

Title	モルフォロジーにおける画像処理構造要素の最適設計法に関する研究
Author(s)	山本, 慎一
Citation	
Issue Date	1998-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/1159">http://hdl.handle.net/10119/1159</a>
Rights	
Description	Supervisor:小谷 一孔, 情報科学研究科, 修士

# 修士論文

## モルフォロジーにおける画像処理構造要素の 最適設計法に関する研究

指導教官 小谷 一孔 助教授

北陸先端科学技術大学院大学  
情報科学研究科情報処理学専攻

山本 慎一

1998年2月13日

## 要旨

モルフォロジーは、画像処理を対象画像と構造要素の集合演算で定義し、基本的な集合演算の組合せと構造要素によって様々な画像処理を実現することができる。

しかし、モルフォロジーの演算式は理論的に設計できても、画像情報のプローブとも言うべき構造要素は集合であるため、対象画像から抽出しようとする特徴が数学的な表現が困難な場合には、その設計は困難なものとなる。

そこで、本研究の目的は、モルフォロジーで定義された画像処理に対し、その処理目的に最適な構造要素を求める手法の獲得とする。

その構造要素の最適化手法として、目的とする画像処理に対してその評価関数を設定し、その評価値が改善されるようにトレーニングを行なうことにより、最適な構造要素を設計する。しかし、構造要素は、2次元もしくは3次元空間における位置ベクトルを要素とする集合であるため、そのような部分集合は無数に存在する。このような、無数に存在する構造要素の中から最適なものを効率良く求めるための最適化アルゴリズムとして、本論文では、2値画像のように狭い探索空間において精度の良い最適構造要素を求めることができる Simulated Annealing(SA) による最適化手法と、濃淡画像のように広い探索空間において効率よく最適な構造要素が求まる Genetic Algorithm(GA) による2つの最適化手法を提案する。

また、提案した手法の有効性を確かめるために、モデル画像に対するエッジ検出への適用や、パターンスペクトルに基づくテクスチャ解析への適用、特に、実画像への応用例として顔画像の表情解析への適用を試みた。

その結果から、対象画像が2値画像のように狭い探索空間においては、SAの方が効率良く最適構造要素が求まることが分かったが、対象画像が濃淡画像のように探索空間が広い場合、GAの方が評価値の良い構造要素を得ることができると明らかとなった。

# 目次

<b>1</b>	<b>序論</b>	<b>1</b>
1.1	はじめに	1
1.2	モルフォロジー	2
1.2.1	モルフォロジーの基本演算の定義	3
1.2.2	モルフォロジーの基本演算の演算例	4
1.2.3	構造要素の違いによる効果	5
1.2.4	パターンスペクトルとエントロピー	5
1.2.5	モルフォロジーの特徴	7
1.3	本研究の目的	7
<b>2</b>	<b>構造要素の最適化手法</b>	<b>8</b>
2.1	SAによる構造要素の最適化手法	9
2.1.1	構造要素のモデル化	10
2.2	GAによる構造要素の最適化手法	11
2.2.1	構造要素の最適化における遺伝的操作	11
2.2.2	構造要素の選択方法	12
<b>3</b>	<b>構造要素の最適化の有効性</b>	<b>13</b>
3.1	エッジ検出への適用	13
3.1.1	実験方法	13
3.1.2	SAによる構造要素の最適化のエッジ検出への適用結果	15
3.1.3	GAによる構造要素の最適化のエッジ検出への適用結果	16
3.1.4	まとめ	17

3.2	パターンスペクトルに基づくテクスチャ解析への適用 . . . . .	17
3.2.1	パターンスペクトルに基づくテクスチャ解析手法 . . . . .	17
3.2.2	顔画像の表情解析への適用 . . . . .	18
3.2.3	構造要素の最適化を行わない場合の実験結果 . . . . .	20
3.2.4	SA による構造要素の最適化の表情解析への適用結果 . . . . .	21
3.2.5	GA による構造要素の最適化の表情解析への適用結果 . . . . .	22
3.2.6	表情の判別精度の比較 . . . . .	23
3.2.7	最適構造要素の汎用性 . . . . .	25
3.2.8	まとめ . . . . .	26
4	むすび . . . . .	27

# 第 1 章

## 序論

### 1.1 はじめに

モルフォロジーは、対象画像と構造要素との集合演算で定義され、基本的なモルフォロジー演算の組合せと構造要素によって様々な画像処理を実現することができる。しかし、演算式は理論的に設計できても、画像情報のプローブとも言うべき構造要素は集合であるため、対象画像から抽出しようとする特徴が数学的な表現が困難な場合には、その設計は困難なものとなる。

そこで、本研究の目的は、モルフォロジーで定義された画像処理に対し、その処理目的に最適な構造要素を求める手法の獲得とする。これにより、対象とするものの特徴が、数学的に表現が困難な場合でも最適な構造要素の設計が可能となる。

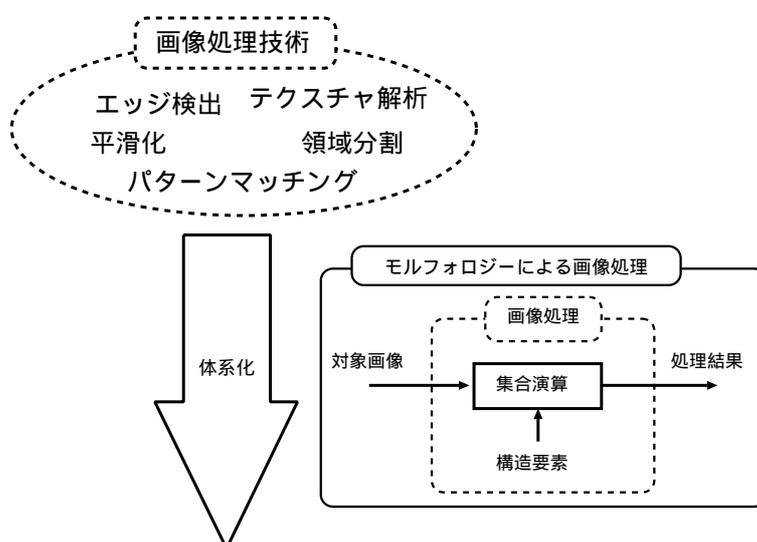
また、その手法の有効性を確かめるために、モデル画像に対するエッジ検出への適用や、パターンスペクトルに基づくテクスチャ解析への適用を試みる。特に、実画像への応用例として顔画像の表情解析への適用を試みる。

本論文では、本章でモルフォロジーによる画像処理について述べ、第 2 章で構造要素の最適化手法を提案し、その手法の有効性を確かめるために、第 3 章でモデル画像のエッジ検出への適用や、パターンスペクトルに基づくテクスチャ解析への適用を試みる。

## 1.2 モルフォロジー

モルフォロジーは、画像処理を対象画像と構造要素の集合演算で定義し、基本的な集合演算の組合せと構造要素によって様々な画像処理を実現することができる。このように、基本的な演算のみで画像処理を表すことにより画像処理が体系化され、処理の論理的妥当性の明示が容易になったり、画像処理ハードウェアのシステムデザインが容易となるといった利点が挙げられる。

以下に、そのモルフォロジーの基本演算の定義と、その特徴について述べる。



- ・画像処理の論理的妥当性の明示が容易
- ・画像処理ハードウェアのシステムデザインが容易

図 1.1: モルフォロジーによる画像処理

### 1.2.1 モルフォロジーの基本演算の定義

モルフォロジーの基本演算として dilation, erosion と、それらを組み合わせた opening, closing があり、対象画像が 2 値画像、濃淡画像の場合、それぞれ以下のように定義されている。

