

## バイスペクトルに基づく2次元及び3次元画像の相似変換に対する不変特徴

堀川 洋<sup>†</sup>

Feature of 2D and 3D Images Invariant to Similarity Transformations Based on the Bispectrum

Yo HORIKAWA<sup>†</sup>

あらまし 2次元及び3次元パターンの相似変換に対して不変な特徴を、バイスペクトルに基づいて構成した。不変特徴は画像中の3組の周波数成分の積のヒストグラムとして、フーリエ変換画像から直接に計算することができる。そして、正弦波パターン、2次元テクスチャ画像、及び、3次元人体頭部画像を対象とした識別実験を行い、不変特徴は、任意の位置ずれ・回転、及び、2倍程度までの拡大・縮小に対して有効であり、また、0dB程度までの雑音に対する頑健性を有することを示した。

キーワード 不変特徴, 相似変換, バイスペクトル, テクスチャ, 3次元画像

### 1. ま え が き

相似変換(位置ずれ, 回転, 拡大・縮小)に対して不変な2次元及び3次元パターンの認識は, 画像処理における困難な問題の一つである[1]。これまで, 構造的及び統計的手法を組み合わせた様々なアプローチがなされているが, 効率の良い識別には変換に対して不変かつ原パターンの情報を十分に保持している特徴を構成することが必要である[2]。相似変換に対する不変特徴としては(1)特徴点や境界線などの特徴成分を抽出して構成するものと(2)直接に積分型変換を用いて構成するものがある(1)としては, 例えば, 座標点の集合から点パターンマッチングを用いるもの, 境界線からフーリエ記述子やGeometric Hashingを用いるものなどがある。一方(2)の場合には, 位置ずれ, 回転及び拡大・縮小の三つの変換に対して同時に不変となる変換は存在しないため(2-1)位置ずれを重心などによって正規化した後にモーメントやフーリエ・メルン変換などによって回転と拡大・縮小に対する不変特徴を構成するものと(2-2)位置ずれに対して不変である相関関数(あるいはそのフーリエ変換

のパワースペクトルやバイスペクトル)などに対して積分型変換を用いて, 回転と拡大・縮小に対する不変特徴を構成するものとに分けられる。

これらの手法はそれぞれに特徴をもつが, 例えば, 雑音の加わったテクスチャ画像などの濃淡画像を対象とする場合には, 特徴成分の抽出や重心の計算が困難であるため(1)及び(2-1)の手法はあまり適さない。また(2-2)においても, 2次相関関数及びパワースペクトルは雑音の影響を受ける。それに対して, 3次の相関関数及びバイスペクトルは, 雑音に対して頑健かつ位相情報を保持しているという利点を有しており[3], [4], それを用いた不変特徴及び認識手法が提案され有効性が確認されている[5]~[17]。

本研究では, バイスペクトル(3次相関関数のフーリエ変換)を用いて, 相似変換に対して不変な特徴を新たに構成した。不変特徴はパターンのフーリエ変換から直接に計算することができ, 高次の特徴量を用いる際の難点であった計算量の問題を緩和している。以下, まず, 不変特徴の導出及び計算方法を述べる。そして, 2次元及び3次元の正弦波パターン, 2次元テクスチャ画像, 及び, 3次元人体頭部画像を対象とした識別実験の結果について示す。最後に, これまでの実験結果との比較及び今後の課題について記す。

<sup>†</sup> 香川大学工学部情報工学科, 高松市  
Faculty of Engineering, Kagawa University, Takamatsu-shi,  
761-0396 Japan

## 2. 不変特徴

$f(x)$  を 2 次元あるいは 3 次元の画像データとし、 $F(\omega)$  をそのフーリエ変換とする。 $f(x)$  のバイスベクトル： $B(\omega_1, \omega_2)$  は、次式で与えられ、パワースペクトルと同様に位置ずれに対する不変量となっている。

$$B(\omega_1, \omega_2) = F(\omega_1)F(\omega_2)F(-\omega_1 - \omega_2) \quad (1)$$

ここで、画像データが任意の点の回りで回転したとき、そのフーリエ変換 ( $F(\omega)$ ) は原点の回りで同じ角度だけ回転したものとなる。また、画像データが拡大・縮小したとき、そのフーリエ変換はその逆数倍に拡大・縮小される。このような原点回りの回転及び拡大・縮小による変化に対して、二つの波数ベクトル： $\omega_1, \omega_2$  の長さの比： $r = |\omega_1|/|\omega_2|$ 、及び、内角： $\cos \theta = \omega_1 \cdot \omega_2 / (|\omega_1||\omega_2|)$  は不変に保たれる。そのため、一般にバイスベクトルの任意の関数において、長さの比及び内角を一定とする（すなわち、原点を含めて相似三角形の頂点を成す）波数ベクトルの組 ( $\omega_1, \omega_2$ ) についてその関数値を足し合わせたものは、回転及び拡大・縮小に対して不変な値をとることになる（なお、拡大・縮小に対しては定数倍の変化が生じるので正規化する必要がある）。そこで、バイスベクトルを  $|\omega_1|/|\omega_2| = r$  及び  $\omega_1 \cdot \omega_2 / (|\omega_1||\omega_2|) = \cos \theta$  の条件のもとで積分した次の量を考える。

$$I'(r, \theta; g) = \iint_{r, \theta} g(B(\omega_1, \omega_2)) d\omega_1 d\omega_2$$

$$(|\omega_1|/|\omega_2| = r, \omega_1 \cdot \omega_2 / (|\omega_1||\omega_2|) = \cos \theta) \quad (2)$$

ここで、 $g(B)$  は、バイスベクトル値の任意の関数である。この  $I'(r, \theta; g)$  は、画像の周波数成分のうち、長さの比が  $r$ 、内角が  $\theta$  であるような波数ベクトルの組のバイスベクトルの大きさを表している。そして、これを正規化することによって、相似変換に対する不変特徴： $I(r, \theta; g)$  が得られる。

$$I(r, \theta; g) = I'(r, \theta; g) / \left( \int_0^\infty \int_0^{2\pi} I'^2(r, \theta; g) dr d\theta \right)^{1/2} \quad (3)$$

