

認識通信における文字・画像入力システム

木田 博巳 荒川 弘熙 中島 信造 印牧 直文
(日本電信電話公社 横須賀電気通信研究所)

1. まえがき

情報化社会の進展にともない、計算機、通信回線網、および端末は多様化、高度化、複合化の傾向にある。高度化の一要素として、認識機能や知識処理機能の活用が人間-計算機向インタフェースを改善する上で重要であり、パターン認識技術は今後ますます期待される技術である。とくに文字認識技術は今や本格的な応用期を迎えており、図形認識と音声認識技術についても、認識対象を限定しつつ具体的な応用が進められている⁽¹⁾。

今後、認識法の進歩による認識対象の拡大と認識精度の向上、およびLSI技術の導入による装置の低価格化が図られ、応用分野の拡大、利用形態の多様化が可能となる。例えば、手書き文字を読む端末OCRの適用域拡大、マーク認識の公衆ファクシミリ通信網への応用⁽²⁾、ファクシミリと文字認識技術を結合し、オンラインまたはオフラインでファクシミリから計算機への文字入力を行うオフィスシステムの開発⁽³⁾などが進められている。さらに、文字に加え図形や画像など処理対象の拡大を可能とするために、イメージOCRを含む帳票処理システム⁽⁴⁾やファクシミリ入力による図面・原稿編集システム⁽⁵⁾などが開発されている。

こうした状況の下、筆者らは複合通信の一環として、認識機能を通信網内に分散配置した認識通信システムを提案し⁽⁶⁾、ファクシミリ入力による手書き文字認識技術⁽⁷⁾⁽⁸⁾を活用して文字と画像を含む帳票をファクシミリから入力し、文字認識結果で画像の登録、編集、通信制御を行う帳票処理方式について

検討を進めてきた。

本稿では、書式の異なる多様な帳票を画像として入力し効率的に処理するために、認識機能を活用した帳票処理システムについて検討する。多様な帳票の処理においては以下の課題がある。(1)帳票の印刷や管理が煩雑である。(特に文字認識を行う場合、文字枠をドロップアウト・カラーで印刷する必要があるのである。)

- (2)帳票対応に処理プログラムを作成することは効率が悪く、効率的な帳票フォーマット定義法が望まれる。
- (3)帳票を画像情報として蓄積する上で、帳票の属性に応じ領域を切り出し、圧縮する技術が望まれる。

こうした課題の解決をめざし、ここでは、簡易な図形認識技術を用い帳票フォーマットを作成する方式、帳票のイメージで処理装置側に蓄積しておき必要に応じてイメージ記録部へ出力し使用する帳票管理方式、帳票フォーマット作成時に帳票処理プログラム用の帳票フォーマット定義テーブルを自動生成する方式、および帳票処理プログラムにより、帳票中の文字領域、画像領域など領域の属性に応じて認識や圧縮を行う方式を結合した帳票処理の形態を提案し、実験的検討を加える。

2. 認識通信システムのご概念

認識通信システムでは認識機能を通信網内に負荷分散または機能分散し、認識装置コストおよび伝送コストに応じて経済的に認識機能を提供することをめざす。例えば、図1に示すように、リモート側に認識部を負荷分散したり、英数字・カタカナ程度の比較的カテゴ

リ数が少なく、かつ使用頻度の高い文字セットはリモート側認識部で、また漢字のようにカテゴリ数が多く認識部が高価になるものはセンタ側認識部といった機能分散を図ることが考えられる。また知識データベースの所在に応じ、センタ側とリモート側で通信しアプリケーションに依存した知識処理（プラグマティズム）を認識、後処理として行うことも考えられる。

帳票入力および出力用の端末としては、電話ファクシミリなど汎用ファクシミリを応用することが可能であり、また構内系では、構内回線帯統により専用イメージスキャナを収容すること一般化するものと思われる。

