

マルチメディア情報検索システム "MIRACLES"

長田茂美 遠藤進 椎谷秀一 上原祐介 増本大器
株式会社富士通研究所 コンピュータシステム研究所 知能システム研究部
〒261-8588 千葉市美浜区中瀬 1-9-3 (幕張システムラボラトリ)

E-mail: {nagata.shigemi, endou.susumu-02, shiitani, yuehara, masumoto.daiki}@jp.fujitsu.com

概要: 本稿では、マルチメディア情報を構成する各メディアの特徴および各メディア間の関連性を有効に活用した新しい検索モデル“パノラミック検索モデル”を提案する。検索対象がテキストと画像とで構成される典型的な場合には、このモデルは、テキストによる意味的な検索と画像による視覚的な検索を兼ね備えた情報検索手段を提供する。また、パノラミック検索モデルに基づく検索システム“MIRACLES (Multimedia Information Retrieval, Classification, and Exploration System)”を構築し、Web 文書検索、デジカメ画像検索、ショッピングポータルにおける画像一覧検索サービスに適用した。それらの応用事例についても述べる。

キーワード: マルチメディア情報検索, パノラミック検索モデル, 3次元情報空間, 視覚化, クロスメディア検索

Multimedia Information Retrieval System "MIRACLES"

Shigemi NAGATA, Susumu ENDO, Syuichi SHIITANI,
Yusuke UEHARA, and Daiki MASUMOTO
Intelligent Systems Laboratory, Computer Systems Laboratories,
FUJITSU LABORATORIES LTD.
9-3, 1-chome, Mihama-ku, Chiba City, Chiba 261-8588

Abstract: We propose a multimedia information retrieval model named "Panoramic Information Retrieval Model" and a multimedia information retrieval system, based on the model, named "MIRACLES (Multimedia Information Retrieval, Classification, and Exploration System)". The system, which typically integrates semantic search by texts and visual search by images, enables "seeing and selecting information". We also describe applications of MIRACLES including web document retrieval, digital camera image retrieval. **Keywords:** Multimedia information retrieval, Panoramic information retrieval model, 3-D information space, Visualization, Cross-media retrieval

1. はじめに

インターネットの爆発的な拡がりとともに、テキスト、画像、音声など多様なメディアが混在したマルチメディア・コンテンツの蓄積、活用が盛んになってきた。これに伴い、膨大な情報の中から、所望のコンテンツを効率的に探索したり、手に入れたりするマルチメディア情報検索技術の重要性も益々

高まっている。

このようなマルチメディア情報の検索技術として、現在、実用に供されているのは、依然として従来からのテキスト検索技術や内容に基づく画像検索技術 [1] などである。これらの検索技術は、検索対象であるマルチメディア情報を構成する各メディアを独立に検索する技術であり、一つの情報表現の纏まりと

してのマルチメディア情報を構成する各メディアの特徴および各メディア間の関連性を有効に活用した検索技術は、未だ大部分のものが研究段階にある。

マルチメディア情報検索という観点から捉えると、従来からのテキスト検索技術は、マルチメディア情報を構成するテキストを検索対象とすることで、マルチメディア情報の意味内容に関する検索手段を提供する。また、内容に基づく画像検索技術は、マルチメディア情報を構成する写真や図などを検索対象とすることで、色や形などの視覚的な特徴に基づいて類似した画像の検索手段を提供する[1]。しかし、これらの検索技術、すなわち、検索者が与える検索条件とデータベース内の情報との照合に基づく検索技術では、検索者が適切な検索条件を与えなければ、高精度の検索は実現できない。一般に、検索者が必要とするマルチメディア情報を特定する詳細な検索条件を想起したり表現したりすることは困難であり、ある程度検索精度を犠牲にした粗い検索条件で検索することになる。その結果、検索者は大量の検索結果を閲覧する多大な労力を要求される。

