

統合ドキュメント処理システム (IDEOS)

一 構成と処理 一

宮井 均, 森下 丈, 首藤正道, 小野田勝洋

(日本電気株式会社 C & Cシステム研究所)

1. まえがき

従来よりオフィス業務の自動化, 省力化が強くいわれており, 昨今の情報処理市場においても, 既にワードプロセッサをはじめとしたオフィスオートメーション(OA)関連製品が多々発表されている。それに伴って従来あい昧だった事務分野からのニーズも, 次第に明確化してきている。

数あるオフィス業務の中で, ドキュメントの作成, 管理に関するものが, 大きな比重を占めるといわれている。事実日本語ワードプロセッサの登場により, ドキュメントは, 以前より効率良く且つ容易に作成, 管理できるようになった。

しかし依然として問題点が残されている。

◎ 高品質ドキュメントの作成

事務分野で用いられるドキュメントの中で, 従来でのEDPSで処理不可能あるいは不十分であった, より広範囲なマルチメディアとしてのドキュメントを作成する必要がある。

◎ ドキュメント作成方式

ドキュメントの作成(編集)操作手順が簡便である必要がある。

◎ 日本語入力方式と入力速度

素人には入力し易さ, 玄人には入力速度を考慮する必要がある。

以上のような観点から, 筆者等はイメージ情報処理システム^[1]等のドキュメント作成関連システムを作成してきた。現在さらに多種のデータを含み, 且つより複雑なフォーマットをもつ事務用ドキュメントを作成することを目的とした統合ドキュメント処理システム(IDEOS)を構築中である。今回その構成と処理について報告する。

2. ハードウェア構成 [2][3]

図1に, ハードウェア構成を示す。

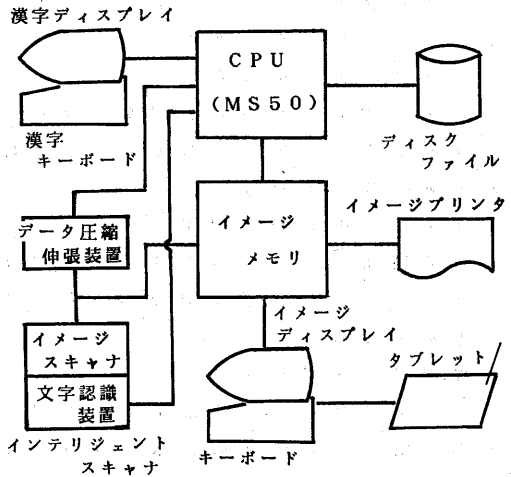


図1

a. インテリジェントスキャナ

文字(英数カナ)認識機能とイメージ入力機能を有し, 入力帳票上の英数カナ文字をコードデータとして, イメージをラスタスキャンドットデータとして, 同時にミニコンピュータに入力することができる。

b. イメージディスプレイ

ドキュメントの編集合成を行なうための大容量ページメモリを備えた高解像度ディスプレイである。特徴を以下に示す。

◎ メモリ容量

2048 X 4096ビット/ページのメモリを2枚(A, B)

◎ メモリ間の重ね合わせ

A, B両メモリの内容の重ね合わせ表示。重ね合わせは, AND, OR, XORである。

◎ 1/2倍縮小表示

隣接4ドットの論理和による縮小表示。

◎ ドット単位のアドレッシング

◎ イメージメモリの入出力

イメージメモリへの入出力は、ディスプレイのリフレッシュの間で行なっている。

◎ カーソルジュネレータ

c. イメージプリンタ

イメージメモリのハードコピーをとるレーザプリンタである。

d. データ圧縮伸張装置

イメージデータの容量削減のため、ファクシミリで用いられるデータ圧縮伸張を行なっている。

e. イメージ拡大縮小回路^[4]

9分割法^[5]、ニアレストネイバ法によるイメージデータの拡大縮小を行なうことができる。

f. 漢字ディスプレイ

コードリフレッシュ漢字ディスプレイで、漢字テキストの表示、イメージディスプレイの補助表示に用いる。

g. その他

操作者との対話用として

○ タブレット

○ キーボード

等が用意されている。

3. データの種類^[6]

ドキュメントは、一般に複数種のデータで構成されている。以下のように分類される。

画像 (イメージ)

