



子・348

領域・空間情報を表現する
グラフ構造を用いた
類似画像検索

北 本 朝 展

目次

1	序論	12
1.1	画像データベースの内容検索	12
1.1.1	はじめに	12
1.1.2	キーワード検索方式	12
1.1.3	新たな検索方式	13
1.2	類似画像検索	13
1.2.1	類似画像検索の課題	13
1.2.2	代表的な類似検索システム	14
1.2.3	類似画像検索システムの問題点	15
1.3	本研究の方針	16
1.4	本研究の目的および成果	16
1.5	本論文の構成	17
2	本研究の対象画像および本研究が用いるモデル	19
2.1	本研究が対象とする画像データベース	19
2.2	衛星画像データベース	20
2.2.1	気象衛星 NOAA および気象衛星ひまわり	20
2.2.2	NOAA 衛星画像データベースの対象	21
2.2.3	GMS 衛星画像データベースの対象	22
2.2.4	衛星画像データベースの諸問題	23
2.2.5	衛星画像の内容とは	24
2.2.6	衛星画像の雲パターンの類似検索	25
2.3	カラー自然画像データベース	25
2.3.1	カラー自然画像データベースの対象	26
2.3.2	カラー自然画像データベースの内容	26
2.4	階層モデル	27
2.4.1	階層モデルの利点	27
2.4.2	各レイヤの概説	28
2.4.3	階層モデルと本研究で用いた手法の対応関係	30
3	L型やJ型・U型の形状をもつミクセルの面積占有率密度	31
3.1	第3章から第5章までの関係	31
3.2	概要	31

3.3	従来研究の問題点	31
3.4	面積占有率密度	32
3.5	正方形モデルから導出した面積占有率密度	33
3.5.1	画素と正方形モデルの位置関係で場合わけした面積占有率密度	34
3.5.2	それぞれの場合の生起確率	36
3.5.3	サイズ密度によるランダム化	37
3.6	十字形モデルから導出した面積占有率密度	38
3.6.1	画素と十字形モデルの位置関係で場合わけした面積占有率密度	38
3.6.2	それぞれの場合の生起確率	46
3.6.3	サイズ密度によるランダム化	46
3.7	フラクタル合成画像から計算した面積占有率密度	46
3.8	4種のアラビア数字の面積占有率密度の比較	47
3.9	まとめ	49
4	期待値に基づいたミクセルの面積占有率の推定	50
4.1	概要	50
4.2	従来研究の問題点	50
4.3	問題の定式化	52
4.3.1	ミクセルの線形モデル	52
4.4	ミクセルの面積占有率の推定	53
4.4.1	面積占有率を固定した場合の同時密度関数	53
4.4.2	同時密度関数の計算方法	54
4.4.3	面積占有率密度	55
4.4.4	ミクセルの期待面積占有率の計算	56
4.4.5	デルタ関数の場合の期待面積占有率	56
4.5	正規密度の場合の期待面積占有率	57
4.5.1	1バンド2クラスミクセルの期待面積占有率	58
4.5.2	1バンド3クラスミクセルの期待面積占有率	59
4.5.3	2バンド2クラスミクセルの期待面積占有率	59
4.5.4	2バンド3クラスミクセルの期待面積占有率	61
4.5.5	p バンド k クラスミクセルの期待面積占有率	61
4.6	実験	61
4.7	まとめ	62
5	ミクセル密度を含めた混合密度推定と統計的な画像分類への応用	64
5.1	概要	64
5.2	従来研究の問題点	64
5.3	問題の定式化	65
5.3.1	ミクセルの線形モデル	65
5.3.2	ミクセル密度を含む混合密度推定	65
5.4	ミクセル密度	66

5.4.1	ミクセル密度の導出	66
5.4.2	面積占有率密度が一様分布の場合のミクセル密度	66
5.4.3	面積占有率密度が一様分布でない場合のミクセル密度	68
5.4.4	ミクセル密度の形状	68
5.5	EM アルゴリズムを用いた混合密度推定	69
5.5.1	ミクセル密度を含む混合密度推定の特徴	69
5.5.2	EM アルゴリズム	70
5.5.3	ミクセル密度を含む場合の EM アルゴリズム	70
5.5.4	初期値推定	71
5.6	実験	71
5.6.1	シミュレーション	72
5.6.2	リモートセンシング画像を用いた実験	73
5.6.3	各クラスの面積比の計算	76
5.7	まとめ	77
6	領域に基づいた変形モデルを用いた形状分解	78
6.1	概要	78
6.2	形状情報の表現	78
6.2.1	領域特徴量による形状情報の表現	78
6.2.2	領域や輪郭線による形状構造の表現	79
6.3	領域に基づいた変形モデルを用いた形状分解	79
6.3.1	変形モデル	79
6.3.2	形状分解に用いる基本構成要素	80
6.3.3	シルエット画像を対象としたエネルギー関数	80
6.3.4	濃淡画像を対象としたエネルギー関数	81
6.4	実験	83
6.4.1	アルゴリズム	83
6.4.2	初期値推定	83
6.4.3	実験結果	84
6.5	まとめ	87
7	引力モデルを用いた群化と階層化属性付き関係グラフの構築	88
7.1	概要	88
7.2	知覚的群化	88
7.3	仮想引力モデル	89
7.3.1	仮想引力モデルの定式化	89
7.3.2	仮想引力モデルによる関係のネットワーク	90
7.4	階層化属性付き関係グラフ	91
7.5	グラフ構造への構造化	92
7.5.1	仮想引力モデルを用いたグラフ構造への構造化	92
7.5.2	隣接関係を用いたグラフ構造への構造化	92

7.6	まとめ	93
8	A* アルゴリズムを用いたグラフマッチングによる画像検索	94
8.1	概要	94
8.2	グラフマッチング	94
8.3	グラフマッチングコスト	95
8.3.1	シンボルコスト	95
8.3.2	属性コスト	96
8.3.3	ヌルシンボルとの距離	97
8.4	グラフマッチング高速化の基本方針	98
8.5	探索アルゴリズムの比較	98
8.6	認容可能条件を満たすヒューリスティック関数	99
8.6.1	ノードのヒューリスティック関数	99
8.6.2	アークのヒューリスティック関数	100
8.7	最大許容コスト	100
8.8	実験結果	101
8.8.1	実験の概要	101
8.8.2	実験の結果	103
8.9	まとめ	105
9	人為選択法に適した遺伝的アルゴリズムを用いた類似検索規準の学習	106
9.1	概要	106
9.2	類似検索規準の学習	106
9.2.1	想起と認識	106
9.2.2	例示検索とブラウジング	107
9.2.3	問題の定式化	107
9.2.4	オフライン問題	108
9.2.5	オンライン問題	109
9.3	人為選択法	109
9.3.1	人為選択法の特徴と利点	109
9.3.2	人為選択法の問題点	109
9.4	ライフサイクル型遺伝的アルゴリズム	110
9.4.1	世代型遺伝的アルゴリズムからライフサイクル型遺伝的アルゴリズムへ	110
9.4.2	人為選択法に応用するライフサイクル型遺伝的アルゴリズム	111
9.4.3	ライフサイクル型遺伝的アルゴリズムのオペレータなど	113
9.4.4	ライフサイクル型遺伝的アルゴリズムの類似検索規準の学習への応用	115
9.5	実験結果	117
9.5.1	オフライン問題	117
9.5.2	オンライン問題	118
9.6	まとめ	118

10	類似画像検索	120
10.1	概要	120
10.2	階層モデルに基づいた画像データベースシステムの実装	120
10.2.1	NOAA 衛星画像の類似検索システム	121
10.2.2	GMS 衛星画像の類似検索システム	123
10.2.3	カラー自然画像の類似検索システム	125
10.3	類似検索結果	126
10.3.1	NOAA 衛星画像データベースの類似検索結果	127
10.3.2	GMS 衛星画像データベースの類似検索結果	134
10.3.3	カラー自然画像データベースの類似検索結果	138
10.4	類似画像検索の評価	140
10.4.1	検索精度の評価	142
10.4.2	検索速度の評価	144
10.4.3	柔軟性の評価	146
10.5	まとめ	147
11	結論	149
11.1	本研究の成果	149
11.2	今後の課題	150
A	類似度と距離	152
A.1	はじめに	152
A.2	客観的な距離	153
A.3	主観的な距離	154
A.3.1	対数距離	155
A.3.2	ヒューリスティックな距離	155
A.3.3	心理学的尺度構成法	155
A.4	特徴量の重み付け	156
B	スケールスペースを用いたヒストグラム解析	158
B.1	スケール	158
B.1.1	画像とスケール	158
B.1.2	スケールスペース	159
B.1.3	基本的な要請	160
B.2	スケールスペースを用いたヒストグラム解析	160
B.2.1	2次微分のゼロ交差を用いた混合密度の初期値推定	161
B.2.2	1次微分のゼロ交差を用いたヒストグラムの多値化	162
C	ボリュームレンダリングを用いた気象衛星画像の3次元可視化	164
C.1	はじめに	164
C.2	本研究の特徴	164
C.3	ボリューム・ビジュアライゼーション	165

C.3.1	ボリュームデータの表現と内部構造	165
C.3.2	ボリュームデータの可視化手順	166
C.3.3	分類	166
C.3.4	シェーディングモデル	167
C.4	本研究で提案する手法	167
C.4.1	雲のシェーディングモデル	167
C.4.2	海のシェーディングモデル	168
C.5	実験	169
C.5.1	前処理およびデータ分類	169
C.5.2	3次元ボリュームデータへの変換	169
C.5.3	可視化の結果および考察	169
	研究発表	172
	参考文献	174
	謝辞	181

目次

2.1	気象衛星 NOAA および気象衛星ひまわり。	19
2.2	NOAA 衛星画像データベースに蓄積されている画像。	20
2.3	GMS 短期型衛星画像データベースに蓄積されている画像。	21
2.4	GMS 長期型衛星画像データベースに蓄積されている画像。	22
2.5	カラー自然画像データベースに蓄積されている画像。	26
3.1	ミクセルを含む画像の統計的な解析手法に用いられる確率モデルの関係。	32
3.2	2 クラスミクセルのモデル。	33
3.3	画素と正方形モデルの位置関係による 5 個の場合への場合わけ。網かけされた正方形は正方形モデルを表し、もう一方の正方形が画素を表している。	34
3.4	Case 5 の場合の面積占有率密度の計算方法。	35
3.5	サイズ密度として用いる対数正規密度。	37
3.6	正方形モデルから導出した L 型面積占有率密度。	38
3.7	画素と十字形モデルの位置関係による 15 個の場合への場合わけ。網かけされた領域は十字形モデルを表し、もう一方の正方形が画素をあらわしている。	40
3.8	Case 6 の場合の面積占有率密度の計算方法。	41
3.9	Case 7 の場合の面積占有率密度の計算方法。	42
3.10	Case 11 の場合の面積占有率密度の計算方法。	44
3.11	十字形モデルから導出した J 型面積占有率密度。	46
3.12	(a) フラクタル合成画像を 2 値化した画像。(b) U 型面積占有率密度。	47
3.13	3 種の面積占有率密度の比較。L 型と J 型では対数正規密度のパラメータを $\mu = 4, \sigma = 2$ とした。	48
4.1	1 バンド 2 クラスミクセルの期待面積占有率。	58
4.2	1 バンド 3 クラスミクセルの期待面積占有率。	59
4.3	2 バンド 2 クラスミクセルの期待面積占有率。	60
4.4	2 バンド 3 クラスミクセルの期待面積占有率。(a)、(b)、(c) : $\sigma_{ij} = 1, \rho_i = 0$ 、(d)、(e)、(f) : $\sigma_{ij} = 1, \rho_1 = 0.8, \rho_2 = -0.8, \rho_3 = 0.8$ 。	61
4.5	面積占有率のシミュレーションの結果。(a) 面積占有率密度が一様分布の場合、(b) U 型面積占有率密度を用いた場合。	62
5.1	面積占有率密度が一様分布の場合のミクセル密度。	67
5.2	4 種の面積占有率密度を用いた場合の 2 クラスミクセル密度。	68

5.3	2クラスミクセル密度のシミュレーションと混合密度推定結果。各パラメータの値は表 5.1を参照。	71
5.4	気象衛星 NOAA 画像と画像分類結果。画像分類には図 5.5(d) の結果を用いる。	73
5.5	画像ヒストグラムに対する混合密度推定結果。	75
5.6	ミクセルの期待面積占有率。(a)~(d) は図 5.5で用いた記号と同一である。	76
6.1	変形モデルのあてはめ。	81
6.2	本研究で用いるメンバーシップ関数の模式図。	82
6.3	中心位置の初期値推定法。	83
6.4	シルエット画像の形状分解におけるペナルティ定数 p の効果。後の実験では $p = 1$ とした。	85
6.5	濃淡画像の形状分解。楕円の場合と超楕円の場合の比較。	86
7.1	二つの仮想引力モデル。(a) r は輪郭間の最小距離、(b) r は重心間の距離から図形の細長さを考慮した長さを引いた距離。	89
7.2	非連結グラフを接続する二つの方法。	90
7.3	プリミティブを構造化して得られた階層化属性付き関係グラフ。	92
8.1	4個のノードを持つ2個のグラフをマッチングした場合の状態空間。実線が最小コストに対応した経路。	95
8.2	A* アルゴリズム・認容可能条件を満たすヒューリスティック関数・最大許容コストのすべての高速化手法を採り入れたグラフマッチングアルゴリズム。優先順位付き待ち行列 OPEN には、改良した2分探索木を用いている。	102
8.3	条件を変えた場合の検索時間の比較。(a) 三種類の探索方法の比較、(b) ヒューリスティック関数の効果、(c) 最大許容コストの効果、(d) 全数検索と上位検索の比較。	104
9.1	人為選択法の問題点の模式図。	110
9.2	ライフサイクル型遺伝的アルゴリズムの模式図。	110
9.3	ライフサイクル型遺伝的アルゴリズムの模式図。点線を境にインタフェース部とサーバ部の2つの部分に分かれる。	112
9.4	ライフサイクル型遺伝的アルゴリズムによる例示検索・ナビゲーションのインタフェース。左下が評価テーブル。	116
9.5	世代型遺伝的アルゴリズムとライフサイクル型遺伝的アルゴリズムの場合の最大適応度の比較。	117
10.1	NOAA 衛星画像データベースで用いる例示画。	121
10.2	被験者によって選ばれた主観的類似画像 29 枚の中の上位 10 枚の画像。主観的類似度が画像の下の数字で示されている。	122
10.3	GMS 衛星画像データベースで用いる例示画。左が短期型データベース用、右が長期型データベース用の画像。	123
10.4	カラー自然画像データベースで用いる例示画。	125

10.5 (実験 1) NOAA 衛星画像データベースの学習用サブセット 163 枚を検索対象とした検索結果。重み係数は表 10.1 を用いた。問い合わせ画像は 1 位に検索されている画像である。	129
10.6 (実験 2) NOAA 衛星画像データベースの全画像 1027 枚を検索対象とした検索結果。重み係数は表 10.1 を用いた。問い合わせ画像は 1 位に検索されている画像である。	130
10.7 (実験 3) NOAA 衛星画像データベースの全画像 1027 枚を検索対象とした検索結果。重み係数は表 10.1 を用いた。問い合わせ画像は 1 位に検索されている画像であり、図 10.6 とは変えている。	131
10.8 (実験 4) NOAA 衛星画像データベースの学習用サブセット 163 枚を検索対象とした検索結果。重み係数は表 10.2 を用いた。問い合わせ画像は 1 位に検索されている画像である。	133
10.9 短句型 GMS 衛星画像データベースの検索結果。重み係数は表 10.5 を用いた。問い合わせ画像は 1 位に検索されている画像である。	136
10.10 長句型 GMS 衛星画像データベースの検索結果。重み係数は表 10.6 を用いた。問い合わせ画像は 1 位に検索されている画像である。	137
10.11 (実験 1) カラー自然画像データベースの全画像 950 枚を検索対象とした検索結果。重み係数は表 10.7 を用いた。問い合わせ画像は 1 位に検索されている画像である。	139
10.12 (実験 2) カラー自然画像データベースの全画像 950 枚を検索対象とした検索結果。重み係数は表 10.7 を用いた。問い合わせ画像は 1 位に検索されている画像である。これは図 10.11 とは異なる画像である。	141
10.13 個々の重み係数を調整するためのユーザインタフェース。	147
A.1 Thurstone の方法の模式図。	157
B.1 スケールスペースの例。下から上へ、スケールパラメータ σ の増加に伴う輪郭の変化。	159
B.2 ガウシアンフィルタを用いてヒストグラムをぼかした例。 τ がガウシアンフィルタの標準偏差を示している。	160
B.3 スケールスペースを用いた初期値推定。	161
C.1 ボリュームレンダリングで用いられる座標系。この例は平行投影のものであるが、本研究では中心投影を用いていることに注意。	165
C.2 サンプル点の色の決定。	167
C.3 雲頂高度と雲底高度の模式図。遮蔽にも注意。	169
C.4 NOAA 衛星画像をボリュームレンダリングにより 3 次元可視化した結果。1995 年 9 月 16 日の台風 12 号。	170

表目次

2.1	衛星画像の対象物とその特徴。	24
2.2	本論文で用いる、画像検索システムの階層的モデル。	28
2.3	本研究で用いた手法と階層的モデルとの対応関係。	30
3.1	正方形モデルを用いた場合の、Case 1 ~ Case 5 の場合の面積占有率密度の導出。	34
3.2	画素と正方形モデルの位置関係による 5 個の場合それぞれの生起確率。表中で面積 S は $S = (b + c)^2$ である。	37
3.3	十字形モデルを用いたときの、Case 1 ~ Case 15 の面積占有率密度の導出。	39
3.4	画素と十字形モデルの位置関係による 15 個の場合それぞれの生起確率。表中で面積 S は $S = (5b + c)(b + c)$ である。	45
5.1	図 5.3 に関するパラメータ一覧。	72
5.2	クラス 1 が占める面積比の推定。	76
7.1	各画像データベースの階層化属性付き関係グラフで用いられている形状特徴量の一覧。	93
8.1	代表的な探索アルゴリズムの比較。最短経路を求める問題の場合。	98
8.2	1027 枚のグラフマッチングの計算時間 (20 回の試行の平均、単位は秒)。(a) 両方の高速化手法を用いた場合、(b) ヒューリスティック関数だけを用いる場合、(c) 最大許容コストだけを用いる場合、(d) 両方とも用いない場合、である。上段はアークのヒューリスティック関数を用いた場合、下段は用いない場合である。	103
10.1	NOAA 衛星画像データベースの検索に用いられた重み係数 (1) の一覧。遺伝的アルゴリズムによって最適化された結果である。	128
10.2	NOAA 衛星画像データベースの検索に用いられた重み係数 (2) の一覧。遺伝的アルゴリズムによって最適化された結果である。	128
10.3	類似検索実験 1 の場合に上位検索された画像と主観的類似度との関係。	128
10.4	類似検索実験 4 の場合に上位検索された画像と主観的類似度との関係。	132
10.5	短期型 GMS 衛星画像データベースの検索に用いられた重み係数の一覧。遺伝的アルゴリズムによって最適化された結果である。	135
10.6	長期型 GMS 衛星画像データベースの検索に用いられた重み係数の一覧。遺伝的アルゴリズムによって最適化された結果である。	135
10.7	カラー自然画像データベースの検索に用いられた重み係数の一覧。遺伝的アルゴリズムによって最適化された結果である。	140

10.8 各画像データベースの場合の再現率と適合率。	143
10.9 重み係数がさまざまに異なる場合の検索時間の平均。	144
10.10 上位 10 枚 / 100 枚 検索の場合の検索時間の平均。	145
B.1 画像のスケール。	159

第1章

序論

1.1 画像データベースの内容検索

1.1.1 はじめに

人間はきわめて視覚的な動物である。ゆえに、絵画から写真、映画、テレビへ — 静止画から動画へと発展しつつあるこの画像情報というメディアは、人間に多量で豊かな情報を与えることができる重要なメディアである。また言葉では表現できず、画像によって初めて表現できる情報があることは、昔からの諺「百聞は一見にしかず」に端的に示されているといえよう。しかし同時にコンピュータにとっては、画像情報は極めて扱いにくいメディアでもある。なぜなら画像情報そのものはコンピュータにとって扱いやすい形に構造化されていないメディアだからである。したがってコンピュータは画像をそのまま「理解」することはできず、画像情報をコンピュータで扱うためには、何らかの方法で画像をコンピュータが扱いやすい表現に変換する処理が必要となる。

このように、画像をコンピュータが扱いやすい形に変換したうえで、またはコンピュータが扱いやすい表現で表された情報を画像にリンクしたうえで、大量の画像を系統的に蓄積したものが画像データベースである。近年の情報処理技術の発達に伴い、さまざまな情報源から得られた画像情報が世界の至るところで膨大に蓄積され始めている。このような画像データベースがインターネットなどのネットワークを通じて接続されることにより、まさに我々は大量の画像情報に満ち溢れた世界と向き合い始めているのである。このような大量の画像情報を有効に活用するための技術が画像検索技術である。つまり画像データベースから何らかの条件を満たす画像のみを検索できる技術が、この大量の画像情報を有効に活用するには必須の技術となるのである。

そこで本研究ではこのような画像検索技術を研究の対象とする。特に近年になって強力な画像検索手法として注目されている、画像内容に基づいた検索 (Content-based Image Retrieval : CBIR) 手法 [1, 2] について多角的に検討を加える。この手法では、ユーザは検索したい画像内容を検索キーに直接指定することで、大量の画像の中から所望の画像を効率的に検索することができる。もはや大量の画像情報の活用は夢ではないのである。

1.1.2 キーワード検索方式

しかし単に画像内容に基づいた検索ということであれば、それほど目新しい技術ではない。例えば従来の画像データベースで用いられてきたように、各画像に人間が前もってキーワードな

どのテキスト情報を与えることによって、キーワードを用いた CBIR が実現できる。テキスト情報は記号的な情報に抽象化されているため、コンピュータによる情報検索に適した表現であり、しかも人間にとっても直観的に理解しやすい有効な表現である。さらに、キーワード付与には通常は人間のもつ強力なパターン認識能力を用いているが、この強力なパターン認識能力には現代のいかなる最先端のコンピュータ技術も太刀打ちすることはできない。つまりこの方法はそれなりに優れた方法なのである。

しかしその最大の欠点は、キーワードなどのテキスト情報を人手で付与しなければならない点にある。近年のコンピュータや記憶装置の発達に伴って扱われる画像の量が飛躍的に増大しつつあり、キーワードの付与処理に人手の介入を必要とする点がキーワード検索方式のボトルネックになりつつある。むろん、画像を自動的に処理してキーワードを付与する試みが対象領域を絞って研究されている [3]。またキーワード間に自動的に類似度を付与するための研究もおこなわれている [4, 5] が、全体的に現在の技術はいまだ未成熟な段階にあるといえる。

また、そもそもコンピュータによって自動的に画像情報を抽出するのであれば、キーワードにこだわる必要はあるのだろうか？むしろ定量的表現に優れた別のモデルを使う方がコンピュータの数値的な処理能力をより有効に活用することができるのではないかと筆者は考える。キーワード検索方式は、もし画像中に出現する事象が人間にとって何らかの意味をもつものであれば、その事象には何らかのキーワードがすでに付与されているはずであるとの仮定のもとで、その有効性が保証されている。しかし画像の場合を考えてみると、すべての意味のある事象に言葉が与えられていると考えるのはいささか楽観的である。実際は形状特徴量や位置関係など、的確な言葉で表しようがない場合も多い。ゆえにキーワードはいつでも有効なわけではなく、やはり画像というメディア特有の情報を表現できる画像表現モデルが必要となるのである。

1.1.3 新たな検索方式

では一体、近年の CBIR 技術の何が新しいのだろうか？これは人間が主観的に感じる画像内容を多変量解析や感性情報処理の枠組に基づいてモデル化する方向や、キーワードによっては表せない画像情報を画像表現モデルによって表現し内容検索のキーとする方向などが新しい技術であると考えられる。特にコンピュータは数値そのものを扱うことにかけては人間よりも優れているため、画像情報のなかでも数値的な情報をコンピュータに処理させることにより、キーワードが苦手とする定量的な情報をキーとして検索することが可能となる。さらにコンピュータの数値処理能力を活用すれば、画像内容の類似度を数値的に定義することで、「画像内容の類似した」画像を類似度に基づいて検索する方式が実現できる。このように画像内容を検索キーとして指定することにより、画像内容の類似した画像を検索できるような高度な画像データベース検索方式を類似検索 (Similarity-Based Retrieval) という [6]。本研究が目標とする検索方式はまさにこの類似検索である。

1.2 類似画像検索

1.2.1 類似画像検索の課題

類似画像検索の研究が盛んにおこなわれるようになったのは最近 10 年ほどの間である。それから現在までの間にかかなり多くの研究が発表されてきた。その全部を網羅してここに紹介するの

は無理であるため、以下では代表的な研究を取り上げてその問題点を明らかにする。まず一般的に類似検索を実現するための技術課題について検討する。それは美濃によると次の3点にまとめることができる [7]。

1. 信号として与えられるマルチメディアデータから検索のためのさまざまなインデックスを作成するパターン認識技術。
2. インデックスが誤りを含んでいても重大な検索ミスを起こさない検索手法。
3. データ間の類似度の定義がユーザによって異なるため、検索の過程で個人に適應できるように類似度計算が定義できるメカニズム。

最初の点は基本的には従来のパターン認識技術の延長線上にある問題点であり、パターン認識技術の向上がまず必要である。しかし画像検索との関連を考えた場合、ある1枚の画像だけに顕著な特徴量を抽出するよりは、データベース中の多数の画像において顕著であり、しかも画像の識別に役立つ特徴量を画像から抽出するという観点が生まれてくる。このような観点に基づいて、例えば画像から容易に抽出できる多数の特徴量の中から、識別に対する有効性という観点で特徴量に優劣をつける、または特徴量ベクトルの主成分を新たな特徴量として考える、などの研究が進んでいる [8, 9]。

2番目の点に対しては、従来のデータベース検索で用いられるように、正確に一致するマッチング (exact matching) ではなく、不一致を許容するマッチング (inexact matching) を用いることが重要である。特徴量ベクトルなどの数値的なモデルはもともと不一致を許容するマッチングであるが、記号的なモデルを用いる場合には、そのどちらを選ぶかは本質的な問題である。不一致を許容するマッチングに対しても、さまざまな研究がなされている [10, 11, 12]。

3番目の点はまさに類似検索の根幹を成す問題である。画像内容の類似度を客観的に不変な形で定めることができる問題は非常に限られている。一般に類似度というものは、問題領域に依存するばかりでなく検索の目的やユーザの知識によっても動的に変化するものであり、さらに個人個人の主観・感性にも依存するものである。したがって類似度を問題領域から独立した形で定式化することは本質的に困難である。しかしこの点についてはもう少し慎重な検討が必要であろう。たとえ処理全体を問題領域から独立した形で定式化できないとしても、その部分問題は問題領域に独立な形で定式化できる可能性があるからである。そこで、類似検索問題をいくつかの部分問題に分割し、問題領域に依存する部分と依存しない部分とを明確に区別できるような方法論が必要となる。本研究ではこのような方法論の具体例として「階層モデル」を後に提案する。

1.2.2 代表的な類似検索システム

類似検索システムの代表的な研究には、まず電総研が中心になって構築した一連の画像データベースである、商標・意匠データベース TRADEMARK [13] や電子美術館 ART MUSEUM [14] などがある。これらの研究では、画像特徴量の抽出から始まり、画像に対する人間の主観的印象を反映した検索を可能とする機構までを総合的に検討している。このような総合性は本研究の目指す方向に近い研究である。この研究では、濃度値分布などの比較的簡単な画像特徴量を用いて検索する方法を提案している。この方法はロバスト性に優れ、また画像の印象を捉えるには巧妙な方法であるといえるが、画像の構造や要素の位置関係などを明示的に指定した検索などには向かない方法である。

次に自然画像データベースとしては、IBMを中心として研究されているQBIC[15,16,17,18,19]がある。この研究は画像の内容検索技術を現段階で実用化した代表的な研究例である。この研究では、画像の領域分割などを半自動でおこなうことにより、精度と速度のトレードオフをうまく解決して実用性を高めている。また特徴量のIndexingや検索候補のFilteringの研究も進んでおり、これらの機構によって大量の画像を高速に検索することに成功している。しかし属性値としては色情報などの比較的簡単なものを用いており、空間的配置をキーにした検索などには限界がある。

またVirage Inc.が開発しているPinPointも実用レベルに達している研究である[20]。まず[21]では画像検索システムの枠組について述べられている。一般に画像データベースの研究は個別の要素技術の実装で終わる傾向があるが、この論文で提案されたXenomaniaシステムでは一般的な画像検索システムの枠組を構築し顔画像検索に応用している。さらにその発展形であるPinPointは、WWW上でのデモも提供している。

このようにいくつかの大規模な研究が進むなかで、個別の研究も多く研究者によって進められている。例えば医療画像を対象としたもの[22]、グラフ構造を用いたもの[23,24]、形状特徴量を用いたもの[25]、カラー画像の色情報をファジーで扱うもの[26]画像要素の位置関係などを指定できるもの[27,28,29,30]などの研究がある。またプリミティブ分解を用いるシステムVPointでは、ユーザの視点に基づいてゆらぎを調整することにより、柔軟な画像検索システムを実現している[31,32]。

1.2.3 類似画像検索システムの問題点

このように類似検索はさまざまな要素技術を必要とする総合的な問題であるが、従来の研究では個別のパターン認識技術の高度化に研究の重点が置かれていた。そのため類似検索システムをどのように構築すればよいかという方法論が大きな構想で論じられることは少なかったように筆者は感じている。またもう一つ別の側面として、画像特徴量の抽出に非常な労力が注ぎ込まれているにもかかわらず、抽出された画像特徴量どうしの類似度を計算するという類似検索の最も本質的な部分では、なぜか拍子抜けするほど単純な計算式を用いている研究が多いとも筆者は感じている。つまり次の2点が問題である。

1. 類似検索システムの枠組について総合的に論じられることが少ない。
2. 類似度の計算方法については注意があまり払われていない。

ここで注意すべき点は、「類似度」という言葉がさまざまな側面を持つことである。一つの側面として、領域情報(1項関係)と空間情報(2項関係など)という種類の異なる特徴量を統合して同時に扱えるような類似度の定義が少ない。どちらかを個別に扱うのではなく、両者を統合して扱えるような画像表現モデルに基づいた類似度が必要である。さらに、類似度という概念には「人間がみて」類似であるという要素が必然的に含まれるが、このような側面までを考慮して類似度を定式化している研究はわずかであり、ART MUSEUM[14]や感性情報処理などのごく一部の画像検索システムで検討されているのみである。

1.3 本研究の方針

このような問題点を踏まえた上で、本研究が問題を定式化し研究を進めていくうえでの方針についてまとめる。

- 個別の問題領域に特化したシステムを目指すのではなく、画像データベースの全体的な枠組を意識しながら、より一般的な観点から構築したシステムを個別の問題領域に応用するアプローチをとる。具体的には、まず画像データベース構築の方法論として6層の階層構造から成る「階層モデル」を第2.4節で提案する。この方法論は特定の対象になるべく依存しないように作られたモデルであり、問題領域が異なる場合には各レイヤの個別の画像処理手法を必要なだけ交換すればよいという枠組が維持できるようなモデルを追究する。このようなモデルによって画像検索システムを統一的に把握することができ、しかも各レイヤを比較検討することでどのレイヤが弱点であるかを判断することもできる。
- 領域情報と空間情報を統合して扱うことができる画像表現モデルとして、グラフ構造を用いる手法を提案する。このモデルは明示的な記号表現であるが、属性として数値的な情報も含む。
- またどの画像特徴をどれだけ重視するかという、各特徴の重要性を表す重み係数を導入し、ユーザの類似性に対する観点を重み係数の値の組によって表現する。この重み係数を調整する機構を備えることによって、多様な検索目的に応えうる柔軟な(主観的な)検索を可能とする。
- できるだけ人手を介さずに画像から情報を抽出し、その情報を自動的に構造化することによって画像検索システムを構築することを目指す。階層モデルに基づき、画素単位のミクロなレベルから画像単位のマクロなレベルへと進む階層的な処理によって構造を備えたインデックスを画像から自動的に抽出する。現時点では技術的にかなり困難な課題ではあるが、大量の画像を処理するためにはこの観点からの研究が必要である。

1.4 本研究の目的および成果

このように本研究では、画素単位の処理から始めてボトムアップ的な処理をおこない、画像の領域情報と空間情報を表現するグラフ構造を自動的に抽出することを目指す。またこのグラフ構造に関係する重み係数の調整を通じて、人間の多様な検索要求に柔軟に応える画像検索システムの構築を目指す。このように本研究の目的は、画像データベースの総合的な研究を通して、個別の要素技術と画像データベースの全体像との橋渡しをおこなうことが一つ、もう一つの目的はこのようなモデルを用いて柔軟で高度な画像検索システムを実現する点にある。また階層的なモデルを用いて画像検索という複雑な問題を整理することで、複数の種類の画像データベースに応用しやすい枠組を作ることのもう一つの目的である。

これは全体的にはいささか野心的な計画ではあるものの、個別の要素技術の追究だけではなく総合的に画像検索という問題を眺めることで見えてくるものがあるはずであると期待している。そして本研究で得られた主な成果を、以下にまとめる。

1. 階層モデルによって画像データベース検索システムを構築するための方法論を示すことで、3種類の異なる画像データベースに本研究のシステムを見通しよく応用することができた。
2. ミクセルに対する総合的な研究によって画像中のミクセルを適切に扱える方法を提案した。研究ではミクセルの統計的性質と、それを導くために必要な基本的な確率密度関数の性質を明らかにした。また従来の方法ではあまり考慮されなかったクラス内分散の影響を取り入れる形でミクセルの理論を構築し、実際の衛星画像に応用した結果、ミクセルを含む画像上で画素をピュアピクセルとミクセルを明確に分類し、さらにはミクセルの面積占有率を計算することが可能となった。
3. グラフ構造を用いた類似検索手法を提案した。グラフ構造は計算量の大きい検索手法であることが欠点となっていたが、類似検索のためのデータ構造としては記述力が高く有望なデータ構造である。この計算量の増大を抑えるための手法として、本研究ではA*アルゴリズムを用いた高速化手法について提案し、実際に検索の高速化に役立つことを示した。
4. 人為選択法に基づいた類似検索規準の学習法についても研究をすすめた。特にこのような手法に適した手法として遺伝的アルゴリズムが提案されているが、本論文ではこの遺伝的アルゴリズムをこの問題に適するように本質的に改良し、このアルゴリズムによって動作速度の向上と収束性の向上とが得られることを確かめた。また重み係数調整の機構により、人間が選んだ主観的類似画像を検索できるような重み係数が得られることを示し、それを応用してさまざまな観点からの類似検索が簡単にできる方法を提示した。

1.5 本論文の構成

本論文は「領域・空間情報を表現するグラフ構造を用いた類似画像検索」と題し、階層モデルに基づいた画像データベース構築の方法論から、各階層に適した種々のアルゴリズムの検討までを総合的に研究し、例示画に類似した画像を検索できるような高度な画像データベースを構築することを目指した一連の研究をまとめたものであり、以下の11章および付録3章から構成されている。

第1章「序論」では、まず画像データベース検索についての関連研究をサーベイしたうえで従来の研究の問題点をまとめ、次にそのような問題点に対する本研究の提案を述べる。そして本研究の目的を明らかにするとともに、本論文の構成について述べている。

第2章「本研究の対象画像および本研究が用いるモデル」では、本研究で画像データベースとする対象画像について簡単に紹介し、次に本研究が画像データベース構築の方法論として用いる階層モデルを紹介する。そしてこの階層モデルを構成する6層のレイヤについて、それぞれの処理の目的とその内容についてまとめる。

ここまですが本論文の導入部である。以下の章では本論文で新たに提案する手法を中心に述べる。

第3章以下の3章は、画像中に存在するミクセルを扱う手法に関して述べている。まず最初の第3章「L型やJ型・U型の形状をもつミクセルの面積占有率密度」ではミクセルの面積占有率密度を計算する手法を述べる。次に第4章「期待値に基づいたミクセルの面積占有率の推定」では、第3章で提案した手法を用いながら、ミクセルの面積占有率を期待値に基づいて計算する

手法を提案する。最後の第5章「ミクセル密度を含めた混合密度推定と統計的な画像分類への応用」では前の2章で提案した手法を応用しながら、ミクセル集合が画像ヒストグラムに形成するミクセル密度という新しい密度を提案し、この密度を含む混合密度推定によってミクセルを含む画像を画素単位で分類する手法について述べる。ここまでの階層モデルでは画素レイヤに相当するものであり、特に衛星画像への応用を考えた研究であり、本論文が提案する手法の中でも一つの中心となっている部分である。

第6章「領域に基づいた変形モデルを用いた形状分解」では、2値画像と濃淡画像のいずれに対しても適用できる形状分解手法を提案する。この手法は階層モデルでは領域レイヤに相当するものである。

第7章「引力モデルを用いた群化と階層化属性付き関係グラフの構築」では、それまでに抽出されたプリミティブをグラフ構造に構造化する手法を扱う。ここまでの画像はグラフ構造に変換されており、そのグラフ構造は領域・空間情報を表現したものとなっている。

第8章「A* アルゴリズムを用いたグラフマッチングによる画像検索」では、いよいよ画像検索の本題に入り、グラフ構造を用いた類似検索の方法について述べる。ここも本研究が提案する手法のもう一つの中心である。

第9章「人為選択法に適した遺伝的アルゴリズムを用いた類似検索規準の学習」では、この類似検索システムをさらに柔軟なものとするための方法としてユーザによって異なる検索目的をうまく類似検索に反映させるための機構を提案する。このようにして、階層モデルに基づいた総合的な類似画像検索システムが完成する。

ここまでの本論文が提案する手法の説明はひとまず終わりであり、以下の章では実験結果と結論について述べる。

第10章「類似画像検索」では各問題ごとに階層モデルをどのように実現したかを簡単にまとめ、そのあとで類似検索の実験結果を示す。この結果により、階層モデルは画像データベースをさまざまな問題領域に適用する際に、問題の構造を明らかにする方法論となりうること、またグラフ構造を用いた類似画像検索によって、形状・空間情報を活用した類似画像検索が可能となるなどの結果を得た。

第11章「結論」では本研究で得られた成果と今後の研究課題をまとめる。

なお付録として以下の3章を加えた。

第A章「類似度と距離」では、本論文が扱う類似度の計算について、本論文の内容を補足する説明を加えた。

第B章「スケールスペースを用いたヒストグラム解析」では、本論文の何ヶ所かで用いられているこの手法について、そのような手法が用いられる背景と手法の具体的な記述をまとめた。

第C章「ボリュームレンダリングを用いた気象衛星画像の3次元可視化」では、気象衛星画像をそのまま見せるのではなく、3次元情報を用いてわかりやすく可視化する方法について述べた。

第2章

本研究の対象画像および本研究が用いるモデル

2.1 本研究が対象とする画像データベース

第1章では画像データベースの内容検索の一般論について述べ、その問題点と本研究の方針についてまとめた。しかし画像の内容検索技術はリアルな対象を得てはじめておもしろく展開していくものである [33]。つまり画像データベースの要素技術の開発に対象画像群の特性を生かしていくところに興味あるシステムの実現が可能となっていく。そこで本章では本研究が画像データベースの研究対象として選んだ画像を紹介する。そしてその画像を対象とする画像データベースを構築する背景や、その画像に固有の問題点などを紹介する。本研究では扱う画像データベースは以下の3種類である。

1. 気象衛星 NOAA 画像データベース
2. 気象衛星ひまわり画像データベース
3. カラー自然画像データベース

以下本章では、最初の二つの画像データベースを衛星画像データベースとして一つにまとめて紹介し、次にカラー自然画像データベースについても紹介する。

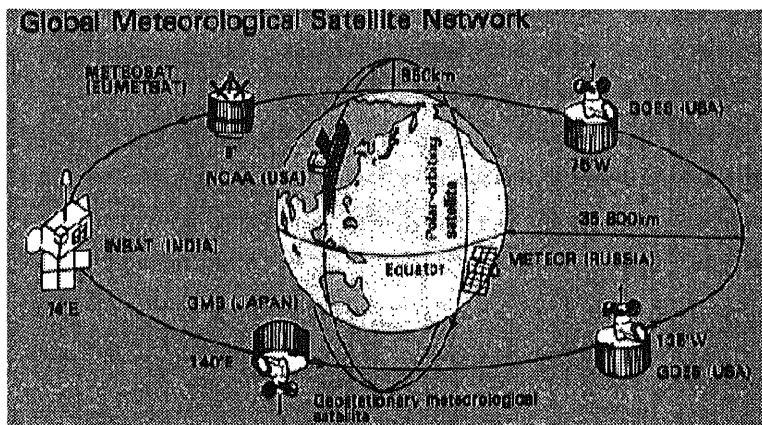


図 2.1: 気象衛星 NOAA および気象衛星ひまわり。

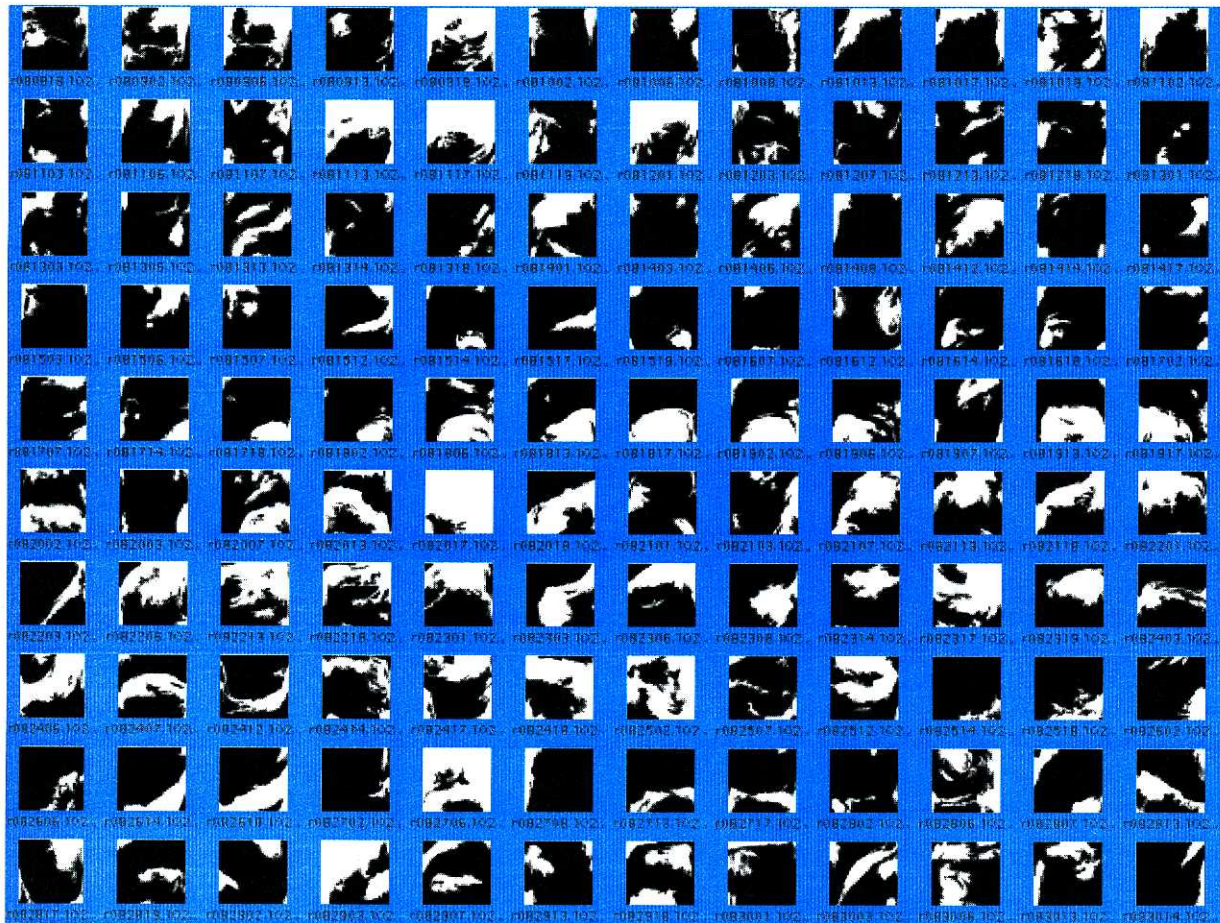


図 2.2: NOAA 衛星画像データベースに蓄積されている画像。

2.2 衛星画像データベース

2.2.1 気象衛星 NOAA および気象衛星ひまわり

まず最初に二つの衛星画像データベースについて述べる。本研究では衛星画像データベースが研究の中心となっている。このように衛星画像データベースを研究の中心に据える理由は、著者が所属する研究室では現在気象衛星 NOAA と気象衛星ひまわり (GMS) からデータを毎日受信・蓄積しており、その画像を効率的に活用できる画像データベースが大きな課題であることがその理由である。まずこの二つの衛星の特徴と違いを簡単に説明する。

まず気象衛星 NOAA は図 2.1 の中で地球を回っている衛星を指し、地球を周回する極軌道気象観測衛星である。現在は NOAA12 号と NOAA 14 号の二つの衛星が稼働中であり、これらの衛星が日本上空を通過し著者が所属する東京大学生産技術研究所¹から観測可能となるのは 1 日 8 回～10 回程度である。ゆえにすべての受信画像を蓄積するならば、画像数は 1 年間では約 3200 枚程度に達することになる。

一方の気象衛星ひまわりは図 2.1 の中で赤道上に描かれている衛星を指し、赤道上空に静止している静止気象衛星 (Geostationary Meteorological Satellite) である。東経 140 度の赤道上空 35,800km から、直下点を中心にして半径 6000km、地球のほぼ 1/4 の広い範囲を絶えず観測し

¹東京都港区六本木 7-22-1

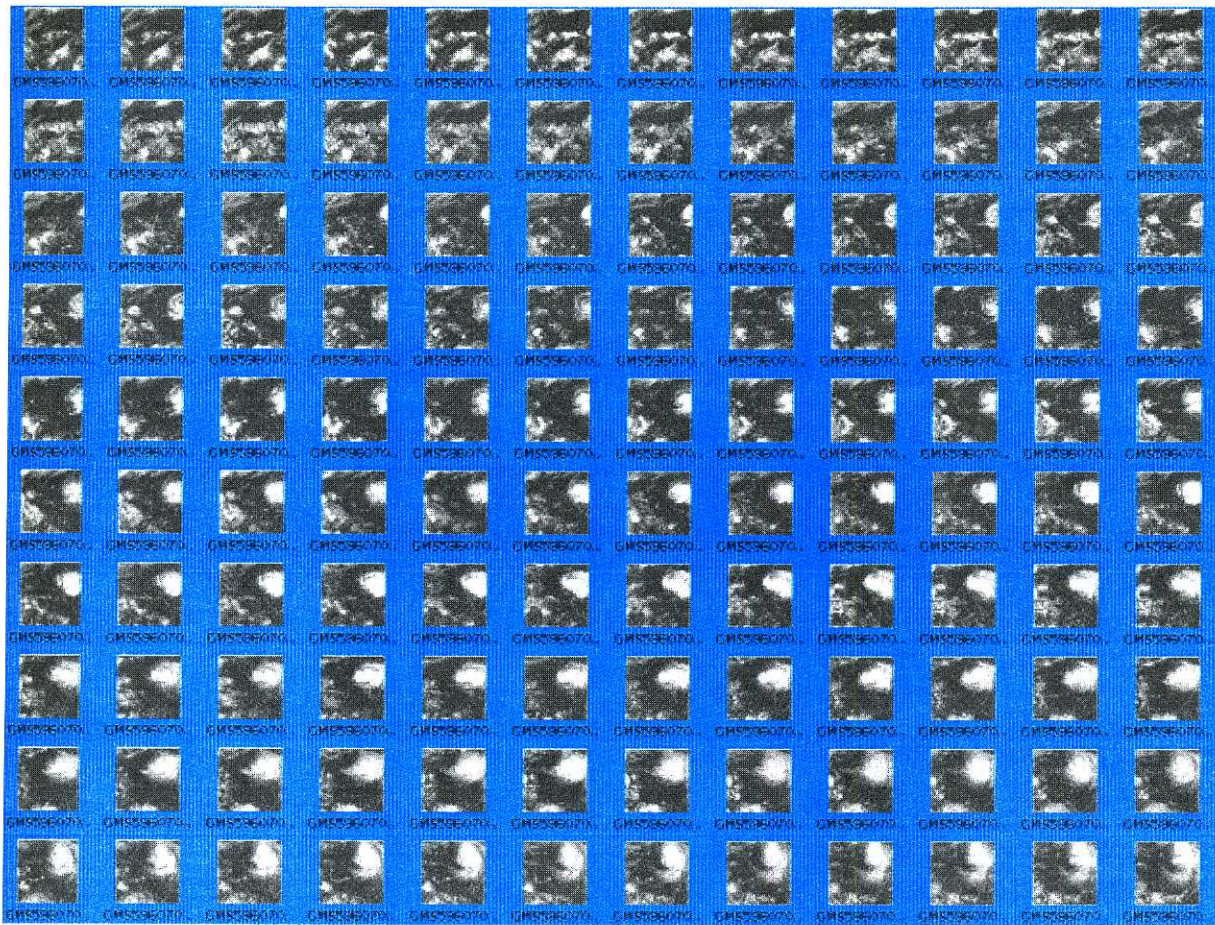


図 2.3: GMS 短期型衛星画像データベースに蓄積されている画像。

ている [34]。この衛星の観測頻度は通常は 1 時間に 1 回の毎時観測である。ゆえにすべての受信画像を蓄積するならば、画像数は 1 年間で 9000 枚弱に達することになる。

要するに気象衛星 NOAA との最も大きな違いは、気象衛星 NOAA が地球を周回する衛星であるのに対し、気象衛星ひまわりは赤道上空に静止している衛星である点である。この違いによって 2 種類の画像では観測地域と受信間隔の面で大きな違いが生じる。まず、気象衛星 NOAA では毎回の受信ごとに観測地域が異なるのに比べ、気象衛星ひまわりではどの受信においても観測地域は同一である。また受信間隔についても、気象衛星 NOAA の場合は不定期なのに比べて気象衛星ひまわりでは毎時観測が可能である。以下ではこのような特徴をどのように考慮するかという点についてももう少し詳細に記述し、また本研究で用いた画像の具体的なデータについても触れる。

2.2.2 NOAA 衛星画像データベースの対象

気象衛星 NOAA で観測した画像を基本的なデータベースとして用いる。この画像は気象衛星 NOAA の AVHRR(Advanced Very High Resolution Radiometer) センサから受信したデータであり、解像度は衛星直下点で約 1km である。AVHRR センサは観測チャンネルとして可視と赤外のチャンネルを持っているが、この中で赤外のチャンネル(チャンネル 4)で観測した画像を画像検索に用いる。赤外チャンネルを用いるのは、夜間でも観測が可能であることが理由である。また撮影された画像は幾何学的に歪んでいるため、その歪みを幾何補正により取り除き、さらに衛星

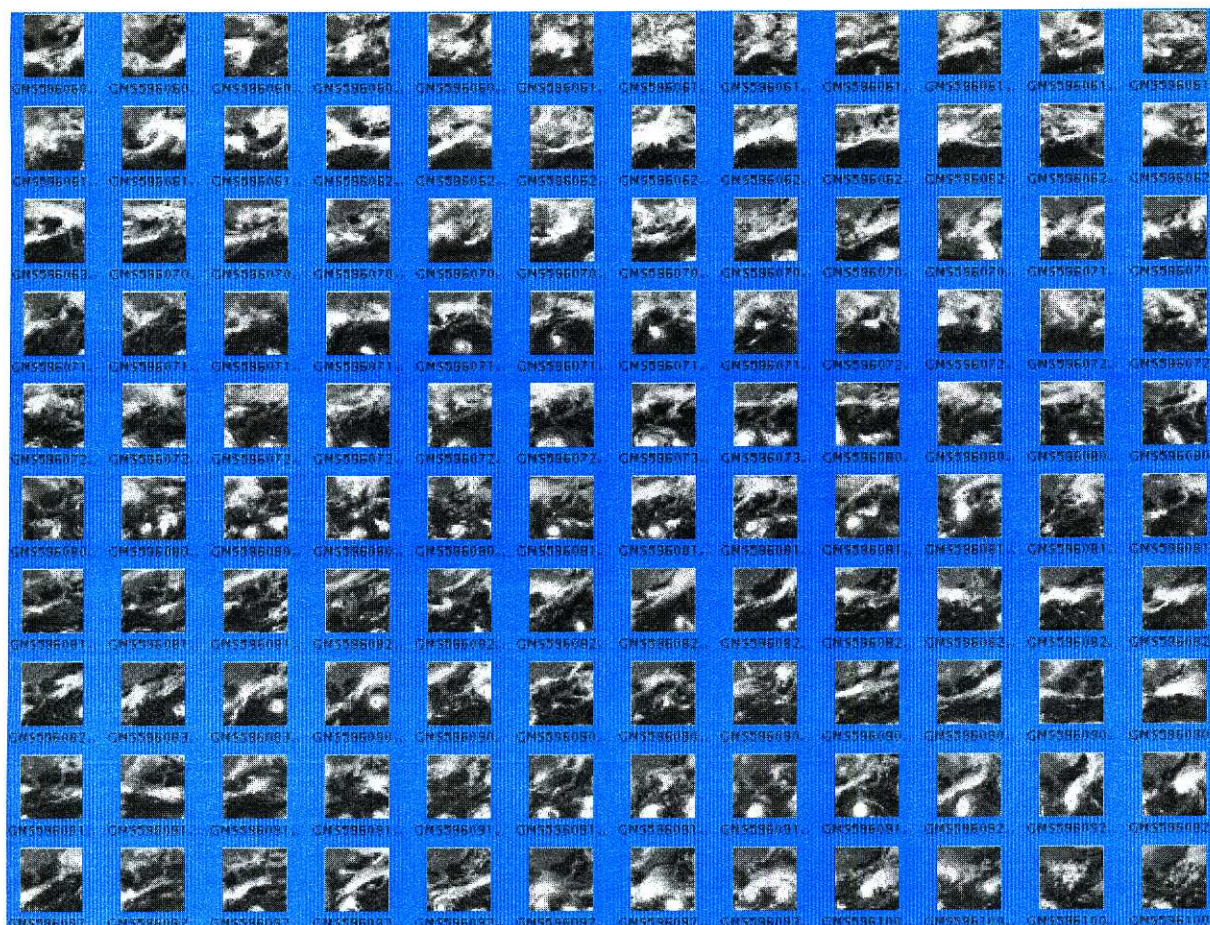


図 2.4: GMS 長期型衛星画像データベースに蓄積されている画像。

データの観測値(カウント値)を放射量補正して輝度温度に変換する。検索対象とするのは、受信画像から日本付近を切り出して縮小画像とした、いわゆるクイックルック画像(512×480)であり、観測日時が1991年6月から1992年2月までの合計1027枚が検索対象である。この画像データベースに登録されている、2値化したクイックルック画像の例を図2.2に示す。

2.2.3 GMS 衛星画像データベースの対象

次に気象衛星ひまわり(Geostationary Meteorological Satellite)から受信したデータを対象とした画像データベースを構築する。こちらも受動型の光学センサによって観測したものであり、受信画像の解像度は可視の場合は衛星直下点で約1km、また赤外の場合は約5kmである。先述したようにGMS衛星画像は、毎回の受信で観測地域が同一であり、しかも定期的に観測できるという、NOAA衛星画像にはない大きな利点を有している。この特徴を活用して、本研究ではGMS衛星画像データベースとしてさらに2種類のデータベースを作成する。

短期型データベース 毎時観測のデータをすべて蓄積することによって、気象状況の細かい変化を追跡できるような検索システムを目的とする。1996年7月3日から9日にかけて毎時観測された132枚の画像をデータベースに登録した。毎時観測する関係で夜間の観測画像も含まれるため、観測チャンネルとしては赤外バンド(チャンネル1)を用いた。観測範囲としては太平洋上の海面を選んだ。この季節のこの海域では、熱帯性低気圧(台風)が頻繁に発生する気象状況となってい

る。この画像データベースに蓄積された画像の例を図 2.3に示す。

長期型データベース こちらは短期型データベースとは異なり、1日に1枚の画像しか登録しない。しかしその画像を約1年分にわたって蓄積することによって、長期にわたる気象の変化を追跡できるような検索システムを目的とする。1996年1月から11月にかけて日本時間正午²に観測された331枚の画像をデータベースに登録した。毎時観測のデータをすべて蓄積すると画像数が多くなり過ぎるという別の問題が生じるため、プロトタイプ画像データベースとしては1日1枚の画像に限定した。ただし同時刻の画像を登録することにより、観測の周期性という特徴を失わないように配慮している。もちろん将来的にはすべての画像を1つの画像データベースに蓄積することが目標である。観測チャンネルとしては短期型とは異なり可視バンドを用いる。これは、1) 観測時間がすべて正午と可視画像にとって最も観測に適した時間であること、2) 可視画像の方が赤外画像に比べて解像度が4倍も優れていること、3) 可視画像の方が雲域の抽出処理が容易であること、などが理由である。冬から夏を過ぎまた冬になるまでの画像を揃えることで、各季節の気象状況をおおよそ網羅した画像データベースとなっている。この画像データベースに蓄積された画像の例を図 2.4に示す。

2.2.4 衛星画像データベースの諸問題

このような衛星画像データベースを構築する際には、次の二点が問題となる。まず最初の問題は、衛星データの大きな情報量(大きな画像サイズ)である。衛星データの1回の受信では、そのデータ量が数十メガバイトから数ギガバイトのオーダーに達するため、このような大量のデータをどこにどのように蓄積すれば効率的に検索できるかという問題が生じる。その対策としては、階層的なファイルシステムを構築して高速化する方法などがすでに提案されている。しかし本研究では、このような保存・蓄積に関する諸問題ではなく、もう一つの別の側面に注目する。すなわち第二の問題は、画像サイズではなく画像数の意味からも衛星画像データベースは大量の画像を蓄積している。先述したように、GMS衛星画像データベースでは1年間に1万枚の規模で蓄積画像が増加していく。そこで、このような大量の画像の中から所望の検索目的を満たす画像を検索するにはどうすればよいか、という問題が浮上する。

現状では蓄積された画像を検索するための手段として、撮影日時による検索手段が提供されている[35]。また気象衛星NOAAのように観測ごとに観測領域が異なる場合には、観測領域を指定した検索手段も提供されている。しかしこれらは画像に対する付加情報を用いた検索手法であり、画像の内容に踏み込んだ検索手段が提供されているわけではない。そのため本当に画像の内容を知る必要がある時には、クイックルック画像という縮小画像を1枚1枚目視確認することによって検索する方法しかないのが現状である。しかしこの方法では、数年分のデータの中から必要なデータを選び出すなどという検索は困難であり、ゆえにせつかく大量に蓄積されたデータも有効に活用されないことになる。そこで撮影日時などの従来型の情報の他に、リモートセンシング画像の内容をキーとした検索を可能とすることが、大量の画像を活用するために必須の技術となるのである。では、このように内容に基づいた衛星画像の検索手段を提供するために、まずは衛星画像の内容とは何かという点を考察する。

²一部の画像については正午のデータが得られず、14時に観測したデータとなっている。

表 2.1: 衛星画像の対象物とその特徴。

対象	陸	海	大気(雲)
遮蔽	雲によって遮蔽される	雲によって遮蔽される	基本的に遮蔽されない
相	固体	流体(液体)	流体(気体)

2.2.5 衛星画像の内容とは

衛星画像は宇宙空間から地球表面を観測した画像である。本研究が用いるような衛星の解像度(1km程度)を考えると、その画像の中で重要と考えられる対象は、地球を構成する3大要素でもある陸、海、および大気(雲)であると考えてよいだろう。この3大要素はさらに細分化でき、たとえば陸の中ではさらに植生や氷雪などが対象となることもあるが、大分類としてはこの3種類に分類するのが適当だろう。この3種類の対象物の特徴を表2.1にまとめた。なお表の中で、大気の情報としては通常は雲のみを考えればよい³。まず陸と海は「遮蔽」という観点から考えると、「雲によって遮蔽される可能性がある」という点で共通している。そのため、陸と海の情報を検索したいユーザにとっては、「ある領域が雲によって遮蔽されているかどうか」という点が第一に重要な情報である。それに対し雲の場合には基本的に遮蔽の問題を考慮する必要がない⁴。次に海と雲に共通する特性はそれらが「流体である」ことである。流体であるという特性により、雲や海流の形状や位置は、静止することなく常に変化し続けている。そのため雲や海の情報を検索したいユーザにとっては、「海流や雲がどこでどのような形をしているのか」という点が第一に重要な情報である。以上をまとめると、衛星画像を検索する際に重要となる画像内容としては、大きく分けて以下の2点が考えられる。

1. 指定した領域が雲に覆われているかどうか
2. 指定した対象がどこでどのようなパターンとなっているのか

まず画像内容1の場合には、指定した領域の雲量をできるだけ高速に計算するアルゴリズムの検討が主題となる。そしてこの種の問題に適した、4分木などの階層的なデータ構造[36]の検討が重要となる。それに対して画像内容2の場合は、内容検索に適したデータ構造やアルゴリズムははるかに複雑なものとなる。なぜなら画像内容2の場合には、各領域の形状や領域間の空間的な位置関係という情報を表現できるデータ構造が必要となるからである。

このように2種類の検索目的は、それを実現するために適した手法やデータ構造が基本的に異なるため、別々に分けて考えるのが適当である。しかしここで、どちらのタイプの検索にしても、各画素に「雲が含まれているのかいないのか」という情報が根本的に重要であることは注意すべきである。雲はそれを検索対象にするにしろ、それを取り除く対象にするにしろ、まず最初にそれが存在するのかわかるかを画素(または画素ブロック)ごとに判別する必要がある。この種の情報の抽出については、その重要性から多くの研究がおこなわれてきたが、本論文でもこの問題の重要性を考慮してその情報を抽出するための手法を第3章～第5章で提案する。この手法では雲の有無という2値情報ばかりではなく、画素が雲をどの程度含んでいるかという雲量を計算する方法を、ミクセルという概念に基づいて計算する手法を提案している。検索目的によって雲域抽

³まれには黄砂などの大気現象が問題になることもある。

⁴高層の雲が低層の雲を遮蔽することがあることに注意。図C.3参照。

出 / 除去と使い分けられるこのような手法の適用が、まずは第一の課題である。さてここで内容検索に関して重要となるのは、雲域抽出 / 除去の後の処理である。

2.2.6 衛星画像の雲パターンの類似検索

本研究では画像内容 2 の場合に着目し、画像内容という言葉が指す範囲を、「衛星画像中の雲パターンの領域・空間情報に関する類似度」という内容に制限する。つまり衛星画像データベースの内容検索の一例として、雲パターンの類似検索を研究目的とする。その応用には気象学への応用がある。雲のパターンは気象学では重要な問題であり、種々の代表的なパターンは名前をつけて呼ばれているものもある [34]。このような雲パターンを衛星画像から自動的に抽出した画像情報で検索するという応用が考えられる。

またこの検索にキーワード検索を併用できればより有効な検索システムとなるが、本研究では基本的に画像情報から画像処理アルゴリズムを用いてキーワードを抽出する方法は採らない。なぜなら、画像からキーワードを直接抽出するよりは、気象情報や気象通報式・天気図などの別の情報源からの情報を統合する方が、的確なキーワードを確実な方法で手にいれることが可能だからである。このように、画像に関連したテキスト情報から自動的に検索に有用なキーワードなどを抽出する手法は近年の研究の大きな流れであり、キャプション付きの画像からキーワードを抽出する方法 [37] や、シナリオをドラマ映像の検索に用いる方法 [38] など、同様の考え方に基づいた研究が多数おこなわれている。

そこで本論文ではキーワードによる検索ではなく、例示画による検索手段を提供することを研究の目的とする。この検索手段では、検索キーとしてある画像を提示すると、その例示画に類似した画像をシステムが検索する。探したい雲パターンを含む画像を提示すれば、その雲パターンを含む画像をシステムが検索するのである。このような場合には、画像を何らかの画像表現モデルで表現し、そのモデルの間の類似性を計算する機構が必要となる。本論文でも第 7 章において、画像の領域情報と空間情報を表現できる画像表現モデルとして階層化属性付き関係グラフを提案する。

2.3 カラー自然画像データベース

以上に述べてきたように、本研究では衛星画像データベースを中心に研究することになる。この画像データベースはどちらかというと学術的な用途に用いられるものであるが、画像データベースのニーズは決してこのような学術的な用途ばかりではない。むしろより一般的な用途でも画像データベースの需要は高まる一方である。例えば個人的に撮影した写真を蓄積した電子アルバムなどの応用では、個人個人が自分自身のために利用する画像データベースなどの利用形態も考えられる。このように蓄積された画像が大量となればブラウジングによる検索にも限界があり、内容をキーとして画像を検索したいという欲求が生じるのである。しかし現在のところ、画像検索機能をもつ市販ソフトウェアの大部分では、画像検索手段はそれぞれの画像に事前に付与しておいたキーワードを用いるものであり、キーワードを手で付与する作業は欠かせないものとなっている。したがってこのような分野でも、今後は画像特徴量を用いた内容検索機能が重要性を増してくるはずである。



図 2.5: カラー自然画像データベースに蓄積されている画像。

2.3.1 カラー自然画像データベースの対象

そこで本研究ではこのような一般向けの用途をも考え、カラー自然画像を対象とした画像データベースも併せて研究をすすめる。具体的に検索対象としたのは、International Microcomputer Software Inc. 社のフォトライブラリー“Master Photo”に含まれる写真である。このフォトライブラリーには合計で約 26,000 枚の写真が含まれるが、この中で Building に関係する写真として製品側で分類されている 950 枚の画像を検索対象とした。この中にはある程度は多様な種類の画像が十分な量だけ含まれているので、まずは初期段階の実験の対象として適当な画像データベースであるといえる。このフォトライブラリーに含まれる画像は JPEG を用いて圧縮されている画像であるためそれほど高品質の画像ではない。大きさは 320×240 である。なおこの中にはカラー画像だけではなく中には白黒画像も含まれている。この画像の例を図 2.5 に示す。

2.3.2 カラー自然画像データベースの内容

カラー自然画像の内容を定義することは非常に難しい。どのような画像が出現するかわからない一般の画像を対象にした場合には、画像の内容を定義することはほとんど不可能であるし、また実用化の域には達している研究もほとんどない。このような画像データベースでは、やはりキーワードはそれなりに重要な役割を果たす。写真に写っている人物名などの情報をパターン認識アルゴリズムで特定するのは不可能だからである。ただしキーワード検索の問題は、キーワー

ドとして何を与えるかという選択が難しいという点がある。衛星画像の場合とは異なり、一般の自然画像の場合には画像を表現するキーワードが非常に多数考えられる場合がある。このような場合には、キーワード付与時とは異なる観点から検索をしたい場合に、キーワードが与えられていない画像は検索することができない。そのため、画像から特徴量を自動的に抽出して画像表現モデルで表現し、そのモデルを用いた検索手段が必要となる。本研究ではカラー画像の内容も、基本的には衛星画像の場合の内容と同様に考え、「カラー画像に含まれる領域の領域情報と空間情報の類似度」を画像内容として考える。ただし雲パターンと異なるのは、画像中のすべての領域を検索対象とする点である。どの領域が注目領域であるかを事前に判断できない以上、すべての領域を検索対象とする必要があるからである。そして衛星画像データベースの場合と同様に、画像の領域情報と空間情報を表現できる画像表現モデルとして階層化属性付き関係グラフを用いる。

2.4 階層モデル

本研究では以上に述べたように、衛星画像データベースとカラー自然画像データベースという、かなり性質の異なる2種類の画像を扱う。このように性質の異なる画像を扱える画像データベースを実現するためには、適切なデータ構造や手法の選択が重要である。しかしすべての画像を適切に表現することができる統一的な画像表現モデルは現時点では存在しない。また将来的には発見されるという保証も全くない。そのため、画像データベースを構築する際にはどうしても対象とする画像の問題領域に固有の表現を用いることになり、したがって画像データベースの構築は個別におこなわれることになる。しかしこの世界に画像の種類は無限に多い。したがって、画像データベースを問題領域ごとに個別に構築するよりは、それを構築するための何らかの統一的な方法論が望まれるのである。

そこで本研究では画像データベースを構築するための方法論として「階層モデル」を提案する。これは画像表現モデルの統一性を目指すものではなく、画像表現モデルや画像処理手法の選択に指針を与えることのできる統一的な方法論として本研究が提案するものである。

2.4.1 階層モデルの利点

「階層モデル」ということばは、一般的に二つの意味に使われている。一つは解像度を変化させた画像をピラミッド状に積み上げる方法である。本研究でいう階層的とはこのような意味ではなく、複雑な問題をいくつかの階層に分割することで問題構造の把握を容易とする方法論を指している。このような階層構造を考えるまず基本的な側面は、画素値配列そのものの解析に起因する膨大な計算量を回避したいという要請にある。しかしそれに劣らず重要な側面は、我々が何を計算しようとしているかを明確にしたいという要請である。

そこで本研究では表 2.2 のように6層から構成された階層モデルを提案する。そして各階層での問題を定式化する際には、Marr が提案した三つの水準 [39] — 原理・アルゴリズム・実装 — を手本とする。まず原理の水準では何をすべきかを明確に定め、それに適切な計算理論を構築する。次にアルゴリズムの水準ではその原理をどのように処理するかを考え、さらに入力と出力の表現を定める。最後に実装の水準ではそのアルゴリズムをどのような物理的装置で実現するかを考える。

そこで本研究の階層モデルでは以下のような枠組みを念頭に置く。まず各階層の計算原理・

表 2.2: 本論文で用いる、画像検索システムの階層的モデル。

レイヤ	目的	規準	出力表現
理解レイヤ	人間の主観を最適に反映する類似度空間の構成	例示データと検索結果との差+学習コスト→最小	主観的な画像間類似度
認識レイヤ	2 画像間の類似度の計算	画像モデルと認知モデルの差+類似度の計算コスト→最小	画像間類似度
関係レイヤ	領域間の関係の構築と構造化	認知と現実の食い違い+認知の複雑さ→最小	階層化属性付き関係グラフ
領域レイヤ	画素集合(領域)の特徴表現(プリミティブなど)	当てはめ誤差+プリミティブの複雑さ→最小	プリミティブ・形状特徴量
画素レイヤ	画素単位分類(物理世界の統計的性質など)	最大尤度+自由パラメータ数→最小	混合密度・しきい値
観測レイヤ	センサを通した物理世界の観測	測定コスト+測定誤差→最小	画素値配列

