

手書き文字認識を利用した 数独の自動正解判定システムの開発

金子 真二郎^{1,a)} 寺沢 憲吾^{1,b)}

概要：数独（ナンバープレース）は世界中で流行しており、数独に関する様々な電子ツールが開発されている。一方で、紙媒体に手書きで解くという数独の解答形式は依然として根強い人気を誇っている。本研究では、こうした手書きでの解答に着目し、手書き文字認識を利用して数独の解答を読み取り、自動で正誤判定を行うことで解答者を支援するシステムを提案する。具体的には、スマートフォンのカメラなどから入力された画像に対し、解答枠領域を抽出し、正方形に正規化した後に各マス毎に分割した上で文字認識を行って解答テーブルを取得し、これに対して正誤判定を行う。このシステムでは、数独ユーザが手書き解答の答え合わせにより負う手間を省くことに加え、数独を解く楽しみを増大させることが期待できる。本稿では、入力から出力までのシステムの概形の実装と、数字認識の改良の成果について述べる。

Automatic Verification System for Sudoku Puzzles Using Handwritten Character Recognition

SHINJIRO KANEKO^{1,a)} KENGO TERASAWA^{1,b)}

Abstract: Sudoku (Number place) is quite popular all over the world, and various electronic tools for Sudoku have been developed. On the other hand, it is still popular to solve Sudoku with handwriting the figures on paper. In this study, we focus on such solving styles and propose a system that assists such answers by automatically verifying the handwritten answers using digit recognition. Specifically, when the image is input from a camera of a smart phone, this system extracts the area of Sudoku answer, normalizes it for a square, divides it into 81 grids, gets the answer table using digit recognition, verifies the answer table, and finally output the verification result. This system can save user verification time and increase the pleasure of Sudoku solving. This paper describes the outline of the system from input to output and the result of improvement of digit recognition.

1. はじめに

本稿では、数独（ナンバープレース）の手書き解答を対象とした数独ユーザの支援システムについて提案する。数独とはペンシルパズルの一種であり、図 1 のようなあらかじめいくつものマスが埋められた 9×9 のマス目状の盤面を問題とし、ルールに従って残りのマスに 1~9 の数字を当てはめるといったものである。数独は世界的に流行して

おり、世界選手権も 2006 年以降毎年開催されている。また、書店のパズル誌のコーナーでは定期発行の数独専門誌が 10 誌以上並んでいるほか単行本も多く出版されており、また 100 円ショップに数独問題集のコーナーが作られていたり、一般の新聞・雑誌に数独問題が掲載されていたりすることなどからも、数独が愛好家から一般ユーザまでの多くの人々の間に幅広く普及し、楽しまれていることが見て取れる。

近年、科学技術の発展に伴い、様々な分野で電子化が進み、多くの電子ツールが実用化されている。これは数独を含むペンシルパズルにおいても同様であり、電子端末上で利用できる数多くのアプリケーションが存在する。最も一

¹ 公立はこだて未来大学大学院システム情報科学研究科
Graduate School of Systems Information Science, Future
University Hakodate

