

## Band 相関法による印刷文字 OCR の設計理論\*

小沢一雅\*\* 田中幸吉\*\*\*

### Abstract

Pattern matching method has usually been used for implementing the printed character oriented OCR. This method, however, may seem to be not so suitable for realizing an OCR with high performance because this method involves such a defect as misreading due to noise and position shift of input characters.

In this paper, a new type of pattern matching method named "Band Correlation Method" is proposed. Especially the technical procedures for implementing an OCR by the method are described in detail. By the proposed method, it is expected that the defect mentioned above is entirely eliminated.

### 1. ま え が き

パターン認識、特に文字図形認識の研究は基礎研究の段階を終え、高性能の実用機を如何に簡潔に実現するかという問題に研究努力が集中されるに至った。筆者等はすでに高性能 OCR 方式としての Band 相関法の原理的な提案を行っているが<sup>1),2),3)</sup>、本稿では、Band 相関法による OCR を具体的に実現するための設計理論<sup>4),5),6)</sup>を中心に述べる。

Band 相関法は、従来のパターンマッチング法を母体として空間フィルタリングによる雑音除去処理等を加味した OCR 方式である。Band 相関法の特長は、スペクトル情報を利用して OCR に不可欠な種々の前処理操作をきわめて簡単に実現するところであり、これによって、原始的パターンマッチング法の欠点をほぼ克服した実用的にして高性能な印刷文字用 OCR 開発への道を開いたものといえる。Band 相関法は、筆者等によって継続的に進められているレーザ OCR<sup>7),11)</sup>の研究過程で着想されたものであるが、その適用範囲は必ずしもレーザ OCR だけにとどまらず、電子回路的手段による OCR をも完全に包含するものである。

とはすでに指摘している<sup>1),4)</sup>。

スペクトルの利用法に関して Band 相関法は大きく3つのタイプに分類することができる。すなわち、(1)複素振幅型、(2)パワー型および(3)振幅型の3つであり、それぞれ複素振幅スペクトル、パワースペクトルおよびその平方根を利用するものである。パワー型と振幅型はある程度の情報欠損はあるが、入力パターンの位置ずれに関するスペクトルの不変性を有効に利用でき、認識対象を選べば実用的 OCR への近道となるものといえる。

一方、複素振幅型はスペクトルが位置ずれによって不変に保たれないなどの不便はあるが、本質的な情報欠損がないため、英数字だけでなくカタカナあるいは漢字などのカテゴリ数が多い対象に対しても高性能な認識能力を期待できるものと考えられる。本稿では、フーリエスペクトルを用いる複素振幅型 Band 相関法を中心に述べることにする。

### 2. 文字認識における雑音

パターンは本来その存在理由すなわち役割をもってしているものと考えられる。文字というパターンの役割は文字概念の伝達にあると考えることができよう。特に文字は他の一般画像と異なり直線や曲線の配置関係によって情報を伝達するという特殊な画像であると考えられる。ここで重要な意味をもつ『線』という文字の構成要素はわれわれが観念的にとらえる場合、『幅が

\* Printed Character Oriented OCR By Band Correlation Method by Kazumasa OZAWA (Osaka Electro-Communication University) and Kokichi TANAKA (Osaka University)

\*\* 大阪電気通信大学工学部

\*\*\* 大阪大学基礎工学部

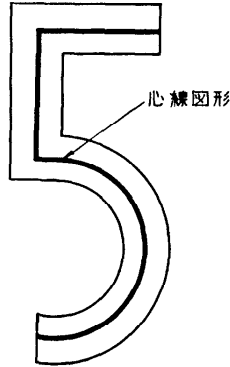


Fig. 1 A character and its skelton

ない』といういわゆる幾何学的線分でありそれによって形づくられる理想的文字はパターン認識の用語で心線図形と呼ぶことができるであろう (Fig. 1 参照)。

心線図形は幅がなく理想的であるが故にそれ自身で物理的に存在し得ず、したがって表象する文字概念を伝達するエネルギーをもたない。これに対して実際に与えられる文字の具体例はペンで書かれたり、あるいは印刷されたものであり、これらは有限な『幅をもった』線分から構成されている。線分が幅をもった瞬間から文字は伝達のためのエネルギーを保有することになるが、同時にたとえば『しみ』や『かすれ』などの雑音の介入が避けられないものとなってくる。

