

特集2： 新しいコードの提案

紙の書類における自動読取システム の構築

(株)アポロジャパン
顧 澤 蒼

BA0806-18 0915-1060/08/ ¥500論文/JCLS

はじめに

2005年に個人情報保護法が発表され、前面実行されてきた。その中で特に個人情報の40%以上が紙媒体の書類からの情報漏洩であるという現状をうけて、複合機のセキュリティ機能向け紙の書類に情報の埋め込む研究が注目されている。

また近年、紙の書類の保存・運用・管理のコストを減らすため、国家政策の推進により、紙の書類から電子化への移行が急速に進んでいる。紙媒体は、直感性、安価性、立証性などの利点を持つため、いかに電子化が進んでも、紙媒体は長く存続していくと思われる。その紙媒体に対して、もっと有効に利用し、効率的に電子化へ移行するためには、大容量の情報を埋め込むことができるような技術が不可欠である。

従来、紙の書類から電子化への処理をするために、紙の書類をスキャナーで、イメージデータとして登録するのが一般的な手法である。この方法は、イメージデータの編集やデータベースへの登録ができないという問題が残っている。

また、読取技術の代表的な手法では、文字認識OCR技術を用いて、紙の書類を自動的に読み取する方法がある。OCR技術は文字パターンを認識し、文字コードを変換することができるが、ドキュメントの組み版情報や、伝票のデータベースフォーマットなど情報を得ないので、必然的に人間の介入しなければならない。

上述の問題を改善するために、紙の書類データを付加情報として地紋の形で紙上に埋め込み、書

類データと同時に印刷することができれば、大容量の情報を埋め込んでいるデジタル印刷物を記録媒体として、何時でもコンピュータに登録することができる。また紙の書類と情報が一体化され、複数の入力は要らないメリットと、登録前の人間の介入は不要、完全な自動読み取りが可能になる。その上、紙の書類を記録媒体として、大量な情報を埋め込むことができれば、紙媒体の利用効率はアップし、紙媒体を減らすことによって環境にやさしい効果も期待できる。

従来の紙媒体情報埋め込みコード

紙媒体の秘密文書の流通経路を追跡する場合は、数十バイトぐらい容量として埋め込むことができればいいので、従来、電子透かし技術がよく使われている。この技術は、信号処理理論に基づいて、複数の異なる搬送方向のシンプルを設けて、情報を記述する。情報記録のシンプルのドットパターンのマトリクス配置によって地紋を構成し、紙文書の背景として印刷する。情報記録のシンプルを隠すために、情報記録のシンプルの周りに大量の関係ないドットの配置の必要はある。これらのドットの配置によって、ノイズの原因になり、情報記録のシンプルにとって誤認識の問題を発生することが指摘されている。また、情報記録のシンプルを無差別で配置しているため、紙文書と重なった時に、情報記録のシンプルの認識ができない場合は多かった。そのために、同様の情報として、数回にわたり繰り返し埋め込む必要があり、埋め込み情報量は実用的に100バイトに留まる。

第1表

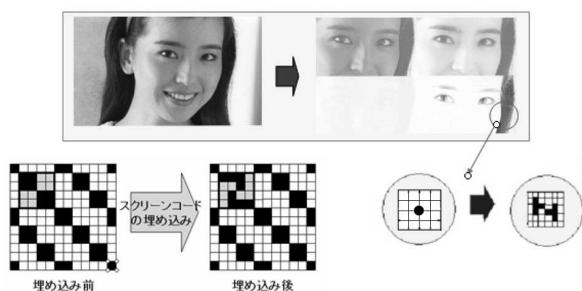
	複写禁止	追跡機能	改竄防止	自動読取
埋め込み容量(以上)	2Bit	30バイト	500バイト	3キロバイト

第1表は、複合機向け自動読取機能とセキュリティ機能が必要とする情報埋め込み容量の比較である。第1表によって、従来の技術では埋め込み量が少なくすむ秘密文書の複写禁止と漏洩経路の追跡には対応ができるが、大容量が必要となる改竄防止や自動読取などの機能には十分に対応できない。

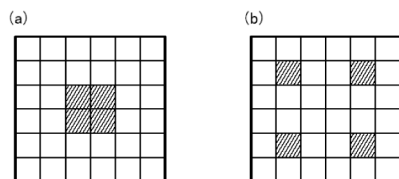
新しい紙媒体情報埋め込みコード

ここで、従来電子透かし技術と全く違う発想で、印刷のスクリーン理論を導入して、情報記録コードと情報埋め込みコードを兼ねてスクリーンコードという新しいコードを提案する。スクリーンコードは、印刷物にとって違和感が生じないように、印刷スクリーンの特徴を利用する。まず、第1図のように、印刷スクリーンの各々網点に対して、各々網点の形を変わって、大量情報記述することは可能である、また、形を変えた網点の階調値が変わってないので、画質を落さない特徴もある。

