

視覚の誘導場理論を用いた文字認識の提案

長石 道博

ATR視聴覚機構研究所

視覚研究室

619-02 京都府相楽郡精華町光台2丁目2番地

あらし

視覚の誘導場理論に基づき文字の周りの「場」を推定し、「場」の違いをエネルギーとして定量化し文字を識別する方法を提案する。「場」の効果に応用した文字認識等が幾つか提案されているが、これらは従来の文字認識手法に「場」の効果の考え方を加えたものである。しかし「場」の効果に応用して柔軟で人間の感覚に近い文字認識を行うには、元になる「場」の物理的な意味から文字認識方法を考え直さなければならない。それには、(1)「場」の物理的な意味とその推定方法、(2)「場」の物理的な意味に立脚した識別理論を明確にする必要がある。本稿では、まず今まで報告されている心理学や生理学的知見を基に「場」の物理的な意味を電磁気学理論として扱う推定方法と、「場」の違いを弾性エネルギーを指標として識別する方法について述べ、最後にパターンについて実験した結果についてふれる。
和文キーワード 視覚の誘導場、場の効果、電磁気学、文字認識、弾性エネルギー

A Proposal of Character Recognition using the Theory of Field of Induction on the Retina

Michihiro NAGAISHI

Visual Perception Department

ATR Auditory and Visual Perception Research Laboratories

2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02, Japan

Abstract

A recognition method is presented for evaluating differences in the field of induction of characters. This method is based on the theory of field of induction on the retina. To develop a character recognition method in applying a physiological phenomenon called the effect of field of induction, we must consider the method from the viewpoint of the physical meaning of the field. Therefore, it is necessary to clarify the following two items: (1) the physical meaning and calculating method of the field of induction; and (2) a recognition theory based on the physical meaning of the field. This paper first describes a calculating method that employs physiological and psychological knowledge and considers the field as electromagnetic. It then presents a recognition method that evaluates the differences in the field of induction of characters. Finally the experimental results are described for some patterns.

英文 key words field of induction on the retina, effect of field, electromagnetic, character recognition, elasticity energy

1. はじめに

文字等のパターンをどのように認識したらよいか工学的な方法について今だ十分には解かっていない。ところが我々は普段何気なく文字を変形させたり略して書き読んでいる。しかし書かれた文字が全くランダムに変形している訳でない。例えば図1のように、変動の大きな自由手書き文字「あ」を大きさの正規化を行なって複数重ねた場合¹⁾、文字の変形がランダムならば一様な分布になるはずである。ところが「あ」に限らず、どの文字でも複数の同じ文字を重ねた場合、必ずその文字が浮かび上がる。まるで各文字が、ある領域に書かれる規制があるかのように見える。このことから、文字についてこのような規制を行う何かしらの法則があるのではないかと考えられる。また、我々は種々の自然現象をパターンとして認識、理解出来る。そして物理的な法則が、これらの自然現象のパターン形成を支配していることが報告されている¹⁴⁾。文字自体は人工的なパターンであるが、手書き文字は手の運動等の過程によって生成されるので、物理的な法則が文字の書かれる領域の規制に関与していると考えられる。

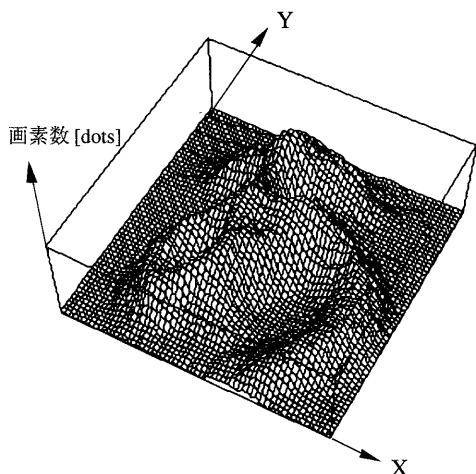


図1 複数の「あ」を重ねた画像の例

文字を複数重ねた際の文字画像の広がりや書道で文字の空間的な配置が重要なこと等から、文字が書かれる領域を規制するものとして文字の周りに醸し出される「場」の性質に着目した。このような図形等の「場」は、生理学、心理学現象として実際に報告されている。図2に横瀬によって報告されている図形の「場」の一例を示す³⁾。図2に示すように「場」は、太い線で描かれた線図形の周囲に、地図の等高線のように分布している。「場」の強さは、図形からの距離に依存し、図形に近い程大きな値をもつ。

