

データベース構築促進及び技術開発に関する報告書

# 連想検索写真データベースの プロトタイプ作成

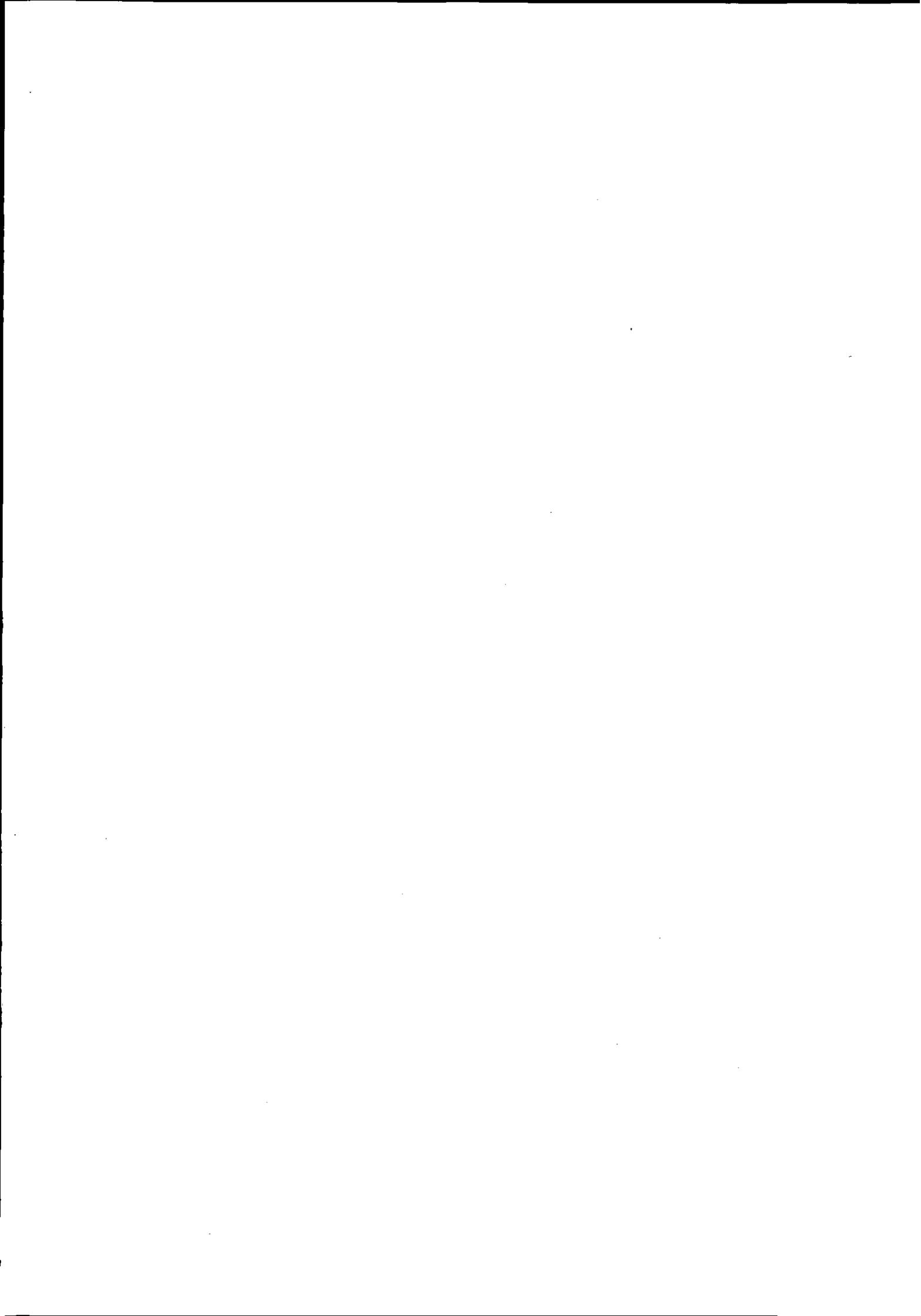
平成7年3月

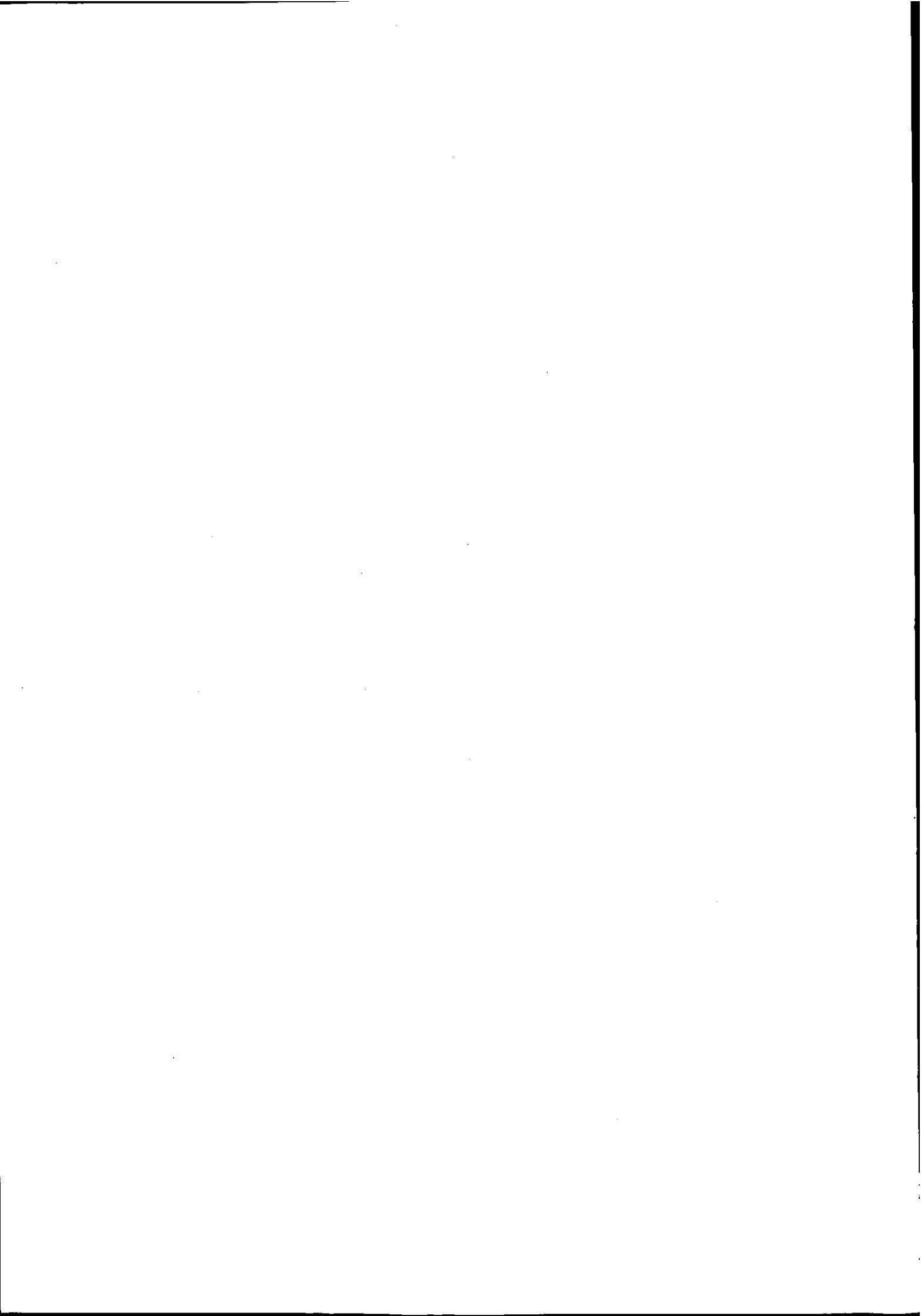
財団法人 データベース振興センター  
委託先 シャープ株式会社

KEIRIN

00

この事業は、競輪の補助金を受けて実施したものです。





## 序

データベースは、わが国の情報化の進展上、重要な役割を果たすものと期待されている。今後、データベースの普及により、わが国において健全な高度情報化社会の形成が期待される。さらに海外に対して提供可能なデータベースの整備は、国際的な情報化への貢献および自由な情報流通の確保の観点からも必要である。しかしながら、現在わが国で流通しているデータベースの中でわが国独自のものは1/3にすぎないのが現状であり、わが国データベースサービスについてはバランスある情報産業の健全な発展を図るためには、わが国独自のデータベースの構築およびデータベース関連技術の研究開発を強力に促進し、データベースの拡充を図る必要がある。

このような要請に応えるため、(財)データベース振興センターでは日本自転車振興会から機械工業振興資金の交付を受けて、データベースの構築および技術開発について民間企業、団体等に対して委託事業を実施している。委託事業の内容は、社会的、経済的、国際的に重要で、また地域および産業の発展の促進に寄与すると考えられているデータベースの構築とデータベース作成の効率化、流通の促進、利用の円滑化・容易化などに関係したソフトウェア技術・ハードウェア技術である。

本事業の推進に当って、当財団に学識経験者の方々に構成されるデータベース構築・技術開発促進委員会(委員長 前山梨学院大学教授 蓼沼良一氏)を設置している。

この「連想検索写真データベースのプロトタイプ作成」は平成6年度のデータベースの構築促進および技術開発促進事業として、当財団がシャープ株式会社に対して委託実施した課題の一つである。この成果が、データベースに興味をお持ちの方々や諸分野の皆様方のお役に立てば幸いである。

なお、平成6年度データベースの構築促進および技術開発促進事業で実施した課題は次表のとおりである。

平成7年3月

財団法人 データベース振興センター

平成6年度 データベース構築・技術開発促進委託課題一覧

分野	課題名	委託先
社 会	1 報道ニュースのマルチメディアデータベース構築 2 携帯型電子新聞プロトタイプの開発 3 SPORTS POWER INDEXのパイロット版データベースの構築 4 画像付き地図資料データベースのプロトタイプの作成 5 パソコンを利用した副作用症例データベースの構築調査	日本電子計算(株)名古屋支店 (株)日本経済新聞社/(株)日経データ (株)ビデオ・リサーチ (財) 地図情報センター 協立医師協同組合
中小企業振興 地域活性化	6 関西イベント&プロジェクトデータベース構築 7 ハイビジョン大型映像ソフト流通促進用データベース構築 8 分散協調型データベースのための高度運用システムの開発 と学習支援環境への応用	(株)京都新聞社 (財) 大阪科学技術センター (株)エマーズ
海 外	9 電子デバイス情報の電子化ドキュメントのプロトタイプ作成 10 海外ユーザ向けの国産データベース検索支援用端末ソフトウェアの開発	電子デバイス情報サービス(株) カテナ(株)
技 術	11 連想検索写真データベースのプロトタイプ作成 12 レーザー研究データベースの構築 13 砂漠化防止技術データベース 14 メディア変換型感性データベースの構築方法の研究 15 フォールトトレランスな多データベースサーバシステムに関する調査研究 16 創出キーワードの自動付与に関する調査研究	シヤープ(株) (財) 応用光学研究所 (株)新産業創造センター (財) イメージ情報科学研究所 (株)シネジャーナルプロダクション (株)エレクトロニック・ライブラリー

