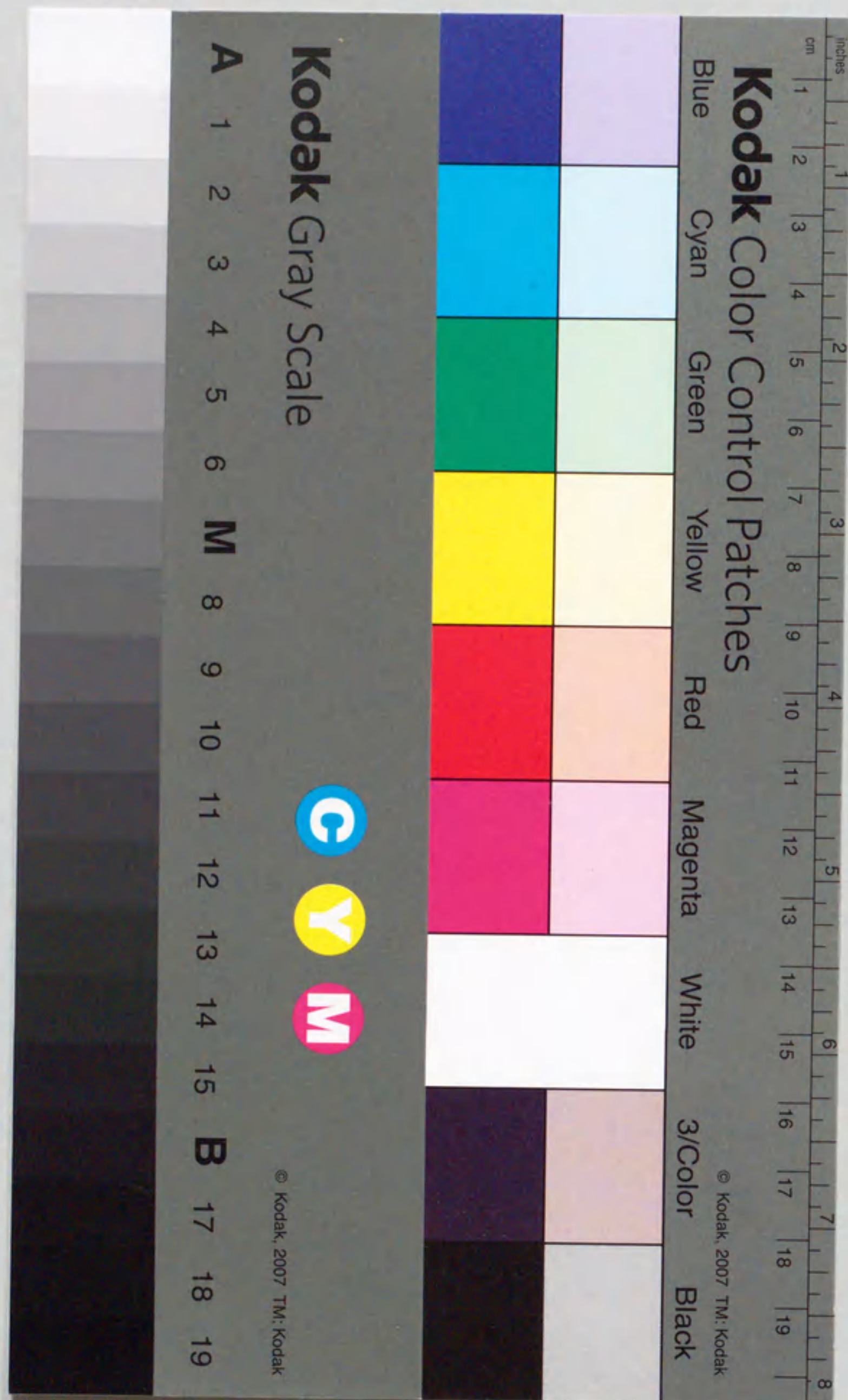


曖昧さを許容する計算機インタフェースの研究

高橋 友一



曖昧さを許容する計算機インタフェースの研究

高橋 友一

目 次

第1章 序論

1.1	研究の目的	1
1.2	研究背景と基本方針	1
1.3	関連研究の動向	3
1.3.1	画像データベース検索方法の研究	3
1.3.2	位置関係に関する研究	5
1.3.3	手振りを利用したインタフェース	8
1.4	研究の概要と本論文の構成	10

第2章 位置関係の表現

2.1	位置関係の特徴	12
2.2	位置関係の表現	14
2.2.1	漠然さの表現	14
2.2.2	位置関係の漠然さに関する主観実験	18
2.2.3	ルールを用いた位置関係の表現	19
2.2.4	合成語の漠然さの表現	28
2.3	位置関係の作成	29
2.4	第2章のまとめ	33

第3章 位置関係を用いた画像データベース検索

3.1	画像データの管理方法	34
3.2	位置関係を手がかりとする検索方法	36
3.3	SPADEシステム	38
3.4	試作データベース	42
3.4.1	浮世絵画像データベース	42
3.4.2	文書画像データベース	48
3.4.3	考察	48
3.5	第3章のまとめ	53

第4章 インタフェースとしての手振り	
4.1 手振りの表現	54
4.2 手形状入力装置による手の形識別	55
4.2.1 主成分分析	61
4.2.2 クラスタ分析	63
4.2.3 コード化方法による手の形表現	66
4.2.4 考察	70
4.3 第4章のまとめ	70
第5章 手振り認識方法	
5.1 手の動きのコード化方法	72
5.2 手形状入力装置を用いた手の形・動きの認識方法	73
5.3 手振り認識実験	75
5.3.1 手振りの区切りを与えた条件下での手の形認識実験	79
5.3.2 手振りの区切りを与えない条件下での手の形と動き認識実験	80
5.3.3 手振りの区切りを与えない条件下での手の動き認識実験	82
5.3.4 考察	82
5.4 第5章のまとめ	83
第6章 マルチモーダルインタフェースへの適用	
6.1 地図案内システム I M A G E の概要	84
6.2 指示内容の統一表現	90
6.3 I M A G E システムの評価	95
6.4 第6章のまとめ	99
第7章 結言	100
謝辞	103
参考文献	104

第1章 序論

1.1 研究の目的

半導体の集積度に代表されるハードウェアの進歩や文字データのほかに音声、画像データ等処理するソフトウェア技術の進歩により、計算機の利用方法は多種、多様な方面に広がっている。情報処理システムの発展、普及に伴い、計算機と人間とのインタフェースの形態も、さらに高度化を目指して、キーボード入力やテキスト出力のほかに音声入出力、画像入出力専用の装置などの新しい装置が開発されている。

人間と計算機とのインタフェースを、人間にとって指示しやすいインタフェースに近づけるためには、装置の開発のほかに、計算機に人間の特徴である「曖昧さ」のモデルを用意する必要がある。このような見地にたって、本研究は、人間にとって使いやすく、情報の伝達効率のよい情報処理システムを実現するために、計算機と人間とのインタフェースの高度化を目指して、以下の2点について研究を行った。

第1の目的は、人間が日常生活に用いている定性的な空間指示の言葉のモデルを評価し、計算機のインタフェースに応用することである。具体的には、「. . の右にある. . .」といった位置関係を用いた指示をもとに、座標値で管理されている対象データのなかから適切な対象を同定する事である。計算機のもつ定量的な数値情報に対し、定性的な言葉を用いて指示する事により、より使い易いインタフェースが実現できる。

第2の目的は、インタフェースに身振り(手振り)情報の利用を検討する事である。人間同士の会話では、言葉によって伝えられる情報に、その時の声の抑揚や身振り(手振り)からの情報を加えて意味を解釈している。手振りにより表現されている情報を認識し、インタフェースに利用する事により、計算機の新しい利用方法がひらける。

1.2 研究の背景と基本方針

計算機インタフェースに関する研究には、入出力機能の効率向上を目指したキーボード入力方法¹⁾およびポインティング装置に関する研究と、アプリケーション

ンにおけるデータ入力形式や方法を論じた研究が多い。後者の研究の例として、データベース検索時の入力インタフェースの研究がある。データベースへの問い合わせを、ディスプレイ上に表示された表形式や自然言語の形で入力し、データベースマネジメントシステム専用の問い合わせ言語に変換し検索する研究はその一つである。

さらに、専門家でなくても利用できるように、音声入力によりインタフェースを向上させる研究や、対話処理から利用者の検索意図の曖昧さを解消し検索効率を高める研究がなされている²⁾。人間にとって使いやすいインタフェースを実現する為に、プログラムで決められたデータ形式に従って入力する方法のほかに、多種多様な方法を用いて入力されるデータを解釈し、全体として必要な情報にまとめる入力方法が検討されている^{3), 4)}。

マルチメディアデータベースは、著者名、キーワードといったテキスト情報と図表、写真等の画像データを含んでいる。マルチメディアデータベースの検索には、テキスト情報をてがかりにする検索に加え、「誰々の絵で、富士山が真ん中にあった絵」のように、利用者が持つ画像イメージに言及した内容検索が必要になる。このような内容検索では、利用者が主観的に感じるイメージを検索の手がかりとするため、内容検索に十分なキーワードをあらかじめ付与する事は、文献検索に対するキーワード付与に比べ難しいといえる。

画像データベースの内容検索を実現するには、キーワードに基づくテキスト情報の検索のほかに、定性的な言葉で表現されるイメージを解釈し、検索手がかりにすることが必要になる。本研究の第1の基本方針は、対象の頂点座標をもとに、「右」といった定性的な位置関係を表現することにある。位置関係のモデルを用いて、定量的な数値ではなく、位置に関する言葉により、対象を指示できる使いやすいインタフェースを検討する。

自然言語理解の研究、音声認識・合成の研究などのほかに、声質や画像処理を用いた手振り、口の動きの解析などの非言語によるインタフェースの研究が行われている⁵⁾。情報処理システムの発達、普及に伴い計算機と人間のインタフェースは、計算機の処理にあった方法から、人間が日常生活で使用している表現や方法を倣う方向にある。人間の指示動作を識別し、その動作の意味を記号で表現できれば、インタフェース設計の役にたつ。人間と計算機のインタフェースにおい

て、使用される言語や動作に関する検討は、今後より一層不可欠になると思われる。本研究の第2の基本方針は、手形状入力装置を用いて、手振りによる指示を計算機インタフェースに適用することにある。その為に、指の曲げや手の向きで手の形状を表現し、手の形、動きを認識する技術を検討する。

1. 3 関連研究の動向

これまで行われてきたインタフェース研究のうち、画像データベース検索方法の研究、位置関係に関する研究、手振りをを用いたインタフェースの研究について概観する。これらの研究においては、人間の指示における曖昧さを扱ったインタフェースが検討されている。

1. 3. 1 画像データベース検索方法の研究

画像データを含んだマルチメディアデータベースの研究は、データベース構造、検索手がかり（検索インデックス）の自動付与方法、検索インタフェース、入出力機器等、多岐にわたっている^{6), 7)}。画像データは、文字データと比較して多くの情報を含んでいる。画像データベースの検索方法として、データベース管理者が付与するキーワード情報を用いた検索方法のほかに、画像の性質を利用した検索インデックスの自動作成法や視覚的インタフェースによる検索方法が研究されてきた。

ここでは、画像データベースの検索方法を検索インデックスから分類する。そして、画像に含まれる対象の位置関係を検索の手がかりにしている画像データベースシステムについて概観する。

(1) 画像内容を反映したキーワードで画像を検索する方法。

画像に付随した情報を、主に人手で該当画像データのキーワードとして入力し、CODASYL型やリレーショナル型データベースで管理する。画像データは2次ファイルに格納する。文献データベースと同様にキーワード情報を用いて検索要求に該当する画像識別子を検索する。検索された識別子と同じ識別子をもつ画像を表示する。内容検索は付与されたキーワードにより限定されるが、LANDSAT画像を、経緯度、撮影日付等のキーワードで管理している画像データベースIMDS等実用例が多い⁸⁾。大川らは、人間が

画像内の対象情報から画像を解釈する時に用いる知識を用い、キーワード情報以上の情報を用いた検索方法を提案した⁹⁾。この方法は、キーワード検索に知識処理を加えた検索方法といえる。

(2) 表示されたガイド画像を用い検索する方法。

静止画検索システム F O R K S¹⁰⁾のように、ディスプレイ上に表示されている複数の縮小画像をもとに、目的とする画像をブラウジング検索する。笠原らは、照明器具画像データベースに対し、ディスプレイ中央に表示された照明器具と、形や構成品の数など特徴の異なる照明器具の画像から構成される画面を用いて、ブラウジングする画面を変化させる事で、画像の内容を絞っていく方法を提案した¹¹⁾。

(3) スケッチ画を検索キーとし、同じ構造または類似した構造を持つ画像を検索する方法。

検索する内容をスケッチした画像とデータベース内の画像との類似性を手がかりに目的とする画像を検索する。長谷川らは、胸部X線写真に対するスケッチ画を予め自動的に作成する事で、検索時処理時間の高速化、要求に合わせた特徴量の計算などの、検索の多様化を図る手法を提案した¹²⁾。林らは、地図画像をプリミティブ(道路、河川、施設)、構造、インデックスで記述し、検索要求に対応するスケッチ画から抽出されるプリミティブと照合する検索手法を提案した¹³⁾。

(4) 画像の画素ラン数等の特徴パラメータ空間で対象画像を検索する方法。

画像に含まれる対象の形状、大きさ、濃淡、色など画像処理で得られる各種統計量の特徴パラメータを、画像の検索インデックスとして検索を行う。

大田らは、名刺画像を対象に、氏名の漢字に対する文字認識の特徴量をパラメータにした手法を提案した¹⁴⁾。山守らの天気図に含まれている気象情報データを符号化し、蓄積、検索する天気図データベースシステムもこの範疇に入る¹⁵⁾。

(5) 定性的な言葉や感性を数量化して人間の抽象過程により作成される情報をキーとし検索する。

感性情報をキーワードとした検索方法で、感性や印象を表現する言葉やイメージを例示することによりデータ検索を行う。笠原らは、意匠デザインにお

いて感覚的、情緒的な印象で表現されるイメージを感性軸上の数値で表現し、一定範囲の感覚にあうイメージを検索する手法を提案した¹⁶⁾。栗田らは、商標・意匠図を対象とした画像データベースにおいて、各利用者が感じる画像間の類似度を反映した特徴空間で検索を行う手法を提案した¹⁷⁾。

この他に、以上の手法を組み合わせた検索方法も提案されている^{18)、19)}。

画像に含まれる対象の空間的位置、相対的位置関係を検索インデックスとする画像データベース検索システムとしては地図を対象とした I I B²⁰⁾、P R O B E²¹⁾や、水墨画を対象とした山本らの研究²²⁾がある。I I Bでは、対象の外接四角形間の相対位置関係を2-Dstring(直交座標系での外接四角形の座標値の大小関係の集合)で表現し、検索キーに用いる。例示画と同じ位置関係を持つ地図画像の検索を、利用者によりディスプレイ上で示される例示モデル画像に対する2-Dstringとデータベース内の2-Dstringを照合する事で実現した。位置関係の照合を、位置関係を表現している2-Dstringを用いて照合するという点で分類(3)のシステムといえる。P R O B Eでは、対象を対象に含まれている領域の集合(point-set)で表現し、point-set に対する和、積、差集合演算を基に対象間の overlaps, left-of, above, adjacent-to の位置関係に対応する検索機能を実現している。山本らは、水墨画の構図に関する問い合わせを、対象の重心位置・面積・形状等の画像の構造情報に変換して検索を行う手法を提案した。

これらのシステムは、画像データベースの検索手がかりとして対象間の位置関係を用いた特徴あるシステムである。しかし、対象間の位置関係は、幾何学的な関係として扱われ、人間が日常生活で用いる位置関係のもつ漠然さなどを考慮した定性的な関係として扱ってはいない。

1. 3. 2 位置関係に関する研究

システムと人間とのインタフェースにおいて、座標位置による対象の指示方法などは直感的でなくわかりにくい。人間にとっては、「上、下、左、右」といった言葉による指示の方が表現しやすく、わかりやすい。対象A、B間の位置関係を、on(A, B)やleft(A, B)のような述語表現で表現することが考えられる²³⁾。Winstonは、「左」といった定性的な位置関係を幾何学的な

位置関係-Aの重心がBの左端より左にあり、Aの右端がBの右端より左にある関係で定義した²⁴⁾。彼の定義では、図1.1の左半分のケースでleft(A, B)が成立する。

人間が日常会話で用いる位置関係の概念は、対象の位置関係のほかにまわりの状況に影響される曖昧さ、漠然さを含んでいる。「AがBの左にある」では、AがBの左にある事を意味しているが、AとBとの距離や方向については表現していない。図1.2の(1)と(2)において、対象A、Bは同じ位置にある。しかし、(1)においては「AはBの上にある」といえるが、(2)では上にあるとはいえない。状況による位置関係の相違や、位置関係の持つ漠然さを、left(A, B)のような記号表現だけで表現することは難しい。また、対象間の距離や方向などの特徴量に対するしきい値で、関係の成立、不成立を決定する方法は、曖昧性、状況依存性を含んだ定性的な位置関係に対しては、不都合な場合が少なくない。

人間の使用する定性的な関係を表現する問題は、パターン情報から人間の認知する内容を記号で表現する際や、逆に文章等の記号で表現された内容をパターン情報に変換する際に生じる問題で、人間と計算機の高度なインタフェースを実現するのに必要な研究課題の一つである。

画像が表現しているパターン情報から画像のもつ構造等を記号表現する。その記号表現を用いて記号とパターン情報を統合して扱う研究が幾つかなされている²⁵⁾。ここでは、物体の構造を記述したり、記述から概念を学習するのに位置関係を用いたシステムや対象間の位置関係をパターン情報である画像情報と記号表現である言語情報の両面から扱った研究を概観する。

Winstonは、積木の世界の線画を対象に、構造概念学習を行うARCHシステムを作成した。ARCHシステムでは、学習させたい概念の例を線画から解析し、対象の構成要素とその間の関連を意味ネットワークで表現する。概念の例を構成する要素間の関連に、left-of, right-of, in-front-of, aboveのような位置関係を用いた²⁴⁾。

安部らは、コンピュータグラフィックスで作成されたアニメーション画像を、自然言語による説明に基づき理解するシステムを作成した。その中では、対象の領域を9領域(左右方向を右, 中央, 左, 上下方向を上, 中央, 下の3分割)に、

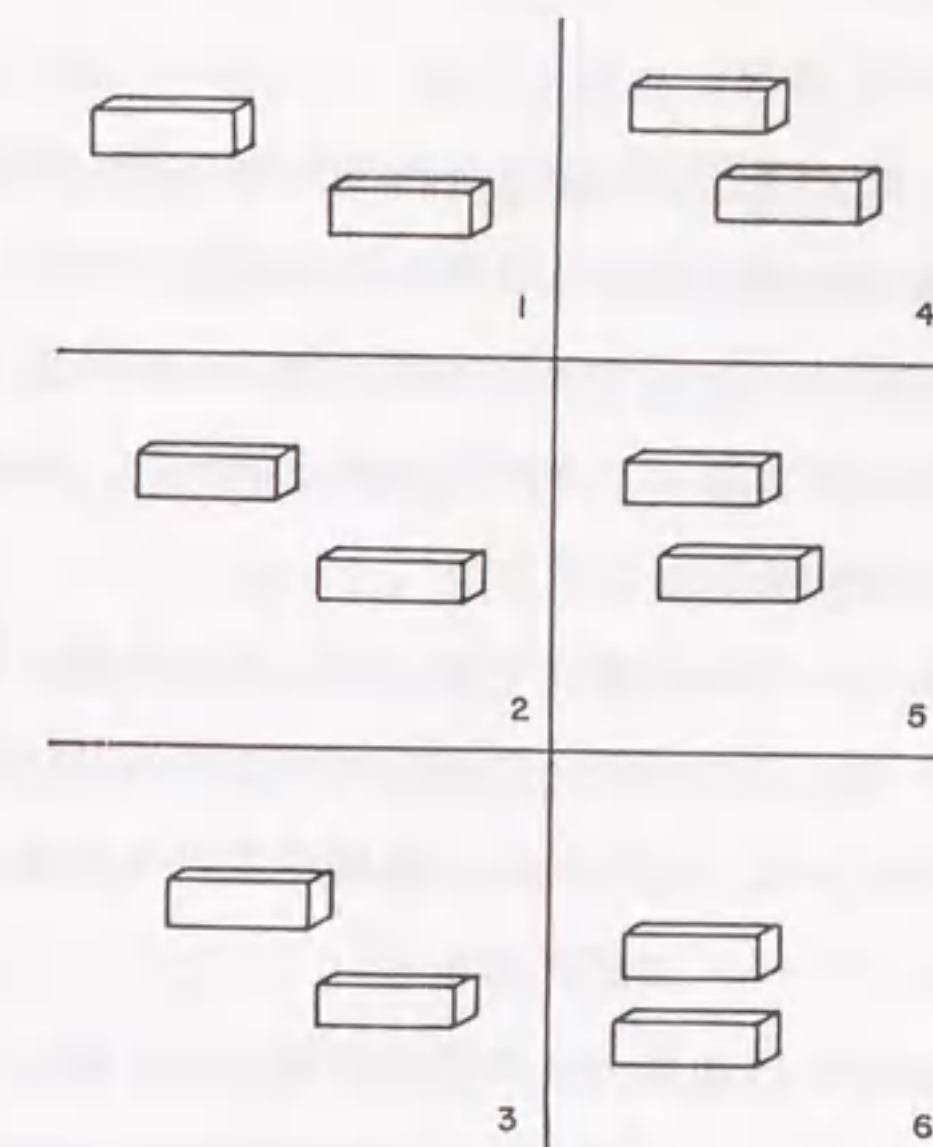


図 1.1 位置関係の持つ漠然性 (“コンピュータビジョンの心理”²⁴⁾ p.186 より)

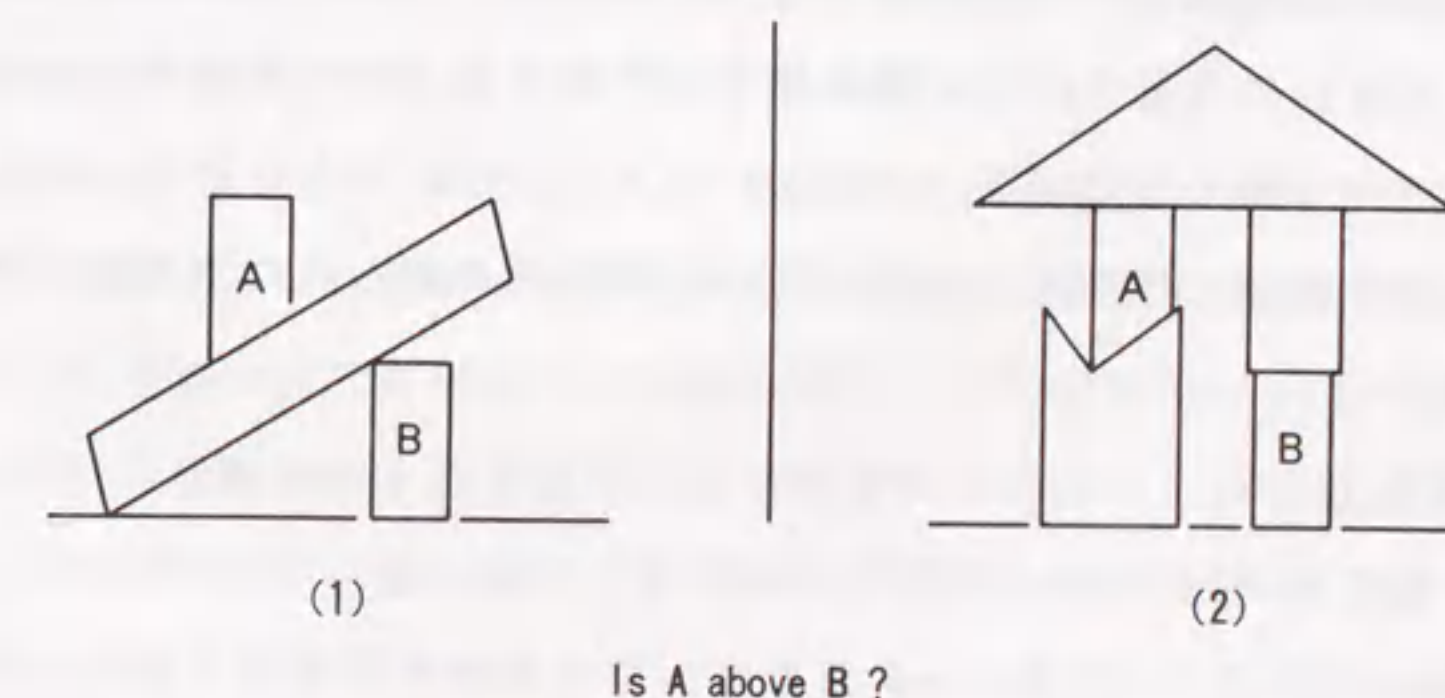


図 1.2 位置関係の状況依存性
(“The Modelling of Spatial Relations”²³⁾ p.161 より)

部品領域相互の関係を4種類の関係(内, 外, 接して内, 接して外)で表現して画像と言語の対応づけを行った^{26), 27)}。

岡田らは, 図形認識に加え, 要素図形から構成される図形の意味解釈と自然語記述, 位置関係に着目した動画像の意味解釈の研究を行った^{28), 29)}。高木らは画像の持つ構造などの情報を, 定型的な処理で自然言語と共通した特徴テーブルを用いて表現し, 画像データから自然言語の文書を作成する研究と, 画像情報と言語情報を突き合わせて理解する研究を行った^{30), 31)}。

これらの研究は, 図形画をもとに文章を作成する, 図形画を言葉で検索するためのインデックスを自動抽出するなどの点で, 先駆的研究であるが, 扱われた位置関係は, 幾何的な関係で漠然さを扱っていない。位置関係のもつ漠然さを扱った研究に, Haar³²⁾と山田ら^{33), 34)}の研究がある。

Haarは, 距離と方向に関する信頼度付きの基本述語-DISTANCE(A, B), BEARING(A, B)-を導入し, 位置関係のもつ曖昧さをファジィ集合で表現した。そして, 「上, 下, 左, 右」といった定性的な位置関係を, 距離と方向に幅をもたせることにより, DISTANCE(A, B)とBEARING(A, B)を用いた定量的な表現に結びつける方法について述べた。

山田らは, 自然言語で表現された情景を対象間の位置の依存関係を利用して解釈し, 解釈した情景の幾何モデルを再現するSPRINTシステムを検討している。ここでは, 定性的に表現された位置関係の漠然さを明示的に表現するためにポテンシャルモデルを導入している。ポテンシャルを最小にすることで, Haarらの推論する存在範囲ではなく, 対象位置の依存関係を最もみたと対象の存在位置を推論する。

1. 3. 3 手振りを利用したインタフェース

人間と計算機との間のインタフェースとして, キーボードや音声を用いた会話形式の利用形態から, 計算機が作り出す人工世界で人間が操作するといった「人工現実感」(Virtual Reality)といった環境まで考えられている^{35), 36)}。そのような人工世界におけるインタフェース環境では, 言葉のほかに物をつかむといった操作者の動作も人間から環境への働きかけ(計算機への入力方法)となる。

体の各部位の3次元的な動きを測定するための技術や装置が開発されている。

そのような装置の一つに, Data GloveTMがある³⁷⁾。Data GloveTMは, 光ファイバセンサを指に沿ってはわせ, 指の曲げによる光ファイバの透過効率の変化から各指の各関節角を測定する手形状入力装置である。

手振りを用いた計算機インタフェースの先駆的な例としてPUT THAT THERE⁵⁾がある。このシステムでは, 3次元位置と方向を測定する3次元磁気センサを用いてディスプレイ上の位置を指定する。そして, 音声による“copy the red triangle there”といった指示と, 手による位置指示を用い, 図形を複写したり, 移動したりする。PUT THAT THEREは, “there”という特定の言葉をもとに, 手による指示対象を認識するといった制約にもかかわらず, 言葉と手振りという複数の入力方法を併用したマルチモーダルインタフェースが, 計算機インタフェースとして有効な事を示した。

手振り, 身振りによるコミュニケーションの代表例として手話があげられる。手話には手振りで表現される単語が3000語近くあり, その中には複数の話言葉の単語に相当する手振りもある³⁸⁾。田村らの手話動画像認識システムは, 連続した画像から画像処理技術によって, 右手の形, 位置と動きに関する特徴量を抽出する。抽出結果を, 手の形, 動き, 位置を記述した手話辞書と照合することで, 「父」, 「母」等の10個の手振りを認識した³⁹⁾。手話の46個の指文字の手の形から41個を, 手の形の輪郭などの特徴を用いた画像処理で, 識別した報告もある^{40), 41)}。

手振りを用いたインタフェースの実現には, 手振りの表現方法と, 手の形や動きをリアルタイムに認識する方法が必要になる。手振りのコード化方法としては, アメリカ手話における手振りの記述に用いられているStokoeらの構動素⁴²⁾や, 藤野らの手話動画画像から手話単語を認識する時に用いられる機械辞書の検討⁴³⁾, 黒川らの視覚的な特徴に基づく記述法等が提案されている⁴⁴⁾。

Sturmanは, 手や指が静止している姿勢(posture)と, 手や指の動きを伴う手振り(gesture)の組み合わせから, 手の動きをbutton, valuator, 3D locator等の用途に分類している⁴⁵⁾。手形状入力装置(Data GloveTM)を用いたシステムとしては, 手の動きに対象タスク固有の意味を定義し, 曲面形状を入力するシステムをはじめ, いくつかの研究が発表されている^{38), 46)}。

視覚的な特徴に基づく手形状のコード化方法では, 手形状入力装置の測定デー

タからは抽出する事が難しい特徴が用いられている。手形状入力装置を用いたシステムでは、指の曲げ角のとりうる範囲を規定する事で、手の形を表現する方法がよく使用されている。この方法には、対象タスク固有の手形状の表現になる、複数の手の形を表現するには曲げ角の範囲の重なりが生じる、時間的に変化する指の動きを表現する事が難しい等の問題がある^{47), 48)}。

1. 4 研究の概要と本論文の構成

1. 3において、視覚的な特徴を定性的な位置関係で表現する研究や、手振りをとり入れたインタフェースに関する研究について概観した。人間と計算機間のインタフェースをより人間にとって使いやすくするには、複数の方法で入力される定量的、定性的なデータをもとに、その指示内容を解釈することが必要になる。本論文では、人間に対する指示と同じように、位置関係に関する言葉と身振りをを用い、計算機に指示できるインタフェースを研究する。

第2章では、自然言語で用いられる定性的な位置関係の持つ漠然さと、位置関係の表現方法の問題について述べる。次に、対象の近似多角形の頂点座標をもとにした、幾何学的な関係から自然言語で用いられている言葉までの位置関係の階層表現について述べる。

第3章では、ルールで表現された位置関係を用いた画像データベースの内容検索方法について述べる。対象の近似多角形の頂点座標を画像インデックス情報とする画像データベースを、書誌情報のほかに位置関係に関して、言葉で問い合わせ、検索する方法を提案する。風景画を中心とする32枚の浮世絵画像データベースと、66件の科学論文の第1頁の文書画像データベースを検索した結果について述べる。

第4章では、インタフェースにおける「手振り」を「話し手の意図を伝える手の形・動き」と定義し、手形状入力装置により測定される手話に指文字の指の曲げ、手の向きデータをもとに主成分分析、クラスタ分析した結果について述べる。さらに指の曲げ、手の向きに着目した手の形のコード化方法を提案する。

第5章では、4章の手形状コード化方法で測定された手の形状コードと、手の向きと3次元位置の変化に基づく動きコードを用いたリアルタイムな手振り認識方法を提案する。手話の指文字を手振りの例として、動きのある手振りの認識実

験結果について述べる。

第6章では、定性的な位置関係を含む言葉による指示と、手振りによる指示という複数の方法を用いたマルチモーダルインタフェースをとりあげ、複数の指示方法の統一表現を提案する。そして、それにもとづき試作した地図案内システムの概要と評価について述べる。

第7章では、研究のまとめと残された課題について述べる。

第2章 位置関係の表現

本章では、位置関係を表現することを目的とする。位置関係に関する定性的な言葉をインタフェースに用いるためには、言葉の持つ漠然さや状況依存性を扱う必要がある。最初に、位置関係の定性的な特徴について述べる。次に、画像処理技術などによる特徴抽出を前提とした、対象の近似多角形の頂点座標を用いた対象間の位置関係の漠然さの表現と、ルール表現を用いた頂点座標と自然言語との対応について述べる。

2.1 位置関係の特徴

一般に、辞書では「右」を「南を向いた時、西にあたる方（広辞苑）」と他の位置関係を用いて説明してある。計算機内では、位置関係を対象の持つ基本的な量・関係を用いて表現する必要があり、人間の常識を計算機で定式する時の一つの問題になっている^{49), 50)}。

位置関係の解釈には、以下の項目が問題になる。

(1) 定量的な画像情報と定性的な言語情報の対応である位置関係には、漠然性が伴う。

1. AがBの右にある場合、以下の関係を満たしている。

a) Bの中心からみて右半面にAがある。

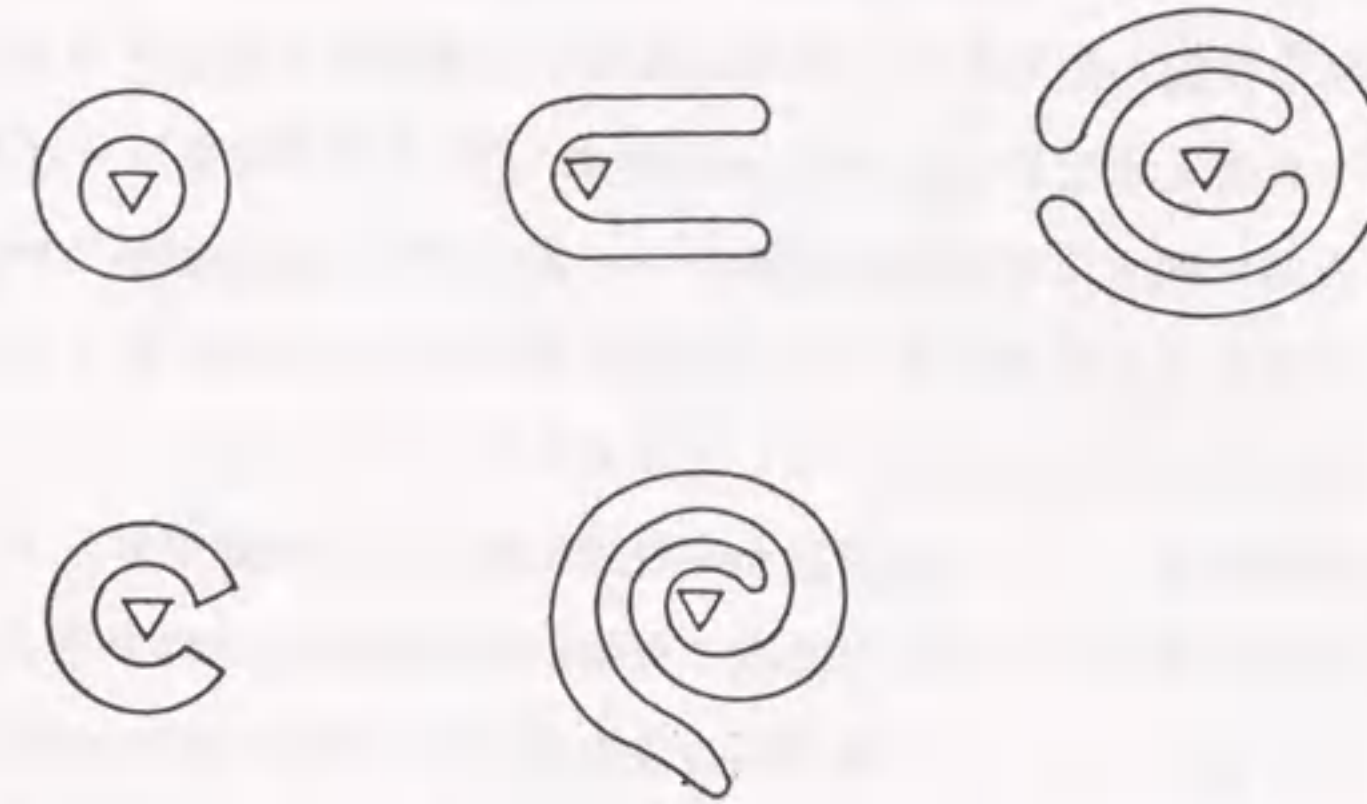
b) AとBとは、お互いに離れている。

条件a), b)を満たしていても「右」より「上」といった他の位置関係で参照される時もある²⁴⁾。(図1. 1)

2. 包含関係については、凹のある図形に対しては幾何学的（鉛直線算法：多角形と1点pが与えられた時、pを始点とし鉛直下方に伸びる半直線と多角形の辺との交点の数が偶数ならpは外部にある⁵¹⁾。）に外側にあっても心理的には内側と参照される時もある²³⁾。(図2. 1)

(2) 自然言語と位置関係は、一対一対応ではない。

1. 同じ「右」という言葉にも、「Aが右にある。」と画像全体における対象の空間的位置を意味する時と、「AがBの右にある。」と対象間の相対的位置関係を意味する時がある^{26), 52), 53)}。



Is one object inside the other(s)?

