

ファクシミリとOCRシステム

Facsimile and OCR System

近年、情報処理システムの一手段として、ファクシミリの出力情報(受信画)をコンピュータに入力することが強く要望されている。この背景のもとに、日立製作所は高速ファクシミリとOCRを利用したファクシミリ及びOCRシステムを、情報処理作業の効率化を目的として開発した。従来のOCRシステムでは、OCRが設置されている場所まで入力伝票を郵送するのが一般的であった。このシステムは、入力伝票の送付手段にファクシミリを使用し、ファクシミリ受信画をOCRで読み取り、コンピュータへ入力する情報処理システムである。

この論文は、ファクシミリとOCRシステムの背景及び位置づけ、並びに両者を結合するための実現手段、システムの特長、運用例などについて述べる。

山尾 強* Tsuyoshi Yamao
内山 元** Hajime Uchiyama
長田耕一*** Kōichi Osada
山内 清**** Kiyoshi Yamauchi

1 緒 言

画像情報の処理に関し、通信分野ではファクシミリがあり、コンピュータ分野ではOCR(Optical Character Reader: 光学文字読取り装置)による方式がある。ファクシミリは書かれているものをそのまま伝送するが、OCRは書かれているものをコード情報としてコンピュータに入力し、処理を行なって出力する機能がある。ファクシミリで送・受信されるデータには各種伝票類など、コンピュータ処理を伴うデータが多くある。これらのデータは、キーパンチなどで入力データを作成し、コンピュータ処理されているのが通常である。これらをOCRに直接読み取らせることができれば、ファクシミ

りとOCRの結合による事務の流れの円滑化を図ることができ、オフィス合理化の推進に大きく寄与する。

2 コンピュータとファクシミリの接続

2.1 背 景

ここ数年来ファクシミリの伸びは著しく、あらゆる企業でファクシミリの導入が急速に進んできた。その理由は、オフィスワークの効率化をねらったものである。このような背景のもとに、ファクシミリの普及につれて、ファクシミリとコンピュータとを結合した有機的なシステム実現の要求が高ま

