

手形・小切手認識技術

高橋 清

佐藤 元

金融機関においては、事務の合理化・省力化・効率化・視認性の向上を目的として、営業店・事務集中センタをはじめとするあらゆるチャネルで、帳票画像で業務を連携するソリューションを拡大する傾向が顕著となっている。

このようなソリューションを支える代表的な技術のひとつとして、文字認識技術がある。帳票をターゲットとした文字認識技術には、帳票規格に沿って作成された専用の定型帳票を読み取る定型帳票認識技術と定型帳票化されていない帳票を読み取る非定型帳票認識技術があり、対象ソリューションの拡大に伴い、認識対象帳票の傾向も定型帳票から非定型帳票への対応が求められるようになっている。

金融機関において、取扱いの多い非定型帳票のひとつとして手形・小切手がある。手形・小切手は、有価証券であるため、記載内容の確認やエントリに当たっては、一般帳票とは違った管理が必要である。手形・小切手の処理に対応した非定型帳票認識技術を利用することで、入力ミス・取扱いミス軽減の効果が期待できる。

本稿では、手形・小切手の管理に必要な、帳票種別、支払人、手形交換所、金額等の情報を、帳票画像から認識する技術を紹介する。

手形・小切手の仕様

手形・小切手認識技術で対象としている帳票は、「小切手」「約束手形」「為替手形」の3種類である。

我々はこの3種類の帳票に適応した文字認識技術を、東京手形交換所規則¹⁾の帳票規格をもとに開発している。これらの帳票は非定型帳票であるが、東京手形交換所規則では、帳票サイズや記載内容および記載位置が明確に規定されている。

規定されている内容は、帳票種別の表示位置や、MICR (Magnetic Ink Character Recognition) 方式による印字を行うクリアバンドの位置、手形交換所名等の表示を行う手形整理番号欄の記載方法と位置、金額欄の位置の様式等である。帳票様式の概略を図1に示す。

MICRとは、磁気インキによる専用の字形の印字を読み

Figure 1 illustrates the general layout of three types of bills: 小切手 (Check), 約束手形 (Promissory Note), and 為替手形 (Exchange Bill). Each bill format includes a header section with fields for bank/branch (銀行/支店), amount (金額), and a serial number (整理番号) box. Below the header is a large MICR band (クリアバンド) used for automated processing. The 約束手形 and 為替手形 formats also include fields for payment date (支払期日), payment location (支払地), and payment place (支払場所).

図1 帳票様式の概略

取る仕組みであり、JIS X 9002 (磁気インキ文字読取用字体および印字仕様) によって、この字形および印字仕様が定められている。この専用の字形が、図2にあるE13Bフォントである。

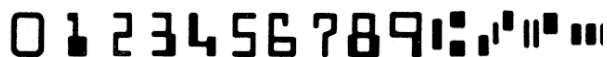


図2 E13Bフォント

手形・小切手の金額欄は、東京手形交換所規則によって、チェクライタによる印字か、手書きによる大字 (漢数字) の記載と規定されている。

チェクライタの字形は、改ざん防止のため、通常の文字と異なり、細い平行線によって構成された刻み文字

¥12,345,678

図3 チェックライタの印字例

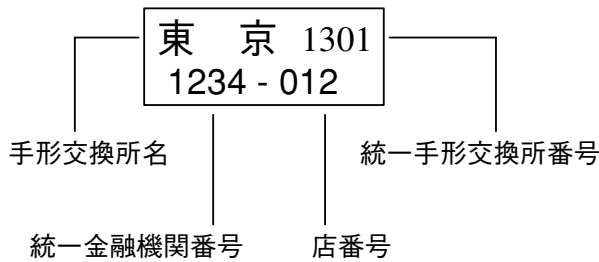


図4 手形整理番号欄の印字例

になっている。チェックライタによる印字の例を図3に示す。

また、手形整理番号欄は、図4に示すように2行に渡って手形交換所名、統一手形交換所番号、統一金融機関番号および店番号が記載されている。

この規格に基づき、手形・小切手の帳票画像を具体的に認識するための仕様として、表1の認識仕様を規定している。

表1 認識仕様

項目	内容
読取媒体	「東京手形交換所規則集」で定められた規格・様式に従う手形・小切手であること。
読取画像	解像度：200dpi～400dpi
読取箇所 読取字種	①証券種類 印刷漢字 小切手、約束手形、為替手形 ②MICR E13Bフォント ③金額欄 チェックライタ印字

手形・小切手認識技術

我々は、帳票画像に対する2段階のアプローチで手形・小切手の認識処理を行っている。それが、帳票識別技術と文字認識技術である。

「小切手」「約束手形」「為替手形」の3種類の帳票は、それぞれ、記載内容が異なっており、営業店・事務集中センタでの業務フローも異なる場合があるため、まず帳票を分類する必要がある。これを実現する技術が、帳票識別技術である。帳票識別を行った後、それぞれの帳票の記載内容を文字データ化する技術が、文字認識技術である。