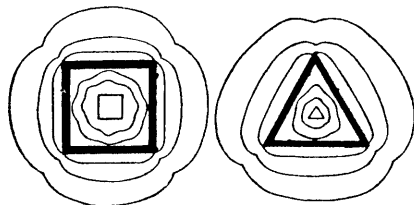


図2 図形の「場」(横瀬 1966)³⁾

認識を含めた人間の知覚理論を「場」の理論で説明する「視覚の誘導場」の多くの研究が行なわれている²⁾³⁾⁴⁾。そして人間の文字パターンの読み方の類似性と文字の「場」の類似性の間に密接な関連があることが示唆されている³⁾⁴⁾。そこで「視覚の誘導場」理論から文字の「場」を推定し、その物理的な性質に基づいた文字認識方法を考案できれば、人間の知覚理論や生理学的知見と対応のとれた、極めて人間の感覚に近い文字認識が可能になると考えられる。

「場」を利用して文字認識を実現するには、次の2つの問題点がある。まず「場」の物理的な意味とその具体的な推定(計算)方法を明確にしなければならない。次に「場」の物理的な意味に立脚した識別理論を構築する必要がある。

本研究は、最初に今まで報告されている心理学や生理学的知見を基に「場」を電磁気学的に扱う推定方法を提案し、「場」を弾性体と見なし「場」の違いを弾性エネルギーを指標として識別する方法について述べる。最後に提案した方法に基づき実際の類似パターンで実験し、文字認識の可能性を示した。

2. 視覚の誘導場の例と推定方法の提案

2.1 生理学的見知からの視覚の誘導場

生理学的な実験事実を元にした誘導場については藤井らのモデル等がある⁸⁾。このモデルでは知覚現象を電流作用が引き起こす電磁気理論にならって構成している。一方、本川らは、鯉の網膜上の電位分布を走査法により映像化することで錯視図形等の場を求めている²⁾。しかしどちらの方法も、複雑な図形の場の分布を求めるのは難しい。

2.2 横瀬の心理的ポテンシャル場理論

横瀬は光刺激閾法を用いた心理実験から、図形の場の強さ M は、図形からの距離 D 、明瞭度 H (図形と背景との明度対比)、固有結構 E (図形の形の因子) の関数であることを見出した³⁾⁴⁾。

$$M = f\left(E \frac{H^b}{D^d}\right) \quad (b, d \text{ は定数}) \quad (1)$$

横瀬は式(1)を元に、場の強さを与える関係式を提案している。図3に示す線分abが任意の点Pに作る場の強さは、線分abを一つの荷電体と見なし、線分上の単位面積が点Pに及ぼす影響を線分abについて積分したものとする。図3において点Pから線分abにおろした垂線の足をO、 $PO=D$ とおく。

$$M_p = \frac{1}{D} \int_{-\theta_2}^{\theta_1} \cos \theta d\theta = \frac{\sin \theta_1 + \sin \theta_2}{D} \quad (2)$$

横瀬は「場」が人間の視覚の事象であることから、式(2)で表される「場」を心理的ポテンシャル場と呼んだ。図4のように線分がab, bcと複数存在する場合は各々の線分について式(2)を計算して加算する。

図4の点Pの場の強さ M_p は、

$$M_p = \frac{\sin \theta_1 + \sin \theta_2}{D_1} + \frac{\sin \theta_3 + \sin \theta_4}{D_2} \quad (3)$$

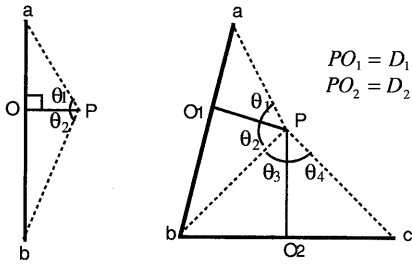


図3 線分

図4 重ね合わせ

横瀬理論では、場の強さは図形を構成する線分の微小線素の積分で、複数の線分がある場合は各線分による場の重ね合わせになっている。更に内山は面図形の場について横瀬の心理的ポテンシャル場が面図形の輪郭の作る場に等しく、またクーロンポテンシャル、即ち電磁気学的な「場」として理解することが適切なことを示した⁹⁾。図形の場について多くの心理実験が行われており、横瀬と同じ現象が報告されている¹⁰⁾。なお横瀬はMüller-Lyerの錯視図形について本川が求めた誘導場の分布²⁾が、式(2)で計算したポテンシャル場の分布と同じであることから、心理学的結果が生理学的事実と一致していることを示した⁴⁾。以上をまとめると電磁気学的な理論式が提案されている横瀬の理論を、視覚の誘導場理論の基礎理論として吟味する余地は十分あると考えられる。