図2.1 対象間の包含関係
(“The Modelling of Spatial Relations”²³⁾ p.158 より)

2. 日本語の「上」は英語の「on, over, above」に対応する。自然言語により表現する範囲が異なるため、幾何的に同じ位置関係であっても、状況により対応する言葉が異なる。

3. 「左」と「上」のように異なり排他的に思える位置関係でも、図形によってはどちらの位置関係が適当であるか識別しにくい。図1. 1の例では、1から6に移るにつれ、2つの対象間の位置関係は「左」から「上」への度合が強くなる。中間の3, 4では、「左」と「上」の両方の関係が成立していると解釈できる。

(3) 対象間の位置関係は、対象の位置だけではなく、まわりの状況に依存して解釈される。

1. 観測者の定める向き、画像全体の定める向き、対象自体の定める向きにより同じ位置にある対象に対しても、参照される位置関係は異なる^{29), 50)}。

2. 2つの対象の位置が同じであっても、他の対象の位置により参照方法が異なる事もある。(図1. 2)

3. 2次元世界における「下」という位置関係は、地図画像に対しては「南」になったり、3次元世界が投影された世界では「手前」と参照されるように、

言語表現は画像の表現世界に依存している。

1. 3. 2で、画像情報から対象の位置や対象間の位置関係を表現する方法として、画像の特徴を利用した2D-string, point-setによる表現、ファジイ集合による位置関係の表現に関する研究例を概観した。2D-string, point-setによる表現では上記の特徴のうち位置関係の持つ漠然さを扱っていない。ファジイ集合による表現では状況の依存性を表現することができない。

本研究では、画像処理などによる図形画から対象の特徴抽出を前提に、2次元図形画に含まれる対象の空間的位置、対象間の相対的位置関係の表現を考える。特徴抽出された対象の頂点座標などの物理的な量を用いて、幾何学的な条件で規定される対象間の位置関係とその漠然さを評価する。そして、その位置関係をもとに、自然言語との対応を持つ定性的な位置関係を、ルールで表現する。

2. 2 位置関係の表現

位置に関する量・関係には、特徴抽出された対象の頂点座標から計算できる対象の中心位置、対象間の距離、角度や包含関係等の物理的な量・関係と、人間が日常用いる上下左右といった定性的な位置関係がある。幾何学的な関係と異なり、定性的な位置関係は曖昧さをもっている。if-then型のルールによる知識表現は、このような人間の断片的な知識を表現しやすく、更新や拡張が容易であるといった特徴をもっている⁵⁴⁾。

2. 2. 1 漠然さの表現

従来、不確実性はベイズ規則に従う確率量で扱われていた。事象Aの生起確率を $p(A)$ とすると、ベイズの規則は加法の規則、

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

$$A \cap B = \phi \text{ なら } p(A \cup B) = p(A) + p(B)$$

を満たし、AのもとでBの条件付き確率を、

$$p(B|A) = \frac{p(B \cap A)}{p(A)} \quad (\text{但し, } p(A) > 0)$$

で与える。一方、人間は、曖昧な状況下でも判断し行動している。曖昧な状況下

での推論が、人工知能の基礎やエキスパートシステム等の応用システムにおいて研究されている。人間の判断をモデル化しようとする知識工学においては、生起確率の他の「曖昧さ」を以下のように分類している^{54), 55)}。

- (1) 制御の非決定性 (non-determinism)
- (2) 多義性 (multiple meaning)
- (3) 不確実性 (uncertainty)
- (4) 不完全性 (incompleteness)
- (5) ファジィ性 (fuzziness, imprecision)

ベイズ規則で扱えない不確実性やファジィ性に対して、エキスパートシステムMYCINのCF (Certainty Factor: 確信度)、Dempster-Shafer理論、ファジィ集合などの方法が用いられる^{56), 57), 58)}。

(1) MYCINのCF

MYCINは、医療の分野における細菌感染症の診断と治療方法の提案を行うエキスパートシステムである。MYCINは、人間の知識をエキスパートシステムで表現する際に現れる不確実さを表現するには、確率量だけでは不十分であるとの見地に立ち、事実やルールを1 (完全な肯定) から-1 (完全な否定) の確信度つきで表現した。証拠となる事実(X, Y)から、結論の事実(A)を導く時の部分問題間にAND, OR関係以外に、複数の証拠により帰結となる事実が独立に支持される結合(Combination)関係を定義した。ルールの前提条件が満足された時、導かれる結論部には以下の式に従って確信度を与える。

AND関係の時:

$$\begin{aligned} CF_{\text{premise}} &= CF[X \& Y] \\ &= \min \{CF[X], CF[Y]\} \end{aligned}$$

OR関係の時:

$$\begin{aligned} CF_{\text{premise}} &= CF[X \text{ or } Y] \\ &= \max \{CF[X], CF[Y]\} \end{aligned}$$

で前提部の CF_{premise} が計算され、ルール自体に付与されている確信度 CF_{rule} とあわせて結論部のAの確信度、

$$CF[A] = CF_{\text{rule}} * CF_{\text{premise}}$$

を計算する。

それぞれ証拠となる事実 (X, Y) から、結論の事実 (A) を導くルールにより確信度 $CF[A]_x$, $CF[A]_y$ で A が推論される時、Combination 関係で導かれる結論に対する確信度 $CF[A]$ を、次の結合関数で計算する。

$$CF[A] = \begin{cases} 1 & \text{if } CF[A]_x = 1 \text{ or } CF[A]_y = 1 \\ CF[A]_x + CF[A]_y - CF[A]_x * CF[A]_y & \text{if } CF[A]_x > 0 \text{ and } CF[A]_y > 0 \\ CF[A]_x + CF[A]_y & \text{if } CF[A]_x * CF[A]_y \leq 0, \\ & CF[A]_x \neq \pm 1 \text{ and } CF[A]_y \neq \pm 1 \\ CF[A]_x + CF[A]_y + CF[A]_x * CF[A]_y & \text{if } CF[A]_x < 0 \text{ and } CF[A]_y < 0 \\ -1 & \text{if } CF[A]_x = -1 \text{ or } CF[A]_y = -1 \end{cases}$$

この結合関数は、合理的な基礎を持っていないがわかり易いので多くのエキスパートシステムで用いられている。

(2) Dempster-Shafer 理論

ベイズの規則では、その加法性の規則から $p(A) + p(\text{not} A) = 1$ である。従って、ベイズの規則では、A でも not A でもない状態—事実を知らない事—を表現できない。事実を知らない事から起こる不確実さを表現するために、 $A = \{a_1, a_2, \dots, a_n\}$ の部分集合 $A_i = \{a_j, \dots, a_k\}$ ($i = 0, 1, \dots$) に対する基本確率を用いて、 A_i の確率の最小と最大を表現する下界確率 $P_*(A_i)$ 、上界確率 $P^*(A_i)$ を定義する。基本確率 $m(A_i)$ は $[0, 1]$ の値をとり、 A_i の部分集合内には存在するが A_i のどの部分集合にいるかは確定できない量で、以下の条件を満たす。

$$\begin{aligned} m(\phi) &= 0 & (\phi: \text{空集合}) \\ \sum_{A_i \subseteq A_0} m(A_i) &= 1 & (i=0, 1, 2, \dots) \end{aligned}$$

事実を知らない程度は、全集合 A_0 のどこに存在するのか定められない基本確率 $m(A_0)$ によって表現される。

下界確率 $P_*(A_i)$ は、基本確率で

$$P_*(A_i) = \sum_{A_j \subseteq A_i} m(A_j)$$

と、部分集合 A_i に必ずある確率として定義される。上界確率 $P^*(A)$ は

$$P^*(A_i) = 1 - P_*(\bar{A}_i)$$

と、部分集合 A_i に少しでもある確率として定義される。

独立な証拠から推論される仮説に対する基本確率が m_1, m_2 で、対応する部分集合が A_{1i}, A_{2j} ($i, j = 0, 1, 2, \dots$) の時、それらを結合した基本確率は以下の式で与えられる。

$$m(A_k) = \frac{\sum_{A_{1i} \cap A_{2j} = A_k} m_1(A_{1i}) * m_2(A_{2j})}{1 - \sum_{A_{1i} \cap A_{2j} = \phi} m_1(A_{1i}) * m_2(A_{2j})} \quad (A_k \neq \phi)$$

(3) ファジィ集合

「old」といっても、40才以上なのか80才以上なのか明確ではない。1965年に Zadeh によって、集合の境界を明確にしないままで、集合を扱うファジィ集合論が提唱された。

対象領域を覆う集合 $U = \{u_1, u_2, \dots, u_n\}$ のファジィ部分集合 F は、メンバーシップ関数 $\mu_F(u)$ で規定される。メンバーシップ関数 $\mu_F(u)$ は、U の要素 u のファジィ集合 F への所属度を $[0, 1]$ の数値で表す。

$$\begin{aligned} F &= \mu_F(u_1) / u_1 + \mu_F(u_2) / u_2 + \dots + \mu_F(u_n) / u_n \\ &= \sum_{i=1}^n \mu_F(u_i) / u_i \end{aligned}$$

ここでの + は、算術の和ではなく集合要素 u_i のメンバーシップ関数付きの和を表現している。ファジィ集合 F は集合であるので、F に対する補集合、和集合、積集合が以下のように定義できる。

$$\bar{F} = \sum_{i=1}^n (1 - \mu_F(u_i)) / u_i$$

$$F \cup G = \sum_{i=1}^n \max(\mu_F(u_i), \mu_G(u_i)) / u_i$$

$$F \cap G = \sum_{i=1}^n \min(\mu_F(u_i), \mu_G(u_i)) / u_i$$

2. 2. 2 位置関係の漠然さに関する主観実験

位置関係の漠然さを決定する要因としては、対象の位置、形、大きさのほかに対象が表示されているディスプレイの形状、配置の安定性、対称性などがある。このうち対象の位置に注目して、相対的な位置関係「右」の持つ漠然さに関して主観実験を行った⁵⁹⁾。

(実験方法)

4名の被験者(A, B, C=男性, D=女性)に対し、図2. 2で示すような3点(原点o, X軸上の基準点a及び比較点b)をディスプレイに表示する。被験者はマウスで比較点bを傾き θ の直線上で、原点oに対する比較点bの「右」の度合が基準点aの「右」の度合と同じになるように動かす。同じ基準点aに対し、複数の比較点bの(x, y)座標を測定する。原点oから基準点aまでの距離Rとして0.15, 0.3, 0.5, 0.7, 0.85の5種類、X軸と直線obの角度 θ として3, 5, 7, 10, 15, 30, 45, 50, 60, 70, 80, 90度の12個の直線上の比較点の(x, y)座標を各々10回測定した。

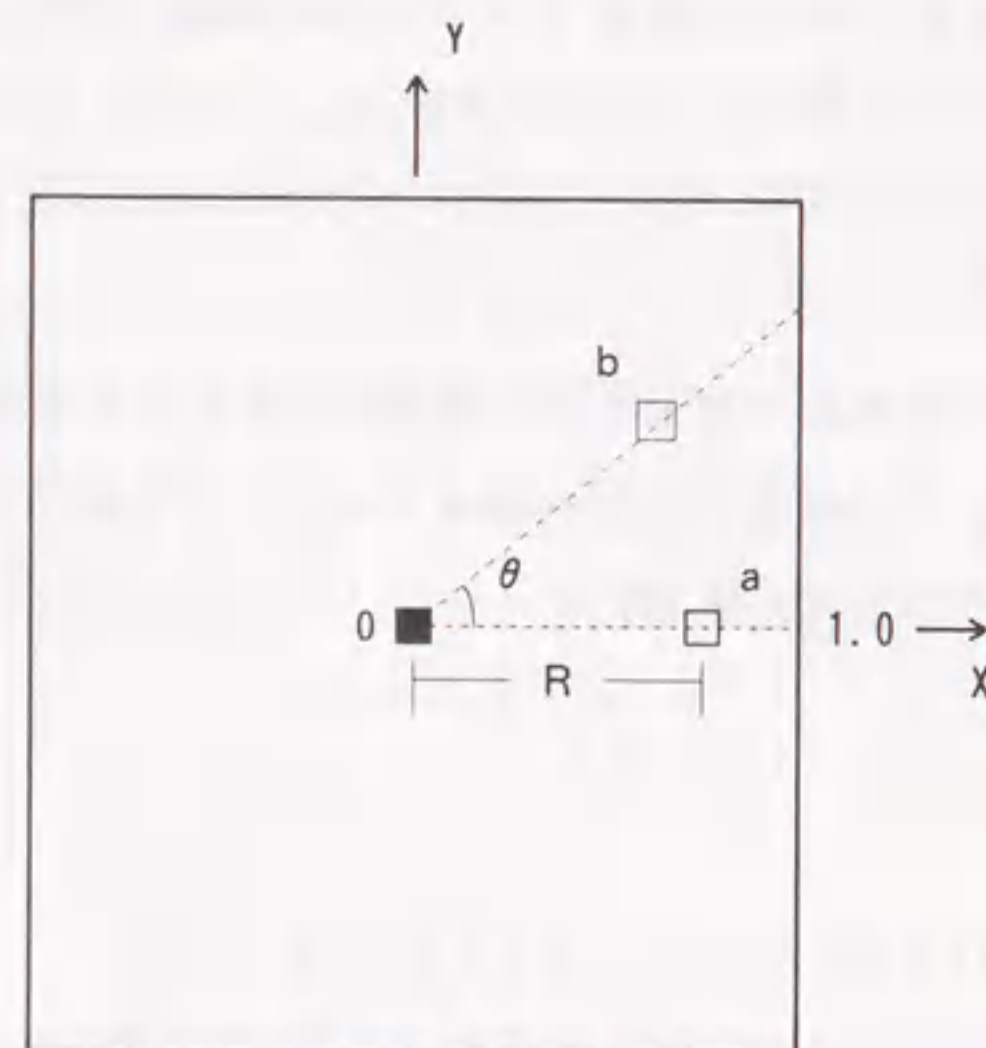


図2.2 主観評価実験のディスプレイ画面

(実験結果)

各(R, θ)に対する各被験者の比較点bの(x, y)座標の平均値を図2. 3に示す。図2. 3において同じRに対する点を結ぶことで、X軸上の基準点aと「右」の度合が同じである等高線を表現することができる。この等高線の曲線近似により、任意の点の「右」の度合がX軸上のどの点と等しいか決めることができる。図2. 3から、関数として、xの値が小さい時は線形に増加し、その後、指数関数的になだらかに減少していることから、等高線を

$$y = x * (a * e^{(-b * x)} - c)$$

で近似した。係数a, b, cを各Rについて求めた結果を表2. 1に示す。

図2. 4に、各(R, θ)に対する比較点bの平均座標を中心とし、偏差の半分を半径とする円で被験者Dの「右」の度合のばらつきを示す。これらの結果から、以下のことがいえる。

1. 位置関係「右」の持つ漠然さには個人差がある。
2. 位置関係の漠然さをRと θ の関数で絶対的な数値としては表現できないものの相対的な尺度として利用できる。
3. 対象の中心を結ぶ直線が水平な程、標準偏差が少ないので信頼性が高い。

この結果は図1. 1の例で、1から6に移るに従い「右」の度合が減少し、「上」の度合が増加することとも一致する。

2. 2. 3 ルールを用いた位置関係の表現

岡田ら²⁹⁾によれば、文献⁶⁰⁾には約500の日本語の静的な位置関係に関する名詞が登録されている。それらは、基本的には約30の位置関係に分類でき、さらに以下のように細分類される。

- 2項関係: A) 向きに関するもの(上, 左など)
 B) 領域に関するもの(内, そばなど)
 C) 複合に関するもの(交差, 接触)
 D) 部分に関するもの(端, 奥など)

- 3項関係: (間)

国語辞典(広辞苑)によるこれら自然言語の関係は、例えば「上」の意味は「物の上部, 高い位置」として他の位置関係で説明されている。さらに、「高い」

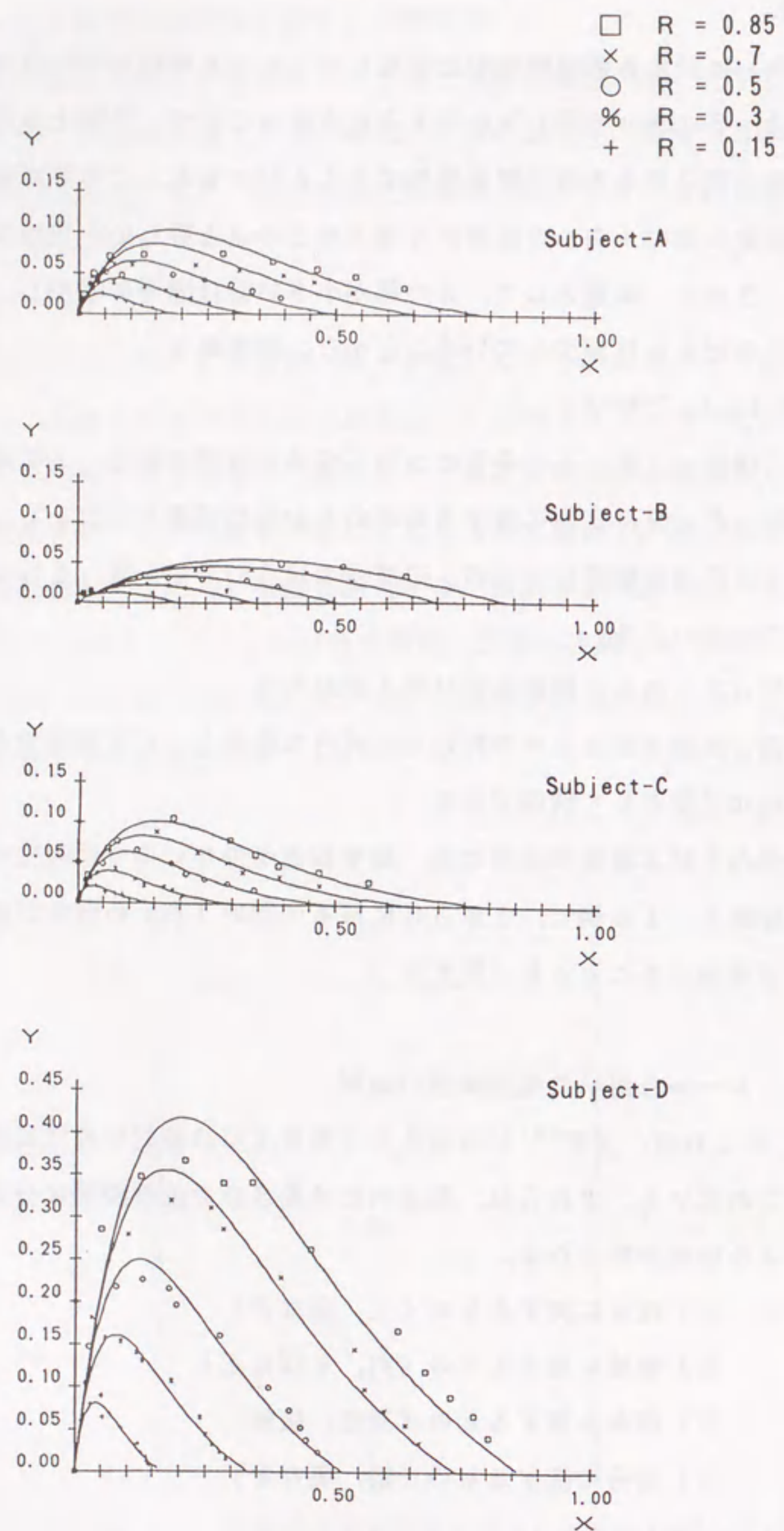


図2.3 各被験者の「右」の度合いの等高線

表2.1 各被験者の係数 a , b , c

R	被験者 A			被験者 B			被験者 C			被験者 D		
	a	b	c	a	b	c	a	b	c	a	b	c
0.15	2.6	31.1	0.01	1.19	26.6	0.03	2.15	35.8	0.00	6.5	27.1	0.04
0.3	1.9	14.9	0.00	0.69	9.2	0.02	1.93	16.9	0.01	5.9	12.8	0.06
0.5	1.5	8.9	0.00	0.52	3.7	0.08	1.79	10.3	0.01	5.5	8.2	0.03
0.7	1.4	6.7	0.00	0.49	2.2	0.11	1.76	7.8	0.00	5.0	5.2	0.09
0.85	1.4	5.5	0.00	0.46	1.7	0.10	1.76	6.4	0.01	5.5	5.1	0.01

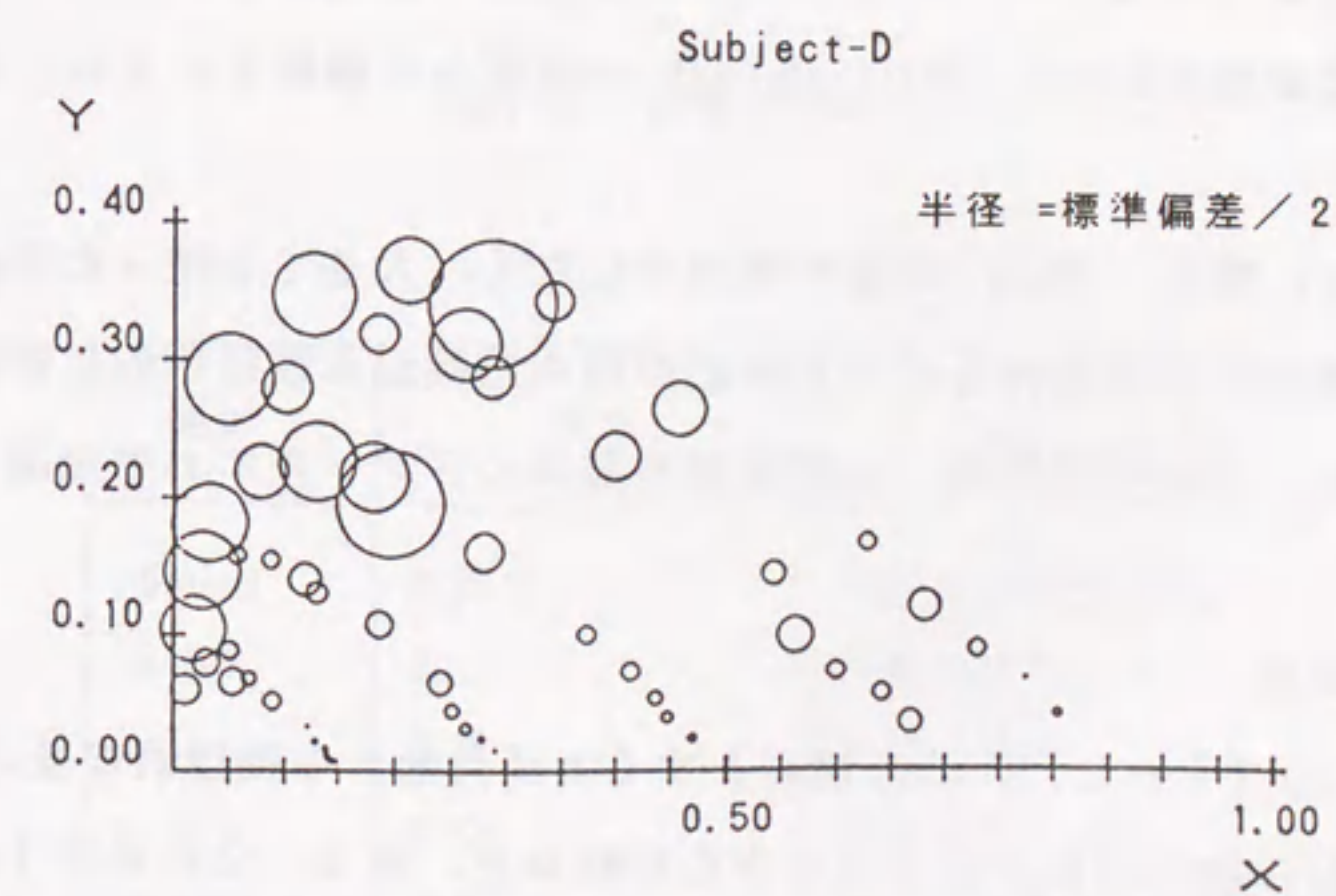


図2.4 被験者Dの「右」の度合いの分散

は「空間的な位置が上方にあって、下との距離が大きい」と「上」の位置関係で説明されている。このように、辞書では、言葉の説明にお互いが使用され、文法に従って、文が単語から構成されるような構造を持っていない。また、日本語の「上」の意味は、英語の「on, over, above」の意味を含んでいるように、同じ幾何的な位置関係に対しても、その関係を表現する自然言語間の位置関係の概念は一対一に対応しない。

一方、2次元図形画の世界は、対象の頂点や辺を基本要素として規定できる。頂点座標値から対象の中心位置、対象間の距離、角度や包含関係等の幾何的な位置関係が計算でき、CADやコンピュータグラフィックスでは頂点や辺の情報を基本情報として、物体のデータは表現されている。

図形画における対象の頂点座標といった幾何的な位置情報から、自然言語の定性的な位置関係までを、図2.5に示すように

1. 自然言語から独立した定性的な基本位置関係を定義する。
2. 基本位置関係を幾何的な位置関係、量で定義する。基本位置関係の持つ漠然さを主観実験にもとづいたヒューリスティックな関数で評価する。
3. 位置関係を基本位置関係のAnd, OrとCombination関係から構成されるルールで表現する。

ことにより結び付ける。図2.5は、対象を質点ではなく、大きさを持った近似多角形で表現し、対象間の位置関係を近似多角形の頂点座標から幾何学的に計算できる物理的量・関係、基本位置関係、位置関係の言葉の3レベルにわけ階層表現した例である。

(1) 物理的な量・関係

対象 $a((x_i, y_i) (i = 1, \dots, n))$ を頂点とする n 多角形の物理的な量として、対象の頂点座標から幾何的に計算される中心座標など、表2.2に示す10個の物理量がある。対象間の物理的な量としては、表2.3に示す各種の対象間の距離と、方位角の物理量がある。対象間の関係としては、完全に含まれている、重なりがある、接している、離れている4種類の包含関係がある(表2.4)。

例：対象の中心座標

$$\text{center}(a) = (\sum_{i=1}^n x_i / n, \sum_{i=1}^n y_i / n)$$

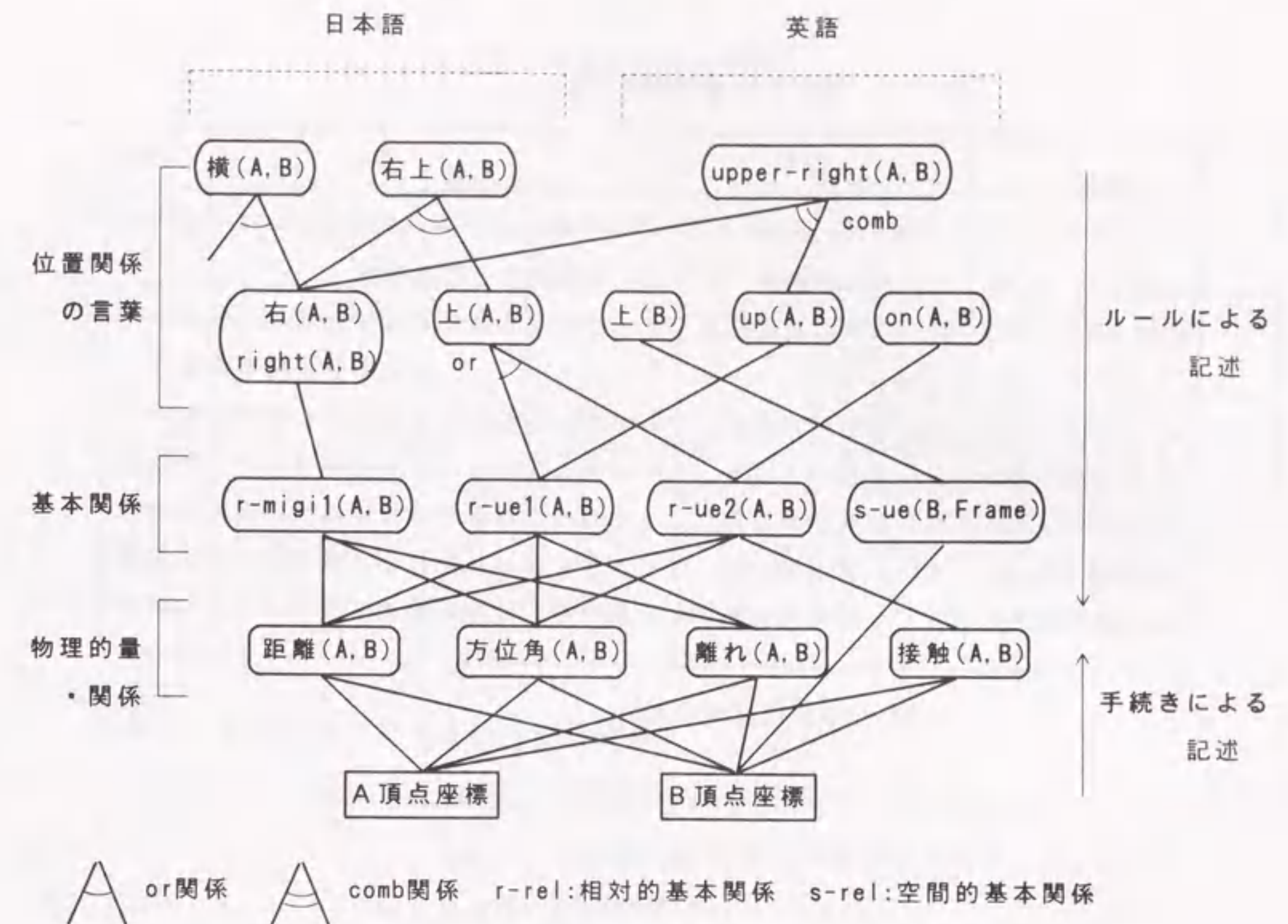


図2.5 位置関係におけるレベル

表2.2 対象のもつ物理量

関係	意味	定義
pbound	境界点	対象の頂点座標のリスト
bound	辺	対象の辺のリスト
area	面積	対象の面積
center	中心	対象の重心座標
radius	半径	重心から最も遠い頂点までの距離
left-end	左端の頂点座標	X座標値の最も小さな頂点座標
right-end	右端の頂点座標	X座標値の最も大きな頂点座標
top	上端の頂点座標	Y座標値の最も大きな頂点座標
bottom	下端の頂点座標	Y座標値の最も小さな頂点座標
lbp	縦横比	上下端距離と左右端距離の比

※ 円も n 多角形で近似する。

表2.3 対象間の物理量

関係	意味	定義
distance1(A, B)	A、B間の距離	A、B間の中心間の距離
distance2(A, B)	A、B間の最短距離	A、Bの中心を結んだ線分とA、Bの辺との交点間の距離
azimuth1(A, B)	A、B間の方位角	Aの中心からみたBの中心との方位角
azimuth2(A, B)	A、B間の方位角	Aの中心からみたBの右端頂点との方位角
azimuth3(A, B)	A、B間の方位角	Aの中心からみたBの左端頂点との方位角
azimuth4(A, B)	A、B間の方位角	Aの中心からみたBの上端頂点との方位角
azimuth5(A, B)	A、B間の方位角	Aの中心からみたBの下端頂点との方位角

※ azimuth1~5の単位はラジアン。

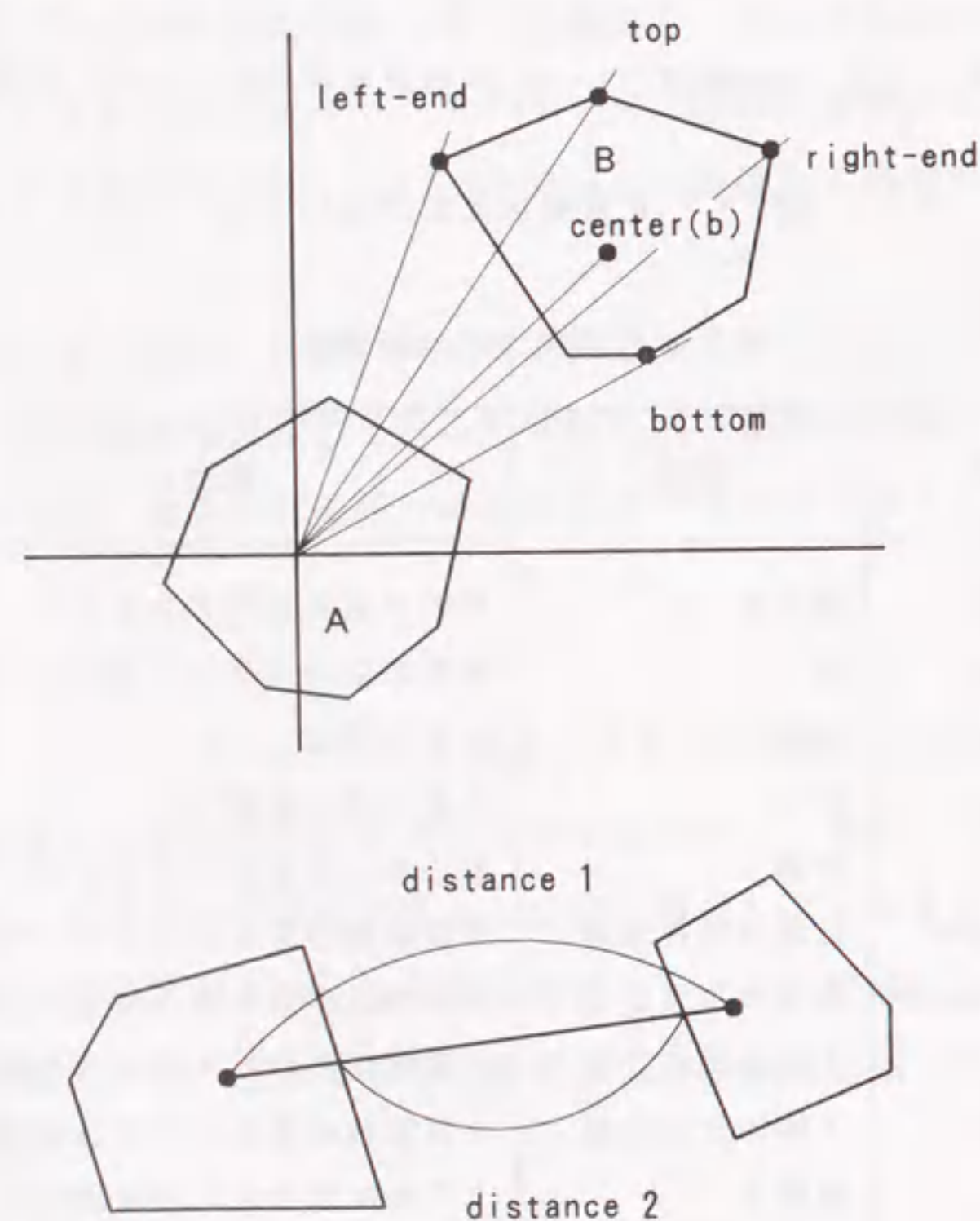


表2.4 対象間の包含関係

関係	意味	定義	
crossing (A, B)	AはBと交差している	辺同士で交差する。	Aのある頂点がBの内部にある。
touching (A, B)	AはBと接している		Aのある頂点がBの辺上にある。 又は、 Bのある頂点がAの辺上にある。
within (A, B)	AはBの中にある	辺同士で交差しない。	Aの面積がBの面積より小さい。 Aの中心がBの内部にある。
remote (A, B)	AはBと離れている		その他。

対象 a, b が離れているという包含関係；

$\text{remote}(a, b) = \text{False}$ (多角形の辺が交差している, または
一方の中心が他方に含まれている。)
 $= \text{True}$ (それ以外。)