近年、このような多大な労力を軽減し、大量のマルチメディア情報を効率的に閲覧する手法を実現すべく、マルチメディア情報検索における情報集合の視覚化の研究が進められており、検索結果としての文書集合や画像集合を2次元空間あるいは3次元空間上で視覚化することにより、検索結果を効率的に閲覧する手法が幾つか提案されている[2][3][4]。しかし、これらの研究では、視覚化された検索結果の情報集合に対して、さらに検索条件を追加して所望の情報を集めたり、逆に、不要な情報を指定する検索条件を追加してそれらを排除したりするといった検索（絞り込み検索と呼ぶ）過程の視覚化は扱われていない。通常、一度の検索で所望の情報が見つかることは稀であり、検索者はインタラクティブに検索を繰り返すことによって徐々に所望の情報を絞り込んでいく。したがって、これらの一連の検索過程における検索者とシステムとのインタラクションの結果得られる検索結果の情報集合や、それらがどのようなインタラクションによって得られたかという検索過程を一覧できることは、所望の情報に効率的にアクセスするための有効な支援となる。

また、もう一つ別の観点として、所望のマルチメディア情報を効率的に検索するためには、マルチメディア情報を構成する各メディアの特徴および各メディア間の関連性を有効に活用し、検索者が必要に応じて各メディアの特徴を適宜使い分けながら検索を進められることが重要な要素となる。例えば、検索対象がテキストと画像から構成されるマルチメディア情報であれば、上述した一連の検索過程において、テキストというメディアの特徴を活かした意味的な検索と画像というメディアの特徴を活かした視覚的な検索とを自由に組み合わせることで検索できるクロスメディア検索（必要に応じて各メディアの特徴を切り替えながら検索、視覚化を進めていくことをクロスメディア検索と呼ぶ）は、検索者にとって極めて有効な支援機能となる。

以下、本稿では、テキストや画像などから構成されるマルチメディア情報の新しい検索モデル“パノラミック検索モデル”とそれに基づく検索システム“MIRACLES (Multimedia Information Retrieval, Classification, and Exploration System)”を提案し、Web 文書検索システム、デジカメ画像検索システム、ショッピングポータルにおける画像一覧検索サービスに適用した応用事例について述べる。

2. マルチメディア情報のパノラミック検索モデル

提案する“パノラミック検索モデル”は、テキストや画像などから構成されるマルチメディア情報を対象とし、1) 検索結果の情報集合と検索過程の3次元空間による視覚化、2) テキスト特徴による検索や画像特徴による検索など、マルチメディア情報を構成する各メディアの特徴および各メディア間の関連性を活用し、各メディアに適した検索手段を有効に組み合わせたクロスメディア検索、を特徴とするインタラクティブな検索モデルである。

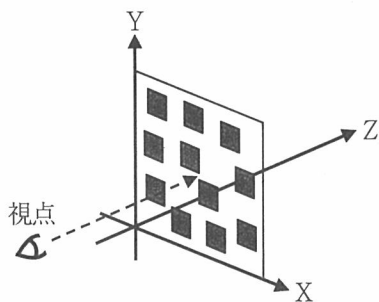
2.1 情報集合と検索過程の視覚化

パノラミック検索モデルでは、図1に示す3次元の情報空間上で検索が進められる。本モデルは広くマルチメディア情報全般に適用可能であるが、ここでは、検索対象となる情報集合をテキストと画像で構成されるマルチメディア情報（一つの情報単位がテキストと画像の対で表現されたもの）に限定して

論ずる。まず、データベースから表示対象となる情報集合をキーワードなどのテキスト特徴や色や形などの画像特徴により検索する。この検索を初期検索と呼ぶ。初期検索により得られた情報集合は、図1に示す灰色の小矩形のようにXY平面上に配置される。このときの配置方法は互いの類似関係に基づく。すなわち、類似した情報同士ほど近い距離に配置される。個々の情報の表示には、例えば、それぞれの情報に含まれる画像を用いる。この配置結果を一覧表示平面と呼ぶ。検索者は、一覧表示平面を垂線(Z軸)方向から眺める。一覧表示平面では情報集合が類似関係に基づいて配置されており、全体的に眺めることで情報集合の大局的な分類構造が一望できる。また、特定の領域に注目することで、興味のある情報とそれに類似する情報を集中して見ることができる。このような一覧表示は、目的の情報を見つけるための一定の手掛かりとなる。しかし、表示される情報の数が多い場合は、さらに条件を加えて目的の情報を選別したり、不要な情報を排除したりする絞り込み検索が必要となる。本モデルでは、絞り込み検索

