

梵字認識システムの開発

西田 茂生 堀崎 遼一

Development of Sanskrit characters recognition system

Shigeki NISHIDA and Ryoichi HORISAKI

梵字は古代インドにおいて使用された過去の文字である。しかし、当時のアジアではこの梵字が共通文字として使用されていた。すなわち、梵字によって古代アジア文明が発展してきたとって過言はないであろう。本研究では遺跡や古文獻に残されている保存状態の悪い梵字を自動認識するシステムの構築を目的とする。その第1段階として個々の梵字を自動認識するアルゴリズムを構築した。本研究では認識の手法として文字画像がもつ光強度分布の相関度を用いる。この手法により文献や木の卒塔婆などに記されている梵字の認識を試みた結果60%以上の認識率を得た。

1. 緒言

文明は文字と共に発達してきた。世界中の遺跡には古代文字が数多く残されている。本研究ではそれらの古代文字の中から古代アジア文明を解明する上で欠かすことのできない梵字（サンスクリット文字）を取り上げる。またこの梵字は仏教の経典、墓地の卒塔婆などで、我々になじみの深い文字である。梵字の認識システムが構築できれば、考古学の発展に貢献でき、また梵字の専門知識がない人でも梵字で書かれた経典を読めるようになるであろう。また、この技術を応用することにより日本語の崩し字の認識に役立つと考えられる。

梵字は古代遺跡、仏教の経典、墓地の石の卒塔婆、木の卒塔婆などに記されており、保存状態が悪く文字が不鮮明であるものが少なくない。そこで現在使われているような、文字の始点、終点、交点等の部分を比較して認識する従来の構造解析による文字認識では、正確さを期待することがきない。そこで本研究では、文字の部分ではなく全体で比較する相関度での文字認識を試みる。すなわち、梵字同士の相関曲線において、最大値および尖頭度を定義し、それによって梵字認識を行った。また認識対象として書物、石および木の卒塔婆を用いた。

2. 原理

2.1 相関度

本研究では、文字を認識するために相関度を用いる。相関度とは(1)式に示される自己相関関数の値および(2)式で示される相互相関関数の値であり、2つの信号がどれくらい似ているかを示す指標である。

$$\Phi(\tau) = \int_{-\infty}^{\infty} g(x)g(x+\tau)dx \quad (1)$$

$$\Phi(\tau) = \int_{-\infty}^{\infty} g(x)h(x+\tau)dx \quad (2)$$

上式は、連続信号の相関関数だが、本研究では離散信号を取り扱うため(1)、(2)式を(3)、(4)式に変形し離散信号の相関度を求める。

$$\Phi[m] = \frac{1}{N} \sum_{n=0}^{N-1} g[n]g[n+m], m = 0, 1, \dots, N-1 \quad (3)$$

$$\Phi[m] = \frac{1}{N} \sum_{n=0}^{N-1} g[n]h[n+m], m = 0, 1, \dots, N-1 \quad (4)$$

さらに本研究では信号を2次元の文字画像として取り扱うため、実際の相関度を求める式は(5)式に示す2次元の相互相関関数となる。

$$\Phi[a, b] = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} g[n, m]h[n-a, m-b], \quad (5)$$

$$a = 0, 1, \dots, M-1, b = 0, 1, \dots, N-1$$

すなわち、2次元の文字画像同士の相互相関関数値を求めてその文字の相関度とする。

2.2 文字認識の原理

次に、梵字認識の原理を述べる。

本研究では取り込んだ2次元の文字画像を256×192画素の配列として扱っているが、ここでは簡単のため2×2画素の配列で説明する。

2つの梵字画像データを図1に示すg(2, 2), h(2, 2)とする。

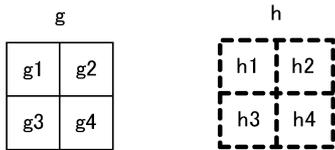


図1 取り込んだ画像の配列gとデータベース内の画像の配列h

配列gが取り込んできた文字画像、配列hがパソコンに格納されているデータベースの文字画像に相当する。配列g, hを図2のように端から順に列をずらしながら重ね合わせた画素同士の光強度をかけ合わせ、それを図3に示すような配列Φの値とする。これを最後の列すなわち図2のΦ(1, -1)の状態まで行う。次に行をずらし2行目で、同様の計算を図2のΦ(1, 1)の状態まで行う。

