

視覚の誘導場モデルによる  
文字パターン認識の研究

2000年6月

博士（工学）

長石道博

豊橋技術科学大学

# 視覚の誘導場モデルによる文字パターン認識の研究

## 論文概要

本論文は、視覚の誘導場によるパターン認識のモデルを提案し、工学的応用が可能なモデルであることを示した。コンピュータに文字を認識させる研究など、コンピュータに視覚をもたせる努力が古くから行われている。しかし、従来の認識方法は、人間がヒューリスティックに考案した特徴量を用といった極めて経験的であったため、限定された条件では実用化されているが、例えば自由に書いた文字を認識することは未だ困難である。そこで、従来の方法よりも科学的なモデルや方法論を与えてくれる認知心理学、生理学的なアプローチが注目されている。しかし、これまでの多くのアプローチは、従来のヒューリスティックな手法に心理・生理学的知見のエッセンスを加えたものであったり、従来手法と比較して十分な性能が出ないなど、心理・生理学的知見をモデルに生かすことが難しかった。

本論文では認知心理学的なアプローチとして、横瀬の視覚誘導場理論に注目した。誘導場はいろいろな視覚現象を説明する心理学的な概念であり、生理学的にもその存在可能性が示唆されている。最近、誘導場のような「場」は重要な視覚の枠組みとして注目されている。そのため、誘導場は文字パターン認識など、いろいろな視覚を統一的に説明できることが期待されている。ところが、これまで視覚の誘導場の研究は、その存在や分布についての議論が主で、具体的な認識モデルはほとんど検討されていなかった。本論文では、誘導場による文字パターンの認識と空間的な配置状態の評価（文字の切り出しや読み易さ）のモデルを提案・検証し、その有効性を示した。

まず、誘導場の計算方法を提案し、いろいろなパターンの誘導場が計算できることを示した。次に、誘導場の分布の違いを弾性エネルギーで評価するパターン認識モデルを提案し、いろいろな文字パターン（人工的な線図形、活字、変形パターン、手書き文字）の類似性が、従来の手法よりも人間に近い評価が可能であることを示した。続いて、誘導場の分布状態を利用した手書き文字切り出し方法を提案し、従来の手法では切り出しが難しい非常に接近した手書き文字が切り出せることを示した。更に、誘導場の強さが、これまで尺度のなかった文字列の読み易さの指標になることを示した。また、個々の文字の誘導場の分布から文字列の読み易さを予測できることを示した。以上のように、主に分布や存在が議論されていた誘導場研究から一歩踏み出し、誘導場による具体的なモデルを提案した。そして、視覚の誘導場により、認知心理学的なアプローチから工学的に応用可能なモデルが構築できることを示した。

# Research on Pattern Recognition by the Induction Field in Vision

## Abstract

This thesis proposes and verifies a new pattern recognition model and evaluation methods of arrangement by the induction field in vision that can explain psychological experimental results. Many efforts have been made on creating a computer featuring pattern recognition. Most of the traditional recognition methods that use heuristic features depending upon experience or inspiration are useful in constrained conditions. However, the recognition abilities of these methods are insufficient for recognizing handwritten characters in free formatted condition, which are primarily used when we express an inspiration, for example, composing a memo. Therefore, cognitive psychological or physiological approaches providing scientific model with theory, and comparing previous heuristic approaches have been applied. Many recognition methods based on psychological or physiological knowledge have been considered. However, most of these methods are made by adding the essence of psychological or physiological knowledge to previous heuristic methods. It is difficult to state whether these methods have enough recognition ability compared with traditional methods.

In this thesis, Yokose's theory of the induction field in vision is applied as a cognitive psychological approach. The induction field is a physiological idea that can explain many psychological phenomena. Some physiological experiments suggest existence of the induction field in the brain. Recently, the field such as the induction field has been regarded as an important frame of human visual perception. Consequently, the induction field theory is expected to uniformly explain phenomena of visual perception. However, previous research on the induction field has been discussed mainly with connection to the existence or distribution of the induction field. Models of the induction field have never been applied to visual perception such as pattern recognition. This thesis proposes and verifies a new pattern recognition model and two evaluation methods of arrangement (segmentation of handwritten character and legible display of Japanese characters) by the induction field that can explain psychological experimental results.

At first, the calculation of the induction field on digital image is proposed assuming that the induction field is similar to electric field. Using this method one can calculate the induction field for various patterns such as handwritten characters. Next, a pattern recognition model that evaluates differences among induction fields in terms of elasticity energy is proposed. It is indicated that the proposed recognition model can explain psychological experiment results of similarity for artificial patterns, printed characters, ambiguously shaped patterns and handwritten characters compared with previous recognition methods.

Furthermore, a segmentation method for handwritten characters using the field distribution around characters is proposed. It is shown that the proposed segmentation method by the induction field can

segment very closely placed handwritten characters.

Finally, it is confirmed that the strength of the induction field can be an index for the legibility of characters. There is no method, however, that provides an index for the legible display of any types of characters. It is described that the strength of the induction field that surrounds each character can be an index for the legibility of characters. It is also shown that the legibility of characters can be estimated by using the field distribution of each character. These proposed methods of the induction field are distinguished from previous methods, and have certain advantage over the traditional methods. Consequently, it is clarified that theory of the induction field can be applied to models and methods, such as pattern recognition. Research on the induction field have progressed towards applied models and methods from argument about the existence or distribution of the induction field. It is also useful to pursue cognitive psychological approach.

# 目次

第 1 章	序 論	1
1.1	背景	1
1.2	本論文の構成	2
第 2 章	文字パターン認知の研究	5
2.1	文字認識	5
2.2	パターンの認知機構に関する研究	6
2.2.1	認知心理学の方法論	6
2.2.2	文字パターン認知の研究	7
2.3	視覚研究における場の意義	8
2.4	横瀬の視覚の誘導場理論	8
第 3 章	視覚の誘導場理論	9
3.1	横瀬の研究	9
3.2	誘導場の追試	9
3.2.1	深水, 吉田らによる誘導場の検証	10
3.2.2	外部散乱光の影響	10
3.2.3	空間異方性	11
3.3	誘導場の計算方法	11
3.3.1	横瀬の基本式	11
3.3.2	深水, 吉田らの計算モデル	13
3.3.3	ニューラルネットによる計算モデル	15
3.3.4	デジタル画像における計算モデルの提案	16
3.4	生理学的な根拠	19
3.4.1	本川の網膜誘導の場	19
3.4.2	本川の研究に対する議論	20
3.4.3	内山らの実験	21
3.4.4	杉江らの実験	21
3.5	誘導場による視知覚現象の解釈	21
第 4 章	誘導場によるパターン認識モデル	23
4.1	従来のパターン認識モデルの研究	23
4.1.1	人工的な線図形	23
4.1.2	活字	25
4.1.3	変形パターン	25
4.1.4	手書き文字	27
4.2	新しい視覚の誘導場によるパターン認識モデルの提案	28
4.2.1	ダイナミクスと場によるモデル	28
4.2.2	誘導場によるパターン認識モデル	29

4.2.3	認識アルゴリズム	30
4.3	検証方法	31
4.3.1	人工的な線図形	32
4.3.2	活字	34
4.3.3	変形パターン	37
4.3.4	手書き文字	40
4.4	検証結果	41
4.4.1	人工的な線図形	41
4.4.2	活字	45
4.4.3	変形パターン	48
4.4.4	手書き文字	52
4.5	文字認識	56
4.5.1	活字認識	57
4.5.2	手書き文字認識	57
4.6	考察	58
4.6.1	誘導場によるパターン認識モデルの有効性	58
4.6.2	ダイナミクスの意義	58
4.6.3	認識性能を向上させるための課題	59
<b>第 5 章</b>	<b>誘導場による手書き文字切り出し</b>	<b>63</b>
5.1	従来の文字切り出し	63
5.1.1	手書き文字の状態	63
5.1.2	文字切り出し	63
5.2	誘導場による文字切り出し方法の提案	65
5.2.1	複数パターンによる誘導場の分布	65
5.2.2	手書き文字列の作成と文字の隣接状態の分析	65
5.2.3	視覚の誘導場による文字切り出し方法	70
5.3	手書き文字切り出し実験	71
5.4	考察	73
5.4.1	誘導場による文字の切り出し	73
5.4.2	手書き文字切り出しにおける誘導場の意義	73
5.4.3	手書き文字切り出し性能向上のための課題	73
<b>第 6 章</b>	<b>誘導場による文字列の読み易さの評価</b>	<b>75</b>
6.1	従来の文字の読み易さの定量評価	75
6.2	文字列における視覚の誘導場	76
6.2.1	文字列の視覚の誘導場の分布と文字配置	76
6.2.2	複数の文字列による誘導場の分布	77
6.3	文字列の読み易さと誘導場の関連性の検証	77
6.4	誘導場による文字列の読み易さに関する検証結果	79
6.5	文字列を構成する文字による文字列の誘導場の予測	81
6.5.1	2文字間の誘導場による予測	81
6.5.2	文字の誘導場の予測	83
6.5.3	誘導場による読み易い文字列予測方法の検証	85
6.5.4	読み易い文字列の状態の決定	86
6.6	考察	86

6.6.1	誘導場による文字列の読み易さの説明	86
6.6.2	読み易さの客観的指標	86
6.6.3	読み易い文字列の状態の決定	87
6.6.4	読み易い文字列状態の予測性能向上のための課題	87
<b>第7章</b>	<b>今後の展望</b>	<b>89</b>
7.1	視覚の誘導場研究の発展	89
7.2	工学的応用	89
7.2.1	表情の定量評価	89
7.2.2	文字パターン認識	90
7.2.3	文字列の読み易さ	91
7.2.4	重力レンズ錯視	91
7.3	視覚の誘導場理論の展開	91
7.3.1	誘導場計算の精密化	91
7.3.2	階調, 色のある場合の視覚の誘導場	91
7.3.3	2次元パターン配置の評価の数学的な取り扱い	92
7.4	生理学からのアプローチ	92
<b>第8章</b>	<b>結 論</b>	<b>93</b>
	謝 辞	95
	参考文献	97
	付 録	105
	業績目録	111



# 目 次

1.1	本論文の構成と各事項の関連	3
2.1	現代心理学の考え方：Brunswik の図式（野澤 1975）	7
2.2	文字パターンの認知のモデル（横澤，梅田，淀川 1986）	7
3.1	心理的ポテンシャル場の計算原理	12
3.2	心理的ポテンシャル場の計算における遮蔽	13
3.3	視覚の誘導場の基本はクーロンポテンシャルである	16
3.4	デジタル画像における視覚の誘導場の計算	17
3.5	誘導場と静電場の違い	17
3.6	誘導場の計算例：正方形，三角形，Landolt 環	18
3.7	誘導場の計算例：活字，手書き文字	18
3.8	誘導場の計算例：文字列	19
3.9	誘導場の計算例：Müller-Lyer，Kanizsa の錯視図形	19
4.1	類似した線図形（Palmer，1978）	24
4.2	単純な線図形（Lupker，1979）	24
4.3	漢字パターン（横澤，梅田，淀川，1986）	24
4.4	基本的な直線セグメントで構成できる線図形	25
4.5	leg; 脚の長さの定義	26
4.6	文字サンプル例：東	27
4.7	誘導場によるパターン認識モデルの概念	30
4.8	認識過程	31
4.9	3つの評価方法による検証	33
4.10	アルファベット大文字	34
4.11	アルファベット小文字	34
4.12	カタカナ	35
4.13	アルファベット大文字（ドットマトリックス）	35
4.14	ひらがな	36
4.15	大きさが異なっている例	36
4.16	アルファベットの変形パターン 1 (R.J. Shillman 1976)	37
4.17	アルファベットの変形パターン 2 (Yasuhara ら 1978)	37
4.18	変形したカタカナ（田中）	38
4.19	変形パターン 1（横瀬）	38
4.20	変形パターン 2（横瀬）	39
4.21	評定 100 文字種	40
4.22	単純な線図形の相関比較	42
4.23	漢字パターンの心理実験と3つの評価方法との比較	43
4.24	類似した線図形の検証結果	45

4.25	心理実験と3つの評価方法との比較(「P」と「i」)	46
4.26	心理実験と3つの評価方法との比較(「Q」と「ク」)	47
4.27	心理実験と3つの評価方法との比較(アルファベット)	49
4.28	心理実験と3つの評価方法との比較(カタカナ)	50
4.29	心理実験と3つの評価方法との比較(横瀬の変形パターン)	51
4.30	心理実験と3つの評価方法との比較(画数の少ない漢字:「小」と「古」)	53
4.31	心理実験と3つの評価方法との比較(「史」と「我」)	54
4.32	心理実験と3つの評価方法との比較(画数の多い漢字:「胃」と「興」)	55
4.33	活字認識で使用了文字の誘導場例	57
4.34	文字サンプル例:次	59
4.35	変位が極端にずれている場合	62
5.1	手書き文字の接触状態の例	63
5.2	手書き文字列の射影	64
5.3	複数の文字による視覚の誘導場の例	66
5.4	手書き文字列周辺の誘導場の分布例	67
5.5	閉曲線の階層構造	68
5.6	文字要素の分離	69
5.7	閉曲線群の形状の推移	69
5.8	等ポテンシャル面からの閉曲線の抽出	70
5.9	閉曲線の包含関係	71
5.10	手書き文書からの文字の切り出し例	72
5.11	手書き文字の切り出し例	72
5.12	手書き文字の切り出し例(失敗した場合)	73
6.1	文字列における視覚の誘導場	76
6.2	呈示した文字列	78
6.3	文字の配置と文字列の変更	78
6.4	検証結果(被験者群,呈示方法の違い)	79
6.5	文字間隔,文字の大きさを一定にした場合の文字列の読み易さの違い	80
6.6	文字列の読み易さと誘導場の強さの相関	81
6.7	円における誘導場の平衡	82
6.8	円の誘導場の予測	83
6.9	円相当径 $d_H$ の測定例	83
6.10	文字の位置関係	84
6.11	実験結果(読み易い文字列の予測)	85
6.12	最適な文字間隔の予測	86
7.1	異なる表情(線画)における視覚の誘導場の例	90
8.1	Hough変換による法線方向の推定	105
8.2	変位方向の求め方	106
8.3	LDCD特徴の求め方	107
8.4	円がつくる誘導場	108

# 表 目 次

4.1	相関をとる変数の一覧（人工的な線図形）	33
4.2	相関をとる変数の一覧（活字）	37
4.3	相関をとる変数の一覧（変形パターン）	39
4.4	正解パターンを除いた相関係数の平均（人工的な線図形）	44
4.5	相関係数の検定結果（人工的な線図形）	44
4.6	正解パターンを除いた相関係数の平均（活字）	48
4.7	相関係数の検定結果（活字）	48
4.8	相関係数の平均（変形パターン）	50
4.9	相関係数の検定結果（変形パターン）	51
4.10	相関係数の例（手書き文字）	56
4.11	相関係数の検定結果（手書き文字）	56
4.12	手書き文字の認識結果例	58
4.13	正規化の効果（手書き文字）	60
4.14	誘導場の弾性エネルギーにおける正規化の効果（活字）	60
4.15	ポテンシャル値による相関係数の違い	61
5.1	手書き文字の接触状態の頻度	64
6.1	文字列の読み易さについての心理学実験結果	80
8.1	相関の強さの判定	109



# 第1章 序 論

## 1.1 背景

21世紀は情報の時代と言われている。1960年代以降の20世紀後半にコンピュータなど情報を扱える機械が飛躍的に発達した。それに伴い、複雑で大量の機械的な計算だけではなく、文書作成などこれまで手作業に頼ってきた、我々の思考活動を支援する知的な作業が徐々にできるようになった。特に1990年代にはインタ-ネットの急速な普及により、ネットワークを利用して膨大な知識の共有と利用、検索ができるようになり、コンピュータはより一層知的活動の支援に使われるようになってきた。

コンピュータを知的活動の支援に利用する上で最も問題になるのは情報の入力である。現在、主として使われているキーボード、マウス、グラフィカル・ユーザ・インターフェース(アイコンなどのメタファ)などは1960年代に発明・開発されたものである。しかし、これらを越える更に使いやすい入力方法は、マウスの発明者、ダグラス・エンゲルバートが指摘しているようにほとんど実用化されていない。例えば、1980年代にはペン入力の検討が盛んに行われ、ペン入力専用OSやベンチャー企業が生まれた。ところが、手書き文字認識性能が今日においてもまだ低いことなどから、一般ユーザーの同意は得られず、現在携帯端末などの一部の機器で利用されている程度である。このように、情報の入力は文字などを中心に認識性能は向上しているものの、なかなか実用化が進んでいない。そこで、入力を文字、音声などに限定せず、例えば、情報を取り込む機器自体を人間に装着して情報を動的に判断するウェアブル・コンピュータ<sup>注1</sup>の検討も始まっている。このコンピュータは装着した人間のおかれた状況を判断して、現レベルではやや信頼性の欠ける文字、音声などの入力を減らそうとしている。しかしながら、人間が外界から何らかの情報をもとに判断、行動する場合、情報は五感、主に視覚や聴覚から得ている。したがって、コンピュータに知的な活動をさせるには、基本的に我々が視覚や聴覚から得る情報を、コンピュータに正しく理解させなければならない。

特に視覚は重要である。人間が得る情報の大多数は視覚から得られていると言われている。視覚情報の中で、とりわけ文字は太古から使われている情報手段である。文字は太古のパピルス、粘土板、紙・本、最近の電子化された媒体など、どの時代においても中心的な情報伝達手段であった。コンピュータが文字を自由に理解できれば、非常に有効な入力の実現できる。そのため、コンピュータによる文字認識の研究は1960年代から行われている。今日までの長い研究の蓄積により、手書きの郵便番号やペン入力認識など一部実用化されている。しかし、まだ人間が自由に書いた文字を認識することは困難である。

文字パターン認識のこのような状況を打破しようと、1980年代から認知心理学、生理学などの知見から、

---

<sup>注1</sup><http://www.wearcam.org/> 解説記事：[http://www.zdnet.co.jp/macweek/9901/25/wi\\_wearable\\_1.html](http://www.zdnet.co.jp/macweek/9901/25/wi_wearable_1.html)

人間のパターン認知機構をモデル化して、その成果を工学的な文字認識手法に適用しようとする方式が盛んになった。視覚、特にパターン認知は認知心理学、生理学で最も研究が進んでいる分野であり、数多くのパターン認識の数値モデルや理論が検討された。従来の認識方法は、人間がヒューリスティックに考案した特徴量を用るといった極めて経験的なものであった。一方、認知心理学、生理学的なアプローチはより科学的なモデルや方法論を与えてくれる。しかしながら、対象パターンが比較的単純な場合のモデルが多く、手書き文字などの複雑で変形のあるパターンなどに有効なモデルまでには至っていない。

## 1.2 本論文の構成

本論文は、認知心理学的なアプローチで文字パターンの認識機構を明らかにして、認識モデルを構築することを目的としている。その手段として横瀬の視覚誘導場理論を用いた。誘導場はいろいろな視覚現象を説明する心理学的な概念である。特に最近誘導場のような「場」は、重要な視覚の枠組みとして注目されている。これまで視覚の誘導場の研究は、主にその存在や分布について議論されていた。しかし、誘導場が文字パターンの認知など、いろいろな視覚現象をどのように説明できるかといった具体的なモデルはほとんど検討されていない。本論文では、誘導場理論による文字パターンの認識と空間的な配置状態の評価（文字の切り出しや読み易さ）のモデルを提案・検証し、その有効性を示した。

図 1.1 は、本論文の構成と各事項の関連を示したものである。まず、これまでの文字パターン認知の研究（第 2 章）や視覚の誘導場理論の概要（第 3 章）、次に文字パターン認知への応用（第 4 章）、そして、2 次元的なパターン分布の評価である手書き文字の切り出し（第 5 章）と文字列の読み易さ（第 6 章）への応用が述べられている。最後に、今後の視覚の誘導場における応用と理論の展望（第 7 章）と、本論文の結論（第 8 章）がまとめられている。

第 2 章では、これまで行われてきた工学的文字認識技術は、実用化が狭い範囲内に限定されているという壁に直面しており、認知心理学的な方法論が、この壁を打破する手段として注目されていることを指摘する。そして、認知心理学における文字パターンの認知の研究を概観し、「場」は視覚を説明する枠組みであるという観点から、視覚の誘導場を用いて視覚モデルを検討する意義を述べる。

第 3 章は、視覚誘導場理論についてのまとめである。まず、横瀬が行った視覚の誘導場の研究と、深水、吉田らが最近行った精密な測定について述べ、現在までの視覚の誘導場に関する心理学的な研究をまとめる。

更に、視覚の誘導場の計算理論を、最初に横瀬の理論式、続いて側抑制モデルから見直した深水、吉田らの計算式について述べる。そして、筆者が提案したデジタル画像の計算式を説明する。また、誘導場に関連する生理学的な研究についてまとめる。

第 4 章では、誘導場による文字パターン認識モデルを提案する。このモデルがいろいろな文字パターン（人工的な線図形、活字、変形パターン、手書き文字）の類似性を説明できることを、心理実験例と比較して検証する。

第 5 章では、誘導場による手書き文字の切り出し方法を提案する。この方法が、実際に自由に筆記され

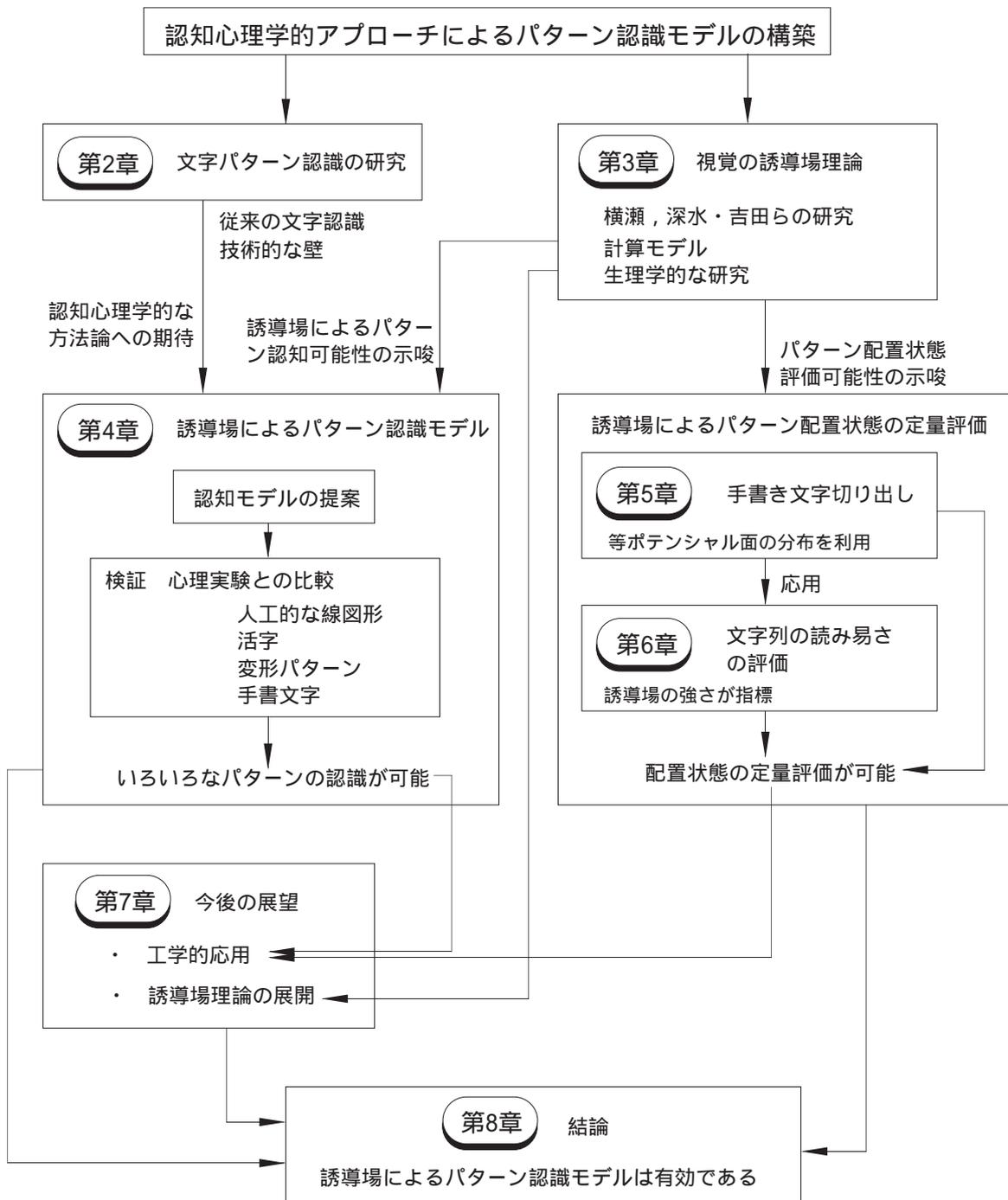


図 1.1 本論文の構成と各事項の関連

た試験答案文書の手書き文字切り出しに有効であることを示す．

第6章では，文字列の読み易さを誘導場によって定量化する方法を提案する．この方法が心理実験で得られた文字列の読み易さを説明できることを示す．また，任意の文字列の読み易さが，誘導場を計算して原理的に予測できることを示す．

第7章では，視覚の誘導場研究について，誘導場の工学的応用と今後の誘導場理論の展開の観点から展望する．

第8章では，第4章から第6章で述べた誘導場の応用と第7章の展望についてまとめ，本論文の結論を述べる．

## 第2章 文字パターン認知の研究

### 2.1 文字認識

文字は人類にとって太古から最もよく使われてきた重要な情報伝達手段である<sup>1)</sup>。コンピュータの本格的な応用がはじまった1960年代から、パターン認識、とりわけ文字認識の取り組みが開始された<sup>2)</sup>。当時、現在のコンピュータで広く使われるアイコン、メタファなどグラフィカル・ユーザ・インターフェース (Graphical User Interface : GUI) の概念は完成されていなかった。また、基本的な技術であるビットマップ、ライトペン、マウスなども開発中であった<sup>3, 4)</sup>。

このようなコンピュータの黎明期にあって、文字認識は有力なコンピュータの入力手段として非常に期待されていた。最初、アメリカでアルファベットの活字を中心に研究がはじまった<sup>2)</sup>。日本は欧米のようにタイプライタ文化が無いこと、更に仮名漢字の存在など、キーボード入力のハンデがあったため、手書き文字認識への要求と期待が非常に高かった<sup>5)</sup>。日本では1968年の郵便番号の自動読み取り機実用化を皮切りに、通産省電子技術総合研究所では、主に手書き文字を中心にデータの収集が行われ、文字データベース ETL シリーズが整備された<sup>6, 7, 8)</sup>。そして、1970年代から1980年代にかけて、ETLシリーズの整備を契機にさまざまな大学、研究機関で非常に多くの認識手法の提案、検討が行われた<sup>2, 5, 9, 10)</sup>。その結果、1980年代からのパーソナルOCRの販売、1999年の郵便番号7桁化などの実用化が着々と進んでいる。現在も、1980年代からの研究成果を更に発展させるべく多くの研究がなされている。例えば、郵政省の手書き郵便番号データベースを用いた文字認識アルゴリズムの性能コンテストが隔年で開催されている<sup>11)</sup>。

このように、一見文字パターン認識の研究や応用技術は非常に進んでいるように見られる。最近、計算機の急速な性能向上のおかげで、複雑な識別関数などが利用できるようになった<sup>12)</sup>。そのような背景もあり、確かに認識性能、速度などは着実に改良が続けられている。しかし、実際には1980年代後半から技術的に大きな壁にぶつかっている<sup>13, 14, 15)</sup>。例えば、技術的に対応できるのは郵便番号や帳票など手書きでも枠などの筆記制限がある場合が多い。発想支援のような場面<sup>16)</sup>などで我々が日常書くような自由手書き文字<sup>17)</sup>や英語などの筆記体の認識は非常に困難である<sup>9, 10, 18)</sup>。これは、文字認識、またはパターン認知の根本的な原理が解明されていないので、認識対象の制限が広がると、現在のヒューリスティックな特徴+統計的な手法ではカバーすることが難しくなるためと考えられている<sup>13, 14)</sup>。

文字パターン認識の「壁」を打破しようと、1980年代後半から人間の情報処理機構に学ぶという新しい研究の方向が注目されている<sup>19, 20)</sup>。例えば、20世紀に航空機が著しく発展したのは、空を飛ぶ「鳥」を単に模倣したのではなく、鳥に学び、鳥の飛行原理が流体力学で記述できることを見出したからである。

従来の工学的な文字認識では、特にパターンから抽出される「特徴」が非常にヒューリスティックなものであった。また、識別関数などは基本的に統計的手法である。一方、文字認識のようなパターン認知は人間が行っている情報処理である。したがって、まず人間の情報処理機構がどのようになっているか、例えば認知心理学的な方法を用いて理解し、鳥の飛行の例でいえば流体力学にあたる理論を見い出す。そして、その成果を工学的な文字認識に反映させれば、文字パターン認知の「壁」を越えることができると考えられている (13, 14, 21)。

以上のような研究が始まった 1960 年当時には非常に期待された文字パターン認識だが、基本的な原理の理解までには至っておらず、実用できる技術も限定されている。その間に、GUI を応用したコンピュータが、特に Windows95 の発売をきっかけに爆発的に普及し、パターン認識技術を使わずとも多くの人々がコンピュータを使うことに慣れてしまった。しかし、GUI の普及と同時にその限界も明確になってきている。最近、更に使いやすいインターフェイスが模索され始めている<sup>22)</sup>。真に使いやすい優れたコンピュータやインターフェイスを実現するには、人間の知覚、知能などをもっとよく研究し、その上に知識、統計的手法など各種情報を組み合わせて実用システムに活用することが王道と言われている<sup>22)</sup>。その意味で、文字パターン認識などの重要性を再び見直す動きが出てきた<sup>9, 10)</sup>。したがって、人間の文字パターン認知機構を解明することは非常に意義深いと考えられる。

## 2.2 パターンの認知機構に関する研究

### 2.2.1 認知心理学の方法論

認知心理学では、人間の情報処理機構を Black Box とみなし、適当な入力（刺激）を与えた時の出力を観測することで Black Box の機構を推測する<sup>23)</sup>。いろいろな入出力の形態や段階などがあるが、Black Box を中心としたこの関係は図 2.1 の Brunswik の図式としてまとめられている<sup>24, 25)</sup>。図 2.1 の 7 本の縦線 c, b, a, O, A, B, C は Black Box をめぐって生じる事象の因果の前後関係を示している。この方法論は、物理学者が伝統的に行う、物をぶつけて物質の状態を推測する方法<sup>注1)</sup>に似ている<sup>26)</sup>。これまで、この方法によって人間の情報処理機構、とりわけ文字パターン認識など視覚は非常に良く調べられ、認知機構のモデルがつけられてきた<sup>23, 27, 28, 29, 30)</sup>。もちろん、Black Box の実体は生理学的に決定されるべきものである<sup>23, 31)</sup>。しかし、認知心理学の方法は、人間の情報処理機構の基本的枠組みなどの重要な概念を提供してくれる上で大変重要である<sup>23, 27, 31)</sup>。この枠組みやアナロジは工学的に応用する際の有効な指針として利用できると考えられる<sup>19)</sup>。

注1) 例えば、ラザフォードは薄い金の板にぶつけた電子の反射から原子核を発見した。

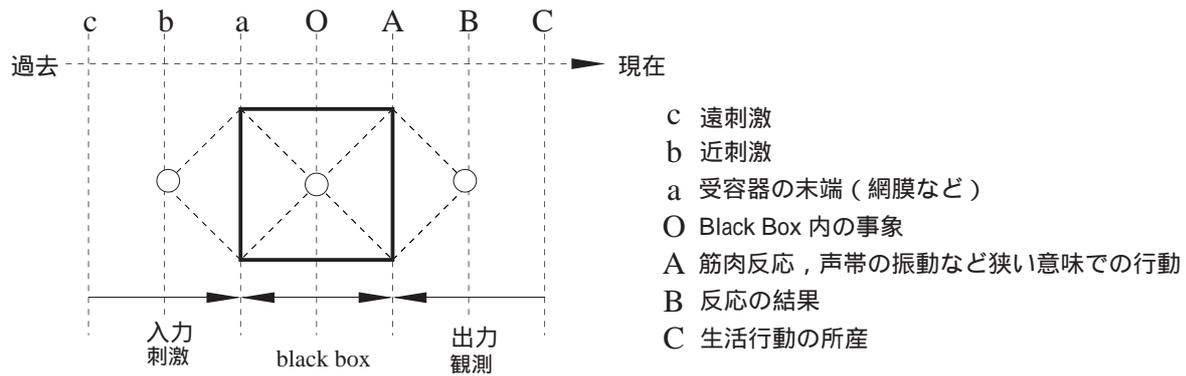


図 2.1 現代心理学の考え方：Brunswick の図式（野澤 1975）

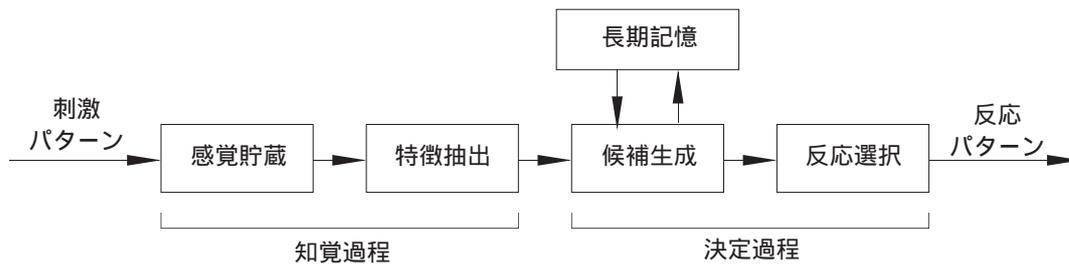


図 2.2 文字パターンの認知のモデル（横澤，梅田，淀川 1986）

2.2.2 文字パターン認知の研究

最初に文字パターン認知のモデルとして明確化されたモデルに、セルフリッジ (Selfridge) のパンデモニアム・モデル (Pandemonium model) がある<sup>32, 28)</sup>。このモデルでは、デーモン (daemon) と呼ばれる個々のモジュールが協力することで、文字パターンなどのイメージから特徴を抽出し、更に様々な情報を統合して認識が行われる概念を示した。今日の工学的文字認識の過程はこのモデルの構造に非常に近い<sup>32, 28)</sup>。

パンデモニアム・モデルは極めて抽象的なモデルである。そこで、図 2.2 の例のような文字パターンの認知モデル<sup>33)</sup>を仮定し、実際に人間が判断する文字パターンなどの類似性の心理実験結果と認知モデルの評価値を比較して、モデルの有効性を検討する数理モデルの研究が行われている<sup>19, 29, 30)</sup>。数理モデルは Brunswik の図式の Black Box を具体化したものである。そして、同じ文字パターンを数理モデルと人間に入力し、双方の出力を比べることで仮定したモデルの内容を検討する。

数理モデルは Townsend らによって初めて明確化された<sup>29)</sup>。その後、様々なパターンの数理モデルが検討された<sup>19, 30)</sup>。しかし、そのほとんどが線分のみで構成された人工的な線図形などのパターンを対象としており、我々が日常使う活字や手書き文字の数理モデルは検討されていない<sup>30)</sup>。また、数理モデルごとに類似性を説明できるパターンが限られており<sup>19)</sup>、様々なパターンの類似性を統一的に説明できる数理モデルはほとんど検討されていない<sup>19, 30)</sup>。

## 2.3 視覚研究における場の意義

前節に述べた、いろいろな文字パターンの類似性を統一的に説明することが難しいのは、人間のパターン認識機構が図 2.2 の例のように、パターンから特徴を抽出しそれを評価する方法ではないからであるという指摘がされている<sup>24, 34, 35)</sup>。また、文字パターン認知に有効な特徴を、例えば文字品質評価の心理学実験<sup>36, 37)</sup>や視線分析<sup>38)</sup>などから明らかにする研究も行われている。文字種によりある程度有効な特徴は示唆されるが、一般的に有効な特徴は明確にされていない。

一方、最近、文字パターン認知など視知覚を統一的に説明する枠組として「場」が注目されている<sup>35, 39, 40)</sup>。もし、「場」による視知覚を統一的に説明できる理論ができれば、これまでの文字パターン認知の数理モデルがいろいろな文字パターンの類似性を統一的に説明できないといった問題を解決できる可能性がある。

視知覚の説明原理としての「場」はいろいろな観点から指摘されている。例えば、柿崎はこれまでの視知覚の心理学的な知見を見直して、「場」的な発想からの理論的枠組みの有効性を指摘している<sup>35)</sup>。また、内藤は「場」を中心とした物理的なアナロジを適用することで、視知覚を統一的に説明することの有効性と理論の見通しの良さを期待している<sup>39, 40)</sup>。実際、これまでの心理学では非常に説明が難しい重力レンズ錯視に一般相対性理論<sup>41)</sup>、視空間の歪みによる錯視<sup>42)</sup>にリーマン幾何学などのように、物理、数学の概念を導入することでうまく説明できることが報告されている。

もちろん、場以外の概念による統一理論<sup>43)</sup>も検討されているが、従来の個々の事例を説明する事象の積み上げであり、場のような理論的枠組み、原理は明確ではない。一方、「場」の概念は物理学的に明確であり、数学的な取り扱いも完備しており、「場」のアナロジが心理学現象にうまく適用できれば、内藤<sup>39, 40)</sup>が指摘しているように視知覚の物理数学モデルが明確になると考えられる。また、Maxwell の方程式がいろいろな電磁界現象を統一的に説明できるように、「場」の概念で構築された数理モデルが、いろいろな視知覚現象を統一的に説明することも期待できる。

## 2.4 横瀬の視覚の誘導場理論

「場」の概念に基づいた心理学的理論として、横瀬は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」を提案している<sup>44)</sup>。横瀬は図 2.1 の入力(刺激)段階を、特に刺激布置(constellation)条件ととらえ、背景と刺激の関係、周囲の状況から「場」の重要性に注目している。そして、誘導場の分布は静電場に似ており、計算で求められることを示した。また、横瀬は多くの心理実験から、この図形の周りに仮定した場が、文字パターン認知、文字の読み易さなどの視知覚現象に関わっていると推測している<sup>44)</sup>。残念ながら、各視知覚現象を具体的に説明する機構の解明、モデルの構築までには至っていない。しかし、場がいろいろな視知覚現象を統一的に説明する有効な概念として注目されていること、そして横瀬が視覚の誘導場はいろいろな視知覚現象に関わっていると指摘していることから、「視覚の誘導場理論」が文字パターン認知などを統一的に説明できる理論の1つとして非常に有効であることが期待される。

## 第3章 視覚の誘導場理論

### 3.1 横瀬の研究

視覚の誘導場は横瀬が独自に発展させた心理学的な概念である。第二次世界大戦前後、主にカモフラージュの研究を行っていた横瀬は、図形と背景と明度比が交通標識などの視認に重要な意義をもっていることを明らかにした<sup>44)</sup>。更に図形と背景の境界線が図形の視認に大きくかかわっていることに気づいている。一方、輪郭図形を瞬間提示後、黒い正方形を提示すると、最初に提示した輪郭図形の輪郭の周辺が白く抜けて見えるという輪郭線効果<sup>30)</sup>という現象が知られている。これは、図形の輪郭線のところだけではなく、輪郭線からやや離れたところまで何らかの影響を与えていると考えられる。そこで、横瀬は図形の輪郭線効果に着目し、特に輪郭図形からの距離によってどのような量的な違いがあるかを実験により明らかにしようとした。図形の輪郭線の重要性はそれまでも指摘されていたが、その効果を定量的に調べ、評価した例はほとんどなかった<sup>45)</sup>。横瀬は、カモフラージュの研究などで行っていた図形と背景と明度比の測定実験方法を元に、図形の周囲のいろいろな場所に小光点を提示した後に、徐々に弱めた小光点が見えなくなる強さを測定する心理実験（光刺激閾法）を考案し、輪郭線の効果を調べた。その結果、同じ強さの小光点の見えなくなる明るさ（閾値）を結んだ等閾値線が、提示した図形の周りに地図の等高線のように分布していることを見出した<sup>44)</sup>。このことから、横瀬は図形の周りにはある種の効果が波及しており、それは一種の場を形成していると考えた。そして、多くの実験データから図形の場の強さ  $M$  は、図形からの距離  $D$  に逆比例し、図形の明瞭度  $H$ （図形とその背景との明度比）に比例し、図形の固有結構<sup>注1</sup>  $E$  の関数である式 (3.1) で表されることを示した ( $a, b$  は定数)<sup>44)</sup>。

$$M = f \left( E \frac{H^b}{D^a} \right) \quad (3.1)$$

### 3.2 誘導場の追試

視覚の誘導場は、横瀬を中心とする名古屋大学グループの小光点を用いた光刺激閾法およびこれを改良した方法（小点変位法、変位 matching 法、小光点減法など）や flicker 法による多くの実験結果が、その存在を支持している<sup>44, 46, 47)</sup>。そして、いろいろな研究者が求めた誘導場の分布は、測定方法の如何を問わず非常に類似している<sup>46)</sup>。このことから、図形の周りに誘導場のような「場」が存在することはほぼ確認されたと言える。

<sup>注1</sup> 図形の形に関する変数で、どんな形の図形を刺激に用いるかにより誘導場の強さが変わる<sup>44)</sup>。W.Köhler らが追及してきたゲシュタルト性質そのものを表現した変数と考えられる<sup>24)</sup>。

### 3.2.1 深水，吉田らによる誘導場の検証

以上の研究の多くは電卓もない古い時代に行われたものである。例えば，横瀬の光刺激閾法では，刺激図形の呈示は60Wの電球やスリガラスを用いており，明るさは光の電圧をテスターで換算しているなど<sup>44)</sup>，光源の均一性や明るさの測定精度などは現在よりかなり悪いことが想像される。したがって，工学的に応用するにはより精密な精度での検証が必要と考えられる<sup>48)</sup>。

最近，深水，吉田らを中心とする拓殖大学のグループが，数学的解析に耐え得る十分な精度の心理実験結果が得られるよう非常に精密な測定装置，環境を整え，横瀬と同じ光刺激閾法を極めて精密に行い，横瀬の誘導場を検証した<sup>48)</sup>。まず，深水，吉田らは，照度ムラが非常に少ない図形を提示する白色のスクリーンを用意した(400[lx]の照明に対し偏差1[lx]以内)。そして，小光点はパーソナルコンピュータの制御により，任意の位置への極めて精密な移動(回転角度の最小値0.0072度)と明るさの調整ができる。被験者はマウスを操作するだけで小光点が消失する明るさを効率的に記録することが可能である。

この精密な装置を使った実験結果は，横瀬の実験結果とおおよそ一致することが示され，横瀬の誘導場の再現性が確認された<sup>48)</sup>。ただし，詳細については横瀬の実験結果と異なる結果が得られている。まず，場の存在領域が図形の比較的近傍に限定されており，横瀬のように図形からかなり離れた場所での場は確認できなかった。また，図形の近傍の場は図形を一巡するような閉曲線になっていなかった。

このような違いを更に詳しく調べるため，深水，吉田らは横瀬の実験をさらに追試した<sup>49)</sup>。その結果，横瀬の実験では場の強さは図形から100mm程度離れて飽和するのに，深水，吉田らの実験ではわずか10mm程度で飽和してしまうことがわかった。これは，閾値の測定方法が異なることによるものと考えられる。また，光源の明るさが横瀬は60[lx]と暗いこと，深水，吉田らは両眼だが，横瀬は単眼のため閾値が低めに出ていると考えられる。したがって，これは実験方法の違いによる閾値の測定範囲の違いであり，場は図形からかなり離れた場所まである程度広がっていると考えられる。

また，追試の結果，図形の近傍の場に関して，単に閉曲線にならないだけでなくクロスや尻切れなどの分布があることがわかった。このような分布は，横瀬の実験などでは報告されておらず，深水，吉田らの精密な測定で初めて示された分布である。ただし，これらの違いは，個人差，疲労などの実験誤差も考えられる。特に光刺激閾法は順次提示される小光点を判断させるため，各人の判断基準の変動，順序効果が影響している可能性も高いと考えられる<sup>49)</sup>。

### 3.2.2 外部散乱光の影響

光刺激閾法は小光点という光刺激図形を用いて行われる実験である。そのため，眼球内の散乱光の分布によって，横瀬の場と同様の結果が得られる可能性が指摘されている<sup>47)</sup>。横瀬は誘導場と外部散乱光とは別な効果であることを主張している<sup>47)</sup>。宗宮<sup>50)</sup>，高橋<sup>51)</sup>は，誘導場の効果は確かに存在し，外部散乱光は小さいことを実験的に示している。一方，野澤はこのような疑問がある小光点を用いず，CRT画面上に