およびその原理を数理的に表現した規準を明確に定め、次に下位レイヤからの入力を制約条件として各階層の規準を最適に満たす解を発見し、最後にその解をレイヤの出力とする。このとき入出力表現の選択は軽視できない問題である。なぜならば、表現モデルで表現できる情報のみが上位レイヤに伝達され、表現できない情報は不要な情報として廃棄されるからである。また各レイヤは比較的独立していると考えてもよいが、完全に独立させることはできない。例えば下位レイヤの入出力表現を変更すれば、上位レイヤも必然的に影響を受けてしまう。しかしできるだけ各レイヤの処理を独立に考えることを基本方針とする。

さて表 2.2では、規準を原理に、方法をアルゴリズムに対応させることを念頭に置きながら、各階層ごとに階層モデルを具体化した。このようなレイヤを構成するために文献 [40] を参考にしている。また表 2.2に規準として書いたものは、パターン認識の一般的な原理である「ケチの原理 (principle of parsimony)」または「オッカムの剃刀 (Occam's razor)」 [41] という考え方を参考にして定式化したものである。

2.4.2 各レイヤの概説

以上のようにして提案された6層からなる階層モデルは、何段階もの処理から構成される複雑な問題構造を明確にするとともに、他のさまざまな問題領域への類似画像検索システムの応用にも指針を与えるモデルである。また下層は画素単位というマイクロなレベルの処理であるが、上層は記号化されたグラフ構造というマクロなレベルの処理であるとも考えることもできる。以下では、本研究が土台としている図 2.2の階層モデルに沿って、本論文の構成を説明する。

観測レイヤ このレイヤはセンサを通して物理世界を観測し、画素値配列としての画像を取得する段階を指す。本研究で扱う衛星画像データベースでは、センサそのものの改良や衛星データの幾何学的補正・大気補正などの話題がある。カラー自然画像の場合にも、撮影条件などが問題となるだろう。しかし本研究では画像の取得の問題は扱わず、すでに得られた画像をいかに画像データベースに蓄積するかという点に問題を絞る。観測に関するデータの一部については、すでに第 2.2節や第 2.3節において述べた。

画素レイヤ 画像を注目領域と非注目領域に分割したり、複数のまとまりのある領域に分割したりすることで、どの画素がおおよそどの分類クラスに属するかを決定することが目的である。この処理は基本的に画素単位であり画素の濃度値を情報として用いる場合が多い。そのために用いられる方法の一つであるエッジ検出は画像空間上の局所的な演算で処理することが可能であり、しかも生理学的知見に基づいた局所的演算により人間の視覚特性を考慮した処理も可能である[39]。しかし本研究では、衛星画像のように物理世界の性質がある程度既知の場合には、画素値の特徴空間の統計的性質を用いた画素単位の統計的な分類手法を用いる。特に第3章～第5章では、「ミクセル」という現象を考慮した統計的分類法を提案し、従来の研究にはない新しい考え方に基づいた分類手法を提案する。一方統計的性質が複雑な一般の自然画像に対しては、濃度ヒストグラムの形状を利用した領域分割手法を応用する。

領域レイヤ 下層の「画素レイヤ」でおおよそ分割された画素集合、すなわち領域に対して、その構造や特徴量を計測して領域情報を得ることが目的である。パターン認識手法で用いられる形状特徴量などが最も単純な出力表現であるが、領域の構造を捉えたい場合にはさらに複雑な処理、例えば形状分解(領域に基づいた方法)や多角形近似(輪郭線に基づいた手法)などの手法が強力である。本研究では変形モデルに基づいた形状分解の手法を第6章で提案し、この手法によって衛星画像をプリミティブの集合として出力し領域情報を得る。

関係レイヤ 下位レイヤで得られた種々の画像特徴を1つの画像表現モデルに統合し、画像全体をこのモデルで統一的に記述することが目的である。そして、下位レイヤで得られたプリミティブ間の関係をどのようにモデル化し計算するかが課題である。もし領域間やプリミティブ間に隣接関係が定義できればそれが最も単純な関係となる。しかし必ずしも隣接していないプリミティブ間にも、何らかの遠隔作用をモデル化した関係を定義することができる。本研究では物理的なアナロジーに基づいたモデルを第7章で提案し、また隣接関係を用いた手法についても検討する。またこのレイヤの出力表現にはモデリング能力に優れたグラフ構造を用いる。

認識レイヤ 下位レイヤで得られた画像表現モデルを用いて画像間の類似度を計算することが目的である。この画像間類似度は画像表現モデルに依存して定義されるものであるが、本研究では画像表現モデルとしてグラフ構造を用いるため、その類似度をグラフマッチングコストと定義する。このグラフマッチングは計算量が大きい問題であることから、その高速化はこのレイヤの重要な課題であるが、第8章で述べるように、最大許容コストやヒューリスティック関数の導入などという種々の工夫によって、マッチングを高速化できることが示されている。

理解レイヤ ユーザの検索目的を類似度に反映させることで柔軟な類似画像検索を実現することがこのレイヤの目的である。言い換えると、ユーザの検索目的を最適に反映するような距離空間の構成法が問題となる。ユーザの検索目的を柔軟に反映させるための方法として、本論文では教師付き学習をおこなう。すなわち人間が類似画像として与えた教師データを類似検索結果が最適に反映するように、検索パラメータを最適化させる問題を解く。この最適化問題の解法として、第9章で述べるような、人為選択法を応用した遺伝的アルゴリズムを用いた。この手法で検索パラメータを最適化した結果、人間の選んだ主観的類似画像を上位に類似検索できるような検索パラメータを学習することができることを示している。

表 2.3: 本研究で用いた手法と階層的モデルとの対応関係。

レイヤ	二値画像 NOAA 衛星雲画像	濃淡画像 GMS 衛星雲画像	カラー画像 自然画像
理解レイヤ	教師付き学習・遺伝的アルゴリズム (第9章～第10章)		
認識レイヤ	グラフマッチング・A* アルゴリズム (第8章)		
関係レイヤ	平面的な引力モデル (第7章)	階層的な引力モデル (第7章)	隣接関係 (第7章)
領域レイヤ	形状分解 (楕円) (第6章)		領域特徴量の計測 (第7章)
画素レイヤ	二値化	ミクセルを含む混合密度 推定による画素分類手法 (第3章～第5章)	Ohlander に基づいた 方法 (第B章)
観測レイヤ	衛星センサ (赤外 / 可視) (第2章)		カメラ

2.4.3 階層モデルと本研究で用いた手法の対応関係

表 2.2 で述べた階層モデルを基本として、各画像データベースに適した手法を個別に検討した。その詳細については以下の章や第 10.2 節で述べることにするが、ここではまず全体像を捉えるために表 2.3 に本研究で用いた手法を階層モデルの形にしてまとめた。下層のレイヤでは問題領域に応じた手法を用いているが、関係レイヤの出力表現である階層化属性つき関係グラフに至り、どの対象画像も共通の画像表現モデルに変換される。このように上位層ではグラフ構造を用いた統一的な扱いが可能である⁵。

また、自然画像データベースと 2 種類のかなりの性質の異なる画像を研究対象とするのは、階層モデルを 2 種類の画像データベースに適用することで、階層モデルの基盤をより強固に形成することも目的の一つである。このように階層モデルをさまざまな具体的な例に適用していくことで、その利点も欠点もより明確になることだろう。

⁵しかし細かい相違点はもちろん存在し、例えば属性値の取り扱いなどは問題領域に依存して定義されねばならない。しかし大きく見れば統一的に取り扱えるとしてよいだろう。

第3章

L型やJ型・U型の形状をもつミクセルの面積占有率密度

3.1 第3章から第5章までの関係

以下の3章ではミクセルを含む画像の統計的な解析手法について、本研究が提案する手法を述べる。この手法では統計学的な理論を展開するが、その過程で必要となるさまざまな確率モデルには似たような関数名が多く出現し、また関数どうしがお互いにどのように依存しているかは、それほど明瞭ではないと感じられる。そこで図3.1には、以下の3章で登場する関数の相互関係を図示し、合わせてそれぞれの関数が何章で取り上げられるかを示した。図3.1では、解析のもとになる確率モデルを短い破線で、それらから得られる確率モデルを長い破線で、本研究で提案する確率モデルを実線で、最後に本研究で得られる最終結果を一点鎖線でしめした。その最終目的とは、各画素をピュアピクセルかまたはミクセルに分類し、なおかつミクセルに関しては面積占有率を計算するという結果を得ることである。そのためにいくつかの中間的な計算が必要となることを、以下の章で説明していく。以下の記述で議論の全体像が見えなくなった場合には、常にこの図を参照していただきたい。

3.2 概要

ミクセルとは一画素の内部に複数の分類クラスを含む画素を指す。そしてミクセルの内部で各分類クラスが占める割合を面積占有率とよぶ。この面積占有率は、画像中のミクセルの集合を考えるとある確率密度(面積占有率密度)にしたがって分布している。この密度関数は、各クラスが画像中に占める面積を求める問題や、個々のミクセルの面積占有率を計算する問題に必要とされる重要な関数である。しかし従来の研究ではこの面積占有率密度は暗黙のうちに一様分布と仮定される場合が多く、その形状についてはほとんど研究されていない。そこで本研究では、まず簡単なモデルを用いて解析的に面積占有率密度を導出し、それが一般に一様分布とはならずL型やJ型の分布となることを示す。次にフラクタル合成画像によるシミュレーションを用いて面積占有率密度を推定し、やはりそれが一様分布とならずにU型の分布となることを示す。最後に4種の異なる面積占有率密度の間の相互関係を考察し、面積占有率密度の応用について述べる。

3.3 従来研究の問題点

混合画素(mixed pixel)またはミクセルとは一画素の内部に複数の分類クラスを含む画素のことを指す用語である。このようなミクセルを扱う研究としては、個々のミクセルの面積占有率

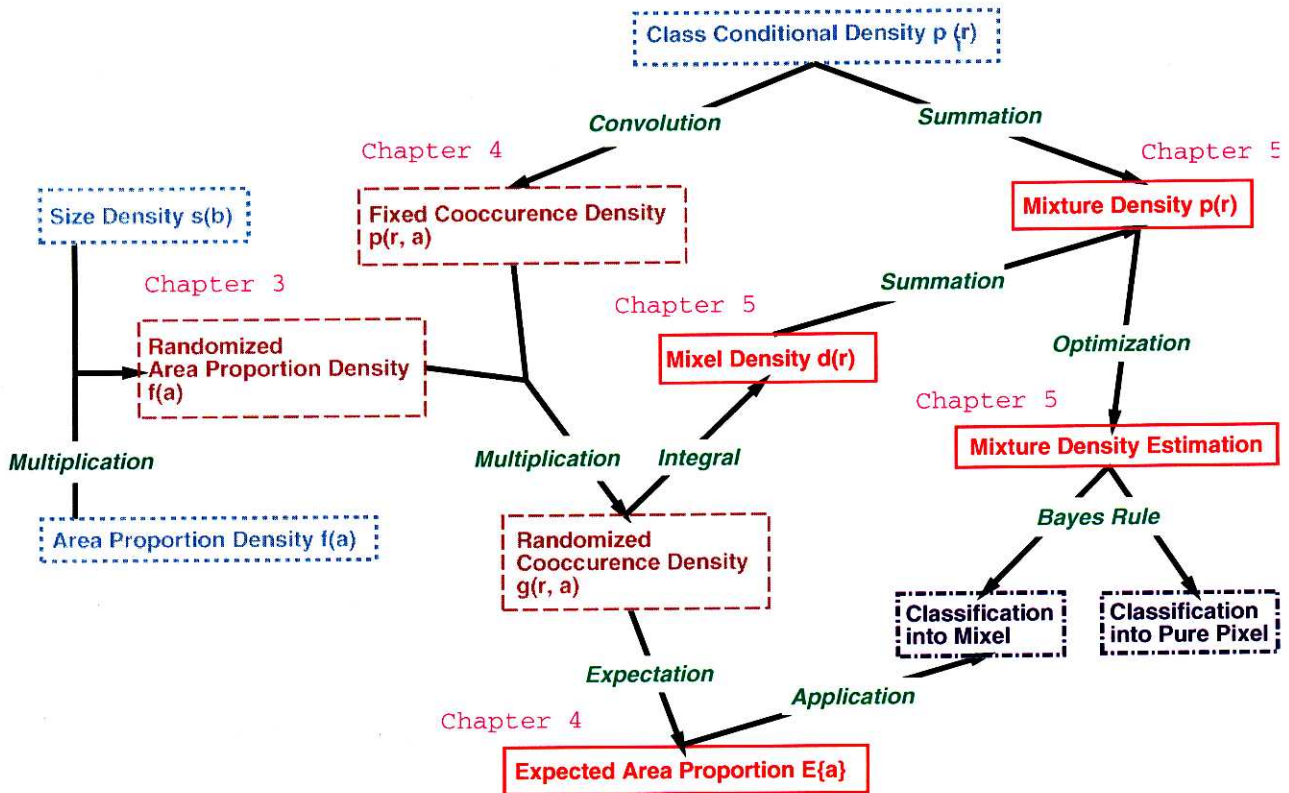


図 3.1: ミクセルを含む画像の統計的な解析手法に用いられる確率モデルの関係。

を求める方法に研究の重点が置かれてきた。この方法については、本論文でものちほど第4章で述べる。しかし一方で、画像全体のミクセルの集合を考えた場合に、ミクセルの面積占有率の分布、すなわち面積占有率密度を求める問題も重要である。この密度関数は、各クラスが画像中に占める面積を求める問題や、個々のミクセルの面積占有率を計算する問題に必要となる。しかし従来の研究ではこの密度関数の重要性がさほど認識されておらず、多くの研究ではその密度関数を暗黙のうちに一様分布と仮定してきた。本研究ではこの仮定が適切であるかを検証する。まず、簡単なモデルから解析的な計算によって導いた面積占有率密度 (3.5章) と、実際の画像に類似したフラクタル合成画像によるシミュレーションから計算した面積占有率密度 (3.6章) とを求め、実際には面積占有率密度は一様分布とならないことを示す。また画像中で各クラスが占める面積比の計算では、面積占有率密度の仮定によって誤差が生じることを指摘する。

3.4 面積占有率密度

図 3.2 にミクセルの例として 2 クラスミクセルを示す。このようにミクセルが k 個のクラスから構成されているとき、これを k クラスミクセルとよぶ。 k クラスミクセルを構成する各クラスの (画素内) 面積占有率を a_i とすると、各クラス i について $0 < a_i < 1$ および $\sum_{i=1}^k a_i = 1$ が成立する。ここで k クラスミクセルの面積占有率の自由度が $k - 1$ であることから、 $a_k = 1 - \sum_{i=1}^{k-1} a_i$ とし $k - 1$ 次元の面積占有率ベクトル $\mathbf{a}' = (a_1, \dots, a_{k-1})^t \in \Psi$ を考える。ここで \mathbf{a}' が存在するパラメータ空間 $\Psi \subseteq R^{k-1}$ は $\Psi = \{\mathbf{a}' \mid \sum_{i=1}^{k-1} a_i < 1, a_i > 0\}$ と定義される。 \mathbf{a}' はこのパラメータ空間 Ψ の内部で、ある確率密度にしたがう確率変数ベクトルであると仮定し、この確率密度を面積占有率密度 $f(\mathbf{a}')$ とよぶ。

一般的にはこの面積占有率密度は暗黙のうちに一様分布と仮定されている。つまり $f(\mathbf{a}') =$

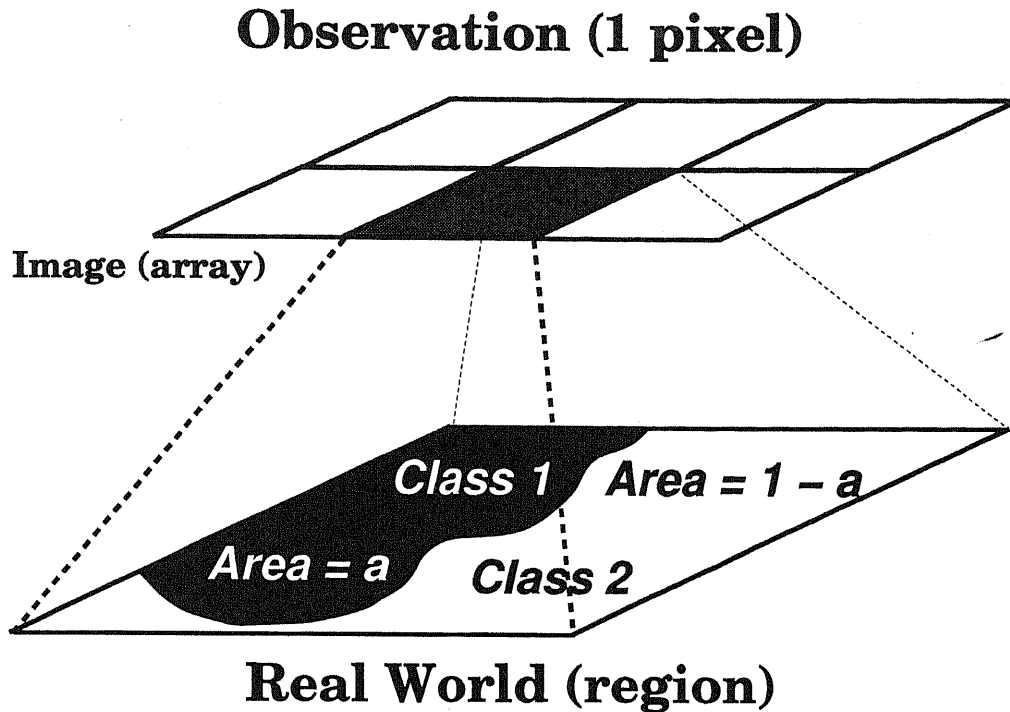


図 3.2: 2 クラスミクセルのモデル。

$1/|\Psi|, |\Psi| = \iint_{\Psi} da'$ という仮定である。例えば a' が一次元ベクトルであれば $f(a') = 1$ である。直観的にはどの面積占有率の組み合わせも等確率で出現し、面積占有率密度は一様分布になるとの仮定は成立するように思える。しかしこの仮定は、実際の画像に出現しているミクセルに対して検証されたわけではない。そして実際の画像に対してこの仮定の正当性を直接検証することは、真のデータが不明である以上かなり困難な課題でもある。そこでその代替手段として、シミュレーションを用いた面積占有率密度の推定が研究されている。例えば雲を想定した面積占有率密度の推定として、正方形を観測平面上にランダムに配置したシミュレーションなどの研究がある [42]。しかし本研究ではまず簡単なモデルを用いて解析的に面積占有率密度を導き、シミュレーションに比べてより精密な議論を展開する。しかし解析的に計算できるのはあくまで単純な場合だけであるから、次に実世界に近い状況を合成したシミュレーションによって、モデルから得られた結論を補強する。これらの結果から得られた面積占有率密度を比較検討し、結論として面積占有率密度の形状は一様分布とならないことを示す。

なお本研究の前提には注意すべき点がある。まず本研究では2クラスミクセルの面積占有率密度のみを扱う。またセンサの検知素子が1画素として観測する瞬時視野の形状を楕円形ではなく正方形と近似することで、解析的な計算を容易としている。

3.5 正方形モデルから導出した面積占有率密度

最も簡単なモデルとして、観測平面上に置かれた大きさ $b \times b$ の正方形領域を、大きさ $c \times c$ の瞬時視野に対応する画素で観測する場合を考える。そして正方形内部の領域をクラス1、それ以外の領域をクラス2とし、クラス1が画素の内部に占める面積の割合を面積占有率 a とする。 $0 < a < 1$ である。 $c < b$ の場合は、正方形を含む画素は、クラス1とクラス2の2クラスミクセルかピュアピクセルのどちらかであり、一方 $b < c$ の場合はすべて2クラスミクセルである。

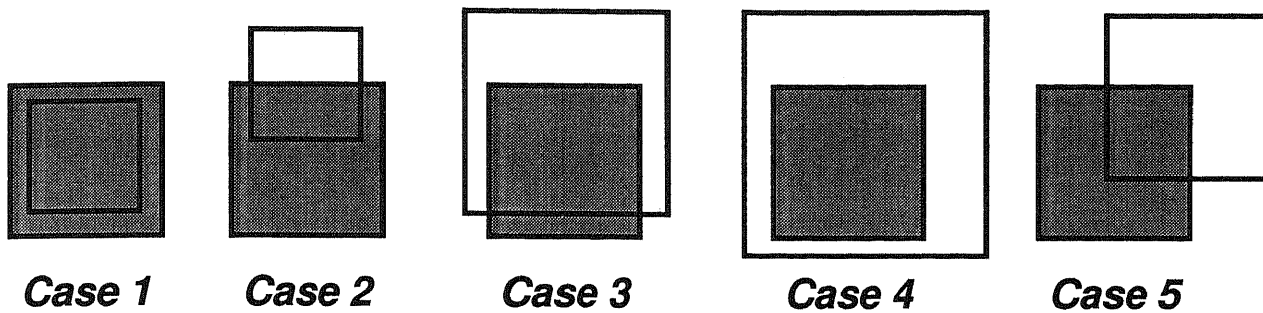


図 3.3: 画素と正方形モデルの位置関係による 5 個の場合への場合分け。網かけされた正方形は正方形モデルを表し、もう一方の正方形が画素を表している。

表 3.1: 正方形モデルを用いた場合の、Case 1 ~ Case 5 の場合の面積占有率密度の導出。

Case	$F(a)$ または $G(a)$	$f(a)$	b の範囲	a の範囲
1	—	—	—	—
2	—	$f(a) = 1$	$c < b$	$0 < a < 1$
3	—	$f(a) = \frac{c^2}{b^2}$	$b < c$	$0 < a < \frac{b^2}{c^2}$
4	—	$f(a) = \delta \left(a - \frac{b^2}{c^2} \right)$	$b < c$	—
5-a	$G(a) = \frac{1}{c^2} \int_0^{c(1-a)} \left\{ \frac{ac^2}{x-c} + c \right\} dx$	$f(a) = -\log a$	$c < b$	$0 < a < 1$
5-b	$G(a) = \frac{1}{b^2} \int_{c-b}^{c-\frac{ac^2}{b^2}} \left\{ \frac{ac^2}{x-c} + c \right\} dx$	$f(a) = -\frac{c^2}{b^2} \log \left(\frac{c^2}{b^2} a \right)$	$0 < b < c$	$0 < a < \frac{b^2}{c^2}$

3.5.1 画素と正方形モデルの位置関係で場合分けした面積占有率密度

さてこの正方形モデルを多数回観測することを考える。そしてその 1 回ごとの観測において、画素の中心と正方形の中心とのずれの分布が一様分布にしたがうと仮定する。ある特定の位置ずれの場合の面積占有率は簡単に計算できるため、この結果と一様分布の仮定とを併せることで、すべての可能な位置ずれを考慮した面積占有率を導くことができる。ここで重要なのは、位置ずれの分布が一様分布であるという仮定が、面積占有率密度も一様分布であるという結論に直ちに結び付くわけではないことである。

そこでまず、画素と正方形モデルとがどのような位置関係になるかを調べる。画素と正方形モデルが重なり合わない場合は画素がミクセルとならないため考慮する必要はなく、それ以外の重なり合う場合の位置関係を整理すると、結局図 3.3 のように 5 個の場合に集約できる。この中で Case 1 はピュアピクセルとなるため考慮する必要はなく、その他の場合について面積占有率密度を計算した結果を表 3.1 に述べた。以下でその計算方法を詳しく述べる。

Case 1

この場合は $c < b$ のときに生起する。この場合は画素が正方形モデルのピュアピクセルとなるため、面積占有率密度の計算からは除外する。画素の大きさに比べて正方形モデルの方が大きい場合には、画素はピュアピクセルかミクセルのどちらかとなる。逆の場合はすべてミクセルである。

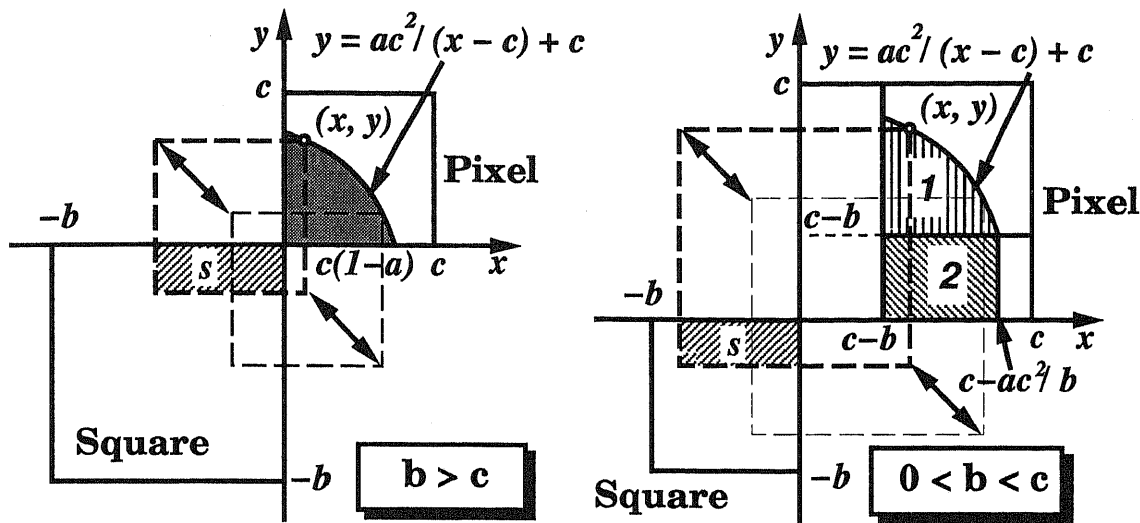


図 3.4: Case 5 の場合の面積占有率密度の計算方法。

Case 2

この場合は $c < b$ のときに生起する。図 3.3 では縦方向の移動のみが面積占有率を変化させるため、面積占有率密度は一様分布となる。面積占有率の定義域が $(0 < a < 1)$ であることから、面積占有率密度は、

$$f_2(a) = 1 \tag{3.1}$$

Case 3

この場合は $b < c$ のときに生起する。Case 2 と同様、図 3.3 の場合は縦方向の移動のみが面積占有率を変化させるため、面積占有率密度は一様分布となる。面積占有率の定義域が $(0 < a < b^2/c^2)$ であることから、

$$f_3(a) = \frac{c^2}{b^2} \tag{3.2}$$

Case 4

この場合は $b < c$ のときに生起する。縦方向、横方向のどちらの移動も面積占有率を変化させないため、面積占有率は一定である。すなわち面積占有率密度は、以下のようなデルタ関数で表すことができる。

$$f_4(a) = \delta\left(a - \frac{b^2}{c^2}\right) \tag{3.3}$$

Case 5

まず座標系を図 3.4 のように設定し、画素の左下隅の座標を (x, y) とする。このとき画素の内部に含まれる正方形モデルの面積は $s = (x - c)(y - c)$ である。ここで画素の面積が c^2 であることから、正方形モデルの面積は面積占有率を a を用いて $s = ac^2$ とも表すことができる。ゆえに $ac^2 = (x - c)(y - c)$ である。この式を y に関して解くと次の式を得る。

$$y = \frac{ac^2}{x - c} + c \tag{3.4}$$

ここで Case 5 をさらに二つの場合に分類して考察する。

$c < b$ の場合 この場合の (x, y) の定義域は $(0 < x, y < c)$ である。さて面積占有率が a 以上であるという事象の確率 $Pr(\{A \geq a\})$ を a の関数とみて、累積分布関数 $G(a) = Pr(\{A \geq a\})$ を定義する。すると $G(a)$ は図 3.4 の灰色部分の面積の割合となるから

$$\begin{aligned} G(a) &= \frac{1}{c^2} \int_0^{c(1-a)} \left\{ \frac{ac^2}{x-c} + c \right\} dx \\ &= 1 - a + a \log a \end{aligned} \quad (3.5)$$

ここで面積占有率が a 以下である事象の確率 $Pr(\{A \leq a\})$ の累積分布関数 $F(a)$ は $F(a) = 1 - G(a)$ である。そして求める面積占有率密度 $f(a)$ は、累積分布関数 $F(a)$ の微分として得られる [43]。

$$f_5^1(a) = \frac{dF(a)}{da} = -\log a \quad (3.6)$$

ただし a の定義域は $0 < a < 1$ である。

$0 < b < c$ の場合 この場合の (x, y) の定義域は $(c-b < x, y < c)$ である。累積確率分布 $G(a)$ は図 3.4 の領域 1 の面積に相当し、これは領域 (1+2) の面積から領域 2 の面積を引いたものである。そこで累積確率分布は

$$\begin{aligned} G(a) &= \frac{1}{b^2} \int_{c-b}^{c-\frac{ae^2}{b}} \left\{ \frac{ac^2}{x-c} + c \right\} dx \\ &= \frac{b^2 - ac^2}{b^2} + \frac{ac^2}{b^2} \log \left(\frac{ac^2}{b^2} \right) \end{aligned} \quad (3.7)$$

この場合も $F(a) = 1 - G(a)$ の微分により面積占有率密度 $f(a)$ を求めることができる。

$$f(a) = \frac{dF(a)}{da} = -\frac{c^2}{b^2} \log \left(\frac{c^2}{b^2} a \right) \quad (3.8)$$

ただし面積占有率の定義域は $(0 < a < b^2/c^2)$ である。

結局、Case 5 の面積占有率密度は以下のようになる。

$$f_5(a) = \begin{cases} -\log a & (c < b) \\ -\frac{c^2}{b^2} \log \left(\frac{c^2}{b^2} a \right) & (0 < b < c) \end{cases} \quad (3.9)$$

式 (3.9) は、面積占有率が小さい領域が高い割合を占める密度関数となっている。この形の面積占有率密度が存在するために、面積占有率密度は一様分布とはならない。

3.5.2 それぞれの場合の生起確率

正方形モデルと重なり合う範囲での画素の移動可能範囲の面積を S とする。また各場合に対する同様の面積を S_i とすると、画素と正方形モデルとの位置ずれの分布が一様分布にしたがうとの仮定のもとでは、それぞれの場合の生起確率は S_i/S として計算できる。このように計算した生起確率を表 3.2 にまとめた。この場合は $S = (b+c)^2$ である。全体の面積占有率密度は、各場合の面積占有率密度を生起確率で重みづけした重みつき平均として得られる。すなわち Case

表 3.2: 画素と正方形モデルの位置関係による 5 個の場合それぞれの生起確率。表中で面積 S は $S = (b + c)^2$ である。

	$c < b$	$0 < b < c$
Case 1 (t_1)	$(b - c)^2/S$	0
Case 2 (t_2)	$4c(b - c)/S$	0
Case 3 (t_3)	0	$4b(c - b)/S$
Case 4 (t_4)	0	$(c - b)^2/S$
Case 5 (t_5)	$4c^2/S$	$4b^2/S$

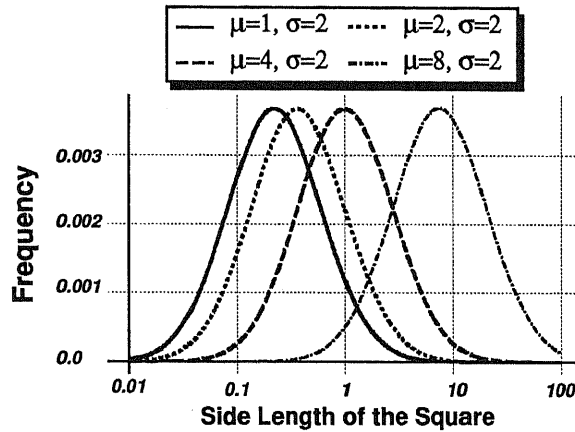


図 3.5: サイズ密度として用いる対数正規密度。

1 から Case 5 の生起確率を t_1, \dots, t_5 とすると、面積占有率密度 $h(a, b, c)$ は以下の式で求められる。

$$h(a, b, c) = \frac{\sum_{i=1}^5 t_i f_i(a)}{\sum_{i=1}^5 t_i} \tag{3.10}$$

3.5.3 サイズ密度によるランダム化

さらに現実に近い状況を想定するために、観測平面上にさまざまな大きさをもつ正方形が分布している状況を考える。つまり正方形モデルの大きさをパラメータとし、このパラメータがある確率密度にしたがって分布していると仮定する。この確率密度をサイズ密度 $s(x)$ とよぶ。サイズ密度に用いる密度関数の候補としては、対数正規密度 [44] やべき乗密度 [45] などが提案されているが、本研究では対数正規密度を用いる。

$$s(x) = \frac{1}{\sigma x \sqrt{2\pi}} \exp \left\{ -\frac{\ln x - \mu}{2\sigma^2} \right\} \tag{3.11}$$

対数正規密度の関数形を図 3.5 に図示する。この図の横軸 (対数尺度) には正方形の 1 辺の長さ b を取っており、したがって図 3.5 は $s(b^2)$ の密度関数となっていることに注意する。また横軸は対数軸となっている。このように正方形モデルの 1 辺の長さは $10^{-2} < b < 10^2$ の間に分布していると仮定した。

このサイズ密度を用いると、正方形モデルの大きさをパラメータとみなした場合の面積占有

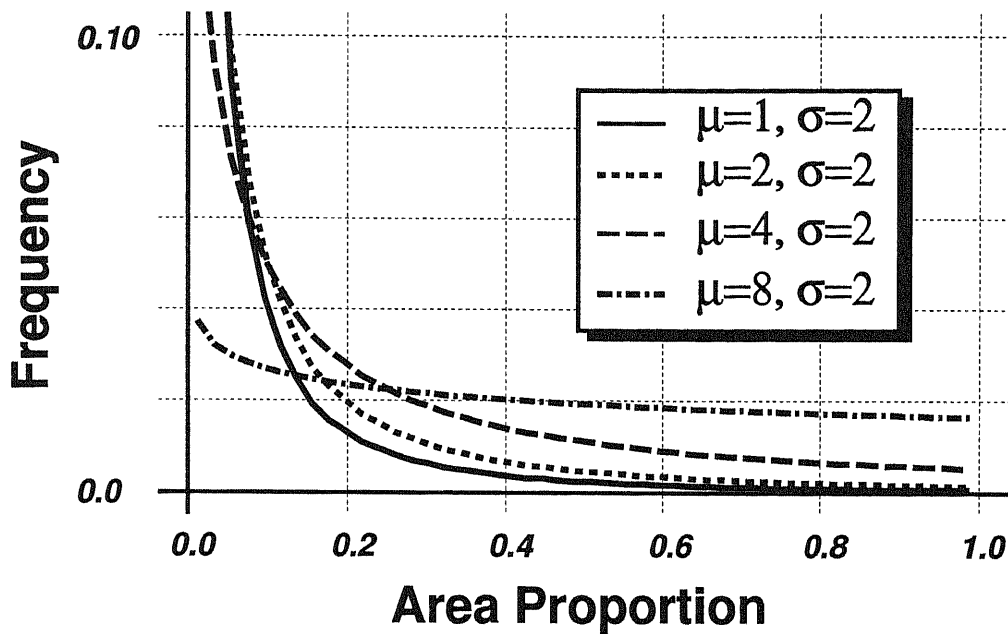


図 3.6: 正方形モデルから導出した L 型面積占有率密度。

率密度 $f(a)$ を以下のようなランダム化 [43] により求められる。

$$f(a) = \int_0^{\infty} h(a, b, 1) s(b^2) db \quad (3.12)$$

この面積占有率密度 $f(a)$ をランダム化された面積占有率密度とよぶ。なお式 (3.12) では画素の大きさ c の値そのものは重要ではなく、 b と c との大きさの比が重要であるため、 $c = 1$ に固定する。式 (3.12) にパラメータを代入し、数値計算して得られた面積占有率密度の形状を図 3.6 に示す。パラメータの値によって形状は多少異なるが、どの曲線も a の増加にしたがって単調減少するという共通の特徴を備えている。この密度関数をその形状から L 型面積占有率密度とよぶ。

ただし以上の計算においてどの正方形も最低 1 画素以上の間隔で隔たっているとの仮定をしており、より現実的なモデルとするには複数の正方形が 1 画素に含まれる可能性を考慮しなければならない。このような場合を解析するためには、正方形モデルの観測平面上での分布が必要となる。この分布はポアソン分布などでモデル化できるが、その解析は非常に複雑となるため今後の課題である。

3.6 十字形モデルから導出した面積占有率密度

正方形モデルは完全に凸図形であり、凹の部分がないという点でモデルとして不十分である。そこで本章では凹凸図形の最も単純なモデルとして、正方形を 5 個張り付けた十字形をモデルとする。

3.6.1 画素と十字形モデルの位置関係で場合わけした面積占有率密度

この場合に画素と十字形モデルとの位置関係を整理すると図 3.3 の 15 個の場合に集約できる。この中で Case 1 はピュアピクセルとなるため考慮する必要はなく、その他の場合について面積占有率密度を計算した結果を表 3.3 に述べた。以下でそれぞれの場合の面積占有率密度の計算方法を詳しく述べる。

表 3.3: 十字形モデルを用いたときの、Case 1 ~ Case 15 の面積占有率密度の導出。

Case	$F(a)$ または $G(a)$	$f(a)$	b の範囲	a の範囲
1	—	—	—	—
2	—	$f(a) = 1$	$c < b$	$0 < a < 1$
3	—	$f(a) = \frac{c^2}{b^2}$	$b < c$	$0 < a < \frac{b^2}{c^2}$
4	—	$f(a) = \delta \left(a - \frac{5b^2}{c^2} \right)$	$b < \frac{1}{3}c$	—
5-a	$G(a) = \frac{1}{c^2} \int_0^{c(1-a)} \left\{ \frac{ac^2}{x-c} + c \right\} dx$	$f(a) = -\log a$	$c < b$	$0 < a < 1$
5-b	$G(a) = \frac{1}{b^2} \int_{c-b}^{c-\frac{ac^2}{b^2}} \left\{ \frac{ac^2}{x-c} + c \right\} dx$	$f(a) = -\frac{c^2}{b^2} \log \left(\frac{c^2}{b^2} a \right)$	$0 < b < c$	$0 < a < \frac{b^2}{c^2}$
5'-a	—	$f(a) = -\log(1-a)$	$c < b$	$0 < a < 1$
5'-b	—	$f(a) = -\frac{c^2}{b^2} \log \left(\frac{c^2(1-a)}{b^2} \right)$	$\frac{1}{2}c < b < c$	$0 < a < \frac{b^2}{c^2}$
6-a	$F(a) = \frac{1}{(c-b)^2} \int_0^{\frac{ac^2}{b^2}} \left\{ \frac{ac^2+b^2}{x+b} - b \right\} dx$	$f(a) = \frac{c^2}{(c-b)^2} \log \left(\frac{ac^2+b^2}{b^2} \right)$	$\frac{1}{2}c < b < c$	$0 < a < \frac{b(c-b)}{c^2}$
6-b	$F(a) = \frac{1}{(c-b)^2} \left[\left(\frac{ac^2+b^2}{c} - b \right) (c-b) + \int_{\frac{ac^2+b^2}{c}}^{c-b} \left\{ \frac{ac^2+b^2}{x+b} - b \right\} dx \right]$	$f(a) = \frac{c^2}{(c-b)^2} \log \left(\frac{c^2}{ac^2+b^2} \right)$	$\frac{1}{2}c < b < c$	$\frac{b(c-b)}{c^2} < a < \frac{c^2-b^2}{c^2}$
6-c	$F(a) = \frac{1}{b^2} \int_0^{\frac{ac^2}{b^2}} \left\{ \frac{b^2+ac^2}{x+b} - b \right\} dx$	$f(a) = \frac{c^2}{b^2} \log \left(\frac{ac^2+b^2}{b^2} \right)$	$0 < b < \frac{1}{2}c$	$0 < a < \frac{b^2}{c^2}$
6-d	$F(a) = \frac{1}{b^2} \left[\left(\frac{ac^2-b^2}{2b} \right) b + \int_{\frac{ac^2-b^2}{2b}}^b \left\{ \frac{ac^2+b^2}{x+b} - b \right\} dx \right]$	$f(a) = \frac{c^2}{b^2} \log \left(\frac{4b^2}{ac^2+b^2} \right)$	$0 < b < \frac{1}{2}c$	$\frac{b^2}{c^2} < a < \frac{3b^2}{c^2}$
7-a	$F(a) = \frac{1}{2(c-b)(2b-c)} \left(-ac - b - \frac{b^2}{c} \right) \left(b + \frac{ac^2}{b} - c \right)$	$f(a) = \frac{c}{c-b} \frac{ac^2 - b(c-b)}{b(2b-c)}$	$\frac{1}{2}c < b < c$	$\frac{b(c-b)}{c^2} < a < \frac{b^2}{c^2}$
7-b	$F(a) = \frac{2ac-b}{2(c-b)}$	$f(a) = \frac{c}{c-b}$	$\frac{1}{2}c < b < c$	$\frac{b^2}{c^2} < a < 1 - \frac{b^2}{c^2}$
7-c	$G(a) = \frac{1}{2(c-b)(2b-c)} \left(c(1-a) - b + \frac{b^2}{c} \right) \left(b + \frac{(1-a)c^2}{b} - c \right)$	$f(a) = \frac{c}{c-b} \frac{(1-a)c^2 - b(c-b)}{b(2b-c)}$	$\frac{1}{2}c < b < c$	$1 - \frac{b^2}{c^2} < a < 1 - \frac{b(c-b)}{c^2}$
8	—	$f(a) = \frac{c^2}{(2b-c)(c-b)}$	$\frac{1}{2}c < b < c$	$\frac{b^2+c(c-b)}{c^2} < a < \frac{(2c-b)b}{c^2}$
9	—	$f(a) = \delta \left(a - \frac{(2c-b)b}{c^2} \right)$	$\frac{1}{2}c < b < c$	—
10-a	—	$f(a) = \frac{c}{c-b}$	$\frac{1}{2}c < b < c$	$\frac{b^2}{c^2} < a < \frac{b^2+c(c-b)}{c^2}$
10-b	—	$f(a) = \frac{c}{b}$	$\frac{1}{3}c < b < \frac{1}{2}c$	$\frac{b^2}{c^2} < a < \frac{b(b+c)}{c^2}$
11-a	$F(a) = \frac{1}{b(c-2b)} \left[\left(\frac{ac^2-b^2}{c} \right) (c-2b) + \int_{\frac{ac^2-b^2}{c}}^{\frac{ac^2-b^2}{2b}} \left\{ \frac{ac^2-b^2}{x} - 2b \right\} dx \right]$	$f(a) = \frac{c^2}{b(c-2b)} \log \left(\frac{c}{2b} \right)$	$b < \frac{1}{2}c$	$\frac{b^2}{c^2} < a < \frac{3b^2}{c^2}$
11-b	$F(a) = \frac{1}{b(c-2b)} \left[\left(\frac{ac^2-b^2}{c} \right) (c-2b) + \int_{\frac{ac^2-b^2}{c}}^b \left\{ \frac{ac^2-b^2}{x} - 2b \right\} dx \right]$	$f(a) = \frac{c^2}{b(c-2b)} \log \left(\frac{bc}{ac^2-b^2} \right)$	$b < \frac{1}{2}c$	$\frac{3b^2}{c^2} < a < \frac{b(b+c)}{c^2}$
12	—	$f(a) = \frac{c^2}{b(c-2b)}$	$\frac{1}{3}c < b < \frac{1}{2}c$	$\frac{b(b+c)}{c^2} < a < \frac{b(2c-b)}{c^2}$
13	—	$f(a) = \frac{c^2}{b^2}$	$b < \frac{1}{3}c$	$\frac{4b^2}{c^2} < a < \frac{5b^2}{c^2}$
14	—	$f(a) = \frac{c^2}{3b^2}$	$b < \frac{1}{2}b$	$\frac{3b^2}{c^2} < a < \frac{5b^2}{c^2}$
15	—	$f(a) = \frac{c^2}{2b^2}$	$b < \frac{1}{2}b$	$\frac{3b^2}{c^2} < a < \frac{5b^2}{c^2}$

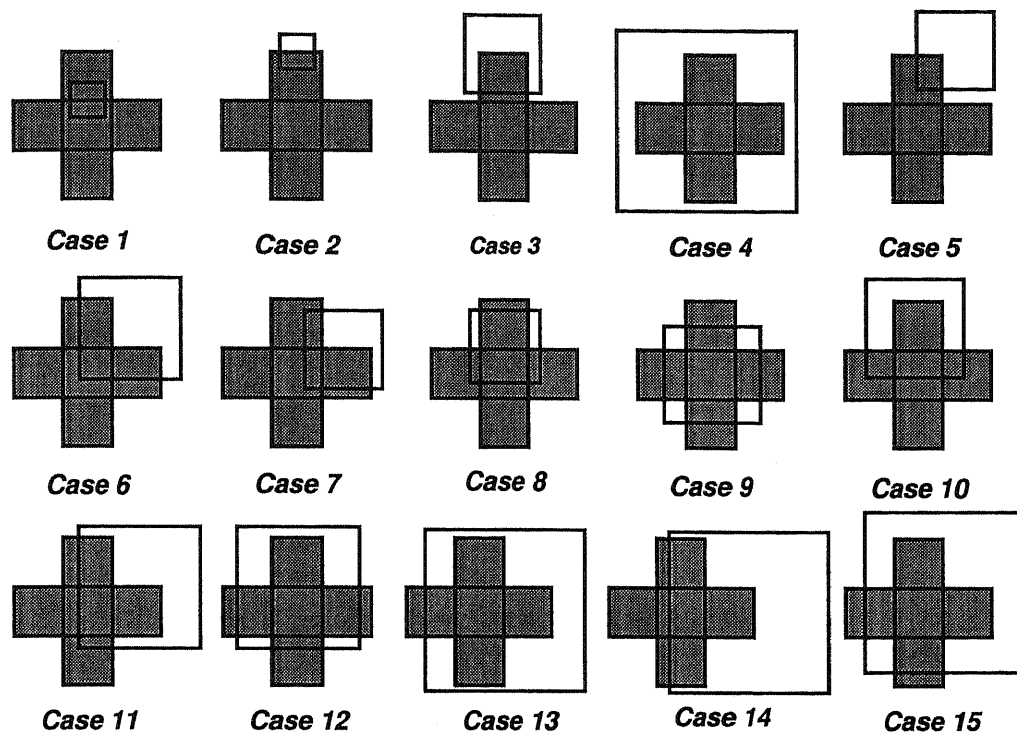


図 3.7: 画素と十字形モデルの位置関係による 15 個の場合への場合わけ。網かけされた領域は十字形モデルを表し、もう一方の正方形が画素をあらわしている。

Case 1 ~ Case 5

基本的に前節の結果をそのまま適用できる。また Case 1 はミクセルではない。以下で前節の結果との相違点を述べる。まず Case 4 の場合は前節とは異なり、 $b < \frac{1}{3}c$ のときに生起する。また面積占有率密度は以下となる。

$$f_4(a) = \delta \left(a - \frac{5b^2}{c^2} \right) \tag{3.13}$$

次に新たな場合として Case 5' を導入する。これは十字形の凹部分に対応する場合であり、Case 5 の場合と比べてちょうどクラス 1 とクラス 2 の立場を入れ替えた場合に相当する。そのためこの場合の面積占有率密度は式 (3.9) の a に $(1 - a)$ を代入すれば求まる。

$$f_{5'}(a) = \begin{cases} -\log(1 - a) & (c < b) \\ -\frac{c^2}{b^2} \log \left(\frac{c^2(1-a)}{b^2} \right) & (\frac{1}{2}c < b < c) \end{cases} \tag{3.14}$$

Case 6

まず座標系を図 3.8 のように設定し、画素の右上隅の座標を (x, y) とする。このとき画素の内部に含まれる十字形モデルの面積は $s = (x + b)(y + b) - b^2$ である。 $s = ac^2$ の関係式を代入してこの式を y について解くと

$$y = \frac{s + b^2}{x + b} - b \tag{3.15}$$

という式を得る。ここで Case 6 をさらに 4 個の場合に分類して考察する。

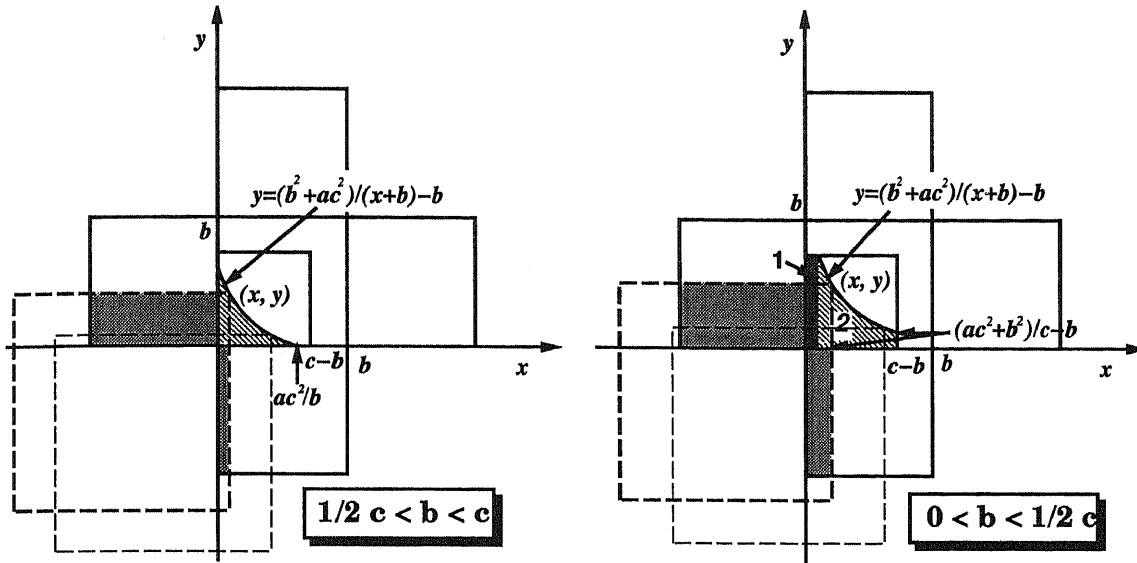


図 3.8: Case 6 の場合の面積占有率密度の計算方法。

$\frac{1}{2}c < b < c$ かつ $0 < a < b(c-b)/c^2$ のとき (x, y) の定義域は $(0 < x, y < c-b)$ である。累積確率分布 $F(a)$ の値は図 3.8 の斜線部の面積である。

$$F(a) = \frac{1}{(c-b)^2} \int_0^{\frac{ac^2}{b}} \left\{ \frac{ac^2 + b^2}{x+b} - b \right\} dx$$

$$f_6^1(a) = \frac{c^2}{(c-b)^2} \log \left(\frac{ac^2 + b^2}{b^2} \right) \tag{3.16}$$

$\frac{1}{2}c < b < c$ かつ $b(c-b)/c^2 < a < (c^2 - b^2)/c^2$ のとき (x, y) の定義域は $(0 < x, y < c-b)$ である。累積確率分布 $F(a)$ の値は図 3.8 の領域 1 と領域 2 の面積の和である。

$$F(a) = \frac{1}{(c-b)^2} \left[\left(\frac{ac^2 + b^2}{c} - b \right) (c-b) + \int_{\frac{ac^2 + b^2}{c} - b}^{c-b} \left\{ \frac{ac^2 + b^2}{x+b} - b \right\} dx \right]$$

$$f_6^2(a) = \frac{c^2}{(c-b)^2} \log \left(\frac{c^2}{ac^2 + b^2} \right) \tag{3.17}$$

以上の計算から、 $\frac{1}{2}c < b < c$ の場合の面積占有率密度は、 $f_6^1(a)$ と $f_6^2(a)$ とを接続した関数となる。なお a の定義域は $(0 < a < (c^2 - b^2)/c^2)$ である。

$0 < b < \frac{1}{2}c$ かつ $0 < a < b^2/c^2$ のとき (x, y) の定義域は $(0 < x, y < b)$ である。

$$F(a) = \frac{1}{b^2} \int_0^{\frac{ac^2}{b}} \left\{ \frac{b^2 + ac^2}{x+b} - b \right\} dx$$

$$f_6^3(a) = \frac{c^2}{b^2} \log \left(\frac{ac^2 + b^2}{b^2} \right) \tag{3.18}$$

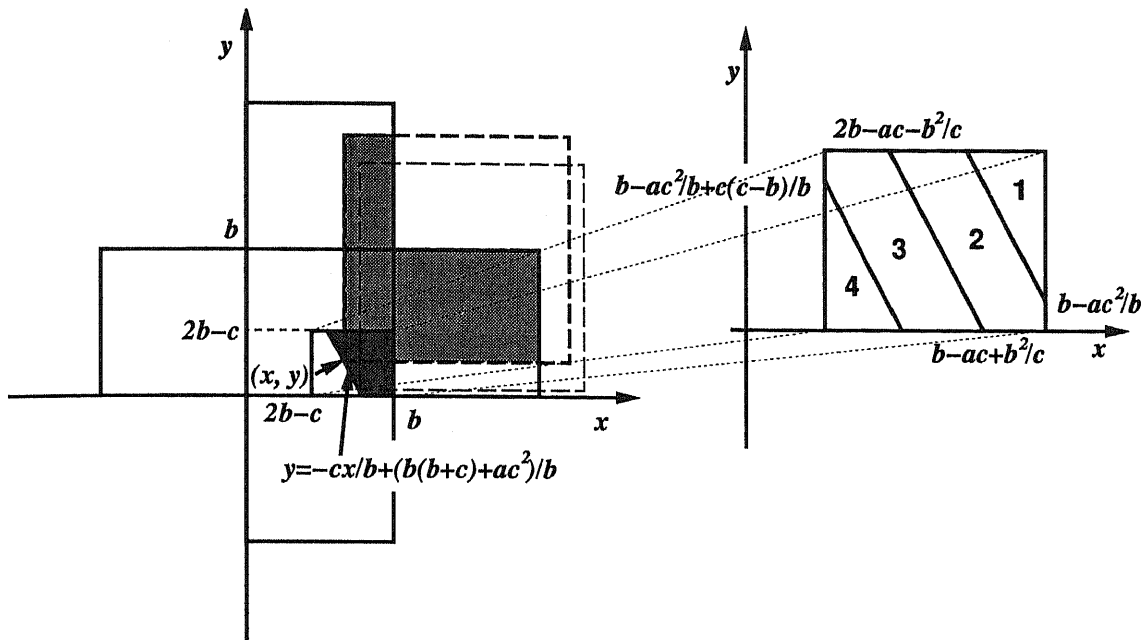


図 3.9: Case 7 の場合の面積占有率密度の計算方法。

$0 < b < \frac{1}{2}c$ かつ $b^2/c^2 < a < 3b^2/c^2$ のとき (x, y) の定義域は $(0 < x, y < b)$ である。

$$F(a) = \frac{1}{b^2} \left[\left(\frac{ac^2 - b^2}{2b} \right) b + \int_{\frac{ac^2 - b^2}{2b}}^b \left\{ \frac{ac^2 + b^2}{x + b} - b \right\} dx \right]$$

$$f_6^4(a) = \frac{c^2}{b^2} \log \left(\frac{4b^2}{ac^2 + b^2} \right) \quad (3.19)$$

以上の計算から、 $0 < b < \frac{1}{2}c$ の場合の面積占有率密度は、 $f_6^3(a)$ と $f_6^4(a)$ とを接続した関数となる。なお a の定義域は $(0 < a < (c^2 - b^2)/c^2)$ である。

なおこのようにして求めた面積占有率密度は、確率密度の積分が 1 になるという条件を満たしている。この性質は、式 (3.16) と式 (3.17) を積分

$$\int_{-\infty}^{\infty} f(a) da = \int_0^{\frac{c^2 - b^2}{c^2}} f(a) da = 1 \quad (3.20)$$

に代入して計算することにより確認できる。

Case 7

この場合は $\frac{1}{2}c < b < c$ のときに生起する。まず座標系を図 3.9 のように設定し、画素の左下隅の座標を (x, y) とする。このとき画素の内部に含まれる十字形モデルの面積は $s = c(b - x) + b(b - y)$ である。 $s = ac^2$ の関係式を代入してこの式を y について解くと、

$$y = -\frac{c}{b}x - \frac{ac^2}{b} + b + c \quad (3.21)$$

という式を得る。 (x, y) の定義域は $(2b - c < x < b)$ および $(0 < y < 2c - b)$ である。ここで Case 7 をさらに 3 個の場合に分類して考察する。

$b(c-b)/c^2 < a < b^2/c^2$ のとき 図 3.9 の三角形領域 1 の面積が $F(a)$ となる。 $x_1 = b - (2b + ac + b^2/c)$ 、 $y_1 = 2b - c - (b - ac^2/b)$ とおくと、

$$F(a) = \frac{1}{(c-b)(2b-c)} \left(\frac{1}{2} x_1 y_1 \right)$$

$$f_7^1(a) = \frac{c}{c-b} \frac{ac^2 - b(c-b)}{b(2b-c)} \quad (3.22)$$

$b^2/c^2 < a < 1 - b^2/c^2$ のとき 図 3.9 の領域 1 と領域 2 を合わせた台形の面積が $F(a)$ となる。上辺の長さを $x_u = b - (2b - ac - b^2/c) = ac + b^2/c - b$ 、下辺の長さを $x_l = b - (b - ac + b^2/c) = ac - b^2/c$ とおくと、

$$F(a) = \frac{1}{(c-b)(2b-c)} \frac{(x_u + x_l)}{2} (2b-c)$$

$$f_7^2(a) = \frac{c}{c-b} \quad (3.23)$$

$1 - b^2/c^2 < a < 1 - b(c-b)/c^2$ のとき 図 3.9 の領域 1 から領域 3 の面積の和が $F(a)$ であり、これは長方形から三角形領域 4 を引いた面積に等しい。 $x_2 = c(1-a) + b^2/c - b$ 、 $y_2 = b - ac^2/b + c(c-b)/b$ とおくと、

$$F(a) = 1 - \frac{1}{(c-b)(2b-c)} \left(\frac{1}{2} x_2 y_2 \right)$$

$$f_7^3(a) = \frac{c}{c-b} \frac{(1-a)c^2 - b(c-b)}{b(2b-c)} \quad (3.24)$$

以上の計算から、全体の面積占有率密度は $f_7^1(a)$ 、 $f_7^2(a)$ 、 $f_7^3(a)$ を接続した関数となる。

Case 8

この場合は $\frac{1}{2}c < b < c$ のときに生起する。密度関数は一様分布である。

$$f_8(a) = \frac{c^2}{(2b-c)(c-b)} \quad (3.25)$$

ただし a の定義域は $((b^2 + c(c-b))/c^2 < a < (2c-b)b/c^2)$ 。

Case 9

この場合は $\frac{1}{3}c < b < c$ のときに生起する。密度関数はデルタ関数である。

$$f_9(a) = \delta \left(a - \frac{(2c-b)b}{c^2} \right) \quad (3.26)$$

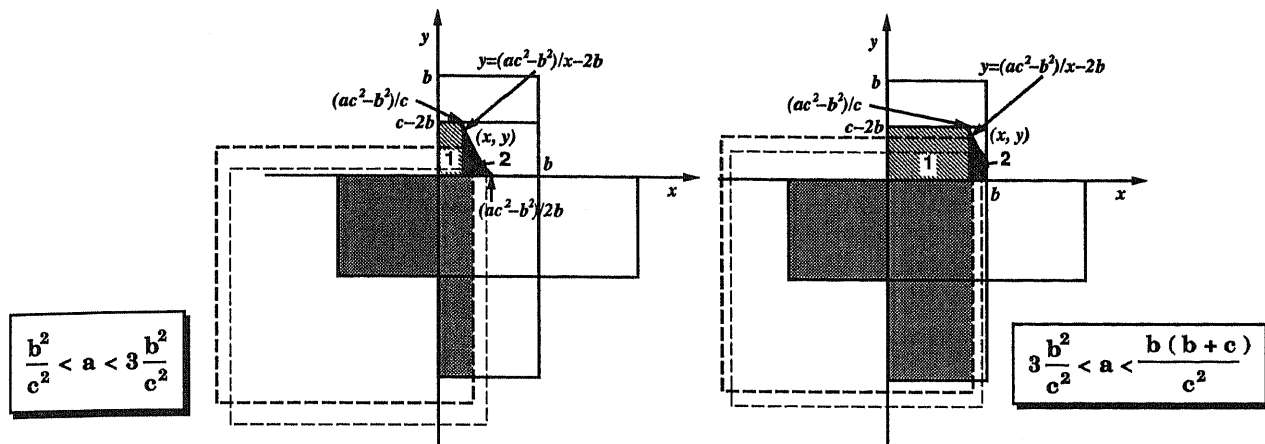


図 3.10: Case 11 の場合の面積占有率密度の計算方法。

Case 10

この場合は $\frac{1}{3}c < b < c$ のときに生起する。密度関数は一様分布である。まず $\frac{1}{2}c < b < cb$ の場合は $(b^2/c^2 < a < (b^2 + c(c - b))/c^2)$ で、

$$f_{10}^1(a) = \frac{c}{c - b} \tag{3.27}$$

次に $\frac{1}{3}c < b < \frac{1}{2}c$ の場合は、 $(b^2/c^2 < a < b(b + c)/c^2)$ で

$$f_{10}^2(a) = \frac{c}{b} \tag{3.28}$$

Case 11

この場合は $b < \frac{1}{2}c$ のときに生起する。まず座標系を図 3.10 のように設定し、画素の右上隅の座標を (x, y) とする。このとき画素の内部に含まれる十字形モデルの面積は $s = x(y + 2b) + b^2$ である。 $s = ac^2$ の関係式を代入してこの式を y について解くと、

$$y = \frac{ac^2 - b^2}{x} - 2b \tag{3.29}$$

という式を得る。このとき (x, y) の定義域は、 $(0 < x < b)$ および $(0 < y < c - 2b)$ である。ここで Case 11 をさらに 2 個の場合に場合に分類して考察する。

$b^2/c^2 < a < 3b^2/c^2$ のとき 図 3.10 の領域 1 と領域 2 の面積の和が $F(a)$ となる。

$$F(a) = \frac{1}{b(c - 2b)} \left[\left(\frac{ac^2 - b^2}{c} \right) (c - 2b) + \int_{\frac{ac^2 - b^2}{2b}}^{\frac{ac^2 - b^2}{c}} \left\{ \frac{ac^2 - b^2}{x} - 2b \right\} dx \right]$$

$$f_{11}^1(a) = \frac{c^2}{b(c - 2b)} \log \left(\frac{c}{2b} \right) \tag{3.30}$$

$3b^2/c^2 < a < b(b + c)/c^2$ のとき 図 3.10 の領域 1 と領域 2 の面積の和が $F(a)$ となる。

$$F(a) = \frac{1}{b(c - 2b)} \left[\left(\frac{ac^2 - b^2}{c} \right) (c - 2b) + \int_{\frac{ac^2 - b^2}{c}}^b \left\{ \frac{ac^2 - b^2}{x} - 2b \right\} dx \right]$$

$$f_{11}^2(a) = \frac{c^2}{b(c - 2b)} \log \left(\frac{bc}{ac^2 - b^2} \right) \tag{3.31}$$

以上の計算から、全体の面積占有率密度は $f_{11}^1(a)$ と $f_{11}^2(a)$ とを接続した関数となる。

表 3.4: 画素と十字形モデルの位置関係による 15 個の場合それぞれの生起確率。表中で面積 S は $S = (5b + c)(b + c)$ である。

Case	$c < b$	$\frac{1}{2}c < b < c$	$\frac{1}{3}c < b < \frac{1}{2}c$	$0 < b < \frac{1}{3}c$
1	$(b - c)(5b - c)/S$	0	0	0
2	$12(b - c)c/S$	0	0	0
3	0	$4b(c - b)/S$	$4b(c - b)/S$	$4b(c - b)/S$
4	0	0	0	$(c - 3b)^2/S$
5	$8c^2/S$	$8b^2/S$	$8b^2/S$	$8b^2/S$
5'	$4c^2/S$	$4(2b - c)^2/S$	0	0
6	0	$4(c - b)^2/S$	$4b^2/S$	$4b^2/S$
7	0	$8(c - b)(2b - c)/S$	0	0
8	0	$4(c - b)(2b - c)/S$	0	0
9	0	$(c - b)^2/S$	$(3b - c)^2/S$	0
10	0	$4(c - b)^2/S$	$4b(3b - c)/S$	0
11	0	0	$8b(c - 2b)/S$	$8b^2/S$
12	0	0	$4(c - 2b)(3b - c)/S$	0
13	0	0	0	$2b(c - 3b)/S$
14	0	0	0	$2b(c - 3b)/S$
15	0	0	$4(c - 2b)^2/S$	$4b^2/S$

Case 12

この場合は $\frac{1}{3}c < b < \frac{1}{2}c$ のときに生起する。密度関数は一様分布であり、 $(b(b + c)/c^2 < a < b(2c - b)/c^2)$ で

$$f_{12}(a) = \frac{c^2}{b(c - 2b)} \quad (3.32)$$

Case 13

この場合は $b < \frac{1}{3}c$ のときに生起する。密度関数は一様分布であり、 $(4b^2/c^2 < a < 5b^2/c^2)$ で

$$f_{13}(a) = \frac{c^2}{b^2} \quad (3.33)$$

Case 14

この場合は $b < \frac{1}{3}c$ のときに生起する。密度関数は一様分布であり、 $(b^2/c^2 < a < 4b^2/c^2)$ で

$$f_{14}(a) = \frac{c^2}{3b^2} \quad (3.34)$$

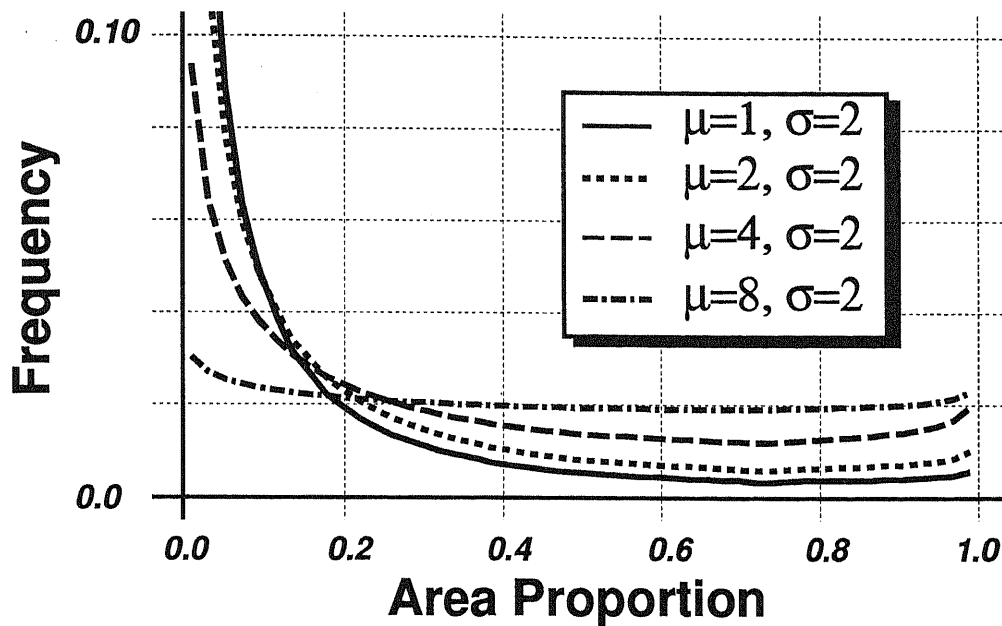


図 3.11: 十字形モデルから導出した J 型面積占有率密度。

Case 15

この場合は $b < \frac{1}{2}b$ のときに生起する。密度関数は一様分布であり、 $(3b^2/c^2 < a < 5b^2/c^2)$ で

$$f_{15}(a) = \frac{c^2}{2b^2} \quad (3.35)$$

3.6.2 それぞれの場合の生起確率

十字形モデルと重なり合う範囲での画素の移動可能範囲の面積は、正方形モデルの場合と同様に考えると $S = (3b + c)^2 - 4b^2 = (5b + c)(b + c)$ となる。また同様に各場合の生起確率を b と c の大小関係により 4 個の場合に分けて表 3.4 に示す。

3.6.3 サイズ密度によるランダム化

十字形モデルの場合も正方形の場合と同様に、式 (3.11) の対数正規密度をサイズ密度として用い、ランダム化した面積占有率密度を求める。その結果を図 3.11 に示す。図 3.6 との最も大きな相違点は、L 型面積占有率密度が a の増加に伴って単調減少していたのに比べ、こちらは途中から増加に転じている点である。この効果は凹の部分の存在によるものであり、その形状からこれを J 型面積占有率密度とよぶ。

3.7 フラクタル合成画像から計算した面積占有率密度

これまでに用いたモデルは、解析的に面積占有率を導くための単純なモデルであった。このモデルから得られた結論を補強するために、現実に近い状況を想定した合成画像を生成するシミュレーションから面積占有率密度を求める。このような現実に近い合成画像として、本研究ではフラクタルを用いて合成画像を生成する。フラクタルによる方法を用いるのは、本研究が主な研究対象とする画像がリモートセンシング画像であり、この種の画像には複雑でフラクタル的な

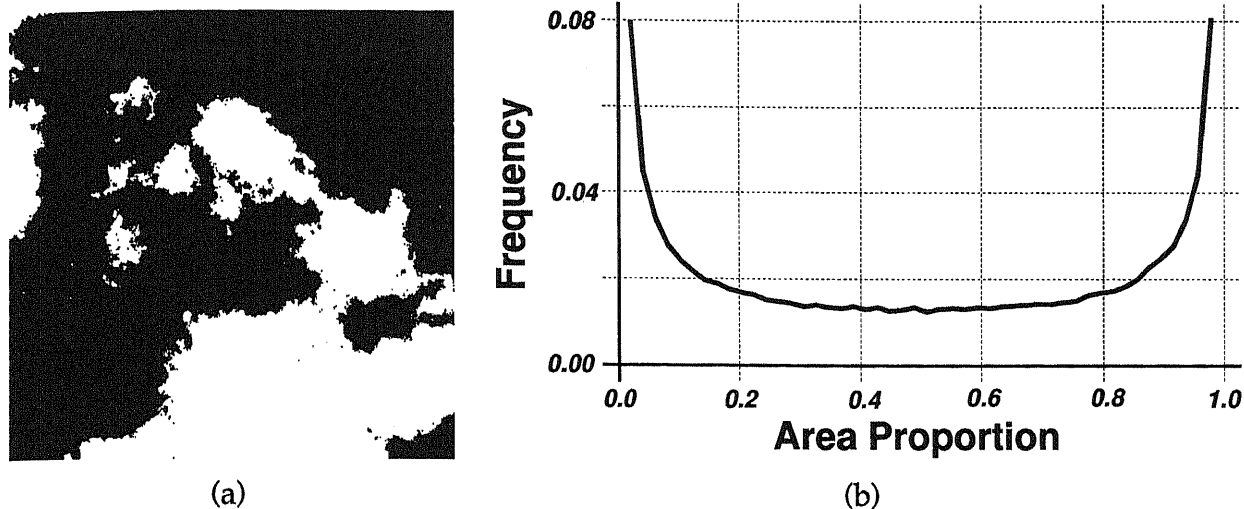


図 3.12: (a) フラクタル合成画像を 2 値化した画像。(b) U 型面積占有率密度。

輪郭形状をもつ領域が多く含まれていることが理由である [46]。フラクタル的な形状を発生させるアルゴリズムとしては、非整数ブラウン運動に基づいた手法を用いる。特に本研究で用いたアルゴリズムは逐次ランダム加算法を 2 次元の画像平面に拡張した方法である [47]。

