



HOKKAIDO UNIVERSITY

Title	任意サンプル数のデータによる二次元高速フーリエ変換法（第3報）：その漢字認識への応用
Author(s)	仲谷, 元; Nakatani, Hajime; 青木, 由直 他
Citation	北海道大學工学部研究報告, 72, 73-81
Issue Date	1974-09-14
Doc URL	https://hdl.handle.net/2115/41236
Type	departmental bulletin paper
File Information	72_73-82.pdf



任意サンプル数のデータによる二次元 高速フーリエ変換法 (第 3 報)

— その漢字認識への応用 —

仲谷 元* 青木由直*

(昭和 49 年 3 月 30 日受理)

Two-Dimensional Fast Fourier Transform with Data of Arbitrary Sampling Number (3)

— Its Application for Chinese-Character Recognition —

Hajime NAKATANI Yoshinao AOKI

(Received March 30, 1974)

Abstract

An application of a two-dimensional fast Fourier transform for Chinese-character recognition is discussed. The character recognition is done by a cross-correlation operation with two sampled and quantized Chinese-character patterns. The cross-correlation is calculated by a two-dimensional fast Fourier transform algorithm to shorten the execution time. It is shown that the mode of the sampling of original patterns has a great influence on the recognition process. By taking a suitable sampling number, the recognition time is shortened without decreasing the recognition accuracy, resulting in an increase of the recognition efficiency.

1. ま え が き

任意サンプル数のデータによる二次元高速フーリエ変換法 (FFT) のアルゴリズム及びそのプログラミングについては、このテーマに関する研究の第 1 報¹⁾、第 2 報²⁾ で論じた。本報告では二次元高速フーリエ変換法の漢字認識への応用について論じている。

計算機により漢字の認識を行なうためには、あらかじめコアメモリや磁気ディスクなどの記憶装置に標準となる漢字を格納しておくことが必要である。標準漢字と入力漢字のパターンマッチングにより認識を行なうことになるが、漢字は教育漢字で約 900 字、当用漢字では約 2,000 字というように莫大なカテゴリを有することから、その方法には工夫の必要がある。方向と大きさが規格化されている漢字の認識には、規格化された相互相関値により認識する方法が有効である。相互相関値は二次元高速フーリエ変換法を用いることで短時間で計算されることから、一回当たりのマッチング計算時間はある程度短縮される可能性をもつ。相互相関値の計算は入力される漢字に対するサンプリングの方法によって影響を受けるが、比較的粗いサンプリングによっても十分認識は可能である。このことによって標準漢字の記憶容量を減らすことができ、同時にマ

* 電子工学科

ッチング計算時間はさらに短縮される。筆者らは「谷」, 「合」, 「中」の三種類の漢字に対してこれらの点に関する認識実験を行なったのでここに報告する。

2. 漢字認識の原理

文字認識を機械により行なう場合にその方法はいろいろ考えられるが、大きく分けて次の二つの方法になる。一方は標準文字と入力文字との一致の度合いから認識する「パターンマッチング法」であり、他方は文字線の端点や交点などの特徴を抽出してその組合わせから認識する「特徴抽出法」である。一般に標準文字の定めやすい印刷文字の認識には「パターンマッチング法」が用いられ、個人差が現われる手書文字の認識には「特徴抽出法」が用いられるようである。本報告では標準漢字の間での認識を論じているところから「パターンマッチング法」を用いている。「パターンマッチング法」は文字の一致の度合いによる認識方法であるから、何らかの手段により文字間の類似性を定量化する必要がある。この特徴量を類似度と呼ぶならば、類似度はパターンの本質を変えない種々の変形、すなわち回転、併進、伸縮、濃淡の変化などによる影響を受けない安定なものであることが望ましい。

通信の分野では二つの信号の関連性を調べるのに相互相関関数を用いることがある。これは一種の重なり量を表現するものであるから漢字パターンの一致を知る場合にも適用することができる。相互相関関数は二つの関数を $f(\xi, \eta)$, $g(\xi, \eta)$ とすると式(1)で定義される。

$$\Phi(x, y) = \iint_{-\infty}^{+\infty} g(\xi - x, \eta - y) f(\xi, \eta) d\xi d\eta \quad (1)$$

