

Studies on discrimination of images by unifying contour information and texture information

矢野, 啓司

<https://doi.org/10.11501/3148620>

出版情報 : 九州芸術工科大学, 1998, 博士 (工学) , 課程博士
バージョン :
権利関係 :



氏名・本籍(国籍) 矢野啓司(大分県)

学位の種類 博士(工学)

学位記番号 甲第 32 号

学位授与の日付 平成 11 年 3 月 18 日

学位授与の要件 学位規則第 4 条第 1 項該当

学位論文題目 輪郭線情報とテクスチャ情報の統合による画像識別に関する研究

審査委員会 幹事 教授 瀧山龍三

委員 教授 長島健次

委員 教授 浦濱喜一

論文内容の要旨

近年、電子技術のめざましい進展に伴い、計算機を用いた画像の識別に関する研究とその産業への応用が盛んに行われている。

画像識別を行なうには画像から特徴を抽出する必要があるが、識別結果は抽出された特徴の性質に大きく左右されるので、特徴を適切に取り出すことが重要である。画像の特徴を表す有効な情報として、その形状に関する情報及び模様に関する情報とがある。形状は多くの場合輪郭線で表され、通常閉曲線となる。画像から輪郭線を抽出・記述して、その性質を把握する方法はこれまでも多くの研究がなされており、実用的な種々の提案もなされている。模様は多くの場合テクスチャとして捉えられる。テクスチャの解析は、画像処理の基本的なものの一つとして、これまで多くの研究がなされており、最近リモートセンシング画像や医用画像の識別、更には 3 次元物体識別等において積極的に利用されている。

輪郭線やテクスチャは、それぞれ画像の識別において重要な情報である。しかしながら、更にこれら両者を統合して識別を行なえば、より良い結果を得ることができることもいくつかの例で明らかになっている。輪郭線情報とテクスチャ情報の統合にとどまらず、異質で有効な複数の情報を統合して識別を行うという研究は、今後より広く進められていく必要がある。

本論文は、輪郭線情報とテクスチャ情報との統合による画像識別に関する研究結果をとりまとめたものであり、7 章から構成されている。

第 1 章は序論であり、本研究の背景と扱っている問題を示し、あわせて論文の概要について述べる。

第 2 章では、輪郭線情報の種々の記述方法に関して概説し、これらの方法の性質を詳細に分析する。この結果、P 型記述子から得られるパワー・スペクトルを用いる方法が、画像の平行移動、拡大・縮小及び回転に対して不変であり、更に輪郭線の始点の位置の移動に対しても不変であることから、最も有効な方法であることを明らかにする。また、P 型

記述子を精度良く求めるためには輪郭線を等辺多角形で近似する必要があるので、その方法についても詳述し、特に EPAC に基づく方法は、輪郭線が再生可能であり、かつデジタル化誤差の影響を受けにくいと、きわめて有用な等辺多角形近似法であることを明らかにする。

第 3 章では、輪郭線情報を安定して抽出する方法としてロバストエネルギー最小化による輪郭線保存平滑化法を述べる。輪郭線情報による識別において、大域的な形状の情報の方が局所的な情報よりも重要な場合が多い。また、輪郭線の等辺多角形近似において、平滑化された輪郭線の方が等辺多角形近似アルゴリズムが収束しやすい。そこで、大域的な情報を損なわず、ノイズ等の不要な局所情報を除き輪郭線を平滑化する方法として、ロバストエネルギー最小化による方法を提案し、この方法が、特に平滑化能力とアルゴリズムの扱い易さにおいて、きわめて有効であることを明らかにする。更に、具体的なエッジ平滑化の例としてノイズの混入した波形データの平滑化実験をも行い、種々の方法と比較した結果、本法が最も有用であることを実証する。

第 4 章では、テクスチャ情報の種々の記述方法に関して、構造的な解析法と統計的な解析法とに大別して概説し、これらの方法の性能を詳細に分析する。この結果、濃度共起行列による方法が、平行移動や 45 度単位の回転に対して不変であり、更に画像の一様性やコントラスト等の多くの性質を表すことが可能であることから、最も有効な方法であることを明らかにする。

第 5 章では、輪郭線情報とテクスチャ情報とを統合して識別を実行するシステムとして、異種情報の統合が可能な回路である階層型ニューラルネットワークについて、その基本的な性質を示すとともにそれによる輪郭線情報とテクスチャ情報との統合方法を具体的に述べる。

第 6 章では、魚画像を対象として行った、輪郭線情報とテクスチャ情報との統合による画像識別実験の結果を示す。本実験では、第 2 章から第 5 章までに述べた方法をすべて適用する。輪郭線情報は P 型記述子のパワースペクトルの低周波成分であり、テクスチャ情報は濃度共起行列の要素である。両者の情報は 3 層ニューラルネットワークで統合される。魚画像は標準画像と標準画像の回転、サイズ変換画像で構成され、それらの半分をニューラルネットワークの学習用の画像とし、残りの半分を識別テスト用の画像とする。情報の統合回路はネットワークの学習過程で形成される。輪郭線情報とテクスチャ情報とを統合して行った画像識別結果が、両者の情報を個別に用いた場合に比べて格段に良いことを明らかにし、両者の情報の統合がきわめて有用であることを実証する。

第 7 章では結論として、本研究で得られた結果を要約するとともに、今後の課題を述べる。

論文審査の結果の要旨

最近の情報処理技術の進展にともない、パターン認識の具体的な応用の可能性が一層拡がりを見せている。なかでも画像認識・識別の技術は、各種生産現場のみならず医療・福祉、通信における多くの場面において、一層の具体化が期待されている。特に簡便で安価な画像識別の方式の開発の要求は、今後更に強まるものと思われる。

一般に画像識別の技術は、対象とする画像の物理的及び情報学的性質に大きく依存する。従って識別の対象によって、適切な“特徴”を見出していく必要がある。2次元静止画像の場合、特徴として輪郭線情報とテクスチャ情報とが一般に有効な情報として取り上げられ、その記述法が様々に検討されてきた。しかしながら多くの場合それぞれの情報単独では十分な結果は得られず、これらを統合して用いることの必要性も指摘されている。

本論文は、輪郭線情報とテクスチャ情報とを統合して用い、安価で効率の良い画像識別システムの開発に関する研究をまとめたものである。すなわち、この分野でのこれまでの研究成果を調査すると共に、輪郭線情報及びテクスチャ情報それぞれの記述法として最良のものを指摘し、その基礎となる輪郭線の抽出法を提案し、情報の統合法として3層ニューラルネットを用いる方法を提案している。更にこれらを総合して用いれば、魚画像を識別するシステムが効率良く構成されることを示している。この方法はパターン認識研究において重要なテーマの一つとされている、異種情報の統合によるパターン識別の例ともなっている。

第1章においては、研究の意義とその背景及び論文の目的を明らかにしている。

第2章においては、輪郭線情報の記述法に関してこれまで提案されている多くの方法の中で、P型フーリエ記述子を用いることの優位性と、前処理として輪郭線の等辺多角形近似が有効であることを示している。この等辺多角形近似法として、著者等が開発したEPAC法が重要な寄与をしている。

第3章においては、第2章の基礎となる輪郭線の抽出が安定して行える方法を提案している。これはロバストエネルギー最小化による輪郭線の平滑化を実行する方法であり、既存の方法に比べていくつかの点において有利であることを実証しており、これ自体の今後の発展も見込むことができる。

第4章においては、テクスチャ情報の記述法に関してこれまで提案されている多くの方法を比較検討しその中で、濃度共起行列の要素を用いることの優位性を示している。

第5章においては、ニューラルネットワークに関する基本事項を述べ、異種情報の統合法として3層のニューラルネットワークが有効に機能することを具体的に示している。

第6章においては、第2章から第5章までの研究結果を総合して、輪郭線情報とテクスチャ情報の統合による魚画像の識別システムを構成し、その結果を述べている。これによると、輪郭線情報とテクスチャ情報をニューラルネットの学習により統合すれば、それぞれの情報を単独で用いるのに比べて、格段に識別率が向上することが分かり、実用上きわめて興味深い。

第7章においては、本論文の結論を述べ、今後の課題について論じている。

以上要するに、本論文は安価で効率の良い画像識別システムの実現に向けて、輪郭線情報とテキスト情報とを統合することによって識別率を向上させる方法を具体的に提案し、魚画像の識別に応用してその有効性を実証したものであり、工学上寄与するところが大きい。よって本論文は博士(工学)の学位論文に値するものであると認められる。

最終試験の結果の要旨

本論文の全般にわたって著者の説明を受けた後、審査委員が口頭によって試問を行った。すなわち、1)P型記述子における等辺多角形近似の意味、2)提案されている輪郭線の平滑化法におけるアルゴリズムの収束性、3)濃度共起行列の回転不変性、4)魚画像における特徴抽出法の妥当性、5)今後に残された問題、などについて質問があり、何れに対しても明らかな解答が得られた。

次に本論文について情報伝達専攻の公開発表会が開かれ、本学関係研究室、他大学及び地元企業から多数の出席者があった。著者の発表の後の討論では、輪郭線記述におけるデジタル誤差の影響、ロバストエネルギー最小化問題におけるパラメータの意味、処理画像の濃度レベル数と共起行列との関係、システムの実現上の問題点等について活発な質問がなされ、著者の説明によって何れに対しても十分な理解が得られた。よって著者は最終試験に合格したものと認めた。

Subject

Studies on Discrimination of Images by Unifying Contour Information and Texture Information

Abstract

Although the contour information and the texture information of an image are frequently used separately, a simultaneous use of the two will lead to more successful results. This thesis describes studies on image discrimination by unifying contour and texture information, which is consisted of 7 chapters.

Chapter 1 is an introduction, which describes background and problems of studies, and summarizes this thesis.

In chapter 2, various methods which describe contour information are examined. As the result, we show that the power spectrum obtained by P-type Fourier descriptor is most efficient. Furthermore, we show that to extract P-type descriptor accurately, the equilateral polygon approximation of contour is needed, and propose an algorithm for equilateral polygon approximation (EPAC).

In chapter 3, the edge smoothing algorithm based on the minimized robust energy is proposed for stable contour information extraction. In many case, for image discrimination by contour, using global information of a figure is more important than using local one. The proposed algorithm can reduce local noise and can smooth edge of a figure without losing global information. Moreover, an experimental result proves that this algorithm is useful.

In chapter 4, various methods which describe texture information are examined. As the result, we show that the co-occurrence matrix is most efficient.

In chapter 5, the basic property of the multi layer neural network which unifies contour and texture information to discriminate images is presented. Moreover, the process of unifying two is shown concretely.

In chapter 6, an examination of fish images discrimination by unifying contour and texture information is shown. In this examination, all methods described in chapter 2 to 5 are applied. Contour features are lower components of the power spectrum by P-type descriptor, and texture features are elements of the co-occurrence matrix. Three layer neural network unifies these features and discriminates images. This experimental result clarifies that the discrimination rate obtained in case of the simultaneous use of two is much better than rates obtained in case of the sole use of two respectively. This shows that the two kinds of information is proved extreme efficient for figure discrimination.

Chapter 7 is a conclusion, which summarizes the result of studies and presents future subjects.

目次

第 1 章 序論	4
1.1 研究の目的とその背景	4
1.2 論文の概要と構成	10
第 2 章 等辺多角形近似による輪郭線特徴の抽出	12
2.1 まえがき	12
2.2 種々の曲線記述法	13
2.2.1 偏角関数	13
2.2.2 曲率関数	15
2.3 フーリエ記述子	17
2.3.1 輪郭線のフーリエ記述子	17
2.3.2 G型記述子	17
2.3.3 Z型記述子	18
2.3.4 P型記述子	19
2.4 E P A C による閉曲線の記述	20
2.4.1 デジタル輪郭線に対する等間隔再標本化	20
2.4.2 再標本化の際の標本点間隔について	24
2.4.3 E P A C によるデジタル閉曲線の記述	26
2.5 閉曲線の前処理と E L P A	28
2.5.1 曲線の単調性	28
2.5.2 閉曲線の前処理と E L P A	31
2.6 まとめ	32
第 3 章 ロバストエネルギー最小化による輪郭線保存平滑化	34
3.1 まえがき	34
3.2 エネルギー最小化による平滑化	35
3.2.1 最小 2 乗エネルギー	35
3.2.2 M 推定エネルギー	36
3.2.3 L M e d S 推定エネルギー	37
3.3 ノイズあり波形に関する平滑化実験	39

3. 3. 1	平滑化実験	39
3. 4	まとめ	44
第 4 章	種々のテクスチャ特徴による画像解析	45
4. 1	まえがき	45
4. 2	Hough 変換を用いる解析法	46
4. 2. 1	Hough 変換による直線検出	47
4. 2. 2	Hough 変換による直線画像の解析	49
4. 3	濃度ヒストグラムを用いる解析法	51
4. 4	自己相関関数を用いる解析法	52
4. 5	フーリエ・スペクトルを用いる解析法	53
4. 5. 1	2 次元フーリエ変換	54
4. 5. 2	フーリエ・スペクトル	54
4. 6	差分統計量を用いる解析法	55
4. 7	濃度共起行列を用いる解析法	56
4. 7. 1	濃度共起行列	56
4. 7. 2	14 種類の特徴	59
4. 8	ランレングス行列を用いる解析法	62
4. 9	フラクタルを用いる解析法	65
4. 9. 1	フラクタル次元	65
4. 9. 2	フラクタル特徴の画像解析への応用	66
4. 9. 3	フラクタル特徴の拡張	68
4. 10	まとめ	69
第 5 章	ニューラルネットワークによるパターン識別	71
5. 1	まえがき	71
5. 2	ニューラルネットワークの基礎原理	72
5. 2. 1	ニューロンとニューラルネットワークモデル	72
5. 2. 2	ニューラルネットワークの学習	73
5. 2. 3	ヘップの学習則	74
5. 2. 4	相関学習則	74
5. 2. 5	デルタルール	75
5. 3	3 層ニューラルネットワークとその学習	75
5. 3. 1	3 層ニューラルネットワークの構造	75

5.3.2	バックプロパゲーションによる学習	76
5.4	情報の統合と識別	78
5.5	まとめ	79
第6章	輪郭線情報とテクスチャ情報の統合による魚画像の識別	80
6.1	まえがき	80
6.2	識別システム	81
6.3	魚画像の特徴抽出	83
6.3.1	魚画像一覧	83
6.3.2	輪郭線特徴の抽出	95
6.3.3	テクスチャ特徴の抽出	97
6.4	特徴の統合と識別実験	99
6.4.1	特徴の統合	99
6.4.2	魚画像識別実験	101
6.5	まとめ	107
第7章	結論	109
	謝辞	111
	参考文献	112
	図目次	120
	表目次	123

第 1 章

序論

1. 1 研究の目的とその背景

近年、半導体技術やプリント基板への電子部品実装技術のめざましい進歩により、計算機を用いた画像処理に関する研究と産業への応用が盛んに行われている。

計算機処理の対象となる画像は、主としてカメラやP S Dなどの光学センサによってメモリに取り込まれ、2次元の直角座標連続空間上に広がる2変数の実数関数で表される。この関数は画像の各点（画素）における強度を表している。一般に、モノクロ画像では可視光の明るさ（輝度、濃度）を強度として用い、カラー画像では赤・緑・青（R G B）という光の3原色の明るさを強度として用いる。ただし、処理する目的や対象により、赤外線、紫外線、X線など非可視領域の光や電磁波を、対象物に照射して得られた強度を用いる場合もある。また、カラー画像の実際の処理においては、R G B画像を一旦、H S Iや $L^*c^*h^*$ 、 $Y P_r P_b$ など、明るさの情報と色の情報が分離したカラーモデルに変換し、それらの強度を画素の値として用いる場合が多く[1][2]、必ずしも人間の目を見たものと同質の画像を処理するものではない[3][4][5]。

画像処理は、3つの分野に大きく分けることができる。第1の分野は画像変換である。この分野では、幾何学変換、フィルタリング、階調補正などの処理を行い、ノイズやぼけなどで劣化した画像を修復したり、人間の目では検出しにくい画像の微妙な変化を強調したりする。こうした処理により、元の画像

を人間の目で特徴づけしやすい画像に変換する。この分野の処理は、テレビ受像機の画質改善や、人工衛星から送られるリモートセンシング画像の処理、医用画像の処理等を中心として発展している。第2の分野は画像符号化である。この分野では画像を何らかのデータで記述したり、記述されたデータから画像を生成したりする。この分野の処理は、通信、データベース化における画像情報の圧縮や、CAD/CAMでの画像生成を中心として発展している。第3の分野は画像計測、画像認識である。この分野では、画像の中から対象物を抽出してその面積、個数、明るさなどを計測したり、対象物の特徴を抽出してその特徴をもとに識別や理解を行ったりする。この分野の処理は産業界の多くで、特に製造現場での生産品の検査や選別などを中心として発展してきた。機械による計測・識別は、人間によるものと比較して高速で信頼性が高い。従って、今後一層進む企業の合理化や品質管理の高度情報化のためにも、計測や識別の機械化はますます進展し、また福祉や交通などの新たな領域にも展開していくであろう。そのためこの分野の研究は更に発展していくと考えられる。

画像識別を行なうには、画像から識別すべき対象物の性質を表わす特徴を抽出する必要がある。画像は画像空間から特徴空間に写像され、特徴の比較によって識別がなされる。ここで、対象物の性質を表す一般的な情報として、その形状や模様が挙げられる。

形状は輪郭線で表され、通常閉曲線となる。画像から対象物の輪郭線を抽出・記述して、対象物の性質や特徴を把握する方法はこれまでも多くの研究がなされており、実用的な種々の提案もなされている。これらの輪郭線記述方法は、記述に用いる変数によって次の3種類に大きく分けられる。

①座標値を直接用いる方法

入力画像をある座標系と考えて、輪郭線上の画素の座標を直接記述する方法である。G. H. Granlund は画像を複素平面にとらえ、輪郭線を実数軸方向と虚数軸方向の座標値の配列で記述し、この方法を文字認識に応用している [6]。

②輪郭線の傾きを用いる方法

これは座標系に対する輪郭線の接線の傾きを用いて記述する方法である。この方法による記述は、対象物の平行移動に対して不変であり、拡大・縮小に対しては輪郭線の長さを正規化しておけば不変となる。回転に対して

は記述それ自体は変化するが、多くは簡単な操作で不変なパラメータを得ることができる。例えば、8連結のデジタル曲線を8方向の傾きコードで記述したものがFreemanのチェインコードである。また安居院、中島らは、閉曲線を偏角関数と呼ばれる、接線の傾きを使った関数で記述する方法を提案しており、湖の識別を行っている[7]。上坂は、偏角関数を複素空間上の位相と考え、その複素数値関数で記述する方法を考案し、このフーリエ・スペクトルをP型記述子と呼んだ。このP型記述子が、輪郭線情報として優れた性質を持つことを理論的、実験的に示している[8]。

滝山、外波らは、横顔シルエットの輪郭線を、EPACと呼ばれる等辺多角形近似関数で表現して、P型記述子を用いて同定実験を行っている[9]。また、滝山と矢野は、輪郭線の大域的な特性を失わずにより辺の数が少ない等辺多角形で輪郭線を記述する方法として、一旦辺の数が多き等辺多角形で近似してP型記述子を求め、そのパワー・スペクトルから高周波成分を取り除いてできた閉曲線を、再び辺の数が少ない等辺多角形で近似するELPA法を提案している[10][11][12]。

③輪郭線の曲がり具合を用いる方法

これは曲線を曲線自身に対する曲がり度で記述する方法である。理論的には、②の記述を微分すればこの記述となる。従って、平行移動に加えて回転に対しても不変である。拡大縮小の変化に対しては、曲線の長さを正規化すれば不変となる。また、閉曲線に対してこの記述は周期関数となる性質を持ちあわせている。このため、このフーリエ・スペクトルを不変なパラメータとして用いることが可能となる。白井は曲線の曲がり具合をラジアンで表わしたものを曲率関数と呼び、曲がりの大きいコーナー部の検出等に用いている[13]。安居院、中島らは偏角関数を微分した記述のパワー・スペクトルを用いて、アニメーションの閉領域同定に応用している[14]。

輪郭線の記述方法について重要な点は3つある。第1に、対象物の幾何学変換に対してその記述が不変であること、第2に、曲線の記述からもとの曲線が再生可能であること、第3に、解析的なレベルではなくデジタル画像に対する処理としてみた場合に、デジタル化誤差の影響を受けにくいものであること、などである。本論文では以上の点を考慮して、平行移動、拡大・縮小、回転に対して不変な特徴である、EPAC法に基づくP型記述子を利用すること

が最も有利であることを示し、実際に用いることにする。

ところで、輪郭線を特徴とした実際の画像の識別においては、輪郭線にノイズ成分が含まれることが多い。ノイズが混じっていると、輪郭線の記述アルゴリズムの収束性が悪化したり、識別の正解率が低下することが多い。このため輪郭線の性質を保存しつつノイズを除去する方法が要求される。輪郭線からのノイズ除去方法については現在までにいくつかの手法が提案されているが[15] [16] [17] [18] [19]、除去できるノイズの性質が制限されていることや輪郭線の保存能力が低い、あるいは輪郭線の復元能力は高いが多くのパラメータを適切に設定する必要があるなど、一般化には難がある。このため、本論文では新たなノイズ除去・輪郭線保存手法としてL M e d S推定法[20]を提案する。L M e d S推定法は処理に用する計算量が多いという短所がある反面、輪郭線の復元能力が比較的高く、かつパラメータ数が少ないという長所がある。このことにより、用いる輪郭線やノイズの種類によらずにほぼ安定して良好な輪郭線の復元が可能であると考えられる。本論文では、ノイズの混入した波形信号について復元実験を行い、L M e d S推定法の有効性を実証する。

模様はテクスチャと呼ばれ、これは局所的にはランダムに見える画像の中にいくつかの基本的な構成要素が存在し、それらがある種の規則のもとに配列されてできる繰り返しパターンを意味する。テクスチャの解析は、画像理解の基礎をなすものとして、これまでに多くの研究がなされている[21] [22]。また、近年ではリモートセンシング画像や医用画像、3次元物体認識などのあらゆる画像処理の手法として応用されている。テクスチャ解析の応用法は以下の4つの分野に大きく分けられる。

①テクスチャの識別

テクスチャ自身が持つ固有の特徴を抽出してモデル化を行い、画像パターンを識別する。リモートセンシング画像や航空写真による砂漠地帯や森林地帯、都市部等の識別や、岩石の写真による花崗岩や大理石などの識別を行なうためにテクスチャモデルを作成することなどがこれに相当する。

②テクスチャによる等質領域の抽出

一枚の画像の中から異なるテクスチャの領域を識別する。航空写真における畑や家、道路などの領域の抽出や、レントゲン写真からの病巣領域の抽出などがこれに相当する。

③テクスチャの変化を利用する境界領域の抽出

②と同様の事象に用いられ、テクスチャの性質が異なる部分を領域の境界として抽出することなどがこれに相当する。

④テクスチャの応用

テクスチャを合成して計算機から自動的に画像を発生させたり、テクスチャのきめの細かさから物体の距離の測定を行なうことなどがこれに相当する。

テクスチャによる画像の解析に関しては現在までに数多くの研究が行なわれ、各種の方法が提案されている。解析方法は大きく分けると2つに分類される。

1つは構造レベルでの解析であり、直線や点などのテクスチャを構成する基本的な要素を抽出し、それらの配列規則を求める方法である。G. Eichmann と T. Kasparis は、Hough 変換を用いて直線で構成されたテクスチャ画像を解析して特徴を抽出する方法を提案している [23]。構造レベルの解析は、航空写真における道路や建造物などの人工物など、単純な形状の要素で構成された画像に適しているが、岩石や芝生、布地、木目など自然発生したテクスチャでは、その構成要素を求め、配列規則を知ることが容易でなく、従って解析できる対象物は限定される。

もう1つは統計レベルでの解析であり、画素の濃度に着目して、その一様性、方向性、コントラスト変化などの画像の性質を表す統計量を求める方法である。統計レベルでの解析方法として、濃度ヒストグラムを求めそのヒスト値の平均や分散から特徴を求める方法や、2次元の自己相関関数を用いる方法、差分統計量から平均やコントラスト、2次モーメントなどを求める方法がある。R. M. Haralick らは特定の相対的な位置関係にある画素対の濃度の累算値から濃度共起行列を求め、更にその行列からコントラストや2次モーメント、相似性などの14種類の性質を計算して特徴として用いる方法を提案している [24]。また、M. M. Galloway は画像の符号化に用いられるランレンジスから行列を求め、その行列をもとに、画素値のラン方向の不均一性やランの長さの不均一性などを計算して特徴とする方法を提案している [25]。ランレンジス法に関して K. Chehdi と Q. Liao は、濃度によって感度が異なるランを作成した識別実験で比較的高い正解率を得ている [26]。金子はコンピュータ・グラフィクスな

どで用いられるフラクタルを応用した特徴抽出方法を提案している [27]。

これらの種々のテクスチャ特徴の性能に関する研究も行なわれており、Weszka や Conners [28] の研究によると、統計的な解析法のなかでは、濃度ヒストグラムによる方法が最も弱く、濃度共起行列による方法が最も強い。その他の方法は同じ程度の強さであるという。画像の幾何学変換については、濃度共起行列とランレングス行列が 45° 、 90° 、 135° の回転に関して独立しているほかは、すべて幾何学変換により値が変化することを示している。幾何学的な変換に対応できる方法として、F. S. Cohen らが提案するガウス＝マルコフ確率場を用いてテクスチャをモデル化する方法があるが [29]、モデル化の過程が複雑で時間も大量にかかるためあまり有効ではない。

近年、梅田は濃度共起行列による方法に関して、14 種類の特徴を用いず、行列の要素の値そのものを用い、更に識別系に階層構造のニューラルネットワークを利用した識別実験で高い正解率を得ている [30]。また、滝山と矢野は、同様に濃度共起行列の要素とバックプロパゲーションモデルを用いた実験を行ない、識別に用いる画像を再量子化して 8 階調程度で十分高い正解率を得られることと、行列を作成する際に用いる画素対の距離を 1 画素にすると最も高い正解率を得られることを報告している。更に、14 種類の特徴を用いる実験では、14 種類の中から 9 ～ 12 個のより有望な特徴を用いて識別を行なう方が 14 種類の特徴を全て用いた識別よりも高い正解率を得られることを提案している [31]。

本論文では以上の内容を考慮して、平行移動および 45° 単位の回転に対して不変であり、一様性やコントラストなど画像の多くの性質を表現できる特徴である濃度共起行列の行列の要素が最も有利であることを示し、実際に用いることにする。

これまで述べたように、輪郭線やテクスチャはそれぞれに画像の性質を表す重要な情報であるが、これらを単独に用いたのでは画像識別がうまく実行できない場合もまた多い。これまでに行われてきた輪郭線やテクスチャに関する研究は、これらの情報を単独で用いたものが多かった。しかしながら、輪郭線情報とテクスチャ情報との双方を統合して用いれば、単独情報では識別が困難であった画像をも正しく識別できる場合が多い。より一般には性質の異なる複数の情報を抽出して統合する識別法の開発が真に望まれる。その中でも輪郭線情

報とテクスチャ情報は基本的情報とみなされるので、両者の情報の統合について論じることは極めて重要である。

本論文では、輪郭線情報とテクスチャ情報の統合の有効性について論じ、実験をも行ってこれを実証する。情報の統合にはニューラルネットワークを用いる。ニューラルネットワークは学習機能を持ち、異種情報の統合には極めて適切な回路である。具体的な研究対象とする画像は魚類図鑑 [32] 中の魚画像である。以上の研究を通して輪郭線情報とテクスチャ情報とを統合する方法を与え、それが画像識別に極めて有用であることを示す。

1. 2 論文の概要と構成

本論文は、1.1 で述べた輪郭線情報とテクスチャ情報とを統合して行う画像識別に関する研究をまとめたものであり、7章から構成されている。

以下に、本論文の概要と構成を示す。

第1章では、序論であり本研究の背景と扱っている問題を示し、あわせて論文の概要について述べる。

第2章では、輪郭線情報の種々の記述方法について論じ、分析を行う。その結果最も有用な輪郭線情報の記述法として、E P A C法に基づくP型記述子を選び、この方法に関して詳細を述べる。

第3章では、第2章で述べた輪郭線情報を安定して抽出する方法として、L M e d S推定法すなわち、ロバストエネルギー最小化による輪郭線保存平滑化法を提案する。これは大域的な情報を損なわず、ノイズ等の不要な局所情報を取り除き輪郭線を平滑化する方法である。その有用性を実験を通じて更に明らかにする。

第4章では、テクスチャ情報の種々の記述方法について論じ、分析を行う。その結果最も有用なテクスチャ情報の記述法として、濃度共起行列の行列の要素を選び、この方法に関して詳細を述べる。

第5章では、ニューラルネットワークを用いて輪郭線情報とテクスチャ情報とを統合するために、ニューラルネットワークの基本事柄をも含めて述べ、これによる情報の統合と識別法を具体的に示す。

第6章では、これまで述べてきた輪郭線情報とテクスチャ情報とを統合して画像識別を行なう方法を、具体的な問題に適用した結果について詳述する。実験の対象とする画像は魚画像である。これらの画像を標準画像とし、標準画像から回転及びサイズ変換画像を作成する。標準画像、変換画像を2つに分け、半分をニューラルネットワークの学習用の画像とし、残りの半分を識別実験用のテスト画像とする。これらの画像に対して第2章から第5章までに述べた方法をすべて適用して、輪郭線情報とテクスチャ情報の統合が画像の不変識別に極めて有効であることを実証する。

最後に第7章では、本研究のまとめと今後の課題を述べる。

第 2 章

等辺多角形近似による輪郭線特徴の抽出

2. 1 まえがき

画像中の対象物の性質を表す一般的な情報として、その形状や模様が挙げられる。本章では、このなかで形状情報を用いた特徴の抽出について説明する。形状は輪郭線で表され、通常閉曲線となる。画像から対象物の輪郭線を抽出・記述して、対象物の性質や特徴を把握する方法はこれまでに多くの研究がなされており[33]、実用的な種々の提案もなされている。これらの輪郭線記述方法は、記述に用いる変数によって次の 3 種類の方法に大きく分けられる。

①座標値を直接用いる方法

入力画像をある座標系と考えると、輪郭線上の画素の座標を直接記述する方法である。G. H. Granlund は画像を複素平面にとらえ、輪郭線を実数軸方向と虚数軸方向の座標値の配列で記述し、この方法を文字認識に応用している[6]。

②輪郭線の傾きを用いる方法

これは座標系に対する輪郭線の接線の傾きを用いて記述する方法である。この方法による記述は、対象物の平行移動に対して不変であり、拡大縮小に対しては輪郭線の長さを正規化しておけば不変となる。回転に対しては記述それ自体は変化するが、多くは簡単な操作で不変なパラメータを得ることができる。例えば、8 連結のデジタル曲線を 8 方向の傾きコードで記述したものが Freeman のチェーンコード

である。また安居院、中島らは、閉曲線を偏角関数と呼ばれる、接線の傾きを使った関数で記述する方法を提案しており、湖の識別を行っている[7]。上坂は、偏角関数を複素空間上の位相と考え、その複素数値関数で記述する方法を考案し、このフーリエ・スペクトルをP型記述子と呼んだ[8]。このP型記述子が、輪郭線の特徴パラメータとして優れた性質を持つことを理論的、実験的に示している。

③輪郭線の曲がり具合を用いる方法

これは曲線を曲線自身に対する曲がり度で記述する方法である。理論的には、②の類の記述を微分すればこの記述となる。従って、平行移動に加えて回転に対しても不変である。拡大縮小の変化に対しては、曲線の長さを正規化すれば不変となる。また、閉曲線に対してこの記述は周期関数となる性質を持ちあわせている。このため、このフーリエ・スペクトルを不変なパラメータとして用いることが可能となる。白井は曲線の曲がり具合をラジアンで表わしたものを曲率関数と呼び、曲がりの大きいコーナー部の検出等に用いている[13]。安居院、中島らは偏角関数を微分した記述のパワー・スペクトルを用いて、アニメーションの閉領域同定に応用している[14]。

輪郭線の記述方法について重要な点は3つある。第1に、対象物の幾何学変換に対してその記述が不変であること、第2に、曲線の記述からもとの曲線が再生可能であること、第3に、解析的なレベルではなくデジタル画像に対する処理としてみた場合に、デジタル化誤差の影響を受けにくいものであること、などである。

2. 2 種々の曲線記述法

2. 2. 1 偏角関数

図 2.2.1 で示すように、平面上の輪郭線 C 上に任意の点 A をとり、 A から C に沿って測った弧の長さを l とする。 l をパラメータとして C を $Z(l)$ で表す。ただし、 l の符号は時計回りを正とする。 C を一周したときの長さを L として、次に定義する $\phi_z(l)$ で閉曲線を表す。

$$\phi_z(l) = \theta_z(l) - 2\pi/L \quad (2.2.1)$$

但し、 $\theta_z(l)$ は距離 l での点の接線の偏角とする。

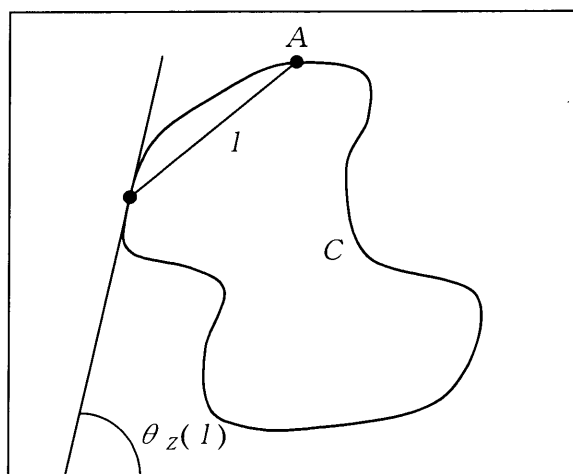


図 2. 2. 1 輪郭線の偏角

この $\theta_z(l)$ の図形的な意味は次のように考えられる。図 2.2.2 のように、閉曲線 C の接線ベクトルを $r(l)$ で表す。また一方、閉曲線 C と周囲長が等しい円を考えて、その円の下点を始点としたときの接線ベクトルを $t(l)$ で表せば、 $r(l)$ と $t(l)$ とのなす角 ϕ がその点における閉曲線 C の偏角関数 ϕ となる。

