

## 手書き漢字認識のための Gabor 特徴

政水克典† 内村俊二‡ 浜本義彦† 富田眞吾†

山口大学工学部† 大島商船高等専門学校‡

あらまし 文字認識において、人間の視覚系を模倣したモデルに基づく特徴を用いることで、人間の認識性能に近い結果が期待できる。本研究ではこのような観点から、人間の視覚系をモデル化した Gabor フィルタに着目し、それに基づく特徴抽出法を検討する。また、本手法と、他の特徴抽出法とを誤識別率の観点から比較し、本手法の有効性を明らかにする。

和文キーワード 手書き漢字認識、Gabor フィルタ、人間の視覚系

## Gabor Features for Handprinted Chinese Character Recognition

Katsunori Masamizu†, Shunji Uchimura‡  
Yoshihiko Hamamoto†, Shingo Tomita‡

Faculty of Engineering, Yamaguchi University†  
Oshima National College of Maritime Technology‡

Abstract A method for handprinted Chinese character recognition based on Gabor filters is proposed. The Gabor approach to character recognition is intuitively appealing because it is inspired by a multi-channel filtering theory for processing visual information in the early stages of the human visual system. The performance of a recognition system using Gabor features is demonstrated. Experimental results show that Gabor features yielded an error rate of 2.4 % versus the error rate of 4.4 % obtained from the other feature extraction method.

英文 key words Chinese character recognition, Gabor filter, human visual model

# 1 まえがき

手書き漢字認識について、これまでに数多くの研究がなされている。しかし、人間の認識性能との間には精度においてかなりのギャップがあるのが現状である。このギャップは、特徴が認識結果に強い影響を与えることから、主に特徴に起因するものと考えられる。従来、手書き漢字認識に用いられてきた特徴は、人間が文字パターンから作為的に生成したものである。この作為的な特徴と人間の視覚系が抽出している特徴との差異により上述のギャップが生じている、と考えられる。そこで、人間の視覚系が抽出している特徴と類似の特徴を手書き漢字認識に用いれば、人間の認識性能に近い結果が期待できる。

人間の視覚系をモデル化したものの一つに Gabor フィルタ [1] がある。Marcelja[2] は、Gabor フィルタは皮質ニューロンの受容野特性を近似することができると指摘した。これまでに Gabor フィルタは、エッジ検出 [3] やテキスチャ解析 [4] などに用いられてきた。また、山田と津雲 [5] は印刷漢字認識に Gabor フィルタを適用し、その安定性を検討した。本研究では、人間の視覚系をモデル化した Gabor フィルタに基づく特徴抽出法を提案する。計算機シミュレーションにより、Gabor 特徴を分析すると共に、Gabor 特徴と他の特徴とを誤識別率の観点から比較し、手書き漢字認識への有効性を検討する。

## 2 Gabor 特徴

### 2.1 Gabor フィルタ

Gabor フィルタは二次元ガウス曲面と二次元平面上を一方方向に伝わる余弦波とを掛け合わせたフィルタである。

Gabor フィルタは、ガウス曲面における標準偏差、余弦波曲面における波の進行方向、波長、位相により決定される。標準偏差を  $\sigma_x, \sigma_y$ 、波長を  $\lambda_m$ 、位相を  $\phi_k$ 、波の進行方向と  $x$  軸とがなす角度を  $\theta_l$  とすると、Jain ら [6] は Gabor フィルタを式 (1) で与えた。

$$f(x, y, \theta_l, \sigma_x, \sigma_y, \lambda_m, \phi_k) = \exp\left(-\frac{1}{2}\left(\frac{(x \cos \theta_l + y \sin \theta_l)^2}{\sigma_x^2} + \frac{(-x \sin \theta_l + y \cos \theta_l)^2}{\sigma_y^2}\right)\right) \times \cos\left(\frac{360(x \cos \theta_l + y \sin \theta_l)}{\lambda_m} + \phi_k\right) \quad (1)$$

