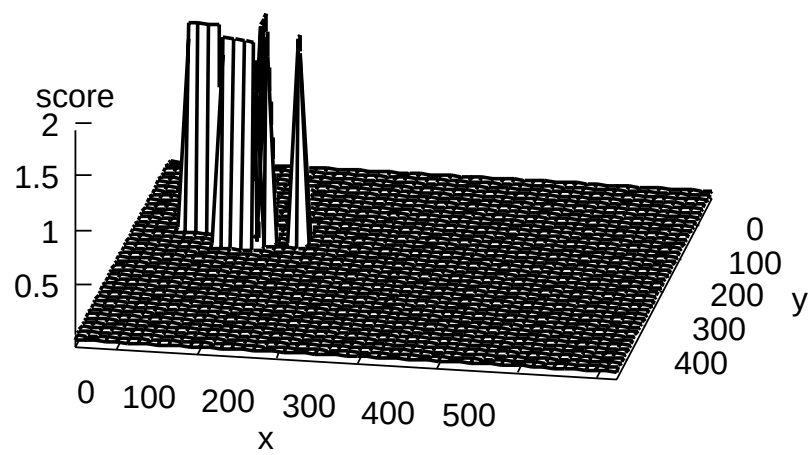


図 3.16: 距離 750mm の入力画像の 900mm の投票空間

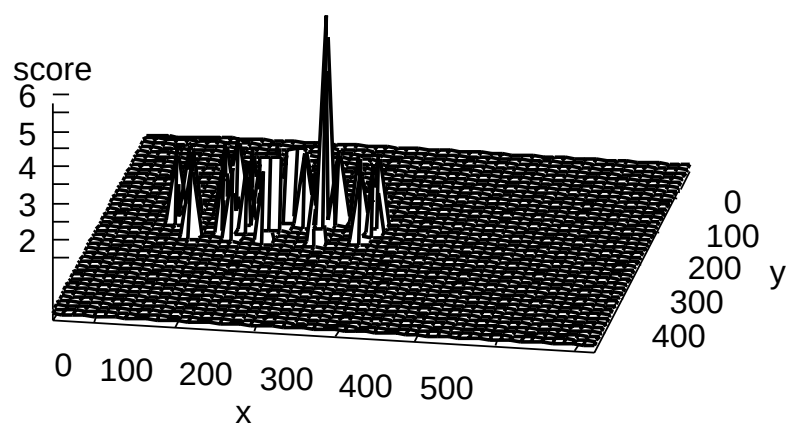


図 3.17: 距離 750mm の入力画像の 1500mm の投票空間



図 3.18: 距離 $900mm$ の入力画像, ランドマークの位置 (196,119)

