視覚の誘導場モデルによる手書き文字の文字品質評価と心理実験結果との比較

Comparison between Handwritten Character Quality by the Induction Field Model in Vision and Objective Evaluation by a Psychological Experiment

正会員 長石 道博†

Michihiro Nagaishi†

Abstract The induction field model can be used to explain Kato's experimental psychological results of objective evaluation of handwritten character quality. A lot of psychological experiments have been carried out to find the recognition process and the features of handwritten characters. However, there is no mathematical model that can explain the results of psychological experiments for handwritten characters. Recently, the focus has been on dynamics as the concept that explains the recognition process. Furthermore, the induction field model based on dynamics can explain some psychological experimental results for some artificial patterns, letters, and ambiguously shaped patterns. This paper shows that this model can explain psychological experimental results for handwritten character quality.

キーワード：手書き文字，視覚の誘導場，数理モデル，パターン認識，ダイナミクス，文字品質

1. ま え が き

人間は手書き文字のような複雑で多様なパターンを識別できる優れた能力がある。この人間の優れた文字認知過程を解明するには、文字パターンの認知過程の数理モデルは有効な手段である。これまで、数多くの数理モデルが検討されてきた。例えば、図 1(a) のような幾何学的な人工パターン^{1)~4)}、図 1(b) の我々が日常使用する活字のアルファベット^{5)~7)}や仮名⁸⁾⁹⁾などの文字パターンを対象にした数理モデル^{1)~4)10)}が検討されている。そして、図 1(c) のような変形のある人工パターンの数理モデルが検討されている^{8)11)~14)}。さらに、人工パターンではなく手書き文字の数理モデルも検討されている。例えば、海保らは漢字の構造、読み易さなど多方面から漢字に有効な特徴を検討している¹⁵⁾。しかし、手書き文字は変形が大きく複雑なので、まず実際に収集した手書き文字サンプルの変形を定量化した上で、検討した特徴が変形を説明できるか検証する必要がある。これまでいろいろな手書き文字変形の定量化手法が検討されている。例えば、萩田らは文字のストロークセグメントの平均位置、ずれの範囲、変化率など¹⁶⁾、岡本らは

図 1 心理実験に使われたパターン例
Examples of patterns for psychological experiment.

文字の外形（多角形近似）、外接矩形の重心、大きさ、面積、慣性軸など¹⁷⁾による定量化を試みている。また、但馬らは文字ストロークから抽出した方向線素を用いた定量化を提案している¹⁸⁾。これら工学的な手書き文字変形の定量化方法は、大量の手書き文字サンプルの変形を機械的に評価できる。しかし、定量化の方法によって同じパターンでも変形の評価結果が大きく異なる上、人間の視感的な直感と異なるという問題がある。

一方、工学的な方法ではなく、手書き文字の変化の違いを心理的な尺度で定量化できることが示唆されている¹⁹⁾。そして、実際に収集された手書き文字を被験者に見せて、変形の違いを手書き文字品質という心理学的な評価値として定量化する方法が試みられている。例えば、安藤らは、手書き文字認識の研究で最も使われる、電子技術総合研究所の大規模な手書き文字データベース ETL9²⁰⁾および、自由に筆記された「東」95 パターンについて、手書き文字品

1999 年 4 月 23 日受付，1999 年 6 月 7 日採録

†セイコーエプソン株式会社

(〒392-8502 諏訪市大和 3-3-5, TEL 0266-52-3131)

†SEIKO EPSON Corporation

(3-3-5, owa, Suwa-shi, Nagano 392-8502, Japan)

図2 文字サンプル例：東
Sample characters : "東".

質を6名に4段階(1:高品質, 2:普通, 3:低品質, 4:読めない)評定させた。そして、評定した文字品質を文字ストロークに着目した特徴(位置, 長さなどとそれらの変動)で説明できるか検討している²¹⁾。