図 1, 2 に Gabor フィルタを示す。図 1 では  $\theta_l = 0^\circ$ ,  $\sigma_x = 2\sqrt{2}$ ,  $\sigma_y = 2\sqrt{2}$ ,  $\lambda_m = 2\sqrt{2}$ ,  $\phi_k = 0^\circ$ , 図 2 では  $\theta_l = 45^\circ$ ,  $\sigma_x = 2\sqrt{2}$ ,  $\sigma_y = 2\sqrt{2}$ ,  $\lambda_m = 2\sqrt{2}$ ,  $\phi_k = 0^\circ$  である。

上述の標準偏差  $\sigma_x, \sigma_y$  はガウス曲面の広がり定めるもので、これらパラメータは  $\lambda_m$  と密接な関係がある。余弦波曲面の波長にあわせて、ガウス曲面の標準偏差を変化

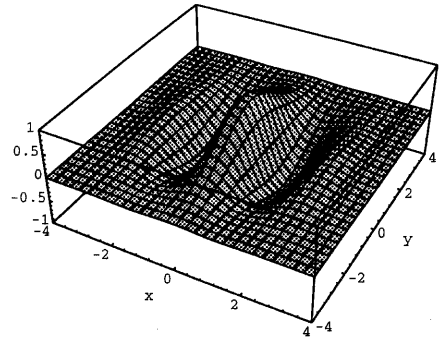


図 1: Gabor フィルタ

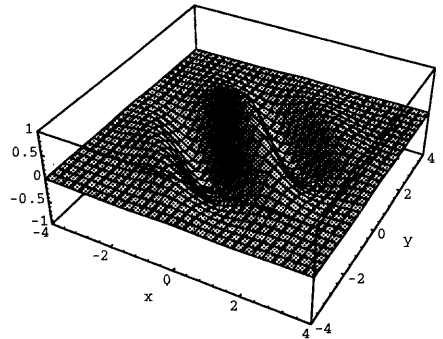


図 2: Gabor フィルタ

させるため、次のように標準偏差  $\sigma_x, \sigma_y$  は  $\lambda_m$  の関数となる [5][6]。

$$\begin{aligned} \sigma_x &\rightarrow \sigma_x(\lambda_m) \\ \sigma_y &\rightarrow \sigma_y(\lambda_m) \end{aligned}$$

すなわち、 $\lambda_m$  の値が与えられると、 $\sigma_x, \sigma_y$  の値が定まるのである。よって、式 (1) は式 (2) のように書き換えることができる。

$$f(x, y, \theta_l, \lambda_m, \phi_k) = \exp\left(-\frac{1}{2}\left(\frac{(x \cos \theta_l + y \sin \theta_l)^2}{\sigma_x^2(\lambda_m)} + \frac{(-x \sin \theta_l + y \cos \theta_l)^2}{\sigma_y^2(\lambda_m)}\right)\right) \times \cos\left(\frac{360(x \cos \theta_l + y \sin \theta_l)}{\lambda_m} + \phi_k\right) \quad (2)$$

## 2.2 Gabor 特徴

Gabor 特徴は Gabor フィルタと文字パターンとの畳み込みから得られる。

文字パターンを大きさ  $N \times N$  の二値画像  $I(i, j)$  ( $0 \leq i \leq N-1, 0 \leq j \leq N-1$ ) とし、位相  $\phi_k$  を要素とする集合  $S_n$  ( $n = 1, 2, 3, 4$ ) を式 (3) のように定義する。

$$\begin{aligned} S_1 &= \{0^\circ\} \\ S_2 &= \{0^\circ, 90^\circ\} \\ S_3 &= \{0^\circ, 60^\circ, 120^\circ\} \\ S_4 &= \{0^\circ, 45^\circ, 90^\circ, 135^\circ\} \end{aligned} \quad (3)$$

文字パターン上のサンプリング点を  $(X, Y)$  とし、特定の  $\theta_l, \lambda_m, S_n$  に対する Gabor 特徴は式 (4) で与えられる。

$$\begin{aligned} g(X, Y, \theta_l, \lambda_m, S_n) \\ = \sum_{\phi_k \in S_n} \left| \sum_{x=-X}^{N-X-1} \sum_{y=-Y}^{N-Y-1} I(X+x, Y+y) \right. \\ \left. \times \hat{f}(x, y, \theta_l, \lambda_m, \phi_k) \right| \end{aligned} \quad (4)$$