(2) 基本位置関係

画像に含まれる対象の位置関係は, 対象の近似多角形に対しても成立する。多角形の頂点座標で計算機に入力される画像データを, 定性的な位置関係で表現するには, 位置関係の持つ漠然さを頂点座標をもとに定量化する必要がある。自然言語から独立した定性的な位置関係の幾何学的な成立条件を, 物理的な量・位置関係を用いて, ルールの形で表現する。さらに, 位置関係の持つ漠然さを, 物理的な量を変数とするヒューリスティックな関数で評価する。

この基本位置関係には, 対象の空間的な位置関係と対象間の相対的な位置関係がある。空間的な位置関係には上下左右の位置関係, 相対的な位置関係には上下左右のほか遠近など表2.5に示す関係がある。

関係が成立する(肯定された)時の関係の持つ漠然さの度合を, 2.2.2の主観実験から得られた性質を反映した関数形をもつヒューリスティックな関数で表現する。関数は, $[0, 1]$ の値をとり, 関数値が1に近いほど, 関係が成立している度合が高いことを示す。

表2.5 対象間の基本位置関係 (1/2)

位置関係	意味	定義	ヒューリスティックな関係の評価関数 (CF)
r-ue1(A, B)	AはBの上にある	$\pi/4 \leq \theta_1 \leq 3\pi/4$ and $\text{center}(A)_x \geq \text{center}(B)_x$	$\exp(-\text{distance1}(A, B)^2)$
r-ue2(A, B)	AはBの真上にある	$\pi/3 \leq \theta_1, \theta_2, \theta_3, \theta_4, \theta_5 \leq 2\pi/3$ and $\text{center}(A)_x \geq \text{center}(B)_x$	
r-ue3(A, B)	AはBの上で接する	$\pi/4 \leq \theta_1 \leq 3\pi/4$ and $\text{touching}(A, B)$	
r-shita1(A, B)	AはBの下にある	$-3\pi/4 \leq \theta_1 \leq -\pi/4$ and $\text{center}(A)_x \leq \text{center}(B)_x$	$\exp(-\text{distance1}(A, B)^2)$
r-shita2(A, B)	AはBの真下にある	$-2\pi/3 \leq \theta_1, \theta_2, \theta_3, \theta_4, \theta_5 \leq -\pi/3$ and $\text{center}(A)_x \leq \text{center}(B)_x$	
r-shita3(A, B)	AはBの下で接する	$-3\pi/4 \leq \theta_1 \leq -\pi/4$ and $\text{touching}(A, B)$	
r-migi1(A, B)	AはBの右にある	$-\pi/4 \leq \theta_1 \leq \pi/4$ and $\text{center}(A)_x \geq \text{center}(B)_x$	$\exp(-\text{distance1}(A, B)^2)$
r-migi2(A, B)	AはBの真右にある	$-\pi/6 \leq \theta_1, \theta_2, \theta_3, \theta_4, \theta_5 \leq \pi/6$ and $\text{center}(A)_x \geq \text{center}(B)_x$	
r-migi3(A, B)	AはBの右で接する	$-\pi/4 \leq \theta_1 \leq \pi/4$ and $\text{touching}(A, B)$	
r-hidari1(A, B)	AはBの左にある	$\theta_1 \leq -3\pi/4$ or $3\pi/4 \leq \theta_1$ and $\text{center}(A)_x \leq \text{center}(B)_x$	$\exp(-\text{distance1}(A, B)^2)$
r-hidari2(A, B)	AはBの真左にある	$5\pi/6 \leq \theta_1, \theta_2, \theta_3, \theta_4, \theta_5 \leq -5\pi/6$ or $\theta_1, \theta_2, \theta_3, \theta_4, \theta_5 \leq -5\pi/6$ and $\text{center}(A)_x \leq \text{center}(B)_x$	

表2.5 対象間の基本位置関係 (2/2)

位置関係	意味	定義	ヒューリスティックな関係の評価関数 (CF)
r-hidari3(A, B)	AはBの左で接する	$3\pi/4 \leq \theta_1$ or $\theta_1 \leq -3\pi/4$ and $\text{touching}(A, B)$	$\exp(-\text{distance1}(A, B)^2)$
near(A, B)	AはBの近くにある	$0.35 < \frac{\text{lmax}(A) + \text{lmax}(B)}{\text{distance1}(A, B)} < 0.5$ and $\text{remote}(A, B)$	$1 - \text{abs}(1 - \frac{\text{lmax}(A) + \text{lmax}(B)}{\text{distance1}(A, B)})$
far(A, B)	AはBの遠くにある	$0.0 < \frac{\text{lmax}(A) + \text{lmax}(B)}{\text{distance1}(A, B)} < 0.3$ and $\text{remote}(A, B)$	$1 - \frac{\text{lmax}(A) + \text{lmax}(B)}{\text{distance1}(A, B)}$
along(A, B)	AはBに沿っている	$\text{touching}(A, B)$ and $\text{CF}_{\text{along}} \geq 1/3$	$1 - 2 * \frac{\text{lmax}(A)}{\text{lmax}(B)}$
s-migi(A, F)	Aは右にある	$-\pi/2 \leq \theta_3 \leq \pi/2$	$(-3 * \text{distance1}(A, F) * (\text{distance1}(A, F) - 1) + 1/4) * (1 - \text{ABS}(\theta_1) / \pi)$
s-hidari(A, F)	Aは左にある	$\pi/2 \leq \theta_2 \leq \pi$ or $-\pi \leq \theta_2 \leq -\pi/2$	$(-3 * \text{distance1}(A, F) * (\text{distance1}(A, F) - 1) + 1/4) * (1 - (\pi - \text{ABS}(\theta_1)) / \pi)$
s-ue(A, F)	Aは上にある	$0 \leq \theta_5 \leq \pi$	$(-3 * \text{distance1}(A, F) * (\text{distance1}(A, F) - 1) + 1/4) * (1 - \text{ABS}(\theta_1 - \pi/2) / \pi)$
s-shita(A, F)	Aは下にある	$-\pi \leq \theta_4 \leq 0$	$(-3 * \text{distance1}(A, F) * (\text{distance1}(A, F) - 1) + 1/4) * (1 - \text{ABS}(\theta_1 + \pi/2) / \pi)$

F : 表示画面の中心, $\theta_1 = \text{azimuth1}(B, A)$, $\theta_2 = \text{azimuth2}(B, A)$, $\theta_3 = \text{azimuth3}(B, A)$, $\theta_4 = \text{azimuth4}(B, A)$, $\theta_5 = \text{azimuth5}(B, A)$

例：空間的な右に対するルール表現；

```
If aの左端が右半面にある.  
then s-ue(a, Frame) = True  
with 漠然さの評価関数.
```

図2.6に、空間的に右らしさを示すヒューリスティックな関数の例を示す。

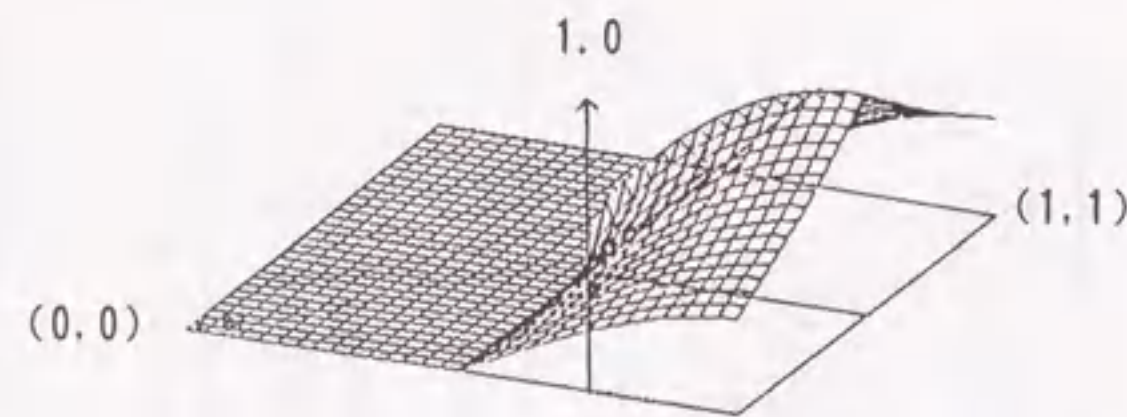


図2.6 位置関係「右」の漠然さの表現例

(3) 位置関係

自然言語の位置関係を、基本位置関係のAnd, Or, Combination関係の組合せにより表現する。英語のonとupは、AとBが離れている事、接触している事が条件になる相対的な基本位置関係r-ue1とr-ue2で、定義する。「Aの上にBがある」といった、英語のonとupの意味を含んだ日本語の相対的位置関係「上(A, B)」は、基本位置関係r-ue1とr-ue2のOr関係で定義する。

例：up(a, b)；

```
If r-ue1(a, b) then up(a, b).  
on(a, b)；  
If r-ue2(a, b) then on(a, b).  
上(A, B)；  
If r-ue1(a, b) or r-ue2(a, b)  
then 上(a, b).
```

2.2.4 合成語の漠然さの表現

自然言語には、「右上」の様に他の位置関係「右」と「上」の合成で定義される位置関係も多い。これらの他の位置関係で定義される位置関係に対しては、個々に基本位置関係で定義する方法と、位置関係のAnd, Or, Combination関係で定

義する方法がある。後者の方法のほうが、自然言語における定義との対応をとり易い。And, Or, Combination関係で定義された位置関係の漠然さは、定義に用いられた関係のもつ漠然さを反映している。

And, Or関係で定義される位置関係の漠然さの評価は、MYCIN⁵⁸⁾のCF(certainty factor)値の計算方法に従い、もとの位置関係のもつ漠然さのMin, Maxで評価する。Combination関係に関しては、証拠の独立性を仮定したMYCINの結合関数による評価では、どちらかの証拠のCF値が1の時、結合結果のCF値は1になる。一方、位置関係においては、位置関係「右」と「上」は必ずしも独立とはいえない。「右」または「上」らしさが増えれば「右上」らしさは増すが、「右」らしさが1でも、必ずしも「右上」らしさは1とは限らない。

このような位置関係のCombination関係で定義される位置関係の漠然さを結合するには、

1. 結合する関係の漠然さが両方とも1の時、合成される関係の漠然さは1になる。
2. 合成された関係の漠然さは、結合する関係の漠然さに応じて単調に変化する。
3. 合成された関係の漠然さは、結合する関係の漠然さよりも小さい。

といった性質を満たす必要がある。この性質を満たすようにMYCINの結合関数を修正した関数を、位置関係の結合関数として、

$$s(\text{comb}(x,y)(a)) = \frac{s(x(a)) + s(y(a)) - \lambda * s(x(a)) * s(y(a))}{2 - \lambda}$$

を用いる。

ここで、s(rel(a))は対象aが関係relらしさである漠然さ、comb(x,y)は関係xとyにより合成された関係である。λは、0 < λ < 1の範囲でヒューリスティックに設定する。λ = 1の時、位置関係の結合関数はMYCINの結合関数に等しくなる。

2.3 位置関係の作成

2.2で述べたif-then型のルールによって表現される位置関係の表現規則を図2.7に示す。幾何的な位置関係である「空間的な右」、「右上」の定義例を図2.8に示す。ルール表現された知識にもとづき、推論する知識ベースシステムにプロダクションシステムがある⁶¹⁾。プロダクションシステムは、知識ベースとして


```

(defrel コメント
  (definition (rel <レベル> 定義する関係名 (対象リスト))
    <= (成立条件の記述)
      (漠然さの評価式)))
成立条件の記述 ::=
  (and {<位置関係> | <属性情報> | <test型>}) |
  (or {<位置関係> | <属性情報> | <test型>}) |
  (not <位置関係> | <属性情報>) |
  <位置関係> | <属性情報> | <test型> | <bind型>) |
位置関係 ::= (rel <レベル> 関係名 (対象リスト) {(値)})
属性情報 ::= (dat attribute 対象名 属性値)
test型 ::= (test S式)
bind型 ::= (bind 変数 S式)

```

図2.7 位置関係のルール表現規則

```

(defrel空間的右の定義
  (definition (rel 2 sright (?A))
    <= (and (rel 1 distance1 (?A Frame) (?Distance))
      (rel 1 azimuth1 (?A Frame) (?Azimuth1))
      (rel 1 azimuth3 (?A Frame) (?Azimuth3))
      (test (and (<= ?Azimuth3 half-pi)
        (<= minus(half-pi) ?Azimuth3))))))
    (cf-calc (* (+ (* -3 ?Distance (- ?Distance 1)) (/ 4))
      (- 1 (/ (abs ?Azimuth1) half-pi)))))

(defrel 右上の定義
  (definition (rel 3 右上(?A ?B))
    <= (rel 3 右(?A ?B)))
  (definition (rel 3 右上(?A ?B))
    <= (rel 3 上(?A ?B)))

```

図2.8 位置関係の定義例

のルールの集合、与えられたデータや推論結果を記録するデータベースとデータベースの内容に応じてルールを実行する推論エンジンから構成される。推論エンジンは、if部に書かれたルールの条件部とデータベースの内容を照合する。条件部をみたすルールのthen部を実行し、データベースを更新する。この照合一実行を繰り返すことで、データベースの内容を目的の状態に変換する。

プロダクションシステムは、画像に含まれる対象の近似多角形の頂点座標を初期状態とするデータベースに、if-thenルールで表現された位置関係の規則を前向きに用いることで、対象の位置記述、対象間の位置関係を導出する。導出された対象の物理量や位置関係 p_i は、

(物理量, 対象 i , 値)

(関係名, 対象 i , [対象], ..., 漠然さの度合) 但し, []内は省略可。の形でデータベースに格納される。画像の内容を、これら物理量や位置関係の集合として規定することができる⁶²⁾。

画像 $\alpha = \{p_i(a, b, \dots) \mid a, b, \dots \in \text{画像}\alpha\} \quad (i=0, 1, \dots)$

図2.9に示す図形画に対し、対象の頂点座標をもとにプロダクションシステムで位置関係を導出した例を図2.10に示す。

自然言語における位置関係は幾何的な位置関係に加えて、対象の意味情報に関する制約が付加される。「ふもと」という2対象間の位置関係は、一方の対象が山であることを意味している。図2.9の図形画の3角形を山、4角形の家、8角形を湖といった意味をもった風景画とみるには、対象の意味情報—attribute(A_1 , 家), attribute(A_2 , 山)—のほか、位置関係の知識ベースに対象の意味を反映したルールが必要になる。図2.11に位置関係「ふもと」を定義した例を示す。図2.12に、対象の意味情報を加え、位置関係を求めた結果を示す。

ヒューリスティックな関数により計算される漠然さの値は、同じ位置関係に対して相対的な尺度を与える。図1.2のような同じ位置関係のA, Bを含む画像(1), (2)に対して、

$$s(r-ue1(A, B))_{at=(1)} = s(r-ue1(A, B))_{at=(2)}$$

であっても、画像(2)には

$$s(r-ue1(C, B))_{at=(1)} > s(r-ue1(A, B))_{at=(2)}$$

となる対象Cが存在する。従って、画像(1)では「AはBの上にある」と参照され

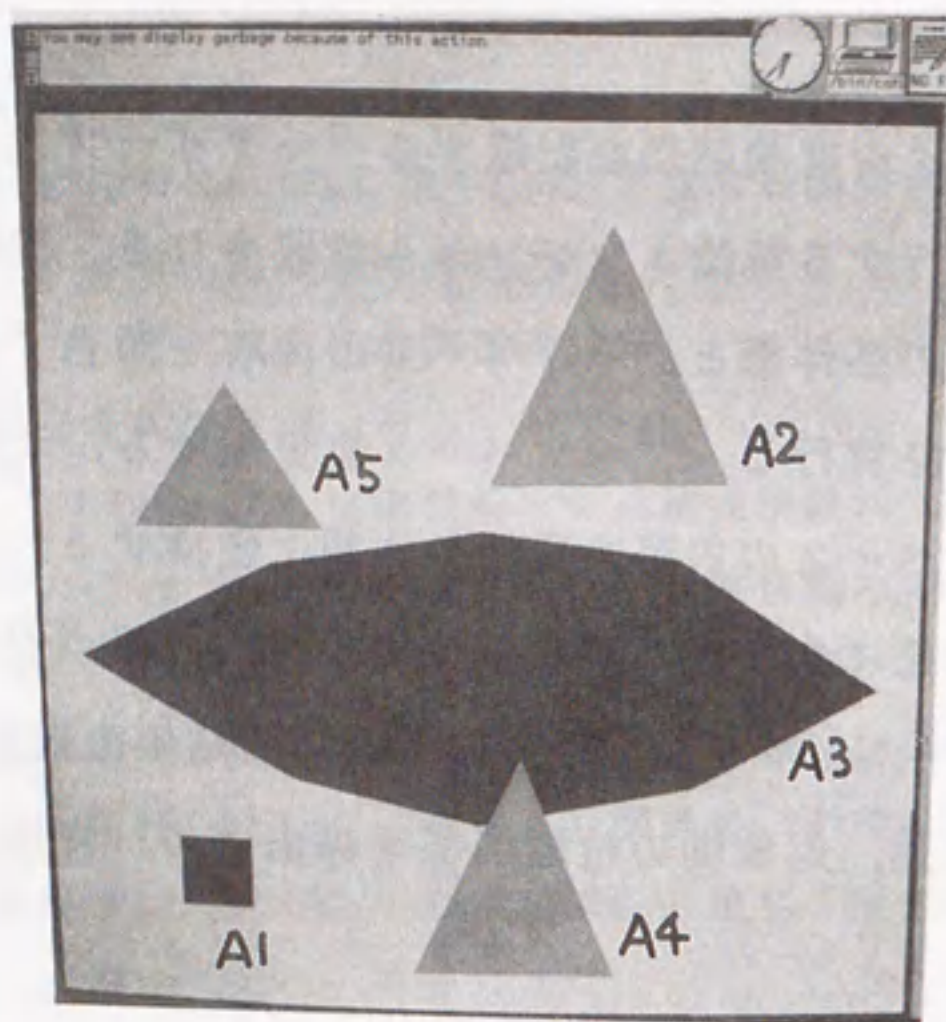


図2.9 図形画

- 0.89 A 3 は A 2 の下にある。
- 0.8 A 1 は左下にある。
- 0.57 A 1 は下にある。
- 0.54 A 1 は左にある。

図2.10 作成された位置関係 (抜粋)

(defrel ふもとの定義例

```
(definition (rel 4 ふもと(?A ?B))
  <= (and (rel 3 下 (?B ?A))
    (dat attribute ?A 山))))
```

図2.11 位置関係の定義例

- 0.89 湖は山の手前にある。
- 0.89 湖は山のふもとにある。
- 0.8 家は南西にある。

図2.12 作成された内容記述 (抜粋)

るのに、画像(2)では「CはBの上にある」の位置関係が、「AはBの上にある」の位置関係よりも適切に参照され、他の対象の存在といった状況依存性にも対応することができる。

2. 4 2章のまとめ

本章では、定性的な位置関係を表現することを目的とし、はじめに自然言語における位置関係の特徴について述べた。位置関係を定量的に表現される基本位置関係と定性的な量である位置関係に階層的にわけ、if-then型のルールで表現することを提案した。人間の持つ位置関係の漠然さを評価するために行った主観実験について述べ、そこで得られた結果をもとにヒューリスティックな関数で位置関係の漠然さを評価する方法について検討した。

さらに、図形画を用い、本方法が位置関係のもつ漠然さや自然言語の位置関係との対応をとり、位置関係の状況依存さがある程度評価できる方法である事を確認した。

第3章 位置関係を用いた画像データベース検索

図表、写真を含む画像データベースの検索には、著者名といった書誌情報のほかに、「北斎の富士山がまん中にあった版画」といった画像の表現しているイメージに関する問い合わせがされる。このような内容検索には、問い合わせに含まれる定性的な位置関係を検索手がかりにして画像を検索することが必要になる。本章では、画像に含まれる対象の近似多角形の頂点座標を画像の管理情報として、問い合わせに含まれる対象の位置関係を、検索インデックス（手がかり）とする画像データベースの検索方法と、試作したSPADE (SPAtial relationships based retrieval/DEscription) システムについて述べる⁶³⁾。

3.1 画像データの管理方法

人間は画像から、描かれている世界に対する背景知識などから、作者名やタイトルなどの書誌情報や属性情報に加えて、自分なりの意味情報を解釈している。そのような人間の解釈した内容を反映した内容検索を実現する方法として、1. 3.1で概観したように多くの方法が検討されている。

内容検索に含まれる位置関係には、対象の幾何的な位置関係のほかに、対象の意味を反映した「ふもと」などの定性的な位置関係や、「2段組」などの画像の内容に固有な言葉が含まれている。内容検索の検索インデックスとして、これらの位置関係の定性的な情報が必要になる。内容検索に十分なインデックスを予めデータベースに入力することは、データ入力者と検索者の主観の相違などにより難しい。

画像の持っている情報には、画像に関する書誌情報や画像に含まれている対象に関する属性情報などの記号情報と、対象の形状、大きさや配置などの特徴に関するパターン情報がある(図3.1)。

テキスト情報：文書画像における雑誌名、題名、著者名や地図画像における建物の名前・種類など対象の属性に関する情報や、画像の表現している世界の情報など。

画像インデックス情報：画像に含まれる対象の形、色、大きさ、位置座標などの情報など。

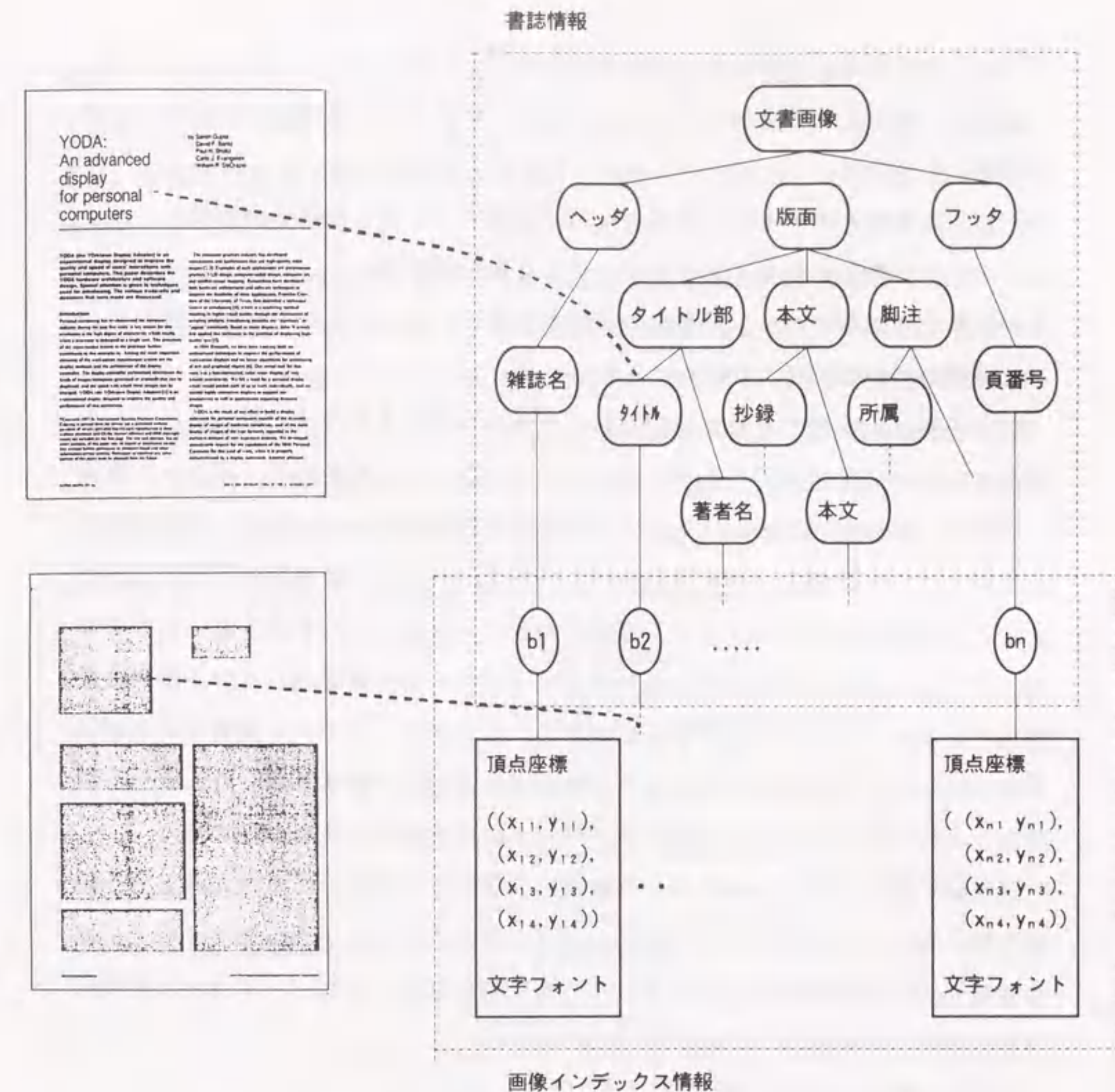


図3.1 文書画像における画像情報

これらのテキスト情報と画像インデックス情報の2種類の情報が、画像データベースの検索インデックスとして要求される。

書誌情報や対象の属性情報は、文字や数値データ形式で表現され文献データベースの形で検索される。一方、対象の色、形や位置関係などのパターン情報には、パターンの特徴やその構造を記述する統一的な方法がない。問い合わせに含まれるパターン情報とデータベースに含まれる情報を照合する方法も、パターン情報の表現方法により色々な方法が提案されている。

3.2 位置関係を手がかりとする検索方法

画像データベースの問い合わせには、画像データベースを管理しているデータの種別に応じて

- 1) 書誌情報などのテキスト情報を検索インデックスにした問い合わせ。
 - 2) パターン情報を検索インデックスにした問い合わせ。
 - 3) テキスト情報とパターン情報の両方を検索インデックスにした問い合わせ。
- の3種類の問い合わせ方法が考えられる。

問い合わせに含まれる書誌情報などを手がかりにする検索は、データベースに管理されている数字や文字のテキスト情報との照合で実現される。パターン情報の検索は、問い合わせにおけるパターン情報の表現方法とデータベースに管理されているパターン情報の表現方法が等しい時は、パターン情報間の類似度を評価することで実現される。テキスト情報とパターン情報の両方を含む問い合わせでは、パターン情報に関する問い合わせとデータベースに管理されている情報の表現方法と異なる形式で照合する必要がある。ここでは、テキスト情報である書誌情報と、パターン情報の一つである対象間の位置関係に関する問い合わせを、同時に、言葉で問い合わせる画像データベースの内容検索方法を提案する。

「北斎の富士山がまん中にあった版画」の問い合わせは、作者が北斎という書誌情報に関する問い合わせと、富士山がまん中にあったという画像イメージに関する問い合わせ内容を含んでいる。この問い合わせを、文献データベース検索用言語SQL的に記述すると以下のようになる。

```
select 画像  $\alpha$  where
    author( $\alpha$ ) = "北斎"
    and
    {  $\exists a \subset \alpha$  such that name(a) = "富士山" and center(a) = True }
```

ここで、author(α)は画像 α の作者、name(a)は対象aの名前、center(a)は対象aがまん中にある時にTrueになる位置関係。

author(α)は画像に関する書誌情報との照合、name(a)は画像のテキスト情報と照合することにより検索できる。一方、内容検索に必要なcenter(a)のような対象の位置関係を用いた検索に関しては、以下の問題点がある。画像に含まれている対象の位置座標のような定量的なデータを入力することができても、

画像に含まれるすべての対象間のあらゆる位置関係を、画像インデックス情報としてデータベースに登録することはできない。さらに、定性的な位置関係に関しては、位置関係自体の持つ漠然さがあり、この漠然さを考慮した照合方法が必要になる。

位置関係に対する検索インデックス作成上の問題点を解決する方法として、対象の頂点座標をインデックスとし、2章で述べた位置関係のif-then型のルールを、検索時に後ろ向きに用いる検索方法を提案する。問い合わせに含まれている位置関係center(a)を、目標として推論を進める。頂点座標が位置関係centerに対する仮説を満たし、名前が富士山である条件を満たす対象を含む画像を検索する。位置関係の持つ漠然さに対しては、ルールに付与されている評価関数を用いて、画像における漠然さを評価する(図3.2)。

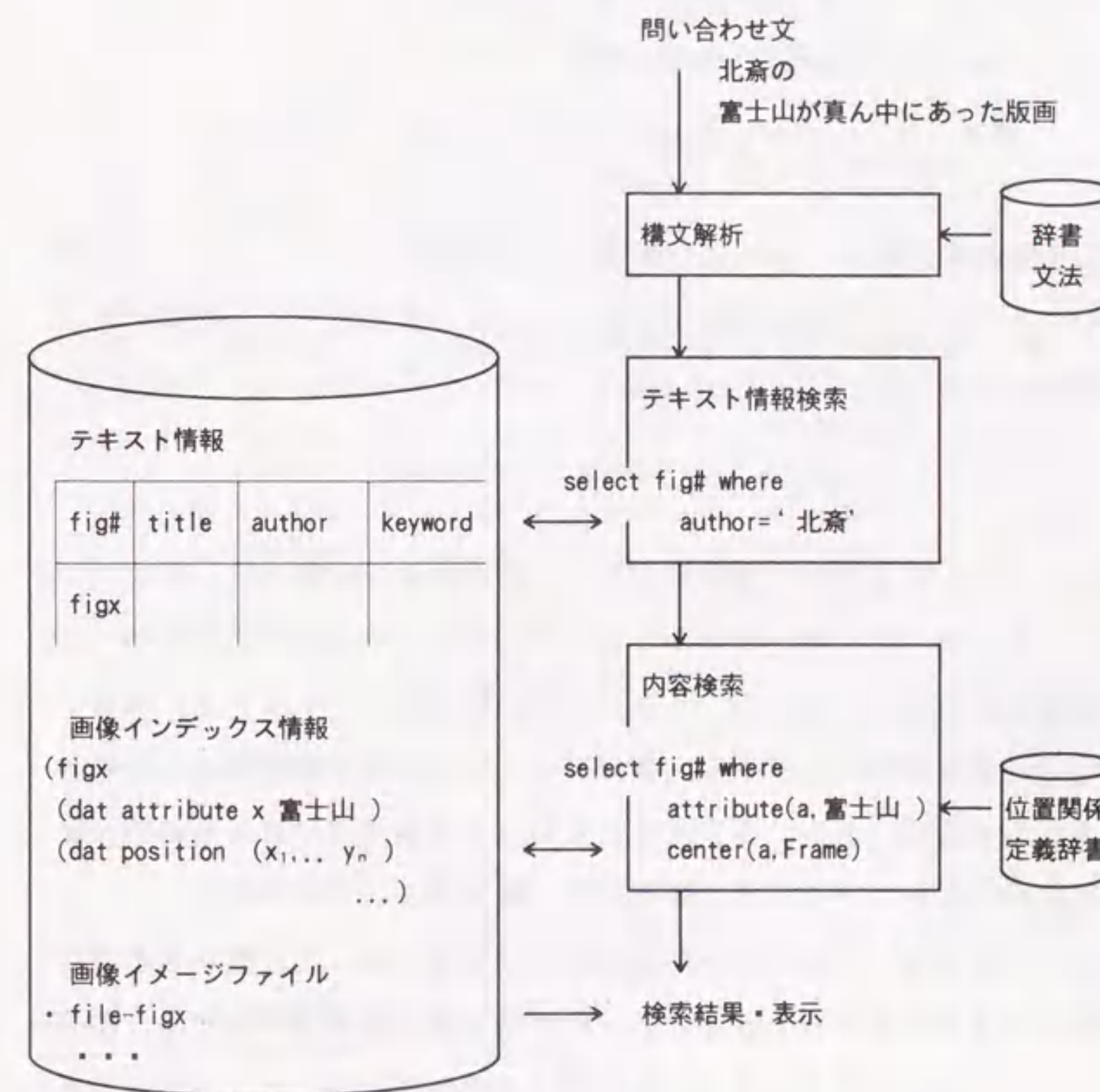


図3.2 位置関係を用いた画像の内容検索

「山のふもとに川のある絵」といった画像内容の問い合わせは、構文解析部により、対象が「山」と「川」で、位置関係「ふもと」を満足しているという3個の検索条件の積集合として表現される。

$\text{attribute}(X, \text{山}) \ \& \ \text{attribute}(Y, \text{川}) \ \& \ \text{ふもと}(X, Y) \ \dots\dots ①$

位置関係「ふもと」に対し、図2. 11の対象の意味情報を含んだ位置関係のルールを後ろ向きに適用することで、位置関係「下」と属性情報「山」が仮説集合に追加される。

$\text{attribute}(X, \text{山}) \ \& \ \text{attribute}(Y, \text{川}) \ \& \ (\text{下}(Y, X) \ \& \ \text{attribute}(X, \text{山}))$

$\rightarrow \text{attribute}(X, \text{山}) \ \& \ \text{attribute}(Y, \text{川}) \ \& \ \text{下}(Y, X) \ \dots\dots ②$

さらに位置関係「下」が表2. 5の定義により、物理関係に展開される。

$\text{attribute}(X, \text{山}) \ \& \ \text{attribute}(Y, \text{川}) \ \& \ (\text{center}(X)_y \leq \text{center}(Y)_y) \ \& \ (-\frac{3\pi}{4} < \text{azimuth1}(X, Y) < -\frac{\pi}{4}) \ \dots\dots ③$

③の仮説集合を満たす対象 x, y を含む画像が、位置関係「下 (x, y)」に付与されたヒューリスティックな関数で評価された漠然さの値とともに、画像データベースから検索される。

3. 3 SPADEシステム

SPADEシステムをSUN3上にLISP, C言語及びORACLEを用いてインプリメントした。ここで、ORACLEは、リレーショナルデータベースシステムで書誌情報の管理に用いた。システムは、構文解析プログラム、検索プログラム、ウィンドウ管理プログラム、画像データベース及び位置関係の定義辞書から構成される(図3. 3)。SPADEにおいて2章で述べた対象の物理量及び基本位置関係をもとに定義した位置関係の一覧を表3. 1に示す。

構文解析プログラムは、自由文法で記述された構文則に従って、問い合わせをテキスト情報に関する検索条件と画像イメージに関する検索条件にわけ、構文解析結果は、書誌情報に関するテキスト条件部の検索結果に対して、パターン条件部の検索を行うSQLの副問い合わせ文的な形式で出力する。