2.3 視覚の誘導場の推定方法の提案

それではどのようにして任意の図形や文字の視覚

の誘導場を求めたらよいのだろうか。計算機で「場」を理論的に求めるには、視覚の誘導場理論を一般化してデジタル画像における場の推定方法を考える必要がある。ここでは横瀬の心理的ポテンシャル場理論を視覚の誘導場理論の基礎理論として一般化した「場」の推定方法を提案する。なお視覚の誘導場を今後は単にポテンシャル場と呼ぶことにする。

まず、2.2で紹介した横瀬理論は次のようなものと解釈し一般化することができる。図5のように複数の曲線 $f_1(s), f_2(s), \dots, f_n(s)$ がある。点Pと各曲線上の任意の点との距離を $R(f_i(s))$ とすると、点Pにおけるポテンシャル場の強さ M_p は、

$$M_p = \int \frac{ds}{R(f_1(s))} + \int \frac{ds}{R(f_2(s))} + \dots + \int \frac{ds}{R(f_n(s))} = \sum_i^n \int \frac{ds}{R(f_i(s))} \quad (4)$$

式(4)は、各曲線上に電荷1の点電荷が等分配されていると仮定して、ポテンシャル場の強さ M_p が、各曲線が点Pにつくるクーロンポテンシャルの重ね合わせであることを示している。

なお文献3及び4から、図6のように点Pから見て曲線 $f_1(s)$ に遮蔽されている曲線 $f_2(s)$ の一部と曲線 $f_3(s)$ の積分は行わない。つまり点Pから発した光が当たる部分のみの積分を行なう。式(4)の積分と和には、このような制約条件がつくので注意が必要である。

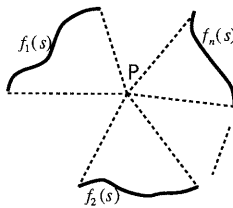


図5 複数の線分

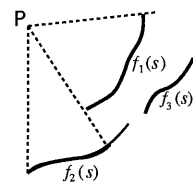


図6 遮蔽される場合

次に、線図形から面図形に拡張する。面図形を導体であると考え。面図形の内部電荷は0で電荷は輪郭線だけに分布するので、線図形と同じに扱える。まず、電荷1の点電荷が面図形の輪郭線上に等分配されていると仮定する。図7において、A, B(斜線部分)は面図形の一部である。A, Bにおいて面図形の外部にある任意の点Pから発した光が当たる輪郭線を各々 I_1, I_2 とする。各輪郭線 I_1, I_2 上の任意の点と点Pの距離を $R_1(s), R_2(s)$ とする。点Pにおけるポテンシャル場の強さ M_p は、

$$M_p = \int_{I_1} \frac{ds}{R_1(s)} + \int_{I_2} \frac{ds}{R_2(s)} \quad (5)$$

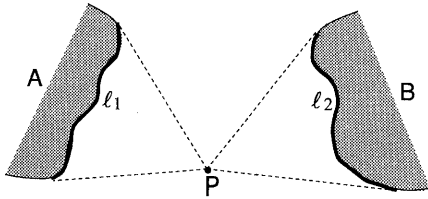


図7 面図形のポテンシャル場

式(5)からポテンシャル場は、電荷1の点電荷と仮定した輪郭線上の各点がつくるクーロンポテンシャルの重ね合わせであることがわかる。

デジタル画像の輪郭線は有限個の点で構成されていることから、デジタル画像のポテンシャル場は次のように拡張することが出来る。デジタル画像の画素は格子状に配列されており、画素間の距離を1とおく。図8のように点列の集まりである曲線 $f(s)$ によって任意の点 P にポテンシャル場が形成されるとする。曲線 $f(s)$ は線図形の線分や面図形の輪郭線にあたる。曲線を構成している各点は、電荷1の点電荷と仮定する。点 P から曲線 $f(s)$ まで走査して曲線 $f(s)$ を構成する n 個の点が見つかり、走査して見つかった各点までの距離を r_i とすると点 P におけるポテンシャル場の強さ M_p は、

$$M_p = \frac{1}{n} \sum_{i=1}^n \frac{1}{r_i} \quad (6)$$

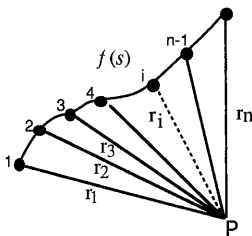


図8 点列のポテンシャル場

式(6)を用いることで、任意のデジタル画像のポテンシャル場を計算することが可能になる。式(6)で全体を走査した本数 n で割っているのは、デジタル画像では画素が有限なので規格化が必要だからである。なお、格子間距離は1なので、式(6)から M_p の範囲は、 $0 \leq M_p < 1$ である。連続画像の場合、式(5)より $r_i \rightarrow 0$ で $M_p \rightarrow +\infty$ なのでデジタル画像の図形のポテンシャル場の分布と異なる。分布の差が生ずるのはデジタル画像の線分近傍（1画素程度離れた部分）で、その大きさよりも計算誤差の方が遥に大きいので、デジタル画像と連続画像の分布の違いによる識別等への問題は無いと考えられる。