このシミュレーションの方法を以下にまとめる。

1. 逐次ランダム加算法により、 $N \times N$ のフラクタル表面を表す濃淡画像を生成する。
2. この合成画像を適当なしきい値で 2 値化し、2 つのクラスをもつ画像を生成する。
3. 2 値化画像の $c \times c$ の画素ブロックを 1 画素として観測し、各画素ブロック中でのクラス 1 の面積占有率を計算する。
4. 1 ~ 3 の過程を M 回繰り返して得られたクラス 1 の面積占有率のヒストグラムを、フラクタル合成画像から得られた面積占有率密度とする。

つまり、実世界に対応する高解像度の合成画像を、粗い解像度の仮想的なセンサによってブロック単位に観測することを模擬する。当然一部の画素はミクセルとなるので、その面積占有率のヒストグラムを計算すれば面積占有率密度が得られる。

アルゴリズムとしては、縮小率が $1/\sqrt{2}$ のアルゴリズムを用いる。画像の大きさは $N = 1025$ 。また $c = 7$ としたので面積占有率は $a = i/49$, ($i = 1, \dots, 48$) の離散的な値をとる。2 値化のしきい値には濃淡画像の最大値と最小値の中点を選んだ。繰り返し回数 $M = 100$ のとき、平均してミクセルは全体の画素の 10.8% であり、また合成画像のフラクタル次元は $D = 1.33$ (標準偏差 $\sigma = 0.032$) であった。合成画像の 2 値化画像を図 3.12(a) に、また得られた面積占有率密度を図 3.12(b) に示す。この場合の面積占有率密度は、面積占有率 0.5 を軸としてほぼ対称な形であり、その形状から U 型面積占有率密度とよぶ。

3.8 4種の面積占有率密度の比較

本研究では、正方形モデルからは L 型、十字形モデルからは J 型、さらにフラクタル合成画像からは U 型の面積占有率密度を導いた。図 3.13 にまとめられたこの 3 種類の密度関数に加え、一様分布までを含めてこれらの関数の関係を分析する。

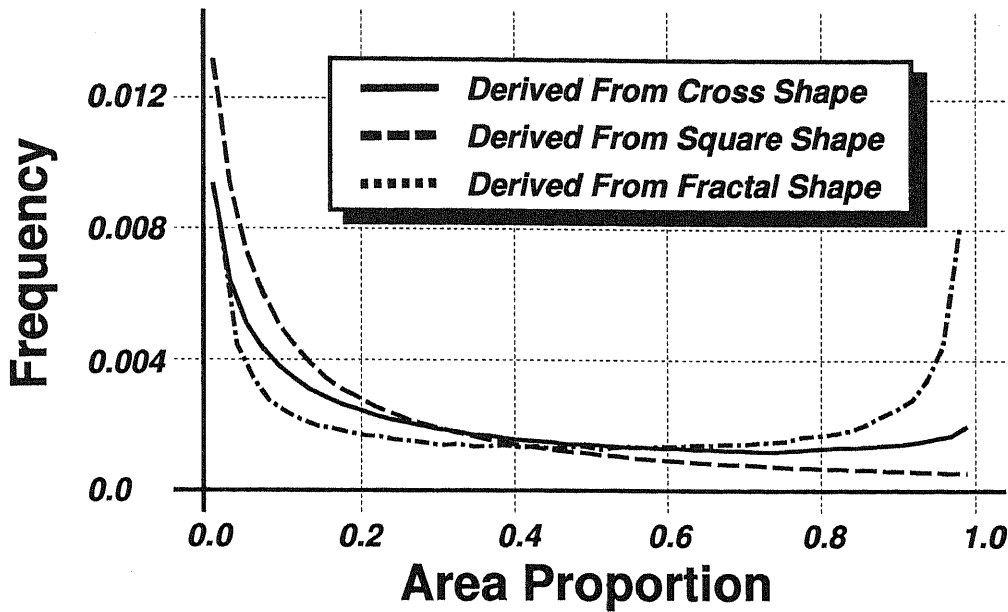


図3.13: 3種の面積占有率密度の比較。L型とJ型では対数正規密度のパラメータを $\mu = 4, \sigma = 2$ とした。

まず面積占有率密度はL型になる傾向がある。それは画素サイズより小さな領域を含む画素の面積占有率が最大でも b^2/c^2 でしかないことが大きな理由である。また正方形モデル Case 5 のように凸の部分もL型となる要因の一つである。しかし十字形モデルのような凹部は凸部と逆の効果をもたらすため、 a の増加にしたがって単調減少する成分に対抗して単調増加する成分も増えてくる。この傾向をさらにすすめるとU型面積占有率密度が得られる。すなわちフラクタル合成画像の場合は、クラス1とクラス2の領域境界の凹凸の程度が全く同等であるため、密度関数は面積占有率0.5を軸として対称な形となる。つまりL型とU型は2クラスの凹凸の程度の極限に位置し、J型はその中間に位置するという関係になる。

一方の一様分布との関係は、図3.6や図3.11においてサイズ密度の中心が画素の大きさに比べて大きくなると面積占有率密度が平坦になることから、画像中に存在する領域境界の空間的変動のスケールが画素の大きさに比べて著しく大きい場合に、面積占有率密度は一様分布に近付くと考えられる。しかしこの仮定とは、まさにミクセルの割合が「少ない」画像に対して成立する仮定である。ゆえにミクセルの面積占有率密度として一様分布は適当なモデルとはいえない。

以上の考察から、実際の画像の面積占有率密度はJ型になっていると考えられる。そして、両方のクラスの凹凸の程度の比率からJ型の左右のピークの高さの比を決定すべきであるが、この高さの比を設定できるような簡単な関係式を著者らはまだ発見できておらず、現在はシミュレーションによって設定している。

なお面積占有率密度の実際の画像分類への応用については第5章で述べる。そこでここでは簡単に面積占有率密度の応用について述べる。画像中でクラス1の領域が占める面積を推定する問題を考える。いま画像中に M 画素のミクセルが存在することがわかれば、このミクセル集合でクラス1が占める面積は面積占有率密度の期待値 $E\{a\} = \int_0^1 af(a)da$ を用いて、 $M \cdot E\{a\}$ によって計算できる。ここで面積占有率密度が対称な密度関数であれば $E\{a\} = 0.5$ となるが、J型のように非対称であれば $E\{a\} \neq 0.5$ となる。つまり画像中で各クラスが占める面積を推定す

る問題では、面積占有率密度の仮定によって誤差が生じる。またミクセルを含む混合密度推定の問題においても、面積占有率密度の仮定によって結果は大きく変わることになるため、実画像の面積占有率密度の推定が重要な問題となるのである。

3.9 まとめ

本研究は、面積占有率密度が一般に思われているように一様分布ではなく、L型やJ型・U型の面積占有率密度になることを示した。まず簡単な図形をモデルに用いた解析的な計算により面積占有率密度がL型またはJ型となることを示し、次にフラクタル合成画像を用いたシミュレーションにより面積占有率密度がU型となることを示した最後にこの3種の面積占有率密度および一様分布との関係を論じた。画像中で各クラスごとの面積を求める場合には、この面積占有率密度の推定が重要な問題となることを述べた。

今後は本研究で用いたモデルをさらに現実的なモデルとすることが課題である。中でも一画素に複数のモデルが含まれる場合の解析が大きな課題である。次に3クラスミクセル以上の面積占有率密度を求めることも課題である。特に2個以上の面積占有率が相関を持つ場合の定式化が重要となる。

第4章

期待値に基づいたミクセルの面積占有率の推定

4.1 概要

ミクセルの観測値からそれを構成する各分類クラスが占める面積の割合(面積占有率)を推定する問題は困難な逆問題である。しかし個々のミクセルに対して面積占有率という情報が得られれば、処理の単位をピクセル単位からサブピクセル単位へと細かくして多くの画像情報を抽出でき、また高次の認識アルゴリズムには個々のミクセルの面積占有率という有用な情報を明示的なしきい値処理なしに渡すことができる。このように有用な情報であるミクセルの面積占有率を推定する問題についてはすでに多くの研究がなされているが、これらの研究では観測値の変動は共通の雑音成分に起因するという仮定が置かれることが多く、各クラスのクラス内分散の存在についてはあまり注意が払われていない。そこで本研究ではこのようなクラス内分散の影響を考慮するため、ミクセルを確率的なモデルとして捉え、ミクセルの面積占有率を期待値に基づいて推定する手法を提案する。次に各クラスの密度関数を正規密度とした場合の期待面積占有率の計算を具体的におこない、クラス内分散がクラスごとに大きく異なる場合には、クラス内分散を考慮しない方法では大きな誤差が生じることを指摘する。最後にシミュレーションによる実験をおこない、本手法が期待値の意味で面積占有率を適切に推定することを確認する。なお本手法は任意の p バンド k クラスミクセルに対して計算可能な拡張性の高い方法である。

4.2 従来研究の問題点

個々のミクセルに対して各クラスの面積占有率という情報を得ることができれば、1画素の観測値からも多くの情報を抽出することが可能となる。このことによる利点には二つの側面がある。まず一つの側面は、ピクセル単位からサブピクセル単位の処理へと移行することで、例えばリモートセンシング画像のように解像度の粗い画像からも高解像度の情報抽出が可能となり、したがって画像から抽出できる情報量が増加するという利点である。もう一つの側面は、しきい値処理などによる画素情報の喪失を防ぐという側面である。例えばしきい値処理による画像の領域分割などの手法 [48] では、基本的にすべての画素をピュアピクセルとみなしている。例えば観測値がしきい値付近の画素はミクセルである可能性があるが、しきい値による分割を通してこのような情報は失われ、ピュアピクセルとミクセルかもしれない画素はそれ以後の処理では全く同等に扱われる。しかし個々のミクセルの面積占有率を求めることができれば明示的なしきい値処理は不要なものとなり、高次の認識アルゴリズムには豊富で有用な情報を渡すことが可能となることがもう一つの利点である。したがって個々のミクセルの面積占有率を求める問題は、画像処理

の高度化にとって重要な課題である。

個々のミクセルの面積占有率を求める問題は、ミクセルに関する研究の中心課題であり、過去にも多くの解法が提案されてきた。本研究ではそれらの方法を以下の5種類に大別する。

1. 各クラスの分光行列の逆行列計算により求める方法 (Unmixing / カテゴリ分解など) [49, 50, 51, 52]
2. 線形回帰モデルやニューラルネットワークなどの非線形回帰モデルにより求める方法 [53]
3. ファジーメンバーシップ関数を面積占有率とみなす方法 [54, 55, 56, 57, 58]
4. 領域境界をパラメトリックな曲線で近似する方法 [59, 60]
5. 確率モデルに基づく方法 [61, 62, 63, 64, 65]

まず最初の方法は逆行列計算によって各クラスの占有率を求める方法である。 k 個のクラスからなるミクセルを考え、その面積占有率ベクトルを \mathbf{a} 、また各クラスの分光行列を M とする。このとき画素の観測ベクトル \mathbf{r} は、以下の線形結合

$$\mathbf{r} = M\mathbf{a} \quad (4.1)$$

で観測されると仮定する。ここで既知の \mathbf{r} と M から未知の \mathbf{a} を求める。 M が正方行列でない場合は通常の意味での逆行列は存在せず、 \mathbf{a} の推定値としては最小二乗の意味で最適な推定値

$$\mathbf{a} = (M^t M)^{-1} M^t \mathbf{r} = M^+ \mathbf{r} \quad (4.2)$$

を用いる。ここで M^+ はムーア・ペンローズ型一般化逆行列とよばれる。ただし式(4.2)を実際の画像に適用すると種々の雑音の影響によって、得られる面積占有率は面積占有率ベクトルのパラメータ空間からはみだすことが多い。そこでこのような制約を考慮した制約付きの最適化問題として式(4.1)を解く解法も提案されている[49, 51]。しかしこの方法は各クラスのクラス内分散を考慮していないため、クラス内分散が大きい画像では適切な面積占有率密度が得られない[66]。次に(2)の方法では、面積占有率が既知のミクセルを教師データとし、その観測ベクトルから面積占有率ベクトルを推定する回帰式を得ることによって面積占有率を推定する方法である。線形回帰モデルだけでなく、ニューラルネットワークなどの非線形なモデルを用いる方法も提案され、こちらの方が線形モデルよりは精度が向上するという報告がある。しかしこの方法で高い精度の回帰式を得るには、面積占有率が既知の画素という大量の教師データを必要とする点が問題である。(3)ではミクセルの概念とファジーの概念とのアナロジーに基づき、何らかの方法によって定めた各クラスのファジーメンバーシップ関数を、面積占有率とみなすことによって面積占有率を得る方法である。しかしファジーメンバーシップ関数の意味を巡っては常に議論があり[67, 68]、これらの方法の定式化が適切であるかどうかには議論の余地がある。(4)は画像から検出されたエッジに、大局的な領域境界を表すパラメトリックな曲線を当てはめ、この曲線から各画素の面積占有率を推定する手法である。この方法は、長方形の田畑が並ぶ地域など、領域境界の形状が前もって知られているごく限られた地域にしか適用できない。

これらの研究に対して、本研究に最も近い研究が(5)の方法である。これらの方法は、確率モデルを用いてミクセルを分析することで、面積占有率を統計的に取り扱うアプローチである。しかし後述するように、基本的な仮定において本研究はこれらの研究とは異なっている。

以下に本章の構成を述べる。まず4.3章において本研究が扱う問題を定式化する。次に4.4章は本研究の主題であるミクセルの面積占有率の推定を扱い、期待値に基づいた面積占有率の計算方法を提案する。4.5章では、提案した手法を実際に応用するための基本として、クラスの確率密度に正規密度を用いた場合の期待面積占有率を具体的に計算する。4.6章ではシミュレーションによって本研究の手法が期待値の意味で面積占有率を正しく推定することを示し、最後に4.7章で本研究をまとめる。

4.3 問題の定式化

4.3.1 ミクセルの線形モデル

まず本研究で用いるミクセルのモデルについて述べる。まず画素の観測値は、 p 個のバンドによって観測された p 次元観測ベクトル $\mathbf{r} = (r_1, \dots, r_p)^t$ であるとする。またミクセルが k 個のクラスから構成されていれば、それを k クラスミクセルとよぶ。例えば図3.2は2クラスミクセルの例を示している。また k クラスミクセルを構成する各クラスの(画素内)面積占有率を a_i 、各クラスからの観測値を \mathbf{x}_i とする。ここで面積占有率は $0 < a_i < 1$ および $\sum_{i=1}^k a_i = 1$ が成立するものとする。

さてミクセルの観測値とは、各クラスの観測値 $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})^t$ が面積占有率の関数で重みづけされて混合したものと考えることができる。そこで本研究では、最も単純なモデルとして以下のような線形モデルを仮定する。

$$\mathbf{r} = \sum_{i=1}^k a_i \mathbf{x}_i \quad (4.3)$$

すなわち、ミクセルの観測値は各クラスからの観測値が面積占有率で重みづけされて混合したものとする。しかし実際の観測では種々の雑音の混入があるため、式(4.3)の代わりに以下のモデルが用いられる。

$$\mathbf{r} = \sum_{i=1}^k a_i \mathbf{x}_i + \boldsymbol{\epsilon} \quad (4.4)$$

ここで $\boldsymbol{\epsilon}$ は雑音成分を表す項である。さて本研究の目的は、式(4.4)から面積占有率を求めるといふ、以下の問題を解くことである。

問題1 ある一つのミクセルに注目する。式(4.3)が成立し、かつ各クラスの確率密度 $p_i(\mathbf{x}_i)$ が既知であるときに、ミクセルの観測ベクトル \mathbf{r} からミクセルの面積占有率 a_i を推定する逆問題を解く。

さてこの問題の定式化と従来の研究の仮定との本質的な相違点を、式(4.3)と式(4.4)を用いて説明する。

従来の研究 \mathbf{x}_i は値が「既知」の分光ベクトル、 $\boldsymbol{\epsilon}$ は正規密度に従う雑音成分と仮定する。

本研究 \mathbf{x}_i は値が「未知」の確率変数と仮定する。ただしクラス確率密度 $p_i(\mathbf{x}_i)$ は既知とする。

つまり x_i を「既知」とするか「未知」とするかという点が最も大きな相違点である。 x_i を既知とすることは、言い換えればあるクラスからの観測値はすべて同一であると仮定することに等しい。つまり、クラス内分散を「無視」するか「考慮」するかという点が相違点と言い換えてもよい。従来の方法では観測値のばらつきの影響を、式(4.4)にあるような、すべてのクラスに共通の雑音成分 ε に押し込めていた。しかし実際の画像では個々のクラスによって異なるクラス内分散が存在し、その変動を共通の雑音成分に押し込めるの方針には無理があると著者らは考える。そこで本研究では、クラスの確率密度を考えることで、クラス内分散を考慮した面積占有率の推定手法を提案する。

4.4 ミクセルの面積占有率の推定

4.4.1 面積占有率を固定した場合の同時密度関数

まず最も単純な問題である、1バンドで観測した2クラスミクセルの問題について考察する。この場合の式(4.4)は

$$r = ax_1 + (1-a)x_2 + \varepsilon \quad (4.5)$$

に等しい。ここで x_i はクラス i の確率密度 $p_i(x_i)$ にしたがう確率変数 X_i の実現値であり、また a はクラス1の面積占有率 ($0 < a < 1$) である。また ε は確率密度 $p_e(x) = N(0, \sigma_e^2)$ にしたがう雑音成分である。ここで $N(\cdot, \cdot)$ は正規密度を表す。

さて x_i が確率変数¹であるというのが本研究の基本的な仮定である。ゆえに式(4.5)では、既知の変数が観測値 r のみであるのに対しそれ以外の4変数はすべて未知であるため、式(4.5)を直接解いて a を求めることはできない。そこで、 x_i がしたがう確率密度 $p_i(x_i)$ を用いて、まず a と r の同時密度関数 $p(a, r)$ を求める。仮に雑音成分を無視すれば、同時密度関数は以下の計算で求めることができる [43]。

$$p(r, a) = \int_{-\infty}^{\infty} q_1(r-t)q_2(t)dt = q_1(r) * q_2(r) \quad (4.6)$$

ここで $q_1(r) = p_1(r/a)$ および $q_2(r) = p_2(r/(1-a))$ である。また $*$ は畳み込み演算を表す。次に雑音成分までを含める場合は、

$$p(r, a) = q_1(r) * q_2(r) * p_e(r) \quad (4.7)$$

と式(4.6)に雑音成分を表す確率密度 $p_e(x)$ を新たに畳み込むことによって同時密度関数を求めることができる。

同様の方法によって p バンド k クラスミクセルの場合の同時密度関数を求めることができる。各 k クラスの面積占有率を表す面積占有率ベクトルを $\mathbf{a} = (a_1, \dots, a_k)^t$ と定めると、観測値 $\mathbf{r} = (r_1, \dots, r_p)^t$ との同時密度関数は

$$p(\mathbf{r}, \mathbf{a}) = q_1(\mathbf{r}) * \dots * q_k(\mathbf{r}) * p_e(\mathbf{r}) \quad (4.8)$$

と計算できる。ただし $q_i(\mathbf{r}) = p_i(\mathbf{r}/a_i)$ であり、また $p_e(\mathbf{r})$ は p 次元の正規密度である。このように k 個の関数および雑音成分の畳み込みによって、任意の p バンド k クラスミクセルに対する同時密度関数を得ることができる。

¹ x_i は確率変数 X_i の実現値であるが、以下では x_i と X_i とを区別せずに用いる。

4.4.2 同時密度関数の計算方法

以下では式(4.6)を実際に計算するための2つの方法について述べる。

特性関数を用いる方法

特性関数

定義1 X は確率密度 f をもつ確率変数とする。 X の特性関数 φ は実数値 ζ に対して以下のように定義される。

$$\varphi(\zeta) = E(e^{i\zeta x}) = \int_{-\infty}^{\infty} e^{i\zeta x} f(x) dx \quad (4.9)$$

もし φ が確率変数 X の特性関数であれば、 $aX + b$ は以下の特性関数をもつ。

$$E(e^{i\zeta(aX+b)}) = e^{ib\zeta} \varphi(a\zeta). \quad (4.10)$$

さらに、二つの確率変数 X_1 、 X_2 を考え、それぞれが確率密度 f_1 、 f_2 と特性関数 φ_1 、 φ_2 をもつとする。もし X_1 と X_2 が独立ならば、指数の乗法性から

$$E(e^{i\zeta(X_1+X_2)}) = E(e^{i\zeta X_1}) E(e^{i\zeta X_2}). \quad (4.11)$$

となる。言い換えれば、二つの独立な確率変数の和 $X_1 + X_2$ はそれぞれの特性関数の積 $\varphi_1\varphi_2$ に対応する。式(4.11)の結果を拡張すると、確率変数の和 $X = \sum X_i$ は以下のように表される特性関数 $\varphi(\zeta)$ をもつ。

$$\varphi(\zeta) = \prod \varphi_i(\zeta), \quad (4.12)$$

ここで $\varphi_i(\zeta)$ はそれぞれの確率変数 X_i の特性関数である。このように特性関数の概念によって確率変数の和の確率密度を計算することが容易となる。

同時密度関数の計算 同時密度関数の基本的な計算方法である式(4.6)と本質的には同等であるがより簡便な方法として、付録第4.4.2節で述べる特性関数を用いる方法がある。各クラスの密度関数 $p_i(\mathbf{r})$ の特性関数が既知であれば、特性関数を用いて簡単に同時密度関数が求められる。

例えば1バンド2クラスミクセルを例とし、各クラスの密度関数 $p_i(\mathbf{r})$ が正規密度 $N(\mu_i, \sigma_i^2)$ で表される場合を考える。正規密度の特性関数は

$$\varphi_i(\zeta) = e^{i\mu_i\zeta - \frac{\sigma_i^2}{2}\zeta^2} \quad (4.13)$$

である。すると付録第4.4.2節の関係から、雑音成分を無視した場合の $p(\mathbf{r}, \mathbf{a})$ の密度関数の特性関数は、

$$\varphi_m(\zeta) = e^{i(a\mu_1 + (1-a)\mu_2)\zeta - \frac{a^2\sigma_1^2 + (1-a)^2\sigma_2^2}{2}\zeta^2} \quad (4.14)$$

となる。これは $p(\mathbf{r}, \mathbf{a})$ が正規密度 $N(a_1\mu_1 + a_2\mu_2, a^2\sigma_1^2 + (1-a)^2\sigma_2^2)$ に等しいことを示している。さらに特性関数を同様に計算することによって、1バンド k クラスミクセルの同時密度関数 $p(\mathbf{r}, \mathbf{a})$ も、

$$p(\mathbf{r}, \mathbf{a}) = N\left(\sum_{i=1}^k a_i\mu_i, \sum_{i=1}^k a_i^2\sigma_i^2\right) \quad (4.15)$$

とやはり正規密度になる。また雑音成分の密度関数は正規密度と仮定されることから、雑音成分を含めた場合の式 (4.15) は

$$p(r, \mathbf{a}) = N \left(\sum_{i=1}^k a_i \mu_i, \sum_{i=1}^k a_i^2 \sigma_i^2 + \sigma_e^2 \right) \quad (4.16)$$

と、分散を雑音成分の分だけ大きくする修正で済む。

このようにクラスの確率密度関数と同時密度関数が同じ形の密度関数となるのは、正規密度が安定分布の一つであるためである。このような安定分布の性質はミクセルのモデルとして好都合である。この安定分布に関する説明を以下に付け加える。

安定分布

定義 2 X, X_1, X_2, \dots は共通の分布 R を持つ互いに独立な確率変数とし、 $S_n = X_1 + \dots + X_n$ とする。このとき分布 R が各 n に対して

$$S_n \stackrel{d}{=} c_n X + \gamma_n \quad (4.17)$$

となるような定数 $c_n > 0$ 、 γ_n が存在していて R が原点に集中していないとき安定であるという。ただし $U \stackrel{d}{=} aV + b$ は U 、 V の分布が位置パラメータだけしか異なっていないことを意味する [43]。

つまり、もし k クラスミクセルに関係するすべてのクラスの確率密度 $p_i(\mathbf{r})$ ($1 \leq i \leq k$) が同じ種類の安定分布であれば、その和の分布であるミクセルの確率密度 $p_m(\mathbf{r})$ は、もとの安定分布と比べてパラメータの値しか異ならないことになる。

通常は確率変数の和の確率密度を初等関数によって表現することはできない。しかし各クラスの確率密度が同じ種類の安定分布ならば、ミクセルの確率密度とも同じ種類の安定分布となる。また、安定分布は積分形やベキ級数展開によって表現され、初等関数によって表現できる場合はほんの少数しか知られていないが、その数少ない例が正規密度やコーシー密度である。よってこれらのモデルは分布の安定性という意味の利点があり、特にミクセルを解析するためのモデルとしては扱いやすい性質をもっている。

高速フーリエ変換を用いる方法

特性関数が簡単な形では表せない一般の確率密度に対しては、式 (4.6) の畳み込みをフーリエ変換を経由して計算する。特に各クラスの確率密度 $p_i(\mathbf{r})$ を離散的に表せれば、式 (4.6) は高速フーリエ変換 FFT (Fast Fourier Transform) を経由して高速に計算することができる [69]。ただし畳み込みに用いられる関数は $p_i(\mathbf{r})$ ではなく $q_i(\mathbf{r}) = p_i(\mathbf{r}/a_i)$ である。もし $p_i(\mathbf{r})$ がパラメトリックな確率密度であればこの変換は容易であるが、そうでない場合には新たなサンプリング間隔 $1/a_i$ に合うように $p_i(\mathbf{r})$ を適切な方法で補間する必要がある。

4.4.3 面積占有率密度

さて多数のミクセルの集合を対象に考えると、面積占有率ベクトル \mathbf{a} はパラメータ空間内でさまざまな値をとる、言い換えれば確率変数ベクトル \mathbf{A} の実現値であると考えの方が適切である。

そこで面積占有率ベクトルとして $k-1$ 次元の面積占有率ベクトル $\mathbf{a}' = (a_1, \dots, a_{k-1})^t$ を考え、このベクトルのパラメータ空間内での密度関数 $f(\mathbf{a}')$ を考えればよいが、この密度関数の導出などの詳しい記述については第3章で述べているので参照していただきたい。このように面積占有率ベクトルを、確率密度 $f(\mathbf{a}')$ にしたがう確率変数ベクトルであるとみなすことにより、面積占有率ベクトルが \mathbf{a}' 、観測ベクトルが \mathbf{r} となる同時確率密度 (ランダム化された同時確率密度) は以下のように求められる。

$$g(\mathbf{r}, \mathbf{a}') = p(\mathbf{r}, \mathbf{a}')f(\mathbf{a}') \quad (4.18)$$

これは式 (4.8) の同時確率密度と面積占有率密度との積である。

4.4.4 ミクセルの期待面積占有率の計算

さて、同時密度関数 $g(\mathbf{r}, \mathbf{a}')$ の情報から面積占有率ベクトル \mathbf{a}' を推定する方法を検討する。その一つの方法は、 $g(\mathbf{r}, \mathbf{a}')$ を最大にする面積占有率ベクトル \mathbf{a}'_e 、すなわち

$$\mathbf{a}'_e = \operatorname{argmax}_{\mathbf{a}' \in \Psi} g(\mathbf{r}, \mathbf{a}') \quad (4.19)$$

を面積占有率の推定値とする方法である。この方法は最尤推定の意味で最適な推定値であることから、これを最尤面積占有率とよぶ。例えば文献 [70] でも最尤面積占有率が提案されている²。しかしこの手法には二つの欠点がある。まず $g(\mathbf{r}, \mathbf{a}')$ の大域的な最大値は必ずしもパラメータ空間の内側にはこないという点である。 \mathbf{r} の変化にしたがって大域的な最大値はパラメータ空間の外側に出てしまい、このような場所で \mathbf{r} の変化に対して \mathbf{a}' が不連続的に変化する。その他にも、バンド数 p がクラス数 k に対して少ない条件不足の時には、ある超平面上の点がすべて最大値となり解が一意に定まらないという欠点もある。

そこで本研究では、 \mathbf{a}' の期待値を面積占有率ベクトルの推定値とする方法を提案する。すなわち面積占有率を以下の式で推定する。

$$\mathbf{a}'_e = E\{\mathbf{a}'\} = \iint_{\Psi} \mathbf{a}' g(\mathbf{a}'|\mathbf{r}) d\mathbf{a}' \quad (4.20)$$

ただし $g(\mathbf{a}'|\mathbf{r})$ は観測ベクトル \mathbf{r} のもとでの条件つき確率密度であり、以下の式で計算できる。

$$g(\mathbf{a}'|\mathbf{r}) = \frac{g(\mathbf{r}, \mathbf{a}')}{g(\mathbf{r})} = p(\mathbf{a}'|\mathbf{r})f(\mathbf{a}') \quad (4.21)$$

$$g(\mathbf{r}) = \iint_{\Psi} g(\mathbf{r}, \mathbf{a}') d\mathbf{a}' \quad (4.22)$$

この推定値 $E\{\mathbf{a}'\}$ を期待面積占有率とよぶ。この方法では \mathbf{r} の変化に対して \mathbf{a}' は滑らかに変化し、しかもいかなるバンド数とクラス数の組み合わせにおいても面積占有率を一意に推定できるという利点がある。

4.4.5 デルタ関数の場合の期待面積占有率

期待面積占有率を具体的に計算する方法について、まずは最も単純な例を取り上げて説明する。再び1バンド2クラスミクセルの場合を考える。問題を単純にするための仮定として、各ク

²面積占有率密度については考察されていない。

ラスからの観測値が全く同一でクラス内分散が0、雑音成分も0の場合を考える。このとき各クラスの確率密度 $p_i(r)$ はデルタ関数で表すことができる。

$$p_i(r) = \delta(r - \mu_i) \quad (4.23)$$

ここで各クラスからの観測値を μ_i としている。すると同時確率密度 $p(r, a)$ は

$$p(r, a) = \delta(r - \{a\mu_1 + (1-a)\mu_2\}) \quad (4.24)$$

として求められる。ここで $\mu_1 < \mu_2$ とし、また面積占有率密度が一様分布 $f(a) = 1$ であると仮定する。この場合は $p(a|r) = p(r, a)$ であるので、式(4.20)により期待面積占有率 $E\{a\}$ は

$$E\{a\} = \int_0^1 ap(r, a) da = \frac{\mu_2 - r}{\mu_2 - \mu_1} \quad (4.25)$$

と求められる。ただし $\mu_1 \leq r \leq \mu_2$ である。

式(4.25)は、観測値 r に対して期待面積占有率が線形に変化する関係を示しており、これは従来の研究でも提案されてきた面積占有率の推定式と一致する。このことは、従来の研究が各クラスの密度関数をデルタ関数と暗黙のうちに仮定し、クラス内分散を0とみなしていたことを意味している。

4.5 正規密度の場合の期待面積占有率

現実の画像では各クラスからの観測値にはクラス内分散に起因するばらつきがあるため、クラスの密度関数としてはデルタ関数は現実的ではなく、より実用的な密度関数として本章では正規密度を用いる。ここで正規密度を用いる利点は以下の点にまとめられる。

- 種々の条件のランダムな変化によりひき起こされる分布は正規分布になるという物理的背景がある。
- 本研究が主な対象とするリモートセンシングのみならず、幅広い分野で最も広く用いられている確率モデルである。
- 分布の安定性などの有用な数学的性質を備えているため解析が比較的容易である。

さてクラス i の確率密度 $p_i(\mathbf{r})$ が多次元正規密度 $N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ であるとする。ここで多次元正規密度は以下のように表される。

$$p_i(\mathbf{r}) = \frac{e^{\{-\frac{1}{2}(\mathbf{r}-\boldsymbol{\mu}_i)^t \boldsymbol{\Sigma}_i^{-1} (\mathbf{r}-\boldsymbol{\mu}_i)\}}}{(2\pi)^{p/2} |\boldsymbol{\Sigma}_i|^{1/2}} \quad (4.26)$$

ただし $\boldsymbol{\mu}_i$ は平均ベクトル、 $\boldsymbol{\Sigma}_i$ は共分散行列である。この確率密度をモデルとしたとき、バンド数 p やクラス数 k を変えた場合に期待面積占有率が示す性質を以下で具体的に計算する。なお本章では議論の繁雑さを防ぐため、雑音成分 ϵ は無視して議論する。ここで得られた結果に雑音成分を含める場合には、式(4.16)による修正を加えればよい。

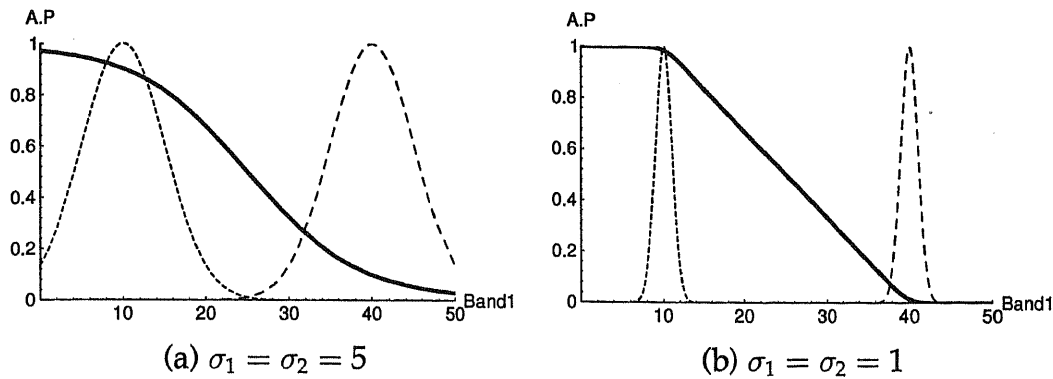


図 4.1: 1 バンド 2 クラス ミクセルの期待面積占有率。

4.5.1 1 バンド 2 クラス ミクセルの期待面積占有率

まずは最も簡単な場合の 1 バンド 2 クラス ミクセルについて計算する。各クラスの確率密度 $p_i(r)$ は、正規密度 $N_1(\mu_i, \sigma_i^2)$

$$p_i(r) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left\{-\frac{(r - \mu_i)^2}{2\sigma_i^2}\right\} \quad (4.27)$$

で表される。式 (4.15) より同時密度関数は $p(r, a) = N(a\mu_1 + (1-a)\mu_2, \{a\sigma_1\}^2 + \{(1-a)\sigma_2\}^2)$ となるので、条件つき密度 $p(a|r)$ として以下の式を得る。

$$p(a|r) = \frac{C}{\sqrt{2\pi}\sigma^2} \exp\left\{-\frac{\left(a - \frac{\mu_2 - r}{\mu_2 - \mu_1}\right)^2}{2\sigma^2/(\mu_2 - \mu_1)^2}\right\} \quad (4.28)$$

ただし $C = \int_0^1 p(a|r) da$ は規格化定数である。

式 (4.20) に式 (4.28) を代入し、さらに面積占有率密度を一様分布と仮定した場合の期待面積占有率の計算結果を図 4.1 に示す。正規密度の分散を小さくするにつれて、期待面積占有率のグラフはデルタ関数の場合へ近づく。正規密度で分散を 0 とした極限がデルタ関数であるから、この結果はデルタ関数の場合の期待面積占有率の結果の自然な拡張となっている。さらに観測値 r の範囲で場合わけしながら、式 (4.28) の性質を定性的に分析する。

$\mu_1 < r < \mu_2$ の場合 式 (4.28) は、 $a = \frac{\mu_2 - r}{\mu_2 - \mu_1}$ のピークを中心にして対称な、正規密度に近い形の関数である。ゆえにもし r が μ_1 または μ_2 付近でなければ a の期待値はピークの値 $a = \frac{\mu_2 - r}{\mu_2 - \mu_1}$ 、すなわちデルタ関数の場合の式 (4.25) とほぼ一致する。しかし r が μ_1 または μ_2 に近づくにつれて関数のピークは a の定義域 $[0, 1]$ の端に近づくため、定義域の内部では関数が非対称な形となり期待面積占有率は式 (4.25) の解から外れてくる。

$r < \mu_1$ の場合 関数のピークは a の定義域 $[0, 1]$ の外側に出る。したがって定義域の内部での最大値は $a = 1$ となるが、定義域の内部に伸びる関数の裾の影響によって、期待値としては 1 よりわずかに小さくなる。この効果によって、 a が 1 付近でも期待面積占有率は r の変化に対して滑らかに変化する曲線となる。また r が μ_1 よりも小さくなるにつれて期待面積占有率は漸近的に 1 に接近する。

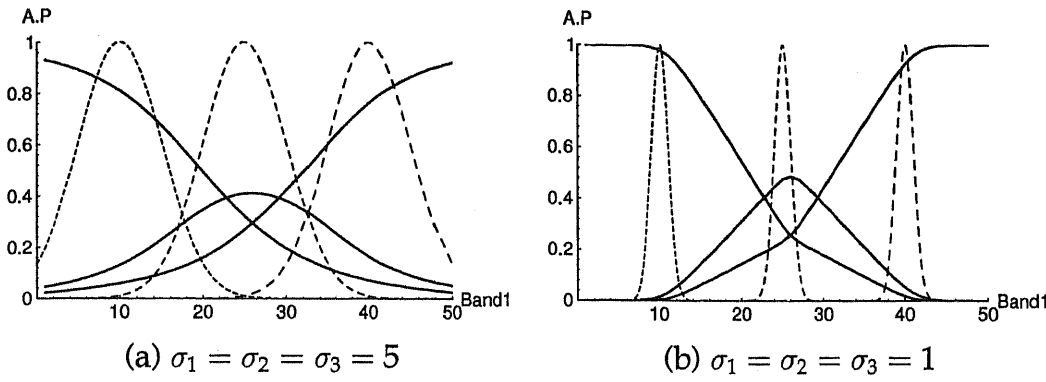


図 4.2: 1 バンド 3 クラスミクセルの期待面積占有率。

$r > \mu_2$ の場合 $r < \mu_1$ の場合と全く逆であり、定義域内部では $a = 0$ で関数が最大値をとる。
 r が μ_2 よりも大きくなると期待面積占有率は漸近的に 0 に接近する。

4.5.2 1 バンド 3 クラスミクセルの期待面積占有率

1 バンド 2 クラスミクセルの場合と同様の方法により同時確率密度を求めることができる。
 $p(r, a_1, a_2) = N\left(\sum_{i=1}^3 a_i \mu_i, \sum_{i=1}^3 (a_i \sigma_i)^2\right)$ ただし $a_3 = 1 - a_1 - a_2$ である。ここで 3 クラスミクセルの場合の期待面積占有率は以下のように計算する。

$$E\{a_1\} = \int_0^1 a_1 da_1 \int_0^{1-a_1} p(a_1, a_2 | r) da_2 \tag{4.29}$$

$$E\{a_2\} = \int_0^1 a_2 da_2 \int_0^{1-a_2} p(a_1, a_2 | r) da_1 \tag{4.30}$$

式 (4.29)、式 (4.30) がそれぞれクラス 1 とクラス 2 の期待面積占有率であり、クラス 3 の期待面積占有率は $E\{a_3\} = 1 - E\{a_1\} - E\{a_2\}$ により求める。

その結果を図 4.2 に示す。ここで仮に左側の密度関数から順番にクラス 1 ~ 3 とすると、クラス 2 の期待面積占有率がクラス 2 の密度関数の中心付近でも小さな値に推定されている点の特徴的である。しかし、たとえ観測値がクラス 2 の確率密度の中心付近であったとしても、そのミクセルはクラス 2 が主要な割合を占めるミクセルか、またはクラス 1 とクラス 3 が主要な割合を占めるミクセルのどちらかであり、そのどちらが正しいかは観測値だけでは判断することはできない、ということを考えれば、このような推定値となるのは自然な結果である。逆に考えると、1 バンドの観測値のみから 3 クラスミクセルの面積占有率を推定する問題は、条件不足のために精度のよい推定は難しいとも考えられる。

4.5.3 2 バンド 2 クラスミクセルの期待面積占有率

次にクラス確率密度 $p_i(\mathbf{r})$ が正規密度 $N_2(\mu_i, \Sigma_i)$ で表される場合を扱う。ここで Σ_i は分散共分散行列

$$\Sigma_i = \begin{pmatrix} \sigma_{i1}^2 & \rho_i \sigma_{i1} \sigma_{i2} \\ \rho_i \sigma_{i1} \sigma_{i2} & \sigma_{i2}^2 \end{pmatrix} \tag{4.31}$$

である。2 バンドの場合でも、特性関数を用いて簡単に同時確率密度 $p(\mathbf{r}, \mathbf{a})$ を得られる。

$$p(\mathbf{r}, \mathbf{a}) = N_2(a\mu_1 + (1-a)\mu_2, a^2\Sigma_1 + (1-a)^2\Sigma_2) \tag{4.32}$$

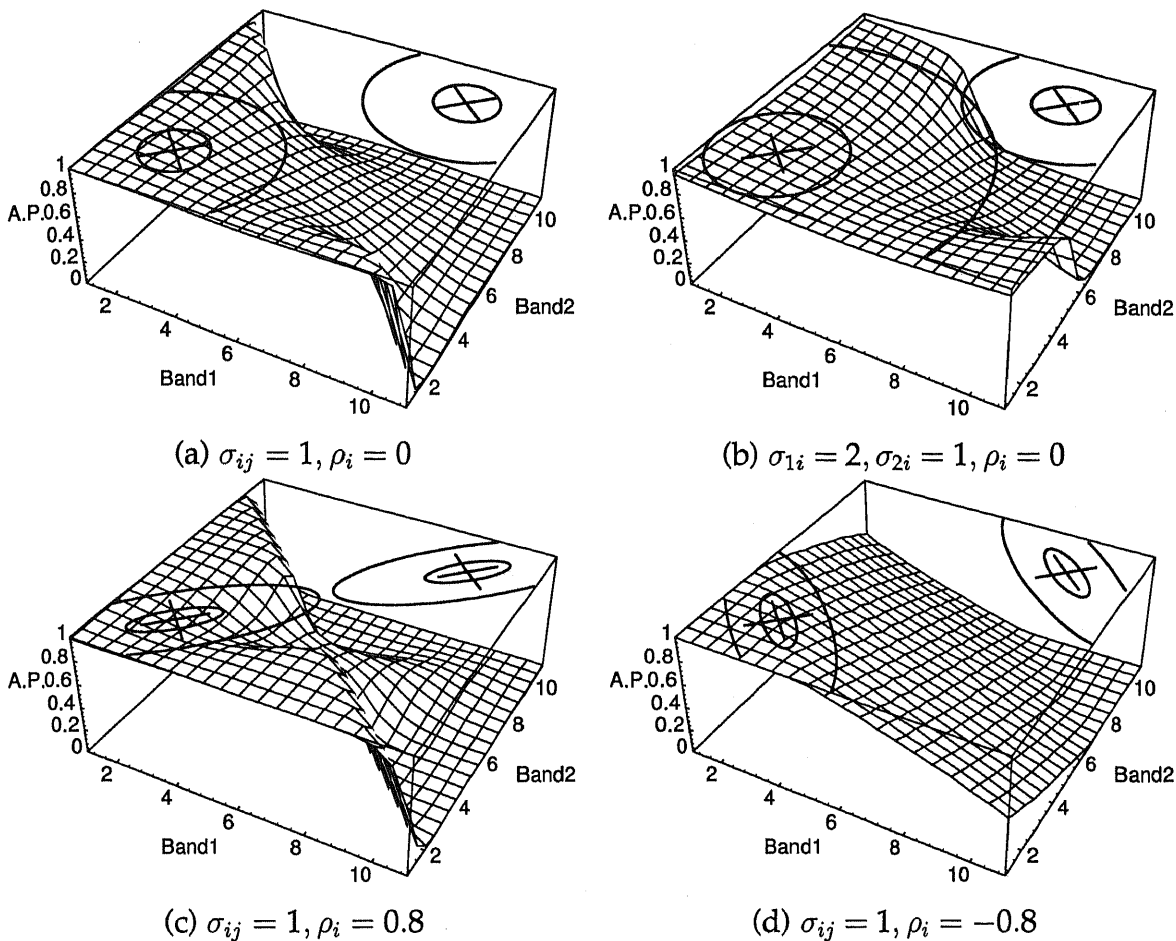


図 4.3: 2 バンド 2 クラス ミクセルの期待面積占有率。

ただし第 4.5.1 節の場合とは異なり式 (4.32) は複雑な式である。この式のみから最尤面積占有率および期待面積占有率の値を推測するのは難しい。そこで各パラメータに具体的な数値を代入し、数値計算により求めた期待面積占有率を図 4.3 に図示した。この図では横軸がバンド 1、縦軸がバンド 2 の観測値を示しており、クラス密度関数 (正規密度) の中心は \times 印で示してある。また上面に描かれた 2 本の等確率密度楕円は、中心からのマハラノビス距離が 1 と 3 の線を表し、おおよそクラス密度関数の広がり の程度を示している。

パラメータを 3 通りに変えた場合のクラス 1 の面積占有率を図 4.3 に示す。2 クラスの密度関数の分散が同一で相関係数が 0 の場合 (a) は、期待面積占有率はクラス 1 の面積占有率が大きい領域と小さい領域とで対称な形になっている。しかし (b) 図のようにクラス 1 の分散が大きい場合には、クラス 1 の面積占有率が大きい領域がクラス 2 の側に食い込むようになる。また (c) 図のように、相関係数の値によっても期待面積占有率の値は大きく変化することがわかる。

各図を比較してみると、各クラスの密度関数の中心付近の期待面積占有率は、どの場合においても大差がない。しかし中心から離れた領域では、パラメータの値によって期待面積占有率は大きく異なる場合がある。極端な場合では、例えば図 4.3 の (10, 4) 付近で (a)、(c) 図と (b) 図では主要なクラスに対する推定で正反対の結論が出るように、どちらのクラスが主要であるかという点に対して大きな誤差が生じることがある。このように、クラス内分散を考慮しない手法では、分散や相関係数などの値によって推定結果が大きく外れる場合があることに注意する必要がある。

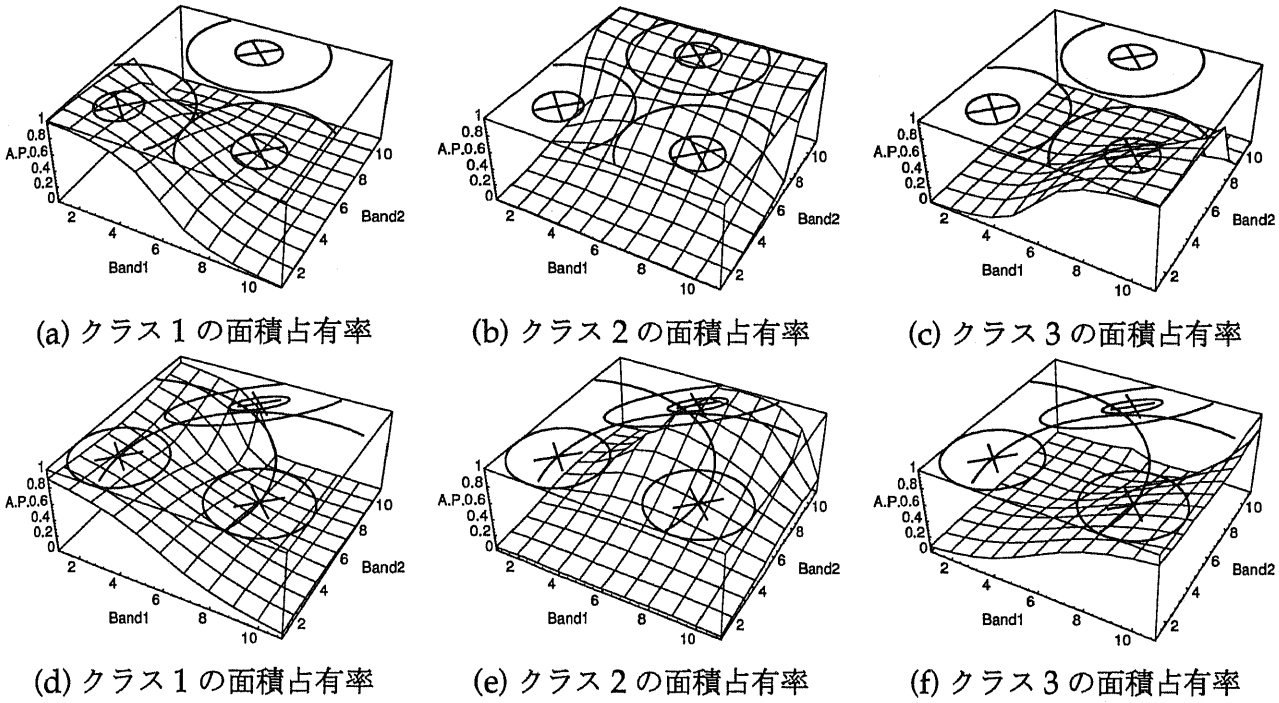


図 4.4: 2 バンド 3 クラスミクセルの期待面積占有率。(a)、(b)、(c): $\sigma_{ij} = 1, \rho_i = 0$ 、(d)、(e)、(f): $\sigma_{ij} = 1, \rho_1 = 0.8, \rho_2 = -0.8, \rho_3 = 0.8$ 。

4.5.4 2 バンド 3 クラスミクセルの期待面積占有率

2 バンド 3 クラスミクセルの場合も、今までと全く同様の方法により期待面積占有率を求めることができる。パラメータを 2 通りに変えた場合の 3 クラスそれぞれの期待面積占有率を図 4.4 に示した。この場合もパラメータの値によっては異なる結論が得られる。例えば左上隅の領域は、(a) 図ではクラス 2 が主要なミクセルとなっているが、(d) 図においてはクラス 1 が主要なミクセルとなっている。この結果からも、クラス内分散を考慮することの重要性が示されている。

4.5.5 p バンド k クラスミクセルの期待面積占有率

以上の議論から、一般の p バンド k クラスミクセルの期待面積占有率の計算方法はほぼ自明である。つまり式 (4.15) により同時密度関数を計算し期待面積占有率を算出する。このように本研究の手法は任意の p バンド k クラスミクセルに自然に拡張できる方法である。

4.6 実験

期待面積占有率が実際に面積占有率を適切に推定できることを示すため、以下のシミュレーションをおこなった。

1. 正規密度に従う確率変数 X_1 および X_2 の実現値を乱数で発生させる。
2. 面積占有率密度 $f(a)$ に従う確率変数 A の実現値 ($0 < a < 1$) を乱数で発生させる。
3. $r = ax_1 + (1 - a)x_2$ を計算し、配列 $h(a, r)$ に登録する。(1) ~ (3) を M 回繰り返す。

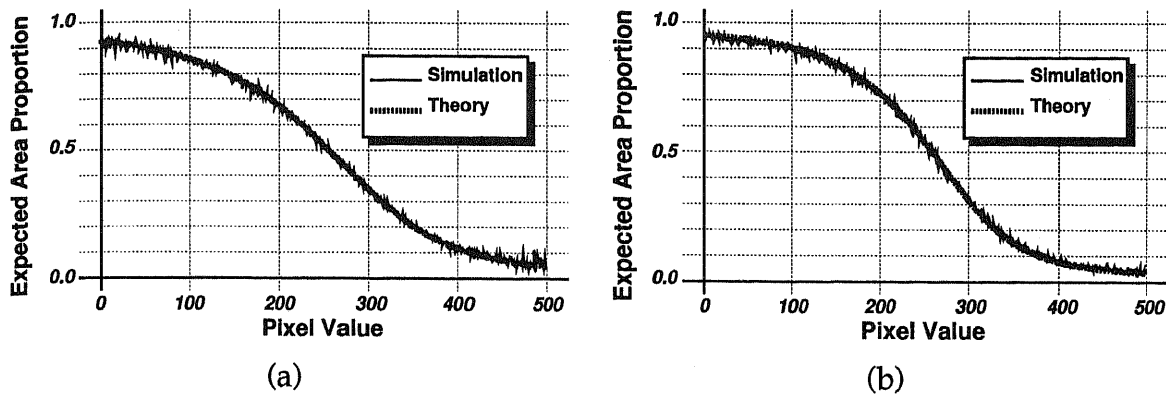


図 4.5: 面積占有率のシミュレーションの結果。(a) 面積占有率密度が一様分布の場合、(b) U型面積占有率密度を用いた場合。

4. 配列 $h(a, r)$ からそれぞれの r に対する条件つき密度 $p(a|r)$ を求め、次に期待値 $E\{a\} = \sum a_i p(a_i|r) f(a_i)$ を計算する。

このシミュレーションでは、クラス1の確率密度を $N(100, 70)$ 、クラス2の確率密度を $N(400, 50)$ とし、また $M = 10^5$ とした。シミュレーションから計算した値と理論値とを比較することで、本手法が面積占有率を適切に推定しているかどうかを検証する。

その結果を図 4.5 に示す。まず図 4.5(a) は、面積占有率密度として一様分布を用いた場合である。この結果では、乱数から計算した期待面積占有率と理論値とがよく一致している。次に (b) では、面積占有率密度として一様分布ではない U 型面積占有率密度を用いた場合を示す [71]。このように、異なる面積占有率密度を用いると異なる期待面積占有率曲線が得られるが、この場合でも乱数からの計算値と理論値とはよく一致している。この結果から、本研究の手法が面積占有率の期待値という意味で、面積占有率を適切に推定できることを確認できた。なお実際の画像への本手法の応用については第 5 章で議論する。

4.7 まとめ

本研究では、ミクセルの面積占有率を期待値に基づいて推定する方法を提案した。各クラスからの観測値を確率密度にしたがう確率変数として捉え、この確率密度をモデルに取り入れてミクセルの期待面積占有率を計算する方法を提案した。まずデルタ関数を確率密度モデルとした場合に従来の研究で用いられる推定式と同等の結果を導き、次に実用的には重要な正規密度を確率密度モデルとして用いた場合の期待面積占有率が、デルタ関数の場合の自然な拡張となっていることを示した。そしてクラス内分散の大きさがクラスごとに異なる場合には、クラス内分散を考慮しない方法では大きな誤差を招く場合があることを指摘した。最後にシミュレーションによって本研究の手法が面積占有率の値を期待値の意味で適切に推定していることを確認し、本研究がミクセルの面積占有率の計算に有用な方法であるとの結論を得た。なお本研究の手法は、任意の p バンド k クラスミクセルに対して計算可能であり、また任意の密度関数に対しても FFT などを用いて計算可能と、拡張性の高い方法として定式化されている。

今後の課題としては、周囲画素の情報(文脈情報)の活用が重要な課題である。確かに期待面積占有率は、多数のミクセルの集合に対する期待値の意味で、面積占有率を適切に推定することができる方法である。しかしこれは画素値の分布という画像の特徴空間の情報を用いた推定方法

であり、これに加えて画像空間上での周囲画素の画素値という文脈情報を総合したうえでの最適な推定方法は何かという問題点については、今後の大きな課題であると考えている。

第5章

ミクセル密度を含めた混合密度推定と統計的な画像分類への応用

5.1 概要

本章の目的は、画像ヒストグラムを混合密度とみなす場合に、ミクセルの影響を含めた混合密度推定の方法を提案することである。ここでミクセルとは一画素の内部に複数の分類クラスを含む画素を指す用語である。このような画素は複数のクラスの統計的性質が混合した画素であるから、ピュアピクセルとは全く異なる確率過程によって生成された画素と考えねばならない。ゆえに、ミクセルが存在する画像の画像ヒストグラムには、ピュアピクセルの集合が形成する密度関数の他に、ミクセルの集合が形成する密度関数が暗に隠されているとの仮説を立てる。このようにミクセル集合が形成する新しい密度関数「ミクセル密度」を、本研究ではミクセルの確率過程をモデル化することによって導出する。ミクセル密度を含めた混合密度推定の手法にはEMアルゴリズムを用い、実際のシミュレーションと衛星画像に対してこの手法を適用する。そして混合密度推定の結果を用いてベイズの決定則により画素単位の分類をおこない、さらに個々のミクセルに対する面積占有率を計算することで、本手法がミクセルを含む画像を有効に画素単位分類できることを示す。

5.2 従来研究の問題点

実世界は連続的な世界であるのに対し、センサを通して実世界を記録した画像は離散的な世界である。したがって、画素とは数学的な意味での「点」のような存在ではなく、その内部には実世界のある領域の情報が非可逆的に埋め込まれている。もしその領域が一つのクラスしか含まないような均質な領域である場合には、そのような画素を純粋画素 (pure pixel) とよぶ。一方その領域が複数のクラスを含むような不均質な領域である場合には、そのような画素を混合画素 (mixed pixel) またはミクセルと呼ぶ。このようなミクセルが出現する割合は、センサの解像度と実世界の領域境界の空間的変動のスケールとの比に依存する [72]。例えばリモートセンシング画像や医用画像の場合はセンサの解像度が領域境界の空間的変動のスケールに比べて粗いため、画像中に高い割合でミクセルが出現する可能性がある。

これらの画像を処理する場合にミクセルの影響を無視することができないのは、ミクセルが複数の分類クラスから構成されている画素であり、各クラスの統計的性質が混合するという、ピュアピクセルとは全く異なる確率過程で形成された画素だからである。ミクセルを形成する確率過程がピュアピクセルと異なるということは、ミクセルを「雑音」が含まれたピュアピクセルとして扱う方法が不適切であることを意味する。これまでに提案されきた多くの統計的手法は、

すべての画素がピュアピクセルであるとの前提のもとに導かれている。例えば統計的手法には判別分析に基づくしきい値選定法 [48] や、画像ヒストグラムをパラメトリックな混合密度とみなして混合密度推定を適用する方法などがある。これらはピュアピクセルに対する「雑音」には対処できるが、ミクセルのような異なる確率過程で生じた画素については考慮していない。したがってこれらの方法をミクセルを含む画像にそのまま適用することはできない。

そこで本研究の目的は、このようなミクセルを含む画像に対しても適用できる統計的手法を提案することにある。特に画像ヒストグラムに着目し、これがピュアピクセルの密度関数ばかりではなくミクセルが形成する新たな密度関数をも含む混合密度になっているのではないか、というのが本研究の仮説である。この仮説に基づき、ミクセル集合が画像ヒストグラムに新たな密度「ミクセル密度」を形成することをミクセルの線形モデルから導く(第5.4節)。また本研究では、混合密度推定によって各クラスの密度関数のパラメータを推定することも目的の一つである。ミクセルに関する従来の研究では、面積占有率を計算するために必要とされる各クラスの分光行列や密度関数のパラメータなどは、既知であるとされるかまたは多くの教師データから回帰的に求めるものであるとされた。しかし多様な撮影条件で得られた大量の画像を自動的に処理するために、EMアルゴリズムを用いた混合密度推定の手法で画像から密度関数のパラメータを自動的に抽出する(第5.5節)。最後に第5.6節では本手法をシミュレーションや衛星画像に適用してモデルの有効性を検証する。さらに個々のミクセルの面積占有率も含めた衛星画像の分類方法を提案する。

5.3 問題の定式化

5.3.1 ミクセルの線形モデル

ミクセルの線形モデルについては、第4.3.1節ですでに述べた。本章で扱う問題は以下である。

問題2 画像中のミクセルの集合に注目する。式(4.3)が成立し、かつ各クラスの確率密度 $p_i(\mathbf{x}_i)$ および面積占有率密度 $f(\mathbf{a})$ が既知のとき、 k クラスミクセルからの観測ベクトル \mathbf{r} が形成する新たな確率密度を求める。

つまり、多数のミクセルが画像中に存在すれば、そのミクセル集合は画像ヒストグラム上に新たな密度関数「ミクセル密度」を形成する、という仮説に基づきこの密度関数を求めることが本研究の目的である。

5.3.2 ミクセル密度を含む混合密度推定

本研究では画像ヒストグラムを有限個の密度関数が混合した有限混合密度であるとみなす。 M 個のクラスに対応した密度関数が混合した有限混合密度は次のように表される。

$$p(\mathbf{r}|\Phi) = \sum_{i=1}^M \alpha_i p_i(\mathbf{r}|\phi_i) \quad (5.1)$$

ここで $p_i(\mathbf{r}|\phi_i)$ は各クラスの密度関数であり、パラメータ ϕ_i の条件つき密度であることを強調して表記している。また α_i は各密度関数の混合比であり、 $\alpha_i \geq 0$ および $\sum_{i=1}^M \alpha_i = 1$ が成立する。この場合 $p(\mathbf{r}|\Phi)$ が画像ヒストグラムに相当する。

ここで、混合比と面積占有率という用語を混同しないように注意する必要がある。面積占有率とは1個のミクセル内で各クラスが占める割合を指すが、一方の混合比は密度関数の割合、言い換えれば画像全体で各クラスに属する画素が占める画素数の割合を指す。また混合密度とミクセル密度も混同しやすい概念である。混合密度は式(5.1)のように各密度関数の和として得られるのに対し、後に述べるように、ミクセル密度は本質的に各密度関数の畳み込みとして得られる。

5.4 ミクセル密度

5.4.1 ミクセル密度の導出

ミクセルの線形モデル式(4.3)がすべての基本である。本研究ではこの関係からミクセル密度を導く。さて第4.4.3節において、ランダム化された同時確率密度を式(4.18)のように求めた。この式を本章では第4章とは別の方向に活用する。さて本研究が求める「ミクセル密度」とは、ミクセルの観測値が観測ベクトル \mathbf{r} となる確率 $p(\mathbf{r})$ を与える密度関数のことを指し、式(4.18)を面積占有率ベクトルのパラメータ空間全体にわたって積分することによって得られる。

$$d(\mathbf{r}) = \iint_{\Psi} g(\mathbf{r}, \mathbf{a}') d\mathbf{a}' = \iint_{\Psi} p(\mathbf{r}, \mathbf{a}') f(\mathbf{a}') d\mathbf{a}' \quad (5.2)$$

このような計算は「ランダム化」とよばれる。式(5.2)のようにランダム化された密度関数 $d(\mathbf{r})$ が本研究で提案するミクセル密度である。

式(5.2)で表される密度関数は、ピュアピクセル密度がたんなる重なり合いで混合したものではなく、式(4.8)でピュアピクセル密度を畳み込み、それを式(5.2)でさらに積分したものである。したがってミクセル密度とは、ピュアピクセル密度とは本質的に異なる確率過程を表す密度関数であるといえる。また式(5.2)で注意すべき点は、ミクセル密度が独自の形状パラメータをもたない点である。式(4.18)の $g(\mathbf{r}, \mathbf{a}')$ は、それを構成するピュアピクセル密度の形状と面積占有率密度の形状から自動的に定まるため、ミクセル密度に関するパラメータはその混合比1個だけである。

5.4.2 面積占有率密度が一様分布の場合のミクセル密度

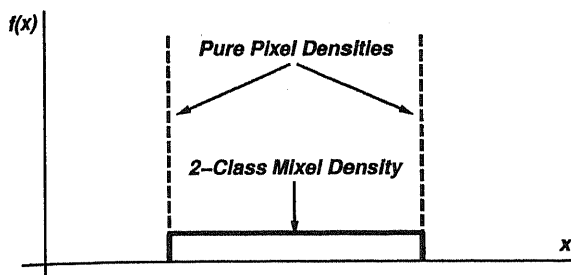
ここでは具体的にいくつかの例についてミクセル密度の計算をおこない、ミクセル密度がどのような特徴的な形状をもっているかを順に示していく。まず面積占有率密度に一様分布を用いた場合を計算する。一様分布は面積占有率ベクトル \mathbf{a}' がパラメータ空間 Ψ 内に一様に分布する場合であり、 $f(\mathbf{a}') = 1/|\Psi|$ で表される。ただし $|\Psi| = \iint_{\Psi} d\mathbf{a}'$ である。

例1 まず最初に1バンド2クラスミクセルの場合を考える。さらに各クラスからの観測値が全く同一でクラス内分散が0の場合を考える。このとき各クラスの確率密度 $p_i(r)$ はデルタ関数で表すことができる。

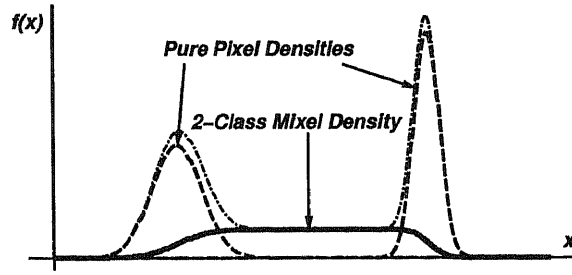
$$p_i(r) = \delta(r - \mu_i) \quad (5.3)$$

ここで各クラスからの観測値を μ_i としている。すると同時確率密度 $p(r, a)$ は以下のように求められる。

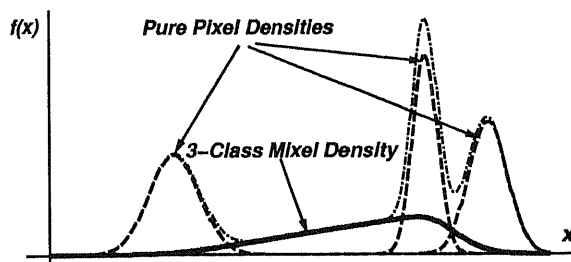
$$p(r, a) = \delta(r - \{a\mu_1 + (1-a)\mu_2\}) \quad (5.4)$$



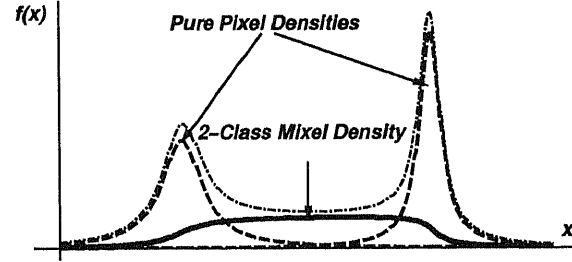
(a) デルタ関数から導いた1バンド2クラスミクセル密度



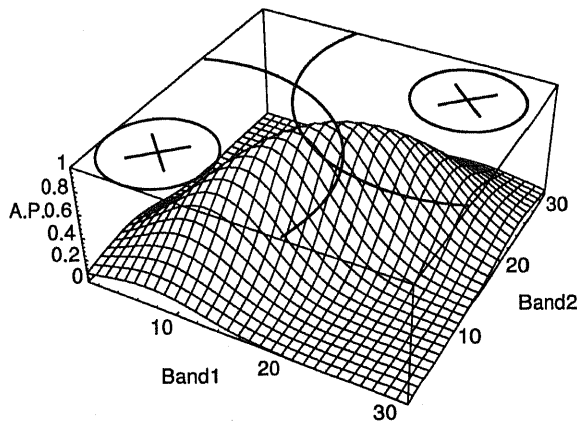
(b) 正規密度から導いた1バンド2クラスミクセル密度



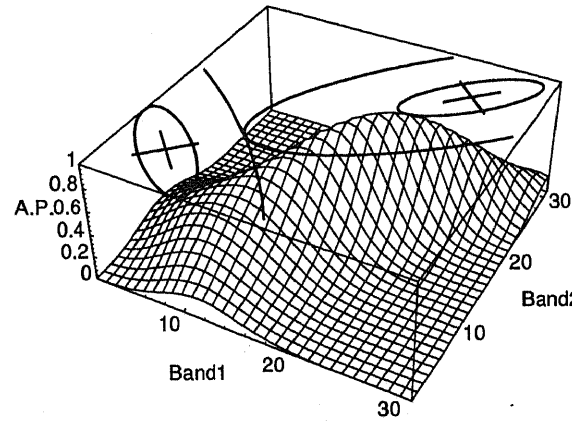
(c) 正規密度から導いた1バンド3クラスミクセル密度



(d) コーシー密度から導いた1バンド2クラスミクセル密度



(e) 正規密度から導いた2バンド2クラスミクセル密度 ($\sigma_{ij} = 5, \rho_i = 0$)



(f) 正規密度から導いた2バンド2クラスミクセル密度 ($\sigma_{ij} = 5, \rho_1 = -0.8, \rho_2 = 0.8$)

図 5.1: 面積占有率密度が一様分布の場合のミクセル密度。

ここで $\mu_1 < \mu_2$ とし、また面積占有率密度が一様分布 $f(a) = 1$ であると仮定すると、式 (5.2) によりランダム化したミクセル密度は以下ようになる。

$$d(r) = \int_0^1 p(r, a) \cdot 1 \cdot da \tag{5.5}$$

$$= \begin{cases} \frac{1}{\mu_2 - \mu_1} & \mu_1 \leq r \leq \mu_2 \\ 0 & r < \mu_1, r > \mu_2 \end{cases} \tag{5.6}$$

つまりミクセル密度は図 5.1(a) のように区間 (μ_1, μ_2) の一様分布となる。

例 2 次に、より実用的な密度関数として正規密度を考える。クラス i の確率密度 $p_i(r)$ が多次元正規密度 $N_p(\mu_i, \Sigma_i)$ であるとする。ここで多次元正規密度は以下のように表される。

$$p_i(r) = \frac{e^{-\frac{1}{2}(r - \mu_i)^t \Sigma_i^{-1} (r - \mu_i)}}{(2\pi)^{p/2} |\Sigma_i|^{1/2}} \tag{5.7}$$

