

漢字システムサポート・ソフトウェア

KANJI System Support Software

近年、情報処理システムのコンピュータ化は、あらゆる分野で急速に拡大している。しかし、コンピュータの入出力機器は英・数字を中心としたものであり、日常の業務処理で用いる漢字、仮名を含むデータの読みやすさ、分かりやすさと比較して、情報処理結果の出力としては十分でない。

このような状況のもとで、日立製作所では日本語情報処理のためのハードウェア、ソフトウェアの研究を行ない、漢字情報処理システムを開発した。本稿では、漢字情報処理システムのソフトウェアについて述べる。これにより、コンピュータの出力を分かりやすい日本語情報で得ることができ、更にコンピュータの利用分野を広げることができる。

吉田浩三* *Yoshida Kôzô*
原田千秋* *Harada Chiaki*
柏木貞一* *Kashiwagi Teiichi*

1 緒 言

コンピュータによる情報処理の適用分野が拡大するにつれて、コンピュータの出力結果を利用する利用者層も広がっている。通常の業務処理のなかで、コンピュータによる情報処理結果を分かりやすいものにするため、情報の利用者からは漢字、仮名を含めた出力の要求が高まってきている。

しかし、この漢字、仮名といった日本語をコンピュータで処理する上で、日本語特有の文字種、パターンの複雑さ、同音異義語など多くの問題点がある。これらの問題点を解決し、コンピュータによる情報処理の効果をより満足できるものに近づけるため、日本語情報処理の分野を四つに分類し、各分野の研究開発を進めてきた。

日本語情報処理分野のうち漢字、仮名を含むデータの入出力及び文章編集機能をもつ漢字システムサポート・ソフトウェアを製品化した。そこで、漢字システムサポート・ソフトウェアの位置づけ及び特長について説明する。

2 漢字システムサポート・ソフトウェアの位置づけ

2.1 日本語情報処理の分類

日本語情報処理の分野は、機能的に表1に示すように分類される。

また、日本語情報処理の機能をコンピュータで扱う上で、既存のコンピュータシステムに蓄積されている情報とシステムの運用形態を大幅に変更せずに導入することが必要である。

日立製作所は、このような背景のもとに、漢字データ処理、簡易編集処理分野のサポート・ソフトウェアを製品化した。特殊編集処理、言語処理分野については、現在更に研究開発を進めている。

2.2 漢字システムサポート・ソフトウェアの構成

漢字システムサポート・ソフトウェアは、HITAC Mシリーズ VOS 2 (Virtual Storage Operating System 2) のもとで動作する。

漢字システムサポート・ソフトウェアのうち、漢字入力編集、漢字ライブラリ保守、外字処理、書式オーバーレイゼネレータ及び漢字ビデオデータシステム用画面制御(MAP/5435)から成る漢字処理用ユーティリティは、漢字、仮名を含むデータの入力から出力まで一貫した処理を可能とし、また漢字編

集プログラムは、理解しやすい日本語を容易に編集出力することができ、コンピュータによる業務処理の適用範囲を一段と拡大することができる。

また、漢字システムサポート・ソフトウェアでは、入出力装置としてH-1811-10漢字入力装置、H-8196漢字プリンタサブシステム(バッチ処理用)、H-8195漢字プリンタサブシステム(オンライン用)及びHITAC T-560/40漢字ビデオデータシステム(オンライン用)をサポートしている。

図1に、VOS 2での漢字システムサポート・ソフトウェアの構成を示す。

表1 漢字情報処理の分類 漢字システムサポート・ソフトウェアは、漢字データ処理、簡易編集処理の分野をサポートしている。

分 類	適 用 業 務	必要なソフトウェア
1. 漢字データ処理	従来、英・数字データで情報処理していたものを、分かりやすい日本語情報の出力を中心とした処理方式に改善する。既存システムへの漢字の導入を容易にする。	人事、販売、在庫管理などの一般事務処理 (1)漢字入力編集 (2)漢字ライブラリ保守 (3)外字処理 (4)書式オーバーレイゼネレータ(帳票フォーマットの作成) (5)COBOL, PL/I言語でのプログラミング
2. 簡易編集処理	漢字、仮名を含む日本語文章を、文の規則あるいは割付けの規則に従って、編集処理を行なう。	会議録、索引編集、ドキュメントの作成など 漢字編集プログラム
3. 特殊編集処理	簡易編集処理の拡張であり、新聞編集、出版などの専門分野の編集技術を要する。	出版、新聞、印刷関係 自動編集、組版システムなど
4. 言 語 処 理	日本語情報の意味解析、構文解析などが中心となる。	日本語研究、翻訳など (1)漢字仮名交じり文変換 (2)同音異義語処理 (3)キーワード自家抽出など

