

非線形フィルタバンクによるテクスチャ画像の クラスタリング

田村仁[†] 阿刀田央一^{**}

[†]東京理科大学経営学部
tamura@ms.kuki.sut.ac.jp

^{**}東京農工大学大学院生物システム応用科学研究所
atoda@cc.tuat.ac.jp

あらまし 本研究はテクスチャ解析のアプローチとして、「ざらざらした」「ごわごわした」などの人間の視覚の感覚量を抽出することを考える。そして均一なテクスチャの画像に対し、人間の感覚量を抽出する非線形フィルタバンクによって、人間の感覚に近いクラスタリングをおこなうことを目的とする。このための非線形フィルタとしての7種類のプログラミックフィルタを設計、実現について述べる。これを一様なモノクロテクスチャ画像に対して適用し、各フィルタに対応した特徴をもった画像のクラスタリングをおこなった。その結果について述べ、さらにいくつかの新たなフィルタについて提案をおこない、フィルタの種類や数が最適なフィルタバンクの構築が必要性を論じる。

キーワード テクスチャ解析, 非線形フィルタ, 画像のクラスタリング, プログラミックフィルタ

A Non-Linear Filter Bank for Clustering of Texture Images

TAMURA, hitoshi[†] ATODA, oichi^{**}

[†]School of Management,
Science University of Tokyo
tamura@ms.kuki.sut.ac.jp

^{**}Graduate School of Bio-Applications and Systems Engineering,
Tokyo University of Agriculture and Technology
atoda@cc.tuat.ac.jp

Abstract We propose a method which clustering texture images. We design and implement non-linear filters which pass some feelings or a specified impression of texture images, for example "coarse texture", "smooth texture" and so on. 7 filters are presented in this paper, and these filters are organized a filter bank for clustering of texture images. We apply the non-linear filter bank to uniform, monochrome and 256×256 pixel, texture images. As a result of this experiment, we obtain clusters which have specified human feeling. But there are many images are not belong to any clusters. We propose ideas of new non-linear filters, and discuss about them.

Keywords texture analysis, non-linear filters, clustering of images, programic filters

1. はじめに

テクスチャ解析は古くからおこなわれており、大きく構造的手法と統計的手法に分けられる。構造的手法はテクスチャ要素を切りだし、大きさや形、配置規則などでテクスチャを記述する[1]。これは人工的なテクスチャには有効であったが、自然のテクスチャからの要素の抽出が困難だった。

後者はパワースペクトルなど統計的性質によってテクスチャを記述する[2]。しかし周波数成分等ではテクスチャの特徴すべてを捉えられない。人間の眼は別の特徴にも影響される。

両者とも小さな物体がランダムに堆積してできたようなガウス性でない雑音画像に対して弱い。しかし人間はそれに対して、ガサガサした、ぱさぱさした等の感覚を得る。また人間の目にはポジ画像、ネガ画像が大きく異なる印象を与えることから、何らかの非線形な処理が必要となる。この人間の感覚を抽出するために本研究では次章で述べるプログラミックフィルタを利用する。

我々はこれまでに非線形フィルタを論理的に構築する枠組として「プログラミックフィルタ」を提案している[3]。本研究ではこれを利用し、テクスチャへの人間の感覚に近い特徴を抽出する非線形フィルタを設計する。例えば、ざらざらした感じを抽出するために、小領域の角を抽出するフィルタなどを用意する。この非線形フィルタをいくつも用意してフィルタバンクを構築する。

本研究の目的は、均一なテクスチャの画像を、非線形フィルタバンクによって人間の感覚に近いクラスタリングすることであり、有効な非線形フィルタの設計、実現をおこなう。

2. プログラミックフィルタ

2.1 プログラミックフィルタの概要

人工視覚系の前処理である平滑化やエッジ抽出などには、線形フィルタリング技術がよく使われているが、概して性能は良くない。それより性能の良い非線形フィルタもかなり提案され、分類されている[4]。それらは特定目的には優れた性能をもつが、線形フィルタとは異なり、多様な画像処理に対して必要に応じて新たなフィルタを簡単に設計活用するような技術にはなっていない。