式(1)は一方の関数の併進に対しても同一の併進を受けることがわかる。すなわち、相互相関関数は併進による影響を受けないと考えられる。

また、可積分であり非負値実数値の関数 g, f に対してはシュワルツの不等式として知られる式(2)の関数が成立する。

$$\int af \leq \left(\int g^2 \int f^2 \right)^{\frac{1}{2}} \quad (2)$$

この関係式は任意の積分領域上で成立し、等号は定数 c に対して $f = cg$ となるときのみ成立する。式(2)を式(1)に適用するならば式(3)のように規格化された相互相関関数が求められる。

$$\Phi(x, y) = \frac{\iint_{-\infty}^{+\infty} g(\xi - x, \eta - y) f(\xi, \eta) d\xi d\eta}{\left\{ \iint_{-\infty}^{+\infty} g^2(\xi, \eta) d\xi d\eta \iint_{-\infty}^{+\infty} f^2(\xi, \eta) d\xi d\eta \right\}^{\frac{1}{2}}} \leq 1 \quad (3)$$

ただし、等号は任意定数 c に対して $f = cg$ となるときのみ成立する。したがって、規格化された相互相関関数を類似度に用いればパターンの併進や濃淡の度合いによる影響を受けない認識が可能となることがわかる。

式(3)は直接に計算することもできるが、高速フーリエ変換法を用いることで短時間での計算が可能となる。すなわち、式(3)の分子を、各関数のフーリエ変換を $F(f_x, f_y)$, $G(f_x, f_y)$ とし、フーリエ変換表示すれば式(4)となる。

$$\begin{aligned} & \iint_{-\infty}^{+\infty} \left\{ \iint_{-\infty}^{+\infty} g(\xi - x, \eta - y) f(\xi, \eta) d\xi d\eta \right\} e^{-j2\pi(f_x x + f_y y)} dx dy \\ &= \iint_{-\infty}^{+\infty} g(\xi', \eta') e^{j2\pi(f_x \xi' + f_y \eta')} d\xi' d\eta' \iint_{-\infty}^{+\infty} f(\xi' + x, \eta' + y) e^{-j2\pi\{f_x(\xi' + x) + f_y(\eta' + y)\}} dx dy \\ &= G^*(f_x, f_y) F(f_x, f_y) \end{aligned} \quad (4)$$

ただし*印は複素共役を表わす。

また、式(3)の分母の一項については、式(4)の特別な場合であり $x=y=0$ として求められるが、これはその最大値になると考えてよい。フーリエ逆変換を \mathcal{F}^{-1} で表わせば式(4)を用いて式(5)のように表わされる。

$$\begin{aligned} \iint_{-\infty}^{+\infty} g^2(\xi, \eta) d\xi d\eta &= \mathcal{F}^{-1} \left[G^*(f_x, f_y) G(f_x, f_y) \right]_{x=y=0} \\ &= \mathcal{F}^{-1} \left[|G(f_x, f_y)|^2 \right]_{\max} \end{aligned} \quad (5)$$

つまり、式(3)をFFTによる計算式で示せば式(6)となる。

$$\Phi(x, y) = \frac{\mathcal{F}^{-1} \left\{ G^*(f_x, f_y) F(f_x, f_y) \right\}}{\left[\mathcal{F}^{-1} \left\{ |G(f_x, f_y)|^2 \right\}_{\max} \mathcal{F}^{-1} \left\{ |F(f_x, f_y)|^2 \right\}_{\max} \right]^{\frac{1}{2}}} \quad (6)$$

結局、漢字パターンの大きさと方向とが規格化されている場合には、標準漢字として $G^*(f_x, f_y)$ を記憶しておき、入力漢字に対して相互相関関数の最大値 Φ_{\max} を式(6)によって求め、十分小さな θ に対して $1 - \Phi_{\max} < \theta$ が成立するかどうかで漢字の認識が可能となる。こうした認識過程をブロック図で示せば図-1のようになる。