* 日立製作所ソフトウェア工場

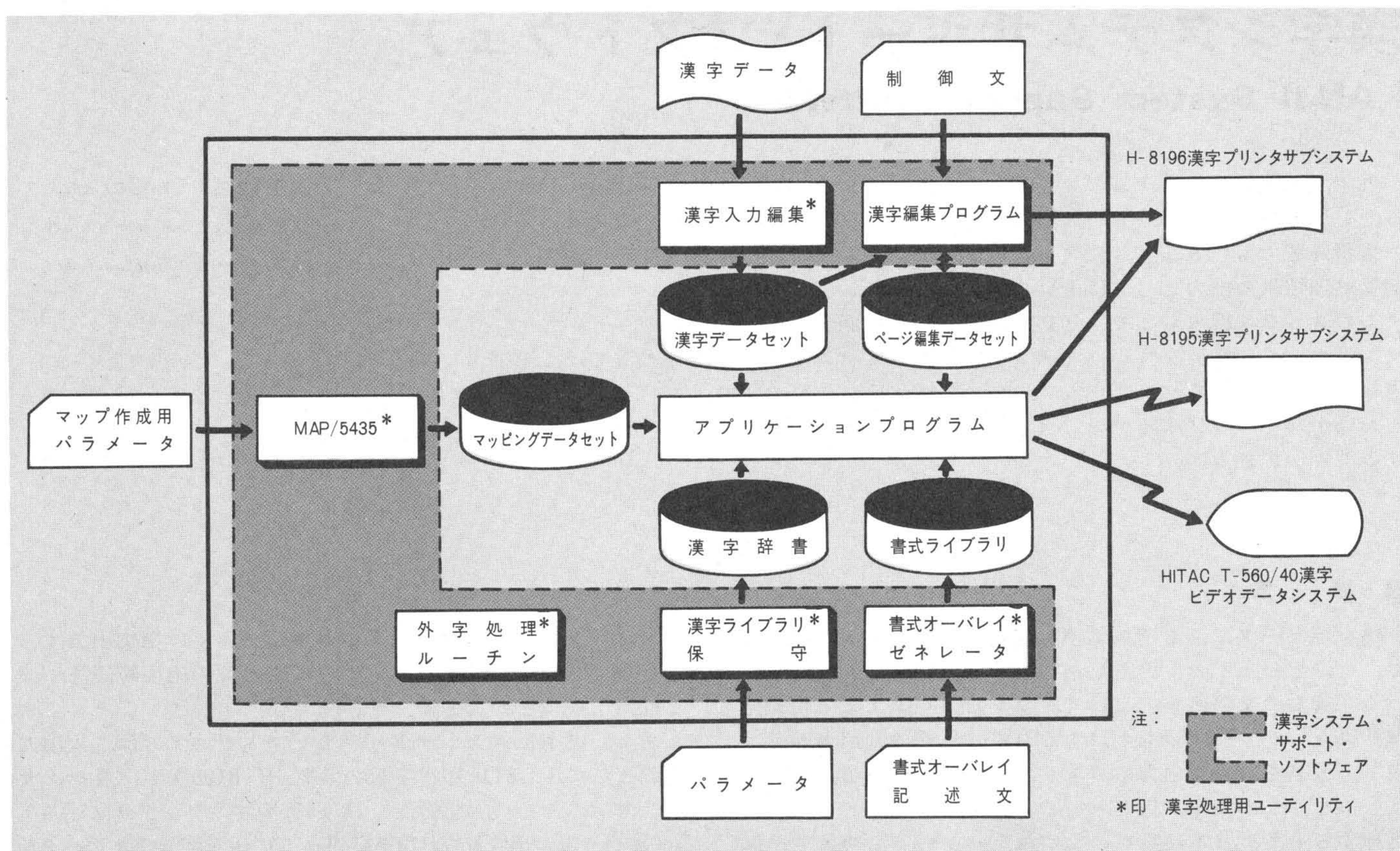


図1 漢字システムサポート・ソフトウェアの構成 漢字仮名交じりデータの入力から出力まで、一貫した処理が可能である。

3 漢字システムサポート・ソフトウェアの特長

3.1 漢字処理用ユーティリティ

3.1.1 漢字データ処理の流れ

VOS 2での漢字データ処理の流れを、図2に示す。

H-1811-10漢字入力装置により、漢字データを紙テープにせん孔し、漢字入力編集ユーティリティにより漢字データセットに変換する。また、既存の仮名データを漢字データに変換して用いてもよい。このようにして作成された漢字データセットは、ユーザープログラムあるいは漢字編集プログラム

により、出力形式を整えて、漢字出力装置に出力される。

VOS 2では、H-8196漢字プリンタサブシステムを、データ管理の順アクセス法SAM (Sequential Access Method) でサポートしている。また、H-8195漢字プリンタサブシステム、HITAC T-560/40漢字ビデオデータシステムを、基本通信アクセス法BTAM (Basic Telecommunication Access Method)、汎用オンラインコントロールプログラムTCS (Transaction Control System) 及びTMS-3V (Transaction Management System-3V) でサポートしている。