図形と小点を提示して場を測定し、横瀬と同様な場を観測している<sup>52)</sup>。以上から、誘導場は確かに存在し、外部散乱光の効果と別の効果であると考えられる。

### 3.2.3 空間異方性

横瀬は暗黙のうちに誘導場が空間的に一様であるとしているが、誘導場の空間異方性を示唆する実験結果が幾つか報告されている<sup>47)</sup>。例えば野澤は、Aを逆さにすると、正立の時と、上下および左右の場の強さが異なる結果を得ている<sup>注2)</sup>。また、深水、吉田らの精密な測定から、誘導場の分布は図形に対して非対称の可能性が示されている<sup>49)</sup>。一方、W.Köhlerは正方形□を傾けて◇にすると角のみかけが著しく異なると述べている<sup>53)</sup>が、野澤の誘導場の測定では角付近の場の強さについて有意差はないと報告している<sup>54)</sup>。なお、空間異方性という場合、暗黙の内に1つの図形の中心を基準に異方性を議論することが多いようであるが、このような仮定自体に問題があるとの指摘もある<sup>47)</sup>。このように、空間異方性はまだ議論すべき点が残されているが、空間異方性が認められる場の強さの違いはわずかであり、図形の場全体としてはほぼ空間的に一様と見なしてよいと考えられる。

このように誘導場の分布は、図形近傍の分布の特異性や空間異方性などが報告されているが、基本的に横瀬の実験結果通りと見て良いと考えられる。

## 3.3 誘導場の計算方法

### 3.3.1 横瀬の基本式

横瀬は、任意の図形の場を数値計算で予測するための計算式を提案、検証している。横瀬はW.Köhlerの「心理物理同型説」<sup>53)</sup>を念頭に、図形の周りに有る「場」のようなものが磁場に似ている<sup>注3)</sup>ことから、実験で得られた式(3.1)の関係を磁界の強さを求めるビオ・サバルの法則のアナロジを利用して、図形の場の強さが予測できる「心理的ポテンシャル場の理論方程式」を提案した<sup>44, 55, 56, 57)</sup>。

まず、単一の線分がつくる場について考える。図3.1(a)のように導線 $ab$ に電流が流れて反時計周りに磁界が生ずると考える。そして、図3.1(b)において、導線 $ab$ を線分 $ab$ に置き換えた場合、線分 $ab$ の付近の任意の点 $P$ における場の強さは、線分 $ab$ 上の単位長が点 $P$ に及ぼす影響を線分 $ab$ について定積分したものとする。点 $P$ から線分 $ab$ に下ろした足を $O$ 、 $\overline{PO} = D$ 、 $\angle aPO = \theta_1$ 、 $\angle bPO = \theta_2$ 、 $H$ を線分 $ab$ の明瞭度とすると、点 $P$ における場の強さ $M_p$ は、

$$M_p = k \frac{H}{D} \int_{-\theta_2}^{\theta_1} \cos \theta d\theta = k \frac{H}{D} (\sin \theta_1 + \sin \theta_2) \quad (3.2)$$

で表される( $k$ は被験者の視力などの個人差の相違を表す)。2値画像など背景との明度差が一定の場合な

注2) F. Nagaoka 測定 1970 (未発表)。

注3) 横瀬は大脳皮質という媒質が何らかの磁気的な場を発生させていると考えていたようである<sup>44, 55, 56)</sup>。

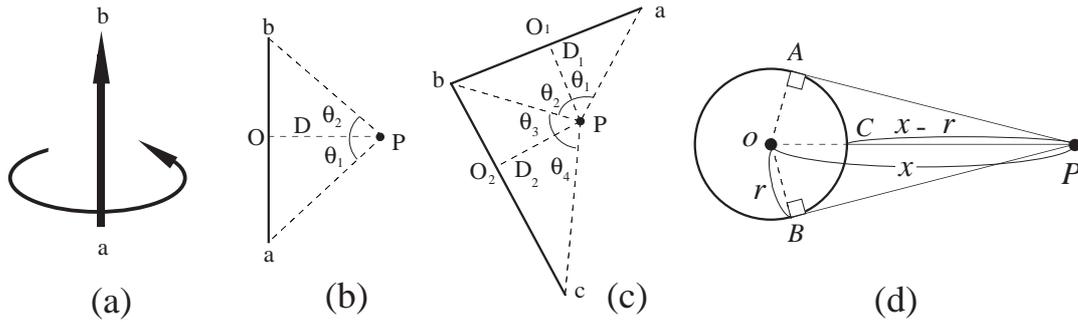


図 3.1 心理的ポテンシャル場の計算原理

どで場の強さを考える場合は、 $H$  と  $k$  は省略できるので、式 (3.2) は次のように表現できる。

$$M_p = \frac{\sin \theta_1 + \sin \theta_2}{D} \quad (3.3)$$

横瀬は式 (3.3) を心理的ポテンシャル場の基本公式と呼び、これを元にいろいろな図形の場の計算を行った。

図 3.1(c) のように 2 つの線分  $ab$ ,  $bc$  が点  $P$  につくる場の強さは、点  $P$  から線分  $ab$ ,  $bc$  にそれぞれ下ろした足を  $O_1$ ,  $O_2$ ,  $\overline{PO_1} = D_1$ ,  $\overline{PO_2} = D_2$ ,  $\angle aPO_1 = \theta_1$ ,  $\angle bPO_1 = \theta_2$ ,  $\angle bPO_2 = \theta_3$ ,  $\angle cPO_2 = \theta_4$  とすると、

$$M_p = \frac{\sin \theta_1 + \sin \theta_2}{D_1} + \frac{\sin \theta_3 + \sin \theta_4}{D_2} \quad (3.4)$$

で表される。このように 2 つの線分がつくる場は各線分がつくる場の重ね合わせになる。線分が複数の場合も、式 (3.4) と同様に計算を行うことができる。このように、式 (3.3), (3.4) を用いることで任意の線図形の場を計算で予測することが可能になる<sup>44, 55, 56, 57</sup>。

なお、円がつくる場の強さは別の公式で計算する。図 3.1(d) のように半径  $r$  の円の中心  $O$  から距離  $x$  の点  $P$  における誘導場の強さ  $M_p$  は、点  $P$  から円に引いた接線を  $PA$ ,  $PB$ , 円と  $PO$  の交点を  $C$  とすると、円弧  $\widehat{ACB}$  からの影響で決まり、次の式で計算される<sup>44, 58</sup>。

$$M_p = \frac{2r}{(x-r)x} \quad (3.5)$$

円場で注意すべき点は、点  $P$  から見えない円弧は場の形成に寄与しないことである<sup>44, 58</sup>。すなわち、点  $P$  から見えない図形部分は誘導場の計算に入れなくてよいことになる。このような「遮蔽」が誘導場にあることを横瀬は明言していないが、円の場合と同様な遮蔽が線図形でもある<sup>44, 55, 56, 57</sup>。例えば、図 3.2(a) のように、点  $P$  から線分  $ab$  に下ろした垂線の足  $O$  が線分  $ab$  上でない場合、まず仮想的に線分  $bO$  の寄与から線分  $aO$  の寄与を差し引いて場の強さを求める。点  $P$  の場の強さは、 $\angle bPO = \theta_1$ ,  $\angle aPO = \theta_2$  とすると、

$$M_p = \frac{\sin \theta_1 - \sin \theta_2}{D} \quad (3.6)$$

で表される。式 (3.6) が示すように、必ず点  $P$  から対象の線分  $ab$  が見える部分からの寄与だけを場として計算する。また、図 3.2(b) の場合は 3 つの線分  $ab$ ,  $cd$ ,  $ef$  それぞれについて式 (3.3) で計算した場の強さ

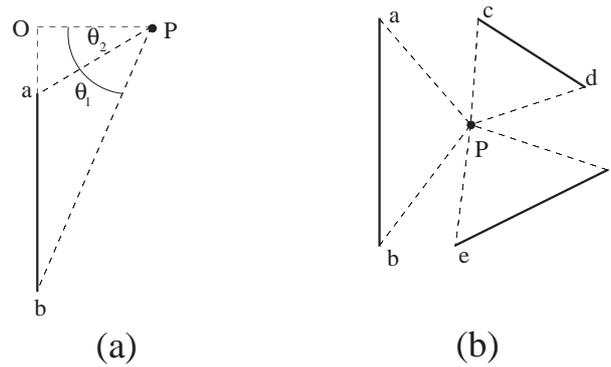


図 3.2 心理的ポテンシャル場の計算における遮蔽

を式 (3.4) の要領で加算する．この時，点  $P$  からは 3 つの線分  $ab$ ， $cd$ ， $ef$  が全て見えていることが前提になっている．以上のように，線分の場合でも必ず点  $P$  から見えている部分だけが場に寄与する．

以上の式 (3.3)，(3.4)，(3.5)，(3.6) のポテンシャル場の公式による場の計算結果と，心理実験による場の測定結果とは一致することが報告されている<sup>59, 44, 55, 56, 57</sup>．

### 3.3.2 深水，吉田らの計算モデル

深水，吉田らは，横瀬の 4 つの場の基本公式 (式 (3.3)，(3.4)，(3.5)，(3.6)) を使って様々な図形の場の数値シミュレーションを試みた<sup>60</sup>．その結果，基本的に場の数値シミュレーションは，横瀬の実験結果と良く一致することを示した．ただし，誘導場の計算はすべて積分演算であることから，線分を太くする，線分先端の計算精度をあげるなど，誤差の累積を防ぐための修正が必要である．

一方，深水，吉田らは 3.2.1 のように誘導場の詳細な再現実験を行ったところ，横瀬の報告にはない図形の近傍の分布が見い出された<sup>49</sup>．この分布は，横瀬の 4 つの場の基本公式の数値シミュレーションに修正を加えても対応できなかった<sup>60</sup>．

そこで，深水，吉田らは横瀬の基本公式の理論的見直しを行った<sup>61</sup>．まず，磁場では場の吸い込みがあるため心理実験結果と矛盾しており，誘導場は静電場ポテンシャルの解釈に近いことを指摘した．しかし，式 (3.7) の Laplace の方程式による静電場ポテンシャル  $\phi$  を計算したところ心理実験結果を再現できなかった．このため，電荷分布が座標  $(x, y)$  に依存する等質でない電荷  $\rho(x, y)$  と仮定し，式 (3.8) の Poisson の方程式でポテンシャル  $\phi$  を計算した結果，図形近傍における場の分布の尻切れ状態のみを再現できた．

$$\frac{\partial \phi}{\partial x^2} + \frac{\partial \phi}{\partial y^2} = 0 \quad (3.7)$$

$$\frac{\partial \phi}{\partial x^2} + \frac{\partial \phi}{\partial y^2} = \rho(x, y) \quad (3.8)$$

このように，純粋な静電場モデルでは場の吸い込みはないが，実験結果を充分再現することは困難である．また，Poisson の方程式の  $\rho(x, y)$  のような電荷分布が等質でないという仮定は，個々の図形ごとに決めなければならない，汎用の誘導場の基本公式として利用するには問題がある．

このような静電場モデルの問題点に鑑み、深水、吉田らは自らが測定した誘導場の分布を説明できる新たな計算モデルを提案している。まず、藤井の側抑制モデル<sup>62)</sup>を利用した計算モデルの検討を行った<sup>63)</sup>。すなわち、網膜上の任意の点の出力  $o(x, y)$  は、刺激図形  $i(\zeta, \eta)$  と結合係数  $w(x, y)$  の畳込み積分で表される<sup>62)</sup>。

$$o(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} w(x, y) i(\zeta - x, \eta - y) d\zeta d\eta \quad (3.9)$$

深水、吉田らは式 (3.9) が図形の場を再現できるように結合係数  $w(x, y)$  を心理実験値から求めた。

$$\begin{aligned} w(x, y) &= \frac{1.6}{3.125 \times 10^{-4} \pi} e^{d_1} - \frac{1.0}{1.8 \times 10^{-3} \pi} e^{d_2} \\ d_1 &= -\frac{x^2 + y^2}{3.125 \times 10^{-4}} \quad d_2 = -\frac{x^2 + y^2}{1.8 \times 10^{-3}} \end{aligned} \quad (3.10)$$

式 (3.9) と (3.10) による側抑制モデル  $F_{(0)}$  は、式 (3.7) や式 (3.8) による、これまでの静電場のモデルよりも心理実験結果を再現できることが示されている<sup>63)</sup>。

そこで、側抑制モデル  $F_{(0)}$  を基本に、視空間の異方性、空間周波数特性、網膜神経密度、図形濃度、更に実験条件の補正を加えて、心理実験の再現性を図った計算モデル  $F_{(1)}$  を提案した<sup>64, 65)</sup>。まず、刺激図形  $i(\zeta, \eta)$  の図形濃度を次のように定義する。

$$0 \leq i(\zeta, \eta) \leq 1 \quad (3.11)$$

式 (3.11) で白の地は  $i = 0$ 、無反射の黒領域は  $i = 1$  である。このとき、側抑制の出力  $o(x, y)$  を用いて、図形近傍の小光点の閾値  $S_r(x, y)$  を次のように表す。

$$\begin{aligned} F_{(1)} &= S_r(x, y) \\ &= K(B, D) o(x, y) + H(B, D) \\ &= K(B, D) \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} w(\zeta - x, \eta - y, D) M(i(\zeta, \eta)) d\zeta d\eta + H(B, D) \end{aligned} \quad (3.12)$$

ただし、 $D$  は視距離、 $B$  は環境照度、 $K(B, D)$  は側抑制の出力から閾値へのスケーリング変数、 $H(B, D)$  は閾値の飽和、 $M(c)$  は刺激強度  $c$  に対する受容強度である。側抑制の結合係数  $w(x, y, D)$  は、

$$\begin{aligned} w(x, y, D) &= \left( \frac{K_1}{2\pi\sigma_1^2} e^{d_1} - \frac{K_2}{2\pi\sigma_2^2} e^{d_2} \right) \Theta(\tan^{-1} \frac{y}{x}) \\ d_1 &= -\frac{X^2 + Y^2}{2\pi\sigma_1^2} \quad d_2 = -\frac{X^2 + Y^2}{2\pi\sigma_2^2} \\ X &= \frac{D_0}{D} x \quad Y = \frac{D_0}{D} y \end{aligned} \quad (3.13)$$

で表される。ただし、 $\sigma_{1(2)}$  は興奮（抑制）部の空間定数、 $K_{1(2)}$  は興奮（抑制）部の応答、 $X, Y$  は基準視距離に変換後の空間スケーリング値、 $D_0$  は基準視距離 (2000mm)、 $\Theta$  は方位別強度関数である。式 (3.12)、(3.13) で表される  $F_{(1)}$  モデルに、実験的に求めた受容強度  $M(c)$ 、結合係数  $w(x, y, D)$ 、方位別強度関数  $\Theta$ 、更に  $K(B, D)$ 、 $H(B, D)$  を用いて、シミュレーションを行った。その結果、 $F_{(1)}$  モデルは心理実験結果と

の整合性が著しく高い (誤差 5%以内) ことが示された<sup>64, 65</sup>).  $F_{(1)}$  モデルは,  $F_{(0)}$  モデルのように図形によって結合係数が異なることがないため, 一般性が非常に高い. 今後, 濃淡や色のある図形などの場のモデルの拡張と図形視などの理論的解析が期待される.

### 3.3.3 ニュ - ラルネットによる計算モデル

中島らは, ニューロンの受容野関数を式 (3.9) のような, 従来のガウス関数の線型結合で表される DOG (Difference Of two Gaussian) 関数ではなく, 視覚の誘導場の効果を入れた新しい関数に置き換えた誘導場の計算モデルを検討している<sup>66</sup>).

横瀬の誘導場の基本式 (3.2) を離散化した空間で表現すると, 点  $P_0$  におけるポテンシャル  $P(P_0)$  は次式で与えられる.

$$P(P_0) = \sum_i^n \frac{1}{r_i} \quad (3.14)$$

ここで,  $r_i$  は点  $P_0$  と線分  $ab$  を構成する  $n$  個の任意の点  $P_i$  の距離である. 式 (3.14) から, ポテンシャルは各構成点からの距離  $r_i$  の逆数である. そこで,  $r_i$  の逆数をニュー - ロンの結合荷重とした心理的ポテンシャル生成ニュー - ラルネット (Potential created Neural Network; PCNN) を考える. そして, 次のような距離の逆数関数を, PCNN の新しい受容野感度関数 DIP (Difference of two Inverse Proportion) とする.

$$g(x) = \begin{cases} b & (|x| \leq \varepsilon) \\ \frac{1}{a|x|} & (\varepsilon \leq |x|) \end{cases} \quad (3.15)$$

ただし,  $\varepsilon = 1/ab$ ,  $a$  は画像のぼかし具合を調節するパラメータ,  $b$  はニュー - ロンの興奮の飽和を規定するパラメータである.

中島らは, 誘導場の分布を直接求めるモデルは提案していない. しかし, 式 (3.15) によるニュー - ラルネットの入出力関係を空間周波数領域, 具体的には, 式 (3.15) をフーリエ変換し, その周波数特性を調べている. 式 (3.15) のフーリエ変換  $G(\omega)$  は,

$$\begin{aligned} G(\omega) &= \int_{-\infty}^{+\infty} g(x)e^{-j\omega x} dx \\ &= -\frac{2}{a} C_i(\varepsilon\omega) + \frac{2b}{\omega} \sin(\varepsilon\omega) + j \left\{ \frac{2}{a} s_i(\varepsilon\omega) \right\} \end{aligned} \quad (3.16)$$

で表される. ここで,  $C_i, s_i$  はそれぞれ,

$$C_i(x) = -\int_x^{+\infty} \frac{\cos t}{t} dt, \quad s_i(x) = -\int_x^{+\infty} \frac{\sin t}{t} dt \quad (3.17)$$

である. 式 (3.16), (3.17) を用いて, シミュレーションした結果, PCNN は, 図形の体制化, ポツゲンドルフ錯視など幾何学的角度方向錯視現象の錯視量の定量的な説明に有効なことを報告している<sup>66</sup>).

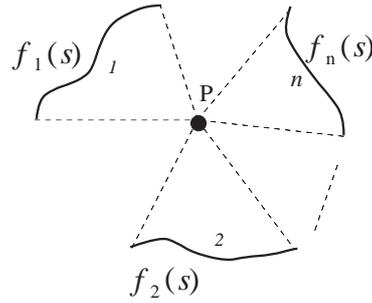


図 3.3 視覚の誘導場の基本はクーロンポテンシャルである

### 3.3.4 デジタル画像における計算モデルの提案

横瀬の4つの場の基本公式(式(3.3), (3.4), (3.5), (3.6))は、直線と円で構成された図形の誘導場計算しかできない問題がある。一方、深水、吉田らや中島らの生体工学的な計算モデルは、デジタル画像のような離散的な画像の誘導場を計算するには不向きである。ここでは、横瀬の基本公式を基に、デジタル画像の誘導場が計算できるモデルを提案する。

深水、吉田らが指摘しているように<sup>61)</sup>、横瀬の4つの場の基本公式(式(3.3), (3.4), (3.5), (3.6))は電磁気学的には誤りである。横瀬は単に電磁気学(ビオ・サバールの法則)のアナロジを利用しただけと考えられる。ここでは、横瀬の場の公式をより一般化した計算方法について述べる。

まず、式(3.3)を純粹に解釈すると、点Pからの距離を積分していることから、ビオ・サバールの法則ではなく、誘導場の強さは単純なクーロンポテンシャルであることがわかる<sup>67)</sup>。実際に Eriksson は誘導場の強さが距離  $r$  の逆数  $1/r$  で表されることを心理実験から示している<sup>68)</sup>。図3.3のように、 $n$ 個の曲線が点Pにつくる場を考える。点Pから各曲線  $j(j = 1, 2, \dots, n)$  上の任意の点  $s$  までの距離の関数を  $f_1(s)$ ,  $f_2(s)$ , ...,  $f_n(s)$  とおく。誘導場はクーロンポテンシャルであるから、場の大きさは距離の関数  $f_j(s)$  の逆数を曲線上の各点について積分したものになる。また、点Pの場は各曲線  $j$  がつくる場の重ね合わせだから、

$$\begin{aligned} M_p &= \int \frac{ds}{R(f_1(s))} + \int \frac{ds}{R(f_2(s))} + \dots + \int \frac{ds}{R(f_n(s))} \\ &= \sum_{j=1}^n \int \frac{ds}{R(f_j(s))} \end{aligned} \quad (3.18)$$

と表される。

次に、白黒2値のデジタル画像における誘導場を計算する。パターンの外郭を構成する画素を点電荷と仮定し、それらがつくるクーロンポテンシャルの集積から、デジタル画像における誘導場の分布を計算する。図3.4(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。曲線を構成する各点を正電荷1の点電荷と仮定し、点Pから曲線  $f(s)$  上の点  $i$  までの距離を  $r_i$  とする。点

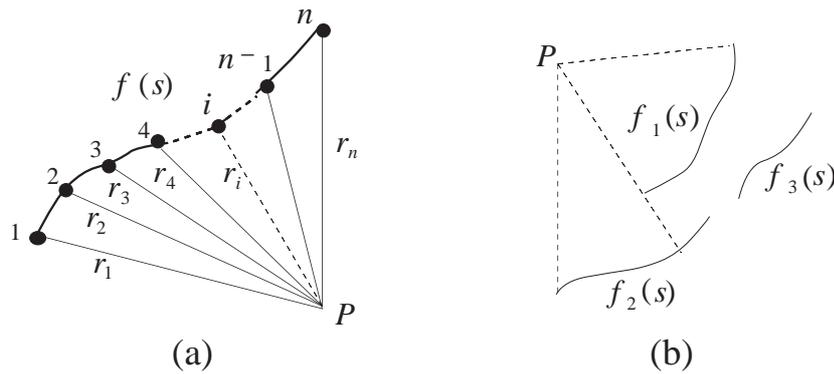


図 3.4 デジタル画像における視覚の誘導場の計算

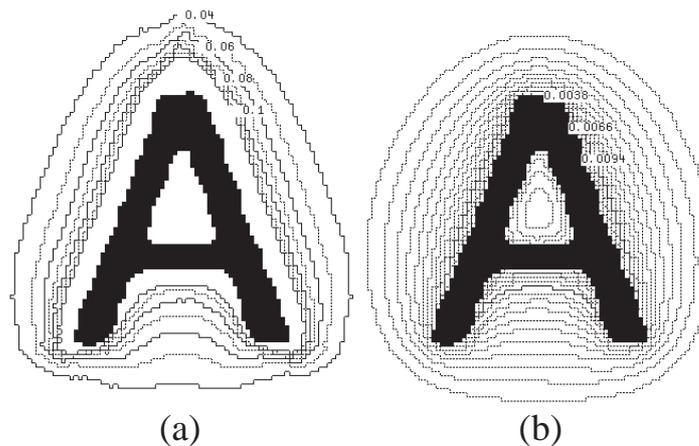


図 3.5 誘導場と静電場の違い

$P$  における誘導場の強さ  $M_p$  を次のように定義する .

$$M_p = \frac{1}{n} \sum_i^n \frac{1}{r_i} \tag{3.19}$$

曲線が複数ある場合 , 式 (3.18) と同様に点  $P$  における誘導場の強さは個々の曲線が点  $P$  につくる誘導場の和になる .

式 (3.19) において , 3.3.1 の「遮蔽」を考慮して , 点  $P$  から見える部分のみの和をとるという制約条件がつく . 例えば , 図 3.4(b) の曲線  $f_1(s)$  に遮蔽されて点  $P$  から見えない曲線  $f_2(s)$  の一部と曲線  $f_3(s)$  の和はとらない . なお , デジタル画像では点  $P$  から見える有効な画素を見つける走査が必要である . 式 (3.14) も  $r_i$  の逆数である点では同じだが , デジタル画像の場合 , 走査は画素が有限というデジタル画像の性質 (画素の連結性) のため , 走査する場所 , 向きにより見つかる画素の密度が変化する . したがって , 式 (3.19) のように走査対象の画素数  $n$  による規格化 ( $1/n$ ) を行うことで単位画素あたりの場の強さを求める .

図 3.5(a) は式 (3.19) で計算した誘導場の例である . 図の「A」周辺に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 0 に近づく . 数字は場の強さを示す<sup>注4</sup> .

図 3.5(a) の「場」の分布の形状・強さは , 横瀬が行った四角形や三角形など , 図形の角付近に関する誘導

<sup>注4</sup> 本論文では , 誘導場の強さの単位を無次元として扱う . 詳しくは付録 F を参照 .

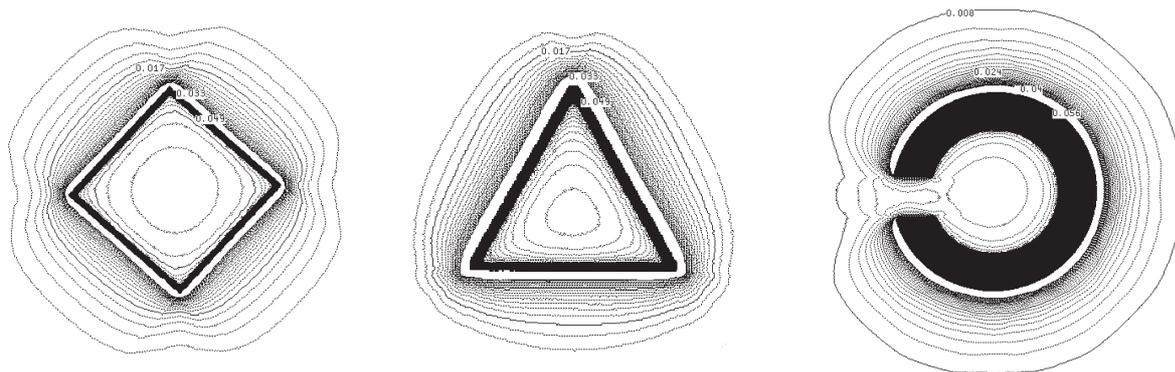


図 3.6 誘導場の計算例：正方形，三角形，Landolt 環

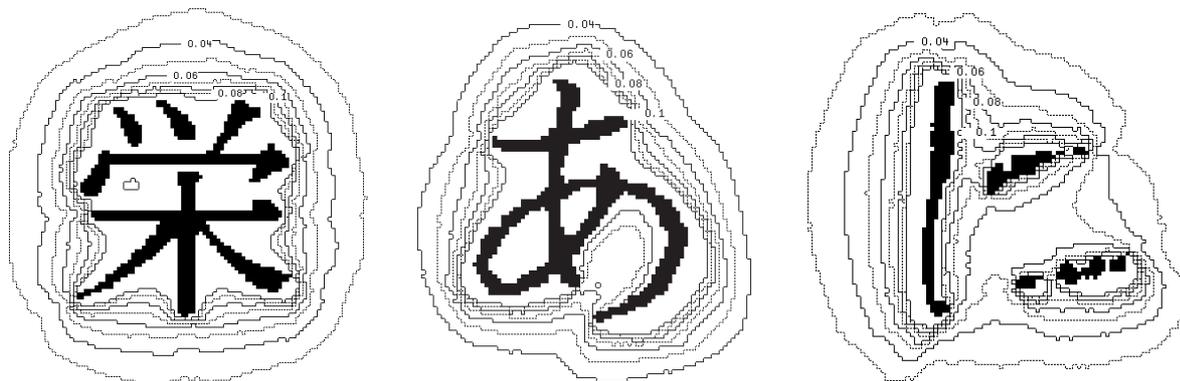


図 3.7 誘導場の計算例：活字，手書き文字

場の分布の心理実験結果<sup>44, 55, 56, 57</sup>)と一致する。一方，図 3.5(b) は遮蔽条件がなく，画素全てを正電荷 1 の点電荷と仮定した静電場の例であるが，「場」の分布は全体的に丸く心理実験結果と異なる。

図 3.6 は式 (3.19) を用いて，正方形，三角形，Landolt 環の誘導場を計算した例である。図 3.6 の例も含め，図形の角周辺の誘導場が他の部分の誘導場に比べて形状や密度などが急激に変化している。一方，視力は Landolt 環の切れ目の幅の逆数で表されるが，図 3.6 からその切れ目の周辺の誘導場が急激に変化している。このように誘導場は図形周辺に図形の輪郭を形どるようにならば一様に分布するのではなく，図形を特徴付ける角や切れ目といった部分の周辺に変化が生じることが実際に計算してみるとわかる。また，図形の内部の場も計算できる。深水，吉田らの行ったような式 (3.8) の純粋な静電場の計算では計算できないが，式 (3.19) では可能である。

式 (3.19) を用いることにより図形パターンのビットマップが得られれば，原理的に任意のパターンの誘導場が計算できる。例えば，図 3.7 のような活字やかすれた手書き文字のようなパターンの誘導場を求めることもできる。更に，図 3.8 の文字列の例のように，単独のパターンだけではなく複数のパターンが存在する 2 次元空間での誘導場も計算できる。また，文字などの図形パターンだけでなく，図 3.9 の Müller-Lyer，Kanizsa の錯視図形<sup>69</sup>) などの特殊な図形についても誘導場を計算することが可能である。

このように，横瀬の理論を拡張した式 (3.19) を用いることで，非常に様々なパターンの誘導場を計算す

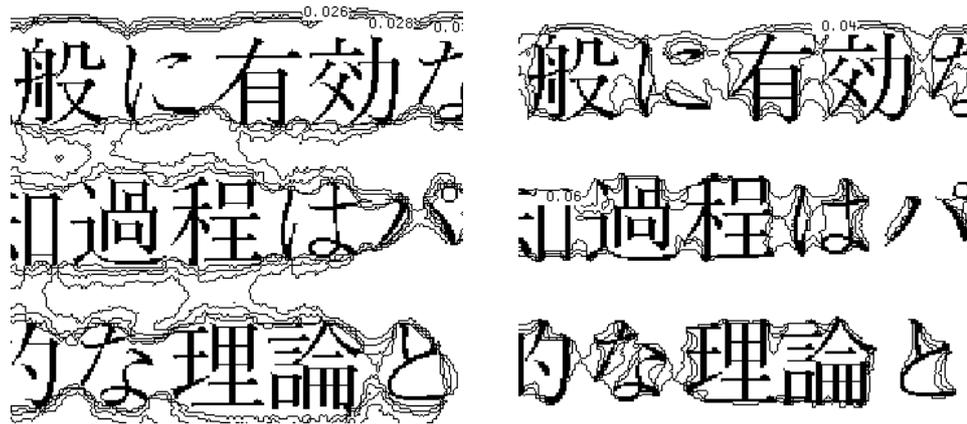


図 3.8 誘導場の計算例：文字列

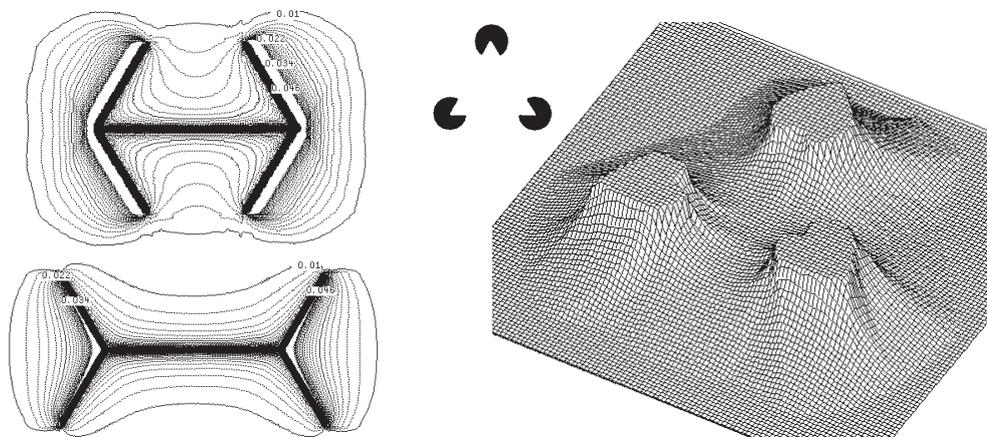


図 3.9 誘導場の計算例：Müller-Lyer, Kanisza の錯視図形

ることができる．そして、式 (3.19) は、深水、吉田らや中島らが側抑制モデルに補正を加えて説明していた、純粋な静電場とは違う遮蔽などの特殊な性質を、点  $P$  から見えない部分を避けるというデジタル画像上での走査を行うことによって比較的簡単に再現している．

### 3.4 生理学的な根拠

誘導場は心理学的実験から推定される概念である．この誘導場の存在を裏づける直接的な生理学的根拠は見つかっていない．しかし、誘導場の存在を示唆する生理学的現象が幾つか報告されている．

#### 3.4.1 本川の網膜誘導の場

横瀬の視覚の誘導場の生理学的根拠として最も引き合いに出されるのが、本川の発見した網膜誘導の場である<sup>24, 46)</sup>．本川は被験者の眼に光を照射した後、眼の付近に接着した電極により 0.1msec で電気刺激を行うと電極間の電位が高まることを発見した<sup>70)</sup>．本川はフィルムの感度に対応して、網膜の光に対する反応を感電性と呼び、電位の高まりで表現した<sup>70, 71)</sup>．すなわち 15 分以上暗順応させた被験者の眼の付近にか

かる電圧を  $E_0$  , 次に2秒間前照射した際の電圧を  $E$  とすると、感電性  $\eta$  は次式で与えられるものとした。

$$\eta = \frac{E - E_0}{E_0} \times 100 \quad (3.20)$$

本川は感電性  $\eta$  は照射する光の波長によって異なることを見出した。この感電性の性質を用いれば、例えば、被験者に図形のある部分を最初黄色で前照射し、次に同じ場所を白色で照らし電気刺激を加えて感電性を測定すれば、黄色の補色である青色を照射した場合に相当する感電性が得られると考えられる。この原理を用いて、例えば円を見たときの感電性を調べることができる。具体的には、黒い背景に円をくり抜いた刺激を黄色で照射、次に白色を照射する<sup>70)</sup>。本川はこの原理を用いて、色々なパターンを提示した際の感電性を求め、等しい感電性の強さからパターンの周りに電磁場のような場(網膜誘導の場)があることを示した<sup>70, 72, 71)</sup>。

更に、本川は Müller-Lyer などよく知られた刺激図形を用いて、鯉の網膜上の電位を測定し、電位がゼロの等電位線が場の分布になっていることを示した<sup>73, 74)</sup>。以上の実験から、本川は誘導の場は網膜で引き起こされる現象であると主張した。

### 3.4.2 本川の研究に対する議論

本川の感電性の実験は、一般には視覚誘発電位として知られている。現在もいろいろなパターン<sup>75)</sup>や線画の顔<sup>76)</sup>を見た場合などの視覚の働きを調べるため、視覚誘発電位を使った実験が行われている。しかし、どのようなパターンを見せると誘発される電位が大きいかなどがわかる程度であり、本川の測定結果に疑問をもつ生理学者も多い<sup>47)</sup>。また、本川の実験は測定に時間がかかるため、実験中に被験者の注視点をどのように維持させたかも疑問がある。

本川は、感電性の実験と同様、鯉網膜で場が測定されたことから、網膜で場が生じていると考えた。しかし、視覚誘発電位は網膜周辺で生じる総合的な現象であって、網膜自体の場ではないと考えられる。したがって、場は網膜で生じていると断言するのは疑問が残る<sup>47)</sup>。

本川の網膜誘導の場と視覚の誘導場は同じものとする立場<sup>注5)</sup>もあるが、果たしてそれが正しいか否かについては問題がある。例えば、横瀬は白色の小光点、本川は多種類の色フィルタ-による散乱光という全く異なる光源を測定に用いていることから、野澤は両者の場が別の効果によるものであるという可能性を指摘している<sup>47)</sup>。

しかしながら、本川は波長による感電性の違いを利用した独自の 방법으로、慎重に網膜誘導を測定している<sup>注6)</sup>。いろいろと疑問や問題はあがるが、本川によって Köhler 以来、心理学的な概念に過ぎなかった「場」の存在が、生理学的に示唆されたことは大変意義があると考えられる<sup>47)</sup>。

注5 文献<sup>46)</sup>では、本川の網膜誘導の場は視覚の誘導場の生理学的測定方法の1つとしてとらえられている。

注6 感電性の実験を行う日は、本川は朝早くから実験する暗室に入り、実験前に充分眼を暗順応させ、実験中は食事、トイレも全て暗室で済ませ、眼の状態を常に一定に保つよう細心の注意を払って実験していたという(聖心女子大 野澤先生談)。

### 3.4.3 内山らの実験

内山らは本川の鯉網膜上の電位分布が、視覚の誘導場に似ていることに注目し、本川の鯉の実験の追試を行っている<sup>44, 77)</sup>。内山らは、本川のように図形を単純に網膜前に提示するのではなく、横瀬の心理実験と同様な小光点を提示し、網膜の表面電位の変化を調べた。その結果、横瀬が心理実験で得た視覚の誘導場と非常に類似した分布が網膜の表面電位から得られた<sup>77)</sup>。このことから、本川の測定した場は視覚の誘導場とは別の現象である可能性が高いこと、また、視覚の誘導場に似た現象が網膜周辺で生じている可能性のあることがわかった。更に内山らは、同様な実験で網膜内部の電位を調べたところ、場の分布の強さは、表面電位の場合に比べ、図形の中心からの距離の効果が、網膜の深いところほど強く現われることを確かめている。これは、網膜周辺および網膜から更に高次の処理をする器官で、視覚の誘導場を生ずるような現象が起きている可能性を示唆するものと考えられる。

### 3.4.4 杉江らの実験

杉江らは、猫の眼にパターン(□, △, 立方体)を提示した場合の外側膝状体(lateral geniculate nucleus)におけるポテンシャルを電気生理的に詳しく求めた<sup>78)</sup>。その結果、提示したパターンの輪郭周辺やエッジ部分のポテンシャルが強くなっている。これは、本川の鯉網膜における網膜誘導の場のような現象が、網膜の情報を大脳に中継すると考えられている外側膝状体においても生じていることを示唆している。以上から、心理学的な視覚の誘導場に関連する現象が、生理学的には網膜周辺で生じていることが示唆される。

以上の実験結果は誘導場の存在を生理学的に直接示すものではない。しかし、誘導場のような現象、機構が脳内に存在する可能性を示唆している。今後、更に研究が進んで誘導場の生理学的実体が解明されれば、誘導場の分布やその利用について大きな進展が期待される。

## 3.5 誘導場による視知覚現象の解釈

最後に、誘導場によってどのような視知覚現象を定量的に説明できると考えられているかについて簡単にまとめておく。

### パターン認知

横瀬は誘導場を提案した当初より、誘導場によるパターン認知の可能性を示唆している<sup>44)</sup>。また、心的回転や年齢による認知の違いなども同様に誘導場で説明できる可能性を示唆している<sup>44)</sup>。ただし、誘導場による具体的な認知機構は明らかにされていない。

## 2次元パターンの広がりの定量化

平面上にある図形の適切な配置<sup>79)</sup>、表札における文字などの適切な配置<sup>80)</sup>、橋など建造物におけるアーチのデザインの見誤り易さ<sup>81)</sup>など、これまで主観的だった平面上のパターン配置に関する評価を誘導場を使って定量化できることが示されている。今後、評価するパターンや図形に応じて、誘導場の分布のどのような側面を評価するかを工夫することにより、多方面で誘導場を2次元パターンの広がりの定量的評価に利用できると考えられる。

## 錯視

誘導場のあるポテンシャル値の分布を調べることで、Müller-Lyerの錯視図形の長さの説明<sup>82, 83)</sup>やポツゲンドルフ錯視<sup>66, 84)</sup>など、実体のある図形の錯視量の定量的説明ができることが示されている。

また、主観的輪郭のような実際にパターンが存在しないような例は直接的な説明はできていない<sup>24)</sup>。ただし、主観的輪郭を引き起こすパターンの位置や背景などを変えると錯視の見え方がかなり異なる。こうした動的な主観的輪郭の見え方の違いを誘導場の分布を利用して説明できる可能性が示唆されている<sup>85)</sup>。

## 第4章 誘導場によるパターン認識モデル

### 4.1 従来のパターン認知モデルの研究

文字パターンの認知過程の数理モデル化は、人間の優れた文字認知過程を解明する上で重要であり、コンピュータを用いて文字認識などへ応用する上でも大切である。これまで多くの数理モデルや心理実験が検討されてきた。以下、主な研究について対象パターンごとに概説する。

#### 4.1.1 人工的な線図形

Townsendらは人工的な線図形を用いて、アルファベット4文字と極めて小数ではあるが、数理モデルによる検証手段を初めて明確にした<sup>86)</sup>。以降、人工的な線図形を対象に、人間の文字認知特性を記述する数理モデルが検討されている。代表的な数理モデルとして、最初に、例えば図4.1の図形のような一本一本の線分（小線分）が知覚単位として抽出され、次に、小線分が組み合わされた、より複雑な特徴が抽出される特徴蓄積モデル<sup>19)</sup>がある。例えば、Palmerは図4.1の標準図形Sと比較対象H、Lの8組について、SとH間、SとL間の類似度を心理実験により測定し、HはLよりもSとの類似性がmatched, unmatched共に高いことを明らかにした<sup>87)</sup>。図4.1のSとH、Lは、3つで1組を構成し、H、Lは、6セグメントから成るSと5セグメントを共有、HはSと接続性、閉合性が一致しているが、LとSは一致しない。これらの組のうち、Sと共有しないH、Lの1セグメントが、Sを基準にした場合、距離、長さ、回転角の絶対値が等しい組をmatched、異なる組をunmatchedの2つに分類する。Palmerの実験結果から、個々のセグメントが独立して知覚されるのではなく、更に高次の構造単位を含むことが示唆されている<sup>87)</sup>。これは、線図形が小線分の単純な和として知覚されないことを示している。

一方、特徴蓄積モデルとは逆に、知覚は最初に全体、次に部分が知覚されるというゲシュタルト的過程を仮定する全体部分モデル<sup>19)</sup>がある。例えば、Lupkerは知覚を焦点決めの過程と考え、まず実際の刺激よりぼけた図形が知覚され、次第にそれが鮮明になって知覚できると仮定した。具体的には、例えば、図4.2のパターン7、8や11、12は、ぼかした場合大きな形状となるので混同しやすい。また、図4.2のパターン5は上、パターン6は下に重心があるので混同しにくい。Lupkerは、このような特性を、図4.2のパターン*i*とパターン*j*間の類似度 $L_{ij}(i, j = 1, 2, \dots, 12)$ を心理実験から求め、 $L_{ij}$ が特徴蓄積モデルよりもうまく説明できることを示した<sup>88)</sup>。

以上の研究は、非常に単純なアルファベットの認識特性の記述に限定されている。これに対し、横澤らは多種類で複雑なパターンとして、漢字の数理モデルを検討している<sup>33)</sup>。このモデルは、セグメントで構

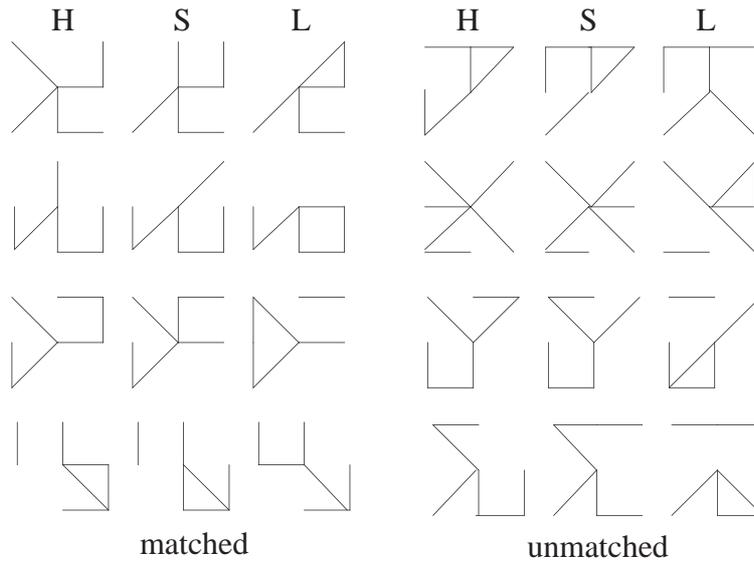


図 4.1 類似した線図形 (Palmer, 1978)

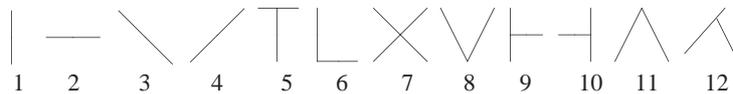


図 4.2 単純な線図形 (Lupker, 1979)