## 目次

1. 研究の目的	1
1.1 写真データベース構築の目的	1
1.2 画像検索技術の課題とキーアイデア	2
1.3 実施内容	3
2. 意味ベクトル技術	4
2.1 意味ベクトルの定義	4
2.2 システム構成	5
2.3 FEATURE の選択とコア単語への FEATURE 付与基準	6
2.4 単語の意味ベクトルのブートストラップ学習	8
3. 意味ベクトルによる画像検索	12
3.1 画像検索システムの概要	12
3.2 意味ベクトルによる画像検索のファンクション	13
4. 写真検索ソフトウェアとユーザーインターフェース	19
4.1 プロセスの概要	19
4.2 ユーザーインターフェース	19
4.3 データベースの構造	23
4.3.1 データベースの概要	23
4.3.2 データベースファイルの構造	23
4.4 まとめ	25

5. 画像検索評価用ベンチマークの作成	26
5.1 画像データベース	26
5.2 画像検索システムの評価	28
5.2.1 概略	28
5.2.2 評価用データ作成の概略	28
5.2.3 評価用データ作成の詳細	28
5.2.4 評価結果	32
5.2.5 問題点	32
5.3 まとめ	33
6. 今後の課題	34
参考文献	35

## 1. 研究の目的

近年、写真データベースの構築が急速に進展している。しかし、これらのデータベースは、一般に、キーワード検索あるいは全文検索技術を用いて構築されるため、写真の内容を詳細に記述する必要がある、データベース作成者の負担が大きく、十分な検索が困難だという問題がある。

本委託課題の目的は、シャープ株式会社情報技術研究所において開発された「意味ベクトルによる検索技術」<sup>1)・2)・3)</sup>を用いて、通常の写真のタイトルやキャプション(説明文)程度の少ない情報を基に、多様な検索目的に対応できる写真検索ソフトウェアのプロトタイプをパーソナルコンピュータ(PC)上に作成することである。PCには、本写真データベースのユーザになると想定されるデザイナーに広く利用されているマッキントッシュを選択した。

### 1. 1 写真データベース構築の目的

写真データベース構築の目的は、次の2種類に大別できる。

#### (1)大量画像の管理

フォトライブラリーのように膨大な数の写真を管理し、必要に応じてストックした写真の中からユーザの要求を満たす写真を探し出すためには、写真番号、ジャンル、撮影地、写真家などで整理して蓄積しておく必要がある。

フォトライブラリーによる大量写真の管理を目的とした写真データベースには、(株)日本フォト・リサーチ・センター((株)JPRC)の写真検索システムがある<sup>4)</sup>。約8万点の写真をレーザディスク2枚に格納し、分類表(シソーラス)に基づく専用の検索システムがPC上で利用できる。

#### (2)画像の検索

画像データベースの中味を知らないユーザを対象に、検索された画像の二次利用を目的にした写真データベースの利用方法がある。フォトライブラリーにとっては、貯蔵写真の電子的なカタログになる。また、企画会議等における発想支援のツールとして写真データベースの利用も考えられる。

デザイナー等のユーザが写真を選択するためのツールとしての利用（写真の検索）を目的とした写真データベースには、（株）JPRCのCD-ROM共同写真カタログがある<sup>5)</sup>。フォトライブラリー40社の写真1万点をCD-ROM2枚に格納し、写真を選択するための1万点の縮小写真（サムネイル）を備えた全文検索による検索ソフトがマッキントッシュ上で利用できる。

本委託課題は、膨大な数の写真が格納された写真データベースの中身を全く知らないユーザが、何等かの目的（DTP、絵葉書の作成、発想支援、図鑑としての利用等）のために、簡単に利用できる写真データベースのプロトタイプを作成することである。

## 1. 2 画像検索技術の課題とキーアイデア

分類表方式やキーワード検索技術の課題を次に示す。

- ・ キーワードを思い出せないと検索できない
- ・ 全ての人を要求を満たすようなキーワード付けや分類は困難
- ・ 分類表の作成困難
- ・ 分類表は1次元の分類のため、多角的な検索ができない

分類表方式やキーワード検索の課題を解決し、キーワードに頼ることなく人間のような連想機能を持った検索手法として、言葉に伴って想起される概念の多次元的表現（意味ベクトル）を用いた連想検索技術を開発した。

基本単語3700語の意味ベクトルを、人手で266種類のFEATURE（画像の特徴を記述する要素）との関係を付けることで作成し、この基本単語のベクトルと百科事典のテキストを用いた統計的（ブートストラップ）学習によって、百科事典に含まれる全単語の意味ベクトルを生成した。この意味ベクトル辞書を用いて、風景、ビジネスシーンなど約170枚の写真を対象に画像データベースを試作し、意味ベクトルを用いた検索を行ったところ、素人によって数語程度の説明文をつけただけでも、多様な検索ができることがわかった。これによって、例えば、キャプションに‘ハワイワイキキ’という単語があるだけで、‘リゾート地’や‘海岸’といった‘ハワイワイキキ’と連想関係の強いキーによってもその写真が検索されるようになる。

### 1. 3 実施内容

百科事典を基に作成した単語の意味ベクトルを写真のキャプション（説明文）に適応させることで、少ないキーワードでも連想的に写真を検索できる写真データベースのプロトタイプをマッキントッシュ上に作成した。実施内容を以下に示す。

#### (1)プロトタイプ詳細設計（平成6年7月～平成6年9月）

4章で報告するように、マッキントッシュ用の写真検索ソフトウェア、検索結果を視覚的に表示するための操作性の良いユーザインタフェースを設計した。

#### (2)プロトタイプ作成（平成6年10月～平成7年2月）

マッキントッシュ（Power MAC）上でC言語を用いて、写真検索ソフトウェアとユーザインタフェースのプロトタイプを作成した。但し、単語の意味ベクトル辞書は、今回の委託課題の対象外とし、ワークステーション上で作成した既存のものを用いた。

#### (3)評価用写真画像の選択（平成6年12月～平成7年1月）

5章で報告するように、画像検索ベンチマークを作成するためのツールをワークステーション上に開発し、画像の内容の主観的印象から14グループに分けられる104件の写真を選択した。

#### (4)写真データベースプロトタイプの検証テスト（平成7年2月）

今回作成した画像検索ベンチマークを用いて評価を行ったところ、上位3位以内に80～100%の確率で少なくとも1つの正解画像が検索されており、意味ベクトルによる検索の有効性を確認した。

## 2. 意味ベクトル技術

本章では、CD-ROMによって提供される百科事典「ブリタニカ小項目事典」<sup>6)</sup>から意味ベクトル辞書を構築する方法について述べる。

### 2. 1 意味ベクトルの定義

意味ベクトルの基本的な考え方は、Waltz, Pollackによって提案されたmicrofeatures<sup>7)</sup>, Gallantによって提案された文脈ベクトル<sup>8)</sup>に基づいている。意味ベクトルは、多くのFEATUREとの意味的な関係をベクトル表現したものである。n個の概念分類をFEATUREとし、各次元が一つのFEATUREに対応したn次元ベクトル空間上の一点で、意味を表現するものである。単語iの意味ベクトル  $X_i = (x_{i1}, x_{i2}, \dots, x_{in})$  の各要素の値は、次のように定義される。

$x_{ij}=0$             if 単語 i がFEATURE j と関係なし  
 $x_{ij}=1$             if 単語 i がFEATURE j と関係あり

例えば、次の6個のFEATUREを選択した場合に、

[人間, 悲しい, 芸術, 科学, 興奮, 政治]

単語「パイロット」の6次元の意味ベクトルXは以下の通りである。

$X = (1, 0, 0, 1, 1, 0)$

上述の論文<sup>7), 8)</sup>では、単語と各FEATUREとの関係を、その強度に応じて多段階で表現している。しかし、本研究では、人手で作成する意味ベクトルに関しては、関係あり、なし、の2値で表現する。FEATUREとの関係を2値で表現することで、分野に依存しない汎用的な単語の意味ベクトル辞書として蓄積できると考えられる。意味ベクトルを単語の意味表現として用いれば、ベクトル演算、例えば、各単語の意味ベクトルの和を文章の意味ベクトルとすることによって、ベクトルの向きで文脈の意味を表現すること

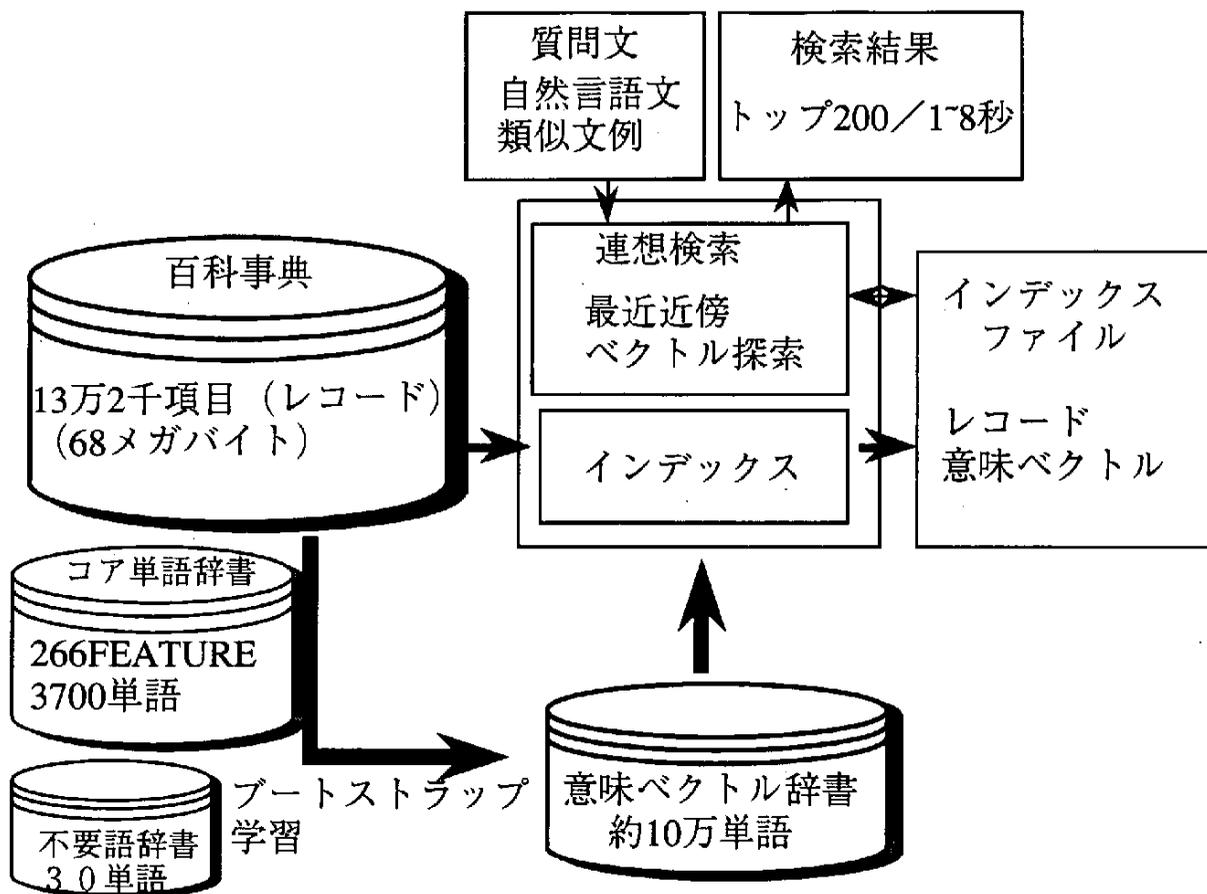


図2・1 システム構成

ができる。

## 2. 2 システム構成

意味ベクトルによる百科事典項目検索システムの構成を図2・1に示す。「ブリタニカ小項目事典」には、約13万2千項目について、各項目1段落からなる説明文が付けられており、総テキスト容量は、約68メガバイトである(項目名とその説名文を併せたものが検索対象になるため、以後レコードと呼ぶ)。百科事典の全テキストを頻度解析し、160レコード以上に含まれている名詞4000語のうち、各レコードの文脈と関係なく出現する単語約30語、例えば、「生年」、「没年」のように「人名」に関するレコードには必ず含まれる単語は、意味ベクトルを与えれば検索にとって悪影響を及ぼすと考えられるので、不要語辞書に登録して削除した。さらに、程度、方向、期間を表す