3. 標準パターンでの認識

図-1で示したような方法で認識を行なうためには、入力漢字をサンプリングと量子化によりデジタル化しなければならない。そこで漢字を含む文字面を縦方向 M 、横方向 N のサンプル数でサンプリングし、文字上であるか、余白上であるかにより二値化するものとする。このような方法でデジタル化された入力漢字 $f(i, j)$ と標準漢字 $g(i, j)$ に対して規格化された相互相関関数を求めるのだが、この場合には積分を和におきかえることで式(7)を計算することになる。

$$\Phi(k, l) = \frac{\sum_i^M \sum_j^N g(i-k, j-l) f(i, j)}{\left\{ \sum_i^M \sum_j^N g^2(i, j) \sum_i^M \sum_j^N f^2(i, j) \right\}^{\frac{1}{2}}} \quad (7)$$

以下に実際に行なった認識実験について報告する。

実験に使用した漢字は「谷」、「合」、「中」の三種類であり、各漢字に対して縦横同数 ($M=N=64$) にサンプル数をとった細かいサンプリングを行なって標準パターンを作成した。これら

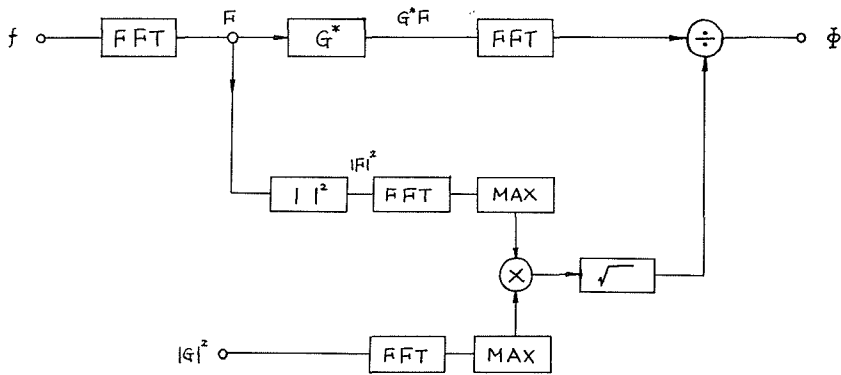


図-1 規格化された相互相関関数の計算

の標準パターンにより、入力漢字を「谷」として、標準漢字「谷」、「合」、「中」との相互相関値を **FFT** を用いた計算によって求めた各マッチング計算の結果は変位 (k, l) に関する関数として得られるが、相関の様子を知るために $X-Y$ プロッターを使用して透視画表示した。図-2は「谷・谷」に対する相互相関を示したものであり、図-3は「谷・合」、図-4は「谷・中」に対する相互相関を示したものである。これらの図をみてわかるように、パターン的一致に対しては他と明瞭に区別されるような非常に鋭い相関が示されている。各相関値の最大値を示せば、「谷・谷」に対しては1.00であり、「谷・合」に対しては0.65、また「谷・中」に対しては0.42である。この計算結果は「中」よりも「合」の方が「谷」によく類似しているという我々の感覚とよく一致するものであり、相互相関値を漢字の認識に使用することの妥当性を示しているといえよう。ここでは三種類の漢字パターンに対してだけ認識実験を行なったのであるが、パターン的一致に対して相互相関値は最大値1をとり、他の場合にはそれ以下の値をとることになり、漢字パターンの数が増した場合においても適当なしきい値を選ぶことで十分認識が可能となるように思われる。

さて、ここでは縦横同数のサンプル数でのサンプリングにより得られた標準パターンを用いて認識実験を試みたのであるが、サンプル数をどの程度に選んでサンプリングを行なうかは検討の価値がある問題である。すなわち、サンプル数を少なくして粗いサンプリングを行なうことによっても十分に認識が可能であるような結果が得られるならば、このことは認識時間の短縮につながり、認識能率が上がることになるからである。そこで、次にサンプル数の減少によるパターンの圧縮が認識に与える影響について論じることにする。

4. 圧縮パターンでの認識

ここでは横方向のサンプル数を $N=64$ と一定に保ち、縦方向のサンプル数 M を減少させた場合について、その認識に与える影響を論じることにする。すなわち、標準パターンの縦方向の圧縮により得られるパターンに対して認識実験を行なうのである。

そこで、縦方向サンプル数 M を減らして M/L にとった場合、図5のように縦方向 L の間の二次元情報が一次元情報にそれぞれ圧縮されるものとする。二値情報をもつ $M \times N$ のサンプル

