

みんなのポジトリ

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

共同研究の概略 一方法と成果一

メタデータ	言語: jpn 出版者: 公開日: 2010-02-26 キーワード (Ja): キーワード (En): 作成者: 洪, 政国 メールアドレス: 所属:
URL	https://doi.org/10.15021/00003574

第 II 部

共同研究の概略

——方法と成果——

洪 政 国*

要旨

民族学研究者にとって重要な研究資料である標本に関する各種情報・データ類を民族学研究者が活用するためのコンピューター応用システムを検討し、民族学研究用画像検索システム、画像類似検索システム、文化項目分類コード概覧システム等の試験システムを構築した。本共同研究は、人文科学系の専門家とコンピューター・情報処理系の専門家による共同研究であると共に、産学共同研究、さらに、人文科学系へのコンピューター応用研究や画像データベースシステムといった、コンピューターの研究事例が少ない分野でもある。我々は、専門分野を異にする参加者にとって興味を共通に持つことができる技術的課題を中心に、コンピューターシステムの最終利用者である国立民族学博物館の民族学研究者にとって役に立つコンピューター応用研究に努めた。本稿では、我々が選択した基本方針とその成果について報告する。

1 はじめに

民族学においては、世界の諸民族の文化的特色を明らかにするため、標本資料、映像音響資料、文献図書資料等が使われ、実証的、比較的に研究が行われる。この内、標本資料とは世界各地の諸民族が使用し、あるいは使用していた「もの」、つまり農業・狩猟・魚労等の生産に関する用具類、衣・食・住等の生活に関する用具類、紡織製品をはじめとした「もの」の製作技術に関する用具類、および冠婚葬祭に関する用具類等様々なものをいう[国立民族学博物館 1985]。一方、民族学研究においては、物理的対象物（もの）と人間の社会的行動（文化）との間には密接な関連があるという立場をとる『物質文化』[SCHLERETH 1985]、ある課題について多くの地域・民族事例を取り上げる『通文化的方法』[MOORE 1961]という基本的な考え方・研究方法があり、「もの」としての標本資料が研究に供される。この際標本は、民族学的見地から意味が付与されているものとみなされている。つまり、個々の標本は、誰もが使うもののインスタンスとして、文化的意味を持つものとして取り扱われ、実

* 日本アイ・ビー・エム(株)東京基礎研究所

世界の認識に使われる。

一般的に、標本資料の種類と数量の増加には際限がなく、用途が多様であるため、コンピューターにより管理が行われ、民族学研究に一層活用されることが期待される。特に、民族学研究では研究者の専門的知識を前提に創造的活動が行われるために、視覚能力の活用が必要であることや、各種研究資料、特に標本資料は「もの」として一般的に利用されること等から、標本を画像データとして利用することと、画像データの効率的活用のための画像データベースの構築の重要性が指摘される¹⁾。

一方、コンピューター技術、画像処理技術およびデータベース管理技術の発展・普及に伴って、画像データベースシステムの研究・発展が近年に至り大きく期待されるようになってきた [横矢・田村 1981; 安田 他 1985; 洪・六川 1985; GARDARIN and GELENBE 1985; CHANG 1985; 中村・石井 1987; 坂内・大沢 1987]。つまり、画像処理技術の研究や実用化の発展は過去15年の間に顕著であり、それに伴い画像を効率的に管理・利用することの重要性が高まってきた。また、データベース管理技術の研究や実用化の発展によって、データベース管理システムの高水準化やデータベースの高度利用化が強調されるようになり、画像情報を新たなデータ型として取り扱うための研究も本格化してきた。

また、これらの状況の中でコンピューターの応用分野が拡大され、いままで機械とか数理といったものから最も縁遠い分野とされてきた人文科学分野にも、コンピューターが普及するようになってきた。このため、理工系分野の利用者を対象とした従来の「開発主導型」の情報処理研究やコンピューターシステム構築に対して、「利用者主導型」のコンピューターシステムの構築や、人文科学分野における各種データに対する情報処理研究の必要性が強調されている [杉田 1987a, b]。

画像データベースの研究は、事務文書、図面、地図、医療画像等個々の事例についてすでに報告され [McKEOWN 1983; 篠田 他 1983; 横矢・田村 1984; 嶋田・江尻 1986; 野口 他 1986]、一部実用化レベルに達するものもあるが、画像データベースの概念や一般的構築法の確立には至っていない。一方、人文科学分野の研究を支援するための画像データベースについてはいくつかの事例が報告されている [小沢 1985; OIKAWA 1987; 打波 1987]。これらは古墳研究を支援するためにコンピューターグラフィックスを組み入れたものや、発掘物を簡易に画像化し検索するもの、現地調査のデータを使ってイベントの全体像を把握するためのマルチメディアデータベース等である。しかし、民族学研究の標本のように各々のデータが多様な意味を持

1) 本書「民族学とコンピューター」および「画像情報の利用と課題」参照。

つ場合のデータの管理や、多様で大量なものをカラー画像として一括して管理する方法、高度で容易な検索条件の入力と検索処理を行う方法等は残された課題といえる。特に、人文科学系では一般的に思弁的な作業が大きな比重を占め、処理作業が定型的ではなく個別的にすすめられ、取り扱う情報や媒体が極めて多岐にわたる等のため、高度で多様な統合化された情報処理技術が要求される。

国立民族学博物館（民博）ではすでに各種の民族学研究用データをコンピュータシステムに入力し、これらの有効活用を図る段階に至っている。標本も同様に各種の型のデータとして入力され、利用可能な状態にあることから、本共同研究においては主として、標本画像の蓄積と検索を中心にコンピュータの応用手法である画像データベースについて、その基本機能の基礎的研究と、それにもとづいたプロトタイプの構築を行うこととした [洪 他 1986; 久保 他 1986; 佐藤・井岡 1986; 洪・井岡 他 1987; 洪・黒川 他 1987; 井岡 1987; 井岡 他 1987; 佐藤 1987; 佐藤 他 1987; 黒川 1987; 黒川 他 1987; 井岡 1988; 佐藤 他 1988; HONG 1988; HONG *et al.* 1988; 橋原 他 1989; KUROKAWA 1989; 黒川 他 1990; HONG 1990]。また、民族学研究において共通の基礎的情報として使用されている HRAF コード²⁾をより有効に活用するために、パーソナルコンピュータ上でのコード概覧機能についても検討した。すなわち、これら民族学研究にとって重要な情報を有効利用するためにコンピュータをいかに応用すべきかについて研究したので、これら各テーマの基本的な考え方や方法、および成果の概要について報告する。なお、個々の技術的内容の詳細は本書で別途報告されている。

2 共同研究における課題と方法

人文科学系研究者とコンピュータの研究者による本共同研究における特徴と課題、およびこれらにもとづいて選択・実施した方法は以下に示す通りである。

2.1 課題

本共同研究は産学共同という特徴を持つが、共同研究を実施するにあたっては、「学」と「産」の違いよりも技術的、学問的観点における共通の課題に取り組むように努めた。この際の課題は次の3点にまとめられる。(1) 人文科学系における情報、

2) 本書資料編B「HRAF/文化項目分類(OCM)コード」、C「HRAF/地域・民族分類(OWC)コード」および本書「標本資料検索コードとしてのHRAFコードの利用について」参照。

データおよびユーザーを対象としている。具体的には、民博が所有する各種の民族学情報とデータ、ユーザーとしての民族学研究者を対象とする。(2) 人文科学系と、コンピューターという理工系分野の研究者による協力作業である。(3) 情報処理・コンピューター技術分野における学問的、技術的課題の解決をめざす。

コンピューターの人文科学分野への応用は、民博においてすでに実施中であり多くの成果を見るに至っているが【国立民族学博物館 1985】、日本アイ・ビー・エム(株)から参加した研究者にとっては初めての経験であり、民族学は勿論人文科学とは馴染みの薄い分野である。また、民博が所有する各種の専門情報・データについてもほとんど何の知識も持っていない状態であった。このように、対象となる専門分野における知識の乏しさをいかに克服するかが大きな課題である。また、この分野の専門家と協力してコンピューターの応用方法を研究するためには、相互の理解が必要である。特に時間に制約されている共同研究の期間内で、いかに相互の専門分野の違いを乗り越えて理解しあえるかが、共同研究成否の鍵ともいえる。

民族学情報、特にその画像情報を民族学研究者にとって効率的に蓄積し、検索するための情報処理技術における課題として次の3点があげられる。すなわち、(1) 画像、特にカラー画像は1点1点のデータ量が大きい(通常は3メガバイト)ことに加えて、データベースの対象となる画像データ全体のデータ量も極めて大きいために(テラバイト単位)、これら大容量のデータをいかに効率的に蓄積し、実用的な時間で検索するかといった、物理的な課題が解決されなければならない。次に、(2) 効率的な画像検索の実現がある。従来のデータベースにおける検索は、統計データベースの例のようにキーワードによるものが主であり、画像のように非コード化情報を扱うものではなかった。画像の検索は、その内容の記述が一様に決まらず極めて複雑であるため、従来のように定型化された文字数値を用いるだけでは効率的な検索が困難である。このため、従来とは異なり画像の特徴を考慮した検索方法が検討されなければならない。最後に、(3) 画像データベースといったコンピューターシステムの利用者が人文科学系の研究者であることから、馴染みやすいマンマシンインターフェースの実現が必要である。コンピューターシステムの機能がいかに優れていても、利用者にとって使い難いものであれば価値の乏しいものになる。特に今回の共同研究は、人文科学系へのコンピューターの応用をめざしていることから、人文科学研究者にとって馴染みやすいユーザーインターフェースは極めて重要な研究課題といえる。

