

みんなくりポジトリ

国立民族学博物館学術情報リポジトリ National Museum of Ethnology

身装関連マルチメディアデータベースの作成と使用 経験

メタデータ	言語: jpn 出版者: 公開日: 2015-03-23 キーワード (Ja): キーワード (En): 作成者: 水田, 忍, 高橋, 晴子, 八村, 広三郎, 久保, 正敏, 大丸, 弘 メールアドレス: 所属:
URL	http://hdl.handle.net/10502/5516

身装関連マルチメディアデータベースの作成と使用経験

水田 忍 高橋晴子 八村広三郎
国立循環器病センター 大阪樟蔭女子大学 立命館大学

久保正敏 大丸 弘
国立民族学博物館

統合検索機能をもつ身装関連マルチメディアデータベースシステムについて述べる。このシステムは、タイプのちがう複数のデータ群を同時に検索することができる。まず、システムの概要について述べる。つぎに、使用経験にもとづきシステムの操作性を検討し、利点と問題点を明らかにする。統合検索機能の利点は、操作の簡略化、検索時間の短縮化がはかれ、また同一画面上での検索結果データの比較検討が可能なことである。問題点は、文献と画像のタイプのちがうデータに対して、共通の分析コードを用いているため、検索の精度に限界のあることである。本システムは、レファレンス用データベースとして考えている「身装専門電子事典」作成のために有効であることがわかった。

Implementation and Evaluation of Multimedia Costume Database

Shinobu Mizuta Haruko Takahashi
National Cardiovascular Center Osaka Shoin Women's College

Kozaburo Hachimura Masatoshi Kubo Hiroshi Daimaru
Ritsumeikan University National Museum of Ethnology

This paper describes the Multimedia Costume Database System which has the capability of handling several kinds of databases at the same time. After the outline and the implementation of this system are described, the result of evaluation of the system is mentioned.

Finally, the use of the system for developing the Digital Costume Cyclopaedia, which is now being developed, is also discussed.

1. まえがき

われわれは、1984年より服装を中心とする身装関連データベース「民博コスチュームデータベース(MCD)」を作成してきた。MCDは9つのデータ群によって構成されており、現在合計約20万件の文字、文書画像および画像のデータが入力されている。

今回、MCDのデータ群のうち、5つのデータ群を対象として、身装関連マルチメディアデータベースシステムを完成した。なお、このシステムは1993年より手がけてきたものである[1][2]。

本システムの特長は、複数のタイプのちがうデータ群を、同時に検索できる統合検索機能をそなえていることである。まず、システムの紹介、および前回報告[2]以後の改良点を述べる。つぎに使用経験にもとづいて、システムの操作性を検討し、利点と問題点を明らかにする。

また、このシステムの、「身装専門電子事典」作成のためのツールとしての役割を検討する。

2. 身装関連マルチメディアデータベースシステム

2.1 対象データ

本マルチメディアデータベースシステムの対象とするデータ群は、つぎの5つのデータ群であり、収録データ件数は合計約56,000件である。

(1) 日本語雑誌論文・記事(1967-)	41,600件
(2) 日本語図書(1868-1993)	10,276件
(3) 外国語民族誌	1,683件
(4) 関連画像資料	663件
(5) 日本語新聞記事	1,846件

関連画像資料と、日本語新聞記事については、現在まだテストの段階であり、それぞれの本来持つデータ件数のすべてをみつかるにはいたっていない。

(1)から(3)は、資料の所在を指示する二次的なデータ群であり、(4)と(5)は、資料の内容をもふ

くんでいる一次的なデータ群である。

データの文字属性情報は、MCDのデータ分析のための共通項目である合計32項目(表1)にそって分析され、情報が付加されている。

表1 データ分析項目

0 データ番号	17 民博標本名
1 著者名	18 現地名
2 論文名	19 民族・地域分類(OWC)
3 掲載誌名	20 特定地域名 (フリーワード)
4 巻号・ページ	21 時代
5 図書名	22 特定期間名 (フリーワード)
6 シリーズ名	23 フリーキーワード
7 版次	24 服装専門分類/身装概念 コード
8 事典項目	25 標本資料分析コード
9 事典整理番号	26 資料の所在
10 引用事典名	27 資料形態
11 出版地	28 使用言語
12 出版社	29 抄録
13 発行年月	30 NDC, UDCあるいは国立 国会図書館分類番号
14 画像番号	
15 所蔵者・機関	31 MCD図書分類番号
16 民博標本番号	

