

光ディスクファイルシステムを支える 情報処理技術の展望

Information Technology Supporting Optical Document Filing Systems

光ディスクファイルは、オフィスシステムの中で着実にその価値を発揮しつつある。多様なシステム形態で、広い応用が可能となるように登録の自動化、検索の高度化、ファイル対象の拡大、パーソナル化、さらにはネットワーク化などのための技術が開発されつつある。高精細な中間調画像のための画像処理技術や文書画像理解技術などにより、将来、カラー文書ファイリングや自動ファイリングなどの機能が実現されると思われる。同義語や異表記語を自動的に生成して、同時に高速検索する文字列照合技術などにより、検索の自由度はさらに増す。これらの情報処理技術と、それを支えるハードウェア技術およびシステム化技術の進展により、いっそう使いやすいシステムへと進歩していく。

藤澤浩道* *Hiromichi Fujisawa*
藤縄雅章** *Masaaki Fujinawa*
黒須康雄*** *Yasuo Kurosu*
和歌森文男**** *Fumio Wakamori*

1 緒言

オフィスでの情報の洪水はOA化によってさらに増している。電子的に作られる情報が増える一方、紙の形の文書の量もますます増加しているのが現状である。このような情報の氾濫(はんらん)をじょうずに管理し、有効な情報と化すために、光ディスクファイルシステムが活躍している。

当初、光ディスクファイルシステムは光ディスクの大容量記憶を生かしたシステムとして、スタンドアロン構成による利用から始まった。情報処理システムに対する期待が高まり、それと同時に光ディスクファイルの役割も拡大してきている。

システムの機能および性能もそれを支えるよう進歩している。しかし、利用範囲が拡大すると、さらに高機能化、高性能化が求められる。本論文では、このように期待が増している光ディスクファイルシステムを支える技術の現状と動向について、主に情報処理技術の観点から解説する。

2 光ディスクファイルの将来展望と技術課題

光ディスクファイルは、使い勝手の向上、ファイリング対象の拡大、大容量化と大規模化、小形化とパーソナル化およびネットワーク化とオープン化の方向に向かって発展していく。

(1) 使い勝手の向上

文書の「ファイリング」はオフィスでの日常的な作業であり、機械であることを意識させない存在になることが理想である。その意味で、使い勝手はきわめて重要な要素である。

文書の登録作業が簡便であること、検索が思ったとおりにすなおにできること、十分な応答速度があること、システム全体として省力化となっていること、しろうとにとっても操作しやすいことなどが求められている。

これらの課題はヒューマンインタフェースの問題ととらえることができるが、多くは機能と性能とに展開できる。なかでも、登録と検索に関する使い勝手については新規機能の開発が必要である。

(2) ファイリング対象の拡大

現在の光ディスクファイルは、文書を画像情報として大量に蓄積することが特徴である。扱える画像はモノクロームの二値画像が中心であるため、ファイリング対象は主に文字や図面から成る文書が中心である。

今後は、多値画像処理の強化により、イラストやカラー写真を含むカラー文書のファイリング機能へと発展することが期待されている。また、ワードプロセッサ(以下、ワープロと略す。)のテキストデータやCADで扱うベクトルデータ、さら

* 日立製作所 中央研究所 工学博士 ** 日立製作所 小田原工場 *** 日立製作所 マイクロエレクトロニクス機器開発研究所
**** 日立製作所 システム開発研究所 工学博士

には音声データのファイリング機能も実現されることが期待されている。

(3) 大容量化と大規模化

光ディスクは大容量性と可換性とに特徴があり、これを生かして、すでに大規模なファイルシステムが構成可能となっている。しかし、省スペースの観点から光ディスク1枚当たりの記憶容量の大容量化が、また検索性能の観点からはデータの読出し速度の高速化が継続して期待される。

現在、新形の12インチ追記形光ディスクは7Gバイトの記憶容量を持ち、約20万ページの文書(標準原稿)を格納可能にする容量に至っている。また、データ読出し速度は最大2.2 Mバイト/sと高速化されている。

(4) 小形化とパーソナル化

光ディスクファイルも小形化の時代を迎えており、デスクトップ形からラップトップ形へと展開しつつある。これにより、個人のファイルを手元に管理したり、あるいは自分の机で大規模な文書ファイルにアクセスしたり、登録したりすることが可能となる。このためには、光ディスクの媒体互換性の確保やネットワークの完備が不可欠である。

(5) ネットワーク化とオープン化

オフィスシステムがネットワーク化されつつあるが、光ディスクファイルもその中に統合化されつつある。今後は、すでに導入されているシステムとの結合を容易にすることがきわめて重要である。ネットワークを介したアクセスのためのプロトコルの標準化や、アプリケーションシステムを構築しやすくするためのミドルソフトウェアの高機能化によるオープン化が進んでいく。

3 登録技術

3.1 自動ファイリング技術

「ファイルする」ということばが、文書を適切に分類して格納するという意味であることからわかるように、文書の登録はファイリングそのものである。したがって、これを自動化する「自動ファイリング」というシステム概念²⁾に沿った技術開発が課題となっている。

このシステム概念は、スキャナから読み取った文書画像の内容を計算機に解析させて、文書名や著者名などの検索のための情報を自動的に取り出し、登録作業を自動化しようというものである(図1)。

(1) 画像品質の最適化

画像の鮮明さを判定して二値化レベルを最適化したり、写真などの多値画像のコントラストを最適設定する機能を実現することにより、スキャナのパラメータ設定のボタン操作をなくすことができる。また、入力文書の位置や傾きなどの検出と自動補正機能により、簡便な入力が可能となる³⁾。