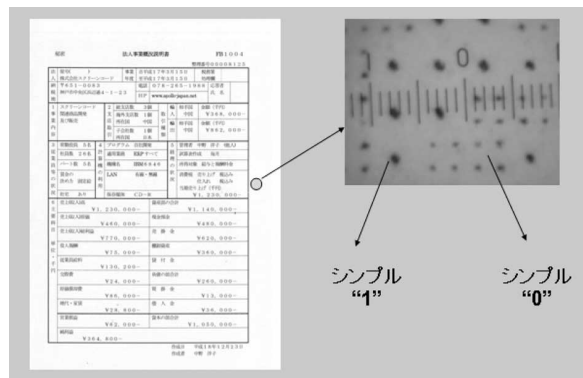
紙媒体の書類にとって、大量情報を埋め込む場合に、地紋方式が一番有効である。第2図の(a)は、網点の大きさによって階調を表わす振幅変調方式(Amplitude Modulation)即ちAMスクリーンをシ



第1図



第2図



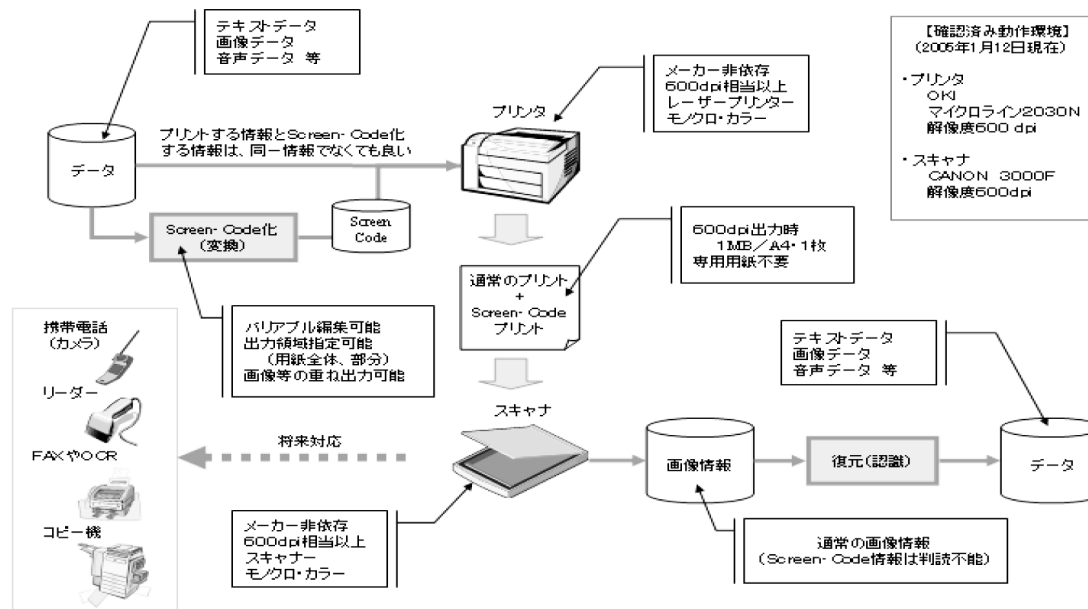
第3図

ンプル“1”とする例である、第4図の(b)は、網点のドットの数によって階調を表わす周波数変調方式(Frequency Modulation)即ちFMスクリーンをシンプル“0”とすることを考える。つまり、異なる変調方式によって情報を記述することができる。異なる変調方式の網点をマトリクスとして配置すれば、地紋を構成することが可能となる。また、異なる変調方式でも、各網点の階調値は変わらないので、一定階調の地紋になることが可能である。

第3図では、異なる変調方式によって紙媒体の書類上に情報を埋め込む例である。第3図のように、異なる変調方式によって構成された地紋を紙伝票の背景として、印刷した。拡大すると、振幅変調方式(AMスクリーン)の一つ大きなドットによって構成した網点をシンプル“1”とし、周波数変調方式(FMスクリーン)の四つ分散している小さいドットによって構成した網点をシンプル“0”とする。また、四つ分散している小さいドットによって構成した網点の大きさを足算すると、一つ大きな網点のドットの大きさと同様であるので、これらの異なる変調方式の網点を印刷スクリーンの特徴に基づいて、一定間隔で、配置すれば、均一な地紋を構成することができる。そのため紙伝票の違和感がない。また、地紋を構成されたすべての網点は情報記録のシンプルとして情報を埋め込むことができるので、数キロバイト以上の大量情報を埋め込むことが可能になった。

