

脳機能と多次元データ解析を利用した 手書き漢字・ひらがな文字認識

鶴岡 信治

鈴鹿医療科学大学 医用工学部 医療健康データサイエンス学科

研究紹介

脳機能と多次元データ解析を利用した
手書き漢字・ひらがな文字認識

鶴岡 信治

鈴鹿医療科学大学 医用工学部 医療健康データサイエンス学科

キーワード： 光学式文字認識 (OCR), 機械学習, 画像認識, 多変量解析, 人工知能 (AI)

要旨

日本の手書き文字認識の研究は、字形が複雑で字種数の多い漢字の認識を目指し、欧米とは異なる手法を研究してきた。その手法は、実際に筆記された文字画像を解析し、独自の機械学習による統計解析手法であり、文字認識手法を取り入れた生産システム、郵便区分機などの製品が世界中で使用されている。この認識手法は、文字認識関連だけでなく、医用画像処理を含む幅広い画像認識の分野で使用され、今日の人工知能の基盤技術の一つとなっている。本原稿では、筆者が三重大学で研究開発し、提案してきた「加重方向指数ヒストグラム法」について、どのような背景で研究開発され、社会実装されてきたのかを具体的に解説する。この研究開発と社会実装の実例は、幅広い分野の方に役立つと考え、異分野の方に興味を持っていただけるように、数式を使用せず、平易に解説する。

1. まえがき

人の文字認識システムは、目から入った光の2次元分布（画像）を水晶体などのアナログ信号処理と神経細胞のデジタル信号処理を巧みに組み合わせて、脳の中で神経回路により抽象的な概念（字種）と結合させている。文字認識システムの開発は、人の文字認識システムを解明し、人工的に作成することを目指した人工知能の代表的な研究分野である。

1.1 光学式文字認識装置（OCR）の歴史

文字を読めるロボットを作成することは、「鉄腕アトム」、「ドラえもん」などの人と協調するロボットのアニメを見て育った筆者にとって、夢の技術である。初期の光学式文字認識装置（Optical Character Reader: 以下 OCR）は、文字型のスリットから出た光の量を電気信号に変換し、字種を推測する電気機械であり、1914年ごろにはゴールドバーグは文字列を電信符号に変換する機械を研究し、世界初の特許は1929年にタウシェクにより出願された¹⁾。その後、OCRはトランジスタの専用回路で開発されたが、1960年代前半のレイ社（米国）のOCRは約3億円であり、高額なために普及しなかった。1966年には東芝（日本）が世界初の郵便番号を自動読み取りできる郵便区分機を発明し、物流の高速化を実現した。1980年代にはOCRを含む第2次人工知能ブームが起き、1990年以降は低価格のOCRを開発するため、ソフトウェアとして実現された。

筆者が手書き文字認識の研究を始めたのは、1977年の大学院生の時からであり、手書き文字データベースが通商産業省の電子技術総合研究所で作成・公開され、手書き片仮名^{2,3)}から、漢字・ひらがな認識^{4,8)}と認識対象を変え、研究してきた。手書き文字の字形は筆記者により異なり（Figure 1）、字種を推定するには新しいデータ分析手法が必要であった。

筆者は、特徴量として「加重方向指数ヒストグラム」、識別関数として「疑似ベイズ識別関数」などを提案し、これらを組み合わせた「加重方向指数ヒストグラム法」を開発し、多数の学術論文を發表し、一連の研究をまとめ、博士論文を作成した⁹⁾。この博士論文には、文字認識方式の主要部分について、だれでも無料で自由に使用できるようにFORTRAN言語のソースコードを掲載し、多方面の方が評価実験に使用できるように公開し、今日でいうオープンソースを実施した。

1.2 脳科学に触発された文字認識方式

脳科学の中で文字認識に関係が深い研究は、ヒューベルとウィーセルにより行われた大脳皮質視覚野における情報処理に関する研究¹⁰⁾である。方位選択性コラムの存在は、文字認識の研究に大きな影響を与え、現在の文字を始めとする画像認識の特徴量として広く使用されている。