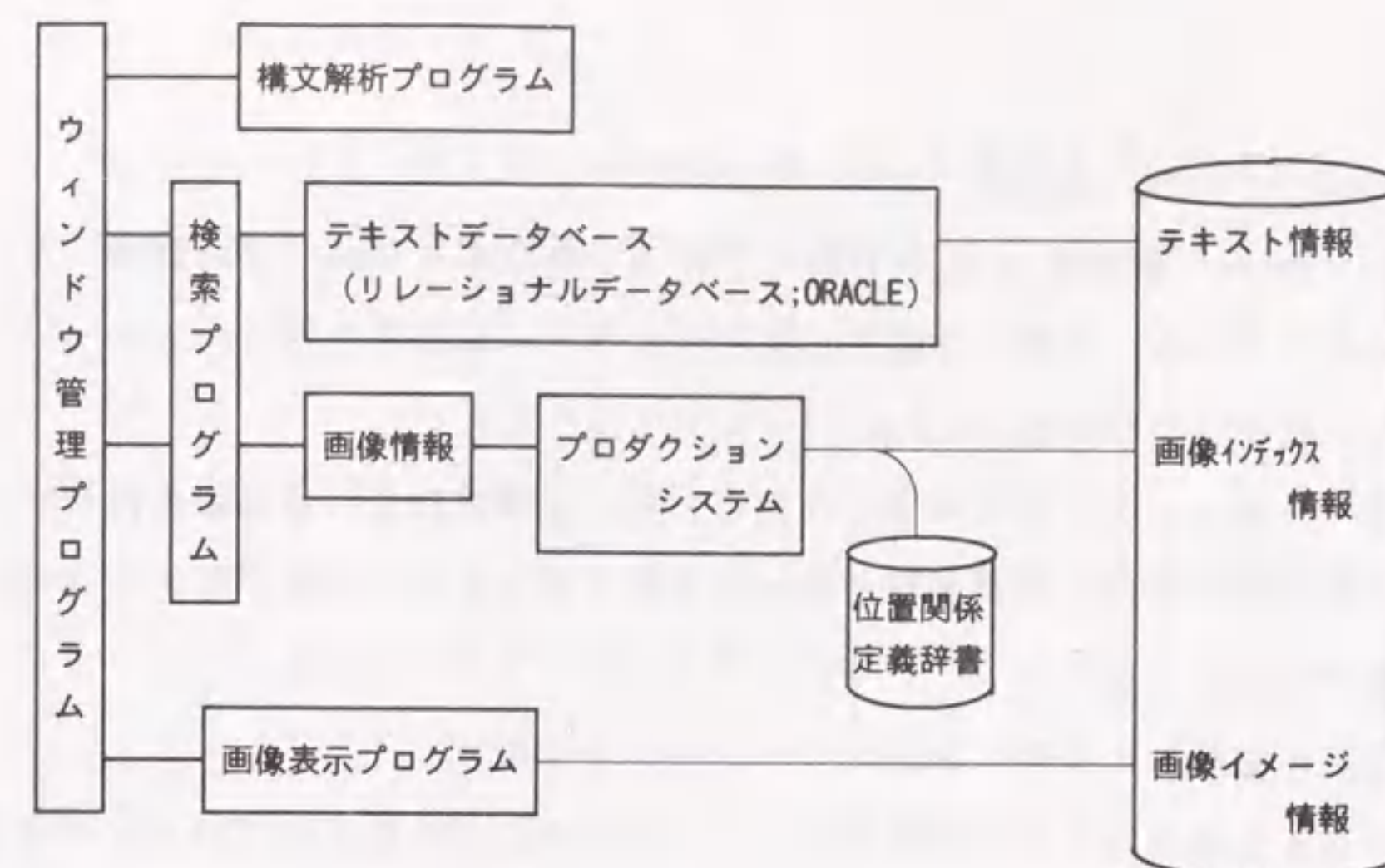


図3.3 SPADEシステム構成

表3.1 SPADEにおける位置関係一覧

レベル	分類*	定義した対象のもつ量と対象間の関係
物理的な量・関係	対象の量 (10)	頂点座標リスト, 辺リスト, 面積, 中心, 近似多角形の外接円の半径, 最左(右, 上, 下)点の頂点, 縦横比率
	対象間の量・関係(11)	対象中心間の距離(distance1)**, 対象間の最短距離, 対象Aの中心と対象Bの中心, 右, 左, 上, 下端との角度(azimuth1~azimuth5)**, 離れている, 重なっている, 接している, 含まれている
基本関係	空間的関係(8)	上[s-ue]***, 下, 左, 右, 中央, 右端, 左端, 隅
	相対的関係(17)	上[r-ue1, r-ue2, r-ue3]****, 下, 左, 右, 遠, 近, 沿う, 間, 囲まれて
位置関係	空間的関係(11)	上, 下, 左, 右, 右上, 右下, 左上, 左下, 右端, 左端, 横
	相対的関係(15)	上, 下, 右, 左, 重なり, 隣接している, 接触, 遠, 近, 上に載っている, 遙か上にある, 手前, ふもと, 二段組, 三段組

*: ()括弧内は定義した量・関係の数。 **: ()括弧内は表2.3における表記。

***: []括弧内は表2.5における表記。

****: 下, 左, 右に対しても同様に3種類の関係が定義されている。

select パターン条件部 from 画像インデックス

where 画像集合 (select テキスト条件部 from 書誌情報)

検索プログラムは、以下の手順で画像データベースの検索を行う。

(手順1) 属性条件の検索 (select テキスト条件部) :

リレーショナルデータベースORACLEで管理されている書誌情報データベースを検索する。検索条件を満たす属性 (タイトル, 作者名等) をもつ画像番号を選択する。

(手順2) 位置条件の検索 (select パターン条件部) :

テキスト条件を満たす画像集合に対し、問い合わせに含まれている位置関係と、画像インデックス情報として登録されている対象位置座標との照合を行う。照合は、位置関係定義ルールを後ろ向きに実行し、位置関係を満たす対象を含む画像を選択する。

(手順3) 検索結果の出力 :

問い合わせに含まれる属性条件と位置条件を満たす画像の番号と、ルールに付与された評価関数で計算された漠然さの値を出力する。指示された画像を表示する。

SPADEの検索画面を図3.4に示す。右半面のウィンドウがメインウィンドウで検索、表示などのコマンドを入力する。左下のウィンドウが検索結果表示ウィンドウで、問い合わせ条件を満たす画像番号と画像に対する位置関係の漠然さの度合を表示する。左上のウィンドウが表示ウィンドウで、表示コマンドにより検索された画像を表示する。SPADEには、問い合わせに含まれている位置関係を検索する過程において、作成された位置関係を画像のインデックス情報として追加登録するコマンド、画像インデックスをブラウズするコマンドがメインウィンドウに用意されている。これらのコマンドを用いて、画像を検索していく過程で作成される位置関係データを、画像インデックス情報として利用することで検索効率を高めることができる。

SPADEの構文解析部として、問い合わせをテキスト情報の問い合わせとパターン情報の問い合わせにわけると簡単な機能のみを実現した。「家のふもとに川のある絵」のように意味のおかしい問い合わせに対しても、構文解析部では単語の意味解析を行わず、定義辞書のルールにしたがって位置関係を展開する。

attribute(X, 家) & attribute(Y, 川) & ふもと(X, Y)

3. 2の①と同様に位置関係「ふもと」の展開を行う。

attribute(X, 家) & attribute(Y, 川) &

(下(Y, X) & attribute(X, 山))

この結果、「山」と「家」の属性を同時に持つ対象xが存在せず、問い合わせを満たす画像が検索されない。SPADEでは、検索要求に該当する画像が検索されないことで、意味のおかしい問い合わせが処理される。



図3.4 SPADE検索画面

3. 4 試作データベース

作者名やタイトルなどの書誌情報による検索以外に、画像の構図や内容に関する問い合わせのある画像の例として、浮世絵と文書画像を取り上げ、SPADEによる内容検索の評価を行った。

3. 4. 1 浮世絵画像データベース

風景画を中心とする浮世絵32枚の画像データベースを作成し、その内容について問い合わせ検索を行った。データ入力者は、ディスプレイ上に表示された浮世絵に対し、入力者が特徴的と思う対象を選択し、その対象の近似多角形の頂点をマウスでクリックする。そして、キーボードで対象の属性情報を入力する事により、浮世絵に対する画像インデックス情報を作成する。図3. 5に、広重の富士三十六景の一つ「東部佃沖之富士」に対し、マウスクリックされた富士山、船、鳥に対する近似多角形を、図3. 6に、対応する画像インデックス情報を示す。浮世絵に関する書誌情報は、浮世絵の書誌情報に関するテーブルと、版画家に関するテーブルの2つのテーブルで構成されるリレーショナルデータベースで管理されている。図3. 7と図3. 8に、浮世絵テーブルと作者テーブルの定義情報と登録内容を示す。

〔検索例1〕「富士山が中央にある絵」といった富士山という名前を持つ対象の空間的な位置関係に着目した問い合わせに対し、検索された3枚の浮世絵を図3. 9に示す。

〔検索例2〕「広重の絵で、橋が富士山の手前にある絵」と書誌情報である作者名と、橋と富士山の相対的な位置関係を含む問い合わせに対し、検索された2枚を図3. 10に示す。

上記の2例とも、検索者の意図に合った浮世絵が検索されている。図3. 10(b)の浮世絵は、検索例1の「富士山が中央にある絵」の問い合わせに対しては、検索されなかった。この浮世絵が検索要求に該当するかどうかは、検索者の主観的な要求にどれだけ合った検索ができるかといった問題を含んでいる。

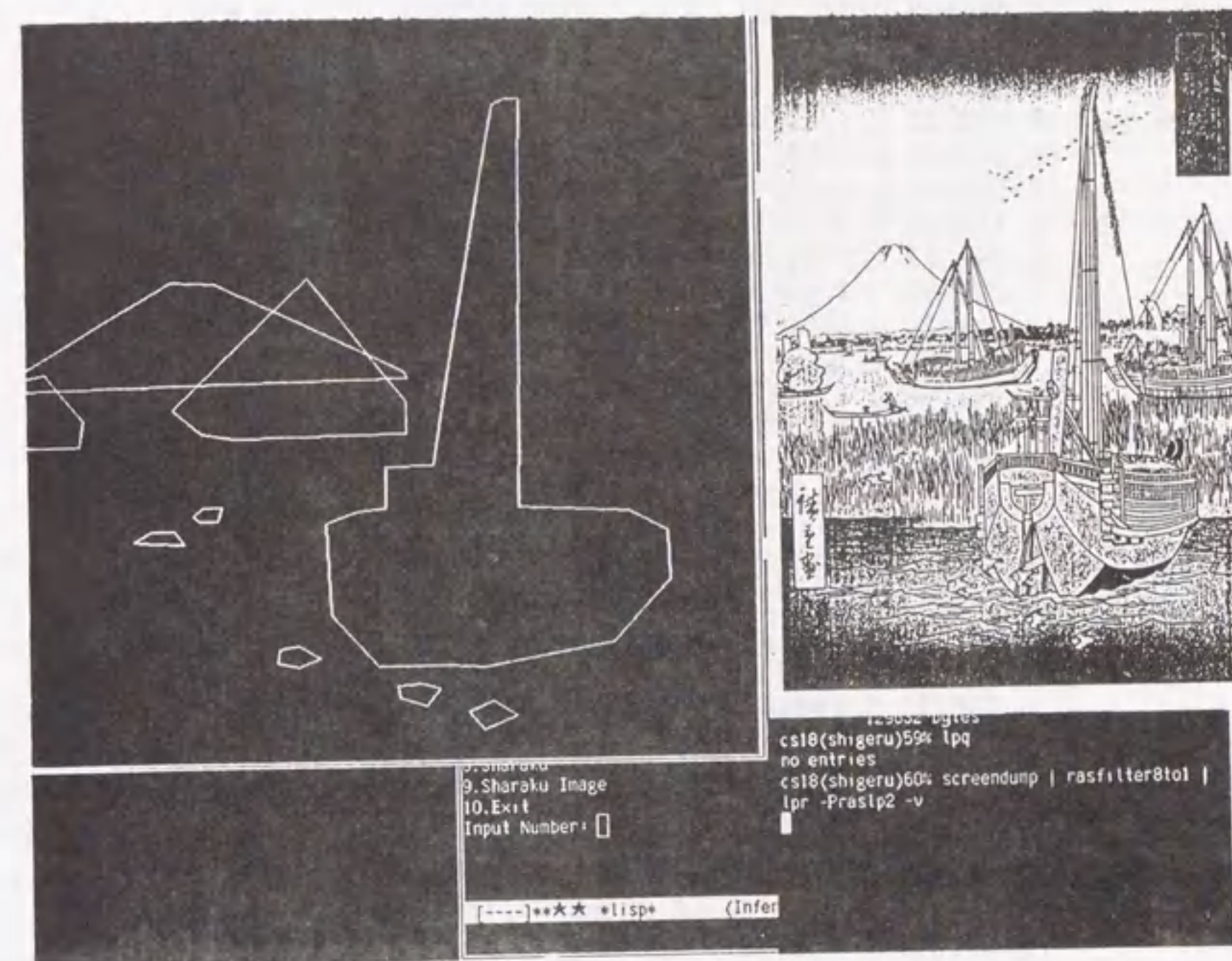


図3.5 浮世絵画像インデックス作成画面


```

(def-figure-data shima1 " "
(dat position Frame ((0.000000 0.000000) (1.000000 0.000000)
(1.000000 1.000000) (0.000000 1.000000)) 1)
(dat position A1 ((0.009585 0.483370) (0.006390 0.430155)
(0.095847 0.430155) (0.102236 0.465632)
(0.076677 0.494457) (0.051118 0.527716)
(0.009585 0.518847)) 1)
(dat position A2 ((0.223642 0.476718) (0.261981 0.443459)
(0.303514 0.436807) (0.533546 0.441242)
(0.533546 0.483370) (0.405751 0.649667)) 1)
(dat position A3 ((0.424920 0.319290) (0.431310 0.197339)
(0.492013 0.128603) (0.587859 0.128603)
(0.635783 0.126386) (0.674121 0.133038)
(0.808307 0.157428) (0.884984 0.239468)
(0.881789 0.305987) (0.830671 0.332594)
(0.683706 0.339246) (0.690096 0.886918)
(0.670927 0.886918) (0.654952 0.878049)
(0.584665 0.492239) (0.568690 0.396896)
(0.504792 0.394679) (0.504792 0.339246)
(0.466454 0.334812)) 1)
(dat position A4 ((0.006390 0.523282) (0.000000 0.505543)
(0.536741 0.514412) (0.530351 0.525499)
(0.281150 0.645233) (0.223642 0.64745)) 1)
(dat position A5 ((0.169329 0.299335) (0.194888 0.297118)
(0.236422 0.294900) (0.223642 0.312639)
(0.191693 0.312639)) 1)
(dat position A6 ((0.249201 0.332594) (0.255591 0.325942)
(0.281150 0.325942) (0.287540 0.343681)
(0.261981 0.343681)) 1)
(dat position A7 ((0.357827 0.135255) (0.386581 0.128603)
(0.415335 0.139690) (0.386581 0.157428)
(0.361022 0.152993)) 1)
(dat position A8 ((0.520767 0.086475) (0.555911 0.075388)
(0.575080 0.097561) (0.549521 0.104213)
(0.517572 0.101996)) 1)
(dat position A9 ((0.613419 0.066519) (0.632588 0.042129)
(0.677316 0.059867) (0.645367 0.079823)) 1)
(dat shape Frame polygon 1)
(dat shape A1 polygon 1)
(dat shape A2 polygon 1)
(dat shape A3 polygon 1)
(dat shape A4 polygon 1)
(dat shape A5 polygon 1)
(dat shape A6 polygon 1)
(dat shape A7 polygon 1)
(dat shape A8 polygon 1)
(dat shape A9 polygon 1)
(dat attribute Frame Frame 1)
(dat attribute A1 船 1)
(dat attribute A2 船 1)
(dat attribute A3 船 1)
(dat attribute A4 富士山 1)
(dat attribute A5 鳥 1)
(dat attribute A6 鳥 1)
(dat attribute A7 鳥 1)
(dat attribute A8 鳥 1)
(dat attribute A9 鳥 1))

```

図3.6 作成された画像インデックス情報

```

create table ukiyoe
(figid char(7),
author char(10),
title char(20),
stitle char(37))

```

(a) 浮世絵テーブルの定義情報

FIGID	AUTHOR	TITLE	STITLE
shima1	立斎広重	富士三十六景	東都佃沖之富士
shima2	葛飾北斎	富嶽三十六景	神奈川沖之浪裏
shima3	立斎広重	東海道五十三次の内	京 三条大橋
shima4	歌川豊春		深川永代橋の景
shima5	立斎広重		東京開化名所両国橋の真景
shima6	三代豊国		東海道関取道中姿
shima7	葛飾北斎	富嶽三十六景	鎧風快晴
shima8	揚州周延	江戸砂子年中行事	元旦の図
shima9	立斎広重	富士三十六景	駿河 薩多の海上
shima10	立斎広重	名所江戸百景	日本橋 雪晴
shima11	歌川国輝		"荒井の渡し"を通過する官軍
shima12	一光斎芳盛		将軍上洛行列の図
shima13	三代豊国	東海道五十三次の内	原
shima14	立斎広重	名所江戸百景	市中繁栄七夕祭
shima15	歌川国利		浅草金龍山境内より隅田川を望む
shima16	歌川国貞		大芝居繁栄之図
shima17	芳艷		東海道名所内 祇園祭礼 長刀ほこの巡幸
shima18	立斎広重	諸国六十余州名所図会	山城 嵐山渡月橋
shima19	立斎広重	東海道五十三次の内	浦原
shima20	立斎広重	東海道五十三次の内	箱根
shima21	立斎広重	東海道五十三次の内	庄野
shima22	歌川国芳		男達五人揃
shima23	立斎広重	東海道五十三次の内	江戸日本橋
shima24	葛飾北斎	隅田川兩岸一覽の内	隅田川 新柳橋の白雨
shima25	国周	新吉原	俄獅子
shima26	国政		歌舞伎
shima27	揚州周延		徳川家康像
shima28	喜多川歌麿		美人花合
shima29	喜多川歌麿	婦女子相十品	ビードロを吹く女
shima30	東州斎写楽		市川海老蔵
shima31	喜多川歌麿		浪華屋おきた
shima32	東州斎写楽	大谷鬼次の奴江戸兵衛	恋女房染分手綱

(b) 浮世絵テーブルの登録内容

図3.7 浮世絵画像の書誌情報(1/2)


```
create table auth
(author char(10),
period char(4),
ybirth char(4),
dbirth char(4),
yomi char(23))
```

(a) 作者テーブルの定義情報

AUTHOR	PERI	YBIR	DBIR	YOMI
立斎広重	江戸	1797	1858	りっさい ひろしげ
葛飾北斎	江戸	1760	1849	かつしか ほくさい
歌川豊春	江戸	1735	1814	うたがわ とよはる
歌川豊国	江戸	1784	1864	うたがわ とよくに
歌川周延	江戸	1836	1912	うたがわ ちかのぶ
歌川国輝	江戸	1830	1874	うたがわ くにてる
芳盛	江戸	1818	1884	よしもり
歌川国利	江戸	1847	1889	うたがわ くにとし
歌川国貞	江戸	1785	1864	うたがわ くにさだ
芳艶	江戸	1822	1866	よしつや
歌川国芳	江戸	1756	1861	うたがわ くによし
国周	江戸	1835	1900	くにちか
国政	江戸	1823	1880	くにまさ
喜多川歌麿	江戸	1753	1806	きたがわ うたまる
東州斎写楽	江戸	1787	?	とうしゅうさい しゃらく

(b) 作者テーブルの登録内容

図3.8 浮世絵画像の書誌情報(2/2)



(a)



(b)



(c)

図3.9 検索例1の検索結果



(a)



(b)

図3.10 検索例2の検索結果

3. 4. 2 文書画像データベース

文書は、内容が理解しやすいようにタイトル、章立てなどの文字や図表が効果的に配列（レイアウト）されている。「著者は誰々で、テーマは何々に関する文献」とか、「あの雑誌に載った誰々の文献」という書誌情報に関する問い合わせのほかに、文献のレイアウトを手がかりに文献を参照される。

科学論文の第1頁の文書画像に関しては、浮世絵のように人手で画像インデックスを入力しなくても、論文レイアウトの知識により論文タイトル、著者名、キーワード等の領域抽出と領域の属性情報を抽出することができる⁶⁴⁾。図3. 11に、入力された文書画像、画像処理により自動抽出された領域及びそれに対応する科学論文の画像インデックス情報を示す。

〔検索例3〕66枚の文書画像に対し、「著者がタイトルの右にある文献」というレイアウト条件で検索した結果を図3. 12に示す。

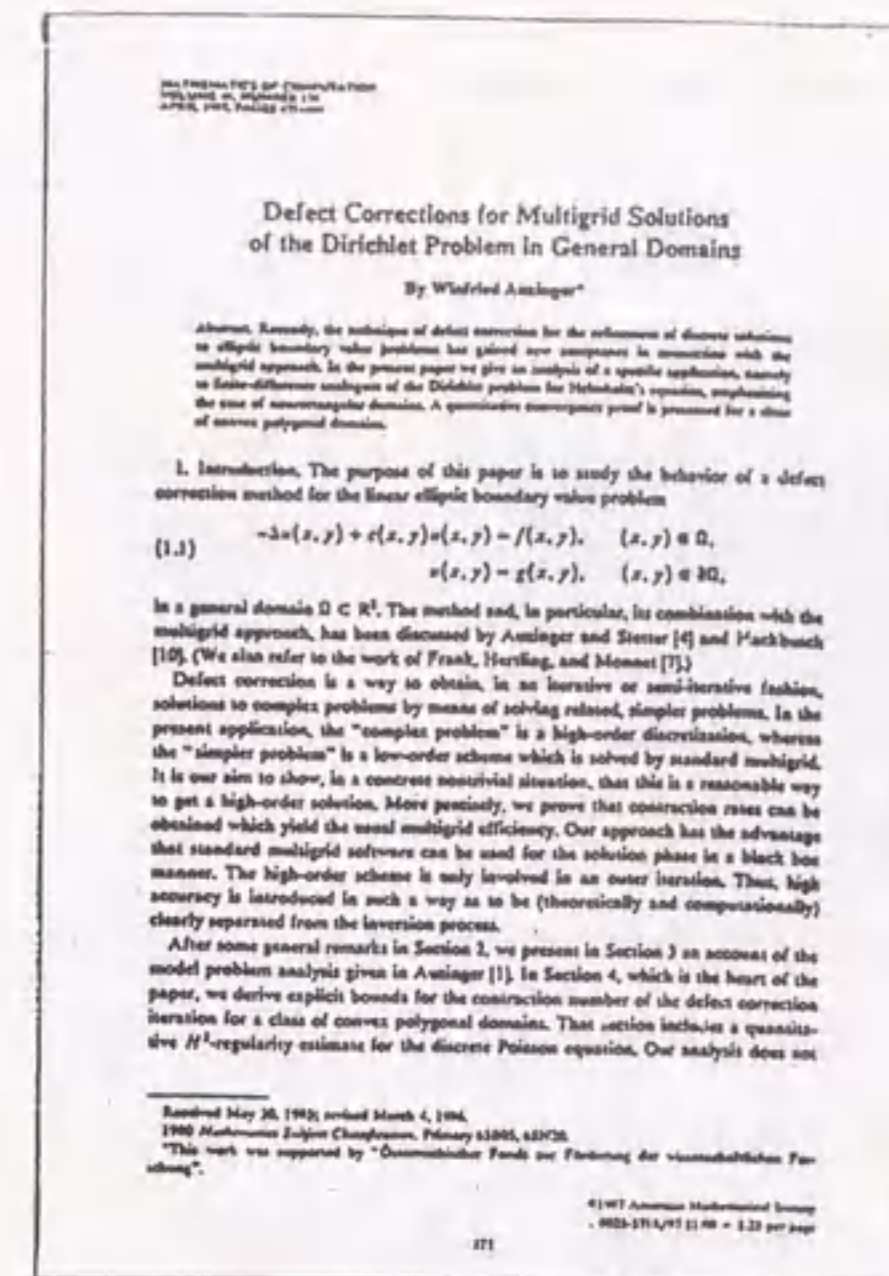
検索された3件のうち(a), (b)の2件は、検索者のイメージしたレイアウト構成を持った文献である。図3. 12(c)の複数著者からなる文献については、「著者がタイトルの下にある文献」の方が適当と思われる。これは、著者属性を持つ領域を独立した領域とし、タイトル属性を持つ領域との間の位置関係「右」で検索するSPADEと、同一属性をもつ領域をまとめ著者全体を1個の領域とする人間との違いによる(図3. 13)。

3. 4. 3 考察

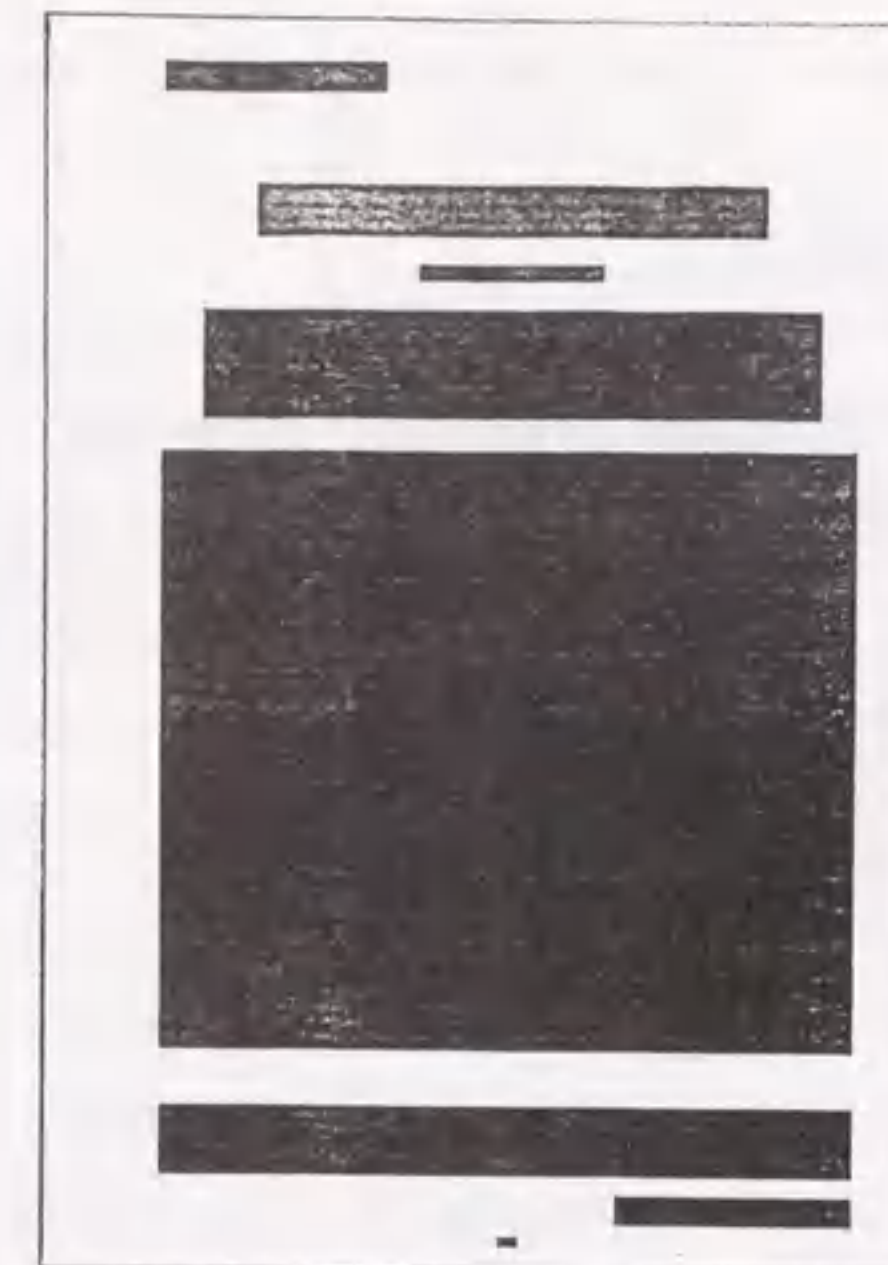
漠然さを持つ位置関係を検索キーにする内容検索の評価に、情報検索システムにおけるキーワード検索の評価に用いられている再現率・適合率を用いる⁶⁵⁾。適合率と再現率は以下の式で定義され、誤りなく検索する、漏れなく検索する尺度を与える。

$$\text{適合率} = \frac{\text{検索した画像のうち検索者の意図に適合した画像の数 (N}_3\text{)}}{\text{検索した画像の数 (N}_2\text{)}}$$

$$\text{再現率} = \frac{\text{検索した画像の数 (N}_2\text{)}}{\text{データベースの全画像のうち検索者の意図に適合した画像の数 (N}_1\text{)}}$$



(a) 入力文書画像



(b) 領域抽出後

```
(def-figure-date26
 (dat position Frame ((0.000 0.000) (0.701 0.000) (0.701 1.000) (0.000 1.000)) 1)
 (dat shape Frame polygon 1)
 (dat position b1 ((0.161 0.798) (0.544 0.798) (0.544 0.839) (0.161 0.839)) 1)
 (dat shape b1 polygon 1)
 (dat attribute b1 TITLE 1)
 (dat position b2 ((0.091 0.911) (0.258 0.911) (0.258 0.935) (0.091 0.935)) 1)
 (dat shape b2 polygon 1)
 (dat attribute b2 NAME 1)
 (dat attribute b2 VOLUME 1)
 (dat attribute b2 NUMBER 1)
 (dat attribute b2 DATE 1)
 (dat attribute b2 PAGE 1)
 (dat position b3 ((0.090 0.074) (0.611 0.074) (0.611 0.127) (0.090 0.127)) 1)
 (dat shape b3 polygon 1)
 (dat attribute b3 FOOTNOTE 1)
 (dat position b4 ((0.346 0.019) (0.360 0.019) (0.360 0.026) (0.346 0.026)) 1)
 (dat shape b4 polygon 1)
 (dat attribute b4 PAGE 1)
 (dat position b5 ((0.434 0.036) (0.611 0.036) (0.611 0.060) (0.434 0.060)) 1)
 (dat shape b5 polygon 1)
 (dat attribute b5 CODE 1)
 (dat position b6 ((0.090 0.170) (0.612 0.170) (0.612 0.634) (0.090 0.634)) 1)
 (dat shape b6 polygon 1)
 (dat attribute b6 TEXT 1)
 (dat position b7 ((0.121 0.659) (0.584 0.659) (0.584 0.743) (0.121 0.743)) 1)
 (dat shape b7 polygon 1)
 (dat attribute b7 ABSTRACT 1)
 (dat position b8 ((0.283 0.764) (0.421 0.764) (0.421 0.777) (0.283 0.777)) 1)
 (dat shape b8 polygon 1)
 (dat attribute b8 AUTHOR 1))
```

(c) 画像インデックス情報

図3.11 文書画像情報

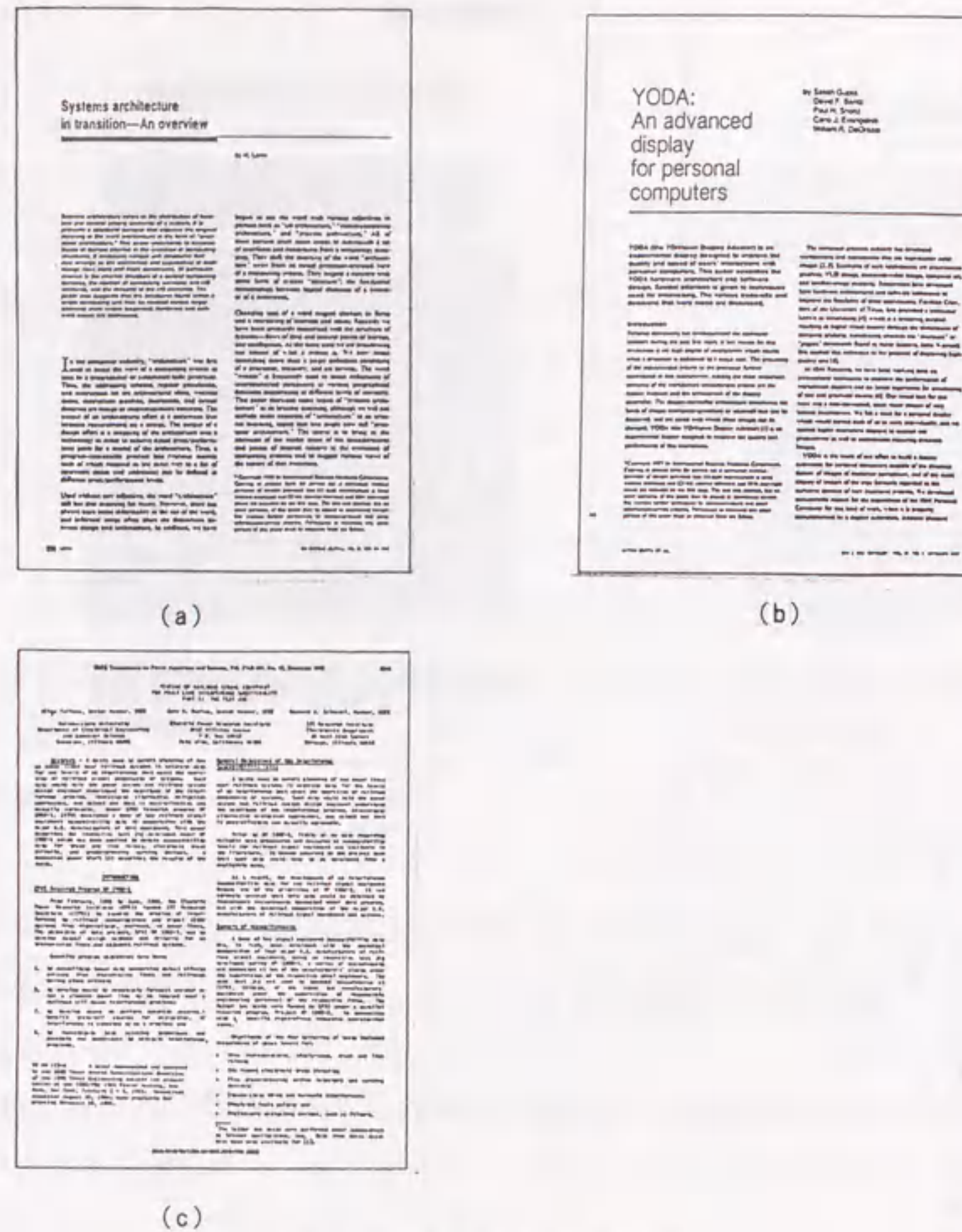


図3.12 検索例3の検索結果

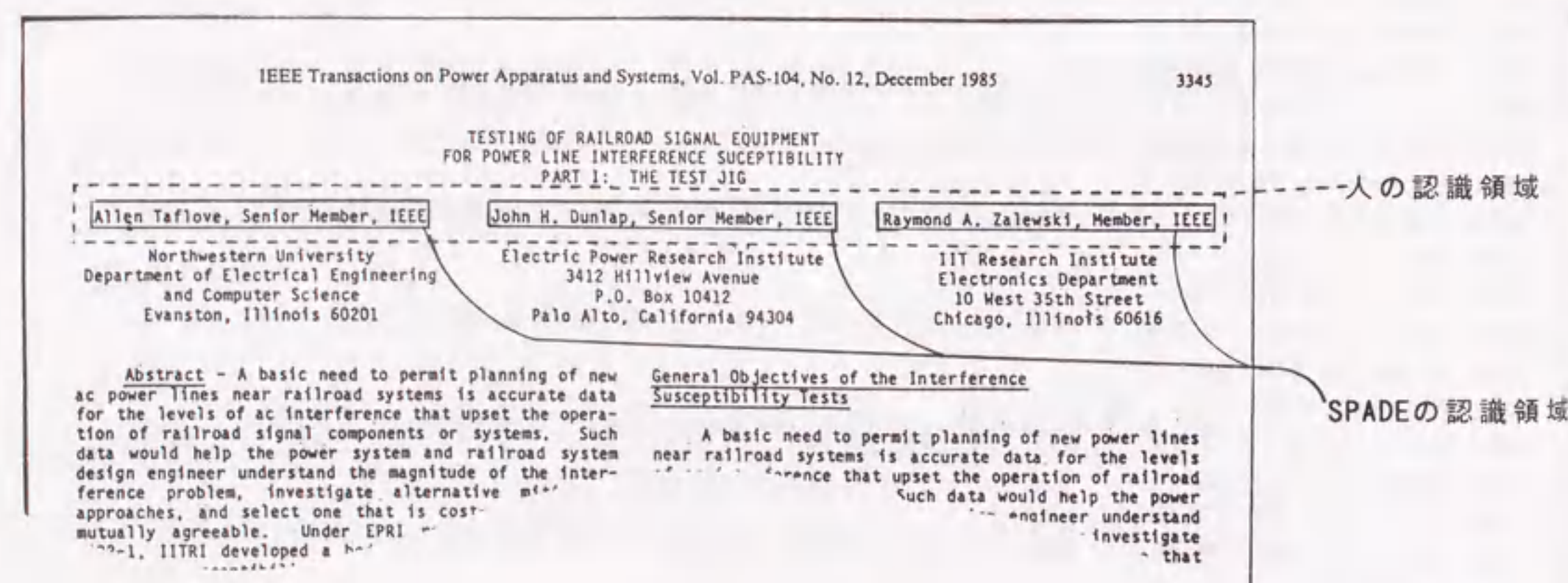


図3.13 著者名のレイアウト

データベース検索に対する要求は、利用者ごとに異なる。少しでも問い合わせに関連のあるものを検索したい利用者は高い再現率の検索をのぞみ、件数が少なくても無駄のない検索をしたい利用者は高い適合率をのぞむ。検索方法を評価する条件が等しければ、適合率と再現率ともに1に近い検索方法が、よい検索結果を与える。

3.4.1と3.4.2.の検索例に対し、書誌情報とテキスト情報を満たす画像数(M)、適合率と再現率の計算に必要な N_1 、 N_2 、 N_3 及び計算された適合率、再現率を表3.2に示す。 N_1 、 N_2 、 N_3 のうち、データベースに入力されている全画像のうち検索者の意図に適合した画像の数 N_1 と、検索した画像のうち検索者の意図に適合した画像の数 N_3 に検索者の主観が入る。検索例1, 2, 3の検索において、一番主観的な要因を含むと思われる検索例1を例に、内容検索に対する適合率と再現率を考察する。テキスト情報「富士山が描かれている」を満たす画像13枚の版面を図3.14に示す。13枚の画像から、再現率の評価に必要な「富士山がまん中にある」条件を満たす画像を選ぶには、位置情報のほかに、「富士山」が浮世絵のテーマになっているなどの判断が含まれると思われる。検索された3枚の浮世絵は、問い合わせの条件にあっている。