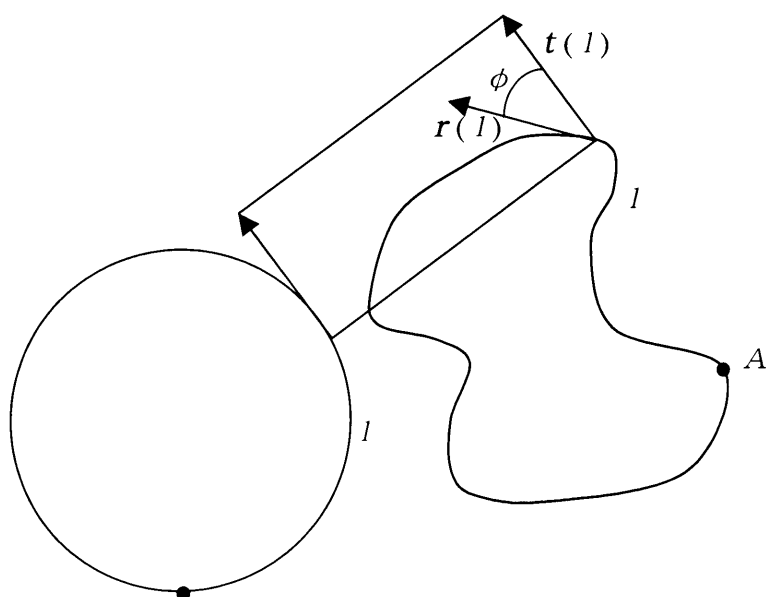


図 2. 2. 2 偏角関数 ϕ

$\phi_z(1)$ は周期 L の周期関数であり、閉曲線の大きさに依存するため、以下のよう
に $\phi_z^*(x)$ を用いて $[0, 2\pi]$ で正規化する。

$$\phi_z^*(x) = \theta_z(Lx/2\pi) - x \quad (2.2.2)$$

このとき、 $\phi_z^*(0) = \phi_z^*(2\pi) = \theta_z(0)$ である。ここで $\phi_z^*(x)$ は、閉曲線の移動や拡大・縮小に対して不変な記述である。

デジタル輪郭線に対する偏角を図2.2.3に示す。

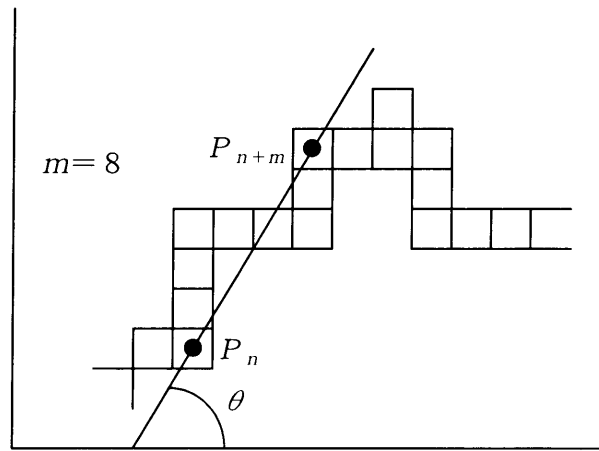


図2.2.3 デジタル輪郭線での偏角

このとき注目すべき画素を P_n とし、これよりも m 個だけ先の画素を P_{n+m} として、この2点を結んだ直線を P_n における接線とする。偏角はこの接線が x 軸となす角であるから以下の式で与えられることになる。

$$\theta_z(n) = \tan^{-1} \left(\frac{y_{n+m} - y_n}{x_{n+m} - x_n} \right) \quad (2.2.3)$$

ここで (x_n, y_n) と (x_{n+m}, y_{n+m}) は画素 P_n 、 P_{n+m} の座標である。また m はステップ数で、スケールに比例させる必要がある定数である。ステップ数は小さすぎると閉曲線の大きな情報に欠け、ノイズやデジタル化誤差に影響されやすく、逆に大きすぎると局所的な情報に欠けることになる。そのためステップ数の設定には十分注意を要する。

2.2.2 曲率関数

曲率関数 $C(1)$ は、閉曲線の接線の偏角 $\theta(1)$ を微分したものである。

$$C(l) = \frac{d\theta(l)}{dl} \quad (2.2.4)$$

デジタル画像の場合、微分の代わりに差分を用いて得られる。すなわち、

$$\begin{aligned} C(n) &= \theta(n) - \theta(n-m) \\ &= \tan^{-1} \left\{ \frac{y(n+m) - y(n)}{x(n+m) - x(n)} \right\} - \tan^{-1} \left\{ \frac{y(n) - y(n-m)}{x(n) - x(n-m)} \right\} \end{aligned} \quad (2.2.5)$$

で得られることになる。図 2.2.4 で示すように、注目する画素 P_n と、これよりも m 個だけ前後の画素 P_{n-m} と P_{n+m} とをそれぞれ直線で結んで、これらが交わる補角を P_n における曲率と定義する。

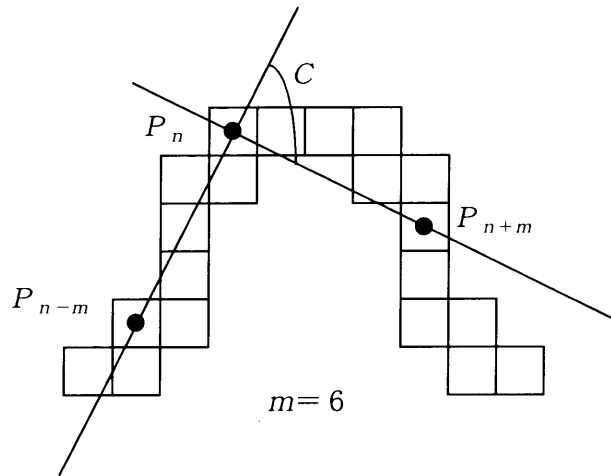


図 2.2.4 曲率 C

曲率関数は、輪郭線が閉曲線であれば周囲長を周期とする周期関数となる。パラメータ n が時計回りを正の方向とするならば、閉曲線の凸の部分に対して負の値、凹の部分に対して正の値（ラジアン）を示す（ n が反時計回りを正の方向とするのであれば正負は逆転する）。また閉曲線の移動、回転に対しては不変であり、拡大・縮小に対しても周囲長を正規化すれば不変となる。更に始点の移動に対しては、その位相がずれるだけであるので、このパワースペクトルを不変な特徴パラメータとすることが可能となる。

また曲率関数は、閉曲線の曲がりの大きいコーナー部で、正負のピーク値をとるので、視覚的に重要な特徴点であるコーナーの検出が容易である。この特徴を利用して、滝山、大石らは横顔シルエットからのあご、口、鼻、額といった対象物の測定に応用している [34]。

2. 3 フーリエ記述子

2. 3. 1 輪郭線のフーリエ記述子

輪郭線を周波数領域で表す方法としてフーリエ記述子がある[6][33]。まえがきで述べたような、座標値、傾き、曲がり具合などで輪郭線を記述したものを被展開関数としてフーリエ変換し、得られるスペクトルの組を輪郭線のフーリエ記述子と呼ぶ。輪郭線記述が離散数値関数であるならば、これを $p(j)$ 、 $j=0, 1, \dots, n-1$ (n は十分大きい正の整数) としてフーリエ記述子 q は次の離散フーリエ変換式で

$$\begin{aligned} q(k) &= \frac{1}{n} \sum_{j=0}^{n-1} p(j) \exp\left(-2\pi i \frac{jk}{n}\right) \\ k &= 0, 1, \dots, n-1 \\ i &= \sqrt{-1} \end{aligned} \quad (2. 3. 1)$$

得られる。

ここでは、G型記述子、Z型記述子、P型記述子などの代表的なフーリエ記述子について述べる。但し説明の便宜上、次のような仮定をしておく。

1. 処理の対象となる輪郭線は、複素平面上に与えられている折れ線図形であるとする。
2. 与えられている折れ線図形は各線分の長さがどれも等しいものとする。
3. 折れ線図形とその記述は一対一に対応しており、変換および逆変換が一意にできるものとする。

2. 3. 2 G型記述子

長さが δ である n 本の線分をつなぎ合わせてできる折れ線図形を C とする。 C は各線分における端点（以後接点と呼ぶ）の集合で表せられる。接点の座標を $\{x(j), y(j)\}$ とすると

$$C = \{x(j), y(j)\} , j = 0, 1, \dots, n \quad (2. 3. 2)$$

を元の輪郭線とみなすことができる。

$x-y$ 座標値を使って複素数値関数 z を次のように定める。

$$z(j) = x(j) + iy(j) , j = 0, 1, \dots, n \quad (2. 3. 3)$$

このようにすると複素数値関数 z は、折れ線図形 C と同一視することができる。

点 $z(0)$ 、 $z(n)$ が C の端点であり、 C が閉曲線の場合には、 $z(0) = z(n)$ である。2 番目の仮定により、折れ線の各辺の長さはすべて等しいので

$$\delta = |z(j) - z(j-1)|, \quad j = 1, 2, \dots, n \quad (2.3.4)$$

が成り立っている。

与えられた任意の輪郭線を上の複素数値関数で近似的に表すことは容易にできるので、一般性を失うことなく式 (2.3.4) を満たす z が輪郭線 C としてはじめに与えられているものとする。

この複素数値関数 z を被展開関数としてフーリエ変換して求められたスペクトル

$$e(k) = \frac{1}{n} \sum_{j=0}^{n-1} z(j) \exp\left(-2\pi i \frac{jk}{n}\right) \quad (2.3.5)$$

$e(k)$ を折れ線図形の「G 型記述子」と呼ぶ。

2.3.3 Z 型記述子

図 2.3.1 に示すように、折れ線図形 C 上に隣り合った線分をベクトル $z(j) - z(j-1)$ と $z(j+1) - z(j)$ とし、これらがなす角（補角）を反時計回りに正に、あるいは時計回りに負に測って、その値を $a(j)$ とする。 $a(j)$ は

$$-\pi \leq a(j) \leq \pi \quad (2.3.6)$$

の範囲にはいるようにとる。但し図 2.3.2 のように、 $a(0)$ はベクトル $z(1) - z(0)$ と x 軸とのなす角（偏角）と定める。この $a(j)$ を点 $z(j)$ における曲率という。

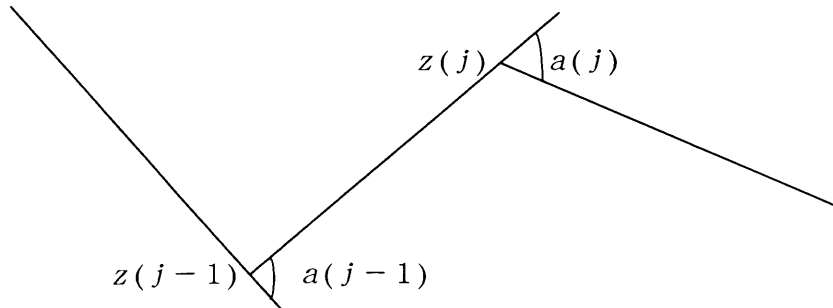


図 2.3.1 補角

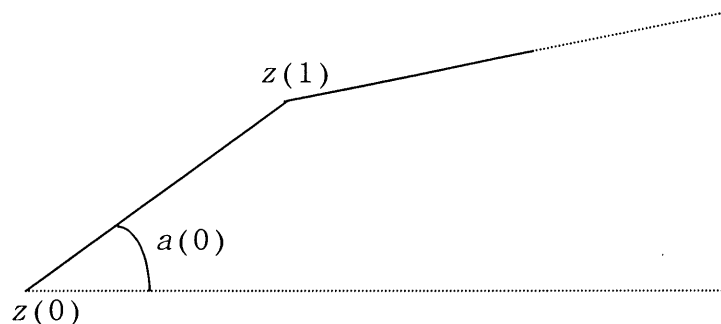


図 2. 3. 2 偏角

次に、曲率関数 a を用いて折れ線図形 C の偏角関数を

$$\begin{aligned}\theta(0) &= a(0) \\ \theta(j) &= \theta(j-1) + a(j), \quad j = 1, 2, \dots, n-1\end{aligned}\tag{2. 3. 7}$$

と定める。偏角関数はもともと滑らかな連続輪郭線上の曲率を輪郭線に沿って積分したものであるが、この概念を折れ線図形にあてはめると式 (2. 3. 7) のようになる。この θ を使って

$$\phi(j) = \theta(j) - \theta(0) - \frac{2\pi j}{n}\tag{2. 3. 8}$$

とおき、このフーリエ変換を

$$d(k) = \frac{1}{n} \sum_{j=0}^{n-1} \phi(j) \exp\left(-2\pi i \frac{jk}{n}\right)\tag{2. 3. 9}$$

とし、 d を折れ線図形の「Z 型記述子」と呼ぶ。

2. 3. 4 P 型記述子

曲率 θ を用いて複素数値関数 ω を

$$\omega(j) = \exp\{i\theta(j)\}, \quad j = 0, 1, \dots, n-1\tag{2. 3. 10}$$

と定め、これを折れ線図形 C の「P 表現」と呼ぶ。式 (2. 3. 10) において、 ω を折れ線図形 C の信号とみたとき、ちょうど位相の場所に相当している。

曲率と偏角関数の定め方から容易に分かるように、ベクトル $z(j+1) - z(j)$ と x 軸とのなす角 $\psi(j)$ は、図 2. 3. 3 に示すように $\theta(j)$ と 2π の整数倍を除いて等しい。従って、

$$\begin{aligned}\exp\{i\theta(j)\} &= \exp\{i\varphi(j)\} \\ &= \frac{z(j+1) - z(j)}{\delta}\end{aligned}\quad (2.3.11)$$

が成り立ち、式(2.3.10)により、

$$\omega(j) = \frac{z(j+1) - z(j)}{\delta} \quad (2.3.12)$$

が得られる。逆に $\omega(j)$ 、 $j = 1, 2, \dots, n$ が与えられれば、座標は

$$z(j) = z(0) + \delta \sum_{r=0}^{j-1} \omega(r), \quad j = 1, 2, \dots, n \quad (2.3.13)$$

によって求めることができる。

このP表現 $\omega(j)$ をフーリエ変換して得られる結果が「P型記述子」と呼ばれるものである。すなわち、折れ線図形CのP型記述子 $c(k)$ は

$$c(k) = \frac{1}{n} \sum_{j=0}^{n-1} \omega(j) \exp\left(-2\pi i \frac{jk}{n}\right), \quad j = 0, 1, \dots, n-1 \quad (2.3.14)$$

によって得られる。この記述子は、輪郭線の平行移動、拡大、縮小に対して不変である。また、対象とする折れ線図形が反時計回りに α だけ回転したときの折れ線図形のP型記述子による結果 $c'(k)$ は

$$c'(k) = \exp(i\alpha) c(k) \quad (2.3.15)$$

となる。従って、両者の絶対値は等しい。始点の移動についても、同様に絶対値を取ることによって不変な結果が得られる。

以上のように折れ線図形をP型記述子によって表し、その絶対値をとれば対象の位置、大きさ、向き及び輪郭線記述の際の始点の位置によらない不変な記述が得られる。滝山、王らはP型記述子のパワースペクトルを横顔シルエット認識に応用して高い識別結果を得ている[35]。

2.4 E P A C による閉曲線の記述

2.4.1 デジタル輪郭線に対する等間隔再標本化

デジタル輪郭線の認識処理などを行う場合、輪郭線を構成しているすべての画素について曲率関数などによって記述する必要はなく、ある適当な数の点を選んで

行えば十分である。すなわち、デジタル輪郭線を再標本化し、新たに定められた標本点に関して曲線の記述を行えばよいことになる。一般に、デジタル情報の処理の多くは、等間隔に標本化されたデータが仮定されており、ほとんどの場合、標本点間隔が等間隔であるほうが都合がよい。このため、与えられたデジタル輪郭線を等間隔に再標本化する必要がある。

輪郭線上の標本点間の距離の測り方には 2 つの方法がある。1 つは、輪郭線に沿った長さを用いるもので、もう 1 つは、標本化後の標本点間の直線距離を用いるものである[36][37]。

①輪郭線に沿って測った標本点間距離が等しい再標本化

輪郭線に沿った長さが等間隔になるような再標本化は、輪郭線を構成している画素の中心を直線で補間することによって比較的容易に実現できる。すなわち、直線補間されてできた輪郭線をアナログ輪郭線として見立てることによって、始点と標本点数とが与えられれば一意に標本化できる。但し、デジタル輪郭線を取り扱うので画素の連結方向に注意する必要がある。

N 個の画素からなるデジタル輪郭線を M 個の標本点に再標本化する問題を与える。まず、周囲長を計算する。その際、画素が水平もしくは垂直に連結している場合には距離を 1 として、斜めに連結している場合には $\sqrt{2}$ として測る。このようにしてもとめられる周囲長は、図 2.4.1 に示すような隣接する画素間を直線で結ぶことによって構成される、いわばアナログ輪郭線の周囲長を表しているといえる。輪郭線の標本化はこのアナログ輪郭線に対して行うことにする。

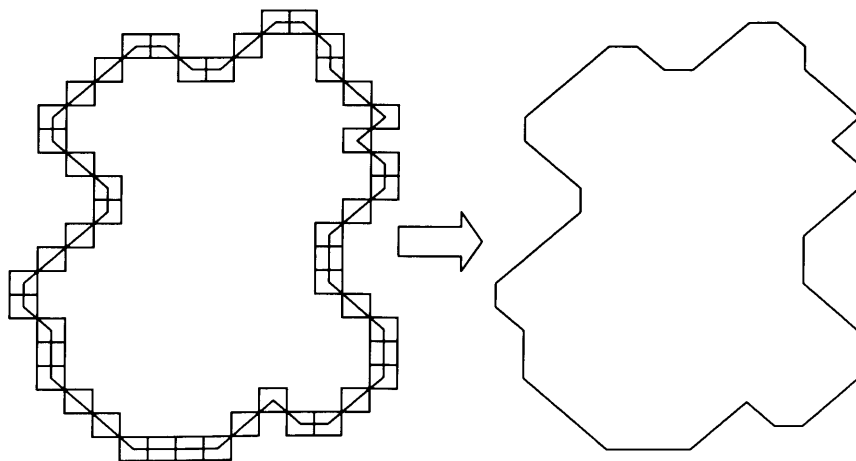


図 2. 4. 1 デジタル輪郭線の補間

入力した分割数と計算された周囲長とから、1区間の距離 R を計算する。始点を決めて、それを第1の標本点とする。第2の標本点は、輪郭線に沿って（補間された直線に沿って）距離 R だけ離れた輪郭線上、すなわち、補間された輪郭線上にとる。以後標本点についても同様に選んでいく。図 2.4.2 に示すように第 $i+1$ 番目の標本点は、第 i 番目の標本点からの距離が輪郭線に沿った長さで R の点として決められる。図 2.4.3 にこの方法による再標本化のためのアルゴリズムのフローチャートを示す。

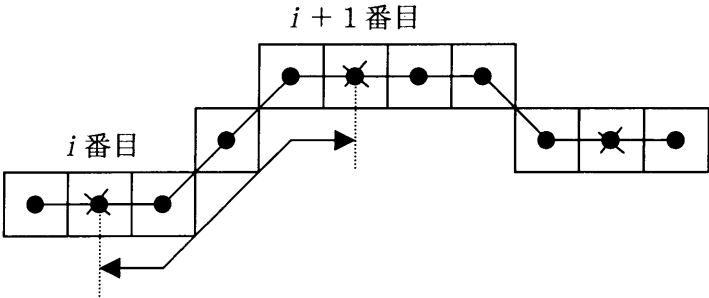


図 2. 4. 2 輪郭線に沿った長さが等しい再標本化

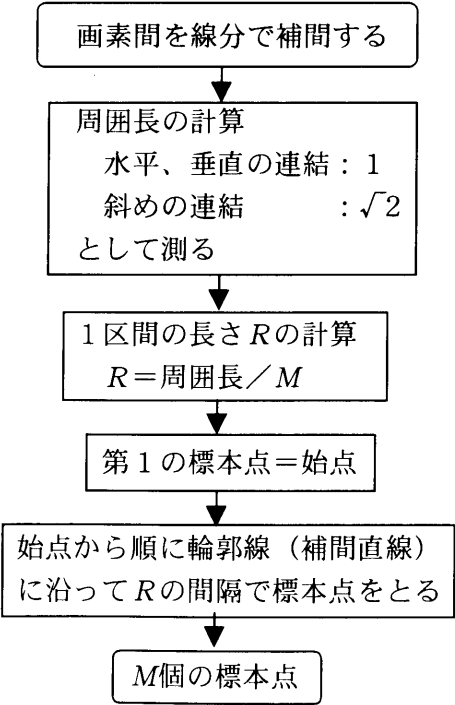


図 2. 4. 3 輪郭線に沿った長さが等しい再標本化のフローチャート

以上の処理によって等間隔標本点が求められる。この処理によって得られる標本点の座標は、画素単位（整数）よりもっと細かい単位（実数）で表されている。そこで、保存に要する資源を押さえることが必要な場合などには、得られた座標値を画素単位の値に丸める操作を行う。一般には、整数値に丸められた座標データを用いても曲率関数などの輪郭線記述の際に、あまり大きな誤差とはならないと考えられる。

②標本点間の直線距離が等しい再標本化

ここでは、輪郭線に沿った長さではなく、標本点間の直線距離が等しい再標本化の方法を説明する。閉曲線に対してこの処理を実行して得られた標本点を結ぶと、等辺多角形が得られる。従ってここで述べる方法は、 N 個の画素からなるデジタル輪郭線を等辺 M 角形で近似するための標本点の決定方法であるといえる。

①で述べた再標本化の方法と同じく、隣接する画素間を直線補間して得られるアナログ輪郭線上に標本点をとることとする。但し、標本化後の標本点間の直線距離が等しくなるように標本点を選ぶ。

まず、標本点間の直線距離の基準とする適当なユークリッド長 R_k ($k=0$) を与える。始点となる画素を決め、始点からの直線距離が R_k で補間された直線上の点を第 2 の標本点とする。以後、同様な操作を行って標本点を決めていく。例えば第 $i+1$ 番目の標本点は、図 2.4.4 に示すように第 i 番目の標本点からの直線距離が R_k の補間直線上の点である。

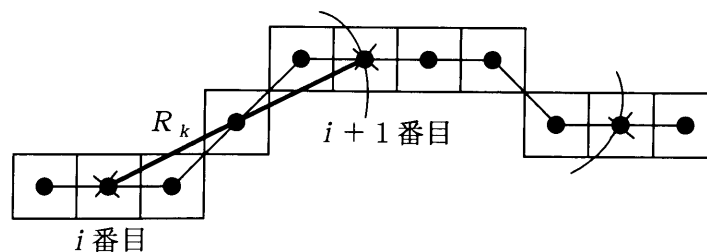


図 2. 4. 4 直線距離で等間隔な再標本化をするときの標本点の決定方法

このようにして標本点を選んでいけば、標本点間の直線距離 r_i ($i=1,$

…， $M-1$ ）はすべて R_k となる。しかしながら、最後の標本点と始点との距離（最後の辺の長さ） $L_k (= r_M)$ は必ずしも R_k とはならず、他の標本点間距離と誤差が生じる。そこで基準となるユークリッド長 R_k を修正する。ここでは最後の辺の長さ L_k と基準ユークリッド長 R_k との誤差を用いた式(2.4.1)によって、基準ユークリッド長 R_k を修正する。

$$R_{k+1} = R_k + \alpha \frac{L_k - R_k}{M}, \quad \alpha > 0 \quad (2.4.1)$$

この式を用いれば、 L_k が R_k よりも小さい場合にはその誤差の大きさに応じて R_k は小さくなり、逆に L_k が R_k よりも大きい場合には R_k は大きくなる。すなわち R_k は、 L_k と R_k との誤差の大きさに比例した大きさの分だけ修正される。この式によって得られた R_k を基準ユークリッド長として再び始点から再標本化を行う。式(2.3.16)で k はアルゴリズムの繰り返し回数を表している。 R_k は k 回繰り返し後の基準ユークリッド長であり、 L_k は k 回繰り返し後の最後の辺の長さである。また、 α はこの繰り返し過程が収束するように与えられる正の実数である。なお、基準ユークリッド長の初期値 R_0 は $R_0 = N/M$ によって与えられる。

この繰り返し過程は、基準ユークリッド長 R_k と最後の辺の長さ L_k との誤差がある値 d より小さくなったときに収束するものとする。そのときに得られる M 個の標本点を、直線距離が等しい標本点間隔の標本点として採用する。図 2.4.5 にこのアルゴリズムのフローチャートを示す。なお、必要であれば、求められた結果を整数値に丸める操作を行う。

2.4.2 再標本化の際の標本点間隔について

デジタル輪郭線を等間隔に再標本化する際、その間隔をどのように決めるかは重要な問題である。間隔が小さいとデジタル誤差が生じやすい。例えば、間隔として 1 画素程度の値を設定すると 8 方向だけしか表現できない。これは 1 周 360° の方向を 8 方向に量子化していることになり、その際の量子化誤差は著しく大きくなる。逆に間隔が大きい場合には、輪郭線の細かい変化が無視されやすくなる。従って、再標本化の際の標本点の間隔は注意して決める必要がある。一般に、輪郭線が複雑に入り込んでいる場合には間隔を小さくして、緩やかな場合には比較の間隔

を大きくするのが適当である。輪郭線の複雑さを表す尺度として、式(2.4.2)で表される複雑度 C がしばしば用いられる。

$$C = \frac{p^2}{S} \quad (2.4.2)$$

すなわち、 C は輪郭線の周囲長 p の自乗を輪郭線図形の面積 S で割ったものである。正円の場合に最も小さい値となり、複雑に入り込んだ輪郭線であるほど大きい値となる。この C に応じて標本点間の間隔を設定すればよい。

一方、デジタル輪郭線を再標本化する目的は、後に続く各種の認識処理などを円滑にするためであるから、標本点の数についても考慮すべきである。輪郭線記述の方法として2.3で述べたフーリエ記述子がある。また、輪郭線を偏角関数や曲率関数などを用いて記述した場合、回転について輪郭線の不変な認識処理を行うためにはフーリエ変換が必要である。従って、データの数 2 のべき乗とした方が後の計算処理の際に有利である。

以上のことから再標本化の際の標本点の数は、輪郭線の複雑度に応じて 2 のべき乗の値に設定すればよい。

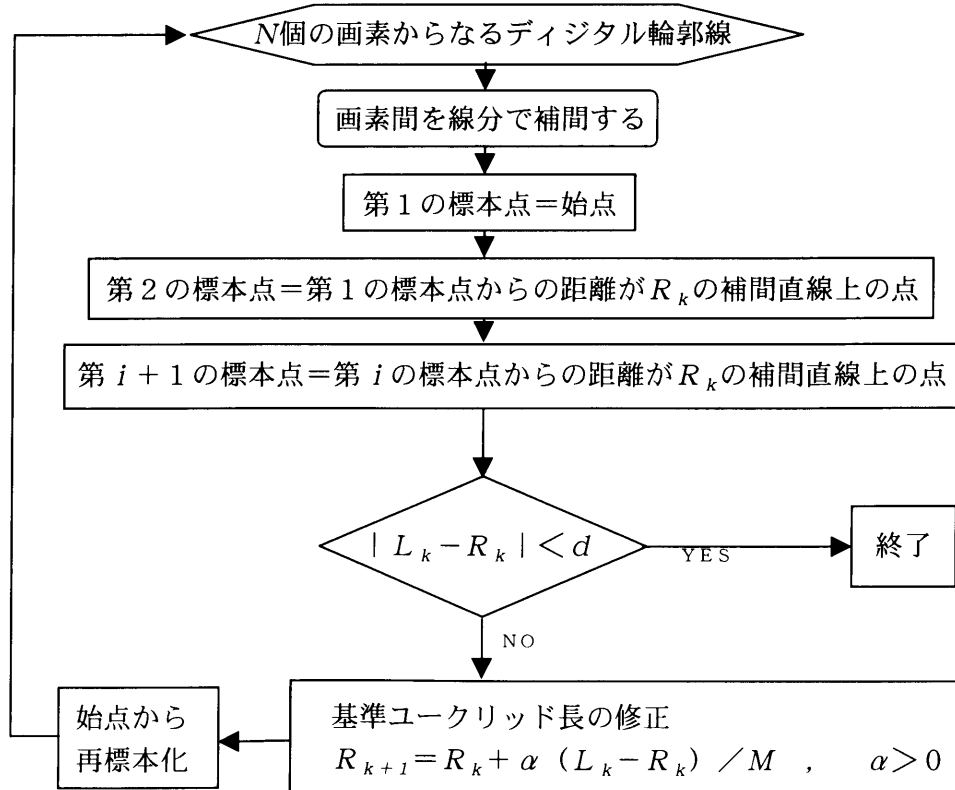


図2.4.5 直線距離で等間隔な再標本化のフローチャート

2. 4. 3 EPACによるデジタル閉曲線の記述

図形をその輪郭線に着目して不変な処理を行う場合、その後の認識処理が円滑に行われるように、デジタル曲線の効率よい正規化と記述法が必要であることは先に述べた。ここでは、正規化および曲線記述過程における実用的なデジタル曲線記述法として、等辺多角形近似曲率関数（Equilateral Polygon Approximated Curvature function：EPAC）[9]について述べ、その有効性を指摘する。

EPACは以下の3つの要求を十分満足するように、白井の曲率関数を改定したものである。第1に、画像（図形）の移動、拡大・縮小、回転に対してその記述が不変であること、第2に、曲線の記述から元の曲線が再生可能であること、第3に、デジタル化誤差の影響を受けにくいものであること、等である。

まず、デジタル曲線からEPACを求めるまでの操作について説明する。

EPACは、デジタル曲線を等辺多角形に近似して、次にこれによって得られる多角形の頂点における曲率を計算するという2段階の操作によって求められる。図 2.4.6 はEPACの原理を示している。多角形（折れ線）に対する曲率関数の抽出は2. 2. 2で述べたので、ここでは主に等辺多角形近似処理を説明する。

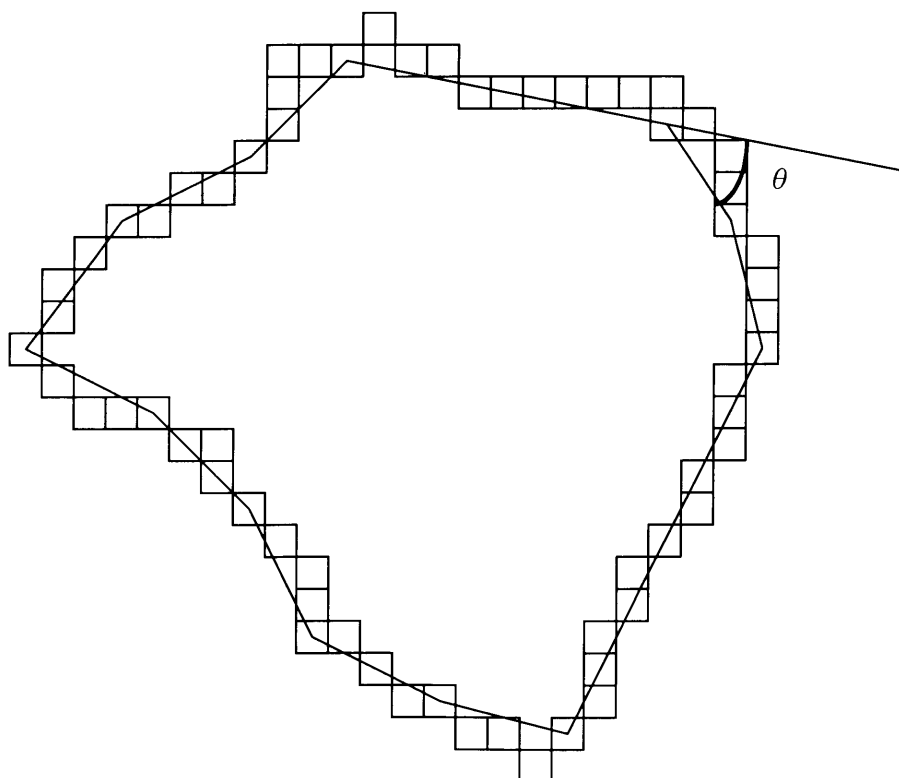


図 2. 4. 6 EPAC

等辺多角形近似はE P A C抽出における骨子となる処理であり、E P A Cの性質を特徴づけるものである。この処理では、多角形における各辺のユークリッド長がすべて等しくなるように標本化している。この操作には、先に述べた直線距離で等間隔な再標本化アルゴリズムを採用している。この場合、標本点数は2のべき乗としてある。その理由は、E P A Cでは後にフーリエ変換を使用することを前提にしているからである。

すなわち、E P A Cの等辺多角形近似は以下の3つの規則のもとで実行される。

- ①頂点の数は2のべき乗となる。
- ②頂点はデジタル曲線を構成している画素の中から選ぶ。
- ③各辺のユークリッド長がどれも等しくなるようにする（各辺の長さの最大誤差が1画素をこえない）。

この中で頂点の数、すなわち標本点数の設定には次の2点を考慮する。1つは、サンプル数は閉曲線の形状を損なわない程度の数が必要であるということ、もう1つは、標本点数によって曲率関数の性質が変わってしまうということである。

前者は、再標本化の作業の時点で当然考慮すべきことである。一方後者は、次のように説明できる。標本点数を設定することによって得られる多角形の一边の長さは一意に決まる。また、曲率関数のステップ数は曲率関数の性質を決定付ける重要な数値である。多角形における一边の長さはすなわちこのステップ数に他ならないので、E P A C（曲率関数）の性質は標本点数の設定に大きく依存することになる。

ここで、サンプル数が大きすぎると、E P A Cは標本化誤差の影響を受けやすい。逆に標本点数が少なすぎると局所的な情報に欠け、形状の特徴自体を失わせてしまうことになる。

先の3つの規則によって、デジタル閉曲線は等辺多角形に近似されるのであるが、この規則では1つの閉曲線に対して画素の選び方によってはいくつかのバリエーションの等辺多角形が存在する。しかし標本点数が十分大きい数であるならば、これらはいずれも視覚的に等価とみてよく、これらの多角形から求められるE P A Cの性質も等価であるとみなせる。