筆者は、大学院生の時には文字の中心線を折れ線近似し、その統計量により手書きカタカナ文字を認識していた^{2,3)}。しかし、同様の方法で手書きひらがなと漢字を

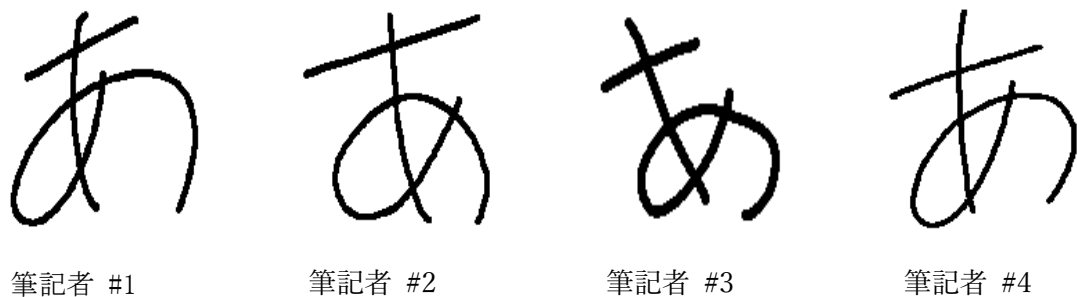


Figure 1 手書き文字の字形は、筆記者により異なり、字種を推定するには新しいデータ分析手法が必要

認識しようとする、中心線が安定して抽出できず、別の特徴量を使用した文字認識方法を模索していた。ちょうどその時（1981年）にヒューベルとウィーセルがノーベル生理学・医学賞を受賞したことにより、大脳の視覚野では輪郭線の方向性¹⁰⁾に注目して物体を認識することを知り、この特徴量を使用した文字認識の研究開発を始めた。研究成果⁴⁾を発表すると、当時だれも輪郭線の方向成分の度数分布を使用した文字認識システムを作成していなかったため、独創的な文字認識方法として注目された。

2. 加重方向指数ヒストグラム法の提案

2.1 特徴量（加重方向指数ヒストグラム）

1977年頃の手書き片仮名文字認識アルゴリズムは、英数字で使用されていた文字の中心線に関する統計的な手法が使用されていた。片仮名の文字線はほとんどが直線であるため、筆記された文字を折れ線近似し、線分の統計量に対して人間の字種識別基準から心理的距離を求め、未学習文字の正読率は99.1%となり、実用レベルを実現できた³⁾。しかし、手書きひらがなや漢字では、曲線や画数が多く、折れ線近似文字は元の文字と大きく異なる形となり、別の特徴量を模索していた。ちょうどその頃、脳の視覚情報処理は、量子化された方向成分の度数の組み合わせにより、画像を識別している¹⁰⁾ことを知った。この特徴抽出方法を使用して、文字認識システムを開発するために、文字領域の外接長方形を縦横のマス目で分割し、輪郭線を構成する方向線分の度数分布を特徴量とする「方向指数ヒストグラム」を考案した。

当初7×7のマス目の小領域で4方向の特徴量（196次元の特徴ベクトル、Figure2(a)）を考えたが、使用する記憶容量と計算量の大きさの制約から、特徴量の情報圧縮を検討した。当時、印刷文字に対する特徴量の情報圧縮技術として、視覚野の処理をモデル化した「ボケ変換」¹¹⁾が提案されていた。この変換（重み係数による圧縮）を使用し、13×13のマス目の小領域を4×4のマス目の小領

域（4×4領域×4方向＝64次元の特徴量）に圧縮する方法を提案し、「加重方向指数ヒストグラム」⁴⁾（Figure 2 (d)）と名付けた。この特徴抽出方法は、文字の複雑性とは関係なく、認識対象とするすべての文字画像に対して適用され、1つの文字画像はすべて64次元の特徴ベクトルで表現され、2.2節で説明する統計的識別関数により、文字を分類する。

