

# 文書管理システムにおけるイメージ処理技術

## Image Processing Techniques for Document Management System

最近、オフィス業務の生産性向上に対する関心が高まり、文書の作成、保管、編集などの文書管理機能の強化とシステム化が強く望まれている。

田畑邦晃\* Kuniaki Tabata

田村秋雄\*\* Akio Tamura

このニーズにこたえ、日立製作所は、イメージ処理技術を応用し、各種の文書情報の保管・検索・編集・複写・伝送が可能な高機能の文書管理システムを試作した。本システムは、英・数字、仮名、漢字などのコードデータに加えて、地図、図面、印影、手書文字などのイメージデータが処理できる。また、対話形式による文書編集とフォーマット定義機能、マルチウインドウ表示方式の導入などにより、マンマシン性を強化するとともに、文書イメージ編集の基礎となる切出合成技術を開発した。

以下、本稿では、イメージ編集処理を中心に文書管理システムの概要と応用例について報告する。

### 1 緒言

近年、オフィス業務の生産性向上に対する関心が高まり、文書の作成、保管、検索、編集、複写、伝送などのシステム化が強く望まれている。日常取り扱う文書情報には、英・数字、仮名、漢字コードなどの符号に変換できる情報と、地図、図面、印影、手書文字などのイメージデータが混在する。従来、オフィスでのコンピュータ化は、前者のいわゆるコードデータ処理を中心に進められてきたが、最近ではイメージデータを含む、より高度の文書管理機能が要求されるようになった<sup>1),2)</sup>。

このようなニーズを背景に、日立製作所は、イメージ処理技術を応用した新しい文書管理システムを試作した。本システムは、各種の文書情報の保管・検索・編集・複写・伝送機能をもつ。その特長は、従来のコードデータに加えて、地図、図面、印影、手書文字などのイメージデータが取り扱える点にある。イメージデータは、切出、合成、拡大、縮小、圧縮、伸張などを施すが、このうち切出、合成とは、文書上の矩形領域に関するイメージデータ(部分図)をメモリ間で転送する処理を言う。すなわち、文書の中から署名、手書メモ、印影、写真などを切り抜き、逆にそれらを他の文書上に重畳するといった処理であり、各種の編集機能を実現するための基礎となる。A4判の書類をイメージデータに変換すれば、データ量は4Mビットにもなり、従来のコードデータ処理に比べて大量のデータを取り扱う必要がある。このため、処理の高速化が重要な課題になる。日立製作所は、文書イメージ編集に適したデータ転送のアドレス制御方式を開発し、この問題を解決した<sup>3)</sup>。

以下、本稿ではイメージ編集処理を中心に、文書管理システムの概要と応用例について報告する。

### 2 文書管理システムの機能と構成

#### 2.1 基本機能

文書管理システムの基本機能を図1に示す。主な機能は、(1)ハイブリッドファイル管理、(2)文書イメージ編集、(3)入出力処理と伝送(メイリング)などである。

##### (1) ハイブリッドファイル管理

本システムは、英・数字、仮名、漢字コードや図表、印影、手書文字などの各種の文書情報をコンピュータファイルに蓄積する。ファイル管理とは、データ登録や検索、追加、削除などを制御する機能を言う。コードデータとイメージデータは性質が異なる情報であるが、これらを統一的に管理するハイブリッド方式を特長とする。ファイル種別には、文書編集のレイアウト情報を格納するフォーマットファイルをはじめ、ドキュメントファイル、データファイルなどがある。