以上、E P A Cによる等辺多角形近似について述べてきた。輪郭線はE P A Cにより等辺多角形近似された後、P型記述子が求められる。このP型記述子はE P A Cの3つの規則から分かるように、図形の移動、回転、拡大・縮小に対して不変な値をとる。更に、等辺多角形近似において始点の位置が移動することによる影響を

受けないようにするために、P型記述子をフーリエ変換してそのパワースペクトルを求め、これを輪郭線の特徴とする。このように、EPACにより求められたP型記述子のパワースペクトルを特徴とすることは、第1に各辺のユークリッド長が等しいためにデジタル化による誤差が発生しにくく、第2にP型記述子を用いているために図形の移動、回転、拡大・縮小に対して不変であり、第3にパワースペクトルを求めているために等辺多角形の始点の位置が移動することによる影響を受けない、という多くの長所を有している。従って、本研究における画像識別で用いる輪郭線情報として、EPACから求められたP型記述子のパワースペクトルを選択する。

2. 5 閉曲線の前処理とELPA

すでに述べたように、与えられた閉曲線をその形状を損なわない範囲であらかじめ等辺多角形近似しておく、情報の圧縮にもなり、また後の認識などの処理が容易になることが多い。しかしながら、適当な標本点数を定めることは簡単ではなく、比較的少ない標本点数で等辺多角形近似を行う場合、式(2.4.1)による繰り返しアルゴリズムが収束しないなど、いろいろと不都合が生じることがある。

ここでは、式(2.4.1)の繰り返しアルゴリズムの収束性が保証されるような閉曲線の性質を説明する。更に処理の対象となる閉曲線に前処理を行い、その図形に繰り返しアルゴリズムを適用して等辺多角形近似する方法を述べる。

ここで、閉曲線Cが与えられているとする。C上にM個の点を取り、隣り合う点の間のユークリッド距離が等しいとき、これらM個の点を直線で結んだ図形 $\tilde{C}$ をCの等辺多角形近似 (EquiLateral Polygon Approximation : ELPA) と呼ぶ [10][11][12]。

2. 5. 1 曲線の単調性

図 2.5.1 に示すように、閉曲線を $x(s)$ ($0 \leq s \leq L$) とする。

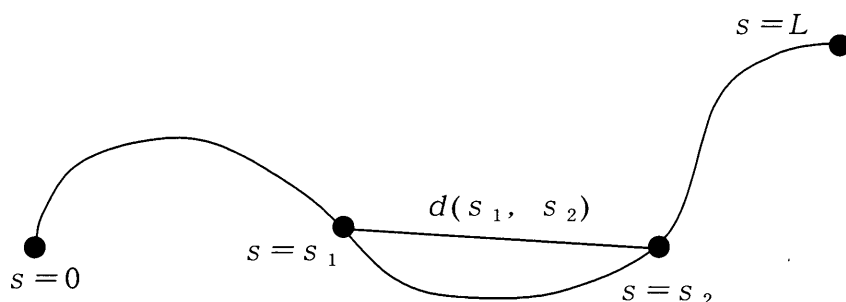


図 2. 5. 1 閉曲線 $x(s)$

$s = s_1$ における点と $s = s_2$ における点とのユークリッド距離を $d(s_1, s_2)$ としたときの曲線の単調性を以下のように定義する。

[定義 2. 5. 1]

$$0 \leq s_1 < s_2 \leq L \quad (2. 5. 1)$$

ならば

$$d(0, s_1) \leq d(0, s_2) \quad (2. 5. 2)$$

であるとき、曲線 $x(s)$ は s 単調であるという。

$s = 0$ の点と $s = L$ の点とを含む直線を x 軸とし、 $s = 0$ の点を原点とすると、曲線を

$$y = f(x), \quad x \geq 0 \quad (2. 5. 3)$$

で表す。このとき、次の定理が成り立つ。

[定理 2. 5. 1]

曲線 $y = f(x)$ が s 単調であるための必要十分条件は

$$f'(x)f(x) \geq -x \quad (2. 5. 4)$$

である。但し $f'(x)$ は $f(x)$ の微分である。

(証明)

曲線上の任意の 2 点を $A(x_1, f(x_1))$ 、 $B(x_2, f(x_2))$ とする。但し、 $x_1 \leq x_2$ 、原点と A の間の距離を D_1 、原点と B の間の距離を D_2 とする。ここで

$D_2 \geq D_1$ ならば

$$x_2^2 + f(x_2)^2 \geq x_1^2 + f(x_1)^2$$

だから

$$(f(x_2) - f(x_1))(f(x_2) + f(x_1)) \geq -(x_2 - x_1)(x_2 + x_1)$$

$x_2 - x_1 \geq 0$ だから

$x_2 \rightarrow x_1$ とすると

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1}(f(x_2) + f(x_1)) \geq -(x_2 + x_1)$$

$$2f'(x_1)f(x_1) \geq -2x_1$$

となる。つまり

$$f'(x)f(x) \geq -x$$

を得る。

逆に、式(2.5.4)が成立しているとする。原点と点 $(x, f(x))$ との間の距離を D_1 、原点と点 $(x + dx, f(x + dx))$ との間の距離を D_2 とすると

$$\begin{aligned} D_2^2 - D_1^2 &= \{(x + dx)^2 + f(x + dx)^2\} - \{x^2 + f(x)^2\} \\ &= 2\{f'(x)f(x) + x\}dx \end{aligned}$$

となり、更に式(2.5.4)を用いると

$$D_2^2 - D_1^2 \geq -2\{-x + x\}dx = 0$$

となる。よって $D_2 \geq D_1$ がいえる。

[系 2. 5. 1]

$y = f(x)$ が s 単調ならば $f(x)$ は 1 価関数である。

<例 2. 5. 1>

中心角が π 以下の円弧は s 単調である。

[定義 2. 5. 2]

閉曲線 $x(s)$ 上の任意の点からはじめて、閉曲線に沿った長さ L でこの閉曲線を分割したとき、各部分が s 単調ならばその閉曲線は $L - s$ 単調であるという。

<例 2. 5. 2>

半径 r の円は $\pi r - s$ 単調である。

[系 2. 5. 2]

$L = ds$ とすれば、全ての閉曲線は $ds - s$ 単調である。

[定理 2. 5. 2]

閉曲線 C が $L - s$ 単調ならば、一辺の長さが l ($< L$) である E L P A、 $\tilde{C}$ を得るアルゴリズムである式(2.4.1)は収束する。

2. 5. 2 閉曲線の前処理と E L P A

次に、任意に与えられた閉曲線の等辺多角形近似 (E L P A) を行う。

大域処理としての E L P A の辺の数 (標本点の数) は、もとの図形の形状を大きく損なわない範囲で少ないほうがよい。しかしながら、与えられた図形に対して辺の長さがある程度大きいと、多角形近似アルゴリズム式(2.4.1)の収束性が悪い。そこでその収束性をよくするために [前処理] を行う。これは閉曲線の局所処理に相当する。この前処理にはやはり式(2.4.1)の繰り返しアルゴリズムと 2. 3. 4 で述べた P 型記述子を利用する。

前処理の考え方は以下の通りである。閉曲線 C が与えられると、辺の数 M_1 を十分大きくして式(2.4.1)を適用する。これは [系 2] によって収束する。この E L P A 曲線を C_1 とする。 C_1 から P 型記述子 D を求める。 D にたいして C_1 の形状を大きく損なわない範囲で高周波成分をカットし、これを逆変換した曲線を C_2 とする。 C_2 は C_1 に対して単調性の高い曲線になっている。更に、 C_2 に式(2.4.1)を適用して $\tilde{C}$ を得る。 C から C_2 を得る過程が最終的な等辺多角形近似図形 $\tilde{C}$ を得るための前処理となる。つまりこの処理は次のような C から $\tilde{C}$ への写像の過程である。

$$C \rightarrow C_1 \rightarrow D \rightarrow C_2 \rightarrow \tilde{C}$$

これを図 2.5.2 にフローチャートで示す。

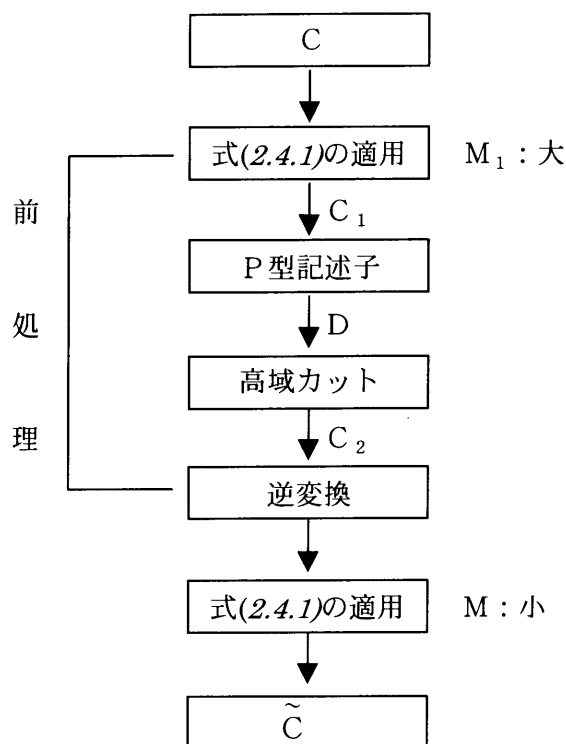


図2. 5. 2 EPACへのアルゴリズム

上の処理において M_1 、 M および高域カットの場合の遮断周波数が処理のパラメータであり、処理の目的、対象となる閉曲線の性質などによって適宜決められるべきものである。

2. 6 まとめ

本章では、画像の性質を表す重要な情報の1つである輪郭線情報について述べてきた。最初に、曲線記述方法として偏角関数による方法と曲率関数による方法について説明した。次に、これらの曲線記述法から輪郭線の特徴を抽出するための手段としてフーリエ記述子を取り挙げ、G型記述子、Z型記述子、P型記述子について述べた。更に輪郭線情報を抽出する総合的な手法としてEPACによる方法について説明した。最後に、等辺多角形近似アルゴリズムの収束性を向上させる方法としてELPA法とその前処理について説明した。これらに関する研究の結果から、E

PACにより求められるP型記述子のパワースペクトルが画像の識別に対して最適の特徴であると判断した。この特徴は、第1に各辺のユークリッド長が等しいためにデジタル化による誤差が発生しにくく、第2にP型記述子を用いているために図形の移動、回転、拡大・縮小に対して不変であり、第3にパワースペクトルを求めているために等辺多角形の始点の位置が移動することによる影響を受けない、という優れた長所を有している。これらの長所は、識別可能な画像の種類が極めて多いことを示している。従って、本研究で行う識別実験に用いる輪郭線情報としてEPACから求められるP型記述子のパワースペクトルを選択する。

第 3 章

ロバストエネルギー最小化による輪郭線 保存平滑化

3. 1 まえがき

前章では、画像の性質を表す重要な性質のひとつである輪郭線情報について論じ、最適な特徴としてE P A Cにより求められるP型記述子のパワースペクトルを選択した。この特徴は、第6章で行う画像識別実験で用いられる。

ここで、現実の世界では識別すべき画像にノイズが混じっていることが多い。ノイズが混じっていると、等辺多角形近似アルゴリズムの収束性が悪化したり、正解率が低下することが多い。前者については、本論文ではE L P A法という方法を用いてアルゴリズムの収束性を確保する方法を示している。後者についてもE L P A法によりノイズ成分を低減させることができるが、逆に輪郭線の性質を表す多くの高周波成分が失われて正解率が低くなることが多い。このため輪郭線の特徴であるエッジを保存しつつノイズを除去（平滑化）する方法が要求される。

エッジからのノイズ除去方法については現在までにいくつかの手法が提案されているが、除去できるノイズの性質が制限されていることやエッジの保存能力が低い等、一般化には難がある。例えば、エッジを保存して、かつインパルスノイズも取り除くことができるフィルタとしてメディアンフィルタが有名であるが、ガウスノイズの平滑化能力が低いので種々の改良が加えられている[15]。ここで、メディアンフィルタは平均フィルタをロバスト化したものである。一方、別の流れとして正

則化フィルタをロバスト化する研究がある。正則化フィルタはガウスフィルタとはほとんど同じ特性をもつ線形フィルタであるが、インパルスノイズの除去やエッジの保存に難がある[16]。そこで正則化にランププロセスを導入し、非線形化してエッジの保存性を高める方法が提案されており[17]、更に同様な 0, 1 変数の追加によりインパルスノイズの除去も行えるように拡張されている[18]。このような拡張は正則化の基礎となる最小 2 乗法をロバスト推定の一種である M 推定に拡張したものと等価である[19]。しかしこの方法にはパラメータの設定が難しいという難点がある。

そこで本章では、M 推定に含まれるしきい値のようなパラメータを含まない方法として L M e d S 推定[20]に基づく平滑化法を提案する[38]。最初に、従来のノイズ除去手法として、最小 2 乗法に基づく方法と M 推定に基づく方法について述べる。次に、今回提案する L M e d S 推定に基づく平滑化方法について詳述する。最後に具体的な例として簡単なデータ波形を用いてのエッジ平滑化実験をも行い、L M e d S 推定に基づく方法の有効性を実証する。

3. 2 エネルギー最小化による平滑化

ここで、簡単のため 1 次元データについて記述する。第 i 番目のデータを d_i 、それからの復元値を f_i とする。ここでは f_i をあるエネルギーの最小解として求める方法を与え、エネルギーとして、まず最初に最も基本的な最小 2 乗法に基づくもの、次に M 推定に基づくもの、最後に L M e d S 推定に基づくものについて示す。

3. 2. 1 最小 2 乗エネルギー

まず最初に、 f_i を次の式から求める方法を考える。

$$\min \alpha \sum_i (f_i - d_i)^2 + \beta \sum_i (f_i - f_{i+1})^2 \quad (3. 2. 1)$$

このエネルギーは下に凸であるから最適解はただ 1 個であり、次の連立 1 次方程式

$$\alpha(f_i - d_i) + \beta(f_i - f_{i-1} - f_{i+1}) = 0 \quad (3. 2. 2)$$

から解は得られる。

これは最も単純な標準正則化法である。

3. 2. 2 M推定エネルギー

式(3.2.1)の2乗エネルギーは変位が大きくなるほど復元力が強くなる。従ってインパルスノイズ付近では復元値が大きくひずみ、またエッジもなまる。これを改善する1つの方法は2乗関数を途中で飽和させて式(3.2.1)を

$$\min \alpha \sum_i \sigma_\alpha(f_i - d_i) + \beta \sum_i \sigma_\beta(f_i - f_{i+1}) \quad (3. 2. 3)$$

と変えるものである。特に

$$\sigma_p(x) = \begin{cases} x^2 & (|x| < t_p) \\ t_p^2 & (|x| \geq t_p) \end{cases} \quad (3. 2. 4)$$

のとき、ロバスト推定の1種類であるM推定、その中でも Hinich-Talwar 推定に対応する[39]。ここで t_p は正定数である。エネルギーの第1項を σ_α にする効果は、 t_α よりも大きいインパルスノイズを取り除けることである。また第2項を σ_β にすることによりジャンプ幅が t_β よりも大きいエッジが保存される。式(3.2.3)のエネルギーは、式(3.2.1)のエネルギーにラインプロセスと呼ばれる0, 1変数を付け加えて、その変数を消去しても得られる。なおエネルギーの第1項に付加される0, 1変数にはノイズプロセスやサンプリングプロセスなどいろいろな呼び方がある。ここで、式(3.2.3)のエネルギーは非凸なので多くの局所最適解をもち、良い解を得るのは容易ではないが、1つの方法としてアニーリングがある。例えば

$$\xi_p(x) = -\frac{1}{g} \ln \left[1 + e^{g(t_p^2 - x^2)} \right] + t_p^2 \quad (3. 2. 5)$$

という関数は $g \searrow 0$ で x^2 と等価となり、 $g \rightarrow \infty$ で式(3.2.4)の σ_p になる。そこで式(3.2.3)の σ_α と σ_β とを式(3.2.5)の ξ_α と ξ_β とに置き換えた式の最適解を $g \searrow 0$ から $g \rightarrow \infty$ まで追跡する。この最適解は

$$\alpha \xi'_\alpha(f_i - d_i) + \beta [\xi'_\beta(f_i - f_{i-1}) + \xi'_\beta(f_i - f_{i+1})] = 0 \quad (3. 2. 6)$$

の解として得られる。 $g \searrow 0$ のとき、この解は式(3.2.1)の解すなわち最小2乗解であり、 $g \rightarrow \infty$ で式(3.2.3)の解、すなわちM推定の解の1つに収束する。このような解追跡法は決定論的アニーリングと呼ばれ、比較的良い解が得られることが経験的

に知られている。

3. 2. 3 LMedS推定エネルギー

M推定による平滑化では t_α や t_β やなどのパラメータを適切な値に設定すればインパルスノイズも取り除くことができ、またエッジも保存される。しかしこれらの適切な値はインパルスノイズやエッジの高さに依存し、また α や β の値によっても変わるので、一般的に設定するのは困難である。そこで、ロバスト推定の別の方法である LMedS (Least Median of Squares) 推定に基づくエネルギーを示す。M推定が最小 2 乗法の 2 乗を飽和関数に変えたものであったのに対し、最小 2 乗平均の平均を中央値 (メディアン) に変えたものが LMedS 推定である。そこでまず式 (3.2.1) をそれと等価な次の式に変形する。

$$\min \alpha \sum_i \text{mean}\{(f_j - d_j)^2 \mid j = i-1, i, i+1\} + \beta \sum_i \text{mean}\{(f_j - f_{j+1})^2 \mid j = i-1, i, i+1\} \quad (3. 2. 7)$$

ここで mean は平均すなわち単純な飽和を表す。なお簡単のためウィンドウ幅は 3 とした。LMedS 推定ではこれを

$$\min \alpha \sum_i p_i + \beta \sum_i q_i, \quad (3. 2. 8)$$

$$p_i = \text{med}\{(f_j - d_j)^2 \mid j = i-1, i, i+1\} \quad (3. 2. 9)$$

$$q_i = \text{med}\{(f_j - f_{j+1})^2 \mid j = i-1, i, i+1\} \quad (3. 2. 10)$$

と変更する。ここで med は中央値を表す。式 (3.2.8) のエネルギーは、各ウィンドウ内の 3 つの変位のうち 2 つが小さくなることを要求し、残る 1 つはいくら大きくてもかまわないということであるから、第 1 項によりインパルスノイズが平滑化され、第 2 項によりエッジが保存される。式 (3.2.8) もまた非凸な最適化問題であり解くのは困難である。そこで、ここでもアニーリングを用いることにする。例えば

$$\sum_{j=i-1}^{i+1} \phi(p_i - (f_j - d_j)^2) = 0 \quad (3. 2. 11)$$

$$\phi(x) = \tanh(gx) \quad (3. 2. 12)$$

の解 P_i は $g \sim 0$ のとき式 (3.2.7) の mean すなわち平均となり $g \rightarrow \infty$ のとき式 (3.2.9)

の med すなわちメディアンとなる [40]。そこで g を 0 から徐々に増加させて式 (3.2.8) をこのようにファジー化した解を追跡することにする。式 (3.2.8) をファジー化した解は

$$\alpha \sum_{j=i-1}^{i+1} \frac{\partial p_j}{\partial f_i} + \beta \sum_{j=i-2}^{i+1} \frac{\partial q_j}{\partial f_i} = 0 \quad (3.2.13)$$

の解として得られる。式 (3.2.13) の中の偏微分は式 (3.2.11) や q_i に関する同様な式を f_i で微分することにより

$$\frac{\partial p_{i-1}}{\partial f_i} = \frac{2\phi'(p_{i-1,i})(f_i - d_i)}{\sum_{j=i-2}^i \phi'(p_{i-1,j})} \quad (3.2.14)$$

$$\frac{\partial p_i}{\partial f_i} = \frac{2\phi'(p_{i,i})(f_i - d_i)}{\sum_{j=i-1}^{i+1} \phi'(p_{i,j})} \quad (3.2.15)$$

$$\frac{\partial p_{i+1}}{\partial f_i} = \frac{2\phi'(p_{i+1,i})(f_i - d_i)}{\sum_{j=i}^{i+2} \phi'(p_{i+1,j})} \quad (3.2.16)$$

$$\frac{\partial q_{i-2}}{\partial f_i} = \frac{2\phi'(q_{i-2,i-1})(f_i - f_{i-1})}{\sum_{j=i-3}^{i-1} \phi'(q_{i-2,j})} \quad (3.2.17)$$

$$\frac{\partial q_{i-1}}{\partial f_i} = \frac{2[\phi'(q_{i-1,i-1})(f_i - f_{i-1}) + \phi'(q_{i-1,i})(f_i - f_{i+1})]}{\sum_{j=i-2}^i \phi'(q_{i-1,j})} \quad (3.2.18)$$

$$\frac{\partial q_i}{\partial f_i} = \frac{2[\phi'(q_{i,i-1})(f_i - f_{i-1}) + \phi'(q_{i,i})(f_i - f_{i+1})]}{\sum_{j=i-1}^{i+1} \phi'(q_{i,j})} \quad (3.2.19)$$

$$\frac{\partial q_{i+1}}{\partial f_i} = \frac{2\phi'(q_{i+1,i})(f_i - f_{i+1})}{\sum_{j=i}^{i+2} \phi'(q_{i+1,j})} \quad (3.2.20)$$

と求まる。ここで

$$p_{i,j} = p_i - (f_j - d_j)^2 \quad (3.2.21)$$

$$q_{i,j} = q_i - (f_j - f_{j+1})^2 \quad (3.2.22)$$

である。これらを式(3.2.13)に代入して f_i に関して形式的に解くと

$$f_i = \frac{\alpha_1 d_i + \beta(r_2 f_{i-1} + r_3 f_{i+1})}{\alpha_1 + \beta(r_2 + r_3)} \quad (3.2.23)$$

となる。ここで

$$r_1 = \frac{\sum_{k=i-1}^{i+1} \phi'(p_{k,i})}{\sum_{j=k-1}^{k+1} \phi'(p_{k,j})} \quad (3.2.24)$$

$$r_2 = \frac{\sum_{k=i-2}^i \phi'(q_{k,i-1})}{\sum_{j=k-1}^{k+1} \phi'(q_{k,j})} \quad (3.2.25)$$

$$r_3 = \frac{\sum_{k=i-1}^{i+1} \phi'(p_{k,i})}{\sum_{j=k-1}^{k+1} \phi'(q_{k,j})} \quad (3.2.26)$$

である。各 g の値で式(3.2.23)の反復計算により解 f_i を求めたその解を $g=0$ から十分大きな g まで追跡すれば LMedS 推定解が得られる。式(3.2.23)の反復の収束性の理論的な保証は難しいが、過去に種々行った実験ではすべて収束した[40]。

3.3 ノイズあり波形に関する平滑化実験

3.3.1 平滑化実験の例

簡単な例として図 3.3.1 のような階段状の波形にガウスノイズと 5 個のインパルスノイズを加えた図 3.3.2 のデータを d_i として以上の 3 種類の方法で平滑化した。エネルギーの係数 α は 1 とした。以下復元値 f_i の図 3.3.1 の値からのずれを誤差と呼ぶ。最小 2 乗法において最も誤差が小さい結果を図 3.3.3 に示す。 $\beta=0.8$ である。また図 3.3.4 に M 推定の誤差が最も小さい結果を示す。 $\beta=1.2$ である。M

推定では β が小さすぎる（0.5 以下）と図 3.3.2 とほとんど同じ復元値となりノイズがとれず、逆に大きすぎる（1.5 以上）と図 3.3.5（ $\beta=2.0$ ）のようにエッジもなまってしまう。従って誤差が小さい解を得るには β を適切な値に設定しなければならない。加えて t_α や t_β も適切な設定を要し、これらのパラメータの設定は困難である。図 3.3.4 と図 3.3.5 では $t_\alpha = t_\beta = 0.5$ である。LMedS推定の結果を図 3.3.6 に示す。 $\beta=1.0$ である。LMedS推定では β の値をこの値より大きくしても図 3.3.6 の波形はほとんど変わらなかった。従ってLMedS推定では β を十分大きく設定しておけば良好な解が得られるものと思われ、 β 以外にパラメータは含まないのでM推定よりも使いやすい。ちなみにメディアンフィルタの結果を図 3.3.7 に示す。それぞれの平滑化法の最良の結果での誤差は、最小2乗法、M推定、LMedS推定、メディアンフィルタでそれぞれ 10.59、1.74、2.18、2.76 となり、M推定が最も平滑化能力が高かったがLMedS推定もM推定に比べてそれほど誤差が大きくない。しかし計算時間はそれぞれ 0.12、8.50、24.20、0.10 でありLMedS推定が最も遅い。

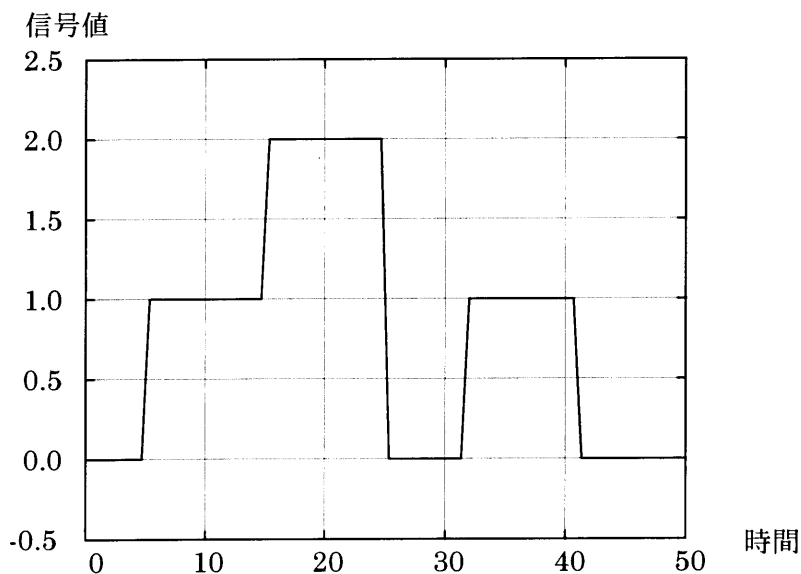


図 3. 3. 1 データ波形

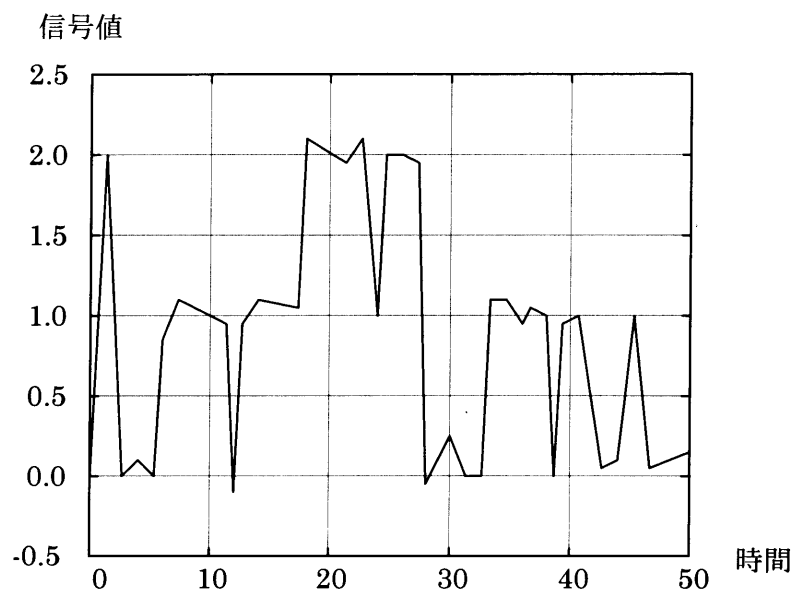


図 3. 3. 2 ノイズが乗ったデータ

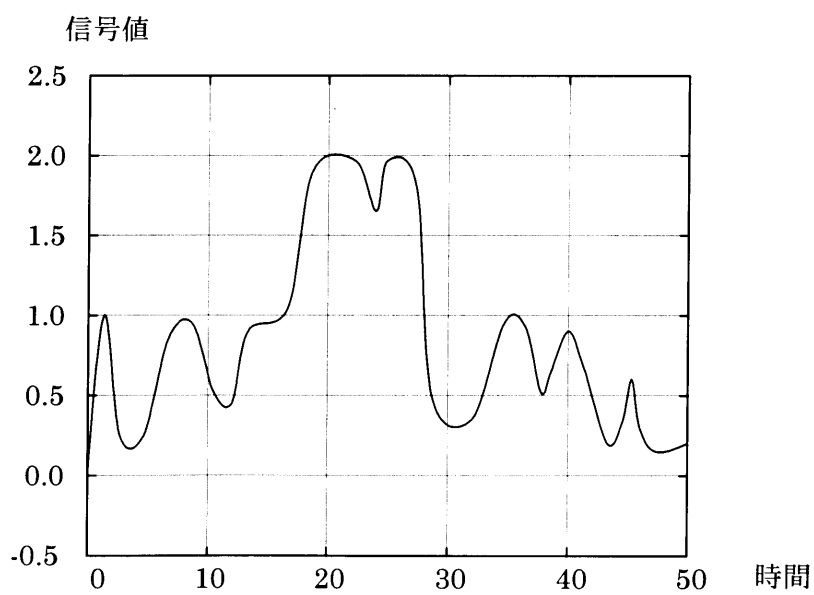


図 3. 3. 3 最小 2 乗法による平滑化

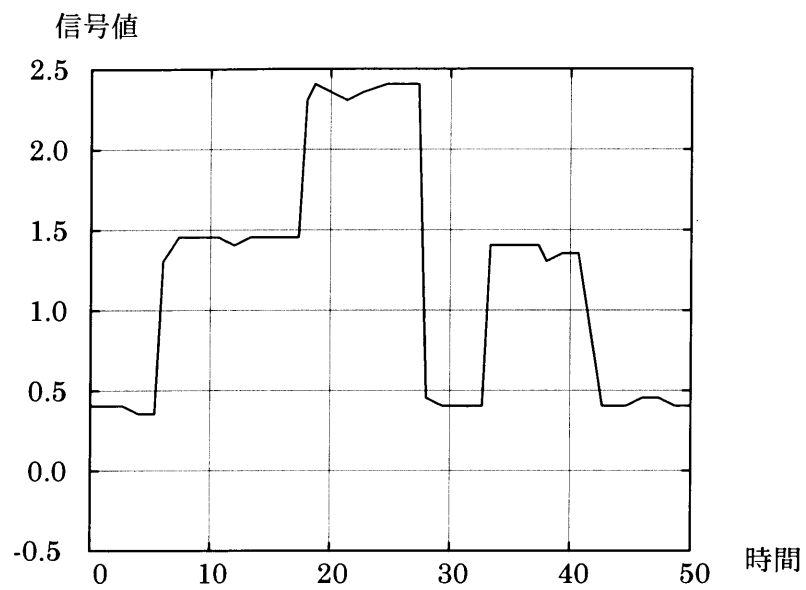


図 3. 3. 4 M推定による平滑化

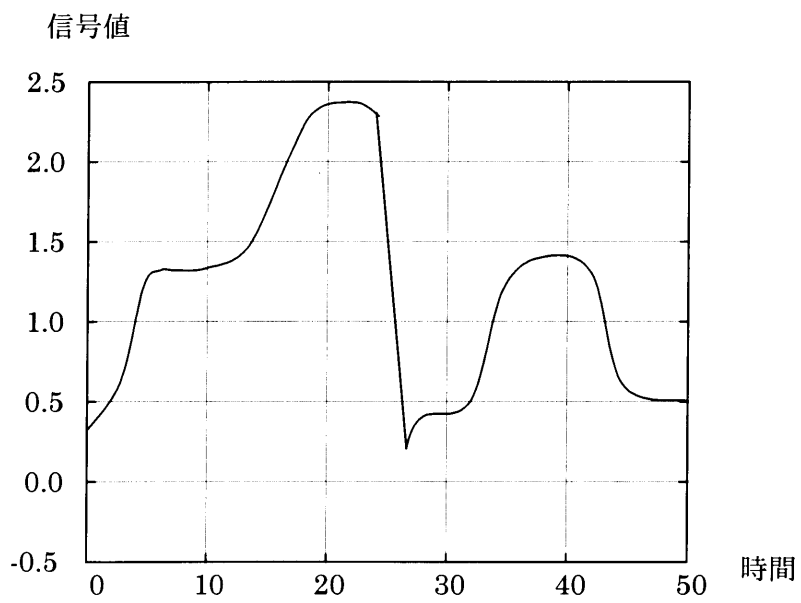


図 3. 3. 5 M推定による平滑化 (β 過大)