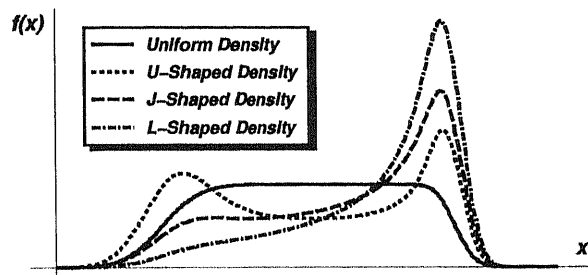


図 5.2: 4 種の面積占有率密度を用いた場合の 2 クラスミクセル密度。

ただし μ_i は平均ベクトル、 Σ_i は共分散行列である。このとき p バンド k クラスミクセルの同時確率密度 $p(\mathbf{r}, \mathbf{a})$ も正規密度となる。

$$p(\mathbf{r}, \mathbf{a}) = N \left(\sum_{i=1}^k a_i \mu_i, \sum_{i=1}^k a_i^2 \Sigma_i \right) \quad (5.8)$$

まず 1 バンド 2 クラスミクセル密度および 1 バンド 3 クラスミクセル密度を図 5.1(b) および (c) に示す。2 クラスミクセル密度は 2 個の密度関数のピークの間には平坦な部分をもつ関数となっている。図 5.1(d) のコーシー密度の 2 クラスミクセル密度も、やはり 2 個の密度関数のピークの間には平坦な部分をもつ関数であるが、コーシー密度の裾が重いためこのミクセル密度も裾が重い関数となっている。また図 5.1(e)、図 5.1(f) のように、2 バンド 2 クラスミクセル密度も全く同様の手法で計算できる。上面の×印が 2 個のピュアピクセル密度の中心を表しているが、ミクセル密度はこれら 2 個の密度関数のピークの間には平坦な領域をもつ関数となっている。同様の手法で任意の p バンド k クラスミクセルのミクセル密度が導出できる。

5.4.3 面積占有率密度が一様分布でない場合のミクセル密度

面積占有率密度が実際の画像においては一様分布ではなく L 型・J 型・U 型の形状となることを第 4 章ですでに論じた。これら 3 種に一様分布を加えた 4 種のアreal占有率密度を用いた場合の 2 クラスミクセル密度の形状を図 5.2 にまとめた。この図からわかるように、ミクセル密度の形状は面積占有率密度の形状に大きな影響を受ける。一様分布の場合はミクセル密度は中心部がほぼ平坦な形状だったが、その他の場合は面積占有率密度の形状を反映した関数となっている。このように、面積占有率密度のモデル化は実際の応用では重要な問題となる。

5.4.4 ミクセル密度の形状

図 5.1 に示されたミクセル密度の形状の中で、特に 2 クラスミクセル密度の形状が最も特異である。これらはどれも 2 個のピュアピクセル密度のピークの間には平坦な部分をもつ形状であるが、この種の形状をもつ関数は混合密度としてほとんど用いられたことがない。このような特異な形状の意味を以下で考察する。

画像ヒストグラムには、複数の鋭いピークの間には「裾の重い」平坦な部分が存在する場合があります。このような場合には正規密度を用いた混合密度推定ではよい適合が得られないことが指摘されていた。この問題の解決策として提案された手法は主に 2 つある。

1. t 分布 [73] やピアソンの分布系 [74] などの裾の重い密度関数を混合密度に用いる。

2. エッジ部分に対応する画素を取り除いて混合密度推定をおこなう [75]。

しかし(1)の方法で代替案として用いる密度関数は、多くの場合ピークの周囲すべての方向に対して裾が重くなる。よってある方向に対する適合を良くすることは、往々にして他の方向に対する適合を犠牲にすることを意味する。またそのような密度関数を用いる根拠も明確ではない。(2)の方法はそのような問題を起こす画素をあらかじめ取り除くという実用的な方法ではあるが、エッジ部分の画素を取り除く処理によって重要な画像情報を捨ててしまう可能性がある。

このような方法に対して、本研究ではピュアピクセル密度とミクセル密度とを組み合わせた混合密度を用いる方法を提案する。本研究の手法では、画像ヒストグラムのピーク部分はピュアピクセル密度が顕著な部分であり、一方の裾の重い部分はミクセル密度が顕著な部分であるとして混合密度を当てはめることになる。この方法ではエッジ画素などを捨てることなしに混合密度推定をおこなうことが可能である。

5.5 EM アルゴリズムを用いた混合密度推定

5.5.1 ミクセル密度を含む混合密度推定の特徴

さて画像の画素値ベクトルを標本 r_1, \dots, r_N と考え、この標本からパラメータ ϕ_1, \dots, ϕ_M および混合比 $\alpha_1, \dots, \alpha_M$ を推定する問題が混合密度推定の問題である。これは対数尤度関数 $L(\Phi)$ の最大化、つまり Φ^* の最尤推定を得ることを目的とする最大化問題に置き換えることができる。ここで対数尤度関数は以下のように定義される関数である。

$$L(\Phi) = \sum_{k=1}^N \log p(r_k | \Phi) \quad (5.9)$$

$p(r|\Phi)$ は式(5.1)で与えられる混合密度であり、 Φ はパラメータ空間 Ω に存在するパラメータである。また N は本研究の場合は画像の画素数と等しい。この最大化問題を解くためのアルゴリズムが、後に述べる EM アルゴリズムである。

しかし EM アルゴリズムについて説明する前に、本研究の混合密度推定が従来の混合密度推定と異なる点を明確にする。最も大きな相違点は密度関数の数 M である。従来の混合密度推定ではクラス数が L のときは密度関数の数も $M = L$ であった。しかしミクセル密度も混合密度に含める場合には密度関数の数は $M > L$ となる。例えば2クラスミクセル密度までを含める場合は、 $M = L + {}_L C_2$ である。このように密度関数の数が大きくなることは本手法の欠点となりうるが、この問題に対しては以下の対策を考えている。

- 通常は2クラスミクセル密度までの混合密度で十分であり、 k クラスミクセル ($k > 2$) の出現確率はミクセル密度を形成するほどには大きくなる。
- 事前知識によってある組み合わせのミクセルの出現確率が小さいと判断できる場合には、そのミクセル密度を混合密度から除外できる。
- ミクセル密度によって裾の重い部分がうまく適合するため、ピュアピクセル密度のみの混合密度よりは少ないクラス数でヒストグラムを表現できる。

5.5.2 EM アルゴリズム

式 (5.9) の最大化問題を解く標準的な手法の一つが EM (expectation maximization) アルゴリズムである [76, 77]。EM アルゴリズムでは、まず E ステップにおいて期待値を求め、次に M ステップでそれを最大にするパラメータの値を求めるのが基本的な処理である。すなわち E ステップでは

$$Q(\Phi|\Phi^{(p)}) = \sum_{i=1}^M \left[\sum_{k=1}^N Z(i|\mathbf{r}_k, \Phi^{(p)}) \right] \log \alpha_i + \sum_{i=1}^M \sum_{k=1}^N \log p_i(\mathbf{r}_k|\phi_i) Z(i|\mathbf{r}_k, \Phi^{(p)}) \quad (5.10)$$

と期待値を求め (expectation)、M ステップでは

$$\Phi^{(p+1)} = \operatorname{argmax}_{\Phi \in \Omega} Q(\Phi|\Phi^{(p)}) \quad (5.11)$$

と最大値を求める (maximization)。ここで p は反復回数を表し、

$$Z(i|\mathbf{r}_k, \Phi^{(p)}) = \frac{\alpha_i^{(p)} p_i(\mathbf{r}_k|\phi_i^{(p)})}{p(\mathbf{r}_k|\Phi^{(p)})} \quad (5.12)$$

は現在の近似的な最尤推定値 $\Phi^{(p)}$ と画素値ベクトル \mathbf{r}_k という条件のもとで k 番目の標本がクラス i に属する事後確率を表す。

このように一般的な EM アルゴリズムでは、E ステップと M ステップの 2 ステップにより Φ を更新する。しかし式 (5.10) の第 1 項は混合比 $\alpha_1, \dots, \alpha_N$ しか含まず、また第 2 項はパラメータ ϕ_1, \dots, ϕ_N しか含まないので、M ステップでは独立に最大化できる。さらに ϕ_1, \dots, ϕ_N が相互に独立なパラメータであると仮定すると、M ステップの更新式は以下のような簡単な式となる。

$$\alpha_i^{(p+1)} = \frac{1}{N} \sum_{k=1}^N Z(i|\mathbf{r}_k, \Phi^{(p)}) \quad (5.13)$$

$$\phi_i^{(p+1)} = \operatorname{argmax}_{\phi_i \in \Omega_i} \sum_{k=1}^N \log p_i(\mathbf{r}_k|\phi_i) Z(i|\mathbf{r}_k, \Phi^{(p)}) \quad (5.14)$$

正規密度などの場合は式 (5.14) を明示的に解くことができ、E ステップと M ステップを実際には 1 ステップに統合できるのは大きな利点である [76]。

5.5.3 ミクセル密度を含む場合の EM アルゴリズム

しかしミクセル密度を含む混合密度に EM アルゴリズムを適用する際には、式 (5.14) で仮定したパラメータ ϕ_1, \dots, ϕ_N の独立性が成立しないという問題が生じる。第 5.4.1 節で述べたように、ミクセル密度はそのミクセルに混合しているクラスのピュアピクセル密度から自動的に決定される。このことを逆に考えると、ミクセル密度を通してすべてのピュアピクセル密度の形状パラメータが相互に影響しあい、各パラメータの独立性が崩れることになる。したがって式 (5.14) の

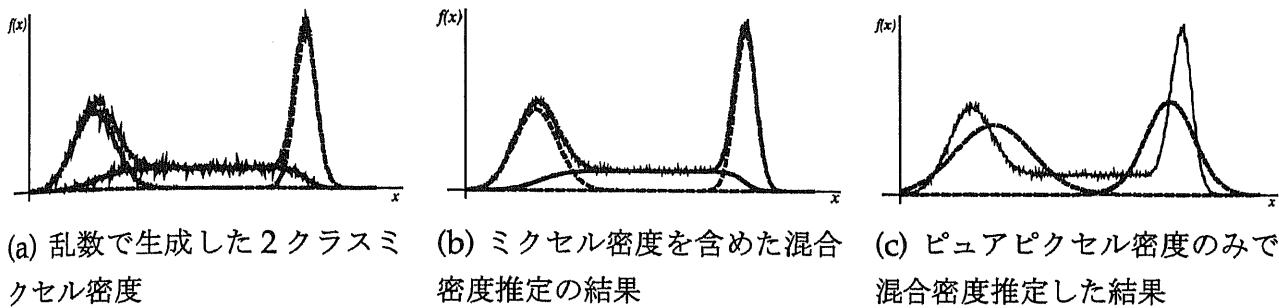


図5.3: 2クラスミクセル密度のシミュレーションと混合密度推定結果。各パラメータの値は表5.1を参照。

ようにMステップで各パラメータを独立に最大化する更新式は使えない。そこで本研究では、 α_i の更新には式(5.13)を用い、その他のパラメータの最大化は式(5.10)の第2項を準ニュートン法(quasi-Newton method)によって最大化する。ただしこのようにパラメータが相互に影響し合うことで問題は非常に非線形となるため、EMアルゴリズムの収束はそれほど早くならず、まれに収束しない場合があることには注意を要する。

5.5.4 初期値推定

EMアルゴリズムではパラメータの初期値推定が重要である。データから混合正規密度のパラメータを推定する方法には、モーメント法[78]などの手法が提案されているが、ミクセル密度を含む場合についての議論はない。また教師データがあればそれから統計量を算出する方法が最も簡便であるが、本研究では画像データから自動的に初期値を推定するための方法として、1バンドの場合にはスケールスペースを活用した方法[79]を用いる。この方法はヒストグラムの形状を活用する方法であり、図B.1に示すようにヒストグラムの2次微分のゼロ交差点を用いて確率密度関数のピーク位置と幅を推定できる。この方法はヒストグラムのピーク付近の形状を見るため裾の重いミクセル密度部分には影響を受けにくいという利点がある反面、あるはずのピークが他のピークに埋もれてしまうとそれを全く検出できないという欠点がある。詳しい説明については第B章を参照してほしい。

次に混合比の初期値推定である。まずピュアピクセル密度の混合比は2個のゼロ交差点の間の面積の割合を推定値とする。次にミクセル密度の混合比の推定は、ミクセル密度の数が組み合わさ的に増える場合には画像ヒストグラムだけでは情報不足である。そこでまず、ピュアピクセル密度のみを用いた混合密度推定により画像を仮分類し、その画像での各クラス間の隣接画素の割合からミクセル密度の混合比の初期値を推定するという方法を探る。

なおクラス数 L 自体を画像データから推定する問題に対しては、情報量規準を用いる方法などが提案されている。しかしこの問題自体が大きな問題であり、しかも本研究ではこの問題は議論の主題ではないことから、人間が前もってクラス数を定めることとした。

5.6 実験

本研究の提案手法の有効性を示すため、シミュレーションと実際の衛星画像を用いた実験をおこなう。

表 5.1: 図 5.3に関するパラメータ一覧。

分布	平均値 μ	標準偏差 σ	混合比率 α
正解			
Pure 1	100	30	0.333
Pure 2	400	15	0.333
Mixel 1 + 2	—	—	0.333
初期値推定			
Pure 1	99.5	29.5	0.397
Pure 2	399.0	17.0	0.420
Mixel 1 + 2	—	—	0.183
ミクセル密度を含めた推定結果 (b)			
Pure 1	102.2	30.37	0.345
Pure 2	399.2	15.15	0.340
Mixel 1 + 2	—	—	0.315
ピュアピクセル密度のみの推定結果 (c)			
Pure 1	134.2	54.31	0.536
Pure 2	383.4	35.20	0.464
Mixel 1 + 2	—	—	—

5.6.1 シミュレーション

以下の簡単なシミュレーションを用いて、本研究が提案するミクセル密度の形状および混合密度推定の精度を確認する。

1. 表 5.1のパラメータをもつ正規密度に従う確率変数 X_1 と X_2 の実現値を乱数で発生させる。
2. 面積占有率密度 $f(a)$ に従う確率変数 A の実現値を乱数で発生させる。
3. 新たな確率変数 $M = aX_1 + (1 - a)X_2$ を計算する。(1) ~ (3) を N 回繰り返す。
4. M のヒストグラムがミクセル密度となり、また X_1 、 X_2 および M のヒストグラムを足し合わせた密度関数が仮想的な画像ヒストグラムとなる。
5. 画像ヒストグラムに対して混合密度推定をおこない、パラメータの推定精度を確認する。

なお面積占有率密度としては一様分布 $f(a) = 1$ を用いた。その結果を図 5.3に示す。まず乱数によって発生させた図 5.3(a) の 2 クラスミクセル密度は、中心部の平坦な形状が 2 クラスミクセル密度に特徴的な形状を示している。次にピュアピクセル密度とミクセル密度を足した分布を画像ヒストグラムとみなし、この画像ヒストグラムに対して混合密度推定をおこなった。その結果を図 5.3(b) に、またパラメータの推定結果を表 5.1に示す。表 5.1の初期値推定では混合比率の推定精度が悪いが、混合密度推定によっていずれのパラメータもほぼ正しく推定されている。

なお図 5.3(a) のヒストグラムに対するピュアピクセル密度のみの混合密度推定では、図 5.3(c) に示すようによい適合が得られない。また表 5.1のパラメータの推定値も、ピュアピクセル密度



(a) 気象衛星 NOAA の赤外バンド (Ch.4) 画像
 (b) ピュアピクセルのみを表示した結果
 (c) ミクセルのみを表示した結果 (濃淡はクラス1の面積占有率に比例)

図 5.4: 気象衛星 NOAA 画像と画像分類結果。画像分類には図 5.5(d) の結果を用いる。

のパラメータとはかなりかけ離れている。この結果を実際の画像ヒストグラムに適用すると、画像ヒストグラムにミクセル密度が含まれている場合にピュアピクセル密度のみを用いた混合密度推定をおこなえば、推定されたピュアピクセル密度のパラメータは、「真の」ピュアピクセル密度のパラメータとは大きく異なった値になる可能性がある。そしてミクセル密度も混合密度に含めることによってはじめて、ピュアピクセル密度のパラメータも適切に推定できるようになるのである。

5.6.2 リモートセンシング画像を用いた実験

対象画像

次に実際のリモートセンシング画像を用いた実験をおこなう。対象画像には気象衛星 NOAA の AVHRR センサで観測した赤外画像 (チャンネル 4) を用いる。実験に用いた画像は NOAA 12 号が 1995 年 2 月 16 日に撮影した画像であり、その中から太平洋上の画素領域 (256×256) を切り取って実験対象画像とした。受信した衛星画像の画素値 r は $0 \leq r < 2^{10}$ のカウント値であるが、それをより物理的な意味をもつ特徴量である輝度温度に変換した値を画素値として用いた。そして温度が低いほど白くなるような濃淡として表示した画像を図 5.4(a) に示す。この領域には冬の季節風の吹き出しによる筋状の雲が並んでおり、ミクセルが発生しやすい気象条件となっている。またこの領域の画像中には雲と海しか存在しないため、クラス数は前もって 2 と設定した。

実験の手順

1. 画像のヒストグラムからそれぞれの混合密度のパラメータを推定し初期値とする。
2. ピュアピクセル密度のみの混合密度に対して EM アルゴリズム [76] を適用する。
3. ベイズの決定則により各画素をピュアピクセルに仮分類する。この分類画像上での各クラスの隣接関係からミクセルの割合を推定し、ミクセル密度の混合比の初期値とする。

4. ミクセル密度も含む混合密度に対してもう一度EMアルゴリズムを適用する。
5. 再度ベイズの決定則を適用し、各画素をピュアピクセルかミクセルに分類する。さらにミクセルに関しては期待面積占有率を計算する。

ベイズの決定則

実験の手順(5)のベイズの決定則とは、画素 i を以下の式を満たすクラス j に分類する方式である。

$$j = \arg \max_{1 \leq k \leq M} \alpha_k p(x_i | \phi_k) \quad (5.15)$$

通常の混合密度推定ならば M はクラス数に等しい。しかし本研究のようにミクセル密度を用いる場合は、 M はピュアピクセル密度の数にミクセル密度の数を加えたものになる。つまりミクセル密度を含む場合のベイズの決定則とは、ミクセルのクラスをピュアピクセルのクラスとは別に設けることで、ミクセルとピュアピクセルとを明確に区別して分類する方法である。例えばしきい値処理による手法や最尤法などの手法ではすべての画素をピュアピクセルに分類するのに対し、ミクセルを扱う手法として提案されてきた多くの手法は、実はすべての画素をミクセルに分類しその面積占有率を求める方法に他ならなかった。後者の手法ではピュアピクセルとはあくまで疑似的な存在であり、ある一つのクラス以外の面積占有率が十分小さな値以下と計算されたミクセルのことを指すものであった。しかし画像には確かにピュアピクセルが存在する。本研究で用いたようなベイズの決定則によって、すべての画素をピュアピクセル、またはすべての画素をミクセルに分類することになるという従来の研究の矛盾点[56]を解決することができる。

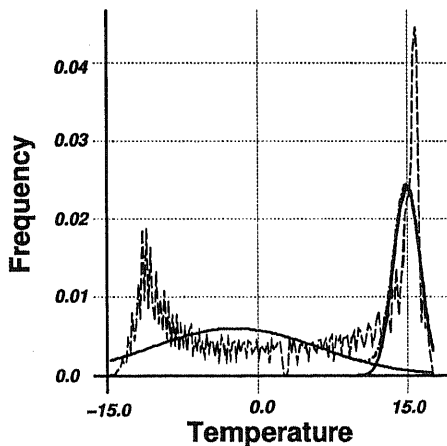
ベイズの決定則によりミクセルに分類された画素に対しては面積占有率を算出する。その計算には期待値に基づいた面積占有率の推定法として第4章で提案した方法を用いる。なおピュアピクセルに対しては面積占有率は定義しない。

混合密度推定の結果

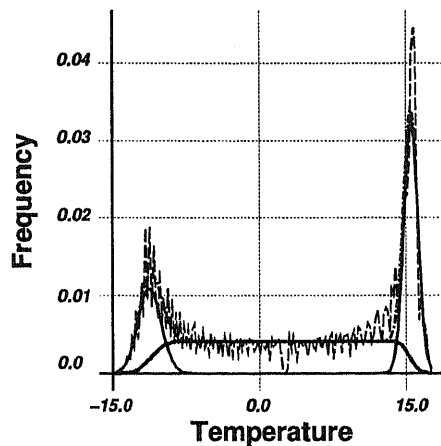
まず混合密度推定の結果を図5.5に示す。図中の点線は図5.4(a)の画像のヒストグラムであり、横軸は輝度温度(摂氏)を表している。低温側のピークが雲、高温側のピークが海に対応している。この2つのピークの間には平坦な部分があることに注目する。

まず、ミクセル密度を用いた場合の混合密度推定の結果を図5.5(b)~(d)に、また用いない場合の結果を図5.5(a)に示す。なお(b)~(d)には第4章で述べた3種の面積占有率密度を用いている。その結果は(a)と(b)~(d)の間の差が顕著に出ている。(a)の場合は低温側の密度関数が非常に幅の広い密度関数として推定されているが、これは明らかにヒストグラムの平坦な部分に影響された推定結果である。それに対し図5.5(b)~(d)では、2つのピークは2つのクラスのピュアピクセル密度として推定され、また中央の平坦な部分は2クラスミクセル密度として推定されている。中でも(c)と(d)の適合が比較的良好であるが、(c)では雲の密度関数の混合比が小さく、また(d)では大きく推定されている点異なる。

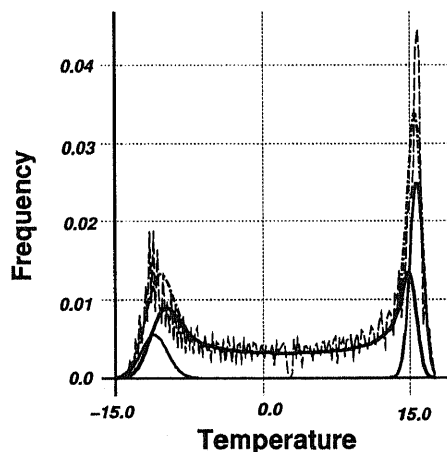
このように、ミクセル密度を用いた混合密度推定によって、裾の重いヒストグラムに対しても混合密度推定で有効な解が得られる。この結果がミクセル密度という概念の有用性を示していると結論できる。



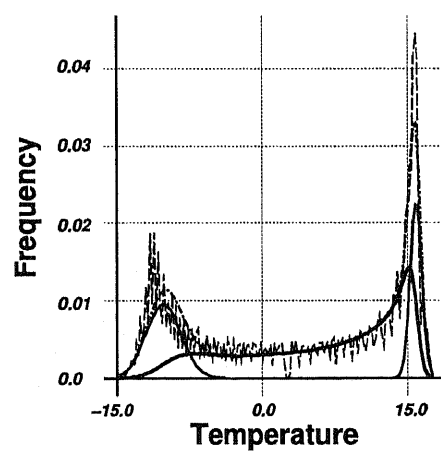
(a) ピュアピクセル密度のみの場合



(b) 一様面積占有率密度のミクセル密度の場合



(c) U型面積占有率密度のミクセル密度の場合



(d) J型面積占有率密度のミクセル密度の場合

図 5.5: 画像ヒストグラムに対する混合密度推定結果。

期待面積占有率および画像分類の結果

図 5.5 の混合密度推定の結果を用いて計算した期待面積占有率の比較を図 5.6 に示す。ここで縦軸の面積占有率は低温側の雲クラス的面積占有率を示している。また期待面積占有率の計算方法については第 4 章で述べている。その結果は、ミクセル密度を用いた場合の (b) ~ (d) では期待面積占有率曲線はほぼ同じ曲線となっており、 0° 付近ではほぼ線形に変化し両端付近で 0 または 1 に飽和する曲線となっている。一方ミクセル密度を用いない (a) の場合のパラメータを用いて計算した期待面積占有率は、雲クラスの密度関数の分散が大きく推定されたことが影響して、かなり非線形で不自然な曲線となった。やはりミクセル密度を用いない混合密度推定では適切な解が得られないことがわかる。なおこの結果から画像を 2 値化するためのしきい値を選ぶならば、面積占有率 50% を一つの合理的なしきい値として選択できる。

次に図 5.4 では、(a) の原画像をベイズの決定則により画素単位で分類した結果を示す。ベイズの決定則に用いたのは混合密度推定図 5.5(d) の結果である。まず (b) にはピュアピクセルに分類された画素を示す (白画素が雲、灰色画素が海)。一方 (c) にはミクセルに分類された画素を示す。ミクセルには図 5.5(d) の期待面積占有率曲線を適用し、濃淡がクラス 1 の面積占有率に比例するように表示した。灰色領域・黒領域はそれぞれ雲・海のピュアピクセルに分類された領域を

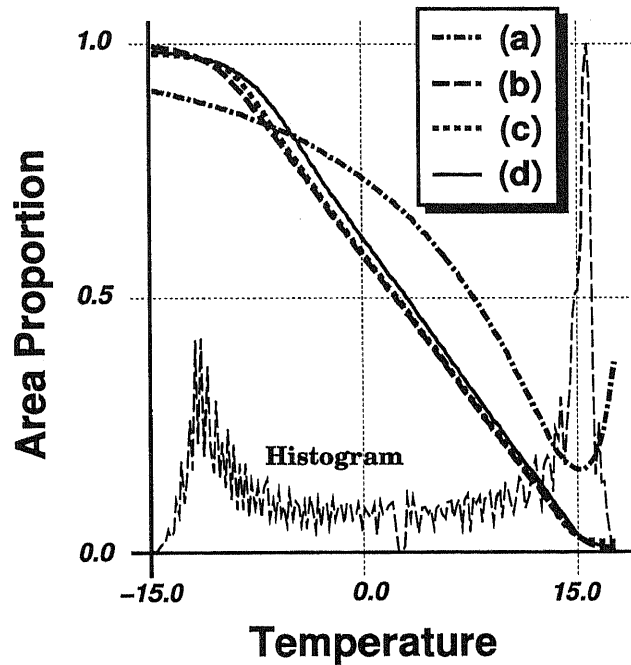


図 5.6: ミクセルの期待面積占有率。(a)~(d) は図 5.5 で用いた記号と同一である。

表 5.2: クラス 1 が占める面積比の推定。

	(a)	(b)	(c)	(d)	Otsu	Kittler
NOAA	0.605	0.422	0.422	0.444	0.422	0.620
Sim	0.480	0.498	—	—	0.498	0.623

示す。この図が示すようにピュアピクセル領域の境界付近にはミクセルが分布している。そしてその期待面積占有率は雲から海へと滑らかに変化している。また特に雲の場合には、雲の厚さが薄い場合は下の海面が透けてみえることがあり、海と雲の性質が混合したものという観点からはこれもミクセルの一種である。ゆえに、ミクセルはいわゆるエッジに沿って並ぶわけではなく、平面的に広がって分布する場合もある。図 5.4(c) でミクセルが平面的に広がっているのはこのような気象条件を反映したものであり、この結果はミクセルがエッジ画素という概念では必ずしも捉えきれないということを示唆している。

5.6.3 各クラスの面積比の計算

このように本研究では、それぞれの画素をピュアピクセルかミクセルに分類し、さらにミクセルに関しては面積占有率を算出するという 2 段階の方法で各画素を分類した。さらに画素ごとの面積占有率を画像全体で合計することで、あるクラスが画像中で占める面積を計算できる。例えば画像全体の雲量を求める問題など、この種の処理は実用的には頻繁に用いられるものである。そこで図 5.4(a) の画像の雲量を、各手法の混合密度推定の結果を利用して計算し、その結果を表 5.2(NOAA) にまとめた。ピュアピクセル密度のみの場合 (a) は 60% となるが、ミクセル密度を用いた場合にはいずれも 40% 強の割合となった。このように推定方法によっては 20% もの誤差が生じるのは、やはりミクセルが画像中に存在することが原因である。

そこで代表的な 2 値化手法 (大津の方法 [48]、Kittler の方法 [80]) を用いた場合の比較をお

こなる。これらの手法によって図 5.4(a) の原画像を 2 値化し、2 値化画像で雲領域の面積を測定した結果を表 5.2 にまとめた。この結果では、Kittler の方法が割合を多めに推定するものの、大津の方法はかなり混合密度推定の結果と近い値を推定している。また図 5.3 で用いたヒストグラムで値が小さい側をクラス 1 とし、そのクラスの割合をヒストグラムから推定する別の実験においても、表 5.2 (Sim) のように大津の方法はほぼ正解 (50%) に一致する推定値を計算している。

このように大津の方法による推定値が混合密度推定による推定値にほぼ一致する理由は以下のように考えられる。すなわち、大津の方法ではしきい値が 2 個のクラスの平均の中間に選ばれるという性質がある [48]。この性質を考えると、この実験で取り上げたように 2 クラスの確率密度が正規密度で十分分離しており、なおかつミクセル密度がほぼ対称形という条件が成立するならば、大津の方法がかなりよい推定値を与えることが予測できる。表 5.2 でも、(b) と (c) がほぼ対称形のミクセル密度であるために大津の方法とほぼ同じ推定値となっているのに対し、左右対称形でないミクセル密度を用いた (d) の場合は、両者の推定値は異なった値である。2 値化手法それ自体はすべての画素をピュアピクセルに分類する手法であるため、そのままミクセルの分類に使うことはできない。しかし以上の実験は、ある条件が成立している場合にはこの種の方法も何らかの推定値を与える手法として活用できるという可能性を示唆している。

5.7 まとめ

本研究では、ミクセル集合が形成する新たな確率密度としてミクセル密度を提案し、特に 2 クラスミクセル密度の形状が非常に特異な形状であることを示した。しかしこの特異な形状は実際の画像ヒストグラムに対応するものであり、実際の衛星画像ヒストグラムに対する混合密度推定でも、ミクセル密度の概念が有効であることを、実験を通して示した。また混合密度推定の結果にベイズの決定則を適用して各画素をピュアピクセルまたはミクセルに分類する方法を示し、また個々のミクセルに対して期待面積占有率を計算した。この結果から各クラスが画像中で占める面積の割合を推定し、従来の 2 値化手法の推定値と比較した。

今後の課題は、第一にはより簡便な近似手法の開発である。混合密度推定は計算量の多い手法であり、高速性が要求される分野への応用には適さない。ミクセル密度の影響を考慮したうえで、判別分析に基づくしきい値選定法 [48] のような高速な近似解法が実用上は求められる。次に密度関数の数を自動的に推定する手法が、多くの画像を自動的に処理するために必要な技術である。この問題には情報量規準などを応用することを検討中である。最後にミクセルの分類手法に共通する問題として、その結果の適切な評価方法が必要である。正答率などの指標はピュアピクセルの概念にかなり偏った評価方法であり、ミクセルに対しても適用できる評価基準の開発が今後の研究課題である。

第6章

領域に基づいた変形モデルを用いた形状分解

6.1 概要

本章は「領域に基づいた変形モデルを用いた形状分解」として、階層モデルの領域レイヤに相当する処理をおこなう。ここでは画素レイヤで画像を画素単位分類した結果を受けて、解析の単位を画素単位から同一の分類クラスに属する画素領域単位へと拡大する。そして領域単位で特徴量を抽出し、計測された特徴を上位レイヤに渡すことがこのレイヤの目的となる。

6.2 形状情報の表現

6.2.1 領域特徴量による形状情報の表現

ここでは領域の形状をどのように計測するかを考える。まず一つの最も簡単な候補は、従来から提案されている形状特徴量やモーメント特徴量である。例えば以下のような特徴量がある[81]。まず面積 A は連結成分内の画素数で表される。周囲長 L は輪郭点数で表されるが、同じ図形に対しても考えている連結性によって値は異なる。面積と周囲長から計算される特徴量には、円形度 ($4\pi A/L^2$) がある。円形度は真円において1となり(デジタル的に計算した面積と周囲長であるため厳密には1とならない)、複雑な形状になるほど値は小さくなる¹。また画像の周上の任意の2点間の距離のうち最大の長さを L_M 、また最大長に平行な2本の直線で図形をはさんだ時の2直線間の最短距離を L_W とすると、縦横比 L_W/L_M が計算できる。

さらに別の重要な特徴量として次式の $(p+q)$ 次の重心回りのセントラルモーメント (central moment) が使われる。ただしデジタル図形を $f(i, j)$ 、その重心座標を (i_g, j_g) とする。

$$M_{pq} = \sum_i \sum_j (i - i_g)^p (j - j_g)^q f(i, j) \quad (6.1)$$

特に細長い図形に対し、図形が伸びている方向(主軸の方向)と i 軸のなす角度 θ は次式で計算でき、方向性に関する特徴量として非常に有効である。

$$\theta = \frac{1}{2} \arctan \left(\frac{2M_{11}}{M_{20} - M_{02}} \right) \quad (6.2)$$

¹なおこの特徴量は名称こそ円形度となっているものの、基本的には領域がどのくらい円に近いかを表す特徴量とするよりは、領域の輪郭の複雑度を表現した特徴量であると考えたほうがよい[82]。

これらの方法は計算が簡単であるため、本研究でも特徴量としてこれらを用いる。しかしこのような形状特徴量だけでは画像検索の目的には不十分である。それは無数の異なった図形が同一の形状特徴量を示すために、画像をうまく区別することができるような検索インデックスとしては不十分だからである。

6.2.2 領域や輪郭線による形状構造の表現

そこで領域の形状の特徴量を計測するだけでなく、その領域形状の構造を計測するための手法がいくつか提案されている。中でも代表的な手法は、領域や輪郭線に基づいた以下のような方法である。

1. 領域に基づいた方法: 骨格線・形状分解・4分木
2. 輪郭に基づいた方法: 多角形近似・チェーンコード・フーリエ記述子・B-スプライン

この中で、本研究が扱う衛星画像の雲パターンの表現に適した方法はどのようなものだろうか。雲パターンの輪郭の形状が非常に複雑なフラクタル的構造であることを考えると、例えば多角形近似などの手法は輪郭の細かい変動に影響されると予想される。そこで本研究では前者の方法の一つである、「形状分解」の手法に着目した。

形状分解とは領域を基本構成要素 (primitive) の集合体として表現するもので、複雑な形状をもつ領域をできるだけ少数の単純な形状の構成要素によって近似しようという方法である。具体的な手法としては、モルフォロジーを用いる方法 [83, 84, 85] や膨張法を用いる方法 [86]、さらに情報量規準を用いて構成要素の分割結果を統合する方法 [87] など提案されている。またこれらの研究とは少し方向性は異なるが、第7章に述べるゲシュタルトの法則に基づいて形状分解を行う手法も提案されている [88, 89]。また少し古いサーベイには [90, 91] がある。

近年のパターン認識の研究では部分 (part) に基づく認識が一つの大きな流れとなっている [92, 93]。物体を眺める際に人間はそれをいくつかの部分に分けて理解する傾向があるため、複雑な形状の領域を部分の集まりとして記述する方法は、人間の思考特性によく適合する。また構成要素として単純な図形を用いたとしても、それらの変形や構成要素どうしの組合せ効果によって、非常に多様な形状の表現が可能である [94]。本研究が対象とする雲パターンの形状も、領域特徴量を単に計測するよりは構成要素の集合として考えた方が、雲領域の形状を表す領域情報としてより有用であると考えられる。

さて本研究で扱う画像は二値画像および濃淡画像である。濃淡画像を形状分解する手法としては、まず濃淡画像を2値化してシルエット画像を作成し、シルエット画像に対して形状分解を適用する方法と、濃淡画像に直接形状分解を適用する方法とが考えられる。しかし2値化という中間処理によって有用な情報を喪失することは好ましくないことから、本研究では濃淡画像にも直接適用できるような形状分解手法を提案する。

6.3 領域に基づいた変形モデルを用いた形状分解

6.3.1 変形モデル

本研究の形状分解手法は、一連の変形モデル (deformable model) を用いた研究に触発されたものである。これは画像からの形状抽出を、力学系でのエネルギー最小化問題のアナロジーと

して定式化するアプローチである。変形可能な形状モデルの振る舞い(変形)を、モデル自体の傾向と外からの制約の線形結合によるエネルギーとして表現し、エネルギー極小の安定状態を見つけることによって画像から対象物を抽出する [95]。このような変形モデルの代表的な手法にはスネーク (snake) がある [96]。また変形モデルを用いた手法には他にもバルーンモデル [97] や Active Net [98] がある。ただしこれらは形状抽出を考えたものであり形状分解のための手法ではない。そこで本研究では、これらの変形モデルを単純化したうえで形状分解に適したエネルギー関数を定め形状分解に応用する。輪郭だけに注目するのではなく領域の内部に注目するという意味では、Active Net に近い面がある。

6.3.2 形状分解に用いる基本構成要素

エネルギー関数の説明を始める前に、まず形状分解に用いる基本構成要素を定める。これはパラメトリックな曲線であれば何でもよく、例えば正方形や三角形でもよい。この基本構成要素を $P(\phi)$ と表記する。ただし ϕ は基本構成要素のパラメータベクトルである。例えば P として楕円を用いると、パラメータベクトルは $\phi = (s_x, s_y, \theta, a_1, a_2)$ と表記することができる。ここで (s_x, s_y) は楕円の中心位置であり、 θ は主軸の傾き、 a_1 と a_2 は座標軸方向の拡大率である。さらに以下の方程式で表される超楕円も基本構成要素として考えることができる。

$$\left(\frac{|x - s_x|}{a_1}\right)^{\frac{2}{\varepsilon}} + \left(\frac{|y - s_y|}{a_2}\right)^{\frac{2}{\varepsilon}} = 1 \quad (6.3)$$

この場合はパラメータベクトルは $\phi = (s_x, s_y, \theta, a_1, a_2, \varepsilon)$ となる。ただし ε は形状を制御する変形パラメータであり、 $\varepsilon = 0$ のとき長方形、 $\varepsilon = 1$ のとき楕円、そして $\varepsilon = 2$ のときは菱形となる。このように一つの方程式で多様な形を扱える点が超楕円の特長である。また式 (6.3) には主軸の傾きは入っていない。

このようにあらゆるパラメトリックな図形の中で、どのような図形が最も雲パターンの表現に適しているだろうか。それは、どのような図形が最適かを判断する規準がない以上は決められないことでもあるが、ここでパターン認識の一般的な法則「ケチの法則」 [41] やゲシュタルトの法則などを考えると、基本構成要素としては単純な図形がよい。それに加えて本研究が対象とする雲パターンの場合には、これが自然現象であることから三角形や四角形などの角が目立つ図形は好ましくない。よって本研究では、楕円と超楕円を構成要素として用いることにする。

6.3.3 シルエット画像を対象としたエネルギー関数

エネルギー関数

本研究で用いる変形モデルは、変形を繰り返しながらエネルギー関数の最大値を探索する動作を続ける。パラメータベクトルの変化によって変形モデルの形状が変化するため、エネルギー関数の最大値を探索する問題は、多変数関数の最適化問題に帰着する。つまり求めたい形状分解の構成要素のパラメータベクトルは

$$\phi_{\max} = \operatorname{argmax}_{\phi \in \Psi} E(\phi) \quad (6.4)$$

によって求められる。ここで Ψ は ϕ のパラメータ空間である。

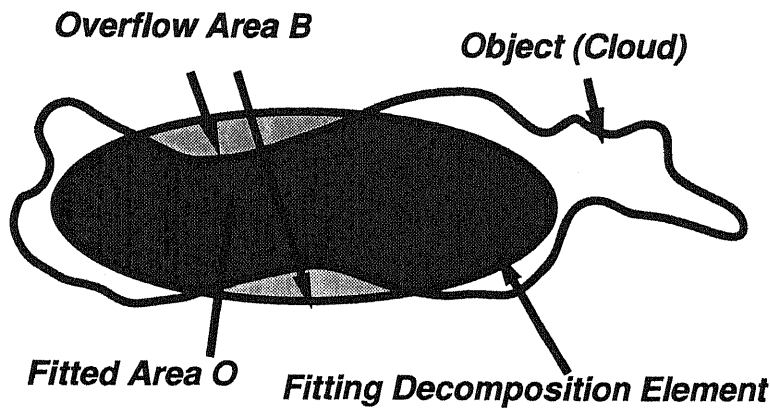


図 6.1: 変形モデルのあてはめ。

そこで次に、変形モデルの本質であるエネルギー関数を具体的に定める。まず二値画像の場合に本研究で用いるエネルギー関数は、領域の面積に基づいた以下のような単純な関数である。

$$E = O - pB \quad (6.5)$$

ここで図 6.1 に示すように、 O は変形モデルが適合している部分の面積、そして B ははみだした部分の面積、 p はペナルティ定数 $p > 0$ である。 O が膨張力として働く一方 B は収縮力として働くため、 O と B がちょうど釣り合う状態が存在する。この状態がエネルギーの最大値であり、それに対応するパラメータが最適な形状分解の構成要素の形状を決定する。

ペナルティ定数

ここではみだしを制御するペナルティ定数 p は、この値を大きくするほど求まる構成要素は領域に適合したものとなる。これは一見大きい値に設定した方がよいようにも思えるが、しかしある程度のはみ出しを許容した方が結果的には直観に合う結果が得られる。またはみ出しの種類をさらに 3 種類に区分することで、より細かく制御できる形状分解も可能である。すなわち、1) 背景画素にかかった場合、2) 画像枠からはみ出した場合、3) すでに形状分解された画素にかかった場合、という 3 種類のはみ出しを定義することで、形状分解構成要素どうしの重なりあいが少ない形状分解なども可能である。

6.3.4 濃淡画像を対象としたエネルギー関数

エネルギー関数

次に濃淡画像の場合のエネルギー関数を定める。濃淡画像の場合は二値画像の場合のように「(関心)領域」と「背景」とを明確に区別せずに形状分解をおこなう必要がある。そこである画素値の画素が領域に属する適合度を示す関数を定義し、これをメンバーシップ関数 μ とよぶ。ただしこのメンバーシップ関数は通常ファジイ理論で用いられるメンバーシップ関数と同じものではない。すなわち値域を $(-\infty, 1]$ とし、画素値 I_x に対して $\mu(I_x)$ が大きいほど領域に属する可能性が高いものとする関数である。するとこのエネルギー関数の場合の変形モデルは、メンバーシップ関数が正の部分をしてできるだけ多く含み、同時に負の値が大きい部分をできるだけ含まないような形状を目指して変形する。具体的に記述すると変形モデルのエネルギー E は、変形モデルの内部

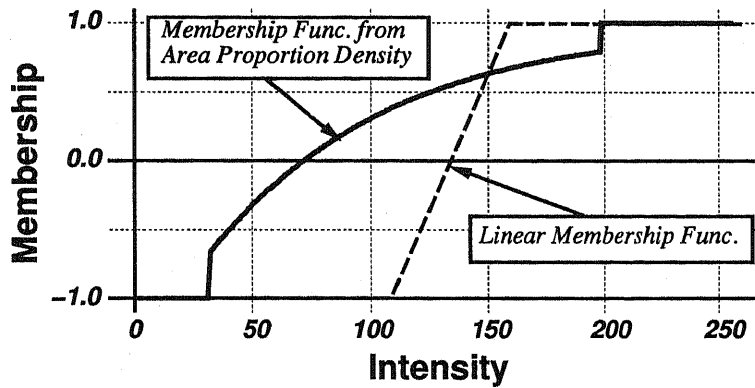


図 6.2: 本研究で用いるメンバーシップ関数の模式図。

O に含まれる画素を $x \in O$ とすると以下のように書くことができる。

$$E = \sum_{x \in O} \mu(I_x) \quad (6.6)$$

すると式 (6.5) とは、式 (6.6) で $\mu(0) = -p$ 、 $\mu(i) = 1 (i \neq 0)$ というメンバーシップ関数を用いた特別な場合に相当することがわかる。ただし式 (6.5) において背景画素の画素値を 0 とした場合である。

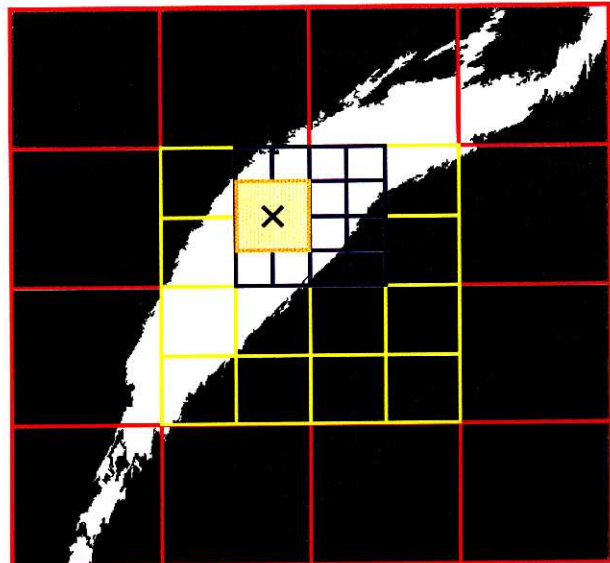
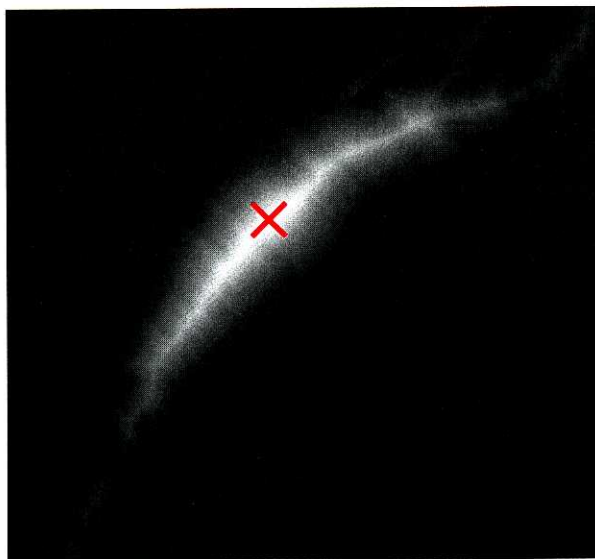
メンバーシップ関数

本研究ではメンバーシップ関数として図 6.2 の 2 種類の関数を用いる。まず単純な関数として線形なメンバーシップ関数を用いる。これは何らかの方法によって画像を 2 値化するしきい値 t を求め、以下の関数をメンバーシップ関数とする方法である。

$$\mu(I) = \begin{cases} -1 & (I < t - w) \\ \frac{I-t}{w} & (t - w < I < t + w) \\ 1 & (I > t + w) \end{cases} \quad (6.7)$$

ただし w はメンバーシップ関数の傾きを決める値であり、領域と背景の分離の度合を考慮して決定する定数である。なおメンバーシップ関数の負の値をあまり大きくすると構成要素が過剰適合するため、メンバーシップ関数の値は最小でも -1 にしている。この方法は直観的で簡便な方法ではあるが、メンバーシップ関数の決定方法としては根拠が薄い方法である。

そこでより根拠のある決定法として、第 4 章で導いた期待面積占有率を用いたメンバーシップ関数を提案する。第 5 章で述べたように期待面積占有率はミクセルに対して定義されるものであり、2 クラスミクセルで片方のクラスの面積が画素内に占める面積を表す値である。そこでメンバーシップ関数にこの面積占有率を適用する。まず面積占有率が 50% となる画素値でメンバーシップ関数が 0 となるようにする。さらに面積占有率が 100% のピュアピクセルで $\mu(x) = 1$ 、面積占有率が 0% のピュアピクセルで $\mu(x) = -1$ となるように面積占有率を線形にスケールアップする。この計算により図 6.2 のようなメンバーシップ曲線が得られる。このメンバーシップ関数は面積占有率という値に基づいて定義した関数であり、根拠が明確な定義であるといえる。



(a) 距離変換を用いる方法：中心位置 (228, 185) (b) 9 分木を用いる方法：中心位置 (208, 163)

図 6.3: 中心位置の初期値推定法。

6.4 実験

6.4.1 アルゴリズム

以下にアルゴリズムの説明を簡単に述べる。

1. メンバーシップ関数の値が正の画素値をもつ画素を抽出し、ラベリングをおこなう。
2. 変形モデルのパラメータベクトルの初期値を設定する。
3. 最適化手法を用いてパラメータベクトルを最適化する。
4. 最適化された変形モデルを取り除き、残りの部分を新たにラベリングする。
5. 最大の面積をもつ領域が、最小許容面積のしきい値より小さければ終了。そうでなければ 2 へ。

ここで最適化手法としては、勾配法的一种である準ニュートン法 (quasi Newton method) を用いた。ここで勾配法を用いる際には初期値推定が重要な鍵を握る。そのため初期値推定について簡単に述べる。

6.4.2 初期値推定

中心位置の推定

パラメータベクトルの初期値推定で最も重要なのは中心位置 (s_x, s_y) の推定である。この中心位置を推定するために、まずメンバーシップ関数の値が正の画素値をもつ画素を抽出し、抽出画素を領域とする 2 値画像を作る。この 2 値画像から初期値として適切な値を得ることを考える。

最も簡単に思い付くのは、領域の重心位置を初期値とする方法である。しかしこの方法は、例えば凹の形をした領域では重心位置が領域の外側にくることがあるために好ましくない。

そこで次に思い付く方法は、領域を距離変換 [81, 99] した距離変換画像の最大値を初期値とする方法である²。この方法ならば、初期値は常に領域の内側であり、なおかつ周囲には領域に属する画素が集中しているところが中心位置として得られる。したがってある程度は滑らかな形状をもつ領域では、この方法が最もよい初期値を与える。なお領域内に含まれる小さな穴は距離変換に大きな影響を及ぼすため、図形の膨張 / 収縮 [81] やモルフォロジーの Opening/Closing [100] などの手法を用いた穴の除去が前処理として望ましい場合がある。この方法を図 6.3(a) に示す。

しかし距離変換にも問題がある。第一の問題点は計算に時間がかかり、初期値推定のわりには計算コストが高い点である。また第二の問題点は、非常に入り組んだ形状の領域では、距離変換といえども適切な初期値を与えないことがあることである。その意味で距離変換はロバストな手法ではない。そこでより簡便でしかもよい初期値を与える方法として、ここでは9分木を用いた方法によって初期値推定をおこなう。この方法では画像を4分割してできる長方形ブロック9個を考える。それぞれの長方形ブロックに含まれる領域の画素数を数え、その中で最も多い画素を含む長方形ブロックを再び4分割して同様の処理を繰り返す。そして長方形ブロックが十分に小さくなって全体が領域に含まれるようになれば、その中心が中心位置の推定値となる。この方法で中心位置を求めた例を図 6.3(b) に示す。この方法は計算が高速であり、しかも領域形状の複雑さとは無関係となる、ロバストな方法である。

今後の実験では両者の長所を比較したうえで、NOAA 衛星画像の場合は距離変換法を、また GMS 衛星画像の場合は9分木法を用いた。

その他のパラメータ

その他のパラメータは、中心位置の初期値に基づいて推定する。中心位置から輪郭までの距離を角度を変化させながら計測し、その距離情報をもとに主軸方向や大きさの初期値を定める。この方法の詳細については省略する。

6.4.3 実験結果

本研究で提案した手法を、シルエット画像と濃淡画像の両方に適用した。シルエット画像としては第 2.2.2 節で述べた NOAA 衛星画像を、また濃淡画像としては第 2.2.3 節で述べた GMS 衛星画像を用いた。

シルエット画像に対する形状分解

シルエット画像の場合はペナルティ定数を調節することによって、形状分解の結果を制御することができる。そこでペナルティ定数を変化させた場合の形状分解の結果を図 6.4 に示す。構成要素としては超楕円を用いている。 p が小さいと領域に過剰適合し、逆に p が小さいと構成要素が広がりすぎている。そのトレードオフを考慮し、最終的には $p = 1$ を最適な値と決定した。このように超楕円を用いることによって、雲パターンの多様な形状を構成要素の集合体として表すことができる。

²厳密には最大値が複数出現する場合があります。この場合は最大値の中から最も適当な最大値を選択する必要があります。これは最大値となる画素の中で連結画素数が最も大きい画素を選ぶことになる。

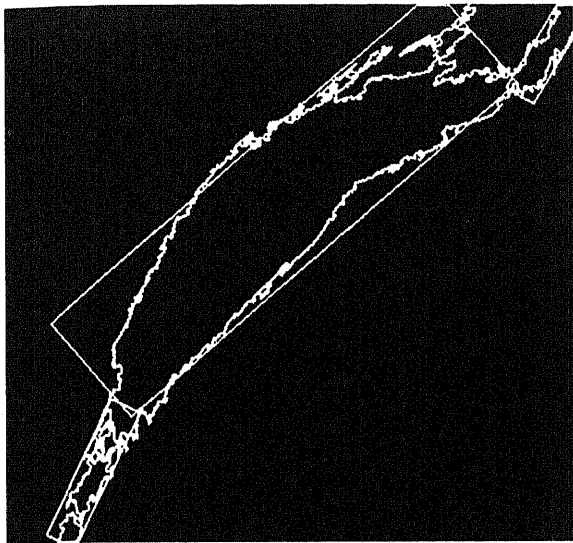
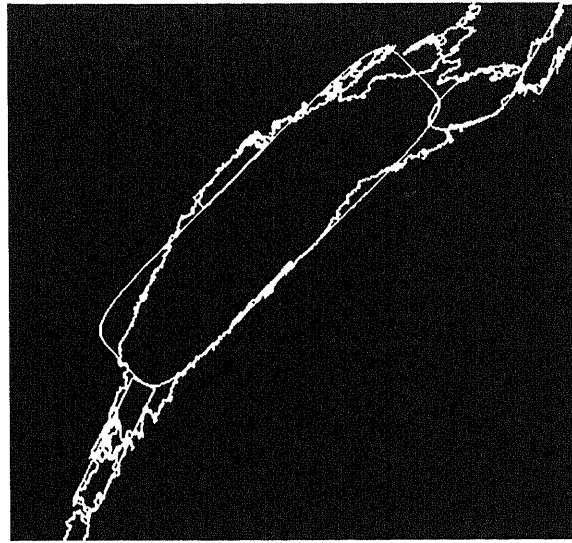
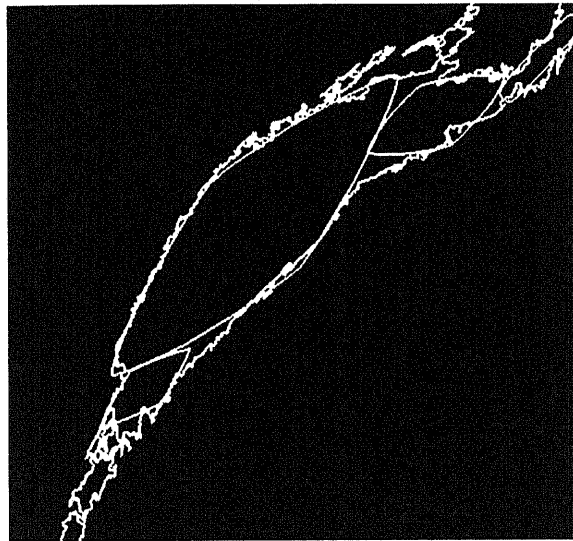
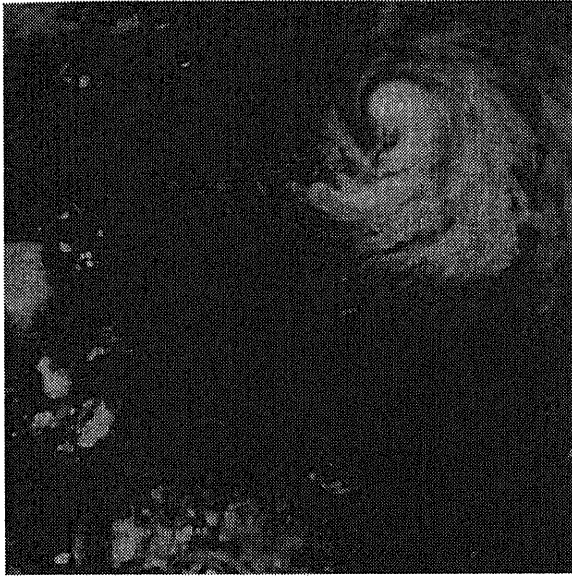
 $p = 0.3$  $p = 1$  $p = 3$

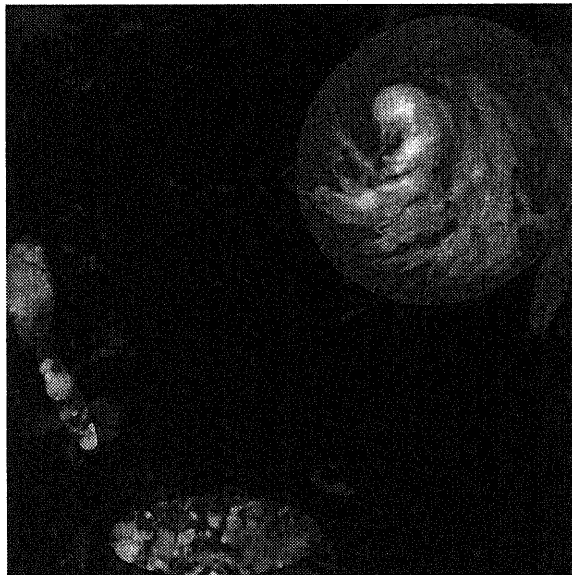
図 6.4: シルエット画像の形状分解におけるペナルティ定数 p の効果。後の実験では $p = 1$ とした。

濃淡画像に対する形状分解

濃淡画像の場合は、メンバーシップ関数に図 6.2 の面積占有率から導出した関数を用いた場合の形状分解結果を図 6.5 に示す。ここでは楕円を用いた場合と超楕円を用いた場合とを比較する。両者を比較すると、楕円を用いた場合の方が、滑らかな形状の構成要素が雲パターンを表現しているという点で、より直観に合う分解結果となっている。逆に超楕円の分解結果が直観と合わない原因は、超楕円の形状パラメータ ε の推定が難しいことが原因である。最適化法が勾配法であるために、どうしても ε の値は定義域 ($0 < \varepsilon < 2$) の端の局所最大値につかまる傾向があるからである。



原画像



楕円による形状分解



超楕円による形状分解

図 6.5: 濃淡画像の形状分解。楕円の場合と超楕円の場合の比較。

6.5 まとめ

以上のように領域レイヤでは、領域に基づいた変形モデルを用いて領域を基本図形の集合として表現する方法を提案した。形状分解の問題を最適化原理に基づいた定式化によりエネルギー関数の最大化問題に帰着させることで、同一の手法で任意のパラメトリック曲線を基本構成要素として扱うことができる。この定式化ではエネルギー関数の定義が重要である。そのためまずはシルエット画像に対する簡単なエネルギー関数を定義し、次にこれを濃淡画像の場合に自然に拡張した。エネルギー関数の計算にメンバーシップ関数という考え方を導入し、線形なメンバーシップ関数と、第4章で述べた面積占有率に基づいた根拠のあるメンバーシップ関数とを用いて形状分解をおこなった。その結果、超楕円は確かにさまざまな形状を表現できる柔軟なモデルであるが、その形状分解結果は雲パターンに対しては必ずしも直観に合うわけではないことがわかり、今後の実験では主に楕円を構成要素として用いることとした。

今後の課題は、形状分解の順番に依存しない方法の開発である。現在の手法では、当てはめた構成要素を順次取り除くことで形状分解を進める方法としているが、この方法では複雑な形状の場合はどこから形状分解を始めるかで形状分解の結果が変わってくる。つまり結果は形状分解の順番に依存し、なおかつこの部分は雑音に敏感である。そこで前もって大雑把に別の方法で形状分解をしておいて超楕円を当てはめる方法や、分散協調処理 [101] の導入などの方法が考えられる。特に分散強調処理の方法は、形状分解の順番に依存しないロバストな処理の方向性を示すものとして、本研究の今後の方向性を示していると考えられる。

第7章

引力モデルを用いた群化と階層化属性付き関係グラフの構築

7.1 概要

本章では「引力モデルを用いた群化と階層化属性付き関係グラフの構築」について述べる。これは表 2.2の階層モデルでは関係レイヤに相当する。下位レイヤで得られた画像特徴を1つのモデルに統合し、画像全体を1つのモデルで記述することが目的である。このとき、下位レイヤで得られたプリミティブ間の関係をどのようにモデル化し計算するかが問題である。

さて「プリミティブ間の関係」と書いたが、その関係として最もよく用いられる概念は「隣接関係」である。2領域が隣接しているとは、2領域からそれぞれ任意の画素を取り出したときに、その組み合わせの中で一つでも4連結または8連結の意味で連結しているものがある場合をいう。このような隣接関係の構築には定式化する上での問題はない。

しかしこのような隣接関係によって関係を定義できない場合がある。例えば本研究が扱う雲パターンを考える。画像中には複数の雲領域が存在しているが、これらは隣接関係という意味では隣接しておらず、海領域に隔てられて散らばっている。しかし人間の視覚情報処理では、厳密に隣接関係ではなくても何らかの関係を知覚することができる。ここで問題なのは、そのような関係が画像上の画素値に反映しているとは限らない点である(例えば主観的輪郭など)。むしろこのような知覚は人間の高度なパターン認識能力によるものであり、したがって情報処理課題として扱うには本質的に困難な問題であるといえる。ゆえに画像上の画素値を単純に解析するだけでなく、プリミティブの関係を計算できる何らかの視覚情報処理モデルや物理的モデルが必要となるのである。

7.2 知覚的群化

このようにプリミティブ間の関係を計算する方法に群化(grouping)の計算理論がある。群化のモデルとしては以下のようなモデルが提案されている。

1. 物理法則のアナロジーとしての仮想的な力や場を用いる方法 [102]
2. ガウス関数などの点広がり関数を適用する方法 [103]
3. ゲシュタルト法則を土台としたルールに基づく方法 [104, 105]
4. 周波数領域上でテクスチャ的な群化を分析する方法 [106]

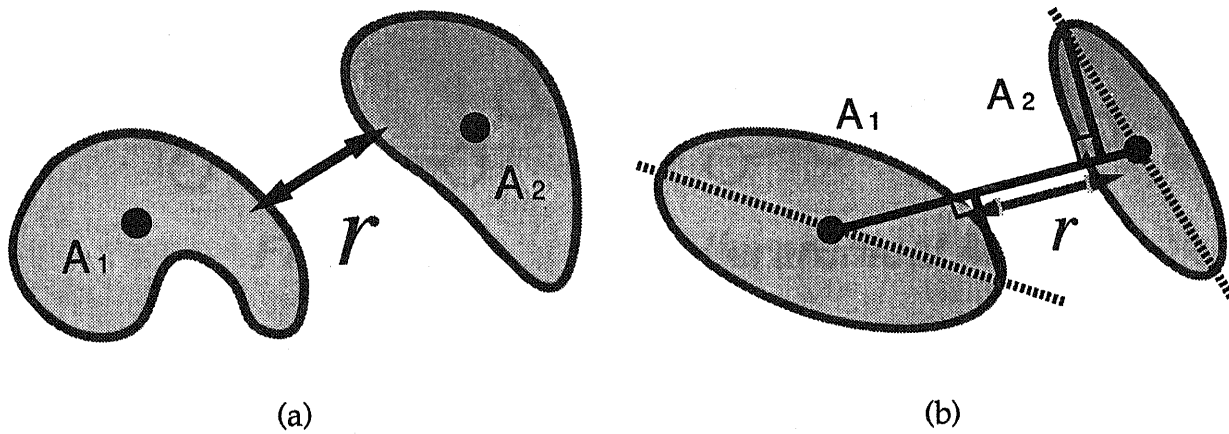


図7.1: 二つの仮想引力モデル。(a) r は輪郭間の最小距離、(b) r は重心間の距離から図形の細長さを考慮した長さを引いた距離。

5. ネットワークを用いる方法 [107]

6. 特徴量の類似性を用いる方法 [108]

この他に人間の視覚生理特性をモデルにした神経回路により群化を研究する流れもあるが、多くの研究は人間の視覚情報処理の結果を観察することで帰納的にルールを発見するアプローチが多い。

そしてほぼすべてのモデルに影響を及ぼしている概念が「ゲシュタルト」と呼ばれるものである。ゲシュタルト法則とは人間の視覚に関する定性的な性質をまとめた法則であり、その本質はPrägnanzと呼ばれるものである [109, 110]。Prägnanzはゲシュタルト心理学者によって、近接の法則やなめらかなカーブの法則・単純性の法則・等質性の法則・閉鎖性の法則などの法則にまとめられた。これらすべての法則は定性的に述べられており、したがってゲシュタルト法則を計算理論として定式化するのは容易ではない。

7.3 仮想引力モデル

7.3.1 仮想引力モデルの定式化

そこで本研究では、ゲシュタルトの法則を忠実に再現するアプローチはとらずに、より簡便な方法である、物理法則のアナロジーに基づいた方法を用いる。基本的な考え方は、二つのプリミティブまたは領域間の関係を考えたとき、両者の面積が大きく両者の間の距離が小さいほど関係が強いと仮定する考え方である。このような関係を表現する方程式は、領域の面積を仮想的な質量とみなせば万有引力の方程式と同様の関係となる。そこで万有引力からの類推で、仮想的な引力 f を以下のように定義する。

$$f = \frac{A_1 A_2}{r^2} \quad (7.1)$$

ここで2つのプリミティブまたは領域の面積がそれぞれ A_1 、 A_2 である。またここでの距離 r は「仮想的な距離」であり、図7.1のように計算される2種類の距離を使い分ける。物理的なイメージに縛られると距離 r に重心間の距離を用いるべきと考えられるが、しかし非常に細長い形や複雑な形をした領域では、重心間の距離よりはむしろ輪郭間の最小距離の方が直観に合った引力を

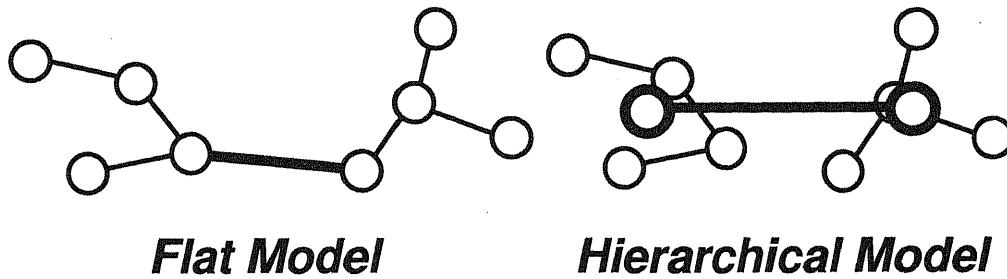


図 7.2: 非連結グラフを接続する二つの方法。

与える。これが図 7.1(a) の場合である。ただし最小距離を発見する計算は、重心間の距離の計算に比べてはるかに計算量が多い。そこで楕円プリミティブのように主軸方向が明確な場合には、重心間の距離から図形の細長さを考慮した長さを引いた量を r とすることによって計算量を少なくする。これが図 7.1(b) の場合である。

このように引力モデルは、直接には隣接していないプリミティブ間に働く、ある種の遠隔作用を直観的にモデル化したものである。このような遠隔作用によって、プリミティブ間がある程度離れていても、その間には何らかの相互作用が隠されていると仮定するのである。こうして離れたプリミティブ間の関係を構築できることになる。

7.3.2 仮想引力モデルによる関係のネットワーク

仮想引力モデル式 (7.1) の計算によって、ある一つの領域 n_i に対する他のすべての領域 $n_j (j \neq i)$ との仮想引力 $G(n_i, n_j)$ を計算することができる。原理的にはすべての 2 領域間に関係が計算できるが、それではモデルが複雑になりすぎる。よってここでは近似として、最大の仮想引力が働く 2 領域間の引力を $G(n_k, n_l) = G_{\max}$ とし、この 2 領域間 (n_k, n_l) に「関係」が存在するとした。それに加えて、 G_{\max} と同程度の仮想引力が働く 2 領域間、すなわち $G(n_p, n_q) > kG_{\max}$ を満たす 2 領域間 (n_p, n_q) に対しても、「関係」が存在するとした。ここで本研究では $k = 0.8$ の値を用いている。

この計算を各領域に対して順番におこなうことにより、各領域間に関係が生成される。しかしこのようにボトムアップに関係を生成していく結果、できあがる関係のネットワークは以下の二つの位相構造をもつ可能性がある。

1. すべての領域 (プリミティブ) が連結グラフとなる。
2. 領域 (プリミティブ) は複数の連結グラフに分割され、全体としては非連結のグラフとなる。

1 の場合には画像が 1 つのグラフとして表現されているため、このままで関係を表現するネットワークは完成している。しかし 2 の場合はこの非連結グラフを一つの連結グラフに結び付ける後処理が必要となる。この後処理には以下の 2 通りの方法を用いる。

平面的引力モデル

ある一つの連結グラフを中心にして、その他の連結グラフに含まれる任意のノードとの間で仮想引力を計算し、その値が最大のノードの組を発見する方法である。このノードの組に関係を張ることによって、2 つの連結グラフ 1 つの連結グラフに統合される。またさきほどと同様に、

同程度の引力をもつノードの組があればそこにも関係を張る。こうしてすべてのグラフが連結となるまでこの処理を繰り返す。このモデルでは新たに張られた関係は他の関係と同等の立場となる。

このモデルを具体的に用いる方法は以下の通りである。まず画像を2値化し、あるしきい値面積以下の小さな領域を併合する。その後2値画像をラベリングして各ラベルに対してそれぞれ形状分解をおこなう。ここで得られたプリミティブは、まず各ラベルごとに平面モデルに基づいた仮想引力によって構造化する。そして各ラベルごとに新たにノードを設け、今度はこのノード間で上位層の関係を構築する。よって平面モデルの場合は、プリミティブを2値化とラベリングという処理でトップダウン的にまとめて下位層におき、領域間の関係という上位層の関係を構築する方法である。

階層的引力モデル

そもそも最初に連結グラフで閉じた領域は、何らかの意味のある一つのまとまりであるとみなし、その連結グラフはそのまま保持する方法である。そして連結グラフごとにその重心に新たなノードを作り、各連結グラフの新しい中心ノード間で再び引力モデルに基づいて引力を計算していく。このようにして、新しく生じた関係を上位の関係とし、もとの連結グラフを下位の関係とすれば、階層的なネットワークが形成される。本研究ではこれを2層以内の階層としている。

このモデルを具体的に用いるのは、濃淡画像を形状分解した場合である。近接したプリミティブのまとまりが下位層、そしてプリミティブのまとまり間の関係が上位層となる階層的な引力モデルにより、関係ネットワークはボトムアップ的に形成される。

7.4 階層化属性付き関係グラフ

以上のように、仮想引力モデルを用いてプリミティブを構造化し関係ネットワークを生成する方法を述べたが、このネットワークを最終的に組み込むモデルである階層化属性付き関係グラフについて次に説明する。

本研究では画像全体を一つの画像表現モデルを用いて記述し、その画像表現モデルには明示的な記号表現を用いることとしている。このような目的に適した画像表現モデルとして、ストリング (string)・ツリー (tree)・グラフ (graph) などのモデルがある。この中でもグラフモデルは解釈に要する計算量が大きい点という欠点はあるが、モデリング能力が強力である [111] という点が魅力的である。そこで本研究では強力なモデリング能力を重視して画像表現モデルとしてグラフ構造を用いる。