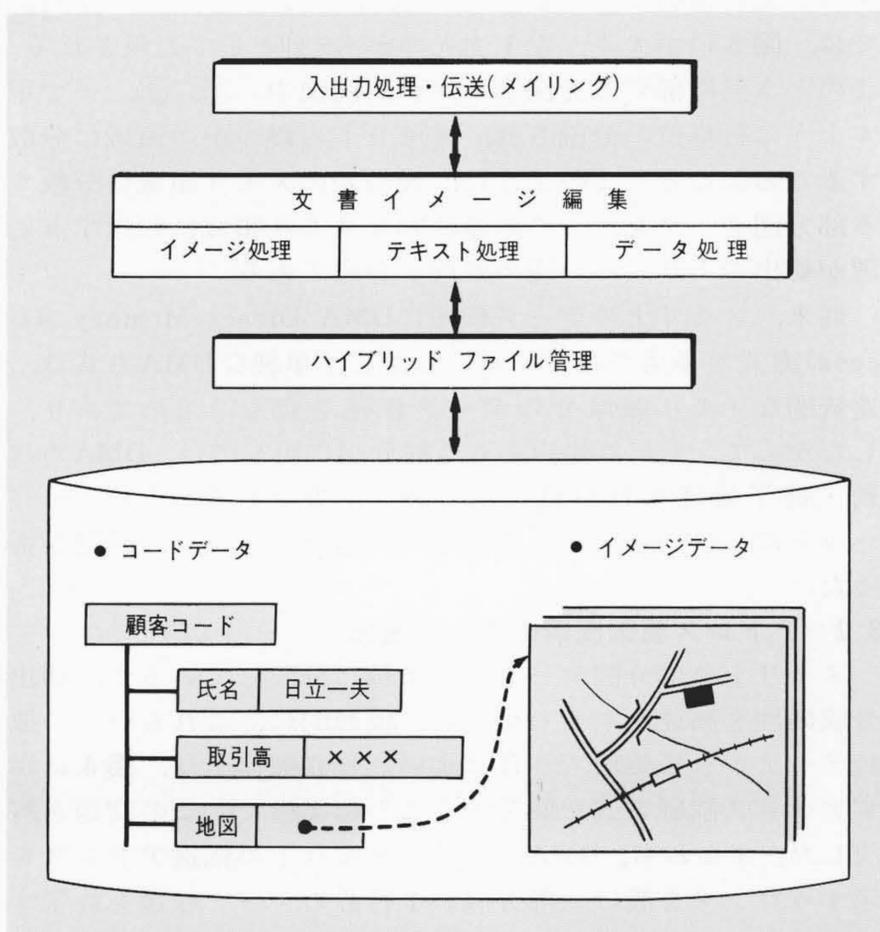


図1 文書管理システムの基本機能 文書管理システムは、英・数字、仮名、漢字コードや図表、印影、手書文字などの各種の文書情報をコンピュータファイルに蓄積し、データの登録、検索、追加、削除、編集、伝送などの機能をもつ。

\* 日立製作所システム開発研究所 \*\* 日立製作所OA事業部

(2) 文書イメージ編集

文書の作成、修正、転記などの機能に必要なイメージ処理、テキスト処理、データ処理を行なう。

地図、印影欄のような文書上のイメージデータを部分図と呼ぶ。文書情報をコンピュータに入力する場合、この部分図を単位に切り出してファイルに格納する。逆にファイルの中から必要な部分図を選択し、数値、文字データや罫線などと合成することにより、目的の文書を作成することもできる。イメージ処理機能としては、この切出、合成のほか、拡大、縮小、圧縮、伸張などがある。

文書管理システムはイメージ処理に重点があるが、従来のワードプロセッサで知られているような、漢字を含むテキスト処理機能をもつ。また、計数データの更新、集計などの簡易データ処理機能を内蔵するとともに、ホストコンピュータ接続のワークステーションとしても利用できる。

(3) 入出力処理と伝送(メイリング)

文書情報をファクシミリで読み取り、イメージデータに変換してコンピュータに入力する。検索用キーデータや仮名、漢字コードなどは、キーボードやタブレットから入力する。また、データの登録、検索、編集などの指示は、イメージディスプレイ上で対話形式により行なう。ディスプレイ上には、上記のイメージ処理・テキスト処理・データ処理の出力結果を同一画面に合成表示できる。また、ディスプレイ画面を「ウィンドウ」と呼ぶ幾つかの表示領域に分割、管理することにより、複数の文書情報を比較照合あるいは転記編集する操作を容易にした(マルチウィンドウ表示方式)。ファクシミリは文書情報の入出力装置であるが、これを用いたメイリングも行なえる。