加藤らも同様に手書き文字品質による手書き文字変形の定量化を行っている²²⁾²³⁾。加藤らはETL9から文字種ごとに無作為に選択した、図2の「東」の例のような25パターンを、ワークステーションのディスプレイに同時に提示した。被験者は提示された25パターンを1パターンずつマウスで選択して、非常にきれい1, きれい2, 普通3, 汚い4, 非常に汚い5の5段階に選別する。そして、10名の被験者の評価結果の平均から、文字種 $j(1, 2, \dots, 100)$ における25パターン内の任意のパターン $i(1, 2, \dots, 25)$ の文字品質の評価値 K_{ji} を求めた。図2の各パターンの下の数字は「東」の評価値 K_{ji} である。加藤らは100文字種, 2500パターンについて得られた文字品質について、被験者を任意の5名ずつ2群に分け、評定結果の各群の平均値より2群の相関を調べた。その結果、平均0.84の高い相関が得られ、評価の安定性が高いことを示した。また、加藤らは評定の再現性を確かめるために、各被験者は日を改めて主観評価を2回行っている。そして、2回評定の相関の各被験者の平均は0.75と高いことから、評定の再現性は高く、文字品質の主観評価実験で得られた10名の平均値 K_{ji} を各文字種の主観評価値とするのは妥当なことを示した²²⁾。

さらに、加藤らは2500パターンの文字品質 K_{ji} が、これまでの工学的な手書き文字認識などを参考に設定した15の特徴を組合せた重回帰の評価式で説明できるか検討した。その結果、10個の特徴(ストロークの分布, 黒画素の分布, 線の太さ, 縦横比, バランスなど)が、比較的文字品質評価に有効であることを示した。しかし、100文字種の文字品質を複数の特徴を組合せた1つの評価式で評価することは不可能であった²²⁾²³⁾。このように、文字品質によって手書き文字の変形は定量化できることが示されている。しかし、手書き文字の文字品質を説明できるような特徴や数理モデルは得られていない。

図3 デジタル画像における視覚の誘導場
Induction field on digital image.

一方、パターン認知は文字パターン共通の特徴ではなく、ダイナミックな過程、ダイナミクスであるという指摘が、従来の視知覚研究に関する知見^{24)~26)}、パターン相互の変換可能性の程度によるパターンの類似性の説明²⁷⁾、パターン変換作用による視覚恒常性の説明²⁸⁾などからなされている。

筆者は、ダイナミクスによる具体的なパターン認知モデルとして、図形の周囲に波及する場を仮定して視覚現象を説明する「視覚の誘導場」¹⁴⁾による認知モデルを提案した²⁹⁾。そして、この誘導場モデルがいろいろな文字パターンの類似性を説明できることを検証してきた。まず、この誘導場モデルは、文字線分などパターンに依存した特徴は一切仮定していないが、全くアプローチが異なる3つの数理モデル^{2)~4)}で説明される、図1(a)などの人工的なパターンをすべて説明できることを検証した²⁹⁾。続いて、図1(b)の活字のアルファベットや仮名など我々が日常使用する多様な文字パターンの類似性も説明できることを検証した³⁰⁾。さらに、図1(c)の人工的な変形パターンの類似性も説明できることを検証した³¹⁾。また、ダイナミクスが手書き文字の認知機構に存在する可能性を、電磁気力を想定した力学的モデルによる手書き文字の識別³²⁾、発振ニューラルネットモデルによる手書き文字変形の再現³³⁾などが工学的に示唆している。以上から、ダイナミクスによるパターン認知モデルである誘導場モデルは、変形パターンの変形の主観評価結果が人工的なパターンだけではなく、実際に我々が日常書いている手書き文字についても、説明できる可能性が高いと考えられる。そこで、本論文は視覚の誘導場モデルを手書き文字に適応し、手書き文字変形を定量化した手書き文字品質に関する心理実験結果を説明できるかどうか検証する。

2. 視覚の誘導場によるパターン認識

デジタル画像における誘導場の分布は、正電荷1の点電荷と仮定したパターンの外郭を構成する各画素がつくるクーロンポテンシャルの集積から計算する²⁹⁾。図3(a)のように n 個の点列から構成される曲線 $f(s)$ によって点Pに視覚の誘導場が形成されるとする。点Pから曲線 $f(s)$ 上の点 i までの距離を r_i とおくと、点Pにおける誘導場の

図4 認識過程
Recognition process.

強さ M_p を次のように定義する.

$$M_p = \frac{1}{n} \sum_i^n \frac{1}{r_i} \quad (1)$$

図3(b)は(1)式による「あ」の誘導場の例である. 地図の等高線状に分布しているのが誘導場の等ポテンシャル線で, 中央から外に行くほど場の強さは弱くなりゼロに近づく.

