1　はじめに

パターンスペクトル [1] は、対象画像をある基本形状 (構造要素) とその相似図形の組合せに分解し、相似図形のスケールの変化に対する分解した体積 (面積) の変化の割合として定義される。パターンスペクトルによる画像識別として、田口らは細かな “きめ” から構成されるテクスチャ画像の識別を入力パターンスペクトルと辞書パターンスペクトルの二乗誤差を用いて行い [2]、結城らは耳介と正方型構造要素で求まる入力パターンスペクトルと辞書パターンスペクトルとの類似度を用いて個人識別を行っている [3]。また、北田らは鼻の三次元形状のパターンスペクトルによる個人識別を行っている [4]。パターンスペクトルは一般に対象画像の振幅に関する特徴をよく表し、位相的なズレに対して不変性を持つが、画像識別の精度は構造要素に大きく依存する。従来、構造要素は対象画像の回転不変性を考慮した円盤型のものなど、適用に応じて工夫されているが、経験的に与えられている為、その画像識別に対して最適である保証はない。最適な構造要素を設計することができれば、識別精度を向上できるだけでなく、システムのデザインが効率よく行える。我々は適切な距離尺度と、辞書間の距離を最大とする構造要素が識別率を最大にすると仮定し、辞書画像のパターンスペクトル間の距離を最大とする構造要素を遺伝的アルゴリズムなどにより設計する手法を検討し、鼻形状による個人識別に適用して、有効性を確認した [5][6]。しかし、この手法では 1 個の構造要素で全ての画像の識別を行っており、辞書数の増加に伴い広げられる辞書間の距離に限界があった。本稿では、画像の識別を複数の構造要素を用いて階層的に行なう手法、及びそれら複数の構造要素の設計法を検討する。これは全ての辞書間の距離を最大にする構造要素を設計するのではなく、辞書を 2 つのグループに分割し、そのグループ間の距離が最大となるよう構造要素を設計する。そして、分割された各グループも同様の処理で分割を繰り返すことにより、階層的に辞書間の距離を広めていく。さらに、本手法をテクスチャ画像識別及び鼻形状による個人識別に適用し、有効性を検討する。

2　パターンスペクトルの導出

モルフォロジーは、集合論的操作からなる 2 値画像または濃淡画像の変形手法の一貫した理論体系で、対象画像と構造要素及びそれらの集合演算で様々な画像処理を実現できる [7][8]。

整数全体の集合を Z、2 次元離散空間を Z^2 とし、各 $(x,y) \in Z^2$,$(u,v) \in Z^2$ が濃淡値 $f(x,y) \in Z$, $g(u,v) \in Z$ を持つとすると、$f(x,y)$ と $g(u,v)$ の Minkowski 和及び Minkowski 差は式 (1),(2) で与えられる。

$$(f \oplus g)(x,y) = \max_{\substack{(x-u,y-v)\in F \\ (u,v)\in G}} \{f(x-u,y-v)+g(u,v)\} \tag{1}$$

$$(f \ominus g)(x,y) = \min_{(u,v)\in G} \{f(x-u,y-v)-g(u,v)\} \tag{2}$$

ただし、F, G は関数 f, g のそれぞれの定義域とする。

また、$\check{g}(u,v) = g(-u,-v)$ とすると、opening,closing と呼ばれる演算は、式 (3),(4) で与えられる。

$$f_g(x,y) = ((f \ominus \check{g}) \oplus g)(x,y) \tag{3}$$

$$f^g(x,y) = ((f \oplus \check{g}) \ominus g)(x,y) \tag{4}$$

そして以上から f の g によるパターンスペクトル $PS_f(n,g)$ は、式 (5) のように定義される。ただし、$A(f) = \sum_{(x,y)\in F} f(x,y)$、$ng = \underbrace{g \oplus g \oplus \cdots \oplus g}_{n}$、$n \geq 0$ とする。

$$PS_f(n,g) = -\frac{dA(f_{ng})}{dn} \tag{5}$$

これは、g をあるスケール倍拡大した ng で分解される f の体積 (面積) から $(n+1)g$ で分解される f の体積 (面積) を引いたものである。さらに、式 (6) によりサイズ n を負の領域へ拡張できる。