成された漢字パターンの各セグメントが独立に知覚され、更にセグメントに接続や閉図形などの構造があると、その単位ごとに従属的に知覚する知覚段階と、認識候補を知覚パターンと候補パターン間とのセグメントの共有状態により選択する決定段階から構成される<sup>33)</sup>。そして、図 4.3 のようにパターンを 9 個ずつ 3 set に分け、set1 の任意のパターン  $i$  をパターン  $j$  と誤読しない認識率  $Y_{1ij}[\%](i, j = 1, 2, \dots, 9)$ 、同様に set2, 3 の認識率  $Y_{2ij}, Y_{3ij}$  を心理実験より求め、提案した数理モデルで、複雑な構造をもつ漢字の認識特性をうまく説明できることを示した<sup>33)</sup>。このように、これまで各心理実験で使用されたいろいろな種類、構造をもったパターンの類似性を説明するには、パターンの種類ごとに、それぞれ全く異なった立場に基づく個々の数理モデルが必要である。したがって、1つの数理モデルで各実験結果を統一的に説明することはできない<sup>19, 29)</sup>。

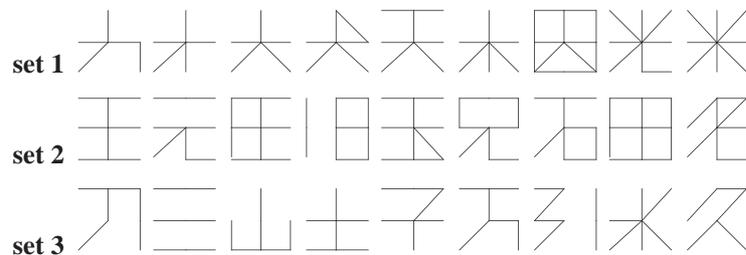


図 4.3 漢字パターン (横澤, 梅田, 淀川, 1986)

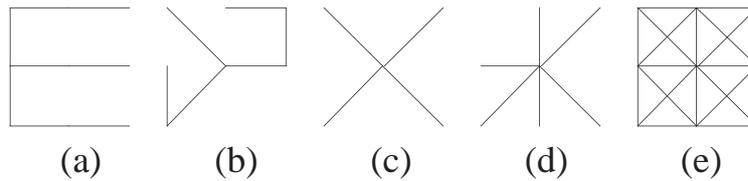


図 4.4 基本的な直線セグメントで構成できる線図形

### 4.1.2 活字

線図形を対象とした数理モデルの研究では、図 4.4(a), (b), (c) および (d) の例のような、いろいろなパターンが扱われているが、ほとんどのパターンは、図 4.4(e) の 20 個の直線セグメント<sup>33)</sup> の組合せで構成される幾何学的な人工パターンである。そして、各数理モデルで有効性が示唆された特徴は、基本的に図 4.4(e) のセグメントの組合せで表現できることから、各数理モデルで有効な特徴は、そのモデルが対象としているパターンの構成要素に還元される<sup>89)</sup>。そのため、各モデルで示された特徴を、我々が日常使用する活字のアルファベットや仮名などの文字パターンの場合に一般化することは難しい<sup>19, 89, 90)</sup>。

そこで、人工的な幾何学パターンではなく、我々が日常使用する文字パターンの特徴を心理実験から直接明らかにする研究が行われている<sup>89, 90)</sup>。例えば、Townsend はタイプライタで打ち出された大文字のアルファベットをタキストスコープで瞬間提示し、アルファベット 26 個 (1, 2, ..., i, ..., j, ...26) の  $i$  番目を  $j$  番目と読み間違え率  $To_{ij}$  を求め、共有する特徴が多い文字は間違いやすいことを示した<sup>91)</sup>。Kuennapas らは、Townsend と同様な実験を行い、w を除く活字のアルファベット小文字 25 個 (1, 2, ..., i, ..., j, ...25) の  $i$  番目と  $j$  番目の類似度  $Ku_{ij}$  を求めた。そして、 $Ku_{ij}$  の因子分析を行い、矩形性、垂直線分などの特徴が有効であることを明らかにしている<sup>92)</sup>。菊地らは発光ダイオードで表示されたドットマトリックスのアルファベット大文字 26 個 (1, 2, ..., i, ..., j, ...26) の  $i$  番目を  $j$  番目と読み間違える割合  $KI_{ij}$  の傾向を階層クラスタ分析し、V, O 字形成分、斜め交差などの基本特徴を明らかにしている<sup>93)</sup>。田中はカタカナ 46 個 (1, 2, ..., i, ..., j, ...46) の  $i$  番目を  $j$  番目と読み間違える頻度  $TA_{ij}$  を調べ、文字の一部を移動または欠損するとほぼ同一の文字になる文字 (ラ, テなど) は見誤りやすいことを示した<sup>89)</sup>。村石らは楷書体の平仮名 46 個 (1, 2, ..., i, ..., j, ...46) の  $i$  番目を  $j$  番目と読み間違える頻度  $MU_{ij}$  を調べ、「ぬ」「ね」「め」など構造が類似した文字が間違いやすいなどの混同傾向を明らかにしている<sup>94)</sup>。以上のように、文字パターンの類似性や読み間違えを説明するのに有効と考えられる、いろいろな特徴の存在が示唆されているが、それらの有効性は充分確められておらず、我々が日常使用するアルファベットや仮名など、文字数が多く複雑な文字パターンの心理実験結果を説明する数理モデルの構築は難しい問題である<sup>89, 90)</sup>。

### 4.1.3 変形パターン

人間には手書き文字のように、手書き変形のため同一カテゴリでも形状がかなり異なり、どのカテゴリに所属するのかあいまいな変形パターンであっても、同じカテゴリであることが正確に識別できる優れた能

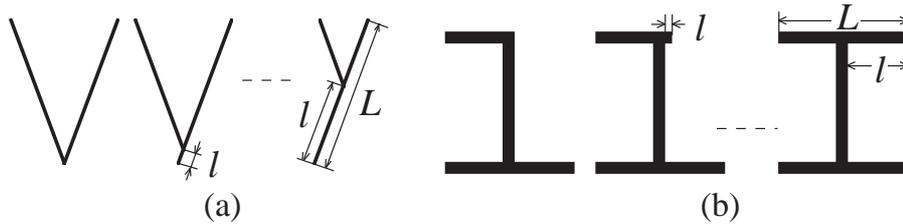


図 4.5 leg; 脚の長さの定義

力がある<sup>89, 90, 95</sup>。

そこで、文字線分の一部を移動または伸縮して、あるカテゴリから別のカテゴリに連続的に変化させた変形パターンの類似性を心理実験から求め、変形パターンを区別する機能的特徴を明らかにする研究が行われている<sup>90, 89</sup>。例えば、Shillmanらは図4.5(a)のように、直線で構成されたVを徐々に6段階でyに変化させたV→yの変形系列において、*i*番目の変形パターンをVと読む程度 $S1_i$  ( $1, 2, \dots, i, \dots, 7$ )、同様に変形系列C→FでC, U→HでUと読む程度 $S2_i, S3_i$ を求めた。 $Sj_i$  ( $j = 1, 2, 3$ )の結果から、パターン変形に伴って、初めはほとんど元のパターン(V, C, U)と読まれるが、変形のある時点でほとんど別のパターン(y, F, H)と読まれるようになる臨界点があることを示した<sup>96</sup>。そして、臨界点における図4.5(a)のV→yの例のような、変化させた線分の変化している部分の長さ*l*と全体の長さ*L*の比*l/L* (leg; 脚の長さ)は、多くのアルファベットで共通であることを示した<sup>97</sup>。

同様な結果が、Vを15段階でXに変化させ、*i*番目のパターンをVと読む率 $Ya_i$  ( $1, 2, \dots, i, \dots, 16$ )を求めた心理実験からも得られている<sup>98</sup>。田中は、Shillmanらと同様の心理実験から、例えば「コ」を10段階に「エ」に変化させた図4.5(b)のような15組のカタカナの変形系列(ラ→ヲ, テ→チのみ7段階)において、*j*番目( $1, 2, \dots, j, \dots, 15$ )の変形系列の*i*番目( $1, 2, \dots, i, \dots, 11$ )のパターンを、変形させた元のカタカナと読む率 $Tj_i$ を求めた。その結果、カタカナもアルファベットと同様な臨界点があるが、図4.5(b)のように計算したlegは文字種によってさまざまに共通でないことから、legはすべての文字共通の機能的特徴ではないことを示した<sup>89</sup>。一方、横瀬は文字線分の一部が伸縮、省略されて別のパターンに連続的に変形させた場合について、E, F, H, Lを4または5段階に変化させた4つの変形系列において、*j*番目( $1, \dots, j, \dots, 4$ )の変形系列の*i*番目( $1, \dots, i, \dots, 4, 5$ )のパターンを元パターンと読む率 $Yoj_i$ を調べた。更にFを20段階「ニ」を12段階、Tを10段階に変化させた*i*番目のパターンをそれぞれF「ニ」、Tと読む率 $Yol_i$  ( $1, \dots, i, \dots, 21$ ),  $Yo2_i$  ( $1, \dots, i, \dots, 13$ ),  $Yo3_i$  ( $1, \dots, i, \dots, 11$ )を調べた。しかし、カテゴリの臨界点やlegのような特徴は明らかにされていない<sup>44</sup>。

以上のように、いろいろな心理実験による変形パターンの類似性が調べられているが、変形パターンを区別する一般的な機能的特徴は明らかにされていない。また、legのような特徴は2つのカテゴリ間で定義される特徴であるため、そのまま複数のカテゴリから正解を見つける、通常のパターン認識モデルの特徴に拡張するのは難しいと考えられる<sup>89</sup>。したがって、所属するカテゴリがあいまいな変形パターンの認知に関する数理モデルはほとんど検討されていない<sup>29, 89, 90</sup>。



図 4.6 文字サンプル例：東

#### 4.1.4 手書き文字

手書き文字の数理モデルとして、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<sup>99)</sup>。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。そこで、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きた。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<sup>100)</sup>、岡本らは文字の外形（多角形近似）、外接矩形の重心、大きさ、面積、慣性軸など<sup>101)</sup>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<sup>102)</sup>。

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<sup>103)</sup>。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よく使われる、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 ETL9<sup>104)</sup> 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<sup>105)</sup>。

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<sup>37, 106)</sup>。加藤らは ETL9 から文字種ごとに無作為に選択した、図 4.6 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  $j(1, 2, \dots, j, \dots, 100)$  における 25 パターン内の任意のパターン  $i(1, 2, \dots, i, \dots, 25)$  の文字品質の評価値  $K_{j_i}$  を求めた。図 4.6 の各パターンの下に数字は「東」の評価値  $K_{j_i}$  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、加藤らは評価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、1回目と2回目の評価結果間の相関について、各被験者の平均が0.75と高いことから、主観評価の再現性は高く、文字品質の主観評価実験で得られた10名の平均値  $K_{j_i}$  を各文字種の主観評価値とすることが妥当であることを示した<sup>37)</sup>。

更に、加藤らは2500パターンの文字品質  $K_{j_i}$  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が示された。しかし、100文字種の文字品質を複数の特徴を組み合わせた1つの評価式で評価することは不可能であった<sup>37, 106)</sup>。このように、文字品質によって手書き文字の変形を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。

## 4.2 新しい視覚の誘導場によるパターン認識モデルの提案

### 4.2.1 ダイナミクスと場によるモデル

このように、いろいろな文字パターン認知に関する数理モデルの検討や人間の類似性判断などの心理実験が行われている。しかし、これまで提案された数理モデルで説明できるパターン種は限られており、いろいろなパターンの類似性を統一的に説明する数理モデルはまだ提案されていない。

今井は、これまで提案された文字パターン一般に有効な特徴の決定やパターンの特徴を用いる数理モデルの構築が困難なのは、パターン認知過程が、例えば2.2.2で述べたバンデモニアム・モデルのような、パターンから「特徴」を抽出、分析する過程ではなく、一方のパターンが他方のパターンに一致するような変換をパターンに施す過程、すなわち、ダイナミクスであることによるとしている。今井は、具体的なダイナミクスとして変換構造説を提案している<sup>34)</sup>。従来の視知覚に関する知見<sup>24, 35, 40)</sup>、年齢別変形文字の読まれ方の研究<sup>44)</sup>、パターン変換作用による視覚恒常性の説明<sup>107)</sup>などが、今井と同様に、パターン認知はダイナミクスであることを示唆している。

また、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<sup>108)</sup>、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<sup>109)</sup>などが、手書きなど複雑な文字パターンの認知機構にダイナミクスが存在する可能性を示唆している。これらの研究から、ダイナミクスによるパターン認知モデルはいろいろな文字パターンを実際に説明できる可能性が高いと考えられる。

そして、2.3で述べたように、最近、文字パターンの認知などの視知覚を統一的に説明する枠組として、「場」が注目されている。その「場」として、2.4で述べたように、図形の周囲に波及する「場」を考えることにより、視覚現象を説明する横瀬の視覚の誘導場があり、パターンの類似性が評価などの可能性が示唆されている。したがって、ダイナミクスによる誘導場のパターン認識が数理モデル化できれば、このモデル

によってこれまで研究されてきたパターン認知の心理実験結果を統一的に説明できる可能性が高いと考えられる。

#### 4.2.2 誘導場によるパターン認識モデル

視覚の誘導場理論で明らかにされているのは、「パターンが、どのように見えるか」だけであり、パターン認識がどのように行われるのかに関してはほとんどわかっていない<sup>24, 35)</sup>。しかし、横瀬は誘導場の等ポテンシャル線の分布の類似性が、パターンを見分ける上で重要なことを示している<sup>44)</sup>。ことから、等ポテンシャル線の分布の違いが評価できれば、誘導場を使ってパターンを認識できる可能性がある。

等ポテンシャル線の分布の評価法として、まず、場を重ね合わせて類似度を求める鑄型照合<sup>29)</sup>が考えられる。しかし横瀬は、片仮名やあるアルファベットの線分の一部を分離するなどして、別のアルファベットに変化させた変形パターンの読まれ方の研究から、等ポテンシャル線の分布の部分的な類似性の違いだけでなく、全体的な分布の違いも評価できなければならないこと、そしてこれは、パターン類似性がゲシタルト心理学のパターンの良さと関連があり、パターンの良さは誘導場の分布で説明できることを示唆している<sup>44)</sup>。ところが、鑄型照合は「場」の強度を座標点に依存する独立変数と考えるため、基本的に分布の微視的な違いしか評価できず、横瀬の示唆する全体の違いやパターンの良さを評価するには問題がある。

一方、横瀬はせん断変形を受けたパターンの読まれ方の研究から、誘導場が一致するようにトポロジカルな変形を行い、同相ならばそのパターンは類似である可能性を示唆している<sup>44)</sup>。これは、「場」におけるダイナミクスによる認識モデルの存在を示唆するものと考えられる。実際、パターン認識などの視知覚を説明する枠組として、ダイナミクスの重要性が指摘されており、「場」はダイナミクスの力学的概念として捉えられている<sup>24, 35, 40, 108, 109)</sup>。特に今井は、ゲシタルト心理学のプレグナンツの原理<sup>注1)</sup>におけるパターンの良さを定量化するという観点から、具体的な認識モデルとして道筋モデルを提案し、パターンの類似性がパターンの良さと関連することを示している<sup>34)</sup>。道筋モデルは、パターン相互の変換可能性の程度によってパターンの類似性を評価するモデルで、パターンの特徴を使わずに1次元点列パターンの類似性を定量的に説明できることが検証されている<sup>110)</sup>。したがって、今井の道筋モデルを誘導場に応用すれば、横瀬が示唆した誘導場の等ポテンシャル線の分布の類似性を評価できる認識モデルを実現できる可能性が高いと考えられる。

道筋モデルは、パターン A に変換を行ってパターン B に一致するまでの変換の手続き（道筋の長さ）をパターン A, B 間の類似性の尺度とする。すなわち、パターン A とパターン B の一致に多くの変換が必要ならば、パターン A, B 間の類似性は低く、変換が容易ならば類似性が高い<sup>34)</sup>。変換を等ポテンシャル線の分布の変位、道筋を分布を一致させるのに必要な変位量とおけば、誘導場の変形による次のパターン認識モデルが考えられる。まず、図 4.7(a) の未知パターン A の誘導場と、複数のパターン B, C, D... の誘導場を準備し、A を変形して他の誘導場と一致するのに必要な変位を求める。図 4.7(b) の A と B, C, D... 間

注1 図、地分化や群化などで、人間が視野内のパターンが最も簡潔で秩序あるまとまりを形成するように認知する傾向。

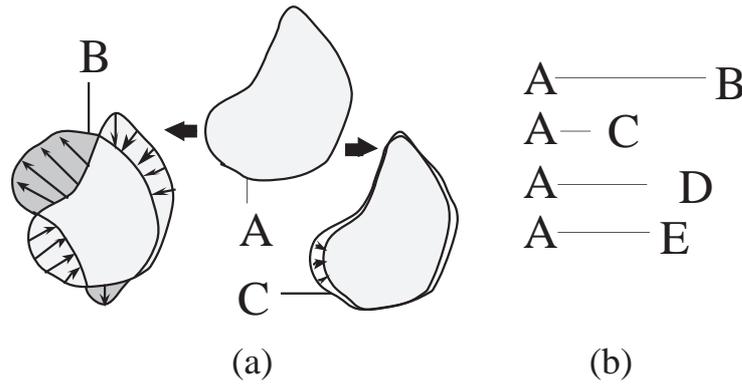


図 4.7 誘導場によるパターン認識モデルの概念

の線分の長さは変位量を示す．例えば，A と C は分布が非常に似ており，少ない変換で一致させられるため，A と B よりも変位量が少ない．したがって，A と C の間の変位量が一番小さい誘導場のパターンが認識結果となる．

### 4.2.3 認識アルゴリズム

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ような弾性体のアナロジで考えることができる．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を用いて評価する．

図 4.8 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し，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．弾性エネルギー  $e_i(p)$  は，輪郭線に張られた膜が平衡状態からある位置まで変位<sup>注2</sup>するエネルギーである，膜の変形に要するエネルギー  $U_1$  と，膜の縁である輪郭線の変形に要するエネルギー  $U_2$  の和である<sup>111)</sup>． $U_1$  は変位を  $u(x, y)$ ，膜の張力  $\tau$ ， $\alpha = \tau/2$  とすれば，

$$U_1 = \alpha \iint \left\{ \left( \frac{\partial u}{\partial x} \right)^2 + \left( \frac{\partial u}{\partial y} \right)^2 \right\} dx dy \quad (4.1)$$

となる． $U_2$  は，膜の張力  $\tau$ ，密度  $\rho$ ，弾性係数は一様で  $C$ ， $\beta = (\tau^2/\rho^2)C$  とすると次式で与えられる．

$$U_2 = \frac{\beta}{2} \iint \left\{ \left( \frac{\partial^2 u}{\partial x^2} \right)^2 + \left( \frac{\partial^2 u}{\partial y^2} \right)^2 \right\} dx dy \quad (4.2)$$

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<sup>注3</sup>．

$$E_{ij} = \sum_p e_i(p) \quad (4.3)$$

注<sup>2</sup>変位の求め方は付録 A を参照．

注<sup>3</sup>差分形式などの詳しい計算方法は付録 B を参照．

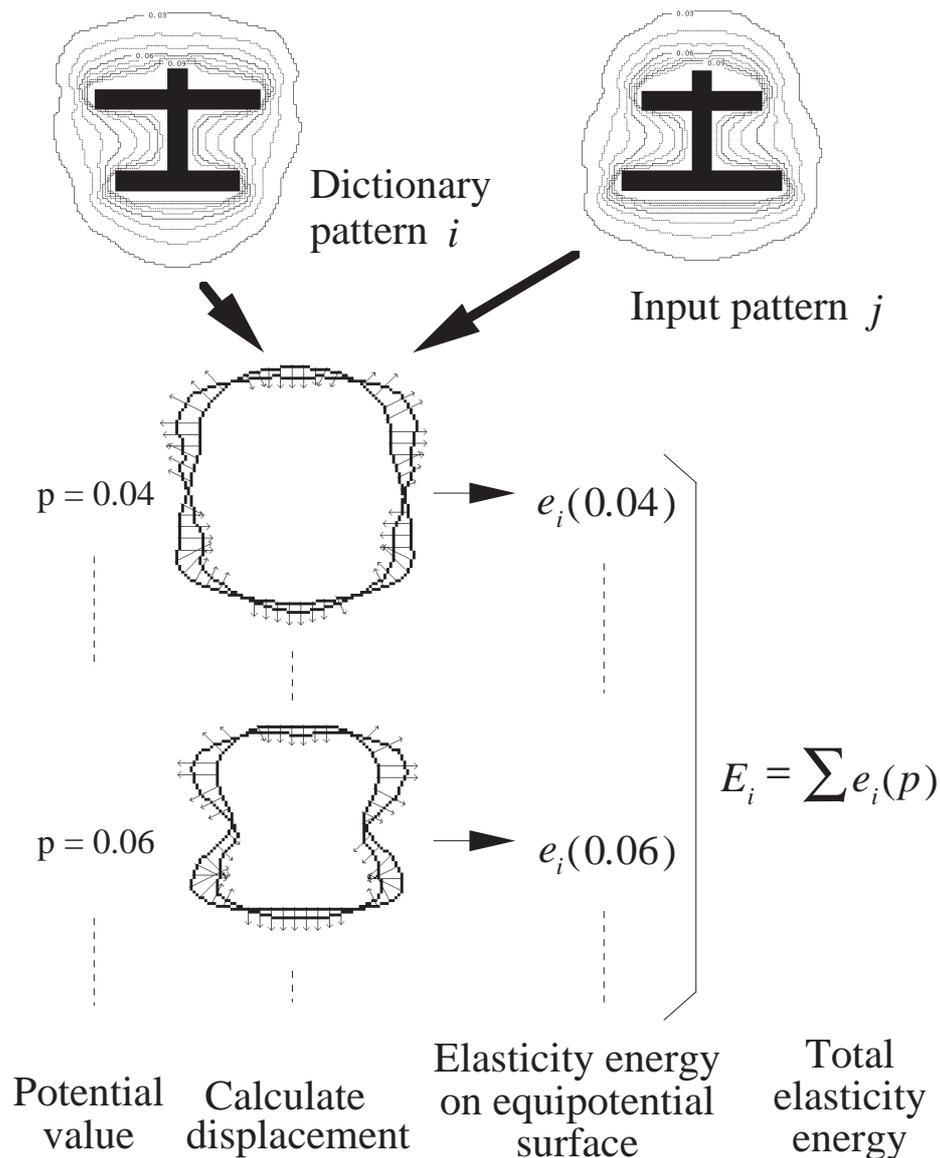


図 4.8 認識過程

### 4.3 検証方法

提案した誘導場のパターン認識モデルが、パターン認知の心理実験結果を、統一的に説明できることを検証するには、複数の心理実験結果と、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。ここでは、人工的な線図形、活字、変形パターン、そして手書き文字ごとに検証を行った。まず、検証の対象とした心理実験および実験パターンについて説明する。次に、各実験パターンについての心理実験結果と、それと比較する誘導場の弾性エネルギーなど工学的評価方法の計算の仕方、変数などについて説明する。最後に心理実験結果と弾性エネルギーなど工学的評価方法との相関の求め方について説明する。最初に人工的な線図形を例に基本的な検証方法を説明する。続いて活字、変形パターン、手書き文字の検証方法を述べる。

### 4.3.1 人工的な線図形

Townsend ら<sup>86)</sup> や横瀬<sup>44)</sup> が人工的な線図形について行った心理実験では、類似性が比較されたパターンが数組程度と極めて少ない。特に横瀬の実験では、変形パターンを変形順に提示してアンケートで可読率を求めているため、文脈性の排除、被験者間のばらつきなどの考慮が不十分である。これらの実験結果は、誘導場モデルの妥当性を検証するには問題がある。ここでは、実験で使われたパターンが多く、それぞれを説明する数理モデルの立場が異なり、パターンの提示条件などが充分考慮されている、Lupker<sup>88)</sup>、横澤ら<sup>33)</sup>、Palmer<sup>87)</sup> の心理実験結果を検証に用いた。

これらの実験で使われた図 4.1, 4.2, 4.3 のパターンの違いを誘導場モデルの弾性エネルギーで評価する。例えば、Lupker の実験の場合、図 4.2 のパターンを  $64 \times 64$  dot の大きさに作成し、パターンの領域を含む  $128 \times 128$  dot の範囲で誘導場を計算する。そして、図 4.2 のパターン  $i, j$  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 から文字線分付近の 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  $e_i(p)$  を式 (4.1), (4.2) より求め、その総和  $E_{ij}$  を図 4.2 のパターン  $i, j$  間の弾性エネルギー  $LE_{ij}$  とする。なお、現時点では弾性エネルギーにおける  $U_1$  と  $U_2$  の最適な割合が不明なため、ほぼ同じパターン間のエネルギーを計算した場合、 $U_1$  と  $U_2$  がほぼ等しい値になるよう、式 (4.1), (4.2) で  $\alpha = \beta/2 = 1.0$  とおいた。以上の処理方法や条件は、以降の実験においてもすべて同じである。

誘導場モデルの弾性エネルギー  $LE_{ij}$  の比較対象として、まず「場」の概念を利用した文字認識で最もよく使われる、「場」の強さを Euclid 距離で評価する方法<sup>112)</sup> を適用した。図 4.2 のパターン  $i, j$  の誘導場間の Euclid 距離  $LFD_{ij}$  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)、パターン  $j$  の場の強さを  $M_j(x, y)$  とすると次式で与えられる。

$$LFD_{ij} = \sqrt{\sum_{x,y} (M_i(x,y) - M_j(x,y))^2} \quad (4.4)$$

続いて、 $LE_{ij}$  の比較用として優れた手書き文字認識性能が報告されており、文字認識の研究のリファレンスとして最もよく使われる LDCD 特徴 (Local Direction Contributivity Density Feature)<sup>113)</sup> を Euclid 距離で評価する方法を適用した。ストロークの連結状態を抽出する LDCD 特徴は、ここで検証に用いる図 4.1~図 4.3 のような直線セグメントで構成されるパターンに用いた場合、最も高い識別能力を発揮すると考えられるからである。LDCD 特徴は、 $64 \times 64$  dot の図 4.2 のパターン  $i$  を、 $8 \times 8$  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  $ver(n)$ 、 $+45$  度  $p(n)$ 、 $-45$  度  $m(n)$  の 4 方向の方向寄与度 (Direction Contributivity Density)<sup>113)</sup> 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<sup>注4</sup>。図 4.2 のパターン  $i, j$  間の LDCD 特徴の Euclid 距離  $LLD_{ij}$  は、パターン  $j$  の LDCD 特徴を  $L_j(n)$  とすると、

$$LLD_{ij} = \sqrt{\sum_n (L_i(n) - L_j(n))^2} \quad (4.5)$$

となる。式 (4.4), (4.5) 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。

<sup>注4</sup>LDCD 特徴の詳しい求め方は付録 C を参照。

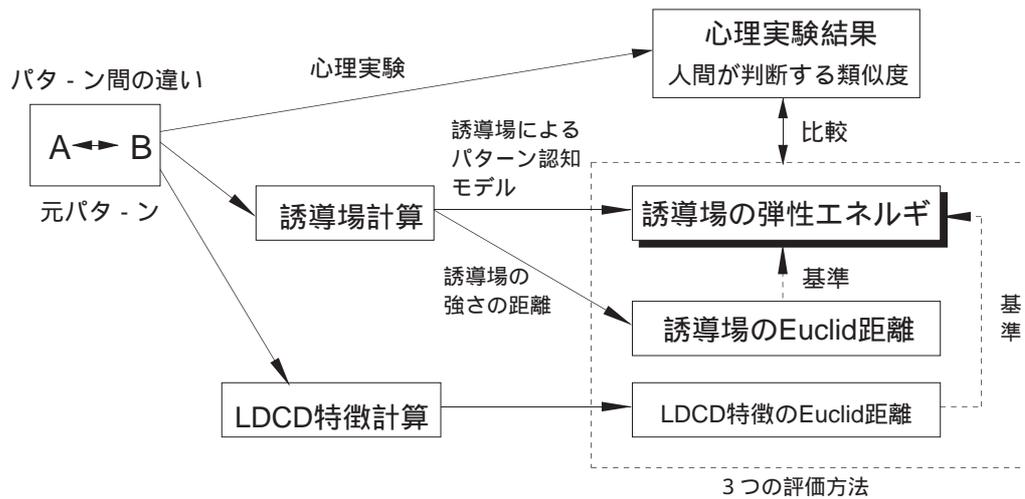


図 4.9 3つの評価方法による検証

表 4.1 相関をとる変数の一覧 (人工的な線図形)

method	pattern			
	Lupker <sup>88)</sup>	set1 <sup>33)</sup>	set2 <sup>33)</sup>	set3 <sup>33)</sup>
experiment	$L$	$Y1$	$Y2$	$Y3$
field(energy)	$LE$	$Y1E$	$Y2E$	$Y3E$
field(Euclid)	$LFD$	$Y1FD$	$Y2FD$	$Y3FD$
LDCD(Euclid)	$LLD$	$Y1LD$	$Y2LD$	$Y3LD$

以上のように求めた誘導場モデルの弾性エネルギー，誘導場の強さの Euclid 距離，LDCD 特徴の Euclid 距離の 3 つの評価方法と心理実験結果との相関を求める．対象パターン，心理実験結果，3 つの評価方法の関係を図 4.9 にまとめた．図 4.9 のように，検証に用いたパターンに対し，心理実験による人間が評価した類似性が視覚の誘導場の弾性エネルギーで説明できるか比較する．その際，誘導場の Euclid 距離と LDCD 特徴の Euclid 距離を比較基準として用いる．活字，変形パターン，手書き文字の場合も同様に検証を行う．

具体的な相関の求め方を述べる．例えば，図 4.2 のパターン  $i$  における Lupker の心理実験による類似度  $L_{ij}$  と弾性エネルギー  $LE_{ij}$  の相関係数を  $R_i(L, LE)$  とする．相関係数は正規分布しない<sup>注5)</sup>ので，12 個の  $R_i(L, LE)$  を正規変換量  $z$  に変換した上で加重平均し，Lupker の全 12 パターンにおける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，図 4.2 のパターン  $i$  における  $L_{ij}$  と誘導場間の Euclid 距離  $LFD_{ij}$ ， $L_{ij}$  と LDCD 特徴の Euclid 距離  $LLD_{ij}$  の相関係数  $R_i(L, LFD)$ ， $R_i(L, LLD)$  から，Lupker の全 12 パターンにおける実験結果と誘導場および LDCD 特徴の Euclid 距離の相関係数の平均を求める．Lupker の実験と同様に，図 4.3 の set1 のパターン  $i, j$  について，誘導場間の弾性エネルギー  $Y1E_{ij}$ ，式 (4.4) から誘導場間の Euclid 距離  $Y1FD_{ij}$ ，式 (4.5) から LDCD 特徴の Euclid 距離  $Y1LD_{ij}$  を求める．そして， $Y1E_{ij}$ ， $Y1FD_{ij}$ ， $Y1LD_{ij}$  と，横澤らの実験結果  $Y1_{ij}$  から，set1 全体について相関係数の平均を求める．set2, 3 も同様である．相関をとる変数の一覧を表 4.1 にまとめた．

注5)付録 E の相関係数の検定を参照．

A B C D E F G H I J  
 K L M N O P Q R S T  
 U V W X Y Z

図 4.10 アルファベット大文字

a b c d e f g h i j  
 k l m n o p q r s t  
 u v x y z

図 4.11 アルファベット小文字

相関をとる変数の名前は次のようにつける．まず，心理実験を行った研究者名の略を心理実験結果とした（例：Lupker L）．次にこの心理実験結果に，誘導場の弾性エネルギーは「E」（例：LE），誘導場の Euclid 距離は「FD」（例：LFD），LDCD 特徴の Euclid 距離は「LD」（例：LLD）を付加したものを変数名とした．活字，変形パターン，手書き文字についても同様な方法で変数に名前を付けた．

パターン  $i, j$  間の類似性が高いほど， $L_{ij}, Y1_{ij}, Y2_{ij}, Y3_{ij}$  は大きく，誘導場モデルの弾性エネルギー，誘導場の強さの Euclid 距離および LDCD 特徴の Euclid 距離が小さくなる．相関が強い評価方法ほど心理実験によるパターンの類似性をよく説明できる．

Palmer の実験の場合，図 4.1 の S と H, L の 8 組，各々における S と H 間, S と L 間について，誘導場モデルの弾性エネルギー，誘導場の強さの Euclid 距離，LDCD 特徴の Euclid 距離を求め，これらの結果を matched, unmatched に分けて平均し，心理実験で測定した類似度（数値が大きいほど類似性が高い）の平均値<sup>87)</sup>と比較した．

### 4.3.2 活字

アルファベット・カタカナ，平仮名など多種類，多数の文字パターンの類似性が調べられており，ランダムに提示するなどパターンの提示条件が充分考慮されている，Townsend<sup>91)</sup>，Kuennapas ら<sup>92)</sup>，菊地ら<sup>93)</sup>，田中<sup>89)</sup>，村石ら<sup>94)</sup>の実験を検証対象に，誘導場モデルの弾性エネルギーで評価した結果と比較した．

各心理実験で提示されたパターンの状態を再現するため，例えば Townsend の実験のアルファベット大文字 26 個に対して，タイプライタで打ち出されたひげのない文字<sup>91)</sup>に近い Times フォントの Bold 体を，文字の大きさを決めるポイント数が 26 個のアルファベット大文字すべて同じで，かつ 26 文字全てが  $64 \times 64$  dot の大きさの外接矩形に収まる最大のポイント数で全文字パターンをパーソナルコンピュータ（Macintosh）上で作成した（図 4.10）．

同様に，Macintosh の 32 種のフォントの内から，Kuennapas の実験のアルファベット小文字 25 個に対して文献<sup>92)</sup>記載の実例に最も近い Geneva フォント（図 4.11），田中の実験のカタカナ 46 個に対して文献

ア	イ	ウ	エ	オ	カ	キ	ク	ケ	コ
サ	シ	ス	セ	ソ	タ	チ	ツ	テ	ト
ナ	ニ	ヌ	ネ	ノ	ハ	ヒ	フ	ヘ	ホ
マ	ミ	ム	メ	モ	ヤ	ユ	ヨ	ラ	リ
ル	レ	ロ	ワ	ヲ	ン				

図 4.12 カタカナ

A	B	C	D	E	F	G
H	I	J	K	L	M	N
O	P	Q	R	S	T	U
V	W	X	Y	Z		

図 4.13 アルファベット大文字 (ドットマトリックス)

<sup>89)</sup> 記載の実例に最も近い Osaka フォント (図 4.12) を選択, 各実験ごとにポイント数を統一し文字パターンを作成した.

菊地らの実験のドットマトリックス (直径 5.1mm の発光ダイオードを, 10cm × 10cm の範囲に中心間距離 6.5mm の間隔で縦 16 個, 横 16 個配列)<sup>93)</sup> 26 個に対して, dot の比率はそのままに 64 × 64 dot の大きさに縮小する. そして, ポイント数が統一された, 英文印刷で一般的な Times フォントの形状になるように Macintosh 上で構成した (図 4.13). なお, 以降の実験において, Macintosh でパターンを作成した際の解像度は全て 72 DPI である.

村石らの実験で, 被験者に提示された楷書体<sup>94)</sup> (手書きの手本) の平仮名 46 個に対して, 現代書道字典 (木耳社刊) の楷書体の見本文字を, 大きさが 64 × 64 dot の外接矩形に収まるようにイメージスキャナ (Ricoh IS-50) の解像度を設定 (400 DPI) して全平仮名を読み取った (図 4.14). 以上のように作成された図 4.10 ~ 図 4.14 のパターンの違いを誘導場モデルの弾性エネルギーで評価する.

図 4.4(e) の基本セグメントの組合せで表現できる人工パターンの場合, パターンの大きさはすべて同じである. 今回の検証実験で用いる図 4.10 ~ 図 4.13 のパターンはポイント数, 図 4.14 のパターンは読み取る解像度を統一している. しかし, 図 4.10 の W と X (図 4.15(a)), 図 4.12 のヤとユ (図 4.15(b)), 図 4.14 の「つ」と「て」 (図 4.15(c)) などの例のように, 文字の大きさが文字種によって異なる場合がある. 通常, 文字認識では, 文字画素ができるだけ重なるように文字の大きさ・位置の正規化を行い, 文字の大きさの違いを吸収する<sup>114)</sup>. そこで, 従来の評価方法である LCD 特徴の Euclid 距離を評価する方法と, 式 (4.4) が示すように画素部分の Euclid 距離も計算するため, 画素の位置ずれを防止する必要がある誘導場の強さの Euclid 距離で, 大きさ・位置の正規化を行った. 具体的には, 図 4.10 ~ 図 4.14 のパターンを, 64 × 64

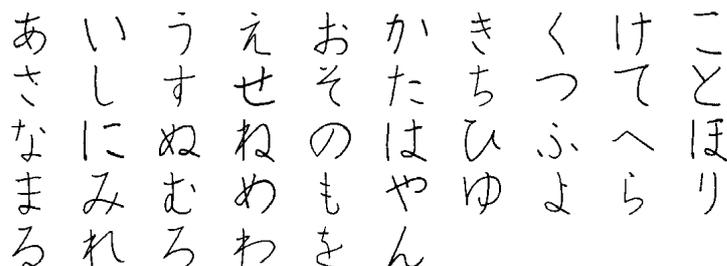


図 4.14 ひらがな

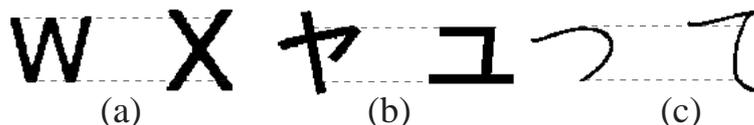


図 4.15 大きさが異なっている例

dot の外接矩形枠に一杯になるようパターンの重心を外接矩形枠の中心に合わせて拡大・縮小したパターンについて、LDCD 特徴や誘導場を計算してから Euclid 距離を求めた。一方、誘導場の弾性エネルギーの計算では、図 4.10 ~ 図 4.14 の状態のまま計算した。これは現時点では、従来の大きさ・位置の正規化が誘導場の弾性エネルギーの場合に有効か不明だからである。

線図形の場合と同様に、誘導場モデルの弾性エネルギー、誘導場の強さの Euclid 距離、LDCD 特徴の Euclid 距離の 3 つの評価方法と、心理実験結果の相関を求める。例えば、図 4.10 のパターン  $i$  における Townsend の実験結果  $To_{ij}$  と誘導場の弾性エネルギー  $ToE_{ij}$  の相関係数を  $R_i(To, ToE)$  とする。人工的な線図形の場合と同様に、26 個の  $R_i(To, ToE)$  を正規変換量  $z$  に変換した上で加重平均し、Townsend の実験の全 26 パターンにおける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、図 4.10 における、Townsend の実験の全 26 パターンにおける実験結果と誘導場および LDCD 特徴の Euclid 距離の相関係数の平均を求める。Townsend の実験と同様に、図 4.11 のパターン  $i, j$  について、誘導場間の弾性エネルギー  $KuE_{ij}$ 、式 (4.4) から誘導場間の Euclid 距離  $KuFD_{ij}$ 、式 (4.5) から LDCD 特徴の Euclid 距離  $KuLD_{ij}$  を求める。そして、 $KuE_{ij}$ 、 $KuFD_{ij}$ 、 $KuLD_{ij}$  と、Kunnapas の実験結果  $Ku_{ij}$  から、図 4.11 の全 25 パターンについて相関係数の平均を求める。同様に、図 4.13 のパターン  $i, j$  について、誘導場間の弾性エネルギー  $KIE_{ij}$ 、誘導場間の Euclid 距離  $KIFD_{ij}$ 、LDCD 特徴の Euclid 距離  $KILD_{ij}$  を求め、菊地らの実験結果  $KI_{ij}$  から、図 4.13 の全 26 パターンについて相関係数の平均を求める。図 4.12 のパターン  $i, j$  については、誘導場間の弾性エネルギー  $TAE_{ij}$ 、誘導場間の Euclid 距離  $TAFD_{ij}$ 、LDCD 特徴の Euclid 距離  $TALD_{ij}$  を求め、田中らの実験結果  $TA_{ij}$  から、図 4.12 の全 46 パターンについて相関係数の平均を求める。図 4.14 のパターン  $i, j$  については、誘導場間の弾性エネルギー  $MUE_{ij}$ 、誘導場間の Euclid 距離  $MUFD_{ij}$ 、LDCD 特徴の Euclid 距離  $MULD_{ij}$  を求め、村石らの実験結果  $MU_{ij}$  から、図 4.14 の全 46 パターンについて相関係数の平均を求める。以上、相関をとる変数の一覧を表 4.2 にまとめた。

表 4.2 相関をとる変数の一覧（活字）

pattern	psychological experiment	field (energy)	field (Euclid)	LDCD (Euclid)
Townsend <sup>91)</sup>	$To_{ij}$	$ToE_{ij}$	$ToFD_{ij}$	$ToLD_{ij}$
Kunnapas <sup>92)</sup>	$Ku_{ij}$	$KuE_{ij}$	$KuFD_{ij}$	$KuLD_{ij}$
菊地 <sup>93)</sup>	$KI_{ij}$	$KIE_{ij}$	$KIFD_{ij}$	$KILD_{ij}$
田中 <sup>89)</sup>	$TA_{ij}$	$TAE_{ij}$	$TAFD_{ij}$	$TALD_{ij}$
村石 <sup>94)</sup>	$MU_{ij}$	$MUE_{ij}$	$MUFD_{ij}$	$MULD_{ij}$



図 4.16 アルファベットの変形パターン 1 (R.J. Shillman 1976)

### 4.3.3 変形パターン

アルファベット・カタカナなど多種類の变形パターンの類似性が調べられており，ランダムに提示するなどパターンの提示条件が充分考慮されている，Shillman ら<sup>96)</sup> (図 4.16)，Yasuhara ら<sup>98)</sup> (図 4.17)，田中<sup>89)</sup> (図 4.18)，横瀬<sup>44)</sup> (図 4.19，図 4.20) の実験パターンを検証対象に用いた．

図 4.16 ~ 図 4.20 の各パターンは，各心理実験で使われたパターンの状態をそれぞれ文献<sup>96)98)89)44)</sup> の記述に基づいてパーソナルコンピュータ (Macintosh) で再現し，すべてが  $64 \times 64$  dot の大きさの文字画像になるよう作成した．そして，各パターン間の違いを誘導場モデルの弾性エネルギーで評価する．

これら変形系列内の類似性の比較は，図 4.16，4.18，4.19 の場合，各系列の一番左のパターンを基準 ( $i = 1$ ) に，同じ系列内の他の任意のパターン  $i$  と比較する．同様に，図 4.17 の場合は左上の V を基準に他の 15 パターンと比較する．図 4.20 の 3 つの変形系列 (set1 ; E，set2 ; 「ニ」，set3 ; T) の場合，各系列で reference と記されたパターンを基準に同じ系列内の他のパターンと比較する．

線図形の場合と同様に，誘導場モデルの弾性エネルギー，誘導場の強さの Euclid 距離，LDCD 特徴の Euclid 距離の 3 つの評価方法と，心理実験結果の相関を求める．例えば，図 4.16 の  $j$  番目の変形系列のパターン  $i$  における Shillman らの実験結果  $S_{ji}$  と誘導場の弾性エネルギー  $S_j E_{1i}$  の相関係数を  $R_j(S_j, S_j E)$  とする．人

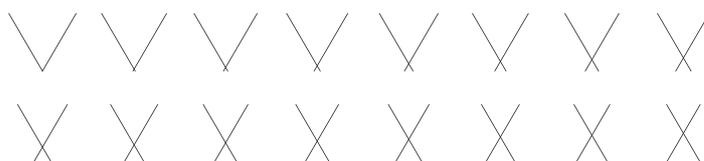


図 4.17 アルファベットの変形パターン 2 (Yasuhara ら 1978)