2.2 方 法

我々が選択した共同研究の実施方法は次の5項目に整理される。すなわち、(1) 応用指向にもとづいたプロトタイプ構築、(2) 民博および民族学研究を知ること、(3) 民族学研究者と情報処理研究者との間の役割分担、(4) 早いサイクルでのプロトタイプ構築と頻繁なフィードバック、(5) 既存の最新技術の応用にもとづいた研究価値の追求等である。

民族学研究者とコンピューター研究者にはそれぞれの専門を尊重し、専門の立場で協力しあうことに努めた。そのため、それぞれの分担を次のようにした。民族学研究者は、基礎データの提供とデータをコンピューターシステムで使用する際の要求内容の提供、およびプロトタイプを実際に使用して、その評価と改善点を指摘することとした。一方コンピューター研究者は要求内容を分析・定義し、コンピューターシステムをデザインすると共に、機能やシステムを実現するための応用技術を研究し、プロトタイプを構築した。

我々は、共同研究を遂行する上で抽象的な論議はできるかぎり避け、具体的な応用分野である民族学分野における「ナマのもの」にできるかぎり直接触れるように努めた。具体的には、標本をどのように収集し、記録し、整理し、利用するのかについては民族学研究者から直接聞くと共に、これら原データをコンピューターに入力する過程（標本画像入力、標本情報カード作成等）や、それらの管理については担当者から直接聞き、見学した。我々は、12名の民族学研究者と各々1回およそ3時間、この他にもおよそ10名の民族学研究者とも散発的であるが聞き取り調査を行った。さらに5名の民族学研究者に対しては、標本資料を検索する際に必要とする29種類の検索項目³⁾についてアンケート調査も行った。

これらの結果、研究者が個人で管理できるカードの点数、すなわち標本の数はおよそ2000点から3000点であり、これらを整理したフォルダーの場合はおよそ100点から200点までであろうことが理解できた。これを、民博がすでに所有する標本画像⁴⁾と、標本資料の情報カード⁵⁾を考慮した時の情報量は、およそ14ギガバイトから21ギガバイトに達する。すなわち、コンピューターを使用した際に個人として取り扱えるデー

3) 本書「属性情報を用いた標本検索」表1参照。

4) カラーおよび濃淡で標本1点につき7メガバイト。本書資料編D「標本画像自動処理装置」参照。

5) 33属性項目が記述され、データ量は標本1点につき平均2.1キロバイト。本書資料編A「情報カード」および本書「標本資料の整理と活用——コンピューターの利用——」参照。

タ量は、少なくともこれら以上であることが必要とされる。また、検索に必要とする項目については、提示した29項目に対する必要度は研究者によって異なるものの、全体的には29項目全てが必要であることがわかった。このことから、利用者全員に対してはこれらの検索項目全てが使用可能とすべきであると共に、個人が融通性を持って自由に項目を選択利用できる環境が必要といえる。これらの値を参考にして、標本画像のデータベースにおけるシステム構成、各種データの蓄積・検索機能、およびこれら操作を容易に行うためのユーザーインターフェースを検討することとした。

また、画像の内容にもとづいた検索方法を研究する上で、標本の実物を民族学研究者はどのように見ているのかを基本にしようとした。形状特徴をもとに内容検索するために「かご」の画像を、また色特徴をもとに内容検索をするために「仮面」の画像を選択した。これら標本の民族学的特徴づけや分類をどのようにするのかを知るために、「かご」を研究対象としている民族学研究者と、色および「仮面」について研究を行っている民族学研究者にそれぞれ直接面談し、聞き取りを行った。

利用者が要求するコンピューターシステムを適切かつ効率的に実現するために、プロトタイピングが有効である。これは、システム利用者とシステム構築者が密に接触し、お互に曖昧な要求定義の内容を「ステップ・バイ・ステップ」でプロトタイプを構築しながら、「トライアル・アンド・エラー」で要求定義を明確にしようとするものである。そのためには、両者が頻繁に接触すると共にプロトタイプ構築のサイクルをできるかぎり短縮し、タイミングの良い改良に努めることが必要である。本共同研究では、目標とする画像データベースシステムの基本的な概念（イメージ）を示すための最初のプロトタイプを共同研究計画作成後5カ月で構築し、使用者に提示した。まず、構築者が実演して全体の操作の流れと基本機能（操作機能、操作速度、操作性、画質、馴染みやすさ等）を一通り見学することから始めた（この際、詳細な説明や仕組みについてはできるかぎり避け、操作の説明に限定することが重要である）。ついで、使用者が直接操作を行ったり、詳細な説明を参考に、使用者側からの具体的な評価、印象、それまで持っていたイメージとの違い等を聴いて、次のバージョン構築に反映させることとなる。最初のバージョン構築からおよそ3カ月で次のバージョンを構築し、その後さらに使用者側から提供される評価・意見等にもとづいておよそ2カ月周期で改良していった。このような短期間でのプロトタイピングを可能にするためには、使用者側の積極的な参加意識と、忌憚のない意見交換に加えて、プロトタイプ構築者のプログラミング能力の高さが不可欠の要因である。

このように我々が採用した方法は、民族学研究者が日常の知的専門活動において

「人手による情報処理」をどのように行っているかを、コンピューターシステムの最終利用者である民族学研究者と直接面談して知ることであり、またプロトタイプの構築と利用者による評価・意見提供の反復を通して、利用者にとって馴染みやすいコンピューターシステムを段階的に実現しようとするものである。この方法は、コンピューターシステム利用者の心的特性を把握するための心理学的方法として、近年特に注目をあびている「思考口述記録法」[ERICSON and SIMON 1984; 加藤 1985; 甲・加藤 1987] に準じるものでもある。また、知的専門活動を支える技術としてカードやフォルダーといった文具類を活用する情報整理の方法 [梅棹 1969] をコンピューターシステム上で電子化し、再現する試みでもある。このような実世界を疑似的に実現することの有用性は他でも指摘されている [CARROL and THOMAS 1982]。

民族学研究者が本共同研究のように、コンピューターや情報処理といった馴染みの薄い分野に積極的に参加し、その専門性にもとづいたコンピューターの応用研究を実現するためには、民族学研究者が本共同研究に対して継続して関心を抱くことが必要である。このために、関心のあるデータを対象とすることは意義が大きい。そこで、本共同研究に直接協力参加する民族学研究者が試験データをそれぞれの関心にもとづいて選択することにした。その結果、標本資料2725点が選ばれた。その内容は刃物類が389点、容器類が759点、玩具類が920点、仮面類が657点である。前三者は東、東南、南アジアの国々で収集されたものであり、仮面類は全世界的に収集されたものである。民博が所有するこれらの原画像データ量（カラーと濃淡）は1点につき7メガバイトであり、試験データ全体で19ギガバイトとなる。また、これら標本の情報カード上に記述されている属性データは、全部で5.7メガバイトになる。これが、画像データベースを研究する際に具体的に扱う原データの量である。これ以外に、民博が対訳として刊行した日本語 OCM (『文化項目分類』) [マードック 他 1987]⁶⁾ も使用した。

最後に、情報処理およびコンピューター技術の研究価値の追求がある。画像データベースと、人文科学系へのコンピューター応用という2つの技術的目標は、情報処理分野やコンピューターの応用分野においても新しい試みであることから、いくつかの点において高度な研究努力が必要とされる。具体的には、大容量の画像データを効率的に蓄積・検索するための画像圧縮方法やマルチウィンドウの活用、高速画像表示方法といったデータ処理技術、大容量外部記憶装置の活用や高速転送、パーソナルコンピューターの活用等システム統合化技術、画像データベースの一般的概念や構築法、