(2) バッチ入力

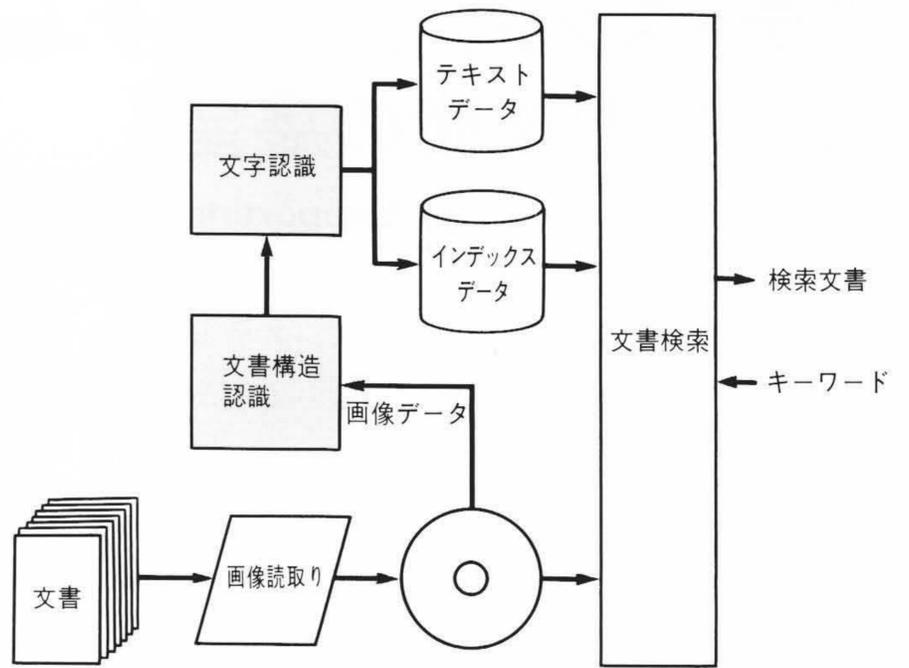


図1 自動ファイリングの概念図 文字認識・文書理解技術を用いて検索情報を抽出しながら、ファイルに自動登録する。

表形式のヘッダシートを準備し、登録のための所定事項を記入して登録すべき一連の文書をバッチ的に入力する方法が考えられる。ワープロなどで作った表の枠構造を画像解析によって抽出し、そこに記入されている文字を認識する技術が開発されている⁴⁾(図2)。現在、文字認識の技術は約2,500字種の手書き漢字が認識できるようになっている。

(3) 書誌情報の抽出と認識

検索情報としての表題、作成者名などの書誌的事項の存在する領域を表紙ページから同定し、文字を認識する。通常、大量の文書は定形のフォーマットを持っており、その構造をあらかじめシステムに知識として与えておくことにより、画像解析を行って重要項目を自動抽出することが可能となりつつある^{5), 6), 7), 8)}。

自動ファイリングは、認識の問題を解く必要があるため、ファイリング対象を特定しながら段階的に実現されていくものと考えられる。

3.2 中間調画像技術

中間調画像技術は、ファイル対象文書を広げるためには欠かせない技術である。

光ディスクファイルが扱う画像は、細かい文字や細い線を表現するため、200 dpi(dot per inch: 2.54 cm当たりの画素密度)ないしは400 dpiの解像度の二値画像である。A4サイズの画像データ量は、500 kバイト/ページないしは2 Mバイト/ページである。この解像度で多値画像を扱うと、256階調の場合、4 Mバイトないしは16 Mバイトにもなる。

したがって、写真などの多値画像は黒画素を空間的に分布させ、その分布密度で濃淡情報を表現するディザ法と呼ぶ方法をとる(図3)。このように多値画像を二値画像に変換することによってデータ量を抑えることができる。この画像はデ

申請書

下記の通り申請します

申請者	日立花子		
申請日	平成 3 年 / 月 5 日		
住所	東京都千代田区 神田駿河台 4丁目 6番地		
筆頭者氏名	日立太郎	続柄	妻
氏名	日立花子	通数	1 通

備考

(a) 読取り帳票

01	02		
03	04		
05	06		
	07		
08	09	10	11
12	13	14	15

枠構造認識結果表示

枠番号	項目名称	データ枠番号
01	申請者	[02]
03	申請日	[04]
05	住所	[06]
05	住所	[06]
08	筆頭者氏名	

(b) 認識結果

図2 表形式文書の理解 枠構造の抽出と文字の認識により、データの種類の記入データを対応づけて認識することができる。

ィザ画像または疑似中間調画像と呼ぶ。

文字などの二値領域の黒画素は、文字どおりその点が黒であることを表しているが、写真などの濃淡領域の黒画素は近傍での分布密度に情報がある。したがって、文字や図面の二値領域と写真などの濃淡領域とでは、異なる画像処理を行う必要がある。このため、従来は写真と文字が混在する文書の高画質入力が困難であったり、画像処理を行う際に写真領域を指定するという必要性があった。

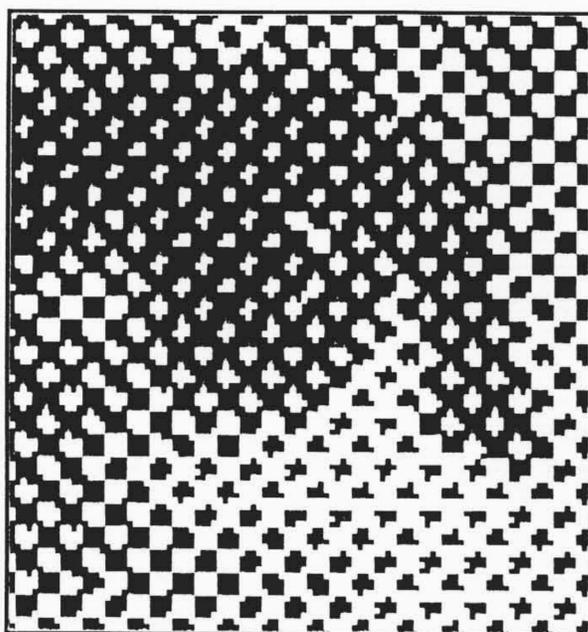
このような問題を解決するために、スキャナの中で文字領域と写真領域を自動的に識別して、文字領域は単純二値化、写真領域はディザ化により、適応的に二値化する方式⁹⁾が開発されている。この単純二値画像とディザ画像が混在する画像を適応形二値画像と呼ぶ。