第4図では、スクリーンコードの処理の流れを示している。

データベースから文書や伝票など電子データを印刷するときに、付加情報として紙の書類データ



第4図

第2表

	記録容量	埋め込み	読取性能	ロバスト性	認識精度
スクリーンコード	20KB	可能	完全自動	高い(5)	高い
電子透かし	100B	可能	(3)	低い	普通
QRコード	60B	要専用場所	(4)	低い	高い
OCR文字認識	(1)	(2)	要人間介入	低い	普通

注：表の中で空欄(1)と(2)はOCR文字認識技術が情報記録コードではないので、比較できない。空欄(3)は電子透かし技術が埋め込み容量が少ないので、自動読取機能を用いる不可であるので比較できない。(4)はQRコードは同上の理由で比較できない。(5)1/3以内破損でも認識は可能である。

をスクリーンコードに変換し、地紋の形で埋め込めば紙文書を電子データとして情報を埋め込んでいる紙の書類が作成できる。上述印刷された紙の書類に対して、必要時に応じて、スキャナーや複合機など、読取機を用いてコンピュータに自動的に再登録することができる。

次は、自動読取システムに関して、従来の技術とスクリーンコード技術との比較は第2表になる。

● 複合機向け自動読取機能

コンピュータ入力会社からの調べによると、入力書類のうち、90%以上がプリンタで出力したものである。即ち現時点では一旦入力し、印刷した書類を、流通経路によって何回も入力しているということである。

そこで、印刷するデータを、紙上に埋め込み、印刷すれば、情報と書類が一体化され、複数入力は不要、必要時に応じて、何時でもコンピュータへ登録することができる。一方、出力とした印刷機能と入力としたスキャナー機能を備えている複合機に対して、情報埋め込み機能と情報読取機能と付加すれば、自動読取システムを構築することができる。また、自動読取システムの構築方法は、ソフトウェアのドライバ方式とICチップ方式の二種類ある。

例えば、保険会社は毎日大量な診断書のような押印の必要な紙の書類を人間で入力している。この膨大な仕事の量を減らすために、次はソフトウェアのドライバ方式を用いて、保険会社向け診断書自動読取システムの構築について、紹介する。

病院から診断書をプリンタで発行する時に、診断書のイメージデータをプリンタドライバから得る。診断内容を付加情報としてスクリーンコードに変換され、マトリクス配列によって構成した地紋を診断書の空白領域に配置してから、印刷書類と一緒に印刷し、第5図のように自動読取可能な診断書出力することができる。

第6図では、普通プリンタの選択画面を示している。第6図のように、普通のプリンタドライバを選べれば、普通の印刷になる。ScreenCode Priterドライバを選択すると、診断書内容を埋め

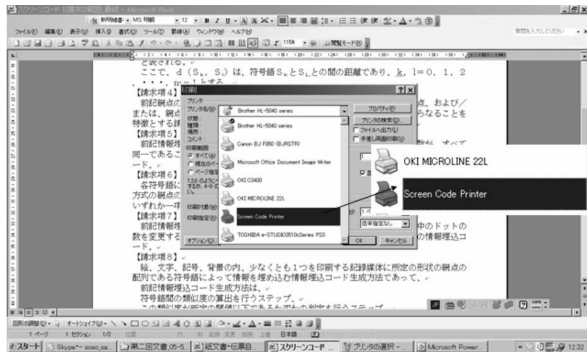
込んだ書類を印刷することができる。

埋め込み診断書データの選択方法は第7図のようになる。

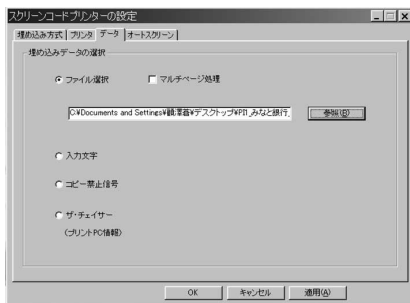
スクリーンコードプリンタの設定画面を用いて、ファイルの選択をチェックし、参照をクリックし、埋め込み診断書のファイルを選択することができる。最近マイクロソフト社から、Officeソフトのフォーマットを公開すると発表されたので、今後、この設定は不要で、WordやExcelなどのようなアプリケーションを使う場合に自動的に印刷のデータを埋め込むことが可能になると想定される。