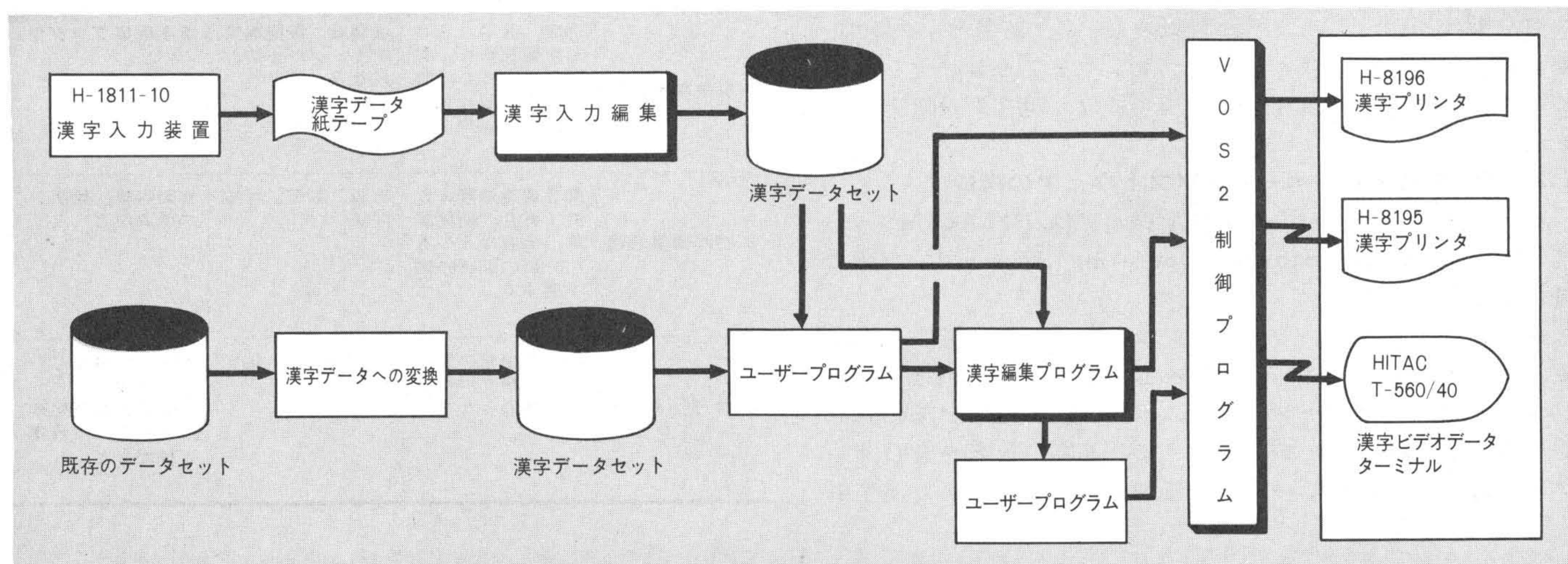


図2 漢字データ処理の流れ VOS 2での漢字データ処理の流れを示す。

3.1.2 漢字処理ユーティリティの特長

漢字処理のために、次に述べるユーティリティがある。

(1) 漢字入力編集

H-1811-10 漢字入力装置で作成した漢字入力データを編集し、漢字データセットへ出力する。コード変換のほか、レコードの編集、修正が可能である。

(2) 漢字ライブラリ保守

日立標準提供漢字辞書を入力として、ユーザーシステム上に漢字辞書を作成する。また、H-8196、H-8195漢字プリンタ制御用の各種のデータを作成する。

(3) 外字処理

漢字出力装置内に常駐していない文字パターンを出力するための処理(外字処理)には、図3、4に示すように2通りの方式がある。

プレロード方式は、漢字ライブラリ保守を用いて必要な外字を、文字パターンデータを格納するライブラリ(イメージライブラリ)に登録しておき、出力処理を行なう前に登録された外字をローディングする方式である。オンデマンドロード方式は、データ出力時に、出力装置によって外字検出が行われた時点で、外字を漢字辞書から取り込んでローディングする方式である。

オンデマンドロード方式を実現するために、外字処理ルーチンがあり、これにより、出力データ中に含まれる突発的な外字の出現に対して対応することができる。

(4) MAP/5435

MAP (Mapping Aid Package)/5435は、HITAC T-560/

40漢字ビデオデータシステム、ビデオデータターミナルの、画面制御機能をサポートするパッケージであり、これにより画面の編集操作を容易に行なうことができる。ターミナルと中央コンピュータ間で送受する複雑な伝送形式のデータを画面の形式に変換することは、ユーザープログラムに大きな負担となるが、MAP/5435により、この変換を容易にすることができる。

(5) 書式オーバーレイゼネレータ

(a) 書式オーバーレイ機構

H-8196漢字プリンタサブシステムには、書式オーバーレイ機構を付加できる。これは、1ページ分の出力データに、あらかじめ用意した書式を重ねて、印字出力できる機構である。従来のラインプリンタでは、書式をあらかじめ印刷した用紙を用いていたのに対して、白紙の出力用紙が使用でき、運用コストの低減を図ることができるという利点がある。

(b) 書式オーバーレイゼネレータの機能

書式は、書式オーバーレイゼネレータにより、書式ライブラリ中に作成される。書式は1ページ分の大きさで、ライン、文字パターン情報、ドット情報から構成される。ラインは、水平、垂直の実線及び点線である。文字パターン情報は、漢字辞書に含まれる文字である。ドット情報は、任意の大きさの矩形に対して、ユーザーの作成できるドットパターンである。

書式オーバーレイゼネレータは、書式オーバーレイ記述文で与えられたこれらの情報に従って、ライン、文字パターン情報及びドット情報に展開して、書式オーバーレイモジュール