$$PS_f(-n,g) = \frac{dA(f^{ng})}{dn} \tag{6}$$

ただし、$n > 0$ とする。また、f を対象画像、g を構造要素と呼ぶ。

3　構造要素の最適化と画像識別手法

3.1　1 個の構造要素の最適化による画像識別

パターンスペクトルによる画像識別は、経験的に与えた構造要素 g を用いて、識別対象である入力画像のパターンスペクトルと最小の距離にある辞書画像のパターンスペクトルのカテゴリを識別結果とする。ここで、パターンスペクトルは構造要素で分解される面積の量のため、識別精度は構造要素 g に大きく依存する。もし、辞書パターンスペクトル間の距離を小さくする g を選択した場合、入力画像のゆれに対する耐性が弱くなり、識別精度が低下するおそれがある為、あらかじめ辞書パターンスペクトル間の距離の最小値を最大化する $\hat{g}$

を式 (7) により与える。ただし、辞書画像を d、識別するクラス数を I、$i,j=1,2,\cdots,I$、$i\neq j$、パターンスペクトル間の距離の最小値を与える関数を D' とする。

$$\hat{g}=g|_{D'\left(PS_{d_i}(n,g),PS_{d_j}(n,g)\right)\to\max} \tag{7}$$

また、入力画像を h とし、識別結果 $\tilde{i}$ を式 (8) により与える。ただし、D はパターンスペクトル間の距離を与える関数とする。

$$\tilde{i}=i|_{D\left(PS_h(n,\hat{g}),PS_{d_i}(n,\hat{g})\right)\to\min} \tag{8}$$

なお、パターンスペクトルの距離を与える D 及び D' には、相関など、適応に応じた距離尺度を与える。この手法による画像識別の例を図 1,2に示す。この図では入力画像及び辞書画像から識別結果を与える過程を概念的に示している。

3.2 GA による構造要素の最適化手法

一般に構造要素内のパラメータに対する偏導関数値の算出が困難なため、最適な構造要素の算出には最急降下法を適用できない。また、全探索は構造要素内のパラメータの取り得る範囲が膨大になり、実際的でない。そこで大域探索に向いた確率的探索手法の 1 つである、遺伝的アルゴリズム (GA) を用いて構造要素の最適化を行なう。アルゴリズムを以下に示す。

アルゴリズム 1 (GA による構造要素の最適化) .

Step 1: (遺伝子型の決定)
構造要素のサイズを $U\times V$ とし、構造要素 g を遺伝子 c とする。

Step 2: (初期遺伝子集団の決定)
世代を $t(t=0)$、個体数を K 個、初期遺伝子集団を $c(t,k)(k=1,2,\cdots,K)$ とし、各遺伝子のパラメータを $0\sim L$ の範囲でランダムに決定する。

Step 3: (各個体の適応度の算出)
各個体の適応度を式 *(9)* から算出する。ただし、$i,j=1,2,\cdots,I$、$i\neq j$ とする。

$$e(k)=D'\left(PS_{d_i}(n,c(t,k)),PS_{d_j}(n,c(t,k))\right) \tag{9}$$

Step 4: (選択、交叉)
適応度 $e(k)$ の低い s 個の個体を淘汰し、残った個体から s 組選択して一様交叉を行ない、淘汰した個体を補う。

Step 5: (突然変異)
K 個の個体に対し *10* %の確立で各パラメータを $0\sim L$ の範囲でランダムに置き換える。

Step 6: (終了判定)
$t=t+1$ とし t が最大ループ回数ならば、最も適応度の高い個体を最適化された構造要素 $\hat{g}$ として処理を終了する。そうでなければ *Step 3* に戻る。

3.3 複数の構造要素の最適化による階層的な画像識別

3.3.1 アルゴリズムの概要

上記では、画像識別に最適な構造要素の設計法を述べた。我々は 1 個の構造要素の最適化による画像識別の効果として、最適化しない場合に比べて１４%識別率が向上したことを報告している [6]。しかし、識別するクラスが多い場合、1 個の構造要素では辞書間の距離が十分広がらず、識別精度が低下するおそれがある。そこで、複数の構造要素による階層的な画像の識別手法を検討する。