6) 本書資料編B「HRAF/文化項目分類 (OCM) コード」参照。

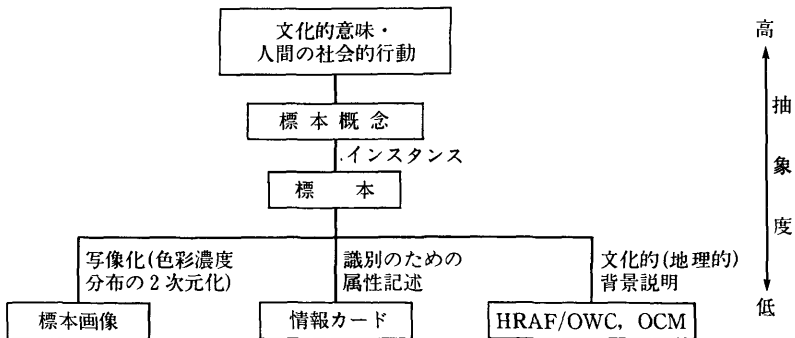
文科系利用者のためのユーザーインターフェースの実現等があげられる。

3 標本資料の情報とデータ表現

ここでは、共同研究における直接の研究対象である標本資料を、情報処理の立場からどのように認識するかを整理する。

3.1 標本資料の情報

標本の種類は、釣針のような小さなものや家屋のように大きなもの、石製のもの、皮製のもの、単調な形態、極めて精巧な細工が施されているもの、素朴な天然の色合を持つもの、一目でなんらかの意図が感じられる人工色で塗り潰されたものまで、様々である。また、収集地は世界全域にわたり、多様な地域や民族で使われているかあるいは使われていたものである。基本的には、これら標本は、民族学研究者による現地調査を通して、文化的意味を持つ概念としての標本集合からインスタンスとして収集された具体例である。例えば、「ここにある“カゴ”」について述べると次のようになる。「この“カゴ”は、『インドネシアのバタック族が収穫に用いる“カゴ”』の“カゴ”」で、固有の文化的意味や人間の社会的行動が説明づけられ、他の「カゴ」と識別されて、文化的特色が実証的、比較的に究明される。民博においては、これら標本の性質や意味づけ、説明づけを次の3種類の形態を用いて行っている。つまり、属性項目別に記述する「情報カード」、汎民族的、汎文化的な分類体系を持つ「HRAFコード」、および標本の写像体としての「デジタル画像」である（図1）。このよう



(民博における標本資料の3種類の情報源)

図1 標本資料の情報構造

に、各々の標本はこれら3種類の情報源（データ源）によってアイデンティティが具現化され、コンピューターシステムでの処理の対象となる。この際、民族学研究者にとって意味のある情報処理をコンピューターシステムで行うためには、これら異なるデータ型をマルチメディアとして統一的に取り扱うことが必要である。

3.2 標本資料のデータ表現

民博の情報カード上で記述されている属性項目を用いて、利用者から要求される利用要件を反映させて概念スキーマを設計する。このため、実世界のデータを表現する上で一般的なモデルとして広く利用され、データ間のマクロな関係をわかりやすく表現する実体-関連モデル [CHEN 1976] を用いることとする。このモデルは、既存のデータベースシステムで実現されている基本的データモデル（関係型モデル、ネットワーク型モデル、階層型モデル等）の記述能力を合わせ持つことから、論理設計や物理設計等を具体的に扱う上で有効である。なお、標本画像が持つ意味については本稿では扱わない。これは、画像がそれぞれの標本の実物写像体であることから、画像をそのまま表示することによってのみ最終的には意味をなすためであり、画像検索に必要とされるものは属性項目としてすでに記述されているためである。また、HRAFコードについてはこの意味づけの体系化はすでにでき上がっていることから、意味の取り扱いにこだわるよりも、このコード体系をコンピューターシステムで具体的に扱う方が重要と思うからである。

それ自体が概念として存在し、他のものと区別できる「実体」をまとめた集合である「実体集合」と、これら実体集合間の関連を表現すると図2のようになる。また、

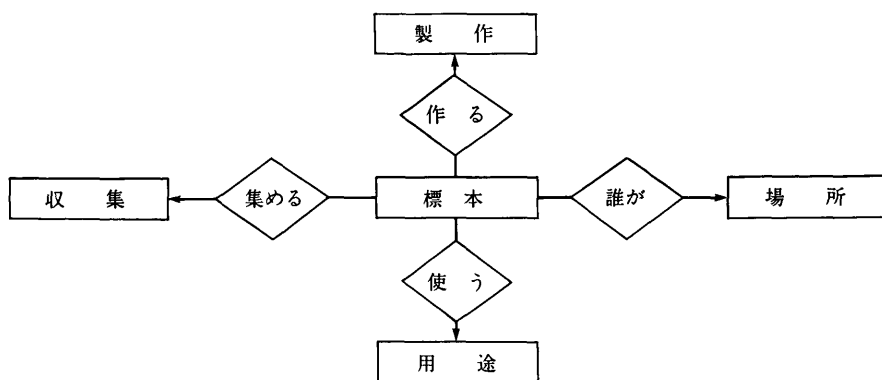


図2 実体-関連モデルによる標本情報カードのデータ表現（実体集合および関連について）

表1 実体-関連モデルにもとづく標本資料の実体集合と属性

実体集合	属 性
標 本	標本番号, 標本名, 標本種類, 大きさ*, 形態*, 色彩*
製 作	製作地, 製作年代, 製作者, 製作法・材料
場 所	現地名, 使用地, 使用年代, 使用民族, 使用者, 変遷・分布
用 途	用途・使用法
収 集	原収集者, 入手状況, 収集年月日, 収集地

(*は民博所有の情報カード上にはない属性項目で、本稿において試験的に追加したものである。)

図中の各実体集合の持つ共通性質、すなわち属性は表1の通りである。

ここでは実体集合は「標本」、「製作」、「場所」、「用途」、「収集」の5つに分類される。実体集合「標本」は、人間の社会的行動や文化的活動が具現化されたものであり、「製作」、「場所」、「用途」といった実体集合は、人間が行うこれら行動や活動を積極的に記述する要素である。また、実体集合「収集」は、同様に要素ではあるが、文化的意味づけという観点から見ると消極的な要素といえる。実体集合「標本」と、これら他の実体集合との「関連」は全て多対1対応にあり、各関連は各々の実体集合の活動を表現している。例えば、各標本は1つの製作法、1つの使用者、1つの用途目的を持つものである、といったものである。

このように、民族学研究の対象である標本は多様な意味づけがなされるものであり、それらを1つの集合として扱わなければならないデータ集合体というべきものである。

4 標本画像データベースシステムの利用者と利用方法

ここでは、本共同研究が対象とする画像データベースといったコンピュータシステムの利用者が、コンピュータシステムを利用する上でどのような特徴を有するかについて、利用者と利用方法とに分けて整理する。

4.1 利用者

まず利用者について整理すると次の通りである。(1) 利用者は一般的には人文科学者であり、(2) 民族学の専門家であると共に、(3) 個々の研究者である。これら利用者は、コンピューターによるデータ処理の専門家もしくはその訓練を十分に受けた人達ではなく、一般にコンピューターシステムには馴染みが薄く、プログラムを自前で

は作成しない半面、コンピューターシステムの使用頻度は比較的高い。このため、コンピューターシステムにおいては利用方法やその内容および操作方法が理解しやすく、機能の習得が容易であることが必須といえる。一方、この *Ease-Of-Use* に加えて、我々が対象とする利用者は民族学分野の専門家としてこれらコンピューターシステムを自己の研究活動に使用するため、創造的活動を支援する高度な機能と個々の専門的利用を満たす差別的機能が必要である。また、利用者は各々個別の研究のための道具として画像データベースシステム等を使うため、個人のコンピューター使用環境は勿論、利用者の研究専門領域と関連領域を幅広く取り扱うためのコンピューターの共有環境が提供される必要があると共に、個人環境と共有環境の両者間での高度な融通性を持った相互切り替えが必要とされる。ここでいう「個人のコンピューター使用環境」とは、1つのコンピューターシステム上で大型計算機やパーソナルコンピューターを使用して物理的に実現されることと同時に、利用時における利用者間での機能の上での差別化を意味する。

このような利用者の要求を満たすためには高度な機能を広範囲に提供すると共に、これら機能を容易に、かつ、利用者独自の方法で使用できることが必要である。

4.2 利用方法

一般に、画像データベースシステムとしては図3のような機能を備えていることが要求される。つまり、実世界にある実物を各種のデータとして入力することや、そのデータの蓄積、検索・表示、検索画像データの特徴抽出・強調処理や統計処理等の各種分析、最終利用形態としての主題図編集、印刷等である。検索や処理されたデータは個人用として保存され、自由に再利用されることが要求される。この内データ入力については、すでに民博で画像化、属性の記述および HRAF コードがふりあてられているので本共同研究ではこれら既存のデータを使用するだけとなる。検索された画像データを処理し、主題図を作成することは利用者によってその必要性が異なり、処理内容が極めて多岐にわたる。一方、画像の検索は利用者全てにとって必要とされる機能であり、画像データベースシステムを利用する上での第一歩といえる機能である。このため、本共同研究では、画像データの蓄積と検索、および検索画像データの保存機能について扱うこととした。

民族学研究においては次の3種類の画像検索が考えられる。すなわち、(1) 属性値から画像の検索、(2) 画像から画像の検索、および (3) 画像から属性の検索である。属性から画像の検索においては、「もの」としての標本とその画像としての特徴にも

とづき画像が検索されるが、ここでいう画像とは「もの」としての標本画像であり、その文化的背景としての背景画（「もの」が使われている様子とか、作られている様子等の風景画）である。画像から画像の検索は、ある参考画像の一部または全体的特徴をもとに類似（内容）検索を行うものと、部分について抽出検索を行うもの、および検索表示された標本画像からその背景画を連動的に検索する場合とに分けられる。いずれの場合でも、「もの」としての標本画像と背景画、および各研究者のフィールドノートに描かれたスケッチ等が対象と考えられる。検索表示された画像から属性を検索することは、民族学研究においては重要である。すなわち、「もの」としての標本とその文化的背景としての背景画（風景画）はどの研究者にとっても客観的なもの