以上、再現率に必要な N_1 の評価は今後の問題であるが、SPADEの方法により適合率の高い内容検索が行えた。

表3.2 再現率と適合率

	M	N_1	N_2	N_3	再現率	適合率
検索例1	13	*	3	3	$\frac{3}{*}$	$\frac{3}{3}$
検索例2	3	2	2	2	$\frac{2}{2}$	$\frac{2}{2}$
検索例3	66	3	3	2	$\frac{2}{3}$	$\frac{2}{3}$



図3.14 富士山を含む版画

3.5 第3章のまとめ

本章では、画像のもつテキスト情報と位置関係に関するイメージ情報を検索手がかりとする画像検索法について述べた。書誌情報と画像に含まれる対象の頂点座標で管理されている画像データベースを、2章で提案した位置関係のルールを用い、位置関係を含んだ言語による問い合わせで検索する方法を示した。浮世絵画像や文書画像を検索するプロトタイプシステムを試作し、対象の意味的情報を含んだ位置関係に関する問い合わせに対して、適合率の高い検索ができる等の有効性を確認した。

今後の課題として、主観的な尺度を含む問い合わせに対する適合率・再現率の検討等がある。

第4章 インタフェースとしての手振り

手話や舞踊においては、手振りや身振りを楽譜の様に記述し^{42), 66)}、それから動きを再現できるような方法がとられている。計算機と人間のインタフェースに手振りを利用するには、手の形をコード化し、その手の形の意味との対応をとる必要がある。本章では、手話の指文字に対応する46個の手形状を、市販の手形状入力装置を用いて測定したデータを用いて、主成分分析、クラスタ分析を行った結果について述べる。次に、指の曲げ角度と手の向きから、手の形をコード化する方法と、指文字の46個の手形状を測定データからのコードから識別した結果について述べる。

4.1 手振りの表現

日常生活において、我々は音声情報と手の動きを同時に用い、音声や手形状による指示に加え、同じ手の形でも動きの違いによる複数の指示形態でコミュニケーションを行っている。

(1) 手の形だけの指示：(ポイント指定)

このビルの名前を教えて欲しい。

(2) 手の動きを伴った指示：(複数ポイント指定)

ここからここへ行く方法を教えて欲しい。

(3) 手の動きを伴った指示：(範囲指定)

この辺りの公園を教えて欲しい。

このような計算機インタフェースの実現には、手形状装置による手の動きの測定データから、予め登録されている手振りを認識する事が必要になる³⁵⁾。手振りの認識には、静止した手の形形状の認識ほか、時間とともに変化する手の形と、位置や向きの変化である手の動き認識及び手振りの終了、開始時点の認識も含まれる。

手振りの表現方法としては、アメリカ手話の身振りの記述に用いられる構動素がある。構動素は、手話の動作が行われる場所を示すTab記号、手の形を示すDez記号、動きを示すSig記号の3つの部分から構成されている。そして、Tab記号、Dez記号、Sig記号をあわせたTD³の形式で、アメリカ手話に用いられてい

る手振り、身振りが表現される。Dez記号には、すべての指を伸ばした形を示す「5」や、cの形のように手の形を示す「C」のように視覚的な特徴をもとにした19種類のコードがある。黒川らは、構動素の考え方を継承して、「指を伸ばしている」状態、「指同士が重なった」状態などの13個のコードを用いた片手の形態表現方法を提案した⁶⁷⁾。

構動素に用いられている視覚的な特徴には、指の重なりや指先の接触など、画像処理で抽出することは難しい特徴が含まれている^{39), 40), 41)}。視覚的な特徴を手形状入力装置からのデータから抽出することは難しく、視覚的な特徴をもとにした手の形状表現を、計算機インタフェースのコードとして使用するには問題が多い。

手形状入力装置を用いたシステムでは、測定される各指の曲げ角度の範囲で、手の形を表現する方法が広く用いられている^{46), 47)}。この方法では、対象タスク固有の手の形表現となり、インタフェースにおいて統一的な扱いができにくい、時間的な変化である手の動きを表現しにくいなどの問題がある。

本研究では、インタフェースに用いられる手振りを、

- (1) ある時点での指の曲げ・重なり、手の向きで決まる静止した「手の形」。
- (2) 連続した手の形データの変化で表現される「手の動き」。
- (3) 「手の動き」を組み合わせで表現される「手振り」。
- (4) 基本的な手振りを組み合わせで表現する「手話」。

の4つに分類する。そして、手形状入力装置を用いて、(1)の手の形と(2)の手の動きを認識するのに必要となる、手の形のコード化方法について考える。

4.2 手形状入力装置による手の形識別

図4.1に手形状の測定に用いたシステム構成図を示す。手形状入力装置としては市販のData GloveTMを用いた³⁷⁾。Data GloveTMは米国VPL社の製品で、指の曲がり具合と、手の3次元位置と手の向きを測定する。親指の第2, 3関節、他の指の1, 2関節の10関節の曲げを、指にそった光ファイバ内の光量の変化から測定する。測定値を、パソコンMacintoshで個人差を補正し、関節の曲げ角度に変換する。測定分解能は1°、測定誤差は5°である。手の甲の部分に米国MacDonnell Douglas社の3次元ディジタイザ3SPACE TRACKERTMが付けられている。

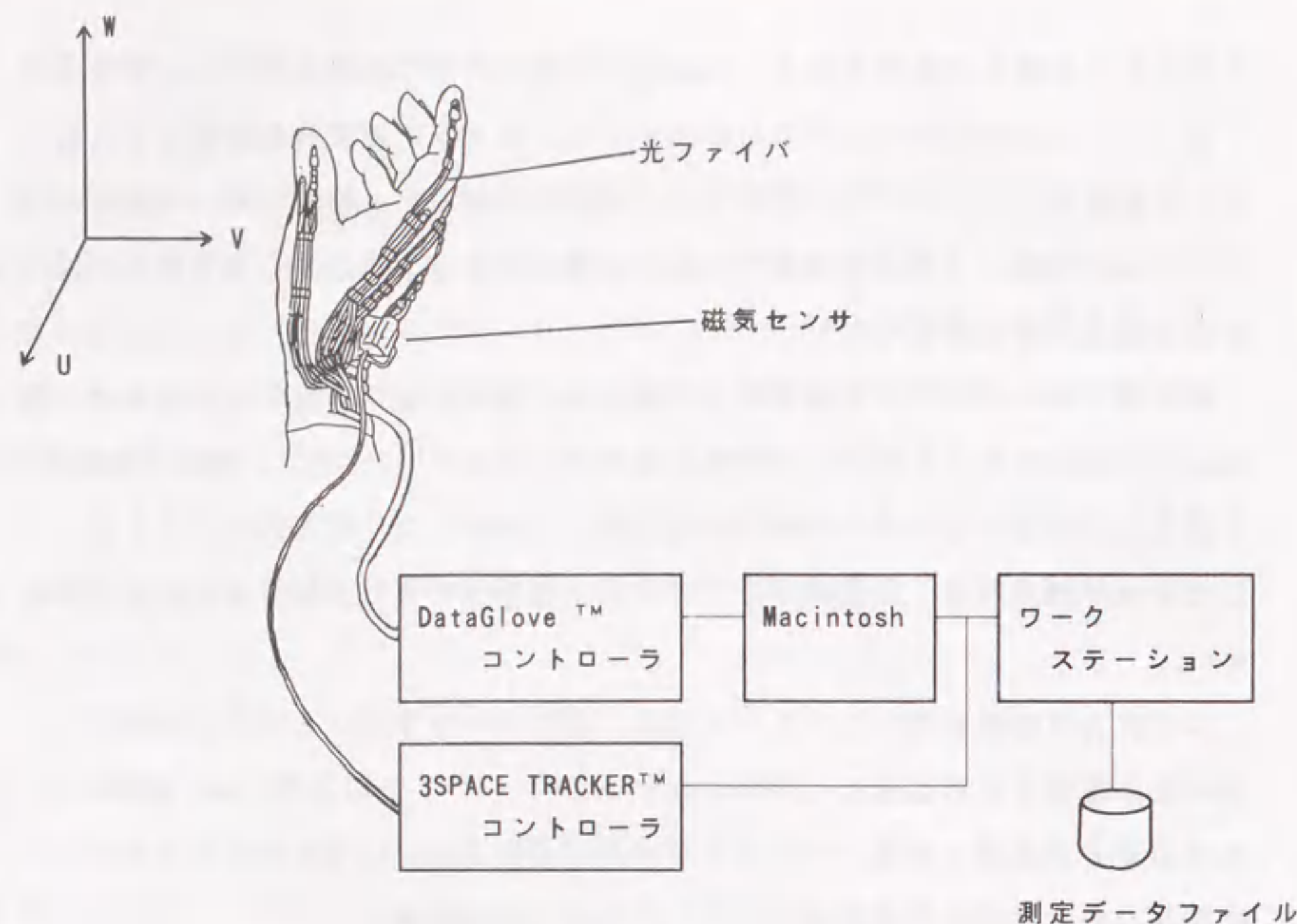


図 4.1 手形状測定システムの構成

3SPACE TRACKER™は、磁気発生器と磁気センサを用い、3次元位置と向き（磁気発生器のもつ座標系と磁気センサの座標系とのなすオイラー角—方位角，仰角，回転角—）を測定する。磁気発生器から4 inch以上28 inchの範囲で、磁気センサの位置に対する測定分解能は0.03 inch，測定誤差は0.1 inch，オイラー角に関しての測定分解能は0.1°，測定誤差は0.5°である。

検討する手振りとして、日常の自然な手振りを含み、人間が容易に異なる手の形と認識している日本語手話の指文字を用いた。日本語手話における指文字の手形状の特徴を以下に示す³⁸⁾。

(1) 清音46文字に対する指の形のうち、41個は手の形だけで文字を表現する。

4文字「の，り，を，ん」に対しては、手全体の動きで表現する。「も」は、親指と人指し指の動きで表現する（図4.2）。

(2) が，ざ，だ，ば行の濁音に対しては、対応する清音の手形を右に動かす事で表現する（図4.3）。



図 4.2 日本語手話の指文字（清音）
（手話入門³⁸⁾ P. 237より）



図4.3 日本語手話の指文字（濁音）
（手話入門³⁸⁾ P.241より）



図4.4 日本語手話の指文字（半濁音）
（手話入門³⁸⁾ P.241より）

(3)ば行の半濁音に対しては、対応する清音の手の形を上にかす事で表現する（図4.4）。

手形状入力装置をつけた一人の被験者の日本語手話の指文字の手形状を測定した。「あ」から「ん」の46個（動きのある5つの指文字に対しては、動きはじめの手の形を用いた）の手形状を、20回測定した。被験者が手話の未経験者である事を考慮し、後半の10回の測定データを解析の対象とした。46個の手形状に対する指の関節曲げ角、手の向きの13種類の測定データの平均値（ ar_{14} , ar_{13} , ..., ar_{43} , ar_{53} , ar_{52} = 関節の平均曲げ角度, $aazi$, $aele$, $arol$ = 手の向きに対する平均オイラー角）と標準偏差（ sdr_{14} , sdr_{13} , ..., sdr_{52} = 各関節の標準偏差, sd_{aazi} , sd_{aele} , sd_{arol} = オイラー角の標準偏差）を表4.1に示す。ここで人指し指から順に小指までを1から4、親指を5で、関節に対して第1関節から順に第3関節に対して4, 3, 2で示し、1番目の添字で指の種類を、2

表4.1 手形状測定データ(1/2)(平均値)

手形状	ar_{14}	ar_{13}	ar_{24}	ar_{23}	ar_{34}	ar_{33}	ar_{44}	ar_{43}	ar_{53}	ar_{52}	$aazi$	$aele$	$arol$
あ	86.6	96.9	85.9	96.9	78.9	90.9	89.5	94.1	21.0	5.6	49.8	16.8	194.1
い	82.3	95.4	68.7	90.7	58.2	75.8	2.3	.0	42.2	46.5	65.9	16.6	178.4
う	19.1	3.8	8.8	.0	59.9	82.1	77.5	69.2	33.9	31.1	59.5	25.9	172.1
え	12.4	77.7	5.7	95.6	6.1	99.4	1.4	92.4	63.4	101.2	63.5	23.0	177.6
お	81.5	64.9	58.1	66.0	49.6	62.1	47.8	55.1	14.6	23.6	61.3	25.9	185.1
か	74.3	.3	90.0	7.1	99.3	79.2	97.1	83.3	42.2	3.0	60.4	5.1	191.4
き	45.8	.0	75.2	42.8	67.6	42.3	8.7	.0	14.9	.4	64.1	36.6	170.5
く	65.7	1.6	45.4	9.8	56.7	15.3	39.6	.6	.6	.0	194.1	31.5	34.6
け	28.5	.0	17.2	4.7	3.9	.0	1.1	.0	74.2	40.2	65.5	15.5	181.3
こ	104.1	5.2	97.8	15.8	77.3	21.8	73.9	23.7	12.5	1.0	63.0	14.2	186.8
さ	79.1	97.2	69.1	102.9	65.4	94.6	68.9	92.4	39.9	47.8	67.5	11.9	182.0
し	78.0	1.3	92.8	8.3	113.8	81.8	99.4	86.3	6.2	.0	192.5	33.9	31.5
ず	61.8	.2	97.5	1.1	114.2	78.6	97.9	82.9	5.2	.0	-68.6	25.5	29.7
せ	76.0	83.8	51.9	.4	64.3	66.2	63.9	75.7	13.4	62.2	59.1	5.5	187.5
そ	36.8	.0	70.3	83.6	82.5	78.6	73.7	79.3	7.3	23.6	-40.4	49.0	59.5
た	93.5	91.8	91.5	93.4	91.3	85.7	91.5	91.1	8.1	1.8	-32.4	67.4	152.8
ち	89.1	35.2	83.0	58.4	72.9	45.5	4.3	.0	10.7	13.7	73.0	23.2	175.6
つ	83.4	56.0	77.4	79.2	63.9	4.3	38.7	.0	33.2	24.5	67.0	11.0	189.3
て	1.1	3.9	.8	2.5	.0	.0	.0	.0	15.1	.3	59.4	25.0	182.9
と	18.7	2.8	17.8	1.5	91.4	65.5	77.0	74.6	33.2	34.7	129.3	31.9	8.0
な	54.9	.0	55.7	.0	88.9	69.8	80.3	74.7	42.6	8.9	-71.2	37.7	20.5
に	59.4	.0	45.1	3.4	83.8	69.2	78.4	74.6	36.9	9.5	197.4	31.4	15.9
ぬ	1.4	78.0	62.5	104.5	88.3	76.5	81.1	83.4	30.1	20.0	49.5	37.2	179.6
ね	40.0	5.3	31.8	11.8	37.9	11.2	24.1	1.0	21.5	.7	-76.7	21.8	25.3
の	28.2	.0	52.1	93.2	70.9	88.3	60.3	72.6	19.7	14.3	-120.8	32.4	21.0
は	12.1	2.3	9.8	1.1	69.2	74.5	61.3	72.3	24.3	27.1	-96.8	50.0	62.0
ひ	23.7	.0	66.4	88.3	69.6	75.6	62.9	82.8	11.2	46.8	44.6	17.6	181.2
ふ	81.9	.0	94.1	95.8	105.7	83.6	90.7	82.6	.0	.0	-75.8	23.5	28.0
へ	105.9	95.4	101.1	92.4	95.9	82.4	60.6	3.8	4.8	.0	-61.0	25.6	26.9
ほ	25.5	4.7	15.4	14.1	26.4	6.2	13.9	.7	25.4	2.5	130.3	25.3	19.3
ま	45.5	.0	32.9	.3	43.3	11.8	54.7	69.3	62.5	7.1	-72.2	28.6	15.3
み	55.9	.0	35.2	.4	46.9	7.9	57.3	67.4	57.0	8.8	202.7	27.4	24.1
む	64.6	.9	74.9	95.8	96.8	88.3	92.0	83.3	.0	.0	188.5	21.5	29.4
め	91.7	42.5	29.2	2.1	16.9	.8	2.3	3.3	1.4	14.9	66.5	23.6	190.6
も	69.0	4.3	93.8	92.0	103.7	84.9	93.7	86.8	.0	.0	174.4	44.1	58.6
や	97.0	92.7	90.3	92.7	77.1	77.2	49.6	4.6	9.0	2.3	53.2	8.3	189.1
ゆ	24.1	.0	13.5	.1	29.3	.5	57.2	69.9	49.7	22.2	130.4	34.4	13.1
よ	79.9	.1	58.9	1.5	66.2	5.5	42.9	.6	68.9	33.5	202.2	35.2	37.2
ら	44.1	12.7	.0	72.5	47.8	84.6	65.6	84.7	36.5	31.7	58.8	14.9	185.0
り	14.7	3.6	18.3	7.3	88.0	76.4	76.8	70.2	29.1	25.2	190.3	55.5	19.0
る	28.9	.1	63.5	12.3	100.2	91.3	98.1	84.4	19.8	.0	49.3	10.7	190.3
れ	25.5	4.0	77.0	100.4	97.0	89.4	95.8	85.4	7.1	.1	47.8	21.9	185.0
ろ	1.2	67.3	20.1	95.1	102.6	83.0	97.6	85.2	41.9	37.4	58.1	30.3	179.0
わ	.0	1.6	.0	8.9	2.1	.0	52.7	75.7	44.7	28.8	53.1	30.7	166.3
を	83.2	51.1	58.7	57.4	55.6	54.2	51.2	45.9	6.9	10.9	58.5	25.8	182.6
ん	36.0	.0	59.5	85.0	69.3	83.9	65.1	76.9	13.6	13.2	-78.5	48.9	27.8

表4.1 手形状測定データ(2/2)(標準偏差)

手形状	sdr ₁₄	sdr ₁₃	sdr ₂₄	sdr ₂₃	sdr ₃₄	sdr ₃₃	sdr ₄₄	sdr ₄₃	sdr ₅₃	sdr ₅₂	sdazi	sdele	sdrol
あ	17.5	15.5	15.5	11.9	21.3	5.3	14.3	6.2	10.8	3.5	3.7	5.0	4.6
い	20.0	18.5	15.0	14.3	17.6	5.3	3.7	.0	8.8	10.4	2.6	2.9	3.2
う	14.3	1.3	6.8	.0	20.2	6.8	17.8	12.0	10.0	5.7	2.7	4.5	4.9
え	23.6	21.4	10.9	12.9	14.3	6.7	4.4	5.9	14.0	20.8	3.3	4.7	4.1
お	19.3	25.8	15.0	17.2	15.5	6.0	16.6	8.7	10.2	12.6	3.8	4.0	4.6
か	25.7	.9	23.3	7.0	14.8	5.5	10.7	6.4	8.5	3.2	3.1	3.6	5.1
き	22.5	.0	15.6	11.6	20.9	3.0	6.1	.0	6.7	.7	3.5	9.5	6.6
く	21.0	1.8	13.7	5.0	17.4	2.6	14.3	1.1	1.9	.0	4.8	5.1	4.9
け	24.7	.0	9.0	3.3	7.0	.0	3.5	.0	16.1	7.0	2.4	4.7	3.2
こ	12.8	8.8	13.5	8.5	21.2	5.7	21.6	11.7	12.1	1.6	2.8	2.3	4.6
さ	18.5	16.6	13.1	9.4	21.1	9.9	20.4	5.9	7.7	6.0	3.1	3.9	5.3
し	21.1	1.8	13.6	7.8	2.3	4.8	8.6	6.1	10.1	.0	4.5	4.9	3.4
す	29.9	.6	14.0	3.5	1.8	4.6	9.8	5.9	8.4	.0	4.4	4.7	5.1
せ	22.2	19.1	12.9	.7	20.8	4.1	20.0	5.3	7.1	7.3	4.5	4.3	4.6
そ	22.9	.0	17.6	17.8	21.3	4.1	20.6	4.5	7.4	9.1	8.9	3.2	8.0
た	15.0	15.2	14.4	13.6	18.1	6.1	12.6	4.9	9.0	2.4	28.0	6.8	27.2
ち	19.8	22.8	18.1	18.4	24.3	10.5	4.5	.0	9.5	22.9	3.8	4.8	5.0
つ	20.0	26.0	14.1	16.5	17.8	2.4	18.6	.0	9.8	21.5	1.5	5.8	4.6
て	2.4	3.3	1.7	3.1	.0	.0	.0	.0	13.4	.7	3.7	9.1	5.4
と	9.3	1.4	10.1	2.8	20.7	4.3	16.3	6.5	10.5	9.9	5.9	4.5	5.0
な	25.1	.0	18.2	.0	21.1	5.2	20.0	5.6	9.3	3.0	7.3	4.9	5.7
に	25.5	.0	10.4	2.8	20.5	5.1	18.5	5.9	9.8	3.2	8.5	3.9	6.6
ぬ	3.1	21.2	15.4	8.7	18.4	4.6	13.1	3.7	6.2	10.3	4.8	8.7	5.6
ね	21.7	5.2	9.5	2.9	13.0	1.8	8.3	2.3	13.1	1.3	2.9	4.0	3.0
の	17.7	.0	15.1	13.7	22.6	8.0	20.3	11.7	9.9	5.1	13.0	5.6	5.4
は	16.4	1.6	8.6	3.5	19.9	5.1	17.4	8.8	9.8	6.2	11.6	9.7	14.4
ひ	13.5	.0	15.0	15.5	21.2	3.4	21.0	7.6	7.9	13.7	5.7	7.6	4.1
ふ	25.8	.0	18.9	12.7	12.3	4.6	13.3	6.6	.0	.0	2.6	6.7	3.0
へ	11.6	18.6	15.5	14.7	19.3	4.8	13.0	5.4	5.5	.0	4.0	7.6	4.2
ほ	19.6	2.8	12.3	6.1	23.1	4.2	18.3	2.2	13.0	4.1	2.7	6.1	5.0
ま	20.5	.0	11.3	.7	11.5	5.8	14.0	6.4	14.8	3.6	3.5	6.3	4.3
み	14.8	.0	10.7	1.3	11.9	4.2	14.2	6.5	12.0	2.4	3.5	2.9	3.9
む	22.1	1.4	15.8	12.3	16.7	5.3	12.2	5.3	.0	.0	7.5	3.6	4.0
め	21.2	22.9	9.0	3.5	10.7	1.1	3.9	2.8	2.3	22.5	4.9	2.8	4.8
も	31.4	11.2	15.0	13.4	12.4	5.4	11.5	6.5	.0	.0	9.2	3.2	10.2
や	14.7	17.0	14.5	13.1	23.1	6.8	12.2	5.9	6.5	3.4	2.9	4.5	5.2
ゆ	15.9	.0	9.7	.3	13.1	1.0	15.9	4.8	8.9	6.9	5.5	5.7	4.8
よ	20.9	.3	12.4	2.2	15.4	2.7	13.8	1.6	14.4	6.7	5.1	8.1	5.2
ら	21.2	4.4	.0	17.3	20.2	7.2	19.7	7.7	10.5	9.5	3.2	4.4	5.9
り	11.2	2.0	11.5	3.4	25.9	5.4	18.9	5.8	8.0	7.7	7.2	8.6	6.1
る	19.5	.3	17.1	5.1	16.3	5.7	9.3	6.9	13.1	.0	2.5	5.9	5.2
れ	15.2	2.1	14.9	10.7	17.5	4.3	10.5	5.1	6.3	.3	2.8	3.7	5.7
ろ	3.2	24.5	8.2	13.7	16.2	4.4	10.0	7.1	14.9	10.0	3.5	4.4	5.9
わ	.0	1.0	.0	3.6	3.9	.0	17.8	7.5	10.0	6.0	4.0	6.9	5.3
を	19.3	26.1	11.0	20.1	13.9	5.9	17.6	7.7	5.0	4.2	3.9	8.2	3.8
ん	18.6	.0	11.4	15.2	20.6	3.6	21.9	4.7	7.1	6.9	12.3	7.5	6.7

番目の添字で関節の種類を示す。従ってar₁₄は人指し指の第1関節の平均曲げ角度を示す。この測定データを用い3種類の識別実験を行った。

- 主成分分析により手の形を識別する特徴的な変量を求める。
- クラスタ分析により手の形を測定データの距離をもとに分類する。
- 手形状コード規則により手の形をコード表現し、識別する。

4. 2. 1 主成分分析⁶⁸⁾

親指や人指し指は、他の指の屈伸状態にかかわらず、屈伸させる事ができる。一方、中指、薬指、小指は、隣の指の屈伸状態によって、曲げ易さや伸ばし易さが異なる。このように、手の形をつくる親指、人指し指、中指、薬指、小指の各指の屈伸状態は、その解剖学的な構成からお互いに独立していない。

測定データから、主要な変動に要約して、データの特徴を把握するための統計的手法に主成分分析がある。主成分分析で、手の形の変化を特徴づける成分の解析を行った。以下に、主成分分析について述べる。

同一のサンプルについて、相関関係があるp個の変量(x_1, x_2, \dots, x_p)($p \geq 2$)のN組のデータが測定されたとする。N組のデータには、p個の変数間の特徴が含まれている。主成分分析は、p個の変量の一次結合

$$z = l_1 * x_1 + l_2 * x_2 + \dots + l_p * x_p$$

を仮定し、 $\sum_{i=1}^p l_i^2 = 1$ の条件のもとで、zの分散が最大になる時のzを第1主成分とする。この時の l_i を l_{1i} ($i=1, 2, \dots, p$)で表すと、

$$z_1 = l_{11} * x_1 + l_{12} * x_2 + \dots + l_{1p} * x_p$$

となる。以下、特徴の説明がつくまで $\sum_{i=1}^p l_{ki}^2 = 1$ の条件下で $k=2, 3, \dots$ と

主成分を順次求める。

ステップ1) 第1主成分 z_1 の分散が、最大になる l_{1i} ($i=1, 2, \dots, p$)を求める。

ステップ2) 第k主成分 z_k は、 z_1, z_2, \dots, z_{k-1} と無相関で、分散が最大になる l_{ki} ($i=1, 2, \dots, p$)を求める。

p変量をもつデータの主成分の個数は、 x_i と x_j との共分散を(i, j)成分とする分散共分散行列 Σ に対する固有方程式

$$|\Sigma - \lambda I| = 0$$

から求まる固有値 $\lambda_1, \lambda_2, \dots, \lambda_m$ ($m \leq p$) の個数と等しい。主成分 z_i の分散は固有値 λ_i に等しく、係数ベクトル I_i は単位固有ベクトルとして求まる。第 k 主成分が測定データを表現している尺度として、全主成分の分散の総和に対する各主成分の分散の比率でその主成分の寄与率と定義される。第 1 主成分から第 k 主成分までの和をそれまでの累積寄与率と呼ぶ。

表 4. 2 に、46 個の指文字に対する手形状の関節の平均曲げ角度 ($ar_{14}, ar_{13}, \dots, ar_{53}, ar_{52}$) に対し主成分分析を行った結果を示す。この結果から、

1. 第 1 主成分の寄与率は約 40% で、中指と薬指の関与が多い。
2. 第 4 主成分までで、累積寄与率は 90% 近くになる。
3. 全体を通して、親指の寄与が少ない。

ことが明らかになった。親指の寄与が少ない理由として、親指の曲げ状態は手首近くにある親指の第 1 関節にかなり依存しているが、用いた手形状入力装置ではこの第 1 関節曲げ角を測定できない事があげられる。

表 4. 2 主成分分析結果

第 k 主成分	1	2	3	4
固有値	4200.	2443.	1824.	757
固有ベクトル				
人指し指 (14)	0.163	-0.378	0.423	-0.367
(13)	0.240	-0.499	-0.320	-0.552
中指 (24)	0.338	-0.197	0.427	0.057
(23)	0.481	-0.334	-0.376	0.589
薬指 (34)	0.366	0.164	0.314	-0.002
(33)	0.456	0.150	-0.179	-0.032
小指 (44)	0.328	0.386	0.164	-0.174
(43)	0.317	0.505	-0.254	-0.248
親指 (53)	-0.138	0.057	-0.195	-0.246
(52)	-0.030	-0.026	-0.368	-0.239
累積寄与率	0.405	0.640	0.816	0.889

4. 2. 2 クラスタ分析^{69), 78)}

測定データの構造を知る手段として、データはある決められた基準 (距離) によって分類する手法がある。その一つにクラスタ分析がある。以下に、クラスタ分析について述べる。

データがもつ p 個の変量 (x_1, x_2, \dots, x_p) ($p \geq 2$) についての値が、 N 組のデータについて測定されている時、クラスタ分析は、以下のステップで組合せ的にデータ分類を行う。

ステップ 1) 各データ間の距離を計算する。初期状態として、各データが 1 つのクラスタを構成しているとし、クラスタの個数 M をデータ数 N とする。

ステップ 2) M 個のクラスタの中で最も類似度の大きい (距離の小さい) 対を求め、それを 1 つのクラスタにマージする。

クラスタ数 M を 1 減らし、 M が 0 に等しいならば終了、そうでなければステップ 3 へ進む。

ステップ 3) 新しく作られたクラスタと他のクラスタとの類似度 (距離) を計算し、ステップ 2 へ戻る。

連続量で表わされるデータに対するステップ 2 の距離としては、各個体を p 次元ユークリッド空間の点としたユークリッド距離、

$$d_{rs}^2 = \sum_{k=1}^p (x_{rk} - x_{sk})^2$$

変量の役割を考慮した重みつきユークリッド距離、

$$d_{rs}^2 = \sum_{k=1}^p w_k (x_{rk} - x_{sk})^2$$

データの分散を考慮したマハラノビス距離、

$$d_{rs}^2 = (x_r - x_s)^t \Sigma^{-1} (x_r - x_s) \quad ; \Sigma \text{ は分散共分散行列}$$

などの個体間の距離や、 x_r, x_s のベクトルの接近している度合を表わす量として、内積による類似度、

$$s_{rs} = \sum_{k=1}^p x_{rk} x_{sk}$$

が用いられる。

ステップ3において、クラスタ C_r と C_s がマージされたクラスタ C_h と他のクラスタ C_i 間の距離として、 C_h に属する x_r と C_i に属する x_s の距離 d_{rs} の最小値をとる最短距離法、

$$D_{hi} = \min_{r,s} \{d_{rs} \mid x_r \in C_h, x_s \in C_i\}$$

C_h に属する x_r と C_i に属する x_s の距離 d_{rs} の最大値をとる最長距離法、

$$D_{hi} = \max_{r,s} \{d_{rs} \mid x_r \in C_h, x_s \in C_i\}$$

や、クラスタ間の距離としてマージされたクラスタ C_h に含まれる変数の平均類似度が用いられる。

関節の曲げ角度と手の向きの角度では、表現している内容が異なる。関節の曲げ角度($ar_{14}, ar_{13}, \dots, ar_{53}, ar_{52}$)で表現された46個の指文字の手形状に対し、平均曲げ角間の距離としてマハラノビス距離、マージされるクラスタ内での平均類似度でクラスタ分析した結果の二分木表現を図4.5に示す。図4.5における①から45の番号は、クラスタ化された順を示す。「し」と「す」、「な」と「に」のように、同じ手の形で手の向きで異なる指文字を丸付き数字で示した。クラスタ分析結果、46個の手形状は大きく

- (1) 人指し指と中指が伸びている。
 - (2) 3本以上の指が伸びている。
 - (3) 人指し指と親指が伸びている。
 - (4) 人指し指と中指が曲がっている。
- の4クラスタに分類された。

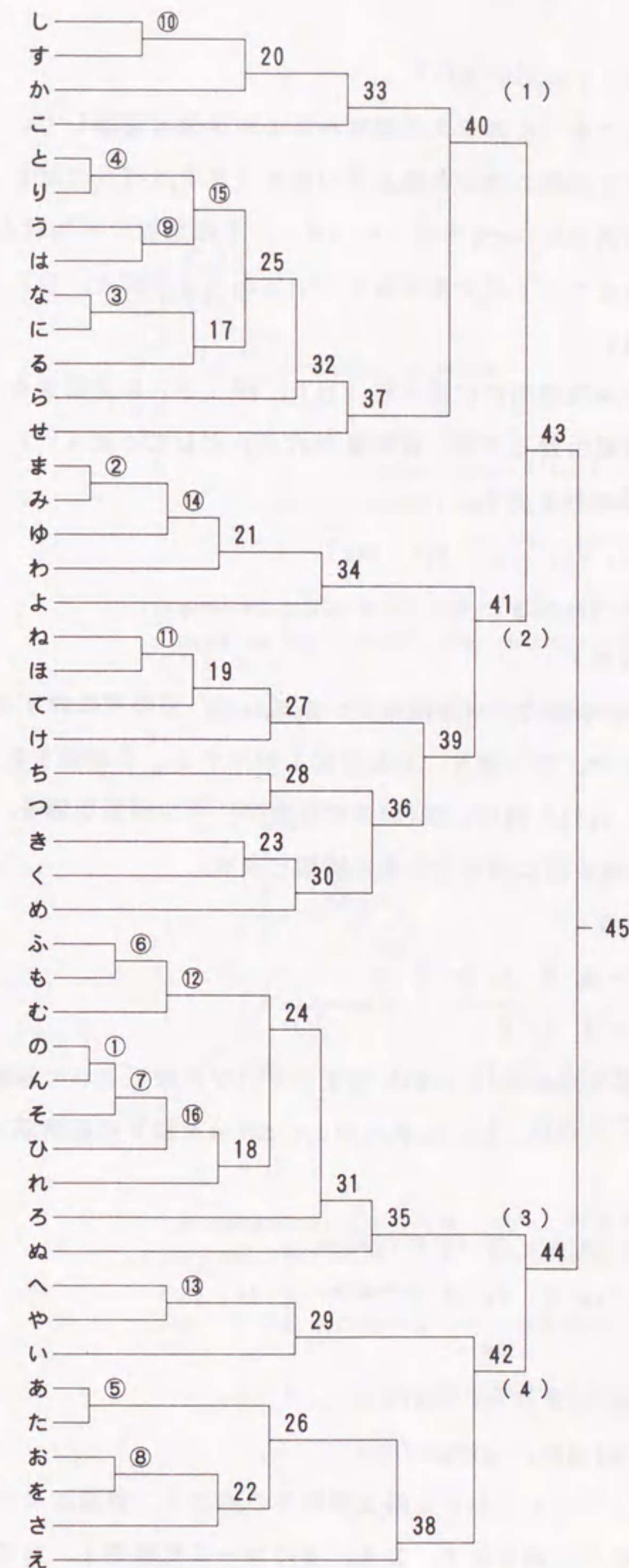


図4.5 クラスタ分析結果の二分木表現

4. 2. 3 コード化方法による手の形表現⁷¹⁾

同じ被験者が同じ手の形をしても、手形状入力装置からの測定値は変動している。ある時間内に測定される10関節の曲げ角度と手の向き（基準座標系に対する手の甲に設定されている座標系のEuler角 ϕ, θ, ϕ ）の測定値の平均値と標準偏差値から、手の形・動きをコード化する方法について述べる（図4. 6）。

（コード化規則1：手の形表現）

曲げ角 45° を閾値とし、関節の屈伸状態を伸（H），屈（B）に分類する。但し、標準偏差が大きい関節に対しては、変動量が大きいため不定（-）とし、10関節の状態での手の形を表す。

手の形表現： $G_a = (P_{14}, P_{13}, P_{24}, \dots, P_{53}, p_{52})$

ここで $P_{ij} = H, B, -$

（コード化規則2：手の向き表現）

基準座標系(U, V, W)から見た中指方向の方向ベクトル(e_1)，手の甲に対する法線ベクトル(e_2)の成分で、手の傾き，回転状態を判断する。手の傾き状態を，上(U)，下(D)，左(L)，右(R)，前(F)，後(B)の6状態で，手の回転状態を，手の掌(H)，甲(T)，横(S)を相手側に向けている3状態で表す。