2.2 システム構成

試作したシステムの構成を図2に示す。16ビットマイクロプロセッサを用い、分散処理形のコンパクトなシステムとした。端末装置は、キーボード(英・数字、仮名)、イメージディスプレイ、ファクシミリ、レーザプリンタ、タブレットなどから成る。ファクシミリ(G III規格、線密度: 8本/mm)はイメージ入出力装置であり、本システムの出力情報を回線経由で遠隔地に伝送することもできる。タブレットは、手書きの文字や図形のオンライン入力用であり、また、漢字入力端末(ペンタッチ)としても用いている。漢字データは、タブレット方式以外に、キーボードからの仮名コード指定の方法も並用した。

イメージ処理部は、切出、合成、拡大、縮小、圧縮、伸張などの機能をもち、1Mバイトのイメージメモリを実装する。

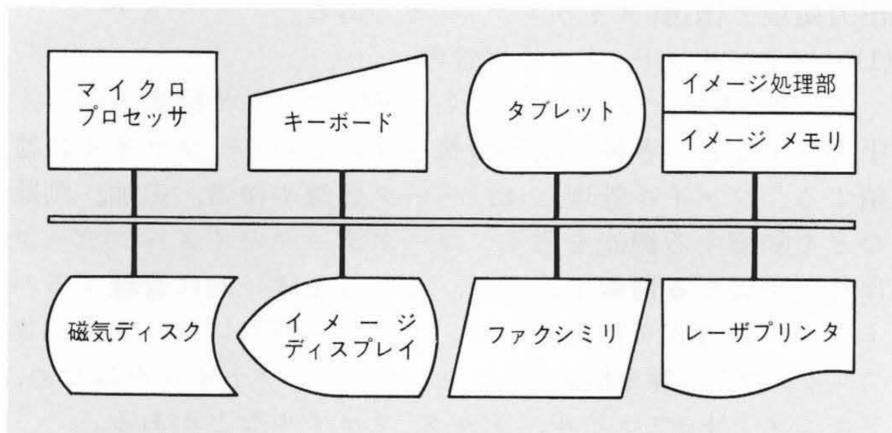


図2 文書管理システムの構成 本システムは、マイクロプロセッサ、イメージ処理部を中心に、キーボード、イメージディスプレイ、ファクシミリ、レーザプリンタ、タブレットなどの端末装置から成る。

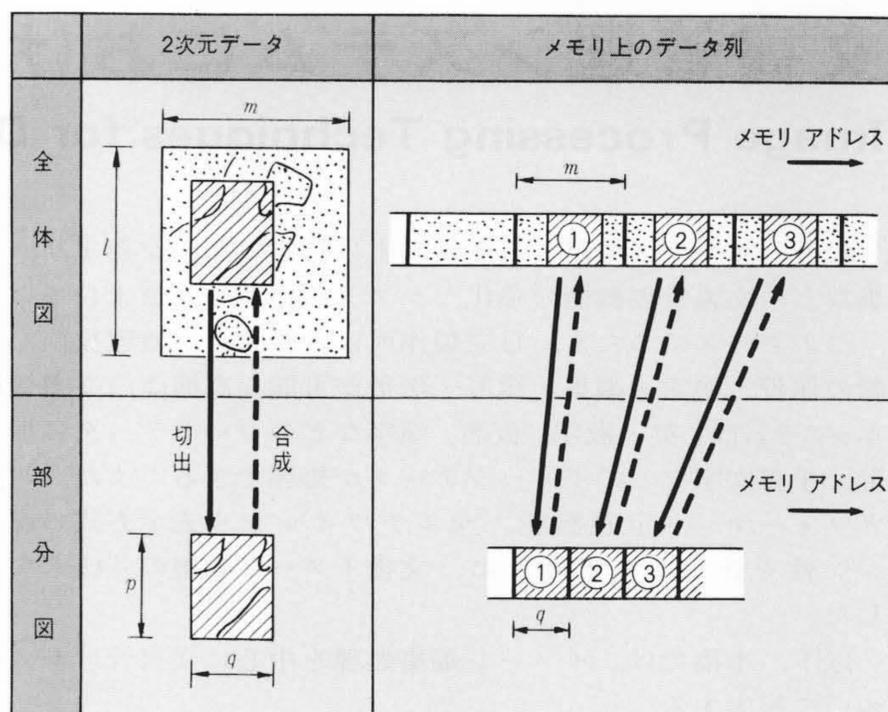


図3 切出合成処理とメモリ上のデータ列 文書情報は2次元の広がりをもつが、コンピュータ内部では1次元のデータ列として表現される。メモリ上に分散する部分図データ(斜線部)を、一つの連続したメモリ領域に転送する処理が切出であり、その逆の処理が合成である。