なお、この不変特徴は、画素値の線形変換に対しても不変である。

次に、バイスベクトルの推定及び補間に伴う計算量の問題を回避するために、以下のような不変特徴の

計算方法を用いる。ここでは簡単のため、2 次元画像データについてのみ示す。

$f(x)$  ( $x = (x, y)$ ,  $x, y = 0, \dots, N-1$ ) を、 $N \times N$  画素からなる画像データとする。そして、そのフーリエ変換： $F(\omega)$  を、ガウス窓： $\exp(-(x-\mu) \cdot (x-\mu)/(2\sigma^2))$  ( $\mu = (N/2, N/2)$ ,  $\sigma = N/3$ ) を用いて FFT によって求める。ここで、すべての周波数ベクトルの組 ( $\omega_1, \omega_2$ ) ( $\omega_i = (\omega_x, \omega_y) - N/2 \leq \omega_x, \omega_y < N/2$ ) について、3 重積： $F(\omega_1)F(\omega_2)F(-\omega_1 - \omega_2)$  の値を計算し、それを関数： $g$  で変換した値を、適当に分割した ( $r, \theta$ ) 平面上の対応する区画に足し合わせていく ( $|\omega_1|/|\omega_2| = r$  及び  $\omega_1 \cdot \omega_2 / (|\omega_1||\omega_2|) = \cos \theta$  である)。その結果得られるヒストグラムを、 $I'(r, \theta; g)$  の推定値として用いる（なお、ガウス窓は画像周辺部の不連続性の影響の軽減のために用いたが、その有無は不変特徴の推定値にはほとんど影響を与えなかった）。本手法においては、時間的には、画像のフーリエ変換から各波数ベクトルの組 ( $\omega_1, \omega_2$ ) についてバイスベクトルの値を求めるために  $N^2 \times N^2$  回に比例する計算が必要である。一方、空間的には、原画像及びそのフーリエ変換を保持できればよいのでメモリサイズは  $N^2$  に比例する程度ですむ。そのため、必要な時間計算量は  $O(N^4)$ 、空間計算量は  $O(N^2)$  のオーダーで押さえられる。なお、一般に、 $d$  次元画像 ( $N^d$  画素) の場合、時間計算量： $O(N^{2d})$ 、空間計算量： $O(N^d)$  となる。

## 3. 計算機実験

### 3.1 正弦波パターン

バイスベクトルに基づく不変特徴は、原画像の位相情報の一部を保持している。その有効性を見るために、次のような位相の異なる 2 次元複合正弦波パターンを対象として識別実験を行った。

$$(a) f(x, y)$$

$$= \cos(2\pi x/N) + \cos(2\pi y/N)$$

$$+ \cos(2\pi(x+y)/N) + \cos(2\pi(x-y)/N)$$

$$+ \cos(4\pi x/N) + \cos(4\pi y/N)$$

$$+ \cos(4\pi(x+y)/N) + \cos(4\pi(x-y)/N)$$

$$(b) f(x, y)$$

$$= \sin(2\pi x/N) + \sin(2\pi y/N)$$

$$+ \sin(2\pi(x+y)/N) + \sin(2\pi(x-y)/N)$$

$$+ \sin(4\pi x/N) + \sin(4\pi y/N)$$

$$\begin{aligned}
 & + \sin(4\pi(x+y)/N) + \sin(4\pi(x-y)/N) \\
 \text{(c) } f(x, y) & = \sin(2\pi x/N) + \sin(2\pi y/N) \\
 & + \sin(2\pi(x+y)/N) + \sin(2\pi(x-y)/N) \\
 & + \cos(4\pi x/N) + \cos(4\pi y/N) \\
 & + \cos(4\pi(x+y)/N) + \cos(4\pi(x-y)/N)
 \end{aligned}$$

ここでは、 $N = 64$  とした。図 1 に、それぞれの不変特徴： $I(r, \theta; g)$  を示す。なお、 $F(-\omega) = F^*(\omega)$  より

$I(r, \theta; \text{Im}(B)) = 0$  となるので、ここでは、実部及び虚部の絶対値： $g(B) = (|\text{Re}(B)|, |\text{Im}(B)|)$  を用いた。また、不変特徴の計算にあたっては、 $(r, \theta)$  平面において、 $r$  については 1.0 から 10.0 の範囲で  $\log r$  のスケールで 0 から 1.0 まで 0.2 ごとに 5 分割、 $\theta$  については  $0^\circ$  から  $180^\circ$  未満まで  $18^\circ$  ごとに 10 分割したものと及び  $180^\circ$  の、 $5 \times 11 = 55$  区画のヒストグラムを求めている。なお、図 1 において、1 番手前の曲線は  $1.0 \leq r < 10^{0.2} \approx 1.6$  の範囲のもの、2 番目のものは  $10^{0.2} \approx 1.6 \leq r < 10^{0.4} \approx 2.5$  の範囲のもの、以下同様である。また、(Re) 及び (Im) はそれぞれ実部及び虚部を示す (図 2 から図 4 においても同様である)。

図 1 から、余弦波からなるパターン (a) では実部にピークを、正弦波からなるパターン (b) では虚部にピークを、また、双方を含むパターン (c) には、実部虚部共にピークを有しており、位相の違いを反映していることがわかる。

そして、表 1 に、原画像と変換画像との間の不変特徴の RMS 距離を示した。RMS 距離の計算には、図 1 に示した  $5 \times 11 = 55$  区画の実部及び虚部の絶対値のヒストグラムの値を用いた。変換画像は、原画像から線形補間によって作成した。ここでは、拡大なし ( $\times 1.0$ ) 及び 2 倍に拡大 ( $\times 2.0$ ) したものについて、0 から 0.5 画素まで 0.1 画素ごとに  $x$  軸方向にずらしたものをそれぞれ  $0^\circ$  から  $45^\circ$  まで  $5^\circ$  ずつ回転させた 60 枚を対象としている。また、原画像に平均：0、分散： $1.0^2$  のガウス性白色雑音を付加したものの (SNR  $\approx 6$  dB) 100 枚についても併せて示している。この場合、すべての変換画像及び雑音付加画像は、原画像の不変特徴との RMS 距離によって正しく識別される。

