

イメージ情報の知的検索インタフェース

7X-8

和歌森 文男 土屋 佳貞

(株)日立製作所システム開発研究所

1. はじめに

高機能ワークステーションの開発、普及に伴い、イメージデータ処理機能の提供が重要な課題となっている。イメージデータは、文字、図形データに比べ、データ量の飛躍的増大をもたらすため、システム構成やソフト設計面に従来システムとは異なる工夫が必要である。ワークステーションの操作者が欲するとき、必要な情報を大量のイメージデータ中から迅速に引き出して処理可能とするために、検索論理のルール化による対象絞り込みと、絞り込まれた候補データのサマリ画像によるビジュアルサーチを組み合わせた知的検索インタフェースの実現方法について報告する。

2. 知的検索インタフェース

従来のデータ検索は、キーワードの論理式入力による問い合わせを基本として考えられてきた。検索対象がイメージデータであるという特徴に合わせ、ビジュアル表示データをキーとして、処理目的や要求の直接的表現を許容する仕組みを提供することが必要である。具体的には、次のような要求への対応を考える。

「表示中のデータAと”関連(類似)”するデータを探す」

上記インタフェース実現のためには、操作者とシステム相互間で検索論理についての合意が成り立たなければならない。そのための知識を、全ての分野、全ての利用者に共通したルールとして蓄積できれば理想であるが、それは不可能である。しかし、ある対象に環境を限定することによって、検索に関わる諸概念(類似性、関連性等)についての法則、ルールが抽出できるものという前提に立つ。そのねらいは次の諸点にある。

- (1) ワークステーションという個人利用を前提としたシステムの特徴を生かして、個人の検索環境を作り易い。このため、従来の検索システムのようにキーワード論理式を毎回の処理において生成する必要がなく、検索負荷が軽減する。
- (2) 類似性や関連性についての定義を明示的に表わさずに行なう検索処理は、人間の頭脳活動としてのイメージ記憶の想起の仕組みと似ており、操作者にとってなじみやすい。
- (3) 処理を行なった結果、利用者とシステム内の知識データの整合性が不十分であることが判明した場合(関連度順位付けの相違等により発生)、ルールの追加、更新を通じて結果をフィードバックすることが容易である。その結果、システムが学習機能を持ち、使いこむうちに利用者の認識に近い検索、処理を行なえるようにすることができる。