次に誘導場による認識アルゴリズムを示す²⁹⁾. 視覚の誘導場は, 基本的にクーロンポテンシャルと解釈できるので, 誘導場の変化を薄膜の弾性体のアナログで考える. ある誘導場の等ポテンシャル線の分布を他方の分布に一致するように変位させると, 弾性体と見なした誘導場に歪みが生ずる. 歪みは弾性エネルギーで一意に表現できるので, 誘導場の変換量を弾性エネルギーで評価する. 例えば図4のように, 辞書パターン*i*「土」と未知パターン*j*「土」の誘導場の重心を一致させ, ポテンシャル値*p*の等ポテンシャル面毎に「土」の誘導場の等ポテンシャル線の輪郭線上のすべての点が, その点の法線方向(輪郭線上の矢印)上にある「土」の等ポテンシャル線の輪郭線に変位したと仮定, 等ポテンシャル面ごとに生ずる弾性エネルギー $e_i(p)$ を計算する*. そして, 全等ポテンシャル面について $e_i(p)$ を加算し総弾性エネルギー E_{ij} を求める. 未知パターン*j*が辞書*i*の誘導場との一致に要する総弾性エネルギー E_{ij} が最小な辞書*i*が認識結果となる.

3. 検証の方法

誘導場モデルが, 手書き文字の変形を定量化した手書き

* 弾性エネルギーの定義と変位計算の詳細は文献29)を参照.

一 二 刀 人 夕 久 万 小 上 川 三 口 山 切 心
収 火 犬 牛 王 日 以 皮 必 左 可 史 世 生 出
玉 古 田 目 次 老 考 成 気 多 回 兩 冷 応 初
我 希 近 園 車 里 述 使 所 非 林 門 東 画 芽
派 級 計 品 首 春 胃 党 夏 高 教 魚 軋 問 然
衆 極 減 專 順 無 齒 葉 買 童 森 教 應 際 疑
歌 閑 潔 質 憲 興 嚴 謝 臨 競

図5 評定100文字種
100 Kanji character categories using in the subjective evaluation.

文字の文字品質に関する心理学実験を説明できることを検証するには, 手書き文字品質に関する心理実験結果と, 実験で使われたパターンの違いについて, 誘導場モデルの弾性エネルギーで評価した結果を比較する必要がある. 手書き文字品質の心理学実験として安藤らの実験があるが, 加藤らの実験のように得られた主観評価の妥当性が検証されていない²¹⁾. 本論文では, 100 カテゴリ, 2500 パターンという非常に多数のパターンについて実験結果が得られている加藤らの実験²²⁾を検証対象に用いた. 加藤らの実験で使われた文字パターンは, ETL9 から抜粋されたもので, 図2の例のように文献23)に記載されている. これらを, 64×64 ドットの大きさ(72DPI)の外接矩形に収まるようにイメージスキャナ(EPSON GT-5500)の解像度を設定(200DPI)して読取り文字画像を作成した. 図5は, 加藤らの実験で評定に使われた100個の文字種について, 各文字種ごとに最も文字品質が良いと評価された文字サンプル例を示した. 図5の左上から下に順に文字種番号 $j(1, 2, \dots, 100)$ をつける. 例えば, 「一」は $j = 1$, 「二」は $j = 2$, 「競」は $j = 100$ である. 作成された各パターンについて, 64×64 ドットの大きさのパターンの領域を含む 128×128 ドットの範囲で誘導場を計算する. まず, ある文字種 j の25パターンにおいて, 基準パターン $i = 1$ を図5に示した各文字種ごとに最もきれいと言われた文字とする. 例えば, 図2の「東」の場合, 左上の評定値1.5の「東」を基準パターンとする. 次に, 基準パターン $i = 1$ と任意のパターン $i(1, 2, \dots, 25)$ の誘導場について, 等ポテンシャル線がデータの枠に接触しないポテンシャル値0.03から文字線付近の0.39まで0.01ずつの各等ポテンシャル面で, 弾性エネルギーの総和 E_{1i} をパターン1, i 間の弾性エネルギー KjE_{1i} とする. そして, 誘導場モデルの弾性エネルギー KjE_{1i} を, 「場」の概念による文字認識で最も使われる「場」の強さをEuclid距離で評価する方法³⁴⁾と, 手書き文字認識研究のリファレンスによく使われるLDCD特徴³⁵⁾をEuclid距離で評価する方法と比較検討した. ある文字種 j における, パターン1, i 間の誘導場間のEuclid距離 $KjFD_{1i}$ は, パターン i の座標 x, y ($1 \leq x, y \leq 128$) に