われわれが文字を正しく判読できるのは、実は与えられた文字からこの心線情報を巧みに抽出する能力を有しているからに他ならない。

このように、実際に与えられる文字はその役割をはたすのに必要な情報成分以外にさまざまな理由にもとづいて介入してくる情報成分を保有している。そこで文字の認識に必要な最小限の情報——本質の情報——を除いた他の情報成分を雑音とよぶことにする。

文字認識がうまく行くかどうかはこの雑音を如何に効果的に除去できるか否かにかかっているといえよう。Band 相関法で空間周波数スペクトルを利用する理由は雑音成分が特定の周波数スペクトル領域に集中するという現象が認められるからであり、適当な空間フィルタを用いて容易に雑音成分を除去することが可能である。たとえば、『しみ』や『かすれ』などのストローク周辺の雑音は高周波領域に顕著に集積し、文

\* ここでは、内積  $(f, \phi)$  をつぎのように定める。

$$(f, \phi) = \int_{R^2} f(x)\phi(x)dx. \quad (\text{—は複素共役})$$

またノルム  $\|f\|$  は  $\|f\|^2 = (f, f)$  で与える。

字の物理的な位置に関する情報など文字概念の差異に無関係な情報は零周波数近傍に集まっている。

### 3. Band 相関法

#### 3.1 類似度 (相関係数)

2次元平面  $R^2$  上で定義される文字パターンを非負値関数  $f(x)$  ( $x \in R^2$ ) で表わす。  $f(x)$  を空間  $L_2$  ( $R^2$ ) の一点とみなし、  $f$  と略記することとする。  $f$  のフーリエ変換  $f^*$  を作用素  $F$  を用いてつぎのように表わすことにする。

$$f^* = F \cdot f \quad (1)$$

ここで、未知パターン  $f$  と照合用パターン  $f_n$  ( $n = 1, \dots, N$ ) の相関係数はよく知られているように、  $L_2$  ( $R^2$ ) の内積とノルム\*を用いて

$$\rho_n = \frac{|(f, f_n)|}{\|f\| \cdot \|f_n\|} \quad (2)$$

と与えられ通常  $f$  と  $f_n$  の類似度とよばれている。  $\rho_n \in (0, 1)$  であり、パターンマッチング法による文字認識では  $\rho_n$  が1に近かければ近いほど  $f$  は  $f_n$  に類似しているものとみなされる。パーセバルの等式から、

$$\frac{|(f, f_n)|}{\|f\| \cdot \|f_n\|} = \frac{|(f^*, f_n^*)|}{\|f^*\| \cdot \|f_n^*\|} \quad (3)$$

がなり立つから式(2)の類似度は空間周波数スペクトルを用いても求め得ることは明らかである。ここで未知パターンに対する雑音除去処理をはじめとするいわゆる前処理を線形かつ空間不変的 (Space invariant) な処理  $\phi$  によって実現するものとすれば、式(2)の代りに

$$f = \phi f \quad (4)$$

で示される処理パターン  $f$  および照合用パターン  $f_n$  を用いることによってつぎのように新たな類似度を定義することができる。

$$\rho_n = \frac{|(f, f_n)|}{\|f\| \cdot \|f_n\|} \quad (n=1, \dots, N) \quad (5)$$

式(5)で示される類似度を Band 相関法では空間周波数スペクトルを用いて算出することになるが、このとき式(4)で示される作用素  $\phi$  を空間フィルタ化する必要がある。これを形式的に

$$\phi^* = F\phi F^{-1} \quad (6)$$

と書くことにする。ここで、処理パターン  $f$  のスペクトル  $f^*$  は  $f^* = \phi^* f^*$  によって求めることができ、照合用パターンについても  $\{f_n^*\}$  を得る。Band 相関法では式(5)の類似度を  $f^*$  と  $\{f_n^*\}$  によって

求めることになる。本稿では  $\{f_n^*\}$  のそれぞれをテンプレートとよぶことにする。

### 3.2 空間フィルタの設計

印圧の不均一さや、紙面の凹凸などによって文字の生成過程では通常『しみ』、『かすれ』などのストローク周辺の雑音が発生する。このような雑音が著るしい低品質の文字でもわれわれは別に不自由なく判読することができる。重要なことはわれわれがもし注意しようとするれば『しみ』、『かすれ』を十分にとらえることができる分解能を保有しているにもかかわらず、判読中はある意味で『意識的』に注意をそらしている事実である。しかも、そのような低品質の文字でも観念的には心線図形としての把握がなされていることも重要である。人間のもつこのような機能に似たものをOCRに付加する試みとして従来より low-pass フィルタリングあるいはボカシ操作などが行われ、実際その有効性が認められている。工学的には雑音除去とよばれているこれらの操作が実は分解能を積極的に落とすことによって、心線図形との区別がつかないという意味で雑音を含んだ文字を一つの心線図形に同定する前処理操作であると考えることができる。Band 相関法でも、空間フィルタの構成要素である low-pass フィルタが『しみ』、『かすれ』などの雑音除去の役割を担っているが、以下にその設計手順について述べることにする。