単語など FEATURE 付けが困難なものを除き、残された約 3700 語をコア単語とした。FEATURE には、2. 3 節で述べるように、266 種類を選択した。このコア単語と各 FEATURE との関係を手で作成し、コア単語辞書に登録した。

コア単語と百科事典の大量テキストを利用した、2. 4 節で述べるブートストラップ学習によって、百科事典中に表われる 10 万単語以上の意味ベクトル辞書を生成した。次に、この意味ベクトル辞書を用いて、各レコードの意味ベクトルを作成し、インデックスファイルに登録した。レコードの意味ベクトルは、レコードに含まれる単語の意味ベクトルの重み付きの和とし、ベクトルの絶対値を正規化したものとする。

連想検索は、利用者が自然言語で質問文を入力すると、最近近傍ベクトル探索によって、質問文と意味的に類似したレコードの上位 200 のランキングが数秒で表示される。質問文の意味ベクトルは、レコードの意味ベクトルと同様に求められ、インデックスファイルに格納された 13 万 2 千レコードの意味ベクトルとの距離（内積）によってランク付けられる。最近近傍ベクトル探索は、各レコードと質問文の意味ベクトルの特徴を概観した概略情報の短縮ベクトルを用いて、高速に候補を絞り込む。

## 2. 3 FEATURE の選択とコア単語への FEATURE 付与基準

FEATURE の組み合わせによって、百科事典の内容を分類できるように FEATURE を選択する必要がある。表 2・1 に示すように、6 種類の大分類、29 種類の上位概念に属する 266 種類の FEATURE を選択した。

表 2・1 FEATURE の分類

大分類 (上位概念数)	上位概念	FEATURE 数	FEATURE 例
人間・生命 (6)	人間	8	人間, 人名, 男性, 女性, 子供, 大人, 老人, 動物, 水棲生物, 鳥類, 虫, 微生物, 植物 生命, 生死, 誕生, 病気, 老い, 殺生, 性
	生物	6	
	生命	7	
	.	14	
人間環境 (5)	人造物	3	道具, 機械機器, 建造物 通信, マスメディア, 交通輸送, 自動車, 会社職業, 教育育児, 福祉年金, 施設設備,
	交通・通信	6	
	社会制度	7	
	.	17	
自然環境 (2)	地域	14	地名, 国名, 日本, 都会, 地方, 海外, 陸地, 山岳地, 天空, 海洋, 気象気候, 環境,
	自然	10	
抽象概念 (4)	精神・心理	5	感覚, 感情, 喜楽, 悲哀, 恐怖 様子様態, 変化, 関係関連, 秩序順序, 数量,
	抽象概念	8	
	.	36	
物理・物質 (8)	動作	2	動作, 行為 運動, 停止, 動的, 静的 蒸発気化, 凝固凍結, 溶解液化, 発熱発光, 鉱物, 資源, 素材材料
	運動	4	
	物理現象	7	
	物質	3	
	.	47	
文明・知識 (4)	人文	20	民族人種, 知識, 言論発話, 書物書籍, 数学, 物理学, 天文学, 地学, 医学薬学,
	学術	6	
	.	36	

この概念分類に基づいて FEATURE をコア単語に付与できる専用ツールによって、能率良くコア単語の意味ベクトルを作成することができる。コア単語への FEATURE 付与基準を次に示す。

(1)論理的関連と連想的関連から、FEATURE を付与する。

[論理的関連] コア単語に対して FEATURE が表 2・2 (a)に示すような直接的関連性を有するものを指す。

[連想的関連] コア単語に対して FEATURE が感覚的関連性、連想により想起される関連性を有するものを指す。例を表 2・2 (b)に示す。

(2)FEATUREの上位概念，大分類は，分類上の目安であり，付与判断の基準はFEATUREそのものである。例えば，FEATURE「温かさ」は上位概念「物理的特性」の下に分類されているが，“心の温かさ”からの連想によってコア単語「愛」に付与する。

表 2・2 FEATURE 付与基準

(a) 論理的関連による FEATURE 付与

	論理的関係	コア単語	FEATURE
1	集合包含	秋	季節
2	同義関係	アイデア	思想
3	部分全体関係	足	人間の身体
4	相補関係	子	親

(b) 連想的関連によるFEATURE 付与

コア単語	FEATURE
愛 アップ 足	優しさ，倫理 経済，映像 自動車，交通輸送

こよのようなFEATURE付与基準からコア単語「愛」に付与されたFEATUREは，[人間，感情，喜楽，関係関連，肯定的，優しさ，温かさ，倫理道德，宗教，心理学]の10種類である。

## 2. 4 単語の意味ベクトルのブートストラップ学習

百科事典中の単語すべてについて，266個のFEATUREとの関連を人手で付けるのは現実的ではない。そこで次の2種類仮説の元に，ブートストラップ学習と呼ぶ手法を開発した。

- (1)レコードの意味ベクトルは、一定数（例えば、5個）以上のコア単語が含まれていれば、そのコア単語の意味ベクトルの加重和によって、適切な文脈情報が表現される。
- (2)単語の意味ベクトルは、その単語が含まれているレコードの意味ベクトルの加重和によって、その単語が使われる適切な文脈情報を獲得できる。

図2・2に、ブートストラップ学習の処理フローを示す。単語バッファには、単語  $j$  が含まれるレコードの意味ベクトルの加重和を計数するバッファ  $BSV_j$  がある ( $BSV_j$  は、各要素が0のベクトルに初期化しておく)。図2・2に沿って、ブートストラップアルゴリズムを説明する。まずは、レコード単位に形態素解析によって単語を抽出する。不要語を除き、一定数以上のコア単語を含むレコードについて、仮説(1)に基づいて、レコード  $ri$  の意味ベクトル  $RSV_{ri}$  を次式によって計算する。

$$(3)RSV_{ri} = k \sum_j W_j P_{ij} (a_1^2 x_{j1}, \dots, a_n^2 x_{jn})$$

但し、 $W_j$  は単語  $j$  の文書データベースにおける重み、 $P_{ij}$  は単語  $j$  のレコード  $ri$  における頻度、 $(a_1, a_2, \dots, a_n)$  は各FEATUREの重み、 $(x_{j1}, x_{j2}, \dots, x_{jn})$  は単語  $j$  の意味ベクトルである。 $k$  はレコードのベクトルを一定の長さにするための係数である。

単語の重み  $W_j$  は、式(4)の文書データベース中の出現確率の逆数によって決める。

$$(4)W_j = \log (N / df_j)$$

但し、 $N$  は全レコード数、 $df_j$  は単語  $j$  が含まれるレコード数である。式(4)は、あるレコードに偏って出現する単語ほど、レコードを検索する上で重要になる（重みが大きくなる）ことを表わしている。

FEATUREの重みは、式(5)のコア単語の各FEATUREの出現確率の逆数である。

$$(5)a_i = \log (T / f_{ci})$$

但し、 $a_i$  はFEATURE  $i$  の重み、 $T$  は対象となる全コア単語数、 $f_{ci}$  はFEATURE  $i$  が付与されたコア単語数である。式(5)は、多くの単語に付与されているFEATUREは、単

語を意味的に分類する上での効果が小さいことを表現している。266次元の意味ベクトルを有効活用するためには、266個のFEATUREが均等に単語に付与されている必要がある。式(3)は、レコード中に含まれているコア単語について、ブートストラップによってFEATUREの偏りが小さくなるようにその意味ベクトルをFEATUREの重みで変調を掛け、文書データベースにおける単語の重みと、レコード内の単語の重みとして単語の頻度を掛けたベクトル和を、一定の長さに正規化したものである。

レコード  $r_i$  から抽出された全ての単語  $j$  (コア単語を含む) について、仮説(2)に基づいて、以下のようにバッファ  $BSV_j$  を更新する。

$$(6)BSV_j \leftarrow BSV_j + P_{ij}RSV_{ri}$$

全レコードについて、式(3)でレコードの意味ベクトルを求め、そのレコードに含まれる全単語の意味ベクトルを式(6)で更新した後、各単語に付与されるFEATURE数が一定範囲内(例えば、8個以上25個以下)に収まるように  $BSV_j$  を正規化(あるいは1ビット量子化)したものを、単語の意味ベクトルとする。