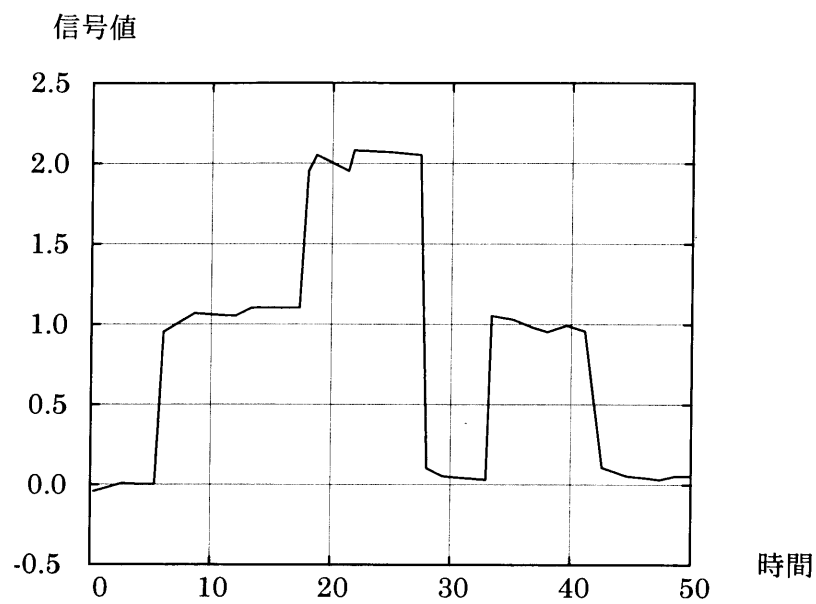


図 3. 3. 6 LMEDS 推定による平滑化

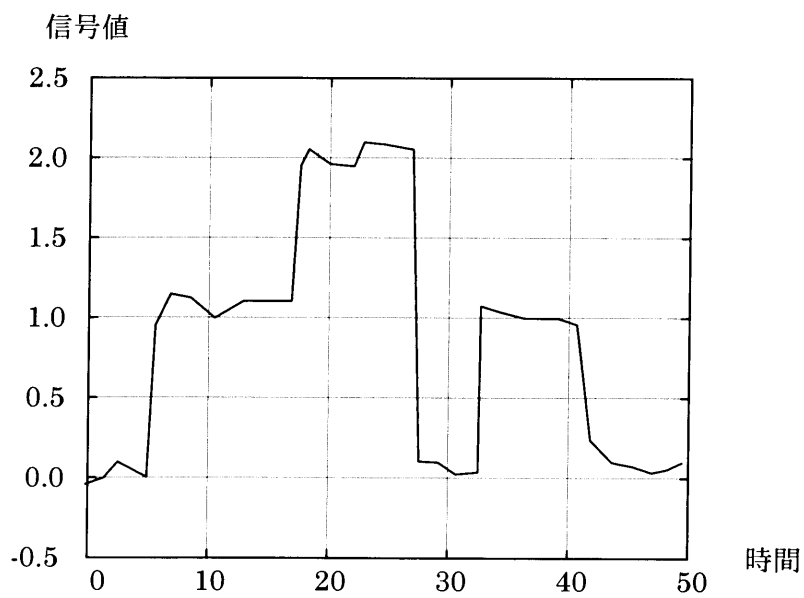


図 3. 3. 7 メディアンフィルタによる平滑化

3. 4 まとめ

本章では、ノイズの混入した輪郭線画像の識別において良好な正解率を得るため、エッジ情報を保存しつつノイズを除去する方法として、LMedS推定に基づく方法を提案し、簡単なデータ波形を用いたエッジ平滑化実験を行い、この方法の有効性を評価した。この結果、本方法はM推定による方法には及ばないが、ガウスノイズとインパルスノイズを取り除いてエッジを良好に復元する能力があることがわかった。ただし、計算量が非常に多いため、処理に要する時間が他の方法と比較してきわめて大きかった。しかしながら、エッジの復元能力が比較的高く、かつM推定のように多くのパラメータを適切に設定する必要がほとんどないことはきわめて重要であり、このことは種々のエッジやエッジを含むノイズに対して、常にある程度適切なエッジ復元が可能であることを意味する。従って、他の方法と比較してLMedS推定法が最も実用性が高いと考える。

第4章

種々のテクスチャ特徴による画像解析

4. 1 まえがき

テクスチャ特徴による画像の解析に関して、今日までに数多くの研究が行なわれ、各種の方法が提案されている。それらのテクスチャ解析の方法は、以下の2つに分類される[24]。

①構造レベルのテクスチャ解析

画像の中から直線や点などの、テクスチャを構成する基本的な要素を抽出し、それらの配列規則を特徴として求める方法である。抽出できる要素の種類は、手法によって異なるが非常に少ない。従って、航空写真やリモートセンシング画像における道路や建造物など、ターゲットとして抽出する構成要素が明確な人工のテクスチャには有効であるが、岩石や芝生、布地、木目など、自然に存在するテクスチャにおいて、その構成要素を求め配列規則を知ることは容易でない。

②統計レベルのテクスチャ解析

画素の濃度に着目して、その一様性、方向性、コントラスト変化などの画像の性質を表す統計量を特徴として求める方法である。ターゲットの形状などが明確でない画像の性質を抽出できるので、構造レベルのテクスチャ解析法より一般的である。

本章では、画像のテクスチャ情報を解析する種々の方法について述べる。構造レベルのテクスチャ解析法については、Hough 変換を利用する方法について説明する。統計レベルのテクスチャ解析法については、1 次統計量である濃度ヒストグラム、2 次統計量である自己相関関数、フーリエ・スペクトル、差分統計量、濃度共起行列、高次統計量であるランレンジス行列による方法について述べる。更に、最近テクスチャ解析の新たな視点として注目されているフラクタル情報を用いる方法についても述べる。最後にそれらを実際に用いる際の問題点について述べる。

4. 2 Hough 変換を用いる解析法

P.V.Hough によって考案され[41]、Duda と Hart らによって改善が進められた Hough 変換は[42][43]、ノイズの混入した画像に対して非常に強力な直線、曲線検出方法として近年盛んに研究が進められている。

ここでは、図 4.2.1 のようなある規則をもつ直線で構成された画像を、Hough 変換によって解析して特徴を抽出する方法について述べる[23]。

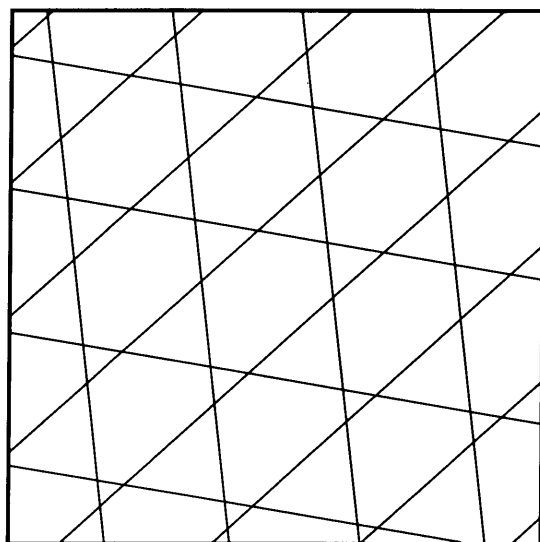


図 4. 2. 1 直線画像

4. 2. 1 Hough 変換による直線検出

Hough 変換の概念は、困難な画像空間における直線検出問題を、より容易なパラメータ空間におけるピーク検出問題に置きかえるということである[44]。これは P.V.Hough らによって提案されてから現在でも変わらない。しかし、直線のパラメータ表現には、Hough の (m, c) 表現や Duda と Hart の (ρ, θ) 表現、R.Warace の (s_1, s_2) 表現[45]など複数の方法が提案されている。これらのなかで現在一般的に扱われている表現方法は、Duda と Hart が提案した (ρ, θ) 表現であり、これは画像を極座標形式で表現する方法である。

なお、本節で述べるテクスチャ解析で用いられる Hough 変換もまた極座標形式で表現されているので、その形式を用いて Hough 変換の手続きを以下に説明する。

2 値画像 $f(x, y)$ 上のある点 (x, y) を通る直線は次のように表される。

$$\rho = x \cos \theta + y \sin \theta \quad (4. 2. 1)$$

ここで図 4.2.2 で示すように、 θ は (x, y) を通る直線の表線の傾きを表し、 ρ は直線と原点との距離を表している。

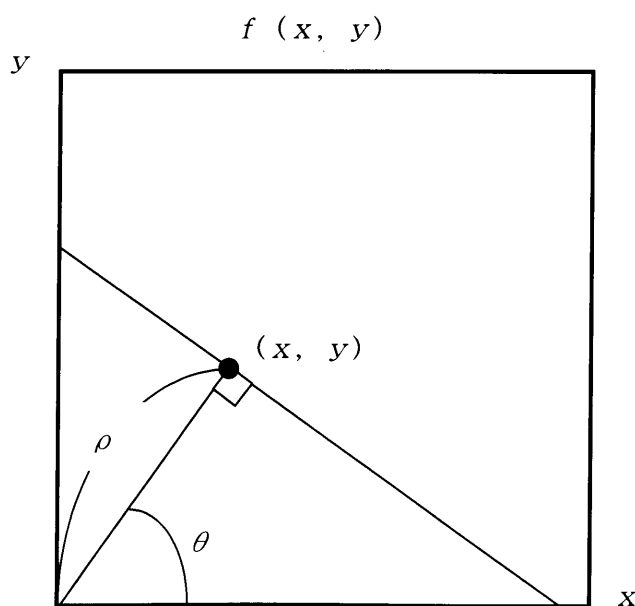


図 4. 2. 2 極座標形式による直線表現

画像空間上の直線のパラメータは一意に定まる。従って式(4.2.1)は画像空間上の直

線（画素の集合）がパラメータ空間上の点に対応することを表している。

また、式(4.2.1)を書き換えると

$$\rho = \sqrt{(x^2 + y^2)} \cdot \sin\{\theta + \tan^{-1}(x/y)\} \quad (4. 2. 2)$$

となる。従って、画像空間上のある点を通る直線の集合はパラメータ空間上では sin 曲線によって表される。このように Hough 変換では、一方の空間のある点がもう一方の空間のある集合に対応する。この関係から、画像空間とパラメータ空間との間での変換、逆変換が可能となる。

次に、画像空間上の異なる 2 つの画素を考える。これらの画素はそれぞれ、パラメータ空間上では異なる 2 本の sin 曲線に対応する。このとき図 4.2.3 で示すように、sin 曲線の交点（共点）の位置によって表されるパラメータは、この 2 点を通る直線（共線）のパラメータとなる。

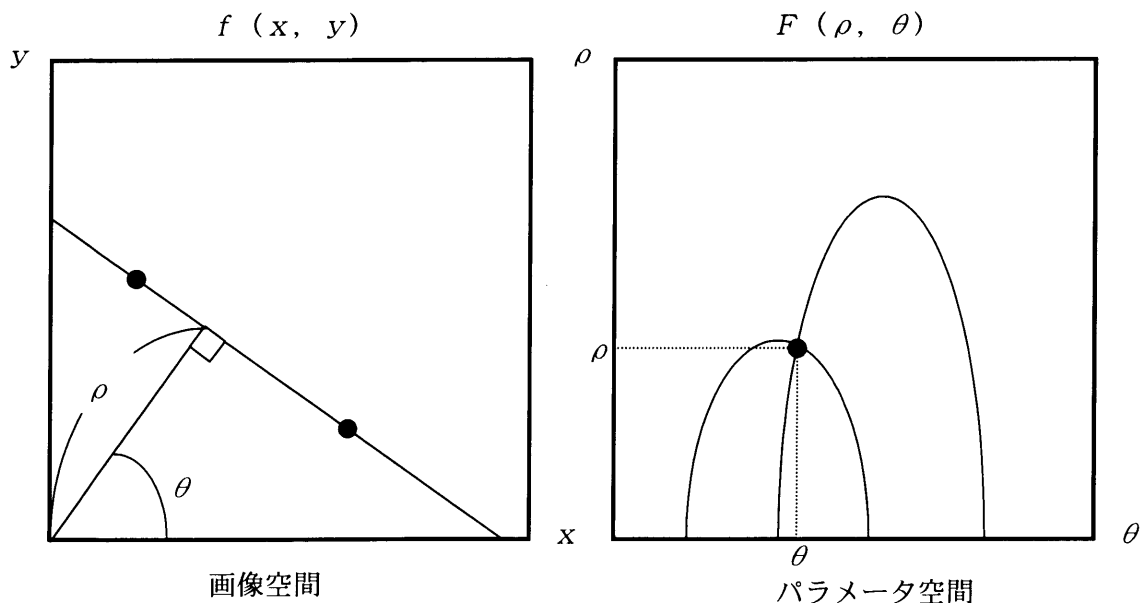


図 4. 2. 3 画像空間とパラメータ空間

従って、線上に含まれる画素の数が多い直線であるほど、パラメータ空間ではその直線のパラメータで表される点で多くの sin 曲線が交差することになる。また、直線が含む画素の数と交差する sin 曲線の数は一一致する。

これらのことから、それぞれの画素に対応する sin 曲線の交点を探し出し、その数をカウントしてピークを求めることによって直線を検出できることがわかる。

このように、Hough 変換の概念は、画像空間における直線検出問題をパラメータ

空間におけるピーク検出問題に置きかえることによって処理を容易にすることである。更に、パラメータ空間の $\sin$ 曲線の交点だけを探すため、ノイズの影響を受けにくい。

現在までに、Hough 変換は処理に要するメモリや高速化の問題について数多くの研究が行なわれているが、近年の半導体技術の発展に従って、もはや処理資源、速度は致命的な問題ではなくなりつつある[46][47][48]。

4. 2. 2 Hough 変換による直線画像の解析

Hough 変換を用いるテクスチャ画像の解析法は、図 4.2.1 のような規則的な直線（分）画像を対象とする。この方法では、次のような手続きで画像の特徴を抽出する。

図 4.2.4 で示すようなある規則性をもつ直線画像において、最初に Hough 変換を用いて直線を検出する。

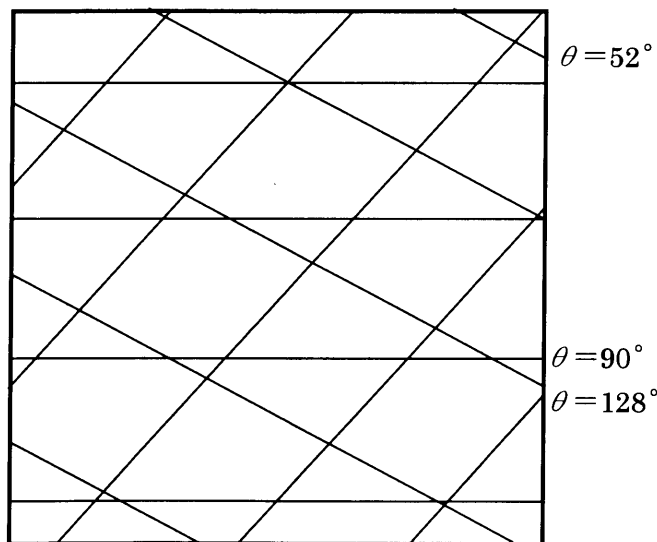


図 4. 2. 4 直線画像

ここで、同じ θ の値を持つ直線群を 1 つのグループとする。このときグループの数を f_1 、直線群の傾きを f_2 とする。図 4.2.4 では直線群は $\theta = 52^\circ$ 、 $\theta = 90^\circ$ 、 $\theta = 128^\circ$ の 3 つのグループに分かれる。従って、

$$f_1 = 3$$

$$f_2 = [52, 90, 128]$$

となる。次に、同じグループに属する直線群のなかでそれぞれ隣あわす 2 本の直線の距離を求めてその平均をとる。距離の平均値を f_3 とする。

これら f_1 、 f_2 、 f_3 をまとめて

$$F = [f_1, f_2, f_3]$$

としたとき、 F は直線画像の特徴となる。 $f_1 = N$ とすると F の次元数は $2N + 1$ である。

F はこのままでも特徴として使えるが、G.Eichmann と T.Kasparis は更に回転や拡大・縮小など幾何学的に変換された画像に対応するために、抽出された特徴 f_2 、 f_3 を正規化している[23]。以下に図 4.2.4 の画像を例とした正規化の方法を説明する。

ここで、図 4.2.4 で表される画像の f_3 を

$$f_3 = [71, 90, 72]$$

とする。

最初に、 f_3 に関してその最大の要素の値を 1 として、すべての要素を正規化する。正規化した f_3 は次のようになる。

$$f_3 = [0.78, 1.00, 0.79]$$

ここで、 f_3 の要素を大きな順に並びかえる。 f_2 の要素の順も f_3 の要素の順に従わせる。すると f_2 、 f_3 は

$$f_2 = [90, 128, 52]$$

$$f_3 = [1.00, 0.79, 0.78]$$

となる。次に、 f_2 のそれぞれの要素の値を最初の要素の値で減算する。このとき引いた後の値が負となる要素には 180 加える。これは直線群の傾きを示す f_2 の要素を 180° を軸として正規化することである。この処理によって f_2 は

$$f_2 = [0, 38, 142]$$

となる。最後に、以下の様に f_2 の要素の値を 180 で除算して正規化する。

$$f_2 = [0.00, 0.21, 0.79]$$

こうして作成された f_2 、 f_3 を新たに正規化された特徴とする。

従って、図 4.2.4 で示される直線画像の特徴 F は

$$F = [f_1, f_2, f_3] = [3, 0.00, 0.21, 0.79, 1.00, 0.79, 0.78]$$

となる。この新たに作成された特徴 F は、回転や拡大・縮小などの幾何学的な変換に対して独立している。

以上に説明した手順によって、直線画像から Hough 変換を利用して特徴を抽出することができる。この方法によって抽出された特徴の特質は、幾何学的な変換に対して独立していることである。この章で説明するテクスチャ解析法によって抽出される特徴の中で、この方法で抽出された特徴のみが幾何学的な変換に対して完全に独立している。また Hough 変換を利用しているのでノイズに対して非常に強力である。

しかし、この方法が対象としているテクスチャ画像は規則性をもった直線(分)画像なので、実際には用途が限られている。

4. 3 濃度ヒストグラムを用いる解析法

濃度ヒストグラムを用いる解析法は、この章の中で紹介する方法の中では、求められる特徴が最も不十分である。しかし、処理に要するメモリや時間は非常に小さく、対象を限定すればある程度のテクスチャを識別することができる[49]。以下に濃度ヒストグラムによるテクスチャ画像の特徴抽出の手順を示す。

全体が 1.0 になるように正規化された濃度ヒストグラム $P(i)$, ($i=0,1,\dots,n-1$) を作成する。そのヒストグラムから次の 2 つの特徴が求められる。

ヒストグラムの平均 μ を

$$\mu = \sum_{i=0}^{n-1} i \cdot P(i)$$

分散 σ^2 を

$$\sigma^2 = \sum_{i=0}^{n-1} (i - \mu)^2 \cdot P(i)$$

とすると、 f_1 、 f_2 を以下のように定める。

f_1 : Skewness

$$S = \frac{\sum (i - \mu)^3 \cdot P(i)}{\sigma^3} \quad (4. 3. 1)$$

f_2 : Kurtosis

$$K = \frac{\sum (i - \mu)^4 \cdot P(i)}{\sigma^4} \quad (4. 3. 2)$$

f_1 は、濃度ヒストグラムの形状が対称な形からどれだけ歪んでいるかの程度を表す特徴であり、 f_2 は、ヒストグラムの分布が平均値のまわりに集中しているか、端のほうへ広がっているかの程度を表す特徴である。

また、濃度の累積ヒストグラム分布を直接用いてもテクスチャを特徴づけられ、テクスチャの類似性の判定として Kolmogorov-Smirnov 検定法が知られている[50]。これは、2つの分布の濃度ヒストグラムを $F_1(x)$ 、 $F_2(x)$ としたとき、適当な定数 ρ に対して

$$\max |F_1(x) - F_2(x)| < \rho \quad (4.3.3)$$

が成り立つときは、2つの分布は同じであるとし、そうでなければその仮説を棄却するものである。

4.4 自己相関関数を用いる解析法

テクスチャ画像の性質を表す2次統計量として、最初に2次元の自己相関関数を示す。

ここでは、次のようにして自己相関関数によるテクスチャ画像の特徴抽出を行う。

あるテクスチャ画像 $f(i, j)$ において、図 4.4.1 に示すようにある点 (x, y) を基準点とする近傍の $N_x \times N_y$ 画素の領域と、これから (m, n) 画素だけずれた領域、つまり点 $(x+m, y+n)$ を基準点とする $N_x \times N_y$ 画素の領域とを対象として、自己相関関数を

$$R(x, y; m, n) = \frac{\sum \sum f(x+i, y+j) f(x+m+i, y+n+j)}{\sqrt{\sum \sum f(x+i, y+j)^2 \sum \sum f(x+m+i, y+n+j)^2}} \quad (4.4.1)$$

により与える。ここで、基準点 (x, y) を固定し、相関をとるもう一方の対象領域を

$$m, n = 0, \pm 1, \pm 2, \dots, \pm N_d$$

の範囲内でずらすことにより、図 4.4.2 に示すようにそれぞれの自己相関関数値を要素とする、 $N_d \times N_d$ の相関関数行列を求めることができる。このとき、相関関数の値が1となる $m=n=0$ の要素を除外する。

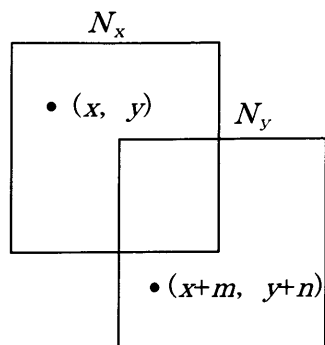


図 4. 4. 1 相関をとる領域

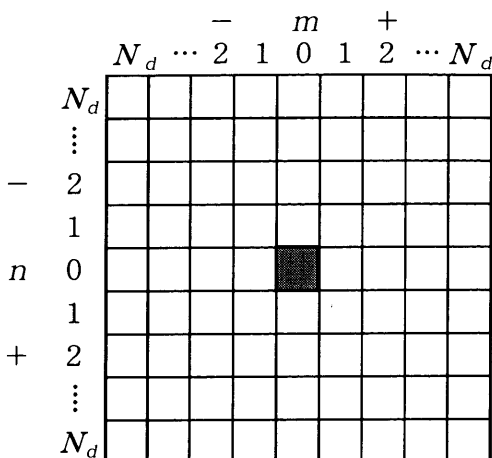


図 4. 4. 2 相関関数行列

更に、作成したテクスチャ画像データの分解能を知るために、前述した自己相関関数において、相関をとるもう一方の基準点をずらすときの画素間隔を表すパラメータを導入し、

$$R(d) = \frac{\sum \sum f(x+i, y+j) f(x+dm+i, y+dn+j)}{\sqrt{\sum \sum f(x+i, y+j)^2 \sum \sum f(x+dm+i, y+dn+j)^2}} \quad (4. 4. 2)$$

として、種々の画素間隔 d に対する自己相関関数を抽出する。

こうして求めた自己相関関数から、例えば、その広がりやテクスチャの粗さや細かさと深く関係し、その周期的な変化などからはテクスチャの周期性や方向性などの性質を知ることができる。しかし、この自己相関関数から、テクスチャの性質を明確に表現する特徴を抽出することは容易ではない。そのため梅田は、抽出した相関関数行列そのものを特徴として、テクスチャ画像の識別を行って高い正解率を得ている[30]。

4. 5 フーリエ・スペクトルを用いる解析法

画像の濃淡に関する情報の1つである空間的な広がりに着目して、空間周波数という概念を用いて、入力画像がどのような空間周波数成分から構成されているかを調べる。ここでは、画像を2次元フーリエ変換によって空間周波数領域に変換して、そのパワー・スペクトルを求めて画像の特徴を抽出する。

4. 2. 1 2次元フーリエ変換

連続画像 $g(x, y)$ に対して

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |g(x, y)| dx dy < \infty \quad (4. 5. 1)$$

とすると、2次元フーリエ変換 $G(u, v)$ は

$$G(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) \cdot \exp\{-j2\pi(ux + vy)\} dx dy \quad (4. 5. 2)$$

となり、逆に $G(u, v)$ から

$$g(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} G(u, v) \cdot \exp\{j2\pi(ux + vy)\} dx dy \quad (4. 5. 3)$$

となり、2次元逆フーリエ変換 $g(x, y)$ が求められる。

ディジタル画像に対しては、 $g(x, y)$ を x, y 方向についてそれぞれ標本化間隔 x_0, y_0 で標本化した $g(m, n)$ 、および $G(u, v)$ を u, v 方向についてそれぞれ $u_0 (= 2\pi/(Mx_0))$ 、 $v_0 (= 2\pi/(Ny_0))$ で標本化した $G(k, l)$ を考えて、

$$\left. \begin{aligned} g(aM+m, bN+n) &= g(m, n) \\ G(aM+k, bN+l) &= G(k, l) \end{aligned} \right\}$$

M, N は横、縦方向の画素数

$$a, b = 0, \pm 1, \pm 2, \dots$$

なる周期性を仮定した上で、2次元離散フーリエ変換 (DFT) 対が定義される。

$$G(k, l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} g(m, n) W_1^{mk} W_2^{nl} \quad (4. 5. 4)$$

$$g(m, n) = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} G(k, l) W_1^{-mk} W_2^{-nl} \quad (4. 5. 5)$$

$$\text{ただし、} W_1 = \exp(-j2\pi/M), \quad W_2 = \exp(-j2\pi/N)$$

4. 5. 2 フーリエ・スペクトル

テクスチャにおける特徴の抽出には、画像をフーリエ変換し、その空間周波数成分の分布から求める方法がある。

画像 $f(x, y)$ のフーリエ変換 $F(u, v)$ のパワースペクトル $P(u, v)$ は、

$$P(u,v) = |F(u,v)|^2 \quad (4.5.6)$$

で定義され、その値は空間周波数(u , v)の強さを表す。 $P(u, v)$ からテクスチャの特徴を計算する方法の 1 つとして、この $P(u, v)$ を極座標形式で $P(r, \theta)$ と表したのちに

$$P(r) = 2 \sum_{\theta=0}^{\pi} P(r, \theta) \quad (4.5.7)$$

$$q(\theta) = \sum_{r=0}^{W/2} P(r, \theta) \quad (4.5.8)$$

を求める。ここで W は $P(u, v)$ の大きさを表している。

更に、 $p(r)$ はパワースペクトル空間において、原点を中心としたドーナツ型の領域内のエネルギーの和を、 $q(\theta)$ は、扇形領域内のエネルギーの和を表している。テクスチャの特徴としては、 $p(r)$ 、 $q(\theta)$ のヒストグラムの位置や大きさ、 $p(r)$ 、 $q(\theta)$ の平均や分散などを用いることが多い。例えば、 $p(r)$ のピークはテクスチャの構成要素の大きさ（テクスチャのきめの大きさ）を表し、 $q(\theta)$ のピークはテクスチャ画素の方向と直角な向きに明確な方向性を持つことを示している。

4. 6 差分統計量を用いる解析法

画像中で、一定の変位 $d = (r, \theta)$ だけ離れた 2 点の濃度の差が k である確率 $P_d(k)$ を求めて、 $P_d(k)$ から次の 4 種類の特徴を計算する。

f_1 : Contrast

$$C = \sum_{k=0}^{n-1} k^2 P_d(k) \quad (4.6.1)$$

f_2 : Angular second moment

$$A = \sum_{k=0}^{n-1} \{k^2 P_d(k)\}^2 \quad (4.6.2)$$

f_3 : Entropy

$$E = - \sum_{k=0}^{n-1} P_d(k) \cdot \log\{P_d(k)\} \quad (4.6.3)$$

f_4 : Mean

$$M = \sum_{k=0}^{n-1} k \cdot P_d(k) \quad (4.6.4)$$

4. 7 濃度共起行列を用いる解析法

R.M.Haralick らによって提案された濃度共起行列 ((gray level) co-occurrence matrix) [51]は代表的なテクスチャ解析法であり、現在も盛んに研究されている [26][28][52][53]。Haralick らは、画像の中の特定の相対的な位置関係にある 2 つの画素の濃度対に関する統計量である 2 次統計量を求める具体的な方法として、濃度共起行列を提案している。更に、その行列を元にして、テクスチャの一様性やコントラスト、方向性などを表すと考えられる 14 種類の特徴を提案し、その抽出方法を示している。また、これらの特徴を用いて顕微鏡写真や航空写真の識別実験を行っている。以下に濃度共起行列と、濃度共起行列から求められる 14 種類の特徴の性質およびその作成方法を示す。

4. 7. 1 濃度共起行列

濃度共起行列 $P(i, j; d, \theta)$ とは、テクスチャ画像 $f(x, y)$ において、図 4.7.1 に示すように、ある特定の相対的な位置関係 (d, θ) にある画素対 (x_1, y_1) と、 (x_2, y_2) においてその濃度対が (i, j) になる、すなわち $f(x_1, y_1) = i$ 、 $f(x_2, y_2) = j$ となる頻度を表したものである。

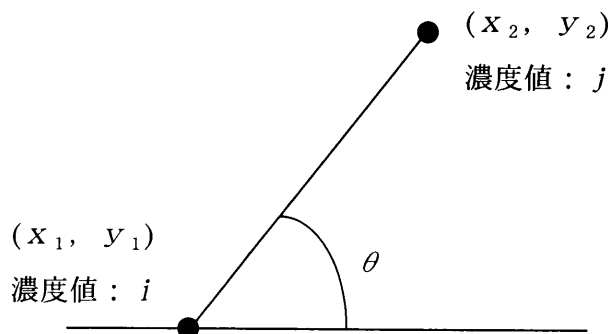


図 4. 7. 1 画素対

ここで、画素対 (x_1, y_1) と (x_2, y_2) との距離 d は

$$d = \max(|x_1 - x_2|, |y_1 - y_2|) \quad (4.7.1)$$

である。

2つの画素を通る直線と水平線とのなす角 θ のレベルは、図 4.7.2 に示されるように、 0° 、 45° 、 90° 、 135° の4レベルである。

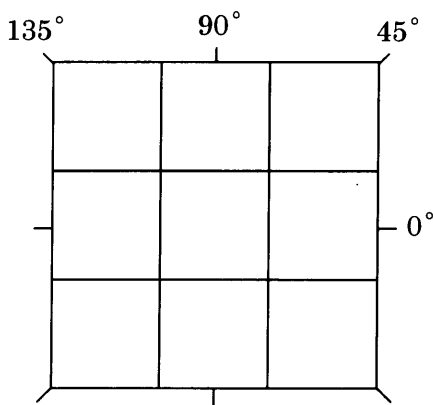


図 4.7.2 θ のレベル

以上に示したパラメータを用いての、濃度共起行列 $R(i, j; d, \theta)$ の作成手順は次の通りである。

ある画素とその画素からみて θ 方向に距離 d だけ離れた画素対とを着目する画素対とする。次に、着目する画素対の濃度値を求め、その濃度が i 、 j であったとすると、濃度共起行列 $R(i, j; d, \theta)$ の対応するようその項に累積していく。このとき、濃度対 i と j は、2つの画素のどちらかがその濃度でもかまわないものとする。濃度共起行列 $R(i, j; d, \theta)$ は、画像の階調数が g であるならば、1つの方向(θ)に対して $g \times g$ の大きさをもつ行列となる。

例として図 4.7.3 に示す2値画像 A に対する濃度共起行列 $R(i, j; d, \theta)$ ($d=1$ 、 $\theta=0^\circ$ 、 45° 、 90° 、 135°) を図 4.7.4 に示す。

図 4.7.4 から分かるように、濃度共起行列 $R(i, j; d, \theta)$ は対称行列となる。

また、各行列の要素の合計値、すなわち画素の結合回数は、例えば画像の大きさが $N_x \times N_y$ で2つの画素間の距離が $d=1$ のとき、

$$\begin{aligned} N_0 &= 2 N_x (N_y - 1) \\ N_{45} &= 2 (N_x - 1) (N_y - 1) \end{aligned}$$

$$N_{90} = 2 N_y(N_x - 1)$$

$$N_{135} = 2(N_x - 1)(N_y - 1)$$

となる。

1	0	1	0
0	1	0	0
1	0	1	1
1	1	1	0

画像 A

4 × 4 画像

2 階調

図 4. 7. 3 2 値画像 A

2	8
8	6

$P(i, j, d, 0^\circ)$

$P(i, j, d, 45^\circ)$

2	9
9	4

$P(i, j, d, 90^\circ)$

$P(i, j, d, 135^\circ)$

6	2
2	8

4	4
4	6

図 4. 7. 4 濃度共起行列の例

テクスチャ解析において、その用途を広げる意味で特徴の正規化は重要な問題である。従って、濃度共起行列の要素を正規化する際には、画素の結合回数が、結合する 2 画素を通る直線の方角(θ)によって変化するということを考慮する必要がある。

以上の手順によって求められた濃度共起行列は、テクスチャ画像の 2 次統計量を表しており、その行列の要素の値の分布状態をもとに、テクスチャの一様性や方向性、コントラストなどの性質を知ることができる。

4. 7. 2 14 種類の特徴

4. 7. 1 に述べた手順によって、テクスチャ画像から濃度共起行列が作成される。Haralick らは更に、この濃度共起行列から新たな 14 種類の特徴を求めることを提案している[51]。以下にその手順を示す。

画素の階調数が g である画像から生成された $g \times g$ の大きさをもつ 4 つの行列、 $P(i, j; d, \theta)$ を、濃度共起行列 (d : 固定、 $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$) とし、4 つの行列の要素を次のように足し合わせ、1 つの行列とする。

$$P(i, j) = \sum_{\theta=0,45,90,135} P(i, j; d, \theta) \quad , \quad i, j = 1, \dots, g \quad (4. 7. 2)$$

更に、 $P(i, j)$ の要素を正規化すると、

$$p(i, j) = \frac{P(i, j)}{\sum_{i=1}^g \sum_{j=1}^g P(i, j)} \quad , \quad i, j = 1, \dots, g \quad (4. 7. 3)$$

$$0 \leq p(i, j) \leq 1 \quad , \quad \sum_i \sum_j p(i, j) = 1$$

となる。

この行列から以下の 14 種類の特徴が生成される。

f_1 : Angular Second Moment

$$f_1 = \sum_i \sum_j \{p(i, j)\}^2 \quad (4. 7. 4)$$

f_2 : Contrast

$$f_2 = \sum_{n=0}^{g-1} n^2 \left\{ \sum_{i,j} P(i, j) \right\}_{|i-j|=n} \quad (4. 7. 5)$$

f_3 : Correlation

$$f_3 = \frac{\sum_i \sum_j (i, j) p(i, j) - \mu^2}{\sigma^2} \quad (4. 7. 6)$$

但し、

$$\begin{aligned} p(i) &= p(j) = \sum_i p(i, j) = \sum_j p(i, j) \quad , \quad i, j = 1, \dots, q \\ \mu &= \sum_i p(i) = 1/g \\ \sigma &= \sqrt{\frac{\sum_i \{p(i) - \mu\}^2}{g-1}} \end{aligned}$$

f_4 : Sum of Squares

$$f_4 = \sum_i \sum_j (i - \mu)^2 p(i, j) \quad (4.7.7)$$

f_5 : Inverse Difference Moment

$$f_5 = \sum_i \sum_j \frac{p(i, j)}{1 + (i - j)} \quad (4.7.8)$$

f_6 : Sum Average

$$f_6 = \sum_{i=2}^{2g} i \cdot p_{x+y}(i) \quad (4.7.9)$$

但し、

$$p_{x+y}(k) = \sum_{i=1}^g \sum_{j=1}^g p(i, j) \quad , \quad k = 2, \dots, 2g \\ i+j=k$$

f_7 : Sum Variance

$$f_7 = \sum_{i=2}^{2g} (i - f_8)^2 \cdot p_{x+y}(i) \quad (4.7.10)$$

f_8 : Sum Entropy

$$f_8 = - \sum_{i=2}^{2g} p_{x+y}(i) \log \{p_{x+y}(i)\} \quad (4.7.11)$$

f_9 : Entropy

$$f_9 = - \sum_i \sum_j p(i, j) \log \{p(i, j)\} \quad (4.7.12)$$

f_{10} : Difference Variance

$$f_{10} = \frac{\sum_{k=0}^{g-1} \{p_{x-y}(k) - \sum P_{x-y}(k)/g\}^2}{g-1} \quad (4.7.13)$$