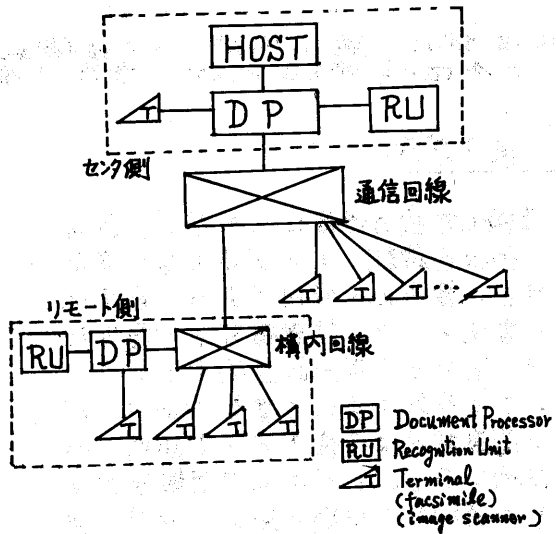


図1. 認識通信システム

3. 認識機能を活用した帳票処理

認識機能を活用し、汎用ファクシミリや専用イメージスキャナ等の画像端末から文字・画像を含む帳票を入力し処理する帳票処理方式を提案する。本方式による帳票処理の基本手順を図2に示し、以下その概要を説明する。

(1) 帳票フォーマットの作成・登録

帳票フォーマットの作成・登録を、ファクシミリ等の画像入力端末、ディスプレイ、ワードプロセッサを用いて行う。まず、初期データとして帳票イメージを画像入力端末から入力し、簡易な図形認識により帳票の枠組みに関する情報を線分リストの形で表現し、これを用いて帳票領域の定義を行う。さらに、ワードプロセッサ、グラフィックディスプレイを用いた会話処理により、帳票上の各領域に対して属性を指定し、帳票フォーマットの詳細な定義を行う。定義された帳票フォーマットには、帳票番号が付与され、フォーマット定義テーブルとして登録される。また帳票イメージも登録して、以後の帳票処理に用いる。

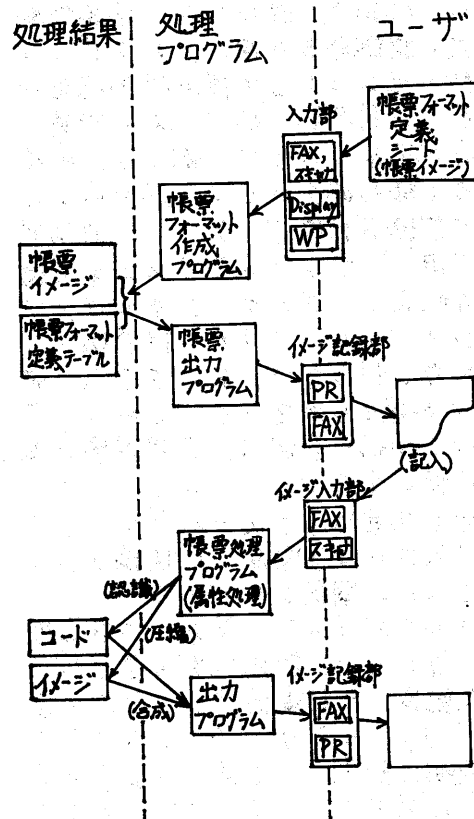


図2. 認識機能を活用した帳票処理

(2) 帳票の出力

センタ側またはリモート側の要求に応じて、帳票イメージをファクシミリ等に画像出力する。出力帳票には、切出し原点を示す基準マークおよび帳票番号を付加する。

(3) 帳票への記入と入力

イメージ記録部に画像出力された帳票あるいはそれを複製したものに所定の事項を記入した後、帳票をファクシミリ等から入力する。

(4) 属性処理

入力帳票上には、文字、記号、画像等の種々の属性をもった情報が混在して含まれる。これらの各属性に応じて、文字、記号等に対しては認識処理によりコード化し、画像情報に対してはイメージとして切出し圧縮処理して蓄積する。各属性の帳票フォーマット作成時に生成したフォーマット定義テーブルに記述されており、この情報に従って属性処理を実行する。