おける場の強さを $M_i(x, y)$ (ただし, 画素部分=1) とすると

$$KjFD_{1i} = \sqrt{\sum_{x,y} (M_1(x, y) - M_i(x, y))^2} \quad (2)$$

となる. LDCD 特徴は, 64×64 ドットのパターン i を, 8×8 メッシュに分割, n 番目 ($1 \leq n \leq 64$) のメッシュについて, 水平 $hor(n)$, 垂直 $ver(n)$, $+45^\circ$ $p(n)$, -45° $m(n)$ の 4 方向の方向寄与度³⁵⁾を求め, 最終的に 256 次元の LDCD 特徴 $L_i(n) = L_i(hor(n), ver(n), p(n), m(n))$ を得る. ある文字種 j におけるパターン $1, i$ 間の LDCD 特徴の Euclid 距離 $KjLD_{1i}$ は

$$KjLD_{1i} = \sqrt{\sum_n (L_1(n) - L_i(n))^2} \quad (3)$$

となる. (2), (3) 式から, Euclid 距離はパターンの類似性が高いほどゼロに近づく.

なお, 文字画像は文献 23) に記載されているままの状態 で 3 つの方法の評価を行った. 文字認識で行われる大きさの正規化³⁶⁾は行っていない. 手書き文字は文字の大きさに違いがあるが, 加藤らの実験では被験者に正規化など一切の処理をしていない状態の ETL9 が提示, 評価されている²³⁾. したがって, 人間による文字品質の主観評価と 3 つの方法の評価能力を比較する場合, 正規化を行わずに文字画像を評価するのが望ましいと考えられる.

以上のように求めた誘導場の弾性エネルギー・Euclid 距離, LDCD 特徴の Euclid 距離と, 心理実験結果の相関を求める. これら 3 つの方法では, 最もきれいな文字を基準 ($i = 1$) に文字品質を評価している. そこで, 加藤らの心理実験結果も最もきれいな文字の評定値 Kj_1 で正規化した評定値 $Knj_i = Kj_i/Kj_1$ と 3 つの方法の評価値と相関をとった. 例えば, ある文字種 j のパターン i における加藤らの実験結果 Knj_i と誘導場の弾性エネルギー KjE_{1i} の相関係数を $R_j(Knj, KjE)$ とする. 相関係数は正規分布しないので, 文字種 j ごとに $R_j(Knj, KjE)$ を正規変換量 z に変換した上で, 加重平均して 100 文字種における加藤らの実験結果と誘導場間の弾性エネルギーの相関係数の平均を求める. 同様に, ある文字種 j のパターン i における Knj_i と誘導場間の Euclid 距離 $KjFD_{1i}$, Knj_i と LDCD 特徴の Euclid 距離 $KjLD_{1i}$ の相関係数 $R_j(Knj, KjFD)$, $R_j(Knj, KjLD)$ から, 加藤らの実験結果と誘導場および LDCD 特徴の Euclid 距離の相関係数を求める. 相関が強い評定方法ほど心理実験による文字品質がよく説明できる. なお, 誘導場の弾性エネルギーは, 指数関数的に値が変化するため, このままでは相関関係が直線的にならない. そこで, 常用対数をとるという変数変換により直線関係にした上で相関係数を求めた.

4. 検証結果

最初に加藤らの心理学実験結果と, 3 つの評価方法 (誘

図 6 心理実験と 3 つの評価方法との比較 (画数の少ない漢字)
Comparison between psychological experiment and other methods (Kanji with some strokes).

導場の弾性エネルギー・Euclid 距離, LDCD 特徴の Euclid 距離) の相関関係を示す. いろいろな文字種の心理実験結果と 3 つの評価方法との相関の様子を示すため, 画数が少なく簡単な漢字から画数が多く複雑な漢字「小, 古, 史, 我, 胃, 興」を代表例として選んだ. 図 6(a),(b),(c) は「小」($j = 8$) について, 横軸に加藤らの実験結果 $K8_i$, 縦軸に (a) は誘導場の弾性エネルギー $K8E_{1i}$, (b) は誘導場の Euclid 距離 $K8FD_{1i}$, (c) は LDCD 特徴の Euclid 距離 $K8LD_{1i}$ をとった散布図の例である. 図中の直線は回帰直線である. 同様に, 図 6(d),(e),(f) は「古」($j = 32$) について, 横軸に加藤らの実験結果 $K32_i$, 縦軸に (d) は誘導場の弾性エネルギー $K32E_{1i}$, (e) は誘導場の Euclid 距離 $K32FD_{1i}$, (f) は LDCD 特徴の Euclid 距離 $K32LD_{1i}$ をとった散布図の例である. 図 6 と同様に, 図 7(a),(b),(c) は「史」($j = 27$), 図 7(d),(e),(f) は「我」($j = 46$), 図 8(a),(b),(c) は「胃」($j = 67$), 図 8(d),(e),(f) は「興」($j = 96$) について比較した散布図の例である. 図 6~8 の例を含め, 他の心理実験結果も同様に誘導場の弾性エネルギーが誘導場と LDCD 特徴で Euclid 距離を使った場合よりも, 回帰直線に分布が集中しており, 相関が強くなっている.