a) g2114007@fun.ac.jp

b) kterasaw@fun.ac.jp

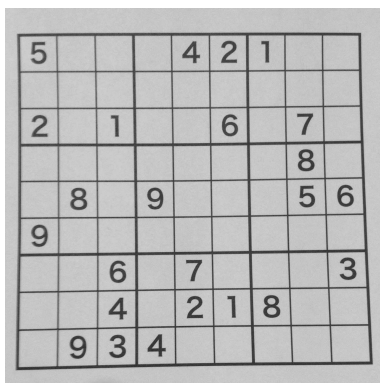


図 1 数独の問題例

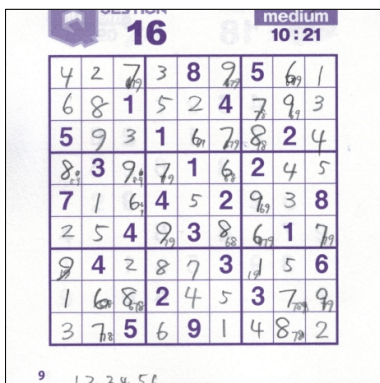


図 2 メモを含む手書きの解答例

一般的なのは、出題された問題をユーザが端末上で解答していく、ゲームタイプのアプリケーションだろう。例えば、パズル制作会社ニコリの Web サイトでは、数独が遊べる Web アプリやスマホアプリが公開されている [1]。その他のツールとしては、問題を入力することで自動で解答を求めるものや、解答過程を入力していくことで各マスに入る数字候補の明示やヒント表示、解答手順の記録といった解答支援を行うものなどが存在している。

また、数独は情報科学の研究対象としてもよく取り上げられる。主な研究テーマとして、自動解法 [2]、難易度判定 [3]、問題生成 [4] などに関する研究が盛んに行われている。これらの研究は、数独問題を数理的なアルゴリズムで処理できる対象として捉え、コンピュータを用いて各作業の自動化や支援を行う手法に関するものである。

こうした電子化の流れが進む一方で、それでも「紙媒体に手書きで解く」という数独の解答形式は依然として高い人気を誇っている。これにはいくつかの理由が考えられるが、特に手書きの「書き込みの自由度の高さ」という利点は重要度が高い要素であると考えられる。数独では、ある程度難易度の高い問題を解答する場合、図 2 のように途中でメモを記入しながら解答を進めるのが一般的である。この際、手書きであれば好きな場所に好きなようにメモや解答を記入できるのに対し、アプリではあまり自由な形でメモが書けないものや、そもそもメモ書き機能がついてい

ないものもある。これは、紙媒体の大きなメリットであると言える。この他にも、問題の紙と鉛筆 1 本でできる手軽さや、ペンシルパズルの名が表すように元々鉛筆で解くものであるというイメージが強いことなどが紙媒体の人気の理由として挙げられる。こうしたことから、現在も多くのユーザが鉛筆を片手に数独を楽しんでいる。

ところが、先に述べた電子ツールや研究の中に、紙に手書きで解答するユーザを意識してこれを支援しようとするものを見つけることはできなかった。もちろん、ユーザが紙の問題を見ながらコンピュータに手入力することで利用できるタイプのツールは存在しているが、これはユーザの負担が大きく、よほどの場合でなければユーザはこれを使おうとはしないだろう。前述の通り、数独は多くの人に「紙に手書き」という形で楽しまれている。したがって、それら手書きでの解答に対する便益向上を目的とした電子ツールを開発することは、多くのユーザの役に立つとともに、数独のファン層の拡大にも貢献するものと考えられる。そこで、数独の手書き解答を支援するシステムのあり方について検討した。

本研究では、手書きの数独の解答に着目し、これを撮影した画像（解答画像）に対して自動で答え合わせ（正誤判定）を行うシステムを提案する。このシステムの狙いは、ユーザが負う面倒を減らし、得られる楽しみを増すことである。一般にパズルというものは、自らの解答が正解であると明確に認められることにより爽快感が得られるものであるが、一方で数独はその特性上、解答後の答え合わせがクロスワードパズルなどと比べて面倒な作業になりがちである。これは、正解が単語で構成されているクロスワードパズルなどと異なり、数独の正解は数字がバラバラに並んだものであるため、1 マスずつすべてのマスを確認することがどうしても必要になるためである。さらに、クロスワードパズルなどではタテ列とヨコ列がクロスするマス目に文字を入れる際は両方の文字列を確認するという作業をしながら解き進めるのが一般的であるのに対し、数独ではタテ・ヨコ・ブロックのいずれか 1 つの制約によりマス目に入る数字が決定する際に、それが他の 2 つの制約を正しく満たしているかの検証は行わないで解き進めることが多い。これは、ミスなく解き進めていれば 1 つの制約が他の制約に矛盾することはあり得ないというパズルの特性に加え、世界選手権では解答時間も競われていたり、一般の問題集でも標準解答時間や目標解答時間が設定されていたりするように解答者は解答時間の短縮を目的とすることがあり、その場合、こうした検証を省略する方が解答時間短縮の上で有利になるからである。その結果、推論にミスがきわめて少ない上級者はともかく、しばしばミスを犯す初級者・中級者においては、解き終わった解答が正解でないことがしばしば起こりえる。そのため、主に初級者や中級者を対象として、その解答画像に対し自動的に正誤判定を行