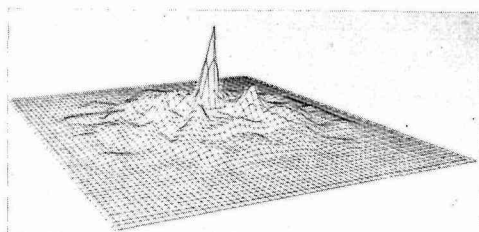


図-2 「谷・谷」の自己相関 ($M=N=64$)

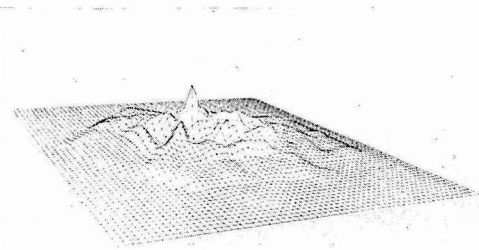


図-3 「谷・合」の相互相関 ($M=N=64$)

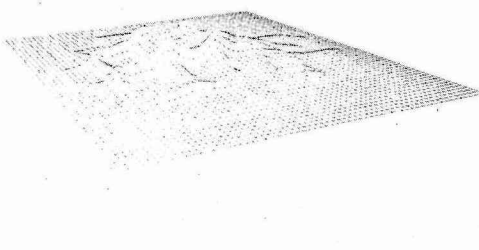


図-4 「谷・中」の相互相関 ($M=N=64$)

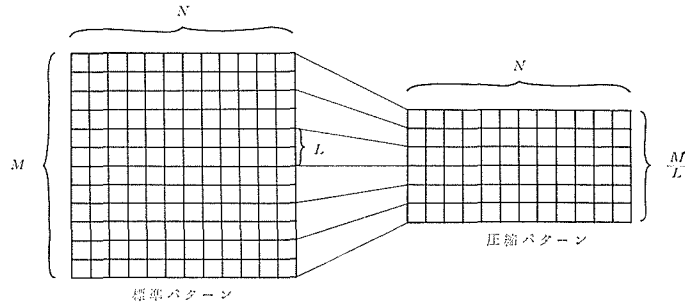


図-5 パターンの圧縮

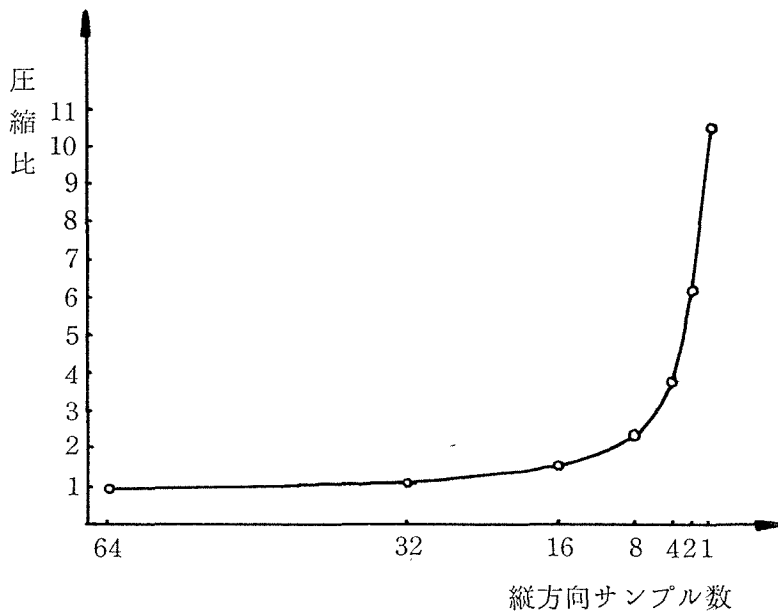


図-6 縦縮比

点により表現されたデジタルパターンのもつ情報量は MN bits であり、図-5 に示すような圧縮を受けた後のパターンのもつ情報量は $(MN/L) \log_2(L+1)$ bits となるから、このような場合の情報量の圧縮比 (CR) は式 (8) で示される。

$$CR = \frac{L}{\log_2(L+1)} \quad (8)$$

ここで M' を圧縮パターンに対する縦方向サンプル数とすると式 (8) は式 (9) となる。

$$CR = \frac{M}{M' \log_2\left(\frac{M}{M'} + 1\right)} \quad (9)$$

$M=64$ にとった場合の M' に対する縦縮比の変化を図示すれば図-6 のようになる。図-6 に示されるように、サンプル数の減少にもなって縦縮比が非線形に変化するが、サンプル数のある程度の減少に対しては縦縮比はそれほど変化せずほぼ一定値を保っている。このことから、サンプル数のある程度の減少に対しては十分な認識が可能であるように思われる。