図 7 心理実験と3つの評価方法との比較
Comparison between psychological experiment and other methods.

文字種ごとの相関係数の例と、100文字種の相関係数の平均を表1に示した。括弧内の数字は説明率(%)である。表1の例を含め他の文字種も同様に、誘導場とLDCD特徴でEuclid距離を使った場合の文字種ごとの説明率は説明率15から50%未満の中程度の相関が多く、説明率50%以上の強い相関は少ない。説明率15%未満の弱い相関もある。一方、誘導場の弾性エネルギーは強い相関が多い。そして、誘導場の弾性エネルギーの説明率は他の2つの評価方法よりも、数%程度から最大44%、100文字種の平均で8%程度大きい。このように、誘導場の弾性エネルギーの説明率は、他の2つの評価方法よりも大きく、心理実験結果との相関が最も強いと考えられる。

誘導場の弾性エネルギーが、他の2つの評価方法と比較して相関係数が大きいかどうか確かめるため、誘導場の弾性エネルギーと誘導場のEuclid距離、誘導場の弾性エネルギーとLDCD特徴のEuclid距離を、100文字種の相関係数の平均に関してt検定した結果を表2に示す。表2の◎は0.1%の危険率で有意差がある場合を示す。数字は標準化した偏差である。表2から、加藤らの実験に関するLDCD特徴のEuclid距離および誘導場のEuclid距離について、

図 8 心理実験と3つの評価方法との比較 (画数の多い漢字)
Comparison between psychological experiment and other methods (Kanji with many strokes).

危険率0.1%で有意なことから、誘導場の弾性エネルギーは、他の2つの評価方法より相関が強いことが示された。したがって、誘導場の弾性エネルギーは、手書き文字に関する加藤らの心理実験結果を最もよく説明できると考えられる。

5. 考 察

5.1 誘導場モデルの検証

検証結果より、誘導場の認識モデルは加藤らの心理実験結果との相関が、ここで検討したモデルのなかでは最も強いことが明らかになった。すなわち、手書き文字の変形を定量化した文字品質が合理的に説明できるモデルとして着目した誘導場の認識モデルは、従来の認識モデルよりも心理実験結果を正しく説明できる(文字種ごとの相関係数の説明率が数%程度から最大44%、100文字種の平均で8%程度以上)ことが示された。

今回検証に用いた図2や図5のような手書き文字は、文字種が多く変形も多様だが、誘導場の認識モデルはこれら手書き文字の変形を定量化した手書き文字品質が説明可能である。このように、誘導場の認識モデルは、人工的なパターンや我々が日常使う文字パターンなど多くの文字種に

表 1 相関係数の例
Example of correlation coefficient.