同様に、3 次元画像として、次のような位相の異なる

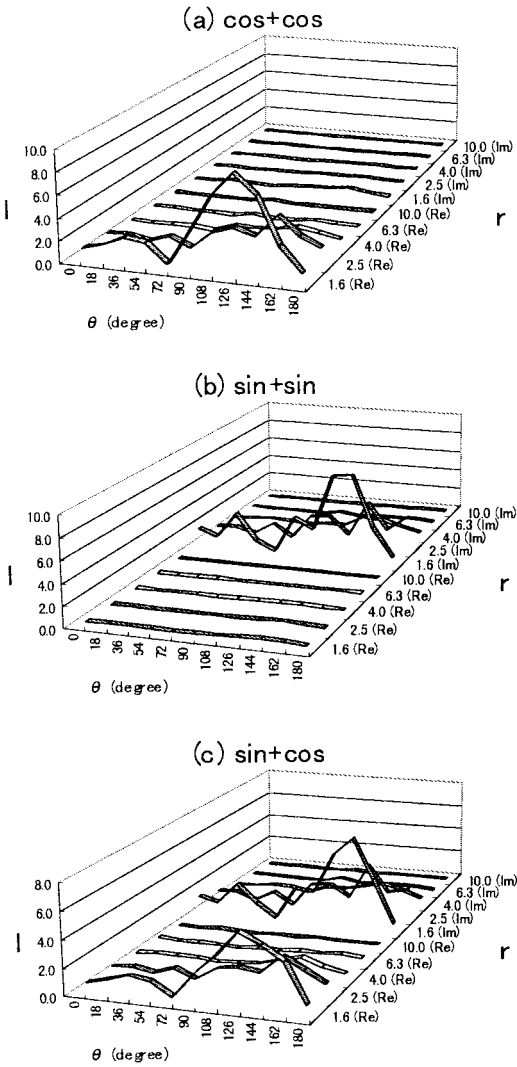


図 1 2 次元正弦波パターンの不変特徴

Fig. 1 Invariant feature  $I(r, \theta; (|\text{Re}(B)|, |\text{Im}(B)|))$  of 2D sinusoidal patterns.

表 1 2 次元正弦波パターンの相変換画像の不変特徴の RMS 距離

Table 1 RMS distance of the invariant feature of transformed 2D sinusoidal patterns.

	(a)			(b)			(c)		
	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.
(a) X1.0	0.12	0.08	0.0	1.41	1.41	1.40	0.90	0.88	0.88
(b) X1.0	1.41	1.40	1.40	0.28	0.15	0.0	0.69	0.68	0.67
(c) X1.0	0.88	0.87	0.85	0.72	0.68	0.67	0.14	0.10	0.0
(a) X2.0	0.38	0.34	0.28	1.41	1.41	1.40	0.90	0.88	0.87
(b) X2.0	1.40	1.38	1.38	0.44	0.41	0.39	0.68	0.68	0.67
(c) X2.0	0.85	0.81	0.76	0.91	0.87	0.84	0.40	0.35	0.33
(a)Noise	0.31	0.29	0.27	1.34	1.33	1.31	0.78	0.77	0.75
(b)Noise	1.37	1.36	1.34	0.22	0.20	0.18	0.60	0.59	0.57
(c)Noise	0.93	0.92	0.91	0.76	0.75	0.74	0.28	0.26	0.24

3 次元複合正弦波パターンについて識別実験を行った .

(a)  $f(x, y, z)$

$$= \cos(2\pi x/N) + \cos(2\pi y/N) + \cos(2\pi z/N) + \cos(4\pi(x+y)/N) + \cos(4\pi(y+z)/N) + \cos(4\pi(z+x)/N)$$

(b)  $f(x, y, z)$

$$= \sin(2\pi x/N) + \sin(2\pi y/N) + \sin(2\pi z/N) + \sin(4\pi(x+y)/N) + \sin(4\pi(y+z)/N) + \sin(4\pi(z+x)/N)$$

(c)  $f(x, y, z)$

$$= \sin(2\pi x/N) + \sin(2\pi y/N) + \sin(2\pi z/N) + \cos(4\pi(x+y)/N) + \cos(4\pi(y+z)/N) + \cos(4\pi(z+x)/N)$$

ここでは,  $N = 16$  とした . 図 2 に, 不変特徴 :  $I(r, \theta; g(B) = (|\text{Re}(B)|, |\text{Im}(B)|))$  を示す . 2 次元の場合と同様に実部と虚部に異なるピークが見られる . そして, 表 2 に, 原画像と, 変換画像及び雑音付加画像との間の不変特徴の RMS 距離を示す . ここでは, 位置ずれは  $x$  軸方向に 0 から 0.5 画素まで 0.1 画素ごとに, 回転は  $x$  軸の回りに  $0^\circ$  から  $45^\circ$  まで  $0.5^\circ$  ごとに行っている . また, 雑音の分散は  $0.5^2$  としている ( $\text{SNR} \cong 10\text{dB}$ ) . この 3 次元画像においては, RMS 距離による識別において 2 倍に拡大したものにいくつかの誤識別を生じる .

そして, 原画像にランダムな相似変換 ( 拡大 :  $[1.0, 2.0]$ , 位置ずれ :  $x$  及び  $y$  軸方向に  $[-0.5, 0.5]$ , 回転 :  $[0^\circ, 360^\circ]$  の範囲で, 各一様分布 ) を行い, 更に白色雑音を加えた画像に対して識別実験を行った ( 3 次元画像に対しては, 位置ずれ及び回転は  $x, y, z$  軸について各々施した ) . 表 3 に, 各 100 枚の画像に対する不変特徴の RMS 距離による正識別率を示す . 2 次元画像では, 0 dB の雑音のもとでも 100% の正識別率が得られている . また, 3 次元画像の場合には, 5 dB 程度までの雑音のもとで 90% 以上の正識別率が得られている .

### 3.2 テクスチャ画像

自然画像として, Brodatz のアルバム [18] から 5 枚のテクスチャ画像 : D4 ( Pressed cork ), D12 ( Bark of tree ), D15 ( Straw ), D17 ( Herringbone weave ), D84 ( Raffia looped to a high pile ) を 256 階調  $260 \times 240$  画素の画像データとして取り込み, そ

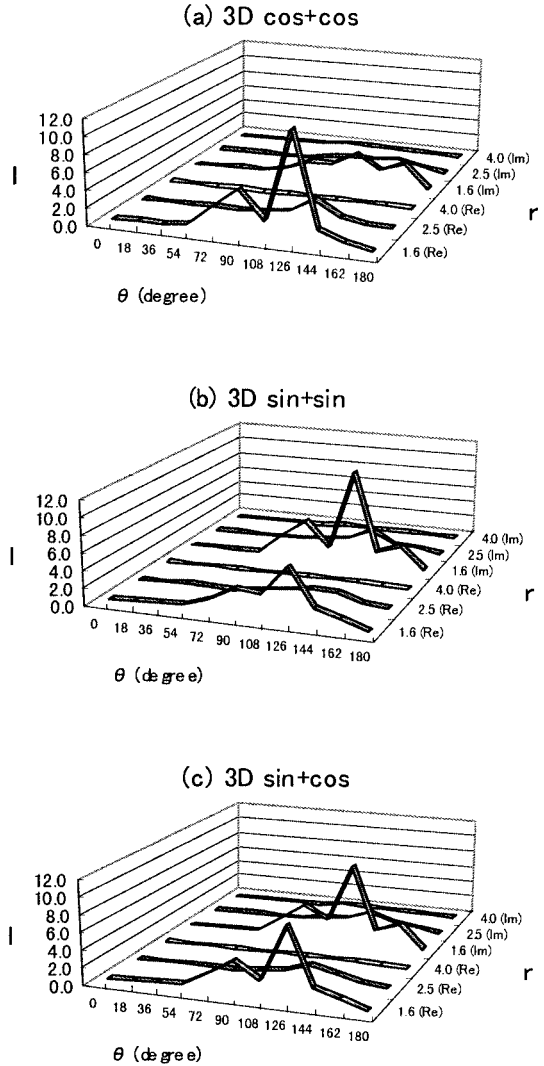


図 2 3 次元正弦波パターンの不変特徴  
Fig. 2 Invariant feature  $I(r, \theta; (|\text{Re}(B)|, |\text{Im}(B)|))$  of 3D sinusoidal patterns.