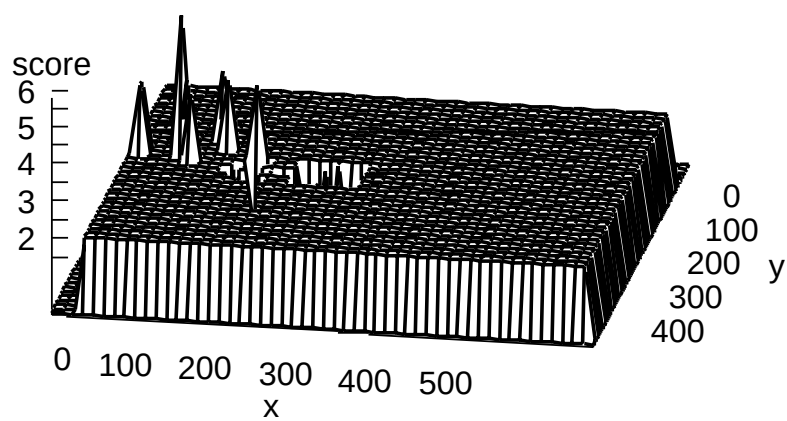


図 3.19: 距離 $900mm$ の入力画像の $750mm$ の投票空間

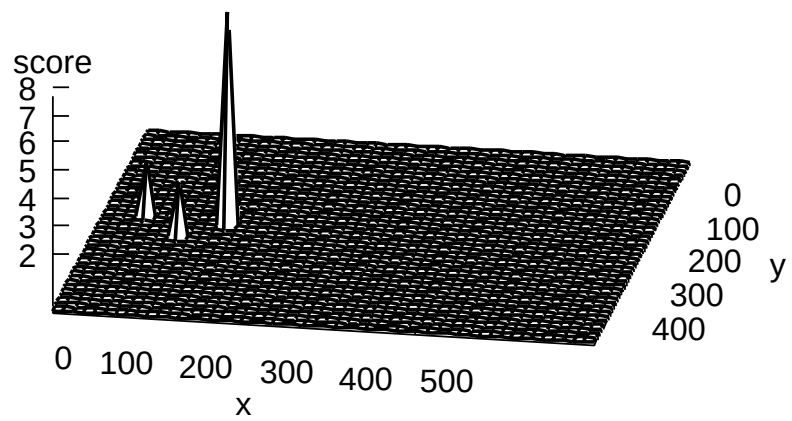


図 3.20: 距離 $900mm$ の入力画像の $900mm$ の投票空間

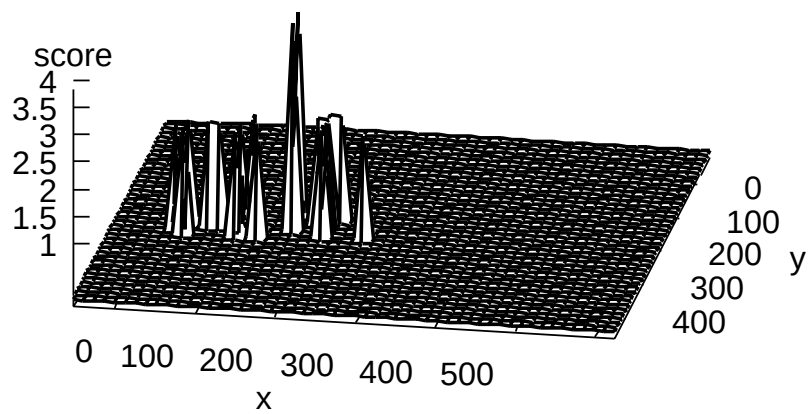


図 3.21: 距離 $900mm$ の入力画像の $1500mm$ の投票空間



図 3.22: 距離 $1500mm$ の入力画像, ランドマークの位置 (180,204)