Kanji	field (energy)	field (Euclid)	LDCC
小	0.793 (62.9)	0.551 (30.4)	0.515 (26.5)
口	0.406 (16.5)	0.226 (5.10)	0.368 (13.5)
心	0.741 (54.9)	0.530 (28.1)	0.579 (33.5)
火	0.802 (64.3)	0.586 (34.3)	0.579 (33.5)
牛	0.788 (62.1)	0.560 (31.4)	0.634 (40.2)
必	0.333 (11.1)	0.206 (4.24)	0.260 (6.76)
史	0.989 (97.8)	0.656 (43.0)	0.681 (46.4)
世	0.670 (44.9)	0.394 (15.5)	0.310 (9.61)
古	0.707 (50.0)	0.508 (25.8)	0.439 (19.3)
次	0.875 (76.6)	0.628 (39.4)	0.675 (45.6)
考	0.631 (39.8)	0.337 (11.4)	0.445 (19.8)
成	0.933 (92.6)	0.570 (32.5)	0.727 (52.9)
冷	0.879 (77.3)	0.652 (42.5)	0.805 (64.8)
我	0.634 (40.2)	0.312 (9.73)	0.276 (7.62)
所	0.717 (51.4)	0.560 (31.4)	0.714 (51.0)
芽	0.685 (46.9)	0.620 (38.4)	0.402 (16.2)
派	0.748 (56.0)	0.596 (35.5)	0.656 (43.0)
胃	0.971 (94.3)	0.804 (64.6)	0.557 (31.0)
極	0.596 (35.5)	0.530 (28.1)	0.465 (21.6)
尊	0.623 (38.8)	0.449 (20.2)	0.611 (37.3)
無	0.509 (25.9)	0.502 (25.2)	0.466 (21.7)
買	0.628 (39.4)	0.420 (17.6)	0.321 (10.3)
感	0.589 (34.7)	0.353 (12.5)	0.444 (19.7)
疑	0.537 (28.8)	0.482 (23.2)	0.499 (24.9)
憲	0.552 (28.8)	0.483 (23.3)	0.487 (23.7)
興	0.824 (67.9)	0.698 (48.7)	0.719 (51.7)
100 Kanji average	0.569 (32.4)	0.486 (23.6)	0.497 (24.7)

表 2 相関係数の検定結果
T-test result of correlation coefficient.

pattern	field (Euclid)	LDCC (Euclid)
100 Kanji	◎ (3.47)	◎ (3.99)

ついて人間が判断するパターン類似性を説明することが可能であるだけでなく、さらに変形パターンの変形の主観評価結果を、人工的なパターンに加え手書き文字についても説明可能である。以上から、誘導場の認識モデルは人間のパターン認知過程を説明する数理モデルのひとつとして有効と考えられる。

5.2 誘導場による認識モデルの応用

人間は文中に誤字があっても、文字の類似性と文脈を利用して正しく読めることが報告されている³⁷⁾。そこで、認識が難しい手書き文字などに対し、識別に使う Euclid 距離などを確信度として認識結果を補正する知識処理が通常行われている。しかし、一般に確信度の信頼性が充分ではない²⁷⁾²⁹⁾ため、通常人間が文脈と類似文字の関連性から正解を予測する際の手順や知識を、認識結果の補正に応用しても好結果が得られにくい。そこで、例えば認識結果を統計的に修正する方法³⁸⁾が検討されているが、カテゴリの違いや文字変形によって確信度の傾向が異なるため、熟語などの照合に悪影響を及ぼす場合があり、認識結果の補正は容易ではない。そのため、信頼性の低い文字の類似性、確信度を使わない、純粋な統計的な知識処理が提案される³⁹⁾など、他の方法も最近検討されている。一方、誘導場の認識モデルが評価するパターン類似性は、心理実験結果に

図 9 文字サンプル例：次
Sample characters: "次".

表 3 正規化の効果
Effect of normalization.

field (energy)	field (Euclid)	LDCC
+0.0191	+0.0378	+0.0254

近い^{29)~31)}。さらに、今回の実験から、実際の手書き文字に対しても誘導場の認識モデルの有効性が示された。したがって、誘導場を用いれば、手書き文字のような変形を伴うパターンについて、人間が行っている文字の類似性と文脈を利用する方法が充分機能し、文脈と類似文字の関連性から正解が、より正確に予測できるようになり、認識率の向上が期待できると考えられる。

5.3 文字の大きさの違い

今回の検証では文字の大きさは正規化してない。しかし、検証に用いた文字画像の中には図 9 の 2, 17, 22, 24 番目の「次」の例のように、同じカテゴリ内に大きさがかなり違う文字が含まれている場合がある。図 9 の各文字の下に数字は文字品質評定値の平均である。図 9 の例を含め他の文字種についても、小さい文字の文字品質の評定が必ずしも低くなることはない²³⁾。このことから、人間は文字の大きさに関係なく文字品質を評定していると考えられる。一方、検証した 3 つの方法は文字の大きさの違いは特に考慮されていないので、人間のように文字の大きさに影響されずに文字品質が評価できるかどうか不明である。そこで、図 9 の 2, 17, 22, 24 番目の「次」のような、他の文字に比べて大きさが異なる文字を除いた場合について、加藤らの心理実験結果と 3 つの評価方法との相関を調べた。小さい文字かどうかは、目視にて同じカテゴリ内で他のパターンよりも文字の大きさが 2/3 程度以下と認められるものを対象とした。表 3 は、3 つの評価方法について、大きさが異なる文字を除いた場合の 100 文字種の相関係数の平均を、除かなかった場合の相関の平均と比較した結果である。表 3 から、どの評価方法も大きさが異なる文字を除くと相関係数が向上することがわかる。このことから、3 つの評価方法による手書き文字品質評価結果は、文字の大きさの違いに影響されると考えられる。したがって、文字の大きさの違いを吸収するような機構を 3 つの各評価方法共に考慮