しかし画像処理に限れば、設計すべきフィルタについて明確な目的と要求性能があり、入力画像の性質や望む出力がわかっている。我々は非線形フィルタに対して、「擬線形性」の制約に基づく形式的な生成規則を与えた。この規則に従って生成されたフィルタを「プログラミックフィルタ」と呼び、非線形な処理を施してもコントラストなど必要な性質を保存することを「擬線形」と呼ぶ。

2.2 フィルタの設計方針

人間がテクスチャを判別する条件を系統的に分析することはかなり困難な作業である。多数の被

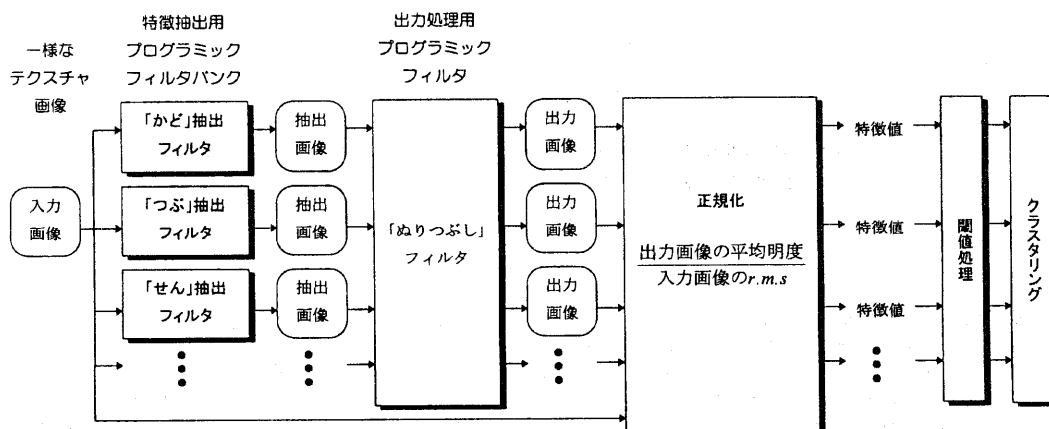


図1 非線形フィルタバンク処理の流れ

験者にテクスチャ画像のクラスタリングをおこなわせる実験で知見を得られるだろうが、時間と手間がかかる。ここではまず手かがりとして、微小領域中の形状や統計量に関するフィルタを設計する。ただし今回は色彩については考慮せず、モノクロ濃淡画像に範囲を絞る。

ここで述べるフィルタはすべて前述のプログラムミックフィルタ生成規則を守っている。

2.3 「かど」フィルタ

人間はある画像に対して「ざらざらした」感覚をもつ。これを抽出するために、角ばった特徴に反応するフィルタを設計する。これが人間の感覚に対応しているかは、今後実験で確かめる必要があるが、仮にこれを「かど」フィルタと呼ぶ。

放射状に8分割するマスクを用意して、マスク領域内で2値化する。8方向のどれか1つに明るい暗い部分が偏った時に、「かど」とみなす。

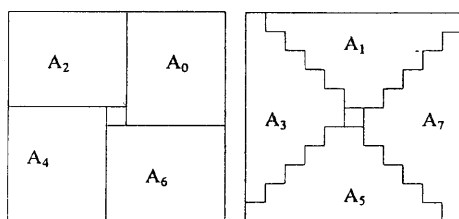


図2 放射状マスク

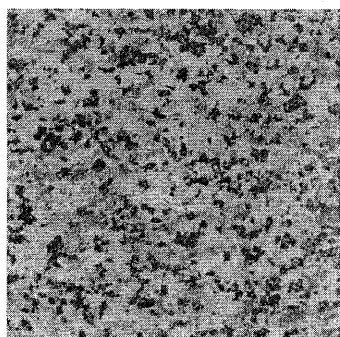


図3 「かど」のある画像

またポジ画像、ネガ画像を区別するために、「白かど」と「黒かど」を別に抽出する。マスク領域内の最大値 x_{\max} と最小値 x_{\min} から閾値 x_t を、