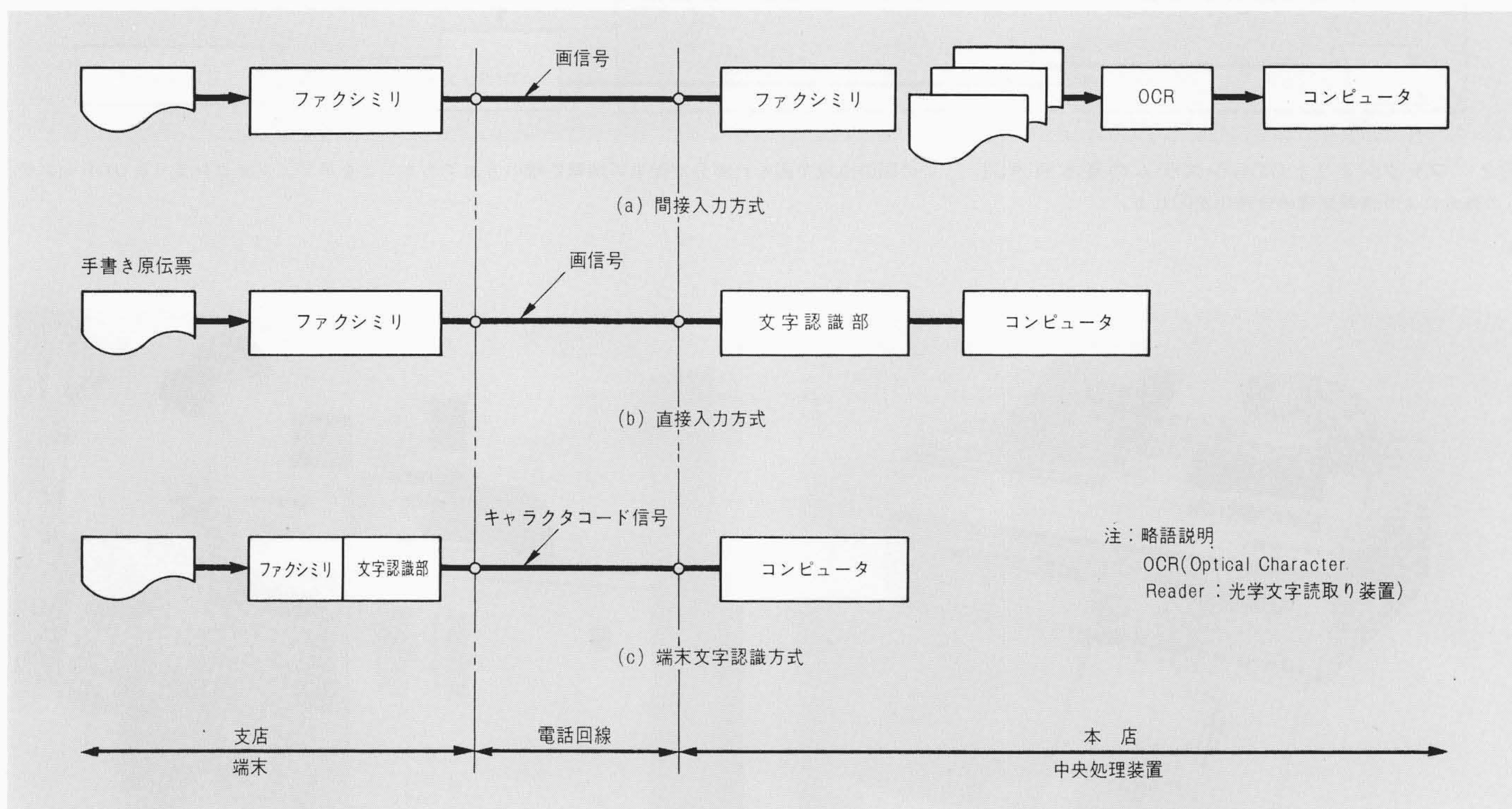


図1 ファクシミリからのデータをコンピュータへ入力する方法 図中(a)は、ファクシミリの情報を直接OCRに入力する方式である。(b)は本論で述べる間接入力方式であり、(c)はOCRそのものの機能を示す方式である。

* 日立製作所戸塚工場 ** 日立製作所小田原工場 *** 日立製作所通信機事業部 **** 日立製作所コンピュータ事業本部

ってきつつある。その背景には、

- (1) ファクシミリが手書きの画像を扱っていること。
- (2) ファクシミリのオペレーションが容易であること。
- (3) ファクシミリ端末の価格は一般的にコンピュータ端末よりも安価であり、一方、国際標準規格化により、各社相互通信が可能となり、今後よりいっそう普及するであろうこと。などの理由が挙げられる。

2.2 接 続

次にファクシミリとコンピュータの接続を分類すると、下記の三通りになる。

- (1) ファクシミリをコンピュータの入力機械として利用すること。
- (2) ファクシミリをコンピュータの出力機械として利用すること(例：本号小特集論文「ファクシミリ応答装置」参照)。
- (3) コンピュータをファクシミリの交換機として利用すること(例：本誌第60巻、第10号掲載「日立パケット交換システム“HIPA-NET”用ファクシミリの端末」参照)。

2.3 コンピュータへの入力

ファクシミリからのデータをコンピュータに入力する方法を図1に示す。同図に示すとおり(a)方式は、ファクシミリの受信紙を媒体としてコンピュータとの接続を行なう方式であり、最も変更の少ない方式である。(b)方式は、電気信号によりコンピュータとの接続を行なうもので、文字認識部が特殊仕様となる。また(c)方式では、遠隔OCRの思想となり、現状

では端末が高価となる。

これらの点から、(a)方式を現実的な解として取り上げた。(a)方式ではこのほかに、文書伝送を併せて行なうことができるため、多目的にファクシミリを使用することができる。

本論文では上述の受信紙を媒体としたファクシミリとOCRシステムについての仕様、運用例などについて概要を述べる。

3 ファクシミリとOCRシステムの概要

図2にファクシミリとOCRシステムの基本原理図を示す。従来、OCR入力帳票は郵送などの手段により、OCRのある場所まで運び処理していたが、ファクシミリとOCRの結合により郵送などの手段をなくし、直接OCRへ入力でき、効率的な処理が行なえる。

これらの製品の外観を図3、4に、装置仕様を表1、2に、OCRのシステム機能構成を図5にそれぞれ示す。

4 ファクシミリとOCRシステムの開発

ファクシミリとOCRは文字を読み取るという点では類似しているが、ファクシミリは出力帳票を人間が見て読み取るため、精度の粗さはある程度許容できるという条件で設計されているのに対し、OCRは機械が読み取るため、帳票精度が要求され、両者をそのまま組み合わせ、ファクシミリで出力された帳票を即OCRで読み取らせることが困難であった。

これらの問題を解決する手段として、

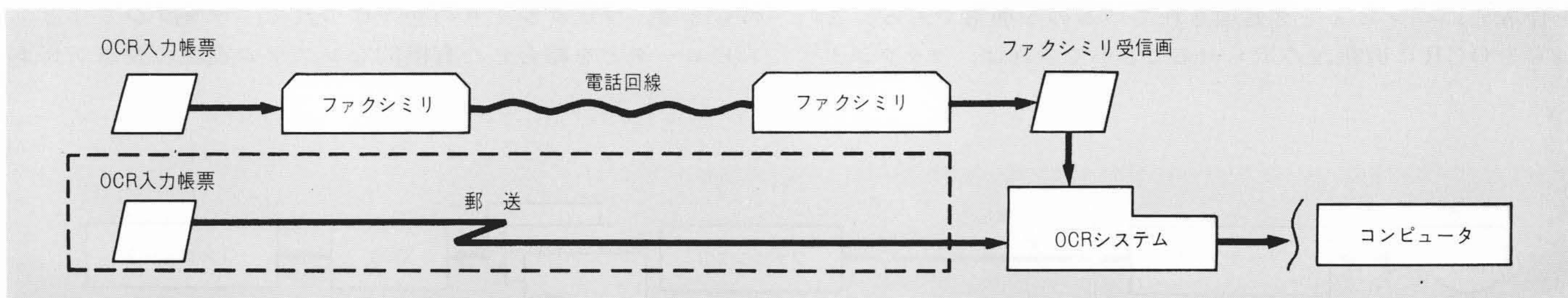


図2 ファクシミリとOCRシステムの基本原理図 この図の点線で囲んだ部分が従来の情報処理の方法であることを示し、ファクシミリとOCRシステムの結合により情報処理の合理化が図れる。