このイメージメモリは、切出、合成を高速化するためのアドレス制御機構を備えている。以下、高速切出合成処理の原理と応用例を重点に説明する。

3 文書イメージ編集における切出合成処理

3.1 技術課題

文書情報をコンピュータに入力する際、地図、印影欄などの部分図を切り出し、逆に、文書作成・編集時にこれを合成する。文書情報は2次元の広がりをもつが、コンピュータ内部では、図3に示すような1次元のデータ列として表現される。このとき斜線部の部分図データは、同図中、①、②、…で示すように行単位に分割され、メモリ上の幾つかの領域に分散することになる。このように、幾つかのメモリ領域に分散する部分図データを、一つの連続的なメモリ領域に転送する処理が切出であり、その逆の処理が合成である。

従来、メモリ上のデータ転送にDMA(Direct Memory Access)方式が多用されている。しかし、単純なDMA方式は、連続的なメモリ領域でのデータ転送を行なうものであり、したがって、上記の制約をもつ部分図に対しては、DMAの起動・終了処理を行単位に繰り返す必要がある。そのオーバーヘッドにより、実効的な処理速度が低下するという問題があった。

3.2 アドレス制御機構による高速化

メモリ上の部分図データは行単位に分散しているが、切出合成処理を高速化するためには、仮想的に、これを一つの連続データとして処理できればよい。日立製作所は、図4に示すアドレス制御方式を開発することによって、この問題を解決した。すなわち、DMA転送時のメモリ上の転送アドレスを示すカウンタを設け、部分図の1行分のデータ転送を終了するたびに、このアドレスカウンタが次行データのアドレスを指すように、アドレスカウンタの内容をスキップさせるものである。同図の加算器がこのスキップ処理を行ない、スキップ幅Iは、全体図の横幅mと部分図の横幅qの差に等しい。同図中のデータカウンタは、qで初期化され、データ転送のタイミングパルスが内容をデクレメントする。データカウンタがゼロになれば、qレジスタの内容を再ロードし、次行の

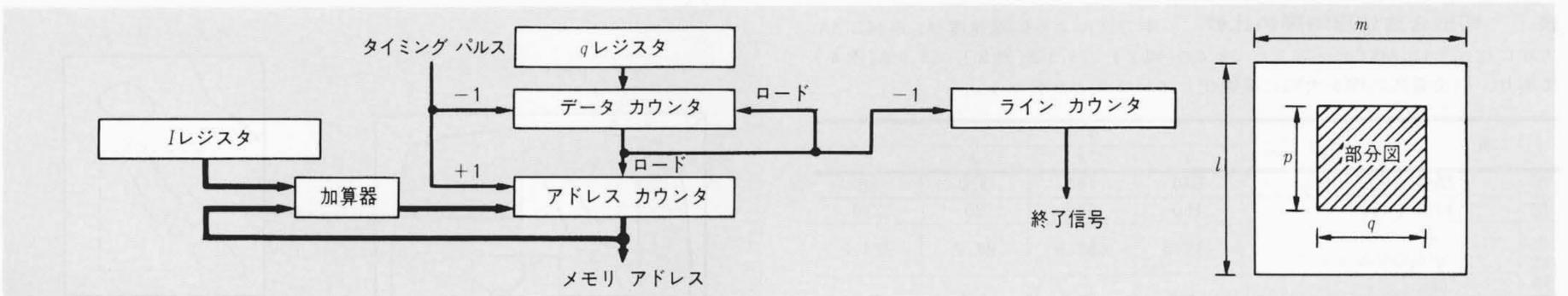


図4 切出合成処理高速化のためのメモリ アドレス制御方式 部分図の1行分のデータ転送を終了するたびに、アドレス カウンタが次行データの先頭アドレスを指すように、アドレス カウンタの内容を自動的にスキップさせる。このアドレス制御をハードウェアで実現することにより、切出合成処理を高速化している。