表 2 3 次元正弦波パターンの相似変換画像の不変特徴の RMS 距離

Table 2 RMS distance of the invariant feature of transformed 3D sinusoidal patterns.

	(a)			(b)			(c)		
	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.
(a) X1.0	0.17	0.13	0.0	0.74	0.71	0.66	0.53	0.51	0.46
(b) X1.0	0.88	0.83	0.74	0.20	0.16	0.0	0.38	0.33	0.23
(c) X1.0	0.57	0.55	0.51	0.26	0.24	0.23	0.15	0.11	0.0
(a) X2.0	0.52	0.48	0.40	0.87	0.69	0.57	0.75	0.56	0.44
(b) X2.0	1.00	0.82	0.71	0.58	0.45	0.40	0.65	0.48	0.41
(c) X2.0	0.75	0.68	0.63	0.49	0.46	0.39	0.46	0.42	0.36
(a)Noise	0.26	0.24	0.23	0.63	0.61	0.60	0.43	0.42	0.41
(b)Noise	0.69	0.67	0.66	0.22	0.21	0.18	0.25	0.23	0.22
(c)Noise	0.56	0.54	0.53	0.32	0.31	0.29	0.20	0.19	0.17

表 3 ランダム相似変換及び雑音付加画像の正識別率 (%)  
 Table 3 Correct classification ratio (%) for randomly transformed and noisy images.

	S.D. of noise (SNR (dB))	correct classification ratio (%)
2D sinusoidal patterns	0.0 ( $\infty$ )	100
	2.0 (0)	100
	4.0 (-6)	63
3D sinusoidal patterns	0.0 ( $\infty$ )	91
	0.5 (10)	97
	1.0 (4.6)	97
2D textures	0.0 ( $\infty$ )	100
	20.0 (6)	98
	40.0 (0)	95
3D brain images	0.0 ( $\infty$ )	87
	25.0 (0)	86

の中心部  $128 \times 128$  画素を対象として識別実験を行った。図 3 に、原画像の不変特徴  $I(r, \theta; g)$  を示す。ここでは、絶対値を用いない方が識別性能が良かったので、 $g = \text{Re}(B)$  としている。不変特徴のグラフは、テクスチャごとに異なる形状を示している。そして、表 4 に、原画像と、変換画像及び雑音付加画像 (分散:  $40^2$ ,  $\text{SNR} \cong 0 \text{ dB}$ ) との間の不変特徴の RMS 距離を示している。変換画像は、3.1 と同様に、拡大なし及び 2 倍に拡大したものについて、0 から 0.5 画素まで 0.1 画素ごとに  $x$  方向にずらしたものをそれぞれ  $0^\circ$  から  $45^\circ$  まで  $5^\circ$  ずつ回転させた 60 枚を対象としている。この場合もすべての画像が RMS 距離を用いて正しく識別される。また、3.1 と同様にして行ったランダム相似変換 + 雑音付加画像: 100 枚に対する不変特徴の RMS 距離による識別結果を、表 3 に示している。拡大率: 2 倍までの相似変換に加えて 0 dB 程度の雑音のもとでも 95% の正識別率が得られている。

次に、類似画像を含むより多数のテクスチャ画像を対象として識別実験を行った。画像データとして、MIT Media Lab の Vision Texture (VisTex) データベース [19] 内のカラーテクスチャ画像 ( $R, G, B$ : 各 256 階調,  $256 \times 256$  画素): Water: 8 枚, Terrain: 11 枚, 及び, Bark: 13 枚, 計: 32 枚を利用した。ここでは、特に、関数:  $g$  として次の 4 種類のものを用いて比較を行い、カラー値の利用の有効性について併せて調べた。

(a)  $g: Y = 0.177R + 0.813G + 0.011B$

(b)  $g: \text{a set of } (R, G, B)$

(c)  $\text{Re}(g): a^* = 500((X/255)^{1/3} - (Y/255)^{1/3}),$

$\text{Im}(g): b^* = 200((Y/255)^{1/3} - (Z/255)^{1/3})$

$(X = 0.49R + 0.31G + 0.20B,$

$Y = 0.177R + 0.813G + 0.011B,$

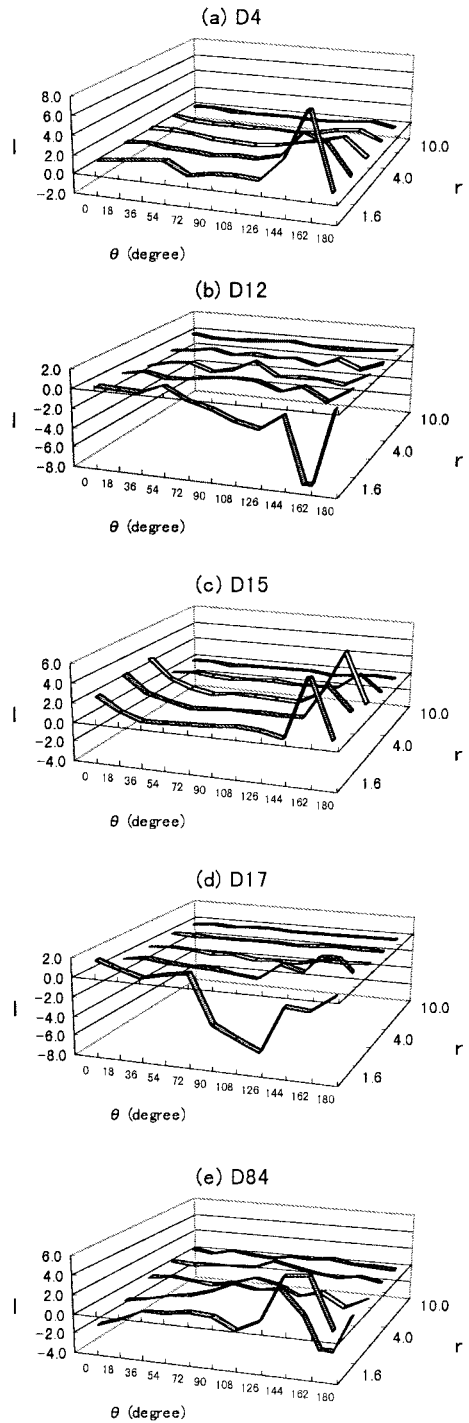


図 3 2次元テクスチャ画像の不変特徴  
 Fig. 3 Invariant feature  $I(r, \theta; \text{Re}(B))$  of 2D texture images.