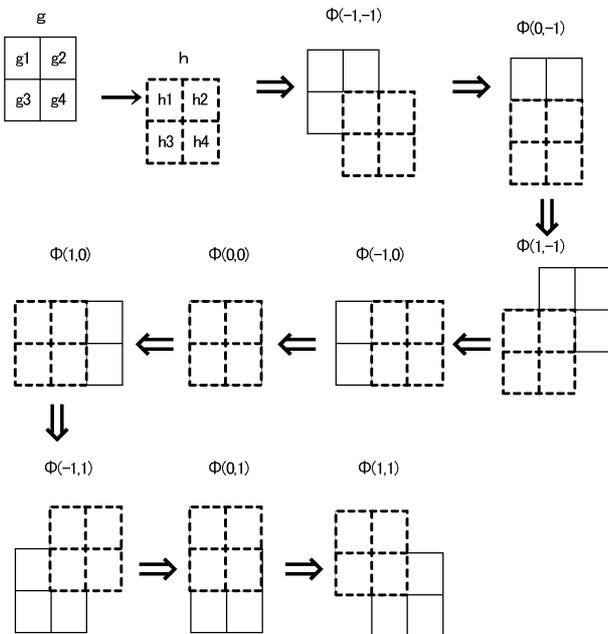


図2 相関関数計算の概念図

Φ(-1,-1)	Φ(0,-1)	Φ(1,-1)
Φ(-1,0)	Φ(0,0)	Φ(1,0)
Φ(-1,1)	Φ(0,1)	Φ(1,1)

※関数Φでは座標が左図のようになる。

図3 相関度の計算結果の配列Φ

結局、配列Φの計算は次式で表される。

$$\begin{aligned} \Phi(-1, -1) &= g_4 \times h_1 \\ \Phi(0, -1) &= g_3 \times h_1 + g_4 \times g_1 \times h_4 \\ &\vdots \\ &\vdots \\ &\vdots \\ \Phi(1, 1) &= g_1 \times h_4 \end{aligned} \tag{6}$$

例として記号の「○」と「◯」の自己相関関数を図4に「○」と「□」の相互相関関数を図5に示す。なお、図では、相関値の大きさをグレースケールで表しており、白くなるほど大きな値を示す。

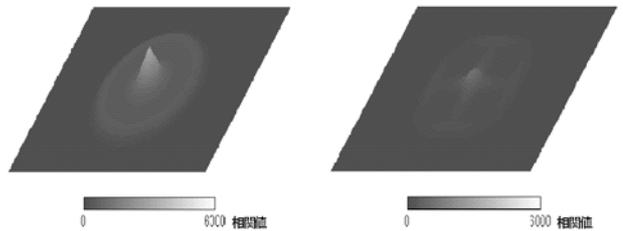


図4 「○」と「◯」の自己相関度 図5 「○」と「□」の相関度

図4, 5からわかるように、よく似ている図形の相関度はピークが鋭く高くなっているのに対し、似ていない図形の相関度はピークが低くなだらかなのがわかる。

本研究ではこのピークの高さと鋭さを文字認識のパラメータとして用いる。つまり、ピークが高いほど、そして鋭いほどよく似ているということがいえる。図6に示すように、このピークの高さを最大値、鋭さを尖頭度と定義する。なお、尖頭度を相関度曲線が1画素だけピークから離れたときの下り勾配と定義する。実際には、精度を上げるため、ピークを中心とした半径20の円の範囲内ですべての下り勾配の平均を尖頭度とした。尖頭度は式(7)で定義したものをを用いる。

$$S = \frac{1}{\pi r^2} \sum_{a^2+b^2 \leq r^2} \frac{p - \Phi(a, b)}{\sqrt{a^2 + b^2}} \tag{7}$$