データ転送に移る。この処理を $p$ 回(部分図の行数回)繰り返し、ライン カウンタがゼロになった時点で、当該部分図の切出合成を終了する。

切出合成処理を一般化すれば図5のデータ転送となり、転送元と転送先の双方に対して、アドレス スキップの制御を行えばよい。この機能を図6に示すメモリ制御装置によって実現した。本装置の各種レジスタあるいはカウンタに設定するパラメータは次のとおりである。

- (1) 転送元先頭アドレス( $A_1$ )
- (2) 転送元データ間隔( $I_1$ )
- (3) 転送先先頭アドレス( $A_2$ )
- (4) 転送先データ間隔( $I_2$ )
- (5) 部分図の横長さ( $q$ )
- (6) 部分図の縦長さ、すなわち行数( $p$ )

以上、概説したアドレス制御をハードウェアで実現しており、ソフトウェアによる起動・終了処理は、各部分図に対して1回行えばよい。このオーバーヘッド削減によって切出合成処理を高速化した。

### 3.3 処理時間の評価

上記のアドレス制御機構による効果を評価するため、切出合成の処理時間を、単純DMA方式のそれと比較した。対象とする部分図のサイズ(横長さと行数)として、4 ケースについて検討した。各ケースの条件と処理時間の実測結果は表1に示すとおりである。同表が示すように、本方式による処理速度は単純DMA方式のそれと比べて、11.8倍(ケース1)、38.4倍(同2)、18.3倍(同3)、55.9倍(同4)であり、切出合成処理を大幅に高速化していることが分かる。

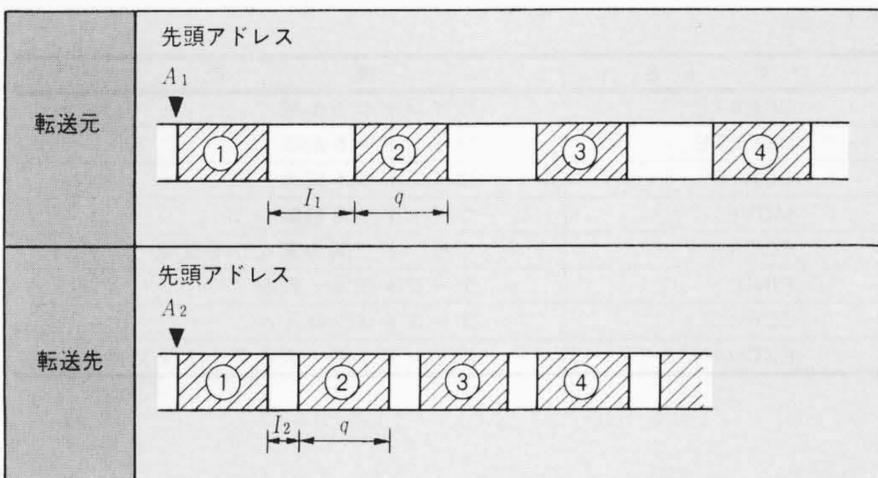


図5 切出合成処理におけるデータ転送 メモリ上の部分図データ(斜線部)は、一般に行単位に分割されている。転送元のデータ間隔( $I_1$ )がゼロの場合が合成であり、転送先のデータ間隔( $I_2$ )がゼロの場合が切出である。