手の向き表現： $G_b = (Q_1, Q_2)$

ここで $Q_1 = U, D, L, R, F, B$

$Q_2 = H, T, S$

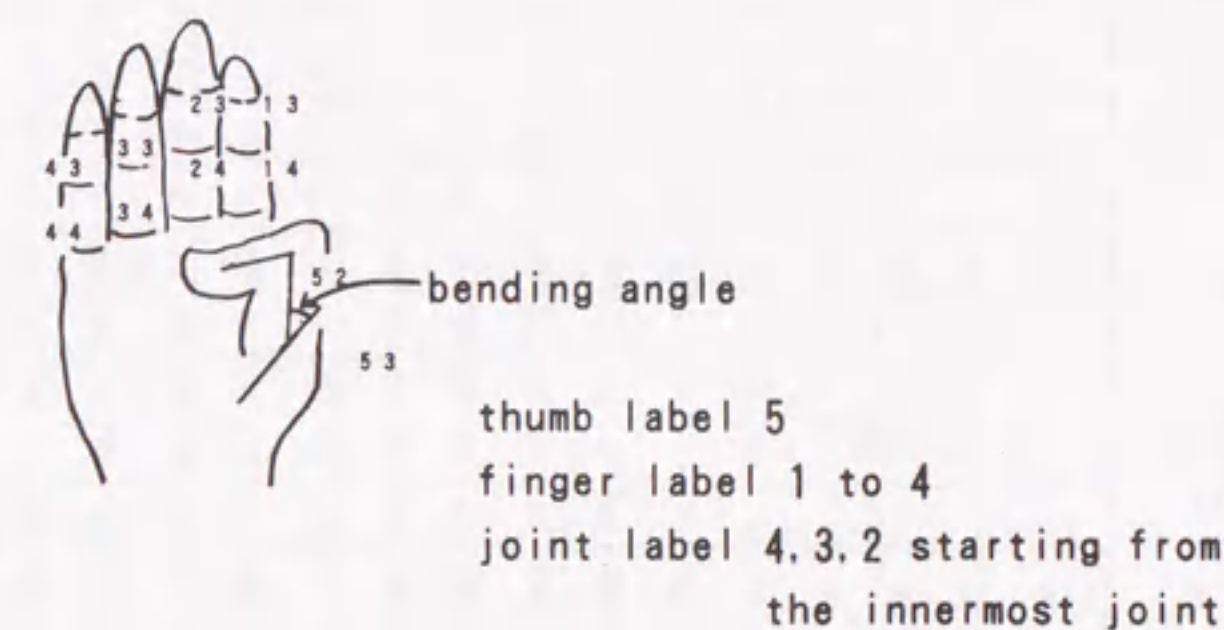
手形状入力装置で測定される基準座標系(U, V, W)に対する手の甲に設定された座標系のオイラー角(ϕ, θ, ϕ)を用い， e_1, e_2 各々のu, v, w成分を以下の変換式で計算する。

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} ca \cdot ce & ca \cdot se \cdot sr - sa \cdot cr & ca \cdot se \cdot cr + sa \cdot sr \\ sa \cdot ce & sa \cdot se \cdot sr + ca \cdot cr & sa \cdot se \cdot cr - ca \cdot sr \\ -se & ce \cdot sr & ce \cdot cr \end{bmatrix} e_i$$

ここで， $ca = \cos(\phi)$ ， $ce = \cos(\theta)$ ， $cr = \cos(\phi)$

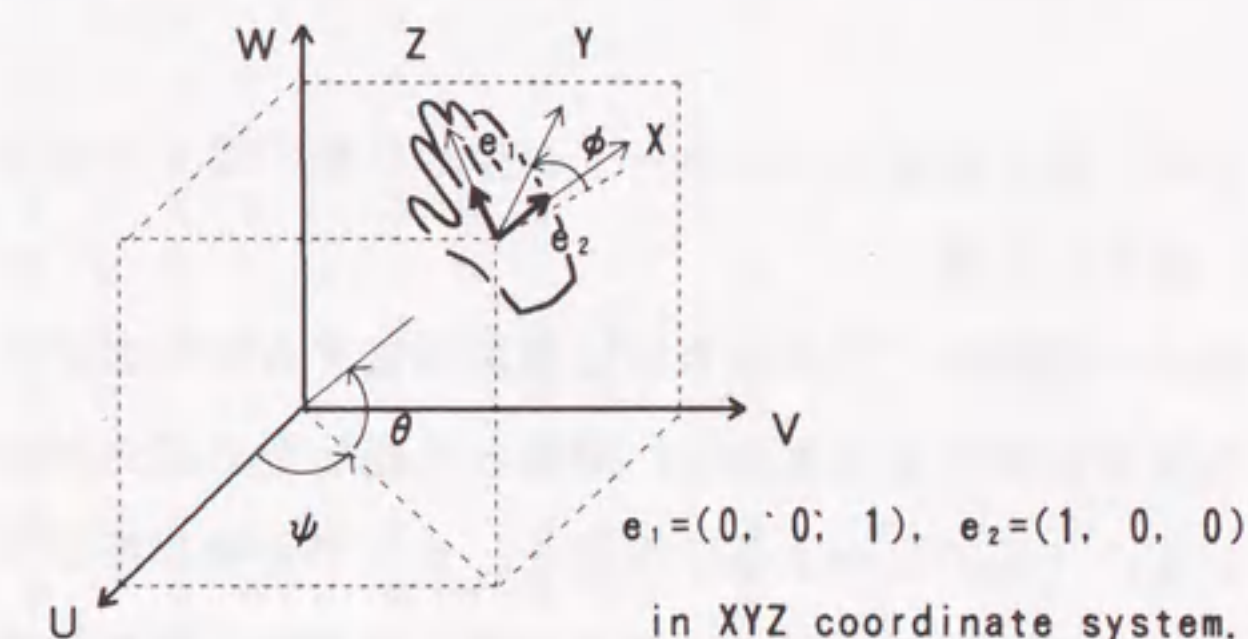
$sa = \sin(\phi)$ ， $se = \sin(\theta)$ ， $sr = \sin(\phi)$

表4. 3に，「ほ」，「う」と「す」に対する指文字の手の形状を，目視によりコード化規則1，2でコード化した例を示す。表4. 4にコード化規則1，2で，表4. 1の測定データを用い46個の手の形をコード化した結果を示す。



if bending angle $\leq 45^\circ$ then status=伸(H)
if bending angle $> 45^\circ$ then status=屈(B)

(a) 指の曲げに関するコード化規則1






if w-component of $e_1 > 1/2$ then 向き=上(U)
if u-component of $e_2 > 1/2$ then 回転=掌(H)

(b) 手の向きに関するコード化規則2

図4. 6 コード化規則の例

表4.3 目視による手形状コード化の例

手形状	関節の屈伸状態 (Ga)										向き (Gb)	
	53	52	14	13	24	23	34	33	44	43	uvw	φ
指文字「ほ」	H	H	H	H	H	H	H	H	H	H	U	T
												
「う」	B	B	H	H	H	H	B	B	B	B	U	H
												
「す」	H	H	H	H	H	H	B	B	B	B	D	T
												

(H/B:関節が伸/屈 U/D:手が上/下向き H/T:掌/甲を向けている状態)

表4.4では、関節を、屈伸状態に不定コードが少ない順、第1主成分の寄与率の多い順に並べた。表4.4は

1. 不定コードに対する関節は、H又はBの状態に対応する指の曲げ角をとりうる。「め」の指文字に対する手形状は、親指と人指し指の指先が接触し、丸い輪の形をしている。「め」に対する手形状と、すべての指が伸びた「て」に対する手形状とが識別できないなど、屈伸状態より指先の接触状態に関係する関節に不定コードが多い。
2. 46個の手形状は、指の屈伸状態(Ga)のコード化規則1により、一つの指文字に対する9個の手形状と、複数の指文字に対する8個のグループに分類される。
3. 屈伸状態(Ga)とコード化規則2による向きの情報(Gb)の12個のパラメータからは、一つの指文字に対する23個の手形状と、複数の指文字に対するAからKまでの11個の手形状の計34個のグループに識別できる。

事を示している。

表4.4 コード化規則による識別実験結果

	関節の屈伸状態 (Ga)										向き (Gb)		グループ
	33	43	53	23	24	52	44	13	34	14	uvw	φ	
め	H	H	H	H	H	-	H	-	H	-	U	H	A
て	H	H	H	H	H	H	H	H	H	H	U	H	1
ね	H	H	H	H	H	H	H	H	H	-	D	T	2
ほ	H	H	H	H	H	H	H	H	-	H	U	T	3
く	H	H	H	H	B	H	H	H	B	-	L	T	4
き	H	H	H	H	B	H	H	H	-	-	U	H	B
こ	H	H	H	H	B	H	-	H	-	B	U	H	5
つ	H	H	H	B	B	-	H	-	B	-	U	H	6
け	H	H	B	H	H	H	H	H	H	-	U	H	7
よ	H	H	B	H	B	H	H	H	B	-	L	T	8
わ	H	B	H	H	H	H	B	H	H	H	U	H	9
ゆ	H	B	B	H	H	H	B	H	H	H	U	T	10
ま	H	B	B	H	H	H	B	H	H	-	D	T	11
み	H	B	B	H	H	H	B	H	B	B	L	T	12
へ	B	H	H	B	B	H	B	B	B	B	D	T	13
ひ	B	H	H	B	B	H	B	B	-	B	U	H	14
ち	B	H	H	B	B	-	H	-	-	B	U	H	C
い	B	H	H	B	B	B	H	B	B	-	U	H	15
は	B	B	H	H	H	H	B	H	B	H	DB	S	16
り	B	B	H	H	H	H	B	H	-	H	B	T	17
と	B	B	H	H	H	H	B	H	-	H	U	T	18
う	B	B	H	H	H	H	B	H	-	H	U	H	D
か	B	B	H	H	-	H	B	H	B	-	U	H	19
し	B	B	H	H	B	H	B	H	B	-	L	T	E
に	B	B	H	H	B	H	B	H	-	-	L	T	20
す	B	B	H	H	B	H	B	H	B	-	D	T	F
な	B	B	H	H	B	H	-	H	-	-	D	T	21
る	B	B	H	H	B	H	B	H	B	H	U	H	G
を	B	B	H	-	B	H	B	-	B	B	U	H	22
お	B	B	H	B	B	H	B	-	B	B	U	H	H
ぬ	B	B	H	B	B	H	B	-	B	H	U	H	
れ	B	B	H	B	B	H	B	H	B	H	U	H	23
も	B	B	H	B	B	H	B	H	B	-	L	T	I
む	B	B	H	B	B	H	B	H	B	-	L	T	24
ん	B	B	H	B	B	H	-	H	-	H	DB	T	25
ふ	B	B	H	B	B	H	B	H	B	-	D	T	J
の	B	B	H	B	B	H	-	H	-	H	D	T	26
そ	B	B	H	B	B	H	-	H	-	-	B	T	27
た	B	B	H	B	B	H	B	B	B	B	B	H	28
あ	B	B	H	B	B	H	B	B	-	B	U	H	29
ろ	B	B	H	B	H	H	B	-	B	H	U	H	K
ら	B	B	H	B	H	H	B	H	-	-	U	H	30
せ	B	B	H	H	B	B	-	B	-	-	U	H	31
ひ	B	B	H	B	B	B	-	H	-	H	U	H	32
さ	B	B	H	B	B	B	-	B	-	B	U	H	33
え	B	B	B	B	H	-	H	-	H	-	U	H	34
	0	0	0	1	1	4	8	8	18	21	不定コードの数		

(H/B:伸屈 -:不定 U/D:上下 R/L:左右 F/B:前後 H/T/S:掌甲横)

4. 2. 4 考察

主成分分析やクラスタ分析結果は、人差し指や中指は、他の指より特徴的な指であることを示している。コード化方法では、これらの指に不定コードの数が多い。この原因としては、

1. 中指の視覚的な屈伸状態は、中指の曲げ角だけではなく、他の薬指の屈伸状態により影響されること。
2. 「お、き、め」に対する手の形のように、屈伸状態より他の指の指先やてのひらとの接触で異なった手の形になる。

など、指の曲げのほかに、視覚的な特徴による手の形が指文字には多い事が考えられる。図4. 7に「め」に対応する手形状(a), 「て」に対応する手形状(d)と、人差し指の第1関節を曲げることで(a)から(d)に変化する途中の手形状(b), (c)を示す。人差し指の曲げに対するコードだけでは、(a)と(d)および(b), (c)に対応する手の形状を識別することは難しい。

これらの結果から、人間が日常生活に用いている手振りの視覚的な特徴を抽出するには、指の曲げ角や向きのほかに、親指と他の指の接触状態などを測定することが必要と思われる。

4. 3 第4章のまとめ

本章では、手形状入力装置で測定した日本語手話の46個の指文字に対する手形状の測定データを用い、統計的な解析結果とコード化方法による手形状の識別実験結果について述べた。統計的な解析結果からは、46個の手形状が4つの主成分で説明できること、4つのクラスタに分類できることが明らかになった。コード化方法では、46個から約30個の手形状が識別できることを示した。

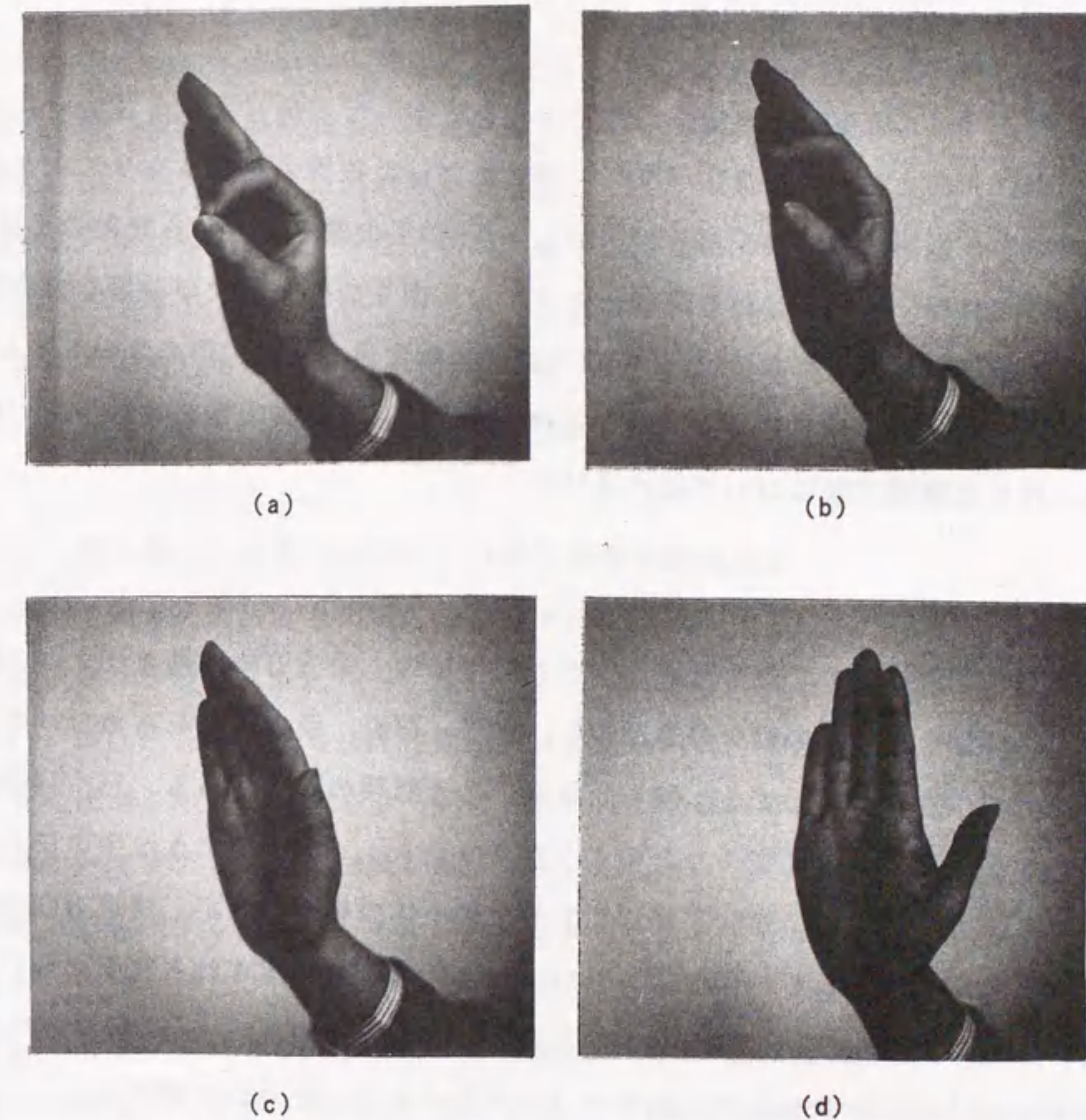


図 4.7 指文字「め」から「て」までの手形状

第5章 手振り認識方法

計算機と人間とのインタフェースに、動きを伴った複数の手振りをリアルタイムに認識し、利用した例は少ない^{39), 46)}。一連の手の動きの中から複数の手振りを認識するには、手振り自体の認識と、1つの手振りの終了と次の手振りの開始を識別することが必要になる³⁵⁾。本章では、4章で述べた手の形のコードをもとにした手の動きのコード化方法、動きを伴った手振りの認識方法および手話の指文字を用いた認識実験について述べる⁷²⁾。

5.1 手の動きのコード化方法

アメリカ手話において動きを表わすSig記号には、垂直方向の動き（上への動き、下への動き、上下の動きの3種類）、水平、前後、手のひらをおおむけにする、下向きにする、回転する、腕を組む等の24種類の記号がある。日本語の指文字の手振りには、指文字「も」のような指の動き、「り」のような手のひらの動きと、濁音や半濁音の指文字に対応する手の全体の垂直、水平、前後方向の動きが含まれている。ある時間内に、Data GloveTMにより測定される関節の曲げ角度と、3SPACE TRACKERTMにより測定される手の3次元位置(u, v, w)と、手の向きに対応するオイラー角から手の動きをコード化する方法について述べる。

(コード化規則3: 手全体の動き表現)

ある時間内の手の全体の動きを、基準座標系(U-V-W)に対する手の3次元位置座標の増減(+, -, 振動, 静止)で示す。

手全体の動き表現: $Gc = (R_u, R_v, R_w)$

ここで $R_i = +, -, 0$, スペース

(コード化規則4: 手の動き表現)

ある時間内の手の動きを、関節の曲げ角度、手の向きに対する測定値の増減(+, -, 振動, 静止)で示す。例えば、手を結ぶ動きは、各関節の曲げ角の増加として表現する。

手の動き表現: $Gd = (S_{14}, S_{13}, \dots, S_{52}, S_u, S_v, S_r)$

ここで $S_i = +, -, 0$, スペースで、 S_{14} から S_{52} の10個は関節の曲げ角度、 S_u, S_v は前後、左右の手の傾き、 S_r は手

の回転に対応する。

対応する清音の手の形のまま、手全体を右に動かす濁音は、清音に対する手の形を表す(Ga, Gb)と動きに対応する $Gc = (, +,)$, $Gd = (, \dots,)$ によって表現する。半濁音、促音や「を」に対する手の動きは、対応する清音や「お」の手の形と手全体の上下 $Gc = (+, ,)$, 前後の動き $Gc = (, , +)$ で表現する。親指と人指し指をつまむ動きに対する「も」は、動き開始時の手の形と人指し指の動きに対する $Gd = (, +, \dots,)$ で表現する。

5.2 手形状入力装置を用いた手の形・動きの認識方法

手の形や動きを認識するには、ある時間周期毎に、手形状入力装置で測定される関節の曲げデータ($X_{14}, X_{13}, X_{24}, \dots, X_{43}, X_{53}, X_{52}$)、手の向きデータ($X_{A21}, X_{ELE}, X_{ROL}$)と3次元位置情報(X_u, X_v, X_w)の測定データ $x(t)$ を、コード化規則1から4によりコード化された形で登録されている手形状パターンとパターンマッチングする必要がある。

手振りを認識する方法として、現在の手振りに対する測定データを、しきい値によりコード化したコードと登録コードとをマッチングする方法が考えられる。同じ手の形をしていても、手形状入力装置からのデータは時間とともに変化している。単純なコードのマッチング方法では、このように常に測定値が変化するデータからは、手の形の認識に失敗したり、複数の手の形を認識することになる。連続した動きの中で、登録されている手の形・動きコード(Ga, Gb, Gc, Gd)に対する各指の屈伸のマッチング度合を評価し、マッチングを行う手の形・動きの認識方法を示す。

屈伸や向きの度合を、関節の曲げ角度、向きベクトルの成分から各状態へのメンバーシップ関数(測定値 $\rightarrow [0, 1]$)で特徴づけられるファジィ集合で表現する。各メンバーシップ関数値の代数和で、登録してある手の形・動きコードと現在の手形状との一致の度合を表現し、手の形・動きの認識を行う(図5.1)。手振り認識の処理が、手形状入力装置の測定周期時間内で終了するように、照合する手の形状の多い手の形マッチングは、最初に類似した複数の手形状から構成されるグループの選択を行い、照合したグループ内で手形状の照合をする2段の構成で高速化を図った。

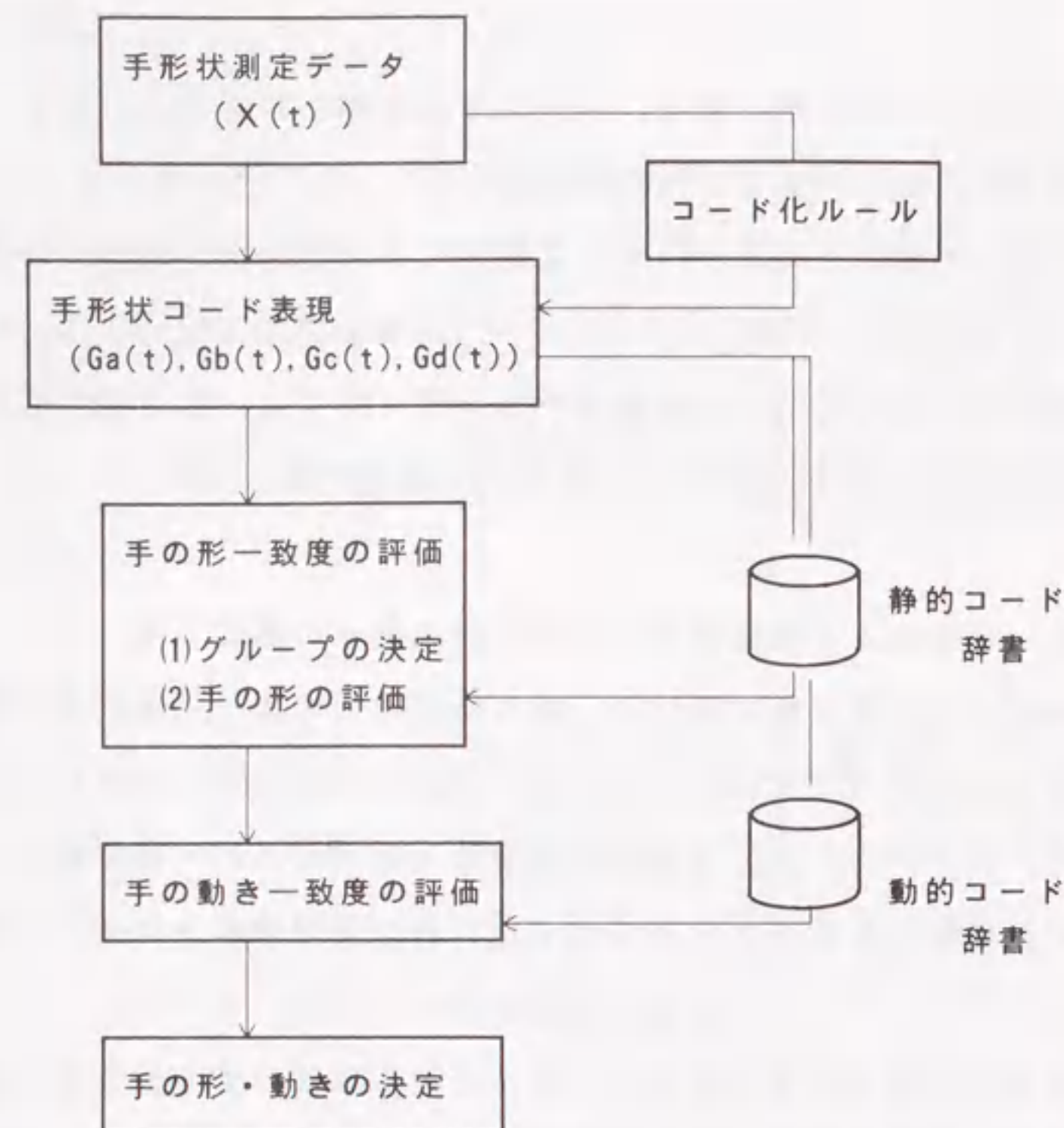


図5.1 手の形・動きの認識フロー

ステップ1)：手の形一致度の計算。

コード化された、指の屈折状態 $P(P_{14}(t), \dots, P_{52}(t))$ からグループの決定をする。決定されたグループに属する手の形状のパターンとの一致度を、現時点での指の曲げ角 $(X_{14}(t), \dots, X_{52}(t))$ と、手のオイラー角 $(X_{A21}(t), X_{ELE}(t), X_{ROL}(t))$ に対するメンバーシップ関数値の代数和で計算する。

但し、「不定」コードをもつパラメータに対しては、測定値に幅があるため計算の対象としない。

ステップ2)：手の動き一致度の計算。

一定時間内の測定データ $x(t)$ ($T_{i-1} < t \leq T_n$) の変化に対応する手の動き、手全体の動きコード $(R_u(t), R_v(t), R_w(t), S_{14}(t), \dots, S_r(t))$ ：ここで $T_{i-n} < t \leq T_n$ で T_n が現時点) の出現頻度を T_n で

の一致度とする。

ステップ3)：手の形・動きの決定。

手の動き開始時の形、動きの対で登録パターンと照合し、両方の一致度が正のパターンを現在の手の形・動きとする。

5.3 手振り認識実験

4章の識別実験と同一被験者を用い、同じ実験システムで以下の手形状認識実験を行なった。

- ・手振りの区切りを与えた条件下での手の形認識。
- ・手振りの区切りを与えない条件下での手の形・動き認識。

複数の認識実験結果を比較するには、同じ手振りデータを用いて評価する必要がある。同じデータで評価するために、「あ」から「ぼ」までの濁音、半濁音を含んだ71個の手の形・動きデータ（関節データ、向きデータ、3次元位置データ）をあらかじめ5回測定し、ファイルに測定時間の情報とともに記録する。認識実験においては、この測定データファイルをもとに、被験者の手の動きデータを発生させた。

表5.1、表5.2に認識実験に用いた清音に対する手の形辞書と、手の動き辞書を示す。手の形辞書は、指の曲げ状態 G_a をもとにした10グループ ($a \sim i, x$) と、清音に対する所属するグループ名と G_a と手の向き G_b から構成されるコードから構成されている。グループに対するコードの「*」は、グループわけに用いない関節を示す。手の動き辞書は、動き開始時の手の形に対応する清音のコードと動きコード G_c, G_d から構成される。

指の屈伸の度合は、連続して状態が変化するメンバーシップ関数を用い評価した。設定するメンバーシップ関数の形の認識率への影響を調べる為に、屈伸度合のメンバーシップ関数として2種類の異なる関数を用いた。図5.2に用いた指の屈伸度合に対するメンバーシップ関数を示す。

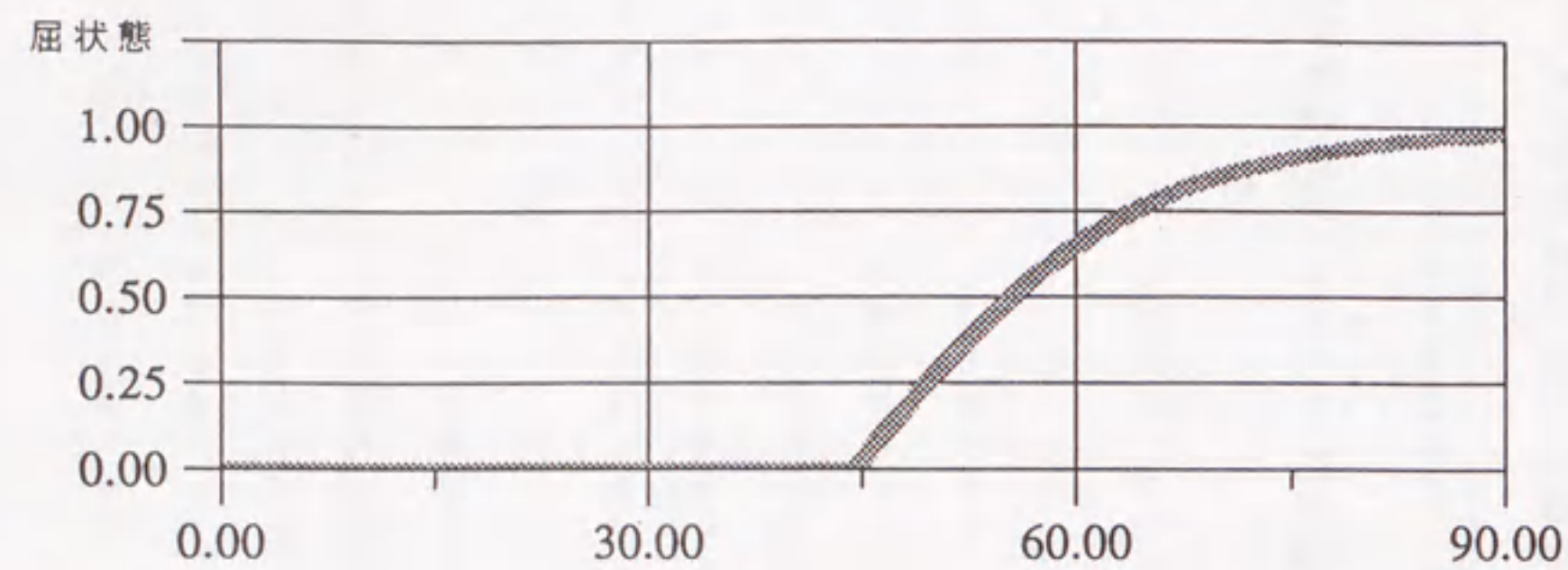
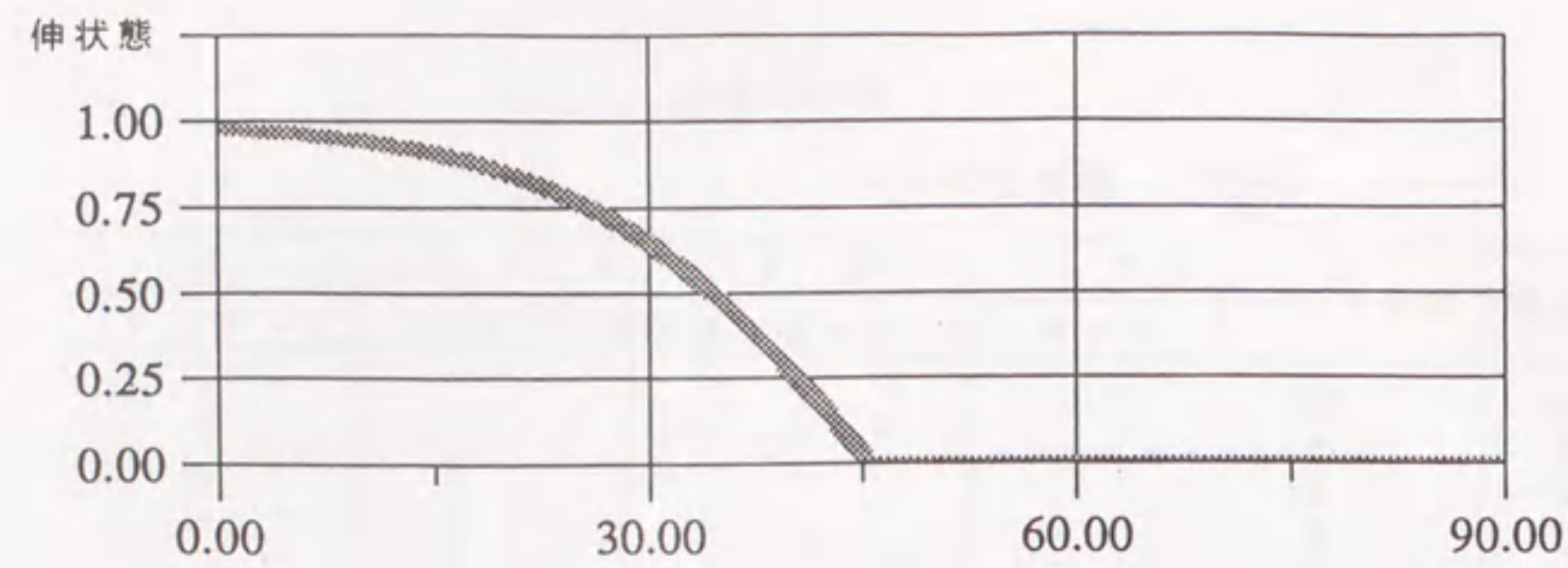
また、ステップ2の手の動きに関しては、指の関節角の変化が 15° 以下、手の向きの変化は方向余弦が0.25以下、3次元位置の変化が5cm未満の場合は、手の動きはなく、静止しているとした。