表 4 2 次元テクスチャ画像 (Brodatz) の相似変換画像の不変特徴の RMS 距離  
Table 4 RMS distance of the invariant feature of transformed 2D texture images.

	(a)			(b)			(c)			(d)			(e)		
	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.
(a) X1.0	0.23	0.16	0.0	1.73	1.70	1.67	0.98	0.93	0.86	1.50	1.44	1.39	1.25	1.22	1.19
(b) X1.0	1.74	1.68	1.60	0.57	0.37	0.0	1.77	1.69	1.59	1.22	1.17	1.09	1.54	1.49	1.43
(c) X1.0	1.00	0.95	0.89	1.80	1.76	1.72	0.51	0.34	0.0	1.44	1.42	1.38	1.53	1.43	1.31
(d) X1.0	1.44	1.42	1.39	1.12	1.09	1.07	1.45	1.44	1.41	0.23	0.17	0.0	1.73	1.72	1.71
(e) X1.0	1.33	1.22	1.16	1.53	1.47	1.37	1.46	1.38	1.33	1.76	1.72	1.68	0.51	0.30	0.0
(a) X2.0	0.52	0.38	0.28	1.74	1.68	1.62	0.86	0.83	0.77	1.51	1.45	1.40	1.24	1.15	1.07
(b) X2.0	1.93	1.85	1.80	1.02	0.96	0.86	1.85	1.74	1.56	1.18	1.06	0.96	1.72	1.61	1.46
(c) X2.0	0.89	0.78	0.66	1.81	1.78	1.75	0.64	0.51	0.44	1.49	1.46	1.41	1.46	1.42	1.36
(d) X2.0	1.47	1.44	1.42	1.17	1.09	1.00	1.48	1.38	1.30	0.49	0.42	0.38	1.83	1.76	1.70
(e) X2.0	1.55	1.28	1.04	1.53	1.46	1.39	1.65	1.36	1.04	1.77	1.52	1.36	1.00	0.89	0.74
(a)Noise	0.25	0.17	0.10	1.78	1.71	1.62	0.97	0.90	0.80	1.56	1.43	1.29	1.35	1.25	1.12
(b)Noise	1.78	1.71	1.64	0.33	0.22	0.13	1.83	1.76	1.68	1.83	1.76	1.68	1.59	1.50	1.40
(c)Noise	0.98	0.89	0.83	1.81	1.76	1.73	0.24	0.15	0.09	1.49	1.44	1.38	1.48	1.40	1.31
(d)Noise	1.50	1.44	1.37	1.15	1.09	1.03	1.50	1.44	1.39	0.23	0.14	0.08	1.73	1.73	1.66
(e)Noise	1.35	1.26	1.11	1.59	1.51	1.40	1.51	1.40	1.26	1.80	1.72	1.62	0.37	0.21	0.11

表 5 2 次元テクスチャ画像 (VisTex) の相似変換画像の正識別率 (%)

Table 5 Correct classification ratio (%) for transformed 2D texture images.

scaling factor	correct classification ratio (%) image (category)			
	(a)	(b)	(c)	(d)
× 1.0	83 (92)	96 (99)	98 (98)	97 (97)
× 1.5	55 (72)	72 (91)	75 (88)	67 (76)
× 2.0	20 (56)	35 (81)	36 (60)	28 (44)
Total	53 (73)	68 (90)	70 (82)	64 (72)

$$Z = 0.01G + 0.99B)$$

(d)  $\text{Re}(g) : R - G, \text{Im}(g) : B - G$

各テクスチャ画像に対して、拡大なし、1.5 倍及び 2 倍に拡大したものについて、位置ずれ：0.0 画素～0.5 画素 (0.1 画素ごと)、回転：0°～45° (5° ごと) を施した各 60 枚の変換画像を対象とした識別実験を行った。ここでも、中心部の 128×128 画素を用いている。表 5 に、画像及びカテゴリー (Water, Terrain, Bark) の正識別率を示す。32 枚の原画像の不変特徴との RMS 距離を求め、カテゴリーの識別は同カテゴリーの原画像との距離が最小であった場合を正識別としている。いずれの場合も、カラー値を利用した方 (b)～(d) が輝度値のみを用いたもの (a) よりも 10～20% 程度高い正識別率が得られている。画像の識別においては、拡大なし (×1.0) の場合正識別率は最大で 98% になるが、2 倍に拡大した画像に対しては、30% 前後に低下する。全体では、良いもの (b)(c) で 70% 程度の正識別率が得られている。一方、カテゴリーの識別では、最も良いもの (b) で、2 倍に拡大した場合で 81%、全体では 90% の正識別が得られている。なお、ここでは示していないが、この場合雑音に対する頑健性はあまりよく

なく、SNR = 20 dB 程度の雑音によって正識別率は 30% 程度に低下してしまう。

### 3.3 人体頭部画像

3 次元のテクスチャ様画像として、人体頭部の MR 画像を対象として識別実験を行った。画像データは、McGill University の BrainWeb [20] 内の 256 階調 181×217×181 画素 (1 mm サイズ) からなる擬似 MR 画像を利用した。頭部内の (a) 右前頭部 (b) 右後頭部 (c) 左前頭部 (d) 左後頭部、及び (e) 中央部の 5 箇所 16×16×16 画素を対象とした。図 4 に、原画像の不変特徴  $I(r, \theta; g)$  ( $g = \text{Re}(B)$ ) を示す。そして、表 6 に、原画像と、相似変換画像 (拡大: 1.0, 2.0, 位置ずれ:  $x$  軸方向に 0 から 0.5 画素まで 0.1 画素ごと、回転:  $x$  軸の回りに 0° から 45° まで 0.5° ごと) 及び雑音付加画像 (分散: 40<sup>2</sup>, SNR ≅ 0 dB) との間の不変特徴の RMS 距離を示す。この場合、3.1 の正弦波パターンの場合と同様に、2 倍に拡大したものにいくつかの誤識別を生じる。