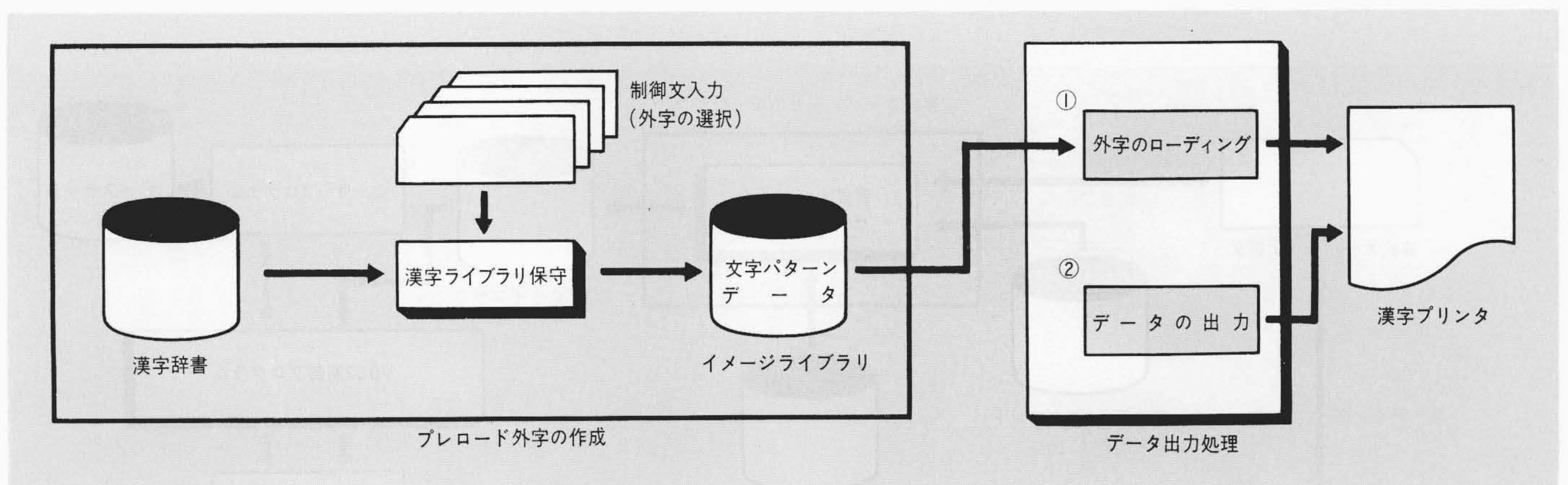


図3 プレロード外字処理方式 プレロード外字処理は、前もって外字を出力装置にローディングしておき、印刷出力時の外字の検出を抑止する方式である。

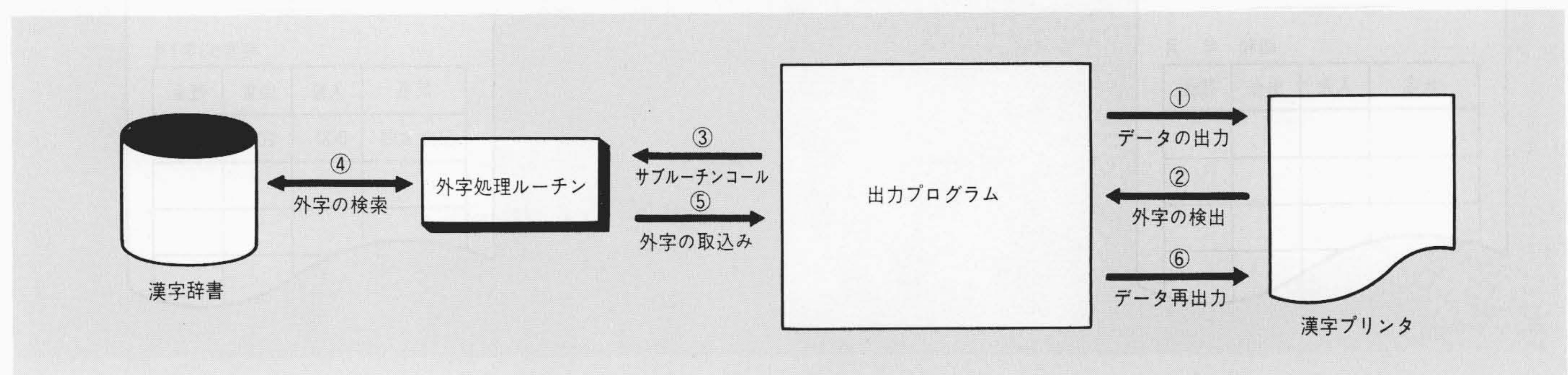


図4 オンデマンドロード外字処理方式 オンデマンド外字処理は、印刷出力時に検出された外字を、出力装置にローディングして再実行させる方式である。

表2 代表的な書式オーバーレイ記述文 書式オーバーレイモジュールは、表のような記述文によって作成される。この例に示す基本的な記述文を組み合わせ、新たに記述文を追加することもできる。

記述文	機能
FORMS	書式の大きさを定める。
SEGMENT	書式部品としてのセグメントの大きさを定める。
LINE	垂直、水平の直線の位置、長さ及び線の種別を定める。
GRAPHIC	書式内にドット情報を置く領域を定める。
DOT	GRAPHIC文で確保した領域に埋め込むドットパターンを与える。
CHAR	書式内におく文字の位置、大きさを定める。
CODE	CHAR文で確保した領域におく文字をコードで与える。

ルを作成する。

表2に、書式オーバーレイゼネレータの代表的な記述文を示す。

(c) 書式作成の手順と運用

書式作成の手順と運用を、図5に示す。

書式の内容が決まると、書式を表現する書式オーバーレイ記述文を作成する。

このとき、書式の部品(セグメント)を定義することにより、同一内容の繰返しや、各書式での共通部品として使用することができる。

作成された書式は、書式ライブラリに格納される。ジョ

ブ制御文又はデータ管理マクロにより、ジョブ開始時あるいは印刷出力時に、必要な書式をH-8196漢字プリンタサブシステムにローディングする。ローディングされた書式は、1ページ分のデータに重ねて出力される。

3.2 漢字編集プログラム

漢字編集プログラムは、従来、写植機や印刷機により行なわれていた各種の通知書や報告書などの簡易印刷物の作成作業を、コンピュータで行なうための漢字アプリケーション・プログラムの一つである。

漢字編集プログラムがサポートする漢字プリンタとして、会議録などの簡易印刷物作成に適したH-8195漢字プリンタと大量データの印刷物作成用としてのH-8196漢字プリンタとがある。