2.2 システム構成

このマルチメディアデータベースは、UNIXワークステーションSUN Sparc Station 10上で動作する。操作は、OpenWindows上に構築されたGUIにより実現している。DBMSには、商用DBMSのMyriad(TransAction Software GmbH)を用いている。

このMyriadは、クライアント/サーバー型のリレーショナルデータベースエンジンである。複数のデータベースの登録が可能で、それぞれのデータベースは計算機ネットワーク上に分散して置くこともできる。

Myriadの特長は、テキストデータの全文検索が可能で、データのすべての文字列を検索の対象と

し典拠ファイルを作成することである。

2.3 ユーザーインターフェース

前回報告 [2] 以後に、改良された点について述べる。

2.3.1 メインメニューの機能

システムを起動するとメインメニューウインドウが表示される(図1)。このウインドウは、検索対象のデータベースを選択し、検索条件を設定するのに利用し、検索実行をおこなう3つのワークスペースから構成されている。

今回、検索条件設定ワークスペースに、事典項目のボタンを追加した。これは、検索用語を設定した場合、その用語が事典項目に登録されている場合は、掲載事典名とページ数を表示するものであり、第4章で述べる「身装専門電子事典」を作成するために必要な機能である。

2.3.2 ウインドウ

従来からのメインメニューウインドウ、検索用語メニューウインドウ、検索結果一覧ウインドウ、詳細内容ウインドウにくわえ、以下のウインドウを追加した。

a) 検索条件一覧ウインドウ(図1)

メインメニュー内で指定された検索対象データベースと検索されたデータ件数を表示する。さらに、データを一覧表示するかどうかを確認するメッセージを表示する。

b) 身装概念コード解説ウインドウ(図2)

身装概念コードは、一種のシソーラスコードである。概念の近縁関係、階層関係を、言葉ではなくコードであらわしたものである。本ウインドウは、属性情報として表示された身装概念コードに対応する用語を表示する。これによってコードの意味が画面上で理解できるようになった。

c) 拡大画像ウインドウ(図5)

一覧表示された画像をクリックすると、詳細表示のウインドウが開き、さらに拡大画像を表示することができる。

拡大画像ウインドウは複数開くことができる。最大5件表示できる詳細内容ウインドウに対応す

る拡大画像を、同時に表示することができる。画像はPhotoCD化したものから、512×768のサイズのものを利用している。

d) エラーメッセージウインドウ

エラーメッセージを表示する。

2.3.3 検索機能

検索は、AND, OR, NOTの論理演算子を用いて作成される論理式による検索文によっておこなう。そして、同じ属性分析項目をもつ、タイプのちがう複数のデータ群にたいしても同時に検索をかけることができ、結果を一覧表示することが可能である。

検索項目のひとつである服装専門分類/身装概念コード(表1の項目24)による検索を、今回つぎのように改良した。この項目は、われわれが独自に作成した専門分類表[3]と身装概念コード表[4]に対応する項目である。

専門分類表は、ふたつの面からなるファセット構造をもっている。布地、ウイメンズウエア、和服などのモノそれ自体をあらわす縦軸(アイテム)と、生産から消費にいたるモノの流れをあらわす横軸(フローステップ)を組み合わせて、分類コードを設定する。分類コードは、アルファベット2桁と、数字1桁の3文字で構成されている。たとえば'EQ1'は、布地の色・文様などの表面特性をあらわす。

いままでは、すべてのデータに付与した、身装概念コードによる検索を優先してきたが、今回、分類コードによる検索機能も整備した。これによって文献索引データ群については、分類コードによる検索が可能となった。