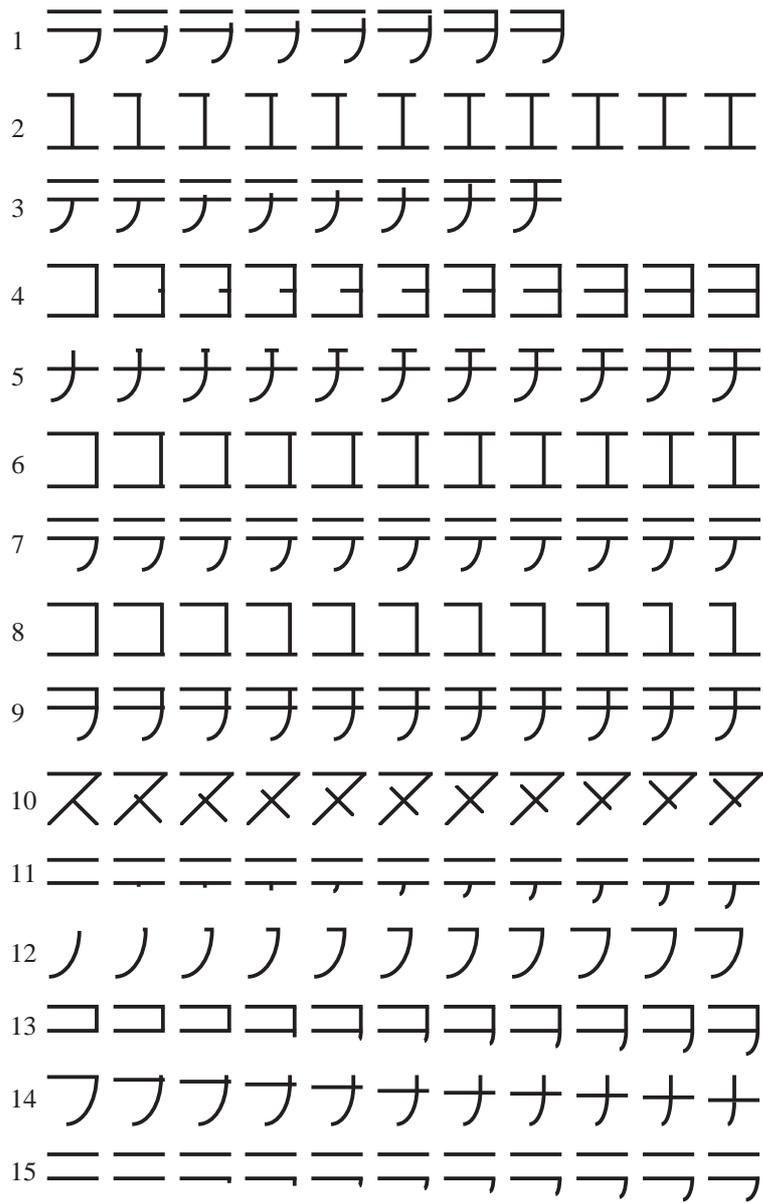


図 4.18 変形したカタカナ (田中)

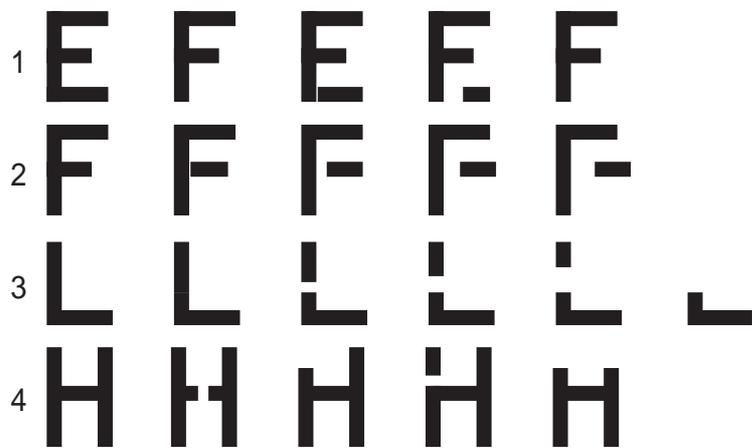


図 4.19 変形パターン 1 (横瀬)

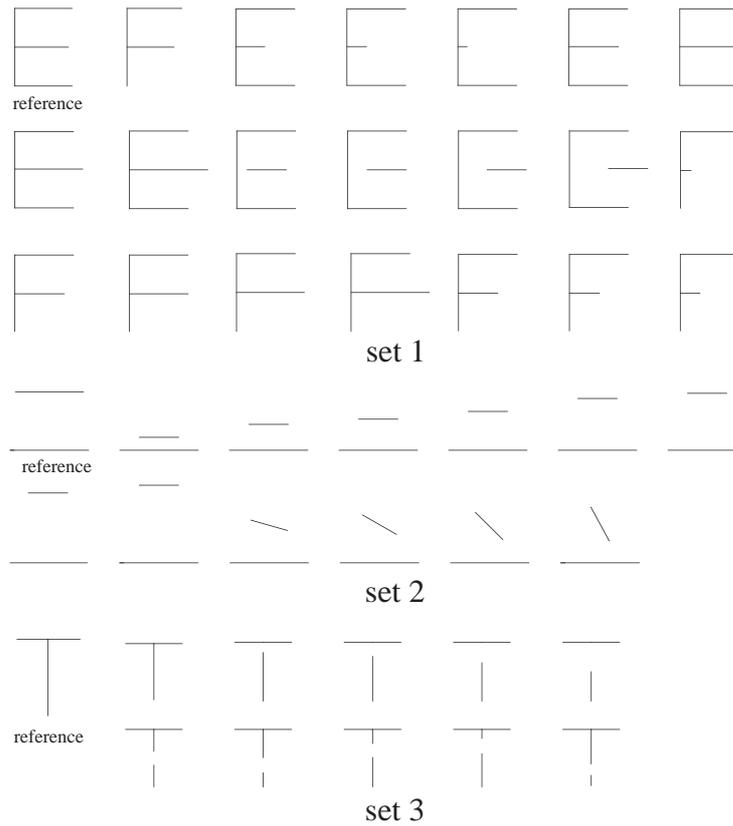


図 4.20 変形パターン 2 (横瀬)

工的な線図形の場合と同様に、3 個の  $R_j(S_j, S_jE)$  を正規変換量  $z$  に変換した上で加重平均し、Shillman らの実験の全 3 変形系列における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、図 4.16 の全 3 変形系列における実験結果と誘導場および LDCD 特徴の Euclid 距離の相関係数の平均を求める。Shillman らの実験と同様に、図 4.17 のパターン  $i$  について、誘導場間の弾性エネルギー  $YaE_{1i}$ 、誘導場間の Euclid 距離  $YaFD_{1i}$ 、LDCD 特徴の Euclid 距離  $YaLD_{1i}$  を求め、Yasuhara らの実験結果  $Ya_i$  から相関係数を求める。図 4.18 の  $j$  番目の変形系列のパターン  $i$  については、誘導場間の弾性エネルギー  $TjE_{1i}$ 、誘導場間の Euclid 距離  $TjFD_{1i}$ 、LDCD 特徴の Euclid 距離  $TjLD_{1i}$  を求め、田中の実験結果  $Tj_i$  から、図 4.18 の全 15 変形系列について相関係数の平均を求める。図 4.19 の  $j$  番目の変形系列のパターン  $i$  については、誘導場間の弾性エネルギー  $YOjE_{1i}$ 、誘導場間の Euclid 距離  $YOjFD_{1i}$ 、LDCD 特徴の Euclid 距離

表 4.3 相関をとる変数の一覧 (変形パターン)

pattern	psychological experiment	field (energy)	field (Euclid)	LDCD (Euclid)
Shillman <sup>96)</sup>	$S_j$	$SjE_{1i}$	$SjFD_{1i}$	$SjLD_{1i}$
Yasuhara <sup>98)</sup>	$Ya_i$	$YaE_{1i}$	$YaFD_{1i}$	$YaLD_{1i}$
田中 <sup>89)</sup>	$Tj_i$	$TjE_{1i}$	$TjFD_{1i}$	$TjLD_{1i}$
横瀬 <sup>44)</sup> (図 4.19)	$YOj_i$	$YOjE_{1i}$	$YOjFD_{1i}$	$YOjLD_{1i}$
横瀬 <sup>44)</sup> (図 4.20)	$Yoj_i$	$YojE_{1i}$	$YojFD_{1i}$	$YojLD_{1i}$

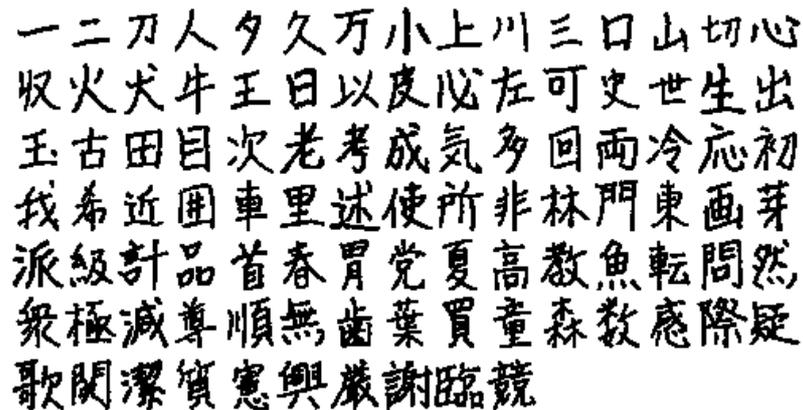


図 4.21 評定 100 文字種

$YO_jLD_{1i}$  を求め、横瀬の実験結果  $YO_j$  から、図 4.19 の全 4 変形系列について相関係数の平均を求める。図 4.20 の set  $j$  の変形系列のパターン  $i$  については、誘導場間の弾性エネルギー  $YO_jE_{1i}$ 、誘導場間の Euclid 距離  $YO_jFD_{1i}$ 、LDCD 特徴の Euclid 距離  $YO_jLD_{1i}$  を求め、横瀬の実験結果  $YO_j$  から、図 4.20 の全 3 変形系列について相関係数の平均を求める。相関をとる変数の一覧を表 4.3 にまとめた。

#### 4.3.4 手書き文字

手書き文字の変形を評定した心理学実験として安藤らの実験があるが、加藤らの実験<sup>37)</sup>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<sup>105)</sup>。ここでは、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<sup>37)</sup>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 4.6 の例のように文献<sup>106)</sup>に記載されている。これらを、 $64 \times 64$  dot の大きさ (72 DPI) 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 (200 DPI) して読み取り、文字画像を作成した。図 4.21 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である。図 4.21 の左上から下に順に文字種番号  $j(1, 2, \dots, j, \dots, 100)$  をつける。例えば、「一」は  $j = 1$ 、「二」は  $j = 2$ 、「競」は  $j = 100$  である。

線図形の場合と同様に、誘導場モデルの弾性エネルギー、誘導場の強さの Euclid 距離、LDCD 特徴の Euclid 距離の 3 つの評価方法と心理実験結果の相関を求める。ここでは、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 4.21 に示した各文字種ごとに最もきれいだと評価された文字とする。例えば、図 4.6 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。

なお、文字画像は文献<sup>106)</sup>に記載されているままの状態ですべての方法の評価を行った。ここで、文字認識で行われる大きさの正規化<sup>114)</sup>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<sup>106)</sup>。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モデルの弾性エネルギー，誘導場の強さの 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．人工的な線図形の場合と同様に，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．

## 4.4 検証結果

検証に用いた各パターンについて，心理実験結果と 3 つの評価方法（誘導場モデルの弾性エネルギー，誘導場の強さの Euclid 距離，LDCD 特徴の Euclid 距離）の相関関係を評価し，どの評価方法が最も心理実験結果をよく説明できるか検討した．

### 4.4.1 人工的な線図形

Lupker，横澤らの心理実験結果の検証

まず，Lupker，横澤らの心理実験結果と，3 つの評価方法の相関関係を示す．図 4.22 は図 4.2 のパターン 2 について，横軸に Lupker の実験結果  $L_{2j}$ ，縦軸に (a) は誘導場の弾性エネルギー  $LE_{2j}$ ，(b) は誘導場の Euclid 距離  $LFD_{2j}$ ，(c) は LDCD 特徴の Euclid 距離  $LLD_{2j}$  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．図 4.23(a),(b),(c) は，図 4.3 の set1 の「天」について，横軸に「天」と比較した set1 の全パターン名，縦軸右に横澤らの実験結果  $Y1[\%]$ ，縦軸左に (a) は誘導場の弾性エネルギー  $Y1E$ ，(b) は誘導場の Euclid 距離  $Y1FD$ ，(c) は LDCD 特徴の Euclid 距離  $Y1LD$  を示した例である．同様に，図 4.23(d),(e),(f) は，図 4.3 の set2 の「王」について，横軸に「王」と比較した set2 の全パターン名，縦軸右に横澤らの実験結果  $Y2[\%]$ ，縦軸左に (d) は誘導場の弾性エネルギー  $Y2E$ ，(e) は誘導場の Euclid 距離  $Y2FD$ ，(f) は LDCD 特徴の Euclid 距離  $Y2LD$  を示した例である．なお，Lupker の心理実験結果の類似度と計算した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，両者の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．以降の検証においても，心理実験結果と誘導場モデルの弾性エネルギーの相関関係が直線的にならない場合，常用対数をとって相関係数を求めた．

図 4.22 のパターン 2 の例を含め，図 4.2 の他のパターンも同様に回帰直線周辺に分布が集中している．ま

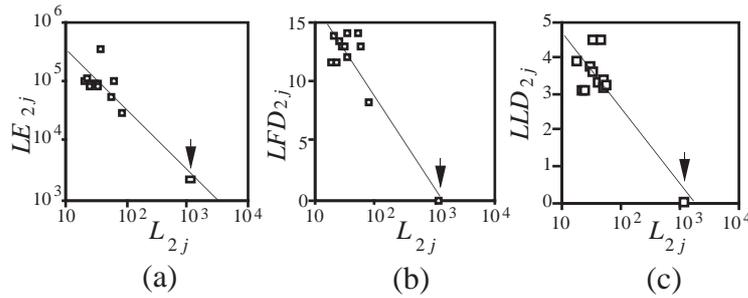


図 4.22 単純な線図形の相関比較

た，図 4.23 の「天」「王」の例を含め，他の set や文字種の例も同様に，心理実験結果の折れ線の変化の仕方が 3 つの評価方法の結果と非常に近いことから，3 つの評価方法すべて心理実験結果との相関が高いと考えられる．相関係数の平均  $r$  を求めた結果，すべての評定方法は説明率 ( $r^2 \times 100$  [%]) 50%以上の強い相関であった<sup>注6</sup>．本論文では，基本的に心理実験結果の全評定値と 3 つの評価方法の結果を比較することを原則とする．検証に用いた心理実験結果は，正解（基準）パターンとそれ以外のパターン間の類似性を人間が評価した結果であり，その全ての評定値に対する相関関係を検証することが妥当と考えられるからである．しかし，3 つの評定方法の違いによる説明率の差は 10%未満と小さく，どの評価方法が最も相関が強いかわからない．それは，図 4.22 の例のように，回帰直線は実質的に正解パターン（図中の矢印）と，それ以外のパターンの 2 つの群を結んだ形になっているのが原因と考えられる．実際，正解パターンの Euclid 距離はすべて 0，弾性エネルギーは他のパターンに比べ約 10 倍小さい．したがって，評価方法の違いを正しく評価するには，正解パターンを除いたパターンの相関を新たに分析する必要がある．

そこで，正解パターンを除外した場合の相関係数の平均を表 4.4 に示した．括弧内の数字は説明率 [%] である．表 4.4 から，第 1 列の誘導場の弾性エネルギーは強い相関，または説明率 15 から 50%未満の中程度の相関である．一方，誘導場と LDCD 特徴で Euclid 距離を使った場合の説明率（第 2，3 列）は，説明率 15%未満の弱い相関か，最大でも中程度の相関で，弾性エネルギーを使った場合に比べ，すべてのパターンで小さい（9%から最大 35%）．このことから，弾性エネルギーが心理実験結果と最も相関が強いと考えられる．誘導場の弾性エネルギーが，他の 2 つの評定方法と比較して相関係数が大きいかどうかを確かめるため，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  $t$  検定した．ただし相関係数は正規分布しないので，表 4.4 で求められた相関係数を正規変換量  $z$  に変換した上で  $t$  検定を行った<sup>注7</sup>．検定結果を表 4.5 に示す．表 4.5 の  $\square$  は 1, 5, 10%すべて， $\circ$  は 5, 10%， $\triangle$  は 10%の危険率で有意差がある場合， $\times$  はいずれの危険率でも有意差がない場合を示す．数字は標準化した偏差である．表 4.5 から，比較対象が誘導場の Euclid 距離で set 3 の場合のみ有意差がない．これは set3 のパターンは set1, 2 に比べ，パターン間の類似性が低く，誤読しにくい<sup>33)</sup> ため，Euclid 距離でも類似性の違いが比較的正しく評価できたと考えられる．しかし表 4.5 から，ほとんどの場合，危険率 5%以上で有意なこと

注6説明率については付録 G を参照．

注7付録 E の相関係数の検定を参照．

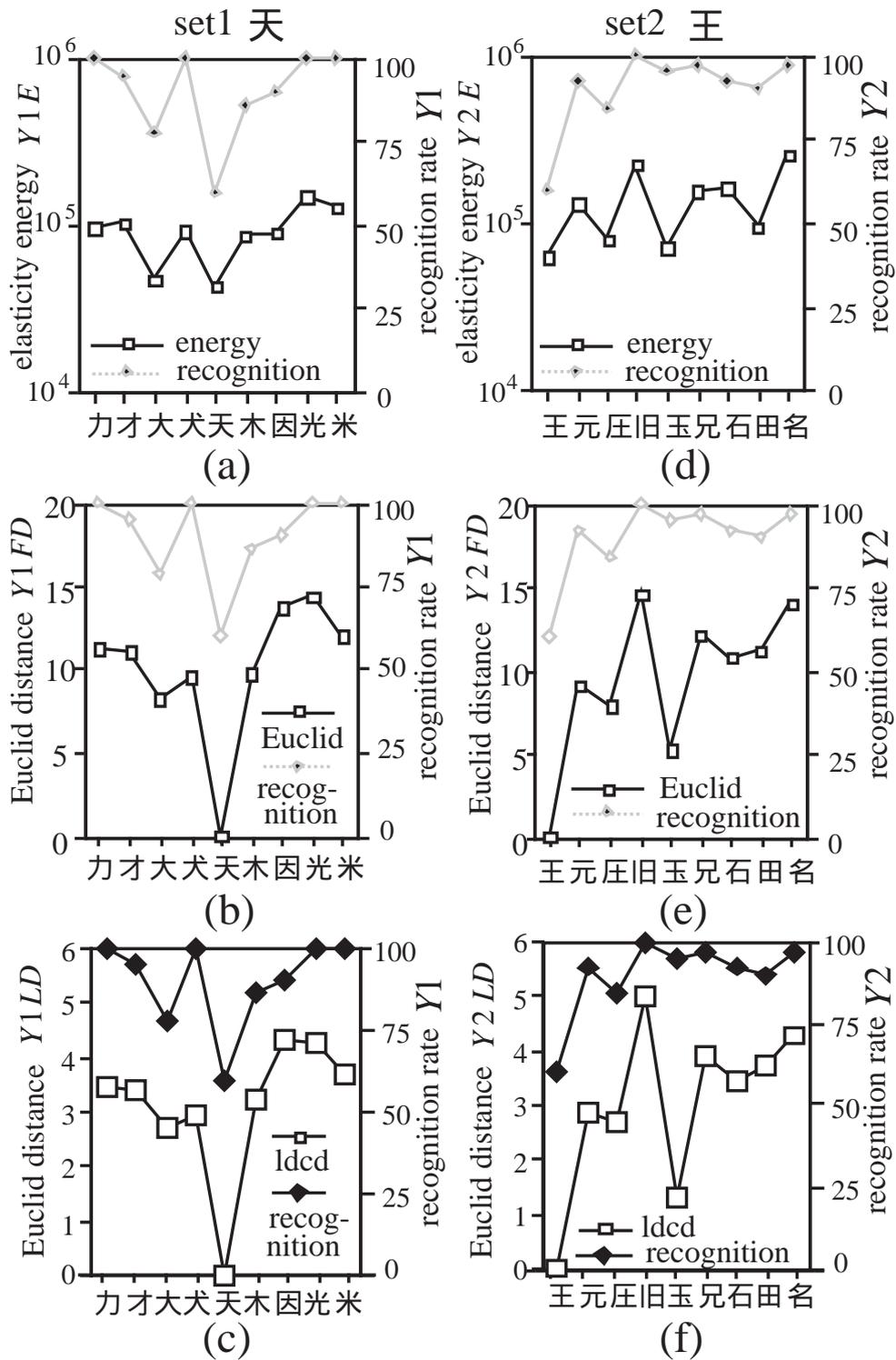


図 4.23 漢字パターンの心理実験と3つの評価方法との比較

表 4.4 正解パターンを除いた相関係数の平均 (人工的な線図形)

pattern	field(energy)	field(Euclid)	LDCD(Euclid)
Lupker <sup>88)</sup>	0.540 (29.2)	0.360 (13.0)	0.380 (14.4)
set 1 <sup>33)</sup>	0.692 (47.9)	0.420 (17.6)	0.351 (12.3)
set 2 <sup>33)</sup>	0.672 (45.2)	0.514 (26.4)	0.463 (21.4)
set 3 <sup>33)</sup>	0.763 (58.2)	0.703 (49.4)	0.627 (39.3)

表 4.5 相関係数の検定結果 (人工的な線図形)

pattern	compared method	
	field(Euclid)	LDCD(Euclid)
Lupker <sup>88)</sup>	(2.15)	(2.09)
set 1 <sup>33)</sup>	(2.92)	(3.46)
set 2 <sup>33)</sup>	△ (1.75)	(2.23)
set 3 <sup>33)</sup>	× (1.07)	(2.09)

から、誘導場の弾性エネルギーは、他の2つの評定方法、特にLDCD特徴のEuclid距離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、Lupker、横澤らの心理実験結果を最もよく説明できると考えられる。

#### Palmerの心理実験結果の検証

図4.1のパターンについて、図4.24(a)はPalmerが心理実験で求めた類似度、図4.24(b)は誘導場の弾性エネルギー、図4.24(c)は誘導場のEuclid距離、図4.24(d)はLDCD特徴のEuclid距離を縦軸に、SとH間、SとL間の平均を横軸に示した図である。図4.24(a)の心理実験の結果(SとH間がSとL間よりも類似性が高い)が、matched, unmatched共に同じ傾向を示しているのは、図4.24(b)の弾性エネルギーの場合だけである。SとH間の弾性エネルギーがSとL間よりも小さいか危険率5%でt検定した結果、matchedは $t_0 = 2.69 > t(3, 0.1) = 2.35$ , unmatchedは $t_0 = 3.32 > 2.35$ で共に有意差があり、SとH間がSとL間よりも弾性エネルギーが小さいことが示された。一方、SとL間、SとH間で誘導場のEuclid距離の差に有意差があるか否かを危険率5%でt検定した結果、matchedは $t_0 = 2.02 < t(3, 0.1) = 2.35$ , unmatchedは $t_0 = 1.99 < 2.35$ で共に有意差はなく、SとL間、SとH間の差は認められなかった。また、SとL間、SとH間でLDCD特徴のEuclid距離の差に有意差があるかを危険率5%でt検定した結果、matchedは $t_0 = 0.652 < t(3, 0.1) = 2.35$ で有意差はなく、SとL間、SとH間の差は認められなかった。しかし、unmatchedは $t_0 = 4.75 > 2.35$ で有意差があり、Palmerの心理実験と逆の結果が示された。以上の結果から、誘導場の強さやLDCD特徴のEuclid距離は図4.24(a)のPalmerの実験結果を説明できないが、誘導場モデルの弾性エネルギーは説明可能であると考えられる。

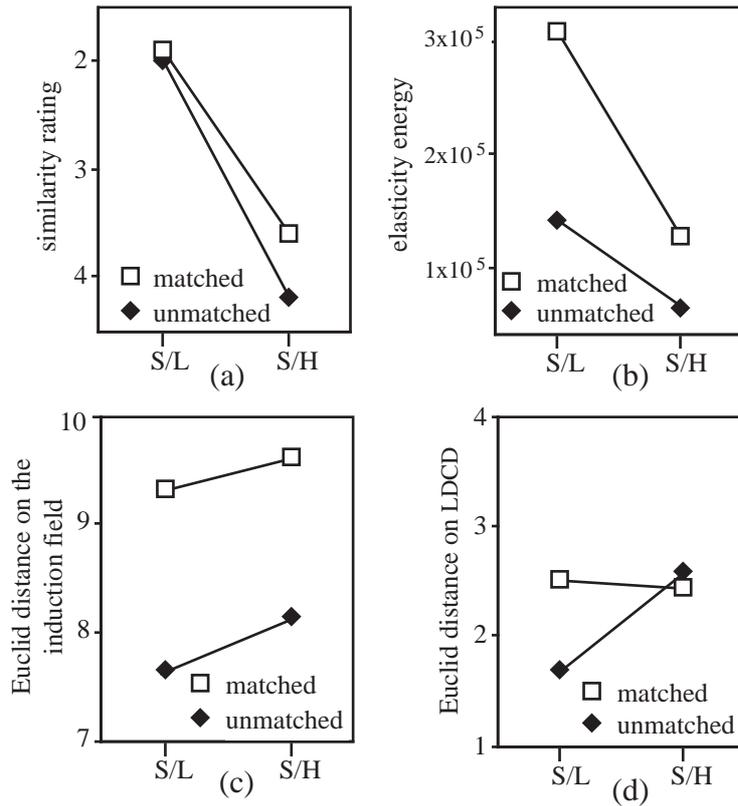


図 4.24 類似した線図形の検証結果

#### 4.4.2 活字

図 4.16 ~ 図 4.20 に関する 5 つの心理実験 (Townsend, Kuennapas ら, 菊地ら, 田中, 村石ら) の結果と, 3 つの評価方法 (誘導場モデルの弾性エネルギー, 誘導場の強さの Euclid 距離, LDCD 特徴の Euclid 距離) の相関係数の平均  $r$  を求めた結果, 全ての評定方法は説明率 ( $r^2 \times 100\%$ ) 50%以上の強い相関であり, 評定方法の違いによる説明率の差は数%程度と小さく, どの評価方法が最も相関が強いかどうか言えない. 相関を求めた際の回帰直線は線図形と同様に, 実質的には正解パターンと, それ以外のパターンの 2 つの群を結んだ形になっているため, 相関が高くなったと考えられる. 実際, 他のパターンと比べた正解パターンの評価値は, Euclid 距離を用いた場合ゼロ, 誘導場の弾性エネルギーは 10 倍程度小さい. 従って評価方法の違いを正しく評価するには, 正解パターンを除いたパターン間の相関を新たに分析する必要がある.

正解パターンを除いたパターン間の相関関係の例として, 図 4.25(a),(b),(c) は, 図 4.10 の「P」について, 縦軸に Townsend の実験  $To_{16j}$ , 横軸に (a) は誘導場の弾性エネルギー  $ToE_{16j}$ , (b) は誘導場の Euclid 距離  $ToFD_{16j}$ , (c) は LDCD 特徴の Euclid 距離  $ToLD_{16j}$  をとった散布図の例である. 同様に, 図 4.25(d),(e),(f) は, 図 4.11 の「i」 ( $i = 9$ ) について, 縦軸に Kunnapas の実験結果  $Ku_{9j}$ , 横軸に (d) は誘導場の弾性エネルギー  $KuE_{9j}$ , (e) は誘導場の Euclid 距離  $KuFD_{9j}$ , (f) は LDCD 特徴の Euclid 距離  $KuLD_{9j}$  をとった散布図の例である. また, 図 4.26(a),(b),(c) は, 図 4.13 の「Q」 ( $i = 17$ ) について, 縦軸に菊地らの実験結果  $KI_{17j}$ , 横軸に (a) は誘導場の弾性エネルギー  $KIE_{17j}$ , (b) は誘導場の Euclid 距離  $KIFD_{17j}$ , (c) は LDCD

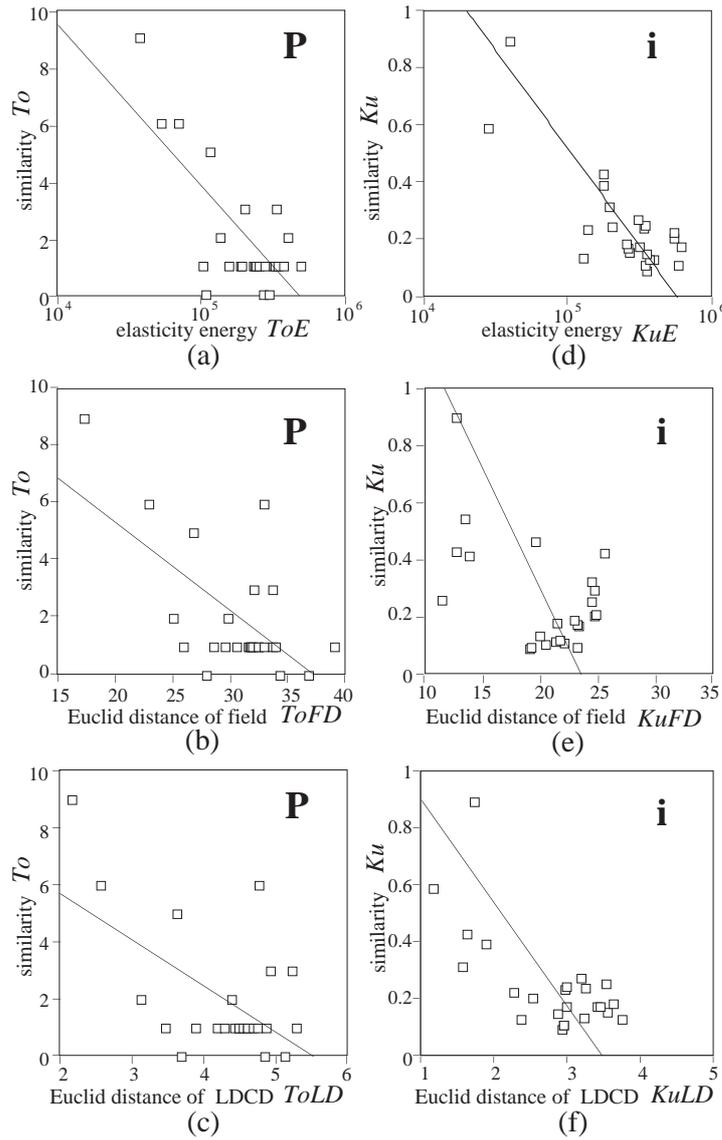


図 4.25 心理実験と 3 つの評価方法との比較 (「P」と「i」)

特徴の Euclid 距離  $KILD_{17j}$  をとった散布図の例である．図 4.26(d),(e),(f) は，図 4.12 の「ク」( $i = 8$ ) について，縦軸に田中らの実験結果  $TA_{8j}$ ，横軸に (d) は誘導場の弾性エネルギー  $TAE_{8j}$ ，(e) は誘導場の Euclid 距離  $TAFD_{8j}$ ，(f) は LDCD 特徴の Euclid 距離  $TALD_{8j}$  をとった散布図の例である．

いろいろな文字の心理実験結果と 3 つの評価方法との相関の様子を示すため，文字の構造が大きく異なっている「P, i, Q, ク」を代表例として選んだ．図 4.25, 図 4.26 の例を含め，他の文字パターンおよびすべての心理実験結果に対しても同様に，誘導場の弾性エネルギーが，誘導場と LDCD 特徴で Euclid 距離を用い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．正解パターンを除外した場合の相関係数の平均を表 4.6 に示した．括弧内の数字は説明率 (%) である．表 4.6 から，第 1 ~ 3 行の Townsend, Kunnapas, 菊地のアルファベットについて，誘導場の弾性エネルギー (第 1 列) は説明率 15% から 50% 未満の中程度の相関，誘導場と LDCD 特徴で Euclid 距離を使った場合の説明率 (第 2, 3 列) は中程度の相関，

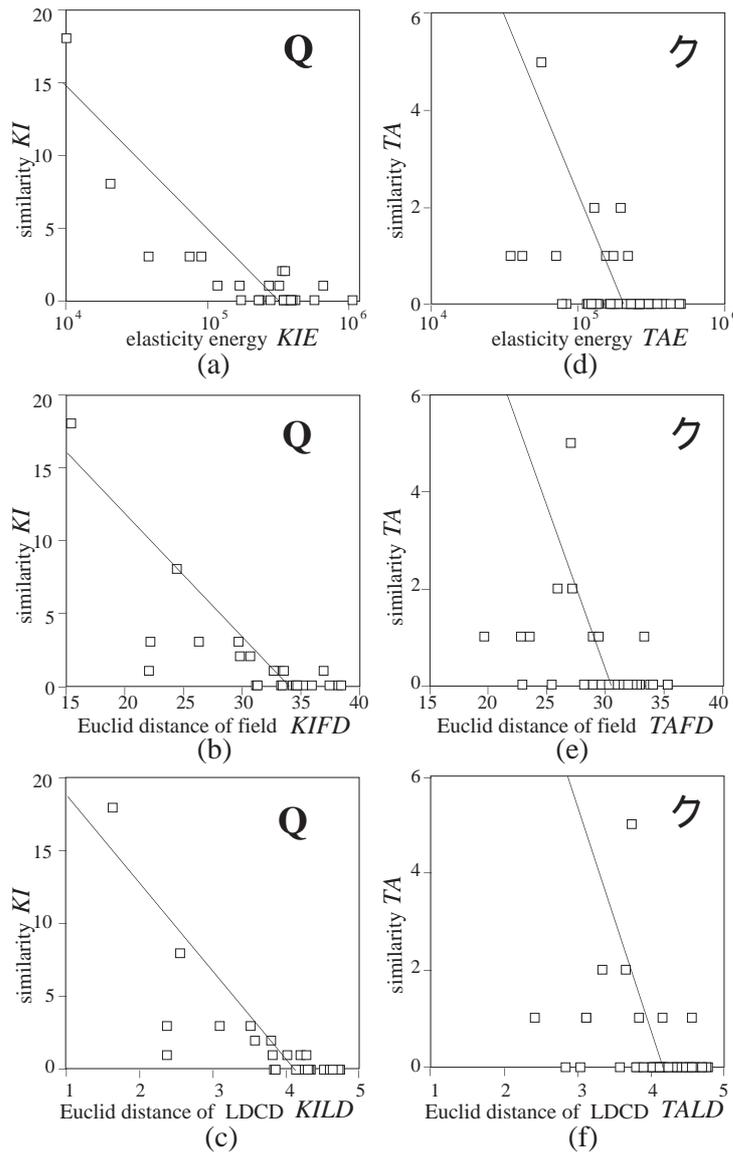


図 4.26 心理実験と3つの評価方法との比較(「Q」と「ク」)

または説明率 15%未満の弱い相関である．第 4，5 行の田中のカタカナ，村石の平仮名について，3つの評価方法ともに弱い相関である．しかし，すべての対象パターンについて，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大きく（5%から 10%），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 距離を正規変換量  $z$  に変換した上で  $t$  検定した．検定結果を表 4.7 に示す．表 4.7 の  $\square$  は 1，5，10%， $\square$  は 5，10%， $\triangle$  は 10%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 4.7 から，Kunnapas，菊地らの実験について LDCD 特徴の Euclid 距離に対して危険率 10%で有意だが，偏差は危険率 5%の水準に近い．それ以外は危険率 5%以下で有意なことから，誘導場の弾性エネルギーは他の2つの評価方法，特に誘導場の Euclid 距離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，5つの心理実験

表 4.6 正解パターンを除いた相関係数の平均 (活字)

pattern	field (energy)	field (Euclid)	LDCD
Townsend <sup>91)</sup>	0.531 (28.2)	0.461 (21.2)	0.464 (21.5)
Kunnapas <sup>92)</sup>	0.584 (34.1)	0.484 (23.4)	0.513 (26.3)
菊地 <sup>93)</sup>	0.436 (19.0)	0.350 (12.3)	0.366 (13.4)
田中 <sup>89)</sup>	0.351 (12.3)	0.305 (9.27)	0.311 (9.70)
村石 <sup>94)</sup>	0.285 (8.13)	0.159 (2.52)	0.190 (3.60)

表 4.7 相関係数の検定結果 (活字)

pattern	field (Euclid)	LDCD (Euclid)
Townsend <sup>91)</sup>	(1.98)	(2.05)
Kunnapas <sup>92)</sup>	(2.61)	△ (1.82)
菊地 <sup>93)</sup>	(2.10)	△ (1.89)
田中 <sup>89)</sup>	(2.45)	(2.19)
村石 <sup>94)</sup>	(4.38)	(2.78)

結果を最もよく説明できると考えられる。

#### 4.4.3 変形パターン

最初に、図 4.16 ~ 図 4.20 に関する 5 つの心理実験 (Shillman ら, Yasuhara ら, 田中, 横瀬) の結果と、3 つの評価方法 (誘導場モデルの弾性エネルギー, 誘導場の強さの Euclid 距離, LDCD 特徴の Euclid 距離) の相関関係を示す。いろいろな変形パターンの心理実験結果と 3 つの評価方法との相関の様子を示すため、各実験において、できるだけパターンの構造が異なっており、線分の伸ばし方、省略などの変形の傾向が異なる「V, y, X, コ, エ, ラ, テ, H, T」を代表例として選んだ。図 4.27(a),(b),(c) は、図 4.16 の 1 番目の変形系列 V → y について、横軸に Shillman らの実験  $S1_i$  (パターン  $i$  を V と読む程度の 5 段階評価, 0; 全く異なる, 5; 一致), 縦軸に (a) は誘導場の弾性エネルギー  $S1E_{1i}$ , (b) は誘導場の Euclid 距離  $S1FD_{1i}$ , (c) は LDCD 特徴の Euclid 距離  $S1LD_{1i}$  をとった散布図の例である。図 4.27 と同様に、図 4.28(a),(b),(c) は図 4.18 の 6 番目の変形系列コ → エ, 図 4.28(d),(e),(f) は図 4.18 の 7 番目の変形系列ラ → テ, 図 4.29(a),(b),(c) は図 4.19 の 4 番目の H の変形系列, 図 4.29(d),(e),(f) は図 4.20 の set3 の T の変形系列について比較した散布図の例である。図 4.27 ~ 図 4.29 の例を含め、他の変形パターンおよびすべての心理実験結果に対しても同様に、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。

一方、図 4.27(b),(e) や 図 4.28(b),(e), 図 4.29(b),(e) の誘導場の Euclid 距離, および図 4.28(c),(f) や 図 4.29(c),(f) の LDCD 特徴の Euclid 距離の例のように、これら 2 つの評価方法による正解パターンの評価結果が回帰直線からはずれることが多い。これは、誘導場の Euclid 距離と LDCD 特徴の Euclid 距離は、誘導場の弾性エネルギーに比べ、正解パターンとその他のパターンの類似性の両方を評価することが難しい

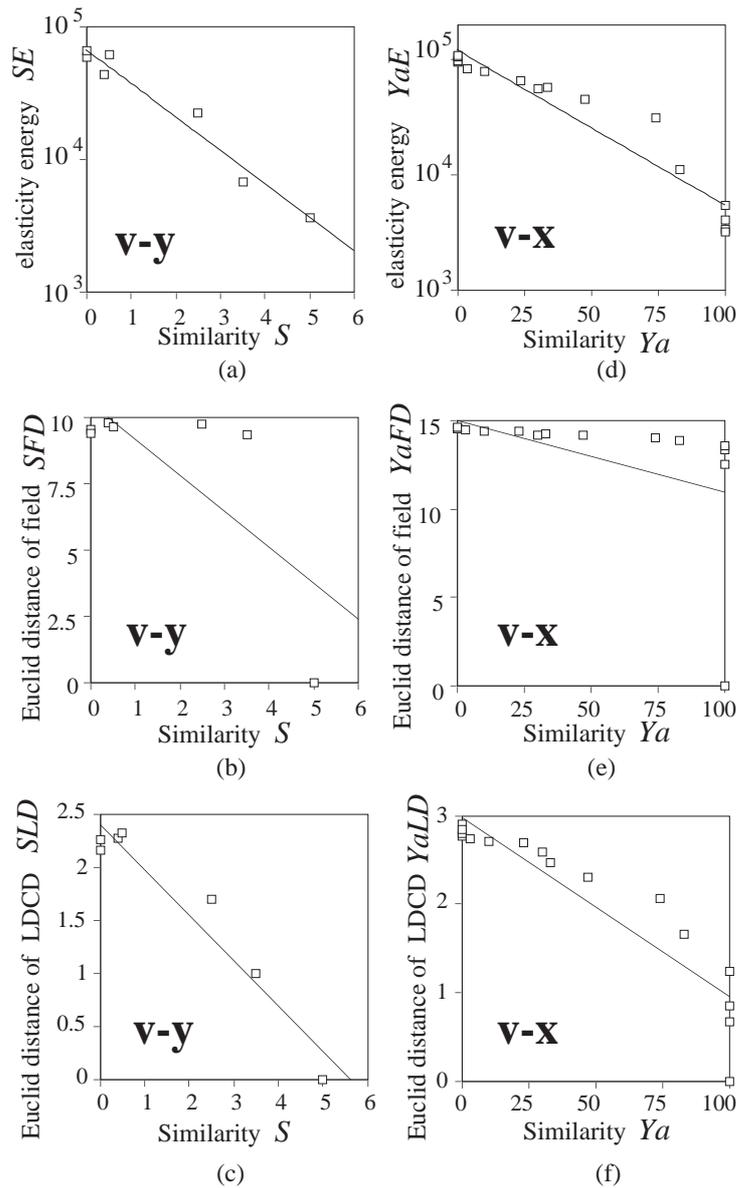


図 4.27 心理実験と3つの評価方法との比較（アルファベット）

ことを示している。

次に、相関係数の平均を表 4.8 に示した。括弧内の数字は説明率 (%) である。表 4.8 から、誘導場と LDCD 特徴で Euclid 距離を使った場合の説明率 (第 2, 3 列) は 15% から 50% 未満の中程度の相関または説明率 50% 以上の強い相関だが、誘導場の弾性エネルギー (第 1 列) はすべて強い相関である。誘導場の弾性エネルギーの説明率は、同じ対象について、他の 2 つの評価方法よりも大きく (15% から 67%)、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 つの評定方法と比較して相関係数が高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  $t$  検定した結果を表 4.9 に示す。表 4.9 から、図 4.20 の横瀬の実験および Yasuhara らの実験に関する LDCD 特徴

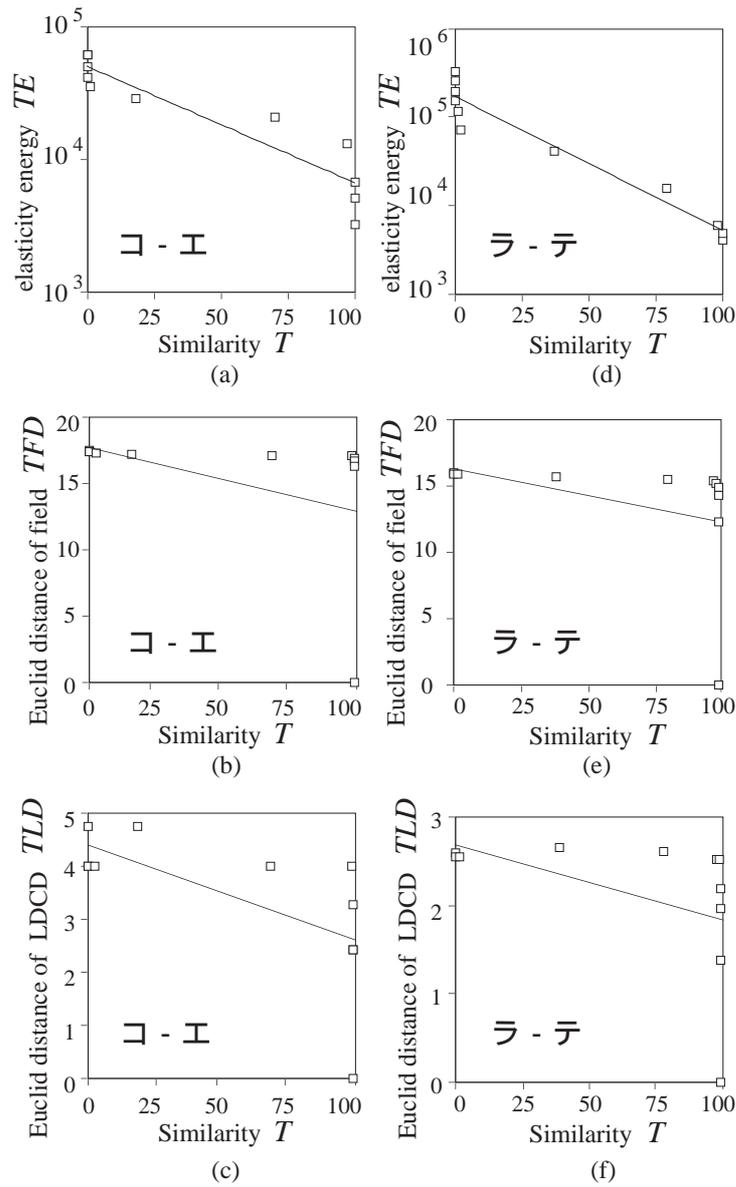


図 4.28 心理実験と3つの評価方法との比較 (カタカナ)