パターン

図形 (ベクトル)

グラフ

テキスト

英数カナ漢字

画像は、手書き文字、図形等の2値画像 (イメージと呼ぶ) で、一般に大容量のためデータ圧縮の必要がある。

パターンは、小容量イメージであり、漢字フォントパターン等の他、ユーザ作成パターンも含める。

図形は、一般に座標点列によって構成されるベクトルであり、表等を構成する

けい線も含めて考える。

グラフは、円、棒、折線グラフ等のビジネス用のものをさす。

テキストは、ワードプロセッサの対象となる文章 (漢字コード列) を考える。

英数カナ漢字は、いわゆるEDP処理で用いられてきたデータであり、英字、数字、カナの8ビット系データと16ビット系の漢字データをさす。

IDEOSでは、英数カナと漢字は、別に扱う。

4. システム機能

IDEOSのもつ主な機能について概説する。

a. 日本語入力機能

各種日本語入力方式の試行

◎ 連想記憶方式

◎ 表示選択方式

◎ カナ漢字変換方式

(音声入力等)

中)

b. 帳票入力機能

◎ イメージ、英数カナの混在した帳票の同時入力、ファイル機能

◎ 入力帳票フォーマット定義機能

c. テキスト入力校正機能

◎ 日本語ワードプロセッシング機能

d. データ作成/登録機能

◎ 各種データ作成/登録機能

◎ 編集データ作成/登録機能

(領域定義、データ指定等)

◎ タブレットによる手書き文字、図形入力機能

e. コマンド入力機能

◎ キーボード

◎ タブレット

◎ 音声認識装置等

f. ファイル管理機能

◎ 各種データを統一的に管理する機能 (ファイル生成、初期化、拡張、アクセス等)

g. ドキュメント編集出力機能

◎ 各種データを編集テーブルに基づい

て編集出力する機能

(英数カナに関しては、語変換機能、テキストデータには、日本語ワードプロセッサのフォーマット機能を用いられる)

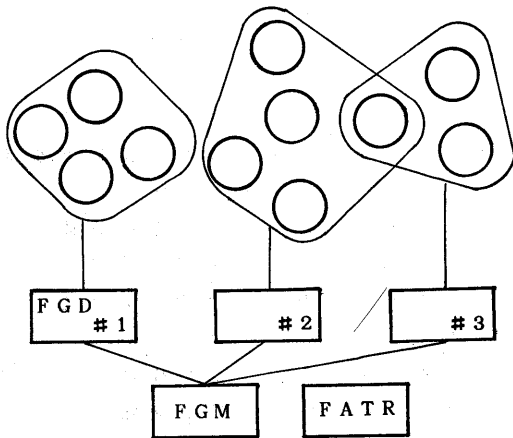
h. その他

◎漢字フォントは、4種類用意されている。(18, 24, 30, 40)

◎日本語入力のための、辞書、テーブル類をもつ

5. ファイル管理サブシステム[7]

既述のような多種多様なデータを統一に取り扱うために、ファイル管理サブシステム(VFAS)では、ファイル群という概念を用いて管理している。利用者は、必要なデータファイルをグループ化して論理体系の中でアクセスすることができる。(図2)



○: データファイル

図2

a. FGMファイル

ファイル群の存在を記述したもので、システム全体のファイル管理を行なうシステムファイルである。ファイル群番号とFGDファイルとの論理結合を行なう。

b. FGDファイル

利用者定義のファイル群に含まれる各データファイルを管理するもので、ファイル群毎に作成される。各データファイルの論理定義を行なっている。

c. FATRファイル

各データファイルの属性(ファイル編成、サイズ等)を記述したものでデータファイルの生成、拡張等に用いられる。

d. システム構成

VFASは、大別するとデータファイルへのアクセス系(VFIAS)と生成系(VFICS)により構成される。(図3)