2.2 識別関数（疑似ベイズ識別関数）

筆者は、当時のパターン認識や画像処理の論文で紹介されている種々の識別関数のプログラムを作成し、通商産業省の電子技術総合研究所から公開された「手書き文字データベース ETL8（教育漢字とひらがな）と ELT9（JIS 第1水準漢字とひらがな）」を対象として、多数のコンピュータ実験を行った。その結果、比較的良好な方法は、主成分分析法であることが実験データにより判明した。しかし、この方法は全クラスのサンプルを1つの特徴空間とし、多数の主成分を求め、サンプルの分布状況からサンプルの性質を要約し、外れ値により異常を発見（不良品を選択）する手法である。また従来の多変量解析の識別理論では、2クラス識別がほとんどで、文字認識のように多クラス識別に関する研究は極めて少なく、独自の研究を発展させる必要があった。文字認識では各字種に対して1つのクラスを設定しており、入力文字の特徴ベクトルについて既存の全クラスに属する確率を推定するために、多次元距離を求め（Figure 3）、入力文字を最も小さな距離値となるクラスに属すると判断している。

文字認識では、自クラス内でのまとまりとクラス間の分離性能が重要で、ベイズ識別関数を主成分で構成した識別関数は、学習サンプルのばらつきが大きい低次の主成分では高い識別性能を示すが、ばらつきが小さい高次の主成分では、大きな推定誤差が発生する。すなわち、学習サンプルでは存在しない変動が入力特徴ベクトルに入っている場合には、距離値が極めて大きくなる。そこで、推定誤差を少なくするために、高次の主成分に対する固有値は大きめの推定値を使用する「疑似ベイズ識別関数」を考案した。この識別関数は、文字認識、医用画像など