### 4 応用例

文書管理システムでのイメージ処理技術の応用例として、(1)文書イメージ編集と(2)マルチウインドウ表示の画面制御について次に説明する。

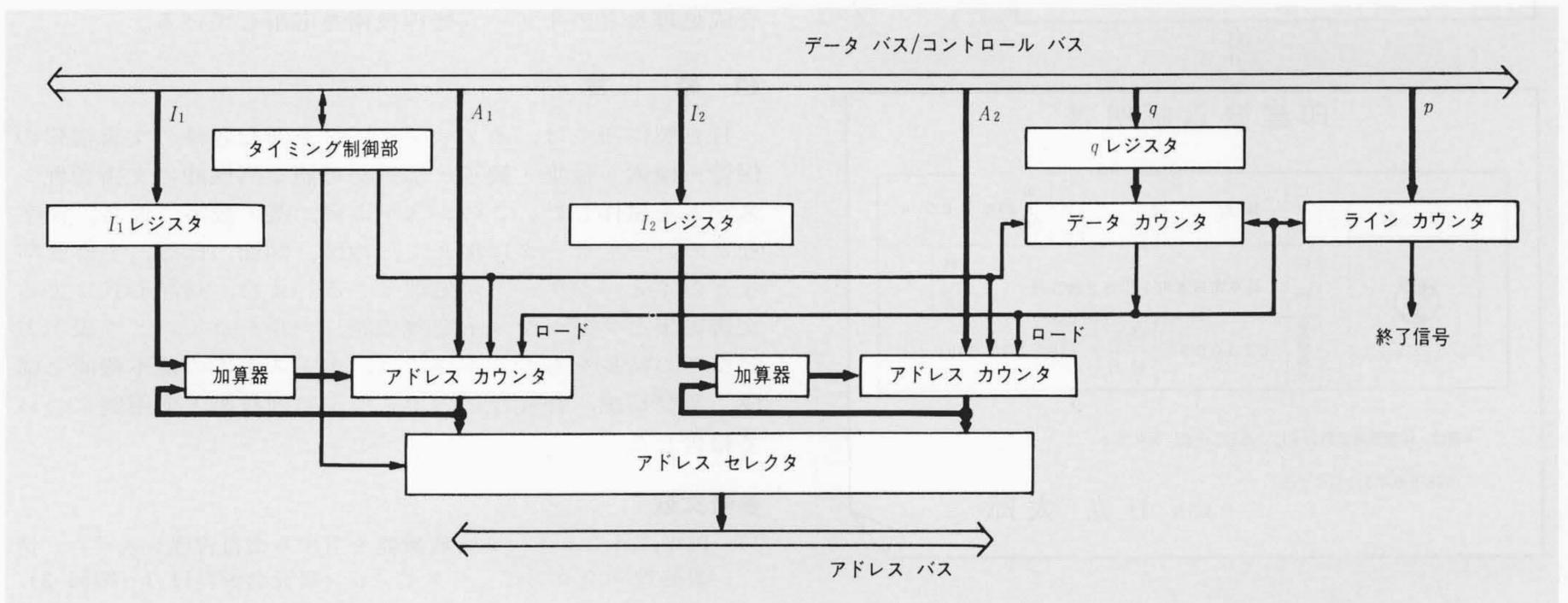


図6 切出合成機能をもつメモリ制御装置の構成 切出合成処理を高速化するために、データの転送元及び転送先のメモリ アドレスを制御している。本装置の各種レジスタあるいはカウンタに設定するパラメータは、転送元先頭アドレス( $A_1$ )、転送元データ間隔( $I_1$ )、転送先先頭アドレス( $A_2$ )、転送先データ間隔( $I_2$ )、部分図の横長さ( $q$ )と縦長さ、すなわち行数( $p$ )などである。

表1 切出合成処理時間の比較 本方式による処理速度は、単純DMA方式に比べて11.8倍(ケース1), 38.4倍(同2), 18.3倍(同3), 55.9倍(同4)であり、切出合成処理を大幅に高速化していることが分かる。

項目	ケース			
	1	2	3	4
部分図				
横長さ(ドット)	640	160	320	80
縦長さ(行)	160	640	80	320
処理時間	(A)			
	単純DMA(ms)	172.8	652.8	84.2
(B)				
	本方式(ms)	14.6	17.0	4.6
(A/B)				
処理時間比	11.8	38.4	18.3	55.9

注:略語説明 DMA(Direct Memory Access)