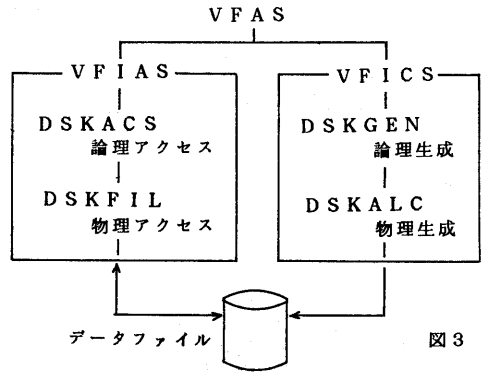


図3

e. ファイル形式

FGM, FGD, FATRファイルは順編成、各データファイルは区分順編成ファイル(図4)である。

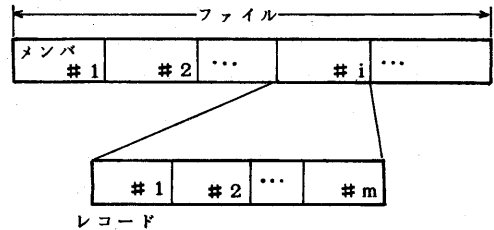


図4

f. アクセス

各データファイルは、FGDファイルにより管理されている論理ファイル名と、指定メンバ名によりアクセスすることができる。アクセスは、制御ブロック(FACB)を介して行なう。

6. 帳票入力サブシステム

帳票上のイメージデータと英数カナデータは、インデリジェントスキャナにより入力される。図5に入力帳票の読取情

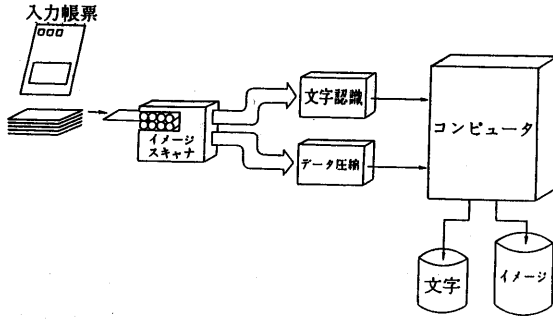


図5

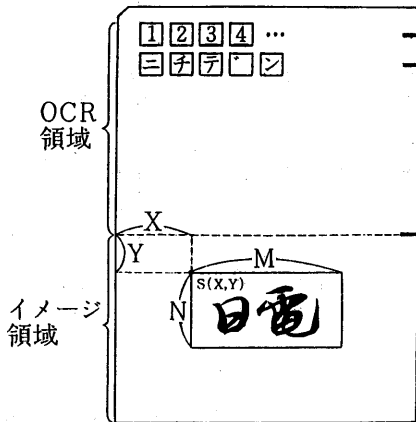


図6

報の流れを、また図6に典型的な入力帳票のフォーマットを示す。

OCR領域に書かれた英数カナ文字は文字認識装置により認識されコード情報として、またイメージ領域に描かれたイメージは、データ圧縮装置により圧縮イメージデータとして、各々コンピュータに入力される。

入力帳票のフォーマットは、あらかじめ設定することができる。特にイメージデータに関して切出位置、サイズ等が必要である。

帳票によるイメージ、英数カナ同時入力は、ファイルされたイメージデータを検索する際に、同時入力した英数カナデータを用いることができる利点がある。当然のことながらイメージスキャナあるいはOCRとして、単独で使うこともできる。

7. NWPサブシステム

NWP(日本語ワードプロセッサ)は大別してエディタ部とフォーマッタ部に分かれており、各々入力校正、編集出力機能をもつ。前述の日本語入力機能により、漢字ディスプレイを用いてテキストの入力校正を行なうことができる(エディタ部)。フォーマッタ部は、ドキュメント編集出力機能におけるテキスト出力に用いられる。

8. データ作成/登録サブシステム

a. イメージ

前述の帳票入力機能を用いて、英数カナと対あるいは単独で入力、ファイルすることができる。

b. パターン

現在漢字パターンファイルとして利用している。

c. ベクトル

座標点列を数値で与えることによりベクトルデータを登録することができる。またタブレットを用いた手書き文字、図形等も同様に扱うことができる。

d. グラフ

グラフ生成に必要な数字データを登録するだけでよい。後述の英数カナデータとしての扱いになる。

e. テキスト

前述のNWPサブシステムのエディタ部により作成される。

f. 英数カナ漢字

従来のEDPSと同じ扱いで作成/登録することができる。通常のデータファイルの他、語変換(英数カナを編集出力時に漢字に変換する)用のファイルも作成する。

9. コマンド入力サブシステム

データ入力用としても用いられるが特に会話型操作を想定した複合入力化を図っている。キーボード、タブレット(カーソル付)の他、音声認識装置、タッチディスプレイの利用も考えている。