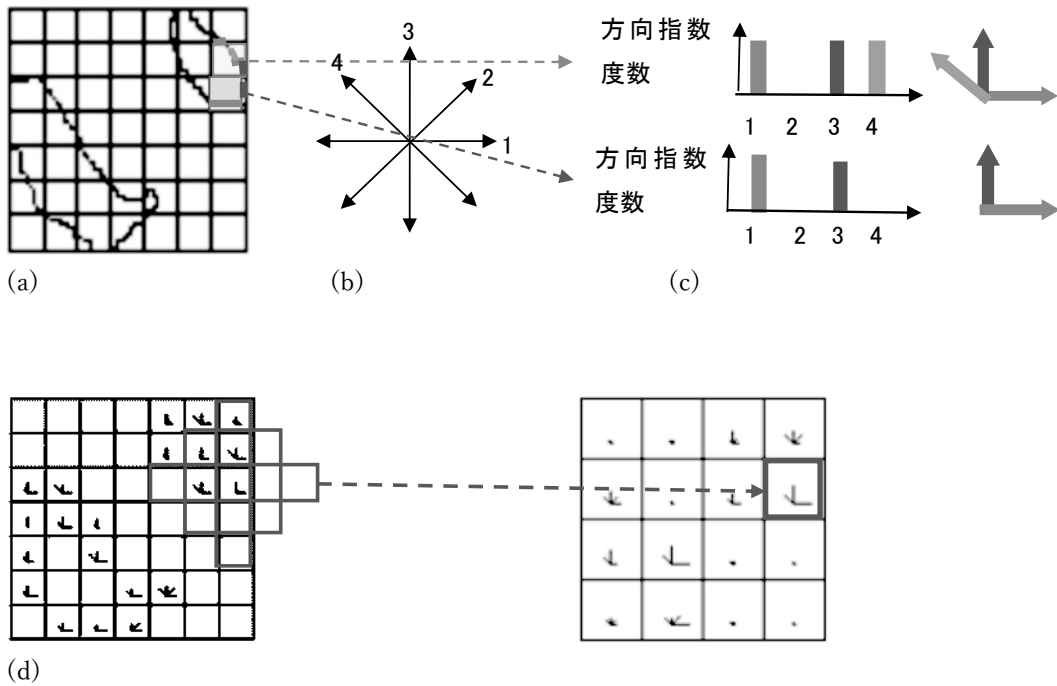


Figure 2 加重方向指数ヒストグラムの作成プロセス

- (a) 文字の外接長方形の 7×7 の小領域
赤枠と青枠で囲んだ小領域内の輪郭線の色は、輪郭線方向指数を示す。
- (b) 各輪郭点の接続方向を4方向に量子化(方向指数)
- (c) 輪郭線方向指数の度数の棒グラフと方向指数ヒストグラム
赤枠の小領域では、1の方向成分は上の横枠線に重なった部分と右下に横線、3の方向成分は右端の縦枠線に重なった部分の縦線、4の方向成分は、小領域内の右上の右下がり線を示す
青枠の小領域では、1の方向成分は下の横枠線の少し上の横線、3の方向成分は右端の縦枠線に重なった部分の縦線を示す。
- (d) 方向指数ヒストグラム (13×13 小領域) から加重方向指数ヒストグラム (4×4 小領域) の作成
 13×13 の小領域の4方向成分にガウス型の重み係数をかけて、 4×4 の小領域の4方向成分に情報圧縮する。

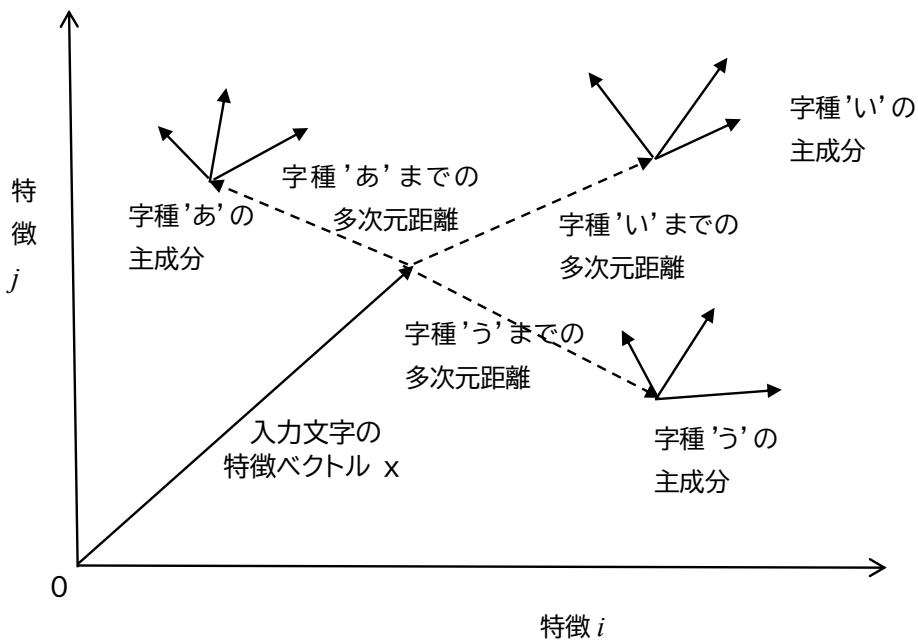


Figure 3 各字種の平均特徴ベクトルから入力文字の特徴ベクトルまでの多次元距離の計測

の多クラスの識別問題において高性能を示し、現在でも広く利用されている。この説明は、ここでは紙面の都合で省略するが、インターネットで全文が無料公開されている博士論文⁹⁾、オープンブック¹²⁾などの多数の学術論文で公開されているので、関心を持った方は参照していただきたい。