の結果を一覧表示平面の垂線方向に移動して表示することで、検索過程を視覚化する。この移動方向の軸(Z軸)を検索過程表示軸と呼ぶ。図2に検索過程の視覚化方法を模式図で示す。初期検索結果の一覧表示平面(R0)に対し、条件を与えて目的の情報を選別した結果の情報集合が、検索者の手前方向に移動して表示される(R1)。このとき、与えられた条件との適合度が高い情報ほど検索者に近い位置に表示される。さらに条件を加えて選別した場合、同様にして選別された情報がさらに手前に表示される(R2)。また、一覧表示平面(R0)に対して不要な情報を排除する条件を与えると、合致した情報は条件との適合度が高いほど奥行き方向に後退して表示される(R3)。また、これらの操作は任意の時点までバックトラックすることで、元の状態に戻すことができる。

以上のような検索過程の視覚化により、各段階の検索の結果得られる情報集合の間の量的な関係や一覧表示平面上での位置的な関係を容易に把握できる。また、検索過程表示軸への展開のバックトラックにより、試行錯誤的な検索が容易になる。



XY : 一覧表示平面
Z : 検索過程表示軸

図1. 情報探索空間の模式図

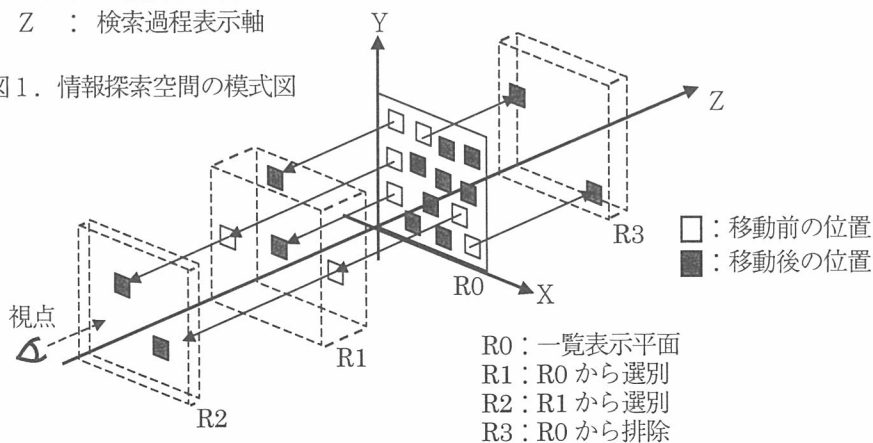


図2. 検索過程の視覚化

込み検索、一覧表示平面の配置のいずれにおいても、典型的には、検索者が、検索対象の意味的な内容に着目している場合（例えば、対象が商品情報の場合、価格やブランド名など）にはテキスト特徴による検索が、検索対象の視覚的な特徴に着目している場合（例えば、対象が商品情報の場合、色や形のデザインなど）には、画像特徴による検索、配置が有効である。本モデルでは、検索過程の各段階でこれらの検索対象の構成要素であるテキスト、画像といった各メディアの特徴を、検索者の目的に応じて適宜使い分けながら検索を進めていくこと、すなわち、クロスメディア検索を実現できる。

3. マルチメディア情報検索システム “MIRACLES”

ここでは、パノラミック検索モデルに基づくマルチメディア情報検索システム MIRACLES の応用事例として、Web 文書検索システム “MIRACLES for WWW”、デジタル画像検索システム、ショッピングポータルにおける画像一覧検索サービスについて述べる。