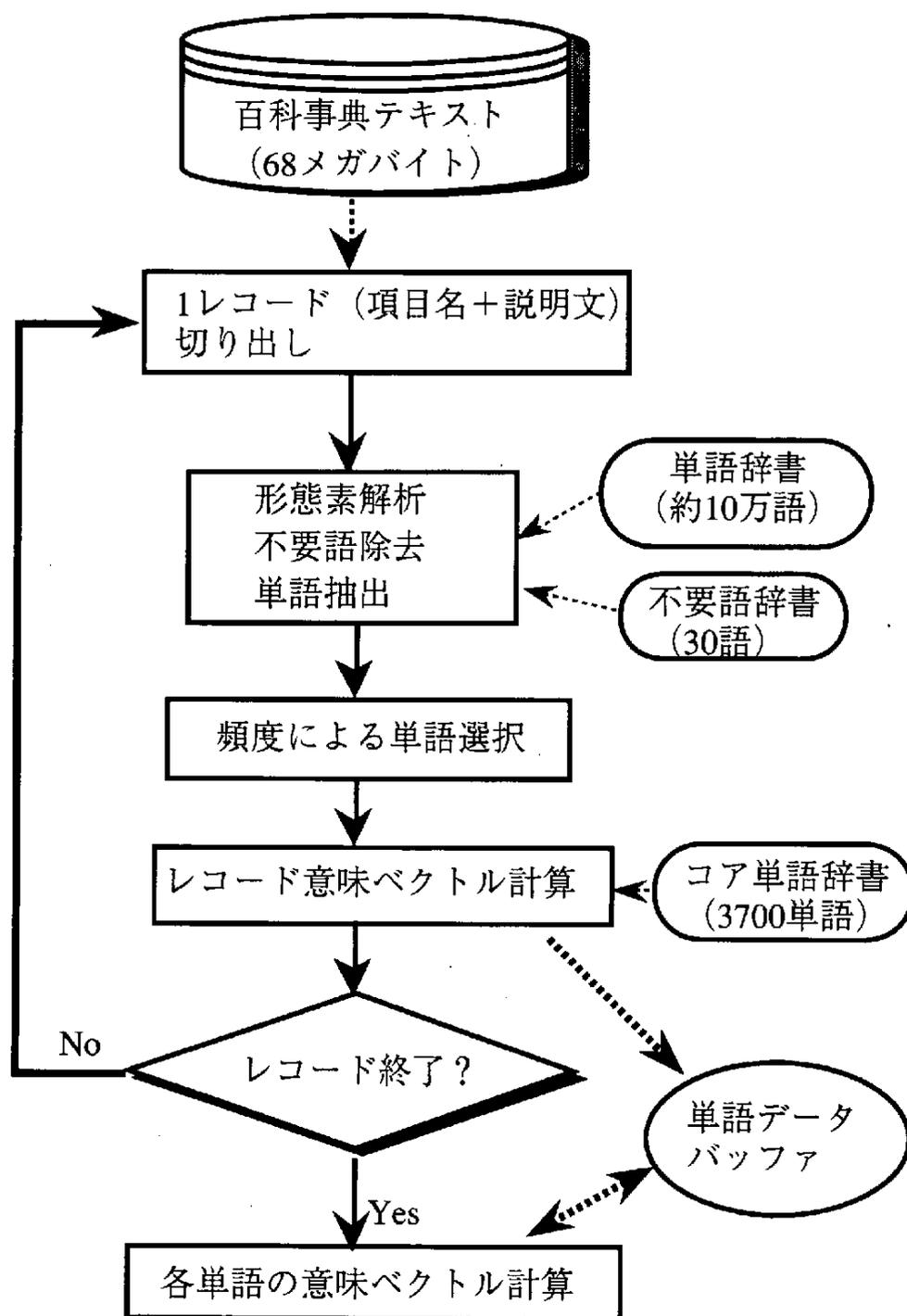


図 2・2 ブートストラップ学習の処理フロー

### 3. 意味ベクトルによる画像検索

本章では、百科事典を基に生成した 10 万単語以上の意味ベクトル辞書を用いて、ワークステーション上に試作した画像検索システム<sup>3), 9)</sup>について述べる。

#### 3. 1 画像検索システムの概要

本システムは、自然言語文による質問文を入力とし、この質問文の意味ベクトルに近い特徴ベクトルを持つ画像を検索して、その結果をランク付けして表示するものである。画像の特徴ベクトルには、画像に付与された説明文から作成した意味ベクトルを使用する。データベースに登録した画像は、自然の風景や都市の景観などの風景写真<sup>10)</sup>と、農作業やビルの建設現場などの産業のイメージ写真<sup>11)</sup>の約 170 枚である。今回利用したこれらの写真には、あらかじめ簡単なタイトルが付けられている。

これらの写真に、一枚につき画像の内容を表わす数語程度の説明文を付与した。説明文は、単語の羅列の場合もあれば、短い文章となっているものもある。説明文を付与するために要した時間は、入力も含めて一枚の画像につき約 1 分という極めて短い時間である。今回付与した説明文の一例を表 3・1(a)に示す。

次に、前述の百科事典テキストから作成した意味ベクトル辞書を用いて、各画像の意味ベクトルを作成し、これを画像の特徴ベクトルとして登録するとともに、画像データの登録も行い、画像データベースを作成した。画像の意味ベクトルの例を表 3・1(b)に示す。

表 3・1 画像の特徴ベクトルの例

(a) 画像に付与した説明文

画像のタイトル	説明文
フランス ベルサイユ宮殿	豪華な宮殿, 広々とした手入れの行き届いた庭。
ビジネスイメージ (接客)	ビジネスマン, 握手, 接客, 接待, 会議

(b) 画像の意味ベクトル

画像のタイトル	意味ベクトル
フランス ペルサイユ宮殿	6ポイント：美しい 5ポイント：住宅 4ポイント：ヨーロッパ 2ポイント：建造物，植物，...
ビジネスイメージ（接客）	5ポイント：共同 4ポイント：経済，商業 3ポイント：住宅，大きい 2ポイント：会社，活動，話す，...

### 3. 2 意味ベクトルによる画像検索のファンクション

画像検索システムをブラックボックスとして、質問文に対して検索された画像とその説明文だけから判断した、実現されたファンクションを次に示す。

#### (1) 単語の意味の類似性に基づく検索

図3・1は、「作物の収穫」という質問文に対し、「農業イメージ」というタイトルが付けられた一連のシリーズが検索されている。表示しているのは、第4位に検索された写真である。この説明文には、質問文中の単語「作物」や「収穫」は含まれていないが、意味的に近い単語が多く含まれている。

#### (2) 連想的検索

「研究」という質問文で検索し、第3位となった画像を図3・2に示す。「研究」という言葉そのものは画像の説明文に含まれていないが、世界各地の風景写真を多く含む本画像データベース中で、「研究」の持つイメージに近いものをうまく「連想」して検索できていると思われる。

#### (3) 百科事典の知識に基づく検索

図3・3は、「リゾート」という質問文で検索を行った結果である。表示しているのは、第3位の画像である。この画像に付けられた説明文には、図3・3左下に示すよう

に、単語「リゾート」が含まれていない。百科事典における「ハワイ」の説明から、「ハワイ」の意味ベクトルが、「リゾート」の意味ベクトルと類似したものになったと考えられる。この画像以外も、リゾート地の海岸が上位に選ばれている。説明文に単語「リゾート」が含まれているのは、第1位となった「アメリカ キーウエスト」だけである。

#### (4)感性検索

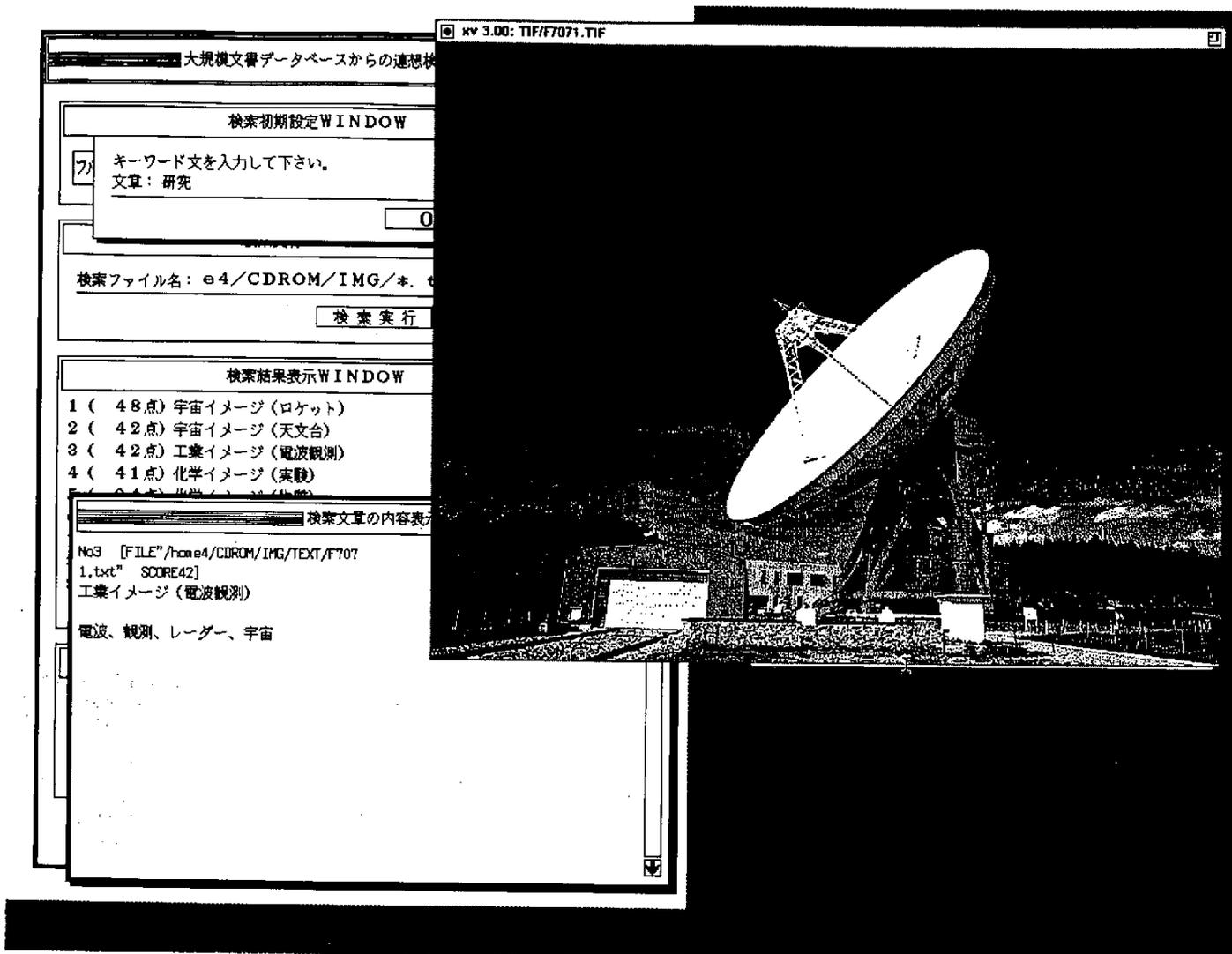
図3・4は、「眺めの良い所」という質問文で検索を行った結果である。表示しているのは、第2位の画像である。「眺め」や「良い」などの単語が含まれていないにもかかわらず、検索されている。この画像以外は、都市を高い位置から撮影したものが上位に選ばれている。第1となった「ブラジル リオデジャネイロ」の説明文は「山の中の首都。高層ビル。遠景。」である。その他の画像の説明文にも、「眺め」や「良い」などの単語は含まれていない。

このように、対象とする画像に関して専門的な知識を持たない素人が付与した簡単な説明を、画像の特徴ベクトルとしても、百科事典を基に生成した意味ベクトルによる連想検索によって、妥当な画像が検索できることを確認した。



