

数独画像に対する画像正規化と認識対象領域の抽出

苗加悠貴† 寺沢憲吾†

公立ほこだて未来大学 システム情報科学部 情報アーキテクチャ学科†

1 はじめに

数独はペンシルパズルの一種であり、 9×9 のマス目にルールに従って 1~9 の数字を埋めていくものである。

近年では電子ツールの普及により、数独の問題を Web サイト上やスマートフォン用アプリなどで提供するケースも増えているが、紙媒体に手書きする形での楽しみ方は依然として高い人気を誇っている。

本研究室では、紙媒体で解かれた数独をスマートフォンなどのカメラで撮影した画像に対して、自動的に正誤判定を行うシステムの開発を行っている。本論文では、その前処理部分にあたる、手書き解答を終えた数独の画像を入力として、 9×9 のマス目の画像を生成（正規化）し、それぞれの画像が最初から印刷された文字であるか、解答者が手で書いた文字であるかを判定するシステムについて述べる。

2 関連研究

紙面の曲がりによる歪みを修正する方法として、[1][2]などがある。これらの手法は紙面の形状を 3 次元的に推定し変換するというものであるが、これは比較的複雑な計算処理を要する。今回のケースでは数独の枠線部分が抽出可能であるという特性を活かした簡易で高速な変換方法を提案する。

3 数独画像の正規化

数独は 1 マスに 1 文字が記入されているため、正誤判定システムを作る上では数独画像のマス目ごとの分割が必要である。また、文字認識システムなどに画像を受け渡すことを考えると、マス目画像は統一された形状と大きさであるこ

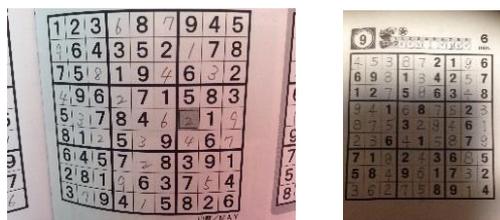


図1 入力画像の例

とが望ましい。この節では、入力された数独画像から、81 個の正規化されたマス目画像を取得する方法について述べる。入力画像については図1のようなものを想定している。

また、このシステム内では二値化処理を行うが、閾値決定には適応的閾値処理を使用する。これは、入力画像が図1右のような陰影の大きいものであっても適切な二値化処理を行うことができるためである。

3.1 数独部分の正方化

まず、撮影時の角度や紙面の反りによって曲がって写っている数独画像を正形状に変形（正方化）することを行う。入力画像の二値化とラベリングで外枠の抽出を行い、双線形補間をベースとした画像変換によって数独部分を正形状に変形させる。前提として、紙媒体での数独問題はほとんどが雑誌や単行本といった形式であるため、枠線のうち、曲がっているのは縦線、横線のどちらか一方であり、もう一方は直線として写っていると考える。

画像変換の手法について、図1左のような横線が曲がっている画像を例に説明する。変換後の画像の点 (x, y) は、外枠の上辺、下辺をそれぞれ $x:1-x$ に内分する点同士を直線でつなぎ、その直線を $y:1-y$ に内分する点を求めることで取得する。曲線の内分点を求めるためには曲線の長さを計る必要がある。これは、曲線を細かい直線の集合と考えて計算する。横線ではなく縦線が曲がっている場合は、 x と y の関係を逆にすることで対応する。

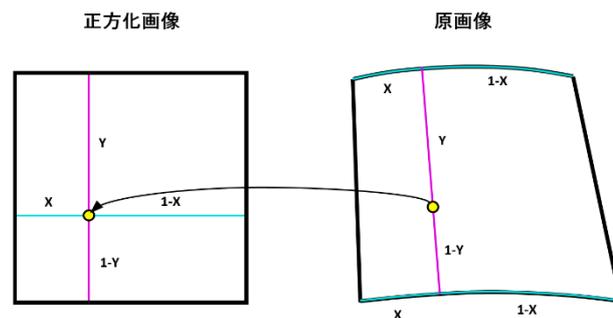


図2 正方化処理の概略

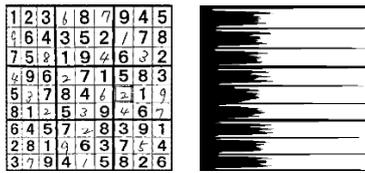


図3 正方形画像と横成分の射影ヒストグラム

3.2 罫線に沿った分割

正方形画像において、罫線は等間隔に位置しているとは限らない。罫線の位置を検出し、それに沿った分割を行う必要がある。

罫線の位置は、正方形画像に膨張処理をかけたのちに縦横それぞれの射影ヒストグラムを取ることによって検出する(図3)。罫線で囲まれた部分を再度正方形に変換することにより、正規化された9×9のマスキ画像を得ることができる。個々の画像サイズは40×40ピクセルとした。