(5) イメージの編集・出力

帳票入力の結果得られた文字コードや画像領域のイメージをアプリケーションプログラムで処理し、処理結果を編集合成しイメージ記録部へ出力する。例えば、画像情報の検索への応用では、コード化した文字データや記号データをキー情報として画像ファイルを検索し、検索画像を原帳票イメージの中に埋め込んで出力することが考えられる。

4. 帳票フォーマットの作成・登録

簡易な図形認識機能を活用した帳票フォーマットの領域の定義とグラフィックディスプレイを用いた台帳形処理による属性の定義の2段階から成る帳票フォーマットの作成法を以下詳説する。

4.1 領域の定義

初期データとして画像入力端末から入力する帳票イメージの例を図3に示す。一般の帳票は、図3に示すように枠付きの帳票形式を持つものが多く、また、枠によって囲まれた領域に1つの属性が対応することが多い。そこで、ここでは、領域を枠情報によって指定することと、入力帳票イメージから枠情報を認識して線分リストの形で表現し、これをもとに領域定義を行う手法を検討した。線分リストの作成には、Hough変換による直線検出の手法を用いた。

(1) Hough変換による直線の検出

ここでは、Duda & Hartによって示された直線検出手法⁽⁹⁾を基本に検討を進める。すなわち、(1)式に示すイメージ空間(x, y)とパラメータ空間(θ, ρ)の変換式を用いる。

$$x \cos \theta + y \sin \theta = \rho \quad (1)$$

θ, ρを固定した場合、式(1)は(x, y)平面上で直線を表現し、ρは原点から直線までの距離を、θは直線の法線方

特 実	公 開	昭	番号	出 願 日	公 開 日	公 告 日	審 査 請 求														
							有 無														
発明の名称																					
出願人																					
概要																					
担当者																					
対 象		特 色	部 門 1	部 門 2	部 門 3	部 門 4	部 門 5														
1	2	3	4	5	6	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4

図3. 帳票イメージの例

向を示す。いまイメージ空間内の n 個の黒点 $\{(x, y), \dots, (x_n, y_n)\}$ が同一直線上の点であれば、各点を式(1)で変換したパラメータ空間内の n 個のカーブは共通の点で交わる。この性質を利用してパラメータ空間内の各点にアキュムレータ $Acc(\theta, \rho)$ を対応させ

$$Acc(\theta, \rho) \geq T \quad (2)$$

(T : 閾値)

を満足する点 (θ, ρ) を抽出することにより直線検出を行うことができる。図3の例のように長方形で構成される帳票イメージの場合、 θ に関し2つのクラスタが出現し、その位相差は $\pi/2$ とがる。帳票入力時の傾きが少ない場合には $\theta = 0$ と $\theta = \pi/2$ 付近に2つのクラスタが出現する。

(2) 線分リストの作成

Hough変換により検出した直線と、帳票イメージとの一致をとることにより、表1の形式で線分リストを作成する。

表1. 線分リスト

	水平線分			垂直線分		
	y 座標	始点	終点	x 座標	始点	終点
H_1	y_1	x_1^s	x_1^t	V_1	x_1	y_1^s y_1^t
H_2	y_2	x_2^s	x_2^t	V_2	x_2	y_2^s y_2^t
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
H_n	y_n	x_n^s	x_n^t	V_m	x_m	y_m^s y_m^t

但し $y_1 \leq y_2 \leq \dots \leq y_n$, $x_1 \leq x_2 \leq \dots \leq x_m$

また、この時 θ の2つのクラスタをそれぞれ水平線、垂直線に対応づけることにより、帳票入力時の傾きを補正する。

(3) 領域分割

線分リスト内の線分 H_i, V_j, H_k, V_l ($i < k, j < l$) が次の2条件を同時に満足するとき、線分 H_i, V_j, H_k, V_l によって構成される長方形領域を帳票上の1つの領域とみ出す。