上式の  $g(\cdot)$  関数は、Gabor フィルタの中心を文字パターン上の  $(X, Y)$  の位置に合わせて畳み込んだものである。Gabor フィルタの中心を文字パターンの  $(X, Y)$  に重ね合わせたとき、Gabor フィルタの正の領域のみに黒画素があるか、あるいは Gabor フィルタの負の領域のみに黒画素がある場合に、 $(X, Y)$  での特徴値  $g(\cdot)$  は大きくなる。Gabor フィルタと文字イメージの畳み込みを図 3 に示す。

位相集合を  $S_n$  に固定し、サンプリング点毎に方向  $\theta_l$ 、波長  $\lambda_m$  の全ての組合せについて式 (4) の  $g(\cdot)$  を計算する。そのため、サンプリング点の数を  $M$  とし、方向  $\theta_l$  と波長  $\lambda_m$  の組合せの数を  $L$  とすると、得られる特徴数は  $M \times L$  で与えられる。

## 3 計算機シミュレーション

Gabor 特徴を定めるパラメータが誤識別率に与える影響、およびノイズやかすれ文字に対する Gabor 特徴の特性を調べるために、計算機シミュレーションを行なった。

認識対象である文字パターンには電総研の手書き教育漢字データベース ETL-8B を用いた。ETL-8B は全 881 字種で一字種当たり 160 文字で構成されており、1 文字は  $64 \times 63$  ピクセルの二値文字である。本研究では ETL-8B の全 881 字種の中から 55 字種 (各字種 160 字) 選び出し、文字の縦横比を同じにするために白画素を下部に付け加え、 $64 \times 64$  ピクセルの二値文字として用いた。

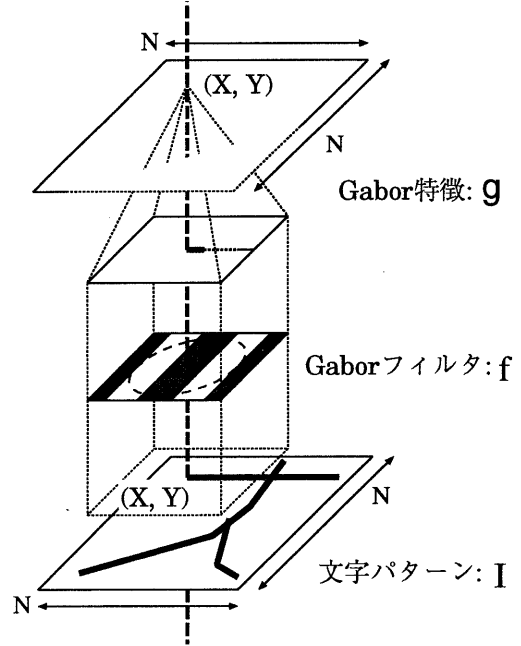


図 3: Gabor フィルタと文字パターンの畳み込み

### 3.1 パラメータの影響

Gabor 特徴のパラメータが誤識別率にどのような影響を与えるかを調べた。さらに、Gabor 特徴を、代表的な特徴抽出法である外郭方向寄与度法 [7] と誤識別率の観点から比較した。

本研究では方向  $\theta_l$  を 4 種類 ( $l = 1, 2, 3, 4$ )、各方向について波長  $\lambda_m$  を 2 種類 ( $\lambda_1, \lambda_2$ ) 用いたので、方向  $\theta_l$  と波長  $\lambda_m$  の組合せの数  $L$  は 8 である。また、サンプリング点は縦方向、横方向とも等間隔に 8 点とり、サンプリング点の総数  $M$  を 64 とした (図 4 参照)。よって、得られる特徴の次元数は 512 となる。