3、ポテンシャル場を用いた識別

3.1 「場」を応用した例

文字画像のつくる「場」を走査して、文字の基本成分である直線、右（左）巻曲線がつくる特徴的な「場」を検出することで文字の基本成分を抽出する研究⁵⁾⁶⁾がある。また誘導場の強さを指標にして、文字画像から抽出した複数の文字線分から適切な文字線分を決定する研究がある⁷⁾。これらの研究は、文字線分等の構成要素もっている「場」の効果に着目したものである。

3.2 ポテンシャル場と文字の類似性

それに対し本論文では、文字のポテンシャル場のもつ性質、特に人間の文字パターンの読み方とポテンシャル場の分布の類似性の間に見られる性質に着目した。文献3及び4では相異なるカテゴリーの等ポテンシャル面間で次のような性質が示唆されている。「ある2つの文字が同一カテゴリーの文字ならば、各々のポテンシャル場の分布が似ており、文字のトポロジー性を失わずに一方の文字の変形を行なって他方に一致させることができる。ポテンシャル場が全く似ていないものは、別カテゴリーの文字である。」

では、等ポテンシャル面であれば、それがもつ性質は、ポテンシャル値にかかわらず同じなのであるうか。ポテンシャル値が高い「場」は、文字成分に近いのでストローク等の文字の詳細な情報を、一方、文字の周辺部等のポテンシャル値が低い「場」は文字の概形の情報をもっていると考えられる。従って、これら識別に関する情報は、ポテンシャル値に伴って図9のように連続的に変化していると考えられる。

以上述べたようなポテンシャル場の性質を用いてポテンシャル場の分布の違いを定量化すれば、文字の識別を行うことが出来ると考えられる。

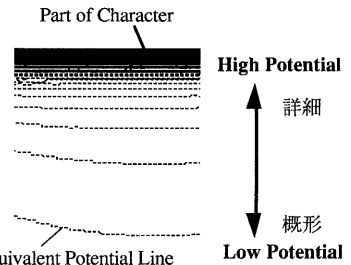


図9 ポテンシャル値の違いによる識別上の意味

3.3 弾性体としてのポテンシャル場

ポテンシャル場の分布状態の違いを考える場合、ポテンシャル場をゴムのような弾性体と考えると3.2の性質を説明するのに都合が良い。相異なる2つの

ポテンシャル場を弾性体と見なしトポロジー数学で行われるように、一方を他方に合わせようとゴムのよう自由に曲げたり伸ばしたりして変形させると弾性体に歪が生ずる。歪は弾性体の変形が大きい程大きいので、歪の量を調べれば両者の違いを定量的に知ることが出来る。図10のように未知パターンと2つの基準パターンA, Bのポテンシャル場が与えられている場合、各基準パターンに対し未知パターンを変位させ、弾性エネルギーの小さかった基準パターンが未知パターンに近いと識別することが出来る。

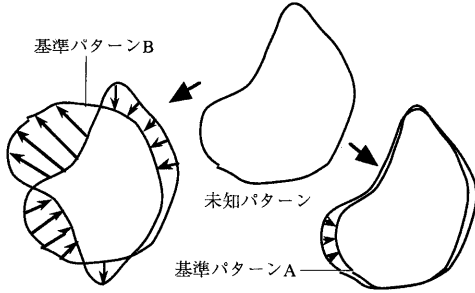


図10 ポテンシャル場を用いた識別

3.4 等ポテンシャル面の利用

ポテンシャル場は2次元上のスカラーポテンシャル場なので、3次元の弾性体として考えることが出来る。図形を弾性体と考えマッチングを行う際、有限要素法を用いて内部応力を求める方法がある¹¹⁾。しかし文字について求めたポテンシャル場は複雑な図形なので、図形分割の困難さや膨大な計算時間等問題が多い。そこで図11のように、3次元のポテンシャル場を2次元の等ポテンシャル面の集まりと考え、各等ポテンシャル面毎の歪を求めて集計する。等ポテンシャル面上では輪郭線のみを考えればよいので、結局1次元点列の問題として扱える¹²⁾。一方3次元対象に網を掛けて網の歪みを扱う方法も考案されている¹³⁾。しかし、3.2で述べたようにポテンシャル値の異なる等ポテンシャル面のもつ識別上の意味が異なっていると考えられるので、ここでは各等ポテンシャル面毎に独自性をもたせられるように、各等ポテンシャル面で歪みを求める形式を採用する。