まず、有限のストローク幅をもつ文字を Fig. 2 のようにセグメントに分解したモデルに立脚して考察する。Fig. 1 に示している心線図形は『広がりがない』理想点の軌跡ないしは配列と解釈できるのに対し、具体的な文字は心線に沿うセグメントの軌跡ないしは配

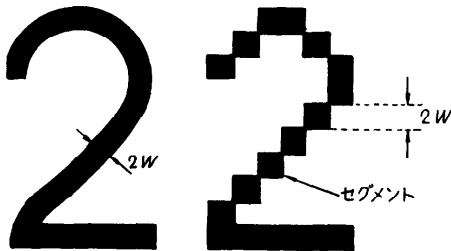
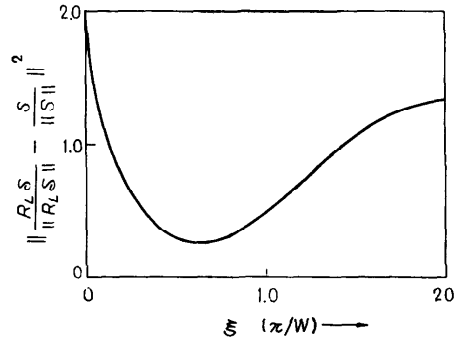


Fig. 2 An array of segments corresponding to a character

\* 式(8)の代りに式(2)に定めた類似度で評価しても結果は全く同じであるが<sup>13)</sup>、筆者等は習慣的にパターン距離を用いている<sup>7),8)</sup>。

\*\* 矩形パルスの幅を変化させながら標準化関数とのパターン距離を計算して求められる。



$$s(x) = \begin{cases} 1 & x \in [x_1 - W, x_1 + W, -W, x_2 + W] \\ 0 & \text{otherwise} \end{cases}$$

Fig. 3 Curve of the proposed measure vs. a parameter  $\xi$  of the square low-pass filter

列と考えることができよう。いま、2次元平面  $R^2$  の一点  $x_0$  におけるセグメントを画像関数  $s(x-x_0)$  で示し、同じく理想点をデルタ関数  $\delta(x-x_0)$  で示すことにしよう。まず、簡単のため一つのパラメータ  $\xi$  によって定まる low-pass フィルタ  $R_L^*$  をつぎのように定める。

$$R_L^*(\xi) f^*(\omega) = \begin{cases} f^*(\omega) & \omega \in D(\xi) \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

ただし、 $\omega=(\omega_1, \omega_2)$  は2次元空間周波数であり、 $D(\xi)$  は零周波数を中心として  $\xi$  にしたがう広がりをもつ low-pass 領域とする。たとえば、 $\xi$  を半径とする円形領域ないしは  $2\xi$  を辺とする正方形領域などが考えられる。筆者等は、上述の low-pass フィルタリングが究極的には理想点とセグメントとの区別がつかないようにする機能であるとみなして、 $R_L \delta$  と  $S$  が最も類似するようなパラメータ  $\xi$  を見出すことによって最適 low-pass フィルタを決定している<sup>7),8)</sup>。すなわち、つぎのように両者のパターン距離<sup>7),8)</sup> を最小にする  $\xi$  を選ぶのである。

$$M_{in} \frac{R_L \delta}{\xi} \frac{S}{\|R_L \delta\| \|S\|} \quad (8)$$

Fig. 3 は1次元問題に帰着して(正方形 low-pass フィルタと正方形パルス形セグメント)求めた計算結果<sup>7)\*\*</sup> である。

Low-pass フィルタの半径は小さければ小さいほどテンプレートに要するメモリを節減できる点でも望ましいことといえるが、式(5)に関連して

$$\lim_{\xi \rightarrow 0} \frac{|(R_L^*(\xi) f^*, R_L^*(\xi) f_n^*)|}{\xi \|R_L^*(\xi) f^*\| \cdot \|R_L^*(\xi) f_n^*\|} = 1 \quad (9)$$