① 条件1 (長方形条件)

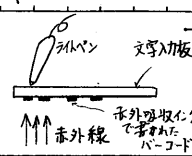
H_i, V_j, H_k, V_l が閉領域を構成する。すなわち、次の(i)~(v)を同時に満足する。

- (i) $x_j \neq x_l$ and $y_i \neq y_k$
- (ii) $x_i^s \leq x_j$ and $x_i^t \geq x_l$
- (iii) $x_k^s \leq x_j$ and $x_k^t \geq x_l$
- (iv) $y_j^s \leq y_i$ and $y_j^t \geq y_k$
- (v) $y_l^s \leq y_i$ and $y_l^t \geq y_k$

② 条件2 (非包含条件)

H_i, V_j, H_k, V_l によって構成される閉領域内に他の閉領域を含まない。すなわち、以下の条件で非包含が定義できる。

- (i) for all k' ($i < k' < k$)
 $x_{k'}^s \geq x_l$ or $x_{k'}^t \leq x_j$
- (ii) for all l' ($j < l' < l$)
 $y_{l'}^s \geq y_k$ or $y_{l'}^t \leq y_i$

③	特公	番号	出願日	公開日	公告日	審査請求		
	実	昭55	925753	7355	123	⑦ 無		
	発明の名称 光学的文字入力装置							
	出願人 電電公社		テンテェンコウシャ					
	概要 文字入力板裏面に赤外線吸収インクでバーコードを描き、表面に通常のインクで母型文字を描く。裏面より赤外線を照射してライトペンにて赤外線を検出し、バーコードによる文字入力を行う。							
								
⑤	担当	対象	特色	部門 1	部門 2	部門 3	部門 4	部門 5
				1 2 3 4 5 6	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
			✓	✓				✓

属性	
①	文字領域(数字)
②	文字領域(カナ)
③	マーク領域(0印)
④	マーク領域(2印)
⑤	スタンプ領域
⑥	イメージ領域
⑦	フル印刷領域

図4. 帳票記入例と属性

4.2 属性の定義

図4は図3に示す帳票内に必要事項を記入した帳票を例示するものである。このように、帳票内の各領域には、文字データ、記号データ、画像データなどが記入される。長方形の領域として定義された各領域に対する属性の定義は、グラフィックディスプレイを用いた会話形処理により実行する。まず表1に示した線分リストの情報から帳票の枠組みをグラフィックディスプレイ上に表示する。次に分割された長方形領域を表示し、この領域に対する属性を属性指定コマンドによりキーボードから入力する。また、帳票上にアレ印刷しておくべき文字情報はワードアセッサを用いて入力する。

4.3 帳票フォーマットの登録

属性の定義が終了した段階で帳票番号を割付け、帳票フォーマット定義テーブルとして登録する。帳票フォーマット定義テーブルの内容は、表2に示すように帳票上の各領域に対する位置情報およびその属性である。

表2. 帳票フォーマット定義テーブル

領域	位置情報	属性
D_1	$x_1^L \leq x \leq x_1^R, y_1^L \leq y \leq y_1^B$	M_1
D_2	$x_2^L \leq x \leq x_2^R, y_2^L \leq y \leq y_2^B$	M_2
⋮	⋮	⋮
D_n	$x_n^L \leq x \leq x_n^R, y_n^L \leq y \leq y_n^B$	M_n

5. 属性処理

フォーマット定義テーブルに記述された情報をもとに、領域単位の切出しを行い、文字領域や記号領域のデータに対して認識処理を施し、画像領域のデータに対しては情報圧縮を行う属性処理について以下に詳述する。