この適応形二値画像は画質の問題を解決するが、従来の方
法で画素密度変換(拡大・縮小処理)を行うと、写真の画質が
低下するという問題がある。例えば、400 dpiで記録した画像
をCRTに表示するために、200 dpiに画素密度を変換すると、
写真画像にモアレが生じるという問題がある。

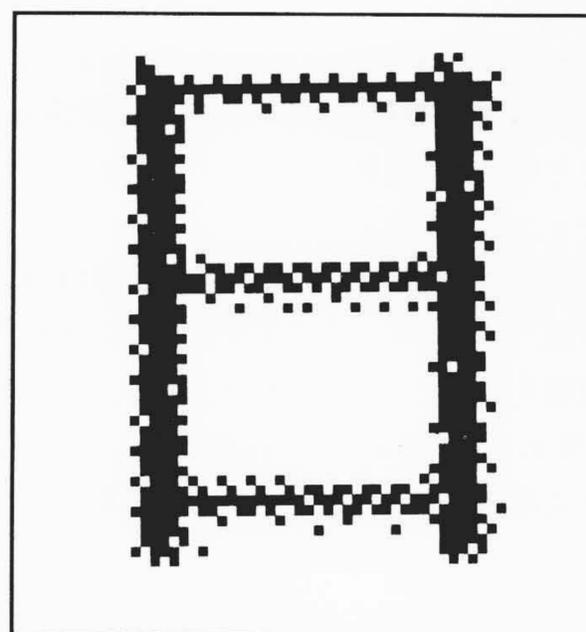
そこで、適応形二値画像の中のディザ領域を自動識別して、
ディザ領域に対しては濃度分布保存形のアフィン変換を行う
方式¹⁰⁾が開発されている。

3.3 カラー文書ファイリング技術

カラー画像の表現には、R・G・B(赤色・緑色・青色)の3
面から成る多値画像が必要である。したがって、従来はアナ
ログ系の技術を用いて、自然画像を対象とした静止画ファイ
ルが製品として開発されている。多くの場合、640×480画素

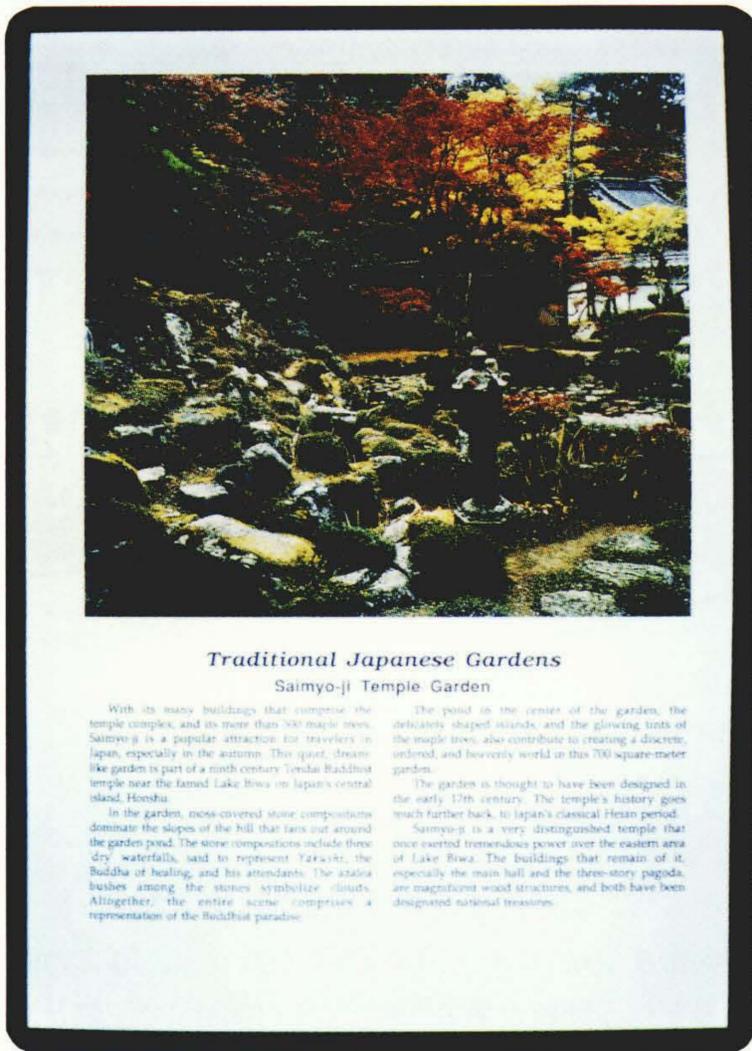


(a) 写真領域のディザ画像



(b) 文字領域のディザ画像

図3 ディザ画像の拡大図 濃淡情報は黒画素の分布密度で表現するが、文字領域にディザ処理を適用すると画像がぼける。



出典：“age of tomorrow”, 112(1989-9)

図4 高精細カラー文書画像表示 高精細カラーCRT(1,872×1,344画素)で、A4サイズのカラー文書画像(150 dpi)を実寸表示することができる。

程度の画像を扱っており、文書画像の表現は困難である。

前節で述べた中間調画像技術は、A4サイズのカラー文書画像を高精細なデジタル画像(1,728×2,400画素以上)として扱い、初めて現実的な規模でのカラー文書ファイリングの実現が可能になる(図4)。

この技術によるカラー文書ファイルは、システムの内部では画像を適応形二値画像の形式で扱うことが特徴となる。光ディスクに格納するときは、二値画像の標準MMR(Modified Modified READ方式：圧縮符号化)方式で圧縮して格納する。通常、ディザ画像をそのまま圧縮符号化すると、かえってデータ量が増えてしまうことがある。