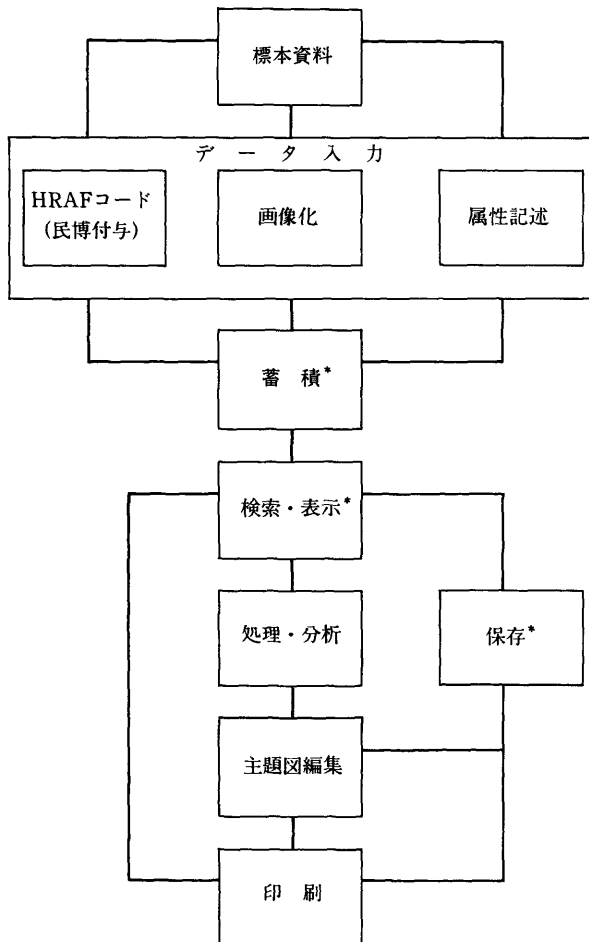


図3 画像データベースの基本機能と民博標本データ（*は本共同研究での対象）

であり、それにもとづいて各研究者が固有の研究目的に利用するために文字・数値として表わされる属性値が必須となるからである。

画像の検索機能は画像の特徴と、検索者が持つ対象画像についての記憶や検索条件の正確さ等に対応する必要がある。画像の内容は非常に複雑であり、画像が持つ情報量は極めて多い。このため画像の内容を一度に記述することは不可能であり、何回も追加記述され、利用者の記憶も曖昧になりやすく検索時の条件設定が曖昧となる。一方画像データベースがこれら検索条件に対応するためには、明確なキーワード処理によるものから曖昧な記憶にもとづいて描かれたスケッチの一部を頼りに検索処理を行う機能まで、多様に提供されなければならない。

各標本のアイデンティティを具現化するものとして、標本情報カードの属性項目がある。この属性データは HRAF コードと異なり、その意味づけや体系化を行う必要のあるものである。これは、標本について利用者がどのように見ているかに係わるものであり、検索のための概念スキーマを設計する上で基本をなす。そこで、民博の利用者が検索時の利用要求としてどのようなものを持っているかについて、アンケート調査と聞き取り調査を通して分析した。その結果、標本を検索する上で情報カード上の属性項目の重要度、必要性は個々の研究者によって異なるものの、全ての属性項目を検索時に使用できるようにすることが必要であることが理解できた。さらに検索照会に見られる特徴として、表2のように検索を繰り返しながら段階的に検索範囲をしばりこみつつ、徐々に検索内容を特定化していく方法をとるようである。この方法は、検索条件の曖昧さにも対応できるものでもある。

これら検索の機能と操作において重要な事柄は、画像データと属性値とが相互に連動されて検索に使われること、および操作時に個人化が提供されることである。異なるタイプのデータが連動される必要があることは、それらが各々の標本資料を補いあって表現していることから容易に理解できる。一方、個人化については、利用者が持っている操作技術のレベル、検索時における興味のおきどころ、また重要度や必要度が各自で様々であることに対応するためである。これに加えて、検索評価としていかなるデータベースシステムにおいてもトレードオフの関係にある、再現率と適合率

表2 検索範囲の段階的なしばりこみ

最初の設問	2回目の設問	3回目の設問
誰が（民族名） 何を 何の目的に使う	どこの誰が どのような何を 何の目的にどのように使う	どこのどのような誰が（社会的身分） どのように作られたどのような何を 何の目的にどのように使う

の内どちらを重視するかを選択することが必要である。本共同研究のように民族学研究者を利用者とする場合、再現率を重視することが適切と思われる。つまり、検索照会処理の曖昧性、入力ミス、各研究者固有の検索、および検索対象画像のデータ量の大量性を考慮すると、その適合率は各利用者に大きく依存するためである。しかし、再現率を高めるためには、雑多な検索結果から検索対象となるデータをしぼりこんだり、画像表示による視認の活用が有効である。

5 標本画像データベースシステム

ここでは、本共同研究における主要な研究対象である標本画像の蓄積と検索の方法を検討し、これにもとづき構築した試験システム「民族学研究用画像検索システム (CIRES: Color Image Retrieval System for Ethnological Studies)」の全体像と基本的な機能、特徴および評価について報告する。ここで対象としている画像はそれぞれ1件全体を1つのデータとして扱うものであり、画像の内容・部分を対象とするものは後節の「6 標本画像類似検索」において報告する。また、画像データベースシステムにとって主要な技術である画像圧縮・表示方法と画像の蓄積方法、属性データベースの作成と検索処理方法、視覚化ユーザーインターフェース等の詳細については本書の中でそれぞれ「画像の表示とその圧縮」、「画像データの蓄積」、「属性情報を用いた標本検索」、「民族学研究用画像検索システムの視覚化ユーザーインターフェース」等別途の論文として報告されている。

さらに、検索のためのシソーラスとして HRAF コードを活用するための最初の試みとして、文化項目分類コードをパーソナルコンピュータ上で扱ったが、これについては第7節においてその概略を、その詳細は別途論文（「HRAF/文化項目分類コードの概覧表示」）で報告する。

5.1 基本概念とシステム構成

画像データベースを研究する上で重要な技術的課題として、次の3点に注目した。つまり、(1) 大容量画像データの効率的蓄積方法、(2) 効果的画像検索方法、および(3) 民博利用者に親しみやすいマンデータベース・インターフェースである。これらは具体的にはさらに次のように細分化され、これらを解決することが CIRES で実現しようとする基本機能といえる。すなわち、(a) 標本資料に関する多様な情報が統一的に扱える。画像、標本情報カードに記載されている文字・数値データ、HRAF

コードが統一的に扱えるようにする。(b) 検索は、標本情報カードに記載されている文字・数値データと共に画像を見ながらでも行えるようにする。(c) 検索条件の入力や検索処理操作が高い融通性を持って行われる。(d) 検索操作方法と内容が理解しやすく、操作性が良い。(e) 一般性が高い大量な画像やデータを蓄積し提供するための共用環境と共に、個々の利用者が自分に合った方法で管理・利用できる個人環境の両方が用意されている。(f) 継続して増加する大容量の画像やデータが蓄積・管理できる。(g) 色、空間解像度等精度の高い画像が表示できる。(h) 操作の応答速度が早い。

ここで (e), (f), (h) が技術課題 (1) に該当し, (b) と (g) が (2) に, そして (a), (c), (d), (e), (h) が (3) に該当する。

以上の課題を解決するための基本的なアプローチは次の4つである。(1) マイクロメインフレーム結合のシステム構成によって、大型計算機側では大規模データベース資源の共有のための共用環境を、パーソナルコンピュータ側では操作の融通性に富む個人環境を実現する (図4)。(2) 関係型データベース処理にもとづく文字・数値