3.2.1 漢字編集プログラムの処理概要

漢字編集プログラムの入力データは、原稿用紙上に記述した文章とこの文章中に挿入する改行や改ページ、行分割禁止などの編集処理を、漢字編集プログラムに対して指示する機能キャラクタと呼ぶ編集指示コードから成る。

原稿用紙上に記述したデータは、H-1811-10漢字入力装置により紙テープにせん孔し、漢字入力編集ユーティリティにより漢字データセットに出力する。

漢字編集プログラムは編集処理に先立ち、用紙の大きさ、用紙の縦組み・横組みの種別、印字文字の大きさなど基本組み仕様と呼ぶ仕様を漢字編集プログラムの制御文により与え、編集処理を行なう上での基本的な編集仕様を定義する。次に入力データ中の機能キャラクタに基づいて編集処理を行ない、編集処理結果を漢字プリンタに出力する(図2参照)。

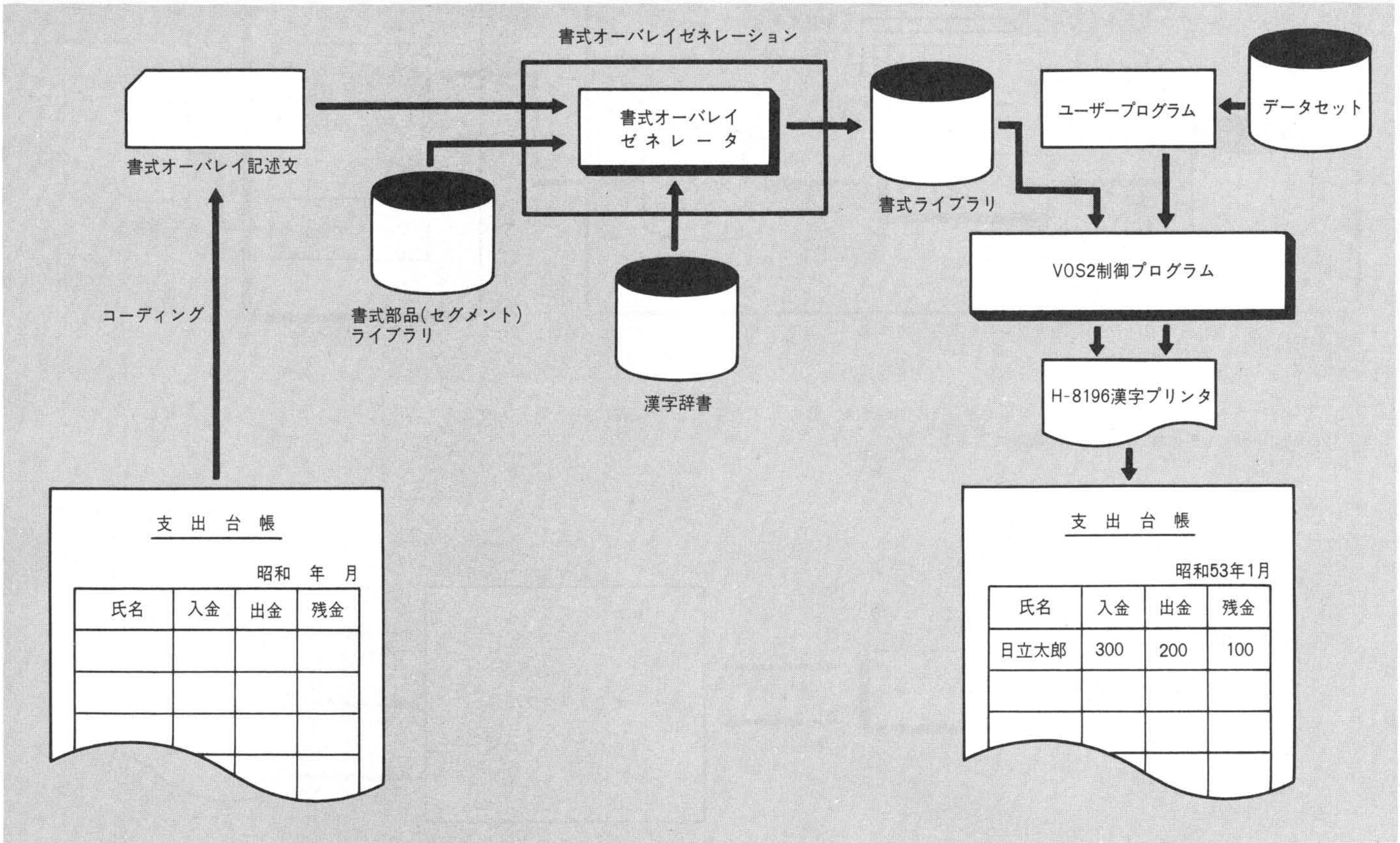


図5 書式作成の手順と運用 書式は、書式オーバーレイゼネレーションを経て、書式ライブラリに書式オーバーレイモジュールとして格納される。データ出力時に、書式ライブラリからプリンタに書式オーバーレイモジュールがローディングされ、データに重ねて印刷される。