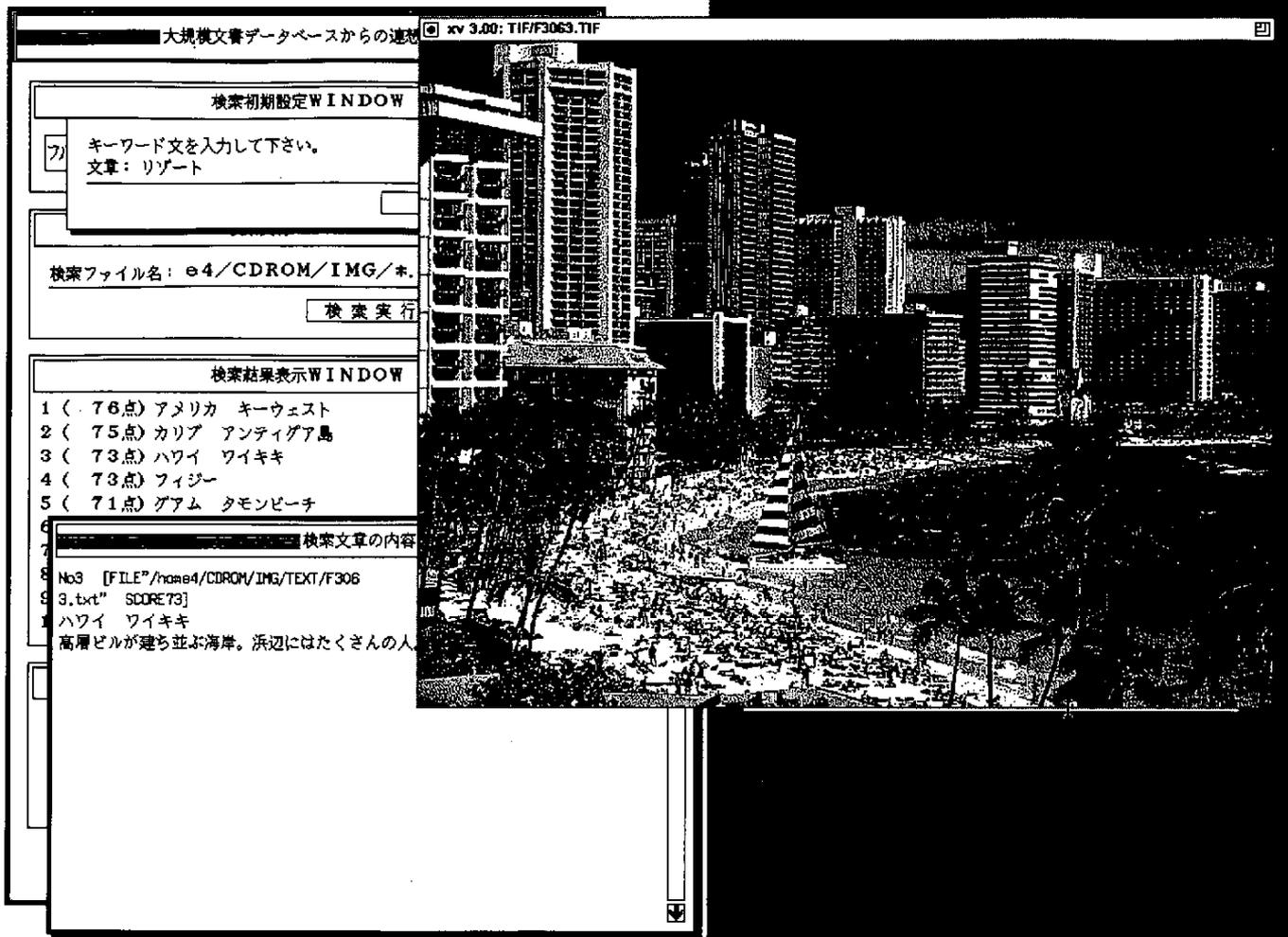
(写真提供 (株)イメージライブラリー)

図3・1 「作物の収穫」で検索の結果、第4位に検索された画像



(写真提供 (株) イメージライブラリー)

図3・2 「研究」で検索の結果、第3位に検索された画像



(写真提供 (株) イメージライブラリー)

図3・3 「リゾート」で検索の結果、第3位に検索された画像



(写真提供 (株) イメージライブラリー)

図 3 ・ 4 「眺めの良い所」で検索の結果、第 2 位に検索された画像

## 4. 写真検索ソフトウェアとユーザーインターフェース

現段階での写真検索を必要とするユーザとして、マッキントッシュ上で作業を行なうデザイナーを想定しマッキントッシュでの写真検索システムを開発した。

### 4. 1 プロセスの概要

図4. 1 にシステム全体のフローを示す。システムはスタート後に単語辞書の読み込み等の初期化処理を行った後、メイン画面とメニューを表示する。検索を行う場合には検索対象となるデータベースを指定し、検索方法を指定した後、質問文を入力し検索を行う。検索結果は上位9画像ずつ出力され、画像での表示と説明文での表示との切り替えが出来る。データベースを登録する場合には後述する登録ウィンドウ上でテキストファイルとイメージファイルを指定して行う。

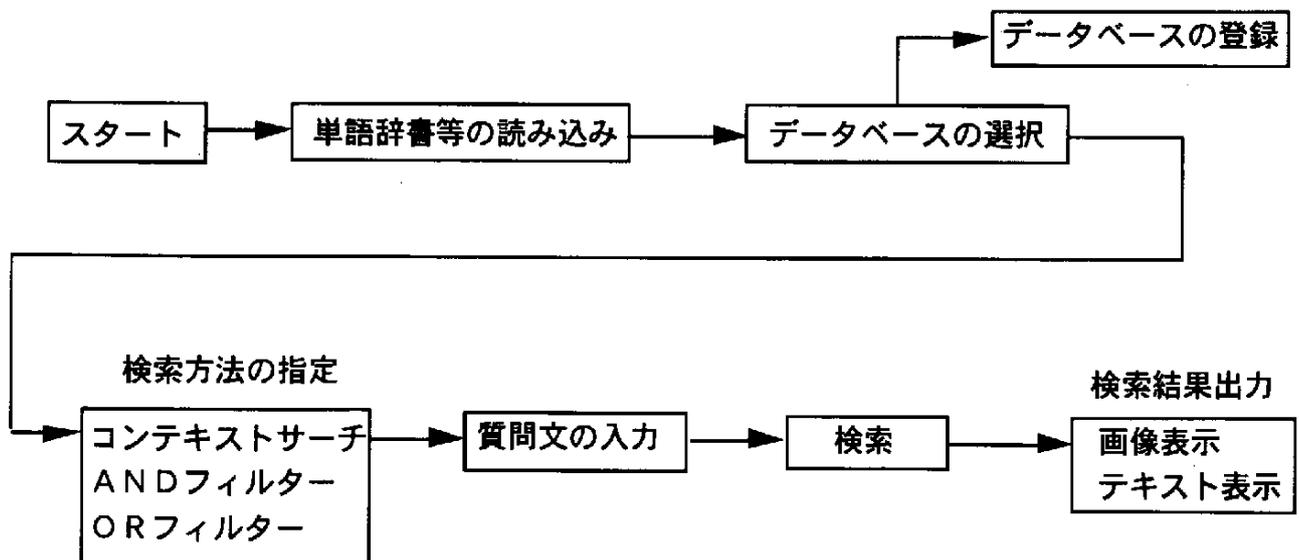


図4. 1 写真検索システムのフロー

### 4. 2 ユーザーインターフェース

図4. 2 はシステムのメイン画面で、左画面上でラジオボタンによる検索方法の指定と質問文の入力を行ない、続いて検索ボタンを押すことで検索が開始される。右画面に

は検索結果として9画像が出力される。出力はラジオボタンにより画像出力と画像の説明文の出力との切替が可能である。従ってこのシステムは説明文の出力に限定すればテキストデータの検索にもそのまま使用できるものとなっている。また検索された画像からのフィードバック検索も可能となっている。各ボタンの役割は下表の通り。

表4. 1 写真検索システムのメイン画面上のボタン

CONTEXT SEARCH	ベクトルのみによる検索の指定
AND FILTER	質問文から抽出された単語の論理積を満たすレコードを対象にベクトル検索
OR FILTER	質問文から抽出された単語の論理和を満たすレコードを対象にベクトル検索
検索	検索の開始
IMAGE	検索結果を画像で表示する。
TEXT	検索結果をテキストで表示する。得点とイメージファイル名も同時に表示する。
FEEDBACK	フィードバック検索の指定

フィードバック検索とは、指定した画像の意味ベクトルに近い画像を順に表示する検索方法である。さらに検索された画像をクリックすることでフルサイズの画像を表示することも可能である。

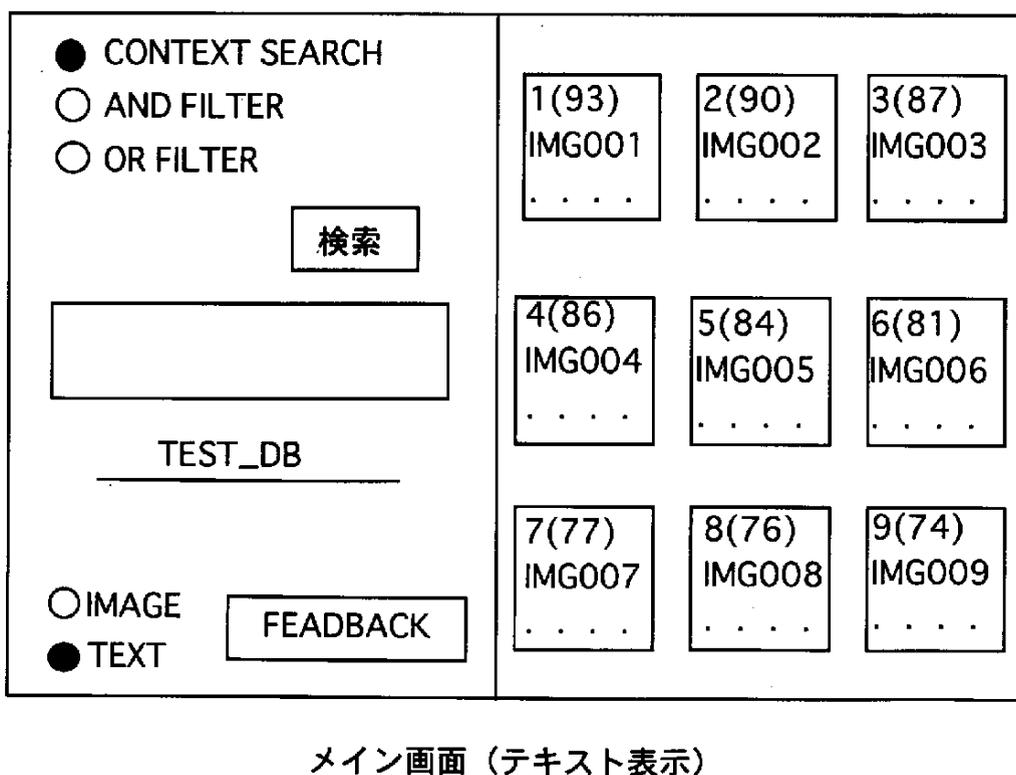
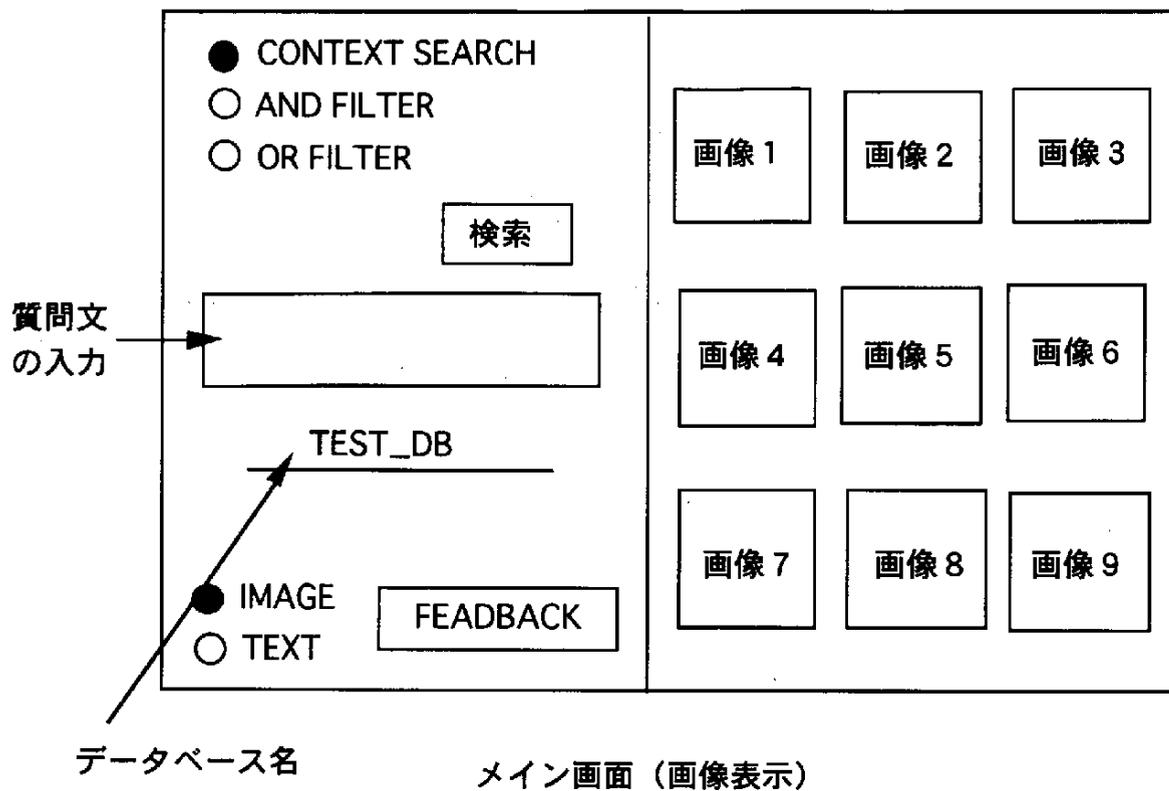