が、すべての  $n$  について成立つから、 $\xi$  が小さ過ぎると入力パターンの識別が困難になることも事実である。したがって、式(8)に与えた  $\xi$  の決定法は low-pass フィルタ設計の一つのめやすを提供するものと考えることができる。

さて、以上に述べた low-pass フィルタに関する問題とは別の観点から空間フィルタリングと文字認識の関係について検討しよう。

いま、認識の対象となる文字集合がたとえば教育漢字あるいは英数字などのように限定されるとパターン集合としての共通の性格が定まるものと考えられる。このような共通の性格がもしスペクトル成分のどれかに反映しているならば、当然削除した方が識別性が高くなると考えられる<sup>1),2)</sup>。Band 相関法では、削除の対象となるスペクトル成分を正規化分散と名づけた一種のスペクトル変動率を調べることによって見出す手順を与えている。

まず、文字集合  $\{f_n\}$  ( $n=1, \dots, N$ ) の平均パターン  $f_0$  をつぎのように与える。

$$\begin{aligned} f_0(x) &= E\{f_n(x)\} \\ &= \frac{1}{N} \sum_{n=1}^N f_n(x) \end{aligned} \quad (10)$$

いま、 $\{f_n\}$  から  $f_n$  を除外した場合の平均パターンを  $g_n$  とし、

$$g_n(x) = \frac{Nf_0(x) - f_n(x)}{N-1}$$

のように表わすこととする。さて、2次元空間周波数平面上のある閉領域を  $\Omega$  としてつぎのような空間フィルタ  $T^*$  を導入する。すなわち、パターン  $f$  について、

$$T^*f^*(\omega) = \begin{cases} f^*(\omega) & \omega \in \Omega \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

なる機能をもつ空間フィルタである。ここで、 $f_0, f_n, g_n$  および  $T^*$  を用いてつぎのような一つの評価基準を設定する。

$$\bar{\rho} = \frac{|E\{(T^*f_n^*, T^*g_n^*)\}|}{\|T^*f_0^*\|^2} \quad (12)$$

$\bar{\rho}$  は空間フィルタ  $T^*$  を作用させた場合の異なる文字間の平均的な類似度を示すものであり、 $T^*$  を与えればただちに値が定まる。また、 $\bar{\rho}$  について、常に  $\bar{\rho} \in (0, 1)$  が成立つ。(証略)

ここで、 $\bar{\rho}$  が1に近くなる  $T^*$  すなわち、それに対応する周波数領域  $\Omega$  を見い出すことができれば、

空間フィルタリング  $f^* = T^*f^*$  によって得られるスペクトル成分  $f^*$  が実は識別に不必要な情報成分—— $\{f_n\}$  の共通の性格を反映する——と判断され、これを除去することによって逆に識別性を高めることができるものと期待される。この評価法を OCR 設計のためにさらに具体化する目的で筆者等によって正規化分散<sup>3)</sup>と名づけられているスペクトル変動率をつぎのように導入しよう。

$$\sigma^2(\omega) = \frac{E\{|f_n^*(\omega) - f_0^*(\omega)|^2\}}{|f_0^*(\omega)|^2} \quad (13)$$

式(12), (13)によって  $\bar{\rho}$  は  $\sigma^2(\omega)$  とつぎのように関係づけられる。

$$\bar{\rho} = \frac{\frac{1}{N-1} \left| \int_{\Omega} |f_0^*(\omega)|^2 (N-1 - \sigma^2(\omega)) d\omega \right|}{\|T^*f_0^*\|^2} \quad (14)$$

式(14)より  $\sigma^2(\omega) \approx 0$  の領域  $\Omega$  を選べば明らかに  $\bar{\rho} \approx 1$  となり平均的な識別性が最も悪くなる。すなわち  $\sigma^2(\omega) \approx 0$  の領域に属するスペクトル成分が空間フィルタリングによる削除の対象となるスペクトル成分であると判断できる。Fig. 4 に示している16種の文字(OCR-B)について計算した正規化分散を Table 1 (次頁参照)に示している。Table 1 は  $64 \times 64$  点のFFTを用いて行った計算結果のうち、零周波数付近の値を表示したものである。Table 1 より、零周波数の近傍で正規化分散が小さいことがわかる。このような零周波数の近傍(零近傍)領域を  $\Omega$  として、つぎのような highpass フィルタ  $R_H^*$  を定める。

$$R_H^*f^*(\omega) = \begin{cases} 0 & \omega \in \Omega \\ f^*(\omega) & \text{otherwise} \end{cases} \quad (15)$$