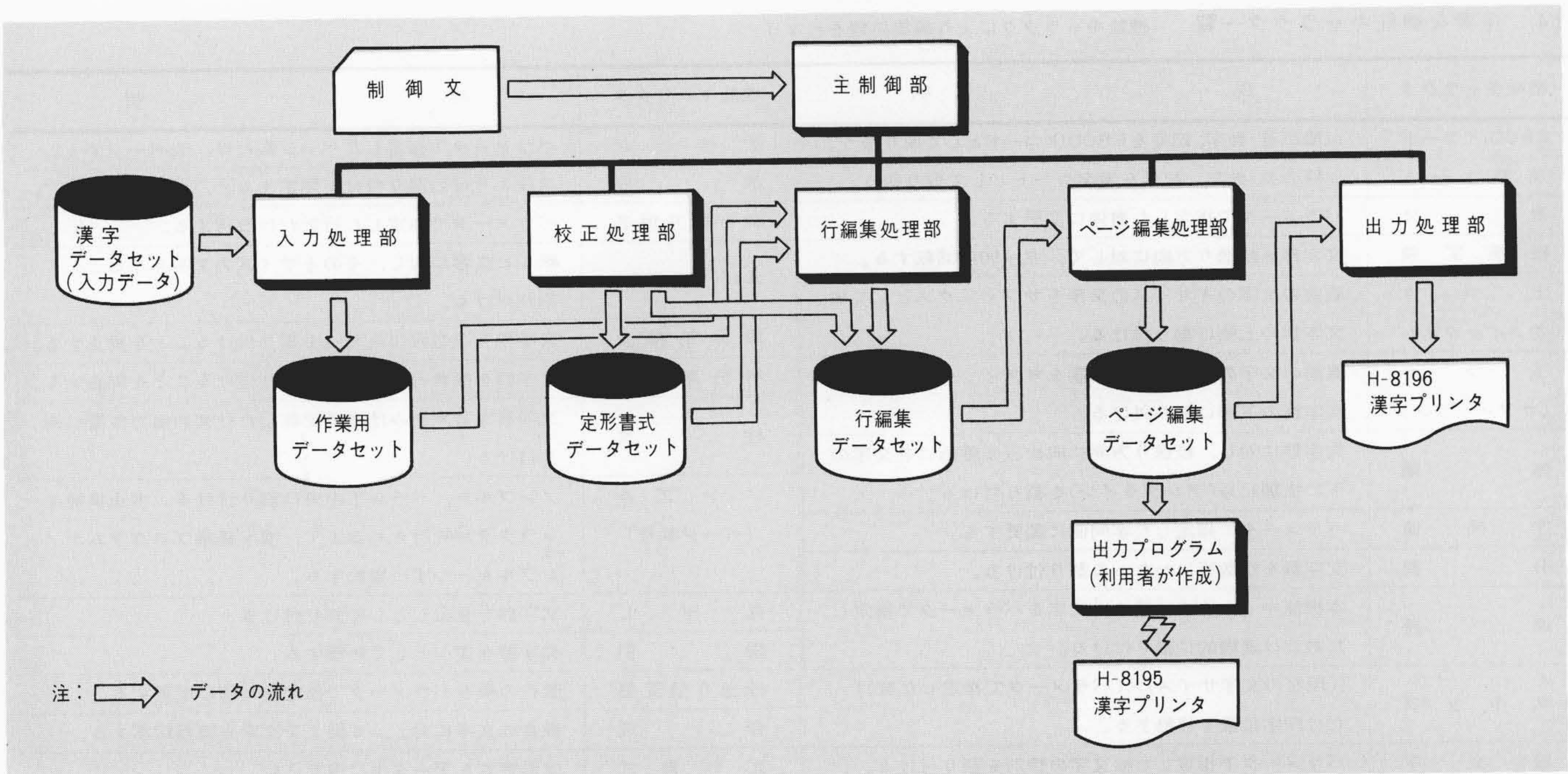


図 6 漢字編集プログラムのプログラム構成 入力データがデータセットを介して各処理部により処理され、漢字プリンタに出力される。

3.2.2 漢字編集プログラムの構成

漢字編集プログラムのプログラム構成は、図 6 に示すように各処理部ごとに独立した構造となっている。

(1) 主制御部

制御文を入力し、この制御情報に基づいて各処理部の実行制御を行なう。

(2) 入力処理部

漢字入力編集ユーティリティにより処理されたデータを入力する。通常、データは章、節などの単位に分割して入力されるが、一度に入力される単位をロットと呼ぶ。入力処理部では、ロット単位に入力されたデータをロットに付加した番号の上昇順に並べ直し、1 ドキュメント単位とする。

(3) 校正処理部

行編集データセット上のデータ又は定形書式データセット上の定形書式データに対して、校正処理を行なう(定形書式：頻繁に使用する文章、表などをあらかじめデータセットとして登録しておき、必要なときに固有の名称により引用する機能)。

(4) 行編集処理部

基本組み仕様及び入力データ中に挿入した行分割禁止、改行、ルビ(振り仮名)など、行の編集に関する機能キャラクタに基づいて行単位の編集処理を行なう。定形書式の引用が指定されている場合には、定形書式データセットから定形書式データを取り込む。

(5) ページ編集処理部

行編集処理部で処理された行編集済みデータを、ページ単位に編集し、ページ番号、柱文(ドキュメントの余白部上方に印刷する見出しの一種)を付与してページとしての体裁を整える。更に指定により、目次及び索引の作成を行なう。ページ単位に編集したページ編集済みデータをページ編集データセットに出力する。

(6) 出力処理部

ページ編集データセットのページ編集済みデータを漢字プリンタへ出力する。

本出力処理部が対象とする漢字プリンタは、H-8196漢字プリンタである。H-8195漢字プリンタへの出力は、ユーザープログラムを介して行なう。

3.2.3 漢字編集プログラムの機能

漢字編集プログラムの機能は、次の二つに大別できる。

(1) 編集処理機能

(2) 校正処理機能

編集処理機能を実現するものとして、基本組み仕様と機能キャラクタがある。それらの一覧を表 3、4 に示す。また、漢字編集プログラムが自動的に行なう編集処理機能として、行頭や行末に位置することを禁止する文字(行頭：右括弧、カンマなど、行末：左括弧)に対し、直前の行へ移動したり、次行へ移動して割り付ける禁則処理がある。