うことは、解答者の答え合わせの手間を省くことができることに加え、解答が正解である場合にはそれを客観的に示すことにより解答者に爽快感を与え、数独を解く楽しみを増大させる効果があることが期待できる。

本システムを実装する上で要となるのが、書かれた解答をコンピュータに読み込ませてその解答内容を認識させるという部分である。なぜなら、各マスに書かれた数字さえ正しく認識できれば、あとの正誤判定処理は単純なアルゴリズムで済むためである。しかし、手書き文字は筆者によって多種多様な形状を持つため、これを正確に自動認識することは必ずしも容易ではない。たとえば、枠内に丁寧に書くという良好な条件の下で記述された郵便番号の認識においても、文献で見られる最高精度は 98.20% にとどまっている [5]。郵便物自動分類の実用においては、郵便番号認識だけでなく住所の文字を読み取った結果と併用することで精度を上げているのが現状である。また、文書画像認識における最大の国際会議である ICDAR2013 において行われた手書き数字認識のコンペティション Handwritten Digit and Digit String Recognition Competition では、孤立文字 (isolated handwritten digits) 部門と、連続する文字列部門に分けて行われたが、孤立文字部門においても参加プログラムの平均精度は 90.68% であり、最高のものでも 97.74% にとどまった [6]。このように、孤立文字の場合ですら、正確な自動認識は依然研究途上の課題である。加えて、今回対象とする数独の画像は、カメラ撮影の画像を前提とするため正規化やノイズ処理などが必要となる上、さらに自由なメモ書きが存在するため、認識を行う上で困難となる要素も多い。数独の正誤判定に実用できる精度の手書き数字認識を実装することは、本研究の大きなテーマとなる。

本稿では、2 章で関連する研究を紹介し、3 章で実装している提案システムについて述べる。4 章でこれまでの成果と今後の課題について述べ、5 章でまとめを述べる。

2. 関連研究

数独に関連する研究としては、前述のとおり文献 [2], [3], [4] のように様々のものが数多く存在する。しかし、手書きの数独を対象としたものや答え合わせをテーマとしたものは皆無であった。これは、本研究の着眼点の新規性を示している。

一方、本研究で必要となる数字認識、文字認識の分野についても、前述の通り、多くの先行研究が存在する。今回、提案するシステムを実装するにあたり、数字認識に関する処理については、若原の公開している資料「手書き数字認識の計算機シミュレーション [7]」を参考に全体を構成した。また、数字認識に使用する特徴量としては、若原が用いた輪郭方向分布特徴の他に、画像認識で広く使われているアルゴリズムである SIFT (Scale-Invariant Feature