この問題は、圧縮符号の長さと同データの長さとを比較し、圧縮効果がないときは、原データをそのまま符号列に挿入する非圧縮モードを実現することによって解決できる。この方式を最適化MMR符号化方式と呼ぶ。この技術により、約40%が写真であるカラー文書画像を約1/20にデータ圧縮することが可能である。

4 検索技術

4.1 同義語・異表記語検索技術

ファイリングシステムとしては、多様な検索方法を提供す

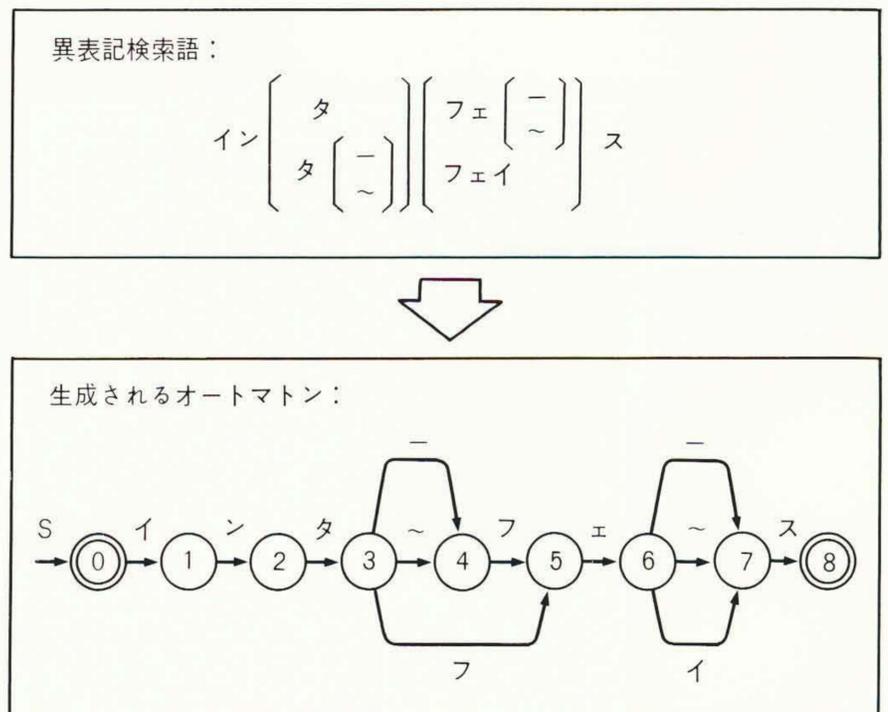
ることが望ましい。日常の文書の出し入れには簡単な分類形でよい。分類されたバインダの中は時系列的に管理されていれば、検索時は高速ページめくりですぐに見つかる。

検索機能の効果は、大量の文書の中から内容を指定して、検索するとき大きい。一般に、欲しい文書に付加したキーワードを正確に思い出すことはできないので、検索システムに投入する検索語は完全な単語ではない部分文字列であってもよいようになっている。この部分文字列と被探索文字列との照合処理に対しては、完全一致、前方一致、後方一致、中間一致、さらには一文字相違などの条件が指定できる。

また、複数の検索語を論理条件と組み合わせることができるので、OR条件により検索漏れがないようにしながら、AND条件によって検索ヒット件数を絞り込むことができる。

しかし、検索漏れがないように同じ意味の単語をすべて思い出すのは困難であるから、システムに辞書を持たせ、これらの単語を自動的に発生する方法¹¹⁾が考えられている。この辞書は、同義語、類義語、省略語、ニックネームなどを含む必要がある。通常の辞書とは異なり、どの表現からでも求める単語集合が出てくる必要がある。

同義語・異表記語の自動生成機能により、今後はきわめて多くの検索語を同時に照合する必要がある。有限オートマトン形の検索原理は、複数の検索語を一括して検索できる特徴がある。最近、特に片仮名の異表記を効率的に検索できるように、複数のトークンを同時に移動できるコンカレントステート形オートマトン方式¹²⁾が考案されている(図5)。このような検索方法をとる全文検索技術¹³⁾も将来重要である。