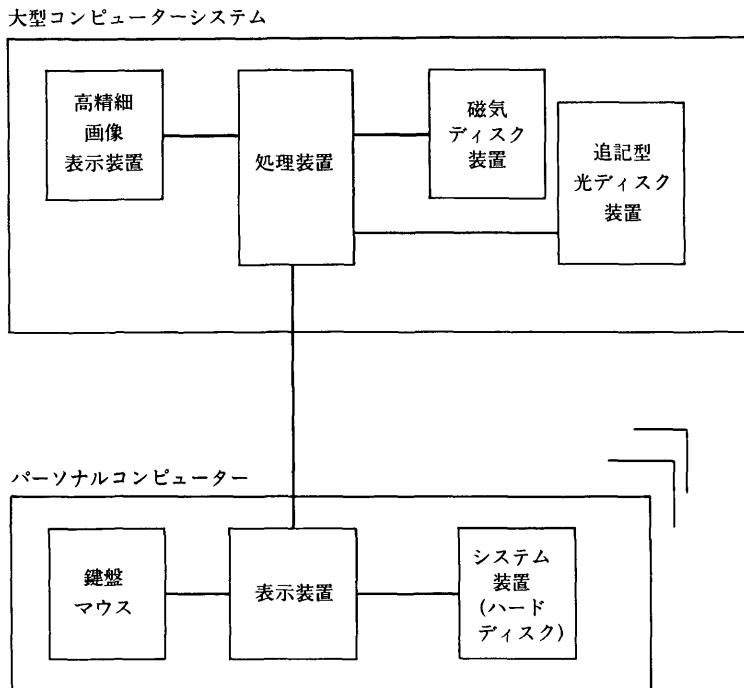


図4 標本画像データベースシステム構成の概要

データの属性検索と、カラー画像の概覧を組み合わせた検索を実現する(図5, 図6)。この場合、属性検索は繰り返し検索を可能とし、候補の対象を大型計算機と対話的に
行いながらしぼりこみ、高い検索効率を実現する。属性検索に続き、パーソナルコン
ピューター上では縮小したカラー画像(カタログ画像)を属性データと共にマルチウ
ィンドウ内で統一的に扱い、多数を同時に高速に表示して概覧できるようにし、検索
候補を視覚にもとづいてしぼりこんで見つけるようにする。視覚にもとづいた分類、
比較、保存ができるようにする。これは、例えば、印刷された商品カタログを見て希
望の商品を探すような感覚を与えることを意図している。あるいは、トランプのカー
ドを切ったり、任意の順番で並べるような感覚ともいえる。(3) 多階層構成による画
像の蓄積と表示によって、高い検索効率と応答性を実現する。このため、できるかぎ
り高速に展開・表示する画像圧縮・伸長方法を実現し、大きさを異にする画像を作成
し、蓄積媒体として磁気ディスクと光ディスクを用いる。ランダムかつ頻繁に、高速
にアクセスされる画像、つまり縮小された小サイズのカラー画像は磁気ディスクに蓄

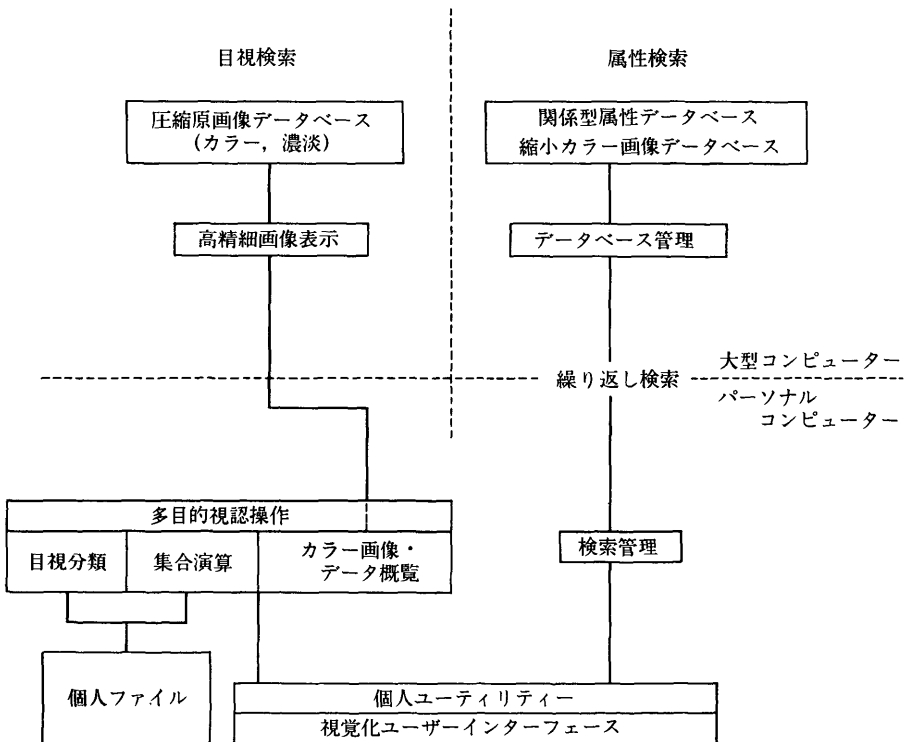


図5 標本画像データベースシステム機能の概要

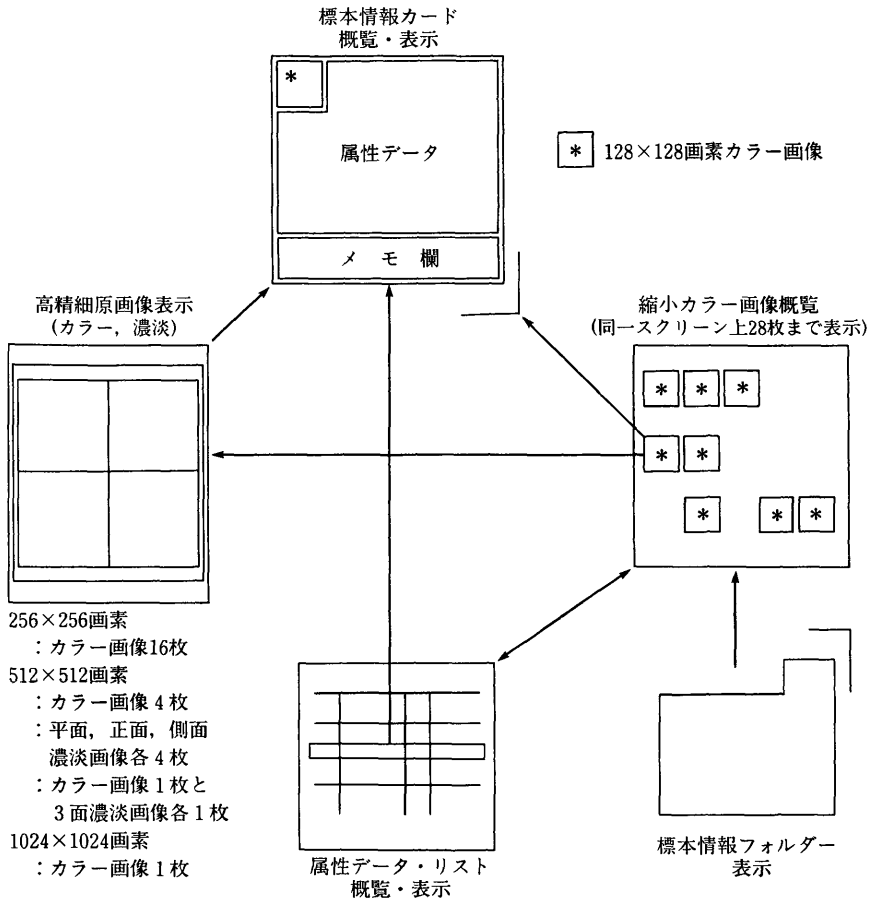


図6 目視検索のための画像およびデータが多階層概観・表示

積する。アクセス頻度が高くなく、かつ転送速度や画像表示に高速性が要求されない反面、色数が多く、空間解像度が高い高精細の画像は大容量の光ディスクに蓄積する。光ディスクは、現物にできるかぎり忠実な高画質画像の保存庫としての役割を担う。画像の表示は、パーソナルコンピュータと高精細な画像表示装置とを併用する。(4) 利用者視野はオブジェクト指向にもとづいた視覚化ユーザーインターフェース、すなわちカード形式によって統一を図ると共に、画像データベースに対する操作を直接的なものとして、操作性の向上を図る。システム操作はポインティング装置による直接操作であるため、操作対象となる全てのデータベース資源はアイコンによって視覚化される。このため、操作は日常馴染みのある方法で行えることとなる(カード、フォルダー等)。

5.2 試験システムの概要

以上の基本概念にもとづいて、必要な基本機能を実現すると共に試験システム CIREs を構築した。最終版の CIREs の概要は次に示す通りである。

5.2.1 ハードウェア構成

システムの構築の容易さとシステム資源の活用を考慮して、入手が容易で用途が汎用的なハードウェアを使用した。大型計算機は汎用型計算機 IBM システム/370 (3090, 3081, 4381, 9370 等) を、パーソナルコンピュータは IBM パーソナルシステム 55 (PS/55) を使用した。高精細画像表示装置として IBM5080 グラフィックス表示装置 (5080) を、画像蓄積のためには大型計算機に接続されている磁気ディスクと追記型光ディスクを、属性データベース蓄積のためには磁気ディスクを使用した。操作は主としてマウスによるが、キーボードも用いることができる。

5.2.2 ソフトウェア構成

大型計算機側では VM/CMS (仮想計算機/会話型モニターシステム) が、PS/55 側では MS-DOS が基本ソフトウェアである。両者間の通信は PC-VM Bond を用いた。大型計算機側では、属性データベースを管理する SQL/DS (構造化照会言語/データシステム) [IBM 1984], 5080 および光ディスクを管理するドライバーは、それぞれ 1つの VM として動作している。一方 PS/55 側ではウィンドウシステム、各種のデータや検索・表示等各種の命令を管理する応用ソフトウェアが稼働している。

5.2.3 文字・数値データと画像の蓄積

属性データは関係型データベース SQL/DS の上で実現されている。検索画像は PS/55 上で概覧するための縮小カラー画像と、確認のために 5080 グラフィックス表示装置に表示する高精細画像 (256色以上のカラーと濃淡) の 2種類であり、それぞれ用途の特性上磁気ディスクと光ディスクに蓄積した。CIREs 上では縮小カラー画像は128×128画素、16色に表現されているが、一方、高精細画像は256色で画素数は1024×1024, 512×512, 256×256の3種類に表現され、圧縮されている。