以上のねらいを持つイメージ情報の知的検索インタフェースの概念説明を図1に示す。

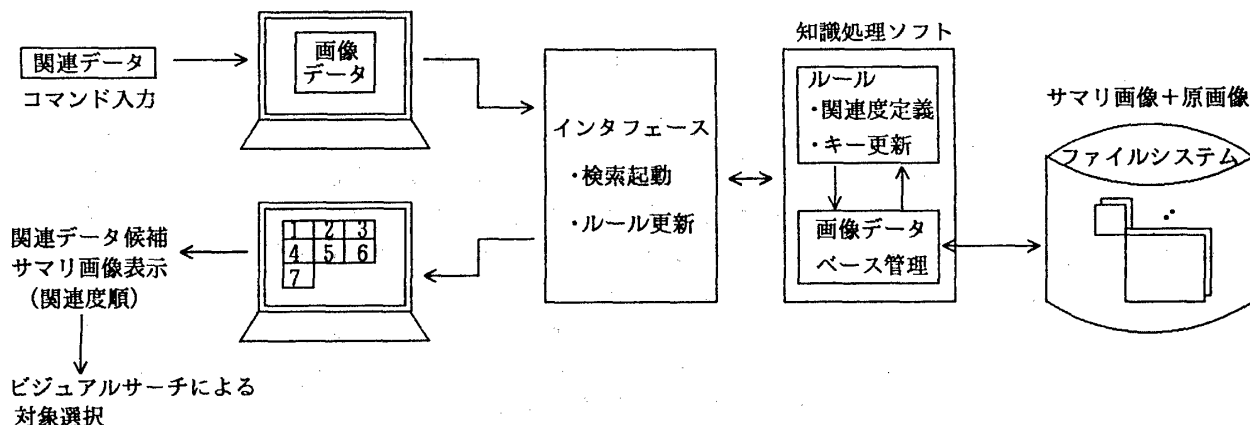


図1 イメージ情報の知的検索インタフェース

3. 実験システムの機能

特許明細書の検索システムを例に実験システムを構成した。

(1) インタフェース仕様

- i) 明細書表示時のコマンドメニューとして「関連特許」を準備する。このメニューを選択することにより、システムに蓄積した関連度についてのルールが調べられ、関連度の高いデータから順次、そのサマリ画像を表示する。
- ii) サマリ画像には、特許の特徴を表わすインデックス情報として、タイトル、登録番号、発明者、出願人、キーワードおよび主図を含める。
- iii) サマリ画像をビジュアルにサーチすることにより、全文明細を調べる特許が決定し、該当明細書全文の表示に移行する。

(2) 知識データ処理

- i) データの関連度定義をルール形式で記述する。共通ルールは、タイトルおよびキーワードの一致度合の高いものの関連度を大とするものである。共通ルールに加え、個人や問題設定に応じて、固有の関連度定義（例えば、出願人の一致を重要視するルール等）ルールを付加することができる。
- ii) 操作結果をルール推論に反映させる。反映法としては、操作履歴に着目し、関連特許として全文明細の参照が行なわれた場合にその情報を記憶し、関連度評価時に利用することがまず考えられる。さらに、システム側が提示した関連度が、操作者の判断と異なる場合には、検索のキーとしたデータの不足を調べることにする。その結果、キーデータの不十分性が見い出された場合にはキーの追加登録、入れ替え等を行なう。実験システムにおける知識データ更新の概念を図2に示す。

以上の実験システムにより、特許明細書データの検索に良好な環境を構築可能なことを確認した。

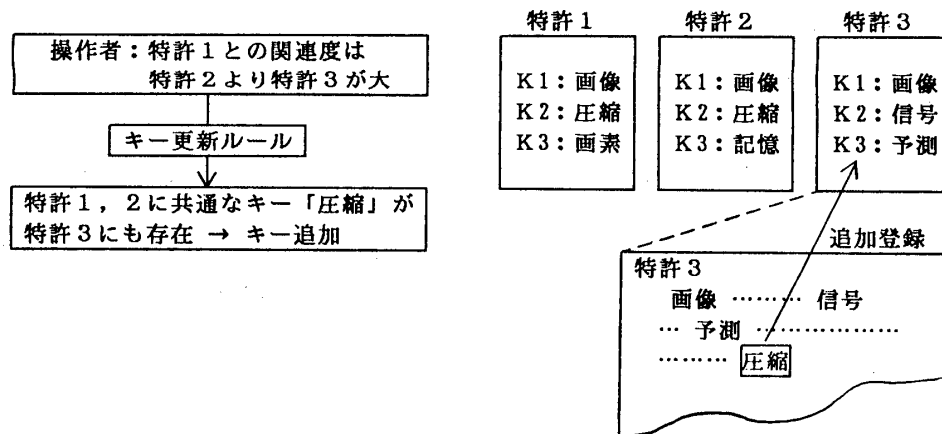


図2 実験システムにおける知識データ更新例

4. まとめ

光ディスク装置等の普及により、画像データの大量蓄積はますます進むものと思われる。それに伴い、効率的な検索方式の開発が求められている。本報告では、(1) ルール化された検索論理にもとづく参照候補データの絞り込み (2) 参照候補データのサマリ画像表示によるデータの特定 (3) 特定結果にもとづく検索論理の逐次追加修正 という3段階による検索インタフェースを提案した。蓄積画像データ量が多くなると、ホストとワークステーションを接続し、大量画像データは、ホスト側で一括管理するシステム構成が多くなるものと考えられる。提案した検索インタフェースに従えば、サマリ画像による対象の特定にもとずき、最終的に参照するデータのみホストからワークステーション側に転送すれば良い。このため、データ転送や表示負荷の軽減が図れ、実用性の高いシステムを構築することが可能である。

参考文献

- 1) N.S.Chang & K.S.Fu : Query-by-Pictorial-Example, IEEE Trans. Software Engineering, Vol SE-6, No.6, pp519-524 (1980)
- 2) 洪, 六川 : カラー画像検索システムへのアプローチ, 情報処理学会第31回全国大会 (昭和60年9月)