そして、原画像にランダムな相似変換 (拡大: [1.0, 2.0], 位置ずれ:  $x, y, z$  軸方向に [-0.5, 0.5], 回転:  $x, y, z$  軸の回りに [0°, 360°], 各一様分布) を行い、更に白色雑音 (分散: 40<sup>2</sup>, SNR ≅ 0 dB) を加えた画像に対して識別実験を行った。表 3 に、各 100 枚の画像に対する不変特徴の RMS 距離による正識別率を示している。0 dB 程度の雑音のもとでも 86% の正識別率が得られている。

## 4. 考 察

### 4.1 テクスチャ画像の識別

2 次元テクスチャ画像の相似変換に対する不変認識に関しては、これまでに様々な手法が試みられてい

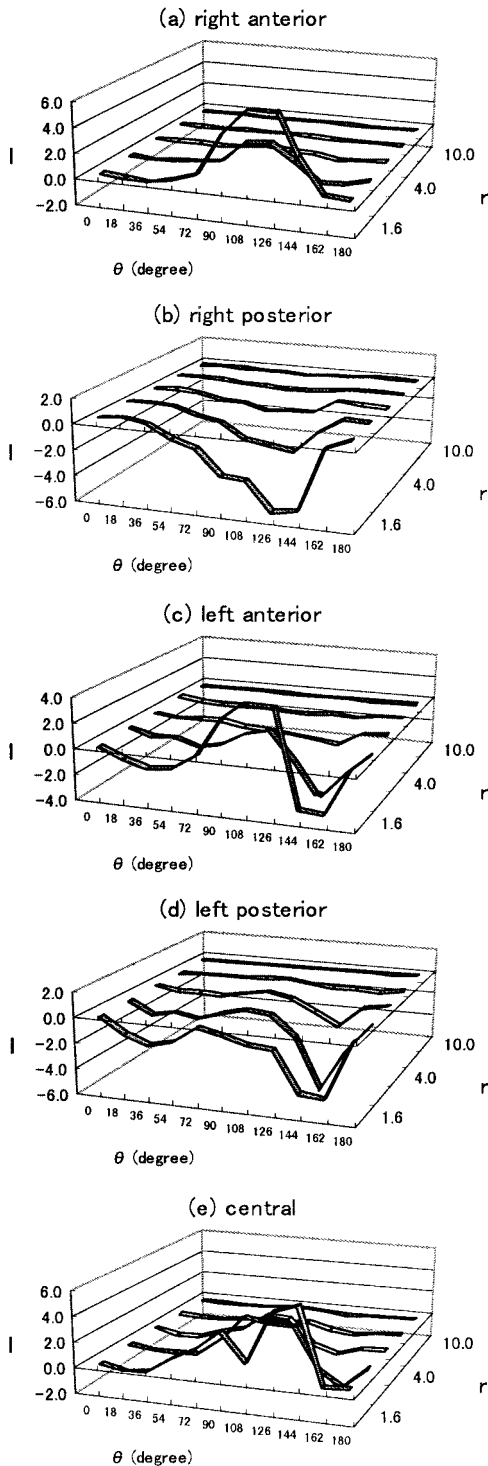


図4 3次元頭部画像の不変特徴

Fig. 4 Invariant feature  $I(r, \theta; \text{Re}(B))$  of 3D brain images.

る [6], [21] ~ [40]. そして, 例えば, 回転のみの場合には 109 種類のテクスチャ画像を対象として 80% 程度の識別率が得られること [39], また, 2 倍までの拡大がある場合でも 16 種類の画像に対して 95% 程度の識別率が得られること [33] などが報告されている. 本手法でもほぼそれらに匹敵する結果が得られているが (表 3 及び表 5), 特に優れたものであるともいえない. しかしながら, 雑音加わった場合についてはこれまであまり調べられておらず, 筆者の知る限りでは, 回転のみの場合で 0 dB 程度の雑音によって識別率が 50% 以下に低下することが示されているのみである [34] (なお, 変形のない場合には, -5 dB 程度までは 100% 近い識別率を保持できるもの [15] などがある). 本手法では, パイスベクトル (3 次相関関数) の雑音に対する頑健性によって, 少数 (5 種類) の画像を対象とした場合ではあるが, 0 dB 程度までのガウス性白色雑音のもとで 95% 程度の識別率が得られている.

一方, 3 次元テクスチャ画像の解析は医用画像の領域抽出などに関連して広く行われているが, 特徴抽出を経ず, 直接に積分型変換を用いた不変認識に関するものはあまり見当たらない. 本研究では, パイスベクトル (3 次相関関数) に基づく不変特徴を 3 次元テクスチャ画像の識別に初めて適用した. そして, 位相の異なる正弦波パターン (3 種類) 及び人体頭部 MR 画像 (5 箇所) を対象として, 任意の位置ずれ・回転及び 2 倍までの拡大のもとで 90% 程度の識別率が得られることを示した.

また, カラーテクスチャ画像に対して, カラー値を利用することによって識別率が改善されることはこれまでにいくつか報告されている [27], [41]. 本手法でも, カラー値を用いることによって濃度値のみの場合に比して 10 ~ 20% 程度識別率が增加することを示した (表 5). この場合, RGB3 値を用いたものが最も性能が良かったが,  $a^*$ ,  $b^*$  の 2 値を用いてもほぼ同じ性能が得られている.

#### 4.2 問題点及び今後の課題

画像パターン認識においては常に問題になることであるが, 本手法においても対象とする画像の性質によって識別性能は異なる. 例えば, 2 次元及び 3 次元の格子点パターンについては, 3 次相関関数 (空間領域) に基づく不変特徴によってうまく識別されたが [10], [13], [14], [16], [17], 本手法のパイスベクトル (周波数領域) に基づく不変特徴にはほとんど差異が生じなかった. また, 3 次相関関数に基づく不変特徴

表 6 3 次元頭部画像の相似変換画像の不変特徴の RMS 距離  
Table 6 RMS distance of the invariant feature of transformed 3D brain images.