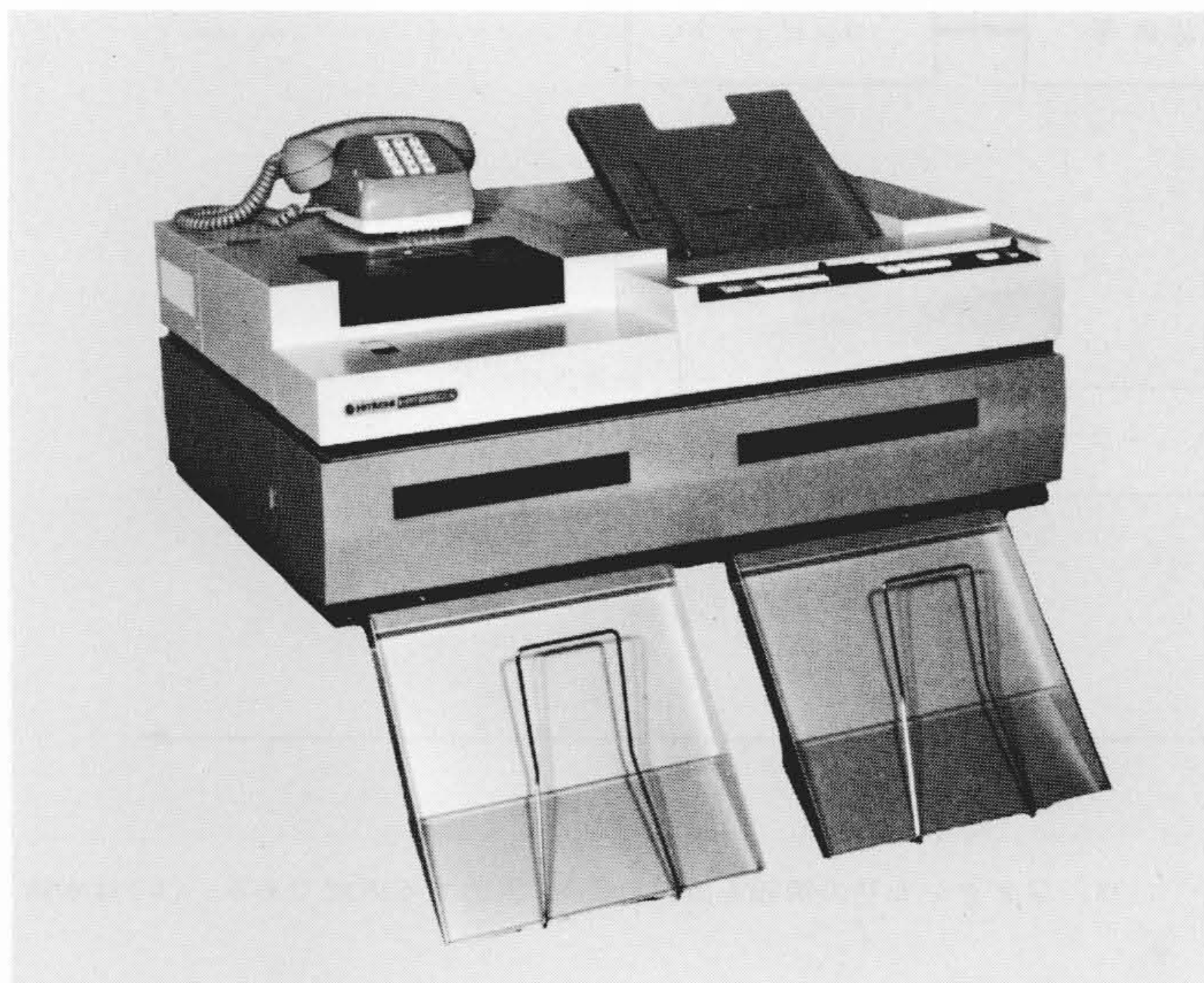


図3 HF550形ファクシミリ装置の外観 右半分が自動給紙装置を含む原稿送信部で、左半分が自動カッターを含む受信記録部である。



図4 HITAC T-550/30分散形OCRシステムの外観 左側半分が紙送り、スタッカを含むOCR部であり、手で持っている箇所がフロッピーディスクである。また、右側はディスプレイと操作部である。

- (1) OCR側では、
ファクシミリを介することによって生ずる原帳票の精度の低下(文字の位置ずれ、ひずみなど)を補正する機構の開発。
- (2) ファクシミリ側では、
OCRに入力するための強度をもつファクシミリ受信紙(静電記録紙)の開発、受信カッタの高精度化、送信の際の帳票の位置ずれ(スキュー)低減及び文字の欠けを防止するための読取りレベルの改善を行なった。