方向  $\theta_l$  は  $\theta_1 = 0^\circ, \theta_2 = 45^\circ, \theta_3 = 90^\circ, \theta_4 = 135^\circ$  に固定し、波長  $\lambda_m$  の値と、波長に対する標準偏差  $\sigma_x, \sigma_y$  の倍率、および位相の集合  $S_n$  を変えてシミュレーションを行なった。波長  $\lambda_1, \lambda_2$  は

$$\sqrt{2}, 2\sqrt{2}, 4\sqrt{2}, 8\sqrt{2}, 16\sqrt{2}$$

の中から選び、標準偏差  $\sigma_x, \sigma_y$  との関係は次の 4 通りそれぞれについて行なった。

- (a)  $\sigma_x(\lambda_m) = 0.5\lambda_m, \sigma_y(\lambda_m) = 0.5\lambda_m$
- (b)  $\sigma_x(\lambda_m) = 0.7\lambda_m, \sigma_y(\lambda_m) = 0.5\lambda_m$
- (c)  $\sigma_x(\lambda_m) = 0.5\lambda_m, \sigma_y(\lambda_m) = 0.7\lambda_m$
- (d)  $\sigma_x(\lambda_m) = 0.7\lambda_m, \sigma_y(\lambda_m) = 0.7\lambda_m$

また、位相については  $S_1, S_2, S_3, S_4$  それぞれについて Gabor 特徴を求めた。

表 1: パラメータの誤識別率に与える影響 (%)

$\lambda_1, \lambda_2$	$S_n$	$\sigma_x = 0.5\lambda_m$ $\sigma_y = 0.5\lambda_m$	$\sigma_x = 0.7\lambda_m$ $\sigma_y = 0.5\lambda_m$	$\sigma_x = 0.5\lambda_m$ $\sigma_y = 0.7\lambda_m$	$\sigma_x = 0.7\lambda_m$ $\sigma_y = 0.7\lambda_m$
$\sqrt{2}, 2\sqrt{2}$	0°	22.3	14.0	21.6	14.2
$\sqrt{2}, 2\sqrt{2}$	0°, 90°	14.8	10.6	14.4	10.7
$\sqrt{2}, 2\sqrt{2}$	0°, 60°, 120°	15.5	10.5	15.1	10.5
$\sqrt{2}, 2\sqrt{2}$	0°, 45°, 90°, 135°	15.0	10.6	14.6	10.5
$2\sqrt{2}, 4\sqrt{2}$	0°	6.4	4.1	7.6	5.3
$2\sqrt{2}, 4\sqrt{2}$	0°, 90°	4.1	2.6	4.8	3.1
$2\sqrt{2}, 4\sqrt{2}$	0°, 60°, 120°	4.2	2.4	4.8	3.1
$2\sqrt{2}, 4\sqrt{2}$	0°, 45°, 90°, 135°	4.1	2.4	4.7	3.0
$4\sqrt{2}, 8\sqrt{2}$	0°	5.9	6.5	9.1	10.2
$4\sqrt{2}, 8\sqrt{2}$	0°, 90°	3.7	4.4	5.1	6.5
$4\sqrt{2}, 8\sqrt{2}$	0°, 60°, 120°	3.6	4.4	5.0	6.4
$4\sqrt{2}, 8\sqrt{2}$	0°, 45°, 90°, 135°	3.6	4.5	5.0	6.4
$8\sqrt{2}, 16\sqrt{2}$	0°	12.8	18.3	17.0	23.7
$8\sqrt{2}, 16\sqrt{2}$	0°, 90°	8.3	15.4	11.3	20.0
$8\sqrt{2}, 16\sqrt{2}$	0°, 60°, 120°	8.6	15.4	11.3	19.9
$8\sqrt{2}, 16\sqrt{2}$	0°, 45°, 90°, 135°	8.4	15.3	11.2	19.8
$\sqrt{2}, 4\sqrt{2}$	0°	6.6	4.2	7.7	5.5
$\sqrt{2}, 4\sqrt{2}$	0°, 90°	4.1	2.6	4.8	3.1
$\sqrt{2}, 4\sqrt{2}$	0°, 60°, 120°	4.2	2.4	4.8	3.1
$\sqrt{2}, 4\sqrt{2}$	0°, 45°, 90°, 135°	4.1	2.4	4.7	3.0
$2\sqrt{2}, 8\sqrt{2}$	0°	7.0	7.9	10.9	12.6
$2\sqrt{2}, 8\sqrt{2}$	0°, 90°	4.0	5.2	5.7	7.7
$2\sqrt{2}, 8\sqrt{2}$	0°, 60°, 120°	3.9	5.2	5.6	7.5
$2\sqrt{2}, 8\sqrt{2}$	0°, 45°, 90°, 135°	4.0	5.2	5.6	7.3
$4\sqrt{2}, 16\sqrt{2}$	0°	18.8	28.2	25.0	35.2
$4\sqrt{2}, 16\sqrt{2}$	0°, 90°	14.1	26.5	19.5	34.3
$4\sqrt{2}, 16\sqrt{2}$	0°, 60°, 120°	14.2	26.4	19.1	33.6
$4\sqrt{2}, 16\sqrt{2}$	0°, 45°, 90°, 135°	14.0	26.4	19.1	33.7