	(a)			(b)			(c)			(d)			(e)		
	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.
(a) X1.0	0.24	0.13	0.00	1.97	1.93	1.89	1.07	0.90	0.76	1.66	1.55	1.47	0.82	0.71	0.62
(b) X1.0	1.99	1.97	1.95	0.47	0.20	0.00	1.77	1.64	1.55	1.29	1.11	1.00	1.96	1.94	1.89
(c) X1.0	1.20	1.02	0.86	1.60	1.49	1.36	0.36	0.16	0.00	0.99	0.83	0.67	1.45	1.34	1.24
(d) X1.0	1.73	1.63	1.57	1.02	0.94	0.76	1.10	0.96	0.89	0.31	0.10	0.00	1.74	1.65	1.60
(e) X1.0	0.75	0.63	0.47	1.99	1.97	1.96	1.44	1.29	1.20	1.74	1.64	1.60	0.34	0.17	0.00
(a) X2.0	1.67	0.92	0.58	1.98	1.84	1.02	1.76	1.53	0.89	1.95	1.77	0.47	1.64	0.78	0.57
(b) X2.0	1.96	1.51	0.48	1.83	0.98	0.32	1.57	1.16	0.51	1.40	1.11	0.64	1.95	1.73	0.95
(c) X2.0	1.90	0.85	0.19	1.97	1.52	0.43	1.45	1.15	0.74	1.82	1.36	0.64	1.87	1.07	0.59
(d) X2.0	1.83	0.76	0.33	1.95	1.64	0.53	1.44	1.16	0.50	1.87	1.42	0.17	1.85	1.05	0.73
(e) X2.0	1.80	1.27	0.15	1.95	1.31	0.69	1.85	1.31	0.60	1.85	1.21	0.12	1.80	1.24	0.57
(a)Noise	0.20	0.10	0.04	1.96	1.94	1.89	1.06	0.93	0.77	1.66	1.57	1.46	0.80	0.70	0.60
(b)Noise	1.97	1.95	1.87	0.27	0.14	0.08	1.66	1.56	1.38	1.21	1.02	0.87	1.97	1.95	1.92
(c)Noise	1.11	0.92	0.72	1.70	1.56	1.42	0.27	0.15	0.07	1.09	0.92	0.73	1.40	1.27	1.10
(d)Noise	1.61	1.59	1.55	1.05	1.00	0.96	0.94	0.91	0.86	0.07	0.04	0.02	1.64	1.61	1.58
(e)Noise	0.94	0.72	0.52	1.97	1.95	1.89	1.45	1.29	1.14	1.72	1.60	1.50	0.33	0.15	0.08

は人間によって認識される特徴とはかなり異なるものと考えられるため、人間にとって容易に識別可能な画像に適用できない場合もあるであろう。逆に、人間にとって識別困難な画像に対する実際の応用例があれば面白い。

また、高次相関特性の計算に伴う時間・空間計算量も問題となる。本手法では、バイスベクトルを求めることなくフーリエ変換画像から直接に不変特徴を計算するため、時間計算量は画素数の 2 乗のオーダーで、また、空間計算量は画素数のオーダーで押さえられる。そのため、メモリサイズに関してはほとんど問題ないが、計算時間に関しては、 $128 \times 128$  画素の 2 次元画像及び  $16 \times 16 \times 16$  画素の 3 次元画像に対して不変特徴の計算にワークステーションで数分程度要する。そのため、依然として実時間処理に利用することには困難がある。

従来の研究では、識別段階においてサンプル画像を利用した学習による特徴選択などの最適化が行われている。それに対して、今回の実験では、識別方法としては、最も単純な形の原画像との RMS 距離によった。それでも、少数の画像の識別には従来の研究に匹敵する性能が得られていることは注目に値すると考えられる。しかしながら、類似画像を含む 32 種類のテキストチャ画像に対しては 70% 程度の識別率しか得ることができず、学習による最適化によってどの程度の性能の向上があるか調べる必要がある。なお、今回の実験では濃度ヒストグラムの均一化は行わなかったが、本不変特徴は画素値の線型変換に対しても不変であるので、照明条件の変化などに伴う濃度分布の変化の影響は小さいものと考えられる。

また、不変特徴の計算における  $(r, \theta)$  平面の区画の大きさの設定も識別性能に影響する。より類似した画像を識別するためには、より細かく分割した区画におけるヒストグラムを得る必要がある。しかしながら、その場合、同一画像の相似変換による変動も大きくなる。この点は、画像データの有限性に伴うトレードオフであり検討を有する。例えば、計算に用いる各区画内の 3 重積の個数が等しくなるように非均一な分割を行うことも考えられる。

併せて、今後、3 次元画像としてテキストチャ動画を対象とした識別実験、2 次元及び 3 次元テキストチャ画像の領域分割への適用、遮へいや欠けなどの影響の評価、及び、人間の認識特性との比較などを行ってみたい。

## 5. む す び

バイスベクトルに基づく 2 次元及び 3 次元パターンの相似変換に対する不変特徴及びその計算手法を提案し、計算機実験によってその識別性能を調べた。そして、2 次元及び 3 次元の 2 次相関特性の等しい位相の異なる正弦波パターンの識別が可能であることを示した。また、2 次元のテキストチャ画像 (5 種類) 及び 3 次元の人体頭部画像 (5 種類) に対して、任意の位置ずれ・回転、2 倍程度までの拡大・縮小、及び、0dB 程度までの雑音のもとで高い識別率が得られることを示した。

## 文 献

- [1] J. Wood, "Invariant pattern recognition: A review," Pattern Recognit., vol.29, pp.1-17, 1996.
- [2] 大津展之, "不変特徴抽出の理論 [I]-[IV]," 信学誌, vol.69, no.5-8, pp.469-475, 585-590, 722-727, 831-837, 1986.