表 4.8 相関係数の平均 (変形パターン)

pattern	field (energy)	field (Euclid)	LDCD
Shillman <sup>96)</sup>	0.900 (81.0)	0.652 (42.5)	0.772 (59.6)
Yasuhara <sup>98)</sup>	0.939 (88.2)	0.455 (20.7)	0.855 (73.1)
田中 <sup>89)</sup>	0.922 (85.0)	0.596 (35.5)	0.693 (48.0)
横瀬 <sup>44)</sup> (図 4.19)	0.752 (56.6)	0.454 (20.6)	0.455 (20.7)
横瀬 <sup>44)</sup> (図 4.20)	0.972 (94.5)	0.819 (67.1)	0.772 (59.6)

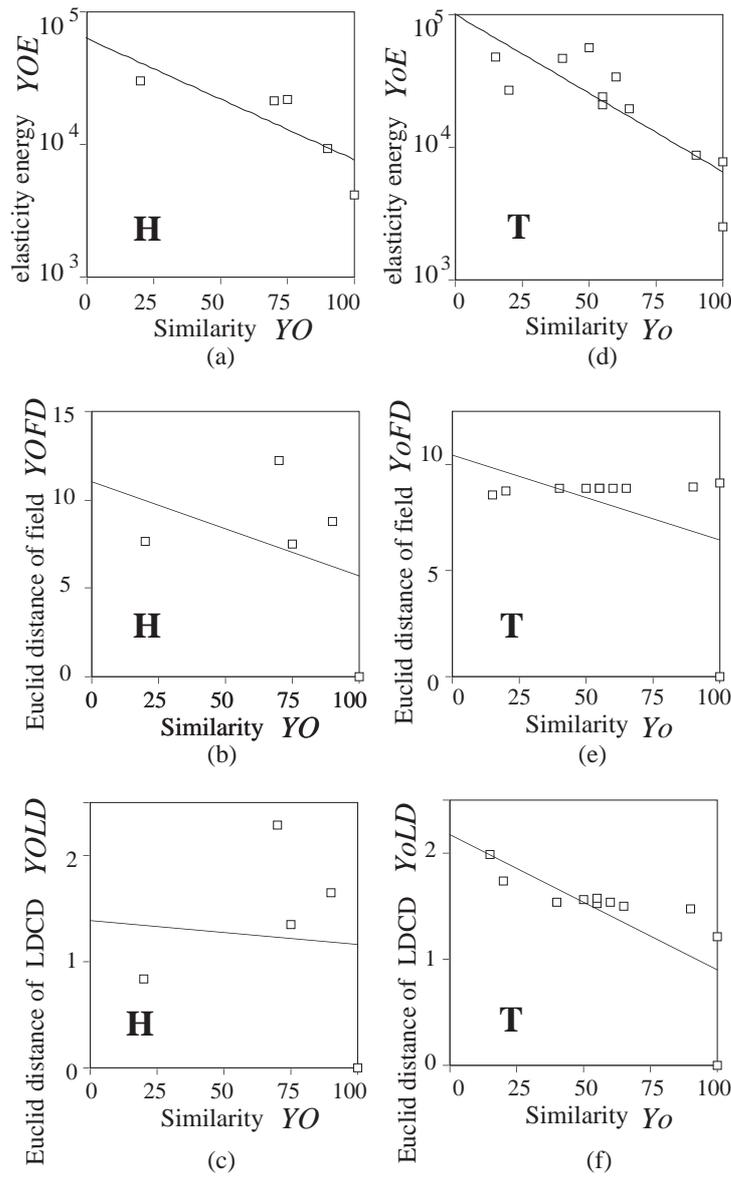


図 4.29 心理実験と3つの評価方法との比較（横瀬の変形パターン）

表 4.9 相関係数の検定結果（変形パターン）

pattern	field (Euclid)	LDCD (Euclid)
Shillman <sup>96)</sup>	(3.93)	(3.38)
Yasuhara <sup>98)</sup>	(6.44)	△ (2.32)
田中 <sup>89)</sup>	(8.21)	(6.37)
横瀬 <sup>44)</sup> (図 4.19)	(3.18)	(2.92)
横瀬 <sup>44)</sup> (図 4.20)	△ (2.33)	△ (2.57)

の Euclid 距離については危険率 5% で有意だが、それ以外は危険率 1% 以下で有意なことから、誘導場の弾性エネルギーは、他の 2 つの評定方法、特に誘導場の Euclid 距離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、5 つの心理実験結果を最もよく説明できると考えられる。

#### 4.4.4 手書き文字

最初に、加藤らの心理学実験結果と 3 つの評価方法（誘導場モデルの弾性エネルギー、誘導場の強さの 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 4.30(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。同様に、図 4.30(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 4.30 と同様に、図 4.31(a),(b),(c) は「史」( $j = 27$ )、図 4.31(d),(e),(f) は「我」( $j = 46$ )、図 4.32(a),(b),(c) は「胃」( $j = 67$ )、図 4.32(d),(e),(f) は「興」( $j = 96$ ) について比較した散布図の例である。図 4.30 ~ 図 4.32 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。

一方、図 4.30(b),(e) や図 4.31(b),(e)、図 4.32(b),(e) の誘導場の Euclid 距離、および図 4.30(c),(f) や図 4.31(c),(f)、図 4.32(c),(f) の LDCD 特徴の Euclid 距離の例のように、これら 2 つの評価方法による基準パターンの評価結果が回帰直線からはずれることが多い。これは、誘導場の Euclid 距離と LDCD 特徴の Euclid 距離は、誘導場の弾性エネルギーに比べ、基準パターンとその他のパターンの主観評価の両方を評価することが難しいことを示している。

次に、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 4.10 に示した。括弧内の数字は説明率 (%) である。表 4.10 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 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 つの評定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 4.11 に示す。表 4.11 から、加藤らの実験に関する LDCD

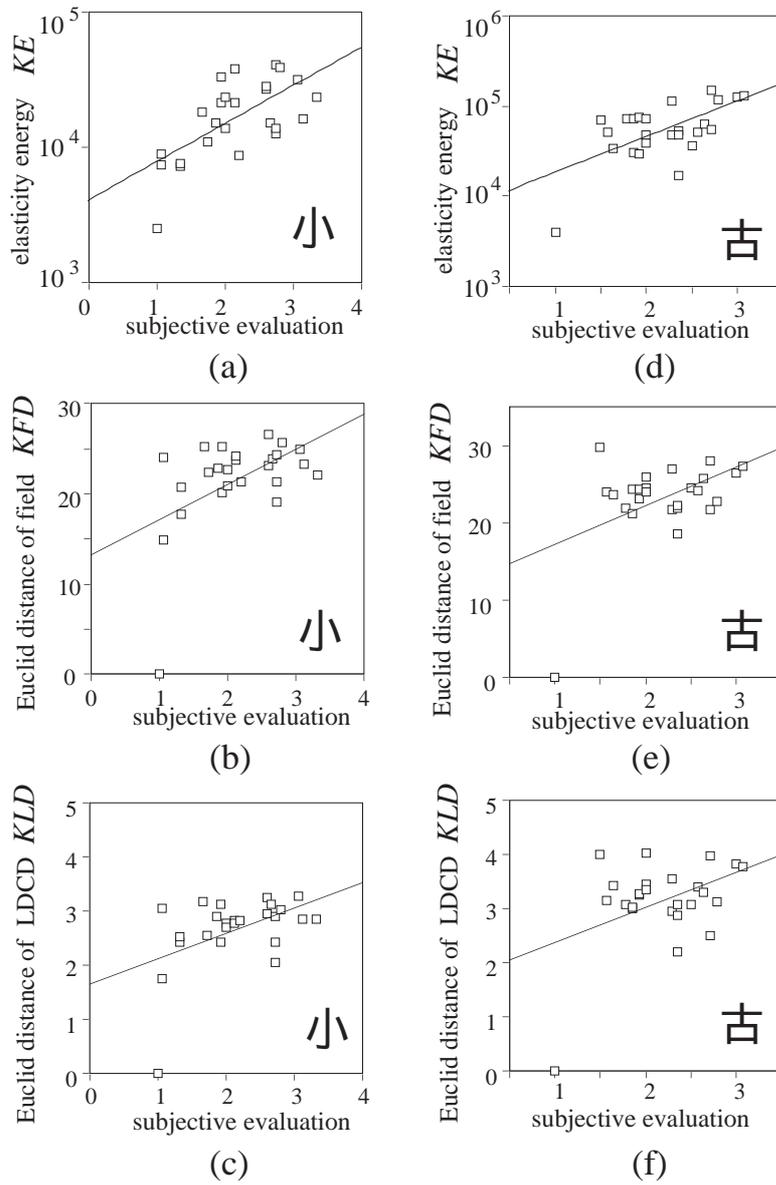


図 4.30 心理実験と3つの評価方法との比較 (画数の少ない漢字:「小」と「古」)

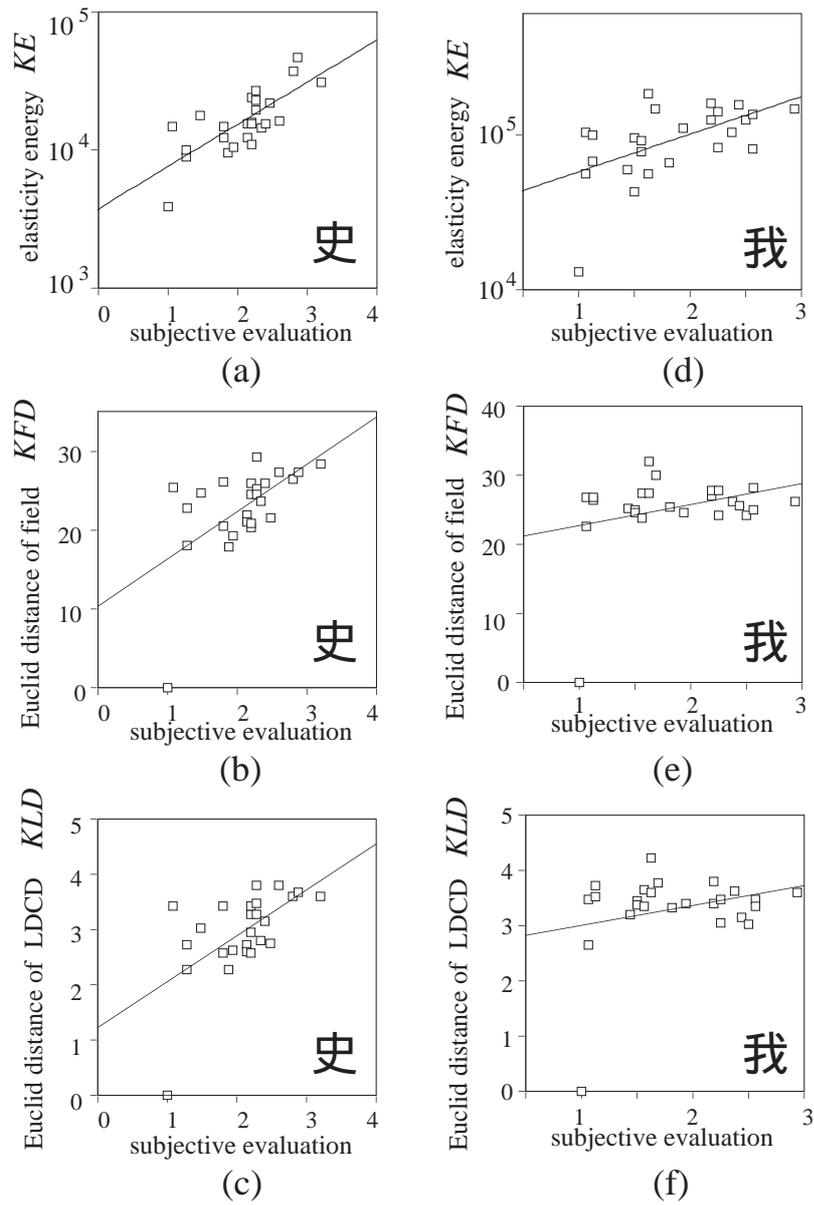


図 4.31 心理実験と3つの評価方法との比較（「史」と「我」）

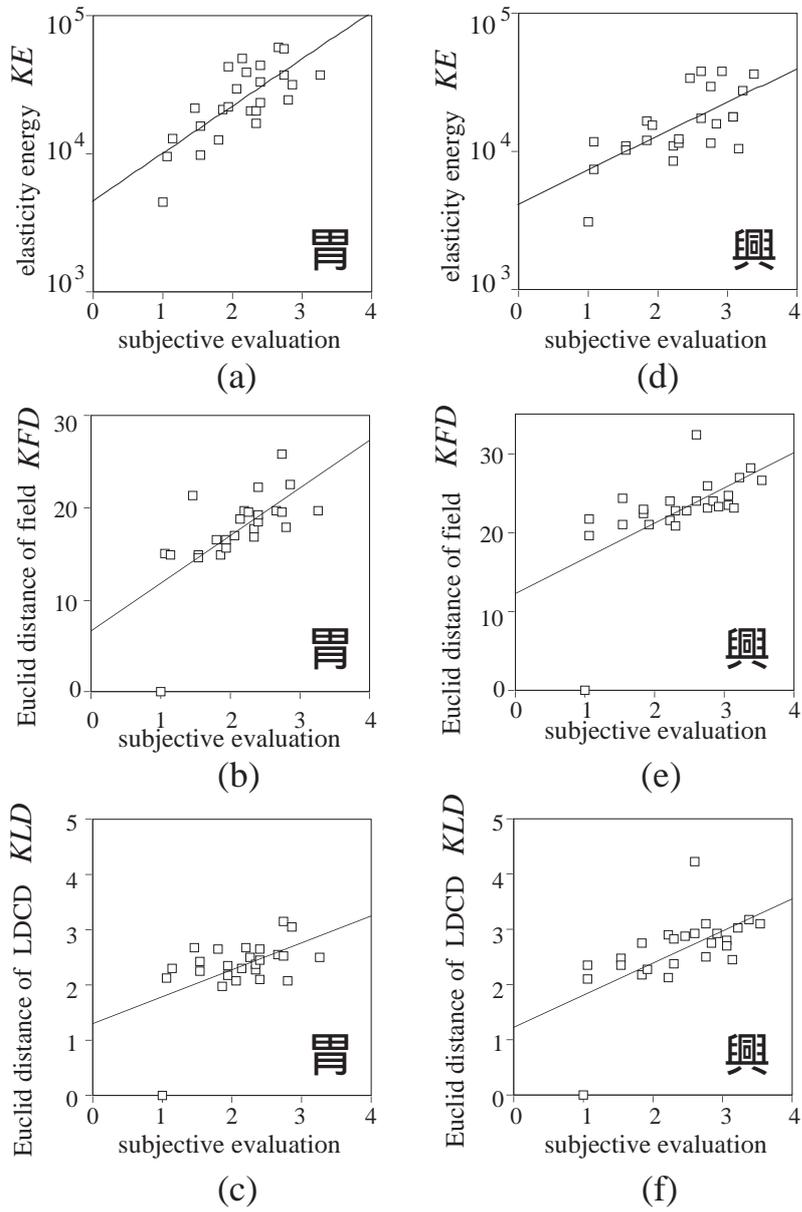


図 4.32 心理実験と3つの評価方法との比較 (画数の多い漢字:「胃」と「興」)

表 4.10 相関係数の例 (手書き文字)

Kanji	field (energy)	field (Euclid)	LDCD
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 4.11 相関係数の検定結果 (手書き文字)

pattern	field (Euclid)	LDCD (Euclid)
100 Kanji	(3.47)	(3.99)

特徴の Euclid 距離および誘導場の Euclid 距離について、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 つの評定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。

以上、人工的な線図形 (表 4.5, 図 4.24)、活字 (表 4.7)、変形パターン (表 4.9)、手書き文字 (表 4.11) について、誘導場モデルの有効性について検証した。その結果、各対象パターン共に誘導場モデルの弾性エネルギーが他の 2 つの評定方法よりも心理学実験結果をよく説明できることが示された。このことから、誘導場モデルは広範囲なパターンの類似性が説明できるモデルとして妥当と考えられる。

## 4.5 文字認識

最後に、誘導場モデルを文字認識に用いた場合の認識能力について述べる。誘導場の計算、弾性エネルギーによる識別は 4.3.1 の検証方法と同じである。また、従来手法として LDCD 特徴<sup>113)</sup> (8x8 メッシュ、水

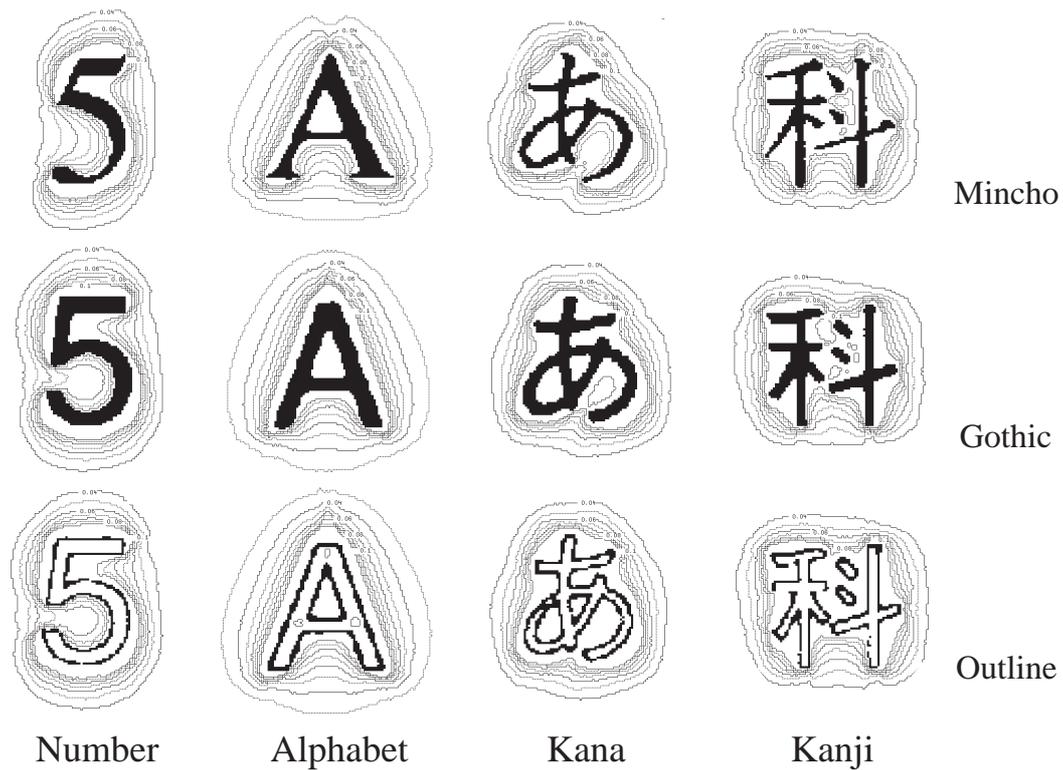


図 4.33 活字認識で使した文字の誘導場例

平・垂直および斜め方向  $\pm 45$  度の 4 方向、256 次元特徴) の類似度と比較した。なお、最適な辞書を決定するための学習は一切行っていない。

#### 4.5.1 活字認識

レーザプリンタ (Apple Laser Writer) で出力した 14 ポイントの数字 (0~9)、アルファベット (26 種)、平仮名 (46 種)、漢字 (50 種) について、各々 3 種類の書体 (細明朝体、ゴシック体、白抜きゴシック体) の活字<sup>注8</sup>の認識を行った。認識実験で使した文字の誘導場例を図 4.33 に示す。3 種類の書体ごとに同じ書体の辞書を用い、全字種・全書体について、誘導場および LDCD 特徴共に 100% 全て認識できた。

#### 4.5.2 手書き文字認識

電子技術総合研究所の手書き数字データベース ETL1<sup>7)</sup> の数字 (0 から 9) から、任意に選んだ 100 文字を、同じ ETL1 のデータベース第 1 番目の文字を辞書として認識を行った。また、同 ETL9<sup>104)</sup> の平仮名 (字種は活字と同じ) から任意に選んだ 20 文字を、「現代書道字典」(阿保直彦編著、木耳社) から、イメージスキャナで読み取った楷書体の手書きの平仮名を辞書として認識を行った。各認識結果の認識率 (%) の一例と全手書き文字の平均認識率を表 4.12 に示した。表 4.12 から、誘導場と LDCD 特徴の認識率は文字

<sup>注8</sup> イメージスキャナ (Ricoh IS-50 400DPI) で読み取り、1 文字づつ手動で切り出した。

表 4.12 手書き文字の認識結果例

数字	0	1	2	3	4	5	6	7	8	9	平均
LDCD	93	77	70	88	68	75	64	77	85	64	76.1
field	77	80	67	87	61	66	80	77	76	81	75.2
平仮名	あ	お	き	く	そ	て	の	る	わ		平均
LDCD	95	65	72	90	70	65	70	50	60		74.2
field	90	75	70	90	75	60	80	55	60		74.8

種により異なるが、範囲と平均値から従来手法と同程度の認識能力があることがわかる。

## 4.6 考察

### 4.6.1 誘導場によるパターン認識モデルの有効性

人工的な線図形、活字、変形パターンおよび手書き文字に関する検証結果より、誘導場の認識モデルは、これらパターンに関する心理実験結果との相関が、検討した3つの評価方法のなかで最も強いことが明らかになった。すなわち、いろいろなパターンの類似性などの心理実験結果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ことが示された。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であると考えられる。

### 4.6.2 ダイナミクスの意義

では、なぜ誘導場の認識モデルが文字パターンの類似性の評価に優れているのであろうか。検証に用いた多くのパターンについて、誘導場の強さおよびLDCD特徴<sup>113)</sup>でEuclid距離を用いた評価方法は、心理実験結果との相関が弾性エネルギーより弱い。実際、誘導場およびLDCD特徴でEuclid距離を用いた場合、図4.27～図4.29、図4.30～図4.32の例のように、正解パターンおよび基準パターンの評価結果が回帰直線からはずれることが多く、パターンの類似性を正しく評価することが難しくなっている。これは、今井<sup>34)</sup>が指摘しているように、Euclid距離などパターン認識の識別に使う距離と、人間が判断するパターンの類似性は必ずしも一致しないことを示していると考えられる。

一方、同じ誘導場でもEuclid距離ではなく、誘導場モデルの弾性エネルギーを用いた評価方法は心理実験結果とよく合致する。誘導場モデルは、4.2.1で述べたようにパターン認識はダイナミクスであるという見解から構築されており、これまでの対象文字パターンの構造などに依存する特徴を元にした方法ではない。そして、ダイナミクスという前提で考案された誘導場モデルが、実際に人工的なパターンから変形のある手書きパターンまでのいろいろな文字パターンについて心理実験との相関が高くなっている。このことは、



図 4.34 文字サンプル例：次

誘導場の認識モデルの枠組みであるダイナミクスは、パターン認識に関して重要な役割をもっており、その結果として、いろいろな文字パターンに対して優れたパターン類似性の評価能力が得られたと考えられる。

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<sup>115)</sup>。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<sup>34)</sup>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<sup>116)</sup>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、純粋に統計的な知識処理が提案される<sup>117)</sup>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、心理実験結果に近いことが示された。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、認識率が向上すると考えられる。

### 4.6.3 認識性能を向上させるための課題

#### パターンの正規化

文字パターンの大きさの違いに対応して、誘導場の正規化も必要と考えられる。ここでの検証では文字の大きさは正規化していない。しかし、検証に用いた文字画像の中には、例えば手書き文字の場合、図 4.34 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 4.34 の各文字の下の数字は文字品質評定値の平均である。図 4.34 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は必ずしも低くなることはない<sup>106)</sup>。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、検証に用いた 3 つの方法は文字の大きさの違

表 4.13 正規化の効果 (手書き文字)

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

表 4.14 誘導場の弾性エネルギーにおける正規化の効果 (活字)

Townsend <sup>91)</sup>	Kunnapas <sup>92)</sup>	菊地 <sup>93)</sup>	田中 <sup>89)</sup>	村石 <sup>94)</sup>
-0.004	-0.090	-0.017	-0.005	-0.035

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 4.34 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 4.13 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係数の平均と比較した結果である。表 4.13 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。

そこで、活字などのパターンについて、誘導場の強さおよび LDCD 特徴で Euclid 距離を用いた評価方法で行った、従来文字認識で一般的に行われている文字の大きさ・位置の正規化が、誘導場の弾性エネルギーを求める場合に有効かどうか調べた。表 4.14 は、元パターンを正規化して誘導場を求めた場合の弾性エネルギーについて、正解パターンを除いた相関係数の平均を、正規化しなかった表 4.6 の場合と比べた結果である。表 4.14 から、単純に元パターンの正規化を行うと心理実験結果との相関が弱くなる (平均  $-0.030$ ) ことがわかる。これは、文字の大きさが異なる図 4.15 のようなポイント数が統一されている状態が、文字の物理的な大きさを正規化した場合よりも、誘導場の弾性エネルギーを求めるのに適している可能性が高いと考えられる。また、ポイント数が同じであっても、物理的な文字の大きさが異なるのは、文字自体の大きさではなく、各文字の醸し出す場の広がりが揃うように、人間が経験的に文字をデザインしている可能性が示唆されている<sup>24)</sup>。これらから、誘導場の正規化は元パターンではなく、誘導場の広がりを揃える方法が適していると考えられる。例えば、各等ポテンシャル面の広がりを同じ面積をもつ円の直径で近似し、同じポテンシャル値ならば円の直径ができるだけ等しくなるように元パターンの大きさを調整して、誘導場の分布を揃えるなどの方法が考えられる。

#### 各等ポテンシャル面の利用

横瀬は変形文字の誘導場の観察から、人間はポテンシャル値が異なる複数の誘導場の分布の違いを総合的に比較することによって、変形パターンの類似性を判断していることを示唆している<sup>44)</sup>。例えば、図 3.5(a)

表 4.15 ポテンシャル値による相関係数の違い

Experiment pattern	potential value range		
	0.03 ~ 0.39	0.03 ~ 0.20	0.21 ~ 0.39
Shillman V-y <sup>96)</sup>	0.983 (96.6)	0.981 (96.2)	0.339 (11.5)
田中 コ - エ <sup>89)</sup>	0.899 (80.8)	0.868 (75.3)	0.892 (79.6)
田中 コ - ラ <sup>89)</sup>	0.928 (86.1)	0.900 (81.0)	0.878 (77.1)
横瀬 <sup>44)</sup> (図 4.19) E	0.966 (93.3)	0.989 (97.8)	0.503 (25.3)
横瀬 <sup>44)</sup> (図 4.20) T	0.834 (69.6)	0.807 (65.1)	0.681 (46.4)

の「あ」の場合、一番外側のポテンシャル値 0.04 の等ポテンシャル面だけでなく、内側の 0.08 や 0.06 など他の等ポテンシャル面も同時に比較する必要性を示している。一方、提案した誘導場の認識モデルは、図 4.8 に示すように複数の等ポテンシャル面における誘導場の歪みの総和を評価している。そこで、誘導場の認識モデルにおいて、弾性エネルギーを計算するポテンシャル値の範囲の違いがパターン類似性評価に影響するかどうか調べた。

表 4.15 は、検証に用いた実験パターンの変形系列 (4.3.3) について、データの枠に接触しないポテンシャル値 0.03 からポテンシャル値 0.20 までの低いポテンシャル値の範囲、およびポテンシャル値 0.21 から文字線分付近の 0.39 までの高いポテンシャル値の範囲の 2 つの場合に分けて、ポテンシャル値 0.01 ずつの各等ポテンシャル面で計算した弾性エネルギーによる相関係数と、ポテンシャル値 0.03 から 0.39 すべての弾性エネルギーを計算した場合の相関係数を比較した例である。表 4.15 の例を含め、弾性エネルギーを計算するポテンシャル値の範囲によって、相関係数が大幅に異なる場合がある。例えば、高いポテンシャル値の範囲における相関係数は、Shillman の  $V \rightarrow y$  (図 4.16) の場合弱い相関 (説明率 15%未満)、横瀬の実験では図 4.19 の変形系列 E、図 4.20 の変形系列 T の場合は中程度の強さ (説明率 15%以上から 50%未満) である。一方、低いポテンシャル値の範囲における相関係数は、ポテンシャル値 0.03 から 0.39 すべての弾性エネルギーを計算した場合の相関係数と同程度の強い相関 (説明率 50%以上) が得られる。このことから、誘導場の認識モデルにおいて、ポテンシャル値 0.03 からポテンシャル値 0.20 までの低いポテンシャル値の範囲が変形パターンなど多様なパターンについて、人間に近い類似性評価を可能にする上で重要と考えられる。したがって、各等ポテンシャル面ごとの評価の仕方を解明すれば、更にパターン類似性評価性能を向上させることが期待できる。

#### 誘導場間の変位計算

誘導場モデルによるパターンの評価結果は心理実験結果と高い相関を示している。しかし、全ての心理実験結果について、必ずしも心理実験結果と強い相関が得られていない。ダイナミクスによる認識モデルにおけるパターンの変形操作は、回転や並進などの複合と考えられている<sup>34)</sup>。しかし、提案した誘導場の認識モデルは、変形を法線方向に限定したため、複雑な誘導場の変形を充分記述することができず、心理実験との違いが生じたと考えられる。実際、変位をベクトルとして表示、観察すると、図 4.35(a) の A, B, C

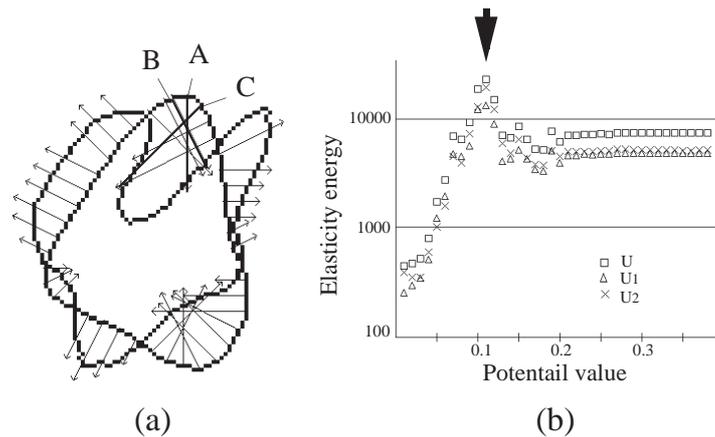


図 4.35 変位が極端にずれている場合

の変位のように、局所的に不適切な変位が生じている場合があり、その部分の弾性エネルギーが図 4.35(b)の矢印のように突出するため、類似性評価を悪くしている。

人間は、網膜に映る図形が大きさ、回転などの変形を受けていても、それらを同じものと知覚することができる恒常性があり、実際、「場」と恒常性の機構との関連性が示唆されている<sup>107)</sup>。そして、「場」はわずかな差異は一様化し、ある程度以上の性質の違いは反発して別になるよう、場全体が自動的により単純で最も安定した状態に遷移するのが基本と考えられている<sup>24)</sup>。したがって、誘導場の認識モデルでは、例えば、文字の大きさ・回転などの変形を通常の文字認識の前処理のように個々に行うのではなく、すべての変形を区別せず一括して誘導場全体を自動的に変形させるのが妥当と考えられる。具体的には、エネルギー最小化原理を用いて、変位で生ずる弾性エネルギーができるだけ小さくなるよう誘導場全体を変形させる変形操作、例えば Snakes<sup>118)</sup>のように、変位方向を図形の形状に対して適応的に決定するコンピュータビジョン的な手法が有効と考えられる。

## 第5章 誘導場による手書き文字切り出し

### 5.1 従来の文字切り出し

#### 5.1.1 手書き文字の状態

文字認識では最初に文字列から個々の文字を切り出す、文字切り出しが不可欠である。しかし、我々が日常的に筆記する手書き文書において、個々の手書き文字が完全に1文字ずつ離して書くことは少ない。したがって、一般に手書き文字列から文字を切り出すことは非常に大変な作業となっている。図 5.1 は、大学の試験答案<sup>注1</sup>の手書き文字について文字線分の接触状況を示した例である。表 5.1 は、図 5.1 の接触状況を分類、その頻度を調べた結果である（調査文書数 100 名分、1 文書当たりの 221 文字）。同様な調査として、目視による A4 の罫線ノートに記入した手書き文字（約 40000 文字）の接触状態が報告されている<sup>119</sup>。この調査によれば、従来の文字画像の射影を利用した方法で切り出せるのは全体の 58.2%に過ぎない。

#### 5.1.2 文字切り出し

5.1.1 で述べたように、手書き文字文書では 20%から 40%もの文字が非常に近接しており、文字の切り出しが難しい状況になっている。したがって、文字間にある有意な間隙を文字画像の射影から推定し、文字領域を矩形で近似する現在の文字切り出し技術では、文字間隔が狭い手書き文字列等から個々の文字を十分な精度で切り出すことは難しい。

印刷文字の切り出し方法として、黒画素の融合と縮退操作を繰り返す方法<sup>120, 121</sup>や、文字が接触している場合、文字間隔を文字行幅で近似して接触していない文字を優先する方法<sup>122</sup>などが提案されている。印刷文字の場合、文字間隔や文字の大きさなどの情報を文字の切り出しに利用することが出来る。

しかし活字に比べ、図 5.2(a) に示すように手書き文字は隣接する文字が接近している場合が多い。例えば「食」と「塩」の文字画像の射影に切れ目がないので、この場合、文字の大きさは同一文字列でほぼ一



図 5.1 手書き文字の接触状態の例

<sup>注1</sup> 静岡大学理学部物理学教室の中山潔教授より提供して頂いた。

表 5.1 手書き文字の接触状態の頻度

接触状況		頻度	図 5.1 の例
周辺分布で切り出せる		83.31%	(a)
周辺分布では 切り出せない	接触なし	直線で切り出せる	(b)
		曲線で切り出せる	(c)
	接触あり	点接触	(d)
		ストローク融合	(e)

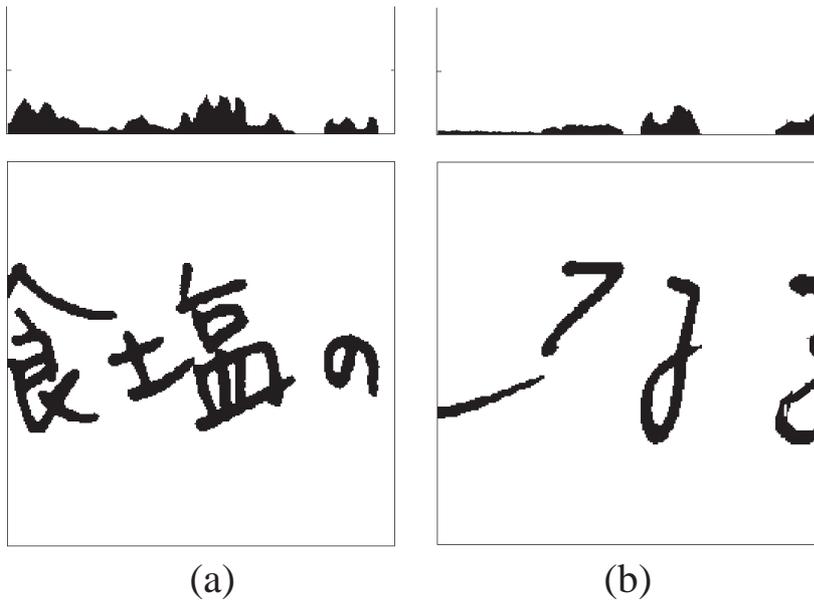


図 5.2 手書き文字列の射影

定で、文字は正方形であるなどの図形的な知識を用いて切り出す方法<sup>123)</sup>等が検討されている。また、図 5.2(b) の「な」の例や、漢字で偏と旁の間隔が開く場合は、射影に切れ目が生ずるので、各々を別々の文字として切り出してしまうこともある。このように、射影の切れ目を文字の間隙として 100%信頼することは出来ない。

そこで認識結果や単語知識等を用いて総合的に切り出す方法が提案されている<sup>124)</sup>。しかし、知識を使う方法では、字種が非常に多い日本語の場合、処理量が著しく増大し適切な結果を得るのは容易ではない。また自由手書き文字は字形や文字の変動が大きく、文字の大きさなどの図形的な性質が仮定できないことが多いので、認識や図形的な知識によって手書き文字の切り出し精度を大幅に向上させることは難しいと考えられる。従って手書き文字を十分な精度で切り出すには、文字の存在領域自体を精密に推定できる方法を検討する必要がある。それには従来使われてきた文字の外接矩形に比べ、一般的な閉曲線は文字の存在領域をより精密に決定する上で有利ではないかと考えられる。

一方、アルファベットの文字間隔が狭くなると、文字と文字との識別が難しくなる現象が、各文字の視覚の誘導場が融合することによって説明されている<sup>125)</sup>。これは視覚の誘導場を用いることにより文字領域を閉曲線によって規定する文字切り出しの可能性を示唆するものと考えられる。

しかし、従来の視覚の誘導場に関する研究では、単一の図形の周りの「場」が、どのように分布しているかに多くの関心が払われていた。そこで、視覚の誘導場による文字切り出しを実現するには、まず複数の文字の作りだす視覚の誘導場について分析を行ない、視覚の誘導場がどのように文字領域を規定するかについて基礎的な検討をする必要がある。

ここでは、最初に複数の文字のつくる視覚の誘導場と文字間隔に関する知見および実際の手書き文字における視覚の誘導場の分析から、視覚の誘導場を使った文字切り出し方法を提案する。そして提案した方法に基づき、実際の手書き文字列からの文字切り出し実験結果を示し、視覚の誘導場による手書き文字切り出しの可能性を示した。

## 5.2 誘導場による文字切り出し方法の提案

### 5.2.1 複数パターンによる誘導場の分布

文字間隔が極めて近くなると、個々の文字が識別しにくくなる現象を、視覚の誘導場を用いて説明できることが報告されている<sup>125)</sup>。例えばアルファベットの'H'と'I'がある程度接近すると、それぞれの文字の視覚の誘導場が融合して分離なくなり、'H'または'I'と識別しにくくなる。そして実験したパターンの大きさの場合、分離できるか否かのポテンシャル値は0.4であることが報告されている。この報告から、視覚の誘導場の分布が相異なる複数の文字の分離に関係していると考えられる。

文字の分離と視覚の誘導場の関係を、逆に文字を書く立場から考えると、読み易い文字というのは、文字の視覚の誘導場の分布がなるべく分離するように心掛けて筆記された文字と考えられる。これは、各文字のつくる視覚の誘導場が干渉して、互いに融合しないように作用しているとも言える。

この関係を手書き文字の切り出しに応用するため、まず日本語の手書き文字に対して同じことが言えるかどうか実際に調べる必要がある。そこで、最初に文字列周辺の視覚の誘導場の分布を調べて(1)文字は、視覚の誘導場によってどのように囲まれているのか(2)文字を囲む視覚の誘導場は、どのように決定されるのか、について検討を行なった。

### 5.2.2 手書き文字列の作成と文字の隣接状態の分析

通常、まず文書画像から文字列が抽出され、次に文字列から文字が切り出される。文字列の存在領域は、手書き文字の場合でも、文書画像の射影等を用いて推定可能なので、既に抽出された文字列に関して視覚の誘導場を用いて文字の切り出しを行なうものとする。今回、視覚の誘導場がどのように文字領域を規定するかについて基礎的な検討を行なうためのデータとして、人為的に作った文字列データと実際の手書き文書からの文字列データを用いた。

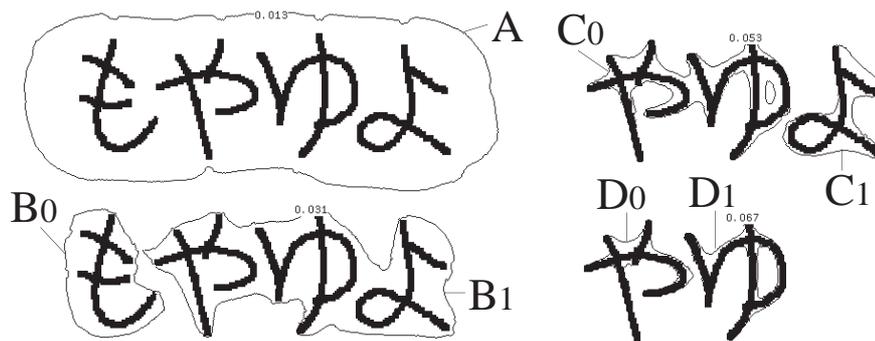


図 5.3 複数の文字による視覚の誘導場の例

#### 人為的に書いた文字列データの分析

まず日本語の文字列がつくる視覚の誘導場はどのような分布になっているのか基本的な解析を行なうため、ワークステーション上で手動で誰でもが読み易いように丁寧に書いた楷書体の平仮名の文字列データを作成した。文字列は、 $512 \times 512$  dot の大きさの枠内の中央に  $120 \times 120$  dot 程度の大きさの無意味な文字を3から5個含んでいる。実際の文字列では文字認識などの高次の情報が切り出しに関連している可能性がある。そのような文字列データのみによる分析では、誘導場による文字切り出しの本質が不明瞭になる恐れがある。そこで実際の文字列データとは対照的なデータとして、このような人為的なデータを用意した。

人為的につくった文字列の誘導場の例として、文字列「もやゆよ」がつくる視覚の誘導場を図 5.3 に示す。視覚の誘導場をポテンシャル値が低いほうから高いほうに見ていくと、まず全文字列を包絡線のように囲む閉曲線  $A$  が得られ、続いて「も」のみを囲む閉曲線  $B_0$  と「やゆよ」を囲む閉曲線  $B_1$ 、そして「やゆ」を囲む閉曲線  $C_0$  と「よ」のみを囲む閉曲線  $C_1$ 、最後に「や」「ゆ」をそれぞれ囲む閉曲線  $D_0$ 、 $D_1$  が得られる。ポテンシャル値を更に高くしても、図 5.3 に示した以上に更に細かい閉曲線に分離することはない。図 5.3 の例を含め、文字列中の個々の文字は全てある閉曲線で囲まれており、それらの閉曲線のポテンシャル値は文字間隔によって異なっている。例えば「も」は、「や」と「ゆ」の間隔に比べ「や」と離れているので、低いポテンシャル値 (0.031) で他の文字から分離する。逆に接近している「や」と「ゆ」は、「も」や「よ」に比べて高いポテンシャル値 (0.067) でないと分離しない。以上のように日本語の場合も、アルファベットと同様に視覚の誘導場の分布から文字領域を知ることができると思われる。

#### 実際の手書き文書からの文字列データの分析

次に実際に書かれた手書き文字の視覚の誘導場を分析するために、5.1.1 と同じ大学の試験答案文書から文字列データを作成した。実際に書かれた手書き文字の切り出しの検討では、文字間隔や隣接文字との接触程度等に意図的な制限が全くないものが理想である。この条件を満たすため、筆記者がデータ収集の目的を意識していない前述の文書からデータを作成した。

データ作成方法は、文献<sup>126)</sup>による方法と同一である。文字の切り出しの際に同時作成される周辺画像

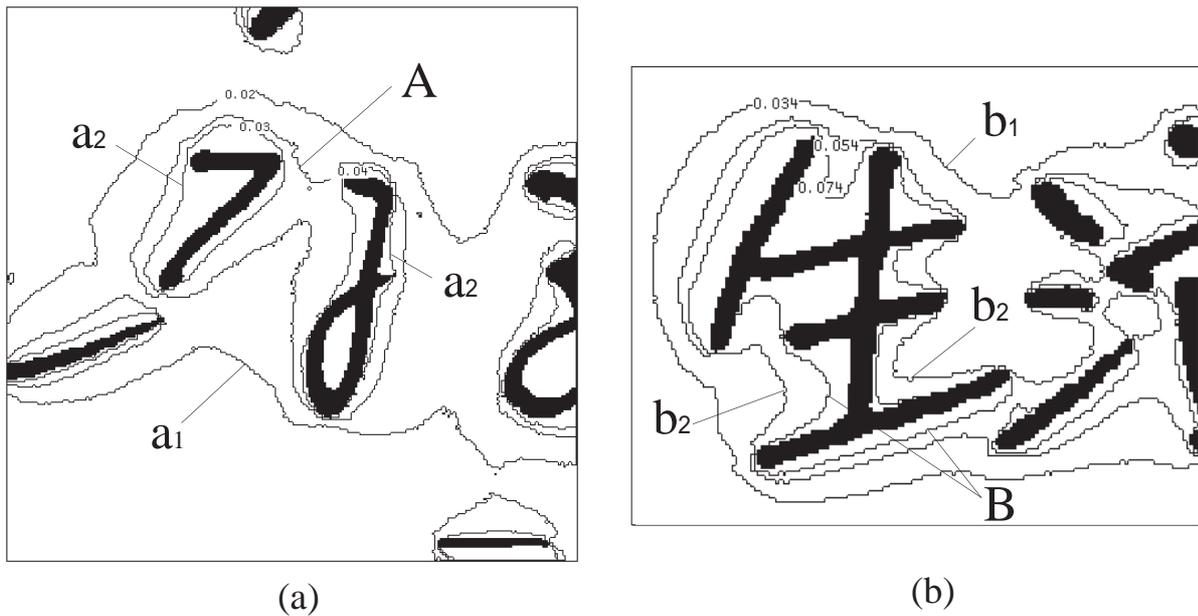


図 5.4 手書き文字列周辺の誘導場の分布例

を文字列データとした．周辺画像の範囲は，切り出した文字の大きさを  $n \times n$  とした場合，元の文字の大きさを中心とする  $2n \times 2n$  の範囲である．文字列データには切り出された文字自身の他に周辺の文字や文字の線分の一部などが含まれる．なお文字の大きさは  $n = 64 \text{ dot}$  から  $256 \text{ dot}$  である．データの大きさの正規化，画像の修正，雑音除去等を行っていない．