辞書 $d_i(i=1,2,\cdots,I)$ のインデックス i をもれなく要素とする集合を $\mathbf{Q}_1(\mathbf{Q}_1\ni i)$ とし、$\mathbf{Q}_p=\phi(p=2,3,\cdots)$ とする。そして、P を $\mathbf{Q}_p\neq\phi$ となる p の最大値として $\mathbf{Q}_p(p=1,2,\cdots,P)$ を 2 つの集合 $\mathbf{Q}_{P+1},\mathbf{Q}_{P+2}$ にコピーする際、式 (10) を満たすような構造要素 $\hat{g_p}$ を与える。ただし、$\mathbf{Q}_{P+1}\cap\mathbf{Q}_{P+2}=\phi$、$\mathbf{Q}_p=\mathbf{Q}_{P+1}\cup\mathbf{Q}_{P+2}$ とする。そして $P=P+2$ として、$\mathbf{Q}_p$ の要素数が 1 になるまでグルーピングを繰り返す。これは、辞書を 2 つのグループ $\mathbf{Q}_{P+1},\mathbf{Q}_{P+2}$ に分ける際、そのグループ間の距離の最小値が最大となるように I 個のインデックスを階層的にグルーピングすることである。ただし、D' はグループ間の距離の最小値を与える関数とする。

$$\hat{g_p}=g_p|_{D'\left(PS_{d_{j'\in\mathbf{Q}_{P+1}}}(n,g_p),PS_{d_{j''\in\mathbf{Q}_{P+2}}}(n,g_p)\right)\to\max} \tag{10}$$

入力 h に対する識別は、構造要素 $\hat{g_p}(p=1,2,\cdots,P)$ を用いて h と最短距離にある辞書が属する集合 $\tilde{\mathbf{Q}}$ を 2 分木探索することとなる。これを式 (11) に示す。ただし、$\mathbf{Q}'=\mathbf{Q}_{r'}\ \text{or}\ \mathbf{Q}_{r''}$、$D$ はパターンスペクトルの距離を与える関数とする。また、$I=5$ としたときの画像の階層的な識別の例を図 3に示す。

$$\tilde{\mathbf{Q}}=\mathbf{Q}'|_{D\left(PS_h(n,\hat{g_p}),PS_{d_{j\in\mathbf{Q}'}}(n,\hat{g_p})\right)\to\min} \tag{11}$$

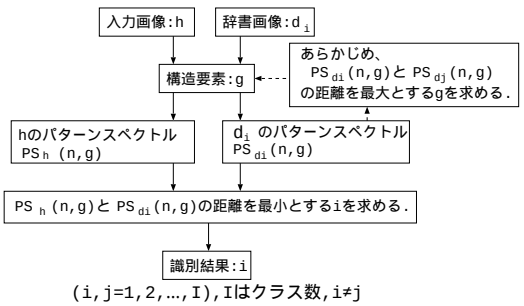


図 1: 構造要素の最適化の概念図

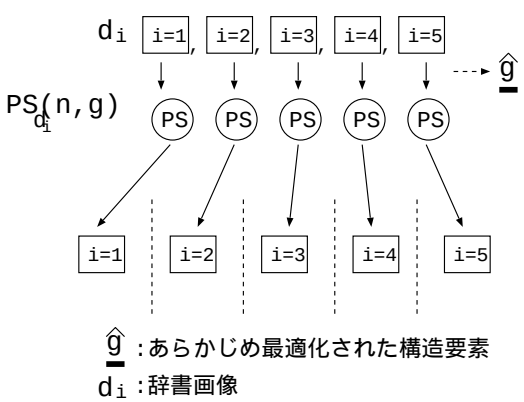


図 2: 一つの構造要素による画像の識別の例

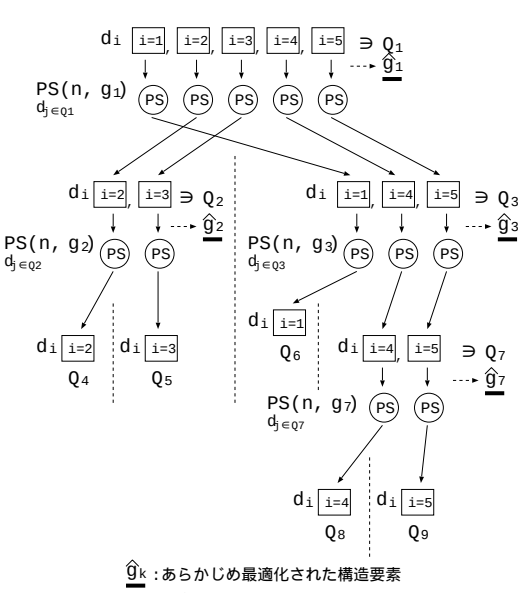


図 3: 複数の構造要素による画像の識別の例

3.3.2 インデックスを 2 つの集合へ分配するアルゴリズム

各 $\mathbf{Q}_p(p = 1, 2, \cdots, P)$ に含まれる辞書をグループ間の距離が最大となるように 2 つのグループ $\mathbf{Q}_{P+1}, \mathbf{Q}_{P+2}$ にコピーするアルゴリズムを詳しく説明する。

アルゴリズム 2 (グループ間の距離の最大化) .