- [3] A.W. Lohmann and B. Wirtzner, "Triple correlations," *Proc. IEEE*, vol.72, pp.889-901, 1984.
- [4] C.L. Nikias and M.R. Raghuveer, "Bispectrum estimation: A digital signal processing framework," *Proc. IEEE*, vol.75, pp.869-891, 1987.
- [5] J.A. McLaughlin and J. Raviv, "Nth-order autocorrelations in pattern recognition," *Inf. Control*, vol.12, pp.121-142, 1968.
- [6] M.K. Tsatsanis and G.B. Giannakis, "Object and texture classification using higher order statistics," *IEEE Trans. Pattern Anal. & Mach. Intell.*, vol.14, no.7, pp.733-750, 1992.
- [7] L. Spirkovska, "Three-dimensional object recognition using similar triangles and decision trees," *Pattern Recognit.*, vol.26, pp.723-732, 1993.
- [8] A. Delopoulos, A. Tirakis, and S. Kollias, "Invariant image classification using triple-correlation-based neural networks," *IEEE Trans. Neural Networks*, vol.5, no.3, pp.392-408, 1994.
- [9] T.E. Hall and G.B. Giannakis, "Bispectral analysis and model validation of texture images," *IEEE Trans. Image Proc.*, vol.5, no.7, pp.996-1009, 1995.
- [10] Y. Horikawa, "Pattern recognition with invariance to similarity transformations based on the third-order correlation," *Proc. 13th International Conference on Pattern Recognition (ICPR '96)*, vol.2, B74.1, pp.200-204, 1996.
- [11] R. Marabini and J.M. Carazo, "On a new computationally fast image invariant based on bispectral projections," *Pattern Recognit. Lett.*, vol.17, pp.959-967, 1996.
- [12] V. Chandran, B. Carswell, and B. Boashash, "Pattern recognition using invariants defined from higher order spectra: 2-D image inputs," *IEEE Trans. Image Proc.*, vol.6, no.5, pp.703-712, 1997.
- [13] 堀川 洋, "3次相関に基づく相似変換に対する不変特徴を用いた3次元画像認識," *情処学コンピュータビジョンとイメージメディア研報*, CVIM109-15, 1998.
- [14] Y. Horikawa, "Invariance of three-dimensional patterns based on the third-order correlation," *Proc. 1st International Workshop on Computer Vision, Pattern Recognition and Image Processing (CVPRIP '98)*, CP-10, pp.307-310, 1998.
- [15] V. Murino, C. Ottonello, and S. Pagnan, "Noisy texture classification: A higher-order statistics approach," *Pattern Recognit.*, vol.31, pp.383-393, 1998.
- [16] 堀川 洋, "3次元パターンの位置ずれ・回転に対して不変な相似三角形ヒストグラム特徴," 1999 信学総大, D-12-69, 1999.
- [17] Y. Horikawa, "Shift and rotation invariant feature of 3D patterns based on the third-order correlation," *Proc. 5th International Symposium on Signal Processing and its Applications (ISSPA '99)*, MPP4-140, 1999.
- [18] P. Brodatz, *Textures: A Photographic Album for Artists & Designers*, Dover, New York, 1966.
- [19] The Vision Texture, <http://www-white.media.mit.edu/vismod/imager/VisionTexture/vistex.html>
- [20] BrainWeb: Simulated Brain Database, <http://www.bic.mni.mcgill.ca/brainweb/>
- [21] R.L. Kashyap and A. Khotanzaed, "A model-based method for rotation-invariant texture classification," *IEEE Trans. Pattern Anal. & Mach. Intell.*, vol.8, no.4, pp.472-481, 1986.
- [22] F.S. Cohen, Z. Fan, and M.A. Patel, "Classification of rotated and scaled texture images using Gaussian Markov random field models," *IEEE Trans. Pattern Anal. & Mach. Intell.*, vol.13, no.2, pp.192-202, 1991.
- [23] M.M. Leung and A.M. Peterson, "Scale and rotation invariant texture classification," *Proc. 26th Asilomar Conf. Signals, Systems and Computers*, vol.1, pp.461-465, 1992.
- [24] J. You and H.A. Cohen, "Classification and segmentation of rotated and scaled texture images using tuned masks," *Pattern Recognit.*, vol.26, pp.245-258, 1993.
- [25] J.-L. Chen and A. Kundu, "Rotational and gray-scale transform invariant texture identification using wavelet decomposition and hidden Markov models," *IEEE Trans. Pattern Anal. & Mach. Intell.*, vol.16, no.2, pp.208-214, 1994.
- [26] H. Greenspan, S. Belongie, R. Goodman, and P. Perona, "Rotation invariant texture recognition using a steerable pyramid," *Proc. ICPR '94*, pp.162-167, 1994.
- [27] R. Kondepudy and G. Healey, "Use of invariants for recognition of three-dimensional color textures," *J. Opt. Soc. Am. A*, vol.11, pp.3037-3049, 1994.
- [28] T.N. Tan, "Noise robust and rotation invariant texture classification," *Proc. 12th European Signal Processing Conf.*, pp.1377-1380, 1994.
- [29] G.M. Haley and B.S. Manjunath, "Rotation invariant texture classification using modified Gabor filters," *Proc. IEEE ICIP '95*, pp.262-265, 1995.
- [30] T.N. Tan, "Geometric transform invariant texture analysis," *Proc. SPIE*, vol.2488, pp.475-485, 1995.
- [31] Y. Wu and Y. Yoshida, "An efficient method for rotation and scaling invariant texture classification," *Proc. ICASSP '95*, vol.4, pp.2519-2522, 1995.
- [32] W.-R. Wu and S.-C. Wei, "Rotation and gray-scale transform-invariant texture classification using spiral resampling, subband decomposition, and hidden Markov model," *IEEE Trans. Image Proc.*, vol.5, no.10, pp.1423-1434, 1996.
- [33] 吉田靖夫, 呉 越, "高次スペクトルモーメントを用いた回転と拡大のあるテクスチャ画像の分類," *画像の認識理解シンポジウム (MIRU '96)*, 1996.
- [34] S.R. Fountain and T.N. Tan, "Extraction of noise robust and rotation invariant texture features via multichannel filtering," *Proc. IEEE ICIP '97*, vol.3,

pp.197–200, 1997.

- [35] S.R. Fountain and T.N. Tan, “Efficient rotation invariant texture features for content-based image retrieval,” *Pattern Recognit.*, vol.31, pp.1725–1732, 1998.
- [36] V. Manian and R. Vasquez, “Scaled and rotated texture classification using a class of basis functions,” *Pattern Recognit.*, vol.31, pp.1937–1948, 1998.
- [37] T.N. Tan, “Rotation invariant texture features and their use in automatic script identification,” *IEEE Trans. Pattern Anal. & Mach. Intell.*, vol.20, no.7, pp.751–756, 1998.
- [38] L. Wang and G. Healey, “Using Zernike moments for the illumination and invariant classification of multispectral texture,” *IEEE Trans. Image Proc.*, vol.7, no.2, pp.196–203, 1998.
- [39] G.M. Haley and B.S. Manjunath, “Rotation-invariant texture classification using a complete space-frequency model,” *IEEE Trans. Image Proc.*, vol.8, no.2, pp.255–269, 1999.
- [40] M. Pietikäinen, T. Ojala, and Z. Xu, “Rotation-invariant texture classification using feature distribution,” *Pattern Recognit.*, vol.33, pp.43–52, 2000.
- [41] G. Van de Wouwer, P. Sheunders, S. Livens, and D. Van Dyck, “Wavelet correlation signatures for color texture characterization,” *Pattern Recognit.*, vol.32, pp.443–451, 1999.

(平成 12 年 7 月 13 日受付, 11 月 20 日再受付)



堀川 洋 (正員)

昭 58 東大・工・計数卒・昭 60 同大学院修士課程了。同年長崎総合科学大・情報制御助手, 平 3 香川大・教育助手, 平 9 同・工助教授。平 7~8 オーストラリア国立大客員研究員。生体情報工学に関する研究に従事。博士(工学)。