5.2.4 視覚化ユーザーインターフェース

操作の対象となるものは、それぞれ民博利用者にとって馴染みやすい形状のアイコ

ンとして PS/55 上に表示する。「情報カード」は、本システムにおけるデータ操作の基本単位であり、縮小カラー画像、属性データ、メモ欄が1枚のカードとして統一的に扱われ、各標本資料を忠実に表現している。「検索カード」は、民博所有の情報カードに類似するように属性項目と条件欄を並べてカード化したもので、検索のために使われる。「情報フォルダー」は、1枚の検索カードと複数枚の情報カードの集合である。「資料室」、「5080」、および「集合演算」はそれぞれ大型計算機で管理されているデータベース、5080 グラフィックス表示装置、情報フォルダー間の集合演算を行う機能を表わしている。これら以外にも、各種のデータベース資源・機能を特異なアイコンで表現した。

5.2.5 属性検索

検索条件の指定は、検索カードの条件欄に文字属性や数値属性を書き込むことによって行われる。この検索条件は、大型計算機のデータベース上で処理され、条件に該当するレコード数を PS/55 側に送り返す。利用者は、検索条件を変更・追加する等しながら、検索処理を何度も繰り返し、段階的に目的とする標本の枠をしぼりこんでいく。文字・数値等の入力にはキーボードを用いて行いか、後述する個人ユーティリティーである検索リストとマウスを用いるメニュー選択によって行う。

5.2.6 概覧・表示

属性検索で充分にしぼりこまれた結果、指定された情報カード（すなわち、縮小カラー画像と属性データ）が大型計算機から PS/55 に転送され、情報フォルダーに蓄えられ、画像と属性データが次のように概覧・表示される。(a) 1つの情報カードを1つのウィンドウ内に表示する。(b) 属性情報をリストとして表示する。(c) 多数の縮小カラー画像（1画面28枚まで可能）を同一のウィンドウに表示する。これらは相互に切り替えながら、単独にまたは色々な組合せでマルチウィンドウを使用して高速に表示するため、日常馴染んでいる商用カタログを概覧するように操作できる。このような概覧操作を通して検索しようとする画像の候補をさらにしぼりこんでいく。

また、PS/55 上で表示されている概覧用の縮小カラー画像を指定することによって、高精細なカラーおよび濃淡画像の単数枚または複数枚を 5080 グラフィックス表示装置に表示することができる。これらの画像を見ることによって詳細が判り、検索の対象画像か否かの確認が可能となる。このための操作はマウスを用いて、全て PS/55 上に表示されるアイコンとポップアップメニューの選択によって行う。

5.2.7 視覚にもとづく比較・分類

画像の内容は非常に豊富で、それを完璧に記述することは事実上不可能であるため、属性にもとづくだけでは十分な検索は不可能か、または検索条件の書き換えを何度も繰り返すこととなり効率的ではない。しかし一方では、表示された画像を見ることによって希望の画像を探すことは非常に容易である。このため、PS/55 上に表示されている画像やデータ、各種のアイコンをスクリーン上で直接移動したり、消去することを可能とした。また、表示されている情報フォルダー間での集合演算を行って、新たに情報カードの集合を作成することもできる。例えば、「丸くて手提のついた籠」の情報フォルダーを作成することは、すでに表示されている籠の画像から「丸くて手提のついた籠」だけを移動しながら直接分類することによって、可能である。また、「タイ王国の籠」の情報フォルダーを新たに作成する場合、すでに作成表示されている「籠」の情報フォルダーと「タイ王国」の情報フォルダーの集合演算を直接とることによって、可能である。

5.2.8 個人ユーティリティー

PS/55 で個人環境を実現するため、いくつかのユーティリティーを提供している。利用者は情報カードや検索カードに個人用のメモ欄を使うこと、属性項目の配列および表示順序を選択・変更すること、属性データの表示内容の変更をすること等が可能である。また、属性検索条件を入力する際には、頻繁に使用する検索条件を最高10件

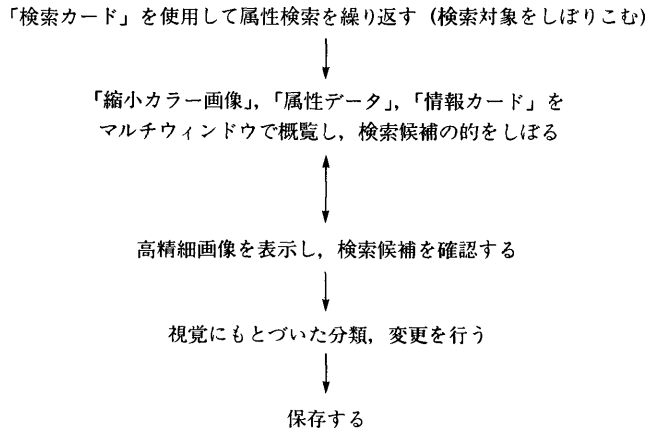


図7 CIRES による標本の検索手順

まで予め登録し、ポップアップメニュー形式でマウスを用い、個人用として使用できるリスト形式の検索用ライブラリーが提供されている。

以上の機能を使って標本を検索する手順の一例を示すと、図7のようになる。

5.3 評価と考察

5.3.1 性能

構築した試験システムには2716点の標本資料が登録されている。これは、属性データとしては2メガバイト、磁気ディスク上の概覧用縮小カラー画像としては4.2メガバイト、光ディスク上の圧縮原画像として3.4ギガバイトに相当する。現在の CIRES の構成では約55000点の標本資料が登録できる能力を持っているため、属性データ量は最大40メガバイト、概覧用縮小カラー画像は最大84メガバイト、圧縮原画像は最大68ギガバイトとなる。

現在の CIRES において2つの属性を指定して属性検索処理を行った場合、SQL の部分文字列マッチング [IBM 1984] によって該当件数19件を得るまでに約22秒の応答時間が必要であり、属性データと概覧用の縮小カラー画像を PS/55 側に転送し、PS/55 上でこれらを表示・概覧できるようになるまで約40秒の応答時間が必要である。PS/55 上で概覧用の縮小カラー画像1件を表示する時間は平均約0.3秒であり、情報カード1件を表示する時間は平均約1.6秒である。これらの応答時間についてはほぼ満足できるものと思われる。特に PS/55 上での表示速度は充分なものと思われるが、応答時間は早ければ早いほどより望ましい。応答時間の向上はハードウェアの性能に依存する割合が多く、高速通信技術の適用等によってかなり改善されるものといえる。特に、光ディスク上の画像アクセス時間とデータ転送速度は画像の蓄積量と密接に対応するため、一層の向上が望まれる。

5.3.2 システム構成

マイクロメインフレーム結合のシステム構成は、利用者にとって標本画像データベースシステムの高い利用価値が実現できると共に、システム資源の有効利用を図る上でも有効である。利用者が検索・利用するデータベースは大きければ大きいほど価値のあるものであり、蓄積容量や応答時間等の性能を考慮すると、大型計算機で管理することが現実的な選択となる。また、このような大規模なデータベースに加えて、

高精細画像表示装置のような高価で頻繁に使用しないシステム資源を大型計算機で管理して、多数の利用者の共有物とすることは画像データベースシステム資源全体の利用効率を高めることになる。一方、大型計算機側では応答時間が遅くても、情報量の多い、大量のデータを利用できる利点がある。これに対し、PS/55 側では扱えるデータ量は少なく、表示される画像の画質が低いけれども、高速で融通性に富む操作が可能であると共に、個人の環境が保証されるために利用価値が高まる。大型計算機、PS/55 共に今後一層の性能向上が期待されることから、それぞれが持つ現在の弱点や制約は改善・縮小され利用効率が飛躍的に向上するものと期待される。

5.3.3 多階層構造による画像の蓄積と表示

画質、すなわち空間解像度や色数が異なる画像をその用途に応じて、機能や性能を異にする外部記憶装置に分散させて蓄積・利用することはデータベースシステム全体の効率と資源の活用面で有効であるだけでなく、検索効率の向上にもつながるものである。本共同研究で使用した原カラー画像は 1024×1024 画素、1670万色を持ちそのデータ量は3メガバイトに達する。これをそのまま蓄積し利用することは、画像化された標本資料の実物に最も近づくことであるが、応答性の観点から頻繁に使用することは望ましいことではなく、またその必要性もない。人間の色に対する識別能力や利用効率および使用する画像表示スクリーンの性能から、適度の大きさと色数に減少させて使用することがより望ましい。いろいろ実験をしたことから考察すると、色の場合、256色から4096色が適当と思われるし、空間の大きさは概覧用には 128×128 画素が、確認のための表示画像では 512×512 画素から 1024×1024 画素が適当と思われる。今回構築した試験システムでは、空間的解像度は満足のいくものであるが、色数の面では決して充分とは思われない。しかし、これはハードウェアの制約であるため、改善は容易である。