作成された文字列データから，今回の検討に適したものを以下のように選択した（１）隣接文字同志の文字線分が接触（重複，交差）していないもので（２）罫線が文字成分を大きく損なわないものから選択した．隣接文字同志の文字線分が接触している場合，接触している文字線分がどちらの文字に所属するかは，文字認識などをしないと判定できない．今回は少なくとも文字同志が空間的に離れているものを対象とした．罫線は複数の文字のつくる視覚の誘導場と関係がないので除去した．なお，同じ筆記者から複数の文字列データが選択されないように注意した．これは同一の筆記者の筆記する文字の間隔が，ある傾向をもっていると考えられるからである．こうして作成した全ての文字列データは文字列画像の射影に切れ目がなく，文字列が短いので，射影や文字列の幅などの情報を使って文字を切り出すことは難しい．

このようにして選択された文字列データについて視覚の誘導場を求め，誘導場の分布と文字領域の関係を分析した．その結果，次の２つのことがわかった．

第１点は，ある単一閉曲線が１つの文字領域を規定している点である．例えば，図 5.4(a) の「な」の構成要素は互いにかかなり離れているので，要素間の関係が希薄のように見える．その上，隣接する文字は「な」に比較的接近しているので，他の文字要素と「な」の文字要素の誘導場が融合するように考えられる．しかし，視覚の誘導場を観察した結果「な」は１文字として単一閉曲線  $A$  で囲まれている．

第２点は，手書き文字の場合でも文字を分離する単一閉曲線のポテンシャル値は画像によって最適値が異なる点である．元のデータの大きさが違う場合，全く相似な図形であっても視覚の誘導場の分布が変わ

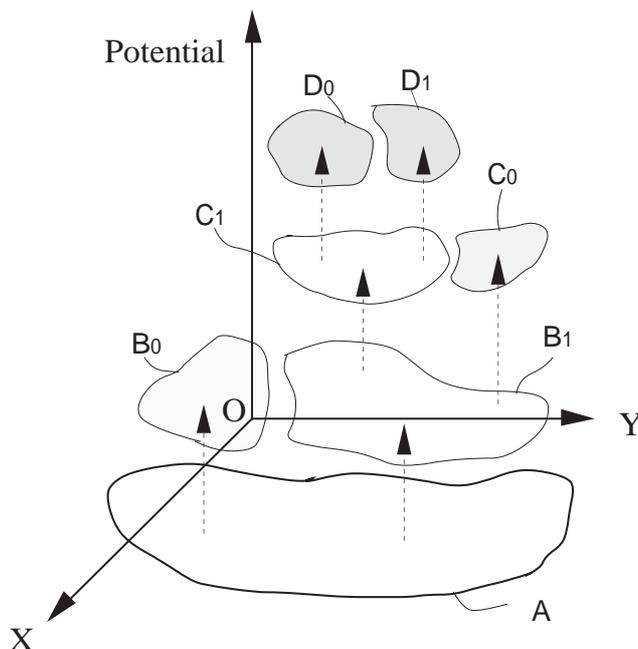


図 5.5 閉曲線の階層構造

る．また文字の構成要素による違いも影響する．例えば，図 5.4(a) の「な」の場合，文字を分離するポテンシャル値は単一閉曲線  $A$  の 0.030 であるが，図 5.4(b) の場合，「生」を囲む閉曲線  $B$  のポテンシャル値は 0.074 である．ポテンシャル値が 0.030 程度の場合，図 5.4(b) で，その値に相当するのは曲線  $b_1$  であり「生」の誘導場が隣接文字の場と融合してしまう．このことから人為的な文字列同様，単一のポテンシャル値で文字を切り出す閉曲線を指定することは不可能である．

#### 文字を分離する閉曲線

人為的に作った文字列と，実際の手書き文字の文字列の視覚的誘導場の分析から，閉曲線は階層構造をもっており閉曲線に囲まれた領域内に必ず文字や文字要素が存在している．図 5.5 は図 5.3 の例について閉曲線の階層構造を示したものである．図 5.5 で同じアルファベットをもつ閉曲線は同じ等ポテンシャル面に存在する．図 5.3 の例のように，ある文字内でその文字の要素が分離していない場合は，図 5.5 に示すような階層構造をもつ閉曲線群の最も末端の閉曲線 ( $B_0, C_0, D_0, D_1$ ) 各々に文字が 1 つずつ存在する．

一方，図 5.4(a) の例 (閉曲線  $a_1, a_2$ ) のように文字内で文字要素が分離して，閉曲線群の最も末端の閉曲線には文字ではなく分離した文字要素が含まれる場合がある．分離した文字要素を囲む閉曲線の形状は元の文字の閉曲線に比べ，細長く複雑になる傾向が見られる．そこで図 5.6 に示す手書き文字列の文字の分離に伴う閉曲線の形状の変化について，複雑度 (Complexity)<sup>127, 128)</sup> と周囲長 (Perimeter)<sup>127, 128)</sup> を尺度に用いて調べた．ある閉曲線の境界点の個数を周囲長  $l_i$ ，閉曲線の内側に存在する画素総数を面積  $S_i$  とすると，複雑度  $C_i$  は次式で与えられる<sup>127, 128)</sup> ．

$$C_i = \frac{l_i^2}{S_i} \quad (5.1)$$

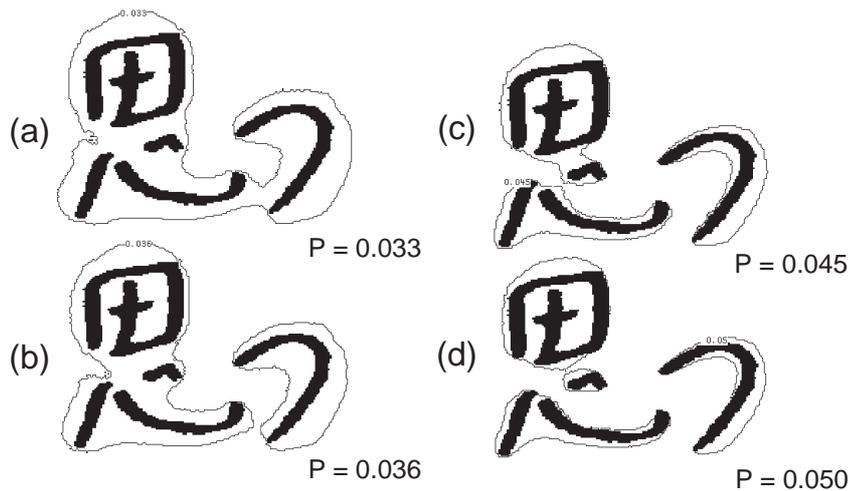


図 5.6 文字要素の分離

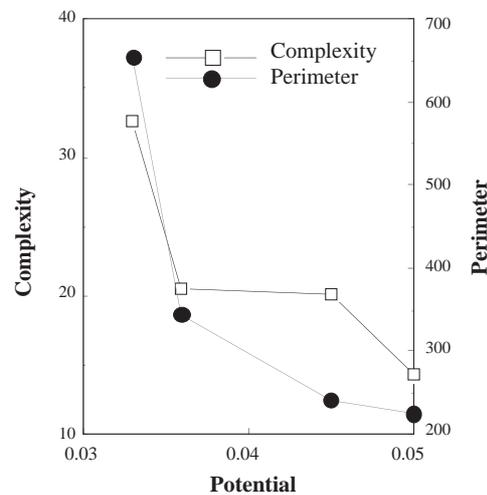


図 5.7 閉曲線群の形状の推移

式 (5.1) より、複雑度の単位は無次元である<sup>128)</sup>。複雑度は図形が円に近いと最も小さな値をとる。図形が複雑になるほど大きな値をとる<sup>127, 128)</sup>

図 5.7 は、図 5.6 の文字列が分離するときの等ポテンシャル面ごとの全閉曲線の複雑度と周囲長の平均を示したものである。図 5.6, 5.7 から、「思」が隣接する文字「つ」から分離したとき、複雑度と周囲長は大幅に減少するが、「思」が文字要素「田」と「心」に分離すると複雑度と周囲長はあまり変化しない。図 5.6 の例を含め、文字から文字要素が分離した場合、閉曲線の複雑度と周囲長は個々の文字に分離する前に比べ大幅に変化せず、ほぼ同じ程度の値で推移する傾向が見られる。このことから、なるべく低いポテンシャル値で円に近い状態の閉曲線が、個々の文字要素ではなく文字全体を規定する閉曲線であると考えられる。円に近い図形の複雑度は 10 から 20 程度、文字列のように長細い場合は 30 から 40 程度であるので、閉曲線の複雑度が 10 以上減少すれば文字列から文字が個々の文字や文字群に分割した状態と考えられる。

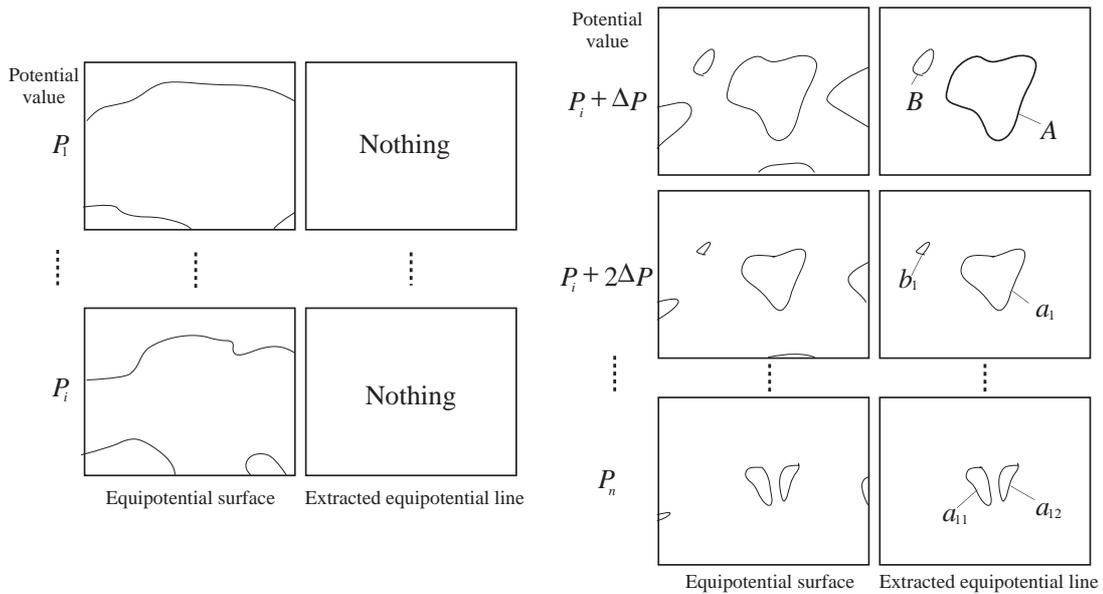


図 5.8 等ポテンシャル面からの閉曲線の抽出

### 5.2.3 視覚の誘導場による文字切り出し方法

以上の分析結果から視覚の誘導場の閉曲線に着目した手書き文字の切り出し方法を提案する．なお文字の切り出しを行なう文字列データは，隣接する文字の線分が接触（重複，交差）していないと仮定する．

まず，文字列の視覚の誘導場を求め，ポテンシャル値が低いところから高いところへ順に等ポテンシャル面上の誘導場の分布を調べ閉曲線を抽出していく．図 5.8 のポテンシャル値  $P_1$  から  $P_i$  のように，等ポテンシャル面が文字画像の枠に接する場合は閉曲線としない．次第にポテンシャル値を高くし，枠に接することなく単独で存在する閉曲線を探す．例えば図 5.8 において，ポテンシャル値が  $P_i$  から  $\Delta P$  増加した  $P_i + \Delta P$  で初めて 2 つの閉曲線  $A, B$  が得られたとする．更にポテンシャル値を徐々に  $\Delta P$  ずつ増加させ，複数の閉曲線が得られる．ここで，図 5.8 上の閉曲線は次のような閉曲線の位置関係にもとづく包含関係をもっている．

$$A \supset a_1, a_1 \supset a_{11}, a_1 \supset a_{12} \quad (5.2)$$

$$B \supset b_1 \quad (5.3)$$

図 5.5 のように，閉曲線の抽出過程で，ある閉曲線が分岐して閉曲線数が増加するたびに次式の包含関係に基づき閉曲線群  $A, B, C, \dots$  について分岐した閉曲線のラベル付けを図 5.9 のように行なう（ $i, j, k$  は整数）．

$$A_i \in A, A_{ij} \in A_i, A_{ijk} \in A_{ij}, \dots \quad (5.4)$$

$$B_i \in B, B_{ij} \in B_i, B_{ijk} \in B_{ij}, \dots \quad (5.5)$$

もし文字要素が分離しなければ，基本的に図 5.9 の斜線で示した分岐末端の閉曲線に個々の文字が存在する．

次に閉曲線が分岐するごとに閉曲線の丸さを評価し，文字要素が分離せずに 1 つの文字として存在する閉曲線を推定する．式 (5.1) の複雑度  $C_i$  を閉曲線の丸さを評価する尺度として，閉曲線  $x_i$  の評価関数

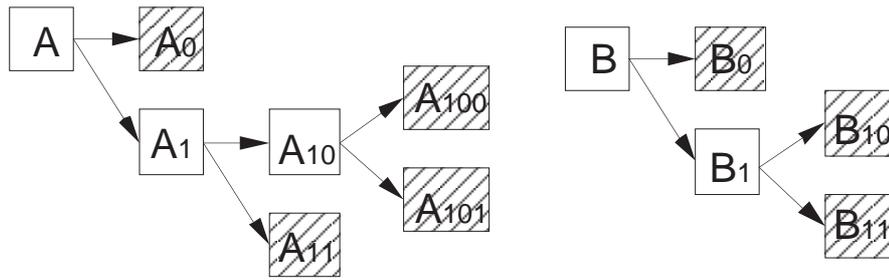


図 5.9 閉曲線の包含関係

$f(x_1, x_2, \dots, x_n)$  を次のように定義する .

$$f(x_1, x_2, \dots, x_n) = \frac{1}{n} \sum_i^n C_i \quad (5.6)$$

文字が1つの文字として存在するかどうかの判定は閉曲線群ごとに行ない、なるべく低いポテンシャル値で丸くなっている閉曲線を決定する。例えば、図 5.8 の閉曲線群  $A$  において、 $A$  が  $A_0$  と  $A_1$  に分岐する際は  $f(A)$  と  $f(A_0, A_1)$ 、 $A_1$  が  $A_{10}$  と  $A_{11}$  に分岐する際は  $f(A_0, A_1)$  と  $f(A_0, A_{10}, A_{11})$ 、 $A_{10}$  が  $A_{100}$  と  $A_{101}$  に分岐する際は  $f(A_0, A_{10}, A_{11})$  と  $f(A_0, A_{11}, A_{100}, A_{101})$  を比較する。分岐によって式 (5.6) による複雑度の評価結果が 10 以上向上した場合は適切な文字領域の分割が行なわれたと考え、更に分岐した場合について評価する。評価の結果の向上が閉曲線の分岐前と同程度 ( $\pm 5$  程度) で、周囲長  $l_i$  が 50% 以下に減少している場合は、複雑度の定義から同程度の大きさの複数の文字に分離していると考えられるので、更に閉曲線が分岐した場合についての評価を続ける。しかし評価の結果が全く向上しない場合、例えば  $A_1$  が  $A_{10}$  と  $A_{11}$  に分岐するとき、分岐するひとつ手前の  $A_0$  と  $A_1$  の閉曲線それぞれに文字が1つずつ存在し、 $A_1$  以降分岐する閉曲線には文字要素のみが存在すると考えて評価を終了する。閉曲線群  $B$  についても同様の評価を行ない文字が存在する閉曲線を推定する。文字の切り出しは、このように推定した閉曲線内に存在する黒画素を求めることによって可能になる。

### 5.3 手書き文字切り出し実験

提案した方法に基づき、5.1.1 と同じ大学の試験答案から作成した文字列データの文字切り出し実験を行なった。データの作成方法は、5.2.2 で行った実際の手書き文書からの文字列データの分析と同様である。これらの文書は今回のような実験を目的に収集されたものではなく、通常筆記される手書き文書である。

実験の結果、55 例中 42 個、切り出し率 76.4% と、多くの文字列データから文字を切り出すことが出来た。その例を図 5.10、5.11 に示す。デジタル画像の視覚の誘導場は凹凸が大きいので、メジアンフィルタを 4 回かける平滑化処理を行なった後に切り出しを行なった。図 5.2 に示したように、射影で切り出すことが難しい文字でも、視覚の誘導場の閉曲線が文字全体を包絡線のように囲み、文字領域が閉曲線によって精密に規定されて文字が切り出されることが解かる。特に図 5.11 で、「食事」の「事」や「し」、そして「な」の細くなっているストロークの先端が他の文字と接近しているにもかかわらず、「食」「し」「な」のそれぞ

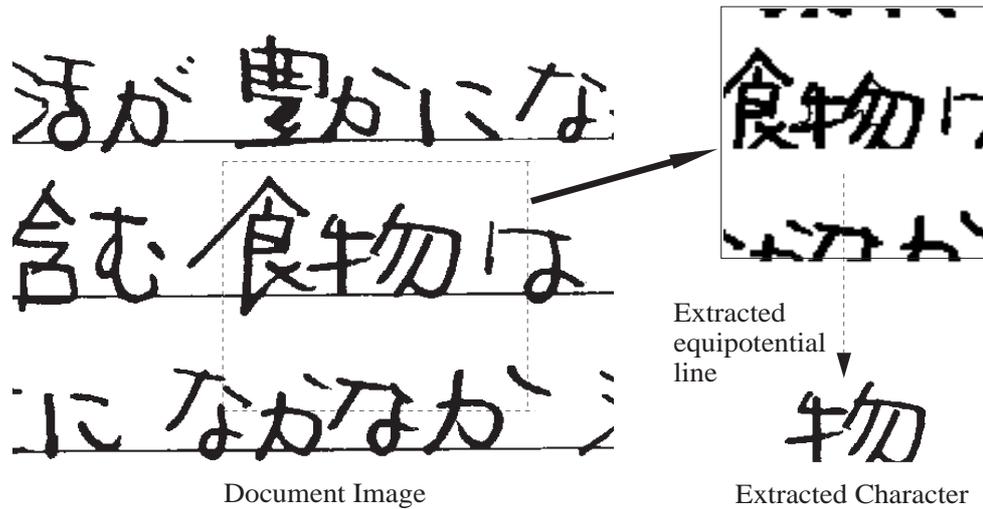


図 5.10 手書き文書からの文字の切り出し例

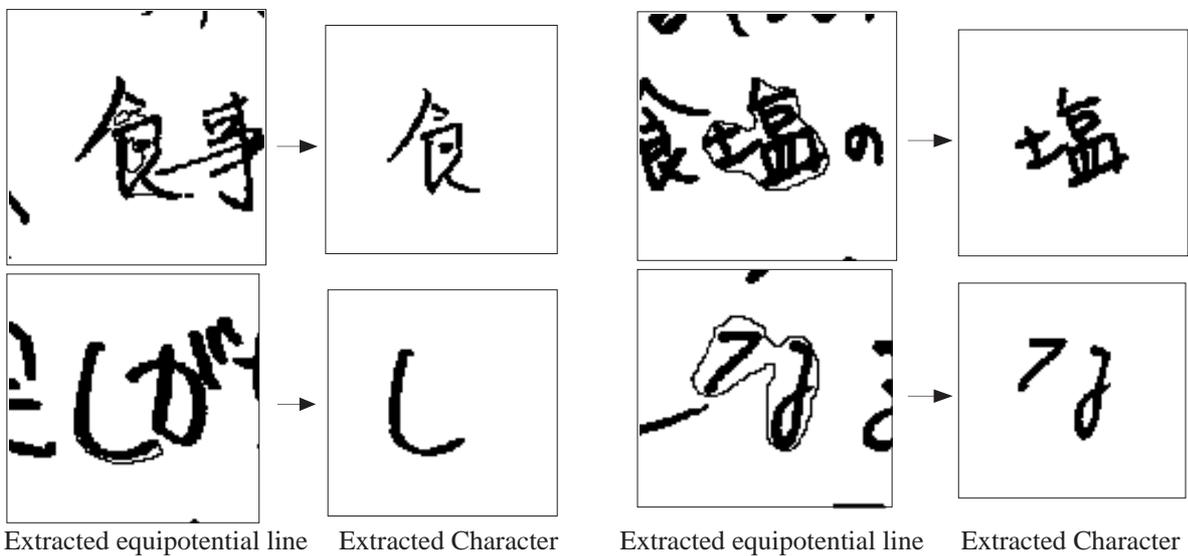


図 5.11 手書き文字の切り出し例

れの文字全体を囲む閉曲線が存在する。

しかし、隣接文字同志の成分が極めて接近していると、個々の文字を分離するような等ポテンシャル線が存在しない場合があり、文字を切り出すことが出来なかった。例えば、図 5.12 の「生」と「活」の視覚的誘導場のように、「生」のストロークの一部が「活」の一部と融合しており、「生」と「活」の各文字のみを囲む閉曲線は存在しない。一方、文字要素が大きく離れている図 5.12 の「い」の場合、文字要素全体を囲む閉曲線が存在しないので、文字「い」として文字を切り出すことが出来なかった。しかし、我々もこのような文字列は非常に読みにくいことも確かである。

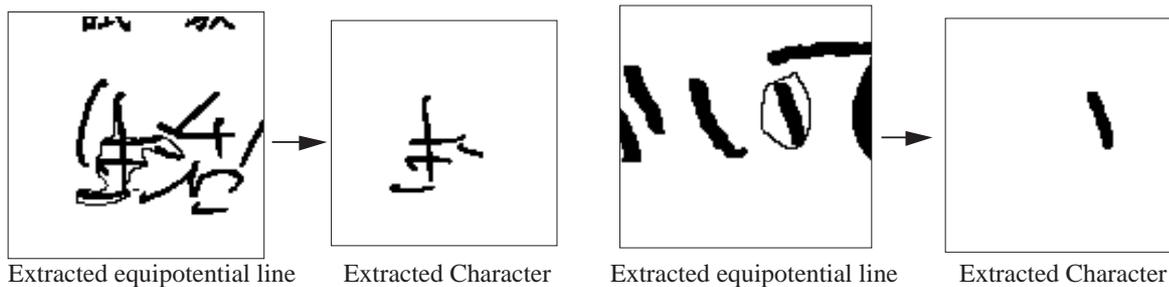


図 5.12 手書き文字の切り出し例（失敗した場合）

## 5.4 考察

### 5.4.1 誘導場による文字の切り出し

手書き文字切り出し実験結果から、複数の文字がつくる視覚の誘導場の階層的な包含関係をもった閉曲線群が文字を包絡線のように囲み、文字の存在領域を精密に規定するので、誘導場の閉曲線が文字を切り出す上で重要な役割をもっていることが明らかになった。

### 5.4.2 手書き文字切り出しにおける誘導場の意義

では、なぜ閉曲線の領域に文字や文字要素があるのだろうか。視覚の誘導場はクーロンポテンシャルとして理解できる。文字が単独に存在する場合、その文字を囲む閉曲線しか存在しないが、周りにその文字以外の文字がある場合、視覚の誘導場はそれらと相互作用する。作用は文字同志が近い場合は強く、離れるに従って小さくなる。5.2.1で、読み易い文字の視覚の誘導場の分布がなるべく分離するように文字が書かれることを述べたが、それが誘導場の融合や分離として現われている。

図 5.11 で「事」「し」「な」の細くなっているストロークの先端が他の文字と極めて接近しているにもかかわらず、「食」「し」「な」のそれぞれの文字全体を囲む閉曲線が存在する。視覚の誘導場はクーロンポテンシャルとしての性質をもっているため、境界条件が急激に変化する細いストローク先端の周辺では、場が非常に強くなる。従って細いストローク同志が対向していたり、細いストロークが他の文字要素の法線方向上にある場合、文字要素間の距離が小さい場合でも文字が分離される。一方、文字要素が平行に分布している場合は文字要素間の場は緩やかに変化するため、それぞれの文字要素はある程度離れていないと分離されない。以上のように文字を囲む閉曲線は、視覚の誘導場のクーロンポテンシャル的な性質から必然的に現われる現象として定性的に理解することが出来る。

### 5.4.3 手書き文字切り出し性能向上のための課題

文字要素が分離する場合、閉曲線の複雑度や周囲長を評価して閉曲線の丸さがあまり変化しない閉曲線を探すことで文字を切り出すことが出来た。誘導場を薄いゴム膜のような弾性体として考えると、なるべく低いポテンシャル値で閉曲線が丸くなるというのは、閉曲線を縁とする薄膜の弾性エネルギーが出来るだけ

小さくならうとしているものと考えられる．従って文字は弾性エネルギーが最小になる状態を実現している閉曲線に含まれる可能性が高いと言える．このことは文字は本来1つ1つ孤立して存在するので、文字を構成する文字要素がなるべく1つの文字としてまとまろうとする性質があるのではないかと考えられる．一方、複雑度による閉曲線の評価は適切な値をどのように決定するかという問題がある．そこで膜の弾性エネルギーを閉曲線の評価に用いれば、複雑度などの直観的な尺度によらず、より誘導場のもつ性質にもとづいた客観的な閉曲線群の評価が可能になると考えられる．

図 5.12 の例のように他の文字の成分と閉曲線が融合または完全に分離している場合、抽出した閉曲線群の評価だけで文字を切り出すことは出来ない．これは、必ずしも全ての文字が 5.2.1 で述べたように、各文字の誘導場が干渉して融合しないように文字が書かれていないことを示していると考えられる．事実、このような文字列は大変読みにくい．しかし少なくとも文字要素は途中で切れたりすることなく、必ずある閉曲線によってその存在領域は規定される．そこで、抽出した閉曲線を文字要素とする候補文字を複数構成し、構成した各文字の閉曲線群のうちで弾性エネルギーが最小になるような閉曲線群を文字として適切であると評価して文字を切り出すなどの方法が考えられる．

## 第6章 誘導場による文字列の読み易さの評価

### 6.1 従来の文字の読み易さの定量評価

文字の大きさや間隔などを調整して、文字列を読み易く表示する一般的な方法は未だ明らかにされていない。1つ1つの文字自体が読み易い条件はこれまで数多く明らかにされている。例えば、活字の場合、書体のデザイン<sup>129, 130</sup>、紙とインキの明度差<sup>130</sup>、可読性と文字の大きさの関連性<sup>130, 131</sup>、2値よりも階調のある文字が読み易い<sup>132</sup>などが明らかにされている。文字列の場合、文字種・文字の大きさ・文字間隔などを調整することによって、文字列が読み易くなると考えられており、例えば、仮名は漢字より間隔を詰めても読み易いなどの傾向が示されている<sup>133</sup>。また、一般に和文印刷では、文字間隔を文字幅の約半分に設定、句読点など一部の文字の間隔を詰めることが多い<sup>134, 135, 136</sup>。欧文の場合、例えば隣り合う文字が、NMやVAのように互いに平行かVNのように1つの線が違っているか、VT、AAのように両方の線が輻輳しているかに従って変えるのがよいとされている<sup>136</sup>。

このように、特定の文字種を中心とした文字間隔の詰め方など、主観的に決められた文字列表示の経験則が存在する。しかし、任意の文字種について、どのくらいの大きさの文字をどのくらい間隔を詰めても読み易いかなどを決める定量的な評価方法は、ほとんど明らかにされていない<sup>136</sup>。したがって、読み易い和文の文字列表示を実現するには、任意の文字種について、読み易い文字の大きさ、文字間隔を定量的に導き出す基準や方法を検討する必要がある。

さて、文字列内の個々の文字が読み易い状態とは、個々の文字が醸し出している文字の周囲に波及する視覚の誘導場が、できるだけ干渉しないような間隔で文字が配置されていることが示唆されている<sup>44</sup>。そして、前章の手書き文字の切り出しで述べたように、個々の文字を囲む誘導場の閉曲線が存在し、そのポテンシャル値が高いと他の文字との分離が難しく、読みにくいことが示されている。また、図形や文字の配置が均整かどうかについても誘導場を用いて評価できる可能性が示唆されている<sup>79, 80, 137</sup>。このことから、誘導場の分布を基準にすれば、人間が判断する文字列の読み易さを定量的に評価できると考えられる。

そこで、最初に文字種、文字の大きさ、文字間隔が異なる日本語の文字列における個々の文字を囲む誘導場の閉曲線の強さが、人間が判断する文字列の読み易さと関連するか調べ、文字列の読み易さの指標としての妥当性を検証する。次に、文字列の誘導場の閉曲線の強さは、個々の文字が単独に存在する場合の誘導場の分布から計算により予測できることを示す。

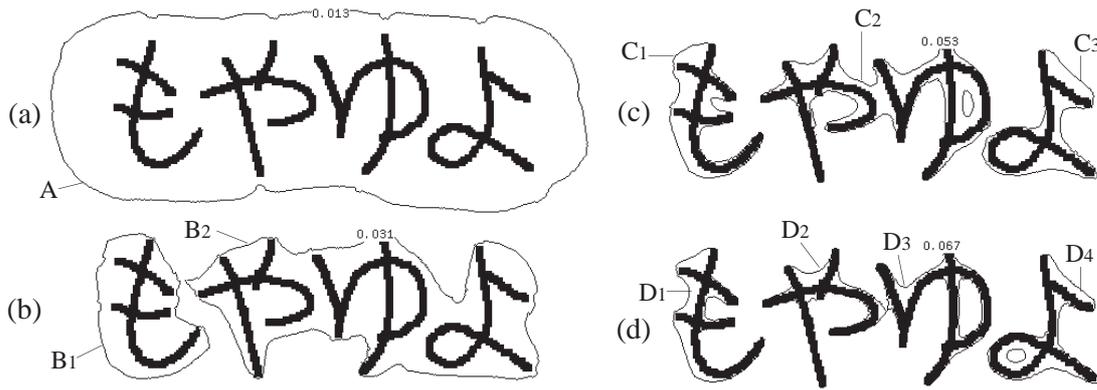


図 6.1 文字列における視覚の誘導場

## 6.2 文字列における視覚の誘導場

### 6.2.1 文字列の視覚の誘導場の分布と文字配置

複数の文字で構成された文字列の誘導場の分布と文字配置の関係を検討する。現在、和文書の大半は横書き文書で、縦書きは文芸書やはがきなどに限られることが多いため、本論文では横書きの文字列について検討する。最初に、図 6.1(a) の文字列「もやゆよ」の例のように上下に別の文字列がなく、一文字列が単独で存在し個々の文字が接触しない文字列について検討する。

図 6.1 は文字列「もやゆよ」が作る視覚の誘導場を、図 6.1(a) はポテンシャル値 0.013, (b) は 0.031, (c) は 0.053, (d) は 0.067 のように徐々に大きくした場合の例である。ポテンシャル値が低い図 6.1(a) の場合、文字列全体を包絡線のように囲む閉曲線 A が存在する。ポテンシャル値が大きくなると、まず図 6.1(b) のように「も」のみ囲む閉曲線 B<sub>1</sub>、「やゆよ」を囲む閉曲線 B<sub>2</sub> が出現する（ポテンシャル値  $\alpha$  0.031）。さらにポテンシャル値が大きい図 6.1(c) の場合、「やゆ」を囲む閉曲線 C<sub>2</sub> と「よ」のみを囲む閉曲線 C<sub>3</sub> が出現する（ポテンシャル値  $\beta$  0.053）。最後の図 6.1(d) の場合、「や」「ゆ」をそれぞれ囲む閉曲線 D<sub>2</sub>, D<sub>3</sub> が出現する（ポテンシャル値  $\gamma$  0.067）。個々の文字を囲む閉曲線のポテンシャル値は文字間隔によって異なっている。例えば「も」は、「や」と「ゆ」の間隔に比べ「や」と離れているので、低いポテンシャル値  $\alpha$  で「や」から分離する。逆に接近している「や」と「ゆ」は、「も」や「よ」に比べ高いポテンシャル値  $\gamma$  でないと分離しない。図 6.1 の例を含め他の文字列の誘導場の分布の観察から、文字間隔が狭く読みにくいと文字列を構成する個々の文字を囲む閉曲線が最初に出現する際のポテンシャル値は高くなる。すなわち、文字の間隔によって誘導場の干渉の強さに違いが生ずる。その違いは文字列を構成する個々の文字が分離する誘導場の閉曲線の強さで表されると考えられる。例えば、図 6.1 の場合、3 つのポテンシャル値  $\alpha$ ,  $\beta$ ,  $\gamma$  の平均を文字列「もやゆよ」の読み易さの指標に使うことができる。そこで、一般に  $n$  個の文字から構成される文字列の隣接する  $i, i+1$  番目 ( $i = 1, 2, \dots, n$ ) の文字が最初に分離する誘導場の閉曲線のポテンシャル値を  $x_i$  とすれば、文字列全体の読み易さは、次のポテンシャル値の平均  $\bar{M}$  を用いて評価する。

$$\bar{M} = \frac{1}{n-1} \sum_i^{n-1} x_i \quad (6.1)$$

### 6.2.2 複数の文字列による誘導場の分布

ある文字列の誘導場の分布に、上下に存在する他の文字列が及ぼす影響を検討するため、印刷に近いコンピュータ組版の和文 $\text{\LaTeX}^{135}$  (文書フォーマット jarticle) で作成した複数の文字列を含む画像 (320×320 dot, 72 DPI) の誘導場を計算し、図 6.1 のようにポテンシャル値を徐々に大きくしながら、各文字列を包絡線のように囲む誘導場の分布と、各文字列内の個々の文字が分離する誘導場の分布を比較した例が図 3.8 である。図 3.8 の例を含め、まず図 3.8 左側の文字列「般に有効な」「知過程はパ」「的な理論と」のように各文字列を囲む誘導場がポテンシャル値 0.02 ~ 0.03 程度で現れ、さらにポテンシャル値が 0.04 程度以上になると、図 3.8 右側の文字列「知過程はパ」の「パ」、「的な理論と」の「と」「な」などのように文字列内の個々の文字が分離する。

このように、文字列が複数の場合、各文字列の分離と文字列内の個々の文字の分離は別々に扱うことができるので、分離した各文字列について、図 6.1 の要領で式 (6.1) による文字列の読み易さを評価できると考えられる。そこで、ここでは文字列の読み易さを、図 6.1 のような単一の文字列について検討する。

## 6.3 文字列の読み易さと誘導場の関連性の検証

式 (6.1) による誘導場の強さが、人間が判断する文字列の読み易さを説明できることを検証する。ここでは、いろいろな文字種、文字の大きさ、文字間隔をもつ文字列を被験者に呈示して、読み易さを 5 段階評価した心理実験結果と式 (6.1) による誘導場の強さの相関を調べた。

最初に、文字の読み易さの心理実験方法について述べる。文字列が単語や熟語を含む場合、可読性が高くなる単語優位効果<sup>115)</sup> や、文字列が長い場合に生ずるサッカードなどの眼球運動<sup>138, 139, 140)</sup> が、文字列の読み易さの評価に影響することを避けるため、眼を動かさず一目で認識できる日本語の文字数の範囲<sup>138)</sup> である 5 文字から成る、熟語を含まない無意味な文字列を呈示した。そして、日本語文書の読みに関する心理実験では、視角を 1 文字あたり  $1^\circ$ <sup>138, 140)</sup> から  $2^\circ$ <sup>139)</sup> 程度に設定すると被験者が読み易いことから、呈示する 1 文字あたりの視角を  $2^\circ$  に設定した。

被験者に呈示した 3 種類の文字列の基本型を図 6.2 に示す。図 6.2(a) はすべて仮名、図 6.2(b) はすべて漢字、図 6.2(c) は仮名漢字混じりの無意味な文字列である。これらは、Macintosh 上で図 6.3(a) のように一辺の長さが 60 ポイント (72DPI) の正方形 5 個を間隔 0 mm で並べ、それぞれの正方形に正方形の一辺の長さと同じポイント数の Osaka フォントの文字を、文字の重心と文字が外接する正方形の中心を合わせて配列した (1 ポイント = 0.3514 mm<sup>130)</sup>)。以降の実験も解像度、文字のフォントはすべて同じである。図 6.2 の 3 つの各基本文字列について、文字の大きさ、文字間隔を変更した文字列を用意する。まず、図 6.3(a) と同じ要領で、文字の大きさを基本状態 (60 ポイント) の他に 56, 52 ポイントの文字列を構成する。さらに、これら大きさの異なる 3 種類の文字列について、文字間隔 0 mm の基本状態の他に、図 6.3(b) のように文字間隔を各文字の間隔を 1 mm 広げる、または、1 および 2 mm 詰める。このように、図 6.2 の 1 つ

ねすにれそ (a) 塩歩晴具堅 (b) ねす晴れそ (c)

図 6.2 呈示した文字列

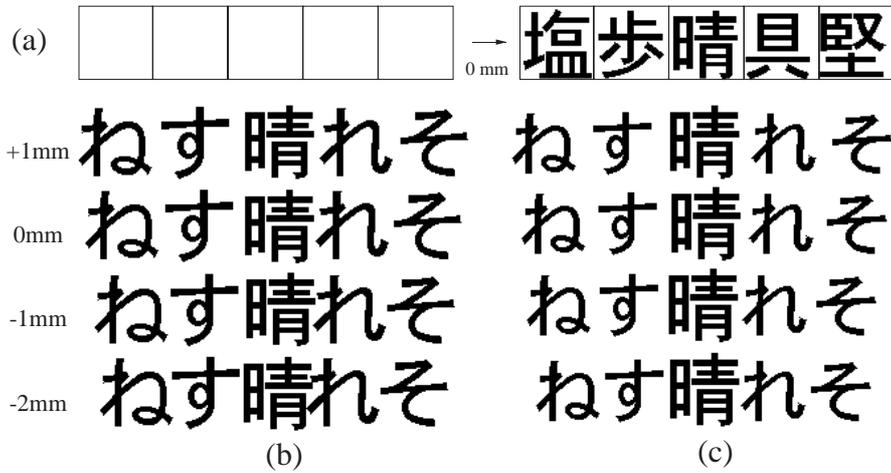


図 6.3 文字の配置と文字列の変更

の基本文字列について、3種類の文字の大きさ（60，56，52ポイント）、4種類の文字間隔（-2，-1，0，1 mm）の12種類、合計36個の文字列を準備した。なお、図6.2(c)の文字列は、仮名を漢字よりも小さく書くことが多い<sup>133)</sup>ことを考慮して、「晴」のみは大きさ60ポイントの外接正方形のままとし、他の仮名の外接正方形の大きさを変更した。図6.3(c)は、図6.2(c)の文字列を仮名の大きさを52ポイントにして文字間隔を変更した例である。

以上のように作成した文字列を、自発光するCRTと反射光を用いる紙の代表的な2種類の方法で、視力が健常な20歳代の男子大学院生19名の被験者に呈示した<sup>注1</sup>。CRTはApple社のMacintosh用12インチカラーモニタ、紙は文字列を印刷したボードを使用した。CRT、紙共に白背景に黒文字である。各方法共に文字列を1文字（60～52ポイント）の視角が2°程度になる視距離1.0mで呈示した。呈示方法による視距離や被験者の姿勢を変化させないため、CRTと紙の文字列の位置が一致するよう、ボードはCRTの直前に置き、被験者の顔を顎台に固定した。実験中、暗幕で外光を遮断して照明条件を一定に保った。そして、文字種、文字の大きさ、文字間隔の異なる36個の文字列をランダムに呈示し、時間制限をせずに各文字列の読み易さを、非常に読み易い5、読み易い4、普通3、読みにくい2、極めて読みにくい1の5段階の絶対評価を口頭で報告してもらった。これを、CRTおよび紙による呈示方法別に行った。評価は、文字列の1文字1文字が何であるか認識するのがどの程度大変に感じるか評価してもらう。読むのが大変に感じるほど数字が低く（2，1）、容易なほど数字は高い（4，5）。例えば、文字間隔が充分広い場合、各文字の認識が容易で読み易い評価になると考えられる。ただし、文字間隔が充分あって文字自体は読み易いが、文字間隔が余り広いのは嫌いだから読みにくいなどといった個人的嗜好は、読み易さの評価に含めないよ

注1 実験は豊橋技術科学大学 情報工学系 白井研究室にて行われた。

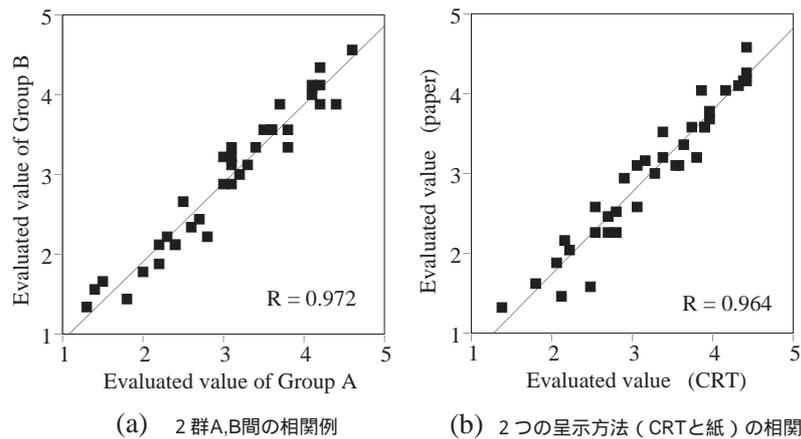


図 6.4 検証結果 (被験者群, 呈示方法の違い)

う指示した。事前に図 6.2 とは文字種が異なる文字列を呈示して評価基準を説明し、それを理解させた後、本実験を行った。

次に、心理実験に使用する 36 個の文字列を文字列の長さに応じて、 $256 \times 256$  dot または  $320 \times 320$  dot の文字画像に変換、式 (3.19) を用いて誘導場を計算した。そして、図 6.1 と同様の要領で、ポテンシャル値を 0.001 ずつ徐々に大きくした誘導場の分布を観察、 $i$  番目の文字を囲む誘導場の閉曲面が初めて現れるポテンシャル値を調べ、式 (6.1) による文字列全体の誘導場の強さの平均を求めた。

## 6.4 誘導場による文字列の読み易さに関する検証結果

最初に、心理実験結果の信頼性を検証するため、19 名の被験者をランダムに 10 名と 9 名の 2 群 (A, B) に分け、評価した 36 個の各文字列について、A, B 群ごとに心理実験で得られた主観評価結果の平均を求め、A, B 群間の相関を調べた。図 6.4(a) は紙に呈示した場合について、横軸に A 群、縦軸に B 群の主観評価の平均をとった散布図の例である。図中の直線は回帰直線で、相関係数は 0.972 と高い。図 6.4(a) の例を含め他の紙に呈示した場合、および CRT に呈示して被験者をいろいろな 2 群に分けた場合を含め、相関係数は 0.925 から 0.972 (平均 0.948) の高い相関が得られた。このことから、任意の文字列に対する読み易さの主観評価は、個人ごとにほぼ一定であると考えられる。

次に、文字列の呈示方法の違いによって文字列に対する主観評価の傾向が異なるかを検証した。図 6.4(b) は 36 個の各文字列について、19 名の被験者全員の主観評価の平均を、横軸に文字列を CRT で呈示した結果、縦軸に紙で呈示した結果をとった散布図である。相関係数は 0.964 と高い。このことから、文字列に対する主観評価は、自発光する CRT、または反射光を用いる紙という光源の違いによって大きく異ならないと考えられる。

続いて、読み易さの主観評価結果を示す。表 6.1 は、文字の大きさ、文字間隔の異なる図 6.2 の 3 つの各基本文字列について、CRT に呈示した場合の被験者全員の主観評価の平均を示した表である (括弧内は紙に呈示した場合)。図 6.5(a) は文字間隔が 0 mm の場合、図 6.2 の 3 つの各基本文字列について、横軸に文