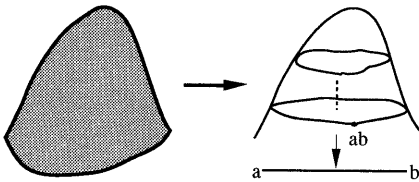


図11 3次元図形を1次元で扱う

3.5 弾性エネルギー

相異なる2つの図形がつくる等ポテンシャル面で、一方の輪郭線が他方の輪郭線に一致しようと変形した場合のエネルギーを求めることを考える。これは各輪郭線に張られた膜が、変形によって生ずる弾性エネルギーを求めることに等しい。膜が平衡状態からある位置まで変位するのに要する仕事 U は、膜の変形に要する仕事 U_1 と、膜の縁である輪郭線が変形するのに要する仕事 U_2 の和である。その他の外力等は一切考えないものとする。

まず U_1 を直交デカルト座標 (X_1, X_2, X_3) で考える。固体の変形で点 P が点 P' に移動した時、ベクトル PP' を変位ベクトル u_i と言う。変位ベクトル u_i は次のように表すことができる。

$$u_i = x'_i - x_i \quad (i = 1, 2, 3) \quad (7)$$

変形するとき u_i は連続的に変化するので、変位ベクトル u_i は座標 x_i の関数である。

次に微小直線の変位を考える。図12に示すように固体内に2点 $P(x_i), Q(x_i + \Delta x_i)$ をとる。変形後、点 P は P' に点 Q は Q' に移動したとし、各々の変位ベクトルを u_i, u_i^Q とする。 PP', QQ' 間の距離は微小として、各々 $\Delta l, \Delta l'$ とする。 u_i^Q は Δx_i が微小であるから、次のように表される。

$$u_i^Q = u_i + \frac{\partial u_i}{\partial x_j} \Delta x_j \quad (j = 1, 2, 3) \quad (8)$$

$$\text{変形前に、} \Delta l = \sqrt{\Delta x_i \Delta x_i} \quad (9)$$

であった距離が、変形後には、

$$\Delta l' = \sqrt{\Delta x_i' \Delta x_i'} \quad (10)$$

になる。ここで、 $\Delta x_i' = \Delta x_i - u_i^Q - u_i$ だから、式(8)より $\Delta x_i'$ は次のように表わせる。

$$\Delta x_i' = \Delta x_i + \frac{\partial u_i}{\partial x_j} \Delta x_j \quad (j = 1, 2, 3) \quad (11)$$

よって、式(10)は、

$$\Delta l' = \sqrt{(\Delta x_i)^2 + (\Delta u_i)^2} \quad (12)$$

ここで、 u_i の偏微分を単に Δu_i とおいた。

従って、2次元平面の微小面素 $\Delta X \Delta Y$ は、 X, Y 方向に Δu 変位した場合、式(12)より各々

$$\begin{aligned} & \sqrt{(\Delta X)^2 + (\Delta u)^2}, \sqrt{(\Delta Y)^2 + (\Delta u)^2} \\ & \text{変化するので、面積素増加は、} \\ & \sqrt{(\Delta X)^2 + (\Delta u)^2} \sqrt{(\Delta Y)^2 + (\Delta u)^2} - \Delta X \Delta Y \\ & = \frac{1}{2} \left\{ \left(\frac{\Delta u}{\Delta X} \right)^2 + \left(\frac{\Delta u}{\Delta Y} \right)^2 \right\} \Delta X \Delta Y \quad (13) \end{aligned}$$

なお式(13)で、 Δu の2階以上の偏微分は小さいので無視した。

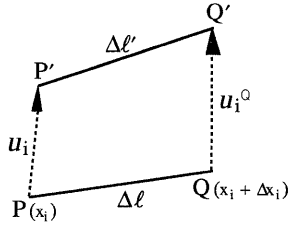


図12 微小直線の変位

よって膜全体の仕事 U_1 は、式(13)を積分したものになるので膜に一様にはたらく張力を τ とすれば、

$$U_1 = \alpha \iint (u_x^2 + u_y^2) dx dy \quad (\alpha = \frac{\tau}{2}) \quad (14)$$