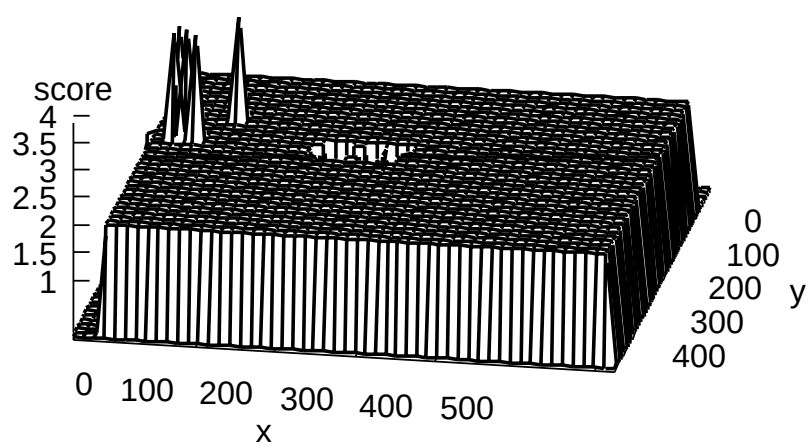


図 3.23: 距離 $1500mm$ の入力画像の $750mm$ の投票空間

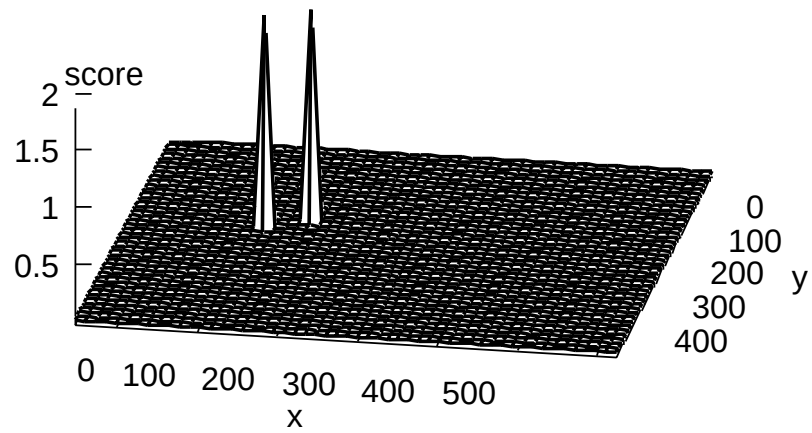


図 3.24: 距離 1500mm の入力画像の 900mm の投票空間

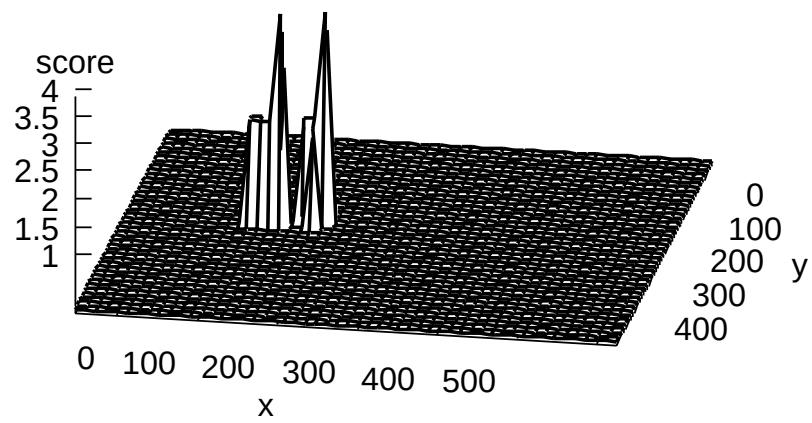


図 3.25: 距離 1500mm の入力画像の 1500mm の投票空間



図 3.26: 遮蔽のあるパラメータ $750mm$ の入力画像



図 3.27: 遮蔽のあるパラメータ $900mm$ の入力画像



図 3.28: 遮蔽のあるパラメータ $1500mm$ の入力画像



図 3.29: 遮蔽のあるパラメータ $750mm$ の入力画像



図 3.30: 遮蔽のあるパラメータ $900mm$ の入力画像

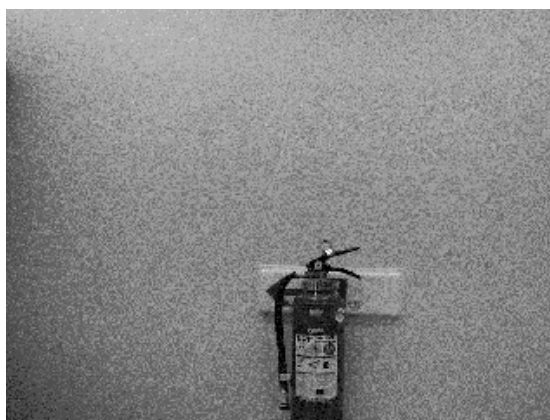


図 3.31: 遮蔽のあるパラメータ $1500mm$ の入力画像

第 4 章

階層的辞書構成

学習ウィンドウ設定基準に基づいて学習ウィンドウを選択し、画像内のランドマークの存在位置をほぼ正しく推定することができた。しかし学習ウィンドウ設定基準には経験的・実験的値を多く含んでいた。そこで経験的・実験的値を減らし、さらに、固有空間法を使用する上で効果的な箇所に学習ウィンドウを設定するために、学習画像の固有空間上の値を用いてクラスタリングし、類似領域を削除した特徴的箇所に学習ウィンドウを設定する。また、認識時の対応ウィンドウの検索回数を削減するために、学習画像を階層的にクラスタリングし、木状に辞書を構成する。

4.1 クラスタリングに基づいた学習ウィンドウの選択

4.1.1 クラスタリング

R 個のデータ $\mathbf{x}_i$ ($i = 1, \dots, R$) に対し、全データの平均ベクトル $\mathbf{h}_T$ 、分散 σ_T^2 は

$$\mathbf{h}_T = \frac{1}{R} \sum_{i=1}^R \mathbf{x}_i \quad (4.1)$$

$$\sigma_T^2 = \frac{1}{R} \sum_{i=1}^R \|\mathbf{x}_i - \mathbf{h}_T\|^2 \quad (4.2)$$

で求まる。この R 個のデータを M 分割し、各クラスを $C_1, C_2, \dots, C_M$ とする。分割位置を $k_1, k_2, \dots, k_M$ とすると、

$$1 \leq k_1 \leq k_2 \leq \dots \leq k_M \leq R \quad (4.3)$$

$$C_m = [k_m, k_{m+1} - 1] \quad (m = 1, \dots, M) \quad (4.4)$$

となる. このとき, クラス C_m の生起確率 ω_m , 平均値 $\mathbf{h}_m$, 分散 σ_m^2 はそれぞれ

$$\omega_m = \frac{1}{R} \sum_{r \in C_m} 1 \quad (4.5)$$

$$\mathbf{h}_m = \frac{1}{R\omega_m} \sum_{r \in C_m} \mathbf{x}_r \quad (4.6)$$

$$\sigma_m^2 = \frac{1}{R\omega_m} \sum_{r \in C_m} \|\mathbf{x}_r - \mathbf{h}_m\|^2 \quad (4.7)$$

で求まる.