3. 1 MIRACLES for WWW

一般に、Web 文書中の画像の周辺には密接に関連するテキスト、例えば、その画像の意味内容を説明するテキストが存在する。このようなテキストを画像から見て“関連テキスト”と呼ぶことにする。図3に関連テキストの例を示す。MIRACLES for WWW では、Web 文書中の画像を検索するために、画像自体の特徴とともに、画像の意味内容を表す関連テキストの特徴も使用する。

図4にMIRACLES for WWWの構成を示す。本システムは、WWWで発信されるテキスト情

報や画像情報を収集する収集モジュール、収集した情報を管理するデータベース、収集したテキストおよび画像から特徴量を抽出する特徴量抽出モジュール、抽出した特徴量に基づいて一覧表示平面上に画像を配置する配置モジュール、配置された情報集合の絞り込み検索を実行する検索モジュール、表示とユーザー入力のための検索GUIから構成される。

画像とテキストをWWWから収集するために用いるクローラは、ページ内に埋め込まれたアンカー情報を解析し、次々とアンカーを辿ることによってページを巡回することができるプログラムである。

まず、検索者はWWW上の画像を収集するためにクローラを起動する。ユーザはクローラに対して二通りの方法で収集範囲を指定するこ

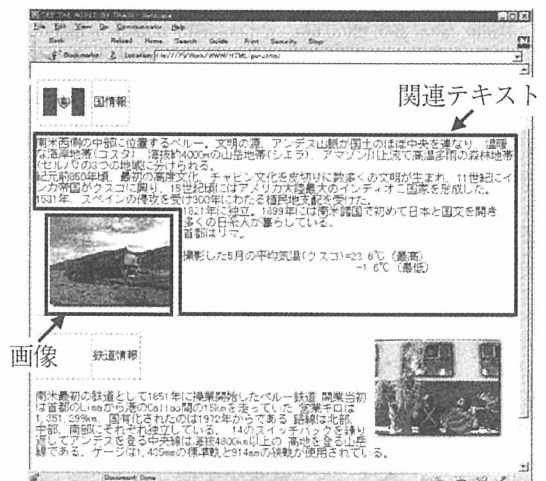


図3. 画像の関連テキストの例

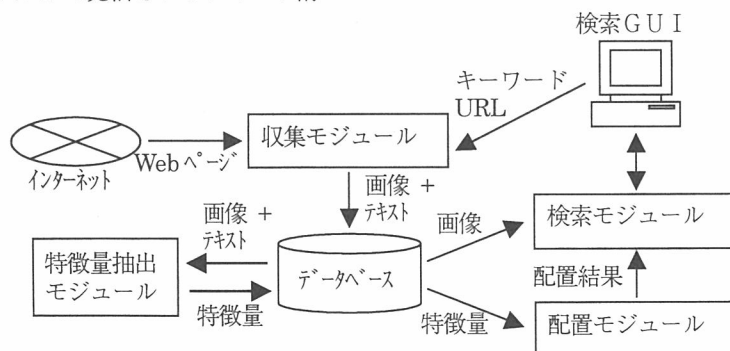


図4. “MIRACLES for WWW” のシステム構成

とができる。一つは、URL (Uniform Resource Locator) を指定する方法であり、クローラは指定された URL を起点に指定された数だけ Web ページのリンクを辿り、ページ内にある画像を収集する(URL 指モード)。もう一つは、収集用キーワードを指定する方法であり、ロボットはそのキーワードを外部のテキスト検索エンジンに渡し、その結果得られる URL のリストを辿って、ページ内にある画像を収集する (キーワード指定モード)。次に、ロボットは収集した画像の関連テキストを収集する[5]。関連テキストは、HTML (Hyper Text Markup Language) ファイルを解析して取得する。一般に、画像とテキストの関連度について、1) 画像と近い位置にあるテキストは画像と内容的に関連していることが多い、2) HTML のタグ構造からも画像とテキストの関連度をある程度推測できる、ということが言える。そこで、基本的には、画像とテキストの関連度をこの2つの観点から定義し、関連度の大きいテキストを関連テキストとして画像とともに保存する。