注：図中-はマイナス、～は片仮名の長音です。

図5 コンカレントステート形オートマトン方式 複数のトークンを同時に移動させることにより、異表記検索のためのオートマトンが簡単になる。

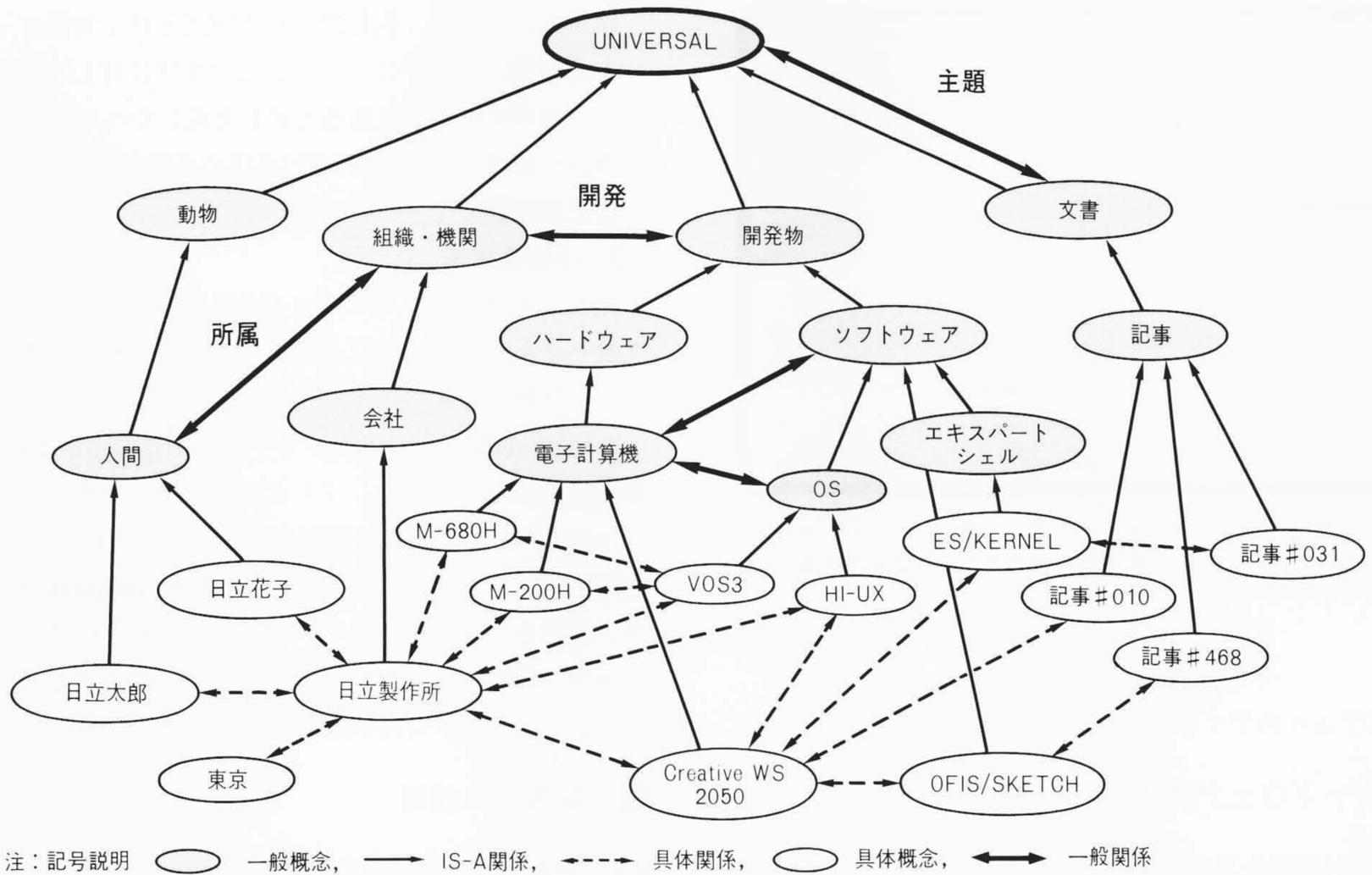


図6 概念ネットワーク 一般概念と一般関係によるネットワークで知識体系を記憶し、具体概念と具体関係によって個々のデータや情報を記憶する。

4.2 知的検索技術とビジュアルインタフェース

情報の分類と整理がファイリングの究極の目標¹⁴⁾であるとするならば、文書画像をキーワードで管理するファイリングシステムに、さらに知識ベースを持たせることが考えられる。単に文書画像を管理するだけでなく、扱う情報の世界の重要な概念(単語)を体系的に分類して、知識ベースに蓄える。われわれは忘れやすいので、単に文書を格納するだけでなく、関連する事実をもいっしょに計算機に格納することにより、いろいろな角度からの検索が可能となる。

概念ネットワークを用いた知的検索技術は、このニーズにこたえるものである¹⁴⁾。概念ネットワークは、図6に示すように、概念ノードと関係リンクとで構成する意味ネットワークである。

従来の情報検索システムでは、文書に付加したキーワードを用いた論理条件(ANDやORなどの条件)で検索要求を記述するが、そのとおりのキーワードを思い出すことが難しい。

概念ネットワークを用いる知的検索システムでは、

(関西)にある(会社)が開発した

(光ディスク)を用いた(検索システム)

のように意味的な関係で結び付けられた複数の単語による検索文が受け付けられるようになる。おぼろげな記憶からの検索では、対象物の記述がこのように上位語を用いた説明的な表現になることはよく経験することである。システムは、こ

の説明的な検索文から分類形の推論を用いて具体的な情報を検索することができる。

知的検索システムで欠かせない機能は、記憶されている情報や知識を拾い読みするためのブラウジング機能である。陽に思い出すのは難しいが、見ればわかるという「記憶の想起と認識」の関係は心理学的によく知られている人間の記憶の特性である。これからのシステムは、この点を勘案してシステムに格納されている情報や知識を、積極的にユーザーに提示しながら検索することができるようになる。

この原理を取り入れた知的検索システムのビジュアルインタフェースの例を図7に示す¹⁵⁾。画面は概念階層木をビジュアルに表示するウインドウ、着目している概念の位置を示すウインドウ、および説明的な検索文を作成するためのウインドウから成る。ユーザーは検索文を組み立てる際に、概念ネットワークを自在に参照し、必要な概念をマウスで選択することによってそれを検索文の中にコピーすることができる。したがって、必ずしも完全に思い出せなくても、表示されている概念ネットワークから単語を選択するだけでよいことになる。

検索結果は、検索文に具体的な情報をはめ込んだ日本語文章、概念を説明するフレーム、あるいは文書画像によって表示することができる。情報をネットワークで表現して、いろいろな角度から検索したり、表示できる点で、知的なハイパーテキストシステムと見ることもできる。今後重要な情報検