#### 2 値画像での定義

2 値画像では、以下のように定義される。

$$dilation : A \oplus B = \{x \in E^N \mid x = a + b, a^\vee \in A, b^\vee \in B\} \quad (1.1)$$

$$erosion : A \ominus B = \{x \in E^N \mid x + b \in A, b^\vee \in B\} \quad (1.2)$$

$$opening : A_B = (A \ominus B) \oplus B \quad (1.3)$$

$$closing : A^B = (A \oplus B) \ominus B \quad (1.4)$$

ただし、 $A, B$  は、それぞれ対象画像と構造要素の図形を構成する画素の座標を表す位置ベクトル  $a, b$  を要素とする集合である。

#### 濃淡画像での定義

濃淡画像では、以下のように定義される。

$$dilation : (f \oplus g)(x) = \max_{\substack{z \in G \\ x - z \in F}} \{f(x - z) + g(z)\} \quad (1.5)$$

$$erosion : (f \ominus g)(x) = \min_{\substack{z \in G \\ x + z \in F}} \{f(x + z) - g(z)\} \quad (1.6)$$

$$opening : f_g = (f \ominus g) \oplus g \quad (1.7)$$

$$closing : f^g = (f \oplus g) \ominus g \quad (1.8)$$

ただし、 $f(x)$  は対象画像、 $g(x)$  は構造要素である。また、 $F, G$  は、それぞれの定義域である。

## 1.2.2 モルフォロジーの基本演算の演算例

1.2.1節で定義されたモルフォロジーの基本演算の演算例を示し。それぞれの演算の特徴を示す。

例えば、図 1.2 に示す 2 値の対象画像と円形の構造要素が与えられた場合、それぞれの基本演算の演算結果は図 1.3 となる。

図 1.3 より、モルフォロジーの基本演算の特徴として以下のようなことが挙げられる。

### 1. dilation

対象とする図形を、構造要素の分だけ膨張させるような演算である。

### 2. erosion

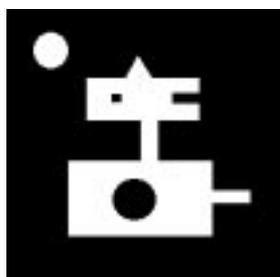
対象とする図形を、構造要素の分だけ縮小させるような演算である。

### 3. opening

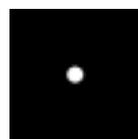
対象とする図形において、構造要素より小さな凸部を削るような演算である。

### 4. closing

対象とする図形において、構造要素より小さな凹部を削ったり孔を潰すような演算である。

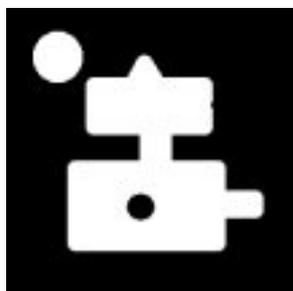


(a) 対象画像:  $A$

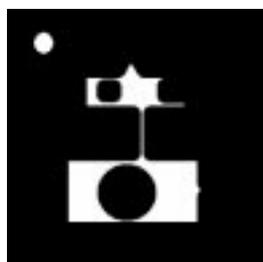


(b) 構造要素:  $B$

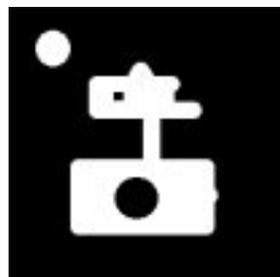
図 1.2: 対象画像と構造要素



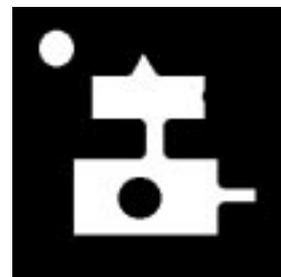
(a) dilation:  $A \oplus B$



(b) erosion:  $A \ominus B$



(c) opening:  
 $A_B = (A \ominus B) \oplus B$



(d) closing:  
 $A^B = (A \oplus B) \ominus B$

図 1.3: 基本演算の演算例

### 1.2.3 構造要素の違いによる効果

モルフォロジーは、基本演算と構造要素の組合せによって画像処理を表す。そのため、同一の演算式であっても、異なる構造要素を用いることにより、異なる演算結果が得られる。

図 1.4は、(a) の対象画像に対して、(b) ~ (d) の構造要素を用いて、同一のエッジ抽出演算 ( $A - (A \ominus B)$ ) を行なった結果を (b') ~ (d') に示している。このように、その構造要素特有の画像特徴を得ることができる。

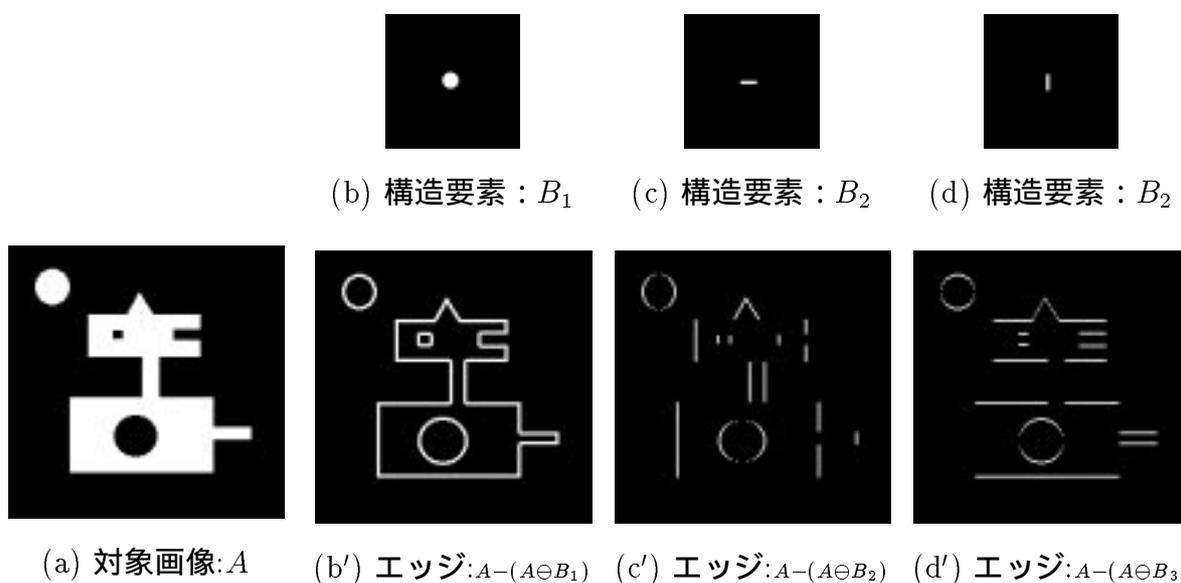


図 1.4: 構造要素の違いによる効果

### 1.2.4 パターンスペクトルとエントロピー

モルフォロジーで与えられるものとしてパターンスペクトルとエントロピーがある。

パターンスペクトル  $PS_f(s, g)$  は、式 (1.9), (1.10) で与えられ、スケール  $s$  の構造要素  $sg$  が、どの程度対象画像  $f$  を表しているかを示す。

$$PS_f(s, g) = A[f_{sg}] - A[f_{(s+1)g}] \quad (1.9)$$

$$sg = \underbrace{g \oplus g \oplus \dots \oplus g}_s \quad (1.10)$$

ただし、 $A[ ]$  は、 $[ ]$  内の輝度値の総和である。

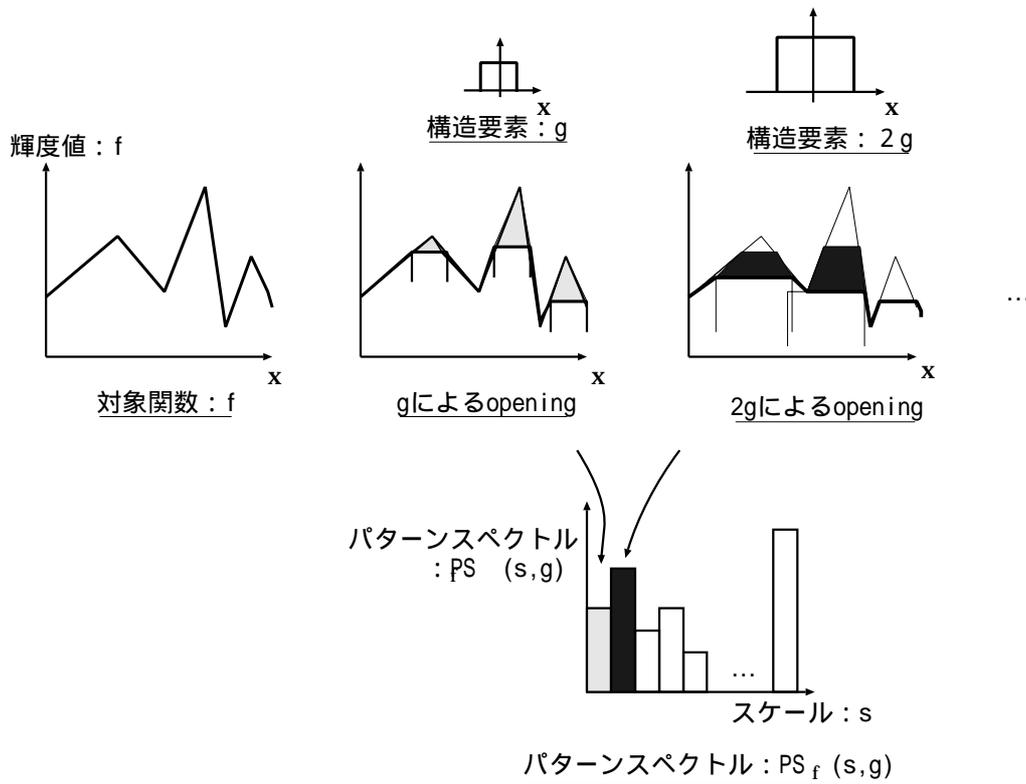


図 1.5: パターンスペクトル

エントロピー  $H(f/g)$  は、式 (1.11),(1.12) で与えられ、構造要素を基準とした画像  $f$  の粗さを表す尺度である。

$$H(f/g) = \sum_{n=0}^N p(n) \log \left[ \frac{1}{p(n)} \right] \quad (1.11)$$

$$p(s) = \frac{PS(s, g)}{S[f]} \quad (1.12)$$

また、エントロピーは、 $PS_f(s, g)$  が広く分布しているとき大きな値を示し、分布が偏っているとき小さな値を示す。

このように、パターンスペクトルより、以下のような画像の特徴をとらえることができる。

- ・スケール  $s$  が小さいところでは、図形の局所的な凹凸の状況を示す。
- ・スケール  $s$  の大きいところでは、図形の概形を表す。
- ・パターンスペクトルの分布は、画像の粗さを表す。