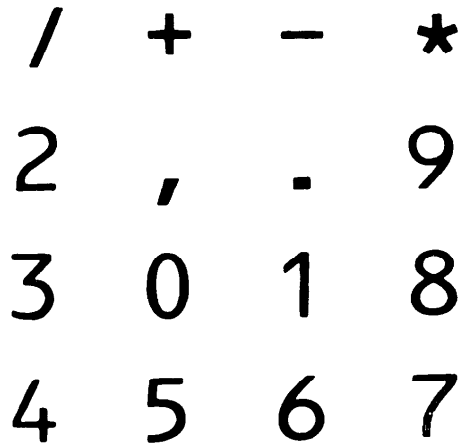
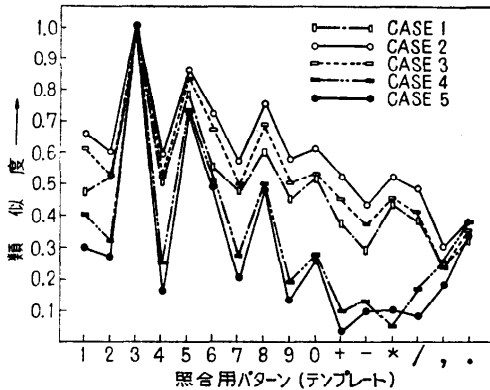


Fig. 4 16 Characters for experiment

\* この考え方は飯島の『正準化』<sup>2)</sup>と対応するものと考えられる。

Table 1 Normalized variance of characters

|    |       |       |      |      |      |       |       |
|----|-------|-------|------|------|------|-------|-------|
| 3  | 14.17 | 11.74 | 8.76 | 6.44 | 4.28 | 4.61  | 7.60  |
| 2  | 4.98  | 12.91 | 4.83 | 6.25 | 7.62 | 7.00  | 9.90  |
| 1  | 2.41  | 2.47  | 0.39 | 0.45 | 0.60 | 4.20  | 7.86  |
| 0  | 48.25 | 0.15  | 0.13 | 0.15 | 0.13 | 0.15  | 48.25 |
| -1 | 7.86  | 4.21  | 0.66 | 0.45 | 0.39 | 2.48  | 92.42 |
| -2 | 9.90  | 7.00  | 7.62 | 6.25 | 4.83 | 12.91 | 94.99 |
| -3 | 7.60  | 4.61  | 4.28 | 6.43 | 6.75 | 11.73 | 14.16 |
|    | -3    | -2    | -1   | 0    | 1    | 2     | 3     |



CASE 1 空間フィルタなし  
CASE 2  $D = \{|\omega||\omega| \leq 5.0[\pi/32]\}$  (low-pass フィルタのみ)  
CASE 3  $D = \{|\omega||\omega| < 1.0[\pi/32]\}$ ,  $D$  同上  
CASE 4  $D = \{|\omega||\omega| < 2.0[\pi/32]\}$ ,  $D$  同上  
CASE 5  $D = \{|\omega||\omega| < 2.0[\pi/32]\}$   
 $U\{(-2\pi/32, 0), (2\pi/32, 0)\}$ ,  $D$  同上

Fig. 5 Some spatial filters and their influences on similarity measures

$R_H^*$  は式 (12) の意味で識別に不必要なスペクトル成分を削除する空間フィルタである。ここで、式 (8) によって決定した low-pass フィルタ  $R_L^*$  との合成によって

$$R^* = R_H^* R_L^* \quad (16)$$

なる band-pass フィルタ  $R^*$  を得る。 $R^*$  の機能を確認するために行った類似度計算の一例を Fig. 5 に示そう。なお  $R^*$  は式 (6) の  $\Phi^*$  を具体化したものである。

Fig. 5 の結果は円形 low-pass フィルタ (半径  $\xi$ ) を採用した場合であり、かつ入力パターンとして雑音のない OCR-B フォントの『3』(Fig. 4 参照) を選んでいる。low-pass フィルタのみを用いた場合は、よく知られているようにフィルタなしの場合と比較して全体として類似度が大きくなるのがわかる。(CASE 1 と CASE 2 を比較) しかし、正規化分散が小さい