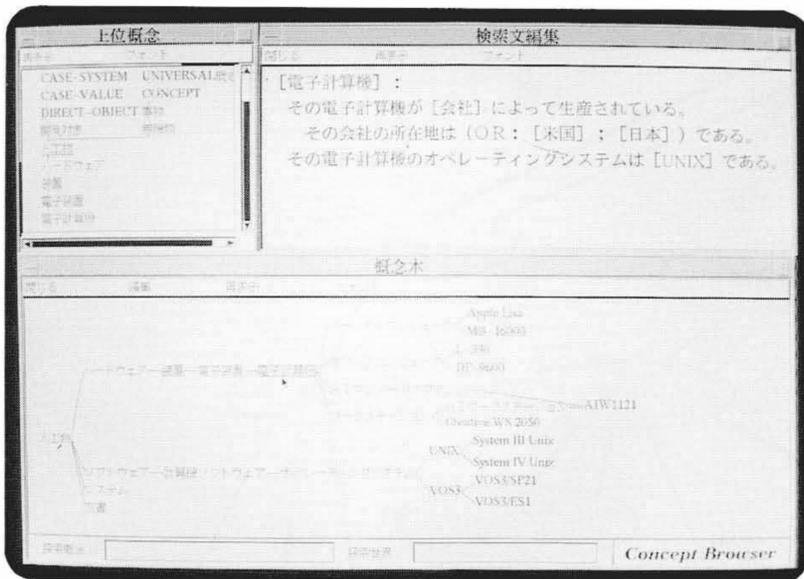


図7 ビジュアルインタフェースの例 概念ネットワークを用いた知識ベースをブラウジングしながら、検索条件を定式化できる。

索システムへのアプローチとなると考えられる。

5 ハードウェア技術

5.1 ハードウェアの構成

光ディスクファイルシステムの特徴は、大きな画像データを高速に処理するための画像処理向けアーキテクチャと、大規模カスタムLSIを採用していることである。

光ディスクファイルは独立した高速画像データバスを持つとともに、高精細な画像を高速処理するための画像処理プロセッサと、あいまい性を許しながら高速に文字列を検索するための検索エンジンを持っている。

画像処理プロセッサは、画像データの圧縮符号化処理と復号化処理、あるいは拡大・縮小・回転などのアフィン変換処理を行う。これまでは文字や図形を中心とした二値画像を主に対象にしてきたが、今後は写真と文字図形が混在する画像データを適応的に処理する。

一方、検索エンジンは、光ディスクあるいはメモリ内に記憶しているキーワードを含む文書名データをスキャンして、指定した検索語の探索を行う。有限オートマトンを検索原理とする検索エンジンは、同義語などの複数の検索語を同時に検索できることが特徴である。

5.2 高速化技術

画像データが大きいことに対処するために、まず画像データバスは32ビットのデータ幅を持ち、最大16 Mバイト/sのスループットを持っている。

さらに、画像プロセッサは、パイプライン処理を行うことによって高速化を図っている。これにより、A4サイズ以上の画像のアフィン変換や圧縮伸張処理を高速に実行する。また、光ディスクからの読み出し、圧縮符号の伸張、および表示処理をパイプライン化することにより、高速なページめくりを実現している。

検索エンジンは、先行フェイル方式と呼ぶ有限オートマトン方式を用いることにより、1入力文字に対して常に1サイクルで処理が終了する高速方式を実現している。これにより、同義語や異表記語から成る最大512語の検索語を同時に検索することが可能となっている。

5.3 小形化技術

光ディスクファイルは、マイクロエレクトロニクス技術の進展とともに、床置き形からデスクトップ形へ、さらにラップトップ形へと着実に小形化してきている。

これは、回路の微細化やチップの大形化などによるLSIの高集積化技術、複数のチップを実装するマルチチップモジュール化技術、あるいは高密度実装技術に支えられている。現在、光ディスクファイルの画像制御部の回路は約200 kゲートの論理で構成されるが、数枚の小形基板で実現されている。

今後、機能の拡大によって回路規模も増大するが、集積度の向上によってさらに小形化が進むと考えられる。

6 システム技術

6.1 システム技術に対する期待

光ディスクファイルシステムは、現在、企業の戦略的意思決定支援という高度な目標に向け新しい展開が求められている。単なる業務文書や設計図面の倉庫としてだけでなく、多様な業務処理要求にこたえて、ネットワークシステムの中の一要素として統合化していくことが求められている。

このような要求にこたえるシステムを最適に構成していくためには、次のような観点からの検討が重要である。

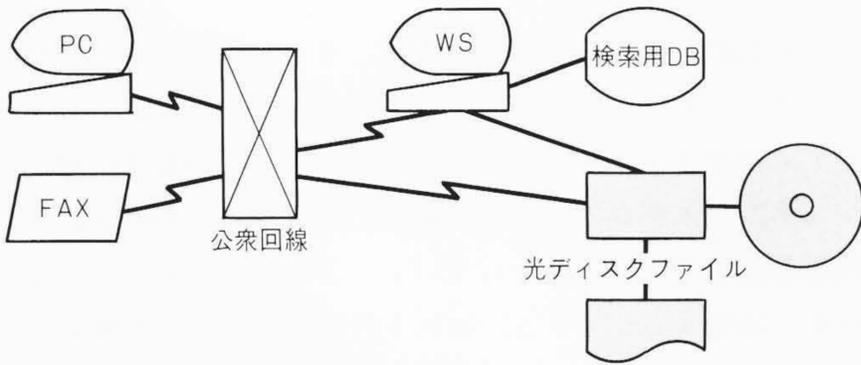
- (1) ファイリングする文書の量、参照頻度、利用形態などに照らして、画像ファイルの設置場所は手元が必須(す)か、離れていてもよいか。
- (2) 画像ファイルを遠隔地に置いてアクセスする場合のネットワークの種別、および適切な回線速度
- (3) コードデータと画像データの混在処理の必要性と、混在する場合の効率の良い処理形態
- (4) 光ディスクファイルとそのほかの関連オフィス機器との役割分担および連携方法

また、これからのオフィスシステムでは、既存のオフィス機器をも取り込んで、全体として有機的な機能をするように構成することが重要である。メインフレーム計算機やパーソナルコンピュータ(以下、パソコンと略す。)、あるいはファクシミリなどとの連携が有効である。