### 1.2.5 モルフォロジーの特徴

これまで述べてきたことから、モルフォロジーの利点と欠点として以下のようなことが挙げられる。

#### 利点

- 画像処理の体系化が可能
  - － 画像処理の論理的妥当性の明示が容易
  - － 画像処理ハードウェアのシステムデザインが容易
- 構造要素を任意に変えることにより、その構造要素特有の画像特徴が得られる。
- パターンスペクトルにより、テクスチャの特徴を捉えることができる。

#### 欠点

- 適切な構造要素を設定しなければ、目的の処理結果が得られない。
- 造要素が数学的に表現が困難な場合、適切な構造要素の論理的設計は困難である。

## 1.3 本研究の目的

1.2.5節で挙げたモルフォロジーの欠点から、本研究の目的は、処理目的を達成する最適な構造要素を求める手法の獲得手法の獲得とし、また、その手法の有効性を確かめるために、モデル画像に対するエッジ検出への適用や、パターンスペクトルに基づくテクスチャ解析への適用を試みる。特に、実画像への応用例として顔画像の表情解析への適用を試みる。

## 第 2 章

# 構造要素の最適化手法

本手法では、まず目的とする処理を実現するようなモルフォロジー演算を論理的に定義する。また、構造要素が適切であるか評価を行なうための評価関数を与える。このとき、評価値が改善されるようにトレーニングを行なうことにより、最適な構造要素を設計する。

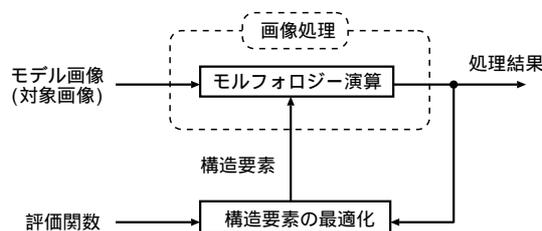


図 2.1: 構造要素の最適化手法

しかし、構造要素は、2次元もしくは3次元空間における位置ベクトルを要素とする集合であるため、そのような部分集合は無数に存在する。

このような、無数に存在する構造要素の中から最適なものを効率良く求めるためのアルゴリズムが必要である。

本研究では最適化アルゴリズムとして、2値画像のように狭い探索空間において精度の良い最適構造要素を求めることができる Simulated Annealing(SA) による最適化手法と、濃淡画像のように広い探索空間において効率よく最適な構造要素が求まる Genetic Algorithm(GA) による2つの最適化手法を提案する。

## 2.1 SA による構造要素の最適化手法

SA のアルゴリズムを、図 2.2 に示す。SA は、評価値が改善されるときは常に構造要素を変形し、評価値が改悪される場合でも、ある確率  $P(\Delta E)$  でその構造要素の変形を認めることによって、局所解から脱出することが可能となり最適な構造要素を求めることができる。 $P(\Delta E)$  は、

$$P(\Delta E) = \exp(-\Delta E/T(t)) \quad (2.1)$$

で与えられる。ただし、改悪量  $\Delta E$  は、次式で与えられる。

$$\Delta E = E(g(t)) - E(g_{tmp}(t)) \quad (2.2)$$

また、 $T(t)$  は、温度とよばれ、探索が進むにつれて小さくするアニーリングと呼ばれる操作を行なうことにより、探索が進むにつれて  $P(\Delta E)$  が減少し、最適構造要素に収束する。しかし、SA の特性は、温度によって左右されるため、適切な温度スケジューリングが必要である。そこで、本研究では、温度スケジューリングとして、S.Geman らによって提案 [3] された、

$$T(t) = \frac{k}{\log(1+t)} \quad (2.3)$$

を用いる。ただし、 $k$  は、正定数である。

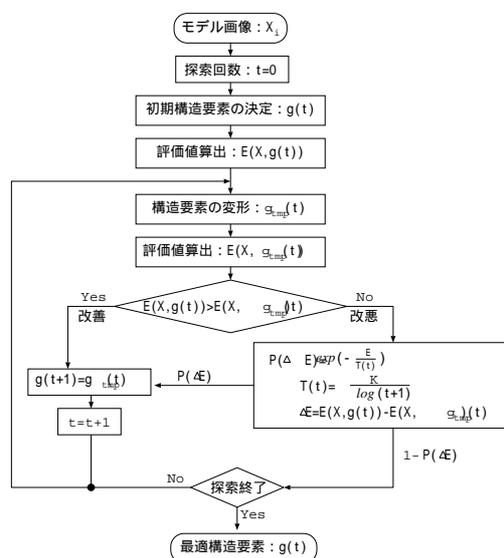


図 2.2: SA アルゴリズム

### 2.1.1 構造要素のモデル化

SA では、現在の構造要素の近傍の構造要素を探索するため、SA の探索範囲は GA に比べて大きくない。そこで、対象画像が濃淡画像の場合、構造要素のとり得るパターンの多さから、SA では、図 2.3 のように水平成分 (HSE) と垂直成分 (VSE) の dilation で与えられる屋根型の構造要素でモデル化し、この構造要素の頂点の位置を変えることにより、構造要素を変形することにより探索空間を小さくし効率的に最適構造要素を求める。

$$HSE_{hp,hl}(x, 0) = \begin{cases} \frac{hp}{hl}x & : 0 \leq x \leq hp \\ \frac{hp}{4-hl} - x + 4 & : hp \leq x \leq 4 \end{cases} \quad (2.4)$$

$$VSE_{vp,vl}(0, y) = \begin{cases} \frac{vp}{vl}y & : 0 \leq y \leq vp \\ \frac{vp}{4-vl} - y + 4 & : vp \leq y \leq 4 \end{cases} \quad (2.5)$$

$$SE = HSE \oplus VSE \quad (2.6)$$

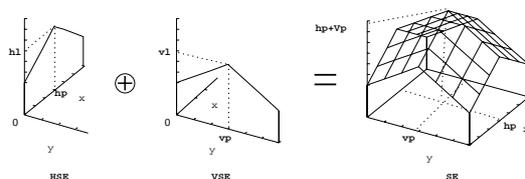


図 2.3: 構造要素のモデル化

## 2.2 GA による構造要素の最適化手法

GA は、生物進化の原理に着想を得たアルゴリズムである。遺伝子型に相当する構造要素の集団に対し、自己再生・交叉・突然変異などの遺伝的操作を繰り返し行ない、評価値の高い構造要素を選んで残していくことにより、構造要素の集団全体の評価値を向上させていく方法である。

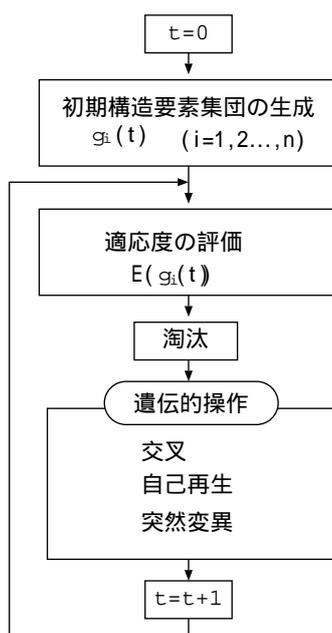


図 2.4: 遺伝的アルゴリズム GA

### 2.2.1 構造要素の最適化における遺伝的操作

通常、遺伝子型は 1 次元配列で表されるが、本手法で求める解は 2 次元配列で表される構造要素であるため、1 つの構造要素を 2 次元の遺伝子型とみなす。そのときの遺伝的操作は、以下のように行なう。

