

自由な筆記にロバストな手書き表認識システム

中洲 俊信 柴田 智行 井本 和範 三原 功雄*

概要. 自由に手書きされた表から、罫線、文字列、行や列などの表構造を安定して認識できる手法を提案する。手書きの表には、線の歪みや線同士の接続ずれなどの変動が生じる。変動にロバストな認識を実現するため、ユーザが手書きした多くの表から主要な変動パターンを洗い出し、それらに基づいて、罫線の抽出、表構造の認識、セル内ストロークの抽出の3段階で表を認識する技術を開発した。手書きの表 880 サンプルを対象とした評価実験で認識率 98.6 % を達成し、提案手法の有効性を示した。また、罫線の直線化や罫線同士が正しく接続するような修正により、表を視覚的に整える表整形インタフェースを試作した。

1 はじめに

近年、ペン入力に対応した携帯端末が注目され、タブレットやスマートフォンに書いた筆記をデジタルデータとして処理することが可能になった。頭の中のアイデアを思いのままに自由に書きとめ、それをすぐに整理・活用できることが価値の高い発想につながると考えられる。本稿では、様々な情報の関係性をわかりやすく可視化し情報整理に役立つツールとして頻繁に使われる「表」に注目し、ユーザが自由に手書きした表から罫線、文字列、行や列などの表構造を認識し、即座に綺麗な表へと変換するシステムを提案する(図1)。

手書きの表は、線に歪みがあったり線同士の接続がずれていたりする。また、ユーザによって書かれる表の形式も様々である。手書き表認識に関する従来研究では、このような手書きの表に含まれる変動を広く洗い出して整理したものは少ない。本稿では、事前検討として自由な手書きに伴う変動のパターンを洗い出し、主要な変動パターンに対してロバストな表認識手法を設計する。具体的には、筆跡の大まかな形状や関係性など特徴的な情報を利用し、単純な表のモデルに対応づけることでロバストに表罫線や表構造を認識する。Linらによる従来技術[2]も手書きに対するロバスト性を考慮しているものの、図1(a)のように外周の罫線が書かれていない表を想定しておらず、認識が困難であった。本稿では、外周の罫線が書かれていない表も主要な変動パターンであると考え、認識できるように対応する。

2 手書きの表の特徴分析

ユーザが手書きした表などを参考に、手書きの表に含まれる変動のパターンを洗い出した。抽出した事例を図2に示す。これらに加え、(i) 複数のストローク(ペンが入力面に接してから離れるまでの1

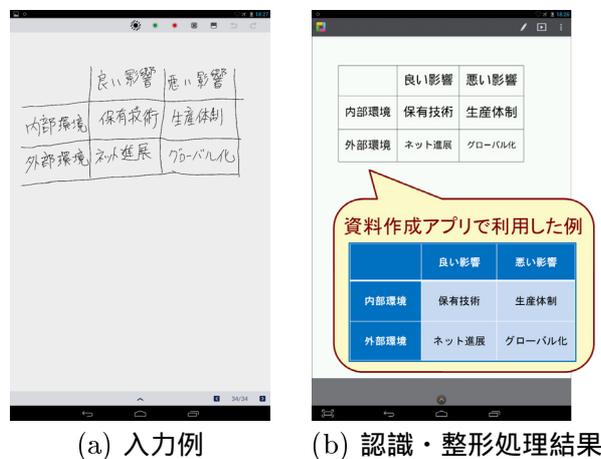


図 1. 筆者らが開発した手書き表認識システム [1]

画の軌跡)を追記して1本の罫線が書かれている、(j) 1本のストロークで複数の罫線が書かれているというパターンを抽出した。ここで、1本の罫線とは、表の左端から右端までを通る各行の境界、および、表の上端から下端までを通る各列の境界である。

3 手書き表認識手法

本提案手法では、ペンデバイスから得られた筆跡を入力とし、罫線の抽出、表構造の認識、セル内ストロークの抽出の3ステップで表を認識する(図3)。

3.1 STEP1: 罫線の抽出

長さが閾値以上で傾きが水平/垂直に近い線分を含むストロークを、罫線を構成するストロークとして抽出する。その後、1画で複数本の罫線が書かれている場合や複数画で1本の罫線が書かれている場合に対応できるように、方向が変化する点でのストローク分割や、同方向で位置的に近いストローク同士の統合を行い、水平/垂直の各罫線を認識する。