零近傍のスペクトルを徐々に削除して行くと次第に相互類似度 (『3』以外との類似度) が全体的に小さくなる band-pass フィルタリングの良好な効果が確認できる (CASE 3~CASE 5)。

Fig. 5 で用いられている low-pass フィルタは、式 (8) の意味で最適半径を有している。low-pass フィルタリングあるいはボカシ操作が実用的 OCR で有効であることは衆知の事実であるが、本稿の最適 low-pass フィルタが実用的意味でも最適であるかどうかは別の問題である。たとえば、認識対象として与えられる文字集合のカテゴリーが少ないときは思いきって半径を小さくしても充分識別できる場合があり、これによってメモリを節減できる利点が出てくる。一方ではこの逆の場合もあり得ると考えられる。

本稿では、この意味で low-pass フィルタについては、半径決定に際して一つのみやすとなる式 (8) を示すことにとどめ、実用的効果に関する定量的な考察は割愛する。

### 3.3 位置正規化操作

複素振幅型 Band 相関法は文字の識別に必要な本質的情報の欠損がないため、高性能の認識能力を有しているにもかかわらず、入力パターンの位置ずれ正規化を抜きにした場合、その能力を最大限に発揮できないことはすでに述べている。ここでは、スペクトル情報を用いる入力パターンの位置推定とそれに伴う位置正規化の方法について述べる。

通常、OCR に用いられる位置ぎめ操作は入力文字の重心ないしは外接矩形枠の中心等にもとづいて行われる。本稿では、重心による位置ぎめを採用することにする。一般に、重心による位置ぎめは入力パターンの各部に一樣に加わる雑音には有効に対処できる反面、入力パターンが濃度傾斜をもつ場合あるいは一部に『欠け』がある場合には誤差が生じるという欠陥があると考えられる。しかし、位置ぎめに大きな誤差を生じた場合、最大類似度が小さくなるなどの現象に着目して入力パターンの印字品質が悪いものと判断し、リジェクトすることによって誤判読を避けることができる。さて、認識対象となるパターンが OCR の観測装置によって切り出され、Fig. 6 (次頁参照) のように切り出し枠の中に定められるものとしよう。位置座標の原点  $0$  を切り出し枠の中心にとったとき、入力パターンが  $f(x)$  と表わされるものとする。

このパターンを重心  $G$  が原点  $0$  に一致するまで平行移動して得られるパターンを  $\tilde{f}(x)$  で示せば、

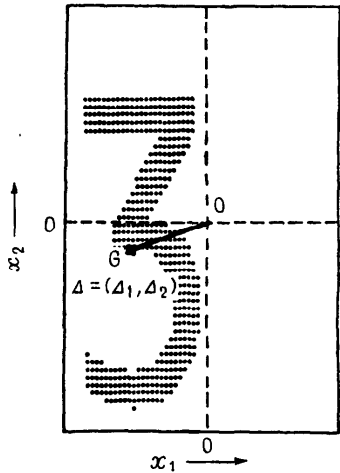


Fig. 6 An input pattern set in the frame of a sensory device

$$f(x) = \check{f}(x - \Delta) \quad (17)$$

と書ける.  $\Delta = (\Delta_1, \Delta_2)$  は  $f(x)$  の重心で,

$$\Delta_1 = \frac{\int x_1 f(x) dx}{\int f(x) dx}, \quad \Delta_2 = \frac{\int x_2 f(x) dx}{\int f(x) dx} \quad (18)$$

のように求められる.

$\check{f}$  を  $f$  の位置正規化パターンとよぶことにする.

式(17)の両辺にフーリエ変換を施して

$$f^*(\omega) = \check{f}^*(\omega) \exp -j\omega \Delta^T \quad (19)$$

を得る. いま, 2次元空間周波数  $\omega = (\omega_1, \omega_2)$  で  $\omega_2 = 0$  において  $f^*(\omega_1, 0)$  を考えよう.

$f^*(\omega_1, 0)$  を  $\omega_1 = 0$  のまわりでテイラー展開し3次以上の項を無視してつぎの式を得る.

$$f^*(\omega_1, 0) = f^*(0, 0) + f'^*(0, 0)\omega_1 + \frac{1}{2}f''^*(0, 0)\omega_1^2 \quad (20)$$

ただし,  $\cdot$  および  $\cdot\cdot$  はそれぞれ1次および2次導関数を示す. 一方,

$$f^*(0, 0) = \int f(x) dx \quad (21)$$

$$f'^*(0, 0) = -j \int x_1 f(x) dx \quad (22)$$

$$f''^*(0, 0) = -\int x_1^2 f(x) dx \quad (23)$$

\*  $\int x_1 f(x) dx \equiv \iint_F x_1 f(x_1, x_2) dx_1 dx_2$  であり,  $F$  は切り出し枠に囲まれる領域とする.