識別器は、ユークリッド距離識別器を用いた。ユークリッド距離識別器では、クラス  $r$  の標本平均ベクトルを  $\mu_r$ 、テストパターンを  $\mathbf{x}$  とすると、

$$(\mathbf{x} - \mu_q)^t(\mathbf{x} - \mu_q) = \min_r (\mathbf{x} - \mu_r)^t(\mathbf{x} - \mu_r) \quad (5)$$

のときテストパターン  $\mathbf{x}$  をクラス  $q$  へと識別する。

誤識別率の推定法には、Leave-one-out 法を用いた。以下、Leave-one-out 法を説明する。いま、利用できる  $s$  個のパターンからなるパターン集合  $W_s$  が与えられているとする。パターン集合  $W_s$  の中からパターンを一つ抜き取り、残りの  $s-1$  個のパターンからなるパターン集合  $W_{s-1}$  で認識系を設計し、抜き取っておいたパターンをテストパターンとして用いる。すべてのパターンが一度限りテストパターンとして用いられるまで、この処理を繰り返し、誤

識別率を推定する。Leave-one-out 法は他の誤識別率推定法に比べて、パターンの利用効率が高いという長所と、 $s$  回認識系を設計しなければならないという短所がある。

実験結果を表 1 に示す。結果より、パラメータの組合せが  $\lambda_1 = 2\sqrt{2}, \lambda_2 = 4\sqrt{2}, \sigma_x = 0.7\lambda_m, \sigma_y = 0.5\lambda_m, S_3$  の組合せと、 $\lambda_1 = \sqrt{2}, \lambda_2 = 4\sqrt{2}, \sigma_x = 0.7\lambda_m, \sigma_y = 0.5\lambda_m, S_3$  の組合せが最小の誤識別率 2.4% を与えた。一方、比較する外郭方向寄与度特徴は、文字のどの部分までを特徴として抽出するかを表す外郭深度により、得られる特徴数が異なる。Gabor 特徴と同じ次元数 512 になる外郭深度は 1 であることから、外郭深度 1 の場合について比較した。Gabor 特徴が 2.4% の誤識別率であるのに対し、外郭方向寄与度特徴のそれは 4.4% であった。一方、実験結果から、位相の数、標準偏差の値に比べて、波長の値は誤識別率に

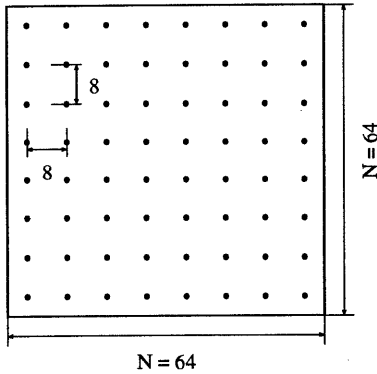


図 4: サンプルング点の位置

大きな影響を与えている。また、標準偏差の値 $\sigma_x, \sigma_y$ が異なっている方が低い誤識別率を与える場合がある。このことから、フィルタ設計の際に $\sigma_x, \sigma_y$ を別々に考慮する必要がある。次に、位相については、位相以外のパラメータが同一の状況下では、位相の数が1つの場合に比べ2つ、すなわち $S_2$ の方が誤識別率が低く、さらに位相の数を増加させても誤識別率に差異はなかった。計算コストの面からは、位相の数は少ないほど良い。従って、位相については $S_2$ が適当と思われる。