判別分析において, 分割箇所の良さを評価する判別基準には次の 3 つがある.

$$\lambda = \frac{\sigma_B^2}{\sigma_W^2}, \quad \kappa = \frac{\sigma_T^2}{\sigma_W^2}, \quad \eta = \frac{\sigma_B^2}{\sigma_T^2} \quad (4.8)$$

ただし,

$$\sigma_W^2 = \sum_{m=1}^R \omega_m \sigma_m^2 \quad (4.9)$$

$$\sigma_B^2 = \sum_{m=1}^R \omega_m \|\mathbf{h}_m - \mathbf{h}_T\|^2 \quad (4.10)$$

はそれぞれ, クラス内分散およびクラス間分散である. σ_B^2 , σ_W^2 , および σ_T^2 の間には分割位置 k_m に関わらず,

$$\sigma_B^2 + \sigma_W^2 = \sigma_T^2 \quad (4.11)$$

の関係が常に成り立つ. つまり, λ , κ , η はすべて互いに同値である. また, σ_T^2 は与えられており, k_m によらず一定である. σ_B^2, σ_W^2 はともに k_m によって決定する値であるが, σ_B^2 は 1 次の統計量のみで求められるのに対して, σ_W^2 は 2 次の統計量も求める必要がある. 従って, η を採用するのが最も効率的であることがわかる. つまり,

$$\sigma_B^2(k^*) = \max_{k \in R} \sigma_B^2(k) \quad (4.12)$$

から最適な分割位置 k^* が求まる.

4.1.2 学習ウィンドウ設定箇所の選択

図 4.1, 4.2 のように, 学習サンプルに局所ウィンドウの x , y 方向サイズ ω の幅をもつ帯状の領域を構成する. ある一定間隔 W ごとに帯状の領域を設定し, 帯状の領域 A_i ($i = 1, \dots, n_x$), B_j ($j = 1, \dots, n_y$) を構成する. A_i , B_j 内には W ごとに学習ウィンドウを設

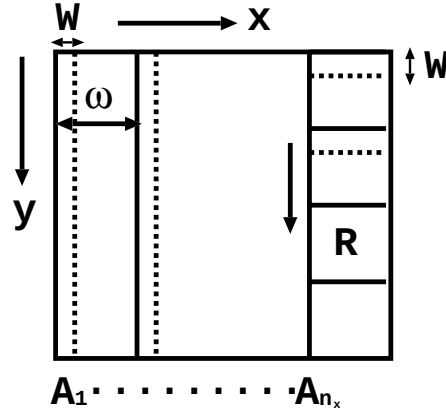


図 4.1: x 軸方向への学習サンプルの分割

定する. A_i を判別基準 η に基づき 2 つの領域に分割する. 各領域に含まれる学習ウィンドウの数がそれぞれ S より多ければ再度 2 分割する. 逆に S より少なくなった領域は分割を終了する. すべての領域に含まれる局所ウィンドウの数が S より少なくなるまで階層的に分割を繰り返す. 求めた各領域に含まれる学習ウィンドウの平均値で代表させ, 各代表値のユークリッド距離が短い領域同士は識別が困難のため, 取り除く.

B_i についても同様の処理を行なう.

A, B 両方の分割方法でも残った領域を特徴箇所として用いることにする.

この方法によって抽出された特徴箇所は, 固有空間上での点が特徴的である箇所を特徴としてとらえており, 輝度値が多方向に変化している箇所ではないが, 類似領域を取り除くことでほぼ同等の領域が抽出されると考えられる.

4.2 階層的クラスタリング

探索回数を減少させるために辞書の構成を各節を固有空間上の点の平均値とした木構造にする.

図 4.3のように, 特徴領域に帯状の領域 C_i ($i = 1, \dots, w$) を設定する. この C_i を階層的に分割し, 各帯状の領域ごとに木を構成する.