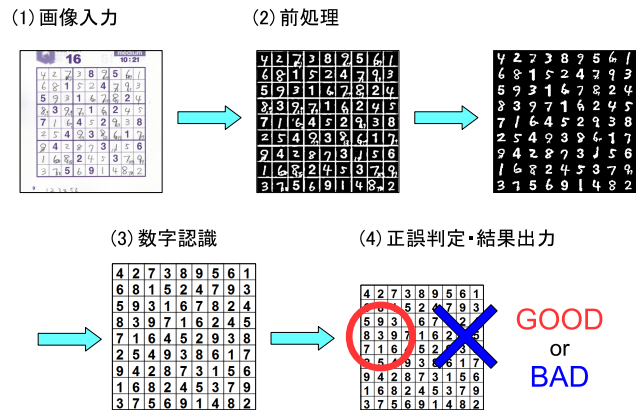


図 3 提案システムの流れ

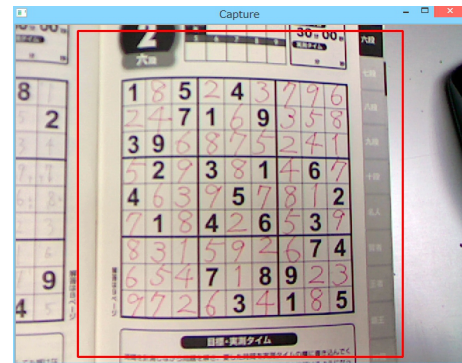


図 4 Web カメラによる入力画面

Transform) の特徴量を用いた [8], [9]。

3. 提案システム

本研究では、数独の手書きの解答を対象として、自動で正誤判定を行うシステムの実装を行う。システム全体の流れを図 3 に示す。それぞれの処理について、以下の節で説明する。なお、本システムを実装するにあたり、画像処理のためのライブラリとして「OpenCV 2.4.2」を、ラベリング処理のために井村のラベリングクラス [10] を使用している。

3.1 画像入力

まず最初に、対象となる数独の解答画像を入力する。入力方法としては、将来的にスマホアプリとして実装することを想定し、ファイル入力及びカメラと連動した直接入力を考えている。図 4 は、USB 接続の Web カメラを用いた入力画面である。

3.2 前処理

数字認識を行うための準備として、入力画像にいくつかの前処理を施す (図 5)。まず 2 値化した入力画像から解答のマス枠領域を抜き出し、さらに画像を各マス毎に分割する。次に分割した各マスの画像から、解答の数字以外の不要な要素 (主にメモ書き) を除去し、マス内での数字の位

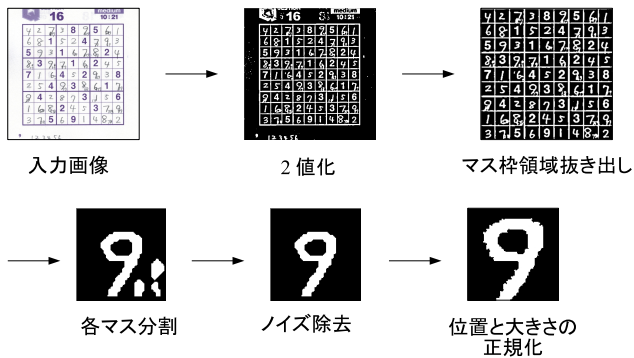


図 5 前処理

位置と大きさを正規化する。これらの処理により、画像を数字認識に使用できる形式にし、また同種の数字ができるだけ似た形状となるようにすることで認識精度の向上を図っている。

マス枠領域の抜き出しにはラベリングを用いた。2 値化した画像の中で最も大きい連結成分が解答枠の外枠を含む要素であると定め、その 4 隅を頂点とする正方形に射影変換している。また、ノイズ除去についても同様にラベリングを利用しており、各マスで最も大きい連結成分が解答を記述した文字であると定め、その他の要素を除去するという方法で行っている。位置と大きさの正規化については、文献 [7] の手法を用いた。

3.3 数字認識

前処理を施した画像から、各マスにそれぞれ何の数字が書かれているかを識別した結果を、解答テーブルとして取得する。この処理は、各マスの画像から特徴量を抽出し、識別器にかけて数字を識別するという流れで行う。以下の項では、特徴抽出及び数字識別の手法についてそれぞれ述べる。

3.3.1 特徴抽出

数字画像から、数字識別に利用するための特徴量を抽出する。抽出する特徴量としては、文献 [7] で紹介されている「輪郭方向分布特徴」、及び SIFT アルゴリズム [8], [9] で用いられている特徴量について実装し、比較検討を行った。比較の結果、4 章でも述べるが、現在は SIFT の特徴量をメインとして開発を進めている。