5.1 領域の切出し

領域の切出しには前述の Hough 変換を用いることも考えられるが、ここでは処理を安定にかつ簡略に行うために先見的に与えられた情報を基本に領域の切出しを行った。すなわち、ここで対象としている帳票は縦横のけい線からなる長方形をその構成要素としており、これらの長方形を抽出することにより領域の切出しが可能である。

これら長方形はあらかじめフォーマット定義テーブル上に線分の始点および終点座標からなる領域位置情報として与えられており、これらの先見的情報から設定される切出し対象の長方形領域と、入力帳票の黒連結成分あるいは白黒の変化点として抽出される長方形領域との整合をとりながら領域の抽出を行う。

5.2 認識処理

(1) 文字の認識

ファクシミリ出力した帳票を使用するターンアラウンド形帳票処理では、文字記入枠にカラーが使用できない。又用紙雑音が問題となる。これに対処する文字切出し法と文字認識法について述べる。

① 文字の切出し

現状の文字認識技術では、ピッチ不定の文字を切出して認識することは困難である。そこで、ドロップアウトカラーの文字記入枠に代えて、文字が一定のピッチで記入されるよう、文字領域の枠に対しては図3に示すようなガイドマークを帳票設計時に付与することとした。また、このガイドマークを枠情報検出時に同時に検出することにより、文字切出し時の文字ピッチ情報として利用することが可能である。文字切出し法としては、従来より文字線の水平軸への射影をとる方法がよく用いられているが、ここでは射影情報とピッチ情報を併用することとする。図5に例を示し、以下処理手順を説明する。

- (i) ピッチ情報が指示する切出し位置Pに文字線の射影がない場合、Pを切出し位置とする。
- (ii) ピッチ情報が指示する切出し位置Pに文字線の射影が存在する場合には、Pの近傍で文字線の射影がない位置を文字切出し位置とする。
- (iii) (ii)で文字切出し位置の決定ができない場合には、Pを切出し位置とする。

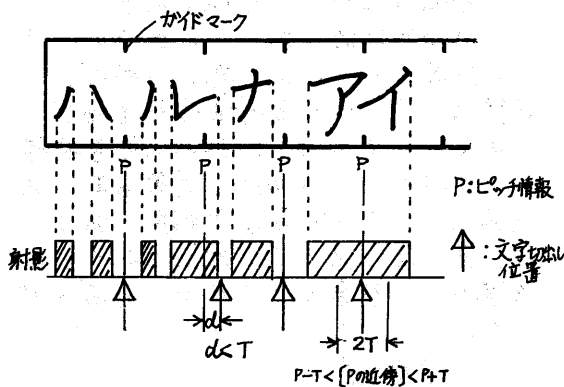


図5. 文字の切出し

②文字認識法

ファクシミリ出力画を(又はそれを複製したものを)、入力帳票として使用するターナランド形帳票処理では用紙雑音が多くなり、文字認識の立場から対策が必要である。ここではファクシミリ入力文字用に設計した文字認識論理を基本に、用紙雑音に起因する不安定な特徴を抑制する識別処理、および識別結果により前処理部に帰還する処理を具備した文字認識系を構成することにより対応した。⁽²⁾

(2) マークの認識

マーク認識の場合、マークの形状よりはむしろその有無が重要となることが多い。従って、文字認識で使用するような複雑なアルゴリズムではなく、連結黒点数や縦横比といった単純な特徴パラメータを用いてマークを記述し識別する方法をとり、処理の簡易化を図った。識別対象には○印と▽印を想定した。

(3) スタンプ・印鑑の認識

日付印等のスタンプあるいは印鑑の認識は基準パターンとの比較照合により行う。この場合、基準パターンと入力パターンの位置合わせの精度が認識精度を決定する主要因となる。