分割箇所は判別基準 η を用いて C_i を m 分割する.

分割数 m の決定法は以下の通りである.

あるクラス C_i に含まれる局所ウィンドウの固有空間上の点 ϕ_j ($j = 1, \dots, n$) を判別基準 η を用いて, 分割数 $m = 2$ における最適な分割箇所 k_2 を求める. k_2 で分割した各領域

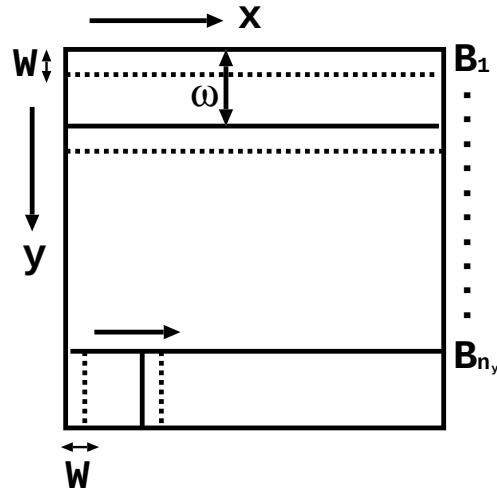


図 4.2: y 軸方向への学習サンプルの分割

$C_{i,1}$, $C_{i,2}$ の代表値 $\hat{\phi}_{i,1}$, $\hat{\phi}_{i,2}$ を計算する. これは, $C_{i,1}$, $C_{i,2}$ それぞれに含まれる学習ウィンドウの固有空間上の点 ϕ_j ($j = 1, \dots, k$), ϕ_j ($j = k + 1, \dots, w$) の平均値で表す. ここで領域 $C_{i,1}$ に含まれる学習ウィンドウは $C_{i,1}$ の代表値に最も類似しており, 領域 $C_{i,2}$ に含まれる学習ウィンドウは $C_{i,2}$ の代表値に最も類似していれば, 認識時に正しく木を辿ることができる. ここで類似度としてユークリッド距離を用いることにし, 距離の短いもの同士を類似していると判断する. そこで, すべての学習ウィンドウがそれぞれ含まれる領域の代表値と最も類似しており, 正しく分類できれば, その時の分割数 $m = 2$ を採用し, その時の分割箇所 k_2 で領域 C_i を分割することにする. しかし, すべての学習ウィンドウを正しく分類できなかった時は分割数 m を増やして再度分割を行ない, 正しく認識できるか調べる. 正しく認識できるまで m を増加させて, 分割を繰り返す. つまり, すべてのウィンドウが正しく分類される最小の分割数が求める分割数 m となる.

新たに得られた領域 $C_{i,m}$ を再度最適分割数を求め, 分割し, 分割した領域の代表値を木の節に持たせる. 1つの領域の学習ウィンドウが1枚になるまで階層的に分割を繰り返す.

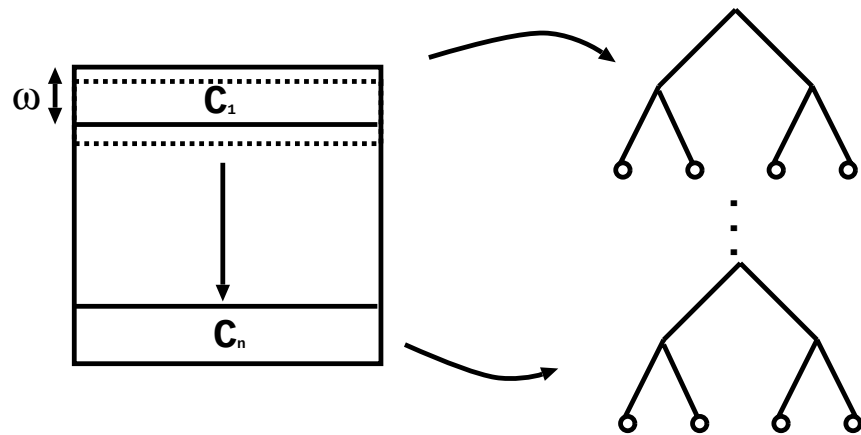


図 4.3: 木構造

4.3 実験

4.3.1 学習ウィンドウ設定箇所の選択

図 3.6 のランドマークとカメラの距離を $750mm$, $900mm$, $1500mm$ として撮影し学習画像とする。各学習画像から特徴箇所を抽出し、その結果を図 4.4～4.6 に示す。学習ウィンドウ数は $750mm$ で 588 枚, $900mm$ で 520 枚, $1500mm$ で 315 枚となった。それぞれ得