Step 1:

$\mathbf{Q}_{P+1}$ に $\mathbf{Q}_p$ の要素のどれか 1 つをコピーし、それ以外の要素を $\mathbf{Q}_{P+2}$ にコピーする。

Step 2:

$k = A$ (A は十分大きな値とする。)

Step 3:

$k = k - \Delta k$ (Δk は十分小さな値とする。)

Step 4:

$if\ \ D'\left(PS_{d_{j\in\mathbf{Q}_{P+1}}}(n, g_p), PS_{d_{j'\in\mathbf{Q}_{P+2}}}(n, g_p)\right) < k$ $then$ $\mathbf{Q}_{P+1}$ に j' を追加し、$\mathbf{Q}_{P+2}$ から j' を削除する。

Step 5:

$if\ \mathbf{Q}_{P+2} = \phi\ then\ goto\ 3$

$else$ 終了

4 識別実験

4.1 実験に用いる画像

提案手法の有効性を確認するために、一例としてテクスチャ画像識別及び鼻形状による個人識別に適用する。

(1) テクスチャ画像として、Brodatz[9] の写真集から 14 種類選んだ画像を識別実験に用いる。これらのテクスチャ画像は細かな “きめ” から構成され、特に規則性は無いが、どの部分を切り出しても同じ模様であると主観的に認められるものである。選んだ各テクスチャの写真集を $650 \times 900[pixel], 256[bit/pixel]$ としてスキャナで取り込み、さらにその画像をランダムに $100 \times 100[pixel]$ の矩形領域として 21 枚 (計 294 枚) 抜き出す。これらの矩形領域の各 1 枚 (計 14 枚) を辞書画像として用い、残りの 20 枚 (計 280 枚) を入力画像とする。

(2) 鼻形状の画像として、レーザーレンジファインダで取り込んだ鼻の奥行き情報を 120×140[pixel],256[bit/pixel] のグレイスケール画像に変換したものを用いる。辞書画像として 10 人の無表情顔の鼻形状の画像各 1 枚 (計 10 枚)、入力画像として 6 表情 (無表情、笑い、悲しみ、怒り、目閉じる、口開く) の鼻形状の画像 (計 60 枚) を用いて実験を行う。

4.2 最適化および識別に用いる距離関数

最適化及び識別に用いる距離関数 D として、式 (12) の相関を用いる。ただし、N は構造要素のサイズ n の定義域とする。

$$D = \frac{\sum_{n \in N} PS_{f_1}(n,g) PS_{f_2}(n,g)}{\sqrt{\sum_{n \in N} PS_{f_1}(n,g)^2} \sqrt{\sum_{n \in N} PS_{f_2}(n,g)^2}} \quad (12)$$

4.3 パターンスペクトル及び最適化の為の各パラメータの決定

テクスチャ画像のパターンスペクトルを算出する為に用いる構造要素のサイズの定義域として、$N = \{-3, -2, -1, 0, 1, 2, 3\}$ を用いる。負のサイズのパターンスペクトルを用いているのは、田口ら [2] が実験的に確かめた、パターンスペクトルの opening 演算と closing 演算の相補的な性質を利用している為である。また、鼻形状の画像では、構造要素のサイズの定義域として、$N = \{0, 1, 2, 3, 4, 5\}$ を用いる。負のサイズのパターンスペクトルを用いないのは、鼻形状は closing 演算で溶融されることが無い為である。

テクスチャ画像識別及び鼻形状による個人識別に用いる構造要素の最適化の為のパラメータは各個体 (構造要素) のサイズ $U = V = 5$、個体数 $K = 15$、各個体のパラメータの範囲 $L = 5$、淘汰する個体数 $s = 5$、トレーニング回数 $T = 30$ とする。なお、上記のパラメータは、従来手法の 1 個の構造要素を用いる場合と、提案手法の複数の構造要素を用いる場合の両方に適用する。