一方、画像を効率的に蓄積するためには、その利用形態を基本にする必要がある。データ量が少なく高速に、かつ頻繁に利用される画像は磁気ディスクに、データ量が非常に大きく、早い応答速度は必要とされず、かつ利用回数が少ない画像は光ディスクに蓄積することはシステム構成上で有効であり、経済的といえる。つまり、磁気ディスク上の画像は自室にあるキャビネット内の、いわばカタログ写真であり、光ディスク上の画像は現物を蓄えている倉庫の役割を果たす。前者には高速性が、後者には大容量性がそれぞれ必須の機能として要求される。これに経済性を加えた3つの要因が全て満たされる技術が開発されれば、この技術だけで対応できることとなる。ま

た、画像やデータの表示装置の場合も同様であり、画質、表示速度、処理能力、容量、経済性等を総合的に考慮すべきである。画像の蓄積と表示についての今回の試験システム上での選択は、現在のハードウェア技術の現状においては現実的かつ適切と思われる。

5.3.4 検索の方法

検索の方法として、(1) キーワード等属性値のみによるもの、(2) パターン認識・画像処理によるもの、および (3) 属性検索と画像の表示とを組み合わせたものが考えられる。今回構築した CIRES では3番目の方法を選択したが、それは検索効率の面で最も現実的かつ適切と思われたためである。属性検索の繰り返し処理を行い、簡略化した画像を高速にかつ融通性を持って表示・概覧すると共に、高画質の画像を同時に表示し検索の対象となる画像を確認できるために、効率的な検索が可能となった。これは、主として高速な画像表示の実現、マルチウィンドウシステムの採用、カード形式によるユーザーインターフェースの実現によるものである。

今後必要とされる機能として検索の際の同義語と画像の内容にもとづく検索がある。現在は、入力者が検索条件として入力した文字についてのみ、それと同一の文字列を持ったデータを全て検索の対象としているが、これでは場合によっては入力すべき文字列が多くなりすぎる。例えば、標本名として「おめん」を指定する場合、「お面」、「めん」、「面」、「かめん」、「仮面」、「マスク」、「かぶりもの」等々を全て指定、入力することとなる。このため、民族学固有のコード体系 (HRAF/OWC, OCM 等) にもとづいたソーラスの利用が考えられる。HRAF/OCM コードの利用については後章で報告する。また一方、画像の内容による検索は属性データベースに記載されていないか、また記載されにくい画像の特徴（色、形状等）を検索時に指定し、その特徴量を抽出したり、パターンマッチングや画像処理によって類似した候補画像を検索するものである。この機能は、属性記述の限界を補う上で必要であり、さらには検索の対象となる候補画像が非常に多い場合、大まかな特徴によって候補画像の範囲をしばりこむことができ、検索の効率を向上させることが期待できる。この技術については、後節で試験した結果を報告する。

5.3.5 視覚化ユーザーインターフェース

今回採用したユーザーインターフェースは有効と思われる。これは、基本的には民博の利用者が日常のデータ・資料整理で馴染んでいる方法をパーソナルコンピュー

ター上で実現したためといえる。利用者に対する直接聞き取りやアンケート調査、プロトタイプに対する評価と分析等によって利用者側の、いわば「なまのこえ」にもとづいた結果である。特に、データ類の表現および取り扱いを全てカード形式にし視覚化したこと、操作方法をマウスによってマルチウィンドウとポップアップメニュー方式としたこと等が、利用者にとって馴染みやすくなり、操作性が向上したためと思われる。

6 標本画像類似検索

前述したように（「4 標本画像データベースシステムの利用者と利用方法」）、画像の検索にはいくつかの方法が考えられるが、画像の類似検索は「画像から画像の検索」に相当し、画像の内容にもとづき、利用者の検索条件を最も融通性に富む方法で処理するために、高度な画像検索といえる。本法の必要性は各所で指摘され、試作が試みられているものの、画像の特徴量の定義、検索時の条件入力方法、候補画像の決定方法、検索処理速度等において未解決の点が多い。本共同研究では、この方法の基本的なアプローチについて整理すると共に、画像の持つ内容特徴を形状と色に分け、具体的に民博の標本画像を使って、検索方法を検討し試験システムを構築した。この詳細は、本書「形状情報を用いた画像の類似検索」と「色情報に基づいた画像の類似検索」で後述する。以下においては、画像の類似検索における基本的な点について整理すると共に、構築した試験システムの概要について報告する。

画像を検索する際、検索者が検索条件を明確にしている程度によって、要求される検索機能の内容や技術的レベルが異なることは前述した（「4.2 利用方法」）。画像の内容にもとづいた画像検索・類似検索は最も曖昧な検索条件をもとにした場合であり、最も高度な検索技術が必要とされる場合である。この場合検索者は、なんとなく「あんなもの」というイメージを頼りに、希望する画像を探すわけである。この際に必要とされる検索技術としては、画像の特徴を抽出し、その画像内容を記述したデータベースを構築することがあり、さらに、検索条件として画像の内容・特徴を入力することがある。検索条件の入力方法として、文字・数値によって画像の特徴を記述する方法と、画像（またはスケッチ画）そのものとして入力する方法がある。後者の場合、入力された画像に対して画像処理を施して特徴を抽出して、初めて入力が完了することとなる。これら入力された画像の特徴は、予め構築されている画像の特徴データベースを使って、類似度を計算・比較する。この結果として、類似度の高い画像が

候補画像として検索者に返される。複数個の候補画像が該当した場合、検索者が希望する画像を正確に選べるような工夫が必要となる。

本共同研究では、形状特徴にもとづく類似検索の研究のためには籠類標本画像の前面濃淡画像を用い、また色特徴にもとづく類似検索のためには仮面類標本画像の鳥瞰カラー画像を用いてそれぞれ試験システムを構築した。実現した機能は次の通りである。つまり、形状特徴による類似検索システムでは (1) フーリエ記述子による特徴量の記述, (2) ユークリッド距離による類似度の計算, (3) 参考画像の指定による検索条件の入力, (4) 汎用型大型計算機 (VM) とグラフィックス表示装置からなるシステム構成等である。また、色特徴による類似検索システムでは (1) $L^*u^*v^*$ 系色空間における統計量としての特徴記述, (2) 斜交軸で定義された距離による類似度計算, (3) 参考画像の指定または参考色の指定による検索条件の入力, (4) 汎用型大型計算機 (VM) とグラフィックス表示装置からなるシステム構成等である。

これら類似検索の機能は、前述した主システムである CIREs とは独立しているが、今後この種の研究の発展に伴い統合化することによって、画像検索機能の強化を図ることができる。

7 文化項目分類コードの概覧

データベースにおいて検索条件を指定する方法として、検索者側から検索条件入力画面に対して文字列や数値が直接入力されたり、画像として入力されたりする方法があるが、検索条件が曖昧であったり複雑であったりするために、これらの方法は検索条件を明確に表現できない場合には有効な方法とはいえない。また、入力する言語の統一が図られないために、実現されているデータベース上で使われている言語との間の突き合せができない場合が生じる。さらに、系統立った検索条件を連続して入力する場合、その正確さと速度等に支障をきたす可能性がある。

これらを解決する 1 つの方法として、絶対多数の利用者によって認知されているコード体系を用いることがある。この場合、コード項目間を自由に航行しながら目的とするコード項目に自由にかつ迅速に到達することが要求される。さらに、操作性の良さ、操作の理解が容易であること等も利用者にとっては重要な事柄といえる。

本共同研究では、民族学分野において最も権威のある HRAF/OCM コード体系『文化項目分類』を民博が日本語訳したものを対象に、パーソナルコンピュータ上で容易に、自由に主題項目間を航行できるシステム機能を実現した。その特徴は、

パーソナルコンピュータ上で主題項目間を自由に航行できると共に、その上位・下位の主題項目を概覧することができること、主題項目とクロスレファレンス間を自由に航行できると共に、それらの内容を表示することができること、そして利用者が希望する航行経路を新たに作成することができること等である。この詳細は別途論文「HRAF/文化項目分類コードの概覧表示」で述べられる。

自由航行機能を検索時に使用することによって、検索機能の強化とシステム全体の性能の向上を図ることができる。

8 おわりに

民族学研究にとって重要な研究資料である標本に関する多様な情報を統一的に活用するため、各種の画像と文字・数値データを蓄積・検索するコンピュータ応用手法を研究し、それらの試験システムを構築した。

本共同研究が産学共同であり、人文科学系と理工系という専門領域が極めて異なる研究者によって構成されていることから、できるかぎり共通する技術的な課題を中心にしようとした。また、コンピュータシステムの最終利用者である民博の民族学研究者にとって役に立つ応用研究に努めた。

このため我々は、民博内で対象となる各種データをどのように収集し、管理しているのかを直接担当者から学ぶと共に、コンピュータシステムの利用者である民族学研究者が日常に行っているデータ整理と利用方法を聞き取り調査によって直接学んだり、アンケート調査を行ってできるかぎり「なま」のものに触れようとした。このようにして得られた、いわば「現場のなまのこえ」を情報学的な観点から整理した。この結果にもとづき、民族学研究者が日頃から実施しているデータ整理の方法を、そのままコンピュータ化するようにした。このような基本方針が妥当か否か、またコンピュータ上での実現機能が正しいか否かについても、利用者である民族学研究者から直接聞くこととした。この場合、利用者側の意見を正確に把握するためには、利用者側が正確に表現すること、聞く側であるコンピュータ専門家が正確に理解することが必須となる。このため、両者にとって共通する具体的なものを通して、具体的に評価・検討することが有効であることから、利用者側からの評価を反映させながら試験システムをできるかぎり迅速かつ高頻度に構築・改良（プロトタイピング）することとした。プロトタイピングを繰り返すことによって、利用者側が抱くシステムのイメージが一層明確になり、評価や改良のための指摘がより具体化して、構築する側の

理解が正確なものになってくる。

この結果、構築した試験システムは利用者側に理解しやすいものとなり、ほぼ満足のいくものと思われた。今後これら成果や経験をもとに、利用者の評価を反映することによって、システムの一層の発展が期待される。

文 献

CARROL, J.M. and J.C. THOMAS

1982 Metaphor and the Cognitive Representation Computing Systems. *IEEE Transactions on Systems, Man and Cybernetics* SMC-12(2): 107-116.