輪郭線の変形に要する仕事 U_2 は、輪郭線 S 上の変位を $u(s)$ 、膜の弾性係数を $\kappa(s)$ すると、

$$U_2 = \frac{1}{2} \int \kappa(s) u(s)^2 ds \quad (15)$$

変位 $u(s)$ は、(膜の張力 τ 、密度 ρ 、 $a = \sqrt{\tau/\rho}$)

$$u(s) = a^2 u_{ss} \quad (16)$$

弾性係数 $\kappa(s)$ は一様で C とすると(15),(16)から、

$$U_2 = \frac{\beta}{2} \int u_{ss}^2 ds \quad (\beta = \frac{\tau^2}{\rho^2} C) \quad (17)$$

3.6 変位の計算と識別

図13のように等ポテンシャル面の境界線1が境界線2に移動する変位 u_i を考える。この変位は境界線1上の各点とその点の法線方向上にある境界線2上に移動する変位である。境界線1上の全ての点について求めた変位ベクトルから弾性エネルギーを計算するには、 U_1 、 U_2 の各偏微分を差分形式に直す必要がある。 $u_i = u(x(i), y(i))$ だから、

$$u_x = \partial u_i / \partial x = x(i) - x(i-1)$$

$$u_y = \partial u_i / \partial y = y(i) - y(i-1)$$

$$u_{xx} = \partial u_x / \partial x = x(i+1) - 2x(i) + x(i-1)$$

$$u_{yy} = \partial u_y / \partial y = y(i+1) - 2y(i) + y(i-1)$$

$$u_{ii}^2 = u_{xx}^2 + u_{yy}^2$$

輪郭線全体について変位 u_i を求め、これらの差分表現を用いて式(14),(17)から変形で生ずる弾性エネルギー $-U = U_1 + U_2$ を求めることができる。

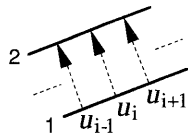


図13 等ポテンシャル面の変位の計算

次に辞書として複数のカテゴリーのポテンシャル場が用意されているとする。ポテンシャル値 P にお

ける等ポテンシャル面上で、ある辞書カテゴリー i と未知パターンが一致しようと変形して生ずる弾性エネルギーを $e_i(p)$ とおく。 $e_i(p)$ はポテンシャル P の関数である。3.2で述べたようにポテンシャル値によって識別に関する情報が連続的に変化すると考えられるので、識別は、ポテンシャル P の関数である、ある評価関数を用いて $e_i(p)$ を評価することで行なわれるのが望ましいと考えられる。またこのような評価関数を用いることで、文字の部分的変形や後処理等にも柔軟に対応することが可能になると考えられる。ここでは、単純にある文字が変形で要した総弾性エネルギーの大きさと識別する方法を示す。2つのパターンが一致するのに要する全エネルギー E_i は、

$$E_i = \int_0^1 e_i(p) dp \quad (18)$$

カテゴリー i が未知パターンに似ている程 E_i は小さくなるので、従って識別は全カテゴリー i に対して、

$$\min E_i \quad (19)$$

なカテゴリー i を決定することに他ならない。

4. 実験結果

4.1 ポテンシャル場の計算結果

文字「E」について式(6)からポテンシャル場を計算した例を図14(a)に、比較のため文献3から引用した例を図14(b)に示す。両者の等ポテンシャルの分布は酷似しており、式(6)による計算方法に問題はないものと思われる。収集した自由手書き文字¹⁾のポテンシャル場を、図15(a)はかすれている「こ」、図15(b)は普通に書かれている「こ」について計算した例を示した。(a)はかすれているものの、(b)と等ポテンシャル面の分布が良く似ている。このように多少品質が悪くても、同じカテゴリーであればポテンシャル場の分布が似ていることがわかる。なお、文字データは2値の64x64ドットの大きさに文字の重心を枠の中心に合わせて拡大縮小した。ポテンシャル場は文字領域を含む128x128ドットの範囲で計算を行なった。