そこで、縦方向サンプル数を32, 16, 8, 4, 2と減少させた場合について前と同様の認識実験を行なった。各場合に対するマッチング計算による相関の様子は図-7(a), (b), (c)~図-11(a), (b), (c)に示すとおりである。各図をみればサンプル数の減少による相関の変化がよくわかるが、各場合の類似度(相関関数の最大値)についてサンプル数の減少による変化を示せば表-1, 図-12のようになる。図-12は漢字の認識におけるサンプル数の選び方を示唆するものであるが、ここに示した類似度の変化は図-6に示した圧縮比の変化に類似していることに気づく、すなわち、パターンの圧縮による情報量の減少が類似度の計算に深い影響をもっているものと考えられる。この点からすると、図-12には三種類の漢字パターンに対する認識結果としてその類似度の変化を示し

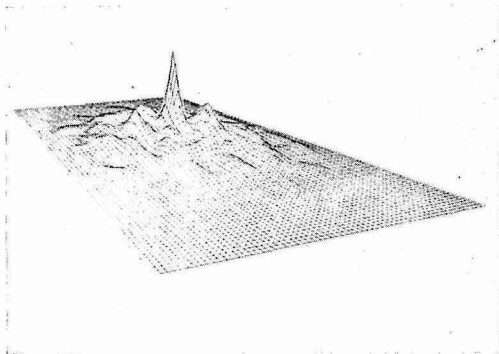


図-7(a) 「谷・谷」の自己相関 ($M=32, N=64$)

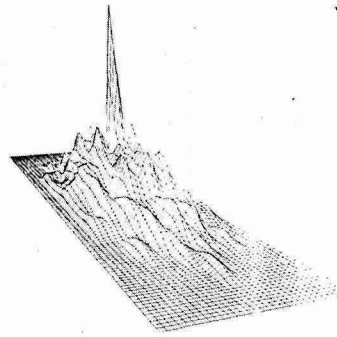


図-8(a) 「谷・谷」の自己相関 ($M=16, N=64$)

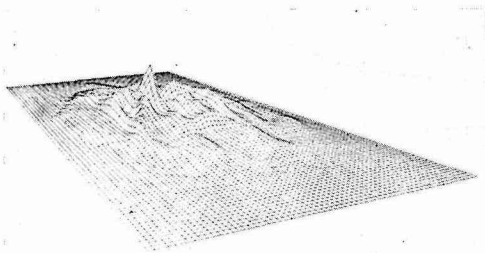


図-7(b) 「谷・合」の相互相関 ($M=32, N=64$)

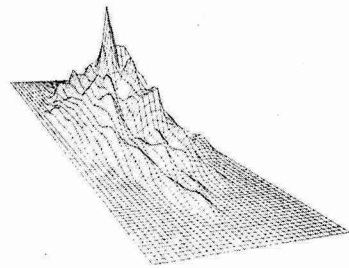


図-8(b) 「谷・合」の相互相関 ($M=16, N=64$)

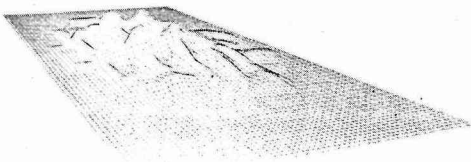


図-7(c) 「谷・中」の相互相関 ($M=32, N=64$)

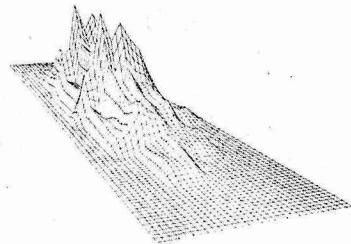


図-8(c) 「谷・中」の相互相関 ($M=16, N=64$)