表1 OCR用ファクシミリ装置仕様 この仕様は標準ファクシミリを基本としており、異なるところは表中に示す(注1)～(注6)の点であり、標準ファクシミリとしても使用できる。

機 種 項 目		A 4 機		B 4 機	
		HF950X-A, OCR	HF550X-A, OCR	HF960X-A, OCR	HF560X-A, OCR
送 信 走 査 方 式		イ メ ー ジ セ ン サ に よ る 固 体 電 子 平 面 走 査			
原 稿 サ イ ズ		幅 120～257mm×長さ 210mm(注1)		幅 120～297mm×長さ 210mm(注2)	
有 効 画 面 幅		最小 201mm		標準 最小 240mm ワイド時 278mm(注3)	
電 送 時 間		約 20 秒	約 40 秒	約 20 秒	約 40 秒
走 査 線 密 度		主走査 8 本/mm, 副走査7.7, 3.85, 2.57本/mm(注4)			
伝 送 速 度		9,600/7,200/ 4,800/2,400 bps	4,800/2,400 bps	9,600/7,200/ 4,800/2,400 bps	4,800/2,400 bps
帯 域 圧 縮 方 式		2 次元逐次符号化方式 (リードコーディング) 1 次元符号化方式 (モディファイドハフマン方式, 日立方式)			
記 録 方 式		乾 式 静 電 記 録			
記 録 紙 サ イ ズ		216mm×65m のロール紙(注5)		257mm×65m のロール紙(注6)	
電 源		AC 100V±10% 50/60Hz			
消 費 電 力		待機時35VA, 送信時350VA, 受信時550VA			
外 形 寸 法		幅780×奥行590×高さ315(mm)		幅950×奥行550×高さ900(mm)	
重 量		7 5 kg		1 1 0 kg	
温 度	OCR 専 用(注7)	動作時: 10～35℃		非動作時: 0～50℃	
	一 般(注7)	動作時: 5～35℃		非動作時: 0～50℃	
湿 度	OCR 専 用(注7)	動作時: 35～70%		非動作時: 10～85%	
	一 般(注7)	動作時: 35～85%		非動作時: 10～85%	

備考: 注説明

注1～6 OCRの入力帳票以外の使い方(一般)の場合、仕様は下記のとおりである。

注1 幅120～257mm, 長さ80～1,500mm

注2 幅120～297mm, 長さ80～1,500mm

注3 B4機にはワイド機能があるが、この場合、画像が縮小されるので

OCR入力帳票の伝送には使えない。

注4 OCR入力帳票の伝送は、副走査7.7本/mmに限る。

注5 216mm×100m(用紙厚50kg相当)のロール紙でも使用可。

注6 257mm×100m(用紙厚50kg相当)のロール紙でも使用可。

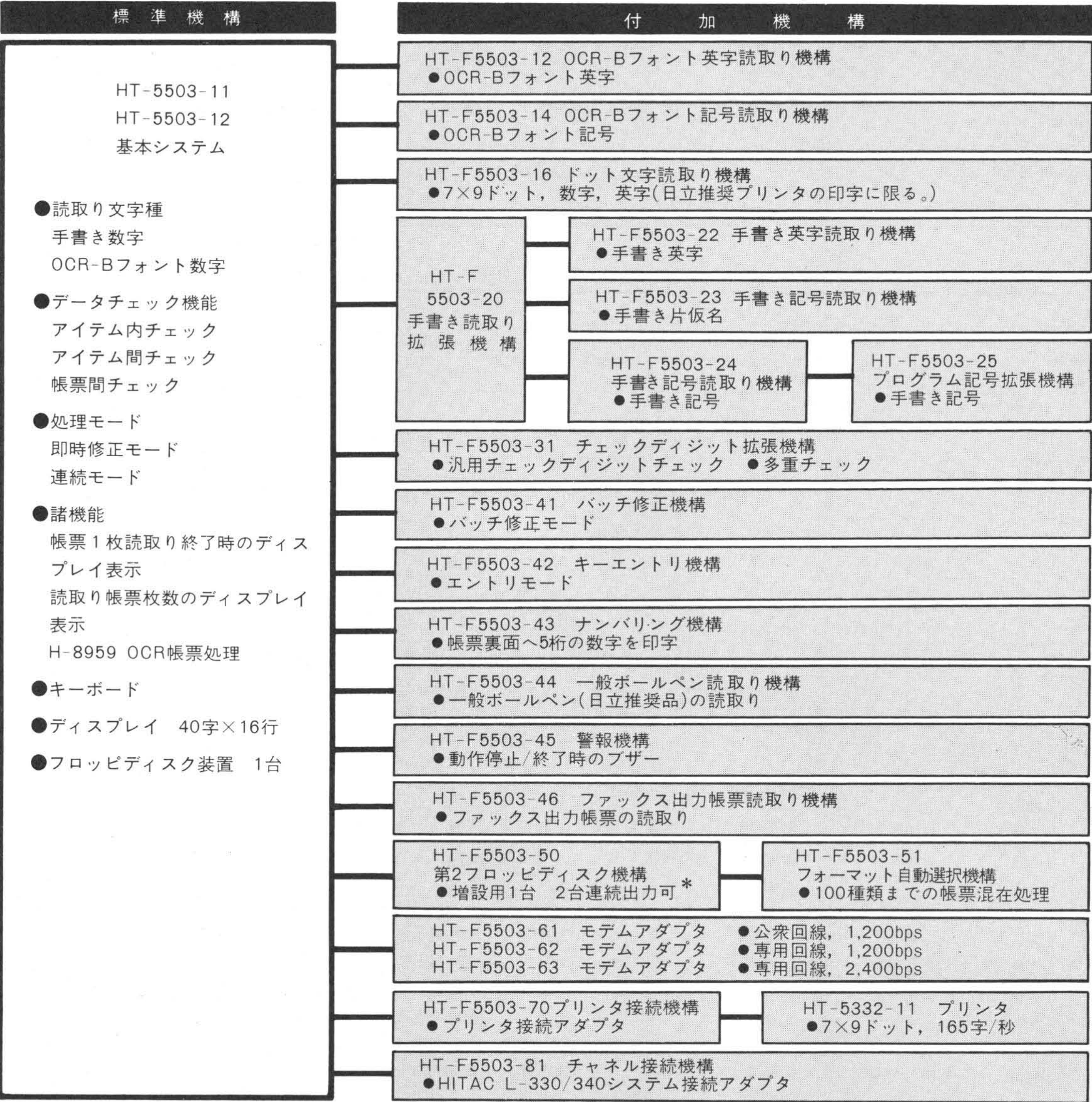
注7 OCR専用とは、OCR入力用帳票の受信側のファックス装置であり、一般とは、その他のすべての場合である。