となるから, 式(18)および(20)~(23)より次式を得る.

$$f^*(\omega_1, 0) = f^*(0, 0) \left[ 1 - \frac{\omega_1^2}{2} \frac{\int x_1^2 f(x) dx}{\int f(x) dx} + j\omega_1 \Delta_1 \right] \quad (24)$$

式(24)より,  $f^*(\omega_1, 0)$  の虚数部が  $\Delta_1$  に比例することがわかる. 同様の結果が  $\Delta_2$  についても得られるから, 位置(重心位置)の推定式を次式のように定めることができる.

$$\Delta_1 = -j \frac{I_m\{f^*(\omega_1, 0)\}}{\omega_1 f^*(0, 0)} \quad (25)$$

$$\Delta_2 = -j \frac{I_m\{f^*(0, \omega_2)\}}{\omega_2 f^*(0, 0)} \quad (26)$$

さて, 式(25), (26)にもとづいて推定された  $\Delta = (\Delta_1, \Delta_2)$  を用いれば, 式(19)から

$$\check{f}^*(\omega) = f^*(\omega) \exp j\omega \Delta^T \quad (27)$$

によって入力パターン  $f(x)$  の位置正規化パターンが複素振幅スペクトル  $\check{f}^*$  の形で得られることになる. 特に, 式(25), (26)で用いる空間周波数は式(20)の仮定から零近傍を選ぶことになるが, これは識別に不必要であるからという理由で零近傍を band-pass フィルタリングによって単に捨て去るだけでなく, 位置正規化に役立てるといふ複素振幅スペクトルの『全体的』な利用法を提起していることになる.

さて, Fig. 4 に示した各文字を用いて行った位置推定の結果を示そう. Fig. 7 は基準位置から人為的に入力パターンを平行移動させたときの移動量と式(25)にもとづく推定位置の関係を示している.

計算機実験では, 各文字を Fig. 6 のようにストロ

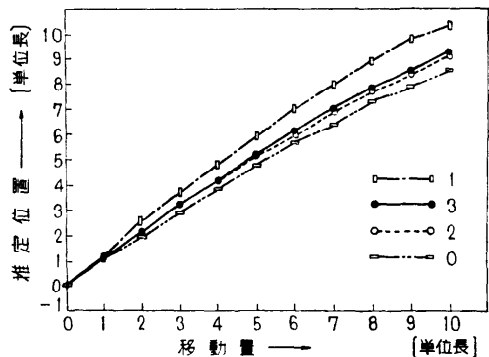


Fig. 7 Experimental results of estimation of position deviation

ーク幅が平均 6 単位長となるよう標準化したものを入力パターンとして用いた。Fig. 7 に示されていない文字についての結果はすべて『上限』1 と『下限』0 の間におさまる。Fig. 7 は、式(25)による位置推定がかなりの精度であることを示している。

### 3.4 テンプレートの設計

式(27)にしたがって入力パターンが位置正規化されるに伴って各テンプレートも位置正規化された照合用パターンにもとづいて設計する必要がある。問題は、認識対象として与えられる印刷文字パターンのフォントについてその設計仕様定められるみかけ上の基準位置と重心とが必ずしも一致しないことである。つまり、手続きとしてはまず各照合用パターンの重心を求めて、これを切り出し枠の中心に位置正規化して後、スペクトルを計算してテンプレートを設計しなければならない。ただし、入力パターンの位置正規化法をテンプレート設計にも適用することによって手続きを単純化することができる。また、テンプレートの形でストアすべきスペクトル情報は band-pass フィルタを通過する成分のみでよいことは明らかである。

さて、テンプレート群に対する入力パターンの類似度を計算し、最大類似度を与えるテンプレートのカテゴリーに同定するというパターンマッチング法の基本原理そのものに問題はないとしても、印刷文字用の実用的 OCR を考える際にはいくつかの技術的問題が派生してくる。まず第一に、たとえば漢字などを認識対象とする場合、ぼう大な数にのぼるカテゴリーに対応するテンプレート群に総当たり式に入力パターンを照合することは実際的でなく、当然階層的照合を考えなくてはならない。このとき、カテゴリーの階層化のアルゴリズムについてはいくつかの提案があるが<sup>9)</sup>、現在のところ必ずしも議論の余地のないものではないと考ええられる。