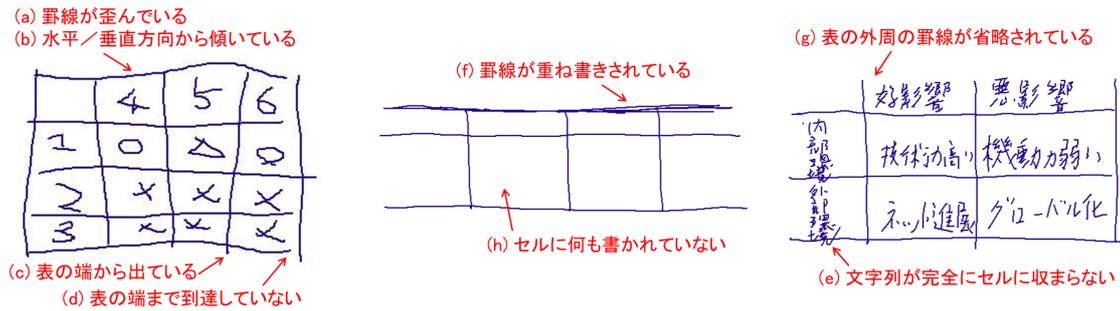


図 2. 手書きの表に含まれる変動パターン

3.2 STEP2: 表構造の認識

抽出した罫線を直線近似し、歪みのある罫線で構成された手書きの表を単純化する。その後、罫線の接続関係や、外周の罫線が省略された意図的な罫線のはみ出しを認識する。意図的に罫線がはみ出して書かれている場合、はみ出し部分が相対的に長かったり、はみ出し部分で構築された領域に文字列が書かれていたりする。この知見から、罫線のはみ出し部分が他のセルのサイズに対して所定割合以上の長さである場合や、はみ出した罫線付近に閾値以上の長さのストローク群が存在する場合に、意図的な罫線のはみ出しと判定し外周罫線を補う。意図的なはみ出しと判定されなかった場合は、垂直（水平）罫線の端点を最も端の水平（垂直）罫線に接続する。

3.3 STEP3: セル内ストロークの抽出

セル内の文字列の一部がセルからはみ出す場合も想定し、各セルに半分以上含まれるストロークをそのセルに内包するストローク群として抽出する。抽出したストローク群は一般の手書き文字認識技術により文字として認識する。

4 表整形インタフェースの試作

ユーザが手書きした表（図 1(a)）に対し、罫線の直線化、外周罫線の補足、罫線の端点での接続処理を施し、即座に視覚的に綺麗な表へと変換する。また、セル内の文字列を所定のフォントに変換する。なお、その場での表編集や、資料作成アプリの表オブジェクトとしての活用も可能である（図 1(b)）。

5 認識性能評価実験

被験者 40 名が作成した 880 サンプルの表を対象に、本提案手法の認識性能を調べた。これらのサンプルは、外周の罫線が無いものやセル内が空欄になっているものなど 11 種類の多様な表を被験者が 2 回ずつ書き写して作成した。また、各表には、行数、列数、罫線を構成するストローク、各セルに含まれるストローク群が特定できるように正解を教示した。

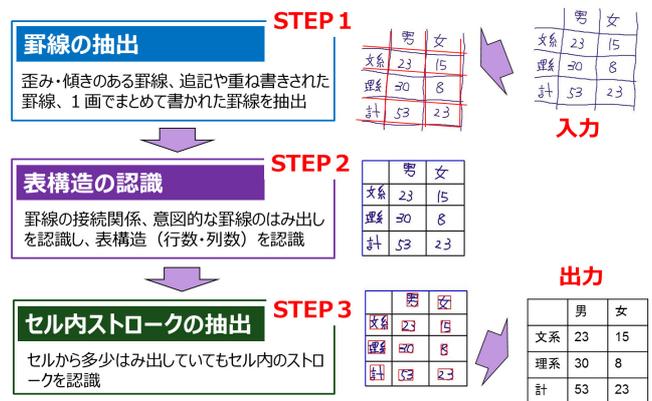


図 3. 手書き表認識システムの処理過程

本提案手法の認識結果の再現率と適合率の調和平均として認識率を算出した結果、98.6% となり、様々な形式の表を高精度で認識できていることがわかった。また、様々な書き方（罫線の歪み、接続のずれ、追記された罫線など）にも対応できていることが確認でき、本提案手法の有効性が示された。

6 まとめ

手書きの表から罫線、文字列、表構造を認識可能な表認識技術を提案し、表整形インタフェースを試作した。なお、本技術はペン入力対応タブレットへの搭載という形で実用化された [1]。今後はシステムの操作感に関するユーザ評価を予定している。

本稿に掲載の商品、機能等の名称は商標として使用している場合があります。

参考文献

[1] "東芝製ペン入カタブレット AT703", 東芝. <http://www.toshiba.co.jp/regza/tablet/>.
 [2] Z. Lin, J. He, Z. Zhong, R. Wang, and H.-Y. Shum. Table Detection in Online Ink Notes. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.28, No.8, pp. 1341-1346, 2006.