このグラフ構造は空間情報の表現には当然のことながら強力であるが、それに加えて領域情報の表現も充実させるため、本研究では属性付き関係グラフを画像表現モデルとして用いる。特に本研究では、過去に提案されてきた属性付き関係グラフ [10, 112] を階層化した、階層化属性付き関係グラフ (hierarchical attributed relational graph) を用いる。このモデルでは、ノード (node) とアーク (arc) はそれぞれ記号的に画像を表現し、一方ノードとアークの属性は画像の数値的な情報を記述する。そしてノードが領域の位置を表す一方、アークは2領域の関係という2項関係を表現する。階層構造は2層構造となっており、上位層ではプリミティブのまとまりの間の関係を表現し、下位層では個々のプリミティブのまとまりの内部の関係または領域内部の構造を表現している。

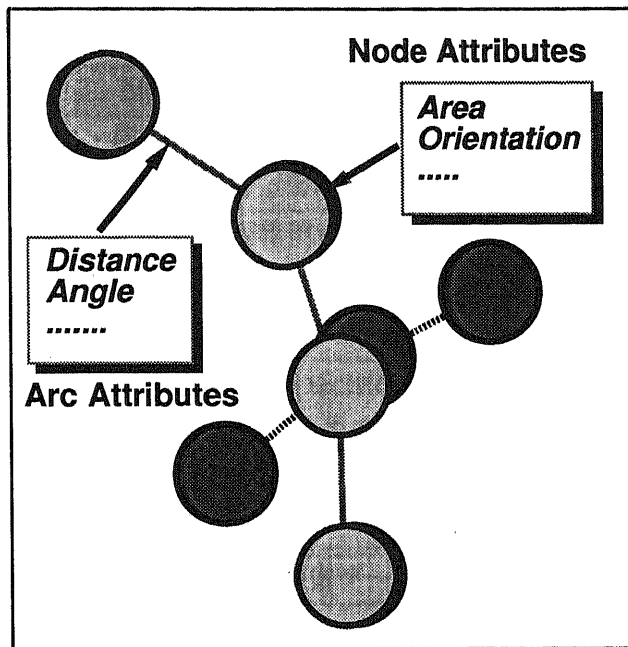


図 7.3: プリミティブを構造化して得られた階層化属性付き関係グラフ。

7.5 グラフ構造への構造化

7.5.1 仮想引力モデルを用いたグラフ構造への構造化

以上に述べてきたように、仮想引力モデルを用いてプリミティブ間の関係を2層構造のネットワークとして表現することができた。このネットワークを階層化属性付き関係グラフに組み込まなければならないが、関係ネットワークの位相構造はそのままグラフ構造に組み込めばよい。あとはノードとアークのそれぞれで属性値などを計算すれば組み込みが完成する。

この仮想引力モデルを用いて構築したグラフ構造の例を図 7.3 に示す。この例は NOAA 衛星画像の場合を示したものであり、上位層のノードは各領域の重心位置、また下位層のノードは形状分解構成要素の重心位置を示している。またアークは仮想引力によって張られた関係を示したものである。また階層化属性付き関係グラフに属性値として組み込まれた属性の一覧を表 7.1 に示す。

7.5.2 隣接関係を用いたグラフ構造への構造化

カラー画像の場合のグラフ構造は、衛星画像の場合と大きく異なる。カラー画像の場合は、領域分割のあとの形状分解を省略するためである。したがってこの場合は、領域それ自体が一つのプリミティブであり、領域間の関係のみが定義される。したがってグラフ構造は1層構造である。またこの場合はすべての領域間に隣接関係が定義できるが、この隣接関係には3通りのラベルを設けた。

1. 周長のうち $S\%$ 以下の長さしか接していない場合は隣接関係を定義しない。
2. 接している長さが $S\%$ 以上 $L\%$ 以下であれば隣接関係を定義する。
3. $L\%$ 以上であれば、特別な隣接関係として「囲まれている」隣接関係を定義する。

表 7.1: 各画像データベースの階層化属性付き関係グラフで用いられている形状特徴量の一覧。

NOAA 衛星画像データベース	
上位層・ノード	主軸方向・縦横比・compactness
上位層・アーク	水平軸に対する傾き
下位層・ノード	主軸方向・縦横比・超楕円の形状パラメータ ϵ
下位層・アーク	水平軸に対する傾き
GMS 衛星画像データベース	
上位層・ノード	主軸方向・縦横比・compactness・円形度
上位層・アーク	水平軸に対する傾き・2領域の主軸方向の角度差・2領域の面積比
下位層・ノード	主軸方向・縦横比・画素値の分散(テクスチャ特徴量)
下位層・アーク	水平軸に対する傾き・2領域の主軸方向の角度差・2領域の面積比
カラー自然画像データベース	
上位層・ノード	主軸方向・縦横比・compactness・円形度・R成分の平均値・G成分の平均値・B成分の平均値
上位層・アーク	水平軸に対する傾き・2領域の主軸方向の角度差・2領域の面積比

このように隣接関係が定義された領域どうしにアークを張れば、1層構造のグラフ構造が構築できる。本研究では $S = 40\%$ 、 $L = 80\%$ としている。

7.6 まとめ

群化の計算にはまだ標準的なモデルと呼べるものがない。また研究の方向性としても、人間の視覚情報処理機構を神経回路によって模擬する方法や、ゲシュタルト法則を計算理論として忠実に翻訳していく方法、さらにはごく簡単な規則をボトムアップに適用していく方法、さらには心理学方面からの研究もおこなわれるなど種々の方法が入り乱れて混沌とした状況であり、今後の研究の方向性もいまだ明確とはなっていない。そこで本章では仮想引力モデルという物理的なモデルに基づいた簡単なモデルによって、プリミティブの集合を構造化する方法を述べた。仮想引力モデルは、2領域の面積が大きく距離が小さいときに関係を張るという直観的で単純なモデルであり、それが長所でもあり欠点でもある。つまり計算は単純でわかりやすいが、しかしなぜそのような計算をするかという根拠が明確ではない。また仮想引力式(7.1)での2領域間の距離計算の方法や、距離が r^2 となっている点など、仮想引力モデルには今後の検討課題が多い。

また画像表現モデルとしてはグラフ構造を用いた。グラフ構造はその記述能力が強力であることが特長であり、また属性付きのグラフ構造とすることにより記号的な情報ばかりではなく数値的な情報も格納することができる。そしてこのグラフ構造によって画像の領域情報と空間情報が表現されるので、問題領域から独立した表現を得ることができる。しかし画像表現モデルに本当にグラフ構造が適切かという問題については、グラフ構造ではないモデルの方が適当な画像もあるかもしれない。どのような新しいモデルが考えられるかという点を今後も検討する必要がある。

第8章

A* アルゴリズムを用いたグラフマッチングによる画像検索

8.1 概要

本章では類似画像検索への応用を目的とした、階層化属性付き関係グラフのマッチングの高速化を主に検討する [113]。類似画像検索では、検索キー画像と蓄積画像との類似度をそれぞれ計算し、類似度の大きい順に整列して出力する処理が基本的である。しかし類似度の計算は一般に計算量が大きいため、画像データベースが巨大となるにしたがって類似度計算の高速化が大きな課題となる。本研究では画像表現モデルとして、モデリング能力が強力であるグラフ構造の一つである階層化属性付き関係グラフを用いている。このグラフ構造間の類似度に相当するグラフマッチングコストの計算は、組合せ的爆発を引き起こす典型的な問題であり、この欠点が画像表現モデルとしてのグラフ構造の利用を難しくしている。そこで本研究はグラフマッチングの高速化を追究する。そのために用いた工夫は以下の3点である。

1. 最良優先探索 A* アルゴリズム
2. 認容可能条件を満たすヒューリスティック関数
3. 最大許容コスト

特に本研究の特徴は、解の最適性を犠牲にせずにグラフマッチングの高速化を達成することを目指しているところにある。従来の手法では高速化の代償として解の質を準最適解に落とすアプローチが多いものの、本研究では厳密な最適解を安定して得たいという必要性があるからである。それは第9章で詳しく述べることになるが、その背景を簡単に述べると、このグラフマッチングの結果を用いてグラフマッチングのパラメータを最適化することを考えていることが背景にある。グラフマッチングの解が準最適解であればグラフマッチングの結果が毎回異なり、したがってパラメータの最適化が難しくなる。そこで最適解を安定して求められ、最適化手法を安定に動作させることのできる方法が必要となるのである。

8.2 グラフマッチング

まず最初にグラフマッチングの基本的な概念を述べる。以下では階層的属性付き関係グラフ H のノード集合を N 、アーク集合を $B \subseteq N \times N$ とする。またマッチングする2個の階層的属性付き関係グラフを H_1 、 H_2 と表記し、そのノード集合を N_1 、 N_2 とする。

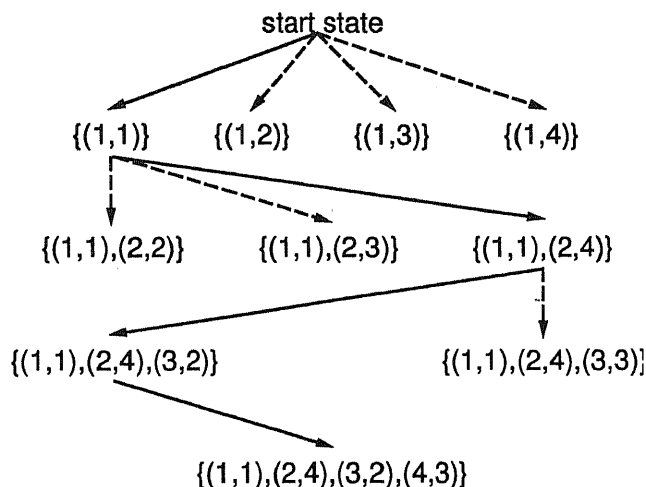


図 8.1: 4 個のノードを持つ 2 個のグラフをマッチングした場合の状態空間。実線が最小コストに対応した経路。

本研究で用いるグラフマッチングの方法は Bunke[10] によるものである。この方法ではグラフマッチングを、代替・削除・挿入という基本的な変形操作により 2 個のグラフを最小のコストで同型とするという、グラフ同型問題に帰着させる。このような 2 個のグラフのグラフマッチング問題は、2 個のグラフからそれぞれ任意に 1 個ずつ取り出した 2 個のノードのペア $(i, j) \in N_1 \times N_2$ を節点とした状態空間の中から、最小コストに対応する経路を探索する問題になる。この状態空間は図 8.1 のような木構造で表現される。ここで節点の展開とは、ある状態を中心にしてそこから次に到達できる状態をすべて生成する処理のことを指す。すなわち状態 S のノードに対して、そこから次に到達できるすべての子ノード S' を生成する処理である。第 8.5 節で述べる探索方法の差異は、生成処理の方法の差異からくるものである。

8.3 グラフマッチングコスト

まずマッチングコストは、シンボルコストと属性コストの和として定義する。そして個々のグラフ変形においてマッチングコストを計算し、最終的にはマッチングコストをすべて加算することで全体のグラフマッチングコストを計算できる。ここでは探索途中の節点 $(i, j) \in N_1 \times N_2$ で生じるマッチングコストについて説明する。

8.3.1 シンボルコスト

シンボルとはノードやアークの特徴を記号的に示すものであり、シンボル間の類似度を代替コストとして定義する。ここでシンボル α とシンボル β の代替コスト $\text{subcost}(\alpha, \beta)$ は、二つのシンボル間で交換則が成立するとして以下のように定義する。

$$\text{subcost}(\alpha, \beta) = \text{subcost}(\beta, \alpha) = g_{\alpha\beta} \quad (8.1)$$

さてここで、ノードの数が異なるグラフ間でもマッチングを可能とするために、特別なシンボルであるヌルシンボル — 大きさがゼロのノードやアーク — を導入する。ヌルシンボルを $\$$ と表記すると、シンボル α との代替コストを以下のように定義する。

$$\text{subcost}(\$, \alpha) = \text{subcost}(\alpha, \$) = g_{\alpha\$} \quad (8.2)$$

8.3.2 属性コスト

本研究では属性として、サイズ属性と形状属性の2種類を考える。これは主にヌルシンボルの導入によって生じる区別である。

サイズ属性 サイズ属性は要素の大きさを示すものであり、ノードの場合は対応する領域の面積、アークの場合は対応する関係の距離などに対応する。他の属性と最も異なる点は、これがヌルシンボルにまで連続的に定義することができるという点にある。

形状属性 形状属性は形状特徴などを数値的に表すものであり、ノードの場合は縦横比、アークの場合は水平方向に対する傾きなどがある。またアークの場合は2項関係に対応する属性として、両端のノードの面積の比や角度差などの2項属性を考えることもできる。ただしこの属性はヌルシンボルには定義することができない。このことがマッチングコストの計算に影響する。

属性コストの計算方法

次に属性コストの計算方法を述べる。以下ではノードの場合だけを取り上げるが、アークの計算方法も基本的に同様である。さて n 個の属性値を並べたベクトルを n 次元の属性ベクトルとしよう。節点 (i, j) でマッチングされる2つの n 次元属性ベクトルを $\vec{x}_i = (x_1^i, \dots, x_n^i)$ 、 $\vec{y}_j = (y_1^j, \dots, y_n^j)$ とする。このベクトルの間の距離がマッチングコストとなる。

一般にこの距離は二つのベクトルを引数とする関数 $p(\vec{x}_i, \vec{y}_j)$ で表すことができるだろう。しかし各成分が全く他と無関係な種類の属性を表現していれば、それらの属性値は別々に取り扱うことができる。これをすべての属性の間で検討していけば、 n 個の属性をお互いに無関係な $L (\leq n)$ 個のグループに分け、距離関数を L 個の部分空間でそれぞれ独立に計算できるようになる。例えば普通の属性はそれぞれが独立しているが、位置属性では2個または3個の属性が関係して1つの距離関数を形成する。そこで属性ベクトルの距離関数として以下のように L 個の距離関数を加えたものを考えよう。

$$\text{cost}(\vec{x}_i, \vec{y}_j) = \sum_{k=1}^L g_k \frac{p_k(\vec{z}_i, \vec{w}_j)}{s_k} \quad (8.3)$$

ここで \vec{z}_i は \vec{x}_i の中で k 番目のグループにグループ分けされた属性値のベクトルであり、 \vec{w}_j と \vec{y}_j の関係も同様である。ただし距離関数は全く自由に定義できるため、これを式 (8.3) において合計する際には、各距離関数の影響がほぼ等しくなるように正規化する必要がある。そこで p_k の値域を同程度にするための定数として正規化係数 s_k を用意する。ここでわかることは L 個の距離関数の具体的な定義がまず第一の問題であるという点である。

さて式 (8.3) でもう一つ重要なのは重み係数 g_k の存在である。この係数はいわば受動的に決定される正規化定数とは異なり、能動的に変化させるための係数である。すなわち、この係数を変化させることによって各距離関数の重要度を変化させると、ベクトル間の距離が変化して類似度も変化する。よってこの値を積極的に操作することによって、検索意図を類似度に反映させることが可能となる。このように検索意図を反映させるための重み係数の設定が、マッチングコストに関係する第二の問題である。

この二点の問題について本研究では考察するが、第二の問題についての議論は第9章に回し、ここでは第一の問題のみを議論する。

距離関数

距離関数の定義は本質的に属性の性質に依存する。この意味では距離関数は問題領域から独立したものではない。では各属性に対してどのような距離関数を考えればよいだろうか。

1. 属性の物理的な性質から距離関数を導ける場合がある。例えばこの世界がユークリッド空間であることから、位置の属性間の距離はユークリッド距離によって計算することができる。このように比較的客観的に定めることができる種類のものがある。ただし本当に人間がこのように知覚しているかという点については保証されない。
2. 物理的な性質からは導けないがそれなりの仮定を置くことによって導ける場合がある。例えば角度に関する属性は、人間の視覚が純粹に角度差のみを見ているとすれば、第 A.2 節で述べるような距離関数が定義できる。またサイズ属性に関しては、人間が長さや面積などの倍率を見ていると仮定すれば、サイズ属性の距離関数を第 A.3.1 節の対数距離とみなし $p(x, y) = |\log(x/y)| = |\log x - \log y|$ ($x, y \gg 1$) と定義することができる。ただしこの場合には、厳密には距離関数の算出に心理学的な実験が必要となる。おそらく距離関数は非線形なものとなり、計算は複雑なものとなるだろう。よって距離関数がよほど非線形なものと思像される場合でなければ、このように簡略化した形の距離関数を用いるのが妥当であると考えられる。
3. どのような距離関数が適当かが直観的にはわからない場合がある。たとえば 0 から 1 の間の値となる円形度の属性値を考えると、その距離関数に適切な表現は直観的にはわからない。この種の属性値が検索に有効であるかどうかは未知数であるが、もし他に適当な表現がないときは、ユークリッド距離や対数距離を距離関数とするしかないだろう。

このように、属性値の距離関数は属性個々の性質に依存していることを議論した。その意味ではグラフ間の類似度は問題領域に依存した表現である。しかしその部分はブラックボックスとして扱うことで、式 (8.3) は形式的には問題領域から独立した表現と捉えることができよう。

8.3.3 ヌルシンボルとの距離

ノード数やアーク数の異なるグラフ間でもマッチングが可能ないように導入したヌルシンボルは、Inexact Matching を実現するための重要なシンボルである。しかしこれはスケール(大きさ)が 0 の要素であるため、ヌルシンボルのノードやアークにはスケール属性は定義できるが形状属性は定義できない。このことは、ヌルシンボルとマッチングしようとするノードやアークとの属性コストを計算する際に、未定義の属性値とのマッチングコストは計算できないという問題を引き起こす。この点は後に述べるように将来のマッチングコストを推定する際に厄介な問題となる。

本研究ではヌルシンボルとのマッチングコストでは、まず式 (8.2) の代替コストを加える。それに加えて面積が大きいシンボルをヌルシンボルにマッチングさせるとマッチングコストが大きくなる項を入れた。その係数を u とすると、

$$\text{cost}(\alpha, \$) = \text{subcost}(\alpha, \$) + u \frac{p_\alpha(A_\alpha, 0)}{s} \quad (8.4)$$

というマッチングコストを定義した。ここで $p_\alpha(\cdot, \cdot)$ は対数距離に基づく距離関数であり、また A_α はシンボル α のスケール属性(面積や長さなど)、そして s は距離関数の正規化係数である。

表 8.1: 代表的な探索アルゴリズムの比較。最短経路を求める問題の場合。

	深さ優先探索	幅優先探索	最良優先探索
判断規準	開始節点から遠い節点を優先	開始節点から近い節点を優先	評価関数値が小さい節点を優先
長所	状態空間木の枝分かれが多いと効率がよい	状態空間木の枝分かれが少ないと効率がよい	優先順位付き待ち行列のおかげで展開する節点の数が少ない
短所	状態空間木の枝分かれが少ないと効率が悪い	状態空間木の枝分かれが多いと効率が悪い	優先順位付き待ち行列の管理に余分な計算コストを必要とする

8.4 グラフマッチング高速化の基本方針

さてここでは今までの導入部を受けて、グラフマッチング高速化のための基本方針をまとめる。図 8.1 のような状態空間の探索では、節点を展開するごとに第 8.3 節のグラフマッチングコストを新たに計算しなければならない。しかしグラフマッチングコストの計算は大きな計算コストを要求するため、まず節点の展開回数をできるだけ少なくすることが第一の目標となる。次に、たとえ有望でない方向に探索が進んだとしても、できるだけ早めに探索経路を変更できれば節点の無駄な展開を抑えることができるだろう。そこで以下のような方針となる。

1. 最も有望な節点を優先的に展開し、状態の無駄な展開を抑える。
2. できるだけ早い段階で、有望でない経路の探索を中断する。

さて 1 を実現するためには、有望な節点から順に展開できるように、展開可能な節点を有望さの程度に応じて整列させておけばよい。次に 2 を実現するためには、探索を続ける価値がある最大のコストをどの状態でも把握できるようにし、探索を続ける価値がなくなれば即座にマッチングを中断できる仕組みが必要となる。そこで以下では、1 に関連した高速化手法を第 8.5 節で、1 と 2 に関連する高速化手法を第 8.6 節で、そして 2 に関連する高速化手法を第 8.7 節で述べる。

8.5 探索アルゴリズムの比較

一般的な探索アルゴリズムでは、展開済みの節点から到達できるがまだ展開していない節点 S の集合について、何らかの評価関数 $f^*(S)$ の値が最大 (または最小) の節点から展開していく方法をとる。代表的な探索アルゴリズムには、深さ優先探索 (depth-first search)、幅優先探索 (breadth-first search)、最良優先探索 (best-first search) の 3 種類があり、その特徴を表 8.1 にまとめた。

ここでの問題は、最良優先探索とその他の手法との比較である。最良優先探索では、優先順位付き待ち行列を活用することで出発点からの距離を問わずに有望な節点から展開していくため、展開する節点の数を他の 2 手法に比べて削減することができる。しかし待ち行列を管理する必要があるため、この部分に余分な計算コストを必要とする。先にも述べたようにグラフマッチングの計算では節点展開の計算コストが大きいと予想されるため、展開の計算コストと待ち行列の管理コストとのトレードオフを実験を通して比較することを目指す。

8.6 認容可能条件を満たすヒューリスティック関数

次に評価関数 $f^*(S)$ について考察する。一般的に評価関数 $f^*(S)$ は、 $f^*(S) = g(S) + h^*(S)$ という形式で表される。ここで $g(S)$ は開始節点 s から節点 S に至る経路のコストであり、 $h^*(S)$ は節点 S から終了節点までの経路の最小コストの推定値である。もし $h^*(S)$ が真の値 $h(S)$ と全く同一ならば、最適経路上以外の節点を展開することなく最適解を発見できる。さて $g(S)$ の小さい順に S を展開する戦略が最良優先探索であるが、さらに探索効率を高めるためにヒューリスティック関数 $h^*(S)$ を活用する。特に、簡単な計算で真の値に近い推定が得られる関数が理想的である。このようなヒューリスティック探索の手法を応用することで、可能な解の数が組合せ的に爆発するのを防ぐことができる。

ここで $h^*(S)$ が認容可能条件 (admissibility condition) を満たせば、得られる解の最適性を保証しながら計算の高速化が可能となることが知られている。このようなアルゴリズムを A^* アルゴリズムという [114]。認容可能条件を満たすためには、 $h^*(S)$ の推定値が真の値 $h(S)$ より小さいかまたは等しいという性質が常に成立しなければならない。そこで本研究では認容可能条件を満たす $h^*(S)$ を導出し、同時にその関数の計算が簡単であり実用的であることを示す。通常は問題領域の知識を用いてヒューリスティック関数を定式化するが [115]、本研究が提案するヒューリスティック関数は系統的に導くことができるものである。なお深さ優先探索と幅優先探索では、節点の展開順序の制御に $f^*(S)$ を使うことはないが、探索の枝刈りとして $f^*(S)$ の値を有効に活用することができる。

8.6.1 ノードのヒューリスティック関数

以下では節点 S までにマッチングされたノードの集合を M とする。すると節点 S の時点でマッチングされていないノード集合は $Y = N - M$ で表せる。

ノードの場合のシンボルコスト まず Y_1 と Y_2 のシンボルの内訳を調べる。そして任意のシンボルの組合せ (α, β) に関して、 Y_1 と Y_2 のノード数の差 $r_{\alpha\beta}$ を記録する。これらが将来の代替コストの原因となることから、このようなノード数の差を最小の代替コストで解消する方法を考える。

まず、任意のシンボルの組合せ (α, β) の中で最もシンボルコスト $g_{\alpha\beta}$ の値が小さい組合せに関するノード数の差 $r_{\alpha\beta}$ を調べ、 $g_{\alpha\beta} \times r_{\alpha\beta}$ を $h^*(S)$ に加える。そしてシンボル α と β に関するすべての組合せの差から $r_{\alpha\beta}$ の値を引く。この操作をすべての差が解消されるまで続けると、シンボルコストの $h^*(S)$ が得られる。

ノードの場合の属性コスト 式 (8.3) の距離関数 $p(x, y)$ は、先述したように人間の知覚を反映した関数であるため関数形は一般に複雑である。またヌルシンボルの導入により、形状属性については統一的な扱いができない。したがってここでは、ヌルシンボルでも属性値が定義されており、かつ距離関数が $p(x, y) = |q(x) - q(y)|$ という差分の形式で表現されている属性に関するヒューリスティック関数を導出する。この形式の距離関数となる代表的な場合は、第 A.3.3 節の間隔尺度と比率尺度の場合である。間隔尺度の場合は属性値そのものの差が距離関数であり、比率尺度の場合は属性値の対数の差が距離関数となるからである。間隔尺度と比率尺度は尺度としては汎用性があり実用上は重要な形式である。

k 番目の属性がこの種類の属性であるとき、そのヒューリスティック関数 $h_k^*(S)$ は

$$h_k^*(S) = \sum_{\Psi} g_k \frac{p_k(x_k^i, y_k^j)}{s_k} \quad (8.5)$$

と定義できる。ここで $\Psi = \{(i, j) | (i, j) \in Y_1 \times Y_2\}$ である。ここで $h_k^*(S)$ が認容可能条件を満たすように、ノードの可能な組合せ集合 Ψ の中から最小の $h_k^*(S)$ を推定する。距離関数が $p_k(x_k^i, y_k^j) = |q(x_k^i) - q(y_k^j)|$ の形式の場合、式 (8.5) は以下のように変形できる。

$$\begin{aligned} h_k^*(S) &= \sum_{\Psi} g_k \frac{|q(x_k^i) - q(y_k^j)|}{s_k} \\ &\geq \left| \sum_{\Psi} (q(x_k^i) - q(y_k^j)) \right| \frac{g_k}{s_k} \\ &= \left| \sum_{i \in Y_1} q(x_k^i) - \sum_{j \in Y_2} q(y_k^j) \right| \frac{g_k}{s_k} \end{aligned} \quad (8.6)$$

ここで $|a| + |b| \geq |a + b|$ という関係を用いている。式 (8.6) は、真のマッチングコストよりも小さい推定値を常に与えるため認容可能条件を満たしている。しかも式 (8.6) は、まだマッチングされていないノード集合の属性値を個々に合計し、その差の絶対値を推定値とするという、単純で実用的な形式になっている。そこで式 (8.6) を認容可能条件を満たすヒューリスティック関数として提案する。

ただし形状属性についてはこの推定ができないことに注意する。なぜならヌルシンボルでの形状属性が未定義であるため、ヌルシンボルとのマッチングコストが式 (8.4) という式 (8.3) とは別の式で計算されることになり統一的な扱いができないからである。この点を改良することがさらに有効なヒューリスティック関数を求めるためには非常に重要な点であり、マッチングコストの定義をヌルシンボルにあわせてうまく定義し直すことが必要となるだろう。

8.6.2 アークのヒューリスティック関数

アークについても基本的にはノードとほぼ同様の方法であり、マッチングされていないアークについてヒューリスティック関数を計算する。しかしマッチングされていないアークを求める計算がノードの場合よりも少し面倒なため、アークのヒューリスティック関数による高速化の効果はノードの場合よりも一般的に劣る。

8.7 最大許容コスト

本研究は類似画像検索への応用を目的として、グラフマッチングの高速化を目指している。そこで類似検索の問題構造をうまく活用して高速化を図る。さてデータベースの N 個のグラフの中から、類似検索で上位 M 個のグラフを検索する上位検索問題を考えよう。

既に検索キーとデータベース中の L 個 ($N > L \geq M$) のグラフとのマッチングが終了しているとし、その時点での検索順位が上位 M 番目のグラフマッチングコストを C_M とする。さて $L+1$ 個目のグラフと検索キーとのマッチングでは、そのマッチングコストが C_M 以内かどうかだけに興味がある。もしマッチングコストが C_M を越えれば $L+1$ 個目のグラフは上位 M 位以

内に検索されず、それならばマッチングを即座に中断すべきだろう。本論文ではこのコスト C_M を最大許容コストと呼ぶ。

最良優先探索の場合は、評価関数が $f^*(S) \leq C_M$ の状態、つまりマッチングを続ける価値がある状態しか優先順位付き待ち行列に登録しないという方法を使う(図 8.2 の 6 を参照)。すると、もし終了状態の一つが待ち行列内になれば、このマッチングの類似順位はたかだか M 位以下であると判定できるため、即座にマッチングを中断することができる。それに対して深さ優先探索と幅優先探索の場合は、探索経路の途中で $f^*(S) > C_M$ となった時点でこの情報を枝刈りに使うことができる。しかしそこでマッチングを中断することはできず、他の経路の可能性が尽きるまで探索し続けなければならない。

また C_M は検索が進むにつれて徐々に小さくなるが、特に検索初期段階で C_M を小さくできれば高速化の効果は大きい。そのため、 N 個のグラフをマッチングする順番によって検索時間は大きな影響を受けることになる。このようにマッチングする順序を順序を記述したリストをマッチングリストと呼ぶが、このリストが実際の類似度の順番に並んでいる時に、類似検索は最も高速になる。

8.8 実験結果

8.8.1 実験の概要

今まで述べてきた、 A^* アルゴリズム・認容可能条件を満たすヒューリスティック関数・最大許容コスト、というすべての高速化手法を採り入れたグラフマッチングアルゴリズムを図 8.2 に示す。このように本論文で提案する図 8.2 のアルゴリズムは、グラフの階層性と高速化の観点から A^* アルゴリズムを改良したものである。また優先順位付き待ち行列には改良した 2 分探索木またはヒープを用い、待ち行列の管理についても高速化のための改良をおこなった。一般的に探索空間が狭い時には並べ替えの必要のない 2 分探索木が高速であり、探索空間が大きくなるにしたがって、左右の枝のバランスがとれたヒープの方が有利となる。

このアルゴリズムを用いて以下の実験をおこなう。まず 1027 枚の画像を階層化属性付き関係グラフに変換し画像データベースとする。次にその中から 1 個のグラフを検索キーとして選ぶ。実験としては検索キー自身を含めて 1027 個の中から、上位 M 個を検索する時間を計測する。ここで計算時間としてはグラフマッチング部分のみに要する計算時間を測定し、入出力に要する時間などは含めていない。また 20 回の試行の平均時間を記録した。実験に用いたコンピュータは IBM SP-2 である。

本研究で用いた階層化属性付き関係グラフの平均的な構造は以下になる。ノードの平均個数：上層 2.41 個、下層 6.32 個に対して、アークの平均個数：上層 1.42 個、下層 3.91 個。次にシンボルはノードもアークも各 1 個ずつ。また属性はノードの個数：上層 1 個、下層 3 個に対して、アークの個数：上層 1 個、下層 1 個である。

ただし以下に述べる実験は本論文を書く数ヵ月前に行われたものであり、第 10 章でおこなわれた実験とは条件が異なっていることに注意していただきたい。

Input: 階層化属性付き関係グラフ H_1 と H_2 、および最大許容コスト C_M 。

Output: H_1 と H_2 のグラフマッチングコスト (非類似度)、およびマッチングされたノードペアを記録した最適な経路。

- Steps:**
1. 出発状態 s を優先順位付き待ち行列 OPEN に登録する。
 2. OPEN に関して
 - (a) 評価関数が最小の状態があればその状態を S としそれを OPEN から削除する。
 - (b) もし OPEN が空ならば失敗として 8へ。
 3. もし S が終了条件を満たすならば、 S から s を指すポインタをたどることによって得られる解とともに成功として 8へ。
 4. 状態 S を展開し S のすべての子ノード S' を生成する。このとき
 - (a) もし S' が下位グラフを持っているならば、
 - i. もし下位グラフのマッチングが初めてであれば、下位グラフどうしのマッチングコスト $d(S')$ を再帰的に計算する。これが失敗ならば $d(S') = C_M$ とする。配列の S' に対応する要素に $d(S')$ を登録する。
 - ii. そうでなければ、配列の S' に対応する要素から過去のマッチングコスト $d(S')$ を読み出す。
 - (b) そうでなければ $d(S') = 0$ とする。
 5. 状態 S' で新たに生じるコスト $c(S')$ を計算し $g(S') = g(S) + c(S') + d(S')$ とする。さらに状態 S' からゴールまでのコスト $h^*(S')$ を推定する。 $f^*(S') = g(S') + h^*(S')$ を計算。
 6. もし $f^*(S') \leq C_M$ ならばそれらを OPEN に登録する。そして OPEN を評価関数の値にしたがって再整列する。
 7. 2に戻る。
 8. 上位グラフがあれば上位に戻る。なければ終了。

図 8.2: A^* アルゴリズム・認容可能条件を満たすヒューリスティック関数・最大許容コストのすべての高速化手法を採り入れたグラフマッチングアルゴリズム。優先順位付き待ち行列 OPEN には、改良した 2 分探索木を用いている。

表 8.2: 1027 枚のグラフマッチングの計算時間 (20 回の試行の平均、単位は秒)。(a) 両方の高速化手法を用いた場合、(b) ヒューリスティック関数だけを用いる場合、(c) 最大許容コストだけを用いる場合、(d) 両方とも用いない場合、である。上段はアークのヒューリスティック関数を用いた場合、下段は用いない場合である。

	最良優先探索		幅優先探索		深さ優先探索	
	$M = 10$	$M = 1027$	$M = 10$	$M = 1027$	$M = 10$	$M = 1027$
(a)	1.724	8.570	2.183	25.490	1.762	8.967
	2.321	9.710	2.693	20.282	2.307	8.873
(b)	8.451	8.462	25.394	25.369	8.863	8.861
	9.607	9.608	20.180	20.174	8.773	8.776
(c)	2.787	9.605	3.050	17.357	2.759	8.393
	2.786	9.636	3.046	17.341	2.721	8.393
(d)	9.483	9.479	17.232	17.122	8.284	8.287
	9.481	9.480	17.142	17.135	8.271	8.280

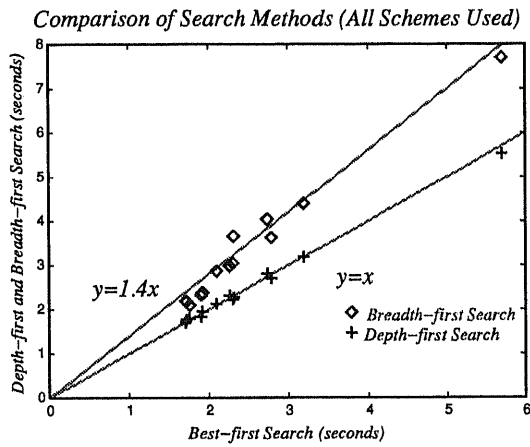
8.8.2 実験の結果

まずある重み係数を与えた時の検索時間を表 8.2 に示す。 $M = 10$ の上位検索では、最良優先探索ですべての高速化手法を適用するという最も速い場合で、約 1.72 秒の検索時間を達成している。しかし重み係数の値によって検索順位が変わるため、この検索時間も重み係数の値によって変わってくる。そこで第 9 章の方法によって最適化した重み係数の組を 13 組用意し、それぞれについて検索時間を計測してグラフにした。

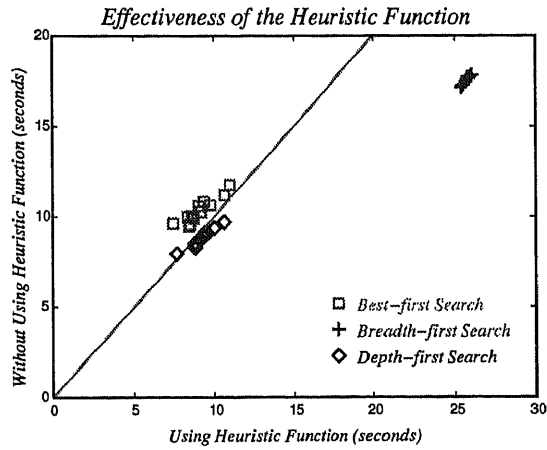
探索方法による比較 まず三種類の探索方法を比較したのが図 8.3 である。その結果、最良優先探索と深さ優先探索が最も高速であり、幅優先探索は最良優先探索の 1.4 倍程度の時間を要するという結果を得た。最良優先探索は、探索する範囲が最も狭いという意味で効率的ではあるが、その分優先順位付き待ち行列の管理に余分な計算コストを必要とする。幅優先探索も同様である。そのため単純なアルゴリズムである深さ優先探索もよい成績を残した。

認容可能条件を満たすヒューリスティック関数による効果 次に図 8.3 ではヒューリスティック関数による推定を用いた効果を見る。その結果最良優先探索では確かに検索時間を短縮できたが、その効果は期待していたほどではない。それどころか、深さ優先探索や幅優先探索では、これを用いることによりむしろ性能が悪化している。このことは、ヒューリスティック関数の効果がそれほど大きくないことを示唆している。これは、ヒューリスティック関数の推定値が真の値よりもかなり低い見積りであったこと、特に形状属性のヒューリスティック関数が使用できないために有効な推定値を得られなかったことが大きな原因である。なおアークのヒューリスティック関数は、表 8.2 の場合はそれを用いた方がより高速となっている。

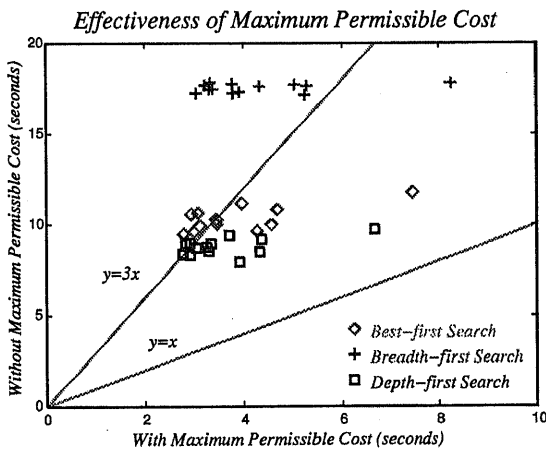
最大許容コストによる効果 どの探索方法を用いた場合でも図 8.3(c) のように、 $M = 10$ の上位検索では最大許容コストを用いることによって大幅に検索時間を削減できた。最大許容コストを用いることによって、検索時間はほぼ 1/3 になっている。ただし重み係数の組によってはそれほ



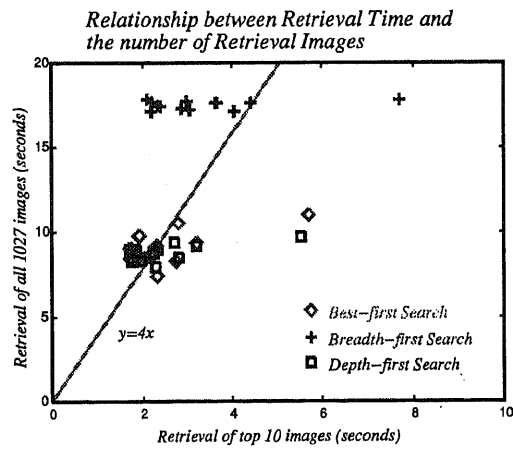
(a)



(b)



(c)



(d)

図8.3: 条件を変えた場合の検索時間の比較。(a) 三種類の探索方法の比較、(b) ヒューリスティック関数の効果、(c) 最大許容コストの効果、(d) 全数検索と上位検索の比較。

どの時間改善効果が見られない場合もある。そして $M = 1027$ のような全数検索では、明らかに最大許容コストの効果はない。またこの結果は、前もって最適に整列されたマッチングリストを用いて測定されたものである。つまり最大許容コストによる効果の上限を示す。一般的には異なる重み係数を用いて検索した前回の検索順位で整列したマッチングリストを用いるため、最大許容コストによる速度向上効果は実際には割り引いて考える必要がある。

全数検索と上位検索の比較 どの探索方法を用いた場合でも図 8.3(d) のように、上位検索 $M = 10$ と比べて全数検索 $M = 1027$ の検索時間は約 4 倍だった。なお、最良優先探索ではヒューリスティック関数を適用した検索時間を採用しているが、その他の探索方法ではヒューリスティック関数を適用しない場合の検索時間を採用している。

8.9 まとめ

本研究で提案した高速化手法をすべて活用することによってグラフマッチングを高速化することができた。最も効果が大きい高速化手法は最大許容コストだったが、他の方法については期待したほどの高速化が達成できなかった。特にヒューリスティック関数にはより有効な表現を与える必要があるが、これはヌルシンボルに対するマッチングコストをどのように定義するかという点が大きく関係してくる。また角度属性に対するヒューリスティック関数の導出も大きな課題である。さらに広く考えれば、探索方法も本研究で検討した範囲のみならず他の手法も幅広く検討する必要があるだろう。最後に、画像数が 1 万枚以上の大規模データベースに「スケールアップ」しても通用する手法とするために、今後とも検索時間の高速化を常に目指す必要がある。

第9章

人為選択法に適した遺伝的アルゴリズムを用いた類似検索規準の学習

9.1 概要

柔軟な画像検索システムとはユーザの多様な検索意図に応えられるものでなければならず、また画像間の類似度を個々のユーザの感性に適応させるためのメカニズムも備えたものでなければならぬ。本章では類似検索を行うユーザの検索意図や感性をまとめて類似検索規準と呼び、この類似検索規準をシステムが自動的に学習する方法を提案する。具体的には第8章で述べたグラフマッチングに関する重み係数を最適化する手法を提案する。まず類似検索規準に対応する教師データをシステムに与えると、システムは教師データに対する最適な重み係数を学習する。最適化手法には遺伝的アルゴリズムを用い、特にブラウジングのような用途に適した遺伝的アルゴリズムとして、「ライフサイクル型遺伝的アルゴリズム」という新しいモデルを提案する。このような重み係数の調整機構を内部に備えることで、画像データベースはより柔軟な類似検索を可能とする。

9.2 類似検索規準の学習

9.2.1 想起と認識

類似画像検索とは、画像データベース中に蓄積された大量の画像の中から、ユーザが提示した検索キーに内容が類似した画像を検索する検索手法を指す。本研究では類似検索に適切な検索パラメータを自動的に最適化する方法を研究する。特に本研究では人間の認知特性、すなわち想起するより (recall) も認識する (recognize) 方が容易である、という特性を画像検索に応用することを考える。例えば親しい人の似顔絵を書こうとしても、細部までをある程度正確に想起しながら書くことは普通の人には難しい。ところがその人の似顔絵を見れば、それがたとえ概略画であっても人間は一目瞭然で認識できる。つまり人間にとっては、想起するよりも認識の方がはるかに楽な作業なのである。類似検索の一つの検索方法として、ユーザが概略画を描いてシステムはそれに類似した画像を検索するという例示検索の方法が広く用いられているが、先述したような人間の認知特性を考えると、この方法は必ずしもユーザにとって楽な方法ではない。では逆に人間に「認識」させながら画像検索することはできないだろうか。このように認識によって検索するアプローチは坂内によると「あんな」検索と呼ばれ、検索意図の解釈、くみとりをつかさ

どるユーザインタフェース部の構成が重要となる [33]。また笠原らは同様の手法をナビゲーションと名付けている [116]。

本研究では「想起」からのアプローチと「認識」からのアプローチを併用する。そして類似検索規準の学習、つまり検索パラメータの最適化には遺伝的アルゴリズム (GA) を用いる [117, 118]。また特に認識に基づいたアプローチであるブラウジングのような用途に適した遺伝的アルゴリズムとして、「ライフサイクル型遺伝的アルゴリズム」という新しい手法を提案する。

9.2.2 例示検索とブラウジング

類似検索手法としては以下の2種類が代表的である。

例示検索 適切な検索パラメータを与えて画像検索する場合であり、ユーザは検索したい画像を前もって想起することができる場合である。この場合は検索パラメータをシステムに与えることで即座に類似検索を実行することができ、したがって高速で実用的な検索手法である。この検索パラメータには、前もって用意された検索パラメータを用いるか、またはユーザ自身が適切な検索パラメータを「想像」して与える必要がある。しかしユーザによって検索目的や感性が多様に異なることを考えると、適切な検索パラメータすべてを事前に用意するのは難しい課題であり、また検索パラメータの数が多い場合には、人手による試行錯誤で適切な検索パラメータを与えるのも困難な問題である。

ブラウジング システムが提示する画像を見ながら検索画像の候補を徐々に絞りこんでいく場合であり、ユーザは検索したい画像を前もって想起できない場合が多い。つまり、検索開始時にはユーザの類似検索規準があまり明確ではなく、むしろ検索しながら類似検索規準を形成していく方法である。この方法では検索に時間がかかることが難点であるが、一方検索しながら検索意図を柔軟に変更できるのは大きな利点でもある。またブラウジング過程を記録することで、次回以降の検索にこの情報を反映させることもできる¹。

このように2種類の方法があるが、もちろん高速な検索手段は前者である。類似検索の研究でも前者に基づいたアプローチが多い。しかし類似検索において適切な検索パラメータを与える問題が人手でいかに困難な問題であるかは、それほど多くの人には認識されていない。人手による試行錯誤で求める方法も不可能ではないが、一般的に画像処理から得られる画像特徴は人間が感じる直観的な画像特徴とは一致しないことが多く、さらに検索パラメータの数が増加すると検索パラメータが相互に影響を及ぼし、その結果は予測できないものとなるのが現実である。

9.2.3 問題の定式化

そこで本研究では、検索パラメータを手で調整するのではなく自動的に調整する方法を考える。その下準備として、ここではまず本研究が扱う類似検索規準の学習問題を定式化する。こ

¹遺伝的アルゴリズムと関連して、ブラウジングが本質的に多点探索であるという性質は興味深い。ブラウジングでは、最初はできるだけ多様な画像をユーザに提供し、提示画像に対するユーザの反応を見ながら徐々に画像探索範囲を狭める。言い換えれば、画像探索空間に最初は複数の探索開始点をばらまき、それぞれの点が局所的に探索しつつも相互に探索結果を交換し、徐々に探索範囲を絞り込む図式となる。この方法は遺伝的アルゴリズムの探索法と表面的にはよく類似している。

これは基本的には第 8.3 節で述べた各コストの重み係数を最適化する問題である。ただし類似検索規準に対応する個々の画像特徴の重要度が重み係数であり、これが検索時の検索パラメータとなる。そこで n 個の画像特徴の重み係数をまとめたベクトル $\mathbf{W} = (w_1, w_2, \dots, w_n)$ を用意する。重み係数の変化が画像間類似度の変化を引き起こすことから、画像 i と画像 j の間の総合的類似度は $S = f(\mathbf{W})$ と重み係数の関数になる。そこで重み係数を積極的に変化させることにより、類似検索規準を最適に反映する重み係数ベクトルを求める問題が、本研究で扱う類似検索規準の学習問題である [119]。

ここで画像特徴が特徴ベクトルで表され、その距離関数がユークリッド距離で与えられる場合については、ユークリッド空間を軸方向に任意に拡大 / 収縮させることで最適化する手法 [9] や、第 A.4 節で述べるような多変量解析を応用する手法 [13] などがすでに提案されている。しかし本研究のようにグラフ構造という複雑な画像表現モデルを用いる場合には、複雑な距離関数が非線形な距離空間を生成し、勾配法などの最適化手法には困難な問題となる。そこでこのような非線形な空間に強い探索手法である、遺伝的アルゴリズムが適当な最適化手法となるのである。

このように自動的に最適な重み係数ベクトルを学習するためには二つの方法が考えられる。

1. 前もって例示画に類似した画像を教師データとして与えておき最適化をおこなう。
2. ブラウジングによって対話的に類似画像を選び最適化をおこなう。

本研究では、最初の方法を「オフライン問題」、2 番目の方法を「オンライン問題」と名付ける。

9.2.4 オフライン問題

オフライン問題では、前もって人間が選んだ主観的に類似した画像を教師データとして与え、システムはその教師データを最適に検索するような重み係数を求める方法である。ここで最適性を判定する評価規準には以下の 2 通りがある。

検索順位優先型 (人間が事前に選んだ) 複数の類似画像の検索順位の重み付き合計を最小化する。

類似度優先型 上位に検索した複数の類似画像の、類似度の重み付き合計を最大化する。

この中で後者の規準を本研究では用いる。

ここでこの方法はやや同語反復的であることに注意する。すなわち、もし例示画に類似した画像を教師データとすることができるならば、すでに類似画像は検索できているのではないかという批判が可能である。したがって 1 の方法は以下のような文脈で用いられるべきである。

1. ユーザではないシステム管理者が、適切な検索パラメータをユーザに提供するための事前準備に用いる手法である。
2. ユーザが一度だけ検索パラメータを学習させておけば、別の検索においても同一の検索パラメータを用いた類似検索が可能である。

この点については第 10 章においてさらに議論を進める。

9.2.5 オンライン問題

ユーザがシステムと対話しながら各画像に主観的類似度を与えていき、徐々に重み係数を最適化していく方法である。ユーザは0～1の範囲で主観的類似度を各画像に与えるが、この類似度は検索途中であっても臨機応変に変更できる。つまり検索途中に検索意図が変わるような場合でも類似検索規準を柔軟に変更することができる。

そこで本研究ではこの方法に注目する。ブラウジングを用いて対話的に最適化ができれば、ユーザにとって直観的で使いやすいマンマシンインタフェースを構築でき、また検索途中での類似検索規準の変更にも柔軟に対応できるからである。

9.3 人為選択法

9.3.1 人為選択法の特徴と利点

本研究では対話的なマンマシンインタフェースとして有望な「人為選択法」とよばれる手法を取り上げる。この手法は主観的な判断規準を自然に取り入れることができる点で注目を集めている手法である。初期の研究としては Dawkins の「バイオモルフ」が代表的である [120]。このシステムでは、繁殖と突然変異による進化を用いて、画像空間の中から自分の好みに合った画像を発見できる。また Sims は「人工的進化」により、遺伝的プログラミングを用いてやはり自分の好みを反映した CG 画像を生成した [121]。その他に犯罪者のモンタージュ作成 [122] や人間の顔の線画の自動作成 [123] などにも対話的な遺伝的アルゴリズムが応用され成果を挙げている。

これらの研究の最大の特徴は、プログラムの出力の「良さ」を人間が直接評価するという枠組が、明確に意識されている点にある。従来の研究ではこのような評価規準を明示的に数式として表現し、システムが自動的に出力の良さを判定できるようにすることが目標であった。しかし評価規準が主観的になればなるほど、このような明示的な評価規準を得るのは難しい。そこで最終判断者としての人間と、人間を助ける役割としてのコンピュータとを一体と考えるアプローチである人為選択法が誕生した。この手法は、出力の良さを判断するのがコンピュータにとっては難しいが人間にとっては簡単であるという問題領域において、最も威力を発揮する。

9.3.2 人為選択法の問題点

人為選択法は遺伝的アルゴリズムと組み合わせて用いられ、人間が各個体に与えた評価値を各個体の適応度とする。しかし人為選択法には図 9.1 に示すような問題点がある。

コンピュータがユーザを待たせる時間(形成) 人間が各個体の表現型の良さを目で見えて評価することから、各個体の表現型を人間が見て理解できる形式に可視化する、すなわち遺伝子型から表現型への形成(formation)が必要である。この形成には大きな計算量を要することがあり、個体数が多い場合には1世代の表現型がすべて揃うまでに長い時間を要する場合がある。そこでの時間は評価者であるユーザには無駄であり、この待ち時間の長さが1世代当たりの個体数を制限してしまう。

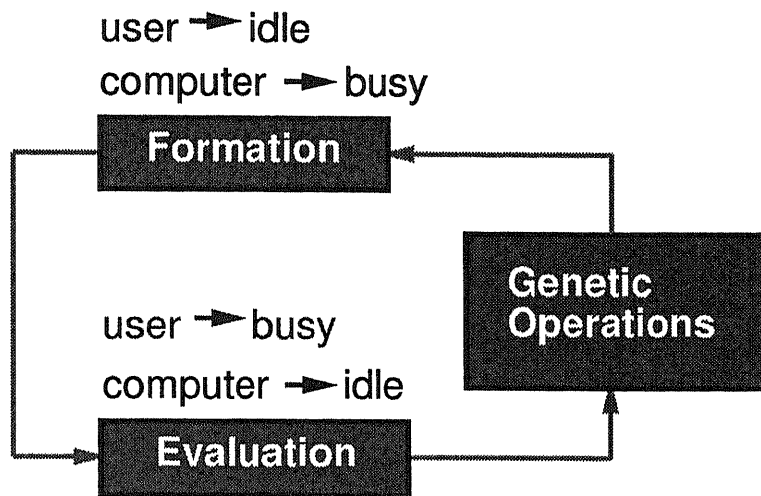


図 9.1: 人為選択法の問題点の模式図。

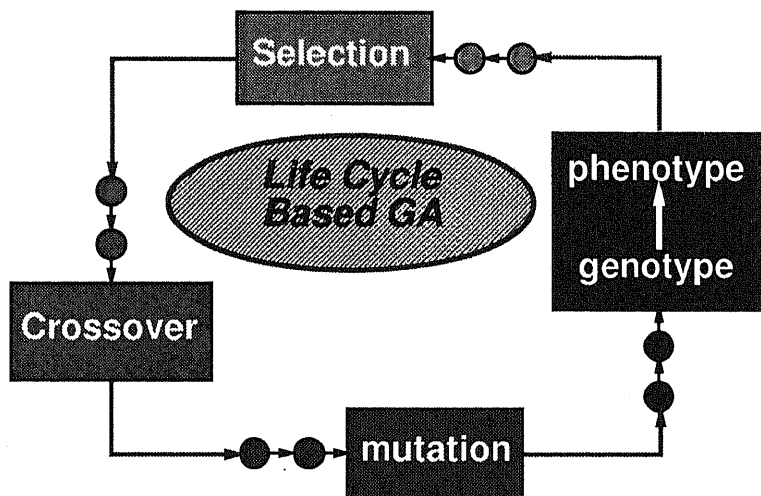


図 9.2: ライフサイクル型遺伝的アルゴリズムの模式図。

ユーザがコンピュータを待たせる時間(評価) ユーザが表現型に評価(evaluation)を与えるためには、表現型に対するある程度の注視時間を必要とする。この待ち時間の間はコンピュータには無駄であり、この待ち時間が全体の高速化を阻んでしまう。

ではこのような問題点をどのように解決すればよいだろうか。その解決法として著者が提案するのが、以下で述べる「ライフサイクル型遺伝的アルゴリズム」なのである。

9.4 ライフサイクル型遺伝的アルゴリズム

9.4.1 世代型遺伝的アルゴリズムからライフサイクル型遺伝的アルゴリズムへ

人為選択法で個体数を小さく制限せざるを得ないのは、ユーザとコンピュータ双方に待ち時間が存在し、効率的な計算ができないことが理由である。言い換えれば「形成」と「評価」という2つの処理が独立しているにもかかわらず、遺伝的操作(genetic operations)という「世代」に基づいた処理があるおかげで、両者を並列に処理できないことが問題となる。それでは遺伝的

操作において「世代」という概念がどのような役割を果たしているかを考えると、それが根本的な役割を果たしているのは適応度の評価だけであることがわかる。つまりその他の部分では世代という概念がない方が効率的なのである。そこで本研究では「世代」という概念に基づかない新しい遺伝的アルゴリズムのモデルとして、「ライフサイクル型遺伝的アルゴリズム (L-GA)」²を提案する。これに対して従来の遺伝的アルゴリズムを「世代型遺伝的アルゴリズム (G-GA)」³と呼ぶ。

従来の世代型遺伝的アルゴリズムでは基本的に個体の集団を単位として処理がおこなわれていた。それに対してライフサイクル型遺伝的アルゴリズムでは、各個体の「ライフサイクル」を重視する。すべての個体が揃って一斉に遺伝的操作を受けるような世代交代の概念を捨て、各個体はそれぞれのライフサイクルに従って生きるというモデルである。ここで個体の寿命はほぼ同じとしているので、これは言い換えれば漸進的な世代交代であるともいえる。しかし連続世代モデル (continuous generation model)[124] などとは根本的に異なる新しい方法を用いている。このモデルを人為選択法に当てはめると、ある個体が人間による評価を受けている間に、他の個体は表現型への形成の段階であったり、遺伝的操作を受ける段階であったり、というように各個体がそれぞれのライフサイクルに応じた段階を過ごしているのである。このようなモデルを機能的に見れば、各段階がパイプライン的に接続され並列的に処理されているモデルであるとも見なすことができる⁴。

ここで注意すべきことは、このようなライフサイクル型のモデルが計算時間の面から有効となるのは、各処理に要する時間がほぼ同等の場合に限られるという点である。本研究のような人為選択法の場合は、形成と評価の双方に多くの時間を要するので、このようなライフサイクル型のモデルが適当となるのである。

9.4.2 人為選択法に応用するライフサイクル型遺伝的アルゴリズム

ライフサイクル型遺伝的アルゴリズムの模式図を図 9.3 に示す。まずシステムの構成は、遺伝子型から表現型への形成を計算する「サーバ」、遺伝的操作やナビゲーションをおこなう「インタフェース」の大きく 2 つに分けられる。一方各個体のライフサイクルは以下のようになる。

1. インタフェース側は適応度が定まった個体に対して遺伝的操作を施し、遺伝的操作が完了した個体 (つまり遺伝子型) をサーバ側に送る。
2. サーバ側は遺伝子型から表現型を形成し、遺伝子型と表現型のセットをインタフェース側に返す。
3. インタフェース側は、受け取った遺伝子型と表現型のセットを、受け取った順にリスト GP List として管理する。
4. インタフェース側は、GP List から表現型を受け取った順序を表すリスト P List を作成し、GP List と P List を関連付ける。

²Life Cycle Based Genetic Algorithm.

³Generation Based Genetic Algorithm.

⁴以前はパイプライン型遺伝的アルゴリズムと呼んでいたが [125]、それはこのような機能的な側面に着目した命名であった。

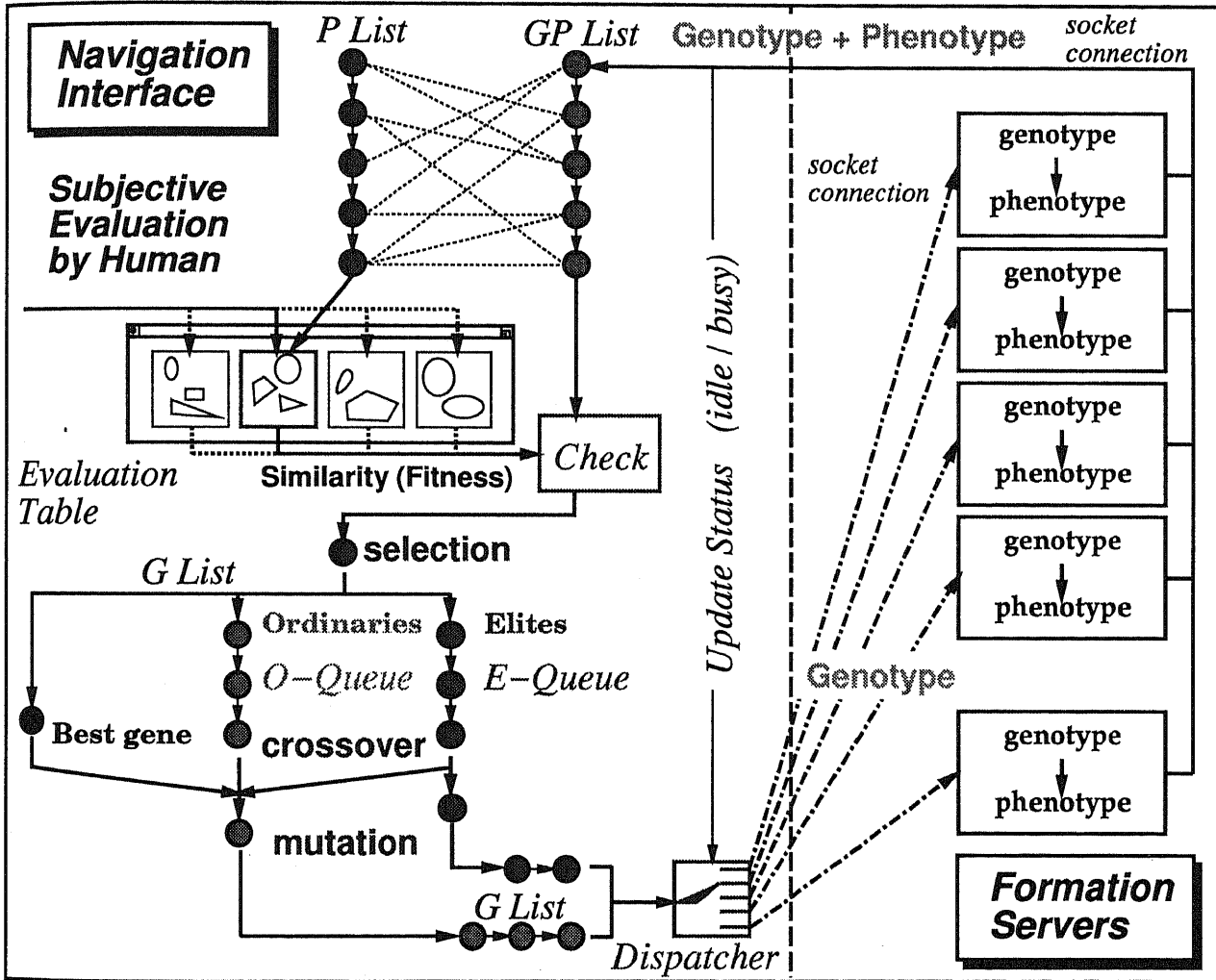


図 9.3: ライフサイクル型遺伝的アルゴリズムの模式図。点線を境にインタフェース部とサーバ部の2つの部分に分かれる。

5. P List の先頭から順番に評価テーブルに表示する。評価テーブルは同時に V 個の表現型を表示できる。
6. ユーザは評価テーブルに表示された表現型に一個ずつ評価を与える。
7. ある個体のすべての表現型が評価を受ければ、その評価を収集して個体の適応度とし 1 へ。

ここで遺伝的操作とは、後に述べる先着順交叉と突然変異を指すものである。

ここで GP List と P List の違いについて述べる。もしある個体の表現型が一つならば、GP List と P List は一致し別々に管理する必要はない。しかし本研究の類似検索問題では、個体の表現型は類似検索で上位に検索された複数の画像であるため、表現型のリストと個体のリストとを別々に管理する必要があるのである。また逆に 1 つの表現型には複数の個体が関係するため、1 つの表現型を評価した結果は複数の個体の適応度に影響を及ぼすことになる。

9.4.3 ライフサイクル型遺伝的アルゴリズムのオペレータなど

1. 表現型の評価

表現型の評価は評価テーブル上でおこなう。ただし P List にある表現型すべてを一度に見せるという方法は、表示スペースの問題や人間が把握できる制約の問題から現実的な方法ではない。そこで P List の先頭から V 枚の表現型を評価テーブルに表示することで、ユーザはその中の任意の 1 個に評価を与えるという方法を用いる。評価を与えられた画像は評価テーブルから消去され、P List の先頭の画像が新たに評価テーブルに表示される。ここで注意すべきことは、このシステムでは評価順序が重要なことである。たとえ同じ評価を与えたとしても評価順序によって交叉の順番が変わるからである。

2. 移動平均適応度比例戦略

このモデルには世代という概念がないため、G-GA で用いられてきた適応度比例戦略を、L-GA に合うように改良する必要がある。そこで、ここでは平均適応度として直前 M 個体の移動平均適応度 \bar{f} を採用する。そしてある適応度 f をもつ個体は、子孫として

$$N = \frac{f}{\bar{f}} = \frac{f}{(\sum_{i=1}^M f_i)/M} \quad (9.1)$$

個の子孫を残すという生殖ルールを用いる。ただし N は小数になるので、実際は子孫の数を $[N]$ か $\lceil N \rceil$ のどちらかに乱数で決定する。そして $N \geq 1$ であれば、その子孫の数だけ交叉待ち行列に同じ個体を並ばせる。なお適応度にはべき乗スケールリングを施した [117]。

一般的には探索が進むにつれて高い適応度を得る個体が増加するので、平均適応度はゆるやかに上昇し N の期待値は 1 以上となる。したがって全体の個体数は探索が進むにつれて徐々に増加していく。しかし L-GA では世代という概念がなくすべての個体をリストで管理しているため、全体の個体数がいくらかは気にする必要がない。

3. 交叉待ち行列

有望な個体をできるだけ多く繁殖させることが遺伝的アルゴリズムの性能を向上させる基本的な戦略である。そこで本モデルでは交叉待ち行列を2本用意し、エリート個体交叉待ち行列 E-Queue (elite queue)、平凡個体交叉待ち行列 O-Queue (ordinary queue) と名付ける。具体的には移動平均 \bar{f} をしきい値として、それを上回る個体を E-Queue に入れ、それ以外を O-Queue へ入れる。今後の遺伝的操作では、基本的に E-Queue に入っている個体を優先する。

4. 先着順交叉

先着順交叉 (First Come First Crossover) とは、交叉待ち行列の先頭にある個体から順番に交叉させる方法である。本研究では2本の交叉待ち行列の間に4種類の交叉を考える。

1. E-Queue 内で先頭から2個体ずつ交叉
2. E-Queue と O-Queue の先頭から1個体ずつの交叉
3. O-Queue 内で先頭から2個体ずつの交叉
4. O-Queue の先頭と最良個体との交叉

ここで交叉させる2個体が同一のものである場合は、異なる個体に出会うまで待ち行列をたどることにする。この交叉をこの順番でおこなう。まず1)と2)をE-Queueに交叉できる個体なくなるまで続け、さらに3)をO-Queueに交叉できる個体なくなるまで続ける。いずれの処理にも交叉率が設定されている。またさらに探索能力を高めるために、O-Queueの交叉を工夫する。先述したようにO-Queueは平凡な個体の集まりなので、O-Queue内での交叉にはあまり期待できない。そこで3)では、O-Queue内の個体とその時点での最良個体とを確率的に交叉させることにする。これが4)であり、最良個体とO-Queue内の遺伝子距離が大きいほど最良個体との交叉確率が高くなるように設定した。

5. 突然変異

交叉を受けた個体はただちに突然変異を受ける。この突然変異は通常遺伝的アルゴリズムと同様である。ただしO-Queueの平凡な個体には高めの突然変異率を設定して大域的探索に向かわせる。

6. 個体のサーバへの送信

交叉と突然変異の遺伝的操作が完了した個体をサーバに送信する際、有望な個体をより多く繁殖させるとの基本方針に基づいて、E-Queueの送信を優先させる。E-Queueを送信後もなおサーバが空いていればO-Queueの先頭からも個体を送信する。

7. 個体の寿命

インタフェースから誕生してサーバに送られ、再びインタフェースに里帰りしてくるまでの時間の長さをその個体の「寿命」と呼ぶ。するとその寿命とは、サーバ側のコンピュータ(同一の場合もある)の負荷の状態に依存するし、またサーバとインタフェースの通信路の混雑状況に

も依存する。このように個体の寿命の長さは、自然環境の変動にも似たこのような環境変動に左右されることになる。

8. 遺伝子の多様性の維持

遺伝子の多様性を維持することは、遺伝的アルゴリズムが局所解につかまらないためには非常に重要な問題である。L-GA では以下の方法により、多様性を維持することを目論んでいる。まず評価テーブル上でユーザがランダムな順番で画像を評価することにより、GP List 上で遺伝子の順序が逆転する⁵。次に個体の寿命が環境変動に左右される効果によりやはり遺伝子の順序が逆転する。このようにL-GA では2通りの効果によって、遺伝子が常に同じ順序でライフサイクルを送るわけではない。こうして従来のような人工的な乱数を用いることなしに遺伝子をかき混ぜることができる点が独特である。

また個体数を無理に一定数に保たずに増やしていくことでO-Queueの長さは徐々に長くなる。このことから、E-Queueの個体が短いライフサイクルを繰り返すのに比べて、O-Queueの個体は長いライフサイクルを繰り返すことになる。すなわちE-Queue個体の側から見ると、O-Queue内には古くて平凡な遺伝子が保存されていることを意味し、これが遺伝子プールの多様性を維持する効果を持っている。つまりエリート主義を用いながらも同時に多様性を維持する効果ももつ機構を実現しているといえる。

9.4.4 ライフサイクル型遺伝的アルゴリズムの類似検索規準の学習への応用

以上のように、ライフサイクル型遺伝的アルゴリズムに特有のオペレータなどを説明してきた。ここではライフサイクル型遺伝的アルゴリズムを、類似検索規準の学習に応用する方法を示す。

遺伝子型

本研究で用いる遺伝子型は13個の重み係数が並んだものである⁶。この重み係数は連続変数であるため、遺伝的アルゴリズムを用いて解く場合には、実数表現を用いるか、量子化してビット列表現を用いる、という2つの選択肢がある。もし実数表現を選ぶならば $0 < w < 1$ の実数値をそのまま遺伝子にできるが、ビット列表現の方では探索の精密さと探索空間の広がりとのトレードオフを考慮して重み係数を量子化する必要がある。一つの重み係数に8ビットを割り当てた実験と実数表現の実験とを比較した結果、実数表現の方がよい解を安定して発見できたため、今後の実験では実数表現を用いる。

遺伝的操作

実数表現に関する遺伝的操作の具体的な手法やパラメータの値は、文献[118]に基づいた。

ソフトウェア

⁵表現型が一つの場合にはこの効果はない。

⁶この実験で用いた重み係数の個数は第10章の実験とは異なっていることに注意。

NAVGA : Status Monitor

File				Help
1	2	3	4	
337	305	252	252	
5	6	7	8	
194	195	179	179	
9	10	11	12	
155	142	133	132	
13	14	15	16	
132	112	99	118	
2932				

NAVGA : Query Window

File Query Search Help

Query

NAVGA : Best Gene

File Weights

Fitness: 2.65500

Genotype: $e+00$ 9.701915e-01 0.000000e+00 6.445914e

Phenotype

0 1

6.420241 6.612565

080403 080207

NAVGA : Navigation Interface

File Show Help

.63 .60 .00 .00 .00

Evaluated 481 : no
 Evaluated 482 : no
 Evaluated 483 : no
 Evaluated 484 : no
 Evaluated 485 : no
 Evaluated 486 : no
 Evaluated 487 : no
 Evaluated 488 : no
 Evaluated 489 : no
 Evaluated 490 : no
 Evaluated 491 : no
 Evaluated 492 : no
 Evaluated 493 : no
 Evaluated 494 : no
 Evaluated 495 : no
 Evaluated 496 : no
 Evaluated 497 : no
 Evaluated 498 : no
 Evaluated 499 : no