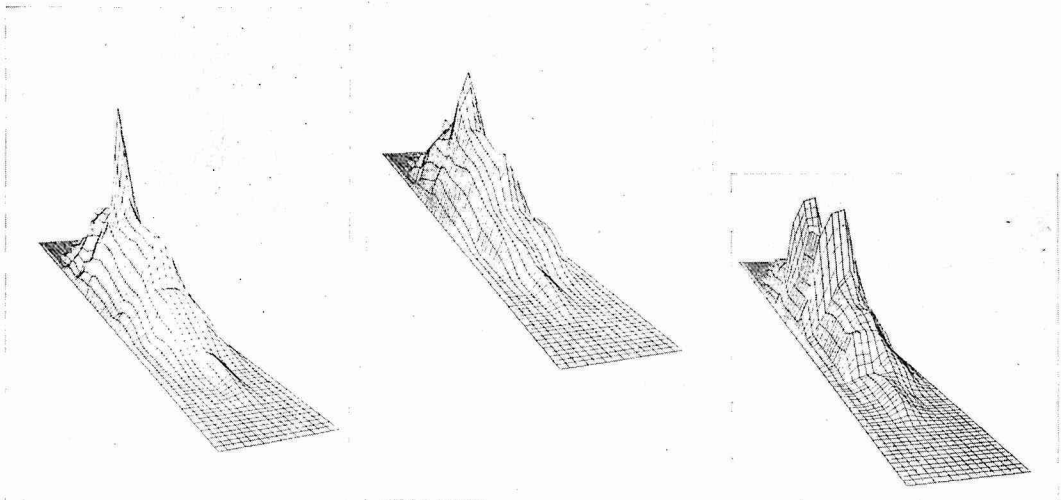


図-9(a) 「谷・谷」の自己相関
($M=8, N=64$)

図-9(b) 「谷・谷」の相互相関
($M=8, N=64$)

図-9(c) 「谷・中」の相互相関
($M=8, N=64$)

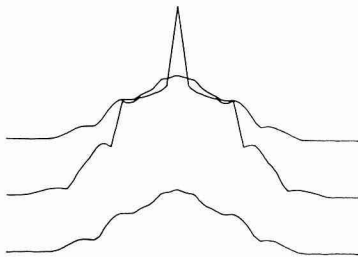


図-10(a) 「谷・谷」の自己相関
($M=4, N=64$)

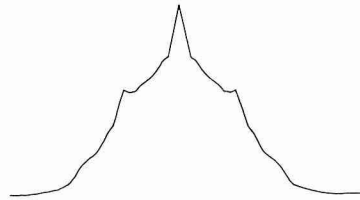


図-11(a) 「谷・谷」の自己相関
($M=2, N=64$)

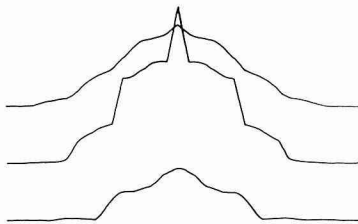


図-10(b) 「谷・谷」の相互相関
($M=4, N=64$)

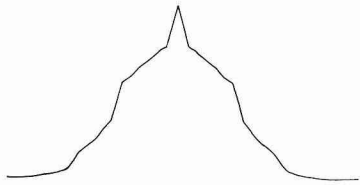


図-11(b) 「谷・谷」の相互相関
($M=2, N=64$)

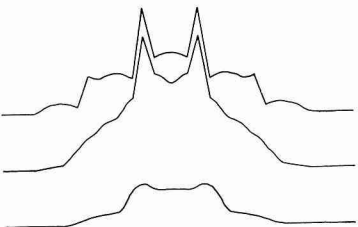


図-10(c) 「谷・中」の相互相関
($M=4, N=64$)