但し、

$$p_{x-y}(k) = \sum_{i=1}^g \sum_{\substack{j=1 \\ |i-j|=k}}^g p(i, j) \quad , \quad k = 0, \dots, g-1$$

f_{11} : Difference Entropy

$$f_{11} = - \sum_{i=0}^{2g} p_{x-y}(i) \log\{p_{x-y}(i)\} \quad (4.7.14)$$

f_{12} 、 f_{13} : Information Measures of Correlation

$$f_{12} = \frac{HXY - HXY1}{\max\{HX, HY\}} \quad (4.7.15)$$

$$f_{13} = (1 - \exp[-2(HXY2 - HXY)])^{1/2}$$

但し、

$$\begin{aligned} HXY &= - \sum_i \sum_j p(i, j) \log\{p(i, j)\} \\ HXY1 &= - \sum_i \sum_j p(i, j) \log\{p(i)p(j)\} \\ HXY2 &= - \sum_i \sum_j p(i)p(j) \log\{p(i)p(j)\} \end{aligned}$$

f_{14} : Maximal Correlation Coefficient

$$f_{14} = (\text{Second largest eigenvalue of } Q)^{1/2} \quad (4.7.16)$$

但し、

$$Q(i, j) = \sum_k \frac{p(i, k)p(j, k)}{p(i)p(k)} \quad , \quad i, j = 1, \dots, g$$

以上が Haralick らの提案した、濃度共起行列から新たに作成する 14 種類の特徴である。C.C.Gotlib と H.E.Kreyszig は、これら 14 種類の特徴が表していると考えられる性質を以下の 4 つに分類している[54]。

①テクスチャの特質を視覚的にあらわすもの： f_1 (全体の均一さ)、

f_2 (局所変化)、

f_3 (方向に対する相関)

②統計量に基づくもの：

f_4 (局所同質性)、

f_5 (複雑さの測度)、

f_6 、 f_7 、 f_{10}

③情報理論に基づくもの：

f_8 、 f_9 、 f_{11}

④相関の情報測度に基づくもの：

f_{12} 、 f_{13} 、 f_{14}

14 種類の特徴を用いたテクスチャ識別に関しては、現在までに多くの研究が行われ、比較的高い正解率を得ている。その中には 14 種類の特徴の中からいくつかの特徴だけを使った識別実験も含まれる[55]。しかしこれらの特徴は、単に上述したテクスチャの性質だけを表しているのではなく、1つの特徴がいくつかのテクスチャの性質を表したり、あるいは他の特徴も同じテクスチャの性質を表すなど、物理的に何を測定しているのか分からないところがある。また、実際の計算に費やす資源や時間などが大きくなることから、実際には入力画像をより少ない階調数へ再量子化する必要もある。

14 種類の特徴に対して、梅田はそれらを用いずに、濃度共起行列の要素を直接の特徴として用い、高い識別正解率を得ている[30]。その実験において、テクスチャ画像はメモリと特徴数を削減するために再々量子化されている。

4. 8 ランレンジス行列を用いる解析法

ランレンジスは、画像情報を圧縮するための符号化法として用いられることが多い[56][57][58]。ここでは、テクスチャ画像のランレンジスを求めることによって特徴を抽出する解析法を説明する。

画像のランレンジスは、図 4. 8. 1 のように、ある方向(θ)に沿って並んだ一列の画素集合に対して、同じ階調をもつ画素の連続する長さを表している。

ある一定の位置関係にある 2 つの画素の濃度の組み合わせに関する統計量である 2 次統計量を表している濃度共起行列やフーリエ・スペクトルに対して、このラン

レンジから求められる特徴は未知の画素を組み合わせてできる高次統計量を表している。

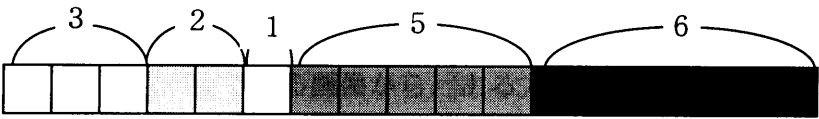


図 4. 8. 1 ランレンジス

ランレンジスを用いた特徴の抽出法を以下に説明する。
 あるテクスチャ画像内で、 θ の方向に関して濃度 i の点が j 個続く頻度（そのランの個数）を要素とするランレンジス行列 $P(i, j; \theta)$ を求める。図 4. 8. 2 に画像とランレンジス行列の例を示す。

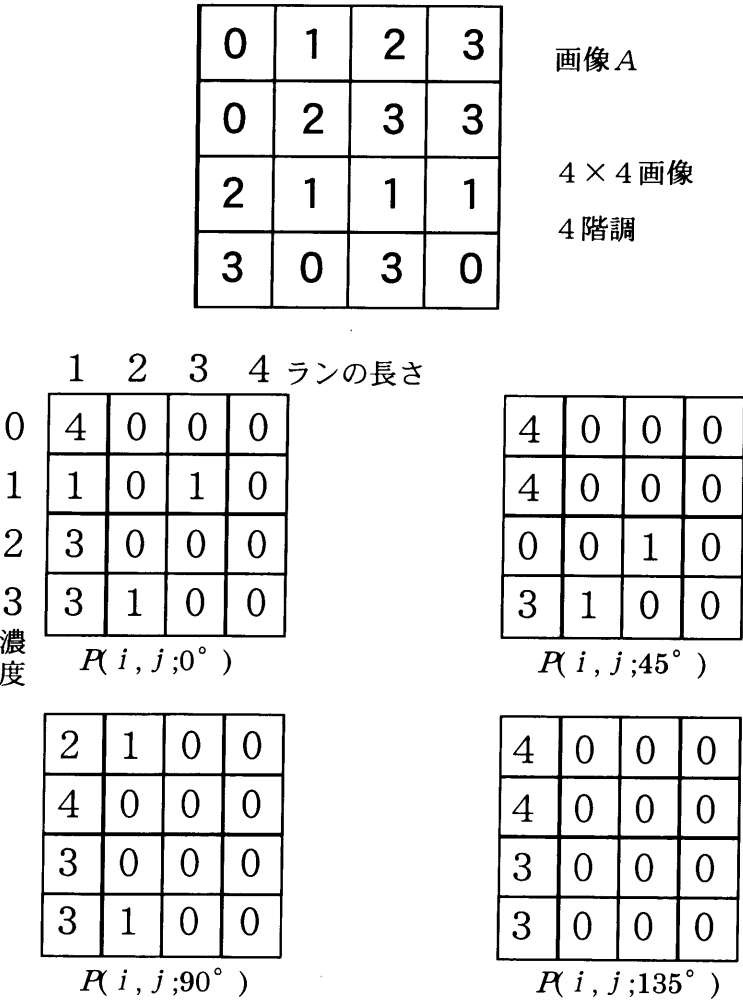


図 4. 8. 2 ランレンジス行列の例

M. M. Galloway は、ランレングス行列から求める 5 種類の特徴を提案している。以下にその 5 種類の特徴を示す。

大きさが $N \times N$ 、階調数が g である画像から、ある方向 θ に関して作成されたランレングス行列の要素の合計値 T を

$$T = \sum_{i=1}^g \sum_{j=1}^N P(i, j; \theta) \quad (4.8.1)$$

とすると、ランレングス行列から作成される 5 種類の特徴は次のようになる。

f_1 : short runs emphasis

$$f_1 = \frac{1}{T} \sum_{i=1}^g \sum_{j=1}^N \frac{P(i, j; \theta)}{j^2} \quad (4.8.2)$$

f_2 : long runs emphasis

$$f_2 = \frac{1}{T} \sum_{i=1}^g \sum_{j=1}^N j^2 P(i, j; \theta) \quad (4.8.3)$$

f_3 : gray level nonuniformity

$$f_3 = \frac{1}{T} \sum_{i=1}^g \left[\sum_{j=1}^N P(i, j; \theta) \right]^2 \quad (4.8.4)$$

f_4 : run length nonuniformity

$$f_4 = \frac{1}{T} \sum_{j=1}^N \left[\sum_{i=1}^g P(i, j; \theta) \right]^2 \quad (4.8.5)$$

f_5 : run percentage

$$f_5 = \frac{1}{N^2} \sum_{i=1}^g \sum_{j=1}^N P(i, j; \theta) \quad (4.8.6)$$

ランレングスを用いて特徴を抽出する方法は、手続きが容易であるので処理に要する時間が比較的少ない。その反面、使用する資源は膨大である。また、階調数が大きくなるとランの長さが短くなる。そのため、入力されたテクスチャ画像を再量子化する、あるいは階調変化に鈍感なランを作成する必要がある。K. Chehdi と Q. Liao は、視覚感度関数を用いて、濃度によって感度が異なるランを作成している[26]。彼らはこの方法によって比較的高い正解率を得ている。

4. 9 フラクタルを用いる解析法

フラクタルは現在、物性物理学からコンピュータ・グラフィクスに至るまで広い範囲で注目を集めている。また近年、画像解析のためのフラクタルの応用についての研究がなされている[27][59]。

4. 9. 1 フラクタル次元

狭義のフラクタルとは、自己相似次元のことである。

相似次元とは自己相似図形に対して与えられる正の値であり、次のように定義される。例として、 D 次元立法体のそれぞれの辺を $N^{1/D}$ 等分すると、もとの立方体は N 個の小立方体に分割される。このとき、もとの立方体と分割後の小立方体との相似比 R は

$$R(N) = \frac{1}{N^{1/D}} \quad (4. 9. 1)$$

と書ける。式(4. 9. 1)を書き換えれば、次元 D は

$$D = \frac{-\log N}{\log R(N)} \quad (4. 9. 2)$$

と書ける。逆に、式(4. 9. 2)の右辺で自己相似図形の相似次元を定義する。 D の値は、線分や立方体などの単純な自己相似図形や、その有限和から構成されるような図形では、位相次元と同じ正の整数値をとる。一方、自己相似図形の極限図形においては通常の次元よりも大きな非整数値をとる。相似次元は複雑さの表現指標と考えられる。相似次元が位相次元よりも大きなとき、Mandelbrot はその図形をフラクタルと呼び[60]、その次元をフラクタル次元と呼んだ。

しかしながら、この自己相似性は多くの事象に対するモデルとしては適さない。従って、狭義のフラクタルの拡張がなされている。以下にそのいくつかを示す。

①確率的フラクタル次元

相似次元の1つの拡張は、確率的（統計的）自己相似性に基づくものである。

確立過程 $X(t)$ が

$$X(t) = t^{D_p} \cdot X(1), \quad t = 1, 2, \dots, n \quad (4.9.3)$$

を満たすとき、 $X(t)$ は D_p を指数とする自己相似な確率過程であるという。式

$$(4.9.3) \quad \text{の両辺の平均をとれば} \quad (4.9.4)$$

$$D_p = -\frac{\log\{E\{X(t)\}/E\{X(1)\}\}}{\log(1/t)}$$

が導かれる。自己相似な確率過程は狭義の自己相似図形の一般化となる。

②ハウスドルフ次元

自己相似性が成り立たない図形に対して、自己相似次元の拡張としてハウスドルフ次元が定義される。

距離空間 S 内の集合 E が、半径 $r_m (m=1, 2, \dots)$ の球 V_{r_m} で被覆されているとする。すなわち、

$$\bigcup_m V_{r_m} \supset E \quad (4.9.5)$$

とする。このとき、 E のハウスドルフ測度 $md(E)$ は次のように定義される。

$$md(E) = \lim_{\max_m(r_m) \rightarrow 0} \inf_{\{V_{r_m}\}} \sum_m r_m^m \quad (4.9.6)$$

ここで、式(4.9.6)の $\inf$ は式(4.9.5)を満たすすべての被覆 V_{r_m} についてとる。このとき、

$$md(E) = \begin{cases} \infty & d < DH \\ 0 & d > DH \end{cases} \quad (4.9.7)$$

ならば、 DH を E のハウスドルフ次元と呼ぶ。

ハウスドルフ次元は相似次元の一般化となる。ハウスドルフ次元には自己相似の概念が入っていないので、任意の図形に対して定義できる。また、ハウスドルフ次元は、曲面の変化の不安定さを表す尺度であると考えられる。以下、相似次元とハウスドルフ次元とをあわせてフラクタル次元と呼ぶことにする。

4.9.2 フラクタル特徴の画像解析への応用

4.9.1 で述べたフラクタル次元では、濃淡画像を取り扱うテクスチャ解析に不十分である。以下では、画像解析への応用を念頭において、さきに述べたフラク

タル次元のデジタル画像上での計算方法について述べる。

①画像面の確率的相似次元

式(4.9.2)、(4.9.4)の比較では、 $1/t$ が相似比に対応する。そこで $r=1/t$ として、 $X(t)=N(1/t)=N(r)$ と定義すると、 $N(r)$ は計測単位 r で図形を計測したときの図形分割数である。これをもとに、画像面が自己相似的であるということは $1/t \times 1/t$ の小領域を単位として計測した画表面分割数 $X(t)$ が自己相似な確率過程であること、すなわち式(4.9.3)が成立することであると定義する。このとき、式(4.9.4)を $N(r)$ で書き換えると

$$D_p = \frac{-\log\{E\{N(r)\}/E\{N(1)\}\}}{\log r} \quad (4.9.8)$$

となる。式(4.9.8)は、その右辺が平均的には r に独立なことを示している。しかし、標本数が有限の画像領域では式(4.9.8)の右辺の計測値は r に依存する。このとき、 $E\{N(r)\}$ が線形回帰モデル

$$\log E\{N(r)\} = -D_p \log r + \log E\{N(1)\} + \varepsilon(r) \quad (4.9.9)$$

に従う($\varepsilon(r)$:攪乱項)とすれば、 D_p は最小二乗法で推定できる。

画像上の具体的な計算では $r \times r$ の小領域を単位として計った画像表面積を $A(r)$ とすると、 $A(r) = r \times r \times N(r)$ なので式(4.9.9)は

$$\log E\{A(r)\} = (2 - D_p) \log r + C \quad (4.9.10)$$

となる。但し、

$$C = \log E\{A(1)\} + \varepsilon(r)$$

とする。

②制限ハウスドルフ次元

先の定義に従って厳密にハウスドルフ次元を計算するのは容易でない。そこで、この定義を変更した制限ハウスドルフ次元を新たに求める。

ある集合 E の制限ハウスドルフ次元 DH_r は、 E を半径 r の閉球で被覆するとき必要とされる球の数 $N(r)$ に対して

$$K = N(r) \cdot r^D \quad (4.9.11)$$

が $r \rightarrow 0$ で0となる最小の D として定義される。これより

$$DH_r = \lim_{r \rightarrow 0} \left\{ \frac{-\log N(r)}{\log r} \right\} \quad (4.9.12)$$

となる。更に、式(4.9.12)はL'Hospitalの定理によって

$$DH_r = \lim_{r \rightarrow 0} \left\{ \frac{-d \log N(r)}{d \log r} \right\} \quad (4.9.13)$$

が得られる。しかし、デジタル画像においては $r \geq 1$ であり、 $r \rightarrow 0$ とはできない。従って

$$DH_r = - \left\{ \frac{\{\log N(2) - \log N(1)\}}{(\log 2 - \log 1)} \right\} \quad (4.9.14)$$

が得られる。以後、式(4.9.13)の右辺をデジタル(制限)ハウスドルフ次元と呼ぶ。

4.9.3 フラクタル特徴の拡張

ここでは、一般的な画像への応用が可能で、豊富な画像解析能力を持ったフラクタル特徴について述べる。

①局所的なフラクタル次元

自己相似次元の概念を一般化して、 r に依存した局所的相似次元を定義する。式(4.9.8)において、 D_p が $(\log r, \log(E\{N(r)\}/E\{N(1)\}))$ の集合の作る曲線の傾きであることに着目する。これをふまえて、局所相似次元 $D(r)$ を

$$D(r) = \frac{-d \log \{E\{N(r)\}/E\{N(1)\}\}}{d \log r} \quad (4.9.15)$$

で定義する。局所フラクタル次元によって、図形の自己相似性の強さを計量することが可能となる。

②2次元フラクタル特徴

局所フラクタルの概念を拡張し、より画像情報の抽出に有用な2次元局所フラクタル特徴の定義を述べる。

曲面 $z = f(x, y)$ を $m \times n$ の矩形で覆い、その図形分割数を $Nf(m, n)$ とする。このとき、2次元局所フラクタル次元 $Df(m, n)$ を

$$Df(m, n) = \left(\frac{dF(m, n)}{d \log m}, \frac{dF(m, n)}{d \log n} \right) \quad (4.9.16)$$

$$F(m, n) = -\log \left\{ \frac{Nf(m, n)}{Nf(1, 1)} \right\} \quad (4.9.17)$$

で定義する。式(4.9.16)を、分割した画面の数 $Nf(m, n)$ を使って書き換えれば、

$$Af(m, n) = mn Nf(m, n)$$

だから

$$Df = \left(\frac{-d \log Af}{d \log m + 1}, \frac{-d \log Af}{d \log n + 1} \right) \quad (4.9.18)$$

が得られる。 $Df(m, n)$ によって、画像の方向性に関する情報の抽出ができる。

4. 10 まとめ

本章では、種々のテクスチャ情報の解析法について述べ、構造的な解析法として Hough 変換による方法を示し、統計的な解析法として濃度ヒストグラムによる方法、自己相関関数による方法、フーリエ・スペクトルによる方法、差分統計量による方法、濃度共起行列による方法、ランレンジ行列による方法、フラクタルによる方法について述べた。ここでは以上の8種類のテクスチャ解析法の性能についてまとめ、最も有用なテクスチャ解析法を選び出す。以下に各方法の性能のまとめを示す。

① 8種類の方法の中では、構造的な解析法である Hough 変換による方法が最も強力である。しかしながらこの方法は対象となるテクスチャの種類が限られる。

② 統計的な解析法の中では、1次統計量である濃度ヒストグラム法が最も弱い。2次統計量と高次統計量に基づく方法では濃度共起行列による方法が最も強力であり、行列から求められる14種類の特徴はコントラストや一様性など画像の多くの性質を表わすことができる。ただし、それぞれの特徴は表わす内容に過不足がある。従って識別など、解析を目的としない場合は行列の要素そのものを用いた法が適切である。

③それぞれの方法によって求められる特徴の中で、Hough 変換による特徴のみがテクスチャの幾何学的な変換に対して独立している。また、濃度共起行列、ランレングス法によって求められる特徴は 45° 単位の回転には独立している。それ以外の特徴はテクスチャの幾何学的な変換に従って値が変化する。

以上の分析から、濃度共起行列の要素をテクスチャ情報として用いる方法が最も有用な方法であると判断する。濃度共起行列の行列の要素に基づくテクスチャ特徴は平行移動および 45° 単位の回転に対して不変である。

第5章

ニューラルネットワークによるパターン 識別

5. 1 まえがき

本章では、輪郭線情報とテクスチャ情報を統合し識別を行う回路として用いるニューラルネットワークについて述べる。ニューラルネットワークは学習機能を持ち、異種情報の統合には極めて適切な回路である。ニューラルネットワークには、階層型モデル[61][62][63][64]や相互結合型モデル[65][66][67]など種々の構造を持つネットワークがあり、近年では多様な分野で応用されている。本研究で行う輪郭線情報とテクスチャ情報の統合には、3層の階層型ニューラルネットワークを用いる。3層ニューラルネットワークは構造が単純であり、またバックプロパゲーションという優れた学習アルゴリズムが利用できる。更に異種情報の統合やパターンの類別能力が高く、極めて有用なネットワークである。

本章では、ニューラルネットワークに関し基本的な事柄を述べ、3層ニューラルネットワークの構造とバックプロパゲーションによる学習アルゴリズムを示す。更に3層ニューラルネットワークによる輪郭線情報とテクスチャ情報の統合方法について詳述する。

5. 2 ニューラルネットワークの基礎原理

5. 2. 1 ニューロンとニューラルネットワークモデル

ニューラルネットワークは、図 5.2.1 に示すように、ニューロンと呼ばれる多数の演算素子と、それらを結合するシナプスと呼ばれるアナログ量の分散記憶素子で構成される。このアナログの結合強度をシナプス荷重あるいはシナプス結合強度と呼ぶ。シナプス結合には正（興奮性）と負（抑制性）の結合がある。

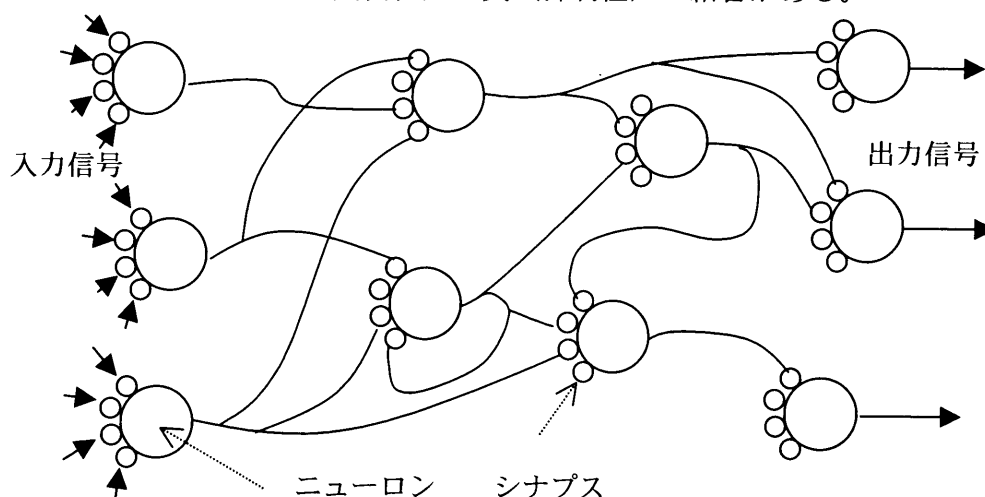


図 5. 2. 1 ニューラルネットワークの構造

ニューロンは多入力ー1出力型の素子であり、図 5.2.2 で示すように入力信号の積和演算とシグモイド型の非線型入出力特性を与えるモデルがよく用いられる。すなわち、あるニューロンの興奮状態（出力）は、そのニューロンに結合されている他のニューロンの興奮状態がシナプス荷重され、非線型入出力特性に従って得られるものである。加算されたシナプス荷重を膜電位と呼ぶ。シナプス荷重は用いるニューラルネットワークモデルによって、あらかじめ計算して与えられる場合と、学習によって決められる場合とがある。

ニューラルネットワークは空間的あるいは時間的なパターンの非線型変換装置である。何らかの空間パターンを入力すると、ニューロン間の配線状態、シナプス荷重、ニューロンの非線型入出力関数によって出力パターンが決まる。

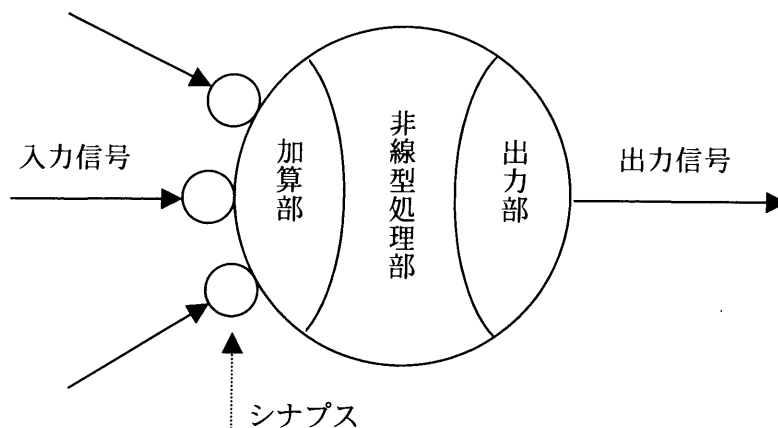


図 5. 2. 2 ニューロンのモデル

5. 2. 2 ニューラルネットワークの学習

ニューラルネットワークにおける学習とは、外界から信号を入力したときにネットワークを構成するシナプス荷重を望ましい値に変化させることである。ニューラルネットワークの学習方法には、1) 入力信号の性質のみに基づく教師なし学習と、2) 望ましい出力が外部から教えられる教師付き学習とに分類される。

学習則の名称	学習信号	代表的なモデル	特徴
ヘップ学習則	v_i	ボルツマンマシン[65]	教師なし学習
相関学習則	d_i	アソシアトロン[62]	連想メモリへの応用
デルタルール	$d_i - v_i$	パーセプトロン [55][56]	パターン分類
一般化デルタルール	$(d_i - v_i) f'(u_i)$	バックプロパゲーション[59][60]	線形分離不可能な問題にも適応可能

表 5. 2. 1 学習則の分類

一般に、 i 番目と j 番目のニューロンを結合するシナプス荷重 w_{ij} の修正量を与える学習は、

$$\Delta w_{ij} = \eta \delta v_j \quad (5. 2. 1)$$

で表される[68]。ここで、 η は学習の速さを表す比例定数で、学習パラメータと呼

ばれる。また、 δ は学習信号と呼ばれる。

学習信号 δ は、一般にシナプス荷重 w_{ij} 、ニューロンの興奮状態 v_i および教師信号 d_i などの関数であり、これによりさまざまな学習方式に分類できる。表 5.2.1 に代表的な学習則を示す。

5. 2. 3 ヘップの学習則

ヘップの学習則は教師なし学習の基本をなすもので[69]、式(4.2.1)において $\delta = v_i$ とする。従って荷重の修正量は

$$\Delta w_{ij} = \eta v_i v_j \quad (5. 2. 2)$$

となる。この学習則では、興奮しているニューロン同士 ($v_i = 1$ かつ $v_j = 1$) のシナプス結合が強化されることがわかる。すなわち、一度信号が伝達されるとその情報伝達路はますます通りやすくなる。但し、このままでは荷重値が発散してしまう場合があるのでシナプス荷重のノルムを一定にしたり、減衰項を付け加えるなどの修正が加えられている。ヘップの学習則は、修正された形でボルツマンマシンなどのモデルに用いられている。

5. 2. 4 相関学習則

学習信号 δ が教師信号 d_i に等しい場合、すなわち

$$\Delta w_{ij} = \eta d_i v_j \quad (5. 2. 3)$$

とするモデルを相関学習と呼ぶ。出力と教師信号の相関が荷重の更新に用いられる。相関学習則では、シナプス荷重は

$$w_{ij} = \frac{\eta}{M} \sum_{m=1}^M d_i^{(m)} v_j^{(m)} \quad (5. 2. 4)$$

に収束することが示されている[68]。ここで M は教師信号の個数である。このシナプス荷重の表式は相関行列と呼ばれ、中野や Kohonen らによって提案され、連想記憶などへの応用が研究されている[70][71]。また甘利らによってその特性が詳細に調べられている[72]。

5. 2. 5 デルタルール

デルタルールは誤り訂正学習法とも呼ばれ、最も用いられている教師付き学習則である。デルタルールでは、式(5.2.1)において学習記号 δ として

$$\delta = d_i - v_i \quad (5. 2. 4)$$

を用いる。ここでは、 d_i は i 番目の出力ニューロンの教師信号、 v_i はネットワークの出力信号である。従ってデルタルールでは、ネットワークが誤った出力をした場合のみ、各々の荷重が教師信号との差に比例して修正される。

5. 3 3層ニューラルネットワークとその学習

5. 3. 1 3層ニューラルネットワークの構造

階層型モデルである、3層ニューラルネットワークの概念図を図 5.3.1 に示す。ネットワークは入力層 (Input: I)、隠れ層 (Hidden: H)、および出力層 (Output: O) で構成される。信号は入力層から出力層に向かって進むので、フィードフォワード型モデルとも呼ばれる。各層のユニット (ニューロン) は、隣り合う層の間でシナプス結合をする。同じ層や入力層と出力層の間では結合はない。階層型モデルは古くから学習機能を持つネットワークとして用いられている。学習では、シナプス結合の荷重を更新させて、出力層のニューロンの出力を最適化する。

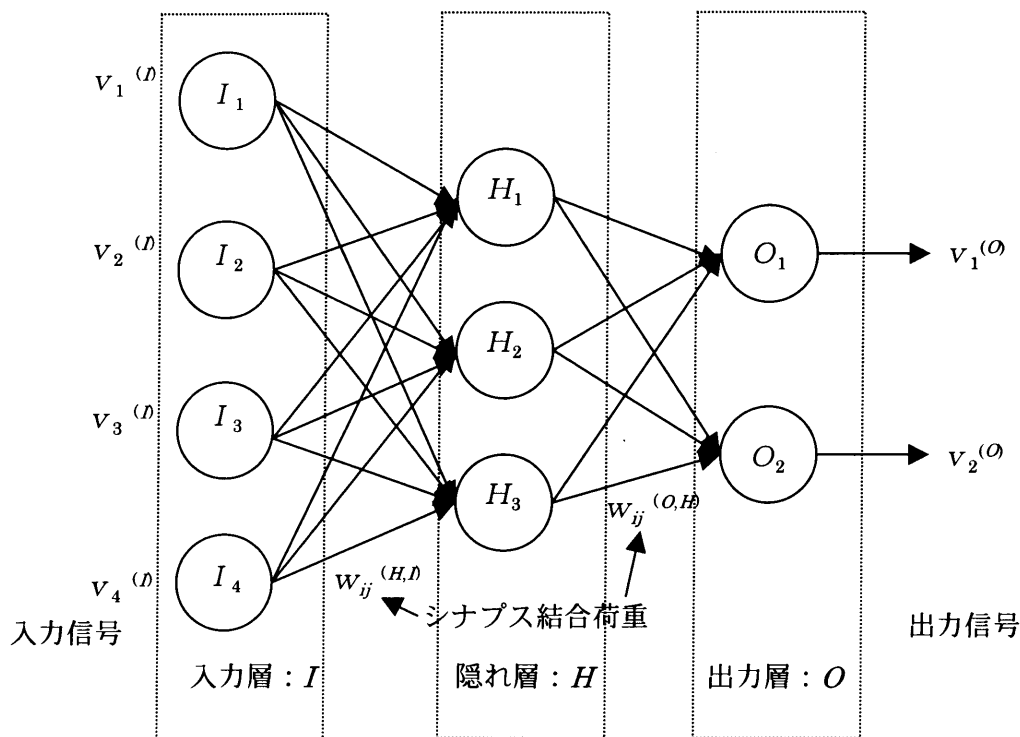


図5. 3. 1 3層ニューラルネットワーク

5. 3. 2 バックプロパゲーションによる学習

次に、ここ十数年注目されている、教師付き学習法の一つであるバックプロパゲーション（Back Propagation：BP）法について詳しく述べる。その学習は、信号の流れから逆方向に、すなわち出力層から入力層に向かって再帰的に計算を行い実行される。BP法と呼ばれる所以である。学習則として、5. 2. 5で述べたデルタルールを拡張した一般化デルタルールを用いる。デルタルールは、出力層と中間層間のシナプス荷重に関してのみ、その修正式を与えている。これを出力層以前の全ての層間のシナプス荷重の修正に拡張したものが一般化デルタルールである。以下に多層ニューラルネットワークに対するバックプロパゲーションアルゴリズムを示す。

バックプロパゲーション法の学習則は次のとおりである。I層を第1層として、 k 層に属する i 番目のニューロンの入出力信号をそれぞれ $u_i^{(k)}$ 、 $v_i^{(k)}$ とすると

$$u_i^{(k)} = \sum_{j=1}^{N_{k-1}} w_{ij}^{(k,k-1)} v_j^{(k-1)} - \theta_i^{(k)} \quad (5.3.1)$$

$$v_i^{(k)} = f(u_i^{(k)}) \quad (5.3.2)$$

で与えられる。ここで $w_{ij}^{(k,k-1)}$ は $(k-1)$ 層内の j 番目のニューロンと k 層内の i 番目のニューロンのシナプス荷重、 N_{k-1} は $(k-1)$ 層に含まれるニューロン数である。また、 $\theta_i^{(k)}$ は k 層内の i 番目のニューロンのしきい値、関数 f はニューロンの入出力関数で、通常

$$f(x) = \frac{1}{1 + \exp(-2x/x_0)} \quad (5.3.3)$$

のシグモイド関数が用いられる。学習は、教師信号に対するネットワークの出力信号の自乗誤差として定義される評価関数が最小値に収束するように実行される。これを各層間に適用して、 k 層における望ましい出力（出力層であれば教師信号）を $v_i'^{(k)}$ ($i = 1, 2, \dots, N_k$)、実際に用いられる出力を $v_i^{(k)}$ ($i = 1, 2, \dots, N_k$) とすれば、評価関数 E は

$$E = \frac{1}{2} \sum (v_i'^{(k)} - v_i^{(k)})^2 \quad (5.3.4)$$

で与えられる。この E を最小にするようにネットワークのシナプス荷重を変化させる。

学習則を求めた結果を示せば、シナプス荷重の修正量 $\Delta w_{ij}^{(k,k-1)}$ は

$$\Delta w_{ij}^{(k,k-1)} = \eta_i^{(k)} v_j^{(k,k-1)} \quad (5.3.5)$$

で与えられる。 η は収束の速さを決める学習パラメータである。 $\eta_i^{(k)}$ は k 層が出力層の場合には

$$\gamma_i^{(o)} = (d_i - v_i^{(o)}) f'(u_i^{(o)}) \quad (5.3.6)$$

である。 d_i ($i = 1, 2, \dots, N_{(o)}$) は教師信号である。一方、 k 層が出力層以外の場合には

$$\gamma_i^{(k)} = \left(\sum_j w_{ij}^{(k+1,k)} \gamma_j^{(k+1)} \right) f'(u_i^{(k)}) \quad (5.3.7)$$

である。式(5.3.5)～(5.3.7)より、任意の k 層の修正信号 $\Delta w_{ij}^{(k,k-1)}$ は一つ前の修正信号 $\Delta w_{ij}^{(k+1,k)}$ から再帰的に計算されることがわかる。

5. 4 情報の統合と識別

これまで述べた3層ニューラルネットワークにより、輪郭線情報とテクスチャ情報は統合される。両者の特徴は画像から抽出され、ネットワークに並列に入力される。最後にネットワークの出力結果から識別を行う。学習では入力された特徴から得られたネットワークの出力値と教師信号を比較し、それらの差をバックプロパゲーション法を用いて修正していく。この学習過程の中で輪郭線情報とテクスチャ情報の統合回路が形成される。図 5.4.1 に3層ニューラルネットワークによる輪郭線情報とテクスチャ情報の統合モデルを示す。