5.3 画像情報の圧縮・蓄積

画像領域に記述された情報に対しては認識処理を施さず、イメージ情報として蓄積する。イメージ情報はその情報量が膨大であるため、冗長度抑圧符号化による情報圧縮を行って記憶する。符号化法としては、ファクシミリの情報圧縮方式として国際標準化されたRLEMD方式を採用した。RLEMD方式は、符号化による画像劣化を伴わない可逆形の符号化方式であるが、画像情報の内容によっては、若干の画像劣化を許容して圧縮効率の向上をねらいとする場合もあり得る。このような情報に対して、属性指定によって非可逆形の圧縮方式を選択する方式も考えられる。

5.4 属性処理による情報圧縮効果

入力帳票上から、コード化できる情報は認識処理によりコード化し、また画像データに関して必要が部分だけを切出して蓄積する属性処理は、情報圧縮の観点から極めて有効な手法と考えられる。ここでは図4の帳票を例にとって属性処理による情報圧縮効果を検討する。図4の帳票の大きさはA5判であり、これを8本/mmの解像度で2値画像とした場合、約250KByteの情報となる。またこの情報はREAD符号化によって14KByte程度の情報に圧縮できる。一方、図4の帳票に属性処理を施した場合、コードデータとして約100Byteイメージデータとして約3KByteの情報に変換でき、原画像情報を約1/80にまで情報圧縮できる。また、圧縮符号化後のデータと比較しても、4~5倍の圧縮性能を有することとなる。

6. 実験システム

6.1 構成

ミニコンピュータベースの帳票処理制御装置と、それに接続されるファクシミリ、午書きカタカナ・英数字用文字認識装置、簡易グラフィックディスプレイ、ディスクパック等から実験システムを構成し、認識機能を活用した帳票処理の検討を行った。実験システムの構成を図6に示す。帳票処理制御装置はチャンネル結合によりホスト計算機と接続されており、午書きカタカナ・英数字以外の認識処理や単語単位の認識処理など各種シミュレーションをホスト計算機で実行できる構成となっている。

(1) 帳票処理制御装置

ファクシミリの入出力制御、帳票フォーマットの作成、領域切出しを含め帳票処理プロセス全体の管理を行う。

(2) ファクシミリ

帳票の入力および出力用の端末として汎用電話ファクシミリ(G3機)を用いる。

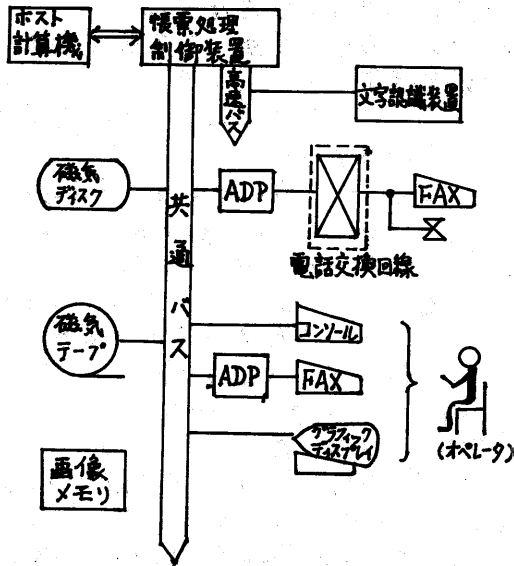


図6 帳票処理実験システムの構成

用いる。帳票処理制御装置には直接あるいは交換回線経由で接続される。

(3) 簡易グラフィックディスプレイ

8色表示、1024×1024のアドレスポイントをもつ。帳票フォーマット作成時、画像ファイルのモニタ表示時に使用する。また認識装置でリジェクトした文字パターンを表示しリジェクト処理を行う上でも使用する。

(4) 文字認識装置

帳票処理制御装置で切出した文字領域に対し、1文字単位または1文字行単位に認識処理を行う。読取り対象は午書きカタカナ・英数字である。

今後、漢字認識機能の導入、高精度ファクシミリの接続、日本語ワードアクセッサの統合を図る計画である。

6.2 処理手順

基本的な帳票処理手順を以下に示す。

(1) 帳票作成

ファクシミリから帳票フォーマットを入力し、簡易グラフィックディスプレイを用い会話形式で帳票フォーマットの作成を行う。生成された帳票のイメージと定義テーブルはディスクパックに格納される。