特徴量抽出モジュールは、画像とテキストから特徴量を抽出する。この特徴量は、一覧表示平面への配置で用いる。画像検索の研究分野では様々な画像特徴量が提案されているが[1][6]、本システムでは、代表的な色と形状に関する特徴量を用いている。色に関する特徴量としては、HSI 色空間を用いたヒストグラム特徴量を[7]、形状に関する特徴量としては、ウェブレット特徴量を[8]、また、テキストの特徴量としては、単語頻度特徴量を用いている。

配置モジュールは、抽出した画像特徴量やテキスト特徴量の類似関係に基づいて、画像集合を一覧表示平面に配置する[9]。すなわち、ベクトル値である特徴量間のユークリッド距離が小さい画像同士ほど平面上の近い距離に配置する。URL 指定モードで <http://www.fujitsu.co.jp> (現在は <http://jp.fujitsu.com> に変更) を指定して収集した情報集合を画像の色 (HSI ヒストグラム) 特徴で配置した例を、図5に示す。同色系のノートPCの画像が中心付近に集まっている。配置方式としては、Kohonen の提案した自己組織化マップ (Self-Organizing Map、以降 SOM と略す) [10]を用いた。SOM はニューラルネット

ワークの一種であり、教師なし競合強化学習モデルである。SOM では、データの分布の把握を容易にするために高次元の特徴ベクトル空間を低次元空間に写像する。その際、高次元空間における距離が近いデータ同士ほど、低次元空間においても近くに配置される。今回用いた SOM は低次元空間として2次元空間 (一覧表示平面) を用い、画像を格子状に配置する。本システムでは、色、形状、テキストの特徴量による3組の配置情報を生成する。検索者は目的に応じて3種類の配置結果を適切替えて表示させることができる。

一覧表示平面に配置された情報集合は、図5の表示例のように検索 GUI を介して検索者に提示される。表示直後の視点は、画像集合全体を見渡せる位置に設定され、その後、検索者はフライスルーによって興味のある画像の付近に視点位置を移動することにより詳細な内容を確認できる。また、画像を指定することにより、その画像の関連テキストを表示できる。図6は、関連テキストの表示例を示す。

検索モジュールは、一覧表示平面の表示内容を対象とした検索機能として、上述した絞り込み検索機能と類似情報リスト表示機能を備える。類似情報リスト表示機能は、検索の過程で目的に近い画像を見つけたときに、その画像と類似する画像を類似度の高い順番で見たいという要求に応えるためのものである。これらの検索機能

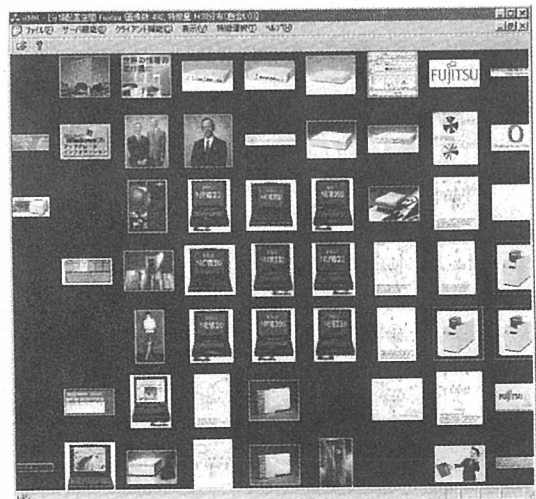


図5. 一覧表示平面の表示例

は、テキストの特徴量、あるいは、画像の特徴量を用いて実行できる。

・絞込み検索： 一覧表示平面に表示された情報集合に対し、検索者がキーワードを入力することで絞込み検索を実行する。検索者は任意の言葉をキーワードとして入力できる。また、図6の例のように表示された関連テキストの一部をマウスでドラッグすることによりキーワードとして入力できる。絞込み検索には、検索者から与えられた条件に合致する情報を選別する方法と合致する情報を排除する方法がある。図7は、条件に合致する情報を選別する絞込み検