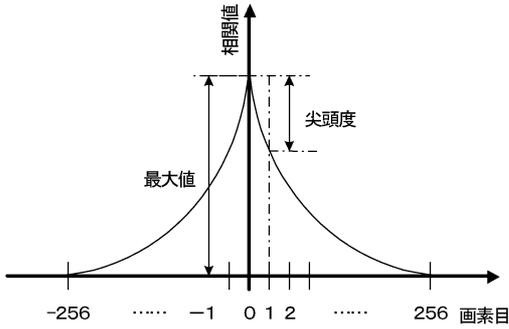


図6 最大値と尖頭度(相関度の断面図)

ここで、 s を尖頭度、 r を円の半径(今回は $r=20$)、 a 、 b を相関値の座標、 p を最大値とする。

また、相関度は、画像が相似でも、画像の大小により計算結果が異なる。例えば図7に示すように「○」と「○」では相似形ではあるが相関関数値が「○」と「○」の時と異なっている。このため文字の大きさをある程度統一する必要が生じてきた。

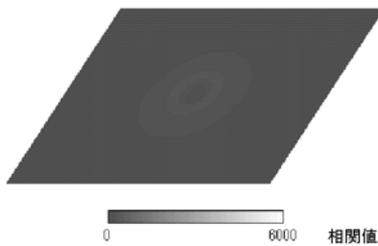


図7 ○と○の相関度

そこで、取り込んだ文字画像を常に同じ大きさ(256×192画素)にそろえることにした。

文字の大きさを統一した後の相関関数値の一例を図8, 9, 10に示す。図8は「○」と「○」の自己相関度、図9は「○」と「□」の相関度を示し、図10は「○」と「○」の相関度を示す。処理の流れを変えたことにより図10から、画像の大小に関係なく文字認識が行う事が可能になったことがわかる。

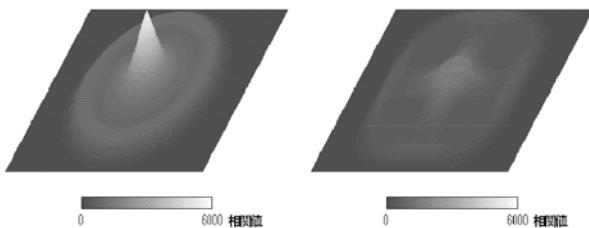


図8 「○」と「○」の相関度 図9 「○」と「□」の相関度

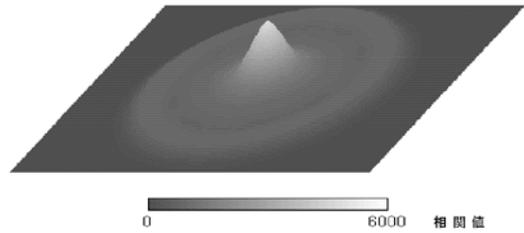


図10 「○」と「○」の相関度

3. 梵字認識実験

3.1 実験手順

次に梵字認識実験の手順を述べる

- (1) 梵字(51文字)のデータベースを作成する。¹⁾データベースの画像には、1~51の数字を順番に振り以後この数字を文字番号と呼ぶ。
 - (2) デジタルカメラ(640×480画素)で認識対象の梵字を撮影し、画像データをパソコンに取り込む。
 - (3) 文字を適当な大きさに切り出し画像を8ビットのグレースケールにし、256×192画素内になるようにサイズを調整する。
 - (4) 画像の明暗を反転させる。
 - (5) しきい値を設定し画像を二値化する。
 - (6) ノイズ、背景を除去する。
 - (7) 文字を256×192画素に拡大する。
 - (8) 取り込んだ画像とデータベース内のすべての文字との相関度を計算し、最大値と尖頭度が最大のものを選ぶ。
- 以上の手順で梵字認識実験を行うが、手順(4)以下はプログラムで自動化している。また、手順(5)におけるしきい値は、あらかじめ梵字画像データを用いて光強度のヒストグラムを制作しておき、それをもとにしきい値を決定している。