一方、本稿で述べた重心にもとづく位置正規化法を含めて、一般に入力パターンの濃度傾斜、欠け等による位置推定誤差は避けられない。この誤差の影響の軽減のためには、たとえば飯島の複合類似度法<sup>10)</sup>の採用等が考えられる。また、筆者等も独自に誤差の影響の少ない類似度の計算法について検討中であるが、これらについては別の機会に報告することとする。

## 4. むすび

本稿では、複素振幅型 Band 相関法による OCR の設計理論を中心に述べた。とりわけ、band-pass フィ

ルタリングがはたす役割が大きいことを指摘し、さらにフィルタの設計が認識対象となる文字集合が与えられれば、確定した手順で客観的に行えることを示した。これは一つの文字集合を認識する OCR に重大な変更を加えずに別の文字集合を認識させることができるという汎用性をもつ OCR を実現する重要な条件の一つである。また、本稿ではスペクトルの『全体的』な利用によって位置正規化が実現できることを示した。筆者等は、Band 相関法による OCR 方式を印刷英数字、カタカナおよび漢字等に適用するための基礎実験を実施中である。それぞれの結果については機会を改めて報告したい。最後に、本稿での計算はすべて大阪大学大型計算機センター NEAC 2200-700 および京都大学大型計算機センター FACOM 230-75 を用いて行ったことを付記する。

謝辞 実験に協力された大阪電通大卒研究生柴田明君(現在日本ユニバック(株))に深謝します。

## 参考文献

- 1) 小沢, 田中: "Band 相関法によるパターン認識" 第 2 回画像工学コンファレンス (昭46-11)
- 2) 田中, 小沢: "空間フィルタリングによるパターン認識" 信学会インホメーション理論研究会資料 IT 71-26 (昭 46-05)
- 3) K. Ozawa: "Pattern Recognition by Spatial Frequency Filtering", Doctoral thesis of Osaka University, 1972
- 4) 小沢, 田中: "Band 相関法による OCR 方式", 信学会パターン認識と学習研究会資料 PRL 73-94 (昭 49-04)
- 5) 小沢, 田中: "Band 相関法による OCR 方式 (その 2)" 同上研究会資料 PRL 74-43 (昭 49-12)
- 6) K. Ozawa and K. Tanaka: "Character Recognition Using Correlation Technique on Spatial Complex-frequency Plane", the 2nd International Joint Conference on Pattern Recognition, 1974
- 7) 小沢, 田中: "コヒーレント光を用いた視覚パターンの特徴抽出(8)" 昭 46 信学会全大予稿, 132
- 8) 小沢, 田中: "文字図形認識のための low-pass フィルタリング", テレビジョン, 27 No. 4 (昭 48-04)
- 9) 山本, 中島, 中田: "階層的パターンマッチング法による漢字認識の実験——印刷漢字認識の研究——" 信学論(D), 56-D, No. 12 (昭 48-12)
- 10) 飯島: "文字読取装置 ASPET/71", テレビジョン, 27, No. 3 (昭 48-03)

- 11) 三木, 河合, 田村, 田中, 小沢: “光学的文字読取に関する研究” 信学会パターン認識と学習研究会資料 PRL 73-74 (昭 48-12)
- 12) 飯島: “パターン認識” コロナ社 (昭 48)
- 13) 山崎, 飯島: “文字図形の標準化について” 信学論(C), 51-C, 9, p. 428 (昭 48-09)

は以下の通りである.

特開昭 48-29326 (公開日昭 48. 4. 18, 出願日昭 46. 8. 19)  
特開昭 49-47041 (公開日昭 49. 5. 7, 出願日昭 47. 9. 13)  
特開昭 48-45141 (公開日昭 48. 6. 28, 出願日昭 46. 8. 19)  
特開昭 49-42363 (公開日昭 49. 4. 20, 出願日昭 47. 8. 9)  
および出願中 2 件 (昭 49. 3. 4, 昭 49. 9. 24)

(昭和 50 年 6 月 10 日受付)

(昭和 50 年 7 月 28 日再受付)

◇  
© 本稿で述べた OCR 方式に関する特許出願状況