表5.1 清音の辞書

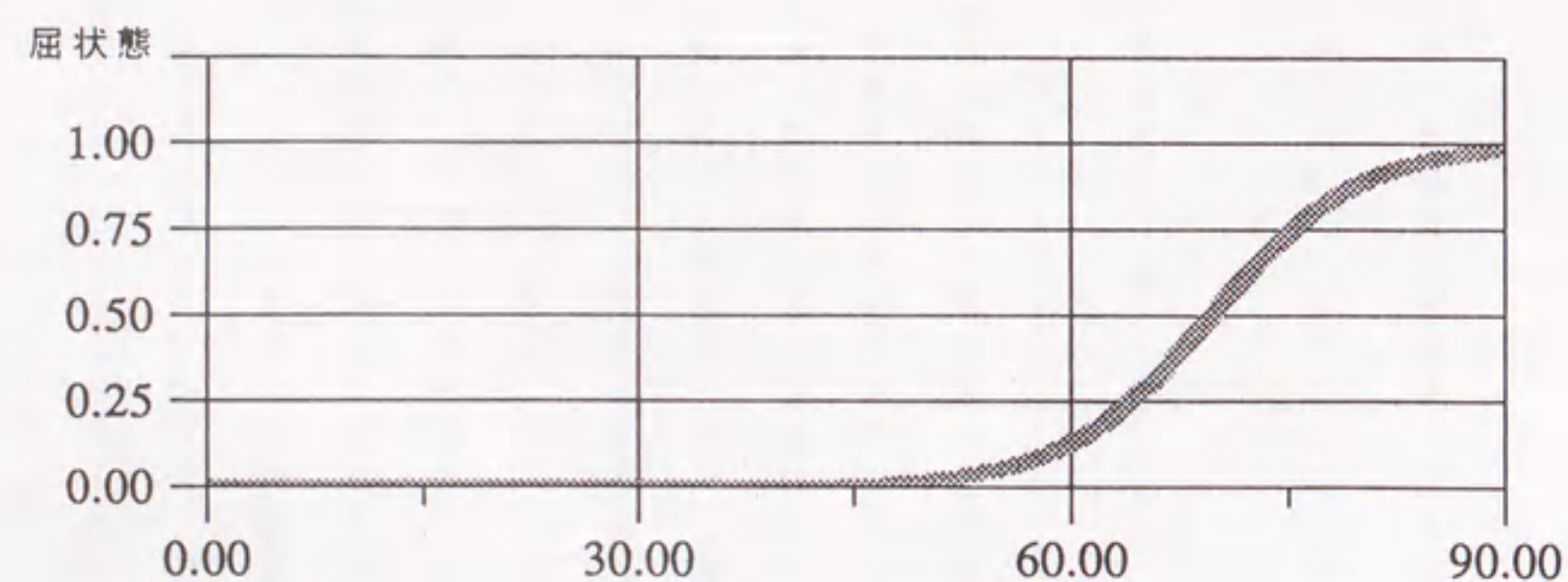
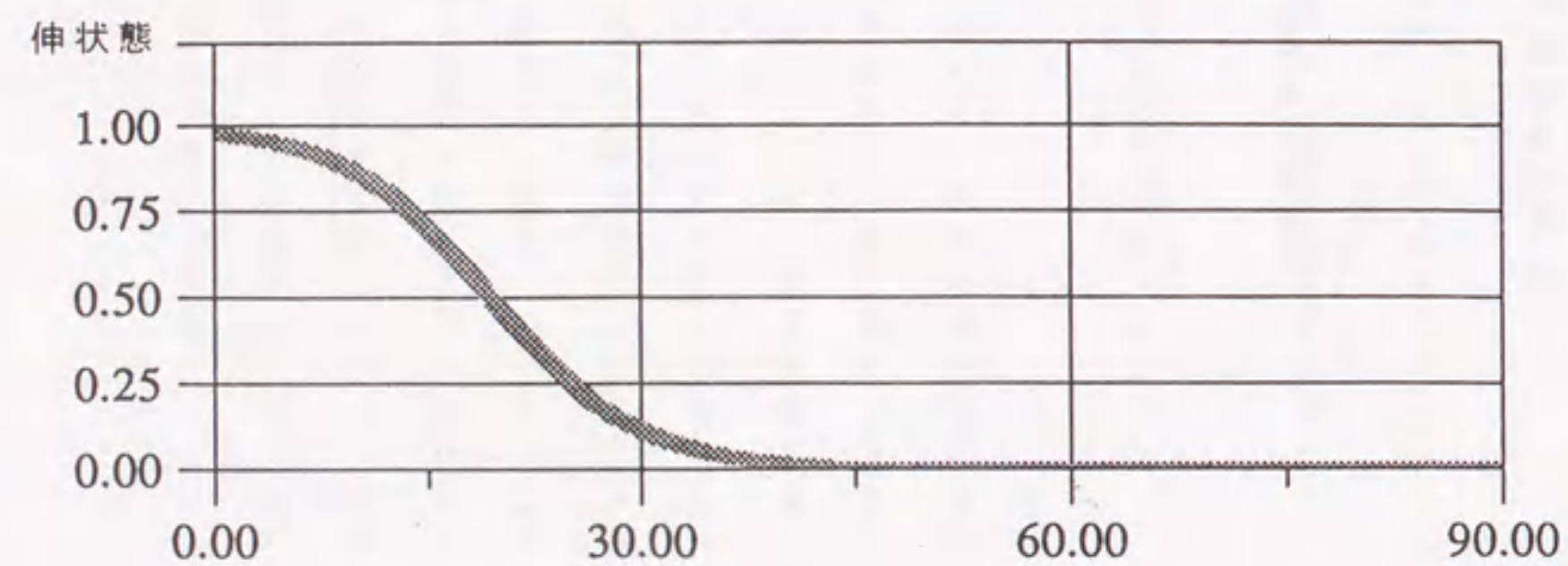
group-id		Ga										Gb	
		P ₁₄	P ₁₃	P ₂₄	P ₂₃	P ₃₄	P ₃₃	P ₄₄	P ₄₃	P ₅₃	P ₅₂	Q ₁	Q ₂
1	a	*	*	H	H	*	H	H	H	H	*		
2	b	*	H	B	H	*	H	*	H	H	H		
3	c	*	H	H	H	H	H	B	B	B	H		
4	d	B	B	B	B	*	B	B	H	H	H		
5	e	*	*	B	B	*	B	H	H	H	*		
6	f	*	H	H	H	*	B	B	B	H	H		
7	g	*	*	B	H	*	B	*	B	H	H		
8	h	*	*	B	B	*	B	*	B	H	H		
9	i	*	*	H	B	*	B	B	B	H	H		
10	x	*	*	*	*	*	*	*	*	*	*		
コード		Ga										Gb	指文字
1	a	-	-	H	H	H	H	H	H	H	-	U	め
2	a	H	H	H	H	H	H	H	H	H	H	U	て
3	a	-	H	H	H	H	H	H	H	H	H	D	ね
4	a	H	H	H	H	-	H	H	H	H	H	U	ほ
5	b	-	H	B	H	B	H	H	H	H	H	L	く
6	b	-	H	B	H	-	H	H	H	H	H	U	き
7	b	B	H	B	H	-	H	-	H	H	H	U	こ
8	x	-	-	B	B	B	H	H	H	H	-	U	ご
9	x	-	H	H	H	H	H	H	H	B	H	U	ざ
10	x	-	H	B	H	B	H	H	B	B	H	L	じ
11	x	H	H	H	H	H	H	B	B	B	H	U	ず
12	c	H	H	H	H	H	H	B	B	B	H	U	ぜ
13	c	-	H	H	H	H	H	B	B	B	H	D	だ
14	x	B	H	H	H	B	H	B	B	B	H	L	ち
15	d	B	B	B	B	B	B	B	H	H	H	D	づ
16	d	B	B	B	B	-	B	B	H	H	H	U	っ
17	e	B	-	B	B	-	B	H	H	H	-	U	ど
18	e	-	B	B	B	B	B	H	H	H	B	U	の
19	f	H	H	H	H	B	B	B	B	H	H	D	ば
20	f	H	H	H	H	-	B	B	B	H	H	U	び
21	f	H	H	H	H	-	B	B	B	H	H	U	ぶ
22	f	H	H	H	H	-	B	B	B	H	H	U	べ
23	g	-	-	-	H	B	B	B	B	H	H	U	へ
24	g	-	-	H	B	H	B	B	B	H	H	L	ぼ
25	g	-	-	H	B	H	-	B	B	H	H	L	も
26	g	-	-	H	B	H	B	B	B	H	H	D	り
27	g	-	-	H	B	H	-	B	-	H	H	D	を
28	g	H	H	B	H	B	B	B	B	H	H	U	ぬ
29	g	B	-	B	-	B	B	B	B	H	H	U	お
29	h	B	-	B	-	B	B	B	B	H	H	U	れ
30	h	H	-	B	B	B	B	B	B	H	H	U	も
31	h	B	-	B	B	B	B	B	B	H	H	U	む
32	h	H	-	B	B	B	B	B	B	H	H	U	ん
33	h	-	H	B	B	B	B	B	B	H	H	L	ふ
34	h	-	H	B	B	B	B	B	B	H	H	L	の
35	h	-	H	B	B	B	-	B	-	H	H	D	そ
36	h	-	H	B	B	B	B	B	B	H	H	D	た
37	h	H	-	B	B	-	B	-	B	H	H	D	あ
38	h	-	H	B	B	-	B	-	B	H	H	D	ろ
39	h	B	B	B	B	B	B	B	B	H	H	F	せ
40	h	B	B	B	B	-	B	B	B	H	H	F	ひ
41	i	H	-	H	B	B	B	B	B	H	H	U	さ
42	i	-	-	H	B	-	B	B	B	H	H	U	え
43	x	-	H	B	B	-	B	-	B	H	B	U	
44	x	H	H	B	B	-	B	-	B	H	B	U	
45	x	B	B	B	B	-	B	-	B	H	B	U	
46	x	-	-	H	B	H	B	H	B	B	-	U	

表5.2 動きのコード

指文字	清音のコード	Gd										Gc		
		S ₁₄	S ₁₃	S ₂₄	S ₂₃	S ₃₄	S ₃₃	S ₄₃	S ₅₃	S ₅₂		S _u	S _v	S _r
がぎぐげござじずぜぞだちづっでどのばびぶべぼもりをん	23												+	
	6												+	
	5												+	
	9												+	
	7												+	
	45												+	
	24												+	
	26												+	
	43												+	
	38												+	
	39												+	
	17												+	
	8												+	
	8													+
	2												+	
	21												+	
	37											-	+	
	19											+	+	
	19											+	+	
	44											+	+	
	44											+	+	
	36											+	+	
	36											+	+	
	15											+	+	
	15											+	+	
	4											+	+	
	4											+	+	
	33										+			
	20											-		
	29													+
	35											0		



(1) メンバースhip関数 1



(2) メンバースhip関数 2

図5.2 屈伸状態のメンバースhip関数

5. 3. 1 手振りの区切りを与えた条件下での手の形認識実験

実験1) 手の形・向きの情報(Ga, Gb)とステップ1の認識アルゴリズムで, 46個の静音に対する手の形の認識実験を行った. 4章の識別実験結果からも明らかのように, 46個の手形状に対する手のコードには同じコードをもつ手形状も含まれている.

本実験では, 被験者がある一つの指文字を示した時点を外部から信号として入力し, 認識処理を開始し正しい手の形を認識しているか否かで手振り認識率を評価した. 46個の手形状に対し, 5回とも正しく対応する指文字が単一に認識できた手形状は, 関数1では18手形状, 関数2では24手形状であった. また, 5回のうちに他の手形状を含んで複数認識した手形状は, 関数1では7手形状, 関数2では10手形状であった(表5.3).

表5.3 手振り認識実験1)の結果

	メンバースhip関数1	メンバースhip関数2
A: 5試行とも 単一認識	あ い え か そ た ち つ て ん ね ふ ほ み む よ れ わ (18)	あ い え け こ さ し す せ そ た ち つ て な ぬ ね ふ へ ほ や よ る れ (24)
B: 他の形も併 せて認識	き け こ さ せ ひ も (7)	お か き く に ひ み む も わ (10)

() 内の数字は認識した手形状の数

5. 3. 2 手振りの区切りを与えない条件下での手の形と動き認識実験

実験2) 手の形・向きの情報(Ga, Gb)とステップ1の認識アルゴリズムで、46個の静音の連続した手の形の認識実験を行った。

本実験では、被験者が「あ」から「ん」までの一連の指文字に対する手振りを行う。認識した手形状が、被験者の示している指文字に等しいか否かで手振り認識率を評価した。実験1とは、手振り間の区切りの認識も必要とされる事が異なる。4. 2. 3の識別結果にもとづき、46個の手形状のコードのうち、単一の指文字に対応する23個の手形状と、複数の指文字に対応する11個の手形状の計34個の手の形に対する認識率を評価した。

コード表現として独立な34個の手の形に対して、5回とも正しく指文字が単一に認識できた手形状は、関数1では8手形状、関数2では11手形状であった。また、5回のうちに他の手形状を含んで認識した手形状は、関数1では9手形状、関数2では10手形状であった(表5. 4)。

表5.4 手振り認識実験2)の結果

	メンバーシップ関数1	メンバーシップ関数2
	① 46個の手形状の内、単一の文字に対応する23個の手形状の認識結果	
A: 5試行とも 単一認識	え せ そ た つ ね ほ (7)	え せ そ た つ ね へ ほ (8)
B: 他の形も併 せて認識	あ け さ ひ み よ わ (7)	あ く け さ ひ み よ わ (8)
	② 46個の手形状の内、複数の文字に対応する11個の手形状の認識結果	
A: 全試行とも 単一認識	I (む、も) (1)	E (し、に) F (す、な) I (む、も) (3)
B: 他の形も併 せて認識	B (き、こ) C (い、ち) (2)	B (き、こ) C (い、ち) (2)

() 内の数字は認識した手形状の数

実験3) 手の形・向き情報に位置・動き情報(Gc, Gd)を加えた、ステップ3までの認識アルゴリズムで動きのある清音、濁音、半濁音を加えた一連の手の形・動きの認識実験を行った。

本実験では、被験者が「あ」から「ぼ」までの76個の手振りを行う。表4. 4では、同じ手形状コードを持っていたが、動き情報を考慮する事により、異なる手振りになる「の／ふ(J)」、を／る(G)」、も／む(I)」を加えた29個の手形状と、複数文字に対応する8個の手形状の計37個の独立な手の形・動きに対して評価を行った。

37個の手の形に対して、5回とも正しく指文字が単一に認識できた手形状は、関数1では7手形状、関数2では10手形状であった。また、5回のうちに他の手形状を含んで複数認識した手形状は、関数1では12手形状、関数2では12手形状であった(表5. 5)。

表5.5 手振り認識実験3)の結果

	メンバーシップ関数1	メンバーシップ関数2
	① 46個の手形状の内、単一の文字に対応する29個の手形状の認識結果	
A: 5試行とも 単一認識	え せ そ た つ ね ほ (7)	え せ そ た つ ね へ ほ (8)
B: 他の形も併 せて認識	あ け さ ひ み よ わ む も ふ (10)	あ く け さ よ わ る む も ふ (10)
	② 46個の手形状の内、複数の文字に対応する8個の手形状の認識結果	
A: 全試行とも 単一認識	(0)	E (し、に) F (す、な) (2)
B: 他の形も併 せて認識	B (き、こ) C (い、ち) (2)	B (き、こ) C (い、ち) (2)

() 内の数字は認識した手形状の数

5. 3. 3 手振りの区切りを与えない条件下での手の動き認識

「の、も、り、お、ん」の動きを伴った清音や濁音、半濁音などの手振りのうち、実験3の認識結果には、実験2でも認識された「も」しか含まれていない。手の動きに着目して、実験3における濁音、半濁音に対する認識率を評価する。

濁音、半濁音をもつか行、さ行、た行、は行の20文字の手形状のうち、実験3において5回とも単独に正しく認識された清音「せ、そ、た、つ、ほ」に対する手の動き認識結果を表5. 6に示す。半濁音における上への動きに対する認識率は、関数1、2とも80%で、濁音における右への動きの認識率は、「ぜ、ぞ、だ、づ、ぼ」の平均で56%、72%であった。

表5.6 手の動き認識の結果

	メンバーシップ関数1	メンバーシップ関数2
横の動き	ぜ (3/5)** ぞ (3/5) だ (2/5) づ (5/5) ぼ (1/5)	ぜ (4/5) ぞ (5/5) だ (2/5) づ (5/5) ぼ (2/5)
縦の動き	ぼ (4/5)	ぼ (4/5)

() 内の数字は正しく認識した回数／試行回数

5. 3. 4 考察

日本語手話における指文字に対する手形状を用いた認識実験から、

1. 指関節の曲げのメンバーシップ関数1と2の認識率に大差がないことから、認識率に対するメンバーシップ関数の影響は少ない。
2. 実験2と実験3における認識率に変化が少ないことから、一連の動きの中で手振り間の区切り認識を含め、複数手振りの認識が行える。
3. 市販の手形状入力装置を用いて、20個程度の手振りが動きの中で認識可能と考えられる。

が明らかになった。

5. 4 第5章のまとめ

本章では、4章での手形状のコード化に加え、時系列的な変化に伴う手の動きのコード化方法、一連の動きの中での手の形・動きのコードを用いた手振り認識方法を提案し、手話の指文字を用いた認識実験について述べた。認識実験結果から20個程度の手振りが認識できることを明らかにし、手振りをを用いた計算機インタフェースの可能性を示した。

1980年にBoltによりMITで開発されたPUT THAT THERE⁵⁾は、言葉と手振りをを用い「青い三角形をそこに移せ」と、複数の入力方法を用いた計算機インタフェースの有効性を示した。このように複数の方法から同時に入力されるマルチモーダルインタフェースにおいては、個々の指示内容を認識し、指示された内容をまとめて一つに解釈することが必要になる。本章では、曖昧さを持った言葉や手振りをを用い、対象を指示できるマルチモーダルインタフェースの地図案内システム (IMAGE: Illustrated map guidance system) の概要とその評価について述べる。

6.1 地図案内システムIMAGEの概要

言葉と手振りで計算機に指示をするPUT THAT THEREでは、言葉での指示—例えば、「青い三角形をそこ(there)に移せ」—に必ず含まれている言葉“there”を合図(voice button)に、手で指示されている対象を同定する。Boltは、このようなマルチモーダルインタフェースの特徴として、適切な入力方法の選択により一つの方法で無理な指示をする必要がない、一つの入力方法では不十分な情報であっても、複数の種類の情報を用いることにより十分で雑音のない情報が得られるなど、unburdening, information summation, redundancyの3つの特徴を挙げている⁴⁾。

IMAGEでは、大型スクリーンに表示されている地図の内容について、利用者が“there”のような特定の言葉によるvoice buttonを必要としない、よりunburdeningな形のインタフェースを検討した⁷⁾。言葉と手振りをを用いたインタフェースにおける指示方法として、以下の3つの形態が考えられる。

1) 言葉による指示:

一番右端にある建物の名前は?

ツイン21ビルの右にある建物の名前は?

2) 手による指示:

なに?

名前は?

(対象を指して)

3) 言葉と手による指示:

この右端にある建物の名前は? (一つの建物を指して)

ここからここへ行く方法は? (二つの建物を指して)

この辺りにある銀行を教えて欲しい。 (ある範囲を指して)

1) の言葉による指示では、建物の名前や「右端」などの、2章で扱った位置関係を手がかりに、地図内の対象を同定する。2) の手による指示では、ポインティング指示の手の形を、5章で述べた認識アルゴリズムで認識し、その時に指示している地図上の範囲の中から、指示対象を同定する。3) の言葉と手振りの両方を用いた指示では、手により指示された範囲内にあり、言葉の指示に含まれる関係を満たす対象を同定する。

IMAGEのシステム構成を図6.1に示す。言葉による指示は、音声認識装置によって認識され、構文解析プログラムPARSERにより解析される。手形状入力装置で手振りによる指示を行う。第5章で述べた手振り認識アルゴリズムを用いたプログラムHAND MOVEMENT DETECTIONが、指の曲げ角、手の向きと3次元位置データから手振りを認識する。プログラムIDENTIFICATIONが、PARSERとHAND MOVEMENT DETECTIONにより認識された言葉による指示と、手振りによる指示の両方の結果を統合する。統合した指示内容を満たす対象の情報を、地図データベースから検索する。そして、検索結果を音声で応答したり、指示内容に応じて表示画像を変更したりする。

地図画像データベースには、地図に含まれている対象の位置や大きさなどの画像情報と、その対象の実世界での意味を表現している意味情報の2種類の情報が含まれている(図6.2)。画像情報として、3章の画像データベースと同様に、地図に含まれている建物の名前、住所や居住者の情報などのテキスト情報と、表示されている画面上の座標位置、色、形、大きさなど画像位置情報の2種類の情報がある。図6.2の地図に対するテキスト情報と位置情報を図6.3、図6.4に示す。地図に含まれている対象の意味構造は、is-a関係で構成した意味ネットワークで表現した(図6.5)。is-a関係は、下位の対象が上位の対象の持つ属性を継承する包含関係である。同じ対象に対する意味情報と画像情報は、共通の識別子(t1,t2,...)で管理されている。

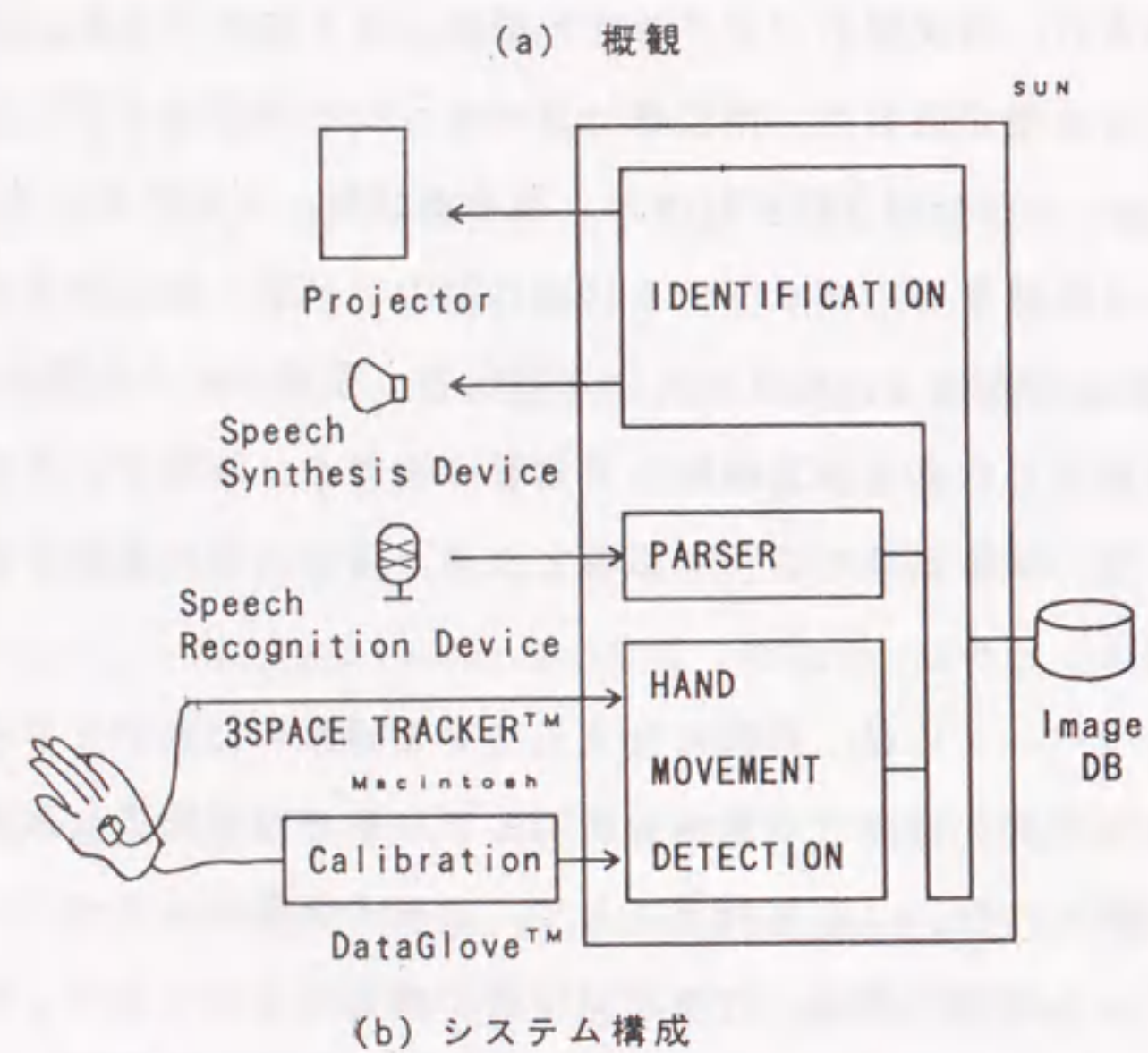
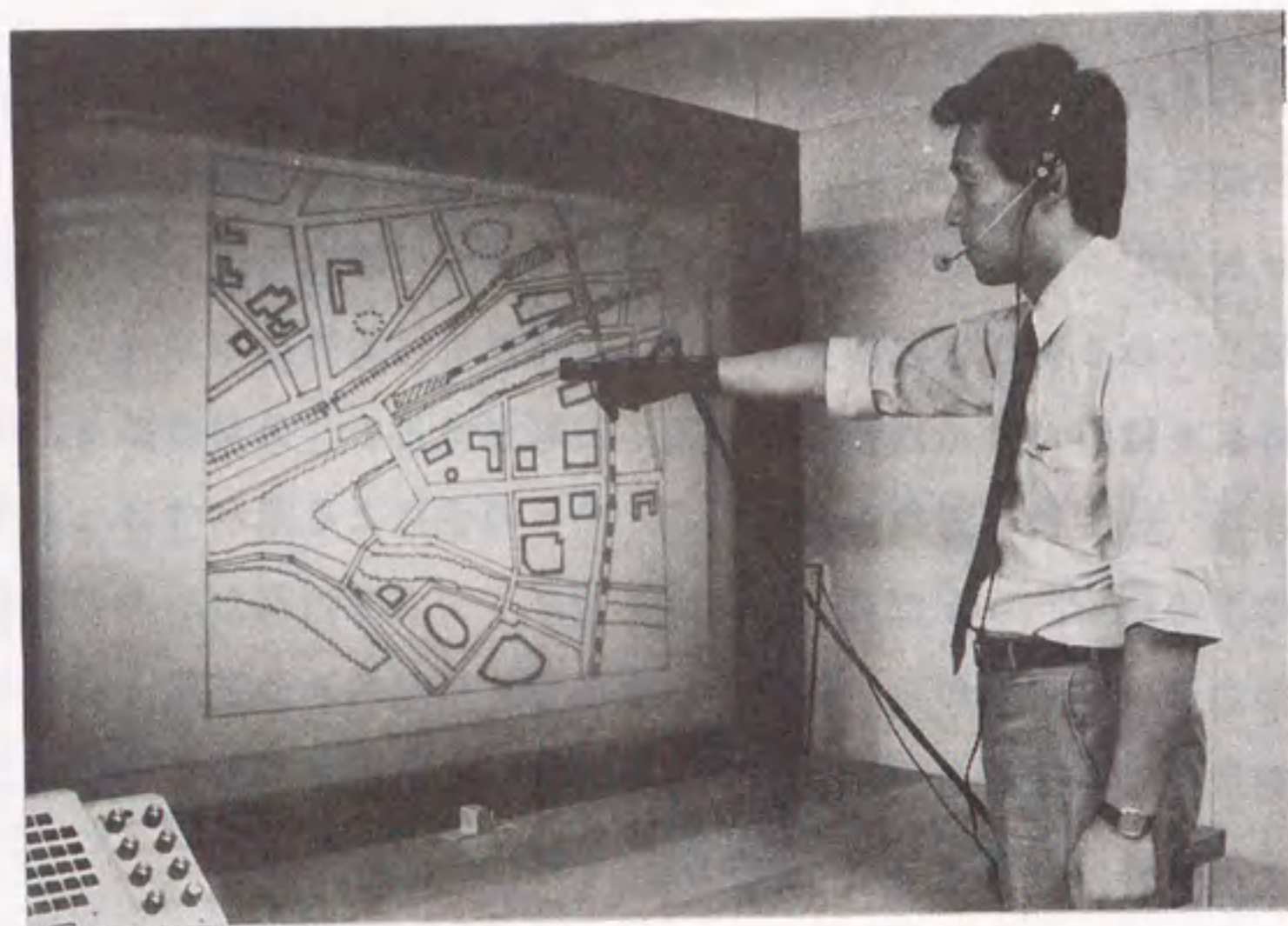
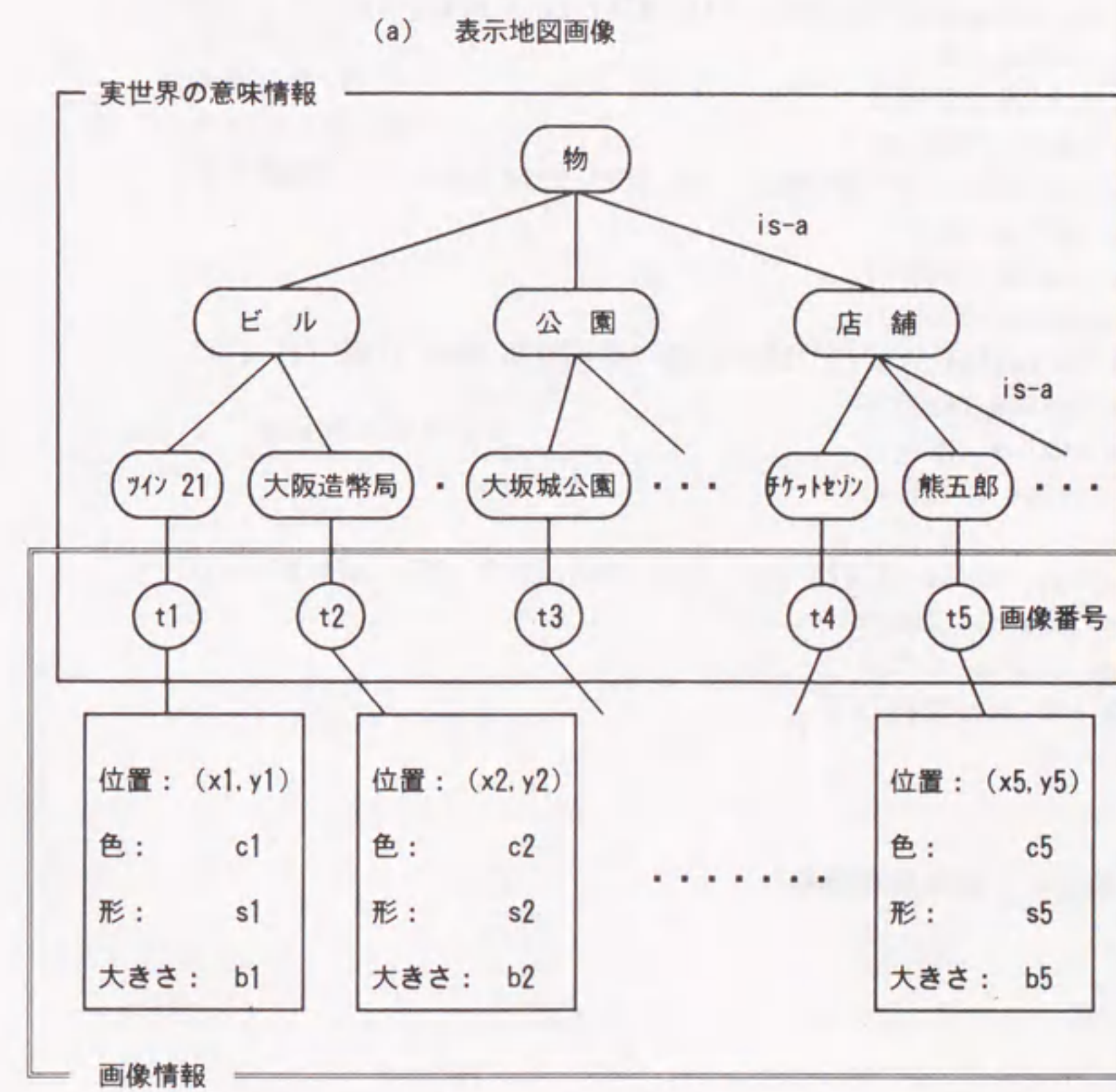
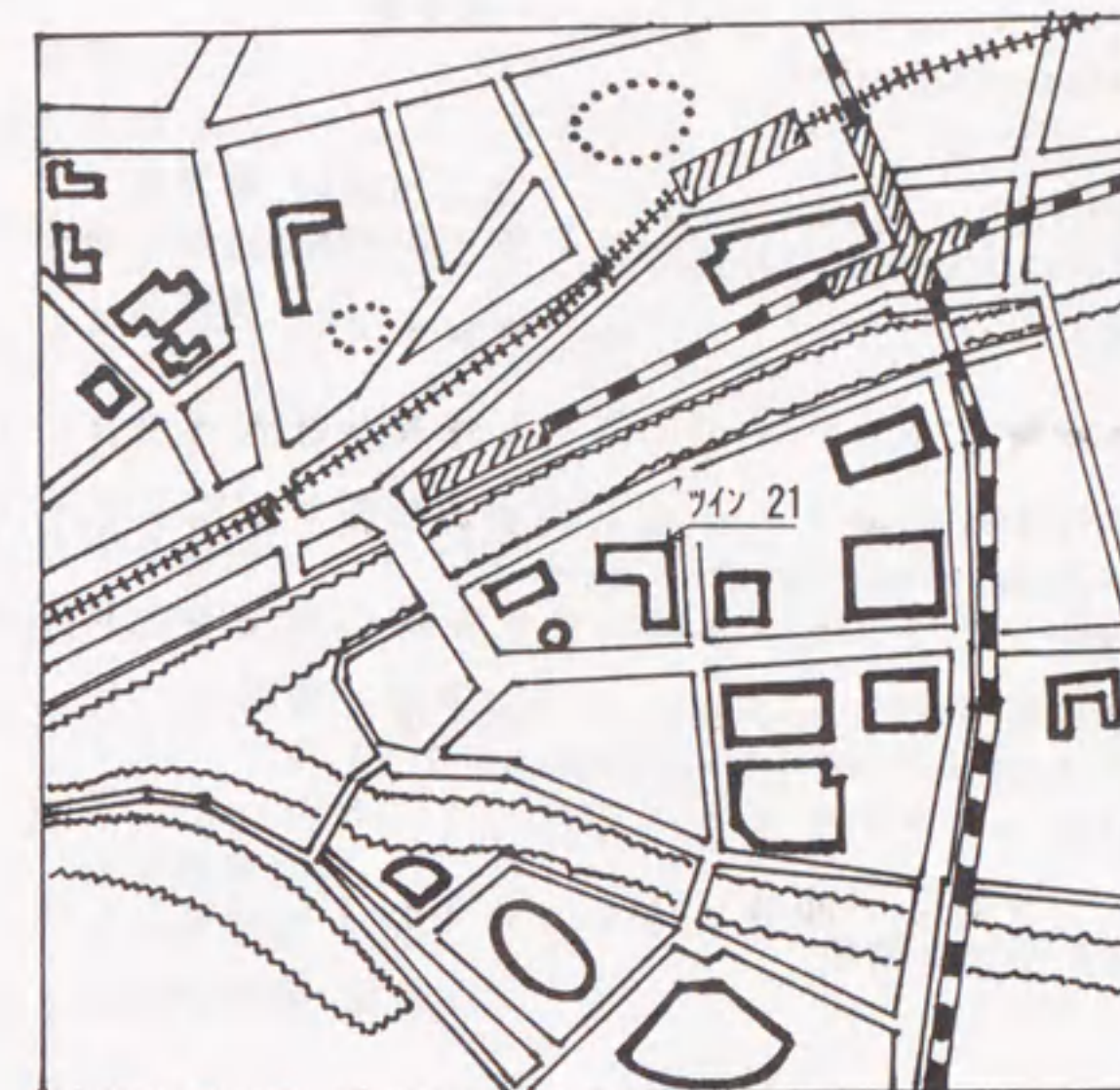


図6.1 IMAGEシステムの概観と構成



(b) 画像情報と実世界情報

図6.2 地図画像の表現


```

t1 name zouheikyoku ; 造幣局
t1 address oosakashikitakutenma1-1
t1 telephone 351-5361
t2 name takigawasyougakkou ; 滝川小学校
t2 address oosakashikitakutenma1-24-15
t2 telephone 351-1582
:
:
t102 name kougaikannshisennta- ; 公害監視センタ
t102 address oosakashihigashinarikunakamichi1-3
t102 telephone 972-1321
t103 name jeia-rumorinomiyaeki ; JR森ノ宮駅
t103 address oosakashihigashikumorinomiyachuuou1-1-33
t103 telephone 941-6124

```

図6.3 画像テキスト情報

```

(
(t4 (position (value (((189 851) (190 826) (219 824) (220 851)))))
(color (value (白))
(shape (value (四角))
(size (value (780)))))
(t5 (position (value (((192 800) (191 758) (215 760) (217 798)))))
(color (value (白))
(shape (value (四角))
(size (value (980)))))
(t6 (position (value (((229 755) (239 688) (276 688) (282 751)))))
(color (value (白))
(shape (value (四角))
(size (value (2921)))))
:
(t39 (position (value (((448 242) (491 182) (520 223) (474 270)))))
(color (value (白))
(shape (value (円))
(size (value (3006)))))
:
)

```

図6.4 画像位置情報

```

( 物
( 建物
( ビル
((雅叙園 t4)
(桜の宮会館 t5)
(太閤園 t6)
:
(ツイン21 t20)
:
(大阪城ホール t39)
(ホテルニューオオタニ t40)
)
)
( 駅
((京阪京橋駅 t11)
(JR京橋駅 t12)
(JR片町駅 t17)
)
))
( 公園
((東野田公園 t9)
(京橋公園 t10)
(大阪城公園 t38)
(野球場 t41)
)
:
)
)

```

図6.5 実世界の意味情報

6. 2 指示内容の統一表現⁷⁴⁾

言葉と手振りを用いた地図上の対象に関する問い合わせは、IMAGEシステムが管理している地図データベースから、問い合わせに含まれている関係を満たす対象の検索とみなせる。

IMAGEシステムでは、指示メディアとして言葉と手振りがある。指示方法として、固有名詞などを手がかりとして対象を指定する絶対指示、画面における空間的な位置関係で指定する空間指示、基準となる対象に対する相対的な位置関係で指定する相対指示の3つの方法がある。これらの方法で指示された内容は、関係の連言形式で以下の形で統一的に表現する(表6. 1)。

select objects Y such that
 $rel_1(x_1, x_1) \wedge \dots \wedge rel_n(x_n, x_n)$

言葉による指示は、構文解析プログラムで地図のもつ情報の関係に表現される。

例えば、「一番右端にある建物の名前は？」に対しては、

$\exists x \{ \text{右}(x, \text{Frame}) \wedge \text{attribute}(x, \text{建物}) \wedge \text{name}(x, A) \}$

表6.1 言葉と手振りによる指示

指示 方法 メディア	絶対指示	空間指示	相対指示
言葉	①ツイン21はどこですか?	②右下の建物は何ですか?	③ツイン21の右の建物は何ですか?
手振り	④これは何ですか?	———	———
言葉と手振り	⑤このツイン21の住所を知りたい。	⑥この下にある建物は何ですか?	⑦ツイン21の右にあるこの建物は何ですか?