身装概念コード表は、概念の近縁関係や階層関係などを表現するために、この服装専門分類表をさらに発展させて作成した一種のシソーラスである。シソーラスとおなじ機能をもっているが、用語に優先語、非優先語の区別をつけず、近い内容をもった言葉をグルーピングしコードづけをおこなっている。コードの最初の3桁は、専門分類コードをそのままつかい、さらに2桁および、小数点以下2桁までの数字で構成されている。

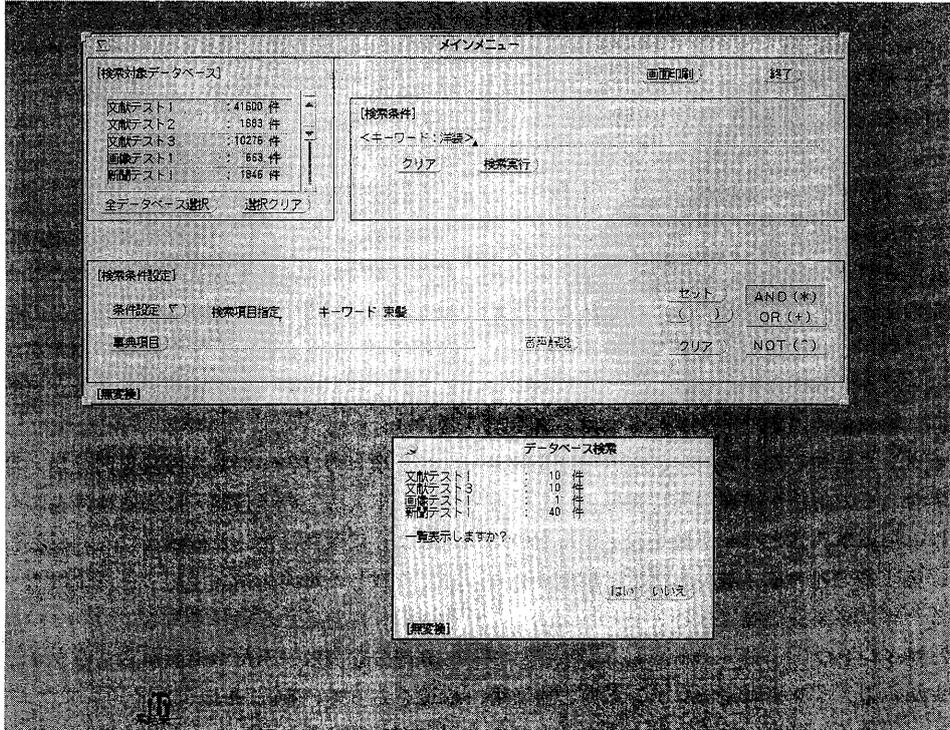


図1 メインメニューウィンドウおよび検索件数一覧ウィンドウ

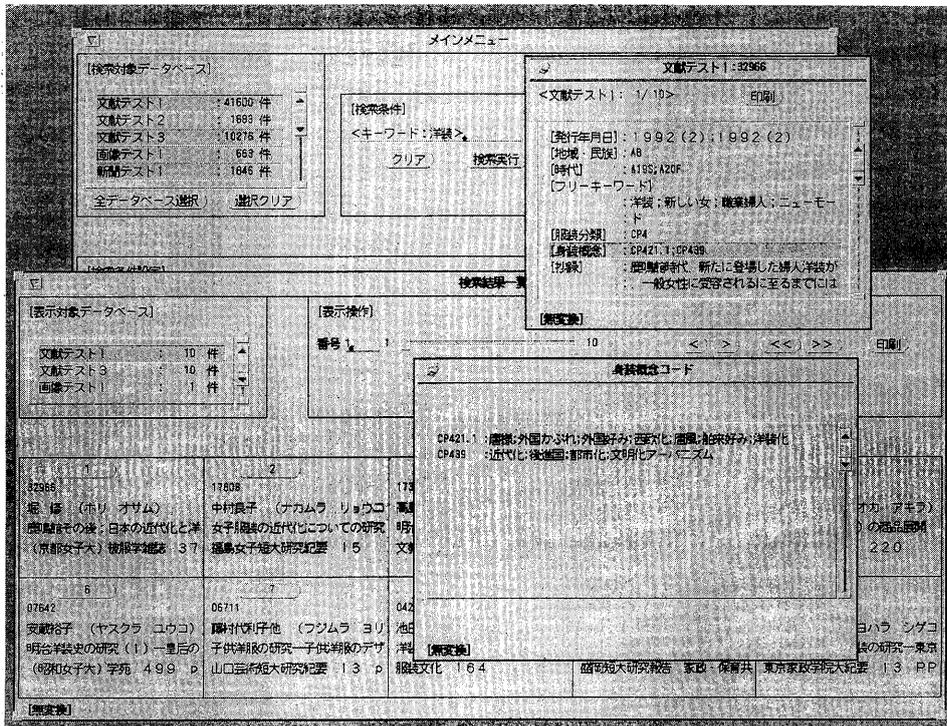


図2 身装概念コード解説ウィンドウ

たとえば、'EQ123' は、紋章と家紋をグルーピングしてつけられたコードである。

従来この身装概念コードを用いて検索をする場合、コードを直接入力して検索を実行してきたが、今回コードに対応する用語によっても検索できるようにシステムを改良した。たとえば、'EQ123' の場合、このコードを直接指示するほかに、「紋章」でも、「家紋」でも検索が可能となった。これによって、検索者はあたかも自然語で検索しているような印象をもつことができる。

同時に、さきにも述べたが、検索結果のデータの属性情報として表示された身装概念コードについては、コードに対応する用語を表示できるウィンドウを作成した。

2.4 検索例

図3、図4および図5は、前述した5つのデータ群のうち、外国語民族誌をのぞく4つのデータ群を対象として、キーワード「洋装」で検索をおこなった結果である。

図3は、雑誌記事索引データの検索結果を表示している。1画面に10件のタイトル一覧を表示し、さらに10件のうち1件の属性情報を表示している。

図4は、新聞記事文書画像の検索結果を表示している。1画面に5件の文書画像を表示し、さらに5件のうちの1件の属性情報、および拡大文書画像をかさねて表示している。

図5は、新聞記事文書画像にくわえて、画像データの属性情報、および拡大画像をかさねて表示している。なお、画像も文書画像とおなじように、1画面に5件の検索結果を表示することができる。