CHANG, S.K.

1985 Image Information Systems. *Proc. IEEE* 73(4): 754-764.

CHEN, P.P.

1976 The Entity-Relationship: Toward a Unified View of Data. *ACM TODS* 1: 9-36.

ERICSON, K.A. and H.A. SIMON

1984 *Protocol Analysis: Verbal Reports as Data*. The MIT Press.

GARDARIN, G. and E. GELENBE (eds.)

1985 *New Applications of Data Bases*. Academic Press.

橋原秀晴・洪政国・杉田繁治

1989 「繰り返し設計法によるユーザ・インタフェースの向上性」『情報処理学会研究報告』89-CH-2, pp.13-19.

洪政国・六川修一

1985 「カラー画像データベースシステム」『アドバンスト・データベースシステム・シンポジウム論文集』情報処理学会, pp.83-91。

洪政国・佐藤真知子・井岡幹博・川辺みどり・香田正人・杉田繁治・久保正敏・山本泰則

1986 「民族学研究のための画像データベース管理システム」『情報処理学会第33回全国大会講演要旨集』4H-10, pp.863-864。

洪政国・井岡幹博・佐藤真知子・黒川雅人・杉田繁治・久保正敏・山本泰則

1987 「民族学研究のためのカラー画像蓄積・検索システム」『情報処理学会研究報告』87-DB-59, pp.1-8。

洪政国・黒川雅人・杉田繁治・久保正敏・山本泰則

1987 「民族学研究における標本資料情報の意味表現」『情報処理学会第34回全国大会講演要旨集』2C-3: 445-446。

HONG, J-K.

1988 Electronic Cataloging and Browsing of Ethnographic Samples. *Proc. 12th International Symposium on the Conservation and Restoration of Cultural Property—Analysis and Examination of an Art Object by Imaging Technique—*, Tokyo National Research Institute of Cultural Properties, Tokyo, pp.119-130.

1990 Image Storage and Retrieval, and Its Applications to Museum. *Proc. 11th International Korean Conf. of Science and Technology*, The Korean Federation of Science and Technology Societies, Seoul, pp.1227-1231.

HONG, J-K., H. HASHIHARA, M. IOKA, M. KUROKAWA, M. SATO, S. SUGITA, M. KUBO and Y. YAMAMOTO

1988 A Color Image Database for an Ethnology Museum—A Multi-Window System for Electronic Cataloging and Browsing of Ethnographic Samples on the PC—. *Proc. Cologne Computer Conf.*, Cologne, pp.B5.1-5.2.

IBM Corp.

1984 *SQL/DS System Application Programming for VM/System Product*. SH24-5068.

- 井岡幹博
 1987 「画像情報システムにおける目視検索について」『情報処理学会第35回全国大会講演要旨集』2K-4, pp.2111-2112。
 1988 「色情報を用いた標本画像の類似度検索」『情報処理学会第37回全国大会講演要旨集』5V-3, p.1575。
- 井岡幹博・佐藤真知子・洪政国・杉田繁治
 1987 「多階層画像表示による画像の目視検索」『第18回画像工学コンファランス論文集』, pp.83-86。
- 加藤隆
 1985 「計算機ユーザのための認知行動原理を探るための一手法」『情報処理』26(9): 1106-1109。
- 甲洋介・加藤隆
 1987 「言語プロトコル分析手法とそのインタラクティブな分析支援ツール」Proc. 3rd Symposium on Human Interface, pp.397-402。
- 国立民族学博物館(編)
 1985 『国立民族学博物館十年史』国立民族学博物館。
 久保正敏・山本泰則・杉田繁治・洪政国・佐藤真知子・井岡幹博・川辺みどり・香田正人
 1986 「民族学研究のための画像検索ワークステーションについて」『情報処理学会第33回全国大会講演要旨集』4H-11, pp.865-866。
- 黒川雅人
 1987 「形状特徴を用いた画像の内容検索」『情報処理学会第35回全国大会講演要旨集』2K-1: 2105-2106。
- 黒川雅人・橋原秀晴・洪政国・杉田繁治
 1987 「民族学研究用画像データベースにおけるユーザインターフェイスの視覚化」『アドバンス・データベースシステム・シンポジウム論文集』情報処理学会, pp.9-16。
- 黒川雅人・洪政国・杉田繁治
 1990 「形状情報を用いた類似検索システムの試作——民族学標本を例にして——」『情報処理学会研究報告』90-CH-5, pp.25-31。
- KUROKAWA, M.
 1989 An Approach to Retrieving Images by Using Their Pictorial Features. Proc. ICIP'89, IEEE, Singapore, pp.588-592。
- マードック, G.P.・C.S. フォード・A.E. ハドソン・R. ケネディ・L.W. シモンズ・J.W.M. ホワイティング
 1987 『文化項目分類 (OCM) 第5版対訳版』国立民族学博物館研究部(訳)(館内資料)。
- McKEOWN Jr., D.M.
 1983 MAPS: The Organization of a Spatial Database System Using Imagery, Terrain, and Map Data. Carnegie-Mellon Univ.
- MOORE, F.W. (ed.)
 1961 Reading in Cross-Cultural Methodology. HRAF Press。
- 中村史朗・石井義興(編)
 1987 「特集: マルチメディアデータベース」『情報処理』28(6): 670-802。
- 野口英男・井上誠喜・浦谷則好・柴田正啓・白田洲宏
 1986 「静止画像検索システムの試作」『電子通信学会技術研究報告』IE86-2, pp.9-16。
- 小沢一雅
 1985 「考古学研究支援型データベースシステムの構成」『情報処理学会論文誌』26(5): 936-945。
- OIKAWA, A.
 1987 Archeological Image Database System. Proc. International Conf. on Data Bases in the Humanities and Social Sciences, pp.157-166。
- 坂内正夫・大沢裕
 1987 『画像データベース』昭晃堂。
- 佐藤真知子
 1987 「画像情報システムにおけるデータの取扱いについて」『情報処理学会第35回全国大

- 会講演要旨集』2K-5, pp.2113-2114。
- 佐藤真知子・井岡幹博
1986 「画像データベースシステムにおけるユーザ・インターフェイス」『情報処理学会第33回全国大会講演要旨集』6P-4, pp.1657-1658。
- 佐藤真知子・洪政国・井岡幹博・橋原秀晴・杉田繁治・久保正敏・山本泰則
1987 「民族学研究のための画像情報システム」『電子情報通信学会技術研究報告』DE87-8~11, pp.23-30。
- 佐藤真知子・橋原秀晴・井岡幹博・黒川雅人・洪政国・杉田繁治・久保正敏・山本泰則
1988 「民族学研究支援のための標本画像検索システム」『情報処理学会論文誌』29(12): 1108-1118。
- SCHLERETH, T.J. (ed.)
1985 *Material Culture—A Research Guide—*. University Press of Kansas.
- 嶋田茂・江尻正員
1986 「日本語インターフェイスを有する知識処理型マルチメディア地図情報システム GENTLE」『情報処理学会論文誌』27(12): 1162-1173。
- 篠田英範・近藤隆志・澤田順夫・沼上英雄・木戸出正継
1983 「ランドサット MSS 画像データベースシステムの開発と評価」『情報処理学会論文誌』24(6): 867-876。
- 杉田繁治
1987a 「現在のコンピュータは文科系の研究に役立つか」『情報処理』28(5): 623-628。
1987b 「人文科学におけるマルチメディアデータベース」『情報処理』28(6): 765-772。
- 梅棹忠夫
1969 『知的生産の技術』岩波書店。
- 打波清一
1987 「フィールドノート調査データ処理におけるマルチメディアデータベース」『情報処理』8(6): 773-789。
- 安田靖彦 他(編)
1985 「画像処理特集」『電子通信学会論文誌』J86-D(4)。
- 横矢直和・田村秀行
1981 「画像データベース研究の動向」『電子技術総合研究所会報』45(9): 38-52。
1984 「病理標本画像データベースシステム PIMAS の開発」『情報処理学会論文誌』25(6): 1055-1063。