10. ドキュメント編集サブシステム

a. ドキュメントの定義

ドキュメントは、

◎ドキュメントを構成する各種データ

◎各種データの編集機能

によって一意的に決定される。

I D E O Sでは、ドキュメントのフォーマットはV F A Sファイルの1つである編集テーブルによって規定される。即ちドキュメントは、上記各種データと編集テーブルの論理的結合によって記述されることになる。このようにすることによって同一データを用いても、編集テーブルを変更することで異なる編集結果のドキュメントを作成することができる。

図7にドキュメントの概念図を示す。

ドキュメント

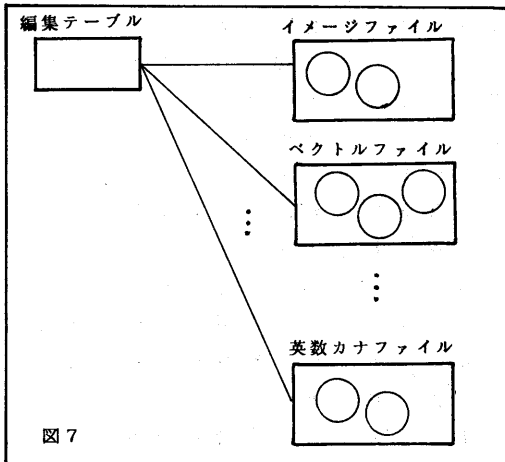


図7

b. ドキュメントの構成

ドキュメントは、編集を容易にするために図8に示すような論理構造をもたせている。

◎ページ

1つのドキュメントは、複数のページに分かれる。

◎サブページ

一般にページ上には、種々のデータが個々の属性および編集情報にしたがって編集されるが、ページを基準としてレイアウトすることは、位置の把握が厄介になる。そこでページ内ページ

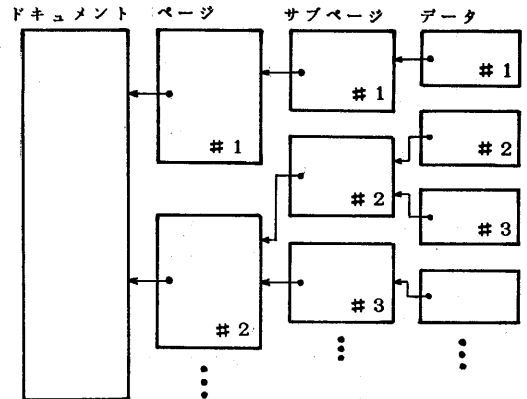


図8

(サブページ)を設け編集作業の便宜を図った。したがって編集のための座標系には、サブページレイアウトのためのページ内座標系と、データレイアウトのためのサブページ内座標系の2つが存在する。またこのようにすることにより、ページ内のサブページレイアウトを変更するだけで、各データレイアウトを変更する必要がない利点もある。