表 6.1 文字列の読み易さについての心理学実験結果

ねすれにそ				
distance	-2 mm	-1 mm	0 mm	1 mm
60 point	2.47 (1.58)	2.53 (2.26)	3.05 (3.11)	3.89 (3.58)
56 point	2.68 (2.26)	3.16 (3.16)	3.89 (3.58)	4.32 (4.11)
52 point	3.63 (3.37)	4.16 (4.05)	4.42 (4.26)	4.42 (4.58)
塩歩晴具堅				
60 point	1.37 (1.32)	2.21 (2.05)	2.68 (2.47)	3.26 (3.00)
56 point	2.11 (1.47)	2.79 (2.53)	3.37 (3.21)	3.74 (3.58)
52 point	3.05 (2.58)	3.58 (3.11)	3.84 (4.05)	4.42 (4.16)
ねす晴にそ				
60 point	1.79 (1.63)	2.16 (2.16)	2.89 (2.95)	3.79 (3.21)
56 point	2.05 (1.89)	2.53 (2.58)	3.37 (3.53)	3.95 (3.79)
52 point	2.79 (2.26)	3.53 (3.11)	3.95 (3.68)	4.37 (4.16)

文字列を CRT で呈示した場合の主観評価の平均。( )内は紙で呈示した場合。

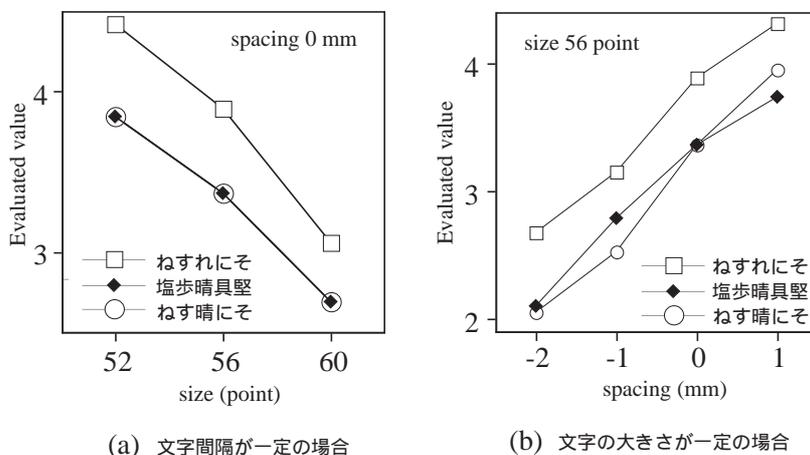


図 6.5 文字間隔，文字の大きさを一定にした場合の文字列の読み易さの違い

字の大きさ（ポイント数），縦軸に文字列を CRT で呈示した場合の被験者全員の主観評価の平均をとって比較したグラフである．表 6.1 および図 6.5(a) を含めた他の例から，文字間隔が同じ場合，文字が小さくなると文字列が読み易くなることがわかる．図 6.5(b) は文字の大きさが 56 ポイントの場合，横軸に文字間隔（mm），縦軸に文字列を CRT で呈示した場合の被験者全員の主観評価の平均をとって比較したグラフである．表 6.1 および図 6.5(b) を含めた他の例から，文字の大きさが同じ場合，文字間隔が広がると読み易くなることがわかる．また，表 6.1 および図 6.5 を含めた他の例から，図 6.2(a) のすべて仮名の文字列は，同じ文字の大きさ，文字間隔の他の文字列より読み易いことがわかる．このように，従来，文字列の状態の表現に使われてきた文字種，文字の大きさ，文字間隔と文字列の読み易さには関連性が認められるので，これら 3 要素の組合せにより文字列のおおよその読み易さを示すことができると考えられる．

図 6.6 は，横軸に被験者全員の主観評価の平均，縦軸に式 (6.1) による文字列全体の誘導場の平均を (a) は紙，(b) は CRT で文字列を呈示した場合をとった散布図である．なお，計算した誘導場の強さと主観評価結果はそのままでは相関関係が直線的にならない．そこで，誘導場の強さの常用対数を取り，両者の相関関係を直線的にした上で相関係数を求めた．図 6.6 より，相関係数は 0.907（紙），0.917（CRT）と高く，

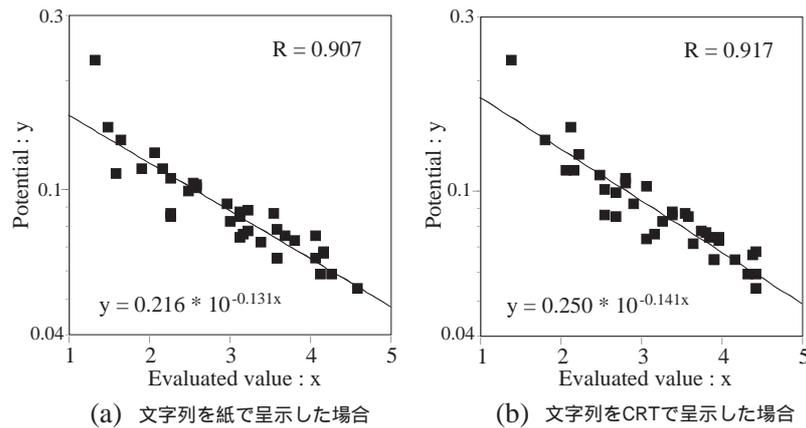


図 6.6 文字列の読み易さと誘導場の強さの相関

式 (6.1) による誘導場の強さの平均は、文字の読み易さの心理実験結果とよく合致している。そして、誘導場の強さの平均は、文字種、文字の大きさ、文字間隔が異なる任意の文字列の状態を 1 つの値で示すことができるので、文字列の読み易さを一意に表現する指標に使えると考えられる。

## 6.5 文字列を構成する文字による文字列の誘導場の予測

### 6.5.1 2 文字間の誘導場による予測

以上のように、いろいろな大きさや間隔の文字列の誘導場を計算し、式 (6.1) による誘導場の強さの平均を求めて文字列の読み易さの心理実験結果と照合すれば、読み易い状態を決定できると考えられる。しかし、式 (3.19) による文字列全体の誘導場の計算は、画素を検出するための走査を広範囲な文字列画像について行う必要があること、また、個々の文字を囲む誘導場の閉曲線の検出において、ポテンシャル値を変化させながら繰り返し行う必要があるため、時間がかかるという問題がある。

複数のパターン全体が形成する誘導場は、基本的に個々のパターンが形成する誘導場の和である。そして、誘導場は遮蔽された部分の影響を受けないので、 $n$  個の文字から構成される文字列における、隣接する  $i, i+1$  番目の文字間につくられる誘導場の分布は、元の文字列から  $i, i+1$  番目の文字 2 つのみを抜き出した文字列の 2 文字間の誘導場の分布で近似できると考えられる。すなわち、個々の文字に分離する誘導場の強さは、隣接する 2 文字の組合せごとに考えればよい。そこで、最初に文字パターンとして非常に単純化した円について検討する。

まず、単独で存在する円の誘導場は、式 (3.19) より円の周辺に同心円状に分布するので、円の誘導場の強さ  $m$  は、円の直径  $d$  と円の中心からの距離  $r$  の関数で表される<sup>注2</sup>。

$$m = f(d, r) \quad (6.2)$$

図 6.7(a) は、式 (3.19) より計算、計測した直径 64 dot の円の誘導場の分布を、縦軸にポテンシャル値、横

<sup>注2</sup>付録 D を参照。

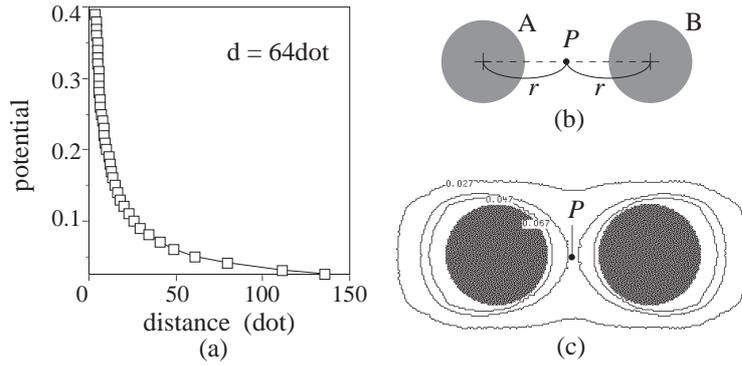


図 6.7 円における誘導場の平衡

軸に円弧からの距離（円の中心からの距離  $r - 32 \text{ dot}$ ）をとって示した例である．図 6.7(a) の例のように，実際にいろいろな直径の円の誘導場を計測して式 (6.2) の関数を決定することができる．

次に，図 6.7(b) のように，等しい直径  $d$  の円 A, B を中心間距離  $2r$  で配置した場合を考える．円 A, B は大きさが同じなので，式 (6.2) は円 A, B 共に等しい．よって，図 6.7(c) の誘導場の分布例のように，中点 P（円 A および B の中心から距離  $r$ ）を境に個々の文字を分離する誘導場の閉曲線が存在する．したがって，円 A, B を個々に囲む誘導場の閉曲線の強さは，中点 P の誘導場の強さ  $M_p$  に等しい．円 A, B が単独で存在する場合の誘導場の強さを  $m$  とすると，誘導場は重ね合わせだから，円 A, B を図 6.7(b) のように並べた場合の中点 P における誘導場の強さは  $M_p = 2m$  である．誘導場の強さは距離の 2 乗に反比例するから，円 A, B を中心間距離  $2r$  で並べた場合， $M_p$  は円 A（または B）が単独で存在する場合の円 A（または B）の中心から距離  $2^2 r$  の誘導場の強さに等しい．

$$M_p = f_A(d, 2^2 r) \quad (6.3)$$

円の大きさが異なる場合（ $d_A > d_B$ ），円 A, B の誘導場の分布は異なる．円 A, B が単独で存在する場合の円の中心から距離  $r$  の誘導場の強さを  $m_A, m_B$  とすれば，誘導場の強さは距離の 2 乗に反比例するから， $m_A, m_B$  の比の平方根だけ中点 P が円 A の中心から距離  $r'$  に移動する．

$$r' = r \sqrt{m_A / m_B} \quad (6.4)$$

したがって，大きさの異なる円 A, B を分離する誘導場の強さ  $M_p$  は式 (6.3) から

$$M_p = f_A(d_A, 2^2 r') \quad (6.5)$$

と表される．式 (6.5) に，式 (6.4), (6.2) を代入して  $r'$  を消去すると， $M_p$  は次式で与えられる．

$$M_p = f_A(d_A, 2^2 r \sqrt{\frac{f_A(d_A, r)}{f_B(d_B, r)}}) \quad (6.6)$$

図 6.8(a) は，円 A の直径が 64 dot，円 B が 48 dot の場合，図 6.8(b) は円 A が 64 dot，円 B が 24 dot の場合について，円 A, B の中心間距離（mm）を横軸に，ポテンシャル値を縦軸にとって，個々の文字を

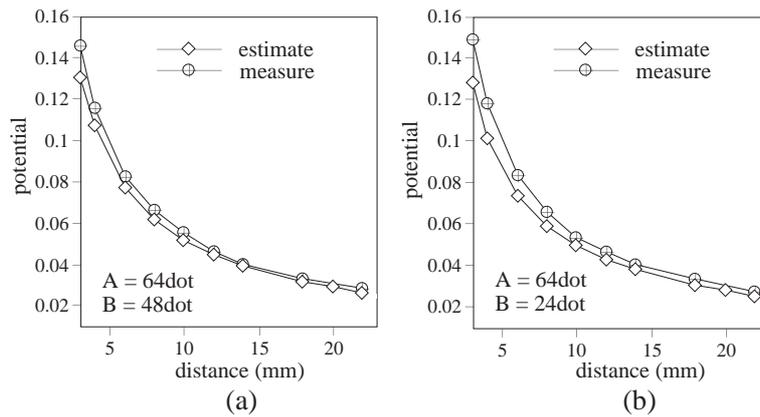


図 6.8 円の誘導場の予測

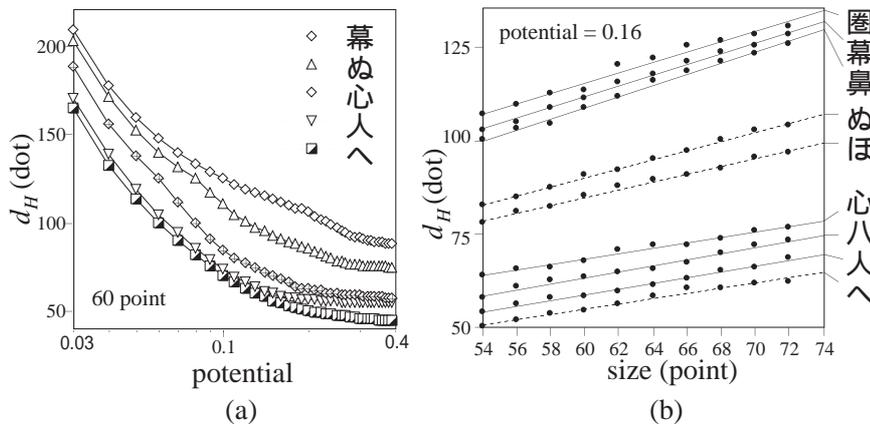


図 6.9 円相当径  $d_H$  の測定例

囲む誘導場の閉曲線が最初に現れるポテンシャル値の式 (6.6) による予測と、円 A, B を配置して式 (3.19) で計算した誘導場の実測を比較したグラフの例である<sup>注3</sup>。図 6.8 の例を含め、他の大きさの円の組合せも予測と実測はよく一致している。このことから、2 つの文字の列が個々の文字に分離する誘導場の強さは、時間のかかる式 (3.19) によって文字列全体の誘導場を計算せずに、あらかじめ求めてある文字が単独で存在する場合の誘導場の分布から、簡単に求めることができると考えられる。

### 6.5.2 文字の誘導場の予測

円の誘導場の分布は円である。そして、文字の誘導場の分布も文字線分に近い部分を除いて、誘導場の分布の大部分は円で近似できると考えられるので、誘導場の分布をあるポテンシャル値における等ポテンシャル面の面積  $A$  と等しい面積をもつ円の直径 (円相当径)  $d_H$ <sup>127)</sup> で表現する。

$$d_H = \sqrt{4A/\pi} \tag{6.7}$$

図 6.9(a) は、文字種別として文字の大きさが 60 ポイントの場合の円相当径 (dot) を縦軸に、ポテンシャル値を横軸に示した例である。図 6.9(a) の例を含め、他の大きさの場合も、ポテンシャル値が大きくなる

<sup>注3</sup>円は Macintosh 上で 72 DPI の解像度で作成した。以下の実験でも同じ条件でパターンを作成した。

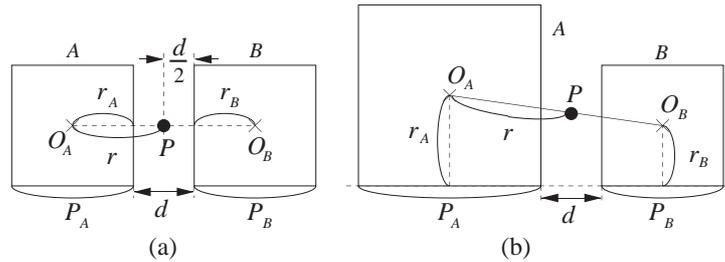


図 6.10 文字の位置関係

と円相当径が急速に小さくなる．図 6.9(b) は文字種別としてポテンシャル値 0.16 の場合の円相当径を縦軸に，文字の大きさを横軸に示した例である．図 6.9(b) の例を含め，他のポテンシャル値の場合も，円相当径は文字の大きさに比例する．このように，ある文字種の誘導場の強さ  $m$  は，文字の大きさ  $P$  (ポイント数) と円相当径  $d_H$  の関数で表現できると考えられる．

$$m = g(P, d_H) \quad (6.8)$$

式 (6.8) は式 (6.2) と同様に，図 6.9 の例のようにいろいろな文字種の誘導場を計算して決定することができる．

次に，隣接する 2 文字の位置関係を外接正方形を基準に定義する．図 6.10(a) のように，2 つの文字 A, B の外接正方形が間隔  $d$  で平行に配列しているとする．文字 A, B の大きさ  $P_A, P_B$  (ポイント数) は，各文字の外接正方形の一辺の長さと同じ．文字 A, B の誘導場の分布が等しいと仮定すると，文字 A, B の外接正方形の中心を  $O_A, O_B$  とすれば，円と同様に  $\overline{O_A O_B}$  の中点 P では文字 A, B の誘導場の強さが等しくなる． $O_A, O_B$  から外接正方形におろした垂線の足までの長さを  $r_A = P_A/2, r_B = P_B/2$  とすると， $O_A$  から中点 P までの距離  $r$  は，

$$r = \frac{1}{2} (r_A + r_B + d) \quad (6.9)$$

と表される．実際には図 6.9(b) の例のように，文字 A, B が同じ大きさでも誘導場の分布は異なるので，円と同様に文字 A, B が単独で存在する場合の，各文字の誘導場の強さに比例して中点 P は移動する．式 (6.6), (6.8) より，移動した中点 P における誘導場の強さ  $M_p$  は，

$$M_p = g_A(P_A, 2r \sqrt{\frac{g_A(P_A, r/2)}{g_B(P_B, r/2)}}) \quad (6.10)$$

として計算できる．

図 6.10(b) のように文字 A, B の大きさが異なる場合 ( $P_A > P_B$ )，一般に外接正方形の底辺はベースラインに揃えることが多いので， $\overline{O_A O_B}$  は外接正方形と平行ではない．文字 A, B の誘導場の分布が等しいと仮定すると， $O_A$  から中点 P までの距離  $r$  は，

$$r = \frac{1}{2} \sqrt{(r_A - r_B)^2 + (r_A + r_B + d)^2} \quad (6.11)$$

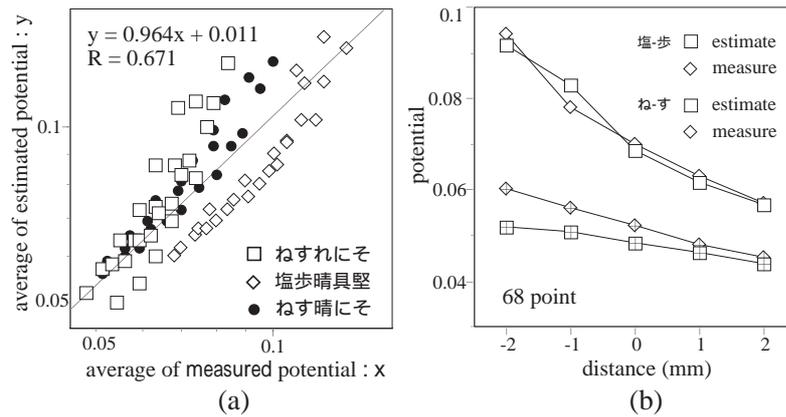


図 6.11 実験結果（読み易い文字列の予測）

と表される．実際には文字 A, B の誘導場の分布は異なるので，図 6.10(a) と同様に，式 (6.10), (6.11) より誘導場の強さを計算する．

### 6.5.3 誘導場による読み易い文字列予測方法の検証

図 6.2 の 1 つの基本文字列について，図 6.3 と同じ要領で，文字間隔が  $-2, -1, 0, 1, 2$  mm の 5 種類，大きさが 72, 68, 64, 60, 56 ポイントの 5 種類の 25 種類，合計 75 個の文字列を用意した．ただし，仮名を漢字よりも小さく書くことが多いことを考慮して，図 6.2(c) の「晴」のみ文字の大きさは 72 ポイントに固定した．文字の読み易さの主観評価実験と同様に，まず，用意した文字列の誘導場を式 (3.19) より計算，個々の文字を囲む誘導場の閉曲面が初めて現れるポテンシャル値を調べ，文字列全体の誘導場の強さの平均を式 (6.1) より求めた．次に，図 6.2 の全文字種について，大きさが 72, 68, 64, 60, 56 ポイントの文字が単独で存在する場合の誘導場を計算し，ポテンシャル値が 0.01 から 0.001 ずつ 0.399 までの円相当径を式 (6.7) より求めて式 (6.8) を決定した．これらのデータを基に式 (6.9), (6.10), (6.11) を用いて，文字 A を  $i$  番目，文字 B を  $i + 1$  番目の文字として個々の文字を囲む誘導場の閉曲線のポテンシャル値を計算し，文字列全体の誘導場の強さの平均を式 (6.1) より求めた．

図 6.11(a) は，75 個の各文字列の式 (6.1) による文字列全体の誘導場の強さの平均について，文字列の誘導場を実測した結果を横軸に，個々の文字の誘導場の分布による予測を縦軸にとった散布図である．図中の直線は回帰直線で，予測と実測が一致する理想の傾き 1.0 に近い 0.964 である．図 6.11(b) は，文字の大きさが 68 ポイントの場合，文字間隔 (mm) を横軸，ポテンシャル値を縦軸にとって，図 6.2(a) の文字列「ねすにれそ」の「ね」と「す」，および図 6.2(b) の文字列「塩歩晴具堅」の「塩」と「歩」を個別に囲む誘導場の閉曲線のポテンシャル値について，個々の文字の誘導場の分布からの予測と実測を比較した例である．図 6.11 の例を含めた，他の条件での予測は実測に近いことから，文字の誘導場の分布を円相当径で表現することによって，個々の文字を個別に囲む誘導場の強さを計算により予測できると考えられる．

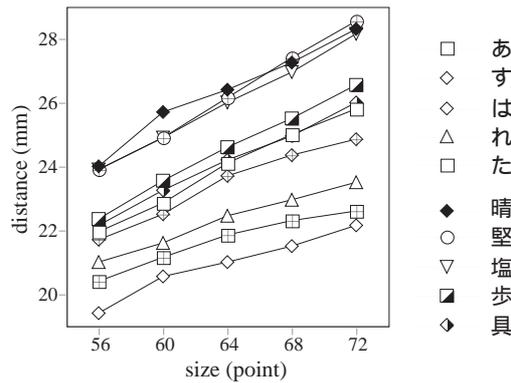


図 6.12 最適な文字間隔の予測

### 6.5.4 読み易い文字列の状態の決定

図 6.12 は、大きさの等しい同じ文字種を 2 個以上、等間隔で並べた文字列を紙に呈示した場合、視距離 1 m で 5 段階評価の読み易い状態 4 にするための文字の大きさ（ポイント数）を横軸に、隣接する 2 文字の中心間距離（mm）を縦軸に示した例である。読み易い状態 4 のポテンシャル値 0.065（図 6.6(a) から計算）を実現する文字間隔を、各文字種、大きさの文字が単独で存在する場合の円相当径の測定結果および式 (6.9)、(6.10) から計算した。図 6.12 から、大きい文字は小さい文字より、漢字の文字列は仮名の文字列よりも広い文字間隔が必要であることがわかる。これは、表 6.1 および図 6.5 の例で示された文字列の読み易さの傾向が、誘導場の分布で説明できることを示している。以上のように、個々の文字の誘導場の分布から、時間がかかる文字列全体の誘導場の計算を行わずに、読み易くするための設定すべき具体的な文字種、文字の大きさ、文字間隔を簡単に計算できると考えられる。

## 6.6 考察

### 6.6.1 誘導場による文字列の読み易さの説明

図 6.6 の結果から、文字列の読み易さの指標として着目した、文字列の誘導場における個々の文字を囲む誘導場の閉曲線の強さが、人間による読み易さとよく一致することが示された。すなわち、個々の文字の誘導場の干渉の強さが文字列の読み易さと関連することが示唆されていたが<sup>44)</sup>、干渉の程度は式 (6.1) による文字列全体の誘導場の強さの平均で表現でき、実際に人間の文字列の読み易さと対応していることが明らかになった。このことは、視覚の誘導場は文字の読み易さを説明できる有効な概念であることが示された。

### 6.6.2 読み易さの客観的指標

表 6.1、図 6.5 から、文字列の状態を決める文字種、文字の大きさ、文字間隔によって、文字列の読み易さをおおよそ推定できると考えられる。しかし、文字列の読み易さを表すには、図 6.5 の例のように、これら文字列の状態を示す 3 要素の組合せが必要であり、一意に読み易さを決めることは容易ではない。特に

フォント、文字飾りなどの要素が増えた場合、文字列の状態の組合せが増大して、読み易さの決定は難しくなると考えられる。

一方、図 6.6 から、誘導場の強さを用いれば、読み易い文字列の状態を一意に表すことが可能であり、誘導場の強さは人間の読み易さとよく一致している。すなわち、文字種、文字の大きさ、文字間隔が異なっても、表示したい文字列の誘導場の強さを計算すれば、文字列の読み易さを誘導場の強さのみで決定できると考えられる。このことから、誘導場は読み易い文字列の表示を行うために有効な客観的指標を提供できると考えられる。

### 6.6.3 読み易い文字列の状態の決定

図 6.11 から、文字列の読み易さの指標となる誘導場の強さは、時間がかかる文字列全体の誘導場の計算をせずに、文字列を構成する 2 文字ずつの組について、各組の文字種が独立に存在する場合の誘導場の分布から、計算によって簡単に予測できることが示された。そして、図 6.12 から、表示したい文字列の文字種の組合せが決まれば、読み易い文字列にするための文字の大きさ、文字間隔が、経験によらず具体的に決定できる。この方法を用いれば、例えば、プリンタなどの印刷をはじめ、電子ブックなどの携帯端末では限られた表示面積で見易さをできるだけ落とさず多くの文字を効率良く、また高精細なハイビジョンでは、より多くの文字をより読み易くするプロポーショナル表示などが可能になると考えられる。

### 6.6.4 読み易い文字列状態の予測性能向上のための課題

図 6.11 のポテンシャル値が大きい場合の仮名などのように、文字の誘導場の分布から予測した個々の文字を囲む誘導場の強さと実測の違いが、漢字の文字列などと比べ大きい場合がある。これは、文字、特に仮名は漢字に比べ長いストロークなどの影響により誘導場の分布が複雑になり、円では誘導場の分布が充分記述できず、誘導場の強さの予測に誤差が生じたものと考えられる。そこで、例えば、等ポテンシャル面の形状が円で近似できるかを円形度や複雑度<sup>127)</sup>で判断し、複雑な場合、楕円で誘導場の分布を近似して、個々の文字を囲む誘導場の強さを予測するなどの方法が考えられる。



## 第7章 今後の展望

### 7.1 視覚の誘導場研究の発展

本論文は、主に図形周辺の分布が議論されていたこれまでの視覚の誘導場研究から一步踏み出し、誘導場がパターン認知などの視覚にどのようにかかわっているかをモデル化することが目的であった。その結果、文字パターン認識や文字の切り出し、文字列の読み易さについて誘導場によるモデルが可能なことが示された。

今後は、まず、これまで提案した誘導場によるモデルの工学的応用を更に進めると共に、誘導場が視覚の機軸であるという仮定のもとに、他の視覚現象に関するモデルも検討して行きたい。そして、モデル化と共に心理学や生理学知見との関連性を検討し、誘導場の科学的な根拠、機構の解明も進めて行きたい。このように、誘導場理論の適応範囲を広げると共に、視覚モデルや機構、特に心理学や生理学との関連性などの解明を進めたい。以下、誘導場の工学的応用が期待できそうな分野と誘導場理論の今後の展開について述べる。

### 7.2 工学的応用

#### 7.2.1 表情の定量評価

顔は人間の複雑で高度なコミュニケーションを媒介する高度な手段であり、それを理解し、工学的に扱うことは非常に意義がある。近年「顔」の研究が非常に盛んになっており、1995年には日本顔学会<sup>注1</sup>が発足した。顔の研究は工学的応用（インタ - フェイス、コンピュータ処理、認識・照合・認証・検索など）<sup>141)</sup>と、心理学的な研究（表情の意義、顔の果たすコミュニケーションなど）<sup>142)</sup>に大別される。

しかしながら、顔は非常に複雑で多義性のあるパターンである上、感情や社会、文化的背景など複雑な条件も付帯する。これまで検討されている顔の工学的応用では、顔から顔の部品を取り出し、それを個々に評価する手法が主流である。部品の抽出方法、扱い方など多くの方法が提案されているが、多様で複雑な顔を理解、表現することは難しい。これは、従来の特徴量を用いる文字認識では、多様な文字パターンの特徴を充分表現できないため、認識が困難であったことに相当すると考えられる。

さて、横瀬は図 7.1 上段の線画の顔について誘導場を計算したところ、誘導場の分布が歌舞伎の「くま取り」の顔に似ていることを報告している<sup>44)</sup>。図 7.1 下段は、式 (3.19) で計算した誘導場の例である。(a) は怒り、(b) は普通、(c) は笑いである。つまり、歌舞伎役者の感情や性格を強調させるための「くま取り」は、

<sup>注1</sup><http://www.hc.t.u-tokyo.ac.jp/jface/>

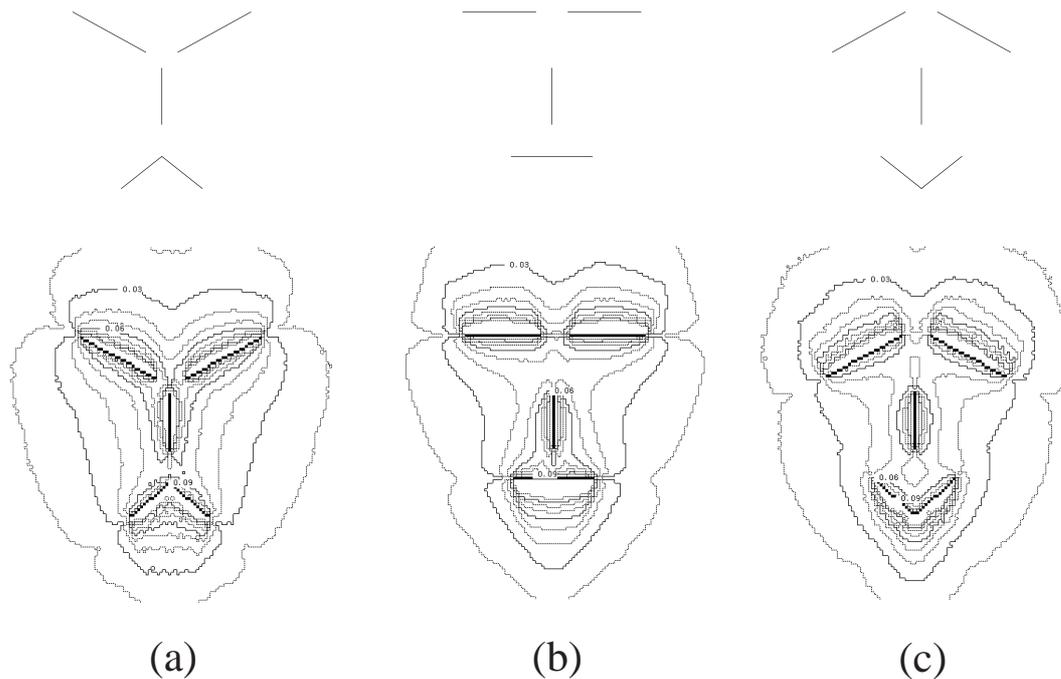


図 7.1 異なる表情（線画）における視覚の誘導場の例

顔の基本要素である，眼，鼻，口から醸し出される誘導場の分布になるように描かれていることを示している．これは，誘導場の分布が顔の表情などの主観評価と結びついていることを示唆するものと考えられる．

本論文で示したように，誘導場は2次元上に分布するパターンの配置状態の定量評価に適している．また，従来の特徴量を用いる文字認識では難しかった，パターン類似性の主観評価が可能である．したがって，顔についても誘導場の分布を解析することで，表情の定量評価などへの応用が期待できる．

## 7.2.2 文字パターン認識

誘導場理論による認識モデルは文字の類似性の評価が優れているという特性がある．また，認識率に関しても従来の手法と同等程度の性能を有している．最近，コンピュータビジョン分野において，Snakesなどの最適化手法の精密化，高精度化を目指した数多くの手法が開発されてきており，一部実用化もなされている．今回提案した誘導場の認識モデルは，4.6.3の誘導場間の変位計算で述べたように，変位計算などの手法が未熟なため，認識性能が下がる問題が残っている．今後，このようなコンピュータビジョン分野の進んだ手法を取り入れることによって，誘導場の認識モデルの性能を向上させることが期待できる．そして，優れた類似性の評価の性質を生かした知識処理を行うことで，総合的に優れた認識システムの構築が期待できると考える．

### 7.2.3 文字列の読み易さ

本論文では、誘導場の分布を用いて、単独の文字列の読み易さを検討した。同じ要領で、複数の文字列の読み易さ（適切な行間）を検討することも可能である。文字列内および文字列間の読み易さを定量評価できれば、これまで印刷物やコンピュータの文字表示で常に悩まされていた、適切で読み易い表示方法の定量的尺度を与えることができると考えられる。また、同様な方法により日本語だけではなく、他言語の文字列の読み易さも評定できると考えられる。

### 7.2.4 重力レンズ錯視

内藤は平行四辺形の頂点に配置した小点の側にやや大きめの円をおくと、小点が平行四辺形の配列から大きくずれて見える錯視を発見した<sup>143)</sup>。この錯視は、天体物理学の重力レンズ理論で説明できることから、重力レンズ錯視と呼ばれる<sup>41, 143)</sup>。重力レンズ錯視は、従来の等方一様を仮定している古典的な場では説明できないため、内藤は一般相対性理論を採用して重力レンズ錯視の錯視量を定量的に説明している。

一方、視覚の誘導場は基本的に静電場だが、式(3.6)が示すような遮蔽があるので小点の側の円の配置方向により、小点付近の場の分布が等方一様ではなくなると考えられる。つまり、重力レンズ錯視を引き起こす円の配置による場の歪みが、誘導場の遮蔽性から説明できる可能性がある。

## 7.3 視覚の誘導場理論の展開

### 7.3.1 誘導場計算の精密化

式(3.19)の計算モデルは、いろいろなパターンのデジタル画像の誘導場について計算できるが、深水、吉田らが報告しているパターン近傍の場の尻切れ状態や左右非対称などの分布は再現できない。これら誘導場の分布状況はまだ充分解明されていない点が多く、今後更なる心理実験や計算モデルの検討が必要である。もし、その結果が計算モデルに正確に反映されれば、文字パターン認識の認識率向上<sup>注2</sup>、更により接近したパターンの切り出しなどの工学的応用に大きく貢献すると考えられる。

### 7.3.2 階調，色のある場合の視覚の誘導場

これまで議論してきた誘導場は背景が白，図形が黒の2値のみである。しかし，現実の世界は色や階調がある。深水，吉田らの側抑制による誘導場のモデルは，階調，色の場の計算可能性を示している<sup>64, 65)</sup>が，実験的に色や階調のある対象について誘導場を求めたり，計算モデルの結果と比較した例はない。

人間が色の三原色をある比率で合成して知覚するモデルが幾つか提案されている<sup>144)</sup>。そこで，デジタル画像における誘導場の計算式(3.19)を，三原色の1つの色の場合と見なして，各色の度合いを色知覚のモ

<sup>注2</sup>例えば，多くの書体，活字が完全な左右対象にデザインされていないのは，誘導場の左右非対称のためデザイナーが本能的に非対称にしているという指摘がある<sup>24)</sup>。

デルの合成比率にしたがって、合成する色の誘導場のモデルが考えられる。

また、階調については、式 (3.19) で仮定している電荷を画像の濃度に対応させる計算式の拡張が考えられる。もし、色、階調のある画像の誘導場が計算できれば、例えば、どのような色の背景で印刷されたどのような輝度と色の文字が読み易いかなどを、誘導場で定量的に決定出来ると考えられる。

### 7.3.3 2次元パターン配置の評価の数学的な取り扱い

手書き文字の切り出し、文字列の読み易さは、共に誘導場の強さを指標にしている。しかし、切り出し、読み易さの判断は人間の主観評価を基にした閾値を設定している。したがって、誘導場を2次元パターン評価で汎用的に利用する上で、このような判断が必要なことは問題がある。

誘導場は物理数学的にほとんどクーロンポテンシャルとして扱うことができる。そして、このようなポテンシャルモデルにおいて、漠然としている2次元平面上の位置関係を数学的に記述する可能性が示唆されている<sup>145)</sup>。具体的には、パターンの布置状態を誘導場の分布に置き換え、ポテンシャル論<sup>146)</sup>の容量(Capacity)<sup>147)</sup>などで純粋に数学的に取り扱い、これまで主観的に行っていた2次元パターン配置の評価を数学的に行うことが考えられる。

## 7.4 生理学からのアプローチ

視覚の誘導場は心理学的な手法で観測される事象を説明する心理学的な概念である。一方、これまで誘導場は、生理学的には網膜周辺で生じている現象との関連が示唆されているに過ぎない。深水、吉田らの側抑制を基にした誘導場のモデルのような、かなり有効な計算モデルも提案されているが、生理学的な誘導場の実体はわかっていない。しかし、誘導場という現象は、人間の脳活動の結果生ずることは確かである。まだ、従来の生理学的観察では発見できなかった、これまで知られていない、何らかの誘導場に関する現象があると考えられる。最近ではfMRIなど人間の脳活動を直接観察できる手法が開発、発展している<sup>148)</sup>。例えば、文字を見せた時の脳活動状況や部位をこれらの装置でより詳しく解析すれば、誘導場を引き起こす機構がわかるかも知れない。これまでは、網膜を中心とした主に誘導場の生成に関する生理学的研究がほとんどであった。今後は、誘導場の生成を含め、パターン認知などより高次の機構についての生理学的探求、解明が重要になると考えられる。そして、誘導場の生理学的な機構の一端でも明らかになれば、これまで提案してきた誘導場モデルの検証やモデルの更なる発展が期待できるだけでなく、誘導場理論そのものの理解が深まるものとする。

## 第8章 結 論

コンピュータは我々にとって非常に身近な存在になっている。一昔前コンピュータは主に文書作成，表計算などに使われてきたが，ここ数年急速に普及しているインターネットから，いろいろな情報を検索して利用するといった知的活動の支援に使われるようになってきた。しかし，これだけコンピュータのハードウェアが普及したにもかかわらず，一般に「コンピュータ」は使うのが難しいとされている。最大の原因は，コンピュータはまだ我々に匹敵するような視覚や聴覚などの五感をもっていないことが考えられる。コンピュータの入力方式は1960年代に発明されたグラフィカル・ユーザ・インターフェースが主体である。コンピュータに知的な活動をさせるには，我々が視覚や聴覚などから得る情報をコンピュータに正しく理解させなければならない。

特に視覚は重要であり，中でも文字は太古から使われている情報伝達手段であった。そのため，コンピュータに文字を認識させる研究は古くから行われてきた。しかし，従来の認識方法は極めて経験的であったため，限定された条件では実用化されているが，人間が自由に書いた文字を認識することは未だ困難である。このような状況において，従来の方法よりも科学的なモデルや方法論を与えてくれる認知心理学，生理学的なアプローチが注目されている。本論文は，認知心理学的なアプローチで文字パターンの認知機構を明らかにして，認知モデルを構築することを目的に，横瀬の視覚の誘導場による文字パターン認識に関する研究を進めた。

第2章では，誘導場はいろいろな視知覚現象を説明する心理学的な概念であり，視知覚を説明する重要な枠組みとしても着目されていることから，視覚の誘導場は文字パターン認知などを説明できることが期待されることを述べた。しかし，これまで視覚の誘導場の研究は，主にその存在や分布について議論されていた。また，誘導場が文字パターンの認知など，いろいろな視知覚現象をどのように説明できるかといった具体的なモデルはほとんど検討されておらず，パターン認識のモデルを構築することは意義深いことを述べた。

第3章では，これまでの視覚の誘導場研究についてまとめた。まず，視覚の誘導場の存在は数多くの心理実験結果から確かめられていることを述べた。また，生理学的にもその存在可能性が示唆されていることを述べた。そして，デジタル画像における視覚の誘導場の計算モデルを提案し，いろいろなパターンの誘導場が計算できることを示した。

第4章では，誘導場による文字パターン認識モデルを提案した。この認識モデルは，一方の誘導場が他方の誘導場に一致するように，仮想的に変形する際に生ずる弾性エネルギーを，パターンの類似性として評価する。そして，この認識モデルにより，いろいろな文字パターン（人工的な線図形，活字，変形パターン，

手書き文字)の類似性が評価できることを示した。

第5章では、誘導場による手書き文字の切り出し方法を提案した。この方法は個々の文字を囲む等ポテンシャル線の分布は、丸く他から孤立する傾向を利用している。そして、この方法が自由に筆記された試験答案のように文字が非常に接近している手書き文字を切り出せることを示した。

第6章では、文字列の読み易さを誘導場によって定量化する方法を提案し、検討した。まず、文字列の誘導場の強さが文字列の読み易さの指標になることを心理実験により明らかにした。続いて、文字列を構成している個々の文字の誘導場の分布があらかじめわかっているならば、文字列全体の読み易さが計算により推定できることを示した。

第7章では、視覚の誘導場研究の今後の展開を展望した。まず、顔表情の定量評価や錯視の説明にも誘導場が期待できることを述べた。そして、視覚の誘導場について、今後、計算精度や色、階調のある対象の計算などの課題があり、それらを解決できれば誘導場の応用範囲がさらに拡大することを述べた。

第3章で述べたように、視覚の誘導場は心理学的裏づけもあり、また直接ではないが生理学的にもその存在の可能性が示唆されており、視覚認知機構のモデルとして有効性が高い。そして、第4~6章に示したように、視覚の誘導場はパターン認識と2次元上のパターン配置の評価にも応用することが可能である。具体的には、パターンの類似性評価や近接した文字の切り出し、文字列の読み易さなど、従来の手法では難しかった定量的評価が可能であることを示した。

これまでも、心理・生理学的知見をヒントにしたパターン認知手法などが検討されてきた。しかし、それらの手法は従来のヒューリスティックな手法に心理・生理学的知見のエッセンスを加えたものであったり、従来手法と比較して十分な性能が出ていないなど、心理・生理学的知見を生かすことが難しかった。一方、本論文で提案した誘導場の工学的応用では、横瀬の理論を基本に従来手法とは全く異なる新しい手法を提案し、有効性も高いことが示された。

また、視覚の誘導場という単一の概念から、パターン認識、2次元上のパターン配置の評価という、これまで別の分野として扱われてきた視覚現象を説明するモデルを構築することができた。これは、誘導場という「場」の概念を視覚の枠組みとすることで、多様な視覚現象を説明できたと考えられる。実際の視覚は、例えば、パターン認識、2次元上のパターン配置などに単純に切り分けられるのではなく、いろいろな機能が複雑に作用している。誘導場を視覚の枠組みに据えることで、このような互い作用している視覚現象を統一的に説明できることが期待できる。

更に、「場」は物理的な性質が十分調べられており、視覚の誘導場は電磁場のような「場」の概念であることから、数学的な取り扱いができる。したがって、「場」による視覚のモデルは、物理数学的な理論の展開、応用、例えば、コンピュータビジョンの計算理論などの応用が容易であるという利点がある。

本論文の成果が、視覚の誘導場の解明や工学的応用の発展に寄与し、真にコンピュータを知的活動の支援に利用できるようになることを期待する。

## 謝辞

本論文をまとめるにあたり、豊橋技術科学大学 臼井支朗教授には丁寧なる御指導と御教示ならびに心理実験の御協力を賜りました。ここに、謹んで深謝の意を表します。また、本論文をまとめるに際し、豊橋技術科学大学 金子豊久教授、山本眞司教授、工学院大学 淀川英司教授、豊橋技術科学大学 中内茂樹講師には、有益な御討論、御助言を賜りました。ここに記し、深く感謝致します。また、第6章の文字列の読み易さの心理実験の御指導、御協力を頂いた、豊橋技術科学大学 石樽康雄氏（現NTTサイバースペース研究所）、愛知県立大学 小濱剛氏、および快く被験者としてご協力頂いた、豊橋技術科学大学 臼井研究室の学生諸氏に感謝します。更に、自由手書き文字の文書を提供して頂いた、静岡大学理学部物理学教室の中山潔教授に感謝致します。

本研究の一部は（株）ATR 視聴覚機構研究所 視覚研究室出向中に行ったものである。ATR 在職中、研究の機会を与えて下さった、元 ATR 視聴覚機構研究所 淀川英司社長（現工学院大学教授）、上野圭一室長（現NTTアドバンステクノロジー株式会社）、および山田光穂主任研究員（現NHK放送技術研究所）に感謝致します。特に山田光穂氏には、研究の開始時より終始有益な御討論および適切な御教示を賜りました。ここで、慎んで深謝の意を表します。

本研究の途上において貴重な御討論および御協力を頂いた ATR 視聴覚機構研究所の諸氏に心から感謝致します。また、文字データの整理、作成などを手伝って頂いた、伊藤薫さん、木山直子さん、柿谷広美さん、ソフト作成を手伝って頂いた（株）SCC の向田茂氏、山上雅之氏に感謝致します。

視覚の誘導場について、有益な御討論、御助言を頂いた、聖心女子大学 野澤農名誉教授、拓殖大学 工学部の野田健一名誉教授、吉田登美男教授、深水義之教授、白石照美助手、熊本電波工業高等専門学校 電子工学科の三好正純助教授、古賀広昭教授、広島修道大学 高木敬雄教授に感謝致します。特に、野澤農名誉教授には、視覚の誘導場に関する幅広い知識を賜りました。ここで、慎んで深謝の意を表します。

最後に、本研究を進めるにあたり、御指導と御教示を賜りました、豊橋技術科学大学 故吉田辰夫教授に感謝致します。



## 参考文献

- 1) ルイニジャン・カルヴェ著, 会津洋, 前島和也訳. 文字の世界史. 河出書房新社, 1998.
- 2) 橋本新一郎編. 文字認識概論. 電気通信協会, 1982.
- 3) 山口富士男. コンピューターディスプレイによる図形処理工学. 日刊工業新聞社, 1981.
- 4) Tekla S. Perry and John Voelcker. Of mice and menu :designing the user-friendly interface. *IEEE SPECTRUM*, pp. 46-51, Sep 1989.
- 5) 山田博三. 文字認識における非線形的手法に関する研究. 電子技術総合研究所報告, Vol.853, 1985.
- 6) 山田博三, 森俊二. 手書文字データベースの解析 ( i ). 電総研彙報, Vol. 39, No. 8, pp. 580-599, 1975.
- 7) 斉藤泰一, 山田博三, 森俊二. 手書文字データベ - スの解析 ( iii ). 電総研彙報, Vol. 42, No. 5, pp. 385-434, 1978.
- 8) 斉藤泰一, 山田博三, 山本和彦. 手書文字データベ - スの解析 ( viii ) - 方向パターン・マッチング法による J I S 第 1 水準手書漢字データベース E T L 9 の評価 - . 電総研彙報, Vol. 49, No. 7, pp. 487-525, 1985.
- 9) 中野康明. 文字認識・文書理解における最近の動向 ( 1 ). 人工知能学会誌, Vol. 11, No. 5, pp. 702-709, 1996.
- 10) 萩田紀博. 手書き文字認識の研究動向. *NTT R&D*, Vol. 45, No. 11, pp. 1189-1194, 1996.
- 11) 松井俊弘, 山下郁生, 若原徹, 吉室誠. 文字認識アルゴリズムの複合化手法の検討 - 第 1 回文字認識技術コンテストの結果より - . 電子情報通信学会技報, Vol. PRU92-33, pp. 65-72, 1995.
- 12) 小川英光. パターン認識のための学習 - 基礎と応用 - 論文小特集の発行にあたって. 電子情報通信学会論文誌, Vol. J82-D-II, No. 4, pp. 573-574, 1999.
- 13) 宇都宮敏男編. 特集 認識と理解. 電子情報通信学会誌, Vol. 71, No. 11, pp. 1111-1256, 1988.
- 14) 小川英光編. パターン認識・理解の新たな展開. 電子情報通信学会, 1993.
- 15) 萩田紀博. 21 世紀の文字認識研究に期待すること. 電子情報通信学会技報, Vol. PRMU98-149, pp. 103-110, 1998.
- 16) 中川正樹. 発想支援手書き環境の硬い技術と柔らかい技術. 情処第 34 回プログラミングシンポジウム報告集, pp. 21-32, 1993.
- 17) 竹下鉄矢, 野澤繁之, 掛布英辰. 日常手書き文字の収集. 電子情報通信学会技報, Vol. PRU88-12, pp. 53-60, 1988.
- 18) 木村文隆. 手書き文字認識の高精度化と文書画像認識. 電子情報通信学会技報, Vol. PRMU98-148, pp. 95-102, 1998.