(a) 交叉

交叉は、縦または横に構造要素を切りつなぎ合わせるにより、新しい構造要素を生成する。

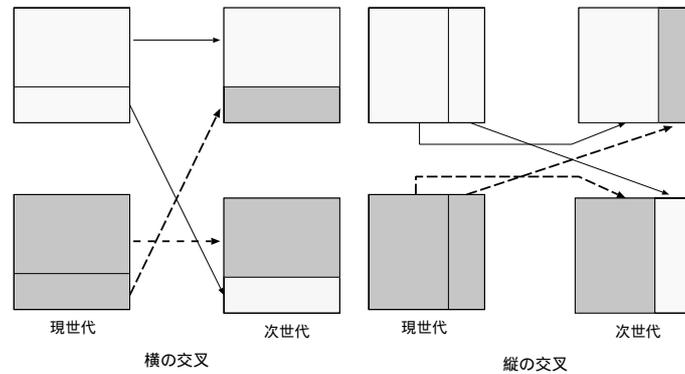


図 2.5: 交叉

(b) 突然変異

突然変異は、乱数によって選ばれた画素の値を変化させることにより、新しい構造要素を生成する。

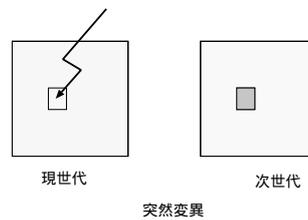


図 2.6: 突然変異

(c) 自己再生

自己再生は、次世代に全く同じ構造要素を残す。

### 2.2.2 構造要素の選択方法

交叉や突然変異を受ける構造要素の選択方法は、良い評価値の構造要素が子孫を残せるように構造要素の評価値  $E(i)$  に比例した確率より選択する。構造要素  $i$  が選ばれる確率  $p(i)$  は次式で与えられる。ただし、 $N$  は集団中の個体数である。

$$p(i) = \frac{E(i)}{\sum_{j=1}^N E(j)} \quad (2.7)$$

また、評価値の良い構造要素を次世代で失うことがないように、構造要素の集団中の評価値が上位のものは、自己再生により次世代に受け継ぐ。

## 第 3 章

# 構造要素の最適化の有効性

### 3.1 エッジ検出への適用

第 2 章で述べた最適化手法の有効性を確かめるために、2 値の対象画像に対するエッジ検出への適用を試みる。

#### 3.1.1 実験方法

まず、処理目的であるエッジ検出に対して、次のモルフォロジー演算を定義する。

$$E_{AB} = A - (A \ominus B) \quad (3.1)$$

ただし、 $A$  は 2 値の対象画像、 $B$  は構造要素である。

次に、入力画像  $A$  に対して、ある構造要素  $B$  で目標とするエッジ画像  $E_{AB}$  を作成する。

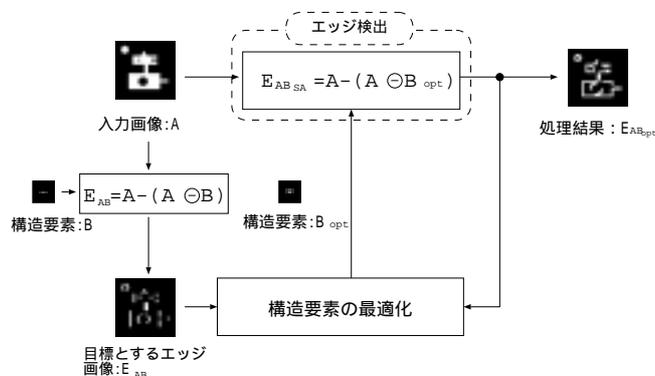


図 3.1: エッジ検出における構造要素の最適化手法の適用

このとき、第2章で述べた最適化手法によって、処理結果が目標とするエッジ画像に近づくように構造要素を変形させ最適な構造要素を求める。

その際の評価関数は、式(3.2)を用い、評価値が高いものを最適構造要素とする。

$$S[E_{AB}, E_{ABSA}] = \frac{\sum_{y=0}^H \sum_{x=0}^W E_{AB}(x,y) E_{ABSA}(x,y)}{\sqrt{\sum_{y=0}^H \sum_{x=0}^W E_{AB}(x,y)^2} \sqrt{\sum_{y=0}^H \sum_{x=0}^W E_{ABSA}(x,y)^2}} \quad (3.2)$$

実験では、図3.2に示す入力画像(64×64pixels)に対して、図3.3(a)~(c)の3種類の構造要素(5×5pixels)によって求められたエッジ画像(図3.3(a')~(c'))を目標とし、第2章で述べたSA及びGAによって構造要素の最適化を行なった。