c. 編集テーブル

編集テーブルは、下記の記述子により構成される。

◎ドキュメント記述子

◎ページ記述子

◎サブページ記述子

◎データ記述子

ドキュメントの編集情報は、上記記述子を用いることにより、図9に示すような一般形で表わされる。以下各記述子の内容(属性)を列挙する。

1. ドキュメント記述子

○ドキュメント識別子

○メンバー名 他

2. ページ記述子

○ページ識別子

○ページ領域

○用紙種別

○フォント、行間、字間 他

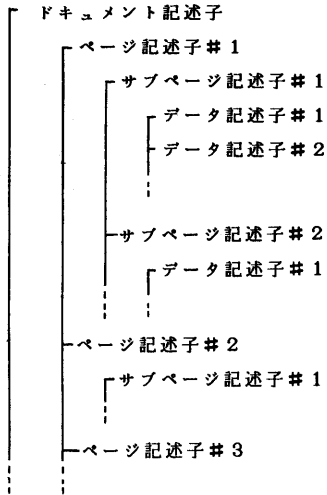


図9

3. サブページ記述子

- サブページ識別子
- サブページ領域
- フォント，行間，字間 他

4. データ記述子

データの種類に応じて6種類の記述子が存在する

4-1. イメージ記述子

- イメージ識別子
- イメージ領域(拡大，縮小)
- 切出位置 他

4-2. ベクトル記述子

- ベクトル識別子
- 描画開始位置 他

4-3. グラフ記述子

- グラフ識別子
- 原点座標
- グラフ種別
- 換算値 他

4-4. テキスト記述子

- テキスト識別子
- フォント，行間，字間 他

4-5. 英数カナ記述子

- 英数カナ識別子
- 編集位置
- 語変換指定
- フォント，行間，字間 他

4-6. 漢字記述子

- 漢字識別子
- 編集位置
- フォント，行間，字間 他

d. 編集処理

編集処理は，上記編集テーブルの記述にしたがって行なわれる。

編集テーブルの中には，原則としてデータファイル名(メンバー名)は記述されない。これは同型フォーマットで，大量ドキュメントを作成する伝票発行業務等を考慮したためである。現在データファイル名は，別途テーブルの形で与えている。

編集テーブルに対する変更，即ちドキュメントのレイアウト変更には，現在タブレットを用いた試みを行なっている。

11. 出力実験と考察

I D E O S が稼動しはじめたことにより，いくつかの実験が行なえるようになった。以下に例を示す。

a. 例1

各種製造業，サービス業並びに広く事務分野で使われるドキュメントは既に定型化されているものが多い。本報告では，企業内業務への応用として欄卸業務を行なった事例^[8]を紹介する。

図10に示すようなイメージ記入領域と英数カナ記入領域の両方をもつ入力帳票を用いた。図の例では1品1葉の帳票であるが，記入者の便宜を図って6品1葉の連記帳票も用意した。

出力帳票としては，図11に示すような処理ドキュメントを，21種類作成した。

処理内容としては，EDP処理(ソート，更新等)に重きがあり，イメージ，ベクトル(けい線)に関する処理は付加的なものである。

b. 例2

オフィスでは，多種データによる複雑なドキュメント即ちレイアウトに重きをおいたものを考える必要がある。

2. 保管品現在数調(期中)

昭和 36 年 7 月 31 日現在

②

研 No. 002443

製造番号	品名	購入又は完成年月	等級	注文数	入庫数	タッグカード NO.	仕掛残高
31	使用中						

数量(1)	金額(1)	数量(2)	金額(2)	数量(3)	金額(3)
2					

数量(1)	金額(1)	数量(2)	金額(2)	数量(3)	金額(3)
31					

(記入コード説明)
 1 完成品
 2 未加工品

(記入例) ① 前回の繰上りの繰下りの記入

図 1 0

図 1 2 には、これらを想定したドキュメントの作成例を示した。マルチメディアとしてのドキュメントが作成できることがわかる。

このような例の場合、表現形式、レイアウトの変更が容易に行なえることが望まれる。即ちグラフの種類、位置関係等の変更し易さを考慮する必要がある。

1 2. あとがき

稼働をはじめた IDEOS の構成と処理について報告した。

今後、試行実験と並行した研究開発が必要なことはいうまでもないが、まえがきで述べたような問題点等についてさらに解決していく予定である。特に、編集における会話方式を重点と考えている。最後に、本研究の機会を与えていただいた荻原周辺機器研究部長、並びにシステム作成、評価に御尽力いただいた C & C ビジネスシステム推進本部柳川氏に感謝致します。

参考文献

1. 宮井他：「漢字、イメージを含む帳票処理オフィスシステムの実験」情処 2 0 全大
2. 森下他：「イメージディスプレイを中心としたイメージ/テキスト入力システム」情処 2 1 全大
3. 森下他：「イメージを含むドキュメント処理システム」画像電子学会 - 1 2
4. 森下他：「9 分割法によるイメージ拡大縮小方式 H/W の試作」情処 2 2 全大
5. 宮井他：「イメージの拡大縮小方式」情処 2 0 全大
6. 宮井他：「統合ドキュメント処理システムの一構成法」情処 2 2 全大
7. 宮井他：「ドキュメント処理におけるファイル管理システム」5 6 年信学会情報システム部門大会
8. 宮井他：「イメージを含むドキュメント処理システムの企業内業務への応用実験」情処 2 3 全大