校正処理機能には、文字の追加、削除及び置換機能がある。

表 3 基本組み仕様 編集処理に先立ち、本表の項目欄に示す編集仕様を定義する。

項 目	説 明
用 紙 サ イ ズ	用紙サイズを定義する。
基 本 組 方	縦組みか横組みかを定義する。
余 白 指 定	漢字編集プログラムの印字領域を余白(天地、左右)により定義する。
基本文字サイズ	文字サイズの指定がないときや文字サイズの指定に誤りがある場合に使用する文字サイズを定義する。
段 数	段組みを行なう場合の段数を定義する。最大 3 段まで定義できる。
段 間 値	段と段との間隔を定義する。
行 送 り 値	行の送り値を定義する。
字 間 値	文字と文字との間隔を定義する。
柱 文	柱文の出力位置を定義する。
段 野	多段組みの場合、段の中央に罫線を付加するかどうかを定義する。

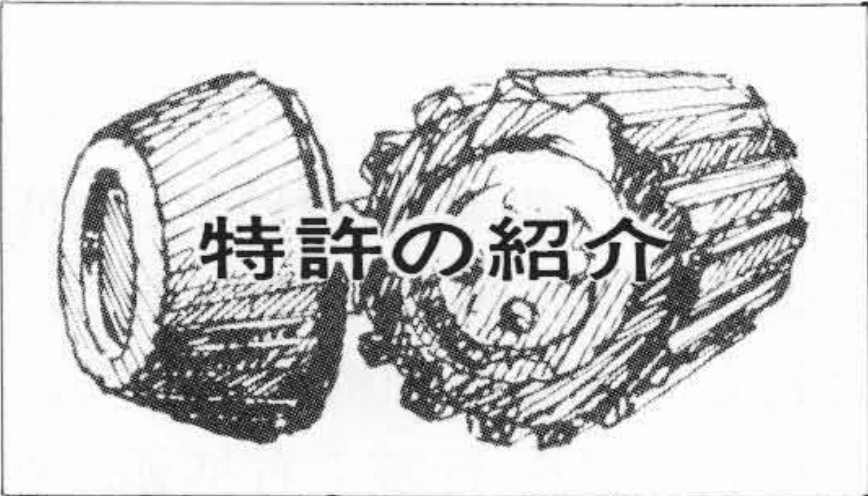
図 4 主要な機能キャラクター一覧 機能キャラクタにより編集処理を行なう。

機能キャラクタ	説 明	機能キャラクタ	説 明
EBCDIK コード	以降の英・数字, 記号をEBCDIKコードとして取り扱う。	改 ペ ー ジ	パラメータで指定したページ数だけ, 改ページする。
漢 字 コ ー ド	以降の英・数字, 記号を漢字コードとして取り扱う。	改 段	次段より行の割り付けを開始する。
書 体	パラメータで指定した書体に変更する。	改行振り仮名	パラメータで指定した行数だけ改行する。
縦 横 変 換	文字群を紙送り方向に対して, 左へ90度回転する。	ルビ(振り仮名)	被ルビ文字に対し, その $\frac{1}{2}$ サイズの文字をルビとして割り付ける。
上 つ き (サフィックス)	直前の文字の $\frac{1}{2}$ サイズの文字をサフィックスとして被文字群の上側に割り付ける。	段 分 割 禁 止	文字群を複数段にわたって割り付けることを抑止する。
下 つ き (サフィックス)	直前の文字の $\frac{1}{2}$ サイズの文字をサフィックスとして被文字群の下側に割り付ける。	行 分 割 禁 止	文字群を複数行にわたって割り付けることを抑止する。
強 調	文字群に対し, 紙送り方向に向かって垂直に各文字の下に強調記号(アンダライン)を割り付ける。	柱 文	文字群を基本組み仕様で定義した柱文の出力位置に割り付ける。
字 間 値	パラメータで指定した字間値に変更する。	ノ ン ブ ル (ページ番号)	ノンブルを, ページ下中央に割り付ける。中止機能キャラクタが発行されるまで, 漢字編集プログラムがノンブルを一つずつ増加する。
分 割	文字群を複数行にわたって割り付ける。	見 出 し	文字群を見出しとして割り付ける。
連 続	本機能キャラクタに続く 1 文字をパラメータで指定した数だけ連続的に割り付ける。	索 引	文字群を索引として処理する。
ス キ ッ プ	{(現在の文字サイズ)×(パラメータで指定した数)}だけ印字位置を移動する。	行 送 り 値 変 更	改行の幅をパラメータで指定した値に変更する。
絵 文 字	パラメータで指定した絵文字の種別を割り付ける。	合 成	被合成文字に対し, 4 回まで文字を重ね印字する。
		定 形 書 式	定形書式をデータ中に取り込む。

4 結 言

VOS 2 漢字システムサポート・ソフトウェアの位置づけ及び特長を中心に説明した。本ソフトウェアにより、コンピュータによる情報処理結果を日本語で出力することができ、日常業務で直ちに役立つ分かりやすい情報を得ることができ

るようになった。
現在、コンピュータの処理は、データベースの多目的利用とオンライン化の方向へ進みつつある。
今後の日本語情報処理も適用分野の拡大のための問題点の解決に加えて、これらコンピュータ利用の動向に沿って発展させてゆく考えである。



画面表示方式

堀野 寛・藤方健一・平野睦房
特許 第885830号 (特公昭52—14052)