索の結果の表示例である。この例は、鞆の商品情報の検索を想定したものである。まず、キーワード指定モードで“バッグ”を指定して画像と関連テキストを収集した。次に、デザインの観点から、画像内容の形状特徴に基づいて一覧表示平面に配置している。ここで、仕事用の鞆の情報のみを選別するために、キーワード“ビジネスバッグ”を入力して絞込み検索を実行した。その結果、関連テキストに“ビジネスバッグ”を含む画像が検索者の手前方向に移動した位置に表示される。図7では、白丸で囲んだ画像が選別結果であり、絞込み前の情報集合との位置関係が一覧可能となっている様子が見える。また、逆に、不要な情報の排除を指定すると、入力されたキーワードを関連テキスト中に含む画像を画面の奥に移動する。図8は、不要な情報を排除する絞込み検索結果の表示例である。この例は、鞆を想定して設定したキーワード“バッグ”により収集した画像の中に車のエアバッグ画像が含まれていたため、キーワード“エアバッグ”により排除した例である。図8では、楕円で囲まれた領域に、奥に後退したエアバッグに関する画像が表示されている。このように、不要な情報を画面上から消去するのではなく、画面の奥に目立たなくして表示しておくことで、検索の履歴情報として利用できる。絞込み検索において、情報集合の移動の過程はアニメ



図6. 関連テキストの表示例

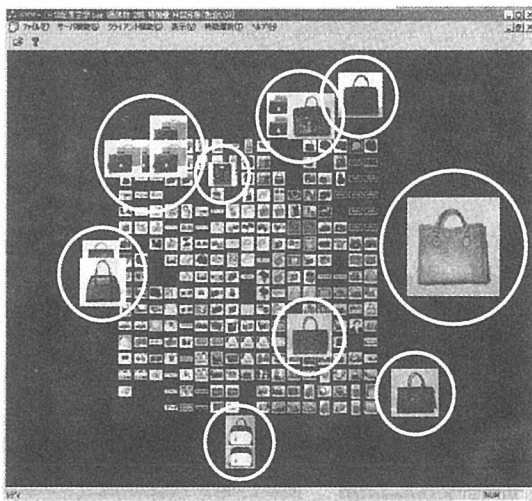


図7. 選別による絞込み結果の表示例

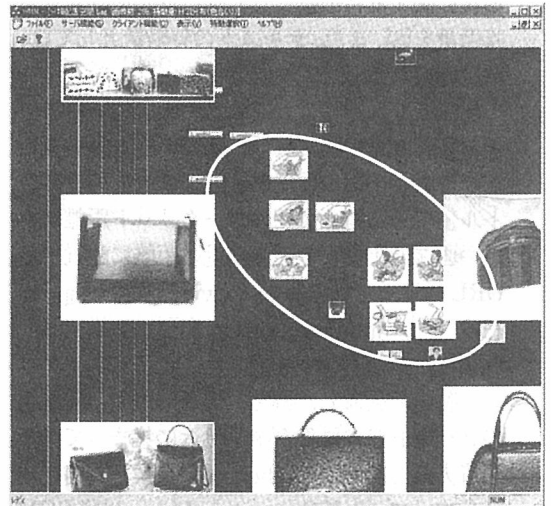


図8. 排除による絞込み結果の表示例

ーションにより連続的に表示される。これは移動する情報の確認を容易にするためである。キーワードの代わりに画像特徴を用いる場合は、検索キーとする画像を選択し、使用する特徴量として色、形状のいずれかを指定する。

・類似情報リスト表示： 検索対象がある程度絞り込まれた段階では、特定の情報に類似した情報を類似する順に閲覧する方が効率的である。この場合、表示された画像の中から検索キーとする画像を指定し、使用する特徴量としてテキスト、色、あるいは、形状のいずれかを選択する。図5の表示例において、ネットワークのハブ装置の画像を検索キーとして指定し、テキスト特徴量を使用したときの類似情報リストを図9に示す。検索キーとして選択した画像がリストの左上に表示され、以下、行先行で類似順に並べられる。