図 4. 2 写真検索システムメイン画面

図4.3はシステムの持つメニュー項目とメイン画面以外のウィンドウを示している。図に示すように検索対象となるデータベースの選択、データベースの登録はメニューから選択し各々のウィンドウ上で行なう。またシステムは過去に入力した質問文を保持しており、その中から選んで質問文として入力することも可能である。

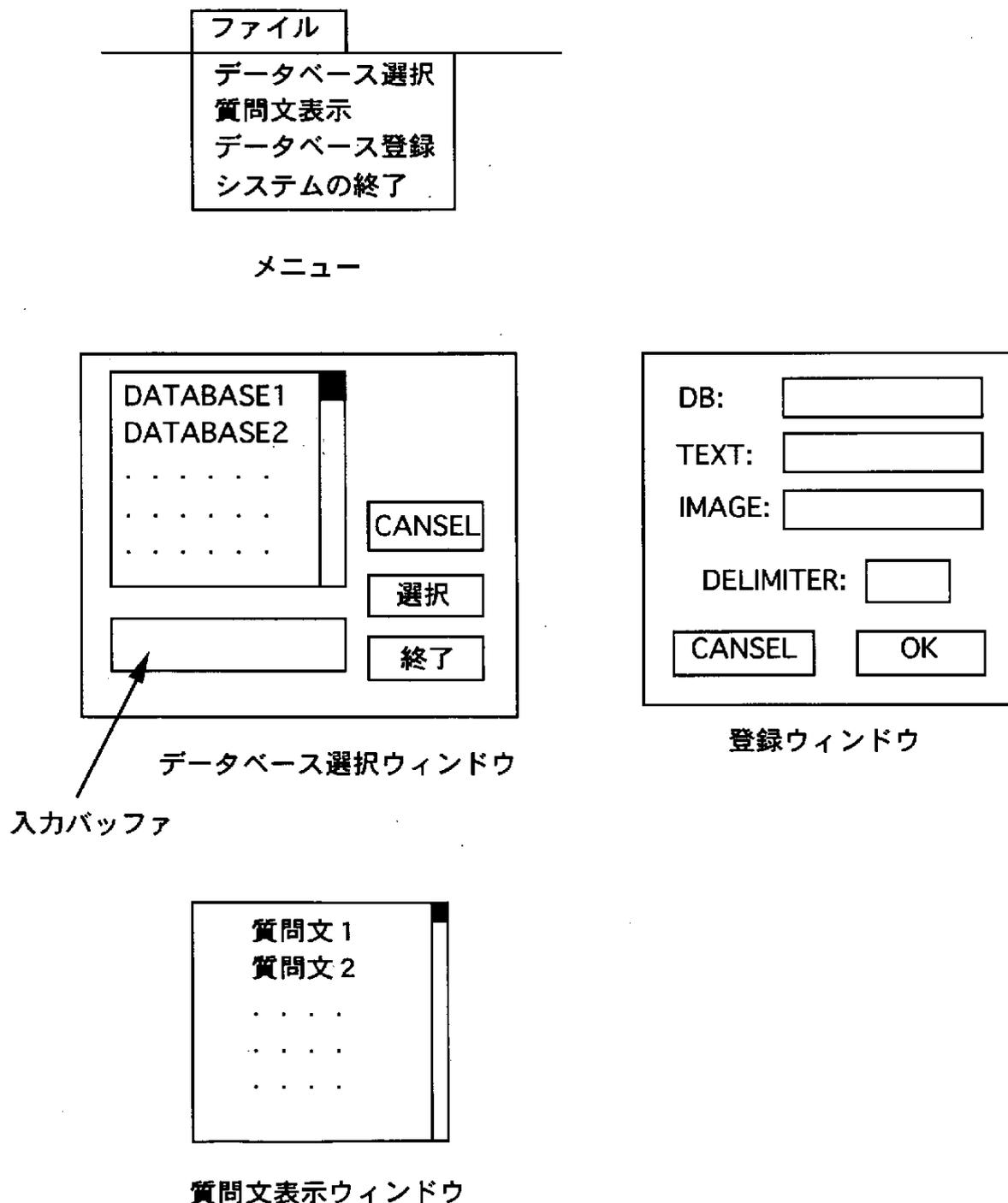


図4.3 写真検索システムメニューと各ウィンドウ

## 4. 3 データベースの構造

### 4. 3. 1 データベースの概要

検索システムのデータベースは各画像のイメージファイル、全画像の説明文をデリミターで区切って保持するテキストファイル、各レコードのベクトル、テキストアドレスおよび各単語のインデックスを保持するデータベースファイルに分かれる。データベースの登録の際に登録ウィンドウ内にテキストファイル、イメージファイルを指定することでデータベースファイルが作成される。以降の検索の際にはこのデータベースファイルを指定する。登録ウィンドウに入力する各項目は下表の通り。

表4. 2 登録ウィンドウの各項目

DB	データベースファイルの指定
TEXT	テキストファイルの指定
IMAGE	イメージファイルを保持するフォルダ名の指定
DELIMITER	デリミターの指定

### 4. 3. 2 データベースファイルの構造

図4. 4にデータベースファイルの構造を示す。データベースファイルにはテキストファイル名とイメージファイルを保持するフォルダ名が保持されており、各説明文のアドレスと長さ、各画像イメージを縮小したサムネイル、各単語に対しそれを含む説明文番号を保持する単語のインデックス情報を内部に保持している。

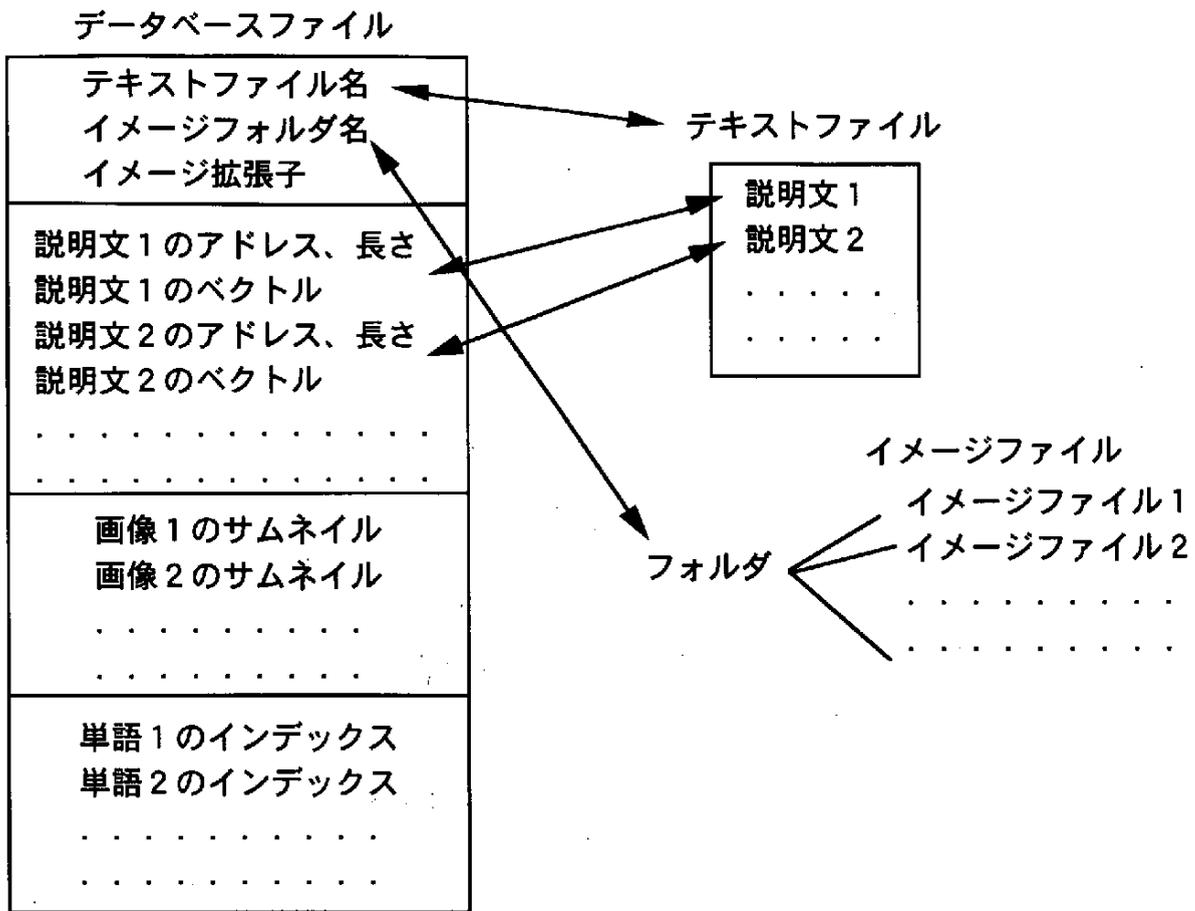


図 4 . 4 データベースの構造

#### 4. 4 まとめ

意味ベクトルによる検索技術の応用として写真検索システムのプロトタイプをマッキントッシュ上に開発し、その概要について説明した。このシステムは上で述べたように写真検索に限らず百科事典などのテキスト検索にも利用できる。また出力方法さえ変更すれば音声、動画等も含むマルチメディアデータの検索にも拡張できる。また利用者にとって使いやすいようユーザーインターフェースに工夫を凝らした。今後の実使用においてさらに操作性が増すようユーザーインターフェースも改良されるべきである。本システムが実際の写真検索の場で生かされ、またユーザーインターフェース部分も含めて、今後の動画等も含むマルチメディアデータ検索装置に応用されるものと期待する。