$$x_t = x_{\min} + (x_{\max} - x_{\min}) \times P_t$$

と決める。 P_t は「黒かど」検出用に 0.7、「白かど」には 0.3 程度とする。各閾値を x_{tw} と x_{tb} と呼ぶ。

次に2値化した画像において「白かど」を検出ならば白い画素が8方向中どれにもっとも偏っているか調べ、「かど」方向を求める。

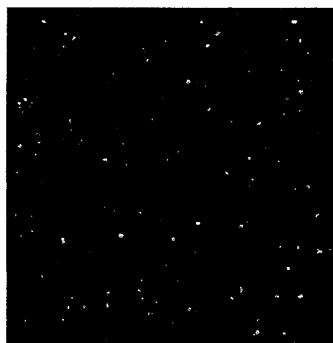


図4 「白かど」検出結果

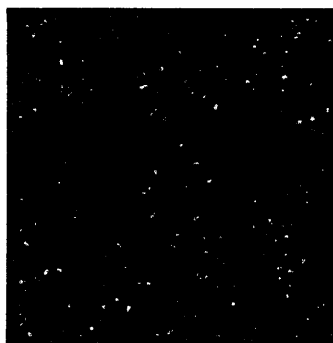


図5 「黒かど」検出結果

ここでマスクサイズに適合した粒子と区別するための条件をつける。2値化画像に対し、マスク中心が「白かど」を抽出するならば白、「黒かど」ならば黒かどうか判定し、中心から塗りつぶしをおこなう。塗りつぶしが「かど」方向の外周まで到達できなければフィルタの値を0とする。

さらに「かど」らしさを判定するためにコントラストを見る。「かど」方向の隣接領域、例えば図2の A_1 に対しては A_3 と A_7 の平均明度を求め、「黒かど」に対してはこの平均明度が x_{tw} より高く、「白かど」では x_{tb} より低くなければならない。これを満たさない場合もフィルタ出力を0とする。

この2つの条件を満たした上で、隣接領域の平均明度を \bar{x}_b 、それに含まれる角成分の画素数を

n_b と同様に角方向の平均明度を \bar{x}_c 、画素数を n_c とする。そして1つの領域の画素数を N_8 として、

$$F_{corner} = \{(n_c - n_b)/N_8\}^2 \times (\bar{x}_c - \bar{x}_b)$$

を出力とする。

2.4 「つぶ」フィルタ

粒状の「つぶ」を検出するフィルタを設計する。

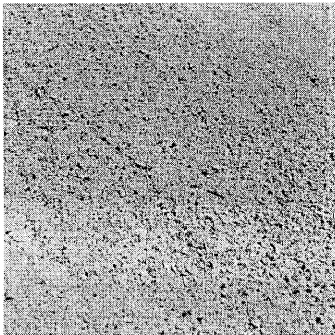


図6 「つぶ」のある画像

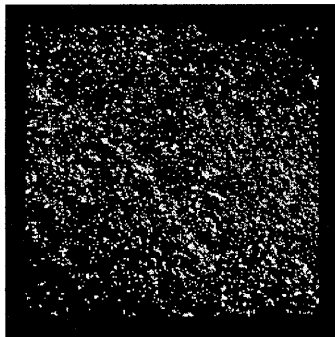


図7 「つぶ」検出結果

まずマスク内の最大値と最小値の間を閾値 x_t とする。「かど」と同様、 x_t でマスク内を2値化し、マスク中心から塗りつぶす。もしマスクの外周に塗りつぶしが到達しなければ、「つぶ」だと判断し、次の式を値とする。それ以外は出力0とする。

$$F_{grain} = |\bar{x}_g - \bar{x}_b|。$$

塗りつぶした領域(2値化前)の平均値とマスク内の塗りつぶし外領域の平均値の差である。

この「つぶ」フィルタでもポジ、ネガを分けるべきであったが、今回の実験では分けていない。

2.5 「せん」フィルタ

次は細長い「せん」を検出する。「つぶ」フィルタと同様に2値化してマスク中央から塗りつぶしをおこなう。この「せん」フィルタではマスク4辺のうち、対向する2辺だけに到達したものを選択する。それ以外はフィルタ出力値を0とする。