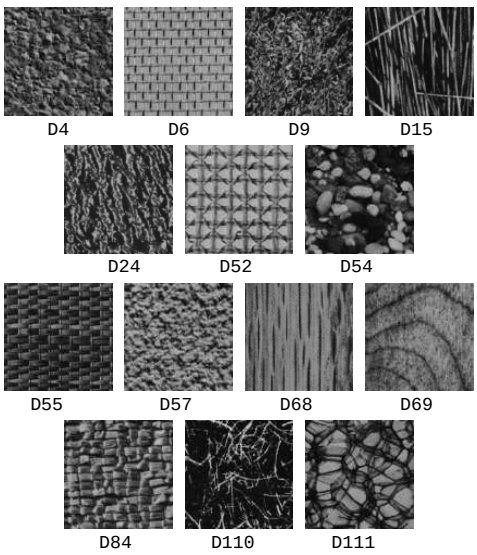
4.4 画像識別実験結果

表 1に従来手法の 1 個の構造要素による識別実験結果及びに本手法の複数の構造要素による識別実験結果を示す。テクスチャ画像識別実験の結果、1 個の構造要素による識別率が 89 %、複数の構造要素を用いた手法の識別率が 99 %となった。鼻形状による個人識別実験の結果は、同様に 73 %と 85 %となった。また、両方の識別実験を複数回繰り返した結果についてもほぼ同様に、1 個の構造要素を用いる場合に比べて 10 ∼ 20 %ほど識別率が向上し、大幅に改善した。

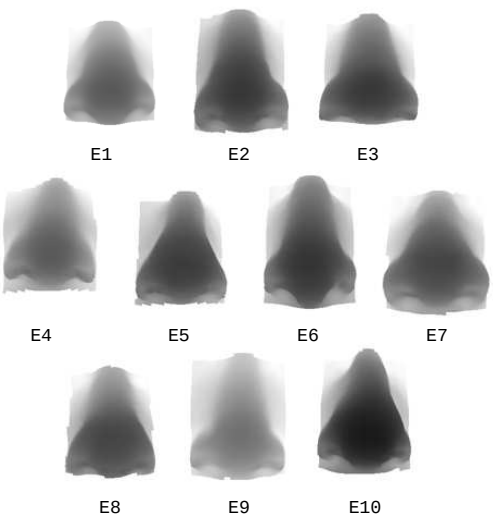
4.5 画像識別実験結果の考察

4.5.1 GA の学習回数に対する識別率の変化

図 5,6にテクスチャ画像識別及び鼻形状による個人識別における GA の学習回数に対する識別率を示す。こ



(a) テクスチャ画像



(b) 鼻の形状

図 4: 識別実験に用いた辞書画像 (Brodatz のテクスチャ画像 14 枚)

表 1: テクスチャ画像識別の実験結果 (表の数値は識別率を表わす。また、括弧内は入力画像の枚数に対する正解した画像の枚数を表わす。)

識別実験に用いる画像	Brodatz のテクスチャ画像	鼻形状
1 個の構造要素による画像識別	89 % (250/280)	73 % (44/60)
複数の構造要素による画像識別	99 % (276/280)	85 % (51/60)

の図から学習回数を増やすと識別率が向上していることが分かる。また、複数の構造要素による画像識別手法は、1 個の構造要素による画像識別手法よりも全体的に高い識別率となっている。学習回数に対する識別率は、20 回ほどの学習で収束しており、学習回数は 30 回程度で十分であることを示している。