3. システムの使用経験と評価

本システムの使用経験のない3名の被験者が、システムの操作性に関する評価をおこなった。

システムの機能、ウィンドウ構成、検索文の作成法について簡単に説明したのち、各自が検索をおこなったが、操作手順については、3人とも容易に理解を示した。むしろ、マウスの左右のボタンの使い分けなど、機器そのものへ馴れることの

ほうに時間を必要とした。

ユーザーインターフェースの機能について、以下のようないくつかの指摘をうけた。

- (1) 検索中のサインが表示されないので、検索をおこなっているのかどうかかわからない。
- (2) 検索文を変更したいとき、検索文設定のための動作を最初から繰り返さなければならないため、操作が煩雑になる。用語選択のためのプルダウンメニューは、検索者が閉じないかぎり、表示されたままになっているほうがよい。
- (3) AND, OR, NOTのボタンが使いにくい。

(1)と(2)については、改善を試みる。(3)については、ボタンの位置について検討する。

つぎに、複数のタイプのちがうデータ群にたいして、おなじ検索文を用い同時に検索をかけた場合の、利点と問題点を記す。

利点はつぎのふたつである。

- (1) 複数のデータ群について、おなじ検索文をもちいて同時に検索をかけ、検索結果を得ることができるため、操作の簡略化、検索時間の短縮化がはかれる。

(2) 得られた結果について、同一画面上で、タイプのちがうデータを同時に表示することができるため、データの比較検討に有効である。

つぎに問題点であるが、身装概念コード(表1の項目番号24)は、タイプのちがうデータ群の共通分析コードとして用いている。最初、この概念コードを作成したときは、まずさしあたりは、文献データを分析するためという目的があったものの、将来、複数の資料・タイプにまたがって利用できる共通分析コードを意図していた。

しかし、経験上、文献データと画像データでは、どうしても概念のズレが生じることがわかった。実際、共通して使用しているコード、すなわち画像の分析にも使えるコードは、約2,000のうち500程度である。たとえば、画像の分析には、「消費者情報」あるいは「繊維産業」などの抽象名詞は不必要である。