3. 文字認識技術によるイノベーション

筆者らが研究開発した文字認識方法は、トヨタ生産方式のかんばんの自動読み取り、郵便物の住所読み取り、各種帳票の自動読み取り、医療用紙カルテから電子カルテへの自動変換¹³⁾などに使用され、日本語だけでなく、英数字、中国漢字¹⁴⁾、タイ文字、アラビア文字、インドの多種類の文字などの世界中の言語で使用されている。提案した画像認識の特徴量「加重方向指数ヒストグラム」は、その後、画像認識の分野では Histogram of Oriented Gradients (HOG) と呼ばれ、幅広い画像認識の分野で使用され、自動車の自動運転¹⁵⁾などで使用されている画像処理の標準的な特徴量となっている¹⁶⁾。

また文字認識のような多クラス識別問題に対する機械学習の手法として提案した「疑似ベイズ識別関数」は、現在もディープラーニングの一部として使用されている。発展分野として、少数の筆記文字から、事前に手書き文字を登録してある登録筆者の中から、類似した文字を筆記する筆者（類似筆者）を推定し、推定した類似筆者の文字識別用データを使用した文字認識システム（パーソナル OCR）が提案され、使用頻度が高い職場ではすでに実用化されている。この考え方は、文字認識だけでなく、AI スピーカーなどの機械学習システムで広く使用されている。ビッグデータの利用により、さらに利用範囲が広がり、現在、クラウドに効率よくデータを記憶する研究開発が行われている。

OCR を利用したシステムは、手書き帳票の自動読み取り、手書き答案用紙の自動採点システム¹²⁾ など、デジタルトランスフォーメーション（DX）を推進する基盤技術として、現在も社会的な価値は高まる一方である。

4. あとがき

人間の脳で行われている文字認識の処理過程は、現時点では画像処理のごく一部が解明されただけで、画像から従来にない仮説を立て、ストーリーを構成するなど、画像の深い意味を理解する機能はほとんど解明されていない。道路に設置してあるナンバープレートの認識装置や産業用ロボットなど特定用途のための人工知能システムでは、人と同じように自分で文字が書かれている場所を見つけ出し、文字を認識することができる。しかし、その文字の意味を人間と同じように理解し、新しい概念（単語）を作り出す機構を解明する研究は、ほとんど進んでいない。文字認識は、例えば古文書を読み下すときに文字が書かれた背景（人間関係、心理状況、生活習慣、歴史の流れなど）を理解する必要がある。この背景を人工知能システムに能動学習させる処理は、人工知能の最も難しい課題であり、今後の研究が必要な研究テーマである。今日、世界的には DX が注目され、DX の基礎技術である文字認識の研究は世界中で活性化している研究分野である。この研究分野の応用範囲は広く、様々な社会問題の解決方法を提供することができる。特に、若い方々は DX を含むデータ・AI などのデジタル技術の利活用により、多くの人々が幸福感を感じることができる未来を創るよう努力していただきたい。

謝 辞

本原稿を執筆するにあたり、コミュニケーションツールとしての文字のデジタル化に関する意見交換、英文原稿の翻訳などご協力いただいた本学 Mark LaForge 准教授に感謝いたします。

文 献

- 1) 橋本新一郎, 文字認識概論, 電気通信協会 (オーム社), 1982.
- 2) 鶴岡信治, 木村文隆, 吉村ミツ, 横井茂樹, 三宅康二, デジタル図形の一細線化法と手書き文字認識