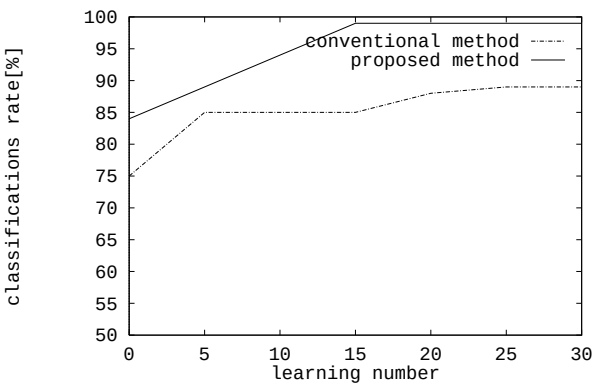


図 5: テクスチャ画像識別における GA の学習回数に対する識別率の変化

4.5.2 グルーピングされる辞書の性質

図 7に複数の構造要素によるテクスチャ画像識別及び鼻形状による個人識別で階層的にグルーピングされる辞書を示す。

(1) テクスチャ画像については、主観的に見て同じ性質を持った辞書がグルーピングされていることが分かる。例えば、図中の D6,D52,D68,D69 でグルーピングされた辞書は全体的に輝度値が高いもので、D4,D9,D24 でグルーピングされた辞書は細かな粒子から構成されるものである。また、その他のグループを見ても似た性質を持つ辞書の集まりであることが分かる。この様にグループ間の相関の距離を最大とする構造要素による画像識別は、同じようなテクスチャ画像をグルーピングする性質を持っており、似た性質を持った辞書のグループであっても、これらの距離を最大とする分割が行われることによって、識別率が向上するといえる。

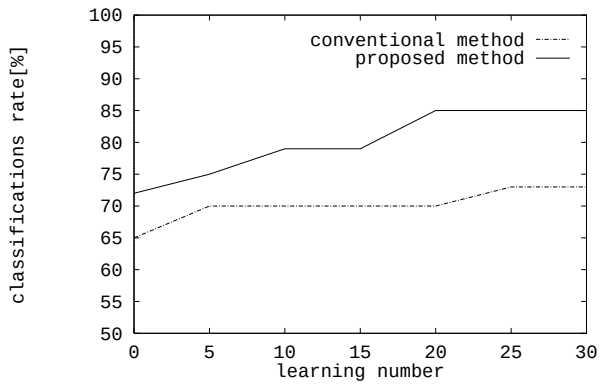


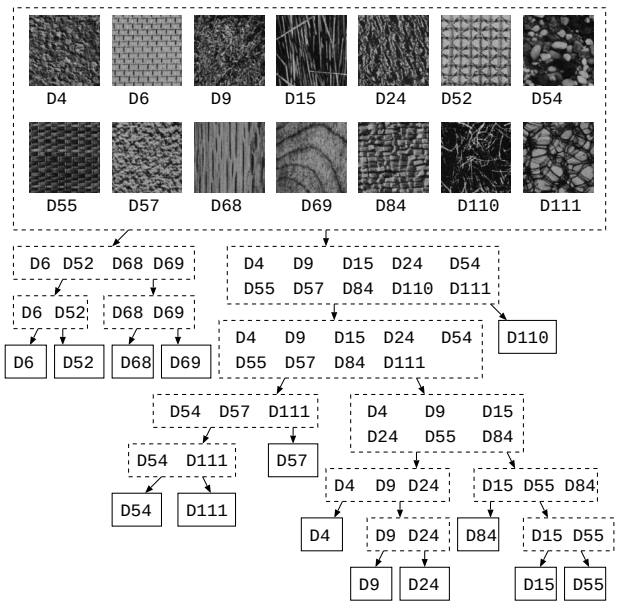
図 6: 鼻形状による個人識別における GA の学習回数に対する識別率の変化

(2) 鼻形状の画像の場合は、一概に同じ性質を持った画像がグルーピングされているとはいえない。例えば、どのグループに属する画像も鼻の幅や高さや奥行きに関して規則性はない。これはパターンスペクトルの相関を距離尺度に用いることにより、グルーピングの際に輪郭などの形状に関する情報が考慮されない為ではないかと考えている。

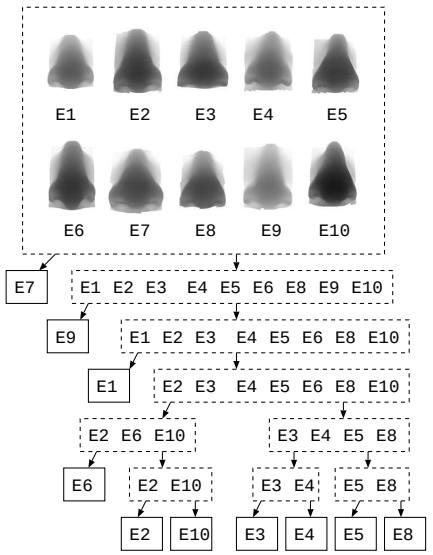
これらの結果から複数の構造要素によるグルーピングでは、画像全体が規則的な “きめ” から構成される性質を持つテクスチャ画像に関しては、各階層で同じ性質を持つようにグルーピングできている。また、全体的な模様ではなく、形状そのものに特徴のある画像は、主観的に見て同じ性質の画像でグルーピングするという特性は見られなかった。