#### 4.1 文書イメージ編集

図7に示す印鑑証明書は、イメージデータを含む文書編集の一例である。同図中、氏名・住所ファイル(コードデータ)と印影ファイル(イメージデータ)を検索し、該当データをイメージメモリ上で編集する。編集のレイアウトや罫線、タイトルなどの書式を示すフォーマット情報は、あらかじめ定義しておく。本システムでは、対話形式によるフォーマット定義機能を備えているので、レイアウトの設定や修正を容易に行なえる。

#### 4.2 マルチウインドウ表示による画面制御

日常のオフィス業務では、複数の文書情報を同時に参照して、比較照合あるいは転記編集する機会が多い。この処理をディスプレイ上で実現するために、本システムでは、図8に示すマルチウインドウ方式を導入した。これは、ディスプレ

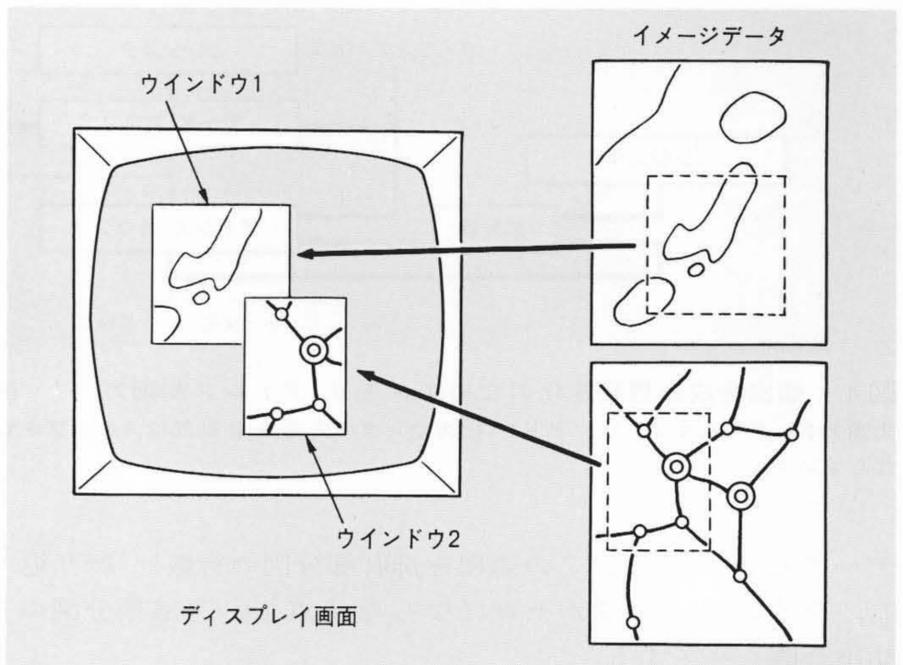


図8 マルチウインドウ方式による画面表示 ディスプレイ画面を、「ウインドウ」と呼ぶ表示領域に分割・管理しているので、複数の文書情報の比較・照合や転記編集が容易に行なえる。

表2 マルチウインドウ方式における画面制御コマンドの例 ウインドウの生成、削除、移動、スクロールなどの画面制御のためのコマンドを示す。

コマンド名	機能
CREATE	ウインドウを生成
DELETE	ウインドウを削除
SCALE	ウインドウを拡大・縮小
MOVE	ウインドウを移動
FLOAT	ウインドウ間の重なりを変更
FIND	データを検索・表示
SCROLL	データをスクロール
EXCHANGE	ウインドウ間で表示データを交換

イ画面を「ウインドウ」と呼ぶ表示領域に分割・管理する方式であり、同図の例では、二つのウインドウに、イメージデータの一部を切り出して表示している。ウインドウの生成、削除、移動、スクロールなどの画面制御は、対話形コマンドにより行なう(表2)。このマルチウインドウ方式には、切出合成処理などのイメージ処理技術を応用している。

### 5 結 言

日立製作所では、イメージデータを含む各種の文書情報の保管・検索・編集・複写・伝送が可能な高機能の文書管理システムを試作した。このシステムは、英・数字、仮名、漢字などのコードデータに加えて、地図、図面、印影、手書文字などのイメージデータが処理できる。また、対話形式による文書編集とフォーマット定義機能、マルチウインドウ表示方式などの特長をもつ。本稿では、本システムの基本機能と構成、及び切出、合成などのイメージ処理技術と応用例について報告した。