する必要がある。

6. む す び

手書き文字変形の主観評価結果を説明できる数理モデルとして視覚の誘導場によるモデルに着目、加藤らの手書き文字の文字品質という主観評価で定量化した手書き文字の変形を、これまでの認識モデルより心理実験結果を正しく説明できることが示された。この結果から、誘導場のモデルは、人工パターンや一般の文字、そして、変形のある人工パターンおよび手書き文字において、従来提案されたパターン認知の数理モデルよりも、さらに正確に人間の認識特性に近い記述として有効であると考えられる。このモデルは人間が判断するパターン類似性に近い類似性が、人工パターン、文字、変形パターンを含めた多くの文字パターンについて得られている。さらに今回、手書き文字でもモデルの有効性が示されたので、人間が文字の類似性と文脈を利用して誤字を正しく読む方法を知識処理に取り入れて、手書き文字などの認識率を向上させる応用が考えられる。今後、手書き文字を含めた広範囲な文字パターンについて、さらに人間の認識特性に近い評価ができる誘導場モデルの検討を行う予定である。

手書き文字品質評価の心理学実験結果について御助言頂いた浜松ホトニクス中央研究所 加藤隆仁氏、本稿をまとめるにあたり御助言頂いたNHK 大阪放送局技術部 山田光穂氏に感謝致します。