この条件を満たせば、次をフィルタ値とする。

$$F_{line} = |\bar{x}_l - \bar{x}_b|。$$

このフィルタも今回はポジネガの区別をしない。

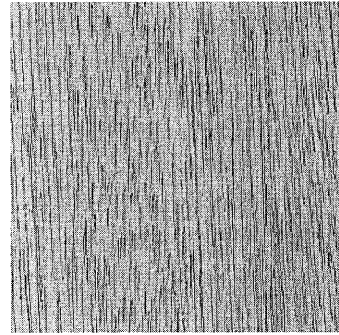


図8 「せん」のある画像

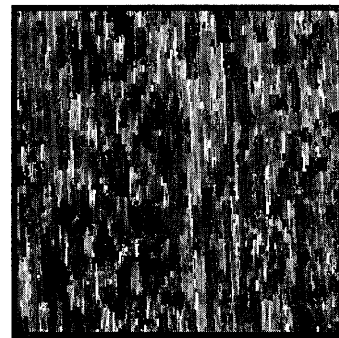


図9 「せん」検出結果

2.6 「たいら」フィルタ

2.3から2.5までは、マスク内の形状を抽出するフィルタである。プログラミックフィルタでは他にマスク内の統計的性質を抽出できる。

ここでは明度変化の少ない「たいら」な面を検出するフィルタを設計した。これはマスク内の平均値 \bar{x} と標準偏差 σ を計算し、明度が $\bar{x} \pm 0.2 \times \sigma$ 以内に収まるマスク内の画素数 n を

求める。マスク内の全画素数 N から正規分布での期待値を求め、 n と差をとる。またコントラスト成分として σ も使い、次の

$$F_{plane} = (n - 0.1554 \times N) \times 4 \times \sigma$$

をフィルタの値とする。これにより、明度変化の少ない「たいら」な領域を抽出する。

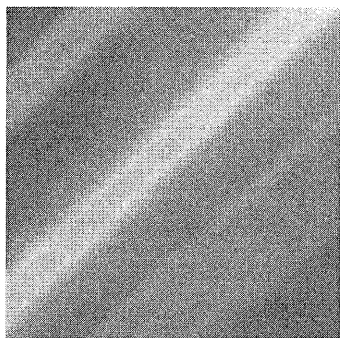


図10 「たいら」のある画像

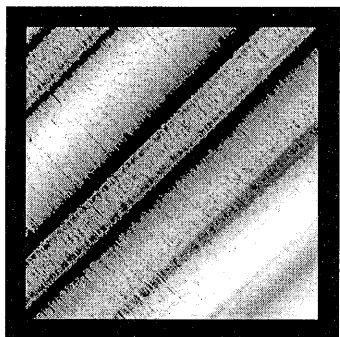


図11 「たいら」検出結果

このフィルタではポジネガの区別は必要がない。ただし「たいら」な分布だけでなく、コントラストが強い分布を抽出することも考えられる。

2.7 「しま」と「こうし」フィルタ

縞模様や、格子模様を抽出するフィルタを作成する。基本的アイデアは明度変化が繰り返し生じた場合にフィルタの出力を高くすることである。この「しま」は、周期が一定でなくともよい。このフィルタの原形は文献[3]に示してある。

このフィルタを縦、横、ななめの「しま」(左右両方)を検出する4つ用意する。そしてそれぞれの処理結果全体の平均明度によって順位づけ、最大

値をもった結果を「しま」フィルタの出力とする。また次に大きい値をもった結果を「こうし」フィルタの出力とする。つまり少なくとも2方向以上で「しま」を有するものを「こうし」と判断する。