第5図



第6図



第7図

次は実際印刷のプリンタの選択方法は第8図のように実際出力したいプリンタを選ぶことができる。

ネットプリンタの場合に上述の方法を用いて、設定できる。

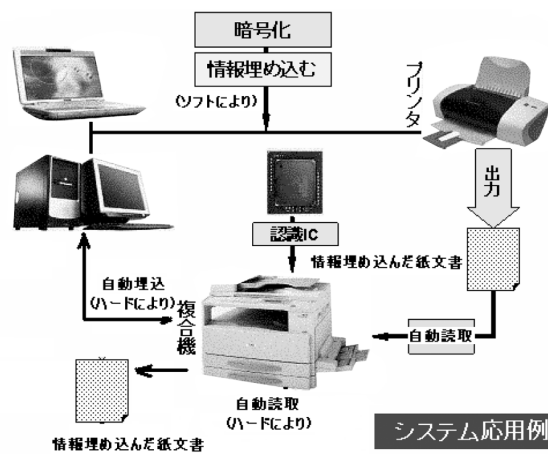
上述自動読取可能な診断書を読み取る場合に第9図のようなソフトを使えば、埋め込んだ診断書データをコンピュータに自動的に再登録することができる。ICチップの方式の場合に、スクリーンコードにより情報の埋め込み処理と認識処理を第10図のようにICチップ化する。



第8図



第9図



第10図

前記プリンタで印刷した自動読取可能な診断書のような紙の書類を、複合機のスキャナーを用いて、大量の書類にとって埋め込んだデータを連続且つ高速にコンピュータへ登録できる。また、複合機で印刷した紙の書類は、書類データを埋め込んでいるので、必要に応じて、何時でも、コンピュータに登録することができる。

OAシステムの革命

常識を超えた自動読取システムの創出は、OAシステムにとって革命的な進歩が可能となる。

まず、自動読取機能を付いた複合機を用いて、情報を埋め込んだ紙の書類をスキャナーで読み取る場合は、紙の書類の表面印刷内容だけではなく、埋め込んだデータを印刷することができる。そうすれば、紙の書類から電子化へ処理する時に、再編集できるようなデータ形式として、人手による介入は不要、完全自動化でコンピュータに登録することが容易になる。

A4サイズの紙一枚で、小説の一冊内容を埋め込むことが可能であるので、複合機の表示ディスプレイの画面は現在より大きくすれば、紙の書類に数百枚以上のデータを埋め込んでいるので、大画面のディスプレイを使って必要な書類を選んで、印刷することができる。また、紙の書類に埋め込んだ情報に対して、暗号キーによって、読む人、場所、時間などのコントロールすることも可能になる。

さらに、証明書を発行するときに、証明内容をすべて書類上に埋め込むことができるので、証明書の表面内容を改ざんしても、埋め込んだ内容を改ざんができないので、完全に改ざん不可能な証明書を発行することが可能である。

印刷された紙の書類では、何時、どのコンピュータ、どのプリンタや書類内容などを追跡情報として、埋め込むことができるので、紙の書類の漏洩管理が強化される。また、上述の追跡情報を利用して、紙書類のタイムスタンプなど認証も可能である。

自動読取システムは、紙媒体の直感性、安価性や立証性などの特徴と、電子媒体の詳細性、検索性や利便性など特徴を統合し、一番効率いい形で構成されている。また、紙の書類に大量な情報を

埋め込むことにより、紙の使用量を大幅に削減でき環境にやさしい技術といえる。

ドキュメントやビジネスフォームのアプリケーションシステムにとって、従来、書類の表面データしか作成できないが、これから、紙の書類に埋め込むデータの処理手順やデータフォーマットの統一などが今後の課題になるといえよう。

今後の展開

印刷スクリーンの特性に基づいて構成されたスクリーンコードは、プリンタの印字精度のアップによって、応用範囲は飛躍的に広がる。従来不可能と考えられてきたことやこれまで思いもなかった問題に対して、いろいろなアイデアや新しいソリューションを開発できると考えられる。印刷の基礎技術によって開発されたスクリーンコードは、紙の書類にとって、画期的なイノベーションを創出できると信じている。

筆者紹介

顧 澤 蒼

(株)アポロジャパン

代表取締役 工学博士

日本電子情報通信学会 (IEICE) と画像電子学会

(IIEEJ) との正員

株式会社アポロジャパン

代表者名 代表取締役 顧 澤 蒼

本社住所

〒650-0044 神戸市中央区東川崎町1-8-4

神戸市産業振興センター7F

URL : <http://www.apollo-japan.ne.jp>

E-mail : infp@apollo-japan.ne.jp

事業内容及び会社近況

弊社が独自開発したスクリーンコードは、日本で出願した二つ特許が既に権利化されており、日本某大手複合機メーカーによりセキュリティ分野での導入に対して日本トップレベルという技術評価も頂いている。

業界で初めての紙上に大量データを埋め込むことができる優位性を生かして、紙の書類から電子化へ進化する膨大なビジネス市場を狙って、自動読取機能付き複合機を開発し、日本大手複合機メーカーをターゲットとして、ライセンス販売事業を展開していきたい。