4.5.3 識別に誤った画像の例

図 8に辞書画像である D68 を誤って識別した D69 に属する 3 枚の入力画像を示す。この図から分かるように D68 と D69 の画像はどちらも木目のある同じ性質を



(a) テクスチャ画像の辞書のグルーピング



(b) 鼻形状画像の辞書のグルーピング

図 7: 複数の構造要素による階層的な辞書のグルーピング (点線の矩形はグループを表わし、実線の矩形は最終的に分割された辞書を表わす。)

持った画像であるため、識別を誤ったものといえる。同様に図 9に辞書画像である E4 を誤って識別した E7,E1 に属する 2 枚の入力画像を示す。鼻形状の場合は、テクスチャ画像の場合と異なり、主観的に見て同じ性質を持つ画像とは言えないが、距離尺度である相関は非常に似た性質を持つことがわかる。これらの結果から、テクスチャ画像では距離尺度の近い画像と主観的に似ている画像が一致するのに対して、鼻形状では必ずしも、主観的に似ている画像が、距離尺度で近い関係にあるとはいえない。

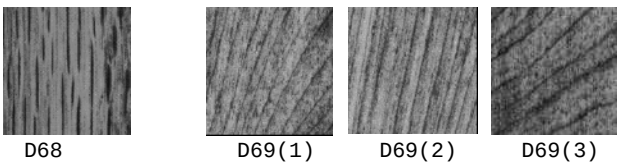


図 8: 辞書画像 D68 を誤って識別した 3 枚の入力画像 D69(1),D69(2),D69(3)

図 9: 辞書画像 E4 を誤って識別した 2 枚の入力画像 E7(1),E1(2)

5 まとめ

本稿ではパターンスペクトルによる画像識別の識別精度を向上させる為、従来の最適化した 1 個の構造要素による画像識別手法を改良して、複数の構造要素による階層的な画像識別手法を提案した。これより従来 1 個の構造要素では十分広めることが出来なかった辞書パターンスペクトル間の距離を、複数の構造要素により階層的に広めることが可能になった。

識別実験の結果、テクスチャ画像及び鼻形状による個人識別の識別率が 10 ～ 20 %向上し、これより本手法の有効性が確認できた。また、解析の結果、グルーピン

グされる画像は各階層に似たものが集まっていることがわかった。

今後の課題としては、今回提案している手法はあらかじめ複数の構造要素を最適化しておかなければならないことから、最適化の処理に時間がかかる問題が存在する。特にグルーピングの際に、辞書が均等に分配されない場合、識別の際に必要な構造要素は (辞書画像数 -1) となる。今後は分割するグループの数を増やすなどして、この問題を解決していきたいと考えている。

参考文献

[1] P.Maragos, "Pattern Spectrum and Multiscale Shape Representation," IEEE Trans., PAMI-11-7, pp.701-716, 1989.

[2] 田口 亮, 轡田 実, "パターンスペクトルによる回転不変テクスチャ識別法に関する一提案," 電子情報通信学会論文誌, Vol.J79-A, No.4, pp.973-980, 1996.

[3] 結城義徳, 大宮邦彦, 篠原克幸, 南 敏, "モルフォロジー演算に基づく耳介の特徴抽出および個人識別への応用," テレビジョン学会技術報告, Vol.18, No.27, pp.13-18, 1994.

[4] 北田基樹, 小谷一孔, "顔部品画像のモルフォロジー解析と個人識別," 情報処理学会研究報告,Vol.98-CVIM-109, No.5, pp.65-72, 1998.

[5] 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応用," 情報処理学会研究報告, Vol.98-CVIM-109, No.5, pp.57-64, 1998.

[6] 竹内伸也, 剣持雪子, 小谷一孔, "パターンスペクトルを用いたテクスチャ画像の判別に最適な構造要素の設計法," 映像メディアシンポジウム (IMPS99) 予稿集, pp.77-78, 1999.

[7] 間瀬 茂, 上田修功, "モルフォロジーと画像解析 [1][2]," 電子情報通信学会誌, Vol.74-2, pp.166-174, 1991. および 74-3, pp.271-279, 1991.

[8] 小畑秀文, "モルフォロジー," コロナ社,1996.

[9] P. Brodatz, "Textures;A Photographic Album for Artists and Designers," Dover Publishing Co., 1966.