また、文献検索時と、画像検索時の検索用語選択についての大きなちがいは、以下のことによる。

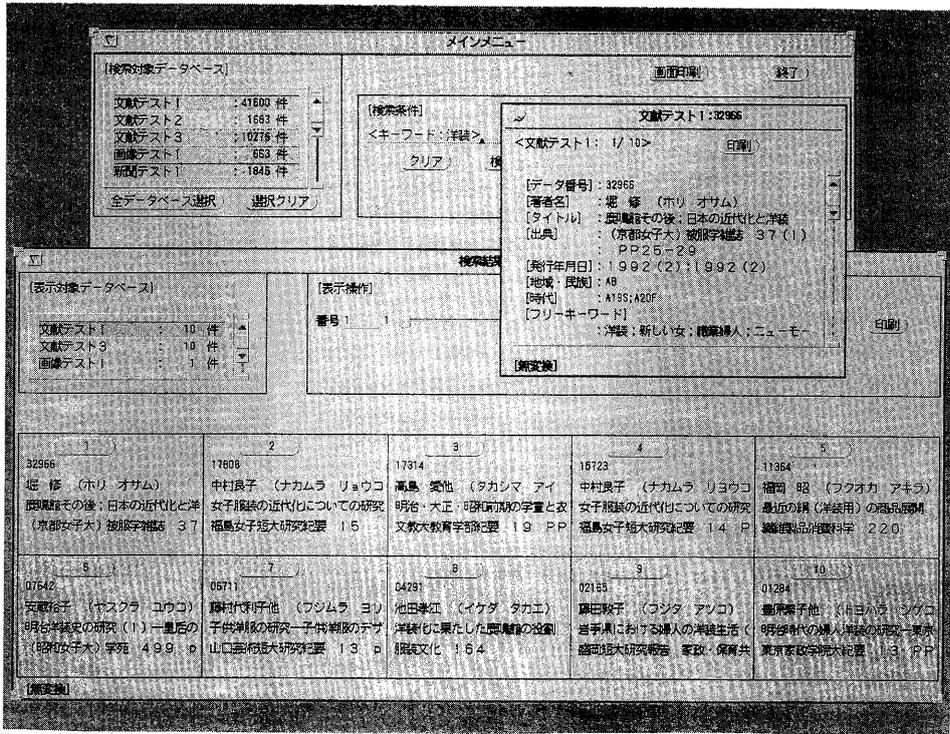


図3 雑誌記事索引データ検索結果

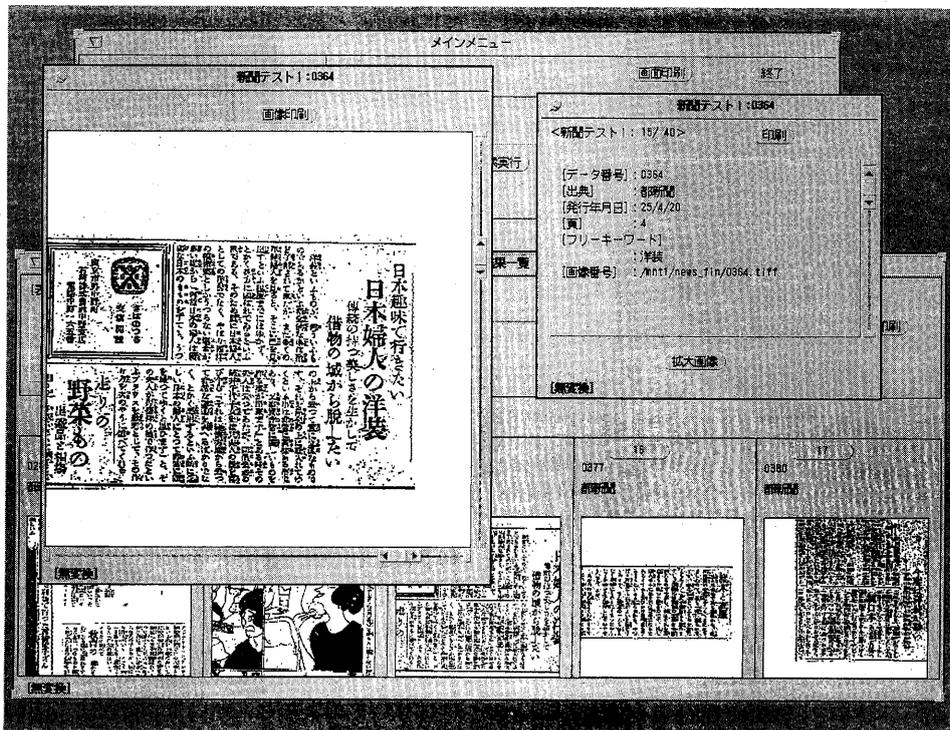


図4 新聞記事文書画像データ検索結果

文献検索の場合は、言うまでもなく、言葉を用いて検索をする。画像の場合は、イメージをイメージで検索する方法をべつにすれば、検索者の頭の中にあるイメージを言葉に変換して検索をおこなう。イメージを言葉に変換する場合、単語をならべるより、ある程度の長さをもつ文章で表現するほうが自然である。

以上のような理由から、文字データと画像データについては、まったくおなじシソーラス、あるいは身装概念コードのようなそれに代わるものを利用するのは、多少の無理があることがわかった。

また上記のことに関連して、画像については、論理演算式による検索法には、今後限界が生じてくるであろうと考える。

つぎに、以上のような利点と問題点をふまえたうえで、本システムを有効利用するための方法について述べる。

4. 「身装専門電子事典」作成のための本システムの利用

「身装専門電子事典」は、身装関連の情報提供者（専門司書や研究者など）が、レファレンスの際のツールとして利用することを想定して、現在作成中の約30,000語を収録予定としている事典形式のデータベースである。

専門家以外の一般の人々は、ある用語に関する説明や画像および文献などについて、特異な例よりも標準的な例を求める場合が多い。たとえば、画像を例にとつていうと、60年代のミニスカートの画像が見たいという場合、ひとりの女性がミニスカートををはき、全身が表現されている画像をまず、好む傾向がある。