図 9.4: ライフサイクル型遺伝的アルゴリズムによる例示検索・ナビゲーションのインタフェース。左下が評価テーブル。

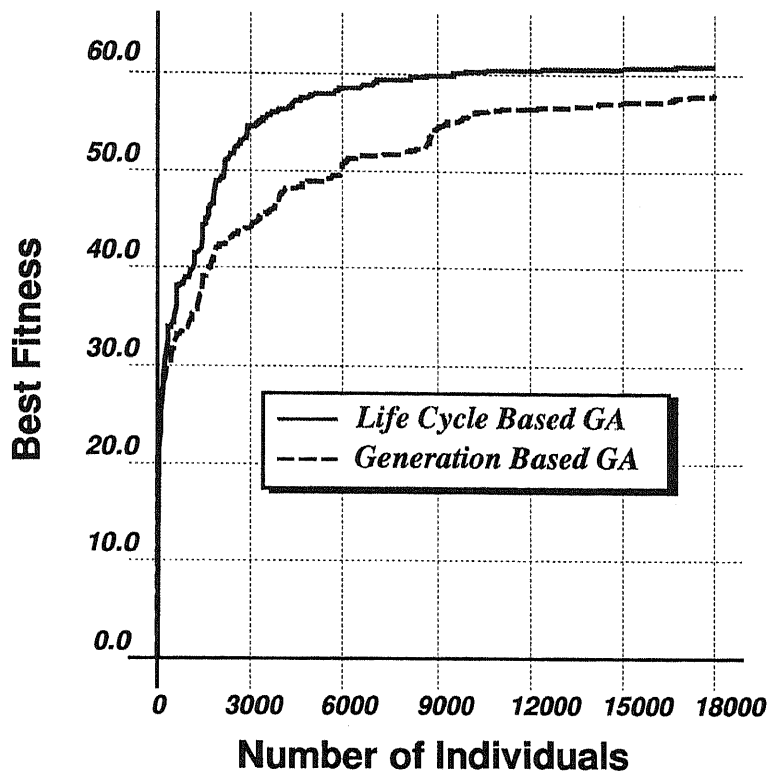


図 9.5: 世代型遺伝的アルゴリズムとライフサイクル型遺伝的アルゴリズムの場合の最大適応度の比較。

図 9.4 のインタフェースをもつプログラムを、UNIX 上の X/Motif を用いて作成した。またサーバとインタフェースとの通信には UNIX のソケット機能を用い、サーバを分散環境で動作させることを考慮した設計となっている。特に本研究では並列処理の効果を活かすため、並列計算機 IBM SP-2 を用いて 16 個のサーバを並列に動作させた。負荷分散の方法は、16 個のサーバを順にチェックし、最初に発見されたアイドル状態のサーバに仕事を任せるといった単純な方法を用いた。

9.5 実験結果

以下ではオフライン問題とオンライン問題の実験結果を説明する。なおこの実験で得られた重み係数を用いた類似検索の実験については、第 10.3 節で述べる⁷。

9.5.1 オフライン問題

オフライン問題では教師データとして主観的類似画像が必要である。ここで用いた主観的類似画像は第 10.2.1 節に述べる方法で選んだ画像であり、29 枚の画像に対して 1 ~ 32 の主観的類似度が与えられている。これを教師データとした最適化問題に、世代型遺伝的アルゴリズムとライフサイクル型遺伝的アルゴリズムの両方の手法を適用し、その性能などを比較する。ただしライフサイクル型遺伝的アルゴリズムでは、人間が実際に類似度を評価する過程をエミュレートする。これは人間が先述の主観的類似度を固定して与えるものとし、また人間の 1 回の評価につ

⁷ただし厳密には、この実験をおこなった当時と現在とでは重み係数の種類が異なっている。

き、画像の注視時間に相当するものとして1秒の待ち時間を設ける方法である。この待ち時間を設けることによってサーバ側の計算時間(約1~2秒)と評価時間(1秒)との釣り合いがとれ、パイプライン型の処理に適した状況となる。一方の世代型遺伝的アルゴリズムでは、個体数30、エリート主義、また交叉率などのパラメータにはよく用いられる値を用いた。

この2種類の手法を比較するために、定期的に最大適応度の変化を記録し、最適解への探索の速さを比較する。世代型遺伝的アルゴリズムでは1世代ごとに最大適応度を記録すればよいが、ライフサイクル型遺伝的アルゴリズムでは、世代型遺伝的アルゴリズムの1世代相当、つまり評価を受けた個体の30個体ごとに最大適応度を記録した。

図9.5には最大適応度の変化を示す。ここでは10回の試行を平均した記録を示した。この図からライフサイクル型遺伝的アルゴリズムの方が世代型遺伝的アルゴリズムより探索性能に優れていること、また18000個体経過時の最適解も上回っていることがわかる。特に初期において優れているのは、この方法がかなり強いエリート主義になっていることが原因である。しかし同時に遺伝子の多様性も保持しているので、局所解につかまる可能性は低い。

以上の結果から、ライフサイクル型遺伝的アルゴリズムは世代型遺伝的アルゴリズムに比べても、性能の面で遜色はないという結論が得られた。ただし、サーバ側とインタフェース側との通信に時間がかかるため、簡単な問題ではL-GAの方が計算時間がかかり遅くなる。よって実用的に用いるときには、通信に要する時間が無視できる本研究のような応用に適したアルゴリズムである。

9.5.2 オンライン問題

実際に図9.4のインタフェースを用いて、ブラウジングにより重み係数の最適化をおこなった。ライフサイクル型遺伝的アルゴリズムを用いることにより評価のためのユーザの待ち時間はゼロとなり、ユーザの負担が少ない効率的なブラウジングが可能になったことがわかった。この点で従来の世代型遺伝的アルゴリズムよりは、はるかにブラウジングに適した手法であるといえる。ただし重み係数をまったくランダムに設定し、そこから最適解を得るためには、かなりの評価回数を必要とし、事実上かなり困難であると言わざるを得ない。よってこの手法を用いる時には、まずオフライン問題を解いて大局的に最適な重み係数を得ておき、ブラウジングはこの重み係数の周辺を局所的に探索するために用いるのが最適な組み合わせであると考えている。

9.6 まとめ

本章では対話的なマンマシンインタフェースとして有望な「人為選択法」に適した遺伝的アルゴリズムとして、個体のライフサイクルを重視した手法である「ライフサイクル型遺伝的アルゴリズム」を提案し、このモデルを類似検索規準の学習問題に応用した。この手法は世代型遺伝的アルゴリズムに比べて、おのおのの処理をパイプライン的におこなえるという意味で人為選択法により適した遺伝的アルゴリズムであり、またアルゴリズムの動作自体も興味深い性質を持っている。このアルゴリズムを類似検索規準の学習問題に応用した結果、探索性能では既存の世代型遺伝的アルゴリズムよりも優れていることを示した。

人為選択法は新しいマンマシンインタフェース技術として有望な技術である。ただし人間の忍耐力を考えると、実際に画像を評価する回数はできるだけ減らしたいのが本音である。そこで、さらに賢い学習能力を備えて人間の意図を汲み取ることができる技術が求められてくるだろ

う。また結果のところでも述べたように、遺伝的アルゴリズムをオフライン問題とオンライン問題の2段階に分けて適用することにより、画像データベースにユーザの負担が軽くて柔軟な類似検索のためのインタフェースを提供できると考えている。

第10章

類似画像検索

10.1 概要

これまでの章で、類似画像検索を実現するためのさまざまな要素技術の提案と階層モデルとの関連について述べてきた。本章ではいよいよ類似画像検索システムを実際に構築し、類似画像検索の実験をおこなう。システムに例示画を示し、システムがそれに類似した画像を検索する類似画像検索の実験を通して、本研究が提案するシステムの有効性を検証する。

本研究では第2.1節で述べたように、3種類の画像データベース — NOAA衛星画像データベース、GMS衛星画像データベース、カラー自然画像データベース — を研究対象とする。最初の二つは衛星画像データベースであるが、最後のカラー自然画像データベースはCD-ROM化された写真集を用いたものである。このように多種類の画像を扱う場合には、いかに画像データベースを構築するかという方法論が重要となってくる。つまりこのような3種類の画像では、それぞれの画像が表現している内容は異なり、また撮影方法が異なるために最適な画像処理方法も異なっている。したがって、このすべての画像に共通する要素は何かを見極め、共通する部分は共用し、共通しない部分は個別に実現していくという方法論が必要となる。

このような方法論として、第2.4節で提案した階層モデルが有用である。つまりこの階層モデルによって問題を整理することで、どの画像データベースを構築する場合にも一定の枠組を考えることができる。また問題の解決に最適なアルゴリズムを決定する問題は、階層モデルにしたがって各レイヤごとに最適な方法を決定する部分問題に分割することができるため、アルゴリズムの取捨選択が見通しよく簡単におこなえる。このような利点をもつ階層モデルに基づき、本章では類似画像検索システムを実現する。

まず最初に第10.2節で、階層モデルに基づいた画像データベースシステムの実装について述べる。3種類の画像データベースで実際にどのように各提案手法を使い分けたかを述べ、その理由についてもできるだけ明確に記述する。第10.3節では類似検索の結果を示し、本研究で提案した手法が類似検索に有効であることを示す。第10.4節では類似検索結果の評価について議論し、最後に第10.5節で本章のまとめとする。

10.2 階層モデルに基づいた画像データベースシステムの実装

階層モデルと本研究で用いた手法の対応関係については、第2.4.3節および表2.3にすでにまとめてある。しかしここまでに本研究の提案手法をいろいろと述べてきたことから、ここでもう一度、各画像データベースの実装に用いられている個別の手法の詳細な内容をまとめる。

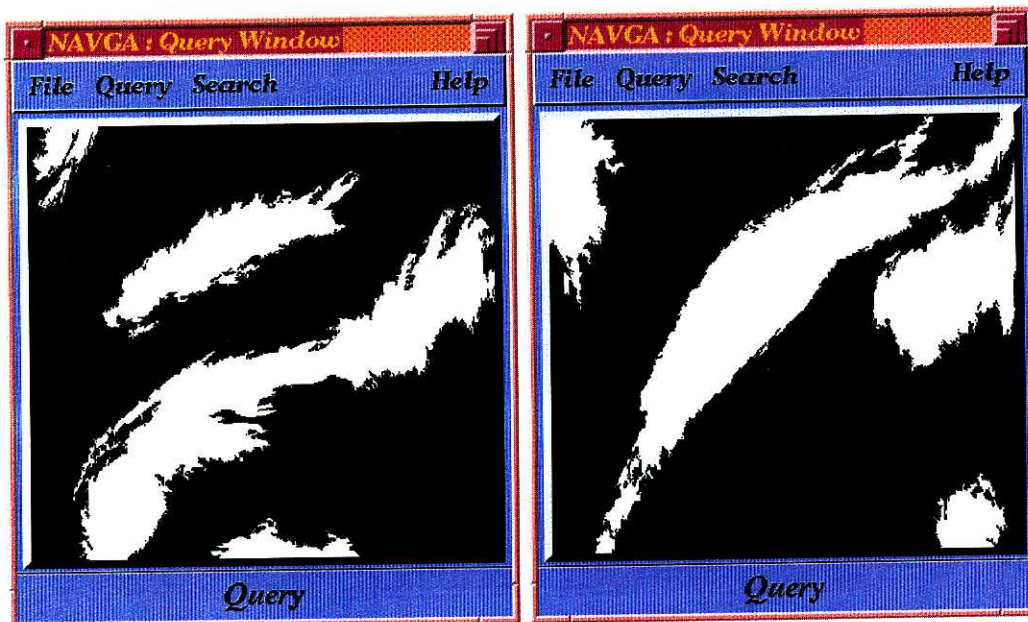


図 10.1: NOAA 衛星画像データベースで用いる例示画。

10.2.1 NOAA 衛星画像の類似検索システム

画素レイヤ

NOAA 衛星画像データベースには NOAA-AVHRR センサで観測された赤外画像 (Ch.4) を用いているため、本来は第3章～第5章で述べたミクセルを含めた統計的分類手法により雲域を抽出すべきである。しかしデータの入手の都合上、本研究では別の方法であらかじめ2値化された雲画像を検索対象とする。そのため画素レイヤの処理は省略している。

領域レイヤ

NOAA 衛星画像では超楕円による形状分解を用いている。2値画像に対するエネルギー関数を用いた形状分解の手法により、ラベリングされた領域ごとに形状分解をおこなう。

関係レイヤ

プリミティブ間の関係の構築および領域間の関係の構築は第7章で述べた平面モデルの手法で構築した。この場合の階層化属性つき関係グラフは24個の重み係数をもっている。

認識レイヤ

階層化属性つき関係グラフのグラフマッチングによって画像間類似度を計算する。このレイヤは画像の種類によらず共通の計算方法である。

理解レイヤ

ここでは重み係数の学習の方法を述べる。重み係数の学習には第9章で述べたように遺伝的アルゴリズムを用いるが、その実験の設定にいくつかのポイントが含まれている。

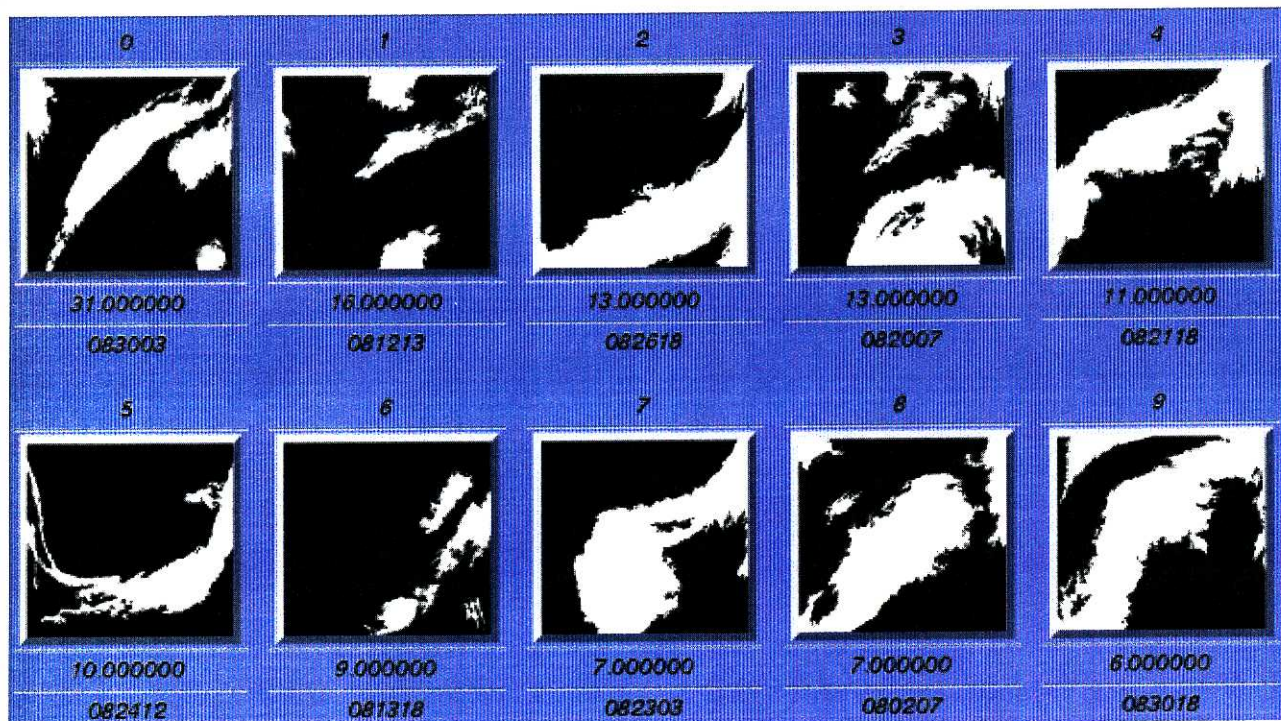


図 10.2: 被験者によって選ばれた主観的類似画像 29 枚の中の上位 10 枚の画像。主観的類似度が画像の下の数字で示されている。

学習用サブセットの選択 まず画像データベースに蓄積された全画像 (1027 枚) の中から、重み係数を学習するための画像を選び出す。この方法は実際の画像データベースの運用形態を考慮した方法である。画像データベースに蓄積されたすべての画像を対象として学習をおこなえば、重み係数にはそれだけ最適な解が得られることが期待できる。しかしこの方法は以下の二つの理由により現実的ではない。

1. 大量の画像を蓄積した画像データベースでは、学習に要する計算量が大きいという面から現実的ではない。
2. 蓄積画像が日々増加していく画像データベースでは、ある時点での全画像が別の時点では一部に過ぎないという面から現実的ではない。

そこで全部の画像の中から少数の画像をうまく抽出してそれを学習用サブセットとして用いる。この際の抽出方法は、理想的には蓄積された画像集合の特性を漏れなく反映した部分集合であるべきであり、統計的な偏りを防ぐためにはある種の無作為抽出法などを必要とする。しかし本研究では簡単のため、本研究で例示画として用いる画像図 10.2.1 が含まれる 1991 年 8 月に撮影された画像 163 枚を学習用サブセットとして選び出した。そして重み係数はこの 163 枚の中で類似検索を繰り返すことを通して学習される。

教師データの選択と主観的類似度の付与 さて次の問題は、最適な重み係数を得るための教師データの選び方である。そこでこの画像データベースの場合には、被験者に主観的類似画像を前もって選んでもらう方法を用いる。すなわち学習用サブセットである 163 枚の画像を被験者 10 人に見せ、各被験者はその中から主観的類似画像 5 枚を選択する。またその 5 枚の画像に合わせて類似順位をつけてもらい、この類似順位を主観的類似度とみなす。このようにして定義した類

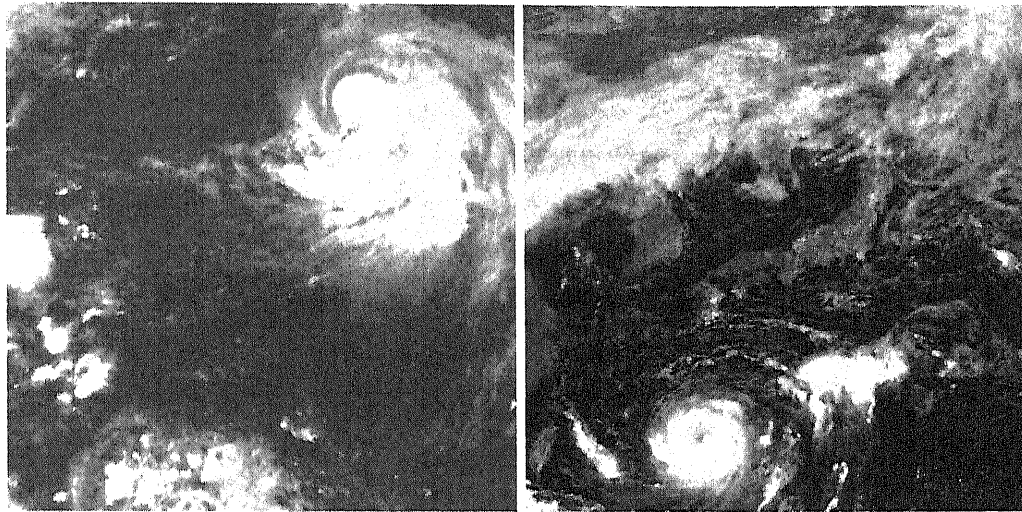


図 10.3: GMS 衛星画像データベースで用いる例示画。左が短期型データベース用、右が長期型データベース用の画像。

似度は、第 A.3.3 節の定義によれば順序尺度を間隔尺度とみなした距離であると考えればよい。10 人の被験者からそれぞれ得られた尺度値を合計することによって、被験者の主観を平均した主観的類似度が得られる。こうして類似画像として選ばれた画像は、全部で 29 枚でありその類似度は 1 ~ 32 となった。その中の類似上位 10 枚の画像を主観的類似度とともに図 10.2 に示す。

この主観的類似画像を、できるだけ上位に検索できるように重み係数を最適化する問題が、この画像データベースに関する類似検索規準の学習問題である。ただし注意すべきことは、被験者が類似画像を選ぶ際には、どのような観点から類似画像を選ぶかという類似規準を特に指定しなかった点である。こうして、主観的類似画像として選ばれた画像には、明らかに類似規準の異なった画像が複数のグループで含まれることになった。このように類似規準が異なることによる影響を考慮した実験は第 10.3.1 節の実験 4 でおこなわれている。

ちなみに本研究で用いる例示画は図 10.2.1 の画像である。この画像の最も顕著な特徴は、南西から北東に伸びる、台風に伴う大きな雲の形状であり、その北西側に平行して並ぶ一本の雲の帯も特徴的である。このような特徴を捉えた画像を検索するのが、この場合の類似検索の目的であり、類似検索規準である。

10.2.2 GMS 衛星画像の類似検索システム

第 2.2.3 節で述べたように、本研究で構築する GMS 衛星画像データベースは、短期型と長期型の 2 種類がある。

画素レイヤ

短期型画像データベースと長期型画像データベースでは用いる方法が異なる。その最も大きな理由は、短期型に用いられる画像が赤外画像であるのに対し、長期型に用いられる画像が可視画像であるという点である。赤外画像が物体の放射温度を測定するのに対し、可視画像は物体の反射率を測定するものである。よって観測される画像の統計的性質も当然異なるものとなる。例えば赤外画像では、画像の南北で温度勾配がある場合には画像の両端で画素の統計的性質は異

なるのに対し、可視画像では画像の南北や東西で太陽高度が異なる影響で画素の統計的性質が異なってくる。このように同じ衛星画像であっても、画像によって異なる扱いが適切となるのである。

そこでこの画像データベースの場合は、まず短期型では第3章～第5章までに述べた、ミクセルを含む混合密度推定による画素単位分類手法を用いて雲域を抽出する。それに対して長期型の場合は、第B.2節で述べる、スケールスペースの1次微分ゼロ交差を用いたしきい値処理によって雲域を抽出する。雲領域は海領域に比べて反射率が大きいいため、画像ヒストグラムでは画素値が大きい方が雲となる。よって画像ヒストグラム上で画素値が大きい方から数えて最初のピークは雲画素から形成されるピークであると考え、その次の谷部分をしきい値とする方法を用いる。もっとも、あるはずのピークが他のピークに埋もれてしまう場合もあるため、しきい値を許容する範囲を前もって定めておく必要がある。

このように二つの画像データベースで異なる手法を適用した。そのためこのレイヤの出力表現も両者で異なり、短期型の方は第4節で述べた期待面積占有率曲線であるのに対し、長期型の方は雲領域を分割するしきい値となっている。

領域レイヤ

領域の構造的な特徴は、第6.3.4節で述べた濃淡画像を対象とした形状分解の手法を用いる。この手法を適用するためにはメンバーシップ関数を定める必要があるが、本研究では短期型の場合は画素レイヤで得られた期待面積占有率曲線を、第6.3.4節の方法でそのままメンバーシップ関数に変換して用いた。一方の長期型の場合には、画素レイヤではしきい値のみが得られているため、第6.3.4節の線形モデルを用いてメンバーシップ関数を計算する。線形のメンバーシップ関数の幅は一定の値として適当に定めればよいが、可視画像のダイナミックレンジが64段階であることを考えて、本研究では $w = 3$ と設定している。なお形状分解の基本構成要素としては超楕円と楕円の場合を比較したが、より自然な分解結果を与える楕円を基本構成要素として用いることにした。

関係レイヤ

濃淡画像の形状分解の結果では、大部分のプリミティブどうしで隣接関係による関係は定義できない。そこで何らかの遠隔作用をモデル化した手法によって関係を構築する必要性が生じるため、この画像データベースの場合には第7.3節で説明した階層的な引力モデルを用いる。このレイヤの出力表現は階層化属性付き関係グラフであり、この時点で短期型データベースも長期型データベースも同一の表現をもつようになる。なおこの場合の重み係数は23個である。

認識レイヤ

階層化属性つき関係グラフのグラフマッチングによって画像間類似度を計算する。このレイヤは画像の種類によらず共通の計算方法である。

理解レイヤ

このレイヤの問題は、最適な重み係数を求めるための教師データの設定である。この画像データベースの場合に用いる例示画は図10.3に示した画像であり、この画像に類似した特徴をもつ画

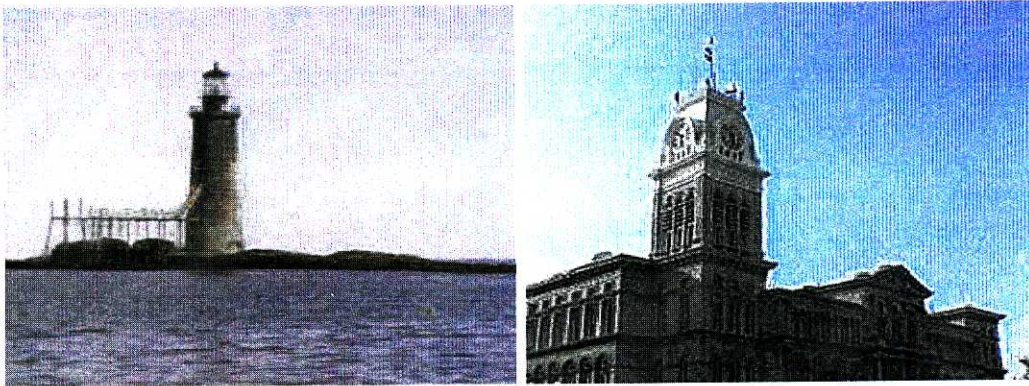


図 10.4: カラー自然画像データベースで用いる例示画。

像を検索することが目的である。以下では短期型と長期型にわけて、実験の設定上のポイントを述べる。

短期型画像データベース この画像データベースには興味深い性質がある。すなわち、短期型では衛星が毎時観測したデータをすべて蓄積しているため、連続した観測で得られる同一領域の画像はお互いに似通った雲パターンを示すことが予想されることである。つまりある例示画に最も類似した画像はその前後の観測画像となるはずであり、しかも観測時間が離れるに従って類似度が小さくなると仮定できる。このような仮定に基づき、教師データとしては例示画の観測直前 10 枚と観測直後 10 枚の計 20 枚の画像を選び、前後の観測画像に類似尺度 10、それから遠くなるに従って類似尺度を 1 ずつ減じるという方法で教師データに類似度を与えた。

このようにして選んだ教師データは、NOAA 衛星画像で用いたような主観的類似画像ではなく、「客観的類似画像」とでも呼ぶべきものである。ここでの教師データの選択規準は画像の観測日時という客観的な規準であり、この選択に人間の主観が介在する余地はない。したがってこのような教師データを用いて類似検索した結果も、やはり純粹に客観的に評価することが可能である。このように客観的に評価することのできる例は類似検索には珍しい。例示画には図 10.3 に示すように、台風が画像の北東方向に写っている点が特徴的な画像を選んだ。

長期型画像データベース 長期型の場合には前後の観測で 24 時間もの間隔が開いていることから、前後の観測で雲パターンが類似しているはず、との仮定はもはや成立しない。そのため、やはり人間が前もって(主観的に)選択した教師データを与える必要がある。例示画には図 10.3 で示すように、太平洋南海上に台風による丸い雲の渦があり、また日本付近には東西に雲が広がっているという、梅雨から夏にかけての典型的な気象状況を表した画像を用いる [34]。したがって、そのような気象状況を満たしている画像を全画像 (331 枚) の中から 12 枚選びだし、これらに適当な類似尺度を与えて教師データとした。

なおどちらの場合も画像データベースに蓄積された画像数がそれほど多くはないため、学習用サブセットは蓄積画像集合と同一のものとした。

10.2.3 カラー自然画像の類似検索システム

画素レイヤ

カラー自然画像の領域分割の問題は、従来から幅広く研究されてはいるものの、依然として困難な問題であることには変わりがない。そこでこの画像データベースの処理では、従来からカラー画像の分割に頻繁に用いられてきた Ohlander の再帰的しきい値処理の考え方に基づいた方法を用いる。その詳しい説明については第 B.2.2 節に述べる。

領域レイヤ

衛星画像に階層モデルを適用した場合には、このレイヤではプリミティブによる形状分解を用いて、領域の構造的特徴を表現していた。しかしこの場合のように自然画像を対象として形状分解をおこなった結果では、直観に合うような結果が得られなかった。このように形状分解がうまく働かなかった原因は、領域分割された各領域の形状が、単純な形のプリミティブの集合では近似しにくい形状となるからであると考えられる。本質的にこの問題を解決するには画素レイヤの改良が欠かせないため、カラー画像の場合の領域レイヤでは、形状分解をせずに種々の領域特徴量を測定するだけでこのレイヤの処理を終えることとする。これらの特徴量では領域の構造を記述するには力不足であるが、しかしある程度の精度での記述ならば可能である。

関係レイヤ

領域レイヤで形状分解をおこなわなかったことにより、このレイヤの出力表現であるグラフ構造は 1 層のグラフ構造となる。そして領域間の関係の構築は、隣接関係を用いたグラフ構造への構造化手法を用いている。この手法については第 7.5.2 節に述べた。

認識レイヤ

階層化属性つき関係グラフのグラフマッチングによって画像間類似度を計算する。このレイヤは画像の種類によらず共通の計算方法である。

理解レイヤ

本研究で用いた例示画を図 10.4 に示す。この左側の図が最初の検索で用いる例示画であり、海に浮かぶ島にそびえ立つ灯台を撮影した写真である。右側の図は 2 番目の検索に用いられる例示画であり、建築物の写真となっている。この場合の教師データは、この画像に類似した画像をフォトコレクションから (主観的に) 選び出す。他に灯台が撮影されている数枚の写真や、写っている建物は異なるが色や空間的配置が類似している 9 枚の画像を選びだし、それらに適当な類似尺度を与えることにより教師データとした。

なお学習用サブセットとして全体の 1/4 に相当する 239 枚の画像を選んだ。これは、全くランダムに抽出した画像に教師データを加えた画像集合となっている。

10.3 類似検索結果

以下では、本研究で扱う 3 種類の画像データベースそれぞれについて類似検索実験をおこなった結果をまとめている。

10.3.1 NOAA 衛星画像データベースの類似検索結果

4種類の実験の内容

この画像データベースに対しては以下の4種類の実験をおこなう。

1. 学習用サブセット 163 枚に対して、例示画を示して類似検索をおこなう実験
2. 全画像 1027 枚に対して、例示画を示して類似検索をおこなう実験
3. 全画像 1027 枚に対して、例示画と類似はしているが異なる画像を新たな例示画として類似検索をおこなう実験
4. 実験 1 で上位に検索されなかった主観的類似画像だけで再学習した重み係数を用いて、学習用サブセット 163 枚に対して類似検索をおこなう実験

どの実験でも、第 9 章で述べたような遺伝的アルゴリズムの手法を用いて最適化された重み係数を用いる。実験 1～実験 3 には表 10.1 の重み係数、また実験 4 には表 10.2 の重み係数を用いている。

実験 1 の結果

この実験の目的は、表 10.1 のように遺伝的アルゴリズムを用いて学習されたこの重み係数を用いることによって、被験者が前もって選んだ主観的類似画像を上位に類似検索できることを確認する点にある。実験 1 の結果を図 10.5 に示す。実験 1 では学習用サブセットを対象として類似検索をおこなったため、重み係数は検索対象に対してほぼ最適な値となっている。そこで表 10.3 では、類似検索順位と主観的類似度との対応をまとめた。主観的類似度で 1 番類似度が大きい画像が類似検索でも 1 位に検索されており、またその他の上位検索画像も主観的類似度は大きいものが多く選ばれている。また目視によっても、特に雲パターンの主軸方向が類似した画像が上位に検索されていることがわかる。このように、人間が前もって選んだ主観的類似画像に対応する重み係数を、最適化の計算によって学習することができる、ということがこの実験の結論である。

実験 2 の結果

実験 2 では実験 1 に比べて検索対象とする画像集合を変えている。あらかじめ学習用サブセットを対象にして学習した重み係数が、全画像に対してどれほど有効に働くかを確認する実験である。もちろん対象画像が増えたわけであるから、新しい画像集合に対しては学習された重み係数は準最適な重み係数であしかない。そこでこの学習された重み係数が新しい画像集合に対してどのくらい通用するのかという汎化能力を確認するのが、この実験の目的である。その結果を図 10.6 に示す。この場合は、実験 1 で類似検索順位が 1 位だった画像が 4 位に検索されており、また 5 位と 9 位の画像も実験 1 で検索されていた画像である。またその他の画像もやはり雲パターンの主軸方向が類似しているという意味で類似した画像が上位に検索されている。これらの結果から、類似検索のサブセットによって学習した重み係数を全画像に適用してもよい検索結果が得られるという結論を得た。

表 10.1: NOAA 衛星画像データベースの検索に用いられた重み係数 (1) の一覧。遺伝的アルゴリズムによって最適化された結果である。

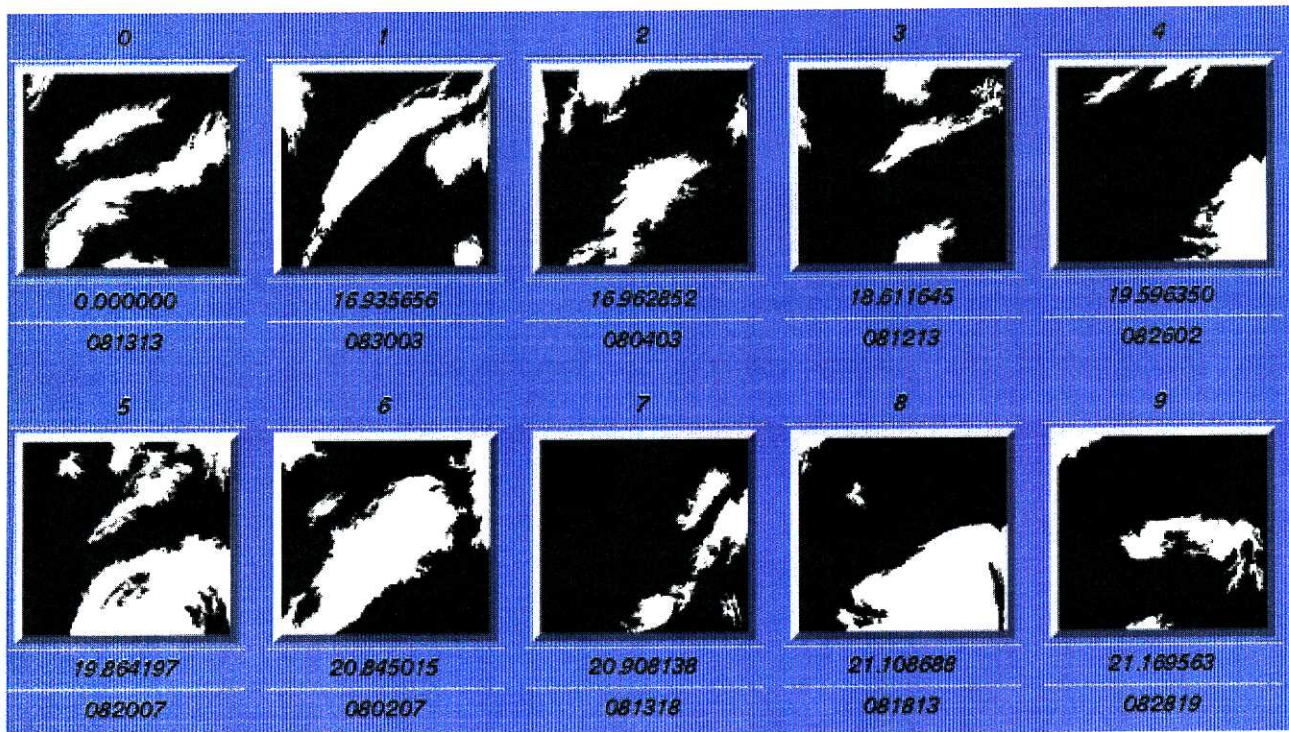
重み係数	重みの値	重み係数	重みの値
ノード代替コスト (上位)	0.2371	アーク代替コスト (上位)	0.0783
ノード代替コスト (下位)	0.7458	アーク代替コスト (下位)	0.7015
ノードスケール (上位)	0.8604	アークスケール (上位)	0.0000
ノードスケール (下位)	0.1332	アークスケール (下位)	0.0000
ノード位置 (上位)	0.4279		
ノード位置 (下位)	0.0000		
ノード主軸方向 (上位)	0.3737	アーク方向 (上位)	0.4327
ノード主軸方向 (下位)	0.8828	アーク方向 (下位)	0.6496
ノード縦横比 (上位)	0.0000	アーク端点領域面積比 (上位)	0.0000
ノード縦横比 (下位)	0.9424	アーク端点領域面積比 (下位)	0.9976
ノード Compactness (上位)	0.5678	アーク端点領域角度差 (上位)	0.4330
ノード円形度 (上位)	0.0000	アーク端点領域角度差 (下位)	0.8663
ヌルシンボル係数 (上位)	0.0568		
ヌルシンボル係数 (下位)	0.0297		

表 10.2: NOAA 衛星画像データベースの検索に用いられた重み係数 (2) の一覧。遺伝的アルゴリズムによって最適化された結果である。

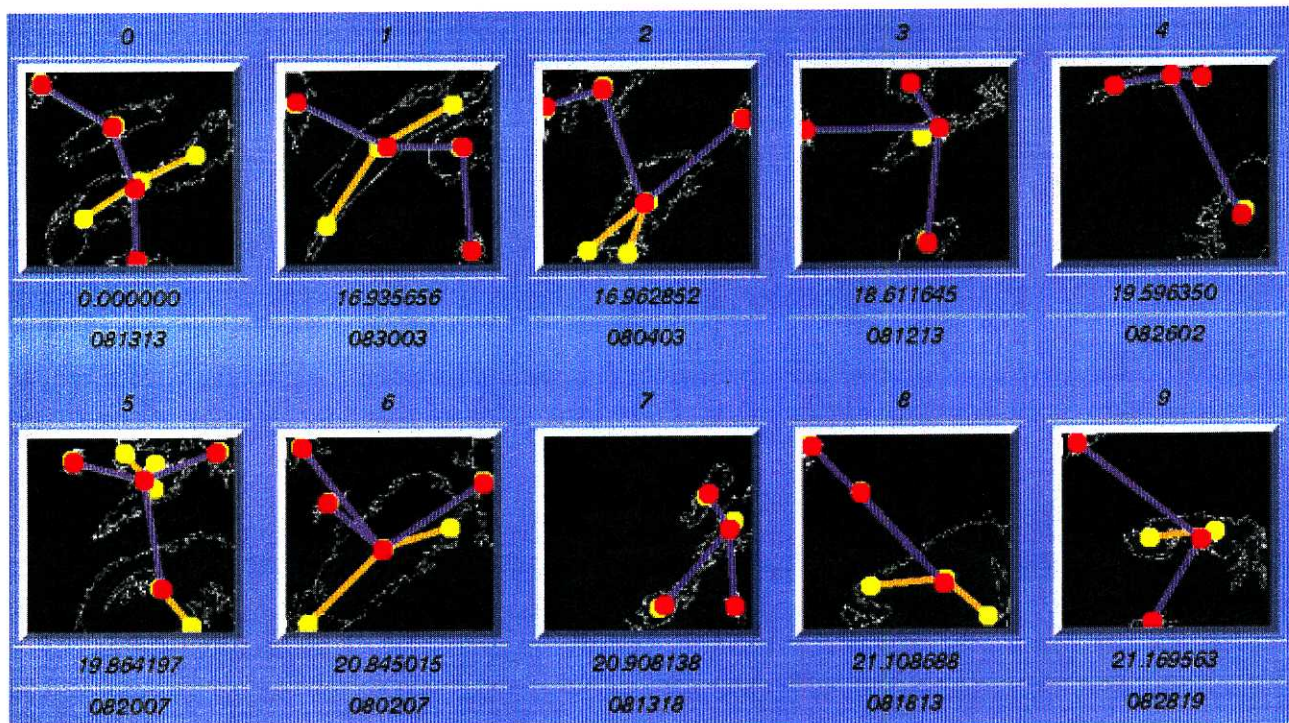
重み係数	重みの値	重み係数	重みの値
ノード代替コスト (上位)	0.6192	アーク代替コスト (上位)	0.0000
ノード代替コスト (下位)	0.0078	アーク代替コスト (下位)	0.0077
ノードスケール (上位)	0.1979	アークスケール (上位)	0.0000
ノードスケール (下位)	0.0250	アークスケール (下位)	0.0000
ノード位置 (上位)	1.0000		
ノード位置 (下位)	0.0000		
ノード主軸方向 (上位)	1.0000	アーク方向 (上位)	1.0000
ノード主軸方向 (下位)	0.4122	アーク方向 (下位)	0.3567
ノード縦横比 (上位)	0.4122	アーク端点領域面積比 (上位)	0.5300
ノード縦横比 (下位)	0.9018	アーク端点領域面積比 (下位)	0.0173
ノード Compactness (上位)	0.6367	アーク端点領域角度差 (上位)	0.9469
ノード円形度 (上位)	0.9368	アーク端点領域角度差 (下位)	0.4087
ヌルシンボル係数 (上位)	0.0051		
ヌルシンボル係数 (下位)	0.0000		

表 10.3: 類似検索実験 1 の場合に上位検索された画像と主観的類似度との関係。

検索順位	0	1	2	3	4	5	6	7	8	9
主観的類似度 (順位)	例示画	31(1)	3(14)	16(2)	—	13(3)	7(8)	9(7)	3(14)	—

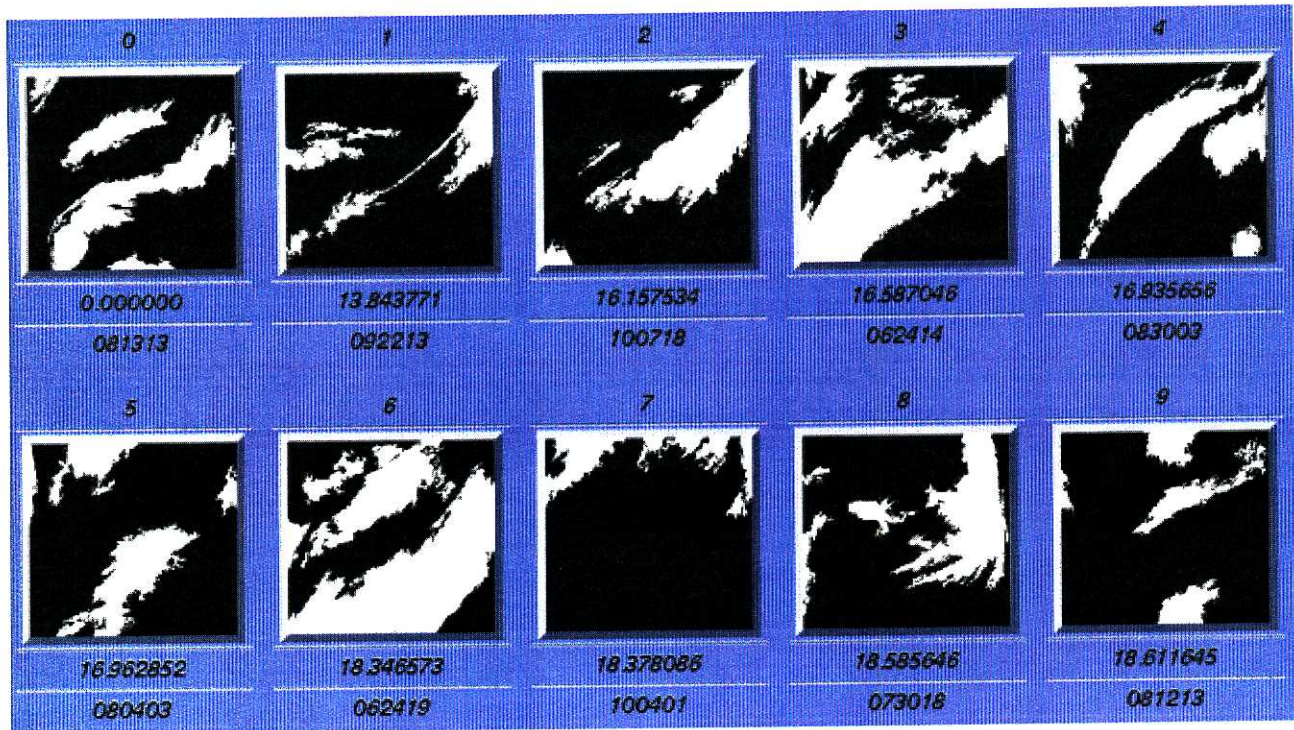


検索結果上位 10 枚の画像

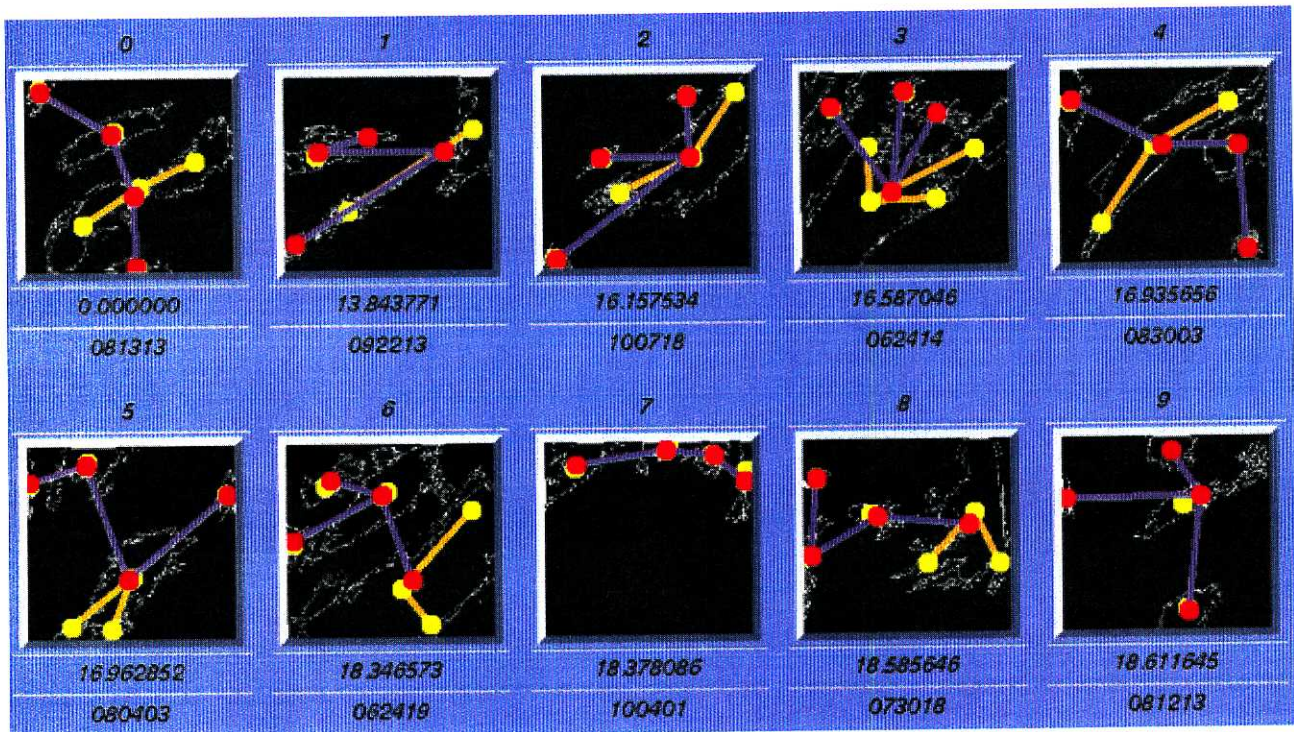


検索結果のグラフ構造表示

図 10.5: (実験 1) NOAA 衛星画像データベースの学習用サブセット 163 枚を検索対象とした検索結果。重み係数は表 10.1 を用いた。問い合わせ画像は 1 位に検索されている画像である。

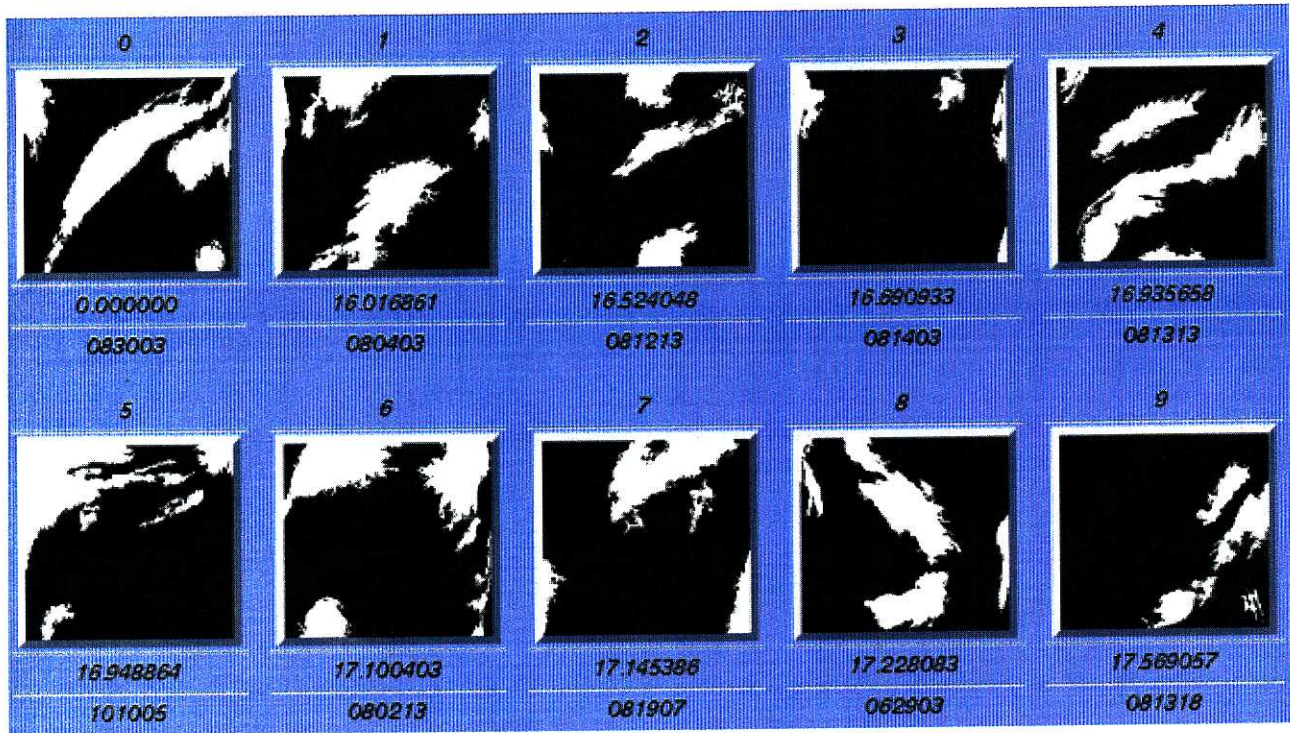


検索結果上位 10 枚の画像

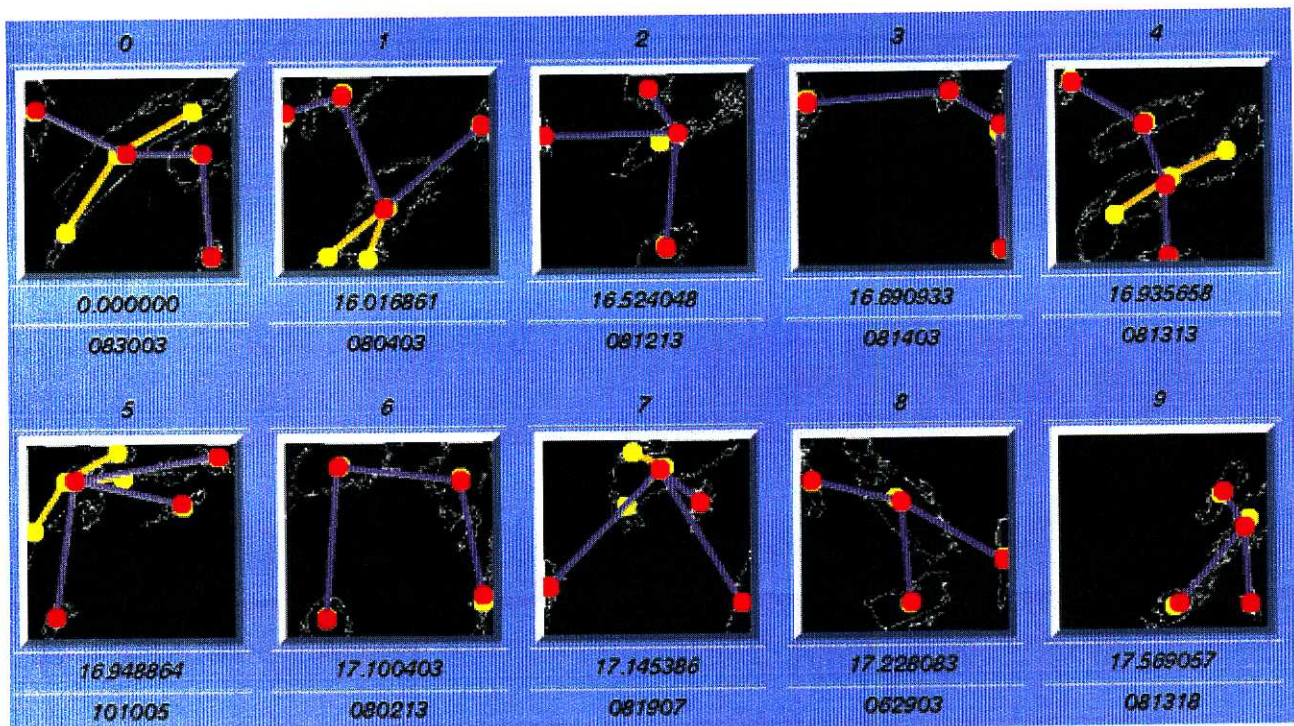


検索結果のグラフ構造表示

図 10.6: (実験 2) NOAA 衛星画像データベースの全画像 1027 枚を検索対象とした検索結果。重み係数は表 10.1 を用いた。問い合わせ画像は 1 位に検索されている画像である。



検索結果上位 10 枚の画像



検索結果のグラフ構造表示

図 10.7: (実験 3) NOAA 衛星画像データベースの全画像 1027 枚を検索対象とした検索結果。重み係数は表 10.1 を用いた。問い合わせ画像は 1 位に検索されている画像であり、図 10.6 とは変えている。

表 10.4: 類似検索実験 4 の場合に上位検索された画像と主観的類似度との関係。

検索順位	0	1	2	3	4	5	6	7	8	9
主観的類似度 (順位)	例示画	5(11)	7(8)	13(3)	6(10)	11(5)	10(6)	4(12)	—	—

実験 3 の結果

実験 3 では実験 2 と比べて例示画を変更している。ここで新しく用いられる例示画は、主観的類似順位が 1 位であり同時に実験 1 でも 1 位に検索されていた画像である。このように例示画を変えると、異なる例示画で学習された重み係数はやはり準最適解でしかない。そこでこの実験の目的は、例示画の小さな変化に対して重み係数がどれほどロバストであるかを確かめることにある。この例示画の小さな変化を実現するために、本実験では主観的類似順位が 1 位の画像を新しい例示画とした。その検索結果を図 10.7 に示す。図 10.5 と比べると、1 位、2 位、4 位、9 位の画像は共通して上位に検索されている画像である。またその他の上位検索画像では、主軸方向による影響が小さくなり、空間的配置の類似性という側面がより強く出てきている。これは図 10.5 の例示画と図 10.7 の例示画とでは、主軸方向の属性値の影響が異なることが原因であると考えられる。いずれにしても、学習された重み係数にはこのように検索意図といったものが埋め込まれており、例示画を少し変えても重み係数はそのまま類似画を検索することができた。

実験 4 の結果

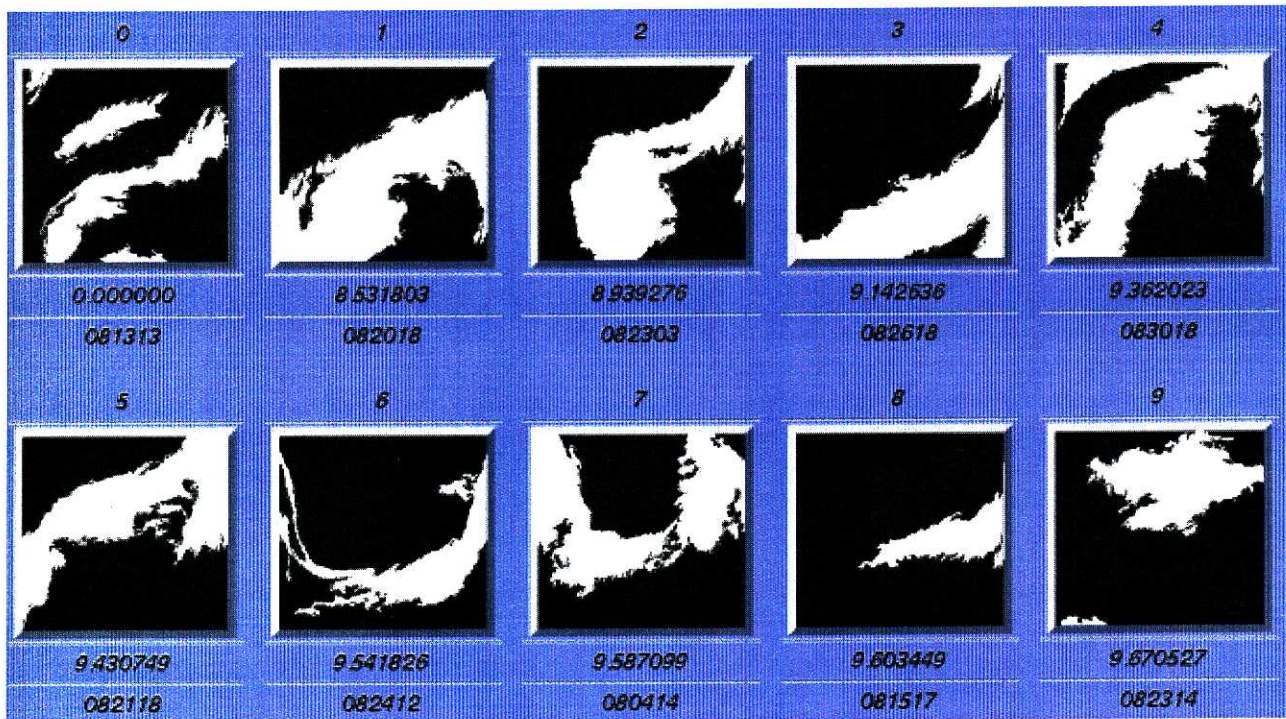
実験 4 では実験 1 と比べて教師データを変えている。すなわち実験 1 とは別の観点からの類似検索をおこなうことを目的とする。実験 1 の結果では、表 10.3 にも現われているように、主観的類似画像の一部はかなり下位に検索されていた。こうなった理由は学習方法の不備ではなく、そもそも主観的類似画像が全く異なる複数の観点から選ばれていたためであると考えられる。ここではその仮説を確かめるため、実験 1 で上位 10 位以内に検索された主観的類似画像を学習用の教師データから取り除き、22 枚の教師データを用いて重み係数を再学習した。そのようにして新たに得られた重み係数を表 10.2 に示す。これを表 10.1 と比較してみればわかるように、この二つの最適な重み係数で、どの属性値がどの程度変化したのかは一見してもわからず、したがってこの重み係数から何らかの結論を引き出すのは難しそうである。

そこでこの重み係数を用いて類似検索した結果を図 10.8 に示す。今度は図 10.5 との違いは明瞭である。この場合は図 10.5 とは全く異なる観点から類似しているとみなされる画像が上位に検索されている。図 10.8 では、いわば「大きな領域の主軸方向が類似している画像」が上位に検索されるような重み係数が学習されたことが推測できる。なお表 10.4 では類似検索順位と主観的類似度との対応をまとめた。1 位から 7 位までは人間が選んだ主観的類似画像 29 枚に含まれるものであるが、どれも表 10.3 には登場しなかったものである。つまり表 10.2 の重み係数は、表 10.1 の重み係数とは全く異なる観点を表現する重み係数であると結論できる。

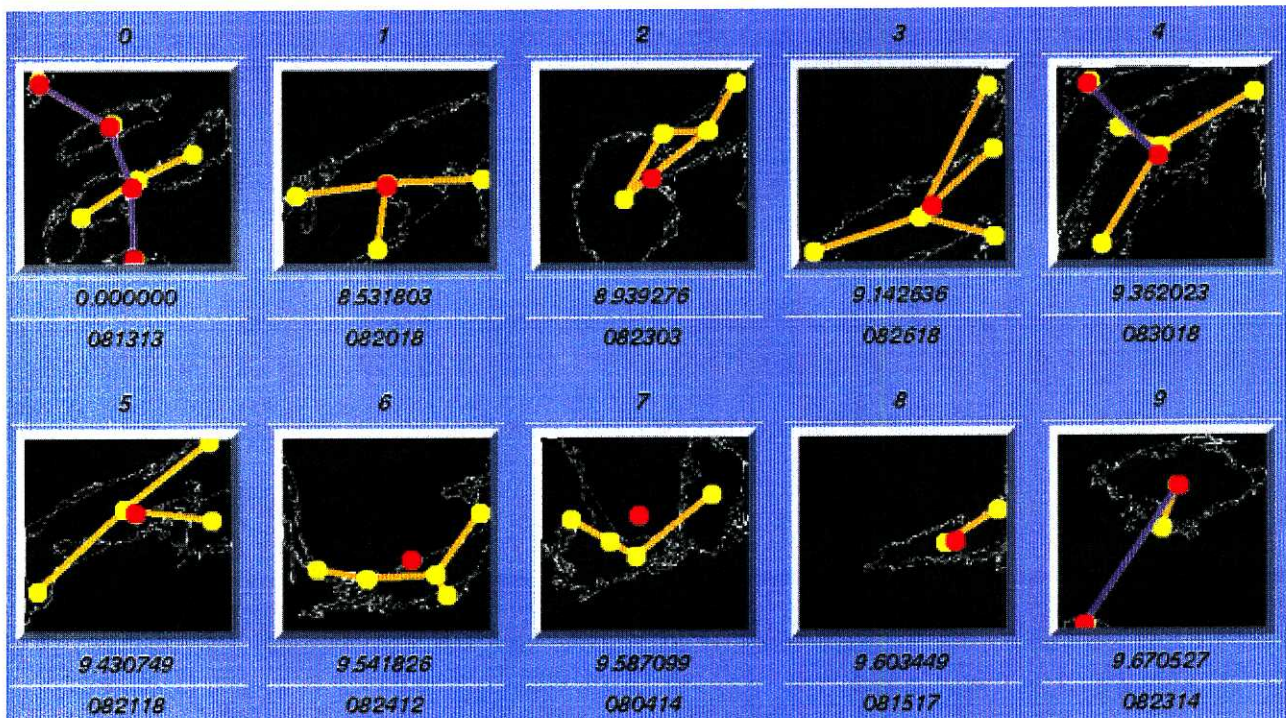
まとめ

以上の実験 1 ～実験 4 の実験をまとめると以下のような結論を得る。

1. 前もって選んだ主観的類似画像を教師データとして類似検索規準を学習させれば、それらを上位に検索できるように重み係数を学習することが可能である。



検索結果上位 10 枚の画像



検索結果のグラフ構造表示

図 10.8: (実験 4) NOAA 衛星画像データベースの学習用サブセット 163 枚を検索対象とした検索結果。重み係数は表 10.2 を用いた。問い合わせ画像は 1 位に検索されている画像である。

2. 検索対象画像集合の変更や例示画の変更によって、学習された重み係数が準最適となる場合でも、学習された重み係数に埋め込まれた検索意図は有効に働く。
3. 異なる観点からの教師データを与えれば、その教師データの類似性の観点を反映した重み係数をそれぞれ学習できる。

結論1はまず第一に確認しなければならない結論である。しかしこれはわれわれが「本当にやりたいこと」ではない。重み係数の学習には例示画に類似した教師データがあらかじめ必要となるが、第9.2.4節で述べたように、教師データが得られているということはすでに検索が完了していることを意味する。したがって「本当にやりたいこと」とは、結論1のように例示画と教師データが対応した場合に最適な重み係数が学習できるのを示すことではなく、それによって学習した重み係数を他の組み合わせに適用した場合にも有効に働くのを示すことである。そしてもし結論2が成立すれば、重み係数は少数の代表的な例のみに対して学習すれば十分となり、あとは重み係数の微調整がユーザにまかされることになる。よって重み係数の調整による柔軟な画像検索の実用化の鍵は、結論2がどの程度成立するかを検証する点にあると考えられる。また結論3を応用することで、異なる観点からの類似画像を別々に教師データとして学習させることにより、複数の観点からの重み係数をそれぞれ学習させることができる。そして検索時には重み係数のリストから適切な重み係数を選ぶことによって、異なる観点からの検索が簡単に実行できる可能性がある。つまり結論2と結論3は、柔軟な画像データベースを構築するための足掛かりとなる結論である。

10.3.2 GMS 衛星画像データベースの類似検索結果

2種類の実験の内容

GMS 衛星画像データベースでは、短期型画像データベースと長期型画像データベースの2種類の画像データベースを構築した。ここではそれぞれの画像データベースに対して類似検索実験をおこなう。

短期型データベースの実験結果

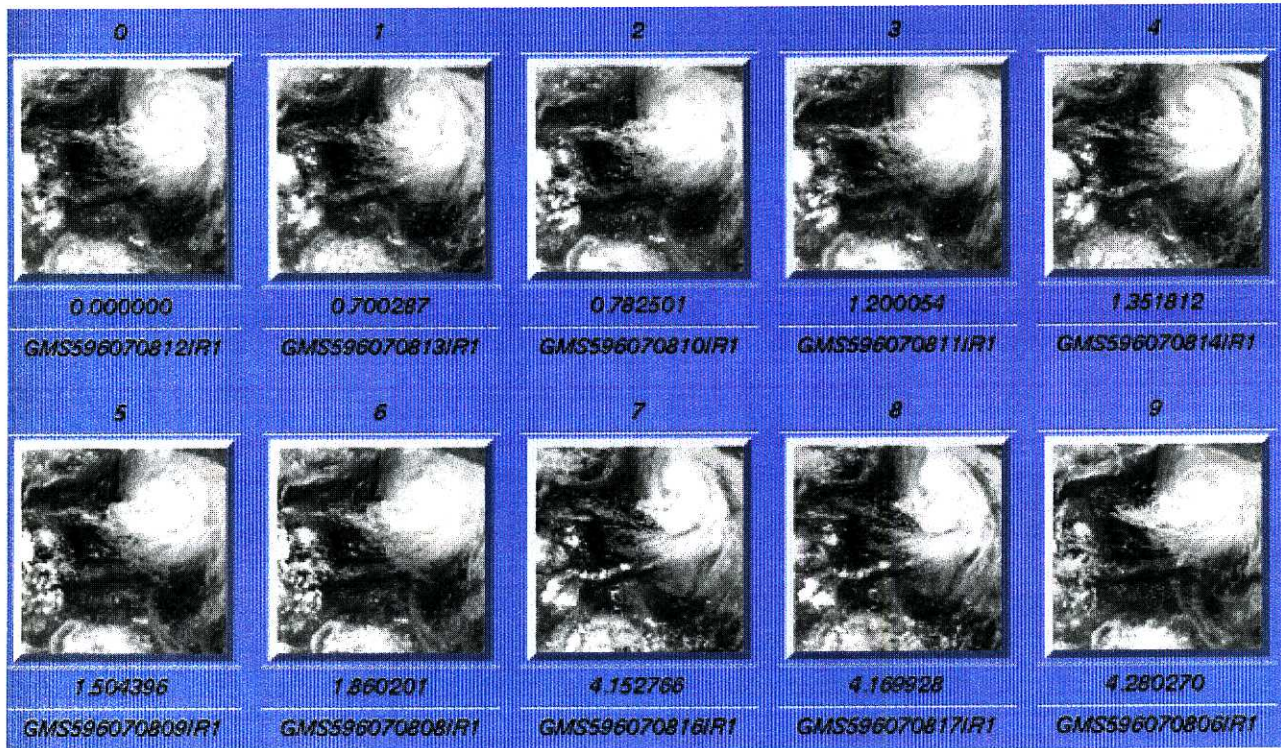
この実験は、第10.2.2節で述べたように客観的に選び出した類似画像を検索する実験によって、類似検索結果を「客観的に」評価することが目的である。つまり他の類似検索ではその結果を客観的に評価することが難しいのに比べて、この場合は観測日時が近い画像が検索されたかどうかという評価規準をもって、類似画像検索の結果を客観的に評価できるのである。重み係数に最適化された重み係数表10.5を用いた場合の類似検索結果を図10.9に示す。この結果では、確かにおおよそ観測日時が近い順に画像が検索されていることがわかる。特に注目すべき点は、7番目の画像に対応するグラフ構造の下位構造が、他のグラフとは異なっているにもかかわらず、これを上位に検索できているという点である。このように形状分解結果によって生成されるグラフが少し異なっても、重み係数の柔軟性がこれを吸収して妥当な類似検索結果を生み出す。このように本研究の手法は、この場合は「客観的に」みても類似画像検索に成功していると評価できる。