表2 HITAC T-550/30 OCRシステム機器仕様 この仕様は、OCRシステムの一般的仕様を示す。

項 目	内 容
帳 票 寸 法	長さ: 88.9～305mm 幅: 58.4～265mm(ただし、長さ0.5≦幅≦長さ2.2) 厚さ: 0.10～0.17mm(70～123kg紙)
読 取 り 文 字 種	手書き文字 数字: 0～9 スペース, 英字: A～Z スペース 片仮名: 「ア」～「ン」 " スペース(ただし「ヲ」は除く。) 記号: =, +, -, /, その他 活字 (OCR-Bフォント) 数字: 0～9 英字: A～Z
読 取 り 文 字 数	手書き: 最大46字/行, 32行/枚 活 字: 最大82字/行, 32行/枚(ただし、1帳票当たり1,600字以内)
処 理 速 度	最高約50枚/分(帳票長さ90mm, 手書き数字10文字で1行の場合)
出 力 媒 体	フロッピディスク 通信回線(公衆: 1,200bps, 専用: 1,200/2,400bps)
ホッパ及びスタック容 量	ホッパ: 約300枚×1 スタック: アクセプト用 約300枚×1, リジェクト用 約100枚×1
ディスプレイ / キーボード	ディスプレイ: 40字×16行 9in CRT キーボード: 整配列(英字, 数字, 片仮名, 記号)

外部条件

寸 法	OCR部: 幅750×奥行870×高さ1,160(mm), 操作部: 幅650×奥行650×高さ1,100(mm)
重 量	OCR部: 300kg, 操作部: 70kg
電 源	電圧 AC100V, 単相 50/60Hz, 所要電力 3kVA
発 熱 量	約1,500kcal/h
周 囲 条 件	動作時: 温度10～35℃, 湿度35～70%RH 停止時(電源オフ): 温度4～43℃, 湿度8～90%RH



注：* HT-F5503-44 一般ボールペン読取り機構を付加した場合、青系統のドロップアウトカラーは使用できない。
* HT-F5503-70 プリンタ接続機構及びHT-5332-11プリンタは、HT-5503-12基本システムにだけ接続できる。

図5 OCRシステムの機能構成 HITAC T-550/30分散形OCRシステムはHT-5503-IIを基本システムとし、付加機能として右側に示したものがある。新たにファクシミリ出力帳票読取り機能を追加した。

4.1 位置ずれの補正

OCRの帳票には図6に示すような各行右端に行位置を示すレファレンスマーク、第1行最左部にスキュー補正、帳票伸縮検出用のベリファイマークがプレ印刷されている。OCR読取時でのスキュー要因としては、OCR内での紙送りスキューに加え、ファクシミリで出力される際に発生する帳票上の画面のスキューがあり、従来以上のスキュー補正の強化が必要となる。このため、第1行のレファレンスマークとベリファイマークの縦方向の位置ずれを検出することにより、文字パターン取込み位置をダイナミックに補正する処理を加えた。一方、ファクシミリ側でのスキュー補正については、多数枚の送信原稿を一枚ずつ自動的に送り出す給紙部の改善を行なった。

4.2 ファクシミリ出力帳票(静電記録紙)