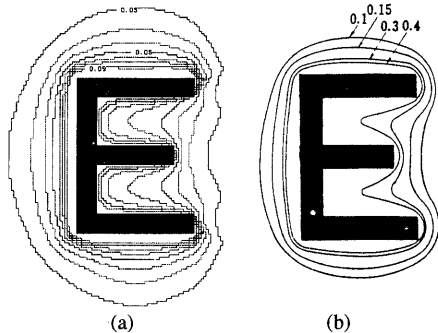


図14 「E」のポテンシャル場の比較

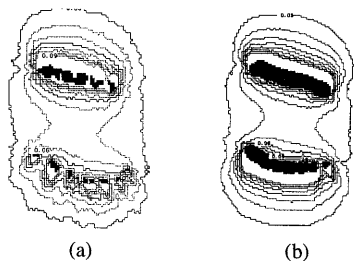


図15 手書き文字「こ」のポテンシャル場

4.2 実パターンによる弾性エネルギー計算

ポテンシャル場による識別能力を見る為に文字パターン「土」「土」を用いて弾性エネルギーを計算した。最初に各々の文字のポテンシャル場の計算結果を図16に示す。「土」「土」は機械認識では類似文字として認識の難しい対象とされている。しかし図16から、両者のポテンシャル場の分布が大きく異なっていることが解かる。次に「土」を基準として「土」に一致させようと変形した場合を考える。図17は、図16で示した各々のポテンシャル場の、ポテンシャル値0.06における輪郭線の「土」から「土」への変位の一例を示したものである。このようにして求めた変位に対し3.6の方法にもとづいて計算した弾性エネルギー $e_i(p)$ を、「土」→「土」、「土」→「土」へ変位させた場合について求めたグラフが図18である。縦軸が弾性エネルギー $e_i(p)$ 、横軸がポテンシャル値 P である。「土」→「土」への変位に要する弾性エネルギーが「土」→「土」よりも10倍程大きく、両者のポテンシャル場の分布の違いがエネルギー差として現われている。

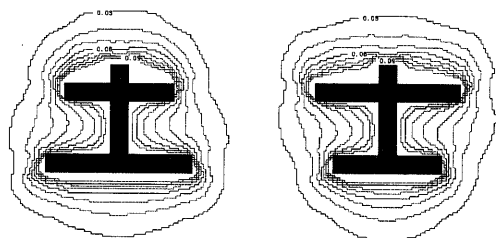


図16 「土」と「土」のポテンシャル場

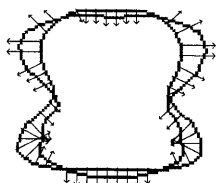


図17 「土」から「土」への変位

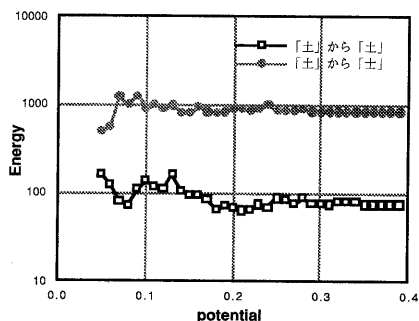


図18 変形に要する弾性エネルギー

式(18)から片仮名「ツ」「ソ」「フ」「ヲ」「ラ」「テ」を「ツ」に一致させようと変形させた場合の総弾性エネルギー E_i ($p=0.05 \sim 0.39$)を図19に示す。また各片仮名のポテンシャル場を図20に示す。我々は、「ツ」は「ソ」と間違え易く「ラ」「テ」とは区別が容易である。図19から弾性エネルギーの大きさも「ツ」「ソ」「フ」「ヲ」「ラ」「テ」の順に大きく、この順序は「ツ」と間違え易いものから容易に区別できる順序であることがわかる。

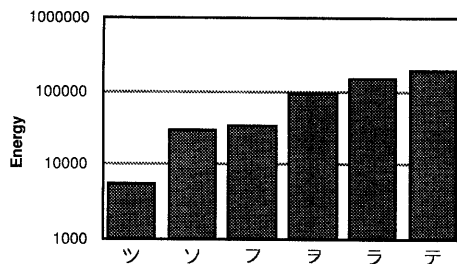


図19 片仮名間の弾性エネルギー

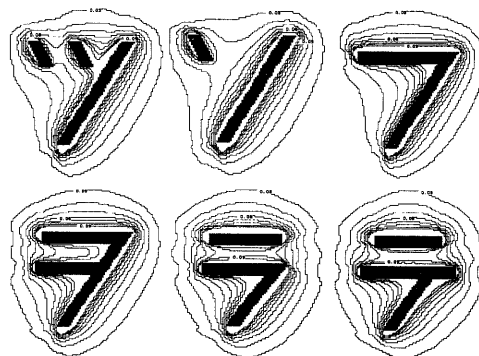


図20 片仮名のポテンシャル場

最後に「E」の線分の一部が変形して「F」に変化する過程⁴⁾を考える。図21に示すようにパターン1から5の順に変形していく線分を中心に、各パター

ンのポテンシャル場の分布が、少しづつ変化していくのがわかる。各パターンをパターン1に一致させようと変形させた際の弾性エネルギー $e_i(p)$ を図22に示す。図22からパターン4,3,2の順に弾性エネルギーが大きくなり、これはパターン1からパターン2,3,4の順に変形していく順に一致している。パターン1と大きく異なっているパターン5は、他のパターンよりも更に大きな弾性エネルギーをもっていることがわかる。また線分の変形による影響が、図22のパターン2から5のポテンシャル0.1以下の弾性エネルギーがパターン1に比べて大きくなっている部分に現われている。特に図22の山aは、パターン4の右下の点が周辺につくるポテンシャル場の影響が現われたものである。以上の実験結果から、ポテンシャル場の分布を見ることで文字の類似性や変形を容易に確認できることがわかる。また、文字の部分的な変形や類似性が弾性エネルギーの大きさに反映されていることから、弾性エネルギーを文字の変形や識別の指標として用いることが可能であると考えられる。