また、SA、GAでの構造要素の探索は、7×7pixelsの範囲で構造要素を探索した。

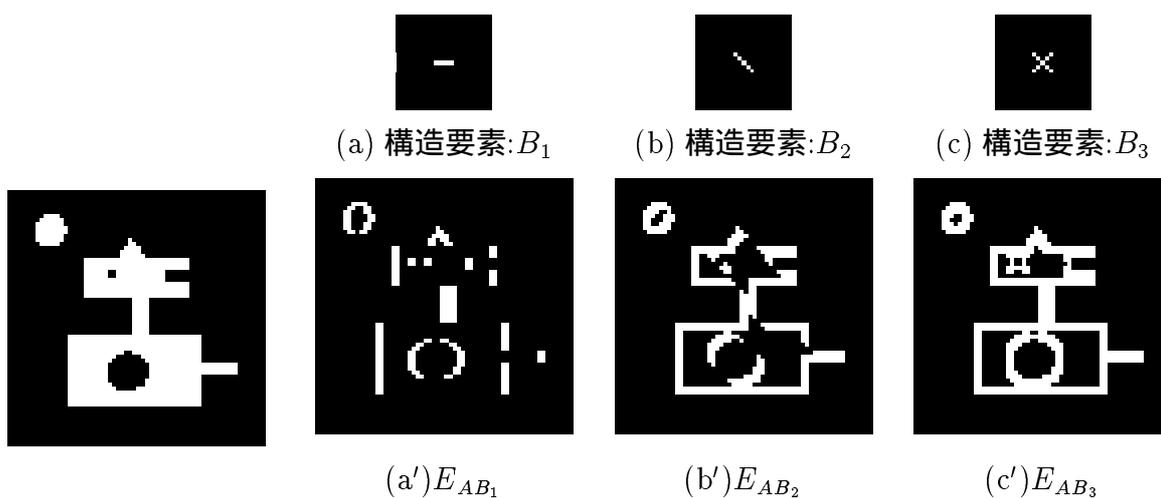


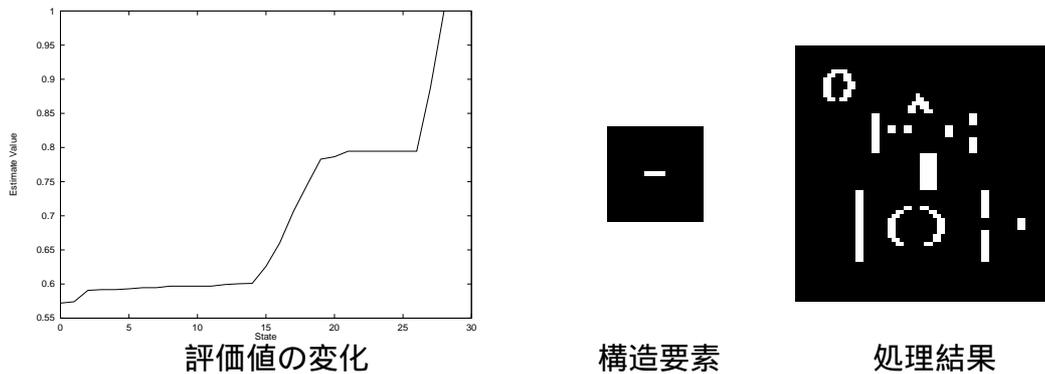
図 3.2: 入力画像

図 3.3: 目標とする処理結果

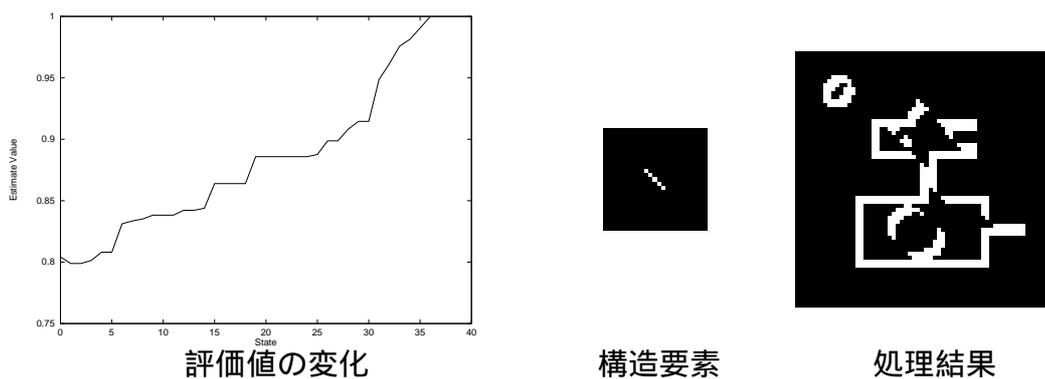
### 3.1.2 SA による構造要素の最適化のエッジ検出への適用結果

以下に、SA による構造要素の最適化の実験結果を示す。

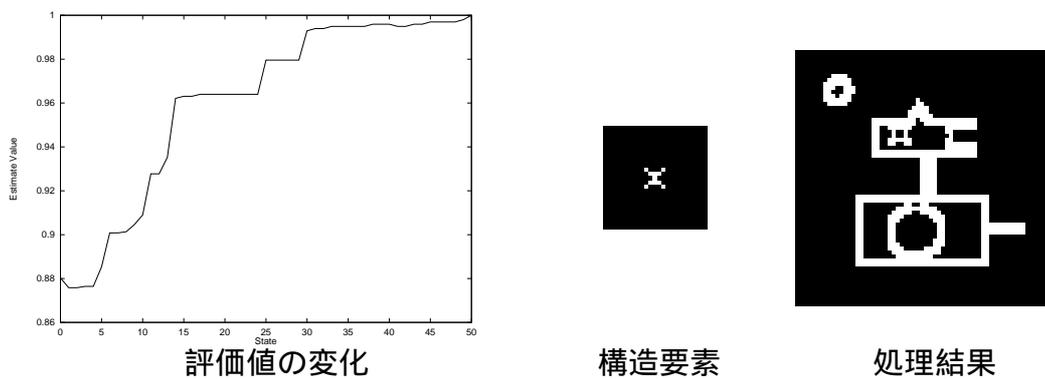
ただし、左図の横軸は探索回数、縦軸は評価値を示し評価値の推移を表している。また、中央は SA によって求められた最適構造要素。右図はそれによるエッジ画像である。



(a)  $E_{AB_1}$  を目標としたときの SA の結果



(b)  $E_{AB_2}$  を目標としたときの SA の結果



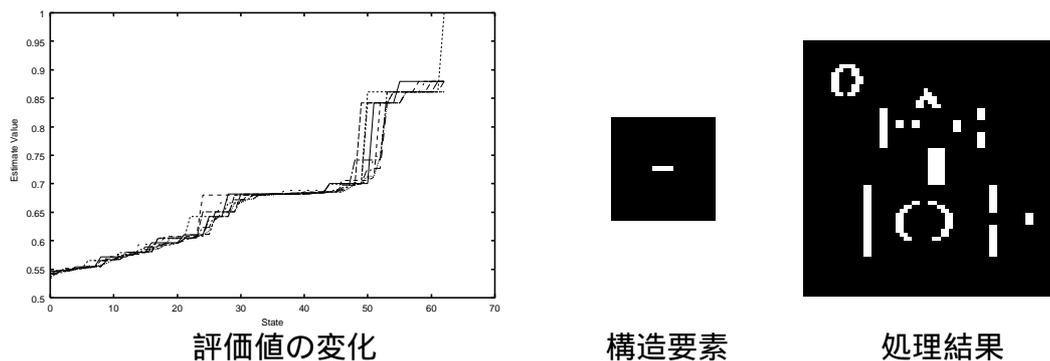
(c)  $E_{AB_3}$  を目標としたときの SA の結果

図 3.4: 実験結果

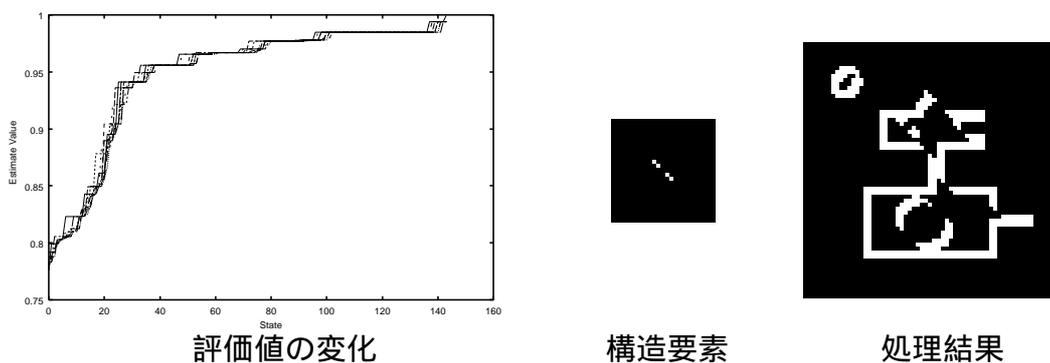
### 3.1.3 GA による構造要素の最適化のエッジ検出への適用結果

以下に、GA による構造要素の最適化の実験結果を示す。

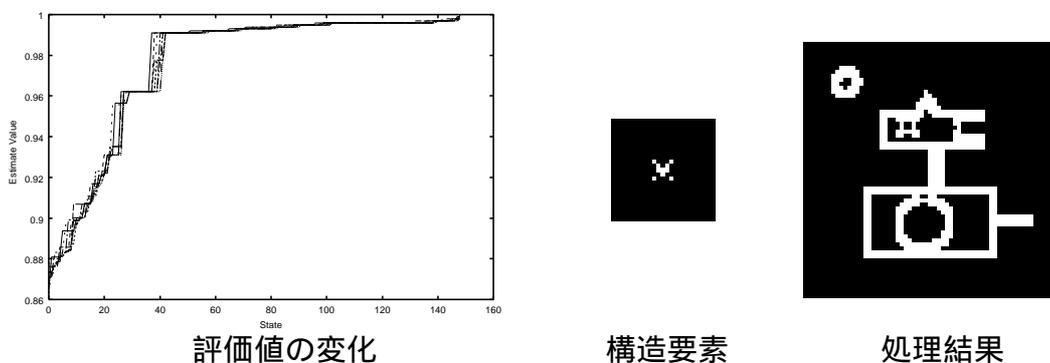
ただし、左図の横軸は世代数、縦軸は評価値を示し、各構造要素の評価値の推移を表している。また、中央は GA によって求められた最適構造要素。右図はそれによるエッジ画像である。



(a)  $E_{AB_1}$  を目標としたときの GA の結果



(b)  $E_{AB_2}$  を目標としたときの GA の結果



(c)  $E_{AB_3}$  を目標としたときの GA の結果

図 3.5: 実験結果

### 3.1.4 まとめ

SAによる構造要素の最適化実験の結果より、いずれの場合も28~50回で評価値が1.0となり、目標の処理結果を得ることができ、SAによって効率良く最適な構造要素が求まることが確認できた。

一方、GAによる構造要素の最適化実験では、SAと比較して探索回数(世代数)が63~145回と多くかかったものの、いずれの場合も評価値が1.0となり、目標の処理結果を得ることができた。

また、初期条件として、初期構造要素の探索範囲を $7 \times 7$ pixelsと、目標とするエッジ画像を作成したときの構造要素( $5 \times 5$ pixels)よりも大きな範囲で探索を行なった。その結果、SA、GAともに最適構造要素は、目標とした $5 \times 5$ pixelsと同等の大きさの構造要素が得られた。これより、最適構造要素の大きさは、トレーニングにより正しく定まることが確認できた。したがって、構造要素の初期条件として大きさは、予め大きめの探索範囲を設定しておけば良いことがわかる。しかし、あまり大きな探索範囲を設定すればその分探索コストがかかることは必要である。

## 3.2 パターンスペクトルに基づくテクスチャ解析への適用

第2章で述べた最適化手法の有効性を確かめるために、パターンスペクトルに基づくテクスチャ解析への適用を試みる。

本研究で扱うテクスチャ解析は、テクスチャが幾つかのカテゴリ( $i: i = 1, 2, \dots, N$ )に分類できるような画像群 $X_i$ を対象とし、入力画像 $X$ がどのカテゴリに属するかについて解析する。

その際に、テクスチャの特徴を捉えるものとしてパターンスペクトルを用い、そのパターンスペクトルに基づいて入力画像のカテゴリを判別する。

### 3.2.1 パターンスペクトルに基づくテクスチャ解析手法

パターンスペクトルに基づくテクスチャ解析手法を、図3.6に示す。予め各カテゴリを代表するモデル画像から、辞書となるパターンスペクトル $PS_{X_i}(s, g)$ を求めておく。そして、入力画像のパターンスペクトル $PS_X(s, g)$ と $PS_{X_i}(s, g)$ の類似度を求め、入力画像がどのカテゴリに属するかを判別する。

ただし、類似度は、式 (3.3) を用いる。

$$S[PS_X(s, g), PS_{X_i}(s, g)] = \frac{\sum_{s=0}^S PS_X(s, g) PS_{X_i}(s, g)}{\sqrt{\sum_{s=0}^S PS_X(s, g)^2} \sqrt{\sum_{s=0}^S PS_{X_i}(s, g)^2}} \quad (3.3)$$

しかし、辞書パターンスペクトル間の類似度が高い場合、誤判定を生じてしまう可能性がある。この誤判定を避けるためには、辞書中のパターンスペクトル間の距離を広げ、分解能の向上が必要となる。

そこで、辞書パターンスペクトルを作成する際に、第 2 章で述べた最適化手法を用いて構造要素の最適化を行ない、各カテゴリ固有のパターンスペクトルの特徴を強調することによって判別精度の向上を計る。