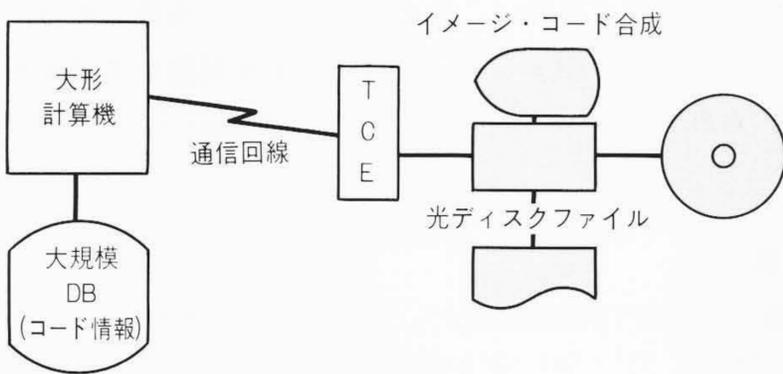
6.2 システム形態

光ディスクファイルを中心に、画像データの取り扱い方法からシステム形態を整理すると、図8に示すように幾つかの特徴的な形態に分類できる。

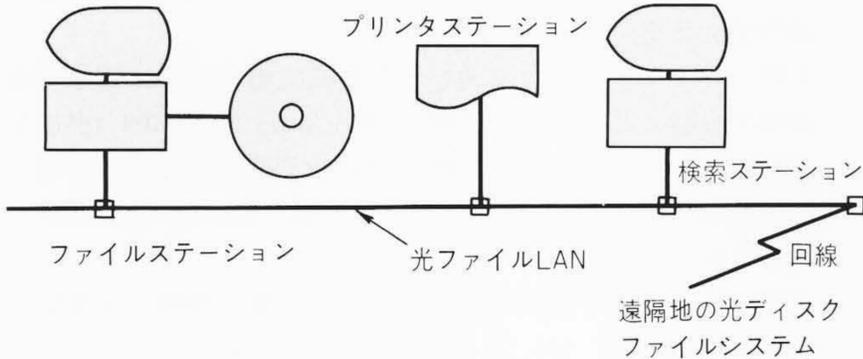
- (1) FAX(ファクシミリ)応用遠隔検索・遠隔入力システム
FAX網による画像情報の受配信機能と、光ディスクファイ



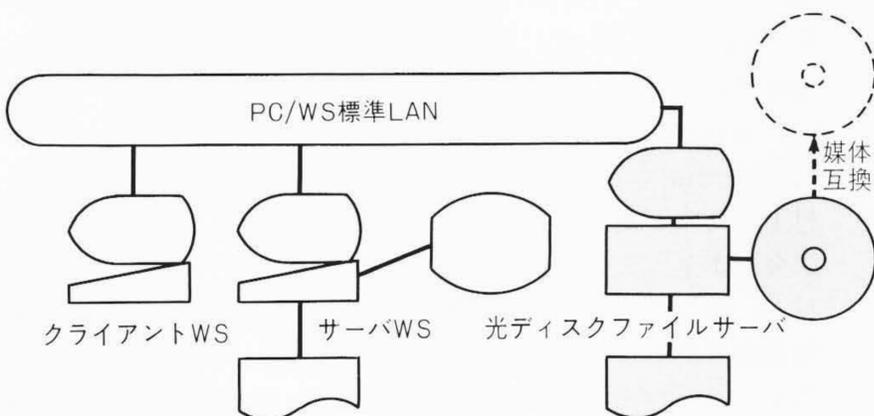
(a) ファクシミリ応用遠隔検索・遠隔入力システム



(b) ホスト計算機応用システム



(c) 光ディスクファイルネットワークシステム



(d) クライアント・サーバ形システム

注：略語説明 PC (Personal Computer), WS (Work Station)
 DB (Data Base), FAX (Facsimile)
 TCE (Terminal Control Equipment)
 FAXを応用した遠隔検索システム, ホスト計算機のデータベースと連携した検索システム, 大量文書に特化したシステム, あるいはクライアント・サーバ形システムなど, 多様なシステムを構築することが可能である。

図8 光ディスクファイルの主なシステム形態 FAXを応用した遠隔検索システム, ホスト計算機のデータベースと連携した検索システム, 大量文書に特化したシステム, あるいはクライアント・サーバ形システムなど, 多様なシステムを構築することが可能である。

ルの大量画像情報蓄積機能を組み合わせて, 遠隔地からの情報検索と情報入力とを実現するシステムである。

例えば, パソコンを用いて遠隔地の光ディスクファイルにアクセスし, 検索結果を手元のFAX装置に送信させて情報を入力したり, それとは逆に手元のFAX装置から文書画像を入力して, 遠隔地の光ディスクファイルに自動的にファイリングすることが可能となり, FAX網の広がり, FAX装置の画質や性能の向上により, ますます応用範囲が広がるものと思われる。

(2) ホスト計算機応用システム

大規模なデータベースを管理するホスト計算機と, それに関連する画像情報を蓄積する光ディスクファイルを有機的に結合したシステムである。大規模データベースの検索結果と, それに関連する画像情報を合成して出力することができる。

大規模なデータの一元管理と, 画像データの高速な表示性能にメリットがある。比較的大規模な組織では, すでにデータベースがホスト計算機上に構築されていることも多いため, 効果的なシステム形態である。

(3) 光ディスクファイルネットワークシステム

光ディスクファイルシステムを構成する要素は, 検索ステーション, ファイルステーション, ラップトップ形検索端末などと多様化し, ネットワークシステムも構築しやすくなっている。技術資料, 設計図面, 特許資料など, 部門間で共通に管理したり参照したりする資料が大量に存在し, 利用目的を限定できる場合は, 光ディスクファイル専用のネットワークを構築することが効果的である。

これにより, 光ディスクファイルシステムの特徴である画像データの高速検索・高速表示, あるいはペーパーレス化に必須な高精細画像の表示が可能となる。

(4) クライアント・サーバ形システム

標準化が進みつつあるワークステーションのオペレーティングシステムやウィンドウシステムによると, 汎(はん)用ワークステーションをクライアントとして, また, それぞれ得意な機能を提供する装置をサーバとして分散処理ネットワークを構築することができる。したがって, 光ディスクファイルをその一つのサーバとしてとらえるシステム形態が将来有望である。

光ディスクの媒体互換性を保つことによって, データの共有化を図るのがその第一歩であるが, 近い将来は標準LANに接続された汎用ワークステーションを介して, 光ディスクファイルをドキュメントサーバとして活用することが多くなる。