### 3.2 耐ノイズ特性

文字パターンにノイズが混入する場合がある。そこで、ノイズ混じりの文字パターンから抽出された Gabor 特徴が誤識別率へどのような影響を及ぼすかを、外郭方向寄与度法と比較した。

原文字パターンからノイズにより劣化した文字パターンを次の手順で作成した。原文字パターンの画素数を P とする。原文字パターンから Q 個の画素をランダムに選び、その画素を反転させたものをノイズにより劣化した文字パターンとした。ノイズに関する SN 比は式 (6) により定めた。

$$SN_1 = 20 \log_{10} \frac{P-Q}{Q} \text{ (dB)} \quad (6)$$

ノイズにより劣化した文字パターンの例を図 5 に示す。

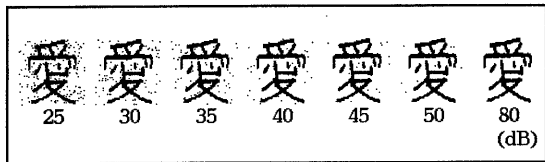


図 5: ノイズにより劣化した文字パターンの例

用いた Gabor 特徴は、3.1 節のシミュレーションにおいて

最小の誤識別率を与えたパラメータ、すなわち $\lambda_1 = 2\sqrt{2}$ ,  $\lambda_2 = 4\sqrt{2}$ ,  $\sigma_x = 0.7\lambda_m$ ,  $\sigma_y = 0.5\lambda_m$ ,  $S_3$ の組合せから得られる特徴である。

認識系を設計する訓練パターンには ETL-8B の原文字パターンを用い、テストパターンにのみノイズを付加した文字パターンを用いた。シミュレーションでは SN 比を変化させ、SN 比毎に誤識別率を求めた。

3.1 節と同様に識別器にユークリッド距離識別器、誤識別率の推定法に Leave-one-out 法を用いた。

テストパターンにはランダムな要素が含まれているので、SN 比、特徴抽出法、次元数が同じであっても試行毎に誤識別率は変化する。試行回数を 5 回として独立にシミュレーションを行なった。それぞれの誤識別率の平均値、標準偏差を表 2 に示す。また、SN 比が 80 (dB) の結果は、ノイズの付加されていない原文字パターンをテストパターンとして用いた結果となる。なお、外郭方向寄与度特徴については、1536 次元の場合も比較した。実験結果より、ノイズが大きくなると、Gabor 特徴に比べ、外郭方向寄与度特徴の方の誤識別率が著しく増加している。

### 3.3 耐かすれ特性

文字パターンがかすれている場合がある。そこで、かすれ文字パターンから抽出された Gabor 特徴が誤識別率へどのような影響を及ぼすかを、外郭方向寄与度法と比較した。

原文字パターンから、かすれ文字パターンを次の手順で作成した。原文字パターンの黒画素の数を R とする。R 個の黒画素の中からランダムに T 個、画素を選びその画素を反転させたものをかすれ文字パターンとした。かすれに関する SN 比は式 (7) により定めた。

$$SN_2 = 20 \log_{10} \frac{R-T}{T} \text{ (dB)} \quad (7)$$

かすれ文字パターンの例を図 6 に示す。

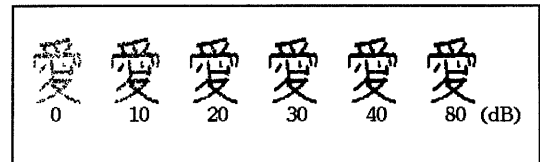


図 6: かすれ文字パターンの例

用いた Gabor 特徴は、3.2 節と同様に 3.1 節において最小の誤識別率を与えたパラメータの組合せで得られた特徴を用い、外郭方向寄与度特徴は、512 次元と 1536 次元の 2 種類の特徴を用いた。また、訓練パターンには ETL-8B の原文字パターンを用い、テストパターンにのみ、かすれ文字パターンを用いた。シミュレーションでは SN 比を変化させ、SN 比毎に誤識別率を求めた。3.1 節と同様に識別