(2) 帳票出力

コンソールから帳票番号を指定し帳票をファクシミリに出力する。またはファクシミリから帳票要求シートで出力ファクシミリの番号と帳票番号を指定し、これを認識して出力制御を行う。

(3) 帳票入力

ファクシミリに出力された帳票、または印刷された帳票に必要事項を午書きし、ファクシミリから直接入力する。処理制御装置側では、まず入力された帳票上の帳票番号を認識する。次に、帳票番号に対応した帳票フォーマット定義テーブルを参照し領域の切出しを行い、領域の属性により文字認識処理、画像圧縮処理等を実行する。

(4) リジェクト処理

文字認識装置でリジェクトが発生した場合、リジェクトされた文字ボタンを簡易グラフィック・ディスプレイに表示し、オペレータがキーボードから代替入力する。

7. おまじ

帳票処理において人間-計算機間インタフェースを改善する一案として、認識機能を活用し、汎用ファクシミリや専用イメージスキャナ等の画像端末から文字・画像を含む帳票を直接入力する方式を提案し検討を加えた。

本方式の特徴を以下にまとめる。

- (1) 簡易な図形認識機能を活用し帳票フォーマットの作成を行い、帳票フォーマット定義テーブルを生成する。
- (2) 帳票イメージを登録しておき、必要に応じて帳票を画像記録部へ出力し、これを用いたターンアラウンド形の帳票入力ができる。
- (3) イメージ入力された帳票に対し、帳票フォーマット定義テーブルに従って領域抽出を行う。文字領域に対しては文字認識し、画像領域に対しては画像圧縮することにより、文字・画像を含む帳票の処理に役立てる。ファクシミリ、帳票処理制御装置、英数字・カタカナレベルの文字認識装置、簡易グラフィックディスプレイ等を結合した実験システムを構成し、本方式による帳票入力の統合動作を確認した。

今後、漢字認識を含む認識機能の拡充、ワードプロセッサとの結合等を図り、構内通信システム等において本方式を適用した帳票処理を実現する計画である。

謝辞

日頃御指導頂く、複合通信研究部の飯村部長、回陽統括役、川野辺分散処理プログラム研究室長に感謝いたします。又、研究の遂行にあたり、御討論御協力頂いた認識グループの諸氏に御謝いたします。

参考文献

- (1) 森俊二, "パターン認識技術の実用化", 情報処理, Vol.22, No.4, pp.259-266, 1981.4, および同号小特集 "パターン認識技術の応用"
- (2) 小川, 小林, 織田, "センタ・エンド形ファクシミリ通信におけるマーク認識のシミュレーションの検討", 昭56信学全大1238
- (3) 荒川, 中島, "ファクシミリ入力文字の認識", 情報処理, Vol.22, No.4, pp.280-285, 1981.4
- (4) 森下, 宮井, 首藤, 小野田, "イメージを含むドキュメント処理システム", 昭55画像電子学会全大12
- (5) 末永, 名倉, 寺西, "文章と図面の編集機能をもつインテリジェントファクシミリシステム", 昭56画像電子学会全大13
- (6) Kaiyo, H., Komori, K. and Nakajima, K. "A Character Recognition Communication System", Proc. of ICCV 78, pp.565-569.
- (7) 中島, 木田, 荒川, "ファクシミリ入力による手書き片仮名の認識", 信学論文誌, Vol. J64-D, No.6, pp.510-517, 1981.6
- (8) 中島, 木田, 荒川, "ファクシミリ入力文字の認識", 沖11回画像工学会誌:72-2
- (9) Duda, R.O. and Hart, P.E., "Use of the Hough Transformation To Detect Lines and Curves in Pictures", Communications of the ACM, Vol.15, No.1, 1972.1