このシステムでは, 光ディスクファイルから呼び出した情報を, 意思決定に役立つ形に加工および編集することが容易となり, 光ディスクの特徴を最大限に生かすことができる。

6.3 システム化を支える技術

多様化するシステムニーズにこたえる以下の技術の開発が

進んでいる。

(1) ユーザーインタフェースの標準化および共通化

標準化はオペレーティングシステムのレベルからユーザーインタフェースのレベルへ進みつつある。パソコンではMS-DOS^{*1)}系、ワークステーションではUNIX^{*2)}系が業界標準であるが、それぞれ画面表示のためのウインドウシステムについても標準化が進行中である。

前者についてはMS-WindowあるいはOS-2のPresentation Manager, 後者についてはOSF(Open Software Foundation)/Motif^{*3)}が世界標準の有力な候補となっている。これらでは画面の見え方と操作方法(Look and Feel)についても統一化が目標になっている。

(2) メディア変換技術

コードデータから画像データへ、あるいはその逆の方向のメディア変換が自由にできることが重要となる。特に後者の方向の変換を実現する文書画像認識技術は近年さらに進歩してきている。認識精度とのトレードオフを考慮することにより、全体として効率のさらに良いものとする事ができる。

(3) オブジェクト指向データ管理技術

扱うデータのマルチメディア化が進む中で、オブジェクト指向のデータ管理方式が注目されている。テキスト、図形、写真画像、音声などのメディアから構成される複雑な構造を持ったオブジェクトを管理する方法が新しいデータベース方式として研究されている。これにより、将来コードデータを中心とするシステムと、画像データを中心とするシステムの真の統合が可能となろう。

これらの技術の進歩により、多様化するニーズに対して使い勝手がよく、かつ性能的に最適なシステムの構築が容易になると思われる。

7 結 言

光ディスクファイルシステムを支える情報処理技術について論述した。文書画像処理技術と高度なパターン認識技術は、文書の登録作業の省力化と自動化、あるいはカラー文書のファイリングを可能にしていく。文書の検索の問題は人間の記憶の特性とも関連が深い問題である。ユーザーが言い足りない情報をシステム側で補足して、おぼろげな記憶からでも検索を可能とする知的検索技術が開発されつつある。同義語や異表記語を自動的に補足して検索したり、システムが持つ知識ベースの中身をビジュアルに表示して、対話的な検索を支

援することが可能になる。

これらの新しい情報処理技術は、さらに高速化と小形化のためのハードウェア技術によって、現実のものとなっていく。光ディスクファイルシステムが広い応用を見いだすに伴い、システム全体を有効的かつ経済的に実現するためのシステム技術もますます重要となっていく。公衆回線を利用した遠隔検索・遠隔入力システム、ホスト計算機と連動した検索システム、光ディスクファイルネットワークシステム、あるいはクライアント・サーバー形システムが、ニーズに最も適した形で実現されていく。

文書ファイリング技術はこのようにトータル技術である。顧客ニーズの発展と相まって、今後ますます発展していくものと考えられる。

参考文献

- 1) 伊藤, 外: HITFILE650光ディスクファイルシステム, 日立評論, 69, 6, 575~581(昭62-6)
- 2) 東野, 外: 自動ファイリングのための文書理解方式—知的ファイリング(その2)—, 情報処理学会第31回全国大会, 2N-2(1985-9)
- 3) 鳴, 外: 文書画像の傾き検出についての実験的検討, 電子情報通信学会秋季全国大会, D-183(1989)
- 4) 中野, 外: 文字認識と協調した表形式文書の理解, 電子情報通信学会論文誌(D), Vol. J69-D, No. 3, pp. 400~409(1986-3)
- 5) 中野, 外: 自動ファイリングのための文書理解の一方式, 電子情報通信学会論文誌(D), Vol. J71-D, No. 10, pp. 2050~2058(1988-10)
- 6) 東野, 外: 書式定義言語を用いた文書画像の理解, 画像電子学会誌, Vol. 17, No. 5, pp. 267~277(1988-5)
- 7) H. Fujisawa, et al.: Document Analysis and Decomposition Method for Multimedia Contents Retrieval, Proc. 2nd Int. Symp. Interoperable Information Systems, ISIS, pp. 231~238, Tokyo(1988-11)
- 8) H. Yashiro, et al.: A New Method of Document Structure Extraction using Generic Layout Knowledge, Proc. Int. Workshop on Industrial Applications of Machine Intelligence and Vision(MIV-89), pp. 282~287, Tokyo(1989-4)
- 9) 村上, 外: 写真混在文書画像の領域分離方式, 電子情報通信学会春季全国大会, D-234(1988-3)
- 10) 村上, 外: デザイン画像を対象としたアフィン変換方式, 電子情報通信学会春季全国大会, D-439(1990-3)
- 11) 畠山, 外: 自由語検索のための同義語・異表記展開方式, 情報処理学会第39回全国大会, 2N-7(1989-10)
- 12) 川口, 外: 自由語検索のための高速文字列検索方式, 情報処理学会第39回全国大会, 2N-8(1989-10)
- 13) 加藤, 外: 全文検索用テキストサーチマシンの開発, 電子情報通信学会研究会, DE89-38, pp. 17~24(1989-12)
- 14) 藤澤, 外: 概念ネットワークを用いた知的ファイリングシステム, 日立評論, 69, 3, 231~238(昭62-3)
- 15) 木内, 外: 知的ファイリングシステムのビジュアルインタフェース, 情報処理学会, 文書処理とヒューマンインタフェース研究会, 21-3(1988-11)

*1) MS-DOSは、米国マイクロソフト社の登録商標である。

*2) UNIXは、米国AT & T社ベル研究所が開発したオペレーティングシステムの名称である。

*3) OSF/Motifは、オープン・ソフトウェア・ファウンデーションの登録商標である。