表 10.5: 短期型 GMS 衛星画像データベースの検索に用いられた重み係数の一覧。遺伝的アルゴリズムによって最適化された結果である。

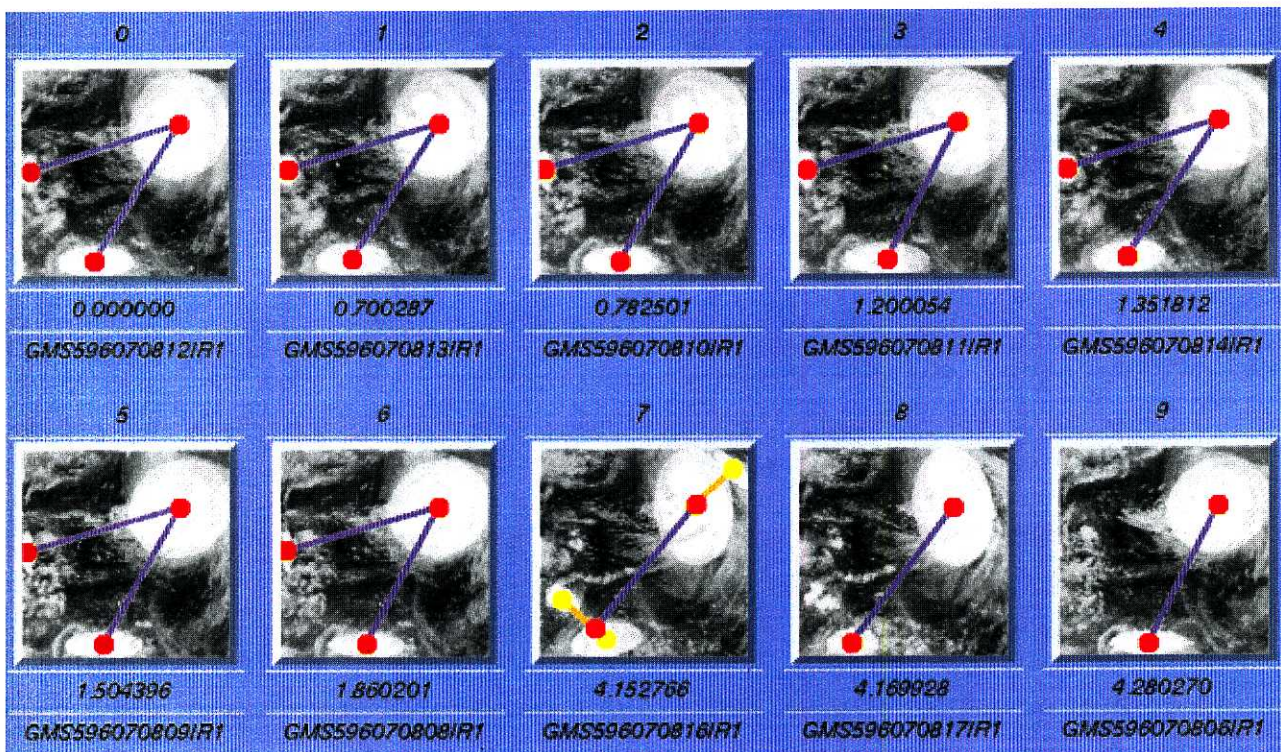
重み係数	重みの値	重み係数	重みの値
ノード代替コスト (上位)	0.1586	アーク代替コスト (上位)	0.5364
ノード代替コスト (下位)	0.0000	アーク代替コスト (下位)	0.0003
ノードスケール (上位)	0.2468	アークスケール (上位)	0.0000
ノードスケール (下位)	0.0000	アークスケール (下位)	0.0000
ノード位置 (上位)	0.4784		
ノード位置 (下位)	0.1643		
ノード主軸方向 (上位)	0.0046	アーク方向 (上位)	0.2252
ノード主軸方向 (下位)	0.9993	アーク方向 (下位)	0.9863
ノード縦横比 (上位)	0.1406	アーク端点領域面積比 (上位)	0.1974
ノード縦横比 (下位)	0.1615	アーク端点領域面積比 (下位)	0.3280
ノード画素値分散 (下位)	0.3181	アーク端点領域角度差 (上位)	0.0000
		アーク端点領域角度差 (下位)	1.0000
ヌルシンボル係数 (上位)	0.0000		
ヌルシンボル係数 (下位)	0.0000		

表 10.6: 長期型 GMS 衛星画像データベースの検索に用いられた重み係数の一覧。遺伝的アルゴリズムによって最適化された結果である。

重み係数	重みの値	重み係数	重みの値
ノード代替コスト (上位)	0.9479	アーク代替コスト (上位)	0.1621
ノード代替コスト (下位)	0.3851	アーク代替コスト (下位)	0.0000
ノードスケール (上位)	0.1526	アークスケール (上位)	0.3416
ノードスケール (下位)	0.2133	アークスケール (下位)	0.0000
ノード位置 (上位)	0.9000		
ノード位置 (下位)	0.0000		
ノード主軸方向 (上位)	0.0000	アーク方向 (上位)	0.4532
ノード主軸方向 (下位)	0.5050	アーク方向 (下位)	0.7006
ノード縦横比 (上位)	0.4975	アーク端点領域面積比 (上位)	0.2242
ノード縦横比 (下位)	0.2529	アーク端点領域面積比 (下位)	0.8390
ノード画素値分散 (下位)	1.0000	アーク端点領域角度差 (上位)	0.0314
		アーク端点領域角度差 (下位)	0.5065
ヌルシンボル係数 (上位)	0.5980		
ヌルシンボル係数 (下位)	0.0492		

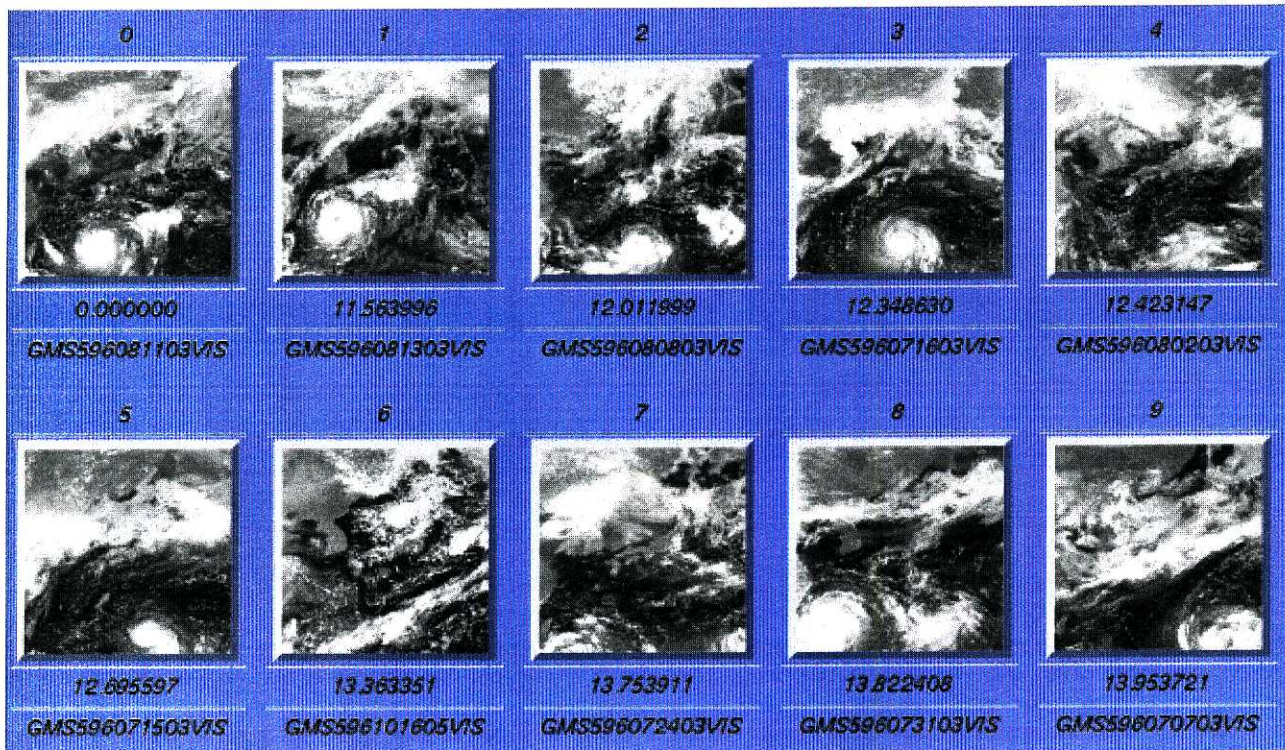


検索結果上位 10 枚の画像

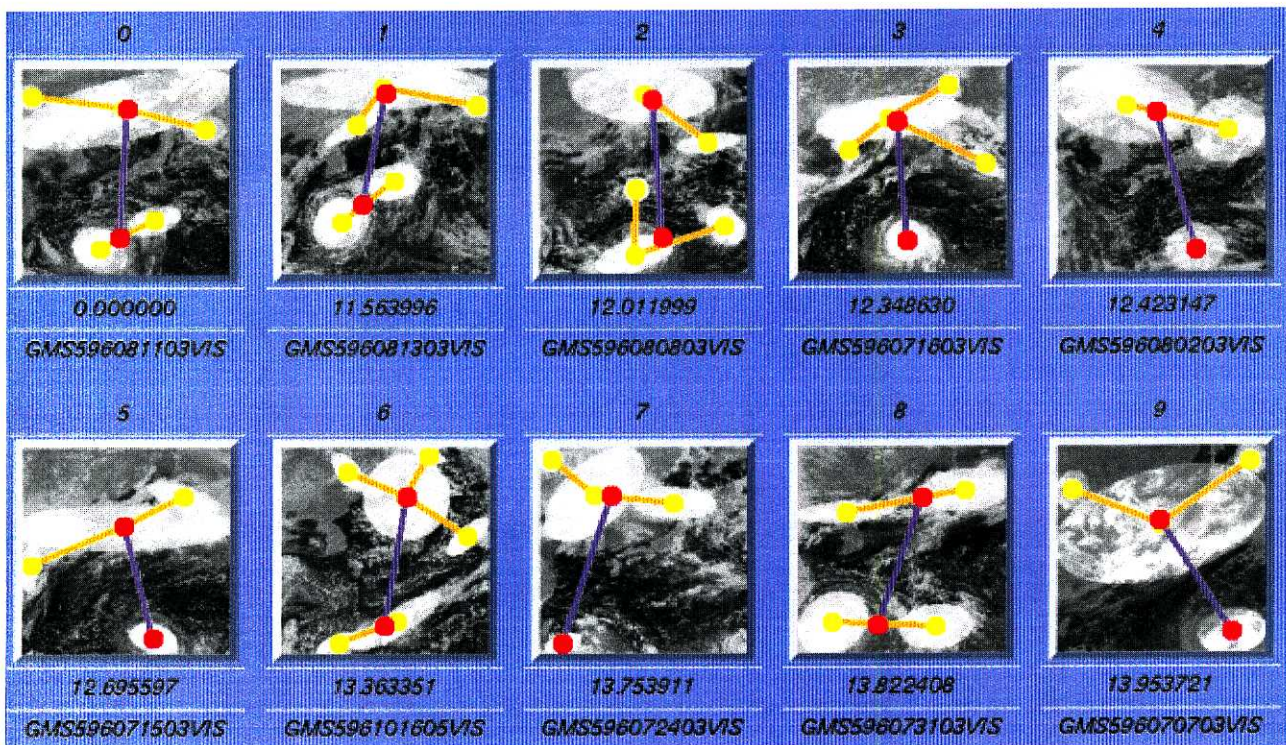


検索結果のグラフ構造表示

図 10.9: 短期型 GMS 衛星画像データベースの検索結果。重み係数は表 10.5 を用いた。問い合わせ画像は 1 位に検索されている画像である。



検索結果上位 10 枚の画像



検索結果のグラフ構造表示

図 10.10: 長期型 GMS 衛星画像データベースの検索結果。重み係数は表 10.6 を用いた。問い合わせ画像は 1 位に検索されている画像である。

長期型データベースの実験結果

このように短期型の検索結果に関しては、客観的にみて成功していると評価できた。しかしこれはやはり我々が「本当にやりたいこと」ではない。「本当にやりたいこと」とは、全く観測時刻の離れた画像から例示画に類似したパターンを発見し、気象学的な意味を探ることにある。そのためには検索対象の画像集合をより広げた長期型の場合にも、類似検索がうまくいくことを示すことが重要である。学習された重み係数を表 10.6 に示し、その重み係数を用いた類似検索結果を図 10.10 に示す。

用いた例示画は 8 月 11 日の画像であることから、1 位と 2 位に検索された 8 月 13 日と 8 月 8 日の画像は、例示画を含む一連の雲パターンであると考えられる。それに対し、3 位と 5 位に検索された画像はそれぞれ 7 月 16 日と 7 月 15 日の画像であることから、例示画とは別の一連の類似した雲パターンを検索した結果である。また 8 位と 9 位にも、それぞれ別の類似した雲パターンを検索している。このように本研究の手法を用いることによって、過去の雲パターンの画像から、例示画と類似した雲パターンを検索することが可能である。

しかしここで注意すべきことは、当然検索されるべき 8 月 12 日の画像が上位に検索されていないことである。目視でみた限りでは 8 月 12 日の画像と例示画はかなり類似している。形状分解の結果をみても、8 月 12 日の画像と例示画はそれほど変わらない。しかし引力モデルの出力として生成されたグラフ構造の段階で 8 月 12 日の画像と例示画とは大きく表現が異なってしまった。具体的に述べると、8 月 11 日および 8 月 13 日のグラフは図 10.10 にあるように上位が 2 ノードのグラフ構造であるが、8 月 12 日の画像は上位が 1 ノードのグラフ構造となってしまう。短期型であったような下位構造の小さな違いは、重み係数の柔軟性で吸収することができる。しかし上位構造に大きな違いが生じると、もはや重み係数の柔軟性では吸収できない。グラフ構造の欠点として、画像上での微小な違いがグラフ構造の位相構造の違いに拡大されて現われる点が従来から指摘されているが、この例はまさにこの畏にはまった例である。この点に注意しながら、引力モデルをよりロバストな手法として発展させていくことが、今後の研究課題となるだろう。

10.3.3 カラー自然画像データベースの類似検索結果

2 種類の実験の内容

カラー自然画像データベースでは 2 種類の実験をおこなった。まず図 10.4 の左側の例示画を用いて、学習用サブセットを用いた重み係数の学習をおこなった。こうして得られた最適な重み係数を用いて以下の実験をおこなう。

1. 検索対象を全体とした場合の類似検索をおこなう。
2. 例示画を別の画像 (図 10.4 右) に変更し、全体を検索対象とした類似検索をおこなう。

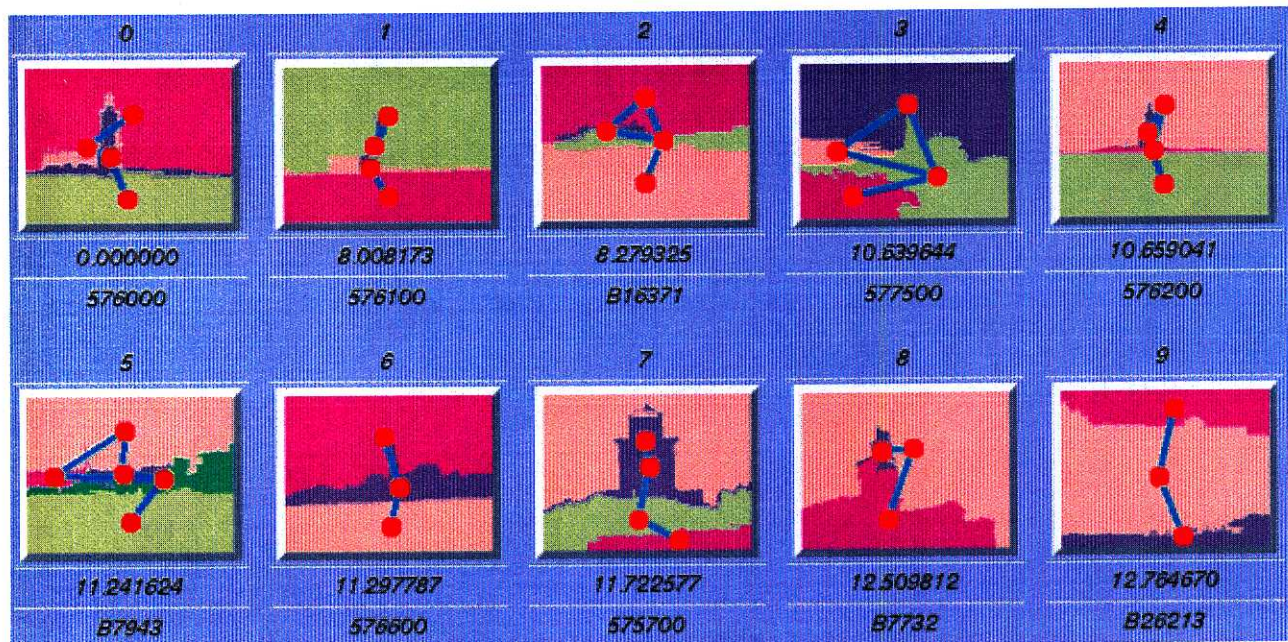
以下ではこの 2 種類の実験の結果について述べる。

実験 1 の結果

実験 1 の目的は、最適化された重み係数がカラー自然画像の場合にも有効に働くのを示すことにある。表 10.7 のように最適化された重み係数を用いて類似検索した結果を図 10.11 に示す。



検索結果上位 10 枚の画像



検索結果のグラフ構造表示

図 10.11: (実験 1) カラー自然画像データベースの全画像 950 枚を検索対象とした検索結果。重み係数は表 10.7 を用いた。問い合わせ画像は 1 位に検索されている画像である。

表 10.7: カラー自然画像データベースの検索に用いられた重み係数の一覧。遺伝的アルゴリズムによって最適化された結果である。

重み係数	重みの値	重み係数	重みの値
ノード代替コスト (上位)	0.4657	アーク代替コスト (S, S) (上位)	0.9799
		アーク代替コスト (S, C) (上位)	0.6367
		アーク代替コスト (S, C) (上位)	0.9487
ノードスケール (上位)	0.5178	アークスケール (上位)	0.0000
ノード位置 (上位)	0.9441		
ノード主軸方向 (上位)	0.1329	アーク方向 (上位)	0.1907
ノード縦横比 (上位)	0.1838	アーク端点領域面積比 (上位)	0.2985
ノード Compactness (上位)	0.9909	アーク端点領域角度差 (上位)	0.0481
ノード平均色 R 成分 (上位)	0.0693		
ノード平均色 G 成分 (上位)	0.0000		
ノード平均色 B 成分 (上位)	0.8988		
ヌルシンボル係数 (上位)	0.0000		

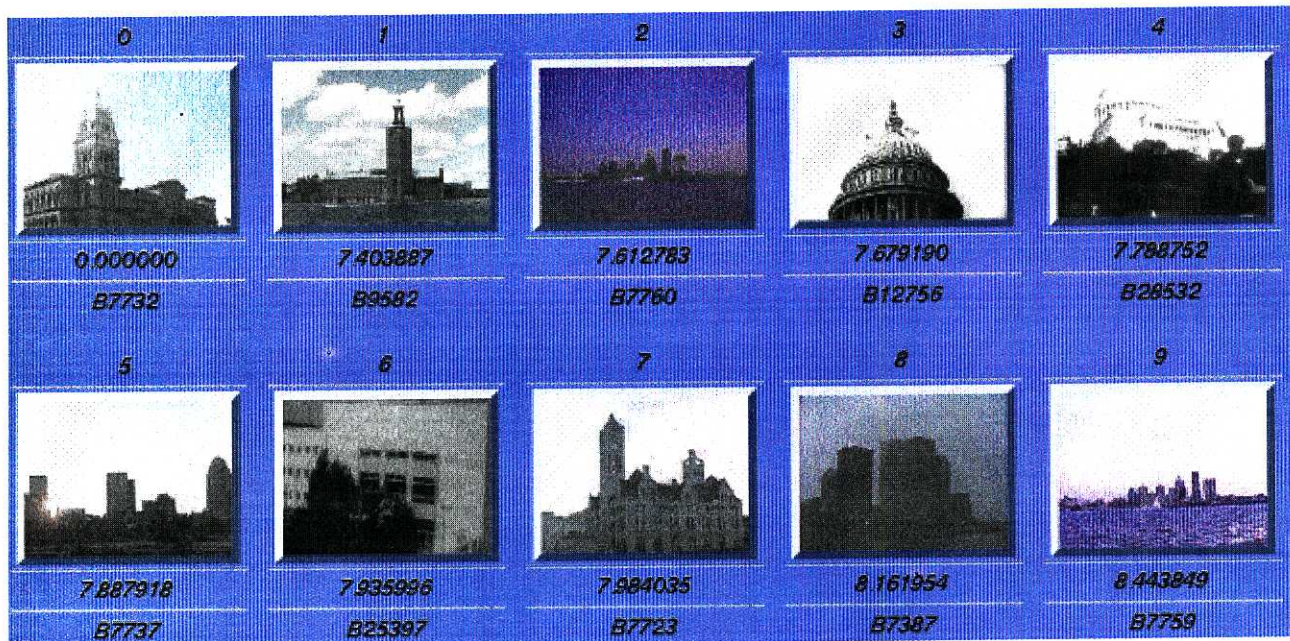
この結果では、例示画として用いた灯台の画像に類似した灯台の画像が、上位に5枚検索されている。また灯台とは全く画像内容の異なる都市の風景(5位)や石造建築物(8位)が、その空間的配置や特徴量の類似性によって検索されている。下のグラフ構造から推測できるように、検索された画像は垂直方向に3個程度の領域が重なり、しかも色が全体的に青みが掛かっているという観点から類似している。それは教師データとして与えた画像が、いずれも海上にある灯台を含む画像であったためであると考えられる。このようにカラー自然画像に対しても、グラフ構造を用いた本研究の手法が有効であることを示した。

実験2の結果

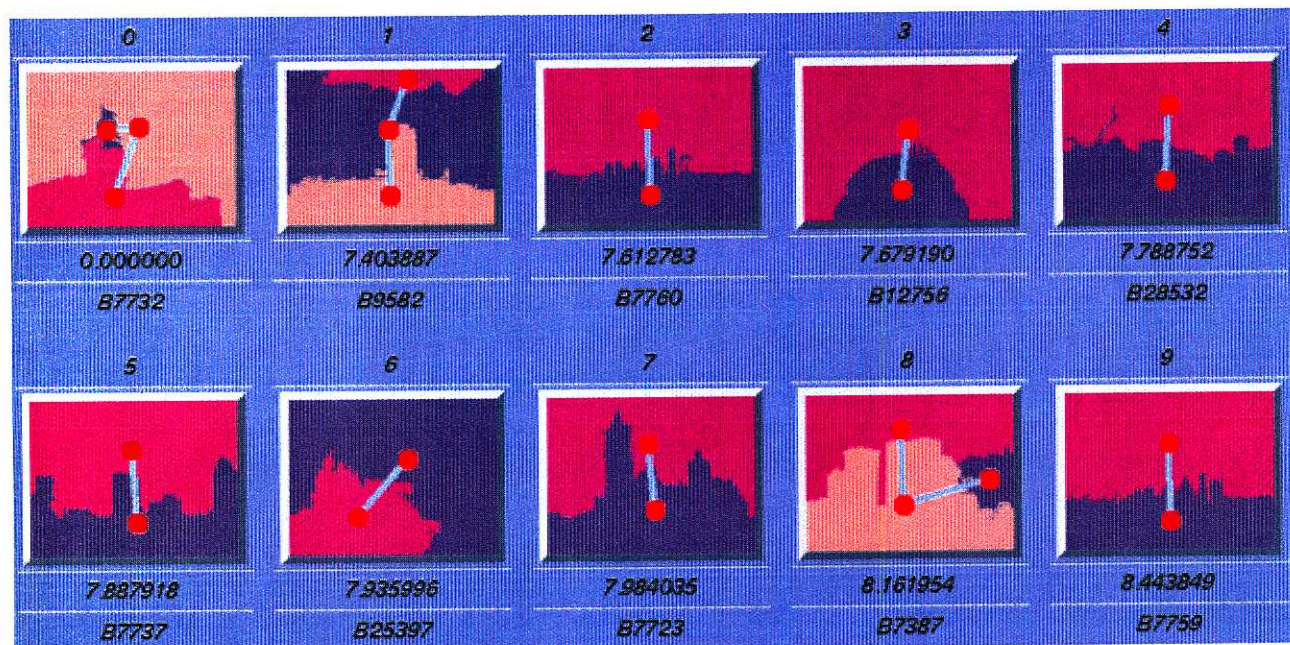
実験2では実験1の結果をさらに発展させる。実験1で検索された画像の中で、灯台ではないが類似した画像として8位に検索された石造建築物の画像に着目し、この画像を新たな例示画にして類似検索を試みる。これは一種の「連想検索」を想定した検索実験である。重み係数には実験1と同一のものを用いた場合の類似検索結果を図10.12に示す。この場合には青い空を背景とした建築物の画像が上位検索画像の多数を占めた。1位や7位の画像は、色の類似性や空間的配置の類似性などの観点から特に例示画に類似している画像である。このように例示画を変更しても類似画像を検索できることはNOAA衛星画像の場合にも確かめたが、これは重み係数のある程度の汎化能力を示すものであると考えることができるだろう。

10.4 類似画像検索の評価

以上の検索結果をまとめると、本研究で構築した3種類の画像データベース、すなわちNOAA衛星画像データベース、GMS衛星画像データベース、カラー自然画像データベースの類似画像検索に対して本研究で提案した手法が有効であることを示した。本節ではこの結果をある程度定量的に評価することによって、さらに多面的に3種類の画像データベースの検索結果をまとめ



検索結果上位 10 枚の画像



検索結果のグラフ構造表示

図 10.12: (実験 2) カラー自然画像データベースの全画像 950 枚を検索対象とした検索結果。重み係数は表 10.7 を用いた。問い合わせ画像は 1 位に検索されている画像である。これは図 10.11 とは異なる画像である。

る。

10.4.1 検索精度の評価

類似画像検索の研究では、検索結果の評価規準が常に問題となっている。例えば、なんらかの数理的な方法で客観的な類似度を与えることができれば、評価規準にあいまいさは生じない。また、いわゆる「専門家の知識」という唯一無二の絶対的な評価規準が存在し、その知識と照らし合わせることで検索結果が評価できる場合にも、評価規準にあいまいさは生じない。しかし現実には多くの場合において、このような数理的な方法で類似度を与えることもできなければ、唯一無二の絶対的な評価規準が存在することもなく、したがって、結果の評価は究極的には検索システムを使うユーザの評価、あえて言えばユーザの主観に帰着してしまう。このような類似度の性質が、類似検索結果を客観的に評価することを難しくする最大の原因である。このような評価法の確立そのものが今後の研究課題であるといえる [6] が、現在のところ理想的な評価規準の代替手段として用いられている評価規準の代表的な例を以下に示す。

1. 検索結果から計算できる数字を用いる方法 最もよく用いられる量は再現率(検索率または呼出率)と適合率である。

再現率 (recall rate) 正しく検索された数 / 検索されるべき数

適合率 (precision rate) 正しく検索された数 / 検索された数

実際にどのような画像集合を考えるかで微妙に計算方法が異なることもあるが、この種の量が最も基本的な評価規準である [6, 3, 23]。その他には以下のような規準も考えられている [8]。

検索ヒット率 検索結果上位 R 位以内に、 r 枚の心理的類似画像が検索される割合。

非類似画像混入率 u 枚の心理的類似画像の非類似度の総和と、上位 u 位以内の検索画像の非類似度の総和の比較

2. 事後調査・アンケート調査による方法 例えば画像データベースのユーザに満足度を5段階評価で示してもらい、ユーザの何%が満足したかを調べる方法が提案されている [126]。

3. 例示画に人工的な変形を加えた画像を用いる方法 例示画をあらかじめ用意し、それに少々の変形を施した画像を生成してデータベースに混入することにより、例示画との類似度があらかじめ判明している画像を用いた類似検索ができる。そして、その変形画像を正しく上位に検索できれば、検索アルゴリズムの動作の正しさとロバスト性を証明することができる。本研究で用いた短期型 GMS 衛星画像データベースは、自然の力によって例示画と類似した画像が生成されている例であり、このような場合を逆に人工的に生成することによってシステムを評価することができる可能性がある。ただしこの方法は、完全にコントロールされた環境のもとでの実験という意味では、他の科学分野ではごく普通におこなわれている実験方法である。

本研究の評価方法は以下の通りである。まず NOAA 衛星画像データベースについては、再現率・適合率を基本とした評価規準を適用する。次に GMS 衛星画像データベースに関しては、短期型の方はすでに客観的な評価規準を示しているのので、ここではあえて評価はしない。長期型およびカラー自然画像データベースについては、教師データに用いた画像をもとにして再現率・適合率を計算する。

表 10.8: 各画像データベースの場合の再現率と適合率。

	再現率	適合率
NOAA 衛星画像データベース (重み 1) (163 枚)	11 / 29	11 / 50
NOAA 衛星画像データベース (重み 1) (1027 枚)	7 / 29	7 / 50
NOAA 衛星画像データベース (重み 2) (163 枚)	16 / 29	16 / 50
NOAA 衛星画像データベース (重み 2) (1027 枚)	6 / 29	6 / 50
GMS 衛星画像データベース (331 枚)	6 / 12	6 / 30
カラー自然画像データベース (239 枚)	7 / 9	7 / 30
カラー自然画像データベース (950 枚)	7 / 9	7 / 30

NOAA 衛星画像データベースの場合

NOAA 衛星画像データベースの場合の再現率および適合率を計算する。まず再現率としては、29 枚の主観的類似画像のうち、上位 M 位以内に検索された枚数を計算した。次に適合率としては、上位 M 位以内に検索された中に含まれる、29 枚の主観的類似画像の枚数を計算した。 M としては $M = 50$ を選んだ。

その結果を表 10.8 に示す。最初の実験では学習用サブセットとして用いた 163 枚の画像を対象として計算し、次に全体の 1027 枚の画像を対象として実験をおこなった。このときの重み係数には表 10.1 を用いた。表 10.8 において再現率が低いのは、そもそも主観的類似画像がある一つの統一された観点から選ばれたものではないことが原因である。重み係数はある一つの観点に対応して最適化されるものであるから、ある一つの観点からみて類似している画像を検索する際には、他の観点からみて類似している画像を上位に検索することは当然むりだからである。

そこで次に表 10.2 の重み係数を用いて実験をおこなった。この場合は、学習用サブセットを対象とした場合には重み係数 1 の場合よりも再現率・適合率ともに良い値となっているが、検索対象を全体とした場合にはその立場は逆転している。なお検索順位の上位に含まれる主観的類似画像は、実際は両方で全く異なっている。これはそれぞれの重み係数が異なる観点からの類似性を表現しているためである。

GMS 衛星画像データベースの場合

短期型の場合は先述したように、観測日時が近い画像が検索されるかどうかを客観的な評価基準とできるから、図 10.9 を満足のいく検索結果と評価できる。次に長期型の場合を検討する。ここでは教師データとして 12 枚の画像を選んだことから、これらの画像を対象として再現率と適合率を計算する。その結果を表 10.8 に示す。

カラー自然画像データベースの場合

カラー自然画像データベースの結果も同時に表 10.8 にまとめたが、こちらの場合は教師データはかなり多数が上位に検索されている。それは教師データとして選ばれた画像がかなり均質なものであり、学習された重み係数がうまく教師データに適合していることがその理由であると考えられる。

表 10.9: 重み係数がさまざまに異なる場合の検索時間の平均。

データベース	数	平均 (s)	1 グラフ (ms)	最大値 (s)	最小値 (s)
NOAA 衛星画像データベース	163	0.561	3.44	1.06	0.11
GMS 短期型画像データベース	132	0.0689	0.522	0.15	0.02
GMS 長期型画像データベース	331	0.403	1.22	1.40	0.12
カラー自然画像データベース	239	0.437	1.83	94.64	<0.01

総合的な評価

このように再現率と適合率を計算することができるが、この値は上位何枚を計算の対象とするかに大きく依存する値である。従って上位何枚が最も妥当かを客観的に決める手段がない以上は、どうしても恣意的な評価規準とならざるを得ない。また大きければ大きいほど良いのは確かだが、ではどの程度大きければよいのか、という値に対しても一般的に受け入れられている値は存在しない。問題の困難さや検索目的によって再現率と適合率の妥当な値は大きく変化するからである。本研究では一般に用いられている評価規準として再現率と適合率を用いたが、この評価規準の開発自体が大きな問題であることには留意しておく必要がある。

10.4.2 検索速度の評価

それぞれの画像データベースで検索速度を測定した。検索速度は以下の2通りの場合で測定した。

1. 重み係数を最適化する過程で得られる情報を用いて、さまざまな重み係数の場合に上位10枚を検索するのに要する時間を測定する。
2. 最適化された重み係数で上位10枚 / 100枚を検索するのに要する時間を測定する。

検索1の方は、ある例示画を用いた場合に要すると見積もられる検索時間の期待値や最良値・最悪値などを測定することによって、検索時間のばらつきを評価することが目的である。それに対して検索2の方は、標準的な検索要求に対する検索時間を測定することで、システムの高速度性を評価することが目的である。

なお本章で述べる検索時間は、第8章で述べた計算時間の測定とは異なる検索条件と異なる計算機上で測定されたものであるため、これらの結果の数値を直接に比較することはできない。

検索1の検索速度の評価

遺伝的アルゴリズムの過程から得られる検索時間の期待値 重み係数の最適化には第9章で述べたように遺伝的アルゴリズムを用いる。この方法は基本的に生成検査法の一つであり、乱数によって新たな重み係数ベクトルを生成することによって、重み係数のベクトル空間内で適度にばらついた重み係数ベクトルが生成検査されることになる。そこでこれらの重み係数ベクトルの場合の検索時間をすべて合計し平均することで、重み係数がさまざまに異なる場合の検索時間の期待値を推定できる。ただし、遺伝的アルゴリズムの特性から最適値に近い重み係数ベクトルは繰り返し評価されるため、その影響で実際の値よりは期待値が小さく見積もられることに注意する。

表 10.10: 上位 10 枚 / 100 枚検索の場合の検索時間の平均。

データベース	数	10 枚 (s)	1 グラフ (ms)	100 枚 (s)	1 グラフ (ms)
NOAA 衛星画像データベース	1027	6.40	6.23	8.34	8.12
GMS 短期型画像データベース	132	0.082	0.62	0.112	0.85
GMS 長期型画像データベース	331	0.761	2.30	1.08	3.26
カラー自然画像データベース	950	0.489	0.51	0.990	1.04

その結果を表 10.9 に示す。なおこの計算に用いた計算機は SPARC CENTER 2000 であり、また表の数値は 50000 回以上の測定で得られた数値を示している。この結果からそれぞれの画像データベースの特徴が見えてくる。ただし用いているグラフ構造はどの場合も階層化属性付き関係グラフであり、グラフマッチングの方法も同一である。するとこのように検索時間に大きな差が生じるのは、1 グラフ当たりのノード数やアーク数が異なることが第一の原因であると考えられる。その他に属性値の個数が多くなると、マッチングコストの計算時間を増加させる要因となる。

平均検索時間 まず平均時間では、NOAA 衛星画像データベースの場合が 1 グラフあたりで最長の検索時間を要している。これは画像から抽出されたグラフ構造が、平均的には NOAA 衛星画像データベースの場合が最も複雑であることが原因である。それに対して GMS 短期型画像データベースは最も単純なグラフであるため、平均時間も短い。

最大値 次に最大値を見ると、カラー自然画像データベースの場合を除けば、どの場合も最大値が最小値の 10 倍程度となっている。つまり重み係数の設定によっては、検索時間に 10 倍程度の差が出ることを想定しなければならない。このデータは最悪の場合を考えたシステム設計に有効に活用することができる。

さらに特徴的なのはカラー自然画像データベースの場合であり、最大値と最小値に著しい差がある。この画像データベースで用いたグラフ構造は他の場合と違って最も単純な 1 層構造であり、本来ならば検索時間は短くなるはずである。その本来の時間は実際のところ最小値に現われており、検索時間最小値が 0.01 秒以下と、他の場合に比べて最も高速である。

これに比べて最大値が著しく大きいのは別の理由による。カラー自然画像データベースの場合、領域分割した結果生まれた各領域にノードを割り当て、それによってグラフ構造を構築した。ここでカラー画像の領域分割はよほどうまく領域併合を工夫しない限り、一般に領域数はかなり多くなる傾向がある。よってカラー画像の場合のグラフ構造は、「平均的には単純」だが「一部は非常に複雑」なグラフ構造となってしまった。最大のノード数は 45 個にも達している。すると重み係数の値によっては、この一部の非常に複雑なグラフとのマッチングに非常に長い時間を要する場合があり、したがって検索時間もこのようなマッチングに引っ張られて大きな値となる。ノード代替コストが 0 に近い値となるとこのような場合が発生する。この結果から、カラー自然画像データベースの検索時間を短縮するためには、領域分割後の領域数が必要以上に増加しないように、後処理としての領域併合が重要となることがわかる。

検索 2 の検索速度の評価

検索2の場合は特に、第8.7節で述べたマッチングリストの順序が検索時間に大きく影響する。マッチングリストの先頭に類似画像が集まるほど検索時間は速くなるため、マッチングリストが異なる並び方をしている場合に検索時間を比較しても意味がない。そこで本研究ではどの画像データベースも同じ条件で測定するために、マッチングリストが最適に整列されている場合の検索時間を測定して比較した。その結果を表10.10に示す。この検索時間は、この例示画と重み係数の組合せとしては最速の検索速度であり、実際の検索ではマッチングリストの整列の影響でこのデータよりは遅くなることに注意されたい。ただし検索を繰り返すことでマッチングリストが徐々に整列されていくため、検索時間は徐々にこの数値に近づくことにはなる。

表10.10ではNOAA衛星画像データベースの場合の検索速度がかなり遅い。これは検索1でも述べたようにこの場合のグラフ構造が最も複雑となっているからである。しかし1000枚で検索時間が約6秒というのは、許容できる検索時間にぎりぎり収まる程度の遅い検索時間である。ユーザの待ち時間や能率などを考えると、目標としては10000枚で5秒～10秒程度が望ましいと考えている。GMS長期型画像データベースにとってもこの目標はかなり苦しい状況であり、やはりグラフ構造の計算量の多さがここに現われている。またカラー自然画像データベースの場合、表10.10を見る限りではかなり高速に検索できているが、これはマッチングリストが整列されているためであり、検索1で述べたように最悪値は著しく遅い検索時間となった。

これらの実験結果を評価すると、やはり心配されたように検索時間の面からはまだまだ不満の残る数字である。そのためには第8章で述べたようなグラフ構造の高速化は重要なテーマであるが、それに加えて検索データのフィルタリングも重要なテーマであると考えられる[17, 127]。非常に簡単な計算によって、明らかに類似していない画像との厳密な類似度計算を省略できれば、検索時間の最悪値を著しく短縮することが可能である。この観点からの研究が今後は重要である。

10.4.3 柔軟性の評価

柔軟性の評価として、重み係数を変えることによって異なる検索意図を表現することができるか、という点を考える。その点についてはNOAA衛星画像データベースに対してすでに実験したように、表10.1と表10.2という2種類の重み係数を使い分けることにより、二つの異なる観点からの類似検索を実現することができた。複数の類似性の観点からそれを満たす教師データを選び、それぞれの場合について別々に重み係数を学習させておけば、実際の検索時にはあらかじめ用意された重み係数リストから適切な重み係数を選ぶだけで、さまざまな観点からの類似検索が可能になると考えられる。

ここで個々の重み係数の値の調整について触れておきたい。本研究の類似画像検索システムは、個々の重み係数の値を調整するための図10.13のようなインタフェースを備えている。従来の研究ではこのような重み係数の調整はすべてユーザに任されており、このような使いやすいインタフェースを備えることがシステムの柔軟性を示すものであると評価されていた。しかし著者の考えでは、それだけでは不十分である。表10.1と表10.2とを比較すればわかるように、どの重み係数をどれだけ変化させれば検索結果がよい方向に向かうのかは決して自明ではなく、実際には非常に難しい問題である。このような困難な作業を、システムに全く知識のないユーザに任せると、ユーザはただ戸惑うのみではないか。重み係数が1個や2個程度の単純なシステムならばそれでもかまわないかもしれないが、本研究のように20個以上もの重み係数が関係する類似検索では、人手で試行錯誤しながら最適な重み係数を発見するのは、ほぼ不可能なことだと言って



図 10.13: 個々の重み係数を調整するためのユーザインタフェース。

よい。

そこで本研究のアプローチのように、重み係数の大きな調整は最適化手法に任せるべきである。図 10.13 のようなインタフェースは、あらかじめ用意された重み係数を微調整するためのインタフェースとして捉えた方が適切であろう。大きな調整は最適化手法が自動的におこない、微調整のみはユーザがおこなう、といった 2 段階の調整機構を備えることで、はじめてユーザにとって使いやすい柔軟な画像データベースが実現すると考えている。

10.5 まとめ

本章では、類似画像検索システムを階層モデルという方法論に基づきながら 3 種類の画像データベースに実装し、それぞれに対して類似検索実験をおこなうことにより本手法の有効性を確認した。まず最初に主観的類似画像が上位に検索できるように重み係数を学習させることができることを示し、その重み係数は異なる教師データを与えると異なる観点を学習すること、また例示画を少々変えても重み係数は有効に働くことを確認した。次に本手法の評価を検索精度の評価、検索速度の評価、柔軟性の評価の 3 点から評価した。検索精度の評価では類似検索の評価法を比較したあとに、再現率と適合率の観点から結果を評価した。また検索速度の面からはグラフマッチングがかなり計算量の多い処理であり、計算量の面からはまだまだ課題が多いことを指摘した。最後に柔軟性の評価では、本研究のように 2 段階で重み係数を調整する機構を備えることにより、ユーザにとって真に使いやすいインタフェースを提供できることを論じた。

今後の課題には、まず重み係数が例示画の変化や類似性の観点の変化にどのくらいロバストであるのかという評価が必要である。この評価をすることによって初めて、多様な観点を網羅するためにはどのくらいの組の重み係数を学習すればよいかを定量的に論じることができる。また重み係数をより高速に学習する方法や、第 9 章で述べたようなナビゲーションの方法などを導入

して、ユーザの観点の変化によりすばやく追従できるような機構も必要である。もう一つの大きな課題は検索時間の短縮である。グラフマッチングを実用的に用いるためには計算時間の短縮が必須である。フィルタリングなどの処理を通じて劇的に計算時間を短縮できるような手法を追究していくことが大きな課題である。

第 11 章

結論

11.1 本研究の成果

本論文は「領域・空間情報を表現するグラフ構造を用いた類似画像検索」に関する研究をまとめたものである。類似画像検索という高度な画像検索手法を実現するために、画像データベース構築の方法論と個々の重要な要素技術に対する提案をおこない、さらには本論文の手法を実際の3種類の画像データベースに適用することで、本論文のモデルがさまざまな画像データベースに対して有効であることを論じた。

本論文ではまず最初に、6層の階層構造で構成された「階層モデル」を提案し、非常に複雑で入り組んだ構造をもつ類似検索問題を6層の階層構造をもって明確に把握できるモデルを示した。これは類似検索問題をいくつかの部分問題に分割し、各レイヤを相互に関連させながら部分問題を整理することで、画像検索システム構築のための有用な方法論となる。

次にこのモデルにしたがって各レイヤに必要となる個別の要素技術について研究を進めた。まず第3章から第5章は「画素レイヤ」に相当し、ミクセルの解析に関する総合的な研究をまとめたものである。リモートセンシング画像の場合に特に重要となるミクセルの概念をより明確にし、「面積占有率密度」や「ミクセル密度」など、従来統計的な画像分類手法では考慮されてこなかった新しい確率モデルを提案した。そしてその手法をリモートセンシング画像に適用することによって、各画素をピュアピクセルまたはミクセルに明確に分類でき、さらにミクセルに関しては面積占有率を計算する方法を示した。

第6章は「領域レイヤ」に相当し、変形モデルを用いた形状分解によって画像をプリミティブの集合として表現する方法を示した。変形モデルのエネルギー関数に用いられるメンバーシップ関数の定義を2値画像から濃淡画像へと自然に拡張することによって、どの画像にも適用できる形状分解の方法を示した。実験の結果、雲領域を表現するプリミティブとしては楕円が最も自然な結果となった。次に第7章は「関係レイヤ」に相当し、引力モデルをもちいたプリミティブの構造化について述べた。グラフ構造への構造化には2通りの方法を示し、さらに階層的に構造化することによって、画像を階層的なグラフ構造として表現する方法を示した。ただしこの2個のレイヤはいわゆるセマンティックギャップ [7] に相当し、パターンと記号の中間レベルに相当する最も定式化の難しいレイヤである。本研究でも十分に追究がなされたとはいえず、今後のさらなる研究が望まれる部分である。

第8章は「認識レイヤ」に相当する部分である。グラフ構造間の類似度であるグラフマッチングコストは計算量の非常に多い処理であり、その計算の高速化を目指して種々の工夫を提案

した。またグラフマッチングコストという類似度の計算について、その属性値の距離関数や重み係数の設定など、第A章も含めて幅広い面から詳細な検討を加えた。これらの高速化手法によって、小～中規模の画像データベースを扱うことができる見通しを得た。しかし1万枚規模の大規模な画像データベース、さらには10万枚規模の超巨大画像データベースにこの手法を適用するには、現状のままでは検索に要する時間が長過ぎる。さらなる高速化が望まれるところである。

第9章は「理解レイヤ」に相当する部分である。ここでは人間の検索意図を柔軟に理解するためのアルゴリズムの一つとして「人為選択法」を応用し、この方法に適するように改良された「ライフサイクル型遺伝的アルゴリズム(L-GA)」を提案した。その結果ブラウジングによって検索パラメータを最適化するための方法を示した。実用的には、まずは遺伝的アルゴリズムをオフラインに適用し、そのあと人為選択法に基づいたブラウジングにより局所的に最適化するという2段階の処理によって、さらに柔軟な画像検索が実現する可能性がある。

以上に述べてきた手法を3種類の画像データベース—NOAA衛星画像データベース・GMS衛星画像データベース・カラー自然画像データベース—に適用した結果を第10章に示した。この3種類の画像データベースの実装には階層モデルを用いたため、共通の手法が適用できるレイヤと個別の手法が適当なレイヤとを分離して効率的に考えることができた。これら3種類の画像データベースで類似検索の実験をおこない、その結果を検索精度の評価、検索速度の評価、柔軟性の評価の3通りの面から評価した。その結果本研究で提案した手法では、1) 最適な重み係数の調整をあらかじめ学習しておくことにより、検索目的に合致した類似度もしくは主観的な類似度を構成できること、2) 複数の重み係数セットを用いることによって、多様な検索目的に対応できる柔軟な画像データベースが構築できること、などを示した。

以上をまとめると、本研究では階層モデルという方法論に基づき、類似検索機能をもつ画像データベースを実現するための要素技術を画素レベルから抽象レベルに至るまで幅広く検討し、さらにそれらの手法を複数の画像データベースの構築に応用した。このように画素レベルから類似度という抽象的なレベルまでを総合的に論じた研究は少なく、従来は個別の要素技術の開発に終わる研究が多かった。このような状況を考えると、類似検索機能をもつ画像データベースに1つのモデルを提供するという面から、本研究は今後の画像データベース研究の発展に貢献するところが大きいと考えている。

11.2 今後の課題

本研究は画像データベースの高度化に重要な技術である類似画像検索という問題について、画素レベルから画像間類似度という抽象的なレベルまでに必要な要素技術を総合的な観点から検討することで、一つの完結した画像検索システムを構築することを目標とした。その目標はかなりの部分で達成されたと評価している。しかしこの計画を完全に実現するというのは、現状の技術レベルから考えるといささか野心的な計画であり、完璧なレベルにまで仕上げることはできなかった。以下では本研究でやり残した部分を明らかにし、今後の課題とする。

まず階層モデル全体を眺め渡してみると、「領域レイヤ」と「関係レイヤ」の部分、いわゆるパターンと記号の橋渡しをする部分が弱点となっていることが見てとれる。この部分は最も定式化の難しいレイヤであり、ゆえにさまざまな手法が提案され混沌としている部分でもある。そのため本研究で用いた手法でも、このレイヤで致命的な誤りが生じることがある。このときに根本的な問題は、本研究のようにボトムアップに画像表現モデルを構築する方法では、下位レイヤ

での誤りは上位レイヤで回復することができず、最終的な検索結果にまで大きな影響が波及するところにある。この問題の解決法としては、以下の二つの方法が考えられる。

1. 下位レイヤにより強力な手法を開発することにより誤りを減らしていく。
2. トップダウン的な処理を加えることにより、上位レイヤが大局的見地から誤りを訂正する。

確かに1の方向性ももちろん重要でありまだまだ発展の余地は大きい。しかし2の方向性もまた重要である。現在のように処理が一方向にしか進まないのではなく処理が双方向に進むことで、階層モデルがより洗練されたものになるのではないかと筆者は考えている。

次の課題はやはりグラフマッチングの高速化である。グラフマッチングは確かに強力な記述能力をもち、小～中規模のデータベースではこのような強力な画像表現モデルが検索の高度化に大きな貢献をするし、また検索時間も実用レベルにすることができる。しかし超大規模な画像データベースを想定すると、グラフマッチングは検索時間の面からかなり厳しい状況である。そこでグラフマッチングをそのまま適用するのではなく、これを「プログレッシブ」的に計算できれば計算の高速化に大きく貢献できるだろう。このような面からの研究は今までのところほとんど見られないが、今後は検討する価値のある課題であると考えている。

さらに重み係数の調整の問題もまだ検討する余地が多く残されている。現在の重み係数調整メカニズムは、遺伝的アルゴリズムをオフラインで適用して最適化したものであり、これは重み係数が「静的な」ものと捉える方法である。しかし実際のところ、人間の感性は確固としたものではなく、うつろいやすいものである。画像検索をする状況によって変わることもあれば、日時が経過するとともに次第にずれていくこともあり、このような重み係数の「動的」な面を実現していくことが今後の課題である。この時に、重み係数を大きく調整しなければならないような感性の「大変化」ではなく、むしろ重み係数の細かな調整が必要となるような感性の「ドリフト」が起こる場面が多いことが予想できる。このような細かな調整に、人為選択法を応用したブラウジングの手法が活用できそうな予感がしているが、それも今後の課題である。

今後も情報処理技術の発達に伴って、画像データベースに関する研究はさらに広がりを見せるだろう。本研究がその発展にいくらかでも貢献できることを筆者としては期待したい。

付録 A

類似度と距離

A.1 はじめに

「これとあれは似ている」という類似性の判定は、「これとあれは同じ」という同一性の判定と比較してかなり高度な情報処理である。しかし情報処理技術の発展に伴って、人間の主観的要素をも含む情報処理系を構築することに次第に関心が高まりつつある。そして客観的な類似性の判定ばかりでなく主観的な類似性の判定までもが、情報処理課題として研究が進められるようになってきた。このような研究はいわゆる「感性情報処理」の分野にも近い。

ここで最終的に求めたいものが類似度であるとしても、実際には「距離」の方が計算しやすいことが多い。距離(非類似度) D から類似度 S を導出する方法は、例えば距離の最大値 D_{\max} が分かっているならば、 $S = D_{\max} - D$ としたり、また $D \neq 0$ であれば $S = 1/D$ とするなどの方法がある。しかし実用上は無理に類似度を計算する必要はなく、距離を非類似度と見做すだけで十分である。そこで以下では、画像間の距離を測定する方法を中心に述べる [128]。

一般的に距離測度 d は次の関係を満足する [81]。

$$d(x, y) = 0 \Leftrightarrow x = y \quad (\text{A.1})$$

$$d(x, y) \geq 0 \quad (\text{A.2})$$

$$d(x, y) = d(y, x) (\text{対称性}) \quad (\text{A.3})$$

$$d(x, y) \leq d(x, z) + d(z, y) (\text{三角不等式}) \quad (\text{A.4})$$

しかしもっと広い意味の距離測度は、2つの対象がたがいにとどのくらい離れているかを表現するために考えられた測度すべてを指す。それらは必ずしも対称性や三角不等式を満足しない。しかしその測度がなんらかの意味で有用であれば類似度とよぶことが多い。このような類似測度 s が満たすべき関係は次のようなものである。

$$s(x, y) = s(y, x) \quad (\text{A.5})$$

$$s(x, x) \leq s(x, y) \quad (\text{A.6})$$

$$(x, x) \geq s(x, y) \quad (\text{A.7})$$

式(A.6)が成立するとき類似度は前向き (forward)、また反対に式(A.7)の方は後向き (backward) であるとよばれる。類似度のもつ数学的条件はゆるく、むしろそれぞれの類似度の定義が個々の問題にとってどのような意味があるかが重要である。つまり類似度の定義は、本質的に問題領域に依存している。特に人間の主観的な類似度を考えた場合、類似度というものは問題領域に依存

するばかりでなく個人個人の感性に依存し、さらには目的や知識によって動的に変化するものである。類似度のこのような性質が問題の定式化を非常に困難なものとしている。

類似度を具体的に計算する際には以下の二点が大きな問題となる。

1. 同種の画像特徴量どうしの距離を計算する関数をどのように構成するか?
2. 複数の画像特徴量がある場合にどのように重み付けすれば(主観的に)適切か?

二つの問題は別々に考慮することもあれば同時に考慮することもある。以下ではこの二つの問題を明確に区別していく。

A.2 客観的な距離

「客観的な距離」とは人間の主観的な要素を含まない距離のことを指す。代表的なものに特徴ベクトルの距離がある。特徴ベクトルとは画像を表現するモデルとしては最も単純な、構造を持たないモデルである。画像の特徴量が K 個であれば、その K 個の特徴量を並べたベクトル $x = (x_1, \dots, x_K)$ が画像特徴ベクトルとなる。最も簡単な画像特徴量は画素値そのものであるが、一般的には画像を何らかの画像表現モデルで表現する過程で有用な画像特徴量を抽出し、それらの特徴量を並べて特徴ベクトルとして用いることが多い。

物理特徴空間上の距離 [81]

以下で述べる距離では、同種の特徴量の差分が距離の基本となっている。

ユークリッド距離 (Euclidean Distance)

$$d_{ij} = \left\{ \sum_{l=1}^K (x_i^l - x_j^l)^2 \right\}^{1/2} \quad (\text{A.8})$$

ユークリッド距離は最も基本的な距離定義であり、 K 次元空間の物理的距離と一致する。

重み付きユークリッド距離 (Weighted Euclidean Distance)

$$d_{ij} = \left\{ \sum_{l=1}^K w_l (x_i^l - x_j^l)^2 \right\}^{1/2} \quad (\text{A.9})$$

重み付きユークリッド距離は K 次元空間の各次元を一様には扱わない。各次元の重要度を変えることにより、物理的距離としてではなく何らかの評価規準に基づいた距離を定める。客観的な距離として用いるのであれば、重みとして例えば分散の逆数を用いることができるし、むしろこの重みを積極的に活用することで、この距離を主観的な距離としても使うことができる。

マハラノビス距離 (Mahalanobis Distance)

$$d_{ij} = \left\{ \sum_{l=1}^K \sum_{m=1}^K (x_i^l - x_j^l) w_{lm} (x_i^m - x_j^m) \right\}^{1/2} \quad (\text{A.10})$$

w_{lm} は分散共分散行列 V の逆行列 $W = V^{-1}$ の (l, m) 要素である。それぞれの特徴量が多次元の正規分布にしたがう場合に最も自然な距離を与える。すべての共分散を無視できる場合には、各特徴量の分散の逆数で重みづけした、重み付きユークリッド距離に帰着する。

市街地距離 (City Block Distance)

$$d_{ij} = \sum_{l=1}^K |x_i^l - y_j^l| \quad (\text{A.11})$$

加算・減算だけで計算できる高速な方法であり、2次元画像空間では4近傍距離ともよばれる。

チェス盤距離 (Chessboard Distance)

$$d_{ij} = \max_l [|x_i^l - y_j^l|] \quad (\text{A.12})$$

2次元画像空間の場合は8近傍距離ともよばれる。

Minkowski 距離 (Minkowski Distance)

$$d_{ij} = \left\{ \sum_{l=1}^K |x_i^l - y_j^l|^r \right\}^{1/r} \quad (\text{A.13})$$

$r = 1$ が市街地距離、 $r = 2$ がユークリッド距離に相当する。また Minkowski 距離で r を無限大にしたときの距離

$$d_{ij} = \max_l [|x_i^l - y_j^l|] \quad (\text{A.14})$$

は Chebyshev 距離 (l_∞ 距離) とよばれる。

特徴量の物理的性質を考慮した距離

一般的には特徴量の差分が距離の基本となるが、すべての特徴量に対してそれが適切な計算方法であるわけではない。例えば角度に基づいた特徴量の距離を考えよう。無向の場合は角度 π で一周するという性質がある。すると角度で π 離れたものの類似度が高いことを考えると

$$d_{ij} = -||x_i - y_j| - \pi/2| + \pi/2 \quad (\text{A.15})$$

という距離を考えることができる。ただし ($-\pi/2 < x_i, y_j \leq \pi/2$) である。このように特徴量の物理的な意味を考えると、こちらの距離の方が単純な差分よりも適切であるといえる。つまりユークリッド距離はあくまで基本に過ぎず、むしろ重要なことは特徴量の性質をよく考慮した距離を考えることである。この過程で距離測度の条件を満たさなくなるかもしれないが、それはやむを得ない。

A.3 主観的な距離

以上のように客観的な距離を計算する方法をいくつか紹介した。しかしいずれも人間が画像を見るという要素を考慮していないため、必ずしも人間の主観的な類似度に対応はしていない。つまり特徴量の大きさと人間の感覚量の違いとを考慮しない限り、人間の主観にあった距離は計算できないのである。さて、もしすでに行われた研究によって人間の認知特性がある程度わかっているならば、その知見から演繹的に主観的な距離空間を構成できる可能性がある。例えば下位レベルの視覚や色彩に関する問題では、主観的な距離をある程度は演繹的に導くことができるだろ

う。しかしさらに上位レベルの問題で人間の認知モデルがまだ知られていない場合には、天降り式に主観的な距離を与えることはできない。結局は人間が与えた教師データから帰納的に主観的な距離を学習するか、または心理学的な実験によって距離関数を構成する必要が生じる。本節ではこのように主観的な距離を計算する方法を述べる。

A.3.1 対数距離

一般に Fechner の法則として知られているように、刺激の強さとそれに対する感覚量の間には対数法則が成り立つ場合がある。例えばデシベルなどの単位はこの性質を根拠とするものである。このような場合の距離は比を基本にして、

$$d_{ij} = \log \frac{x_i}{y_j} \quad (\text{A.16})$$

と定義できる。ここで $(x_i, y_j > 0)$ が任意の特徴量に対して成り立つとする¹。式 (A.16) を変形すると

$$d_{ij} = (\log x_i - \log y_j) \quad (\text{A.17})$$

となる。この形式は特徴量の差分ではないが、特徴量の対数の差分である。一般的に、特徴量の関数の差分をユークリッド距離の拡張として捉えることができる。

A.3.2 ヒューリスティックな距離

ヒューリスティックな距離とは、問題領域に関する事前の知識を用いて、問題を解くために適切な距離関数を作り出す方法である。たとえば各属性値の差異が大きいほど心理的距離が指数関数的に大きくなると仮定すれば、次のような距離関数を考えることができる [129]。

$$d_{ij} = \sum_{l=1}^K w_l \exp \left\{ -c_l (x_i^l - y_j^l) \right\} \quad (\text{A.18})$$

ここで w_l は重み付けであり、また c_l は減衰の速さを決める定数である。このように、ヒューリスティックな距離関数は自由に作り出すことができるものであり、どちらかという理論的な根拠が薄いモデルである。しかし逆に物理的な空間の制約に縛られることなく、問題の解決に最適な距離を自由に決めることができるし、実際の問題ではこのようなヒューリスティックな距離が必要とされ重要な役割を果たすことも多い。

A.3.3 心理学的尺度構成法

名義尺度の記号間で類似度を計算したい場合、または順序尺度だけが得られていてそれを間隔尺度に直したいという場合に有効な手法が心理学的尺度構成法である。もともとは心理学の分野で主に用いられてきた手法である。

¹特徴量がゼロになる場合はそのままでは定義できないが、ゼロ付近を線形に近似することで実用的に定義することができる。

Stevens の 4 分法 心理学的尺度構成法 (psychological scaling) について、1次元尺度構成 (uni-dimensional scaling) に限って述べる。まず直接法と間接法の二つに分類する方法がある。すなわち直接法とは刺激の物理量に対応する心理量を直接に数で表現するという方法であり、間接法とは心理量が正規分布するという統計的仮定に基づいて間接的に尺度を構成する方法である。ここで重要な概念が Stevens の 4 分法である。

1. 名義尺度 (nominal scale) : 分類するためだけに数値が用いられている尺度
2. 順序尺度 (ordinal scale) : 数値の順序情報に意味をもたせた尺度
3. 間隔尺度 (interval scale) : 数値と数値の間隔に意味をもたせた尺度
4. 比率尺度 (ratio scale) : 数値と数値の比に意味をもたせた尺度

類似度として用いるためには、尺度が間隔尺度か比率尺度になっている必要がある。そこで間隔尺度の構成法として有名な Thurstone の方法について述べる。

Thurstone の尺度構成法 まず刺激を提示する方法について述べる。刺激二つを対にして示し、そのいずれが強いか、あるいは好きかを比べさせるのが一対比較法 (method of paired comparison) であり、一つずつ刺激を示してはその心理量を評定させるのが単一刺激法 (method of single stimuli) である。前者は比較判断であるため評定が容易だが、比較せねばならない対の数は nC_2 と多い。一方後者は比較でないため適切な評定がしにくい、刺激数は n と少なくすむ。

Thurstone の方法では一対比較法を用いる。そして、人間の差の判断が正規分布するという仮定がこの方法の要点である。 j なる対象が k なる対象よりも好きという被験者の比率を P_{kj} としよう。 $P_{jk} = 1 - P_{kj}$ 、 $P_{jj} = 0.5$ とする。またこれから求める尺度値をそれぞれ S_j 、 S_k とする。ここで尺度値は、大きいほど好きという尺度になっている。さて各個人による尺度値のばらつきを考慮すると、被験者全員に対する $S_j - S_k$ の分布は中心値のまわりにばらつくことになる。ここで $S_j - S_k$ の分布は正規分布すると Thurstone は仮定する。そこで $S_j - S_k$ の分布の平均を X_{kj} とおく。通常は $S_j - S_k$ の分布の分散が 1 であると仮定する²ので、この正規分布は $N(X_{kj}, 1)$ となる。さて比率 P_{kj} は $x > 0$ の範囲にある $S_j - S_k$ の分布の面積と解釈できるので、図 A.1 を参考にすると、 P_{kj} の値から X_{kj} を求められることがわかる。このようにしてすべての j と k の組合せに対して X_{kj} の値を求めれば、最終的に各対象の尺度値は、全体の対象の数を n とすると

$$S_j = \frac{1}{n} \sum_{k=1}^n S_k + \frac{1}{n} \sum_{k=1}^n X_{kj} \quad (\text{A.19})$$

によって求めることができる。さらに式 (A.19) の第 1 項はすべての尺度値に共通であるため、通常は尺度値の原点をずらすことで第 1 項を消去する。

A.4 特徴量の重み付け

第 A.1 節でも述べたように、類似度には二つの側面がある。ここでは特徴量の重み付けの問題を扱う。人間の主観を適切に反映する重み付けを求める問題自体が、実は大きな問題である。

²このように仮定する理由は、このように仮定しないと問題が解けないからである (ケース Va の仮定とよばれる)。

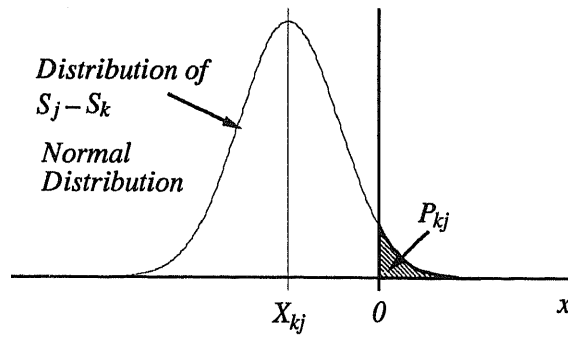


図 A.1: Thurstone の方法の模式図。

なぜなら、画像特徴量として通常用いられるのは画像処理が容易という観点から選ばれた特徴量であることが多く、したがって人間にとって有効な特徴量とは異なるからである。そのため、人間の頭の中にある特徴量と抽出された画像特徴量との間にギャップが生じてしまう。

このような場合には、物理的画像特徴パラメータから主観的類似度を計算によって求めるメカニズムが必要になる。そのための一つの方法は、比較的少数のキーパターン間の類似度をシステムに示し、システムに物理的画像特徴のパラメータ空間上で例示学習(統計的学習)させる方法がある。そして学習によって、物理的画像特徴パラメータ空間から主観的類似尺度を反映した主観的画像特徴パラメータ空間に変換する写像を構成する。もし画像が特徴ベクトルで表されており、しかも距離空間がユークリッド空間ならば、基本的にはこの問題を多変量解析を用いて解くことができる。

利用者の主観的な類似度を反映する空間を判別分析を用いて構成する研究が提案されている[6]。いま対象画像の部分集合 $G = \{g_i | i = 1, \dots, N\}$ が主観的類似度にしたがって K 個のグループに分類されているとする。また各対象画像に対して N 次元の特徴ベクトルが得られている。このとき線形判別分析は、特徴ベクトルの線形変換

$$y = {}^t A x \quad (\text{A.20})$$

により、同じグループの画像はなるべく近く、異なるグループの画像はなるべく遠くに配置する空間(判別空間)を構成する手法である。変換行列 A は以下の固有値問題から求まる。

$$\Sigma_B A = \Sigma_W A \Lambda \quad (\text{A.21})$$

$${}^t A \Sigma_W A = I \quad (\text{A.22})$$

ただし Σ_B 、 Σ_W はそれぞれグループ間共分散行列、グループ内共分散行列である。

また対象画像の部分集合 $G = \{g_i | i = 1, \dots, N\}$ に対して、任意の二つの画像の類似度を利用者が与える。いま特徴ベクトル x_i と x_j の類似度を s_{ij} とする。再び特徴ベクトルを式(A.20)によって線形変換する。そして利用者が与えた類似度 s_{ij} と、変換後の特徴ベクトル y_i 、 y_j の内積 $y_i {}^t y_j$ の積の期待値を最大にするように設定する。変換行列 A は以下の固有値問題から求められる。

$$\Sigma_E A = \Sigma_X A \Lambda \quad (\text{A.23})$$

$${}^t A \Sigma_X A = I \quad (\text{A.24})$$

ただし Σ_E は重み付き分散共分散行列、 Σ_X は分散共分散行列である。このように多変量解析の手法を用いて最適な重み付けを発見し、主観的類似度空間を構成することができる。

付録 B

スケールスペースを用いたヒストグラム解析

本研究の随所には「スケールスペースを用いたヒストグラム解析」が用いられているが、今まで特に説明は加えていない。この方法は今までそれほど広く用いられているわけではないと考えられるため、本章ではこの手法の説明から背景までを簡単に紹介する。初めに導入として、「スケールスペース」とは何かを第 B.1 節で説明し、合わせてスケールスペースの特徴についても述べる [130]。次にその特徴を利用した手法として、第 B.2 節ではスケールスペースを用いたヒストグラム解析の二つの応用について述べる。

B.1 スケール

B.1.1 画像とスケール

人間がものを見る時、どのように見るだろうか？その方法を分析してみると、おおよそ以下のようにしていることが推測できる。すなわちまず対象を大まかに眺めて全体の様子をとらえ(大局的観測)、必要に応じて部分的な特徴に注意を向けて詳しく観察する(局所的観測)というように、観測するスケールを適応的に変化させているだろう。このような階層的な観測パターンは、人間の高度な視覚情報処理にとって基本的なものである [131]。

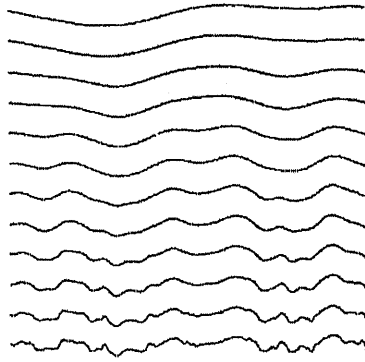
さてこれから画像とスケールの関係について述べていくが、基本的には画像は表 B.1 に示すような 2 つのスケール — 外部スケールと内部スケール — に制約を受けている [132]。そして画像中に存在する「重要な」構造は、その中のある限られたスケール範囲に存在することが多いが、しかしそのスケール範囲についての知識は事前に知られていないのが普通である。

それでも人間の視覚システムは、適当なスケール範囲に「ズームイン」する能力を持っているので、通常は特に問題は生じない。しかもさらに優れたことに、画像を同時にいくつかのレベルの解像度で観測し処理する能力も備えている。それに対し、計算機に画像認識をさせる場合には、何らかの方法で観測に最適なスケールを画像そのものから決定しなければならない。さらには、異なる解像度で得られた情報を統合できるような表現も必要である。

このようなシステムでは、画像中の大きな構造を捉えるために、画像のぼかし (blurring) / 平滑化 (smoothing) という操作が基本的な操作となる。さらにそのぼかしの程度を表すスケールパラメータを導入する。ぼかしを大きくすることで画像中の小さな構造は平滑化され、大きな構造のみが残ることになる。このようにスケールパラメータを可変とし、多重スケールの観測をおこなうことが、優れた視覚システムの基本的なアイデアなのである。

表 B.1: 画像のスケール。

スケール	解釈	人間の視覚システム
外部スケール (outer scale)	画像の広がり (窓) が限られていること	網膜の広がり
内部スケール (inner scale)	画像の解像度が限られていること	網膜受容野の離散的構造



(a) 波形の変化

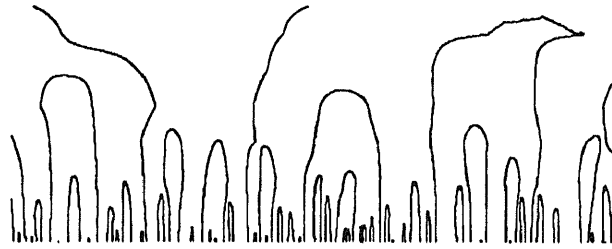
(b) $F_{xx} = 0$ の変化

図 B.1: スケールスペースの例。下から上へ、スケールパラメータ σ の増加に伴う輪郭の変化。

B.1.2 スケールスペース

従来から多重スケールを扱った研究は散発的に見られたが [133]、スケールスペース (scale space) という概念を提案し、それに統一的な表現を与えたのが Witkin である [134]。スケールスペースでは波形を階層的に表現するために、スケールパラメータが可変のガウシアンフィルタで原波形をさまざまなスケールにぼかすことによって得られる波形集合を考える。ここで1次元信号 $f(x)$ のスケールスペース画像 $F(x, \sigma)$ を求めよう。これは信号 $f(x)$ とガウス関数 $g(x, \sigma)$ のたたみこみであらわされる:

$$F(x, \sigma) = f(x) * g(x, \sigma) = \int_{-\infty}^{\infty} f(u) \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-u)^2}{2\sigma^2}} du \quad (\text{B.1})$$

上式で*はたたみこみを表す。このとき、 (x, σ) 平面をスケールスペース、また関数 $F(x, \sigma)$ をスケールスペース画像とよぶことにする。さて図 B.1(a) のように、ガウシアンフィルタは波形を平滑化する効果をもつため、大きな σ に対応する波形では細かい変化が消えている。このことから σ は、波形を観測するスケールに対応するパラメータ — スケールパラメータであることがわかる。これを連続的に変化させることにより、多重スケールにまたがる顕著な波形構造を発見する、というのがスケールスペースの基本的な考えである。