3. 2 デジカメ画像検索システム

MIRACLES をベースにデジカメ画像検索システムを構築した。本システムでは、一つあるいは複数のフォルダに格納された画像を、画像の属性（色、色分布、色レイアウト、撮影/変更された日時などの情報）を用いて、その属性が類似した画像が近くに集まるように3次元空間上に配置して一覧表示し、検索者は、3次元空間内をフライスルーしながら、画像を検索・整理することができる。また、キー画像を指定して、類似画像を検索することもできる。

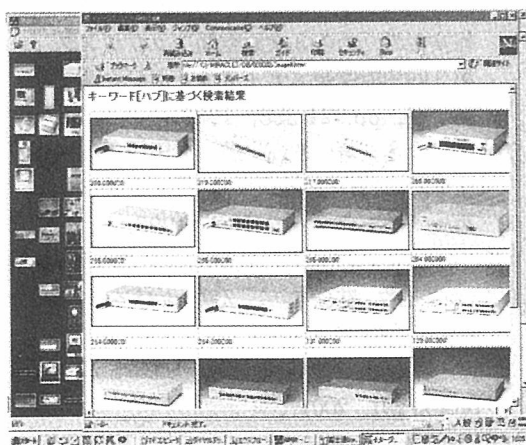


図9. 類似情報リストの表示例

図10に、PC向け画像管理検索ソフト“みよう絵”[11]の表示画面例を示す。

3. 3 ショッピングポータルにおける画像一覧検索サービス

MIRACLES の機能をショッピングポータルでの画像一覧検索サービスにも適用した[12]。検索者は、キーワードで商品を絞り込んだ後、商品画像を一覧表示し、これらの画像を色、価格、商品名などで並べ替えながら、ウインドウショッピング感覚で欲しい商品を選ぶことができる。

図11に、Shopping@niftyにおける商品画像一覧サービス[13]の表示画面例を示す。

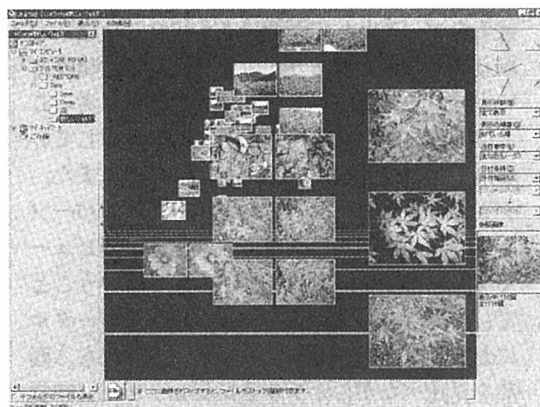


図10. “みよう絵”の表示画面例

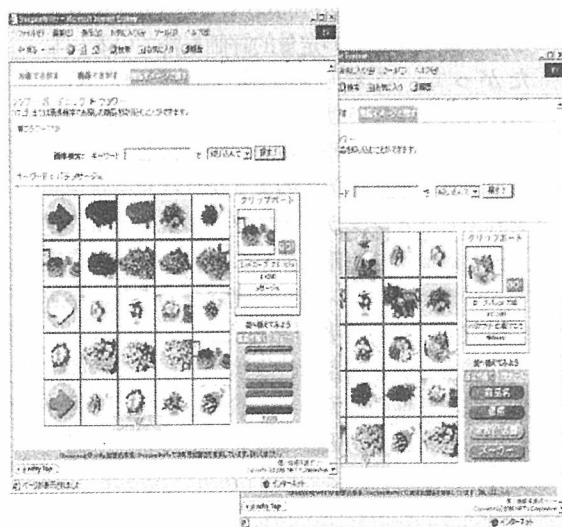


図11. Shopping@niftyにおける画像一覧検索サービスの表示画面例

4. おわりに

本稿では、マルチメディア情報を構成する各メディアの特徴および各メディア間の関連性を有効に活用した新しい検索モデル“パノラミック検索モデル”を提案した。また、パノラミック検索モデルに基づくマルチメディア情報検索システム“MIRACLES”を構築し、Web 文書検索、デジカメ画像検索、ショッピングポータルにおける画像一覧検索サービスに適用した。