金融機関では業務フロー上すべての手形・小切手の記載内容を必要とはしていない。優先度の高い項目は、手形交換所や金融機関の情報、金額なので、MICR、金額欄、手形整理番号欄の文字認識を行っている。手形・小切手は、帳票識別技術と3種類の文字認識技術の計4つの技術により、図5に示すフローで処理される。

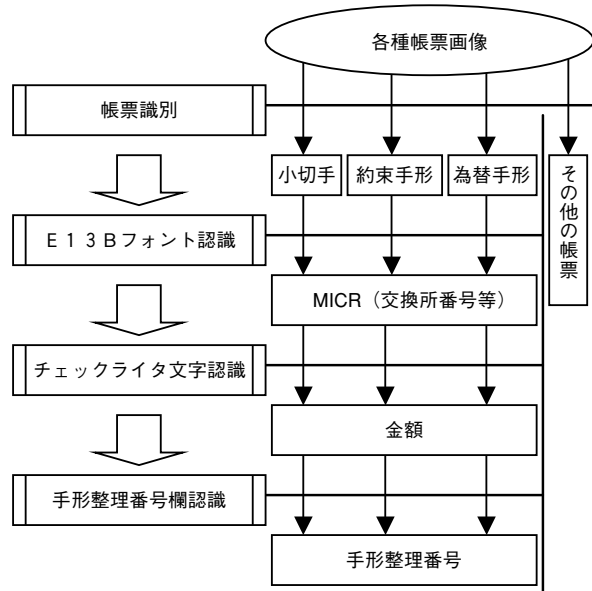


図5 手形・小切手認識フロー概略

その4つの技術が、帳票種類の文字認識を行うキー項目識別技術を用いた帳票識別技術、MICRの文字認識を行うE13Bフォント認識技術、金額欄のチェックライタ文字認識技術、手形整理番号欄認識技術である。個々の技術について説明する。

帳票識別技術

営業店や事務集中センタでは、手形・小切手に限らずに、多くの種類の帳票を扱っている場合が多くある。このような場合、複数の帳票識別技術を用いることで、分類を行うことができる。

現在、帳票画像を扱うソリューション向けに提供している技術として、シートID識別、キー項目識別、罫線特徴識別、領域特徴識別といったものがある。それぞれの識別技術の概略を表2に示す。

手形・小切手の帳票識別には、キー項目識別技術を用いている。東京手形交換所規則によって、証券の種類である「小切手」「約束手形」「為替手形」という記載を必ず行うことが規定されているため、この文字列を認識した結果は、帳票のユニークな情報として適している。

表2 帳票識別技術

名称	概略
シートID識別	帳票上の定められた位置に印刷してある帳票固有の番号を読み取る事で識別を行う。
キー項目識別	帳票毎にある固有の場所に固有の文字列を読み取る事で識別を行う。
罫線特徴識別	帳票毎にある固有の罫線を抽出する事で識別を行う。
領域特徴識別	帳票毎にある特定部分の形状を利用して識別を行う。

E13Bフォント認識技術

営路店や事務集中センタの業務負担の削減には、MICRの情報が必要となる。

本来は、MICR専用の磁気リーダによって読み取るものだが、我々は帳票画像で業務を連携するソリューションとして、文字認識技術によって読み取る機能を実装している。

MICRで使用しているE13Bフォントは、通常の印字フォントとは異なっているため、専用の認識辞書によって実現している。

この方式のメリットは、MICR専用の磁気リーダが必要ないことと、画像ファイリング済みのデータや、事務集中センタへ集められた画像データに対して、後から認識を行うことができることである。

このMICRを読み取ることで、手形交換所番号等の情報が取得できる。

支店番号や口座番号等の情報は、金融機関ごとの記述ルールに従って記載されているため、認識の後処理技術を行い、金融機関ごとのカスタマイズをすることで、利用できるようにしている。

チェックライタ文字認識技術

現状流通している多くの手形・小切手が、チェックライタによる印字を行っているため、チェックライタ印字が行われているときのみに金額欄の認識を行っている。

チェックライタ文字認識を行うためには、下記のような3つの課題があり、それぞれに対応する技術によって、克服している。

- ①印字位置のバラツキ
- ②金額欄への手書き・スタンプの侵入
- ③刻み文字（チェックライタ印字）

1つめの課題は、印字位置が変わることである。発行す

る金融機関ごとに手形・小切手のレイアウトは異なっており、チェックライタを印字すべき金額欄の位置も若干のバラツキがある。しかし、それ以上に問題となるのは、チェックライタでの印字を行う場合、チェックライタの装置ごとに位置合わせが異なるため、印字位置が厳密には決まらないことである。