〔文 献〕

- 1) J.T.Townsend, F.G. Asbby: "Experimental Test of Contemporary Mathematical Models of Visual Letter Recognition", *Journal of Experimental Psychology: Human Perception and Performance*, **8**, pp.834-864 (1982)
- 2) S.E.Palmer: "Structural Aspects of Visual Similarity", *Memory & Cognition*, **6**, 2, pp.91-97 (1978)
- 3) S.J.Lupker: "On the Nature of Perceptual Information during Letter Perception", *Perception & Psychophysics*, **5**, 4, pp.303-312 (1979)
- 4) 横澤一彦, 梅田三千雄, 淀川英司: "人間の漢字パターン認識特性 — 線図形知覚特性に基づく認識モデルの提案とその有効性の検討 —", *信学論*, **J69-A**, 5, pp.600-611 (1986)
- 5) J. T. Townsend: "The Theoretical Analysis of an Alphabetic Confusion Matrix", *Perception & Psychophysics*, **9** (1A), pp.40-50 (1971)
- 6) T. Kuennapas, A. J. Janson: "Multidimensional Similarity of Letters", *Perceptual and Motor Skills*, **28**, pp.3-12 (1969)
- 7) 菊地正, 山下由己男, 佐川賢, 和気典二: "触覚と視覚におけるアルファベット判別の比較", *製品科学研究所研究報告*, **81**, pp.13-20 (1977)
- 8) 田中平八: "文字パターン認知過程の構造について", *東京都立大学人文科学報*, **133**, pp.53-74 (1979)
- 9) 村石昭三, 天野清: "幼児の読み書き能力", *国立国語研究所*, pp.130-143, 東京書籍 (1988)
- 10) 行場次朗編: "認知心理学重要研究集1—視覚認知", pp.64-105, 誠信書房 (1995)
- 11) R.J. Shillman, T.Kuklinski, B.A. Blesser: "Psychophysical Techniques for Investigating the Distinctive Features of Letters", *International Journal of Man-Machine Studies*, **8**, pp.195-205 (1976)
- 12) Mary J. Naus, Robert J. Shillman: "Why a Y Is Not a V: A New Look at the Distinctive Features of Letters", *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 3, pp.394-400 (1976)
- 13) Makoto Yasuhara, Theodore T. Kuklinski: "Category Boundary Effect for Grapheme Perception, Perception & Psychophysics", **23**, 2, pp.97-104 (1978)
- 14) 横瀬善正: "形の心理学", 名古屋大学出版会 (1986)
- 15) 海保博之, 野村幸正: "漢字情報処理の心理学", 教育出版 (1983)
- 16) 萩田紀博, 増田功: "手書き文字の位置ずれと傾き変動の定量化法", *信学技報*, **PRL85-75**, pp.57-68 (1985)
- 17) 岡本正行, 押木秀樹, 中村健二, 伊東一典, 米沢義道: "手書き文字評価方法の検討", *信学技報*, **PRU90-64**, pp.47-68 (1990)
- 18) 但馬文昭, 呉健: "方向線素パターンを用いた手書き文字の変動評価の一手法", *信学論*, **J81-DII**, 6, pp.1115-1122 (1998)
- 19) 高橋ゆみ, 上笹恒, 山本毅雄: "法帖文字の形状認知", 第2回ヒューマンインターフェースシンポジウム, **2414**, pp.491-494 (1986)
- 20) 斉藤泰一, 山田博三, 山本和彦: "JIS第1水準手書き漢字データベース ETL9 とその解析", *信学論*, **J68-D**, 4, pp.757-764 (1985)
- 21) 安藤陽介, 木田博巳: "ストローク変動に基づく文字パターンの分析", *信学技報*, **PRU92-22**, pp.31-36 (1992)
- 22) 加藤隆仁, 横澤一彦: "手書き文字品質の定量評価", *信学論*, **J75-D-II**, 9, pp.1573-1581 (1992)
- 23) 加藤隆仁, 森吉弘, 横澤一彦: "手書き文字品質の評価", *ATR Technical report*, **TR-A-0074** (1990)
- 24) 野澤辰: "空間と時間", 第5章, 彰国社 (1975)
- 25) 柿崎祐一: "心理学的知覚論序説", 培風館 (1993)
- 26) 小川英光編: "パターン認識・理解の新たな展開", pp.46-51, 電子情報通信学会 (1993)
- 27) 今井四郎: "パターン認知の変換構造説 (心理学モノグラフ17)", 東京大学出版会 (1986)
- 28) 小山隆正, 山崎敏正, 山ノ井浩洋: "流れの場のパターン弁別モデル", *信学技報*, **PRU88-64**, pp.33-39 (1988)
- 29) 長石道博: "視覚の誘導場モデルを用いたパターン認識時の心理実験結果の検証", *テレビ誌*, **50**, 12, pp.1965-1973 (1996)
- 30) 長石道博: "視覚の誘導場モデルによる文字パターンの類似性に関する心理実験結果の検証", *映情報学誌*, **52**, 4, pp.594-601 (1998)
- 31) 長石道博: "視覚の誘導場モデルによる変形パターンの類似性に関する心理実験結果の検証", *映情報学誌*, **53**, 3, pp.458-465 (1999)
- 32) 久保田正, 藤村貞夫, 中路幸謙: "場による図形変形を用いた手書き文字の識別", *信学論*, **J55-D**, 5, pp.299-306 (1972)
- 33) 林幸雄: "人工ニューラルネットワークによる時空間パターン処理に関する研究", *ATR Technical report*, **TR-H-064** (1994)
- 34) 岡隆一: "セル特徴を用いた手書き漢字認識の研究", *電子技術総合研究報告*, **834** (1983)
- 35) 萩田紀博, 内藤誠一郎, 増田功: "大局的・局所的方向寄与度特徴による手書き文字認識方式", *信学論*, **J66-D**, 6, pp.722-729 (1983)
- 36) 国際電気通信基礎技術研究所編, ニューラルネットワーク応用, 第3章 視覚情報処理への応用, オーム社 (1995)
- 37) K. Yokosawa, M. Shimomura: "On the Role of Stimulus Similarity and Segmentation in Misprint Detection", In D.Brogan (Ed.), *Visual Search 2*, chapter29, pp.369-376, Taylor & Francis (1993)
- 38) 中山由美, 横塚志行: "文字認識結果の確信度に関する一検討", *信学総大*, **D-541**, p.267 (1995)
- 39) 永田昌明: "文字類似度と統計的言語モデルを用いた日本語文字認識誤り訂正法", *信学論*, **J81-D-II**, 11, pp.2624-2634 (1998)

ながいし みちひろ
長石 道博 1987年、静岡大学理学部物理学科卒業。同年、セイコーエプソン(株)に入社。1990年より(株)ATR視聴覚機構研究所に転出。現在、セイコーエプソン(株)正会員。