MIRACLES の厳密な意味での評価は今後の課題であるが、幾つかの応用例を通じての評価について考察を加えておく。極めて主観的かつ定性的な評価になるが、まず、MIRACLES for WWW は、従来のテキストによる意味的な検索をベースとして画像による視覚的な検索を取り込んだものであり、少なくとも検索精度という観点からは、従来のテキスト検索に劣ることはなく、“画像の一覧性の良さ”を活用している点で、むしろ優れていると考えられる。また、デジカメ画像検索システムについては、パーソナルユースに限定すれば、3次元空間のビューで楽しみながら画像の整理ができるということで、好評を博している。ショッピングポータルにおける画像一覧検索サービスについても、商品を見比べながらオンラインショッピング感覚で楽しみながら商品を選択でき、従来のテキストベースの検索に比べて、商品の詳細を掲載しているショップの最終的なページにアクセスする率が向上したとの定量的評価も得ている。したがって、これらの応用例を通じて、MIRACLES の有効性を確認できたと結論づけても問題ないと思われる。いずれにしても、パノラミック検索モデルのような3次元情報空間を利用したインタラクティブな検索モデルの定量的な評価手法を確立すること自体が極めて重要かつチャレンジングが研究課題であり、早急に確立していく必要があろう。

今後も、引き続き、映像検索などへの機能拡張、改良、他のメディアの取り込みなどを図りながら、MIRACLES を発展させていく予定である。

参考文献

[1] 串間和彦, 赤坂樹, 紺谷精一, 山室雅司: “色

や形状等の表層的特徴に基づく画像内容検索技術”, 情処論, Vol.40, No.SIG3(TOD1), pp.171-184, 1999.

- [2] 波多野賢治, 佐野綾一, 段一為, 田中克巳: “自己組織化マップと検索エンジンを用いたWeb 文書の分類ビュー機構”, 情処論, Vol.40, No.SIG3(TOD1), pp.47-59, 1999.
- [3] 武者義則, 広池敦, 杉本晃宏: “類似画像検索における特徴量空間の可視化インタフェース” 信学論, J82-D-II, 10, pp.1626-1833, 1999.
- [4] 木本晴夫: “3次元探索空間を用いるインタラクティブな画像検索システムとその検索特性”, 情処論, Vol.41, No.3, pp.638-657, 2000.
- [5] 遠藤進, 指田直毅, 増本大器, 長田茂美, 棚橋純一: “画像情報とテキスト情報を統合的に利用したインタラクティブなWeb 検索システム”, 第5回知能情報メディアシンポジウム予稿論文集, pp.163-170, 1999.
- [6] Flickner, M., et al.: “Query by Image and Video Content: The QBIC System”, IEEE Computer, Vol.28, No.9, pp.23-32, 1995.
- [7] 高木幹雄, 下田陽久監修: “画像解析ハンドブック”, 東京大学出版会, 1991.
- [8] 村尾晃平, 安藤淳禎: “画像をキーとする類似画像検索システム”, 1998年電子情報通信学会 情報・システムソサイエティ大会, D-11-60, p.175, 1998.
- [9] 上原祐介, 柿元俊博: “マルチメディア情報検索のための3次元空間ブラウジング”, 第2回知能情報メディアシンポジウム予稿論文集, pp.221-228, 1996.
- [10] T.コホネン著, 徳高平蔵, 岸田悟, 藤村喜久郎訳: “自己組織化マップ”, シュプリンガー・フェアラーク東京, 1996.
- [11] <http://www.fosl.co.jp/miyoue/index.html>
- [12] 新井正敏, 椎谷秀一: “ブロードバンド時代の新サービス—インタラクティブCM/クロスメディア検索—”, FUJITSU, Vol.52, No.4, pp.285-291, 2001.
- [13] <http://shop.nifty.com>