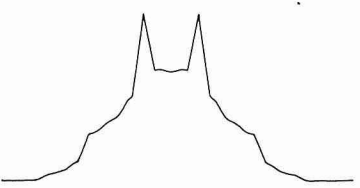
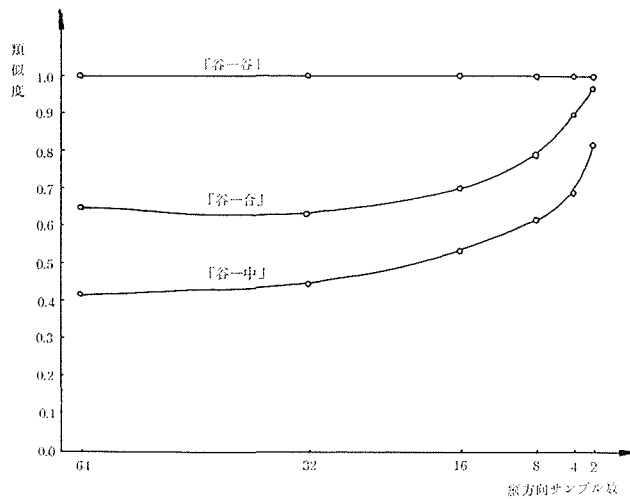


図-11(c) 「谷・中」の相互相関
($M=2, N=64$)

表一 類 似 度

	サ ン プ ル 数					
	64	32	16	8	4	2
「谷・谷」	1.00	1.00	1.00	1.00	1.00	1.00
「谷・合」	0.65	0.61	0.70	0.79	0.90	0.97
「谷・中」	0.42	0.45	0.54	0.62	0.69	0.82



図一12 縦方向サンプル数に対する類似度の変化

ているのだが他の漢字パターンに対しても同じようなサンプル数の減少に対する変化を示すものと考えられる。パターン的一致に対してはサンプル数によらずに鋭い相関を示し、類似度は1.00をとっている。しかし、異なるパターンの間での類似度はサンプル数の減少にともなって非線形に増加しており、その値は上限値1.00に接近することになる。パターンの圧縮がなされない情報量の多い場合においては、類似度はそれぞれある極限值に収束しているが、これは二つのパターンを連続関数と考えたときの積分で与えられる相関値に収束しているのである。この点からすると、サンプル数を大きくとって微細なサンプリングを行なうことは、冗長な情報を増すだけであり、本来の認識に対する効果は少ないと考えられよう。また、パターンを圧縮して二次元情報を完全に一次元情報にしてしまった場合には、いわゆる縦方向の周辺分布を用いた認識^{3),4)}を行なっているのだが、類似度はパターンの過度の圧縮による情報量の減少から上限値付近に集中してくることが予想される。また、この場合には全く異なる漢字パターンであってもその構造によっては圧縮後のパターンが類似したものになることも考えられ、良好な認識は期待できそうにない。しかしながら、多くなく少なくとも中間のサンプル数に対しては、注目に値する類似度の変化を示している。すなわち、これらのサンプル数においては、多くサンプル数をとった場合の類似度と比較してそれほど変わらなく、その集中もみられない。したがって、サンプリングを細かく行なったパターンにおいての認識と同じような認識が、パターンの過度の圧縮による情報量の減少をとまわらない適切なサンプル数によってサンプリングを行なったパターンにより可能であるといえる。このことによって、計算機において標準漢字の格納に必要とする記憶容量を減少させ

ることができ、同時に一回当たりのマッチング計算時間を短縮できることになる。それゆえ、漢字を分類してマッチングをとる漢字のカテゴリを減少させることができるならば、ここに示したような漢字認識の方法はきわめて有効なものとなるであろう。

5. あとがき

任意サンプル数のデータによる二次元高速フーリエ変換法の漢字認識への応用について論じた。二次元 **FFT** を使用してパターン間の相互相関を計算することによる認識方法を示し、筆者らが行なった認識実験について報告した。ここではサンプル数を 2^n ($n=1, 2, \dots$) にとって実験を行なっているが、サンプル数は任意に選ぶことができる。

二次元 **FFT** は、この他計算機ホログラフィなどの計算機光学の分野で使用され成果をあげている。

文 献

- 1) 青木由直：工学部研究報告, 67 (昭48), pp. 73-83.
- 2) 青木由直：工学部研究報告, 69 (昭48), pp. 125-135.
- 3) 中野, 中田：“周辺分布による漢字の分類と認識” 信学会, オートマトン研資 (昭46-09).
- 4) 中野, 中田, 他：“周辺分布のスペクトルを用いた印刷漢字の認識” 信学会, パターン認識と学習研資 (昭47-05).