- システムへの応用, 電子通信学会論文誌, 1983; 66; 525-532.
- 3) 鶴岡信治, 村瀬晶彦, 木村文隆, 横井茂樹, 三宅康二, 人間の字種識別基準を用いた自由手書き片仮名文字認識, 電子通信学会論文誌, 1985; 68; 781-788.
- 4) 鶴岡信治, 栗田昌徳, 原田智夫, 木村文隆, 三宅康二, 加重方向指数ヒストグラム法による手書き漢字・ひらがな認識, 電子情報通信学会論文誌, 1987; 70; 1390-1397.
- 5) 三宅康二, 栗田昌徳, 鶴岡信治, 横井茂樹, 光学式文字認識, 公告平 2-59507 (出願 1983), 日本国特許庁, 1990.
- 6) 鶴岡信治, 森田裕之, 木村文隆, 三宅康二, 筆者に対して適応機能をもった自由手書き文字認識, 電子通信学会論文誌, 1987; 70; 1953-1960.
- 7) F. Kimura, K. Takashina, S. Tsuruoka, Y. Miyake, Modified Quadratic Discriminant Functions and the Application to Chinese Character Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 1987;9; 149-153.
- 8) S. Tsuruoka, H. Morita, F. Kimura, Y. Miyake, Handwritten Character Recognition Adaptable to the Writer, IAPR Workshop on CV, 1988; 179-182.
- 9) 鶴岡信治, 手書き文字の機械認識における特徴抽出と識別関数の研究, 博士論文 (名古屋大学), 1987.
- 10) D. H. Hubel, T. N. Wiesel: Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, J. Physiol. (London), 1962; 160; 106-154.
- 11) 飯島泰蔵, 文字読取装置 ASPET/71, テレビジョン, 1973; 27; 157-164.
- 12) S. Tsuruoka, M. Hattori, Y. Miyake, H. Takase, H. Kawanaka, Usefulness of Only One User's Handwritten Character on Offline Personal Character Recognition, Advances in Character Recognition, Intech Open, 2013; 211-232.
- 13) 梶原宏, 川中普晴, 山本皓二, 高瀬治彦, 鶴岡信治, 医療文書検索を目的としたシェーマ認識手法の提案とその評価, 医療情報学, 2014; 34, 65-73.
- 14) M. Cheriet, N. Kharma, C-L. Liu, C. Y. Suen, Character Recognition Systems A Guide for Students and Practitioners, WILEY, 2007.
- 15) F. Suard, A. Rakotomamonjy, A. Bensrhair, A. Broggi, Pedestrian Detection using Infrared images and Histograms of Oriented Gradients, Proc. of Intelligent Vehicles Symposium, 2006.
- 16) 内田誠一, AIによるくずし字認識の展望 文字認識の過去・現在・未来, 日本文化とAIシンポジウム 2019.

— プロフィール —

鶴岡 信治 鈴鹿医療科学大学医用工学部医療健康データサイエンス学科・教授 工学博士

[経歴] 1979年名古屋大学大学院工学研究科博士前期課程修了, 1989年三重大学工学部助教授, 1991-1992年ミシガン大学 (米国ミシガン州) 客員助教授, 2000年三重大学工学部教授, 2009-2013年三重大学大学院地域イノベーション学研究科初代研究科長, 2015-2019年三重大学理事 (研究・社会連携担当)・副学長, 2019年鈴鹿医療科学大学医用情報工学科・教授。[専門] データサイエンス (文字・文書認識, 画像診断支援システム, 人工知能, 地域イノベーション学など)。

Handwritten Kanji and Hiragana Character Recognition Using Brain Function and Multidimensional Data Analysis

Shinji TSURUOKA

Department of Health Data Science, Faculty of Medical Engineering,
Suzuka University of Medical Science

Key words: optical character recognition (OCR), statistical data analysis, image recognition, multivariate analysis, artificial intelligence (AI)

Abstract

Research on handwritten character recognition in Japan has been developing methods different from those in Europe and the United States, aiming at recognition of Chinese characters with complicated character shapes and many character types. The method is a statistical analysis method based on original machine learning that analyzes written character images, and products such as production systems and postal sorting machines that incorporate character recognition methods are used all over the world. This recognition method is used not only in character recognition but also in a wide range of image recognition fields including medical image processing, and is one of the basic technologies of artificial intelligence today. In this manuscript, I will explain concretely the background of research and development and social implementation of the “Weighted Direction Index Histogram Method (WDIHM)” that I researched and developed at Mie University. I believe that this example of research and development, and social implementation will be useful to people in many fields, and I will explain it in a straight forward manner without relying on mathematical formulas so that people from a wide variety of fields will be able to access this information.