## 5. 画像検索評価用ベンチマークの作成

### 5. 1 画像データベース

ベンチマーク作成に使用した画像データベースは（株）JPRC 提供の写真データ<sup>4)</sup>から2万件を選び出したものである。各写真にはタイトル、分類コード、地域コードの他、「豪華」「ロマンティック」「驚き」などのイメージコード、画像の主題を表わすテーマコードなどが付加されている。下表に分類コード、地域コードの一部を示す。

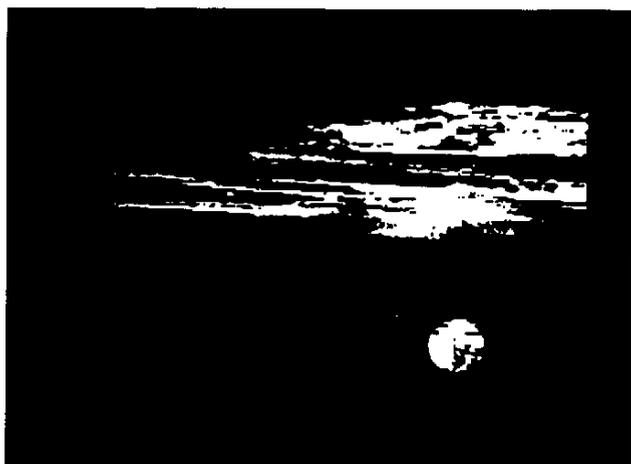
表5. 1 コードの例

1 1 1 1 0 0	風景、自然	1 1 1 0 0	北海道
1 1 1 1 1 1	山	1 1 2 0 0	東北
1 1 1 1 1 2	火山	1 1 2 1 1	青森
1 1 1 1 1 3	岩、岩山	1 1 2 1 2	秋田
1 2 1 1 0 0	太陽（日の出、夕日）	2 1 1 0 0	ヨーロッパ
1 2 1 2 0 0	月	2 1 1 1 1	アイスランド
1 2 1 3 0 0	青空	2 1 1 1 2	アイルランド
4 1 1 7 0 0	観葉植物	3 1 5 0 0	中南米
4 1 1 8 0 0	キノコ	3 1 5 1 1	アンティグアバーブーダ
4 1 1 9 0 0	木、樹木	3 1 5 1 2	エルサルバドル

分類コード（一部抜粋）

地域コード（一部抜粋）

各写真に付加されたタイトルと全てのコードをテキストに置き換えたものを画像の説明文として、検索システムを構成した。画像と説明文の一例を図5. 1に示す。



夕日,西ドイツ,朝夕,オレンジ,モノトーン,  
太陽(日の出夕日),シルエット,ロマンテ  
イック,静かな 静止,立木,シルエット,ロマ  
ンティック,静かな 静止,雲,



住宅のデザイン,オランダ,ロッテルダム,家  
(洋風),パターン 模様,街角(風景 街),パター  
ン 模様,家 塀 壁(風景 建物),パターン 模様,  
建物のデザイン

(写真とその説明文の提供 (株) JPRC)

図5. 1 画像と説明文の例

## 5. 2 画像検索システムの評価

### 5. 2. 1 概略

画像検索システムの評価方法の概略は次の通りである。まず画像データから一定数抜き出したものを評価用のサンプルデータとし、サンプルデータに対して複数個の質問文と正解のペアを予め人手によって作成する。サンプルデータを対象に同じ質問文をシステムに入力し、得られた検索結果と人手で作成された正解データとの一致度で評価を行なう。ここで予め正解データを作成しておくのは、検索の度に出力データに対し正誤判断を行なった場合には、人間の判断が検索の時々で一定せず従って評価が一定しない、あるいは評価の自動化が出来ないなどの理由による。

### 5. 2. 2 評価用データ作成の概略

画像の持つ意味ベクトルは画像に付加された説明文から機械的に作成されたものであるため、画像の主観的特徴を正しく表しているとは限らない。従って画像検索システムを評価する場合、画像から得られる主観的類似度に基づいた評価が不可欠であると考えられる。そこでサンプルデータおよび質問文と正解のペアの作成を説明文を考慮に入れず画像のみを見て行なった。以下にその作成方法の詳細と作業の効率化のために作成したツールについて述べる。

### 5. 2. 3 評価用データ作成の詳細

予め質問文を用意してからそれらに合致する画像を選別する、あるいはある画像に対応する質問文を作成した後その質問文に合致する画像を選別する等の作業は、自然言語である質問文と画像から得られる主観的印象とを常に対応させながら行なう必要があるため比較的複雑で効率が悪い作業である。それに比べ画像から主観的に類似性を判定することは容易である。従って評価用の質問文と正解のペアを作成する際、まず画像を見て似ていると判断したものが同じグループに属するようにグループ分けを行ない、その後各グループに対応する質問文を作成した。またグループ分け作業を効率的に行なう

ため以下で説明するグループ分け補助ツールを使った。

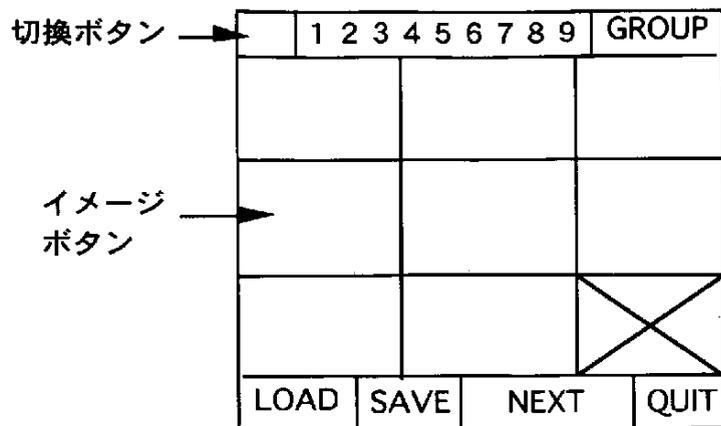
### グループ分け補助ツール

グループ分け補助ツールは、画面上に複数の画像を表示し利用者がそれらを見て対応するグループの番号をマウス操作により入力することを繰り返すことでグループ分けを行なうGUIツールである。図5. 2に本ツールのコントロールウィンドウと動作例を示す。画像は全画像データからランダムに選び一度に8画像を表示する。作業終了時にグループ分け情報をファイルに書き込む。各ボタンの役割は以下の通り。

表5. 2 グループ分け補助ツールのボタン

1～9	グループ番号の指定、切替ボタンで10番以降の指定も可能。
GROUP	グループ番号の指定後クリックすることで対応グループに属する画像が出力される。
イメージボタン	グループ番号の指定後クリックすることで対応する画像がその番号のグループに登録される
LOAD	グループ分け情報をファイルから読み込む。
SAVE	グループ分け情報をファイルに書き込む。
NEXT	次の8画像をランダムに選び出力する。

あらかじめ各グループ番号に対応する画像イメージは想定せず、表示された画像に対し主観的にグループ番号を決めながら、対応するグループに属する画像の個数に応じて調整を行なっていく。画像の中にはどのグループにも属しにくいもの、画像自体の意味が良く分からないものが多数あり、それらにはグループ番号は付与しない。またグループ分けの過程においては画像の説明文は一切参照していないため、画像の意味を取り違えて間違ったグループに分けてしまう可能性もあるが、上で述べたように画像の主観的類似度に基づいた評価としては差し支えない。



コントロールウィンドウ

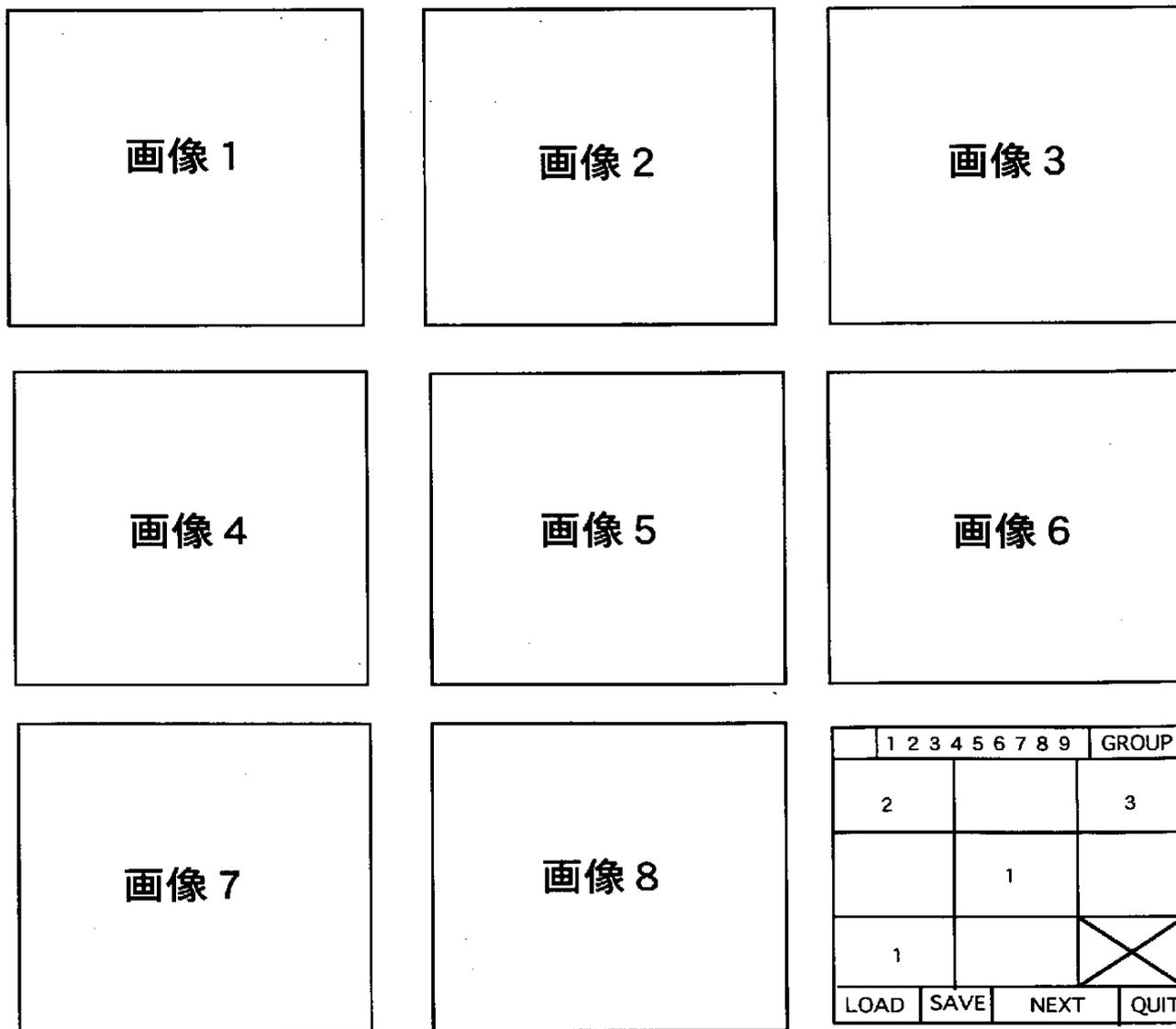


図 5. 2 グループ分け補助ツール

## 質問文の作成

作成した各々のグループに対して合致する質問文を作成する。質問文によって評価結果が大きく変わる可能性があるため3組の質問文を用意した。グループ分けは画像からの主観的印象のみで行なったものであるため、それらに丁度合致する質問文の作成は少し困難な作業である。