OCRではOCR用紙(JIS C-6253で規定)を入力することを前提に紙送り機構が設計されているので、ファクシミリ用記録紙をそのまま入力するとミスフィード率が高くなるという問題が発生した。このため、ファクシミリ受信紙の紙質の改善、すなわちOCR用紙にできるだけ紙質を近づけることが必要となった。性能上特に紙厚、腰の強さに重点を置き開発

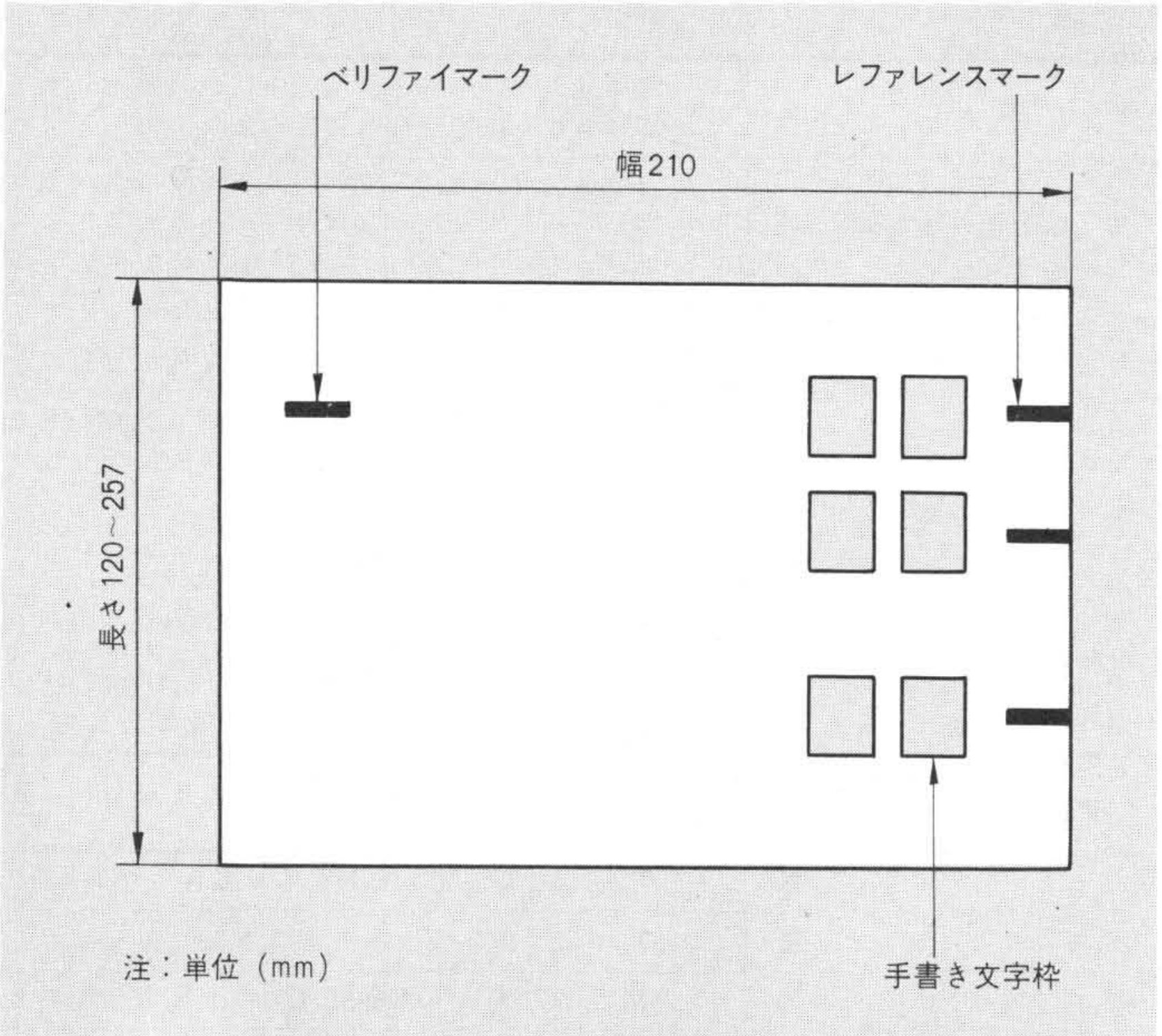


図6 入力帳票仕様 OCR入力用ファクシミリ送信原稿の帳票フォーマットの例を示したものであり、手書き文字枠の中に伝達したい文字、記号を記入する。

表3 紙質特性表 OCR用紙とファクシミリ記録紙の紙質特性を示す。
上表中のOCR用ファクシミリ記録紙が、OCRへの入力としてこのたび開発したものである。OCR用紙の特性は、概略JIS C6253と同じである。

項 目		単 位	70kg用 OCR用紙	OCR用 ファクシミリ 用記録紙	標 準 ファクシミリ 用記録紙
坪	量	g/m ²	83.7	97.3	59.5
厚	さ	mm	0.105~0.120	0.110	0.07
平	滑	S	20~60	33	32
剛 度	M.D	mg	150以上	256	37.7
	C.D	mg	100以上	138	26.5
破 裂	強 さ	kg/cm ²	2以上	3.6	0.8
引張り強さ	M.D	kg/15mm	9~11	11	4.1
	C.D	kg/15mm	6~8	7	3.5
摩 擦	係 数	—	0.3~0.6	0.62	0.5
不 透 明	度	%	80~90	85	73