本発明は、電子計算機から出力される図形情報に基づいて、塗りつぶし図形を表示するラスタ走査表示装置に関するものである。図1のような三角形 $p_0 p_1 p_2$ を塗りつぶしてCRTで表示する場合について以下に説明する。
図2で、線⑤を介して頂点 p_0 , p_1 の座標値がタイミング信号発生器⑥に入力される。タイミング信号発生器⑥は、CRT上に線分 $p_0 p_1$ を表示すべき時刻を上記頂点の座標値から各ラスタごとに決定し、その時刻ごとにタイミング信号⑦を発生し、フリップフロップ⑧をセットする。同様に、線⑤を介して頂点 p_0 , p_2 の座標値がタイミング信号発生器⑥に入力される。タイミング信号発生器⑥は、CRT上に線分 $p_0 p_2$ を表示すべき時刻を上記頂点の座標値から各ラスタごとに決定し、その時刻ごとにタイミング信号⑦を発生し、フリップフロップ⑧をリセットする。フリップフロップ⑧の出力⑨は、ビデオディスク⑩を塗りつぶす信号として用いられる。これらのタイミング信号の発

生は、ビデオディスク⑩からの同期信号⑪に同期して行なわれる。このようにして、ビデオディスク⑩には3角形 $p_0 p_1 p_2$ の塗りつぶし表示のためのラスタ信号が記憶される。次に線⑤を介して、タイミング信号発生器⑥に頂点 p_1 , p_2 の座標値を与え、先に頂点 p_0 , p_2 の座標値を入力されているタイミング信号発生器と協働して、三角形 $p_1 p_2 p_1$ の塗りつぶし信号をフリップフロップ⑧より発生させ、ビデオディスク⑩に記録する。こうしてビデオディスク⑩には図1の三角形 $p_0 p_1 p_2$ の塗りつぶし表示のためのラスタ信号が記憶される。このビデオディスク⑩から、記憶されたラスタ信号⑫を順次読み出し、CRT⑬で表示すると所望の三角形 $p_0 p_1 p_2$ の塗りつぶし図形が表示される。
上述の手順を繰り返すことにより、単に三角形に限らず、任意の多角図形若しくは任意の曲線図形の多角近似図形を、塗りつぶし表示することができる。しかも、このために用いる装置は極めて安価な装置である。

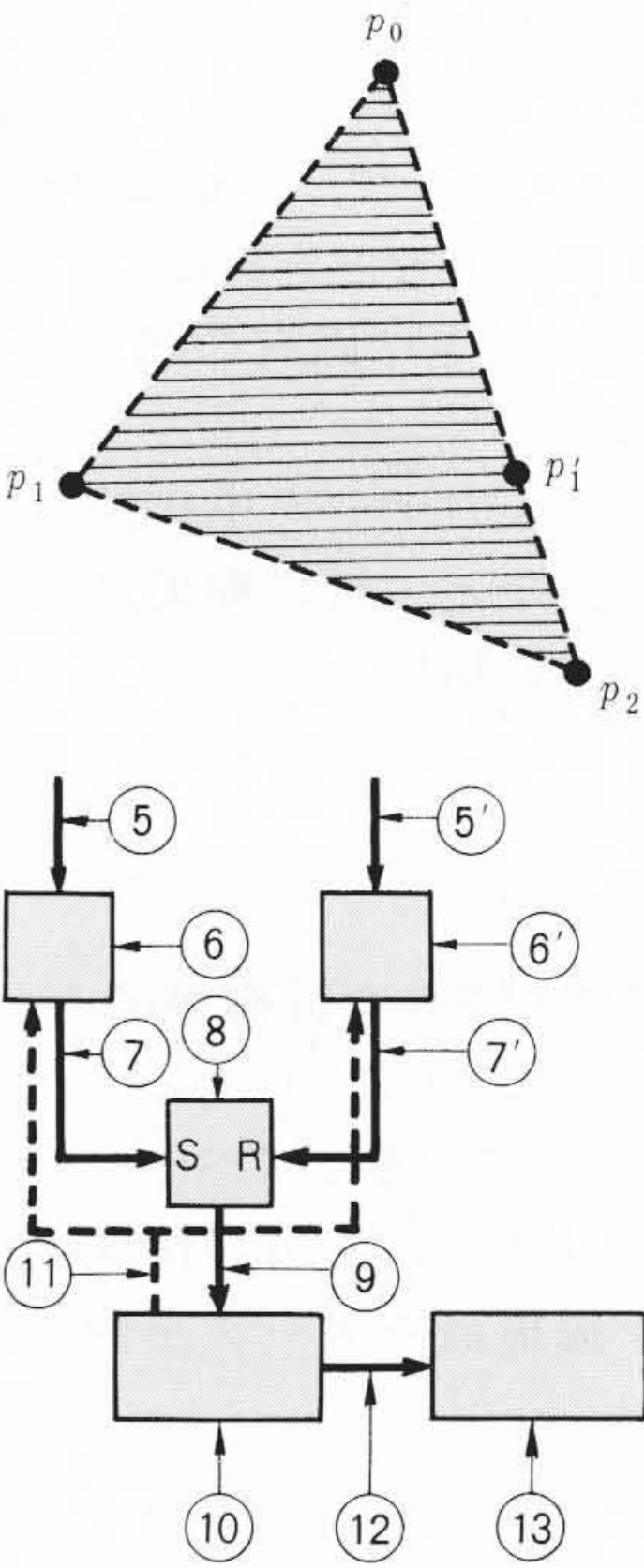


図1 塗り
つぶし図形
の一例

図2 表示
装置のプロ
ック構成図