C20S-17 重点管理資産消化計画管理表 (56年度期中概報) 昭和36年 7月 31日 2頁

製造番号	品名	購入又は完成年月	等級	注文数	入庫数	タッグカード NO.	仕掛残高
817-60	電算マシンのセンター		3			002436	0
32	4	0	31	4	0	使用中	0
817-60	定盤(小)		3			002444	0
32	1	0	31	1	0	使用中	0
817-60	路線経歴明用表		3			002443	0
32	2	0	31	2	0	使用中	0

図 1 1

日本電気

[1] コンピュータと通信

55年度の全社売上高を8,900億円台に乗せ、56年度には1兆円の大台に乗せる可能性を示したのが、日本電気である。創立80周年の54年度にかかげた「C&C」のキャッチフレーズは、イコールNEC、NECといえばC&Cとハネ返ってくるまでになった。

☆C&C— C&Cをキャッチフレーズだけでなく、事業として推進するための組織「C&Cシステム本部」を55年7月に設立したが、56年5月には、これを発展的に解消して「C&C推進企画室」、「C&Cビジネスシステム推進本部」、「C&Cパブリックシステム推進本部」を新設した。C&C事業を進めるには、多くの事業グループを横断的に調整、支援、統括する必要があるため、その機能強化が狙いである。



コンピュータその他電子機器部門の対象商品

- 汎用コンピュータ (ACOSシリーズ)
- オフィスコンピュータ
- 制御用コンピュータ
- 各種汎用・専用ターミナル
- 日本語ワードプロセッサ
- 通信制御装置
- 大気汚染監視システム
- テレメータシステム
- 郵便自動化装置
- 数値制御装置
- 医用電子装置
- データ復旧装置
- 超音波装置

統合ドキュメント処理システム
 複雑な文書を作ることができ、システム
 C&Cシステムには、「C
 周辺機器Cパブリック
 &C事業を進め
 的に調整、支援、
 機能強化が狙いである。

この文書の内容は、雑誌「コンピュータピア」

日本電気のコンピュータその他電子機器部門売上高の推移

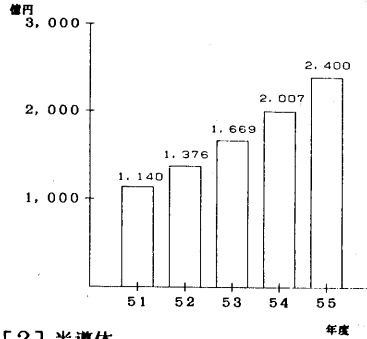


図12

この文書の内容は、雑誌「コンピュータピア」11月号より一部引用したものです。

[2] 半導体

わが国半導体トップメーカーの日本電気は、米データクエスト社の調査によれば、55年度のIC生産実績では、米国のテキサス・インスツルメンツ (TI) 社に次いで全世界で第2位、個別半導体を加えると、TI、モトローラ社の後、第3位を占めている。

日本電気の電子デバイス部門生産推移 (単位: 億円)

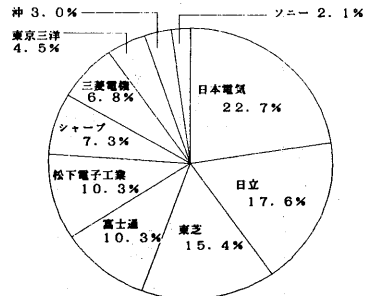
	54年度実績	55年度実績
IC	1,030	1,535
半導体	570	645
その他	505	630
合計	2,105	2,810

5. トップメーカーの日本電気は、米データクエスト社の調査によれば、55年度のIC生産実績では、米国のテキサス・インスツルメンツ (TI) 社に次いで全世界で第2位、個別半導体を加えると、TI、モトローラ社の後、第3位を占めている。

日本電気の電子デバイス部門生産推移 (単位: 億円)

	54年度実績	55年度実績	56年度実績
IC	1,030	1,535	1,900
半導体	570	645	750
その他	505	630	750
合計	2,105	2,810	3,400

トップ10社半導体生産規模別シェア分布 (56年度)



コンピュータその他電子機器部門の対象商品

- 汎用コンピュータ (ACOSシリーズ)
- オフィスコンピュータ
- 制御用コンピュータ
- 各種汎用・専用ターミナル
- 日本語ワードプロセッサ