注：略語説明 M.D (Machine Direction)
 C.D (Cross Direction)

を進め、製紙メーカーと共同検討を行ない、表3に示す紙質の記録紙を実用化した。

4.3 ファクシミリ受信画の品質

ファクシミリ受信画で文字欠け、文字抜けはOCRの読取り率に大きく影響する。OCR帳票に記入された文字、記号などをファクシミリで正確に読み取るようにするため、ファクシミリ送信部の原稿読取りレベルを改善した。

4.4 ドロップアウトカラーの選定

OCRは帳票上のどの部分を読み取るかをあらかじめ規定して、文字を読み取っていく。そのため、帳票上にはOCRでは検出できないドロップアウトカラーで読み取り領域及び文字枠を印刷し、その枠の中にデータを記入する。一方、ファクシミリは原則として帳票中のどのような色でも送信できるほうがよいことが多い。OCRとファクシミリとでは読取りレベルが異なるため、OCR指定のドロップアウトカラーの中から、表4に示すファクシミリにも適用できる色を選択した。

表4 ドロップアウトカラー 手書きOCR帳票の文字記入枠部に使用し、OCR及びファクシミリでは読み取らない色である。

項 番	ファクシミリ	株式会社日本ビジネス コンサルタント	大日本印刷株式会社	トッパンムーア株式会社	小林記録紙販売株式会社
1	A4、B4判ファクシミリ 共通ドロップアウトカラー	5106-H30 ブライトグリーン	D251-H30 マリンブル	75-H30 グリーン	1301-H30 パルグリーン
		0502-H30 スカイグリーン	D221-H30 パステルグリーン		1101-H30 グリーン
					6201-H30 ライトブルー
2	A4判ファクシミリだけの ドロップアウトカラー (項番1も使用可能)	2301-H30 ブラウン	D114-H30 レディッシュオレンジ	51-H30 オレンジ	7501-H30 イエローオレンジ
		2310-H30 ペールブラウン	D104-H30 オレジジ		
3	B4判ファクシミリだけの ドロップアウトカラー (項番1も使用可能)	7021-H30 ブルー			

5 検討結果

4章で述べたように、ファクシミリとOCRを結合するために種々の改良を行ない、実用化できる結果を得た。

実用化に関し最も重要なものは、ファクシミリの受信画をOCRへ入力したときの読取り率である。ファクシミリ伝送による画質低下は、ファクシミリそのもの及び電話回線の品質により影響され、結果としてOCRでの読取り率を悪くする。

ファクシミリ受信画をOCRへ入力したときのOCRの読取り率は、OCR帳票へ文字を記入するとき、標準字体を見ながら記入したときのレベルとほぼ同じである。電話回線の品質による影響は、ファクシミリの電送速度を遅くすることにより改善される。

6 ファクシミリとOCRシステムの特長

(1) きめ細かく奥行の深いネットワークの実現

日本全国、ファクシミリを設置できる場所ならどこからでもOCRの入力データを送信することができる。

(2) ターンアラウンドタイムの短縮

OCR帳票の送付方法が2~3日かかっていた郵送からファクシミリ伝送に切り換えることにより、データ処理が行なえ、ターンアラウンドタイムが短縮できる。

(3) システムの柔軟性の発揮

OCRのデータエントリ用原帳票には、配送先の地図などは記入できなかったが、ファクシミリを利用することにより、システム範囲が拡大され、多様化に対処できる。

(4) 操作性の向上

ファクシミリそのものの操作が簡単のため、容易に入力データが伝送でき、人的要因による入力隘路が解消できる。

7 ファクシミリとOCRシステムの運用例

図7に「受注出荷システム」の例について示す。このシステムは、得意先から受注した営業所で、受注伝票（ファクシミリ送信紙）に商品コード、数量などの必要事項を記入し、日立ファクシミリ“HF550”で配送センタ（出荷部門）に伝送する。ファクシミリ送信紙は、HITAC T-550/30 OCRの入力帳票形式でフォーマットされているため、受信部署では、受信紙になんら手を加えることなくそのままOCRで読