データベース内の梵字画像を用いて文字認識を行った結果、認識率は100%であった。

3.2 梵字認識実験例

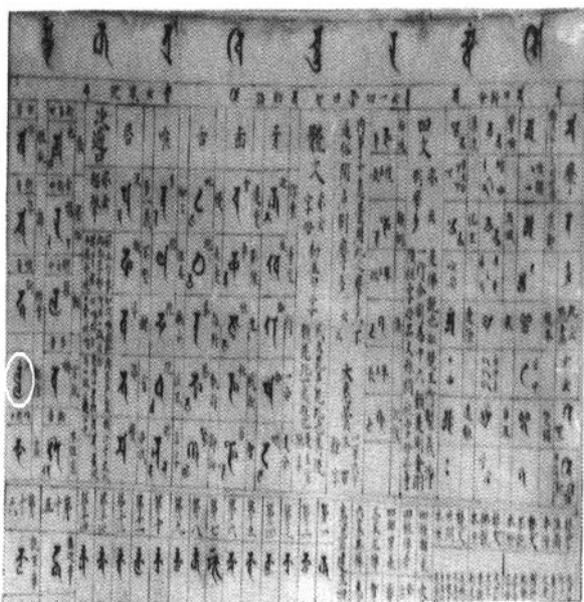
以下に梵字認識実験例を実験の流れに沿って示していく。

- (1) 一例としてデータベース内の文字番号50の梵字を図11に示す。図中、白が光強度255、黒が0である。



図11 文字番号50の文字の画像

(2) デジタルカメラで撮影した書物中の梵字画像を図12示す。



邦教筆「悉曇字母表」(部分、埼玉県熊谷市・常光院蔵)

図12 デジタルカメラでの画像

次に、図12内の白丸で示した梵字の認識を行う。この梵字はデータベース内文字番号50の梵字である。

(3) 適当な大きさに切り出した画像を図13に示す。

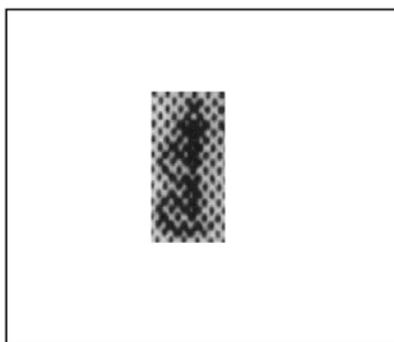


図13 切り出した画像

(4) 明暗を反転させた画像を図14に示す。

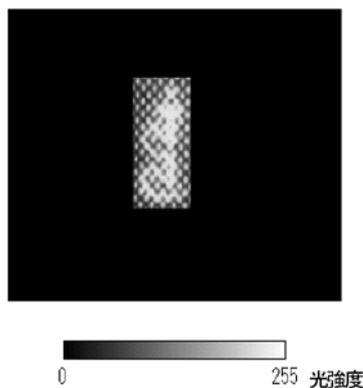


図14 明暗を反転させた画像

(5) しきい値を220に設定して二値化する。

(6) ノイズ、背景をしきい値220のフィルタで除去する。

(7) 画像を256×192画素のサイズに拡大する。これを図15に示す。

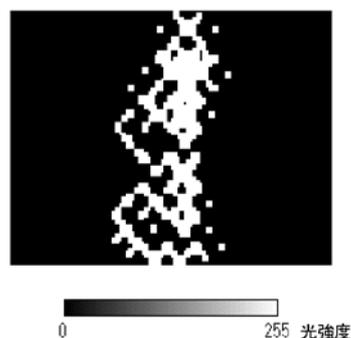


図15 拡大画像

(8) 図15とデータベース内の文字番号1～51との相関度を計算した結果を図16に示す。

なお、計算に要した時間は約5分間であった。図16の横軸は文字番号であり、縦軸は最大値および尖頭度を示す。また、尖頭度は見やすくするために50倍している。図よりデータベース内の文字番号50番の梵字(図中の↓)すなわち図11に示す梵字が、最大値1542、尖頭度26.2と共に一番大きいことがわかる。図11の梵字画像と認識対象梵字画像の相関結果を視覚化したものを図17に、また比較のため図19に示す文字番号51の梵字画像との相関結果を図18に示す。図より明らかに認識対象梵字は文字番号50の梵字であることがわかる。また、この図では認識時間短縮のために相関値をピーク周辺のみで計算した。