輪郭方向分布特徴は、画像を複数のブロック領域に分割し、各ブロック領域に含まれる文字領域の輪郭画素の方向を 4 方向（水平、垂直、右斜、左斜）に分類して集計し、それらの比率をブロック毎に算出して特徴ベクトルとして並べたものである。詳しくは参考文献 [7] を参照されたい。

SIFT のアルゴリズムは、特徴点の検出 (detection) と特徴量の記述 (description) の 2 段階からなる。本研究では認識対象部分は正規化領域として得られているため、このうち特徴量の記述部分の手法のみを採用し、正規化領域を記述する特徴量として用いた。具体的には、正規化領域

を 4×4 の 16 の小領域に分割し、各領域ごとに勾配ベクトルを 8 方向に集約して、128 次元のベクトルを作成し、特徴ベクトルとする。なおこの際、元々の SIFT で行っているような特徴点を中心とした重み付けは行わない。SIFT 特徴量の詳細については参考文献 [8], [9] を参照されたい。

3.3.2 数字識別

各マスの画像から抽出した特徴量を識別器にかけ、書かれた数字を識別する。ここでは、識別器に使用した学習用データと識別手法について述べる。

学習用データとしては、文献 [7] にて紹介されている「IPTP CDROM1B」というデータベースを用いた。このデータは、年賀はがきに記入された 3 桁郵便番号から採取した手書き数字の 2 値画像を整備したデータベースであり、0~9 の各数字について 2500~4000 枚程度の画像が用意されている。この中から数独に用いられる 1~9 の数字について、それぞれ無作為に 500 枚ずつ画像を抽出して、学習用データとした。また、印刷された数独の問題には、必ず初期配置として活字の数字が存在する。そこで、それらを上手く認識するため、手書き画像に加え、活字の数字画像 (MS ゴシック) を各数字について 1 枚ずつ学習用データに追加した。以上の総計 4509 枚分の数字画像について、輪郭方向分布特徴、SIFT の特徴量それぞれの特徴量データを作成した。

識別手法としては、各マスの画像と学習用データの画像の特徴ベクトル間のユークリッド距離を算出し、最も距離の近いデータが属する数字クラスを認識結果とする方法 (NN 法) を用いた。なお、距離の近い上位データのクラスラベルがバラけている場合についてはリジェクト処理を行い、書かれている文字の種別をユーザに問い合わせるという機能を将来的には検討しているが、現時点では未実装である。

3.4 正誤判定・結果出力

数字認識で取得した解答テーブルから、それが正答であるかを判定し、結果を出力する。判定処理については、2 通りの方法が考えられる。1 つは認識した解答テーブルを正答テーブルと比較する方法、もう 1 つは解答テーブルが数独の制約条件を満たしているかを調べる方法である。前者は使用するために正答データを必要とし、後者は解答テーブル単体で適用可能である。

解答の正誤を判定するには、基本的に後者の方法があれば十分である。ではなぜ前者の方法について検討しているかといえば、より柔軟な結果出力を行えるようにするためである。本システムは数独の答え合わせをするというものだが、結果をどのようにユーザに伝えるかは非常に難しい。なぜなら、どういった支援を求めているかはユーザによって異なるためである。制約条件を満たさないマスをも 1 つだけ教えてもらいたいユーザもいれば、正答と異なっている

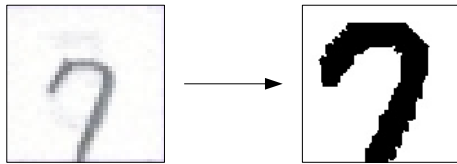


図 6 ミスしたマスの元画像と前処理済み画像

マスすべてを知りたいユーザもいるだろう。そのため、最終的に結果の出力は、ユーザの好みに合わせて設定できるものであることが望ましい。その際に、正答がない場合の判定だけでは、ユーザの要望に答えられないことが起こりえる。したがって、できるだけ多様なニーズに応えられるように、現時点では両方の可能性をにらんで研究を進めている。