4 印刷文字と手書き文字の判定

得られた9×9の計81個の正規化マスキ画像が、問題として最初から印刷されていた文字なのか、解答による手書きの文字なのかを判定する。

ここでは、「人が手で書く文字は、同じ文字であってもそれぞれの形が少しずつ異なるはずである一方、印刷文字の場合は同じ文字同士は形が同一である」という特性を利用する。81個の正規化マスキ画像同士を比較し、形がほぼ同一である画像が見られたときに両者を印刷文字であると判定する、という手法により印刷文字と手書き文字を区別していく。この手法は理論上、同じ印刷文字が2個以上存在する場合に有効である。

4.1 一致率、孔の数

2つの文字が同じ形かどうかを判定するために、「一致率」、「孔の数」という指標を用いる。2つの画像の一致率が大きく、かつ孔の数と同じである場合に「2つは印刷文字である」と判定する。印刷文字だと判定されなかった画像は手書き文字と判定する。

一致率は2つの画像を重ね合わせたときの、(両方が黒である画素数÷どちらかが黒である画素数) = (#AND/#OR)と定める。文字の位置ずれも考慮し、一方の画像を縦横に1ピクセルずつ動かしながら重ね合わせ、#ANDの値が最大となるときの一致率を求める。また、一致率に加えて画像の孔の数を比較することにより判定の精度の向上をねらう。

4.2 手書き文字“1”に対する処理

また、例外的に、縦棒のみの構造の“1”に対し

ては「手書き文字である」と判定する。これは、解答者が縦棒のみの構造の“1”を書いた場合、手書き文字であっても個々の差が出ず一致率が高くなり、「印刷文字である」と誤判定されてしまうことを防ぐためである。今回は文字の横幅が常に6ピクセル以下の画像に対しては「手書き文字である」と判定する。なお、ほぼ全ての数独画像において印刷文字の“1”は上部にひげ飾りが付いているため、この条件にはかからない。

4.3 判定システムの実験

数独の正規化マスキ画像81個の中から2つを比較し、「一致率が0.65以上であり、かつ孔の数が同じであり、かつ縦棒のみの構造ではない場合に両者を印刷文字と判定する」というシステムを実装し、実験を行った。

正常な正規化処理を完了できた、36枚の数独画像をサンプルとし、文字判定システムの実験を行った。文字の合計は2916個であり、そのうち印刷文字が1155個、手書き文字が1761個である。実験の結果、システムが印刷文字と判定した画像の数は1174個、印刷文字を正しく印刷と判定した数は1123個だった。これらの結果から、再現率と適合率を求めた結果、以下のとおりとなった。

$$\text{再現率} = \frac{\text{印刷文字を正しく印刷と判定した数}}{\text{印刷文字の数}} = 0.9723$$

$$\text{適合率} = \frac{\text{印刷文字を正しく印刷と判定した数}}{\text{印刷文字と判定した数}} = 0.9566$$

再現率、適合率ともに0.95を上回る高い数値を得ることができ、有用な判定であったといえる。

5 おわりに

本論文では、数独問題の手書きの解答に対して、数独部分の正規化と、認識対象となる箇所の抽出を行うシステムの実装とその有効性を確認した。今後は文字認識部の実装と、システム全体の改良を行っていく。

参考文献

- [1] Y. He, P. Pan, S. Xie, J. Sun, S. Naoi, “A book dewarping system by boundary-based 3D surface reconstruction”, ICDAR 2013.
- [2] 藤本克仁, 武部浩明, 諏訪美佐子, 直井聡, “平行測地線からの形状復元によるデジカメ文書画像の歪み補正”, 電子情報通信学会技術研究報告, PRMU2006-252, pp. 103-108, 2007.

“Image Normalization and Region Extraction for Handwritten Sudoku Images”

Yuki Noka†, Kengo Terasawa†, †School of Systems Information Science, Future University Hakodate.