表 2: ノイズの誤識別率に与える影響 (%)

$SN_1$ (dB)	25	30	35	40	45	50	80
Gabor 特徴 (512 次元)	3.2	2.9	2.7	2.6	2.6	2.5	2.4
標準偏差	0.1	0.1	0.1	0.1	0.0	0.0	0.0
外郭方向寄与度特徴 (512 次元)	64.4	30.3	10.2	5.7	4.8	4.5	4.4
標準偏差	0.3	0.2	0.1	0.1	0.1	0.1	0.0
外郭方向寄与度特徴 (1536 次元)	81.0	52.5	14.2	2.7	1.5	1.2	1.2
標準偏差	0.2	0.2	0.2	0.1	0.0	0.1	0.0

表 3: かすれの誤識別率に与える影響 (%)

$SN_2$ (dB)	0	10	20	30	40	80
Gabor 特徴 (512 次元)	24.8	4.8	2.9	2.6	2.5	2.4
標準偏差	0.2	0.1	0.1	0.1	0.0	0.0
外郭方向寄与度特徴 (512 次元)	78.3	34.2	9.1	5.2	4.7	4.4
標準偏差	0.3	0.3	0.2	0.1	0.1	0.0
外郭方向寄与度特徴 (1536 次元)	72.0	33.8	5.2	1.6	1.3	1.2
標準偏差	0.3	0.2	0.2	0.1	0.0	0.0

器にはユークリッド距離識別器を、誤識別率の推定法には Leave-one-out 法を用いた。3.2節と同様にテストパターンにはランダムな要素が含まれているので、試行毎に誤識別率は変化する。試行回数を 5 回とし、誤識別率の平均値、標準偏差を表 3 に示す。実験結果より、誤識別率の増加の度合は Gabor 特徴の方が小さいことがわかった。

## 4 むすび

Gabor 特徴の手書き漢字認識への有効性を検討し、ノイズ、かすれのある文字に対して効果的であることを示した。今後の課題として、サンプリング点の取り方、変形文字に対する Gabor 特徴の特性について検討する必要がある。

## 謝辞

本研究に際し、実験に用いた手書き教育漢字データベースを作成、提供して頂いた電子技術総合研究所の関係諸氏に深く感謝致します。

## 参考文献

- [1] D. Gabor: "Theory of communication", J. Inst. Elect. Engr., **93**, pp.429-457(1946).
- [2] S. Marcelja: "Mathematical description of the responses of simple cortical cells", Journal of the Optical Society of America, **70**, pp.1297-1300(1980).

- [3] R. Mehrotra, K. R. Namuduri and N. Ranganathan: "Gabor filter-based edge detection", Pattern Recognition, **25**, 12, pp.1479-1494(1992).
- [4] M. R. Turner: "Texture discrimination by Gabor functions", Biol. Cybern., **55**, pp.71-82(1986).
- [5] 山田敬嗣, 津雲淳: "Gabor 特徴の安定性考察と文字認識", 信学技報, **PRU92-112**, pp.75-82(1993-01).
- [6] A. K. Jain and S. K. Bhattacharjee: "Address block location on envelopes using Gabor filters", Pattern Recognition, **25**, 12, pp.1459-1477(1992).
- [7] 萩田紀博, 内藤誠一郎, 増田功: "外郭方向寄与度特徴による手書き漢字の識別", 信学論 (D), **J66-D**, 10, pp.1185-1192 (1983-10).
- [8] A. Khotanzad and J. -H. Lu: "Shape and texture recognition by a neural network", Artificial Neural Networks and Statistical Pattern Recognition Old and New Connections, I. K. Sethi and A. K. Jain eds., Elsevier Science, pp.109-131(1991).
- [9] 乾敏郎: "視覚情報処理の基礎", サイエンス社 (1990).