図 4.4: 学習ウィンドウ設定箇所 ($750mm$)

られた特徴画像をクラスタリングし、木構造の辞書を作成する。



図 4.5: 学習ウィンドウ設定箇所 (900mm)



図 4.6: 学習ウィンドウ設定箇所 (1500mm)

4.3.2 実験

ランドマークとカメラの距離をパラメータとして位置・パラメータ推定結果を表 4.1～4.3に示す。辞書を木状に構成した時と、全探索行なった時の両方の結果を示す。ただし、全ウィンドウ数は1423枚である。構成された木の本数は62本である。ここで、各入力画像は3章と同様の画像を用いた。

表 4.1: パラメータ 750mm の推定結果

入力画像の ランドマークの位置	推定結果 (木構造)				推定結果 (全探索)		
	位置	パラメータ	照合回数	投票数	位置	パラメータ	投票数
(10, 10)	(0, 0)	750mm	938.5	37.5	(0, 0)	750mm	37.5
(30, 40)	(26, 39)	750mm	938.0	43.7	(26, 39)	750mm	43.7
(60, 70)	(52, 65)	750mm	939.0	43.5	(52, 65)	750mm	45.5
(100, 100)	(91, 91)	750mm	939.4	37.7	(91, 91)	750mm	39.6
(133, 129)	(130, 130)	750mm	938.5	21.8	(130, 130)	750mm	23.8

解像度を下げているため、厳密にランドマークの位置を推定することはできなかったが、正しく位置・パラメータ推定を行なえたといえる。さらに、探索回数を1423から平均940.7に減少させることができた。

遮蔽のある画像（図 3.26～3.31）についても同様に推定結果を示す（表 4.4～4.5）。これより、ある程度の遮蔽には対処できるが、大部分特徴領域が隠されると、認識できない

表 4.2: パラメータ $900mm$ の推定結果

入力画像の ランドマークの位置	推定結果 (木構造)				推定結果 (全探索)		
	位置	パラメータ	照合回数	投票数	位置	パラメータ	投票数
(10, 10)	(0, 0)	$900mm$	939.8	31.7	(0, 0)	$900mm$	33.6
(30, 40)	(26, 39)	$900mm$	940.2	27.6	(26, 39)	$900mm$	27.6
(60, 70)	(52, 65)	$900mm$	940.1	29.7	(52, 65)	$900mm$	29.7
(100, 100)	(91, 91)	$900mm$	940.3	29.8	(91, 91)	$900mm$	29.8
(196, 119)	(182, 104)	$900mm$	942.1	9.8	(182, 104)	$900mm$	9.8

表 4.3: パラメータ $1500mm$ の推定結果

入力画像の ランドマークの位置	推定結果 (木構造)				推定結果 (全探索)		
	位置	パラメータ	照合回数	投票数	位置	パラメータ	投票数
(10, 10)	(0, 0)	$1500mm$	942.6	17.7	(0, 0)	$1500mm$	19.7
(30, 40)	(26, 39)	$1500mm$	942.4	21.8	(26, 39)	$1500mm$	21.8
(60, 70)	(52, 65)	$1500mm$	942.6	27.7	(52, 65)	$1500mm$	27.7
(100, 100)	(91, 91)	$1500mm$	942.6	21.8	(91, 91)	$1500mm$	23.8
(180, 204)	(169, 208)	$1500mm$	944.3	9.8	(169, 208)	$1500mm$	9.8

ことがわかる.

表 4.4: 遮蔽のある入力画像の推定結果

入力画像		推定結果（木構造）				推定結果（全探索）		
位置	パラメータ	位置	パラメータ	照合回数	投票数	位置	パラメータ	投票数
(193, 157)	750mm	(182, 143)	750mm	927.0	21.7	(182, 143)	750mm	19.7
(165, 229)	900mm	(156, 221)	900mm	931.6	21.7	(156, 221)	900mm	21.7
(166, 258)	1500mm	(156, 260)	1500mm	938.3	21.5	(156, 260)	1500mm	21.5