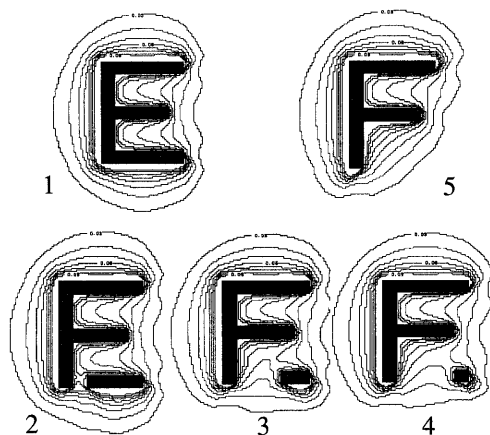


図 2 1 各パターンのポテンシャル場

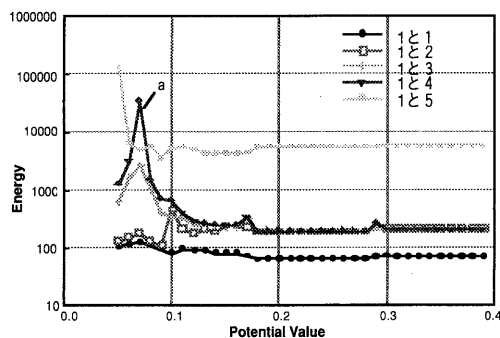


図 2 2 各パターン間の弾性エネルギー

5、まとめ

視覚の誘導場理論を用いた文字認識方法を提案した。そして誘導場を弾性体と見なし、文字の違いを一方の文字の「場」が、他方の「場」に一致しようとする変形によって生ずる弾性エネルギーで記述し、「場」の差異を評価することで「場」による文字認識の可能性を示した。今後、手書き文字等の識別実験を行ない実用上の問題点を検討する予定である。

謝辞

研究の機会を与えて下さったATR視聴覚機構研究所 淀川英司社長、視覚研究室上野圭一室長に感謝致します。そしてソフト作成等に御協力して頂いた三島薫さん、(株) SCC 向田茂氏に感謝します。

参考文献

- (1) 長石:"自由手書き文字の変動解析", 1991年テレビジョン学会年次大会, 10-10, pp.195-196 (1991).
- (2) 本川:"視覚の中枢機序 Pattern認識の問題を中心として", 生体の科学(東京医学書院), Vol.18, No.3, pp.27-38(1967).
- (3) 横瀬:"図形認識と機械的読み取り", 人間工学, Vol.2, No.3, pp.10-16 (1966).
- (4) 横瀬:"視知覚と場", 現代数学3(4), pp.16-22(1970).
- (5) H.Kazmierzak:"The potential field as an aid to character recognition", Proc. International Conference on Information Processing, UNESCO, Paris, pp.244-247(1959).
- (6) 森見徳, 森俊二, 山本和彦:"場の効果法による特徴抽出-閉じ状態の抽出-", 信学論, Vol.57-D, No.5, pp.812-819 (1974).
- (7) Shin-Liang CHANG, Yung-Sheng CHEN & Wen-Hsing HSU:"Advanced Line Continuation Model Design upon Selective Attention and Closure", Proc. The Third International Conference of Visual Search, U.K. (1992).
- (8) 藤井, 松岡, 森田:"Lateral Inhibition による錯視現象の解析", 医用電子と生体工学, Vol.5, No.2, pp.25-34 (1967).
- (9) 内山道明:"面図形上の場の強さについて", 名大文研論集XXX, 哲学, 10, pp.37-47 (1962).
- (10) 伊東:"視知覚における形の場の理論の実験的分析-横瀬のポテンシャル場の理論式の検討-", Jap.J.Psychol., 29, pp.11-19 (1958).
- (11) 志沢:"弾性体モデルを用いた図形のマッチング法", CV53-1, pp.1-8 (1988).
- (12) Michael Kass, Andrew Witkin & Demetri Terzopoulos:"Snakes: Active Contour Models", International Journal of Computer Vision, Vol.1, No.4, pp.321-331 (1988).
- (13) 積山, 坂上, 山本:"Active Net:動的な網のモデル", CV63-2, pp.1-16 (1989).
- (14) "パターン形成", 数理科学 (サイエンス社), No.350 (1992).