経験上からいうと、このような単純な要求に答えられる画像はみつからない場合が多く、また、たくさんの画像のなかのひとつであったりする。これは、画像のデータベースを作成する側の問題であるが、作成者は普通、ある分野の画像を集めようという意図をもって収集する。しかし、ある用語に該当する適当な画像を収集するということは、能率が悪いので、なにか理由がないかぎり普

通このようなことはしない。そのために、汎用の画像データベースでは、ある用語についての標準的な画像のデータをふくんでいる確率は低い。

以上のような理由から、ある一定の基準により、ひとつの用語についてあらかじめ、標準的な画像、文献等を選択して入力しておくほうが、レファレンスツールとしては、有効であると考えた。それには、事典形式、すなわちひとつの用語についての説明、用例、関連する画像・図版および文献を対象としたデータベースが適当である。

現在MCDには、20万件のデータを対象とした9つのタイプのちがうデータ群のほかに、国内外の辞書30種類を対象に、関連用語を収集した「コスチュームサイクロペディア」というデータ群がある。この「コスチュームサイクロペディア」の用語を基本にして、ほかの文字データ群、および画像・文書画像データ群を有機的に結びつければ、事典形式のデータベースは作成可能である。

この「身装専門電子事典」を作成するためには、統合検索機能をそなえた、本身装関連マルチメディアデータベースシステムが利用できると考えている。3章で述べたようないくつかの問題点が残ってはいるものの、画像の分析・検索が用語のレベルで処理できるレファレンス・ツール用のデータベースの作成には、有効に利用できるであろう。

現状では、システムの作成と、用語の整理等のデータ加工をおこなっている。実際にシステムにデータを入力し、運用することは、今後の課題としたい。

5. あとがき

統合検索機能をもつ、身装関連マルチメディアデータベースのシステムの概要、および前回報告〔2〕以後の改良点について述べた。つぎに使用経験にもとづいてシステムの操作性を検討し、利点と問題点をあきらかにした。