そのときの評価関数として式 (3.4) とし、辞書中のパターンスペクトル間の最も高い類似度が、最小となる構造要素を最適構造要素とする。

$$E[g] = \max_{\substack{i,j=0,1,\dots,N \\ j \neq i}} S[PS_{X_i}(s, g), PS_{X_j}(s, g)] \quad (3.4)$$

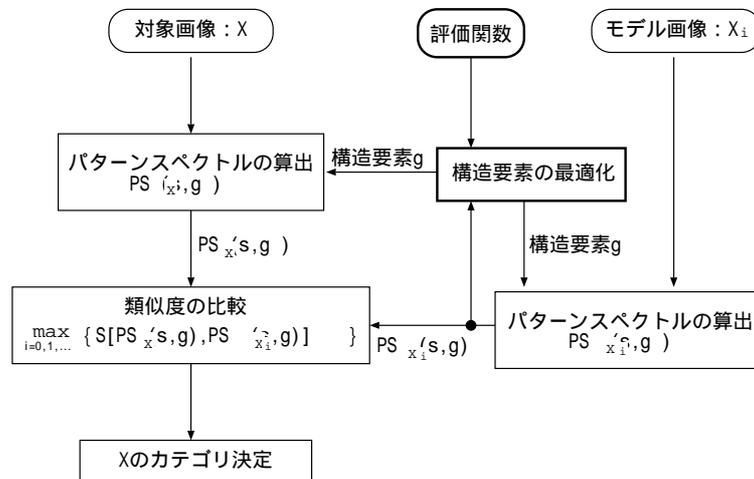


図 3.6: パターンスペクトルに基づくテキスト解析手法

### 3.2.2 顔画像の表情解析への適用

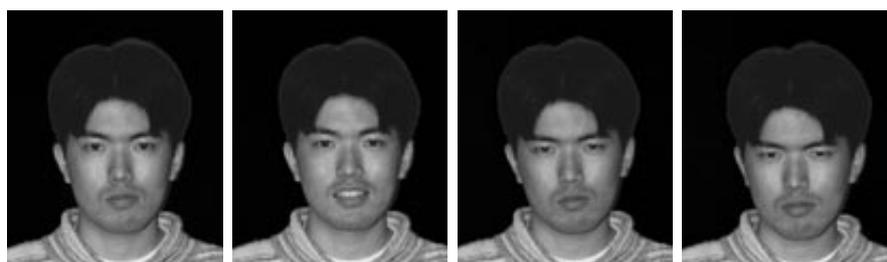
本手法の実画像への応用として、顔画像の表情解析へ適用する。笑い、怒り...など数種の表情カテゴリのうち、入力された顔画像がどの表情カテゴリに属するかを判別する。

今回実験では、無表情・笑い・悲しみ・怒りの4つの表情カテゴリに対して、図 3.7の各カテゴリを代表する顔画像 (100 × 120 pixels) に対して、それぞれの人物に対する辞書パターンスペクトルを作成し、オープンデータとして図 3.7とは別に撮影した顔画像を入力画像として与えたとき、それぞれの表情の顔画像がどの表情カテゴリに属するかを判別する。

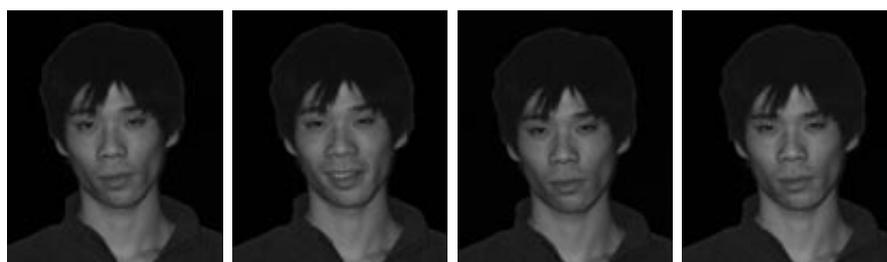
また、辞書パターンスペクトルを作成する際に、第 2章で述べた最適化手法を用いて構造要素の最適化を行ない、各表情カテゴリ固有の特徴を強調することによって判別精度の向上を計ることにより、本手法の有効性を確かめる。

また、実験では比較として、構造要素を最適化しない場合についても表情の判別実験を行なった。その際に、実験で使用した構造要素は、一般的によく用いられる式 (3.5) の SQUARE を用いた。

$$g(x, y) = 0 \quad : \quad \{-2 \leq x \leq 2, -2 \leq y \leq 2\} \quad (3.5)$$



(a) 人物 1



(b) 人物 2

図 3.7: 辞書に用いた画像 (無表情、笑い、悲しみ、怒り)

また、実験では SA、GA による構造要素の探索範囲は、5 × 5 pixels とし、GA においては、30 世代で最も良い評価値を示した構造要素を最適構造要素とする。

### 3.2.3 構造要素の最適化を行わない場合の実験結果

SQUARE を用いて、表情解析を行なった結果を以下に示す。

表 3.1に、SQUARE を用いて作成した辞書パターンスペクトル間における、各表情のパターンスペクトル相互の類似度を示す。また、そのときの評価値（最も高い値を示すもの）をアンダーラインで示す。

表 3.1: 辞書中のパターンスペクトル間の類似度  
(a) 人物 1 (b) 人物 2

	笑い	悲しみ	怒り		笑い	悲しみ	怒り
無表情	0.8927	0.8943	0.9587	無表情	0.9849	0.9932	0.9764
笑い		<u>0.9786</u>	0.9505	笑い		0.9858	<u>0.9934</u>
悲しみ			0.9536	悲しみ			0.9856

表 3.2に、別シーンで撮影した各表情の顔画像をオープンデータとして与えたとき、各表情カテゴリの辞書パターンスペクトルとの類似度を示す。また、その入力画像に対して最も高い値を示した表情カテゴリに対してアンダーラインで示す。

表 3.2: 対象画像と辞書画像のパターンスペクトル間の類似度  
(a) 人物 1 (b) 人物 2

辞書画像	無表情	笑い	悲しみ	怒り	辞書画像	無表情	笑い	悲しみ	怒り
入力画像					入力画像				
無表情	<u>0.9908</u>	0.9671	0.8900	0.9574	無表情	0.9863	0.9821	<u>0.9961</u>	0.9858
笑い	0.8482	<u>0.8874</u>	0.8102	0.8396	笑い	0.9731	0.9958	0.9819	<u>0.9974</u>
悲しみ	0.7511	<u>0.8675</u>	0.8555	0.8266	悲しみ	0.9824	0.9791	<u>0.9915</u>	0.9863
怒り	0.9112	<u>0.9321</u>	0.8255	0.8831	怒り	0.9845	<u>0.9964</u>	0.9913	0.9950

### 3.2.4 SA による構造要素の最適化の表情解析への適用結果

(a) 人物 1

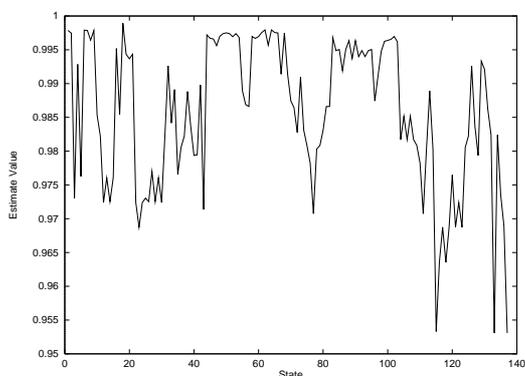


図 3.8: 評価値の推移

(b) 人物 2

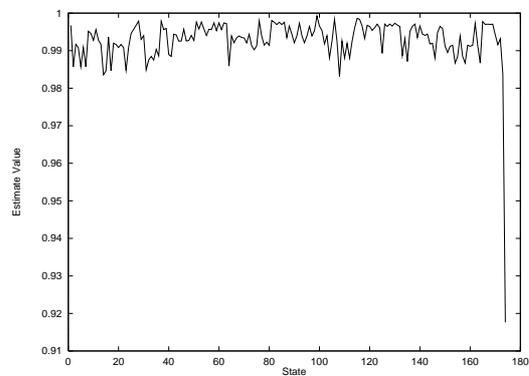


図 3.10: 評価値の推移

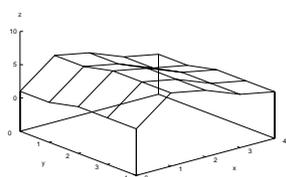


図 3.9: 最適構造要素

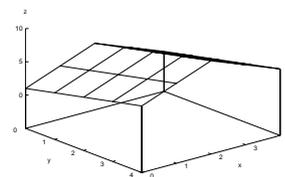


図 3.11: 最適構造要素

表 3.3: 最適構造要素による辞書パターンスペクトル間の類似度

	笑い	悲しみ	怒り
無表情	0.9277	<u>0.9530</u>	0.9220
笑い		0.9401	0.8401
悲しみ			0.7637

表 3.5: 最適構造要素による辞書パターンスペクトル間の類似度

	笑い	悲しみ	怒り
無表情	0.7960	0.8770	<u>0.9176</u>
笑い		0.7333	0.7286
悲しみ			0.6688

表 3.4: 入力画像と辞書画像のパターンスペクトル間の類似度

辞書画像 入力画像	無表情	笑い	悲しみ	怒り
無表情	0.9867	0.9847	<u>0.9972</u>	0.9756
笑い	0.8425	<u>0.8653</u>	0.8590	<u>0.8653</u>
悲しみ	0.8277	0.8496	<u>0.8527</u>	0.8396
怒り	0.7276	0.7016	0.7405	<u>0.7690</u>