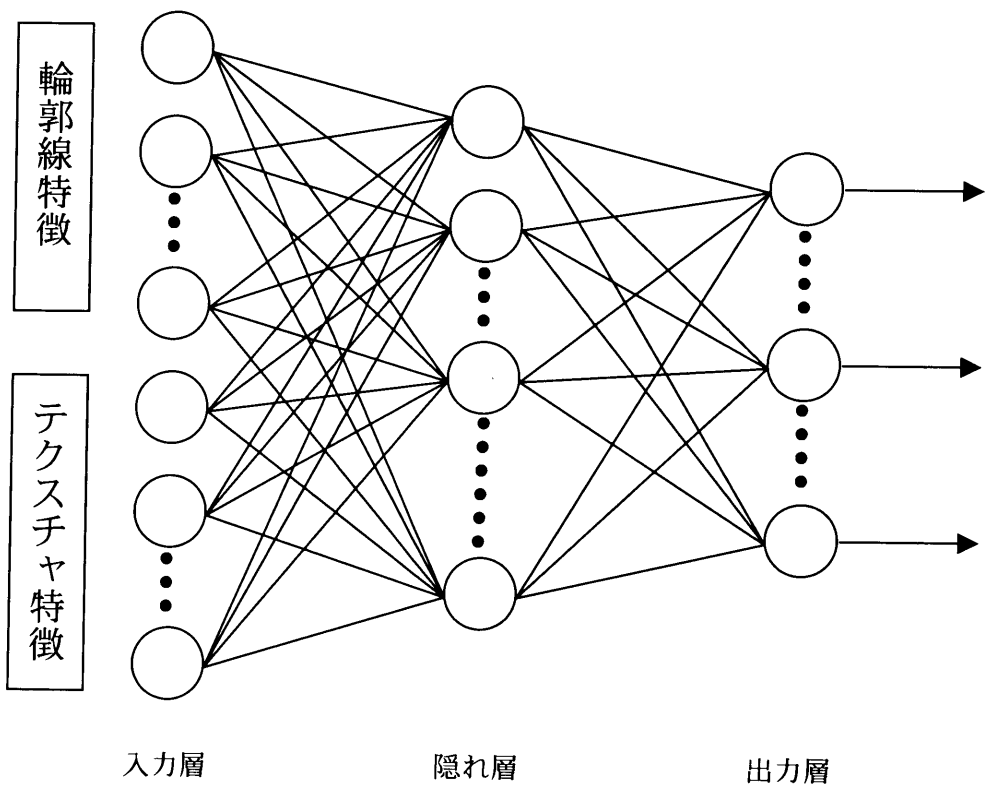


図 5. 4. 1 3層ニューラルネットワークによる情報の統合

5. 4 まとめ

本章では、輪郭線情報とテクスチャ情報とを統合し識別を行う回路として用いるニューラルネットワークに関して述べた。ニューラルネットワークは種々の構造のものがあるが、その中でも3層ニューラルネットワークは構造が単純であり、またバックプロパゲーション法により容易に学習を行うことができる。更に高い異種情報の統合能力と識別能力を持つので極めて有用な回路である。本章ではこの3層ニューラルネットワークによって輪郭線情報とテクスチャ情報とを統合する方法を示した。この方法により両者を統合した結果の有効性を次章の実験で実証する。

第 6 章

輪郭線情報とテクスチャ情報の統合による魚画像の識別

6. 1 まえがき

画像認識を行う多くの場合、画像は画像空間から特徴空間に写像され、特徴の比較によって認識がなされる。画像認識のための特徴として何を用いるかは、対象となる画像の性質に依存する。例えば横顔シルエットや縄文土器画像の形状認識などにおいては、もっぱらその輪郭線に基づく特徴が用いられる。一方、Brodatz のテクスチャ画像[73]の識別やじゅうたん模様の識別などでは輪郭線情報は役に立たず、テクスチャ構造を表現する特徴が用いられる。しかしながら、多くの場合は輪郭線情報とテクスチャ情報がともに有効である。

本論文では、これら輪郭線情報とテクスチャ情報に注目して、両者を統合して用いることが画像識別に対して極めて有効であることを論じてきた。第 2 章では、輪郭線の作成手段や特徴の表現方法について述べ、E P A C に基づく P 型記述子のパワースペクトルが最も有効な輪郭線情報であることを示した。第 3 章では、輪郭線の大域的な情報を安定して取り出すために輪郭線を平滑化する方法として L M e d S 推定法の有効性を述べた。第 4 章では、代表的なテクスチャ解析法について、構造的な解析法と統計的な解析法の 2 つに分けて、さまざまな解析法を詳述し、その中で濃度共起行列の要素を用いる方法が最も有用であることを示した。第 5 章では、輪郭線、テクスチャという 2 種類の異なる情報を統合し識別を行う手段として、3

層ニューラルネットワークに関して述べ、それによる情報の統合方法を示した。

以上の事柄から、輪郭線情報とテクスチャ情報を統合することが画像識別に関して有効な方法であることを明らかにすることができた。更に本章では具体的な例として画像識別実験を行い、情報の統合の有効性を実証する。識別の対象とする画像は図 6.1.1.で示すような魚類図鑑[32]の魚画像である。

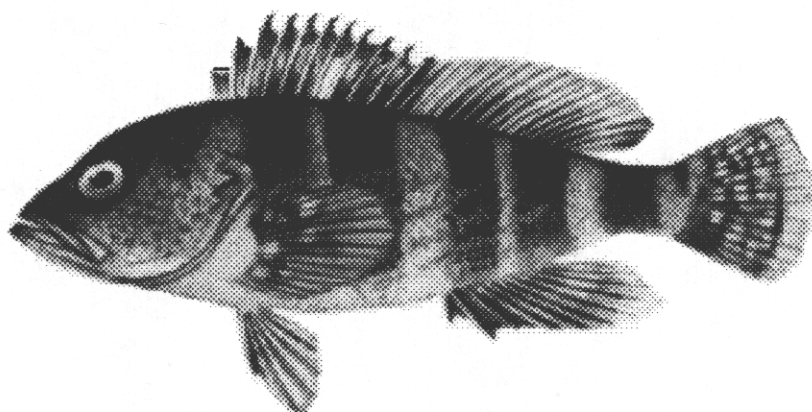


図 6. 1. 1 魚画像の例

6. 2 識別システム

図 6.2.1 で示すように画像識別の過程は、取り込んだ画像から特徴を抽出する処理と、抽出した特徴をあらかじめ定めたクラスに分類する処理に分けられる。前者は与えられた画像空間から他の空間への射影であり、後者は変換された空間の分割を中心とした処理である。本章における実験では、魚画像から輪郭線特徴およびテクスチャ特徴を抽出することが前者に相当し、ニューラルネットワークによる統合と類別が後者に相当する。更に、特徴抽出処理を細かく分けると、カメラやスキャナーなどによって取り込まれた画像に対して特徴抽出が容易となるように画像変換を行う前処理部と、変換された画像から特徴を抽出する本処理部とに分類される。前処理部では、例えば輪郭線特徴の抽出のためには2値化やエッジ抽出を行い、テクスチャ特徴の抽出のためには再量子化や重心検出を行う。実際の識別実験においては、あらかじめ画像取り込みと特徴抽出を行い、特徴データベースをすべての画

像に関して作成した後、識別を行う。

また、本実験では識別手段としてニューラルネットワークを用いるため、事前の学習が必要となる。学習を行う場合においても、その過程は識別の過程とほぼ同等であり、識別を行った後に識別結果と教師信号を比較し、その差をもとにネットワークの荷重を修正する処理が加わる。図 6.2.2 に識別処理のより詳細なフローを、図 6.2.3 に学習のフローを示す。

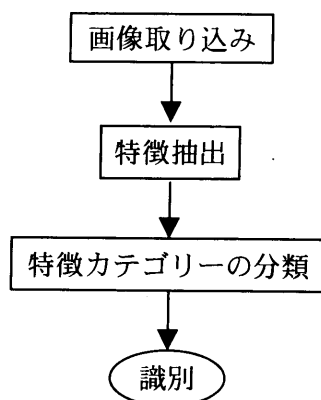


図 6. 2. 1 画像識別の過程

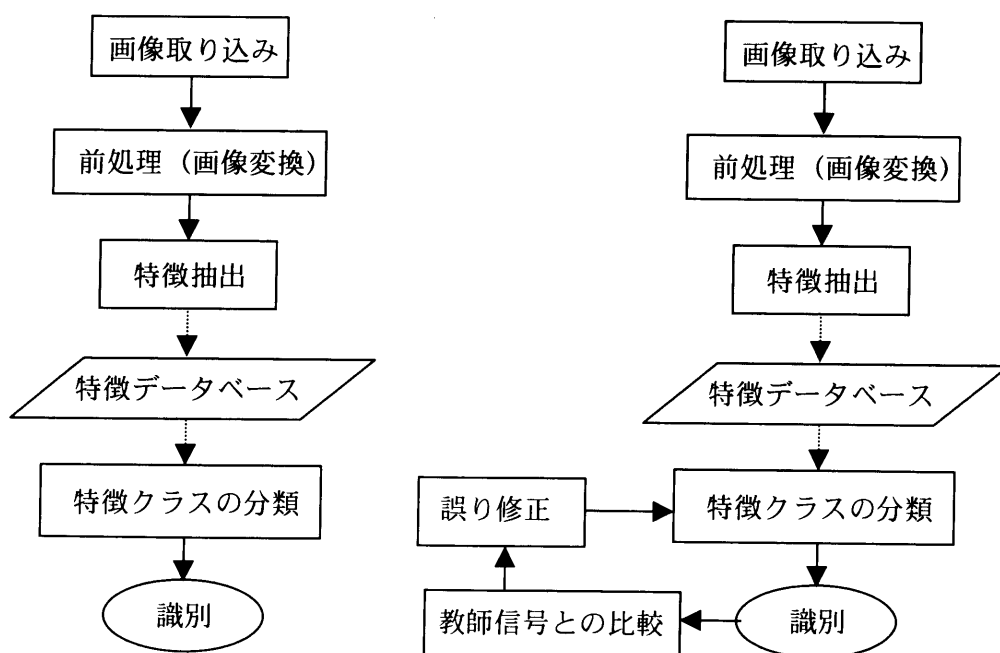


図 6. 2. 2 識別処理のフロー

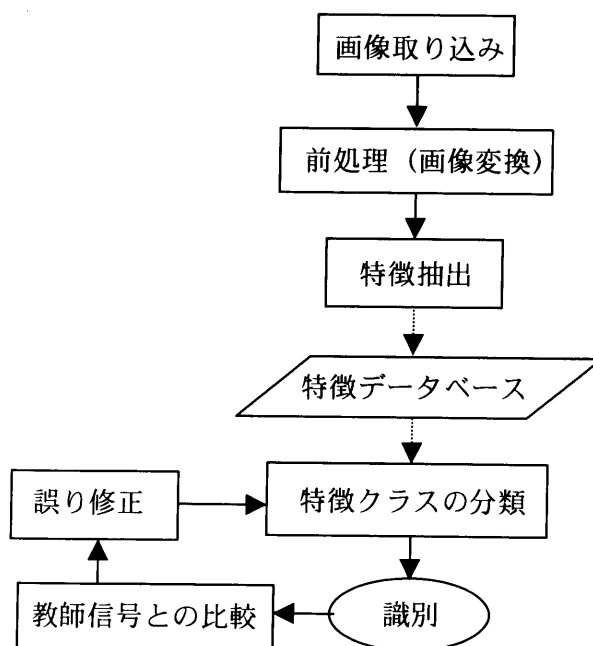


図 6. 2. 3 学習のフロー

6. 3 魚画像の特徴抽出

6. 3. 1 魚画像一覧

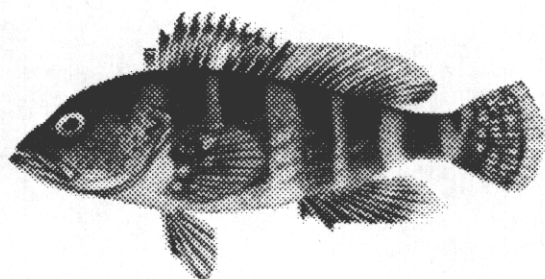
本実験では、入力画像として魚類図鑑からスキャナによって取り込んだ魚画像を用いる。魚は10科選び出し、各科5種類、合計50枚の魚画像を標準画像としている。表6.3.1に魚画像の一覧を示す。更に、これら50枚の標準画像に対して、4種類の拡大縮小（110%、100%、90%、80%）と、4種類の回転（0°、45°、90°、135°）を行い、変換画像を作成した。この結果、対象となる画像は $50 \times 4 \times 4 = 800$ 枚となった。これらの画像の中で、拡大縮小が100%、回転が0°の画像が標準画像である。図6.3.1～図6.3.10に元画像を示す。

科	魚 名	科	魚 名
ハタ科	アオハタ ホウセイハタ モヨウハタ オオスジハタ クエ	アイゴ科	ハナアイゴ ゴマアイゴ ヨコジマアイゴ マジリアイゴ ムシクイアイゴ
カレイ科	アサバガレイ ウマガレイ シモフリガレイ ベロガレイ ツマリクキノクガレイ	ニザダイ科	ニジハギ ヒラニザ イレズミニザ ナンヨウニザ ヒレナガニザ
チョウチョウオ科	カガミチョウチョウオ アミメチョウチョウオ ベニチョウチョウオ ニセフウライチョウチョウオ ゴマチョウチョウオ	マフグ科	メフグ コモnfグ ムシフグ ホシフグ サザナミフグ
モンガラカワハギ科	モンガラカワハギ キベリモンガラ クマドリ モラサメモンガラ アミメモンガラ	ニシン科	ホシヤマトミズン ミズン ギンイワシ ドロクイ リュウキュウドロクイ
ベラ科	シチセンベラ タヌキベラ クロフスキベラ ヤマブキベラ ノドグロベラ	イサキ科	コロダイ ナンヨウダイ ムスジコショウダイ アヤコショウダイ オシャレコショウダイ

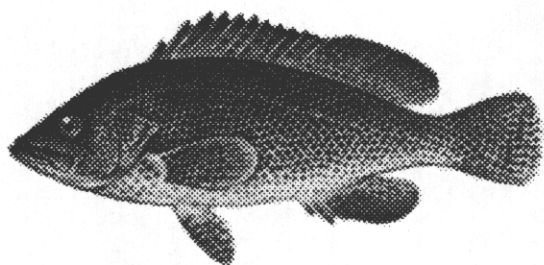
表6. 3. 1 実験に用いる魚画像

画像はすべてグレイレベル 256 階調である。

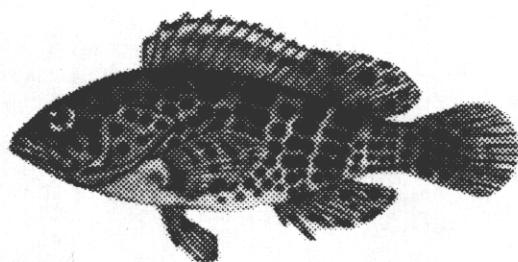
なお、本実験で用いる情報の統合・類別器はニューラルネットワークである。従って、ネットワークの重みを学習によって決定する必要がある。そこで、800 枚の画像のなかで 400 枚を学習データとした。残りの 400 枚をテストデータとして実験を行った[74][75][76]。