## サンプルデータ

グループ分けされた全てのデータをまとめたものをサンプルデータとする、従ってサンプルデータ内の全ての画像は必ずあるグループに属している。

## 作成結果

以上の作業により作成された評価用データは、全画像数104、グループ数14となった。各グループに属する画像数とグループの概要は下表の通り。

表5.3 グループ分け結果

グループ番号	画像数	グループの概要
1	10	都市の夜景
2	9	ビルの建ち並ぶ大都市
3	14	山村風景
4	10	野鳥
5	15	草原の野生動物
6	11	女性
7	8	若いカップル
8	7	町の商店
9	14	城の様な建物
10	10	祭り
11	7	川と小舟
12	8	ビーチ
13	5	花
14	11	部屋の中のペット

## 5. 2. 4 評価結果

システムが出力する検索結果は画像を質問文に近いものから順に得点付けて出力するものであるため、従来の検索で用いられた含有率、適合率による評価法を直接適応することは出来ない。また従来の検索とは異なり曖昧性を許すものであり、また検索された画像をもとにしたフィードバック検索も可能なため、検索結果の上位に少しでも正解画像があれば良いとした評価を用いる。

そこでシステム出力の上位3位以内に1つでも正解画像があれば検索が成功したものとみなし、14個の質問文の中で検索が成功する割合を求めた。以下が3組の質問文で評価した結果である。

1組目の質問文：13 / 14 = 93%

2組目の質問文：14 / 14 = 100%

3組目の質問文：11 / 14 = 79%

以上の結果から正解の画像がほぼ確実に上位に検索されることがわかる。よって意味ベクトルによる検索の有効性が確認された。

## 5. 2. 5 問題点

上述の評価方法における問題点を以下にまとめる。

評価用データの作成に関して

○グループ分け作業において作業者が独自にグループ番号とそれに対応する画像イメージを想定しなければならいため、作業は単純なものとはならずまた作成されたデータは作業者に強く依存する。

○出力する画像は全画像からランダムに選ばれているため、全体の中で多数派で

あるイメージを持った画像が中心に選ばれる。

- 画像のみからグループ分けを行なっているため実際に画像が意図するものとは異なったグループに属している可能性がある。
- グループ分け作業は質問文を意識せずに行なうため、後でグループ内の全画像に対して合致する質問文を作成するのが困難である。

#### 評価方法に関して

- サンプルデータは小規模なものであり、また利用者の入力するあらゆる質問文に対してどの程度の検索精度が得られるかについては評価出来ない。
- 評価結果として得られた数値とシステムの実使用での有効性との関係が不明である。

### 5. 3 まとめ

以上で述べてきた評価方法は、画像検索において利用者が大まかな印象を入力した場合に妥当な画像が検索されるのかを評価する場合を想定したものである。評価結果から上位3位以内に80～100%の確率で少なくとも1つの正解画像が検索されており、意味ベクトルによる検索の有効性が確認された。しかし実際の検索では例えば絵画の検索を行なう場合に作者名を明示的に指定する場合もある。このような場合も含めた評価方法を確立すべきである。またこの検索方式は入力キーワードを設定せず任意の言葉で検索可能なものであるため、本来は利用者が入力すると想定される全ての質問文に対しての評価が必要であろう。従って今後百科事典等の電子化辞書を用いた大規模な評価が必要であると考えられる。本章で提案したものは主に画像から得る主観的な印象に基づいて検索を行なう場合の1評価方法である。

## 6. 今後の課題

写真データベースの利用者は、デザイナーが多く、連想的、感性的に写真を検索したいとの要求が強い。現在、実用化されている写真データベースは、単純なキーワード検索方式が主流であり、デザイナーの要求を満たしていない。本プロトタイプのように、デザイナーが利用するPC（マッキントッシュ）上で連想的に写真を検索できるツールが提供されれば、我が国の写真データベースの普及に大いに寄与すると考えられる。

今後の課題を以下に記す。

(1) フォトライブラリー業者と協同で、膨大な数の写真を格納した写真データベースを開発し、デザイナーや出版社に提供し、ユーザの意見を集めて検索技術やユーザインタフェースに反映させる

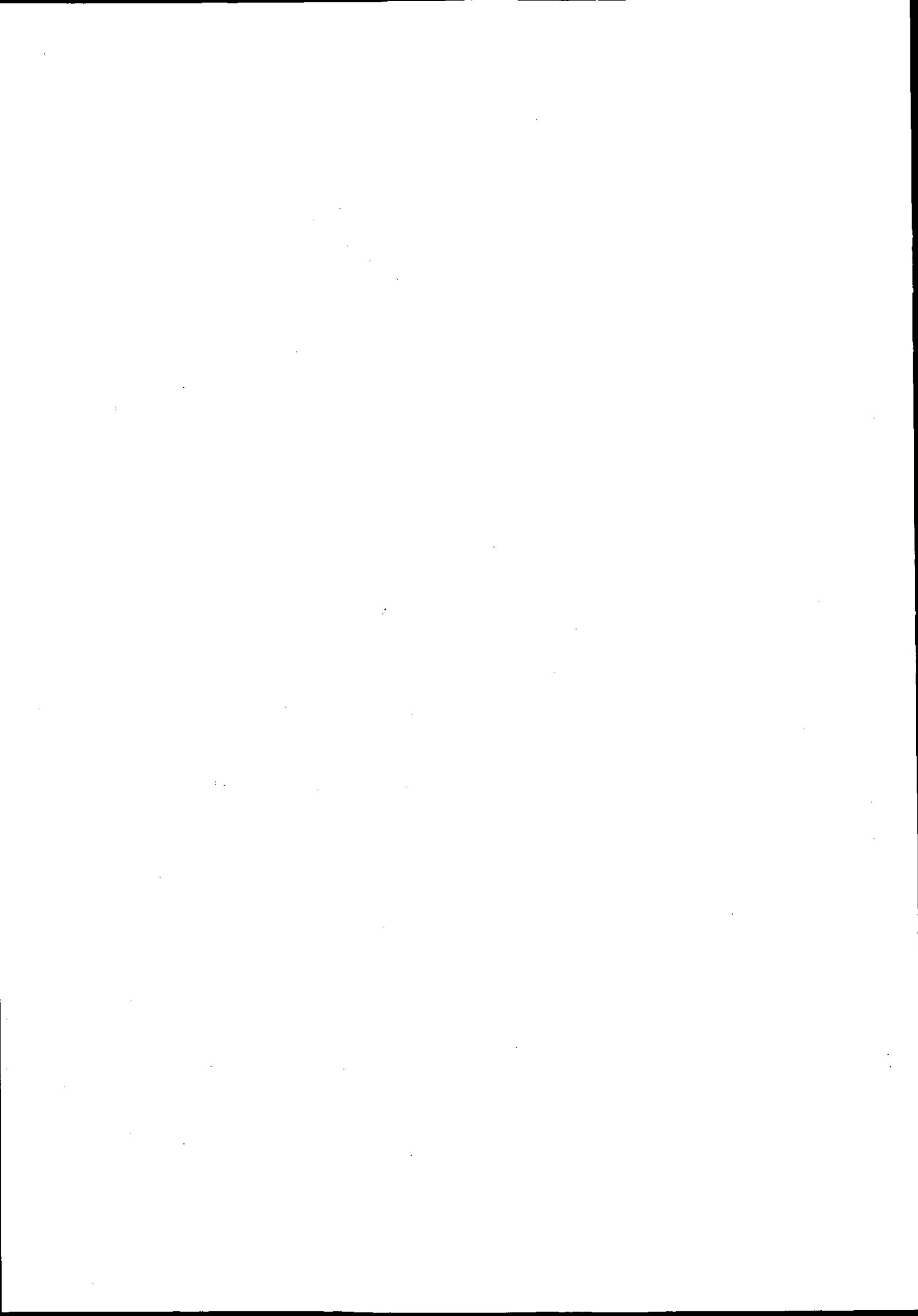
(2) 本プロトタイプシステムを利用して写真データベースを構築するためには、ワークステーションでアプリケーション用の意味ベクトル辞書を作成する必要がある。汎用的な意味ベクトル辞書を対象テキストに適応させる学習機能をマッキントッシュ上のツールとして作成する必要がある。

(3) 意味ベクトル技術は、言語への依存性が少なく、日本語以外の言語への変換も可能である。次のステップとして、英語への対応を行なう。

今後このような研究を実施することによって、写真データベースによる地域紹介、旅行案内、美術鑑賞、図鑑としての利用、家庭での写真の整理など、各種の目的に応じた写真データベースの構築が促進されると考えられる。

## 参考文献

- (1)芥子, 乾, 石鞍: 大規模文書データベースからの連想検索, 電子情報通信学会技術報告, AI92-99. pp.73-80 (1993).
- (2)芥子, 池内, 小淵: 意味ベクトルによる百科事典テキストデータベースの構築, Proceedings of Advanced Database System Symposium'93, pp.227-234 (1993).
- (3)池内, 芥子: 意味ベクトルによる画像検索の試み, 日本ソフトウェア科学会第11回大会論文集, pp.413-416 (1994).
- (4)JPRCフォトディスク, 株式会社日本フォト・リサーチ・センター, (1998, 1991)
- (5)Image Highway ver. 1.0 (サムネイル盤, カンプ盤A・B), 株式会社日本フォト・リサーチセンター, (1994).
- (6)ブリタニカ小項目事典 CD 版, TBS ブリタニカ.
- (7)Waltz, D. L., and Pollack, J. B. : Massively parallel parsing: A strongly interactive model of natural language interpretation, Cog. Sci. 9, 1985.
- (8)Gallant, S. : A Practical Approach for Representing Context and for Performing Word Sense Disambiguation Using Neural Networks, Neural Computation 3, 1991.
- (9)田中, 芥子, 池内: 意味ベクトルによる画像検索, 情報管理, pp.579-585 (1994).
- (10)世界の風景総集編, Vol.3, 株式会社イメージライブラリー (ImageCD).
- (11)ビジネス・産業総集編, Vol.7, 株式会社イメージライブラリー (ImageCD).



————— 禁無断転載 —————

平成 7年 3月発行

発行 財団法人 データベース振興センター  
東京都港区浜松町二丁目4番1号  
世界貿易センタービル7階  
TEL 03-3459-8581

委託先 シャープ株式会社  
大阪市阿倍野区长池町22番22号  
06-621-1221

印刷所 シズミ印刷所  
奈良県天理市櫛本町