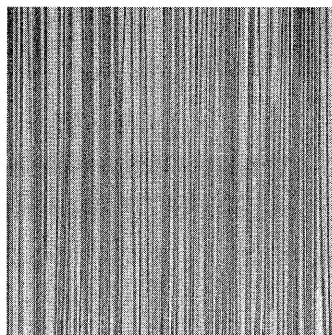


図12 「しま」のある画像

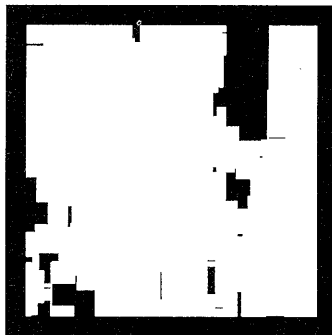


図13 「しま」検出結果

2.8 フィルタの出力処理

各フィルタはの出力結果は、「かど」や「つぶ」に反応した部分が白い点として現れている。このフィルタ出力を、さらに塗りつぶすことで、フィルタに反応している部分を領域化する。

この塗りつぶしも、プログラミックフィルタとして実現する。「かど」フィルタで使用した放射状8方向マスクを使い、出力結果の画像に対して8方向それぞれ最大値を求める。この8つの最大値のなかから4番めに大きい値をフィルタ出力とする。つまり、フィルタに反応した点がマスク範囲で4つ以上存在する頻度で分布していれば、その領域を塗りつぶす。今回は一様なテクスチャ画像に適用したため、画像全体が塗りつぶされる。



図14 図9を出力処理した画像

ただし、「しま」フィルタと「こうし」フィルタ、「たいら」フィルタは、フィルタ出力自体が塗りつぶしを含むため、この出力処理はおこなわない。

3. クラスタリング実験

3.1 実験の対象

サンプル画像として、著作権フリーの CD-ROM からの 1200 枚を使用した。これを 256×256 画素に切り出し、平均明度 128、分散 32 になるよう処理する。「擬線形」性を有するプログラミックフィルタは、入力の前平均明度や全体のコントラストに影響されないが、多数の出力を同列に比較するため、前もって全画像を正規化する。

この 1200 画像に対して「黒かど」「白かど」「つぶ」「せん」「しま」「こうし」「たいら」の 7 つのフィルタを適用する。マスクサイズは「かど」「せん」に 11×11、それ以外を 33×33 とした。

そして 1200 画像の中から一様なテクスチャ画像 724 枚を選択し、クラスタリングをおこなった。一様なテクスチャ画像とは各フィルタ出力を出力処理した結果に対し、上下左右に 2 分割ずつ 4 分割して各領域で出力がばらつかないもの、4 つの領域の平均明度の最大が 20 以上のとき、最小が最大の 60% 以上であることを基準とした。

3.2 実験結果

724 枚の入力画像に対し、図 1 の処理をおこなう。出力画像から得た特徴値を、ある閾値によって各フィルタに反応したかどうかの 2 値にする。

「かど」「せん」「つぶ」など形状に関するフィル

タの閾値には、入力画像と周波数成分だけがよく似たガウス性雑音画像に対する各フィルタの特徴値を使用する。次のように生成した雑音画像からは微小領域の形状などの情報は落とされている。

ある入力画像を 2 次元 FFT によって周波数画像に変換する。実部と虚部をそれぞれ独立に処理し、それぞれの周波数画像に対して 9 近傍の移動平均をとって周波数分布をなめす。そしてそれぞれの周波数画像に -1 から 1 の間の乱数を乗じる。ただしフィルタのマスクサイズより大きい低周波の部分には、例外的に値が小さくなるように処理する。この結果の周波数画像を逆 FFT によって画像に変換し、ガウス性の雑音画像を生成する。

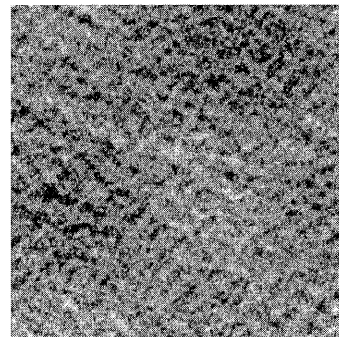


図15 図3から生成したガウス性雑音画像

このガウス性雑音画像を 724 枚生成し、各フィルタに適用した結果の最大特徴値の 150% の値を閾値とする。また、「たいら」フィルタに関しても、この雑音画像に対する特徴値を閾値にできる。

次に、「しま」「こうし」フィルタに対しては適当な基準が存在しないため、724 画像中フィルタの出力が最大の値を基準に、その 50% 以上のものをフィルタに反応したと仮定する。