正答テーブルを作成しようとする場合、問題の初期配置の数字を取得することが必要となる。そのためには手書き文字と印刷文字（活字）を見分けることが必要となるが、この方法については検討中であり、色情報を用いる方法や、パターンの同一性を利用する方法などについて調査している。

4. 成果と課題

4.1 これまでの研究成果

これまでの成果としては、まず画像入力から結果出力までに必要な各処理を実装し、それらを一連の処理として実行できるシステムの概形を完成させた。そしてそれを基に評価・検証をしながら、数字認識の精度向上のために各処理の改良を行った。数字認識の改善については、まずは図 2 の画像を入力とした実行について可能な限り認識精度を向上させる、という方針で行った。また、認識精度は全 81 マス中何マスを誤って認識したかを基準とし、ミス数をいかに減らすかを検討した。

具体的に行った改良としては、学習用データへの活字画像の追加、ノイズ除去処理の実装、SIFT の特徴量の実装、学習用データの利用方法の変更が挙げられる。先の 3 つの項目については、3 章で述べた通りである。学習用データの利用方法の変更については、元々は各クラスについて 500 枚の手書き画像の平均特徴ベクトルを数字識別に用いていたが、これをそれぞれ個別の特徴ベクトルとして利用するようにしたというものである。その他、プログラム上の細かい数値の調整なども行った。これらの改良の結果、初期の実装では 81 マス中 23 マスあったミス数を、輪郭方向分布特徴を用いた場合には 6 マスまで、SIFT の特徴量を用いた場合には 1 マスまで減らすことができた。また、輪郭方向分布特徴より SIFT の特徴量を用いた場合の方が総じて精度が高かったため、現時点では SIFT の特徴量を主として使用している。

SIFT の特徴量で残った唯一誤認識となった画像は、図 6

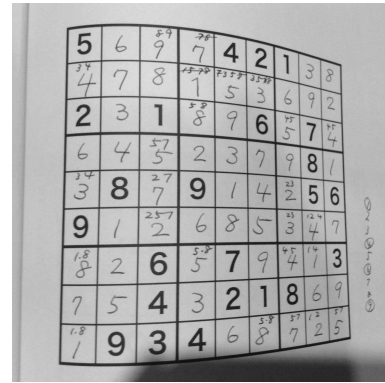


図 7 盤面の歪んだ入力画像の例

の「7」のマスである。プログラムでは、これを「9」と誤認識している。このマスと特徴ベクトルの距離が近い順に上位 10 個のデータのクラスラベルを並べると、{9, 9, 7, 7, 9, 7, 7, 9, 9, 7} という結果であった。このように、距離の近い上位のデータに複数のクラスラベルが混在しているような場合には、前述の通りリジェクト処理を行うことで誤認識を減らすことができると考えているが、その詳細については現在検討中である。

4.2 今後の課題

これまでの研究成果として、図 2 の画像を入力とした認識の精度に関しては、単純な NN 法で 98.8%(= 80/81) の精度を達成した。また、上位データのクラスラベルが混在している場合にリジェクト処理を行うことで、誤認識を減らすことができる可能性を示した。これらは、本システムの実用可能性を示唆するには十分なものであると考えている。しかし、任意の入力画像での使用についてはまだ検証しておらず、またその際にはさまざまなノイズや悪条件の影響で精度が低下する可能性がある。したがって、様々な画像に対応できるよう検証と改良を続けていくことが今後の課題となる。

任意の入力を用いる際に生じる問題の一例として、盤面の歪みが考えられる。今回用いた入力画像は、解答枠がおよそ正方形から歪みのないものであったため、各マスの分割は等分割で行うことができた。しかし、任意の入力では、カメラの角度や本の膨らみなどにより、図 7 のような歪んだ画像が入力されることもありえる。そのため、あまりに大きな歪みの場合は処理をリジェクトし、ある程度以下の歪みの場合はこれを補正して解答枠領域を抜き出し、マス分割を行えるような手法を現在検討中である。