表 3.6: 入力画像と辞書画像のパターンスペクトル間の類似度

辞書画像 入力画像	無表情	笑い	悲しみ	怒り
無表情	<u>0.9005</u>	0.8617	0.8496	0.8274
笑い	0.7288	0.8605	<u>0.8871</u>	0.7557
悲しみ	0.8349	0.8313	<u>0.8992</u>	0.8879
怒り	0.8088	0.8056	0.8450	<u>0.8743</u>

### 3.2.5 GA による構造要素の最適化の表情解析への適用結果

(a) 人物 1

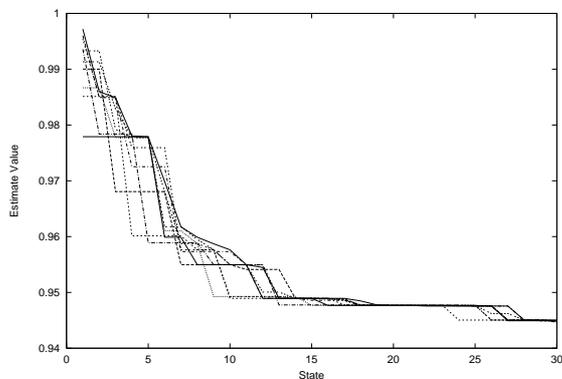


図 3.12: 評価値の推移

(b) 人物 2

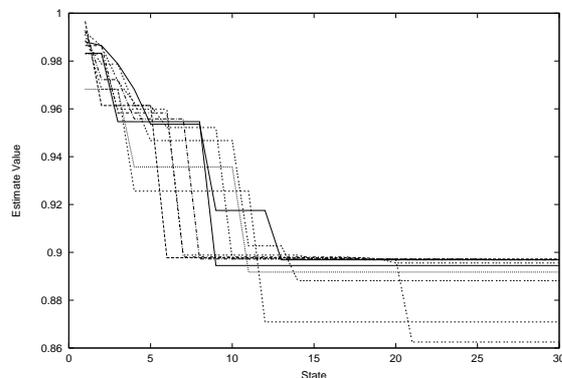


図 3.14: 評価値の推移

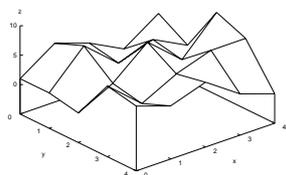


図 3.13: 30 世代で最も良い評価値を示す構造要素

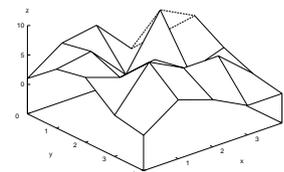


図 3.15: 30 世代で最も良い評価値を示す構造要素

表 3.7: 最適構造要素による辞書パターンスペクトル間の類似度

	笑い	悲しみ	怒り
無表情	0.6674	0.7371	0.7990
笑い		0.7894	0.8729
悲しみ			<u>0.9450</u>

表 3.8: 入力画像と辞書画像のパターンスペクトル間の類似度

辞書画像	無表情	笑い	悲しみ	怒り
入力画像				
無表情		<u>0.9978</u>	0.8966	0.9091
笑い	0.8152		<u>0.9262</u>	0.7832
悲しみ	0.8003	0.9404		<u>0.9407</u>
怒り	0.8374	0.7735	<u>0.9664</u>	

表 3.9: 最適構造要素による辞書パターンスペクトル間の類似度

	笑い	悲しみ	怒り
無表情	0.7153	<u>0.8625</u>	0.7149
笑い		0.6953	0.8104
悲しみ			<u>0.6577</u>

表 3.10: 入力画像と辞書画像のパターンスペクトル間の類似度

辞書画像	無表情	笑い	悲しみ	怒り
入力画像				
無表情		<u>0.9611</u>	0.8569	0.7137
笑い	0.8074		<u>0.9771</u>	0.6966
悲しみ	0.8105	0.7777		<u>0.9915</u>
怒り	0.7680	<u>0.8659</u>	0.6668	

### 3.2.6 表情の判別精度の比較

図 3.7の人物 1、人物 2 に加えて全部で 6 人の人物に対して、ことなるシーンで 3 シーン撮影した各表情の顔画像を与えた時のそれぞれの場合の表情の判別結果を表 3.11～表 3.13に示す。

また、図 3.16, 図 3.17に、各人物に対して、表情判別に用いた構造要素の評価値と、判別率を示す。

表 3.11: SQUARE を用いた場合の表情の判別率

人物	正答数				判別率 [%]
	無表情	笑い	悲しみ	怒り	
人物 1	2	3	2	1	66.6
人物 2	1	2	2	1	50.0
人物 3	1	1	2	2	50.0
人物 4	1	1	2	1	41.6
人物 5	3	2	1	0	50.0
人物 6	3	3	0	2	66.6
判別率 [%]	61.1	66.6	50.0	38.8	54.1

表 3.12: SA で求めた構造要素を用いた場合の表情の判別率

人物	正答数				判別率 [%]
	無表情	笑い	悲しみ	怒り	
人物 1	2	3	2	1	66.6
人物 2	2	2	2	3	75.0
人物 3	3	1	2	2	66.6
人物 4	1	2	3	2	66.6
人物 5	2	2	1	3	66.6
人物 6	3	3	3	0	75.0
判別率 [%]	72.2	72.2	72.2	61.1	69.4

表 3.13: GA で求めた構造要素を用いた場合の表情の判別率

人物	正答数				判別率 [%]
	無表情	笑い	悲しみ	怒り	
人物 1	3	3	2	1	75.0
人物 2	2	3	3	2	83.3
人物 3	2	3	2	3	83.3
人物 4	2	3	2	2	75.0
人物 5	3	1	3	2	75.0
人物 6	3	2	2	2	75.0
判別率 [%]	83.3	83.3	77.7	66.6	77.7

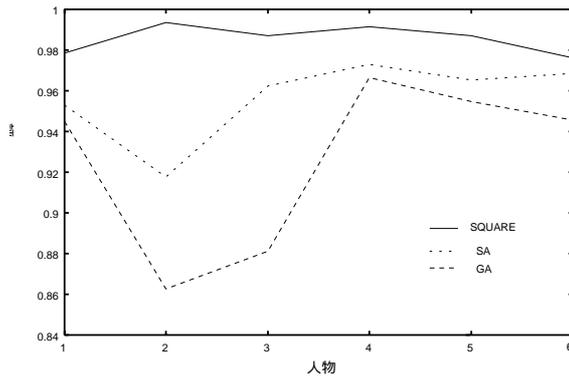


図 3.16: 人物-評価値

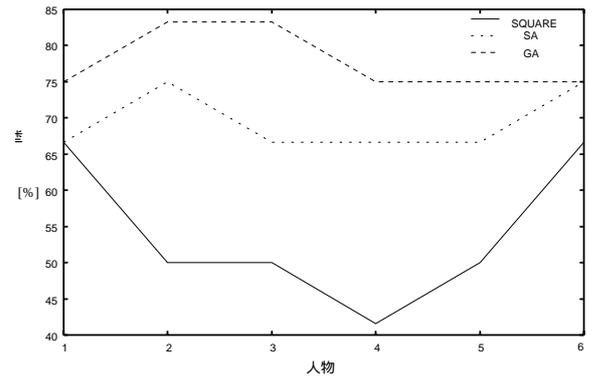


図 3.17: 人物-判別率

### 3.2.7 最適構造要素の汎用性

これまで、ある人物に対して、その人物の顔画像をトレーニングデータとして構造要素を最適化し、その構造要素を用いてその人物に対する辞書の作成及びオープンデータに対する表情の判別を行なった。

しかし、このように特定の人物の表情のみを判別するのでは、あまり実用的ではない。そこで、3.2.4節、3.2.5節において、SA、GA で求められた構造要素が、他の人物に対してどの程度有効かを調べるため、人物 1、人物 2 に対して以下のような実験を行なった。