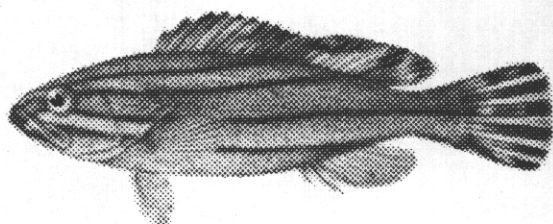
アオハタ



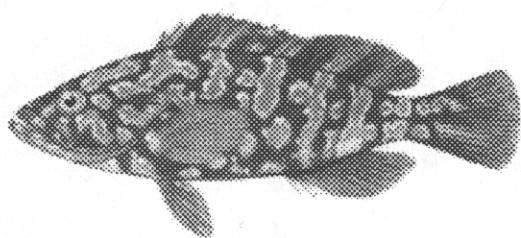
ハウセイハタ



モヨウハタ

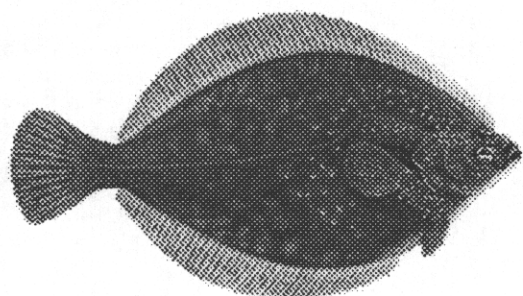


オオスジハタ

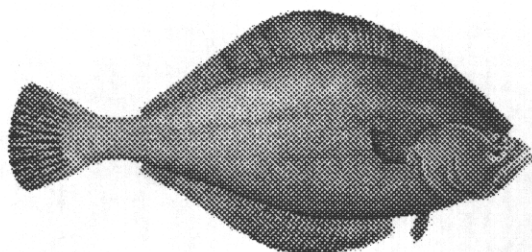


クエ

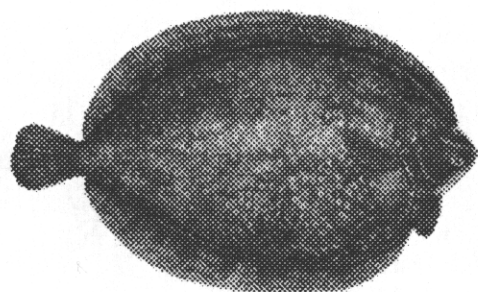
図6. 3. 1 ハタ科



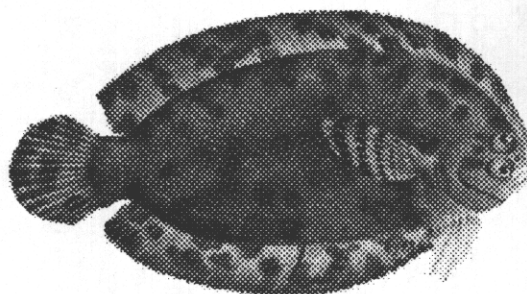
アサバガレイ



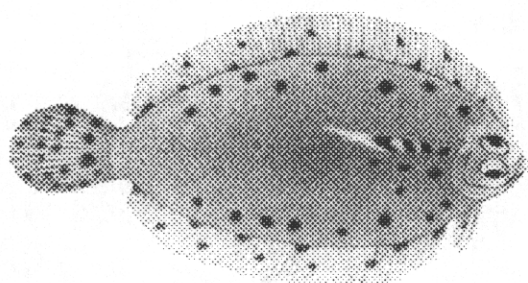
ウマガレイ



シモフリガレイ

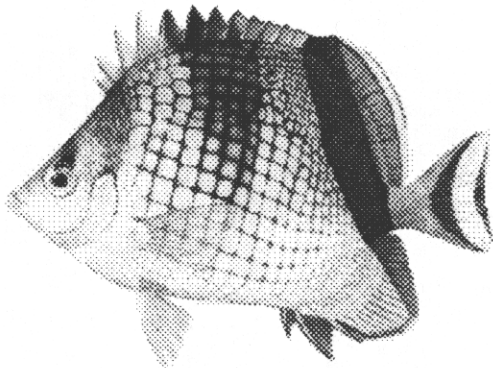


ベロガレイ

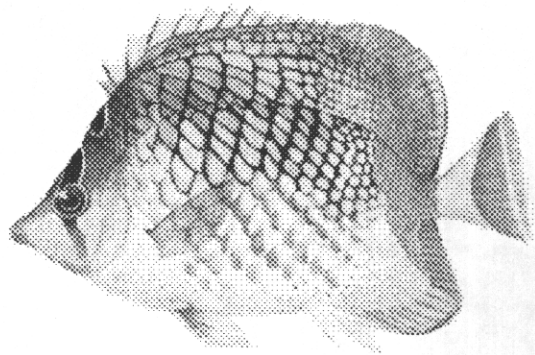


ツマリツキノクガレイ

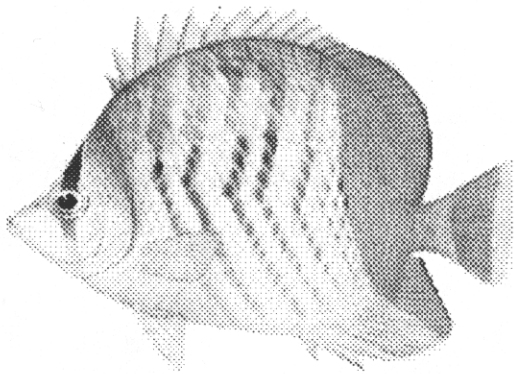
図6. 3. 2 カレイ科



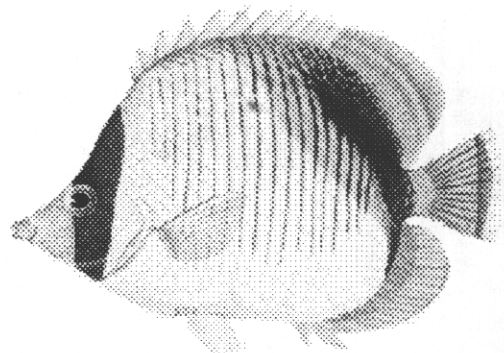
カガミチョウチョウオ



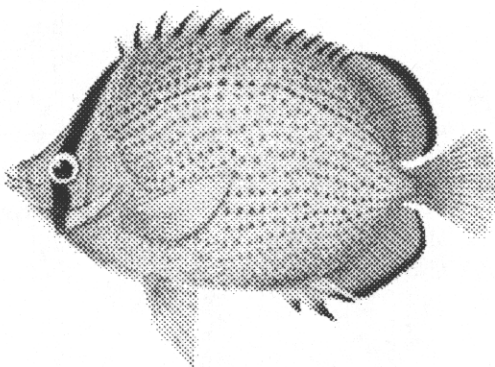
アミメチョウチョウオ



ベニオチョウチョウオ

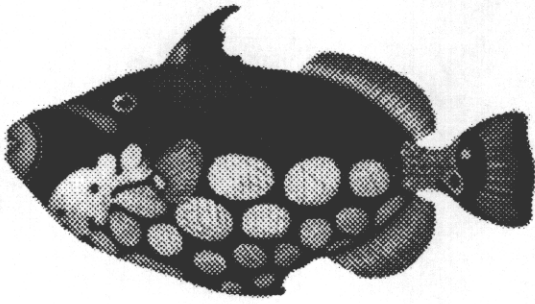


ニセフウライチョウチョウオ

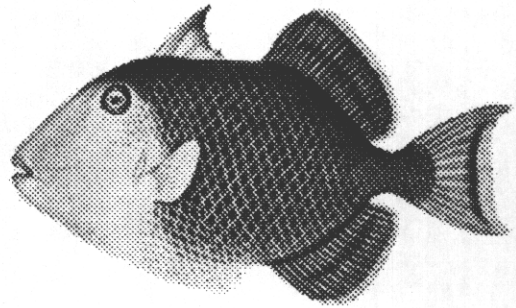


ゴマチョウチョウオ

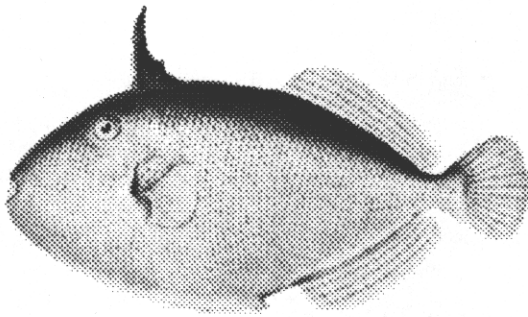
図6. 3. 3 チョウチョウオ科



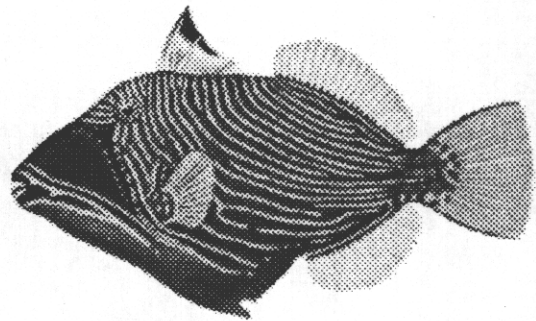
モンガラカワハギ



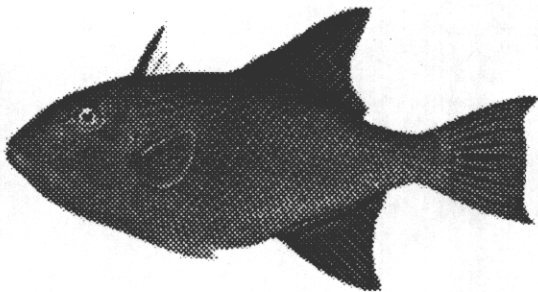
キベリモンガラ



クマドリ

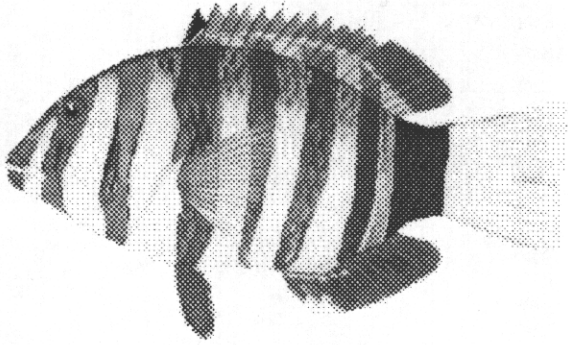


モラサメモンガラ

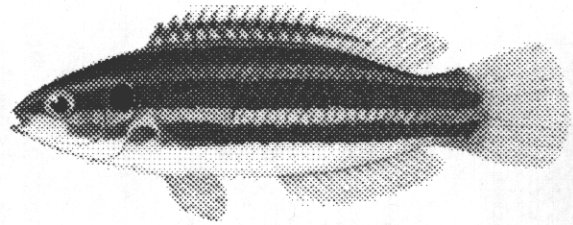


アミメモンガラ

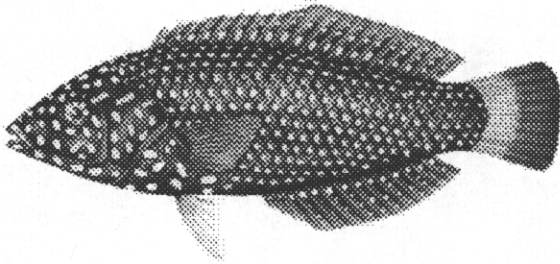
図6. 3. 4 モンガラカワハギ科



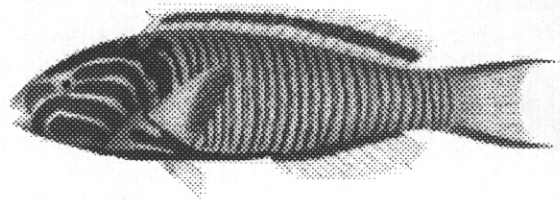
シチセンベラ



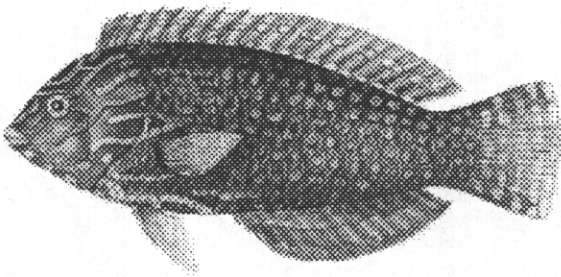
タヌキベラ



クロフスキベラ

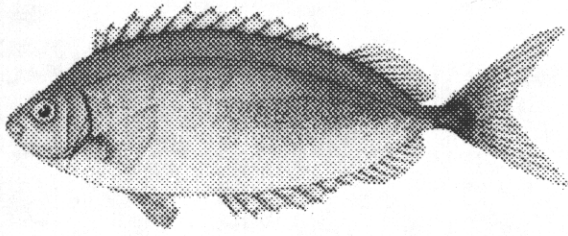


ヤマブキベラ

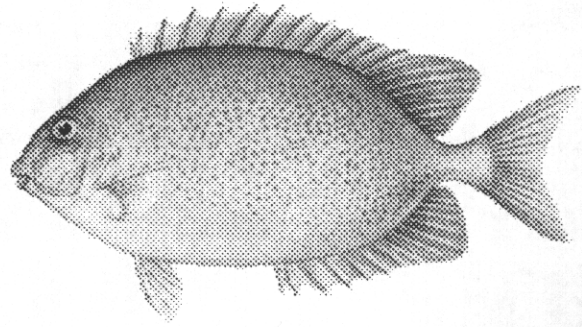


ノドグロベラ

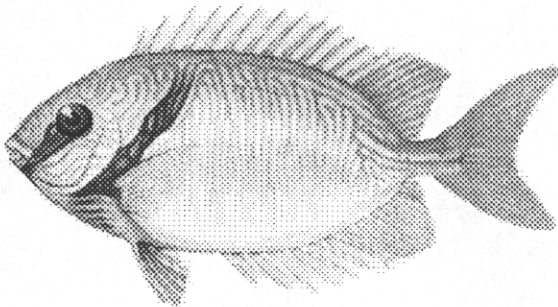
図6. 3. 5 ベラ科



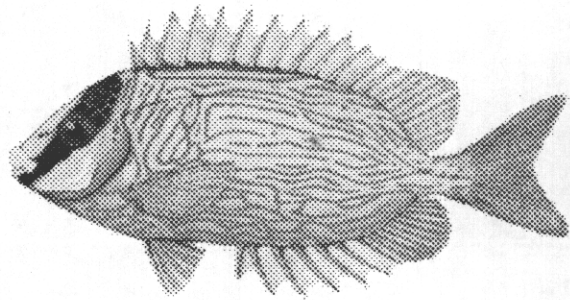
ハナアイゴ



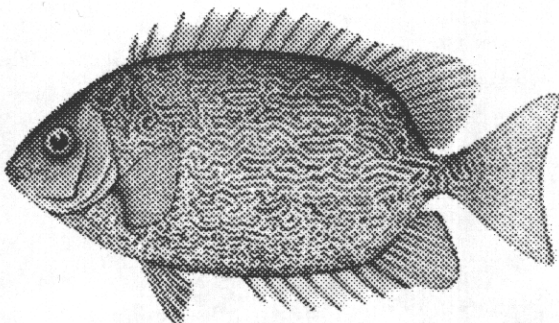
ゴマアイゴ



ヨコジマアイゴ

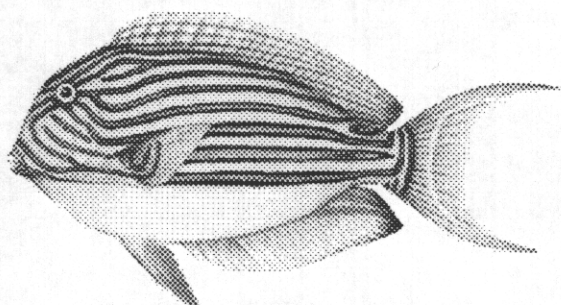


マジリアイゴ

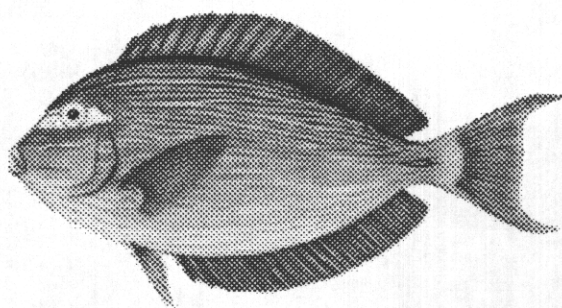


ムシクアイゴ

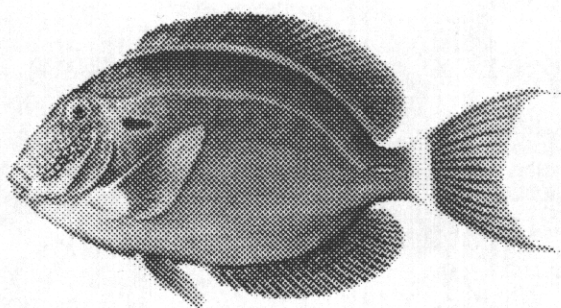
図 6. 3. 6 アイゴ科



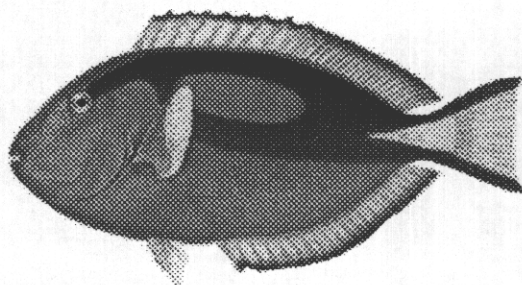
ニジハギ



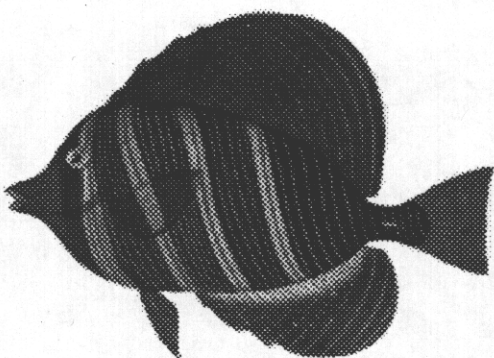
ヒラニザ



イレズミニザ

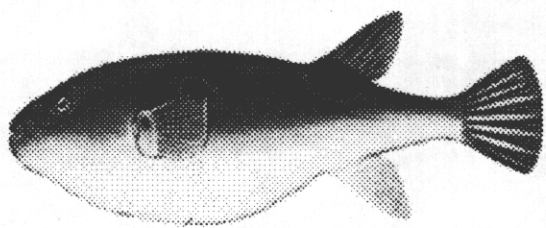


ナンヨウハギ

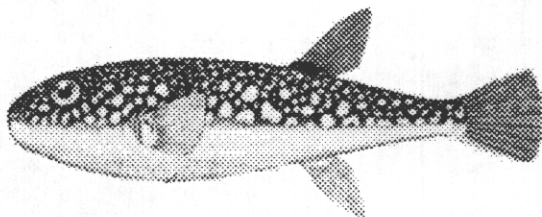


ヒレナガハギ

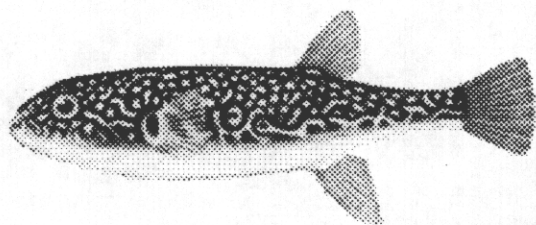
図6. 3. 7 ニザダイ科



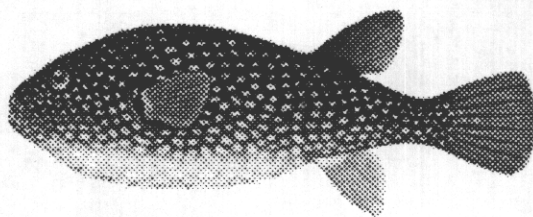
メフグ



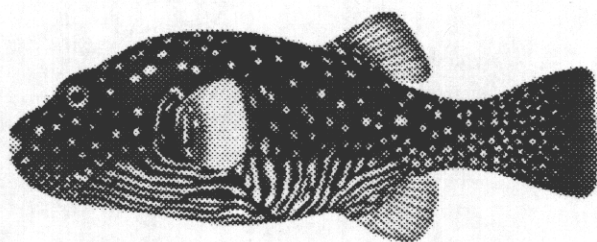
コモンフグ



ムシフグ

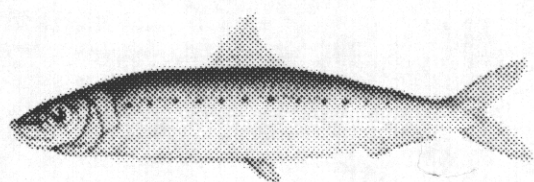


ホシフグ

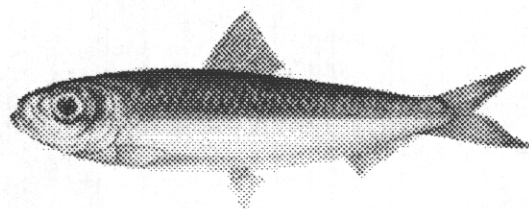


サザナミフグ

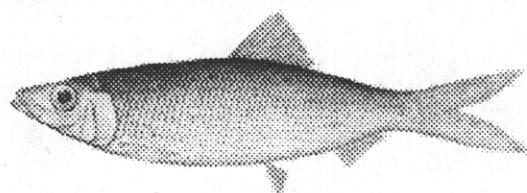
図6. 3. 8 マフグ科



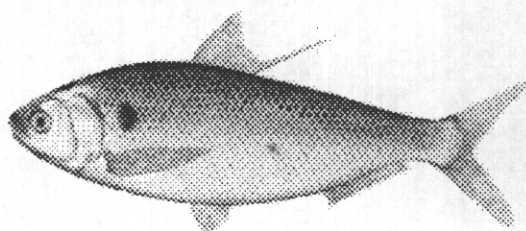
ホシヤマトミズン



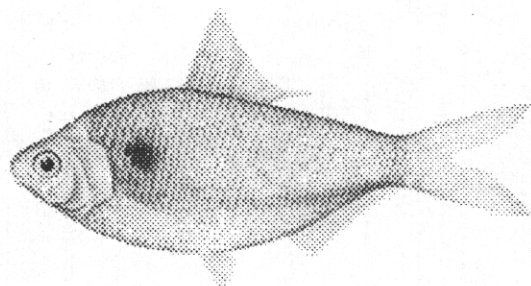
ミズン



ギンイワシ

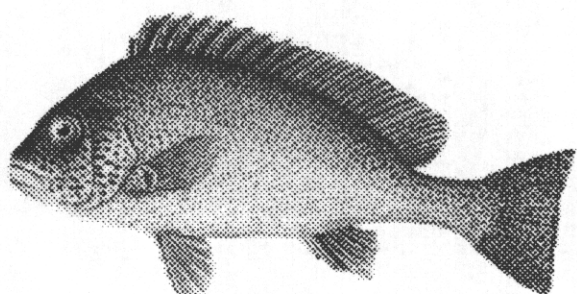


ドロクイ

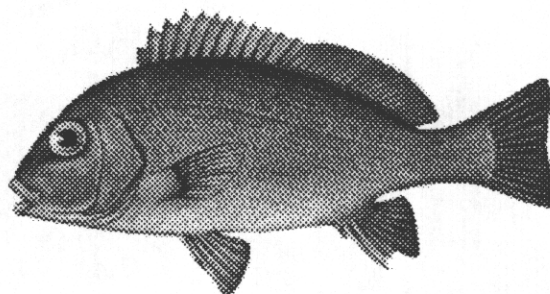


リュウキュウドロクイ

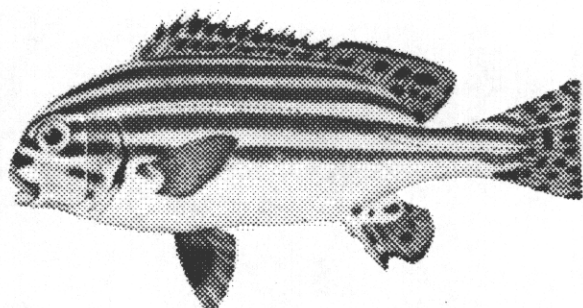
図6. 3. 9 ニシン科



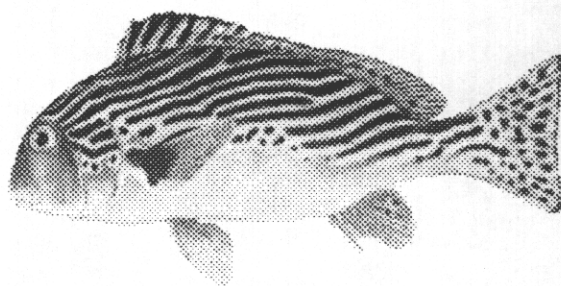
クロダイ



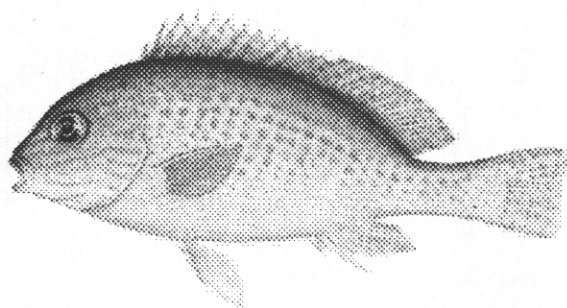
ナンヨウダイ



ムスジコショウダイ



アヤコショウダイ



オシャレコショウダイ

図6. 3. 10 イサキ科

6. 3. 2 輪郭線特徴の抽出

本実験における輪郭線特徴の抽出手順を説明する。

輪郭線特徴である P 型記述子のパワースペクトルを抽出するため、前処理としてグレイレベル 256 階調の画像を 2 値化し、エッジを検出する。この結果、画像は閉曲線となる。求められた閉曲線に対し、式(2.4.1)のアルゴリズムを適用し等辺多角形近似する。ここで、精度の高い P 型記述子のパワースペクトルを求めるため、十分に等辺多角形近似しておくことと、高速フーリエ変換を用いてパワースペクトルを求めるために頂点の数が 2 のべき乗となることにより、等辺多角形の頂点の数を 256 個としている。

次に、求められた等辺多角形を P 型記述子で記述し、そのパワースペクトルを求める。

以上の手続きによって P 型記述子のパワースペクトルは求められる。この中で、スペクトルの高周波成分には原画像の取り込み・作成の際に生じた画素単位の丸め誤差やノイズが含まれている。従って画像の大域的な情報を持つスペクトルの低周波成分をより重視し、周波数が低い順に 50 個の成分を取り出し、これらの成分を輪郭線の特徴とした。図 6.3.12 に輪郭線特徴の抽出フローを示し、図 6.3.11 に各過程での処理結果の例を示す。更に図 6.3.13 にパワースペクトルの例を示す（横軸の中央部が低周波成分である）。抽出された輪郭線特徴は特徴データベースに蓄えられる。

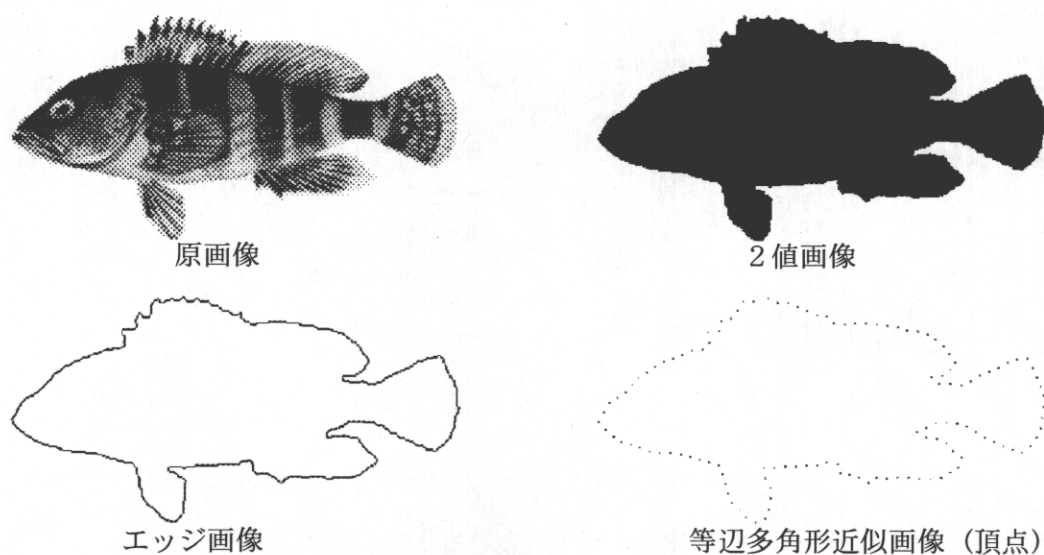


図 6. 3. 1 1 輪郭線特徴抽出の処理例

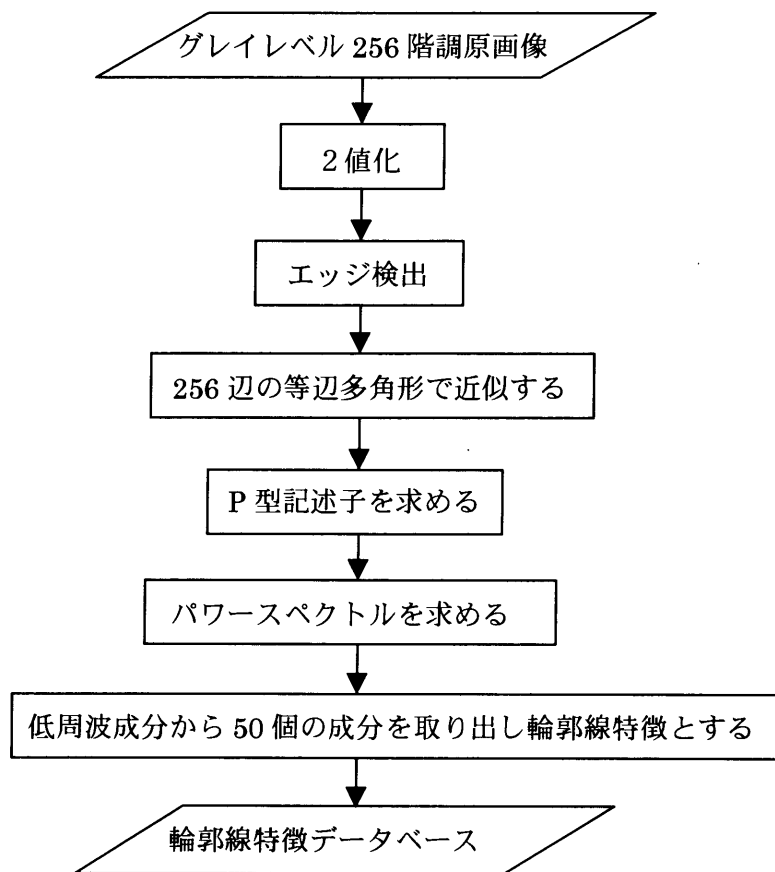


図 6. 3. 1 2 輪郭線特徴抽出のフロー

スペクトル正規化値

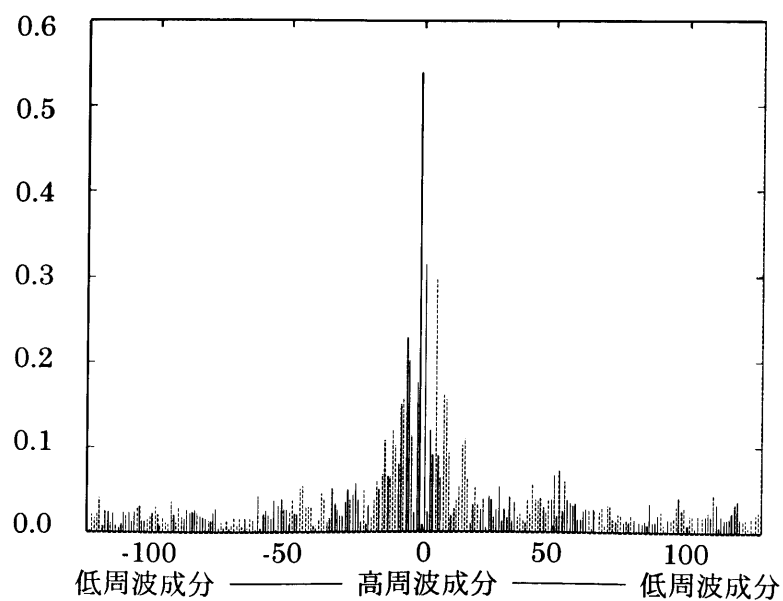


図 6. 3. 1 3 パワースペクトル

6. 3. 3 テクスチャ特徴の抽出

次に、テクスチャ特徴の抽出手順を説明する。

識別の対象となる画像において、必ずしもその画像全体に一様にテクスチャが存在することはない。例えば本実験で対象としている魚画像では、腹部と背の部分、また頭部と尾の部分では大きく模様が異なっている。特に腹部は白っぽく、殆ど模様がなない魚もある。このような領域はその魚のテクスチャ特徴となり得ない。従って、対象画像の特徴が十分取り出せるように、抽出すべき領域を事前に決めておく必要がある。またその位置決め手順は汎用性が高いほどよい。本実験では、魚画像を対象としているが、目視による判断から、魚画像の中心部であれば比較的安定してテクスチャ特徴を抽出できると考えた。ここで、6. 3. 2で求めた等辺多角形を利用して、等辺多角形の重心をテクスチャ特徴の抽出領域の中心としている。また、領域のサイズは魚画像の面積の $1/12$ の正方形とした。図 6.3.14 にテクスチャ特徴の抽出領域の例を示す。

本実験で用いるテクスチャ特徴は、濃度共起行列の要素である。濃度共起行列は対称行列であるため、テクスチャ特徴となる行列の要素の数は

$$\frac{1}{2}g(g+1) \quad (6. 3. 1)$$

となる。但し、 g は階調数である。魚の原画像は256階調であるが、このままでは情報量が多すぎるばかりではなく、各々の要素の値が小さくなるため認識能力が低下することが考えられる。そのため、濃度共起行列の要素を取り出す前に画像を再量子化して8階調に変換する。以前の実験において、4～16階調程度の画像で高い認識正解率を得ていることから[31][77]、8階調に再量子化しても十分な認識能力を得ることができると考えられる。画像を8階調に変換した結果、濃度共起行列の要素の数は $8 \times (8 + 1) / 2 = 36$ となる。図 6.3.15 に抽出領域における256階調原画像と8階調再量子化画像の例を示す。

このようにして画像変換と抽出領域の位置決めを行い、 0° 、 45° 、 90° 、 135° の角度に関して濃度共起行列を求める。更に求められた4個の濃度共起行列について式(4.7.2)、式(4.7.3)により合成し正規化する。なお、行列を求めるために用いる画素対の距離は $d = 1$ とする。

以上に説明した手続きで求められた行列の要素36個を取り出し、これをテクス

チャ特徴として、特徴データベースに蓄える。図 6.3.17 にテクスチャ特徴の抽出フローを示し、図 6.3.16 に濃度共起行列の要素のヒストグラム例を示す。

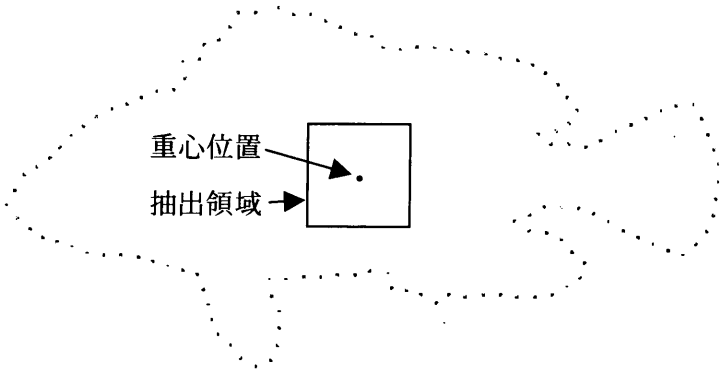
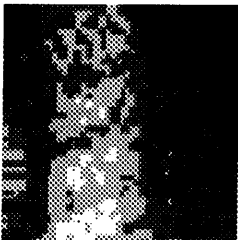


図 6. 3. 1 4 テクスチャ特徴の抽出領域の例



256 階調原画像



8 階調再量子化画像

図 6. 3. 1 5 原画像と再量子化画像の例

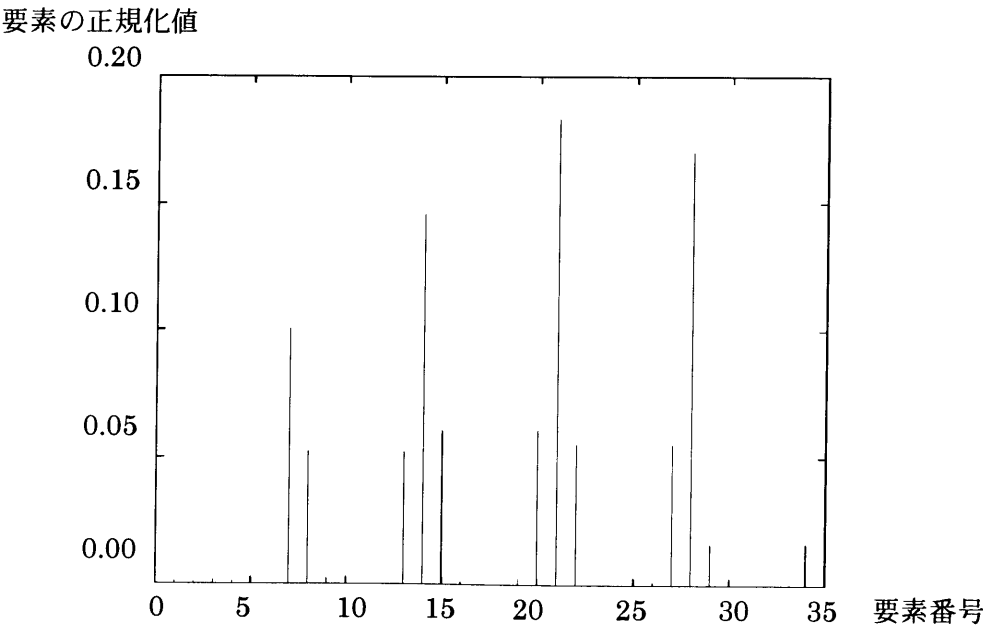


図 6. 3. 1 6 濃度共起行列の要素データ例

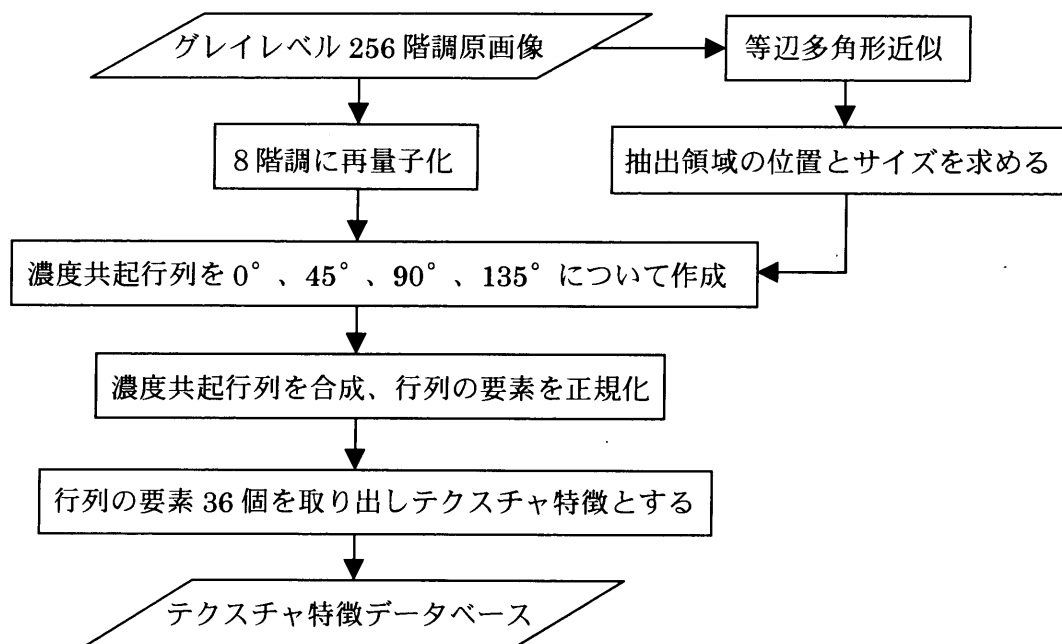


図 6. 3. 17 テクスチャ特徴の抽出フロー

6. 4 特徴の統合と識別実験

前述したように本実験では、輪郭線特徴として P 型記述子のパワースペクトルを用い、テクスチャ特徴として濃度共起行列の要素を用いる。更に両者の特徴を統合し魚画像を識別する手段として、3 層ニューラルネットワークを用いる。以下に 3 層ニューラルネットワークによる特徴の統合方法と、特徴を統合して魚画像を識別する実験について説明する。

6. 4. 1 3 層ニューラルネットワークによる特徴の統合

輪郭線特徴及びテクスチャ特徴を統合して画像を識別するために用いるニューラルネットワークは、入力層、出力層、1つの隠れ層からなる 3 層ニューラルネットワークである。図 6.4.1 で示されるように、輪郭線特徴とテクスチャ特徴が並列にネットワークに入力され、出力された結果から画像が識別される。ニューラルネットワークの学習則として、バックプロパゲーション法を用いる。ネットワークの各

ユニットの出力関数としてはシグモイド関数

$$f(x) = \frac{1}{1 + \exp(-\beta x)} \quad (6.4.1)$$

を用いる。本実験では、シグモイド関数の係数 β を学習の回数に応じて漸次大きくしていく。この方法はアニーリング法（焼きなまし法）と呼ばれており、これによりネットワークの状態を最小点に収束させることができる。

本実験で用いるネットワークの入力層のユニット数は、入力される特徴の数であり $50 + 32 = 82$ 個である。出力層のユニット数は、分類すべき魚の種類であり 50 個である。隠れ層については、現在までにユニット数の最適値の決定法が確立していないため、40、50、60、70、80 と 5 種類のユニット数を選択してネットワークを構成している。従って、5 種類の隠れ層のユニット数が異なるネットワークを用いて識別実験を行っている。

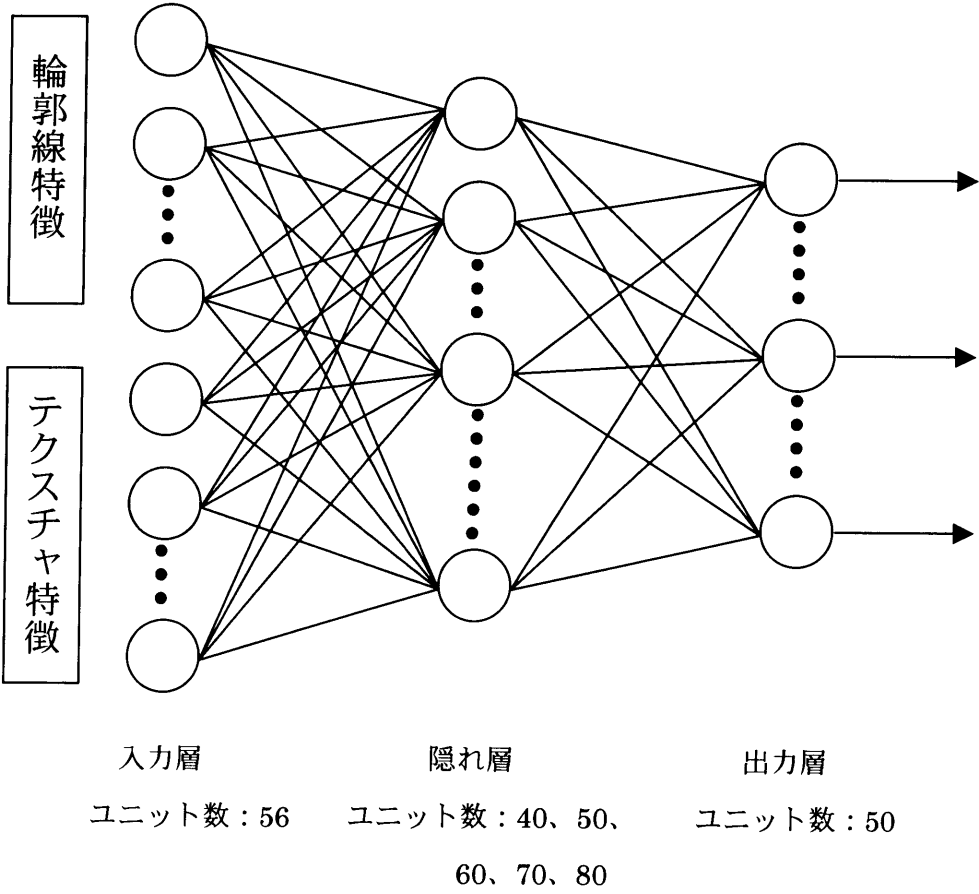


図 6. 4. 1 実験に用いるニューラルネットワーク

6. 4. 2 魚画像識別実験

前述したように、輪郭線特徴としてP型記述子のパワースペクトルを低周波成分から 50 個、テクスチャ特徴として濃度共起行列の要素を 36 個、合計 86 個を魚画像を表す特徴として3層ニューラルネットワークに入力して2種類の特徴を統合し識別実験を行う。但し識別実験を行う前に、図 6.2.3 で示した手順に従ってバックプロパゲーションによる学習を行いネットワークの結合の重みを決定付ける。以下に学習の手続きについて述べる。

対象となるニューラルネットワークは5種類あり、隠れ層のユニット数が 40、50、60、70、80 と異なるネットワークである。これらのネットワークの学習で用いる各々の係数は次の通りである。重み修正、しきい値修正の初期値はそれぞれ 0.4、0.3 としている。また、重み係数の初期値は-0.1 から 0.1、しきい値の初期値は-0.05 から 0.05 の範囲でランダムに与えている。シグモイド関数の係数 β は学習の回数に応じて 0.5 から 9.5 まで漸次増加させている。

6. 3. 1で述べたように、全画像 800 枚の中から 400 枚を学習用の画像とした。これらの画像の特徴は予め求められ、データベースに蓄積されている。特徴データベースとそれぞれの画像に対応した教師信号を用いて学習を進めていく。400 枚の画像について順次、対応する特徴をネットワークに入力し、出力された結果と教師信号を比較し誤差を修正していく。400 枚の画像についてすべて、入力—出力—修正作業が終了した時点で、学習が1回終了する。本実験では学習回数を 5000 回とした。

更に、本実験では輪郭線特徴とテクスチャ特徴の統合の有効性を実証するため、両者の特徴を単独で用いる実験をも行う。そこで単独で特徴を用いる場合のニューラルネットワークも用意し学習を行った。ネットワークの構成は、入力層のユニット数が入力する特徴の数にあわせて変化する以外に違いはない。また学習の手続きも変わらない。

図 6.4.2 に、輪郭線特徴とテクスチャ特徴を統合する方法において、隠れ層のユニット数が 60 個で、重み、しきい値修正の加速係数がそれぞれ 0.4、0.3 であるネットワークを用いた場合の学習曲線を示す。また、図 6.4.3 から図 6.4.7 に隠れ層のユニット数が 60 で、重み、しきい値修正の加速係数をともに 0.1、0.3、0.5、0.7、

0.9 とした場合の学習曲線を示す。

表 6.4.1、表 6.4.2 に輪郭線情報及びテクスチャ情報を単独で用いた場合、表 6.4.3 に両者の情報を統合した場合の識別実験の結果を示す。この結果から情報を単独で用いたものに比べて統合して用いた方が、画像識別の正解率が格段に高いことが分かる。