数字認識の処理についてもいくつか改善案が考えられる。現在の識別器では最も素朴な方法とも言える NN 法を使用しているが、例えばサポートベクターマシンのような高度な手法に変更することで、認識性能の向上が見込める。また、同じ筆跡の同じ数字が複数存在するという数独の特徴に着目し、クラスタリングと組み合わせた数字認識手法に

についても検討中である。こういった純粋な認識精度の底上げは、最終的なシステムの有用性に大きく影響するだろう。

一方で、手書き文字認識という問題の性質上、単一文字の認識精度を 100%にすることは現実的には不可能である。そこで、単一文字の誤認識があったとしても、パズル全体の正誤判定の誤りを極力抑制する工夫を考えたい。これまでに何度か述べているリジェクト処理とユーザへの問い合わせはこれの一種である。また、単一数字の認識結果を確率的な出力として、それらのパズルとしての制約充足性を認識結果にフィードバックさせることで、結果的に認識の精度を高められる可能性もある。ただし、これはやり過ぎると誤解答を正しく検出できなくなる可能性もあるため注意が必要である。こういった面からも検討を重ねることで、システムをより実用的なものに近づけていきたいと考えている。

5. おわりに

本稿では、手書き文字認識を用いた数独の自動正誤判定システムの提案を行った。手書き数独の答え合わせというシステムの着眼点の新規性と有効性について述べるとともに、初期段階の実装の例と、その精度および将来的な改善可能性について示した。また、本システムにおける数字認識処理の重要性和難しさについても述べた。初期段階の実装の例を示すにあたっては、数独の解答画像を入力し、入力画像に前処理を施し、各マスに書かれた数字を認識し、正誤判定の結果を出力する、という一連の処理の意図と実装について記述した。

今後は、任意の入力画像に使用できるよう検証と改良を行う。現在考えているのは、歪んだ入力画像への対応、数字認識精度の底上げ、誤認識への対策についてなどである。また、将来的な実用化を視野に入れて、スマホアプリの開発やユーザインタフェースの整備についても考えていきたい。

参考文献

- [1] WEB ニコリ [パズル通信ニコリ] : <http://www.nikoli.co.jp/ja/index.html> (2014/9/28)
- [2] 佐藤裕二, 長谷川直広, 佐藤未来子, 並木 美太郎: メニューコアプロセッサによる進化計算を用いた数独解法の高速度化, 情報処理学会研究報告, Vol. 2011-HPC-130, No. 4, pp. 1-7, 2011.
- [3] 小場隆行, 中所武司: 数独の難易度判定アプリケーションの提案と評価, 情報処理学会研究報告, Vol. 2011-GI-25, No. 8, pp. 1-6, 2011.
- [4] 那須律政, 松崎公紀: モンテカルロ木探索による数独少数ヒント盤面の生成, 第 54 回プログラミングシンポジウム予稿集, pp. 173-180, 2013.
- [5] 堤田敏夫: 郵便番号データにみる手書き数字認識技術の現状, 郵政研究所月報, No. 103, pp. 38-52, 1997.
- [6] M. Diem, S. Fiel, A. Garz, M. Keglevic, F. Kleber, and R. Sablatnig: ICDAR 2013 Competition on Handwritten Digit Recognition (HDRC 2013), International

- Conference on Documents Analysis and Recognition, IC-DAR2013, pp. 1422-1427, 2013.
- [7] 若原徹, 手書き数字認識の計算機シミュレーション : http://cis.k.hosei.ac.jp/~wakahara/Introduction_to_Handwritten_Character_Recognition.pdf (2014/9/28)
- [8] D. G. Lowe: Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, Vol. 60, No. 2, pp. 91-110, 2004.
- [9] 藤吉弘亘: Gradient ベースの特徴抽出 — SIFT と HOG —, 情報処理学会研究報告, 2007-CVIM-160, Vol. 2007, pp. 211-224, 2007.
- [10] 井村誠孝, ラベリングクラス : <http://oshiro.bpe.es.osaka-u.ac.jp/people/staff/imura/products/labeling> (2014/9/28)