まず、図 3.7 の顔画像に対して、SA、GA で求められた構造要素より次の 4 つの辞書を作成する。ただし、トレーニングデータとして人物 1(図 3.7(a)) 用いて最適化を行なって得られた構造要素を構造要素 1(図 3.9、図 3.13)、人物 2(図 3.7(b)) を用いて最適化を行なって得られた構造要素を構造要素 2(図 3.11、図 3.15) とする。

1. 人物 1 の顔画像から、構造要素 1 を用いて辞書を作成
2. 人物 1 の顔画像から、構造要素 2 を用いて辞書を作成
3. 人物 2 の顔画像から、構造要素 1 を用いて辞書を作成
4. 人物 2 の顔画像から、構造要素 2 を用いて辞書を作成

これらの辞書に対して、人物 1 と人物 2 のことな 3 シーンの各表情の顔画像を入力画像として与えたときの、判別率と各辞書の評価値を表 3.14、表 3.15 に示す。

表 3.14: SA で求めた構造要素の汎用性

辞書に用いた画像	人物 1		人物 2	
	構造要素 1	構造要素 2	構造要素 1	構造要素 2
辞書の評価値	0.9530	0.9986	0.9962	0.9176
人物 1 に対する判別率	66.6%	41.6%	25.0%	25.0%
人物 2 に対する判別率	25.0%	25.0%	50.0%	75.0%

表 3.15: GA で求めた構造要素の汎用性

辞書に用いた画像	人物 1		人物 2	
	構造要素 1	構造要素 2	構造要素 1	構造要素 2
辞書の評価値	0.9450	0.9985	0.9934	0.8625
人物 1 に対する判別率	75.0%	50.0%	25.0%	25.0%
人物 2 に対する判別率	25.0%	33.0%	58.0%	83.3%

### 3.2.8 まとめ

3.2節では、第2章で述べた最適化手法を、パターンスペクトルに基づくテクスチャ解析に適用し、特に顔画像の表情解析に応用することによりその有効性を確かめた。

図 3.17 に示すように、構造要素の最適化を行わないときよりも、最適化を行なった方が良い表情の判別率が得られた。また、図 3.16 と図 3.17 を見比べると、各人物に対して、構造要素の最適化による評価値の改善により、表情の判別率が向上していることがわかる。したがって、目的の処理結果に対して、式 (3.4) の評価関数は正しく設定されていることが確認できる。また、SA と GA を比較した場合も、GA の方が図 3.13 , 図 3.15 に示すように、SA の最適構造要素 (図 3.9 , 図 3.11 ) より複雑な形状の構造要素が得られ、図 3.16 に示すように構造要素の評価値も良くなり、その結果、SA よりも GA の方が高い判別率を示した。

しかし、その最適構造要素は、表 3.14、表 3.15 に示すように、ある人物の顔画像で作成された辞書を他の人物の表情判別に用いても、4つの表情カテゴリに対して 25% ~ 33% の判別精度しか得られず、表情判別が行なえていないことがわかる。

また、他の人物の顔画像をトレーニングデータとして構造要素の最適化を行なって得られた構造要素から辞書を作成し、同一の人物のオープンデータを与えたときの判別率は、41.6% ~ 58.0% と構造要素の最適化をしない場合 (SQUARE) の判別率とほぼ等しかった。したがって、ある人物に対して最適化された構造要素は、他の人物に対しては最適化されていないことが分かる。

この原因は、人物間の個人性によるものと考えられる。3.2.1節に述べた手法では、ある人物に対して、各表情カテゴリ間の距離を広げるように構造要素の最適化を行なった。しかし、この場合他の人物の表情を判別することはできない。

そこで今後、構造要素の最適化を行なう際に、複数人の顔画像をトレーニングデータとして与え、各人物間の個人性を無くし、かつ各表情カテゴリ間の距離を広げるような評価関数の検討が必要である。

## 第 4 章

### むすび

本論文では、モルフォロジーによる画像処理において、その処理目的を達成する最適な構造要素を求める手法を提案し、さらに、その手法の有効性を確かめるために、モデル画像に対するエッジ検出への適用や、パターンスペクトルに基づくテクスチャ解析への適用を試みた。

#### 1. 最適な構造要素を求める手法について

本論文では、図 2.1 のように、画像処理目的に対する評価値が改善されるようにトレーニングを行なうことにより、最適な構造要素を設計する。しかし、構造要素は、2 次元もしくは 3 次元空間における位置ベクトルを要素とする集合であるため、そのような部分集合は無数に存在する。

このような、無数に存在する構造要素の中から最適なものを効率良く求めるためのアルゴリズムとして、Simulated Annealing(SA) と Genetic Algorithm(GA) を用いて構造要素の最適化を行なう手法を提案した。

#### 2. モデル画像に対するエッジ検出への適用について

SA による構造要素の最適化を、2 値のモデル画像に対するエッジ検出へ適用した実験結果より、探索回数が 28 ~ 50 回で、いずれの場合も評価値が 1.0 となり、目標の処理結果を得ることができ、SA によって効率良く最適な構造要素が求まることが確認できた。

一方、GA による構造要素の最適化実験では、SA と比較して探索回数(世代数)が 63 ~ 145 回と多くかかったものの、いずれの場合も評価値が 1.0 となり、目標の処

理結果を得ることができた。

### 3. パターンスペクトルに基づくテクスチャ解析への適用について

SA 及び GA による構造要素の最適化を、パターンスペクトルに基づくテクスチャ解析に適用し、特に顔画像の表情解析に応用することによりその有効性を確かめた。

顔画像の表情判別実験の結果から、構造要素の最適化により、最適化を行わないときよりも表情の判別率の向上が見られた。

また、SA と GA を比較した場合、濃淡画像に対する構造要素のように広範囲な探索空間の場合、GA の方が SA で求めた構造要素より複雑な形状の構造要素が得られ、その分評価値の良い構造要素が得られ、その結果、SA よりも GA の方が高い判別率を示した。

また、最適化された構造要素の他の人物に対する汎用性は、構造要素に関係なくある人物の顔画像で作成された辞書を他の人物の表情判別に用いても、4つの表情カテゴリに対して25%~33%の判別精度しか得られず、表情判別は行なえなかった。

また、他の人物の顔画像をトレーニングデータとして構造要素の最適化を行なって得られた構造要素から辞書を作成し、同一の人物のオープンデータを与えたときの判別率は、構造要素の最適化をしない場合 (SQUARE) の判別率とほぼ等しく、人物間の個人性により、ある人物に対して最適化された構造要素は、他の人物に対しては最適化されていないことが確認できた。

以上の結果から、対象画像が2値画像のように狭い探索空間においては、SAの方が効率良く最適構造要素が求まることが分かったが、対象画像が濃淡画像のように探索空間が広い場合、GAの方が評価値の良い構造要素を得ることができると明らかとなった。

今後の課題としては、SAによる構造要素の最適化については、濃淡画像に対する構造要素の探索においても、より複雑な形状の構造要素が求まるように、自由度の高い構造要素のモデル化が必要である。GAによる構造要素の最適化については、より効率良く探索を進めるために、構造要素の最適化に適した遺伝的操作の検討が必要である。構造要素の最適化の顔画像の表情判別への適用においては、人物間の個人性を無くすような評価関数の検討が必要である。

最後に、本研究では、トレーニングによる構造要素の最適化を行なったが、画像処理を数学的に表現し、その展開・理解が容易になるというモルフォロジーの利点を考えると、最適構造要素の論理的な設計手法の開発が望まれる。そこで、トレーニングによって最適化された構造要素の特徴を解析することにより、最適構造要素の論理的設計における拘束条件が得られることを期待する。

# 謝辞

本研究をすすめるにあたり日頃から熱心に御指導頂きました本学 小谷一孔 助教授に深く感謝致します。終始貴重な御意見、御鞭撻を頂きました本学 宮原 誠 教授に深く感謝致します。

また、日頃から御助言、激励頂いた斎藤 康之氏をはじめとする 本学 像情報処理学講座の学生諸氏、顔画像データベースの作成に協力して下さった本学学生、職員の皆様に深く感謝致します。

## 参考文献

- [1] 間瀬 茂、上田 修功: “モルフォロジーと画像解析 [I]”, 信学誌, Vol.74, No.2, 1991, pp116-174
- [2] 間瀬 茂、上田 修功: “モルフォロジーと画像解析 [II]”, 信学誌, Vol.74, No.3, 1991, pp271-279
- [3] 北野 宏明: “遺伝的アルゴリズム”, 産業図書, 1993 年
- [4] 小畑 秀文: “モルフォロジー”, コロナ社, 1996 年
- [5] 山本 慎一、小谷 一孔: “モルフォロジーにおける最適構造要素によるテクスチャ解析”, 第 2 回映像メディア処理シンポジウム資料, Vol.2, 1997, pp105-106
- [6] 山本 慎一、小谷 一孔: “モルフォロジーにおける構造要素の GA による最適化とテクスチャ解析への適用”, 情処研報, Vol.98, No.5, 1998, pp57-64