- 19) 横澤一彦, 淀川英司. 文字・図形の認知. テレビ誌, Vol. 40, No. 4, pp. 244-250, 1986.
- 20) 淀川英司. 知覚と認知. 電子情報通信学会誌, Vol. 71, No. 11, pp. 1163-1169, 1988.
- 21) 石井健一郎. 認知科学への期待. 認知科学, Vol. 4, No. 2, pp. 1-2, 1997.
- 22) 竹林洋一. ヒューマンインタフェースの観点から見た気の利いた情報システム. 電子情報通信学会誌, Vol. 82, No. 4, pp. 310-318, 1999.
- 23) フランシス・クリック著, 中原英臣, 佐川峻訳. DNAに魂はあるか - 驚異の仮説 -. 講談社, 1995.
- 24) 野澤農. 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, pp. 329-331. 培風館, 1993.
- 26) Tony Hey and Patrick Walters. 目で楽しむ量子力学 (大場一郎訳), 第4章. 丸善, 1989.
- 27) マイケルI. ポズナー, マーカスE. レイクル著, 養老孟司, 加藤雅子, 笠井清登訳. 脳を観る - 認知心理学が明かす心の謎 -, 第2章. 日経サイエンス社, 1997.
- 28) D.E. ルーメルハート著, 御領謙訳. 人間の情報処理, 第2章. サイエンス社, 1989.
- 29) 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.
- 30) 大山正, 今井省吾, 和気典二編. 新編 感覚・知覚心理学ハンドブック. 誠信書房, 1994.
- 31) 臼井支朗編. 脳・神経システムの数理モデル. 共立出版, 1997.
- 32) 行場次朗編. 認知心理学重要研究集1 視覚認知, pp. 18-21. 誠信書房, 1995.
- 33) 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. 電子情報通信学会論文誌, Vol. J69-A, No. 5, pp. 600-611, 1986.
- 34) 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.
- 35) 柿崎祐一. 心理学的知覚論序説. 培風館, 1993.
- 36) Takahiro Kato and Mitsuho Yamada. Quality factors of hand-written characters based on human visual perception. *SPIE*, Vol. 1453, pp. 43-50, 1991.
- 37) 加藤隆仁, 横澤一彦. 手書き文字品質の定量評価. 電子情報通信学会論文誌, Vol. J75-D-II, No. 9, pp. 1573-1581, 1992.
- 38) 加藤隆仁, 山田光穂. 視線検出による手書き文字品質の決定要因の分析. 電子情報通信学会秋季全国大会, Vol.D-356, 1990.
- 39) 内藤誠一郎. 初期視覚系情報処理と物理数学モデル. 電子情報通信学会技報, Vol. HC94, No. 19, pp. 23-30, 1994.
- 40) 小川英光編. パターン認識・理解の新たな展開, 第2.6.4章. 電子情報通信学会, 1993.
- 41) 内藤誠一郎. 物理数学モデルによる錯視分析の試み. 画像応用技術, Vol. 10, No. 1, pp. 8-19, 1995.
- 42) 山崎敏正. 錯視現象に着目した視覚情報処理モデル. 電子情報通信学会技報, Vol. AI91, No. 38, pp. 67-73, 1991.
- 43) 森周司. 精神物理学の統一理論. 心理学評論, Vol. 36, No. 2, pp. 244-264, 1993.

- 44) 横瀬善正. 形の心理学. 名古屋大学出版会, 1986.
- 45) 野澤晨. 輪郭線図形について - 2 知覚の実験観察における輪郭線 -. 聖心女子大学論叢, Vol. 76, pp. 63-121, 1991.
- 46) 大山正, 今井省吾, 和気典二編. 感覚・知覚心理学ハンドブック. 誠信書房, 1969.
- 47) 野澤晨. 図形残効の研究における場の理論. 心理学評論, Vol. 9, pp. 68-97, 1965.
- 48) 吉田登美男, 深水義之, 野田健一. 横瀬の場の理論の実験的検証. デザイン学研究, p. 61, 1993.
- 49) 深水義之, 吉田登美男, 伊藤明, 小谷津孝明. 横瀬の実験の再実験とその検討. デザイン学研究, Vol. 111, pp. 25-32, 1995.
- 50) 宗宮保. 光覚閾で測定される場について - (1) 理論的考察 -. 心理学評論, pp. 163-174, 1965.
- 51) 高橋滯子. 光覚閾が及ぼす近接光の影響 - 横瀬の potential 公式と眼球内散乱光の問題をめぐって -. 心理学評論, Vol. 14, No. 2, pp. 368-386, 1971.
- 52) 野澤晨. 小点の消失閾による図形の場の強さの再検討. 日本認知科学会第 12 回大会, Vol. PD.6, pp. 202-203, 1995.
- 53) W. ケーラー著, 相良守次訳. 心理学における力学説. 岩波書店, 1951.
- 54) Sin Nozawa. The measurement of neighbouring field strength of a figure and the spatial anisotropy. *Abstract guide of XXth International Congress of Psychology*, Vol. IP 2-4-2, p. 331, 1970.
- 55) 横瀬善正. 図形認識と機械的読み取り. 人間工学, Vol. 2, No. 3, pp. 10-16, 1966.
- 56) 横瀬善正. 視知覚と場. 現代数学, Vol. 3, No. 4, pp. 16-22, 1970.
- 57) Zensho Yokose. The law of the field in visual form perception (i). *Japan Psychological Research*, Vol. 1, pp. 55-64, 1954.
- 58) 横瀬善正, 後藤倬男. 円および円弧図形の場の力の大きさの測定. 名大文研論集 XLV, 哲学, No. 15, pp. 35-44, 1967.
- 59) 伊東三四. 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 -. 心理学研究, Vol. 29, No. 3, pp. 11-19, 1958.
- 60) 深水義之, 吉田登美男, 小谷津孝明, 野田健一, 諸角建. 横瀬「場の理論」のコンピュータシミュレーション 景観の図形認識の立場からの解明と評価に関する研究 (1). デザイン学研究, Vol. 92, pp. 43-50, 1992.
- 61) 深水義之, 吉田登美男, 伊藤明, 小谷津孝明. 電磁気学による横瀬の場の理論的見直し - 景観の図形認識の立場からの解明と評価に関する研究 (3) -. デザイン学研究, Vol. 42, No. 6, pp. 39-46, 1996.
- 62) 藤井克彦, 松岡章, 森田龍彌. Lateral inhibition による錯視図形の解析. 医用電子と生体工学, Vol. 5, No. 2, pp. 117-126, 1967.
- 63) 深水義之, 伊藤明, 吉田登美男, 白石照美, 小谷津孝明. 側抑制に基づく図形の解析 - 景観の図形認識の立場からの解明と評価に関する研究 (4) -. デザイン学研究, Vol. 43, No. 2, pp. 69-76, 1996.
- 64) 深水義之, 伊藤明, 吉田登美男, 白石照美. 心理実験に基づく視空間伝達モデルの提案 - 景観の図形認識の立場からの解明と評価に関する研究 (6) -. デザイン学研究, Vol. 45, No. 4, pp. 75-82, 1998.

- 65) 吉田登美男, 深水義之, 白石照美, 伊藤明, 小谷津孝明. 心理ポテンシャルの心理物理的伝達関数の誘導と錯視への応用. ラウンドテーブル No.28, 日本心理学会第 62 回大会, pp. 23-32, 1998.
- 66) 中島和樹, 菅沼義昇, 伊藤正美. 心理的ポテンシャル場とニューラルネットによる図形認識. 電子情報通信学会技報, Vol. PRU90-155, pp. 79-86, 1990.
- 67) ファインマン著, 宮島龍興訳. ファインマン物理学 III. 岩波書店, 1990.
- 68) E. Sture Eriksson. Field effects and two dimensional form perception. *Scand. Journal of Psychology*, Vol. 8, pp. 218-242, 1967.
- 69) G. カニツツァ著, 野口薫訳. カニツツァ視覚の文法 - ゲシュタルト知覚論 -. サイエンス社, 1987.
- 70) 本川弘一. 感覚の生理学的基礎. 科学, Vol. 18, pp. 526-537, 1948.
- 71) 本川弘一. 生理学と心理学との境. 脳と神経, Vol. 10, No. 1, pp. 5-23, 1958.
- 72) Koiti Motokawa. Field of retinal induction and optical illusion. *J. Neurophysiol.*, Vol. 12, No. 475, pp. 413-426, 1949.
- 73) 本川弘一. 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. 139-150, 1967.
- 74) Koiti Motokawa and Tetsuro Ogawa. The electrical field in the retina and pattern vision. *Tohoku J. Exper. Med.*, Vol. 79, pp. 209-221, 1962.
- 75) 伊藤元雄. 幾何学的パターンに対するヒトの視覚誘発電位. 愛知学院大学文学部紀要, Vol. 23, pp. 1-15, 1993.
- 76) 鈴木竜太, 時田学, 山田寛, 巖島行雄. 顔面表情刺激と視覚誘発電位. 電子情報通信学会技報, Vol. HCS99-10, pp. 25-30, 1999.
- 77) 内山道明 (代表). 統合の場としての脳 知覚系 - 行動系の統一的理解への基礎的研究. 昭和 58 年度科学研究費補助金一般研究 (A) 研究報告書, 研究番号 56410001, pp. 13-30, 1984.
- 78) Shinji Kaji, Shigeru Yamane, Masaaki Yoshimura, and Noboru Sugie. Contour enhancement of two-dimensional figures observed in the lateral geniculate cells of cats. *Vision Research*, Vol. 14, pp. 113-117, 1974.
- 79) 巖壮志, 八村広三郎. 視覚ポテンシャルを用いる図形の最適配置. 情報処理学会第 46 回全国大会, Vol. 9C-8, 1993.
- 80) 三好正純, 下塩義文, 古賀広昭. 視覚的に均整のとれた図形配置の視覚の誘導場による評価. 平成 10 年電子情報通信学会総合全国大会, Vol. A-15-5, 1998.
- 81) 扇長洋子, 深水義之, 吉田登美男, 白石照美. アーチライズの見誤りと心理ポテンシャルの関連について. デザイン学研究, Vol. 43, No. 2, pp. 77-84, 1996.
- 82) 長石道博. 錯視図形がつくる視覚の誘導場の解析. TV 学会年次大会, Vol. 7-14, 1993.
- 83) 白石照美, 伊藤明, 深水義之, 吉田登美男, 小谷津孝明. 心理ポテンシャルモデル  $f(1)$  による M・L 錯視図形の解析. 日本心理学会全国大会, Vol. 3-108, p. 518, 1998.
- 84) 細木寛志, 黒田巧, 東恒人, 入江浩一. 図形の誘導場の簡易計算法. 電子情報通信学会総合大会, Vol. A-385, p. 386, 1996.

- 85) 高木敬雄, 小村緩岳. 主観的輪郭線の効果. ラウンドテーブル No.28, 日本心理学会第 62 回大会, pp. 2-10, 1998.
- 86) J.T.Townsend and F.G. Asbby. Experimental test of contemporary mathematical models of visual letter recognition. *Journal of Experimental Psychology : Human Perception and Performance*, Vol. 8, pp. 834-864, 1982.
- 87) S.E.Palmer. Structural aspects of visual similarity. *Memory & Cognition*, Vol. 6, No. 2, pp. 91-97, 1978.
- 88) S.J.Lupker. On the nature of perceptual information during letter perception. *Perception & Psychophysics*, Vol. 5, No. 4, pp. 303-312, 1979.
- 89) 田中平八. 文字パターン認知過程の構造について. 東京都立大学人文学報, Vol. 133, pp. 53-74, 1979.
- 90) 大山正, 今井省吾, 和気典二編. 新編感覚・知覚心理学ハンドブック, pp. 946-981. 誠信書房, 1994.
- 91) J. T. Townsend. The theoretical analysis of an alphabetic confusion matrix. *Perception & Psychophysics*, Vol. 9, No. 1A, pp. 40-50, 1971.
- 92) T. Kuennapas and A. J. Janson. Multidimensional similarity of letters. *Perceptual and Motor Skills*, Vol. 28, pp. 3-12, 1969.
- 93) 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, Vol. 81, pp. 13-20, 1977.
- 94) 村石昭三, 天野清. 幼児の読み書き能力. 国立国語研究所 (東京書籍), 1988.
- 95) B.Blesser, R.Shillman, T.Kuklinski, C.Cox, M. Eden, and J.Ventura. A theoretical approach for character recognition based on phenomenological attributes. *International Journal of Man-Machine Studies*, Vol. 6, pp. 701-714, 1974.
- 96) R.J. Shillman and B.A. Blesser T.Kuklinski. Psychophysical techniques for investigating the distinctive features of letters. *International Journal of Man-Machine Studies*, Vol. 8, pp. 195-205, 1976.
- 97) Mary J. Naus and 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*, Vol. 2, No. 3, pp. 394-400, 1976.
- 98) Makoto Yasuhara and Theodore T. Kuklinski. Category boundary effect for grapheme perception. *Perception & Psychophysics*, Vol. 23, No. 2, pp. 97-104, 1978.
- 99) 海保博之, 野村幸正. 漢字情報処理の心理学. 教育出版, 1983.
- 100) 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. 電子情報通信学会技報, Vol. PRL85-75, pp. 57-68, 1985.
- 101) 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. 電子情報通信学会技報, Vol. PRU90-64, pp. 47-68, 1990.
- 102) 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. 電子情報通信学会論文誌, Vol. J81-DII, No. 6, pp. 1115-1122, 1998.

- 103) 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, Vol. 2414, pp. 491-494, 1986.
- 104) 斉藤泰一, 山田博三, 山本和彦. J I S 第1水準手書き漢字デ - タベ - ス E T L 9 とその解析. 電子情報通信学会論文誌, Vol. J68-D, No. 4, pp. 757-764, 1985.
- 105) 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. 電子情報通信学会技報, Vol. PRU92-22, pp. 31-36, 1992.
- 106) 加藤隆仁, 森吉弘, 横澤一彦. 手書き文字品質の評価. *ATR Technical Report*, Vol. TR-A-0074, 1990.
- 107) 小山隆正, 山崎敏正, 山ノ井高洋. 流れの場のパターン弁別モデル. 電子情報通信学会技報, Vol. PRU88-64, pp. 33-39, 1988.
- 108) 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. 電子情報通信学会論文誌, Vol. J55-D, No. 5, pp. 299-306, 1972.
- 109) 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*, Vol. TR-H-064, 1994.
- 110) 市川伸一, 行場次朗. パターンの精神物理学における方法論的諸問題の検討. 心理学評論, Vol. 27, pp. 132-157, 1984.
- 111) 恒藤敏彦. 弾性体と流体, 第2章. 岩波書店, 1995.
- 112) 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, Vol. 834, 1983.
- 113) 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. 電子情報通信学会論文誌, Vol. J66-D, No. 6, pp. 722-729, 1983.
- 114) 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. オーム社, 1995.
- 115) K. Yokosawa and M. Shimomura. On the role of stimulus similarity and segmentation in misprint detection. In D. Brogan (Ed.), *Visual Search 2*, Taylor & Francis, pp. 369-376, 1993.
- 116) 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. 電子情報通信学会総合大会, Vol. D-541, p. 267, 1995.
- 117) 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. 電子情報通信学会論文誌, Vol. J81-D-II, No. 11, pp. 2624-2634, 1998.
- 118) Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, Vol. 1, No. 4, pp. 321-331, 1988.
- 119) 藤井俊治, 大森健児. 文字同士の接触パターンに基づく文字切り出し処理を用いた手書き文字列認識システム ~文字コード列生成器の開発~. 画像の認識・理解シンポジウム (MIRU'94), Vol. I, pp. 123-130, 1994.
- 120) 中村納, 鈴木弘之, 南敏. 横書き日本語文書における個別文字の抽出. 電子情報通信学会論文誌, Vol. J68-D, No. 11, pp. 1899-1909, 1985.
- 121) 西村康, 野口要治, 豊田順一. 新聞記事の本文を構成する文字の切りだし. 情報処理学会第24回全国大会, Vol. 3E-7, pp. 695-696, 1982.

- 122) 秋山照雄, 内藤誠一郎, 増田功. 非接触文字優先切出しによる印刷物からの文字切出し法. 電子情報通信学会論文誌, Vol. J67-D, No. 10, pp. 1194–1201, 1984.
- 123) 馬場口登, 塚本正敏, 相原恒博. 手書き日本文字列からの文字切り出しの基礎的考察. 電子情報通信学会論文誌, Vol. J68-D, No. 12, pp. 2123–2131, 1985.
- 124) 北村正, 仲林清, 大光明直孝, 中村修. 単語知識を利用した手書き文字列処理方式. *NTT R&D*, Vol. 39, No. 3, pp. 429–436, 1990.
- 125) Zensho Yokose. A study on character-patterns based upon the theory of psychological potential field. *Japanese Psychological Research*, Vol. 12, No. 1, pp. 18–25, 1970.
- 126) 長石道博. 自由手書き文字の効果的な収集法の提案と文書内における変動解析. ヒューマン・インタフェース研究論文集, Vol. 2, No. 1, pp. 19–26, 1993.
- 127) 長谷川純一, 輿水大和, 中山晶, 横井茂樹. 画像処理の基本技法 技法入門編 . 技術評論社, 1986.
- 128) 田村秀行監修, 日本工業技術センター編. コンピュータ画像処理入門. 総研出版, 1985.
- 129) 塚田敢, 湊幸衛. レタリングと文字認識. 心理学評論, Vol. 11, pp. 140–155, 1968.
- 130) 和田正信編. 印写工学 III 画像形成論, pp. 335–340. 共立出版, 1969.
- 131) 和田正信編. 印写工学 VI 視覚と心理, 記録, pp. 24–26. 共立出版, 1974.
- 132) 米山恒雄, 神山忠信, 井沢孝次. 図形パターンの検出感度によるディスプレイ表示文字の評価. テレビ学会誌, Vol. 50, No. 11, pp. 1768–1774, 1996.
- 133) NHK 放送技術研究所. より多くの情報を読みやすく.  
<http://www.strl.nhk.or.jp/open97/ex/z503/index-j.html>, 1997.
- 134) コマーシャル・フォト編集部. コマーシャル・フォト・シリーズ クリエイターのための印刷ガイドブック 1 基礎編, pp. 48–49. 玄光社, 1993.
- 135) 奥村晴彦. L<sup>A</sup>T<sub>E</sub>X 美文書作成入門, pp. 1–10. 技術評論社, 1991.
- 136) 大西誠一郎. 異なった文脈の中に位置づけられた文字の認知について. *Jap. J. Psychol.*, Vol. 38, No. 1, pp. 25–32, 1967.
- 137) 三好正純, 下塩義文, 古賀広昭, 井手口健. 視覚の誘導場理論を用いた感性にもとづく文字配置の設計. 電子情報通信学会論文誌, Vol. 82-A, No. 9, pp. 1465–1473, 1996.
- 138) 池田光男. 眼はなにを見ているか 視覚系の情報処理 自然叢書 8, 第 3 章. 平凡社, 1988.
- 139) 神部尚武. 読みの眼球運動と読みの過程. 国立国語研究所報告, Vol. 85, 研究報告集 7, pp. 29–66, 1986.
- 140) 芋阪良二, 中溝幸夫, 古賀一男編. 眼球運動の実験心理学, 第 8 章. 名古屋大学出版会, 1993.
- 141) 原島他. 顔論文特集. 電子情報通信学会論文誌, Vol. J80-D-II, No. 8, pp. 2029–2258, 1997.
- 142) 吉川佐紀子, 益谷真, 中村真. 顔と心 顔の心理学入門. サイエンス社, 1996.
- 143) Seiichiro Naito and James B. Cole. The gravity lens illusion and its mathematical model. *Contributions to Mathematical Psychology, Psychometrics, and Methodology edited by Gerhard H. Fischer and Donald Laming Recent Research in Psychology, Springer-Verlag New York, Inc.*, pp. 39–50, 1994.

- 144) 納谷嘉信. 色知覚と色彩メディア処理 (iii) 色の見え方の複雑さとそのモデリング . 電子情報通信学会誌, Vol. 81, No. 10, pp. 1057-1062, 1998.
- 145) 山田篤, 西田豊明, 堂下修司. 2次元平面におけるポテンシャルモデルを用いた位置関係の推定. 情報処理学会論文誌, Vol. 29, No. 9, pp. 824-834, 1988.
- 146) 佐藤正千代, 新濃清志. 現代数学レクチャーズ C - 4 ポテンシャル. 培風館, 1984.
- 147) 二宮信幸, 岸正倫. 総合報告 最近のポテンシャル論 I , I I . 数学, 10, 1958.
- 148) 中井敏春, 松尾香弥子, 加藤知佳子, 守谷哲郎, 岡田智久. 核磁気共鳴法を用いた脳機能測定法 ( f M R I ) の方法論入門. 認知科学, Vol. 5, No. 3, pp. 100-118, 1998.
- 149) 松下博, 森吉弘, 乾敏郎. スケルトンを用いた階層的形状記述およびニューラルネットによる階層的形状識別. 電子情報通信学会論文誌, Vol. J73-D-II, No. 6, pp. 812-819, 1990.
- 150) M. Nagaishi. Image processing system. *U.S. Patent, 5365600*, 1994.
- 151) スケデカー, コ克蘭著, 畑村又好, 奥野忠一, 津村善郎訳. 統計的方法 原書第6版, pp. 166-190. 岩波書店, 1973.
- 152) 大村平. 統計解析のはなし データに語るテクニク , pp. 211-213. 日科技連, 1989.
- 153) 石井健一郎, 上田修功, 前田英作, 村瀬洋. パターン認識. オーム社, 1998.
- 154) 田中敏, 山際勇一郎. ユーザーのための教育・心理統計と実験計画法. 教育出版, 1990.

## 付 録

### A 法線方向の決定

誘導場の認識モデルにおいて，式 (4.1)，(4.2) の弾性エネルギー  $U_1, U_2$  を求めるには，最初に図 4.8 のように等ポテンシャル線の輪郭線を求め，変位方向である法線方向を決定する．しかし，デジタル画像の誘導場から，等ポテンシャル面の輪郭線を直接求めると凹凸が大きく法線方向に大きな誤差が生ずる恐れがある．そこで，誘導場自身を平滑化後，誘導場からポテンシャル値ごとに等ポテンシャル面を得て輪郭線を抽出する．誘導場の強さは指数関数的に変化するので，最初にメジアンフィルタ，続いて 8 近傍濃度の荷重平均により平滑化した．

法線方向を決定する際，注目点近傍の輪郭線が滑らかな必要があるため，抽出した輪郭線自体をガウスフィルタによる畳込み積分で平滑化した<sup>149)</sup>．輪郭線の元々の形状が失われない範囲で経験的に  $\sigma = 4$  (dot) とした．

法線方向を注目点近傍のみを参照する差分法により高精度で求めるのは難しい．そこで，図 8.1 のように，輪郭線  $S$  上の注目点  $P$  を中心に  $5 \times 5$  dot の領域の輪郭線に接する接線を，点  $P$  を原点とする Hough 変換で推定する方法<sup>150)</sup> を用いて求め，法線方向  $n$  を決定した．

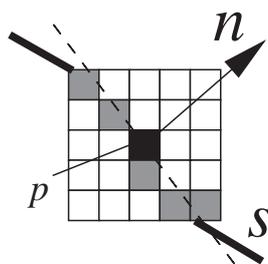


図 8.1 Hough 変換による法線方向の推定

### B 弾性エネルギーの計算

ある等ポテンシャル面における，式 (4.1)，(4.2) の弾性エネルギー  $U_1, U_2$  の計算方法を図 8.2 を例に説明する．図 8.2(a) のように，反時計回り並んだ  $n$  個の点で構成される等ポテンシャル線 A を等ポテンシャル線 B に一致させるとする．図 8.2(b) は点  $i$  近傍の変位を示した図である．図 8.2(b) の要領で，等ポテンシャル線 A の点  $i$  から法線方向上にある等ポテンシャル線 B までの変位  $u(i)$  を求める．変位  $u(i) = u(x(i), y(i))$

だから，式 (4.1)，(4.2) の偏微分は次の差分に直して計算できる．

$$\frac{\partial u}{\partial x} = x(i) - x(i-1) \quad (8.1)$$

$$\frac{\partial u}{\partial y} = y(i) - y(i-1) \quad (8.2)$$

$$\frac{\partial^2 u}{\partial x^2} = x(i+1) - 2x(i) + x(i-1) \quad (8.3)$$

$$\frac{\partial^2 u}{\partial y^2} = y(i+1) - 2y(i) + y(i-1) \quad (8.4)$$

$$\left(\frac{\partial^2 u}{\partial i^2}\right)^2 = \left(\frac{\partial^2 u}{\partial x^2}\right)^2 + \left(\frac{\partial^2 u}{\partial y^2}\right)^2 \quad (8.5)$$

式 (4.1) および式 (8.1)，(8.2) から  $U_1$ ，式 (4.2) および式 (8.3)，(8.4) から  $U_2$  が求められる．

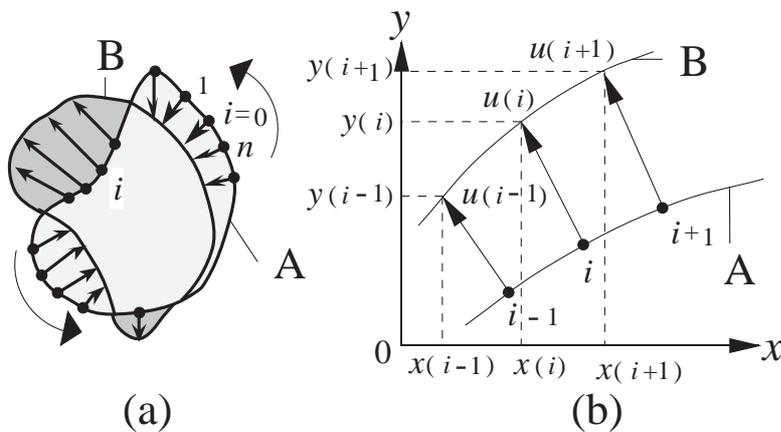


図 8.2 変位方向の求め方

## C LDCD 特徴 (Local Direction Contributivity Density Feature)

まず文字線構造 (方向、接続関係等の幾何学的形状) を 4 次元ベクトルで記述する方向寄与度 (Direction Contributivity Density) <sup>113)</sup> を説明する．方向寄与度  $d_m (m = 1, 2, 3, 4)$  は，図 8.3(a) に示すように，文字線分内にある点  $P1$  から 8 方向に触手を伸ばして決まる各方向の黒点連結長  $l_i (i = 1, 2, \dots, 8)$  を用いて次のように表される．

$$d_m = \frac{l_m + l_{m+4}}{\sqrt{\sum_{i=1}^4 (l_i + l_{i+4})^2}} \quad (8.6)$$

$d_1$  は水平方向， $d_2$  は右斜め 45 度方向， $d_3$  は垂直方向， $d_4$  は左斜め 45 度方向の方向寄与度成分である．

$d_1$  を具体的に計算すると，式 (8.6) から次式を得る．

$$d_1 = \frac{l_1 + l_5}{\sqrt{(l_1 + l_5)^2 + (l_2 + l_6)^2 + (l_3 + l_7)^2 + (l_4 + l_8)^2}} \quad (8.7)$$

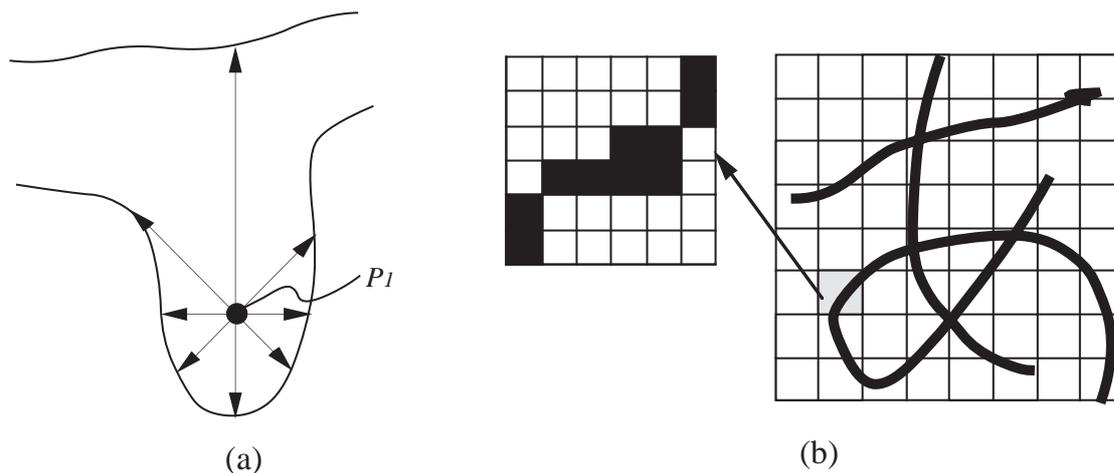


図 8.3 LDCD 特徴の求め方

例えば、点  $P_1$  は垂直方向の直線成分の割合が強いため  $d_3$  が大きくなる。このように方向寄与度は文字線分方向と更に接続関係を定量的に評価できる。LDCD 特徴は図 8.3(b) に示すように文字を  $m \times m$  の粗いメッシュに分割し、各メッシュ領域内の全黒画素についての方向寄与度を求め、それらをメッシュ毎に 4 方向別に投影した平均である。従って各メッシュにつき 4 個、文字全体では  $m \times m \times 4$  個の特徴量が得られる。

## D 円における視覚の誘導場の計算式

円の誘導場は、式 (6.2) の形式で表現できることを示す。図 8.4(a) のように、半径  $r$  の円の中心  $O$  から距離  $x$  の点  $P$  における誘導場の強さ  $M_p$  は、点  $P$  から円に引いた接線を  $PA, PB$ 、円と  $PO$  の交点を  $C$  とすると、点  $P$  から発した光が当たる円弧  $\widehat{ACB}$  のみで決まるから、文献<sup>44)</sup> より次式で与えられる。

$$\begin{aligned}
 M_p &= \int_{PA}^{PB} \frac{1}{x} dx \\
 &= 2 \int_{PA}^{PC} \frac{1}{x} dx = 2 \left[ \log |x| \right]_{x-r}^{\sqrt{x^2-r^2}} \\
 &= 2 \log \sqrt{\frac{x+r}{x-r}} \quad (8.8)
 \end{aligned}$$

円の直径  $d = 2r$ 、 $x - r = l$  とおくと式 (8.8) は、

$$M_p = 2 \log \sqrt{1 + d/l} \quad (8.9)$$

で表される。図 8.4(b) は、式 (8.9) で計算した  $d = 64$  における  $M_p$  を縦軸に、円弧からの距離  $l$  を横軸にとったグラフの例で、 $d = 64$  dot の円を実測した図 6.7(a) の分布に近い。このように、円の誘導場は円の直径と距離の関数で表される。

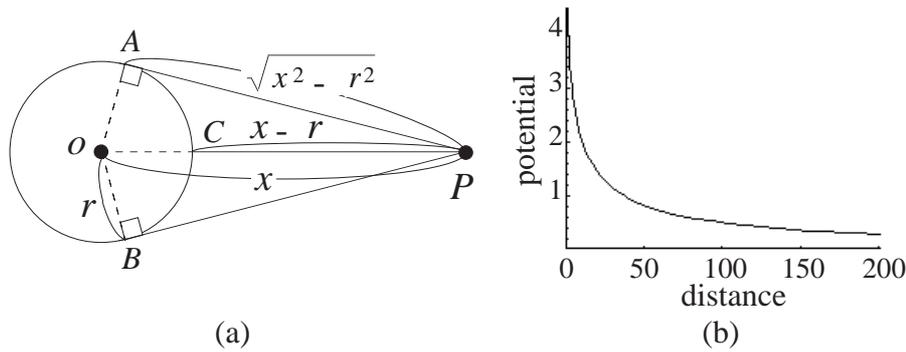


図 8.4 円がつくる誘導場

## E 相関係数の検定

2つの変数  $X_1, X_2$  間の相関係数を  $r$ ,  $X_1, X_2$  の要素数が共に  $n$  とする．母相関係数  $\rho$  は変数の要素数  $n$  について,  $\rho = 0$  のときほぼ正規分布する<sup>151)</sup>．しかし,  $\rho \neq 0$  のとき  $r$  は正規分布しないため, このままでは検定や推定ができないことが知られている<sup>151, 152)</sup>．そこで, 一般には  $r$  を次式を用いてほぼ正規分布する量  $z$  に変換した上で, 母相関係数  $\rho$  の t 検定や区間推定を行う<sup>151)</sup>．

$$z = \frac{1}{2} [\log_e(1+r) - \log_e(1-r)] \quad (8.10)$$

## F 視覚の誘導場における単位

本論文では, 以下の理由から誘導場の強さや弾性エネルギーを無次元として扱っている．

まず, 心理学的に誘導場の強さを表す物理的な単位が明らかにされていない．誘導場の強さは, 被験者に呈示した小光点の明るさが基準になっているが, その単位は研究者によってまちまちである．例えば, 横瀬は誘導場の強さを小光点の明るさを決める電圧値<sup>44)</sup>, 野澤は CRT の階調を決めるプログラムで設定する値<sup>52)</sup>, 深水, 吉田らは照度<sup>48)</sup> を用いている．このように, 誘導場の強さを表す単位は統一されていない．

誘導場の計算においても物理的な単位が定義されていない．横瀬のポテンシャル場の公式(式(3.3), (3.4), (3.5), (3.6))では, 誘導場の強さの単位は無次元である<sup>44)</sup>．例えば, 距離の単位が cm なのか触れられていない<sup>44)</sup>．一方, 深水, 吉田の  $F_{(1)}$  モデルは距離の単位は mm である<sup>65)</sup>．しかし, 誘導場の強さは横瀬の心理実験結果に近い値になるよう, 計算式に補正を加えて無次元化しており, 誘導場の物理的な単位は明確に定義されていない<sup>64, 65)</sup>．デジタル画像の場合, 三好らは式(3.19)について, 誘導場の強さを正規化しているが, 単位については定義していない<sup>137)</sup>．このように, 基本となる心理学実験で誘導場の強さの物理的単位が明確になっていないため, 誘導場の強さは無次元として計算されている．

本論文で提案した, 式(3.19)における距離  $r, n$  の単位はデジタル画像なので画素数(dot)である．しかし, この「画素数」というのはデジタル画像で仮定する仮想的なもので, 実際の長さ(mm など)には対応していない．そのため, 例えば, 本論文で用いた LDCD 特徴など, デジタル画像上の文字認識で使われる

特徴量や距離などには単位はつけない<sup>113)</sup>。また、式 (3.19) による計算や本論文で述べた誘導場によるパターンの違いは、それぞれ物理のクーロンポテンシャル、弾性エネルギーのアナログを利用しているだけである。実際にクーロンポテンシャルや弾性エネルギーを計算しているわけではない。弾性エネルギーのアナログは最近コンピュータビジョンでよく使われるが、エネルギーの単位はつけない<sup>118)</sup>。一般に、デジタル画像で計算を行う文字認識やコンピュータビジョンでは、特徴量や弾性エネルギーなどを無次元として扱うことが多い (153, 128)。

以上のように、心理学実験で誘導場の強さの物理的単位が明確化されていないこと、そしてデジタル画像は特徴量やエネルギーなどを無次元で扱うことが一般的なため、本論文では誘導場の強さおよび弾性エネルギーを無次元として扱った。なお、本論文では誘導場の再現や強さの比較ができることを考慮して、実験に使用したパターンの大きさ (dot)、パターンを作成した際の解像度 (DPI) などを示した。

## G 説明率による相関の強さの判定

相関の強さは、一般に相関係数の大きさによって経験的に表 8.1 のように「弱い相関」「強い相関」などと判定する<sup>154)</sup>。しかし、表 8.1 から、相関係数の値と相関の強さは直線関係がない。例えば、「弱い相関」の段階では相関係数の幅は 0.2 だが、「中程度の相関」の段階は 0.3 である。このように、相関係数の値が 2 倍になれば相関の強さも 2 倍であるとは言えない。相関係数  $r$  で相関の強さを理解しようとするときこのようなイメージをもちやすいので、説明率 ( $r^2 \times 100$  [%]) という指数を用いて相関の強さを判定する<sup>154)</sup>。例えば、 $r = 0.35$  の説明率は 12% である。 $r = 0.71$  (説明率 50%) と比べると相関係数は半分だが、説明率は 4 分の 1 以下の弱い相関であることがわかる。説明率と相関の強さの判定の対応を表 8.1 に示す。

表 8.1 相関の強さの判定

相関の強さの判定	ほとんど相関がない	弱い相関	中程度の相関	強い相関
相関係数の絶対値	0.2 未満	0.2 ~ 0.4	0.4 ~ 0.7	0.7 ~ 1.0
説明率	5% 未満	5% ~ 15%	15% ~ 50%	50% 以上



## 業績目録

### • 学術論文

1. 長石道博: "視覚の誘導場を用いた手書き文字の切出し", 電子情報通信学会論文誌, J 76-D-II, 9, pp.1948-1956 (1993)
2. 長石道博: "自由手書き文字の効果的な収集法の提案と文書内における変動解析", 計測自動制御学会: ヒューマン・インタフェース部会研究論文集, vol.2, No.1, pp. 19-26 (1993)
3. 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", テレビジョン学会誌, vol. 50, No.12, pp.1965-1973 (1996)
4. 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", 映像情報メディア学会誌, Vol. 52, No.4, pp.594-601 (1998)
5. 長石道博: "視覚の誘導場を用いた読みやすい和文プロポーション表示", 映像情報メディア学会誌, Vol. 52, No.12, pp.1865-1872 (1998)
6. 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", 映像情報メディア学会誌, Vol. 53, No.3, pp.458-465 (1999)
7. 長石道博: "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", 映像情報メディア学会誌, Vol. 53, No.8, pp.1176-1182 (1999)

### • 国際会議

1. Michihiro Nagaishi: "Analysis of individual variation in handwritten chinese characters in documents", The Third International Conference of Visual Search, U.K. (1992.8)
2. Michihiro Nagaishi: "Identifying ability of a recognition method based on the field of induction", Proc. International Conference on Document Analysis and Recognition, Tsukuba, Japan, pp. 926-929 (1993 10)
3. Michihiro Nagaishi: "A consideration of segmentation of handwritten characters using potential energy in induction field on the retina", 6th International Conference on Human-Computer Interaction, Yokohama, Japan, pp.94 (1995 7)
4. Michihiro Nagaishi: "Estimation of similarity for patterns by the induction field model in vision", The Second International Conference on Cognitive Science and The 16th Annual Meeting of the Japanese Cognitive Science Society Joint Conference, Tokyo, Japan, P1-22, pp.543-546

(1999.7.27)

● シンポジウム

1. 長石道博: "自由手書き文字の効果的な収集法の提案と文書内における変動解析", 第8回ヒューマンインタフェースシンポジウム 論文講演, 1224, pp.89-96 (1992.10)
2. 長石道博: "視覚の誘導場によるパターン認識の類似性に関する解析", 画像の認識・理解シンポジウム (MIRU'96), pp. I-73-78 (1996.7)
3. 長石道博: "視覚の誘導場を用いたパターン認識 - 視覚の誘導場の工学的応用 -", ラウンドテーブル No.28, 日本心理学会第62回大会, pp.14-22 (1998.10.10)
4. 長石道博: "文字列の読み易さを科学する - 視覚の誘導場による文字列の読み易さの定量評価 -", 平成10年電気関係学会関西支部連合大会 シンポジウム, S11-5 (1998.11.6)

● 特許 (登録)

1. Michihiro Nagaishi: "Image Processing System", U.S. Patent, 5365600 (1994.11.15)
2. 長石道博: "文字認識装置", 特開平 06 - 162261, 登録 No.2088667 (1996.9.2)
3. 長石道博: "画像処理装置", 特開平 06 - 274619, 登録 No.2117207 (1996.12.6)
4. 長石道博: "文字切り出し装置およびこれを用いた文字認識装置", 特開平 06 - 215182, 登録 No.2592756 (1996.12.19)
5. 長石道博: "文字認識装置", 特開平 06 - 131500, 登録 No.2611904 (1997.2.27)
6. Michihiro Nagaishi: "Character Recognition Apparatus", U.S. Patent, 5845020 (1998.12.1)

● 研究会

1. 長石道博: "自由手書き文字の変動分析", 1991 TV 年次大会, 10-10 (1991.7)
2. 長石道博: "自由手書き文字の筆記形態と変動の分析", 1991 信学秋全大 D-234 (1991.9)
3. 長石道博: "自由手書き文字データベース作成条件の検討", 1992 TV 年次大会, 11-7, (1992.7)
4. 長石道博: "視覚の誘導場理論を用いた文字認識の提案", 信学技報 PRU 92-46, pp.7-14 (1992.11)
5. 長石道博: "視覚の誘導場を用いた手書き文字切り出しの検討", 信学技報 PRU 92-120, pp.17-24 (1993.1)
6. 長石道博: "錯視図形がつくる視覚の誘導場の解析", 1993 TV 学会年次大会, 7-14, (1993.7)
7. 長石道博: "視覚の誘導場理論によるパターン認識の検討", 信学技報 IE95-30, pp.45-52 (1995.6)
8. 長石道博: "自由手書き文字における線形変形の分析", 信学技報 PRU95-119, pp.97-102 (1995.9)
9. 長石道博: "視覚の誘導場を用いたプロポーショナル表示の検討", 1997 映像情報メディア学会冬季大会, 7-3, (1997.12)