表 1 に各フィルタの出力の最大値等を示す。

表1 出力結果

フィルタ	特徴値の最大値	ガウス画像最大値	閾値	反応数
黒かど	171.8	20.7	31.05	70
白かど	193.4	18.6	27.921	55
つぶ	90.1	36.6	54.942	64
せん	175.6	149.4	87.8	23
たいら	144.1	51.6	77.327	50
しま	193.5	126.9	96.75	45
こうし	185.0	185.0	92.5	11

注) 「せん」の閾値は「しま」に反応しないガウス画像の最大値

「しま」の中で周期性のある縞模様は周波数成分が特徴の1つなため、ガウス性雑音画像を作っても元画像と同程度に反応する。実は「せん」フィルタもこの周期性の縞模様に反応するが、閾値としては、「しま」に反応しない画像から生成したガウス性雑音画像からの特徴値を使用する。

3.3 画像のクラスタリング結果

各フィルタに反応したかどうかを1と0として本来ならば7次元の多変量解析をおこなう。使用するクラスタリング手法については、フィルタ出力の性質や、入力画像の性質などと絡んだ議論が必要になるが、今回は比較的次元数が少ないため、ベクトルが完全に一致したものだけを単純にクラスタとしてまとめた。

表2 クラスタリング結果

No	L	BC	WC	G	S	C	P	n
1								472
2							○	46
3				○				53
4			○					27
5		○						41
6		○		○				2
7		○	○					21
8	○							11
9	○					○		7
10	○						○	2
11						○		15
12				○		○		3
13			○			○		3
14		○				○		2
15	○			○		○		2
16		○	○		○	○		2
17		○	○			○		2
18				○	○	○		2
19					○	○		7

注) No:クラスタ番号、L:線、BC:黒角、WC:白角、G:粒、S:縞、C:格子、P:平面、n:クラスタ画像数

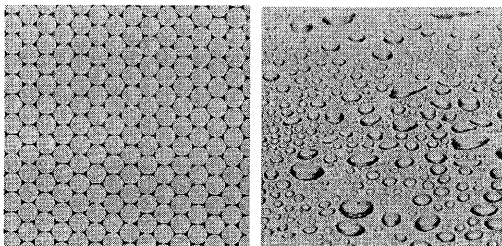


図16 クラスタになった画像例(1)

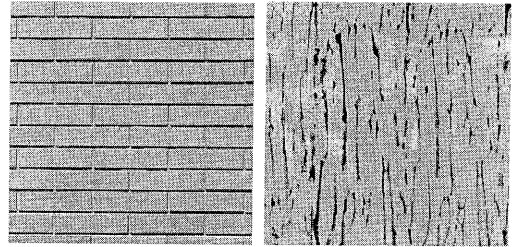


図 17 クラスタになった画像例(2)

4. 実験結果の検討

実験の結果、人間の目から見てある程度似ている画像をクラスタリングできた。ただし、どれにも反応しない、特徴を抽出できなかった画像が724画像中472枚も存在する。これは主にフィルタの種類不足が原因だと考えられる。今後フィルタの数を増やして問題が解決するかわからないが、まだ特徴の数が足りないことは確かである。

2.で述べたようにポジ画像とネガ画像に対するフィルタを「かど」検出フィルタ以外はまだ用意していない。しかし「かど」の結果からみて両者の区別は有効である(図18と図19参照)。

「かど」など形状に関するフィルタはこれ以上種類は増やせない。複雑な形状に反応するフィルタは、パターン認識の手法に近くなり、解像度や傾いたパターンへの対応が必須となる。低水準なテクスチャ解析では必要以上に重い処理となる。

対して、「たいら」のような統計的処理を使ったフィルタは、さらに応用が可能である。また「しま」では、黒地に細い白縞など非線形な特徴を抽出するフィルタが考えられる。

また今回の実験では各フィルタに対して適当な1種類のマスクサイズしか適用しなかったが、マスクサイズをかえることで、また別の特徴を捕らえられる可能性もある。ただし、マスクサイズには入力画像の大きさから制限が付き、また画像の解像度の問題をはらむ。現在は入力画像中の特徴がマスクサイズに適合することを条件とした。