表 4.5: 遮蔽のある入力画像の推定結果

入力画像		推定結果（木構造）				推定結果（全探索）		
位置	パラメータ	位置	パラメータ	照合回数	投票数	位置	パラメータ	投票数
(68, 197)	750mm	(156, 182)	750mm	928.2	6.0	(156, 182)	750mm	6.0
(196, 119)	900mm	(13, -13)	750mm	935.1	6.0	(13, -13)	750mm	6.0
(180, 204)	1500mm3	(0, 169)	750mm	940.1	7.9	(0, 169)	750mm	7.9

第 5 章

まとめ

本稿では局所固有空間法を用いて入力画像内のランドマークの存在位置の認識を行なった。認識時の処理の軽減を目指して、認識精度を低下させることなく学習ウィンドウを削減するための設定基準について検討し、更に、学習画像を判別分析を用いて階層的にクラスタリングし、辞書を木状に構成することにより、探索の効率化を図った。

精度を下げずに探索の効率化を図るために学習ウィンドウ設定基準を設け、それに基づいて辞書を作成した。入力画像内のランドマークの存在位置を推定する実験を行ない、ほぼ正しい推定結果を得ることができ、設定基準の有効性を確認した。しかし、入力画像内のランドマーク位置と推定結果のランドマークの位置の間のずれや、パラメータ推定における誤認識を引き起こすものが多少含まれた。学習ウィンドウ設定基準には経験的・実験的値を多く含んでおり、この値の決定の仕方によって、辞書に登録される学習ウィンドウは大きく変化し、推定精度にも影響を及ぼす。この経験的・実験的値を更によく検討することによって、より良い推定結果を得られると考えられる。

一方、学習ウィンドウを設定する際、経験的・実験的値をできるだけ減らし、固有空間法を使用する上でより効果的な箇所に学習ウィンドウを設定するために、学習画像の固有空間上の値を用いてクラスタリングし、類似領域を削除した特徴的箇所に学習ウィンドウを設定した。そして学習画像を階層的にクラスタリングすることによって、辞書を木状に構成した。この時、階層構造の各段階において間違えることなく認識できるよう、学習ウィンドウをクラスタリングし、精度を落さず認識時の処理量の削減を図った。入力画像内のランドマークの存在位置を推定する実験を行ない、1423 枚の学習ウィンドウに対して約 940 回の比較で対応ウィンドウを探索することができた。そして、ランドマークが大きく遮蔽されない限り、ほぼ正しく入力画像内のランドマークの存在位置の推定ができた。

今後の課題として、更に探索回数を減少させるために辞書の構成を検討し、実際にロボッ

トの自己位置推定を行なうことがあげられる.

参考文献

- [1] 村瀬 他, “2 次元照合による 3 次元物体認識ーパラメトリック固有空間法”, 信学論 D-II, vol.J77-D-II,no.11,pp.2179-2187,1994
- [2] 大場 他, “局所固有空間手法による金属物体の安定認識”, 信学論 D-II, Vol.J77-D-II,no.4,pp.1933-1940,1994
- [3] 小迫,A.Ralescu, 武川,” ベイズ推定に基づく剛体の局所情報の統合ーパラメトリック固有空間表現によるパラメータからの物体抽出”, 信学技法,PRMU95-75,1995.
- [4] C.Tomasi,et al., “Shape and Motion from Image Streams:a Factorization Method-Part 3.detection and tracking of point features”, CMU-CS-91-132.Carnegie Mellon University,Jan.1991
- [5] H.Murakami and B.V.K.V.Kumar,”Efficient calculation of primary images from a set of images”,IEEE Trans.Pattern Anal. & Machine Intell.,vol.4,no.5,pp.511-515,Sep.1982.
- [6] H.Murase and S.K.Nayar,”Visual learning and recognition of 3-d objects from appearance”,Int.J.Comput.Vision,**14**,1,pp.5-24 1995.
- [7] 大津, 栗田, 関田,” パターン認識”, 朝倉書店,1996.
- [8] T.Kanbara,et al., “Selection of Efficient Landmarks for an Autonomous Vehicle”, Proc.IROS'93, Vol.2,pp.1332-1338,1993

謝辞

本研究を行なうにあたり、御指導, 御鞭撻をいただいた北陸先端科学技術大学院大学 阿部 亨 助教授に厚くお礼申し上げます。

また、様々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 岡本 栄司 助教授には、サブテーマで熱心に御指導いただき、深く感謝申し上げます。

また、日頃よりお世話になった当研究室の皆様に深く感謝致します。