また、レファレンス用データベース「身装専門電子事典」の作成には、本システムが有効に利用できることがわかった。

なお、この電子事典のデータは、その素材がか

なり蓄積されてはいるが、収録用語約30,000語に対応する適当なデータが、十分そろっているかといえ、著作権の問題もありそうとはいえない。今後も、データの蓄積を心がけるとともに、これにともなう著作権についても対処する方法を検討していく予定である。

[参考文献]

- [1] 八村広三郎、高橋晴子、久保正敏他“民博コスチュームデータベース”, 情処研報 人文科学とコンピュータ, 20-5, pp. 37-44(1993)
- [2] 高橋晴子、八村広三郎、久保正敏他“身装関連マルチメディアデータベースの構築”, 情処研報

人文科学とコンピュータ, 29-14, pp. 79-84(1996)

- [3] 高橋晴子“服装専門分類表の試み; 大阪樟蔭女子大学の場合”, ドキュメンテーション研究 27(8), pp. 369-374(1977)

- [4] 大丸弘、高橋晴子“服装専門検索語辞書 (MCDシソーラス) の構造” 国立民族学博物館研究報告, 10(3), pp. 681-723(1986)

なお、本研究は一部、文部省科学研究費重点領域研究1(課題番号 09204108)、および公開促進費(データベース)(申請番号7)の補助を受けていることを附記しておく。

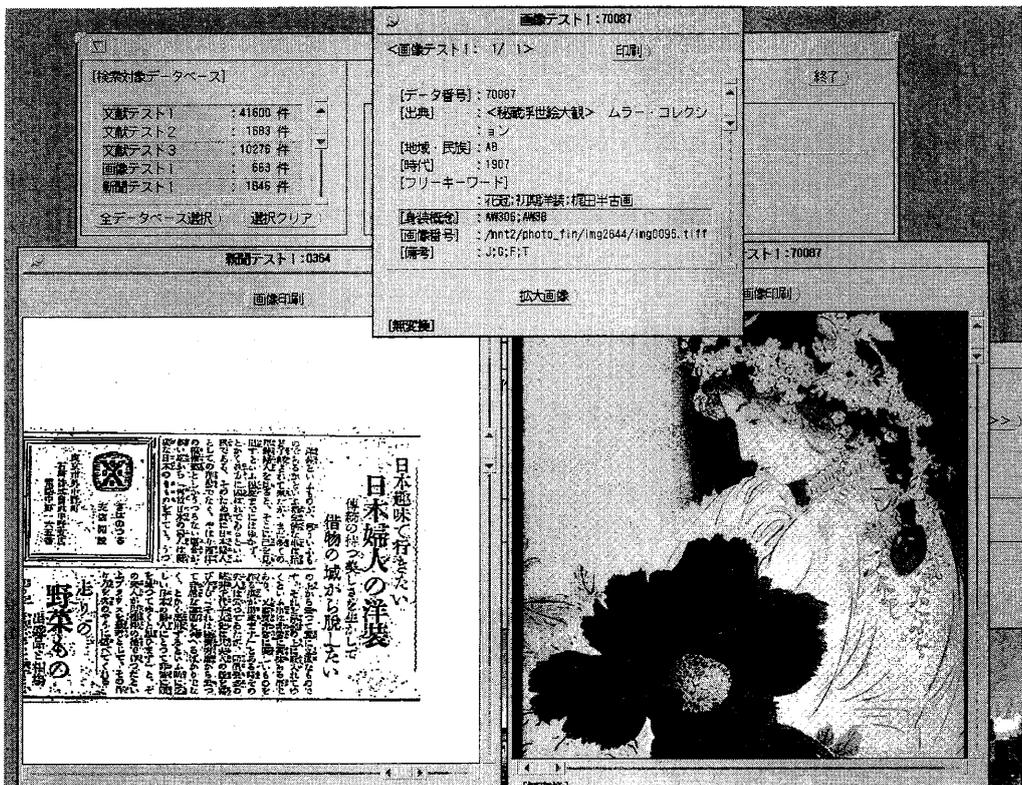


図5 画像データ検索結果