実はすでにくつか別のマスクサイズで適用した結果で次元数を増やし、クラスタリング実験もおこなったが、結果がかなり荒れてしまった。人

間から見ると1つのクラスタになるべき画像群がいくつかのクラスタに分裂してしまうのである。

このようにフィルタ数が多くなると別の問題が発生する。これに関してクラスタリング手法の検討が必要である。現在は完全に一致したベクトルだけをまとめたが、次元数が増えれば「近い」ベクトルをクラスタリングする方法が必要である。

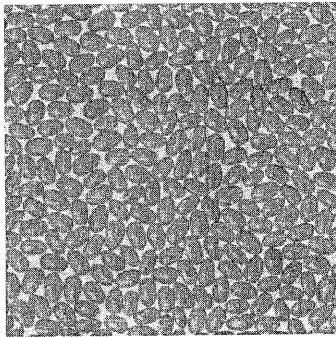


図18 白かどの例

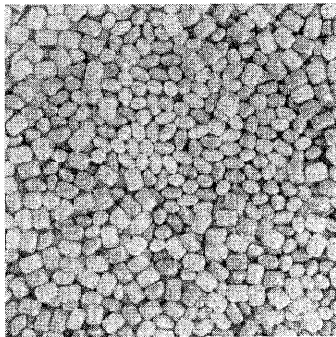


図19 黒かどの例

ただし各次元の重みを検討しなければならない。例えば図18と図19では、「かど」よりも「丸いもの」に人間が大きく影響を受けると仮定したら、この両者の距離は近いと判断すべきである。フィルタが捉えた各特徴の重みを考慮して、ベクトル間の距離を正しく計算することが必要になる。また、フィルタの種類を増やしても他のフィルタと似た特徴を抽出しては意味がない。

このため、十分な数に人間に対してテクスチャ画像に対する感覚量を測定する実験が必要である。そして実験結果を主成分分析等で検討し、フィルタの種類が最適なテクスチャ判別用プログラ

ミックフィルタバンクを構築する必要がある。

また、このプログラミックフィルタでは抽出できない特徴に大域的な情報がある。人間は、図18のような画像に対して、後ろに白地があることを認識する。しかしマスク処理ではその情報は得られない。他方式との連携を検討する必要がある。

5. まとめ

本研究の目的は、均一なテクスチャ画像を人間の感覚に近くクラスタリングすることであった。このテクスチャ解析のアプローチとして、「ざらざらした」等の人間の感覚量を抽出することを考え、これに有効な非線形フィルタを実現した。

非線形フィルタは、プログラミックフィルタと呼ぶ枠組みから構築し、7つのフィルタを実現した。これを一様なモノクロテクスチャ画像724枚に対して適用し、各フィルタに対応した特徴をもった画像のクラスタリングに成功した。

しかし、まだ特徴を抽出できない画像が多く残り、今後フィルタの数を増やすことが必要であった。このためには、十分な数に人間に対して、テクスチャ画像に対する感覚量を測定するための実験などを通して、フィルタの数や種類が最適なテクスチャ判別用プログラミックフィルタバンクを構築する必要があることがわかった。

参考文献

- [1] 富田ほか:構造解析によるテクスチャの分類, 信学論(D),Vol.J61-D,No.1,pp457-464,1978.
- [2] 金子ほか:マルコフ場モデルによるテクスチャ解析: 信学論(D),Vol.J67-D,No.9,pp.1005-1012,1984.
- [3] 阿刀田ほか:プログラミックフィルタ-高性能低水準画像処理用非線形フィルタ作成のための一方式, 信学論(D-II),vol.J79-D-II,No.6,pp. 1034-1045,1996.
- [4] 荒川ほか:非線形デジタルフィルタとその応用,信学誌,Vol.77,No.8,pp.844-832,1994
- [1] 五十嵐ほか:堆積型テクスチャの性質とその解析,計測自動制御学会論文集,Vol.21,No.6, pp.592-599,1985.