さらにスケールスペース画像の2次微分のゼロ交差 (zero crossing)¹を考える。ゼロ交差のスケールに対する変化を追跡したのが図 B.1(b) である。最も粗いスケールまで続いている輪郭が波形中の最も顕著なエッジに対応し、スケールパラメータを小さくするにしたがって次第に細かいエッジに対応する輪郭が現れてくる。こうして画像中の大きな構造と小さな構造を統一的に扱うことが可能となった。

¹信号 $f(x)$ の1次微分のゼロ交差 $f_x = 0$ は信号の極値に対応し、2次微分のゼロ交差 $f_{xx} = 0$ は信号の変曲点に対応し、一般にエッジ部分にも対応する。

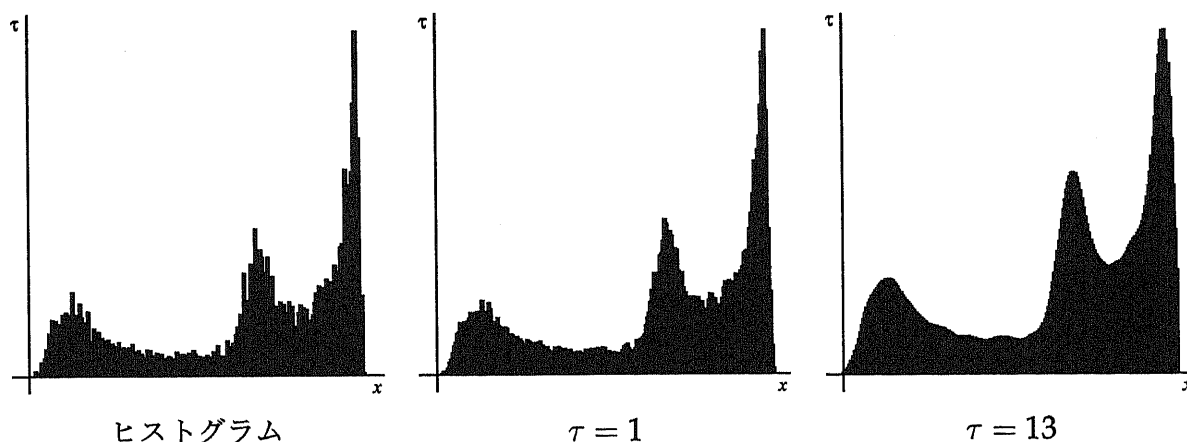


図 B.2: ガウシアンフィルタを用いてヒストグラムをぼかした例。 τ がガウシアンフィルタの標準偏差を示している。

B.1.3 基本的な要請

このように分析ができるスケールスペースを構成するためには、いくつかの基本的な要請がある。その中でも最も本質的な要請は因果律 (causality) である [132]。

因果律 スケールスペース画像を細かいスケールから粗いスケールへとたどっていったとき、その途中で新たに輪郭を生成してはならないこと。言い換えれば、粗いレベルの解像度に存在する特徴は、必ず細かいレベルの解像度に原因があること。

スケールスペースを構成する手法は、どの手法でも必ず因果律を満たす必要がある。なぜならこの因果律によって、

1. まず粗い解像度で特徴を発見 (identification)
2. 次に細かい解像度で位置を決定 (localization)

いう処理 (coarse-to-fine) の妥当性が保証されるからである。その他にも同質性 (homogeneity) や等方性 (isotropy) が要請されることも多いが、この要請は必ずしも本質的なものではないことが Perona らによって指摘されている [135]。

B.2 スケールスペースを用いたヒストグラム解析

このようなスケールスペースを用いたヒストグラムの解析法が、Carlotto によりすでに提案されている [79]。ヒストグラムを混合密度とみなしてこの混合密度を推定する問題に Carlotto はこの手法を応用し、衛星画像から得られた反射スペクトルのヒストグラムを分析して地表面の分類に用いている。

以下では具体的にヒストグラムを眺めながらその方法を詳しく説明する。まずスケールスペースを用いてぼかしたヒストグラムの例を図 B.2 に示す。もとのヒストグラムでは多数の小さな山や谷が密集しているが、 $\tau = 13$ 程度のガウシアンフィルタ²でぼかすと、3 個の大きな山が明瞭に現れてくる。ここでの重要な性質はスケールスペースの因果律である。つまり因果律が成立す

²なおガウシアンフィルタの核 (kernel) としては、ガウシアンフィルタの離散的な核を使用している [136]。

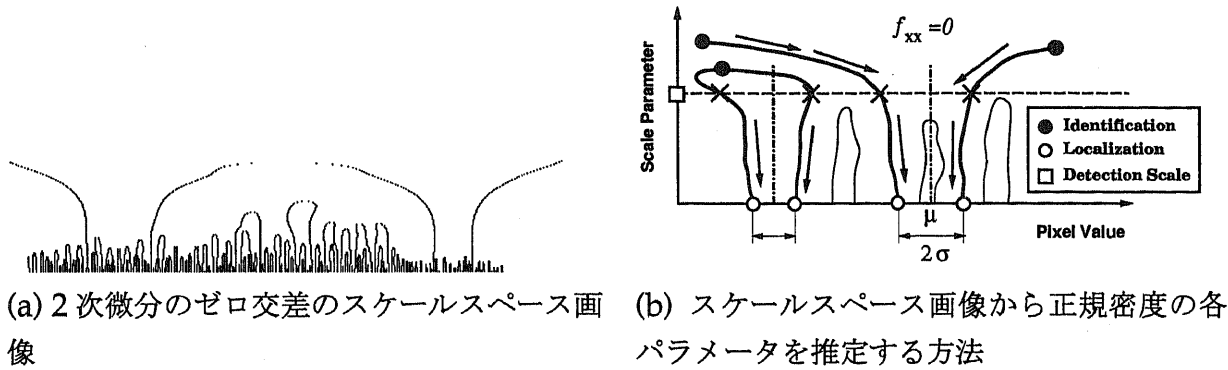


図 B.3: スケールスペースを用いた初期値推定。

ることにより、ヒストグラムをぼかしても新たな山が出現しないことが保証されている。そのためここで存在している山は、細かいスケールからずっと残っている顕著な山に対応するピークであると結論できるのである。またこのようにぼかして検出することによって、雑音に影響されにくいロバストな手法とすることができる。

一般的にスケールスペースを用いた方法では、波形から何らかの特徴点を抽出し、その特徴点のスケール変化を追跡することで顕著な特徴点を抽出することを目的としている。この特徴点としては

1. 1次微分のゼロ交差 ($f_x = 0$)
2. 2次微分のゼロ交差 ($f_{xx} = 0$)

という特徴点が代表的である。本研究では処理の目的によってこの二つの特徴点を使い分ける。

B.2.1 2次微分のゼロ交差を用いた混合密度の初期値推定

まず画像のヒストグラムが混合密度であるとし、またそれぞれの密度関数がピークの両側に1個ずつの2次微分ゼロ交差をもつと仮定する。例えば正規分布やコーシー分布はこの仮定を満たす。次にそのゼロ交差の位置をスケールパラメータ τ を変化させながら記録していく。その例を図 B.3 に示しながら。スケールパラメータ τ の時のゼロ交差の個数を $z(\tau)$ とすると、 $z(\tau)$ は τ が大きくなるにしたがって小さくなる。つまり顕著な密度に対応するゼロ交差に対応する軌跡が大きな τ まで残るのに比べ、顕著でない密度に対応するゼロ交差の軌跡は小さな τ で消えてしまう。例えば図 B.2 では、 $\tau = 13$ の時点で2次微分のゼロ交差がすでに6個に減少している。

この性質を利用して、顕著な分布の順に K 個の分布のパラメータを推定することができる。 K 個の分布には $2K$ 個のゼロ交差に対応するため、大きなスケールから小さなスケールへと τ を小さくしていき、 $z(\tau) = 2K$ となる τ を発見する。次にこのスケールで左から右に向かって隣り合うゼロ交差を組み合わせることで K 個のゼロ交差のペアを得る。さらにこれら $2K$ 個のゼロ交差の軌跡を最も小さい τ まで追跡するという coarse-to-fine に基づいた処理によってゼロ交差の位置を特定する。

以下では正規密度の場合の推定方法を示す。まず正規密度 $N(\mu, \sigma)$ の2次微分のゼロ交差は、

$$x = \mu \pm \sigma \quad (\text{B.2})$$

で与えられる。さてこの正規密度をスケールパラメータ τ のガウシアンフィルタでぼかすと、ぼかされた密度関数は $N(\mu, \sqrt{\sigma^2 + \tau^2})$ となる。この性質と式 (B.2) により、ペアとなった2個のゼロ交差の間隔 d は $d = 2\sqrt{\sigma^2 + \tau^2}$ となる。この間隔は τ が大きくなるにしたがって広がるため、スケールスペース上ではゼロ交差のペアは煙突状の形状を示すのが特徴的である。また $\tau \sim 0$ ではゼロ交差の間隔 d が 2σ で近似できる。以上の考察より以下のパラメータ推定法を得る。

1. $z(\tau) = 2K$ となるスケールを発見し、そのスケールで隣接するゼロ交差をペアに組み合わせる。次にペアのゼロ交差の位置 x_1 と x_2 を細かいスケールで特定する。さらにゼロ交差間の距離 $d = (x_2 - x_1)$ を測定する。
2. 平均値 μ については、 $\mu = (x_1 + x_2)/2$ により計算する。
3. 標準偏差 σ については、 $\sigma = d/2$ により計算する。
4. 混合比率 α については、 $\text{Area}(\mu - \sigma \leq x \leq \mu + \sigma) = 0.683\alpha$ を利用して、 $\alpha = \text{Area}(x_1 \leq x \leq x_2)/0.683$ により計算する。

B.2.2 1次微分のゼロ交差を用いたヒストグラムの多値化

1次微分のゼロ交差点は、ヒストグラムの谷または山に対応する部分である。特に谷の部分には、1次微分がマイナスからプラスへと変化するゼロ交差点が存在する。ここで、ヒストグラムの谷の部分にしきい値を設けてヒストグラムを多値化することを考えると、このようなゼロ交差点を粗いスケールで発見し、そのゼロ交差点を細かいスケールまで追跡することで、顕著な谷を発見してしきい値を定めることができる例えば濃淡画像のヒストグラムを多値化する際には、ヒストグラムを適当なスケールパラメータでぼかし、ぼかしたヒストグラム上で谷に対応する1次微分のゼロ交差を発見し、そのゼロ交差を細かいスケールまでたどることによってしきい値を決定する。

本研究ではさらにカラー画像の領域分割に、この手法を応用する。それは Ohlander の方法に基づいた領域分割の手法である。Ohlander の再帰的しきい値処理とは、カラー画像を画素ごとの色情報を用いて領域分割する手法であるが、特徴空間でいきなり分割するのではなく、分割しやすい領域から逐次抽出していく手法である。本研究で応用した Ohlander の方法の具体的な処理の流れは以下のようなになる。

1. カラー画像から3個のヒストグラムを計算する。最初は全画面に対してヒストグラムを計算する。
2. ヒストグラムの中で双峰性または複数のピークをもつヒストグラムを探し、そのヒストグラムの顕著な谷の部分分割のための(複数の)しきい値とする。
3. このしきい値によって画像を複数の連結領域に分割する。そしてこの分割によって新たに生まれた連結領域、およびその他の残りの連結領域それぞれに対し、すべてのヒストグラムが単峰性となるまで分割をおこなう。

まずカラー画像の色空間としては RGB 色空間を用いる。他の色空間を用いた実験、例えば HSI 空間や HSV 空間、 $L^*a^*b^*$ 空間、 $L^*u^*v^*$ 空間なども試してみたが、平均的に最もよい結果となったのは RGB 色空間であった。そこで今後の実験では RGB 色空間を用いることにする。

そして分割に適したヒストグラムを発見する手法に、スケールスペースを応用した手法を用いる。それぞれのヒストグラムを同一のスケールパラメータでぼかし、平滑化されたヒストグラム上で顕著な谷の個数を数える。そして色空間の3個のヒストグラムの中で、顕著な谷の個数が最も多いヒストグラムを分割に適したヒストグラムとし、その谷に対応するしきい値を用いて画像を分割して次の繰り返しに移る。ちなみにガウシアンフィルタのパラメータである分散の値は $\sigma^2 = 64$ とした。

ただし Ohlander の方法はあくまで特徴空間での分割であるため、画像によっては画像空間上には小領域が多数出現することがある。そのため、あるしきい値以下の面積しかもたない小領域を、隣接する大領域に併合する処理も同時におこなう。この併合の規準は、小領域と8隣接関係にある領域を調べ、その中で最大の面積をもつ領域に小領域を併合するという、単純な評価規準を用いている。もちろん色情報も使えばさらに併合結果が向上する場合もあるが、著者らの実験ではほとんど結果の差が観察されなかったため、簡便な前者の方法で十分であると考えた。

付録 C

ボリュームレンダリングを用いた気象衛星画像の3次元可視化

C.1 はじめに

画像データを有効に活用するためには、画像を生データそのままの形で提供するだけでなく、画像に含まれる情報をさまざまな形で可視化 (visualization) した上で提供する技術が近年になって注目されている。例えば気象衛星画像は2次元の画像として受信されるが、多様な観測チャンネルを活用して受信画像を処理すれば雲の立体構造を把握できることが知られている。そこで本研究では、気象衛星画像上の雲を3次元的に可視化する手法を提案する。このように3次元的に可視化することによって、これまで2次元的に眺めていた画像からは得られない新たな知見が得られることから、本研究は画像情報を新たな視点から活用するための研究と位置付けられる。本研究が可視化する対象は実世界の観測データであり、したがって本研究は、従来のコンピュータグラフィクス (CG) のように仮想的な雲を合成するアプローチとは異なり、実世界のボリュームデータを可視化するボリュームビジュアライゼーション (VV) の手法を用いる [137, 138]。

C.2 本研究の特徴

気象衛星画像は、厳密な意味では3次元のボリュームデータを記録したものではない。気象衛星が地球上空からしか観測できない以上、観測データは必然的に2次元データとなる。しかし赤外領域で観測した画像では、画素値が対象領域の放射温度を表すという特性がある。そこで、大気の垂直温度分布を仮定すれば雲頂温度と雲頂高度 (雲表面) を対応づけることが可能となる。この意味で赤外面像は距離画像 (range image) の性質を備えており、この性質をうまく利用して雲表面を3次元可視化することが本研究の目標である。

さて、地形図などの3次元可視化の技術はすでに一般的な技術といえるが [139]、雲の3次元可視化は以下のような困難な問題を抱えている。

1. 地形図ならば、事前に標高データが用意できるので、ある地点の標高を容易に知ることができるが、雲の場合は、画像ごとに雲の位置も高度も異なる。
2. 地面は不透明な物体であるのに対し、雲は半透明な物体である。
3. 地面は数学的には3次元空間内の一価関数曲面とみなすことができるが、多くの雲は地上と接せずふわふわと浮いているため、ある (x, y) の点で2個以上の値をもつ関数曲面と考える必要がある。

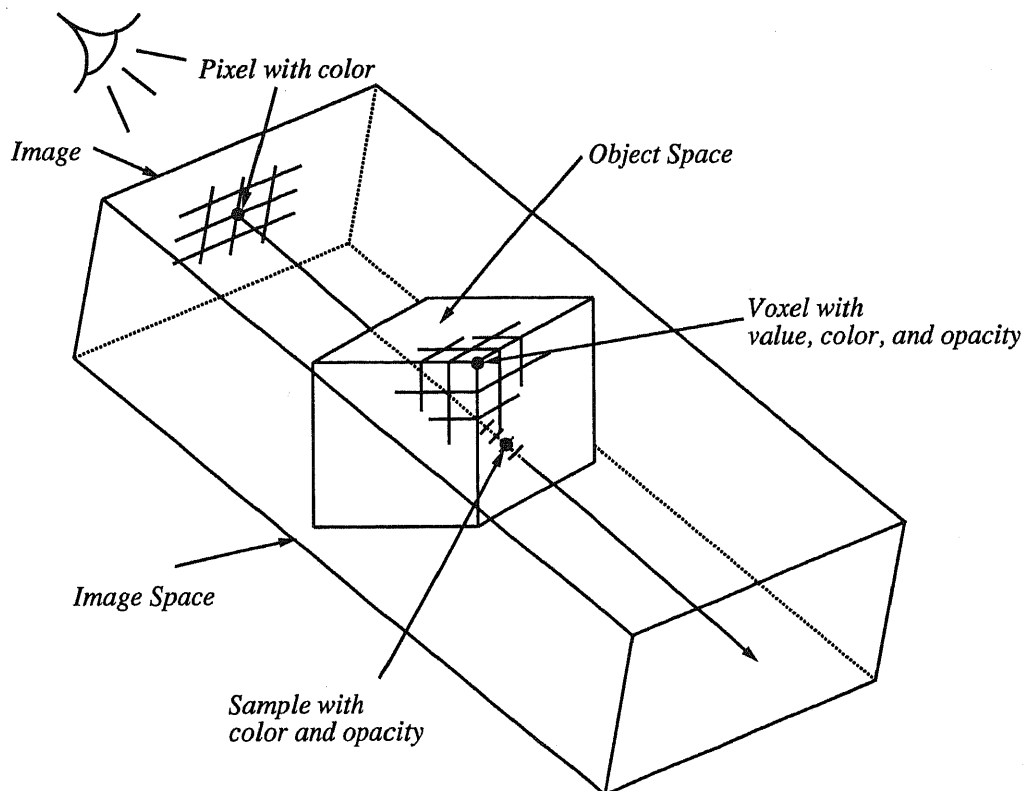


図 C.1: ボリュームレンダリングで用いられる座標系。この例は平行投影のものであるが、本研究では中心投影を用いていることに注意。

4. 雲と地面を画像上で区別する必要があるため、何らかの分類処理が必要である。

雲を立体表示する従来の方法では、特に3の処理がほとんどなされていなかった。しかしこの処理を欠くと、温度画像が連続的に変化する場合に雲が地面に接することになってしまい、立体的ではあるが不自然な画像しか生成できない。そしてこの問題をうまく解決するためには、根本的に4の分類処理が重要となるのである。

雲の3次元可視化に関する従来の研究では、赤外面像で雲頂高度を算出し、計算された雲表面に可視画像をテクスチャマッピングして陰影をつける方法が採られた [140]。それに対して本研究は画像データとして赤外面像のみを用い、ボリュームレンダリング (VR) の手法で3次元可視化する。この方法では、実際の太陽の位置とは異なる場所から異なる光源色で照明するシミュレーションを繰り返しながら、雲のいろいろな部分の立体構造を強調することが可能となる。

C.3 ボリューム・ビジュアライゼーション

C.3.1 ボリュームデータの表現と内部構造

3次元空間の至る所に存在するデータをボリュームデータと呼ぶ。これらのデータの多くは六面体格子構造を用いて表現するため、図 C.1中の各格子点を、ボリュームを構成する基本要素という意味からボクセルと呼ぶ。そして各ボクセル上で離散的に定義される対象データをフィールドデータ (field data) と呼び、特にフィールドに指定されるデータが一つだけの場合をスカラーフィールドと呼ぶ。さて、各軸方向に隣接する計8個のボクセルは、セル (cell) と呼ばれる六面

体の部分空間を構成する。このセル内の任意位置のフィールド値は **trilinear** 補間¹により計算されるのが普通である [141]。ただし本研究では別の方法を用いた。これは、セル内の任意位置から近傍の 8 個のボクセルまでの距離をそれぞれ計算し、その距離の逆数を重みとした重みつき平均を、任意位置でのフィールド値とする方法である。

C.3.2 ボリュームデータの可視化手順

この手順はインダイレクトアプローチとダイレクトアプローチの 2 種に大別される。ここでは特にダイレクトアプローチについて説明する。

ボリュームレンダリングでは、視点と各ピクセルとを結ぶ射影光線に沿って、その光線 (ray) が交差するセルの各点における光学的フィールド値や、再生された微小サーフェスの勾配値などを再サンプリング (resampling) する。このとき各ボクセルのフィールド値として不透明度 (opacity) が定義されており、再サンプリングは不透明度が 1 に到達するまで続けられる。そして、各サンプル点でのシェーディング結果を合成することで、対応ピクセルの色を決定する手法がよく用いられる。このような原理に基づく手法をボリューム光線追跡 (volume ray-tracing) と呼ぶ。ただし、サーフェスレンダリングにおける精密な光線追跡とは異なり、反射・屈折までは通常は考慮しない。その代わりに光線方向のフィールド値の線積分を実行し、ボリュームを通過する光線の減衰をなるべく忠実に再現する。しかし計算量が膨大となるため、ボリューム光線追跡の効率的なアルゴリズムも提案されている [142]。

本手法も基本的に先に説明したアルゴリズムに基づいている。すなわち、光線に沿って一定間隔でサンプリングを繰り返していき、それによって各ピクセルの色を決定する方法である。ある光線 \mathbf{u} 上のサンプル点 \mathbf{U} での色と不透明度をそれぞれ $C(\mathbf{U})$ 、 $\alpha(\mathbf{U})$ とすると、透明度方程式 (transparency formula) は以下のようなになる。

$$\hat{C}_{\text{out}}(\mathbf{u}; \mathbf{U}) = \hat{C}_{\text{in}}(\mathbf{u}; \mathbf{U}) + \hat{C}(\mathbf{U})(1 - \alpha_{\text{in}}(\mathbf{u}; \mathbf{U})) \quad (\text{C.1})$$

$$\alpha_{\text{out}}(\mathbf{u}; \mathbf{U}) = \alpha_{\text{in}}(\mathbf{u}; \mathbf{U}) + \alpha(\mathbf{U})(1 - \alpha_{\text{in}}(\mathbf{u}; \mathbf{U})) \quad (\text{C.2})$$

ただし、 $\hat{C}_{\text{in}}(\mathbf{u}; \mathbf{U}) = C_{\text{in}}(\mathbf{u}; \mathbf{U})\alpha_{\text{in}}(\mathbf{u}; \mathbf{U})$ 、 $\hat{C}_{\text{out}}(\mathbf{u}; \mathbf{U}) = C_{\text{out}}(\mathbf{u}; \mathbf{U})\alpha_{\text{out}}(\mathbf{u}; \mathbf{U})$ 、および

$$\hat{C}(\mathbf{U}) = C(\mathbf{U})\alpha(\mathbf{U}) \quad (\text{C.3})$$

このように各光線に沿って不透明度が 1 に達するまで²サンプリングを続ける。そして各点での色を式 (C.1) に従って次々と計算していくことにより、各ピクセルの最終的な色が定まる。通常の VV では興味ある部位を選択的に表示するために不透明度を積極的に用いるが、本研究ではむしろ雲の 3 次元像の再現を念頭に置いて不透明度の値を決定する。

C.3.3 分類

可視化する前には、あるボクセルがどのような物質で構成されているかを知っておく必要がある。そうでないと、各ボクセルに対して適切なフィールド値を割り当てることができない。そ

¹ $f(x, y, z) = c_0 + c_1x + c_2y + c_3z + c_4xy + c_5yz + c_6zx + c_7xyz$ の各係数を、隣接する 8 個のボクセルでのフィールド値を用いて決定する方法。

²計算量を削減するためには、不透明度が 1 に十分近い値に達した時点で打ち切ってもよい [142]。

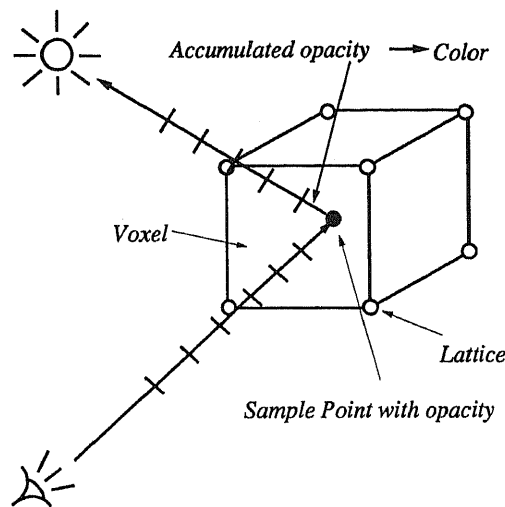


図 C.2: サンプル点の色の決定。

の際、あるボクセルを単一の物質に無理に分類する必要はなく、むしろいろいろな物質がどのくらいの割合で混合しているか、という情報が失われない方がよい。そのような目的に適した分類法が確率的分類法 (probabilistic classification) であり、またその後のレンダリングで人工的な特徴が出現しないようにするには、2 値的な分類法ではなく確率的な分類法を使うべきである [143]。

このような分類には、本研究の第 5 章で提案した、ミクセルを含む混合密度推定による画像分類の手法が非常に適している。なぜなら各画素はピュアピクセルかミクセルとして表現することができ、これがちょうどボリュームレンダリングに適した表現となっているからである。

C.3.4 シェーディングモデル

各ピクセルの色は基本的にレンダリング方程式によって決定される [144]。

$$\text{色} = \text{周囲光} + \text{拡散光} + \text{鏡面反射光} + \text{鏡面透過光} \quad (\text{C.4})$$

ここで周囲光は、通常どこの部分でも一定値とする。また拡散光とは、面のあらゆる方向に一樣に反射される光で、視線の向きとは無関係に反射されるものである。それに対して鏡面反射光とは、面の一部がキラッと輝く現象 (ハイライト) に対応する光で、その強さと色は面法線・光源座標・視点座標から計算する。この計算のモデルとして、Phong のモデルや Blinn のモデルなどが提案されている [145]。

C.4 本研究で提案する手法

C.4.1 雲のシェーディングモデル

電磁波の波長と、それを散乱させる粒子の半径が同じ程度の大きさである時の散乱を、ミー散乱 (Mie scattering) という。このときの散乱光は、入射した太陽光と同じように白色光に近くなる。さて雲粒の大きさはおよそ $1 \sim 100 \mu\text{m}$ であり太陽光線の波長³と同程度となるので、太

³可視光とは、およそ $0.39 \mu\text{m}$ から $0.74 \mu\text{m}$ の間の電磁波を指す。

陽光線は雲粒によってミーン散乱されることになる。そのために雲は白く見える。また雲粒の密度は、大部分の海洋上の雲では 100 個 cm^{-3} より少ないが、これに反して大陸上の雲では、 1 cm^3 に数百個の雲粒、場合によっては 900 個 cm^{-3} を越えることもある [146]。このような微小な雲粒の集合体によって雲が形成されている。その密度が高くなればなるほどそれに比例して大気の不透明度が高まり、散乱光が増加して雲の白い色が濃くなると考えられる。すると雲の色は、雲粒の密度に依存していると仮定してもよいだろう。

本研究では式 (C.1)・式 (C.2) において、雲の不透明度は 0.4 の一定値となるように設定した。つまり雲に分類されたすべてのボクセルには、不透明度として 0.4 の値を与えることにする。また雲がないボクセルの不透明度は当然 0.0 である。このように各ボクセルの不透明度は単純な形式で与えた。

次に各サンプル点での色の決定が問題となる。各サンプル点での色は本研究でも式 (C.4) に基づいて決定されている。しかし雲の場合は表面というものを考えられないため、鏡面反射光はゼロと考えてもよいだろう。そこで周囲光および拡散光のみを考慮する。

次に雲の内部のあるサンプル点での拡散光を決定する方法を考察する。拡散光とはそのサンプル点に到達した太陽光線を拡散した光であるので、拡散光の強さはそのサンプル点に到達した太陽光の強度に比例するだろう⁴。そしてまた拡散光の強さは雲粒の密度にも比例することから、そのサンプル点での不透明度にも比例するだろう。この関係は式 (C.3) に示された関係と同じである。さてそのサンプル点に到達する太陽光の強度を知るのに最も単純な方法は、そのサンプル点から光源方向に新たに光線を発生させ、その光線を追跡しながら不透明度を線積分する方法だろう。この方法を採用すると、光源方向に向かう新たな光線に沿った不透明度 α_L を計算する必要があるが、この部分には第 C.3.2 節の方法をそのまま適用できる。

最終的には、次式によってサンプリング点 U の色 $C(U)$ を決定する。

$$C(U) = (1 - \alpha_L)C(L) + C(E) \quad (C.5)$$

ここで $C(L)$ は光源の色、そして $C(E)$ は周囲光の色である。この手法では、通常の光線の他に光源方向に向かっても大量の光線を発するため、計算量が膨大となる欠点がある。もっと効率の良い手法が望まれるところである。

C.4.2 海のシェーディングモデル

雲とは異なり海は不透明物体であると仮定する。すなわち海に対応するボクセルの不透明度は 1 である。このとき海面に達した光線の不透明度は 1 となるので、その時点でただちにサンプリングは終わる。

次に海ボクセルの散乱光の色には近似モデルを用いる。すなわち純粋な海水に対する散乱光スペクトルの測定値から $(R : G : B) = (1 : 4 : 9)$ と近似する。さらに静かな海面を仮定して、Phong のモデルに基づいた鏡面反射成分も海ボクセルの色に加えるが、相互反射は含めない。海ボクセルの不透明度は 1 とする。また海の場合でも式 (C.5) の計算をおこなうため、光源方向に雲が存在する場合には $\alpha_L = 1$ となってそのピクセルは影となる。影の存在によって、画像の立体感はさらに強力となろう。ただし海面への雲の写り込みについては省略している。

⁴こう考えると、雲の表面に対応するサンプル点での色は、その点に達する太陽光量が多いので明るい色となる。反対に雲の内部のサンプル点での色は、その点に到達する太陽光が弱ければ弱いほど暗い色となる。

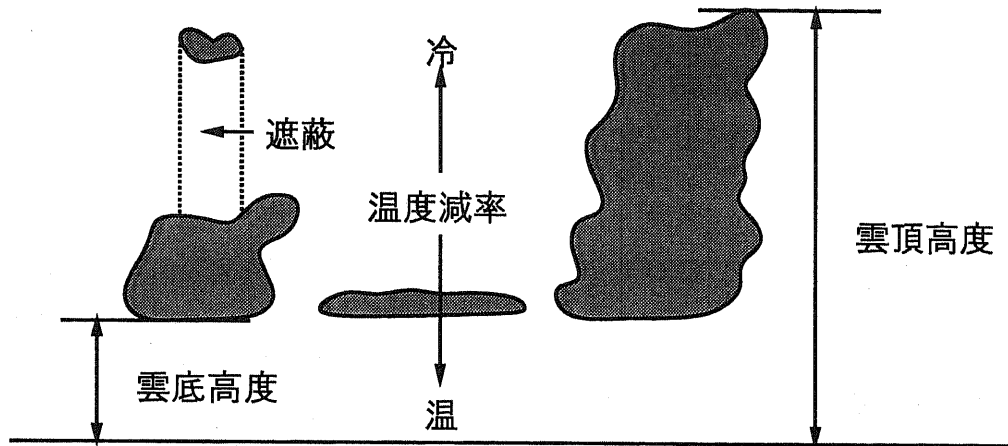


図 C.3: 雲頂高度と雲底高度の模式図。遮蔽にも注意。

C.5 実験

C.5.1 前処理およびデータ分類

対象には、気象衛星 NOAA の AVHRR センサで観測した熱赤外面像 (Ch.5)、および気象衛星ひまわりの赤外面像 (Ch.1) を用いた。実験には幾何補正と放射量補正が済んだ画像を用いる。次に画像の中から海域を切り出し、本研究で述べたミクセル密度を含む混合密度推定の手法により画素単位で分類する。

まず気象衛星 NOAA の画像としては 1995 年 9 月 16 日に撮影された画像を用いる。この画像には 1995 年台風 12 号が撮影されている。この 1 枚をさまざまな角度から眺めた画像を作成する。次に気象衛星ひまわりの画像としては 1996 年 7 月 3 日～7 月 9 日の間に観測された 133 枚の画像を用い⁵、これを 1 枚ずつ可視化して連続再生することで雲のダイナミックな動きを捉えることができる。こちらは固定視点で画像を生成する。

C.5.2 3次元ボリュームデータへの変換

まず雲に分類された画素について、画素値から雲頂高度を推定する。本研究では大気の大気垂直温度分布を最も単純な「線形モデル」で近似する。すると雲頂高度 H_t は、地上温度 T_g ・雲頂温度 T_t ・温度減率 C を用いて、 $H_t = (T_g - T_t)/C$ と計算できる。ここで温度減率には標準的な値 $-6.5 \text{ }^\circ\text{C}/\text{km}$ を採用し、 T_g には海面温度を用いる。また雲底高度 H_b には、画像中の雲画素のうち最も温度が高い画素に対応する高度を採用した。

次に 2 次元画像データを 3 次元ボリュームデータへと変換する。まず画素 $I(x, y)$ が雲画素と判定された場合、ボクセル $V(x, y, z) (H_b < z < H_t)$ を雲ボクセルとする。次に任意の画素 $I(x, y)$ に対して、ボクセル $V(x, y, z) (z = 0)$ を海ボクセルとする。最後にその他のボクセルを大気ボクセルとする。

C.5.3 可視化の結果および考察

⁵実はこの画像は短期型 GMS 衛星画像データベースとして用いた画像集合と同一のものである。

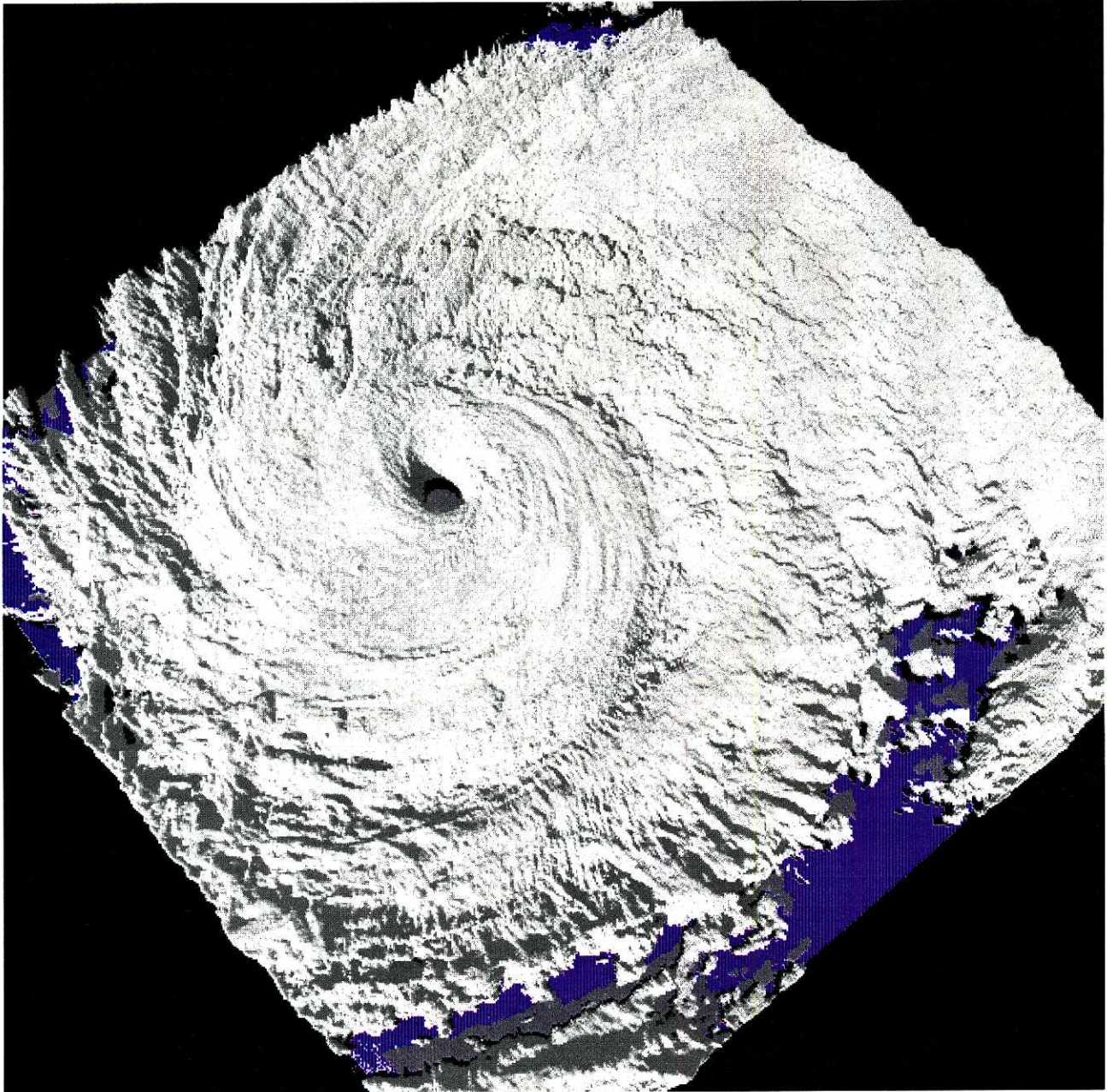


図 C.4: NOAA 衛星画像をボリュームレンダリングにより 3 次元可視化した結果。1995 年 9 月 16 日の台風 12 号。

NOAA 衛星画像を可視化した結果を図 C.4 に示す。このような 3 次元可視化画像を用いることによって、台風の眼などを立体的な構造として容易に捉えることができる。さらに視点や照明を任意に動かすことによりアニメーションも作製できる。

さてこの画像で最も不自然な点は雲の厚さが厚すぎる場所だろう。地球上空からしか衛星は観測できないため、原理的には雲の厚さを衛星データだけから知ることは難しい。そこで、知識を援用した手法や他の情報源からの情報を統合する手法を用いて雲の厚さを推定することが、今後の重要な課題となる。

次にここでは結果を示すことができないが、気象衛星ひまわりの画像 133 枚を用いて動画を作成し、雲のダイナミックな動きを示した。この動画は平面的な衛星画像を連続再生するのとは全く異なる印象を見る者に与える画像であり、したがってこのような視点から画像を可視化することは、画像から新たな視点を獲得するために有効な方法であると考えられる。

このように雲を立体的に表示しいろいろな角度から眺めることによって、今まで気づかなかった新たな発見をすることも可能となる [140]。さらにこのような手法とバーチャルリアリティを組み合わせると、観測者みずからが雲の中を「歩き回る」ことも可能となろう [147]。このように本研究を発展させれば、今まで見えなかった画像情報が見えるようになることだろう。

研究発表

- [i] 北本朝展, 高木幹雄. L型・J型およびU型の面積占有率密度を用いたミクセルの解析. 電子情報通信学会論文誌, 1997. 投稿中.
- [ii] 北本朝展, 高木幹雄. ミクセルの面積占有率推定およびミクセル密度を含む混合密度推定. 電子情報通信学会論文誌, 1997. 投稿中.
- [iii] 北本朝展, 高木幹雄. ミクセルの面積占有率推定およびミクセル密度を含む混合密度推定の一般化. 電子情報通信学会論文誌, 1997. 投稿予定.
- [iv] 北本朝展, 高木幹雄. 遺伝的アルゴリズムを用いた類似画像検索パラメータの最適化. 情報処理学会論文誌, 1997. 投稿予定.
- [v] 北本朝展, 高木幹雄. 領域・空間情報を統合したグラフ構造を用いた類似画像検索. 情報処理学会論文誌, 1997. 投稿予定.
- [vi] Kitamoto, A., Zhou, C., and Takagi, M. **Similarity Retrieval of NOAA Satellite Imagery by Graph Matching.** In *Proc. of SPIE 1908 Image Storage and Retrieval Systems*, pp. 60-73, San Jose, 1993.
- [vii] Kitamoto, A. and Takagi, M. **Retrieval of Satellite Cloud Imagery Based on Subjective Similarity.** In *The 9th Scandinavian Conference on Image Analysis*, pp. 449-456, Uppsala, 1995.
- [viii] Kitamoto, A. and Takagi, M. **Mixture Density Estimation Under the Existence of Mixels.** In *1995 International Geoscience and Remote Sensing Symposium*, Firenze, 1995.
- [ix] Kitamoto, A. and Takagi, M. **A Stochastic Model of Mixels and Image Classification.** In *Proc. of 13th International Conference on Pattern Recognition*, pp. 745-749, Wien, 1996.
- [x] 北本朝展, 高木幹雄. 最適化原理に基づいた衛星雲画像の類似検索. 電子情報通信学会技術報告, Vol. PRU94-49, pp. 15-22, 1994.
- [xi] 北本朝展, 高木幹雄. ミクセルが存在する場合の混合密度推定. 電子情報通信学会技術報告, Vol. PRU95-202, pp. 33-40, 1996.
- [xii] 北本朝展, 高木幹雄. パイプライン型遺伝的アルゴリズムによる模擬育種法を用いた類似画像検索規準の学習. 電子情報通信学会技術報告, Vol. HIP96-4, pp. 17-22, 1996.
- [xiii] 北本朝展, 高木幹雄. 類似画像検索への応用を目的とした階層化属性付きグラフマッチングの高速化. 画像の認識・理解シンポジウム (MIRU'96), Vol. II, pp. 331-336, 1996.
- [xiv] 北本朝展, 周長明, 高木幹雄. 属性付きグラフ構造を用いた NOAA 衛星画像の類似検索. In 1993 年電子情報通信学会春季大会, pp. D-522, 1993.
- [xv] 北本朝展, 高木幹雄. 属性付き関係グラフを用いたヘルマート不変の類似検索. In 1993 年電子情報通信学会秋季大会, pp. D-296, 1993.

- [xvi] 北本朝展, 高木幹雄. 遺伝的アルゴリズムを用いた類似検索規準の学習. In 1995年電子情報通信学会総合大会, pp. D-296, 1995.
- [xvii] 北本朝展, 高木幹雄. ミクセルが存在する場合の混合分布推定. 第51回情報処理学会全国大会, Vol. 1, pp. 17-18, 1995.
- [xviii] 北本朝展, 高木幹雄. FFT畳み込みを用いたミクセル分布の計算. In 1996年電子情報通信学会総合大会, pp. D-488, 1996.
- [xix] 北本朝展, 高木幹雄. ボリュームレンダリングを用いた気象衛星画像の3次元可視化. In 1996年テレビジョン学会年次大会, pp. 114-115, 1996.
- [xx] 北本朝展, 高木幹雄. ミクセルの面積占有率の算出. In 1996年電子情報通信学会秋期大会, pp. D-382, 1996.

参考文献

- [1] Narasimhalu, A.D. **Special Section on Content-Based Retrieval.** *Multimedia Systems*, Vol. 3, pp. 1-2, 1995.
- [2] Gudivada, V.N. and Raghavan, V.V. **Content-Based Image Retrieval Systems.** *Computer*, Vol. 28, No. 9, pp. 18-22, 1995.
- [3] 小野敦史, 天野督士, 斗谷充宏, 佐藤隆, 坂内正夫. 状態遷移モデルとシーン記述言語による自動キーワード付与機能をもつ画像データベースとその評価. 電子情報通信学会論文誌, Vol. J79-D-II, No. 4, pp. 476-483, 1996.
- [4] 芥子育雄, 池内洋, 黒武者健一. 百科事典の知識に基づく画像の連想検索. 電子情報通信学会論文誌, Vol. J79-D-II, No. 4, pp. 484-491, 1996.
- [5] 柴田正啓, 井上誠喜. 画像データベースの連想検索方式. 電子情報通信学会論文誌, Vol. J73-D-II, No. 4, pp. 526-534, 1990.
- [6] 加藤俊一. 類似検索機能を持った画像データベースとその設計法. *O plus E*, Vol. 142,143, pp. 141-165, 1991.
- [7] 美濃導彦. 知的映像メディア検索技術の動向. 人工知能学会誌, Vol. 11, No. 1, pp. 3-9, 1996.
- [8] 田辺勝義, 大谷淳, 石井健一郎. 多次元心理空間を用いる類似画像検索法. 電子情報通信学会論文誌, Vol. J75-D-II, No. 11, pp. 1856-1865, 1992.
- [9] 長嶋秀世, 土方洋一. 人間の主観を重視した類似商標図形の検索の基礎的検討. 電子情報通信学会論文誌, Vol. J74-D-II, No. 3, pp. 311-320, 1991.
- [10] Bunke, H. and Allermann, G. **Inexact Graph Matching for Structural Pattern Recognition.** *Pattern Recognition Letters*, Vol. 1, pp. 245-253, 1983.
- [11] Tsai, W.H. and Fu, K.S. **Error-Correcting Isomorphisms of Attributed Relational Graphs for Pattern Analysis.** *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-9, No. 12, pp. 757-768, 1979.
- [12] 田中栄一. 構造をもつものの距離と類似度. 情報処理, Vol. 31, No. 9, pp. 1270-1279, 1990.
- [13] Kato, T. **Database Architecture for Content-Based Image Retrieval.** In Jamberdino, A.A. and Niblack, W., editors, *Proc. of SPIE 1662 Image Storage and Retrieval Systems*, pp. 112-123, 1992.
- [14] 加藤俊一, 栗田多喜夫. 画像の内容検索 — 電子美術館への応用 —. 情報処理, Vol. 33, No. 5, pp. 466-477, 1992.
- [15] Niblack, W., Barber, R., Equitz, W., Flickner, M., Glasman, E., Petkovic, D., Yanker, P., Faloutsos, C., and Taubin, G. **The QBIC Project: Querying Images By Content Using Color, Texture, and Shape.** In *Proc. of SPIE 1908 Image Storage and Retrieval Systems*, pp. 173-187, San Jose, 1993.

- [16] Faloutsos, C., Barber, R., Flickner, M., Hafner, J., Niblack, W., Petkovic, D., and Equitz, W. **Efficient and Effective Querying by Image Content**. *Journal of Intelligent Information Systems*, Vol. 3, pp. 231–262, 1994.
- [17] Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., and Yanker, P. **Query by Image and Video Content: The QBIC System**. *Computer*, Vol. 28, No. 9, pp. 23–32, 1995.
- [18] Hafner, J., Sawhney, H.S., Equitz, W., Flickner, M., and Niblack, W. **Efficient Color Histogram Indexing for Quadratic Form Distance Functions**. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 7, pp. 729–736, 1995.
- [19] **QBIC Home Page**. <http://www.qbic.almaden.ibm.com/>.
- [20] **Virage Incorporated: Visual Information Retrieval**. <http://www.virage.com/>.
- [21] Bach, J.R., Paul, S., and Jain, R. **A Visual Information Management System for the Interactive Retrieval of Faces**. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 5, No. 4, pp. 619–628, 1993.
- [22] 長谷川純一, 岡田憲武, 鳥脇純一郎. スケッチを利用した胸部 X 線像データベースの知的検索. 電子情報通信学会論文誌, Vol. J71-D, No. 3, pp. 549–559, 1988.
- [23] 黒川雅人, 洪政国. 形状情報を用いた画像の類似検索システム. 情報処理学会論文誌, Vol. 32, No. 6, pp. 721–730, 1991.
- [24] 木下宏揚, 岡野栄司, 宮本敦司, 酒井善則. 構造化情報のグラフマッチングに基づく画像検索法. テレビジョン学会誌, Vol. 49, No. 7, pp. 913–922, 1995.
- [25] 田辺勝義, 大谷淳. 形状類似画像検索における類似尺度の検討. 電子情報通信学会技術報告, Vol. PRU88-68, pp. 65–72, 1988.
- [26] Binaghi, E., Gagliardi, I., and Schettini, R. **Image Retrieval Using Fuzzy Evaluation of Color Similarity**. *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 8, No. 4, pp. 945–968, 1994.
- [27] 高橋友一, 島則之, 岸野文郎. 位置情報を手がかりとする画像検索法. 情報処理学会論文誌, Vol. 31, No. 11, pp. 1636–1643, 1990.
- [28] 美濃導彦, 岡崎洋, 坂井利之. 対象物の属性特徴による画像検索法. 情報処理学会論文誌, Vol. 32, No. 4, pp. 513–522, 1991.
- [29] Chang, S.K., Shi, Q.Y., and Yan, C.W. **Iconic Indexing by 2-D Strings**. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-9, No. 3, pp. 413–428, 1987.
- [30] Yamamoto, A. and Takagi, M. **Extraction of Object Features and Its Application to Image Retrieval**. *The Transactions of the Institute of Electronics, Information, and Communication Engineers*, Vol. E-72, No. 6, pp. 771–781, 1989.
- [31] 金原史和. プリミティブ分解によるユーザの視点に基づいた画像検索. PhD thesis, 東京大学, 1996.
- [32] 金原史和, 佐藤真一, 浜田喬. プリミティブ分解による多様な検索条件を扱えるカラー画像検索. 情報処理学会論文誌, Vol. 37, No. 11, pp. 1989–2000, 1996.
- [33] 坂内正夫. 画像検索技術. 電子情報通信学会学会誌, Vol. 71, No. 9, pp. 911–914, 1988.

- [34] 日本気象協会(編). 「ひまわり」で見る四季の気象. 大蔵省印刷局, 1993.
- [35] **Satellite Image Archive for Network.** <http://www.tkl.iis.u-tokyo.ac.jp/SatIAN/>.
- [36] Samet, H. **A Review of Hierarchical Representation of Shape and Some Applications.** In *Visual Form*, pp. 451-467. Plenum Press, 1992.
- [37] Srihari, R.K. **Automatic Indexing and Content-Based Retrieval of Captioned Images.** *Computer*, Vol. 28, No. 9, pp. 49-56, 1995.
- [38] 柳沼良知, 坂内正夫. DP マッチングを用いたドラマ映像・音声・シナリオ文書の対応付け手法の一提案. 電子情報通信学会論文誌, Vol. J79-D-II, No. 5, pp. 747-755, 1996.
- [39] デビッド・マー. ビジョン. 産業図書, 1987.
- [40] 君山博之, 清末梯之, 大庭有二. 動画像の自動記述方法の検討. 電子情報通信学会技術報告, 1991.
- [41] 小川英光. パターン認識・理解の諸問題. 電子情報通信学会学会誌, Vol. 76, No. 6, pp. 615-631, 1994.
- [42] Luo, G. **Simulations Exploring the Dependence of Cloud-Cover Frequency Distribution on Cloud Size and Image Pixel Resolution.** *Journal of Atmospheric and Oceanic Technology*, Vol. 12, pp. 712-720, 1995.
- [43] フェラー, W. 確率論とその応用 II. 紀伊国屋書店, 1969.
- [44] Crapper, P.F. **An Estimate of the Number of Boundary Cells in a Mapped Landscape Coded to Grid Cells.** *Photogrammetric Engineering and Remote Sensing*, Vol. 50, No. 10, pp. 1497-1503, 1984.
- [45] Zhu, T., Lee, J., Weger, R.C., and Welch, R.M. **Clustering, Randomness, and Regularity in Cloud Fields: 2. Cumulus Cloud Fields.** *Journal of Geophysical Research*, Vol. 97, No. D18, pp. 20537-20558, 1992.
- [46] Lovejoy, S. **Area-perimeter relation for rain and cloud areas.** *Science*, Vol. 216, pp. 185-187, 1982.
- [47] Voss, R.F. **Random Fractal Forgeries.** In Earnshaw, R.A., editor, *Fundamental Algorithms for Computer Graphics*, pp. 805-835. Springer-Verlag, 1985.
- [48] 大津展之, 栗田多喜夫, 関田巖. パターン認識, Vol. 12 of 行動計量学シリーズ. 朝倉書店, 1996.
- [49] Settle, J.J. and Drake, N.A. **Linear Mixing and the Estimation of Ground Cover Proportions.** *International Journal of Remote Sensing*, Vol. 14, No. 6, pp. 1159-1177, 1993.
- [50] 稲村実. カテゴリー分解に基づくリモートセンシング画像データの解析. 電子情報通信学会論文誌, Vol. J70-C, No. 2, pp. 241-250, 1987.
- [51] 伊藤直史, 藤村貞夫. 画素のカテゴリー分解による面積比推定. 計測自動制御学会論文集, Vol. 23, No. 8, pp. 800-805, 1987.
- [52] Goetz, A.F.H. **Imaging Spectrometry: An Image Analysis and Interpretation Challenge.** In *The 9th Scandinavian Conference on Image Analysis*, pp. 619-631, Uppsala, Sweden, 1995.
- [53] Foody, G.M. **Relating the Land-Cover Composition of Mixed Pixels to Artificial Neural Network Classification Output.** *Photogrammetric Engineering and Remote Sensing*, Vol. 62, No. 5, pp. 491-499, 1996.

- [54] Wang, F. **Fuzzy Supervised Classification of Remote Sensing Images.** *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 28, No. 2, pp. 194–201, 1990.
- [55] Kent, J.T. and Mardia, K.V. **Spatial Classification Using Fuzzy Membership Models.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 5, pp. 659–671, 1988.
- [56] Caillol, H., Hillion, A., and Pieczynski, W. **Fuzzy Random Fields and Unsupervised Image Segmentation.** *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 31, No. 4, pp. 801–810, 1993.
- [57] Foody, G.M. and Cox, D.P. **Sub-Pixel Land Cover Composition Estimation Using a Linear Mixture Model and Fuzzy Membership Functions.** *International Journal of Remote Sensing*, Vol. 15, No. 3, pp. 619–631, 1994.
- [58] Foody, G.M. **Approaches for the Production and Evaluation of Fuzzy Land Cover Classifications from Remotely-Sensed Data.** *International Journal of Remote Sensing*, Vol. 17, No. 7, pp. 1317–1340, 1996.
- [59] Tabatabai, A.J. and Mitchell, R. **Edge Location to Subpixel Values in Digital Imagery.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-6, No. 2, pp. 188–201, 1984.
- [60] 吉野邦彦. ミクセルにおける瞬時視野内カテゴリ面積率の推定. 写真測量とリモートセンシング, Vol. 33, No. 2, pp. 28–33, 1994.
- [61] Santago, P. and Gage, H.D. **Quantification of MR Brain Images by Mixture Density and Partial Volume Modeling.** *IEEE Transactions on Medical Imaging*, Vol. 12, No. 3, pp. 566–574, 1993.
- [62] Choi, H.S., Haynor, D.R., and Kim, Y. **Partial Volume Tissue Classification of Multichannel Magnetic Resonance Images — A Mixel Model.** *IEEE Transactions on Medical Imaging*, Vol. 10, No. 3, pp. 395–408, 1991.
- [63] Lei, T. and Sewchand, W. **Statistical Approach to X-Ray CT Imaging and Its Application in Image Analysis — Part II: A New Stochastic Model-Based Image Segmentation Technique for X-Ray CT Image.** *IEEE Transactions on Medical Imaging*, Vol. 11, No. 1, pp. 62–69, 1992.
- [64] Solberg, A.H. Schistad, Bosnes, V., and Storvik, G. **Tissue Classification in MR Images Based on a Mixed Pixel Model.** In *The 9th Scandinavian Conference on Image Analysis*, pp. 113–119, Uppsala, 1995.
- [65] Bosdogianni, P., Petrou, M., and Kittler, J. **Mixed Pixel Classification in Remote Sensing.** In *SPIE Proceedings of the EUROPTO series, Image and Signal Processing for Remote Sensing*, Vol. 2315, pp. 494–505, 1994.
- [66] 新井康平. カテゴリ分解による画素内被雲率推定. 写真測量とリモートセンシング, Vol. 31, No. 5, pp. 4–10, 1992.
- [67] Hisdal, E. **Are Grades of Membership Probabilities?** *Fuzzy Sets and Systems*, Vol. 25, pp. 325–348, 1988.
- [68] Bezdek, J. **Fuzziness vs. Probability – Again (!?).** *IEEE Transactions on Fuzzy Systems*, Vol. 2, No. 1, pp. 1–3, 1994.
- [69] Brigham, E.O. **The Fast Fourier Transform.** Prentice-Hall, 1974.

- [70] 松本雅雄, 藤九浩基, 土屋清, 新井康平. 最尤推定法に基づくカテゴリー分解. 写真測量とリモートセンシング, Vol. 30, No. 2, pp. 25-34, 1991.
- [71] 北本朝展, 高木幹雄. ミクセルが存在する場合の混合密度推定. 電子情報通信学会技術報告, Vol. PRU95-202, pp. 33-40, 1996.
- [72] Raffy, M. **Introduction - Change of Scale Theory: a Capital Challenge for Space Observation of Earth.** *International Journal of Remote Sensing*, Vol. 15, No. 12, pp. 2353-2357, 1994.
- [73] 中村永友, 小西貞則, 大隅昇. 混合分布モデルを用いた分類法とデータ構造の色彩表示. 統計数理, Vol. 41, No. 2, pp. 149-167, 1993.
- [74] Quelle, H.C., Delignon, Y., and Marzouki, A. **Unsupervised Bayesian Segmentation of SAR Images Using the Pearson System Distributions.** In *IGARSS'93*, pp. 1538-1540, Tokyo, 1993.
- [75] 小林幸夫, 稲村實. EM アルゴリズムによるリモートセンシング画像のカテゴリー分類におけるミクセルデータの影響と改善. 情報処理学会論文誌, Vol. 37, No. 1, pp. 39-48, 1996.
- [76] Redner, R.A. and Walker, H.F. **Mixture Densities, Maximum Likelihood and the EM algorithm.** *SIAM Review*, Vol. 26, No. 2, pp. 195-239, 1984.
- [77] 宮川雅巳. EM アルゴリズムとその周辺. 応用統計学, Vol. 16, No. 1, pp. 1-19, 1987.
- [78] Fukunaga, K. and Flick, T.E. **Estimation of the Parameters of a Gaussian Mixture Using the Method of Moments.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-5, No. 4, pp. 410-416, 1983.
- [79] Carlotto, M.J. **Histogram Analysis Using a Scale-Space Approach.** In *CVPR'85*, pp. 334-340, 1985.
- [80] Kittler, J. and Illingworth, J. **Minimum Error Thresholding.** *Pattern Recognition*, Vol. 19, No. 1, pp. 41-47, 1986.
- [81] 高木幹雄, 下田陽久 (編). 画像処理ハンドブック. 東京大学出版会, 1991.
- [82] 本谷秀堅, 出口光一郎. 正則化を用いた **scale-space** 上の一般化図形の塊化手法. 電子情報通信学会技術報告, Vol. PRU92-67, pp. 71-78, 1992.
- [83] Pitas, I. and Venetsanopoulos, A.N. **Morphological Shape Decomposition.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 1, pp. 38-45, 1990.
- [84] Pitas, I. and Venetsanopoulos, A.N. **Morphological Shape Representation.** *Pattern Recognition*, Vol. 25, No. 6, pp. 555-565, 1992.
- [85] 木本伊彦, 浅井基博, 安田靖彦. だ円体によるシルエット図形の記述と表現. 電子情報通信学会論文誌, Vol. J76-D-II, No. 6, pp. 1159-1167, 1993.
- [86] 原潤一, 加藤博一, 井口征士. だ円膨張法を用いた動物シルエット画像の記述と認識. 電子情報通信学会論文誌, Vol. J74-D-II, No. 3, pp. 366-375, 1991.
- [87] 堀越力, 鈴木智, 中根一成. **AIC (赤池の情報量規準)** を用いた最少部品による形状記述法. 電子情報通信学会論文誌, Vol. J77-D-II, No. 9, pp. 1691-1700, 1994.
- [88] 島谷明, 鎧沢勇, 小杉信. 図形分節における文節候補の自動抽出法. 電子情報通信学会論文誌, Vol. J73-D-II, No. 9, pp. 1555-1562, 1990.

- [89] 島谷明, 鎧沢勇, 小杉信. 人間の視覚特性に基づき図形分節を予測するモデル. 電子情報通信学会論文誌, Vol. J74-D-II, No. 6, pp. 778-785, 1991.
- [90] Shapiro, L.G. **Recent Progress in Shape Decomposition and Analysis**. In *Progress in Pattern Recognition 2*, pp. 113-123. Elsevier Science Publishers B.V., North-Holland, 1985.
- [91] Pavlidis, T. **Survey : A Review of Algorithms for Shape Analysis**. *Computer Graphics and Image Processing*, Vol. 7, pp. 243-258, 1978.
- [92] Pentland, A.P. **Perceptual Organization and the Representation of Natural Form**. *Artificial Intelligence*, Vol. 28, pp. 293-331, 1986.
- [93] Pentland, A.P. **Recognition by Parts**. In *Proc. of 1st International Conference on Computer Vision*, pp. 612-620, 1987.
- [94] Biederman, I. **Human Image Understanding: Recent Research and a Theory**. *Computer Vision, Graphics, and Image Processing*, Vol. 32, pp. 29-73, 1985.
- [95] 横矢直和, 坂上勝彦. 画像理解と最適化原理. 電子情報通信学会学会誌, Vol. 74, No. 4, pp. 326-334, 1991.
- [96] Kass, M., Witkin, A., and Terzopoulos, D. **Snakes : Active Contour Models**. In *Proc. of 1st International Conference on Computer Vision*, pp. 259-268, 1987.
- [97] Cohen, L.D. and Cohen, I. **Finite-Element Methods for Active Contour Models and Ballons for 2-D and 3-D Images**. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 11, pp. 1131-1147, 1993.
- [98] 積山洋子, 坂上勝彦, 山本和彦. **Active Net : 動的な網のモデル**. 情報処理学会研究報告, Vol. 63, No. 2, pp. 1-8, 1989.
- [99] Thiel, E. and Montanvert, A. **Shape Splitting from Medial Lines Using the 3-4 Chamfer Distance**. In *Visual Form*, pp. 537-546. Plenum Press, 1992.
- [100] 間瀬茂, 上田修功. モルフォロジーと画像解析 [II],[III]. 電子情報通信学会学会誌, Vol. 74, No. 2,3, pp. 166-174,271-279, 1991.
- [101] 和田俊和, 野村圭弘, 松山隆司. 分散強調処理による画像の領域分割法. 情報処理学会論文誌, Vol. 36, No. 4, pp. 879-891, 1995.
- [102] Zhang, Y., Shigeta, K., and Shimizu, E. **Data Clustering Using the Concept of Psychological Potential Field**. *The Transactions of the Institute of Electronics, Information, and Communication Engineers*, Vol. E77-D, No. 11, pp. 1198-1205, 1994.
- [103] Compton, B.J. and Logan, G.D. **Evaluating a Computational Model of Perceptual Grouping by Proximity**. *Perception & Psychophysics*, Vol. 53, No. 4, pp. 403-421, 1993.
- [104] 曾我真人, 新田祐介, 今中武, 豊田順一, 柳田益造. 群化に基づく図形パターンの階層表現の生成手法. 電子情報通信学会論文誌, Vol. J75-D-II, No. 2, pp. 275-287, 1992.
- [105] 平塚賢, 大西昇, 杉江昇. 群化を用いたグローバルな構造の抽出. 電子情報通信学会論文誌, Vol. J76-D-II, No. 1, pp. 74-83, 1993.
- [106] Reed, T. and Wechsler, H. **Spatial/Spatial-Frequency Representations for Image Segmentation and Grouping**. *Image and Vision Computing*, Vol. 9, No. 3, pp. 175-193, 1991.

- [107] Sarkar, S. and Boyer, K.L. **Integration, Inference, and Management of Spatial Information Using Bayesian Networks: Perceptual Organization.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 3, pp. 256-274, 1993.
- [108] Sarkar, S. and Boyer, K.L. **A Computational Structure for Preattentive Perceptual Organization: Graphical Enumeration and Voting Methods.** *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 24, No. 2, pp. 246-267, 1994.
- [109] Sarkar, S. and Boyer, K.L. **Perceptual Organization in Computer Vision: A Review and a Proposal for a Classificatory Structure.** *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 23, No. 2, pp. 382-399, 1993.
- [110] メッツガー, W. 視覚の法則. 白揚社, 1962.
- [111] Bunke, H. and Sanfeliu, A. **Introduction.** *Pattern Recognition*, Vol. 19, No. 4, pp. 249-254, 1986.
- [112] Shapiro, L.G. and Haralick, R.M. **Structural Descriptions and Inexact Matching.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-3, No. 5, pp. 504-519, 1981.
- [113] 北本朝展, 高木幹雄. 類似画像検索への応用を目的とした階層化属性付きグラフマッチングの高速化. 画像の認識・理解シンポジウム (MIRU'96), Vol. II, pp. 331-336, 1996.
- [114] Nilsson, N.J. **Principles of Artificial Intelligence.** Springer-Verlag, 1982.
- [115] Stewart, B.S., Liaw, C.F., and White, C.C. **A Bibliography of Heuristic Search Research Through 1992.** *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 24, No. 2, pp. 268-293, 1994.
- [116] 笠原久嗣, 岸本登美夫. 画像データベースナビゲーション. 電子情報通信学会論文誌, Vol. IE-46, pp. 71-78, 1988.
- [117] 北野宏明 (編). 遺伝的アルゴリズム. 産業図書, 1993.
- [118] Michalewicz, Z. **Genetic Algorithms + Data Structures = Evolution Programs.** Springer-Verlag, 1992.
- [119] 北本朝展, 高木幹雄. 遺伝的アルゴリズムを用いた類似検索規準の学習. In 1995年電子情報通信学会総合大会, pp. D-296, 1995.
- [120] Dawkins, R. **The Blind Watchmaker.** Longman, 1986.
- [121] Sims, K. **Artificial Evolution for Computer Graphics.** *Computer Graphics*, Vol. 25, No. 4, 1991.
- [122] Caldwell, C. and Johnston, V.S. **Tracking a Criminal Suspect through Face-Space with a Genetic Algorithm.** In *Proc. of Int. Conf. on Genetic Algorithm*, pp. 416-421, 1991.
- [123] Baker, E. **Evolving Line Drawings.** In *Graphics Interface*, 1994.
- [124] Syswerda, G. **A Study of Reproduction in Generational and Steady-State Genetic Algorithms.** In Rawlins, G.J.E., editor, *Foundations of Genetic Algorithms*, pp. 94-101. Morgan Kaufmann Publishers, San Mateo, 1991.
- [125] 北本朝展, 高木幹雄. パイプライン型遺伝的アルゴリズムによる模擬育種法を用いた類似画像検索規準の学習. 電子情報通信学会技術報告, Vol. HIP96-4, pp. 17-22, 1996.

- [126] 西山晴彦, 松下温. 画像の構図を用いた絵画検索システム. 情報処理学会論文誌, Vol. 37, No. 1, pp. 101-109, 1996.
- [127] 平田恭二, 原良憲. ハイパーメディアにおけるメディアナビゲーション. 情報処理学会研究報告, Vol. HI-47, pp. 69-76, 1993.
- [128] 北本朝展. 画像の類似度. 大学院論文輪講, 東京大学大学院工学系電子工学専攻, 1996.
- [129] 水野りか, 菅沼義昇. 属性値の差異に基づくカテゴリー形成モデルの実験的検討. 人工知能学会誌, Vol. 8, No. 3, pp. 357-366, 1993.
- [130] 北本朝展. スケール. 大学院論文輪講, 東京大学大学院工学系電子工学専攻, 1994.
- [131] 佐藤誠. 認識と理解のためのパターン解析. 電子情報通信学会学会誌, Vol. 71, No. 11, pp. 1225-1230, 1988.
- [132] Koenderink, J.J. **The Structure of Images.** *Biological Cybernetics*, Vol. 50, pp. 363-370, 1984.
- [133] Marr, D. and Hildreth, E. **Theory of Edge Detection.** *Proc. Royal Society of London*, Vol. 207, pp. 187-217, 1980.
- [134] Witkin, A.P. **Scale-Space Filtering.** In *Proc. of 8th International Conference on Artificial Intelligence*, pp. 1019-1022. 1983.
- [135] Perona, P. and Malik, J. **Scale-Space and Edge Detection Using Anisotropic Diffusion.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 7, pp. 629-639, 1990.
- [136] Lindeberg, T. **Scale-Space for Discrete Signals.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 3, pp. 234-254, 1990.
- [137] 藤代一成, 茅暁陽, 國井利泰. ボクセル指向3次元データ表現とその表示技術. 情報処理, Vol. 34, No. 3, pp. 285-298, 1993.
- [138] 藤代一成. ボリューム・ビジュアライゼーションの本質・概論. *PIXEL*, Vol. 116, pp. 144-157, 1992.
- [139] 建石隆太郎. 鳥瞰図作成のアルゴリズム. 写真測量とリモートセンシング, Vol. 22, No. 3, pp. 14-25, 1983.
- [140] Hibbard, W. and Santek, D. **Visualizing Large Data Sets in the Earth Sciences.** *IEEE Computer*, Vol. 22, No. 8, pp. 53-57, 1989.
- [141] 藤代一成, 茅暁陽. ボリューム・ビジュアライゼーションの基本アルゴリズム. *PIXEL*, Vol. 119-121,, 1992.
- [142] Levoy, M. **Efficient Ray Tracing of Volume Data.** *ACM Transactions on Graphics*, Vol. 9, No. 3, pp. 245-261, 1990.
- [143] Drebin, R.A., Carpenter, L., and Hanrahan, P. **Volume Rendering.** *Computer Graphics*, Vol. 22, No. 4, pp. 65-74, 1988.
- [144] 高桑昌男. レンダリングの基礎から最新情報まで. *PIXEL*, Vol. 114, pp. 35-42, 1992.
- [145] 横井茂樹, 安田孝美, 岡田稔, 石井智海. CGによるレンダリングのための基本モデル. *PIXEL*, Vol. 113-119,, 1992.
- [146] 小倉義光. 一般気象学. 東京大学出版会, 1984.
- [147] Bryson, S. and Levit, C. **The Virtual Wind Tunnel.** *IEEE Computer Graphics and Applications*, Vol. 12, No. 4, pp. 25-34, 1992.

謝辞

まず指導教官の高木幹雄教授には研究を進めるうえで多大な御指導を賜りました。また、衛星画像を直接受信できるという環境や充実した計算機環境など、本研究はこのような素晴らしい研究環境に支えられて完成したものです。あらためて高木幹雄教授には心から感謝の意を表します。また喜連川優助教授にも御指導および御助言をいただきました。心より感謝いたします。曾根光男氏(現富士通)には、研究を進める上での貴重な御助言や励ましを頂き、深く感謝いたしております。

さらに研究室の方々には、いろいろと本当にお世話になりました。坂元宗和助手には研究室打ち合せという討論の場で数々の鋭い御意見を頂き、研究を進めるうえで貴重な刺激を受けました。また研究室の先輩である根本利弘氏(現東京大学生産技術研究所)や高村誠之氏(現NTT)には、研究はもちろんのこと、研究以外の面でもいろいろと教えていただきました。また会津昌夫氏(現大日本印刷)にもカラー画像の取り扱いなどについていろいろと御指導いただきました。ここにすべての名前を挙げることはできませんが、研究室の多くの方々の支えによって本研究は完成したものです。ここに深く感謝いたします。

また高木研究室が誇る有能な秘書の方々には、学生が研究を進めやすいようにいろいろと配慮して頂きました。また研究とは直接関係はありませんが、六本木での昼食が充実していたこともよい気分転換となりました。心より感謝申し上げます。

最後に、今までの著者の人生を支えてくれた両親および妹に感謝します。