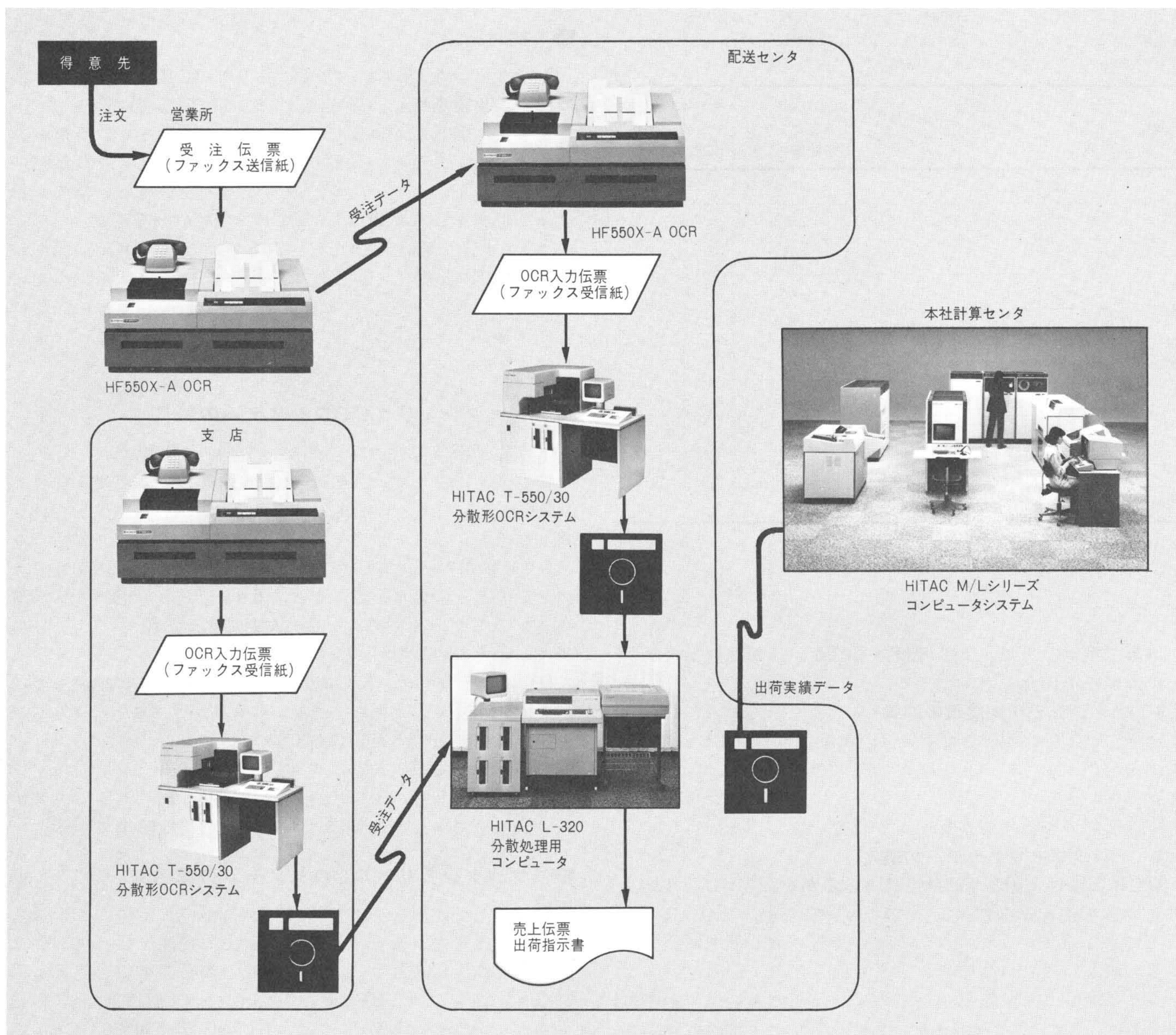


図7 ファクシミリとOCRシステムの運用例 この運用例は、受注から出荷までのシステムについての例を示したものである。

み取らせることができる。このとき、HITAC T-550/30 OCRでは受信紙のプログラムの指示に従って自動的にプログラムがロードされ、帳票データを読み取る。フロッピディスクに蓄えられた各支店のOCR出力データは一括して配送センタに伝送したり、また直接HITAC L-320分散処理用コンピュータにフロッピディスクをセットして、売上伝票及び出荷指示書の発行を行なう。また、先に配送部内で受信した受信紙そのものは、配送先略図やデータ処理の対象とならない特別のコメントが記載されているため、出荷指示書に受信紙を添付して出荷係へ回付する。営業所に対しては出荷報告として売上伝票をファクシミリで返送する。本社計算センタには、HITAC L-320分散処理用コンピュータから出荷実績としてデータを一括伝送する。

このシステムの特長として、

- (1) ファクシミリにより、スピーディで確実なデータ収集ができ、各営業所、支店で発生する伝票を合理的に収集できる。
- (2) ファクシミリにより受信された伝票を、OCRシステムでそのまま読み取れるので、業務の簡素化が図れる。

- (3) OCRを設置した配送センタなどを中心として、ファクシミリ通信により各営業所を結ぶという経済的で大規模なネットワークを構成することができる。

8 結 言

以上述べたように、ファクシミリ受信紙を媒体としたファクシミリとOCRシステムの実用化を実現したが、更に高品質のファクシミリとOCRシステムを実現するために、図1の(b)方式が現在考えられている。ファクシミリとOCRのデータエントリシステムは、情報化分野での技術革新の大きなインパクトとなりつつあり、今後ますます情報化時代の進展につれて、ファクシミリとOCRシステムが必要になるものと予想される。日立製作所は、今後よりいっそう効果的な情報処理技術の開発を進めてゆく考えである。

参考文献

- 1) 田村, 外: 日立パケット交換システム“HIPA-NET”用ファクシミリ端末, 日立評論, 60, 721~725 (昭53-10)