#### 参考文献

- 1) 田畑, 外: イメージ処理機能を有する文書管理システム, 情報処理学会コンピュータビジョン研究会資料11-1 (昭56-3)
- 2) 津原, 外: タブレットによる追記修正および手書メモ交換の原理実験, 情報処理学会第22回全国大会 (昭56-3)
- 3) 町田, 外: 文書管理システムにおける文書イメージ処理方式, 情報処理学会第23回全国大会 (昭56-10)

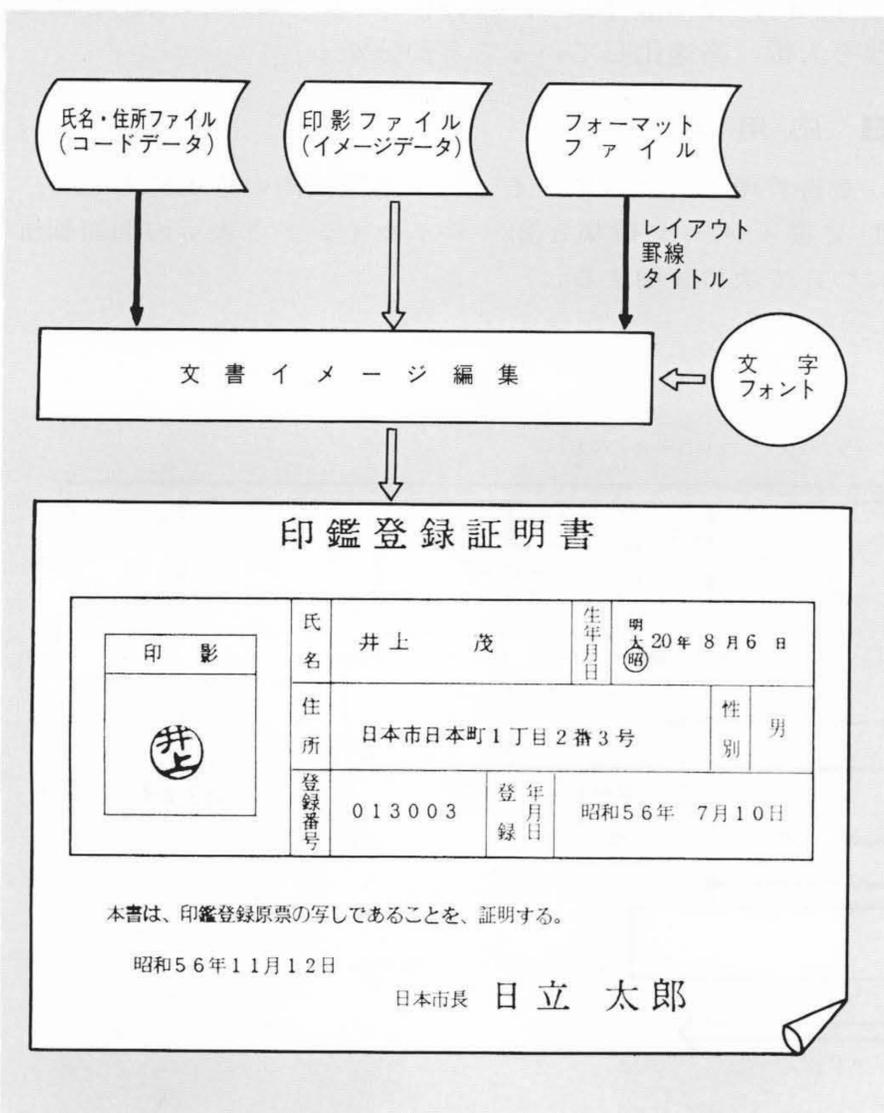


図7 イメージデータを含む文書編集の一例 氏名・住所ファイル(コードデータ)と印影ファイル(イメージデータ)を検索し、該当データを所定のフォーマットに従って編集することにより、印鑑登録証明書を作成している。