図6に示すように小切手の金額欄に対して、チェックライタ印字の位置は、大きなバラツキが発生している。

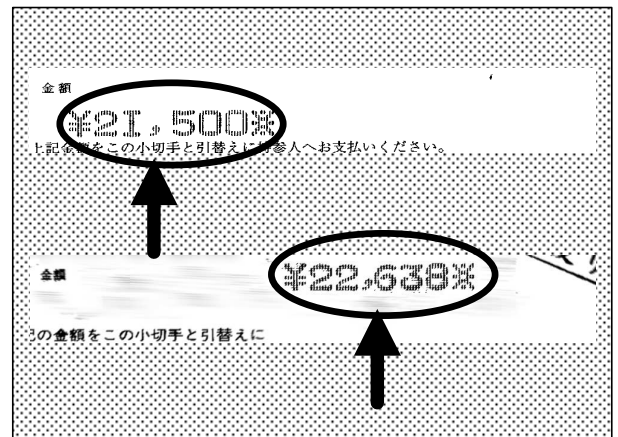


図6 印字位置のバラツキの例

この課題に対しては、非定型文字認識技術の応用として、周囲に印字してある文字列も含めて、すべて文字認識を行い、その認識結果から、金額欄の内容として最も妥当な結果を選択している。

また、妥当な認識結果が存在しない場合は、金額欄の結果を不読としている。

2つめの課題は、手書き・スタンプなどによる侵入のため、予期しないものが、金額欄周辺に重なってくることである。

図7の例では、スタンプ・2次記入の手書きの金額、検印等が金額欄の領域内に存在している。

これは、侵入している領域の画像の特徴を取り出す技術によって、文字列としての特徴を持っているかを判断し、除去を行うことで実現している。

3つめの課題は、チェックライタ印字は独特の刻み文字になっていることである。

また、単一の字形ではなく、チェックライタの装置ごとに、さまざまな字形が存在している。そこで、多くのチェックライタの装置で印字した字形に対して学習した専用の認識辞書を開発することで、チェックライタ文字の認識を実現している。

ただし、チェックライタ装置の種類だけ字形のバリエーションが存在するため、その種類は非常に多く、未学習

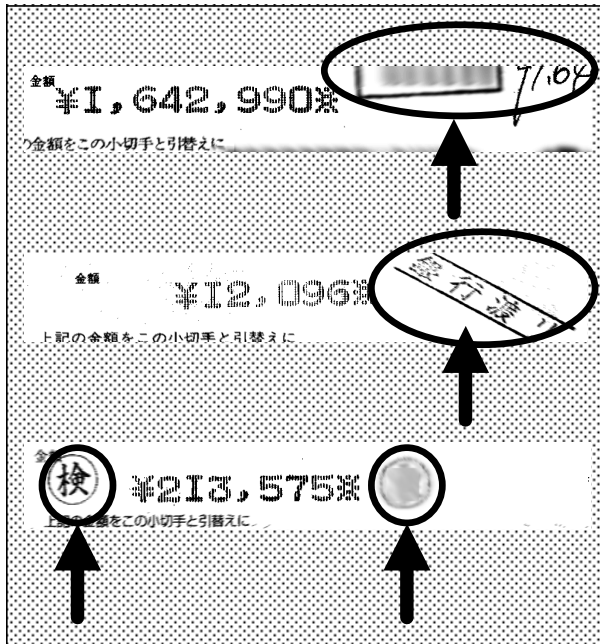


図7 金額欄への侵入の例

のチェックライタの字形に対しては、学習を続けていく必要がある。

手形整理番号欄認識技術

手形整理番号欄は、東京手形交換所規則によって、明確に規定されているが、実際の印字位置はバラツキが発生している。

そこで、記載領域の周囲にある罫線を検出し、位置の補正を行うことで、印字位置のバラツキを吸収している。次に検出した罫線を除去し、領域内の文字列だけを取り出して、文字認識を行っている。

手形整理番号の認識結果は、営業店・事務集中センタでの手数料自動計算などに用いることができる。また、MICRの認識結果とのペリファイを行うことで誤認識を低減できる。

帳票識別技術、E13Bフォント認識技術、チェックライタ文字認識技術、手形整理番号欄認識技術により構成されている手形・小切手認識技術は、Image/OCRコンポーネントに実装している²⁾。

あ と が き

帳票画像で業務を連携するソリューションにおいて、手形・小切手という非定型帳票を扱うための認識技術を紹介した。

営業店・事務集中センタの業務での事務負担軽減のため、今後、手形・小切手認識技術に対する要求は更に増

えてゆくと考えられる。

このような要求に応えるために、更なる認識性能の向上、読み取り範囲の拡大といった技術開発を進めていく予定である。 ◆◆

参考文献

- 1) 東京手形交換所規則 社団法人 東京銀行協会
- 2) 能埜, 橋本: “Image/OCRコンポーネント”, 沖テクニカルレビュー191号, Vol.69 No.3, p.98-101, 2002年7月

筆者紹介

高橋 清: Kiyoshi Takahashi.金融ソリューションカンパニー 金融ソリューション開発本部 イメージソリューション開発部 開発第二チーム

佐藤 元: Hajime Sato.金融ソリューションカンパニー 金融ソリューション開発本部 イメージソリューション開発部 開発第二チーム