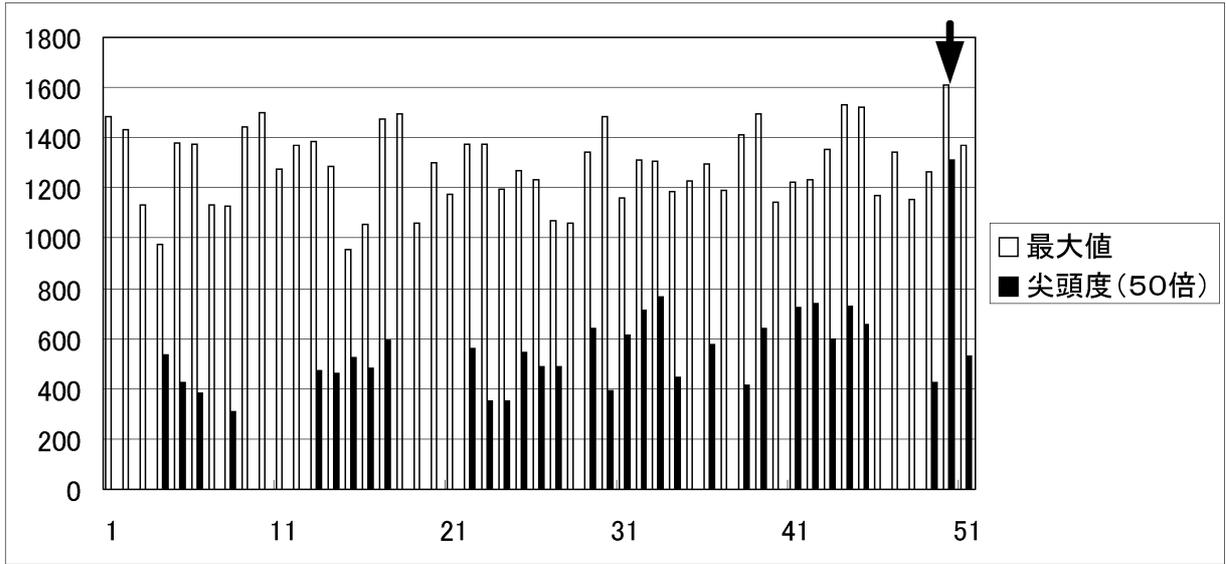


図16 図13に示す梵字画像とデータベース1～51の梵字画像との相関値

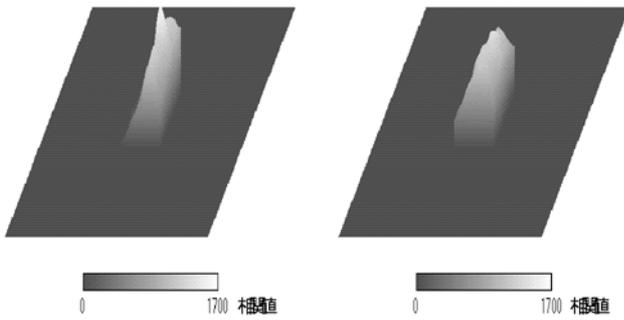


図17 図11と13の相関度 図18 図13と19の相関度



図19 文字番号51の梵字画像

塔婆などに書かれた様々な梵字を対象にして認識実験を行った。卒塔婆に書かれた梵字の一例を図20に示し、認識実験結果を表1に示す。なお、表1の認識率は次式で定義したものをを用いる。

$$\text{認識率}(\%) = \frac{\text{認識文字数}}{\text{合計文字数}} \times 100 \quad (8)$$



図20 木の卒塔婆に書かれた梵字画像(矢田寺)

また、認識の評価の仕方としては、最大値と尖頭度でそれぞれ順位付けし、順位の平均値の低い方の文字を認識対象の梵字であると決定している。

4. 文献および卒塔婆の梵字認識実験

前章の実験例に示したような認識実験を図12に示した文献中の文字や、その他の紙面、また、墓地の木の卒

表1 様々な梵字の認識実験結果

材 質	認 識 (文字)	認識不可 (文字)	認 識 率 (%)
紙 (合計=43文字)	26	17	60.5
紙 (合計=15文字)	10	5	66.7
合計=58文字	36	22	62.1