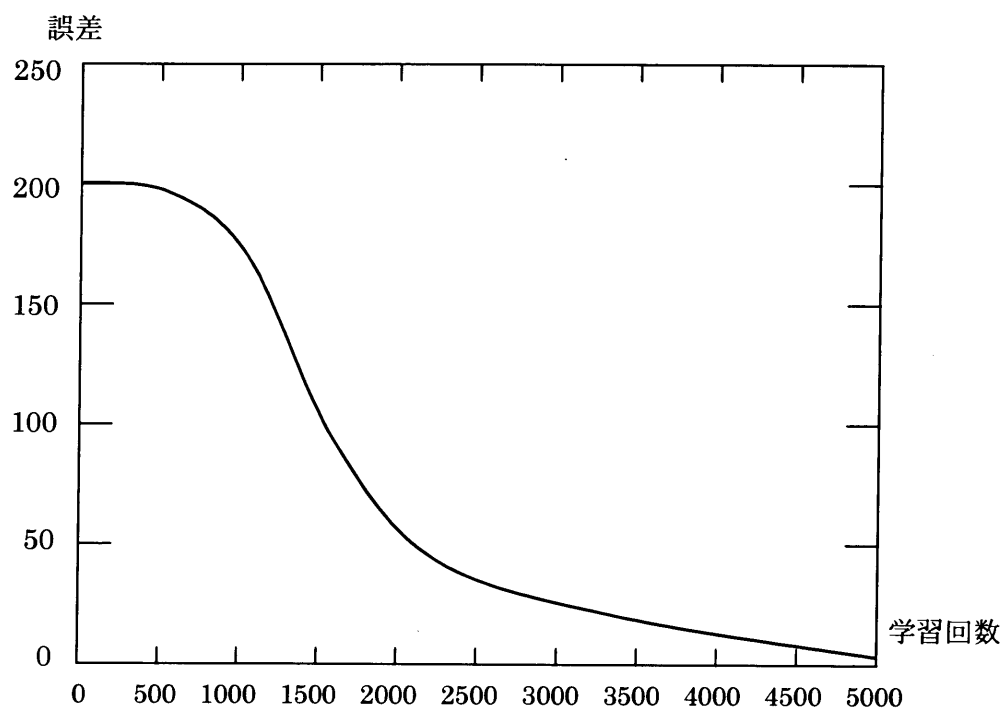


図 6. 4. 2 学習曲線（重み修正加速係数：0.4、しきい値修正加速係数：0.3）

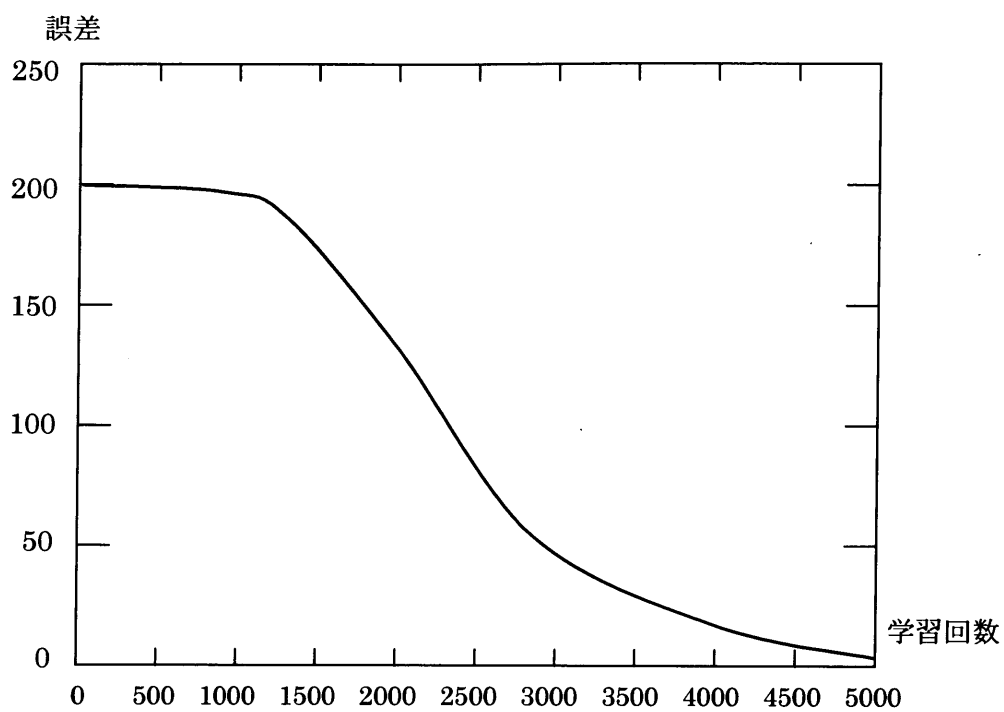


図5. 4. 3 学習曲線（重み修正加速係数：0.1、しきい値修正加速係数：0.1）

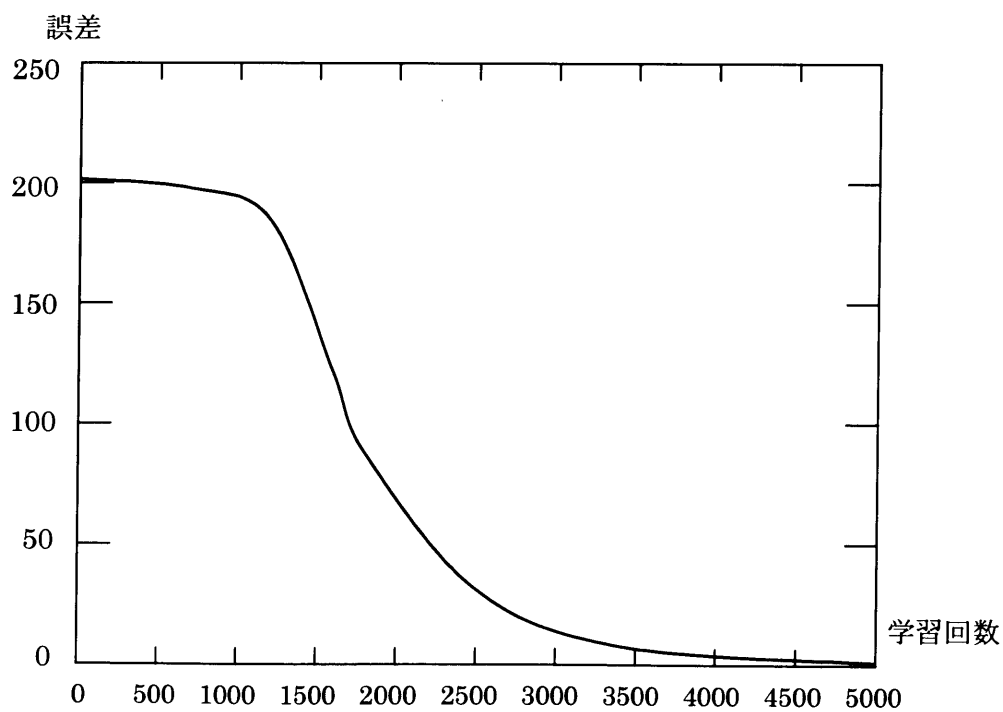


図6. 4. 4 学習曲線（重み修正加速係数：0.3、しきい値修正加速係数：0.3）

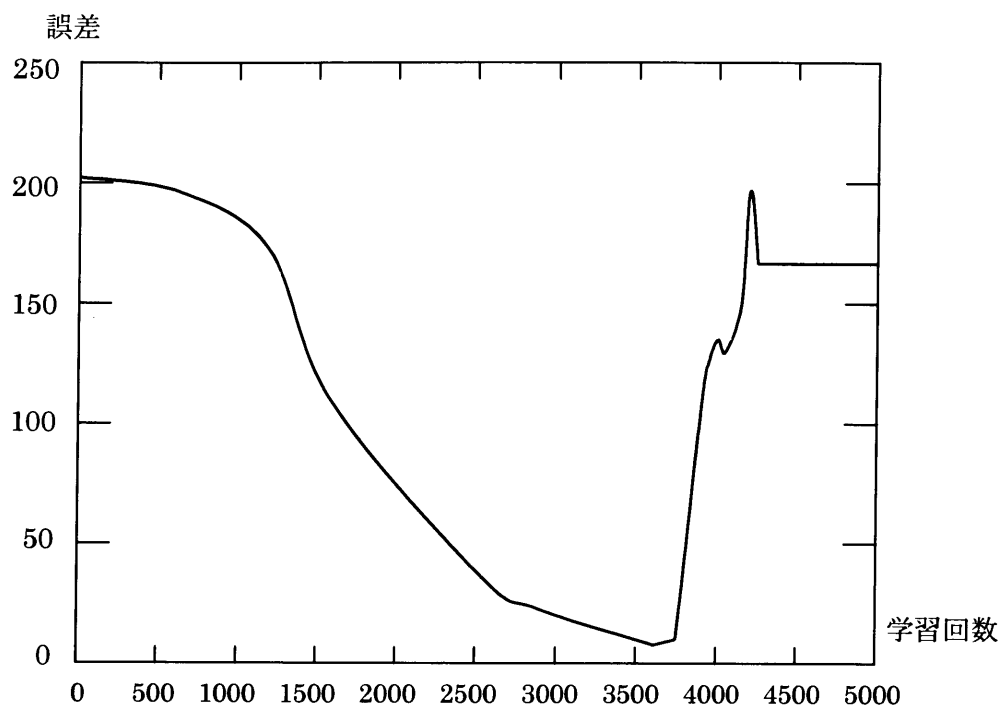


図 6. 4. 5 学習曲線（重み修正加速係数：0.5、しきい値修正加速係数：0.5）

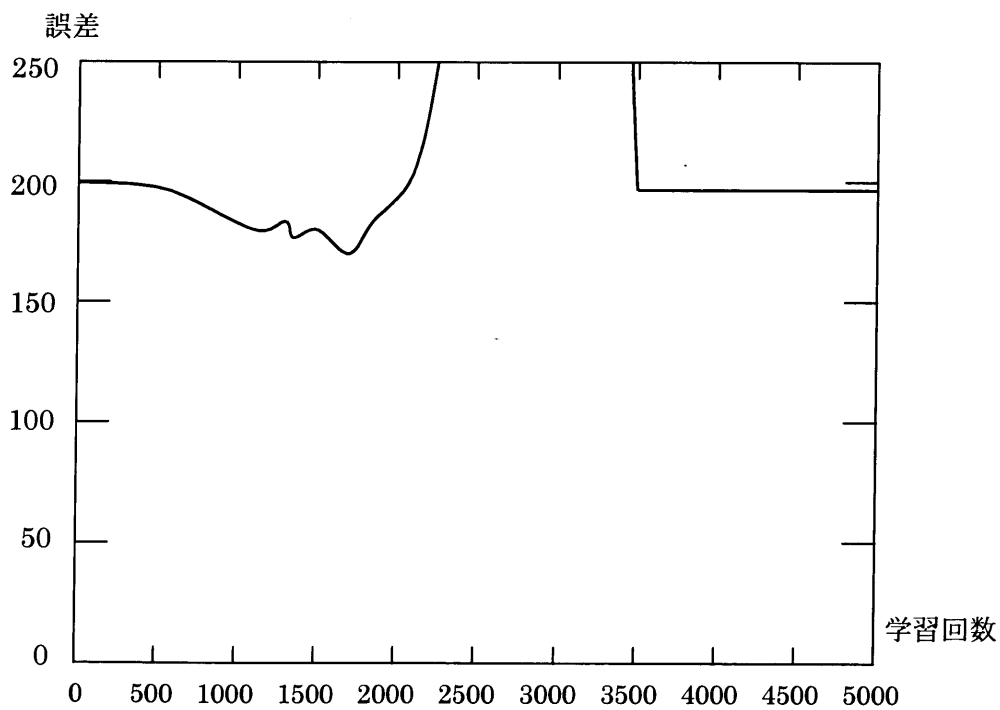


図 6. 4. 6 学習曲線（重み修正加速係数：0.7、しきい値修正加速係数：0.7）

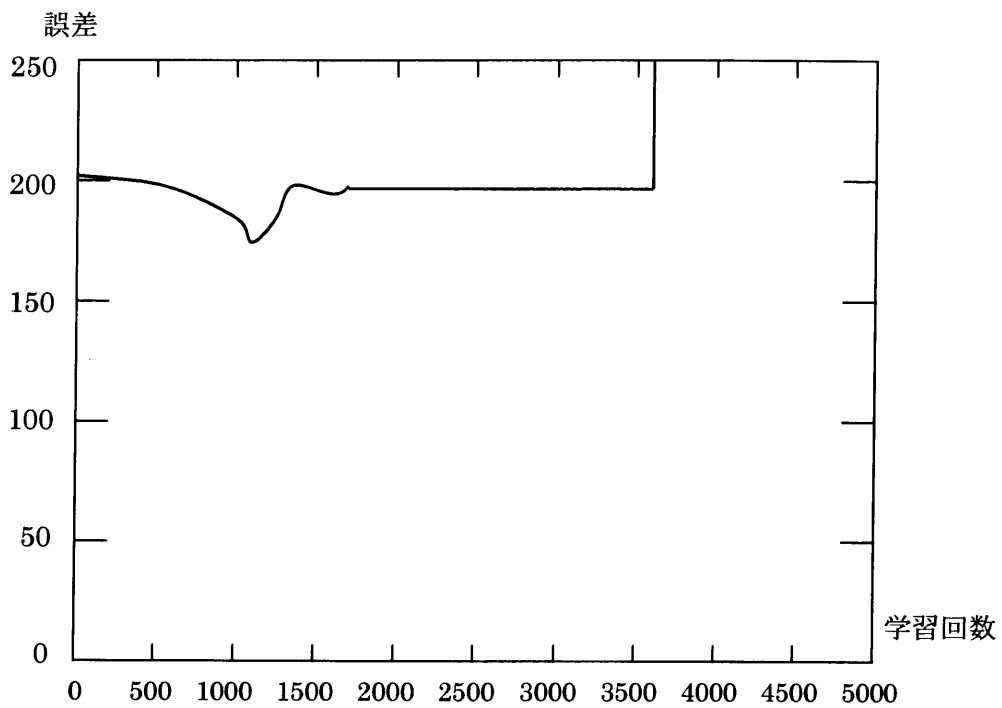


図 6. 4. 7 学習曲線（重み修正加速係数：0.9、しきい値修正加速係数：0.9）

隠れ層のユニット数	学習データの正解率	テストデータの正解率
4 0	8 6. 5 % (346／400)	7 2. 8 % (291／400)
5 0	9 1. 5 % (366／400)	7 7. 8 % (311／400)
6 0	9 1. 5 % (366／400)	7 9. 0 % (321／400)
7 0	9 1. 3 % (365／400)	8 2. 8 % (331／400)
8 0	9 0. 3 % (361／400)	8 2. 0 % (328／400)

表 6. 4. 1 輪郭線情報を単独で用いた場合の識別結果

隠れ層のユニット数	学習データの正解率	テストデータの正解率
4 0	5 9 . 8 % (239 / 400)	5 9 . 3 % (237 / 400)
5 0	6 5 . 5 % (262 / 400)	6 6 . 3 % (265 / 400)
6 0	6 2 . 5 % (249 / 400)	6 3 . 0 % (252 / 400)
7 0	6 0 . 8 % (243 / 400)	6 1 . 0 % (244 / 400)
8 0	5 5 . 8 % (223 / 400)	5 7 . 0 % (228 / 400)

表 6 . 4 . 2 テクスチャ情報を単独で用いた場合の識別結果

隠れ層のユニット数	学習データの正解率	テストデータの正解率
4 0	9 8 . 0 % (392 / 400)	9 4 . 5 % (378 / 400)
5 0	9 8 . 5 % (394 / 400)	9 5 . 3 % (381 / 400)
6 0	9 8 . 8 % (395 / 400)	9 7 . 3 % (389 / 400)
7 0	9 8 . 0 % (392 / 400)	9 5 . 7 % (383 / 400)
8 0	9 6 . 8 % (387 / 400)	9 4 . 5 % (378 / 400)

表 6 . 4 . 3 輪郭線情報とテクスチャ情報とを統合して用いた場合の識別結果

6. 5 まとめ

本実験では、輪郭線情報とテクスチャ情報を統合することの有効性を具体的な画像を用いて実証していった。

魚画像は 10 科各 5 種、合計 50 枚を選び出し、これらを元画像として、サイズ 4 種、回転 4 種の画像変換を行い最終的には 800 枚の画像を作成した。更に本実験では特徴の統合と識別に学習機能を持つ 3 層ニューラルネットを用いるため、800 枚の画像をランダムに 400 枚ずつに分け、一方をネットワークの学習に用い、もう一方を識別テストに用いた。

輪郭線特徴では、E P A C に基づく P 型記述子のパワースペクトルを求め、その中から低周波成分 50 個を取り出した。

テクスチャ特徴では、魚画像を 8 階調に再量子化した後、魚画像の等辺多角形近似図形の重心を中心として、魚画像領域の $1/12$ の面積を持つ正方形を設置してその中から濃度共起行列を求め、その行列の要素 36 個を取り出した。

このようにして求めた特徴を統合し識別するために、ニューラルネットワークを用いた。ネットワークは、隠れ層のユニット数を 40、50、60、70、80 としたもの 5 種類を用意し、それぞれ 400 枚の画像を用いて学習を 5000 回行った。また隠れ層のユニット数 60 のネットワークについては、重み。しきい値修正の加速係数とともに 0.1、0.3、0.5、0.7、0.9 と変化させ、それぞれの学習の収束状態を調べた。

このようにして、輪郭線特徴とテクスチャ特徴を個別に用いた場合と統合して用いた場合に関し、それぞれのネットワークを学習させた後魚画像の識別実験を行った。以下に、本実験の結論を示す。

- ①テストデータの認識に関し、輪郭線特徴を単独で用いた実験の正解率が最高で 82.8%、テクスチャ特徴を単独で用いた実験の正解率が最高で 66.3%であったのに対し、両者を統合した実験では最高 97.3%の正解率を得た。このことから、輪郭線とテクスチャという、異なる 2 種類の特徴を統合することが、画像認識の正解率の向上に寄与しているといえる。
- ②特徴を単独で用いる場合、テクスチャ特徴を用いる場合よりも輪郭線特徴を用いる場合のほうが正解率が高い。これは、認識画像の中に元画像のサイズ変換画像が含まれており、サイズ変換に対しては輪郭線特徴である P 型記述子のパ

ワースペクトルは影響を受けないが、テクスチャ特徴である濃度共起行列の要素は影響を受け、正解率が低下したものと考えられる。

③今回の実験においては、隠れ層のユニット数は輪郭線特徴のみを用いた場合で 70 個、テクスチャ特徴のみを用いた場合で 50 個、両者を統合した場合で 60 個の場合が最も正解率が高かった。隠れ層のユニット数が小さすぎるとネットワークが情報を記憶できず、逆に大きすぎると不必要な情報までも記憶してしまうように考えられる。

④重み修正としきい値修正の学習係数は 0.3 または 0.4 の場合が良い。加速係数が小さすぎると学習の収束が遅く、逆に大きすぎると学習がローカルミニマムから抜け出さなくなり、収束しない。

以上のことから本実験における魚画像の識別において、輪郭線特徴とテクスチャ特徴を統合が極めて有効であったことを実証できた。

第7章

結論

本論文は、画像の性質を表す有力な特徴である輪郭線情報とテクスチャ情報に着目し、両者を統合して画像識別に用いることの有効性について研究した結果をまとめたものである。本研究によって得られた研究成果の概要は以下の通りである。

第1章では、本研究の背景と扱っている問題を示し、あわせて論文の概要について述べた。

第2章では、輪郭線情報の種々の記述方法、主に輪郭線を等辺多角形で近似して求める記述方法について概説した。特に画像の平行移動、回転、拡大・縮小に対して不変な特徴である、EPACに基づくP型記述子から選ばれるパワー・スペクトルが最も有用であることを示し、その詳細を述べた。

第3章では、ロバストエネルギー最小化による輪郭線保存平滑化法について詳述した。輪郭線情報による認識において、大域的な形状の情報の方が局所的な情報よりも重要な場合が多い。また、輪郭線の等辺多角形近似において、平滑化された輪郭線の方がアルゴリズムが収束しやすい。そこで、大域的な情報を損なわず、ノイズ等の不要な局所情報を除き輪郭線を平滑化する方法として、ロバストエネルギー最小化による方法を提案した。更に最小2乗法、メディアン法、M推定法、ロバストエネルギー最小化法について信号の平滑化実験を行い。平滑化の程度および実用性からロバストエネルギー最小化による方法の有効性を実証した。

第4章では、画像の模様情報であるテクスチャ情報の種々の記述方法について、構造的な記述方法と統計的な記述方法に大別して述べた。特に画像の平行移動、45°単位の回転の影響を受けにくい濃度共起行列の要素が最も有用なテクスチャ特徴で

あることを示し、その詳細を述べた。

第5章では、輪郭線情報とテクスチャの情報とを統合し、類別を行うシステムとして用いるニューラルネットワークについて、基本的な事柄を含め、これによる情報の統合方法を述べた。

第6章では、輪郭線情報とテクスチャ情報を統合の有効性を、具体的な実験を行って実証した。対象とした画像は50種類の魚画像である。50種類の魚画像は回転・サイズ変換を行い、1枚の元画像について16通りのバリエーションを作成した。合計800枚の画像のうち400枚をニューラルネットワークの学習データとし、残りの400枚を識別テストデータとした。それぞれの魚画像の輪郭線は等辺多角形で近似され、P型記述子とそのパワー・スペクトルを求めた後、そのなかの低周波成分を輪郭線情報として取り出した。テクスチャ情報の抽出では、画像を8階調に再量子化して濃度共起行列を求め、行列の要素をテクスチャ情報として取り出した。両者の情報は3層ニューラルネットワークで統合される。400枚の学習データ画像により学習を行った後、識別データ画像により識別実験を行った。また、輪郭線とテクスチャの情報について個別に用いた場合での同様の実験を行なった。それらの結果から、両者の情報を単独で用いるよりも統合した方が識別の正解率が格段に高くなり、従って両者の情報の統合が有効であることが実証された。

現在までに、画像処理・画像認識は製造業の生産現場を中心として産業応用が進められたきた。多くの分野で成功を収めてきたとはいえ、産業用に開発された視覚は依然、人間の視覚に遠く及ばない。なぜならば、これらの視覚の多くは対象物や環境を限定し、それに応じてアルゴリズムを工夫した上で実現されているからである。一方、機械による視覚は疲労することなく、人間の視覚で起こる心理作用による判断の変動が発生しないことから、今後ますます産業への応用が進められていくであろう。また、社会構造の変化から交通や福祉など、これまでほとんど実現されなかった分野へも、機械による視覚が利用されていくであろう。従って、機械による視覚に対する要求は更に高度になっていくことが想像される。そのため、本研究で行ったような、複数の異種の情報を統合して行う方法の開発が真に望まれる。

本研究で実施した事柄が、今後の画像処理・画像認識の産業界への応用に関して何らかの形で貢献できれば幸いである。

謝辞

本研究を遂行し、また、まとめるにあたり、多くのかたがたに後指導、御助言および御鞭撻をいただきました。深く感謝いたします。

九州芸術工科大学芸術工学部画像設計学科瀧山龍三教授には、本研究を進めるにあたり、終始御指導いただきました。

本論文の作成にあたっては、九州芸術工科大学浦浜喜一教授、長島健次教授には懇切な御指導と御鞭撻を賜りました。

九州芸術工科大学坂本博康講師、小野直樹助手には、有効適切な御意見と御助言を賜りました。

九州松下電器株式会社F A研究所北原明所長、楠木弘典部長には、本研究の機会を与えていただきましたことを厚く御礼申し上げます。

御指導、ご協力くださった方々に対し、あらためて深謝いたします。

参考文献

- [1]会津, 中川, 高木:” 色空間分布モデルに基づく物体領域抽出”, 信学技報, PRU92-146, pp.17-23, 1993.
- [2]碓井, 田所:” $L^*a^*b^*$ 均等色空間における特徴を用いた文字重畳カラー画像からの文字抽出”, 信学技報, IE94-107, pp.85-90, 1994.
- [3] 福元, 福島:” 色情報に基づく道路情景画像の領域分割と理解”, 信学技報, IE94-33, pp.41-48, 1994.
- [4]原, 大西, 杉江:” カラー画像からの物体抽出について”, 信学技報, PRU94-150, pp.35-42, 1995.
- [5]富永, 今村:” 映り込みを含むカラー画像の解析”, 信学技法, PRU92-145, pp.9-16, 1993.
- [6]G.H.Granlund: "Fourier Preprocessing for Hand Print Character Recognition", IEEE Trans. Computers, vol.C-21, no.1, pp.195-201, 1972.
- [7]安居院, 中嶋, 伊藤” 閉曲線情報処理手法による湖の識別”, テレビジョン学会誌, no.33, pp.210-215, 1979.
- [8]上坂:” 閉曲線にも適用できる新しいフーリエ記述子”, 電子通信学会論文誌(A), J67-A, pp.167-173, 1984.

- [9]瀧山, 小野, 外波: ” 等辺多角形近似曲率関数－E P A C－を用いた横顔シルエットの同定”, 信学技報, MBE87-83, 1988.
- [10]矢野, 小野, 瀧山: ” 閉曲線の等辺多角形近似とその前処理について”, 電子情報通信学会技術報告, IE90-86, 1990.
- [11]矢野, 小野, 瀧山: ” 閉曲線の等辺多角形近似のための前処理について”, 電子情報通信学会春季全国大会, vol.7, pp.280, 1991.
- [12]矢野, 小野, 瀧山: ” 閉曲線の等辺多角形近似アルゴリズム”, 電気関係学会九州支部連合会大会論文集, 1991.
- [13]Y.Shirai: "Analyzing Intensity Arrays Using Knowledge about Scence", The Psychology of Computer Vision, pp.96, Winston, P.H.ed.Mcgraw-Hill, 1975.
- [14]安居院, 中嶋, 岩田: ” 輪郭スペクトル情報を用いた閉領域同定に関する研究”, 信学技報, PRL80-100, 1980.
- [15]I.Pitas, N.Venetsanopoulos: "Order Statistics in Digital Signal Processing", IEEE Proc., vol.80, pp.1893-1921, 1992.
- [16]V.Torre, T.A.Poggi: "On Edge Detection", IEEE T-PAMI, vol.6, pp.721-741, 1986.
- [17]C.Koch, J.Marroquin, A.Yuille: "Analog Neuronal Networks in Early Vision", Proc. Nat. Acad. Sci. USA, vol.83, pp.4263-4267, 1986.
- [18]D.Geiger, A.Yuille: "A Common Framework for Image Segmentation", Int.J.Comput. Vision, vol.6, pp.227-243, 1991.
- [19]S.Z.Li: "Markov Random Field Modeling in Computer Vision", Springer-

Verlag, 1995.

- [20]矢野, 岡田, 松永, 浦浜: "ロバストエネルギー最小化によるエッジ保存平滑化", 電子情報通信学会論文誌, vol.J80-D-II, no.1, pp.386-389, 1997.
- [21]J.S.Weszka, C.R.Dyer & A.Rosenfeld: "A Compratitive Study of Texture Measuresfor Terrain Classification", IEEE Trans. on SMC, vol.SMC-6, no.4, pp.269-285, 1976.
- [22]富田, 白井, 辻: "テクスチャの解析", 情報処理, vol.19, no.6, pp.610-621, 1973.
- [23]G.Eichmann and T.Kasparis: "Pattern Classification Using A Linear Associative Memory", Pattern Recognition. vol.22, pp.733-740, 1989.
- [24]R.M.Haralick: "Statistical and Structural Approach to Texture", Proc. IEEE, vol.67, no.5, pp.786-804, 1979.
- [25]M.M.Galloway: "Texture Analysis Using Gray Level Run Length", Comput., Graphics Image Processing, vol.4, pp.172-199, 1975.
- [26]K.Chehdi and Q.Liao:"A New Approach to the Improvement of Texture Modelisation", DKIS-EIS, pp.24, 1991.
- [27]金子: "フラクタル特徴とテクスチャ解析", 信学論 (D), vol.43, no.12, pp.1361-1369, 1988.
- [28]R.W.Conners and C.A.Harlow: "A Theoretical Comparison of Texture Algorithms", IEEE Trans., vol.PAMI-2, no.3, pp.204-222, 1980.
- [29]F.S.Cohen, Z.Fan and M.A.Patel: "Classification of Rotated and Scaled Textured Images Using Gaussian Markov Random Field Models", IEEE

Trans., vol.PAMI-13, pp.192-202, 1991.

[30]梅田：“PDP モデルによるテクスチャ画像の識別”，信学技報，PRU-12, 1990.

[31]矢野，小野，瀧山：“共起行列に基づくテクスチャ識別”，電子情報通信学会技術報告，PRU91-148, 1992.

[32]阿部 他：“原色魚類検索図鑑Ⅰ、Ⅱ”，北隆館，1989.

[33]C.T.Zahn and R.Z.Roskies: "Fourier Description for Plane Closed Curves", IEEE Trans. Computers, vol.C-21, pp.269-281, 1972.

[34]瀧山，羅，大石”輪郭線図形を用いた横顔の認識”，電気関係学会九州支部大会講演論文集，pp.590, 1991.

[35]瀧山，王：“等辺多角形近似とP型記述子による横顔の同定”，電気関係学会九州支部大会講演論文集，pp.769, 1992.

[36]瀧山，石田，新谷：“輪郭線情報による縄文土器画像の計測と処理について”，信学技報，PRU86-62, 1986.

[37]小野，瀧山：“デジタル輪郭線の等辺多角形近似について”，電子通信学会論文誌(A)，J67-A, pp.166-173, 1989.

[38]P.J.Rousseeuw, A.M.Leroy: "Robust Regression and Outlier Detection", Wilkey, 1987.

[39]P.W.Holland, R.E.Welsch: "Robust Regression Using Iteratively Reweighted Least-Squares", Commun. Stat., vol.6, pp.813-827, 1977.

[40]K.Urahama, T.Nagao: "Direct Analog Rank Filtering", IEEE Trans. Circuits

- Syst., I , vol.42, pp.385-388, 1995.
- [41]P.V.C.Hough: "Method and Means for Recognition Complex Patterns", U.S.Patent 3,069,654, 1962.
- [42]R.O.Duda and P.E.Hart: "Use of the Hough Transformation to Detect Lines and Curves in Pictures", Communs Ass., Comput., Mach., vol.15, pp.11-15, 1972.
- [43]J.Sklansky: "On the Hough Technique for Curve Detection", IEEE Trans., Comput., vol.C-24, pp.923-926, 1978.
- [44]矢野，小野，瀧山：” 非線型な拘束を持つ繰り返し法による楕円及び双曲線の検出”，電気関係学会九州支部連合会大会論文集, 1990.
- [45]R.S.Wallace: "A Modified Hough Transform for Lines", IEEE CVPR Conf., San Francisco, pp.665-667, 1985.
- [46]興水：” Hough 変換に関する最近の研究動向”，コンピュータビジョン， no.51, 1987.
- [47]興水：” Hough 変換に関する最近の研究動向(2)”，信学技報， PRU-12, 1991.
- [48]小野，瀧山，矢野：” 直線検出の一手法”，1990 年電子情報通信学会春季全国大会， vol.7, pp.277, 1990.
- [49]高木他 監修：” 画像解析ハンドブック”，東京大学出版会，1991.
- [50]J.L.Muerle: "Some Thoughts on Texture Discrimination by Computer", Picture Processing and Psychopictorics, Academic Press, 1970.

- [51]R.M.Haralick, K.Shanmugam and I.Dinstein: "Textural Features for Image Classification", IEEE Trans., vol.SMC-3, no.6, pp.610-621, 1973.
- [52]C.Sun and W.G.We: "Neighboring Gray Level Dependence Matrix", Comput., Vision, Graphics, Image Processing, vol.23, pp.341-352, 1982.
- [53]D.C.He, L.Wang and J.Guilbert: "Texture Feature Extraction", Pattern Recognition Letters, no.6, pp.269-273, 1987.
- [54]C.C.Gotlieb, H.E.Kreyszig: "Texture Descriptors Based on Co-occurrence Matrices", Comput., Vision. Graphics, Image Processing, vol.51, pp.70-86, 1990.
- [55]A.Visa: "Identification of Stochastic Textures with Multiresolution Features and Self-Organizing Maps", 10th ICPR-B, pp.518-522, 1990.
- [56]W.K.Pratt: "Digital Image Processing", Wiley-Interscience, 1978.
- [57]安居院, 中嶋” コンピュータ画像処理”, 産報出版, pp.126-136, 1979.
- [58]加藤 : ” 画像符号化アルゴリズム (IV)ーエントロピー符号化”, テレビジョン学会誌, vol.43, no.12, pp.1361-1369, 1989.
- [59]金子 : ” フラクタルとその応用”, テレビジョン学会誌, vol.41, no.4, pp.359-366, 1987.
- [60]B.B.Mandelbrot: "The Fractal Geometry of Nature", Freeman, 1982.
- [61]F.Rosenblatt: "Principles of Neurodynamics", Spartan, 1961.
- [62]F.Rosenblatt: "The Perceptron: A Probabilistic Model for Information Storage

- and Organization in the Brain", Psychological Review, no.65, pp.386-408, 1958.
- [63]S.Amari: "Theory of Adaptive Pattern Classifiers", IEEE Trans., EC16, pp.299-307, 1967.
- [64]D.E.Rumelhart, G.E.Hinton, R.J.Williams: "Learning Representaions by Back-Propagation Errors", Nature, vol.323, pp.533-536, 1986.
- [65]J.J.Hopfield: "Neural Network and Physical Systems with Emergent Collective Computational Abilities", Proc. Natl. Acat. Sci. USA, vol.79, 1982.
- [66]D.H.Ackley, G.E.Hinton, T.J.Sejnowsky: "A Learning Algorithm for Boltzmann Machine", Cognitive Science, vol.9-1, pp.147-169, 1985.
- [67]G.E.Hinton and T.J.Sejnowsky: "Learning and Relearning in Boltzmann Machines", In Paralle Distributed Processing, vol.1, Cambridge,MA:MIT Press, 1986.
- [68]甘利：” 神経回路網の数理”，産業図書，1978.
- [69]D.O.Hebb: "The Organization of Behavior", New York, Wilkey, 1949.
- [70]K.Nakano: "Associatron-A Model of Associative Memory", IEEE Trans., SMC-2, pp.380-388, 1972.
- [71]T.Kohonen: "Correlation Matrix Memories", IEEE Trans., C-21, pp.353-359, 1972.
- [72]S.Amari: "Neural Theory of Association and Concept-Formation", Biol. Cybernetics, vol.26, pp.175-185, 1977.

[73]P.Brodatz: "Textures", Dover, New York, 1966.

[74]矢野，白仁田，瀧山：”輪郭線とテクスチャの統合情報による画像認識”，映像情報メディア学会誌，vol.51，no.6，pp.935-939，1997.

[75]矢野，平岡，瀧山：”輪郭線及びテクスチャ情報に基づく画像認識法 —魚画像認識への応用—”，電子情報通信学会技術報告，PRMU96-148，1996.

[76]K.Yano, T.Hiraoka, R.Takiyama: "Image Recognition by Unifying Contour and Texture Information", International Technical Conference on Circuits/Systems, Computers and Communications 1997, pp.225-228, 1997.

[77]瀧山，矢野：”共起行列に基づくテクスチャ画像の識別”，信学技報，PRU91-148，1992.

図目次

- 2. 2. 1 輪郭線の偏角 14
- 2. 2. 2 偏角関数 ϕ 14
- 2. 2. 3 デジタル輪郭線での偏角 15
- 2. 2. 4 曲率 C 16
- 2. 3. 1 補角 18
- 2. 3. 2 偏角 19
- 2. 4. 1 デジタル輪郭線の補間 21
- 2. 4. 2 輪郭線に沿った長さが等しい再標本化 22
- 2. 4. 3 輪郭線に沿った長さが等しい再標本化のフローチャート 22
- 2. 4. 4 直線距離で等間隔な再標本化をするときの標本点の決定方法 23
- 2. 4. 5 直線距離で等間隔な再標本化のフローチャート 25
- 2. 4. 6 EPAC 26
- 2. 5. 1 閉曲線 $x(s)$ 29
- 2. 5. 2 EPACへのアルゴリズム 32
- 3. 3. 1 データ波形 40
- 3. 3. 2 ノイズが乗ったデータ 41
- 3. 3. 3 最小2乗法による平滑化 41
- 3. 3. 4 M推定による平滑化 42
- 3. 3. 5 M推定による平滑化 (β 過大) 42
- 3. 3. 6 LMedS推定による平滑化 43
- 3. 3. 7 メディアンフィルタによる平滑化 43
- 4. 2. 1 直線画像 46

4. 2. 2	極座標形式による直線表現	47
4. 2. 3	画像空間とパラメータ空間	48
4. 2. 4	直線画像	49
4. 4. 1	相関をとる領域	53
4. 4. 2	相関関数行列	53
4. 7. 1	画素対	56
4. 7. 2	θ のレベル	57
4. 7. 3	2 値画像 A	58
4. 7. 4	濃度共起行列の例	58
4. 8. 1	ランレングス	63
4. 8. 2	ランレングス行列の例	63
5. 2. 1	ニューラルネットワークの構造	72
5. 2. 2	ニューロンのモデル	73
5. 3. 1	3 層ニューラルネットワーク	76
5. 4. 1	3 層ニューラルネットワークによる情報の統合	78
6. 1. 1	魚画像の例	81
6. 2. 1	画像識別の過程	82
6. 2. 2	識別処理のフロー	82
6. 2. 3	学習のフロー	82
6. 3. 1	ハタ科	85
6. 3. 2	カレイ科	86
6. 3. 3	チョウチョウオ科	87
6. 3. 4	モンガラカワハギ科	88
6. 3. 5	ベラ科	89
6. 3. 6	アイゴ科	90
6. 3. 7	ニザダイ科	91
6. 3. 8	マフグ科	92
6. 3. 9	ニシン科	93
6. 3. 10	イサキ科	94
6. 3. 11	輪郭線特徴抽出の処理例	95

6. 3. 1 2	輪郭線特徴抽出のフロー	96
6. 3. 1 3	パワースペクトル	96
6. 3. 1 4	テクスチャ特徴の抽出領域の例	98
6. 3. 1 5	原画像と再量子化画像の例	98
6. 3. 1 6	濃度共起行列の要素データ例	98
6. 3. 1 7	テクスチャ特徴の抽出フロー	99
6. 4. 1	実験に用いるニューラルネットワーク	100
6. 4. 2	学習曲線（重み修正加速係数：0.4、しきい値修正加 速係数：0.3）	102
6. 4. 3	学習曲線（重み修正加速係数：0.1、しきい値修正加 速係数：0.1）	103
6. 4. 4	学習曲線（重み修正加速係数：0.3、しきい値修正加 速係数：0.3）	103
6. 4. 5	学習曲線（重み修正加速係数：0.5、しきい値修正加 速係数：0.5）	104
6. 4. 6	学習曲線（重み修正加速係数：0.7、しきい値修正加 速係数：0.7）	104
6. 4. 7	学習曲線（重み修正加速係数：0.9、しきい値修正加 速係数：0.9）	105

表目次

- 5. 2. 1 学習則の分類 73
- 6. 3. 1 実験に用いる魚画像 83
- 6. 4. 1 輪郭線特徴を単独で用いた場合の識別結果 105
- 6. 4. 2 テクスチャ特徴を単独で用いた場合の識別結果 106
- 6. 4. 3 輪郭線特徴とテクスチャを統合して用いた場合の識別結果 106