- ① $\exists A$ 住所(ツイン21, A)
- ② $\exists A \exists x$ 右下(x, 枠) \wedge 名前(x, A)
- ③ $\exists A \exists x$ 右(ツイン21, x) \wedge 名前(x, A)
- ④ $\exists A \exists x$ hand(x, P) \wedge 名前(x, A)
- ⑤ $\exists A \exists x \exists y$ hand(x, P) \wedge 名前(x, ツイン21) \wedge 住所(ツイン21, A)
- ⑥ $\exists A \exists x$ hand(x, P) \wedge 下(x, 枠) \wedge 名前(x, A)
- ⑦ $\exists A \exists x$ hand(x, P) \wedge 右(x, ツイン21) \wedge 名前(x, A)

「ツイン21ビルの右にある建物の名前は？」に対しては、

$\exists x \exists y \{ \text{name}(x, \text{ツイン21}) \wedge \text{右}(y, x) \wedge \text{attribute}(y, \text{建物}) \wedge \text{name}(y, A) \}$

など、意味情報と位置情報の関係を満たす対象x, yの検索と、そのname属性の表示として表現される。

手によるポインティング動作は、地図に含まれている対象を手振りで限定する操作である。ポインティング動作を、指示の手の形を認識した時の手の向きと3次元位置で決められる円錐体と、ディスプレイ面で決定される指示範囲pに、対象xが含まれることを意味する関係hand(x, p)を用いて表現する。

手による指示「なに？」または「名前は？」は、関係hand(x, p)を用いて以下の形式で表現できる。

$\exists x \{ \text{hand}(x, p) \wedge \text{name}(x, A) \}$

言葉と手振りの両方を用いた指示も、言葉だけ、手振りだけの指示と同様に、関係の連言形式で統一的に表現できる。一つの建物を指しながら「この右端にある建物の名前は？」は、

$\exists x \{ \text{hand}(x, p) \wedge \text{右}(x, \text{枠}) \wedge \text{attribute}(x, \text{建物}) \wedge \text{name}(x, A) \}$

と、手振り動作—hand(x, p)—を1つ含む形で表現される。

二つの建物を、順次、指示して「ここからここへ行く方法は？」は、

$\exists x \exists y \{ \text{hand}(x, p_1) \wedge \text{hand}(y, p_2) \wedge \text{path}(x, y) \}$

と、2つの指示動作—hand(x, p₁), hand(y, p₂)—と対象x, yの位置座標から経路を求める関係—path(x, y)—の連現で表現される。

手を動かして、ある範囲を指しながらの指示「この辺りにある銀行を教えて欲しい。」は、

$\exists x \{ \text{hand}(x, p') \wedge \text{attribute}(x, \text{銀行}) \wedge \text{name}(x, A) \}$

と、複数の指示点で決められる指示範囲p'を用いた関係—hand(x, p')—で表現できる。

関係attributeは、意味ネットワークで表現された意味構造をもとに対象と結合する。attribute(x, 建物)は、「建物」のis-a関係の下位の「ビル」や「銀行」対象を含んだ、attribute(x, ビル)やattribute(x, 銀行)の関係も含んでいる。

IMAGEシステムにおける関係の連言形式の検索と、対象の名前や属性など関係の成立が明確な関係文献データベースの検索では、以下の点が異なっている。

1. 関係 rel_i には、位置関係や手振り指示に対応するhand関係のような曖昧な関係がある。
2. 2点の指示や範囲の指示に対する関係handには、指示された時間の異なる手振りの認識が含まれる。

曖昧な関係のうち位置関係に対しては、各関係の成立の度合をヒューリスティックに評価する。問い合わせに含まれる関係の代数和で、問い合わせ全体の成立度合を評価する。「ツイン21ビルの右にある建物の名前は？」の問い合わせに含まれる位置関係「右(y, ツイン21)」に対しては、関係「右」に付与されているヒューリスティックな評価関数でツイン21からみて、右にある対象の中から、評価値の高い対象を選択する。

言葉による指示と手振りによる指示内容の認識結果は、関係の連言で表現される。手振り動作と言葉の対応には、「この」のような指示語と一つの手振りが対応するケースと、「この辺り」のようにいくつかの手振りと対応するケースがある(図6.6)。「この右端にある建物の名前は？」は、言葉と手振りが1対1に対応する。「ここからここへ行く方法は？」は1対1の対応を複数含むケース。「この辺りにある銀行を教えて欲しい。」は、言葉と手振りが1対多に対応するケースである。以下の2方法で、いくつかの手振り認識結果を含む問い合わせを、関係の連言の形に統合する。

- 1) 「この」、「ここ」といった手振り動作と一緒に用いられる指示語を音声認識した時に、HAND MOVEMENT DETECTIONで認識されている手振りによる指示対象を確定する。
- 2) 音声による指示文全体の解析結果から、必要な手振りを記録されている手形状入力装置のデータからトップダウンに求める。6.1の3) 言葉と手による指示の3例では、言語解析結果に基づき、一つの指示点、二つの指示点および指示範囲に対応する手振りで指示された対象を、音声入力継続時間に対応する手の測定データから確定する。

試作した地図案内システムにおいては、音声による指示が、単語単位で認識された時は方法1)で、指示文単位で認識された時は方法2)で、手振りの認識結果と音

声の認識結果を結合する。手振りの認識結果と、言語の指示内容を統合した関係の連言を満足する対象を地図データベースから検索し、問い合わせされた内容を応答する(図6.7)。

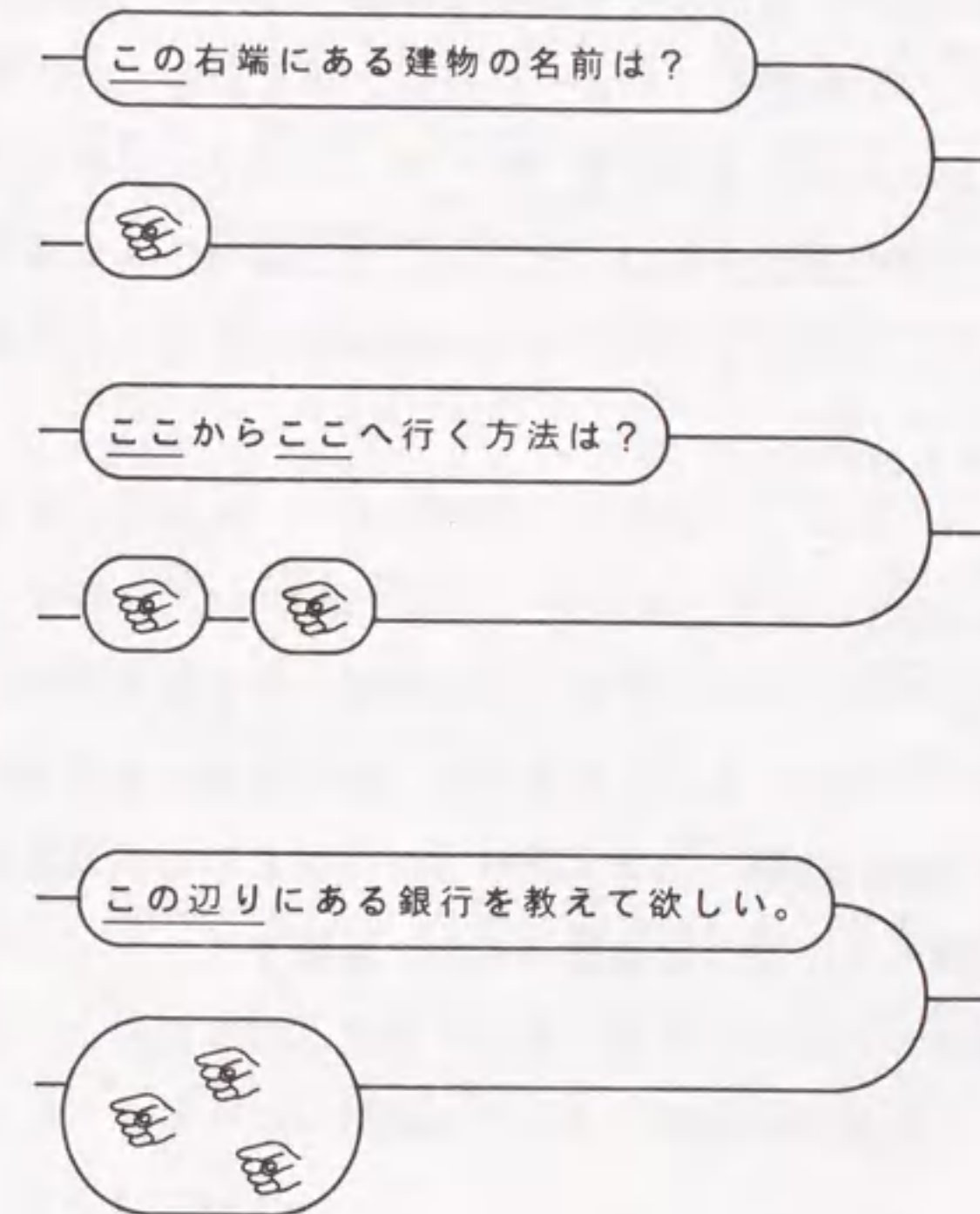


図6.6 言葉と手振り指示の統合

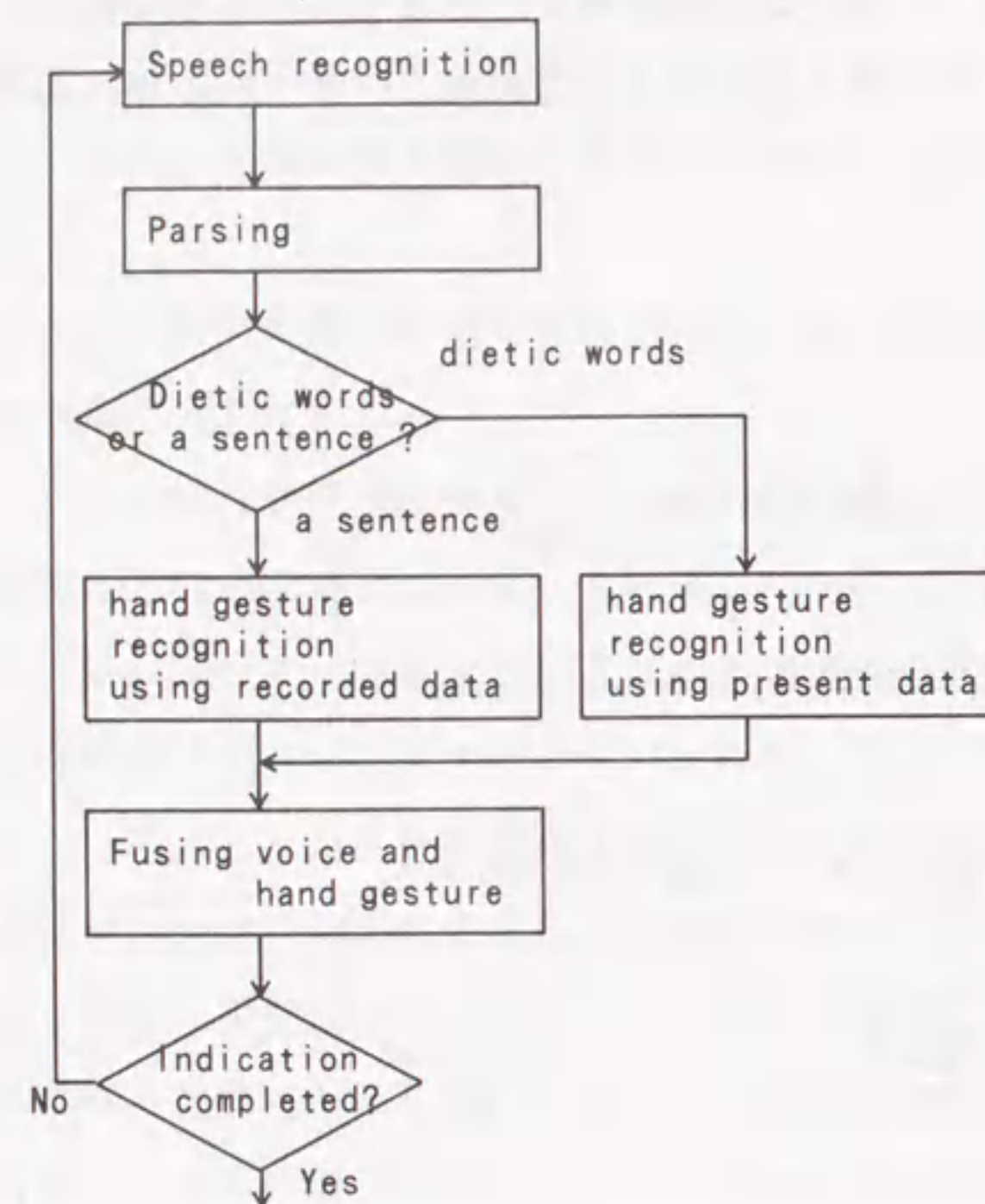


図6.7 情報統合のフロー

6.3 IMAGEシステムの評価

IMAGEシステムにおける音声と手振りによる対象の指示効率を決定する要因としては、音声認識率 (r_1)、手の形の認識率 (r_2)、手のポインティング精度 (r_3)、位置関係の同定率 (r_4) が挙げられる。各々の要因について、考察する。

音声認識率は、使用した認識装置に依存した値で、実用システムにおいても音声認識装置が広く使用されているので実用上問題はない。Data Glove™を用いた手の形の認識率についても、5.3.2の動きの中からも約20個の手形状が認識できる実験3)の結果と5.3.3の動きの認識率の考察から、数個のポインティング動作の手の形の認識率については動きも含めて80%近くの認識率が予想できる。

磁気発生器と磁気センサ (3SPACE TRACKER™) を用いたポインティング精度については、栄藤ら⁷⁵⁾の実験がある。彼らは、指先で指示方向を表現し肘を体側に付け指先は視野周辺とする体側指示、指先で指示方向を表現し指先を視野中心において方向確認する眼前指示、指先と眼の2点で指示方向を表現し指先は視線と一致させる視線指示の3方法にわけ測定した (図6.8)。IMAGEシステムにおけるポインティング姿勢である眼前指示に対して、観測される指示方向はある方向を中心に円錐をなし誤差角度が正規分布に従うと仮定すると、カーソルなどのフィードバックがない状態で3.9°程度の領域が70%程度の確率で指示できることを示している。

残された4番目の位置関係の同定率 (r_4) の評価について述べる。3章の画像データベース検索システムSPADEの評価と同様、指示に用いられる位置関係は、人間の主観を反映した関係である。そこで、図6.2(a)の地図に含まれる複数の対象を指示した文章を、人間とIMAGEシステムが同定した結果を比較することで、位置関係の同定率の評価を行った。

指示文の作成： 6人の被験者に、使用可能な単語、言い回しなどについて制約を課さずに、図6.2の地図に含まれる対象を指示させることで145個の指示文章を作成した。そのうち、構文解析プログラムPARSERで解析できる117個の指示文書を評価実験に用いた (表6.2)。

人間による同定実験1) 指示文作成の6人と異なる3人の被験者に、図6.2

の地図の建物に番号をつけた地図と117個の指示文を与え、指示対象を回答させた。回答は、「適当な対象が同定できない時はその旨を明記」,
「複数の対象が候補になった時はすべてを回答」の条件下で行った。

IMAGEによる同定実験2) 図6.2の地図に含まれる103個の建物に関して、建物の名前や位置などの画像情報と実世界の意味情報を入力し、人間と同じ指示文の同定を行った。

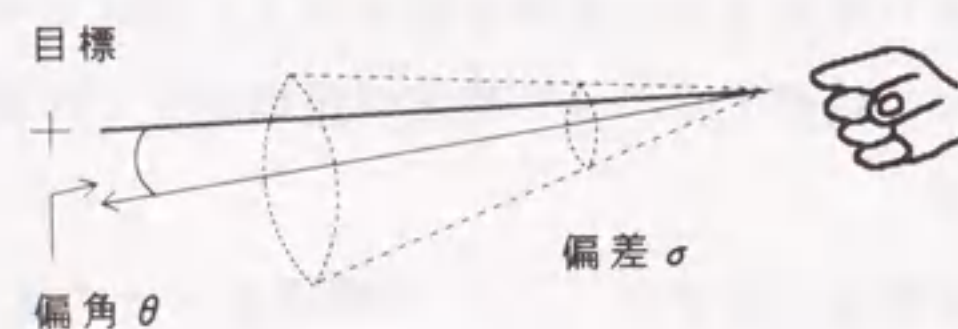
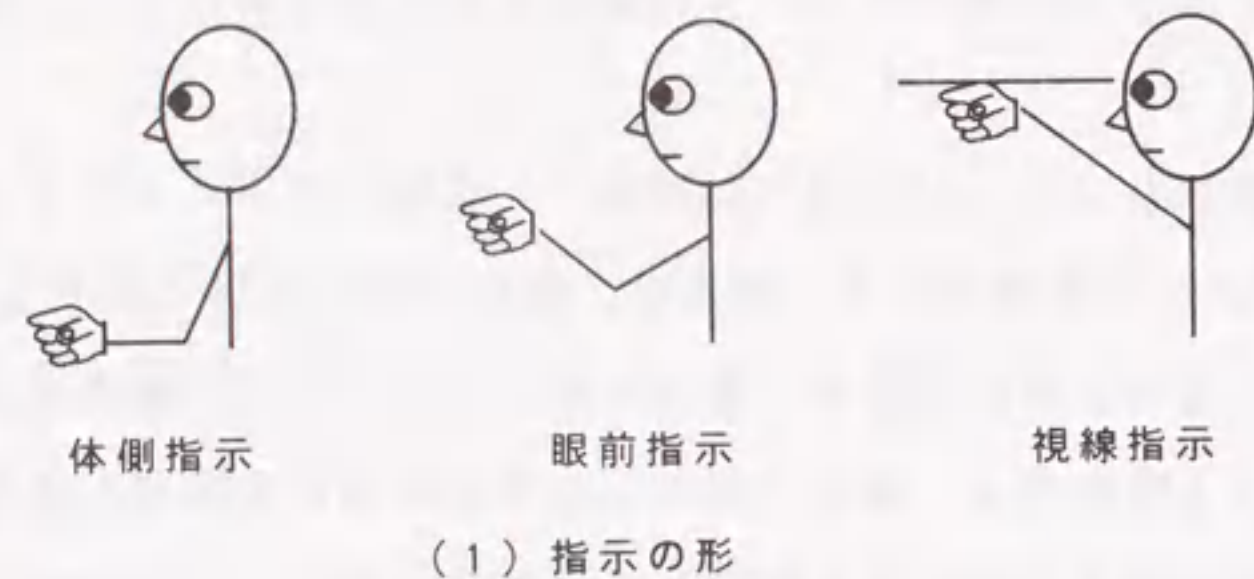


図6.8 指示特性評価実験

(“指の方向検出を用いたヒューマンインタフェース”⁷⁵⁾より)

表6.2 実験に用いた指示文章の例

分類	実験に用いた指示文章の例
空間指示	<ul style="list-style-type: none"> ・下の丸いもの ・下の方にあるだ円形 ・一番右の凹形の建物 ・南の方にあるだ円形 ・下の方にある楕円形のもの ・真中のカギ形のもの ・まん中にあるかぎ形 ・下の丸いもの ・中央にあるし字形 ・真中あたりのカギ形の建物 ・真中のカギ形のもの ・左上のカギ形のもの ・左上の四角 ・上の丸いもの ・真中上にある点線で囲まれただ円形 ・北にある点線で囲まれただ円形 ・一番東にある建物 ・右はしのコの字形のたてもの ・東にあるコの字形の建物 ・右端の凹形の建物 ・右はしのコの字形のもの ・右端にあるコの字形の建物 ・中央にある一番小さい正方形 ・真中にある小さな正方形 ・北西側にある長方形の建物 ・左端の四角の小さな建物 ・左上の小さな正方形のもの ・下の方にあるだ円形のもの ・下の方にあるだ円形 ・南の方にあるだ円形 ・左端の斜めの四角 ・右端の凹字形の建物
相対指示	<ul style="list-style-type: none"> ・南側の川の南にあるだ円 ・ツイン21の西に隣接する四角の建物 ・一番下にある扇形の隣のだ円 ・ツイン21の左側にある四角い建物 ・中央にある円形(のもの)の隣のし字形 ・線路の間のもの ・北側の川の南のカギ形のもの ・私鉄の駅の北側の点線のだ円 ・私鉄の駅の北に隣接する点線で囲まれた部分 ・太閤園の左上の建物 ・川の上にある建物 ・私鉄の駅とJRの駅のもの ・私鉄の駅の北側のもの ・国鉄の線路の北側の建物 ・真中の十字路からみて東南のもの ・駅と駅の間にあるもの ・中央にあるし字形のものの右隣りの正方形のたてもの下にある長方形の建物 ・ツイン21と道をへだてて東向いのもの ・JRの駅と私鉄の駅のもの ・真中の十字路の右上のもの ・中央付近の東西の通りの北側の小さな正方形のもの ・真中の交差点の右下のもの ・一番下の扇形の対岸のもの ・右上のJRの駅の近くにある建物 ・川の下にある丸いもの ・真中のカギ形の建物と交差点の対角にある四角の建物 ・Hotel ニューオータニの斜め向かいにあるだ円の建物 ・ツイン21の西側の四角 ・南側の川の北に面したもの ・私鉄の駅と隣接しているJRの駅の近くにある建物 ・川にかかっている橋のたもとにある建物 ・京阪京橋駅の北側の公園
空間指示 + 相対指示	<ul style="list-style-type: none"> ・右のへりにあってJRの東側の建物 ・JRと川ではさまれたブロックの中の右上のブロックでまん中の交差点に面している建物 ・川の中洲でカギ形の建物の右に隣接する建物

同定実験1)と2)の比較を表6.3に示す。同定実験1)では、117個の指示文書のうち96個(82%)に関して、3人の被験者全員が指示対象物と同一の対象を回答した。残り21個(18%)に関しては、3人の被験者の少なくとも1人は指示対象物とは異なる対象を同定したり、複数の対象を候補にしたり、該当する対象はないと回答した。

同定実験2)は、実験1)において3人が一致する96個の指示文を用いた。IMAGEは、96個のうち76個(79%)の指示文書に対して3人と同じ対象を同定した。残り20個(21%)の指示文書については、人間と異なる結果となった。その内訳は、指示対象物以外の対象を同定したのが10個(11%)、複数の対象を同定したのが8個(8%)、対象を同定できなかった指示文が2個(2%)であった。

表6.3 同定実験の結果

実験結果	比率(文章数)
3人の被験者全員が指示対象物を同定した [集合Aと呼ぶ]	82% (96個)
IMAGEも指示対象物を同定した	79% (76個)
IMAGEは指示対象物を同定できなかった	21% (20個)
異なる対象物を同定した	11% (10個)
複数の対象物を同定した	8% (8個)
同定できなかった	2% (2個)
異なる対象物・複数の対象物を同定するか、対象物を同定できなかった人が、少なくとも一人以上いた	18% (21個)

以上、手の形の認識、ポインティング精度、位置関係の同定率と70%程度の認識率の要素を組み合わせたIMAGEシステムで、

1. 全体としての認識率は、各要素の認識率の代数積程に低下しない。
2. 複数の手振り指示を加えることで、問い合わせの表現方法が広がる。
3. 言葉と手振りを用いた相対指示においては、言葉と手振りの対応に、複数の解釈が存在する。表6.1の「ツイン21の右のこの建物」の代わりに、「このツイン21の右の建物」のとき、手振りによる指示が「ツイン21」を指しているケースと、「ツイン21の右の建物」を指している2つのケースが考えられる。この問題に対しては、対話の流れの中から、指示された対象を同定する必要がある。

ことを確認した。

6.4 第6章のまとめ

本章では、音声と手振りといったマルチモーダルインタフェースをもつ地図案内システムIMAGEについて述べた。複数の方法から入力される問い合わせを、システムの持つデータベースに対する検索とみなし、関係の連言で統一的に表現する方法を提案し、IMAGEにおける指示効率の評価について述べた。マルチモーダルインタフェース全体として定量的な認識率の評価、対話モデルを用いた応答性や認識率の向上については、今後の課題である。

第7章 結言

1. 1章において本研究の目的として、人間が、計算機とのインタフェースに用いる定性的な位置関係のモデルを評価する事と、インタフェースに身振り(手振り)情報の利用を検討する事の2点を挙げた。

第1の目的に関しては、主に第2章と第3章で、定性的な位置関係をif-then形式のルール表現を用い定量的な対象の頂点座標で表現する方法を提案し、画像データベースの内容検索への適用例について述べた。

第2の目的に関しては、第4章と第5章の中で手形状入力装置を用いた手の形・動きのコード化方法を提案し、日本語手話における指文字を例に用いた認識実験の評価について述べた。

第6章では、定性的な位置関係の表現と手振り情報を、言葉と手振りによる指示といった異なる複数の方法を同時に使用するマルチモーダルインタフェースに適用することを検討した。地図案内システムIMAGEを試作し、計算機の新しいインタフェースについて検討を行った。

本研究で得られた主要な結果と今後の課題を以下に示す。

1. 位置関係のモデル評価

自然言語における位置関係の定性的な特徴について述べた。位置関係を定量的な基本位置関係と定性的な位置関係に分けた。位置関係のもつ漠然さに関する主観実験を行い、漠然さには個人差はあるものの、基本的な性質には共通性があることを示した。人間の持つ位置関係の漠然さを、表現するヒューリスティックな関数を加えたif-then型のルールで、位置関係を対象の頂点座標をもとに表現する方法を提案した。

イメージ情報を検索手がかりとする画像データベースの内容検索に、位置関係の知識ベースの利用を検討した。検索時に位置関係のルールに従い、言語による問い合わせに含まれる位置関係と、画像データベースを管理している対象の頂点座標とを照合する方法を提案した。実際に、浮世絵画像や文書画像を用いたプロトタイプシステムSPADEを試作し、対象の意味的情報を含んだ位置関係に関して、適合率の高い内容検索ができる等の有効性を確認した。

2. インタフェースとしての手振り

手形状入力装置を用いた計算機インタフェースが、人工現実感(Virtual Reality)の技術とともに検討され始めている。日本語手話の46個の指文字に対する手形状を、手形状入力装置で測定した関節曲げデータを用い、統計的手法で分析を行った。

主成分分析では第4主成分までで90%近くの累積寄与率があること、クラスタ分析では大きく4つのクラスタに分類できることが明らかになった。指の曲げ角や手の傾きにより手の形をコードで表現する方法では、46個の手形状が30個程度に分類できることを示した。

手形状のコード化に加え時系列的な変化に伴う手の動きのコード化を用いた手振り認識方法を提案した。手形状入力装置を用い、一連の手話の指文字の認識実験から、手の動きを含め20個程度の手振りが動きの中で認識できる事を確認し、手振りの計算機インタフェースへの適用可能性を示した。

3. マルチモーダルインタフェースへの適用

音声と手振りといった複数の入力方法を同時に用いたマルチモーダルインタフェースが、新しい計算機インタフェースとして考えられ始めている。地図案内システムを例にとり、言葉による位置関係の指示と手振りによるインタフェースをもつ地図案内システムIMAGEを試作した。その中で、複数の方法により入力される問い合わせを統一的に表現する方法を提案した。IMAGEのインタフェースにおける指示効率を決める要因を解析し、一つの要因である位置関係の対象同定実験について述べた。

今後の研究課題として、

1. 検索者のイメージする内容を手がかりに、画像データベースを検索する内容検索を評価する適合率・再現率の検討。
2. 日常生活で用いられる手振りには、手の屈伸状態のほかに指先やてのひらとの接触など視覚的な特徴が多い。指の曲げ、手の向きのほかに、指先の接触状態に関する特徴抽出方法の検討。

3. 手振り認識のパターンのもとになる指の曲げ, 手の向きなどには個人差がある。特定操作者の手振り認識から, 複数の操作者の手振り認識を行う不特定操作者の手振り認識方法の検討。
などがあげられる。

謝 辞

本論文をまとめるにあたり, ご指導いただきました名古屋大学工学部杉江昇教授に厚くお礼を申し上げます。また, ご助言いただきました伊藤正美教授に深く感謝いたします。鳥居達生教授にはご助言いただくとともに, 本論文の作成の機会をいただきました。ここに厚く感謝いたします。

本研究は, 著者がATR通信システム研究所での研究業務として実施したもので, 本研究をご指導いただきました山下紘一ATR通信システム研究所社長, 小林幸雄元ATR通信システム研究所知能処理研究室長(現NTTヒューマンインタフェース研究所視覚情報研究部部長), 岸野文郎ATR通信システム研究所知能処理研究室長に厚くお礼を申し上げます。また, 本研究の機会をいただきました葉原耕平ATR通信システム研究所会長, 鈴木敏正元NTT研究開発技術本部副本部長(現NTTエレクトロニクステクノロジー株式会社), 佐々木幹生元NTT技術情報センタ担当部長(現NTTソフトウェア株式会社)に深く感謝いたします。酒井高志NTTヒューマンインタフェース研究所グループリーダーには本論文をまとめる機会をいただきました。

本研究の実施中に, ご助力, ご討論いただいた伯田晃氏(現富士通研究所), 西村康氏(現松下電器産業株式会社), 島則之氏(現住友電気工業株式会社)をはじめとするATR通信システム研究所の方々に心から感謝申し上げます。

参考文献

- 1) 白鳥：左右対称形キーボードの試作と打けん操作特性の評価，計算機システムのヒューマンインタフェースシンポジウム，情報処理学会，pp.113-122 (1989)
- 2) 栢原：日本語文献データベースへの知的アクセス，電子情報通信学会誌，Vol.72, No.7, pp.797-806 (1989)
- 3) 渡辺，松井，廣瀬，神部，井越：3次元対話環境の将来，情報処理，Vol.30, No.12, pp.1517-1533 (1989)
- 4) R. A. Bolt: The Integrated Multi-Modal Interface, 電子情報通信学会論文誌 D, J70-D No.11 pp.2017-2025 (1987)
- 5) R. A. Bolt: The human interface, van Nostrand Reinhold, N.Y. (1984)
- 6) 坂内：画像検索技術，電子情報通信学会誌，Vol.71, No.9, pp.911-914 (1988)
- 7) 木戸出，恒川：画像情報処理におけるマルチメディアデータベース，情報処理，Vol.28, No.6, pp.756-764 (1987)
- 8) Y. E. Lien, D. F. Utter: Design of an Image Data-Base, PDDM '77, pp.131-136 (1977).
- 9) 大川，馬場口，手塚：知識利用型検索システムの構成，情報処理学会研究会知識工学と人工知能 56-9 (1988)
- 10) 浦谷，柴田，野口，相沢：静止画検索システム F O R K S の試作，情報処理学会論文誌，Vol.28, No.7, pp.758-767 (1987)
- 11) 笠原，岸本：画像データベースナビゲーション，信学技報，IE88-46 (1988)
- 12) 長谷川，福村，鳥脇：胸部X線写真データベースのためのスケッチ画像の作成と利用，電子通信学会論文誌，Vol. J65-D, No.9, pp.1121-1128 (1982)
- 13) 林，河合，上原，豊田：画像データベースの意味検索の試みー道路地図データベースの略地図による検索ー，情報処理学会研究会，知識工学と人工知能 44-2 (1986)
- 14) 大田，森，坂井：漢字パターン列の特徴パラメータによる検索ー名刺画像における姓名の場合，電子通信学会論文誌，Vol. J64-D, No.11, pp.997-1004

- (1981)
- 15) 山守，吉田，福村：天気図の符号化とその天気図データベース・システム実現への応用，情報処理学会論文誌，Vol.26, No.6, pp.1049-1056 (1985)
- 16) 笠原，的場，平林：商品企画コンセプトデザインシステム，情報処理学会グラフィックスとCADシンポジウム，pp.159-165 (1987)
- 17) 栗田，下垣，加藤：主観的類似度に適応した画像検索，情報処理学会論文誌，Vol.31, No.2, pp.227-237 (1990)
- 18) 杉田：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，Vol.28, No.6, pp.765-772 (1987)
- 19) 黒川：形状特徴を用いた画像の内容依存型検索，情報処理学会第35回（昭和62年後期）全国大会，2K-1, pp.2105-2106
- 20) S.K.Chang, C.W.Yan, D.C.Dimitroff, T.Arndt: An Intelligent Image Database System, IEEE Trans. Softw. Eng. Vol. SE-14, No.5, pp.681-688 (1988)
- 21) J.A.Orenstein, F.A.Manola: PROBE Spatial Data Modeling and Query Processing in an Image Database Application, IEEE Trans. Softw. Eng. Vol. SE-14, No.5, pp.611-629 (1988)
- 22) 山本，高木：認識対象の特徴に基づく画像検索，情報処理学会第35回（昭和62年後期）全国大会，2K-2, pp.2107-2109
- 23) J. Freeman: SURVEY The Modelling of Spatial Relations, COMPUTER GRAPHICS AND IMAGE PROCESSING 4, pp.156-171 (1975)
- 24) ウィンストン.P.H. 編：コンピュータビジョンの心理（白井，杉原訳），産業図書(1979)
- 25) 松原，坂上，横矢，山本：概念学習を題材とした画像理解と記号処理の統合の試み，人工知能学会誌，Vol.3, No.5, pp.572-579 (1988)
- 26) 安部，曾我，辻：言語・画像を利用した行動の解釈（1）ー発話指示による対象の同定ー情報処理学会論文誌，Vol.23, No.2, pp.124-132 (1982)
- 27) 安部，曾我，辻：言語・画像を利用した行動の解釈（2）ー粗筋抽出と質疑応答ー情報処理学会論文誌，Vol.23, No.2, pp.133-141 (1982)
- 28) 岡田，田町：動画像の意味解釈とその自然語記述ー意味分析ー，電子通信学

- 会論文誌, Vol.J59-D, No.5, pp.331-338 (1976)
- 29) 岡田, 田町: 図形の意味解釈とその自然語記述—要素的図形認識と構造分析—, 電子通信学会論文誌, Vol.J59-D, No.5, pp.323-330 (1976)
- 30) 高木, 伊東, 六沢, 北岡, 清水, 小原: 2次元図形世界における視覚情報からの日本語文の生成, 電子通信学会論文誌, Vol. J67-D, No2, pp216-233 (1984)
- 31) 伊東, 高木, 六沢, 清水, 北岡, 小原: 2次元図形世界における日本語文と視覚情報の付き合わせ処理, 電子通信学会論文誌, Vol. J69-D, No7, pp1083-1094 (1984)
- 32) R. L. Haar: Sketching: Estimating Object Positions from Relational Descriptions, COMPUTER GRAPHICS AND IMAGE PROCESSING 19, pp.227-247 (1982)
- 33) 山田, 西田, 堂下: 2次元平面におけるポテンシャルモデルを用いた位置関係の推定, 情報処理, Vol.29, No.9, pp.824-834 (1990)
- 34) 山田, 網谷, 星野, 西田, 堂下: 自然言語における空間描写の解析と情景の再構築, 情報処理, Vol.31, No.5, pp.660-672 (1990)
- 35) J. D. Foley: 近未来のインタフェース, サイエンス 17 [12], pp.76-85
- 36) 廣瀬: 人工現実感の生成, システム/制御/情報, Vol.33, No.11, pp.40-47 (1987)
- 37) T. G. Zimmerman, J. Lanier, C. Blanchard, S. Bryson, Y. Harvil: A HAND GESTURE INTERFACE DEVICE, CHI + GI 1987, pp.189-192
- 38) 伊藤, 竹村: 手話入門, 廣済堂出版 (1988)
- 39) 田村, 川崎: 手話動画像認識システム, 情報処理学会コンピュータビジョン研資, 44-1 (1986)
- 40) 萩原, 脳: 輪郭特徴を用いた指文字認識方法, 昭和60年度電子通信学会情報・システム部門全国大会, 114
- 41) 浅井, 落合, 鎌田: パソコンによる指文字認識法, 昭和63年電子情報通信学会春季全国大会, D-406
- 42) W.C. Stokoe, D.C. Casterline, C.G.Groneberg: A DICTIONARY OF AMERICAN SIGN LANGUAGE ON LINGUISTIC PRINCIPLES,

- Linstok Press(1976)
- 43) 藤野, 鎌田, 薄井, 中山, 森: 機械辞書のための手話構造の検討, 情報処理学会第39回(平成元年後期)全国大会, 4E-9
- 44) 黒川, 紙谷: コンピュータ身振り辞書の試作, Human Interface News and Report, Vol.4 pp.87-92 (1989)
- 45) D. J. Sturman, D. Zelter, S. Pieper: Hands on Interaction With Virtual Environments, UIST '89: ACM SIGGRAPH/SIGCHI Symposium on User Interface Software and Technology, Nov. 13-15 (1989)
- 46) D. Weimer, S. K. Ganapathy: A SYNTHETIC VISUAL ENVIRONMENT WITH HAND GESTURING AND VOICE INPUT, CHI '89 proceedings pp.235-240 (1989)
- 47) DataGlove Model2 GESTURE EDITOR SOFTWARE Operating Manual, VPL Research Inc. (1988)
- 48) DataGlove Model2 Operating Manual & Test CALIBRATION SOFTWARE Operating Manual, VPL Research Inc. (1988)
- 49) 中島秀之: 知識表現の基礎(I), 人工知能学会誌, Vol.4, No.4, pp.383-388 (1989)
- 50) A.Herskovits: Language and spatial cognition, Cambridge University press(1986)
- 51) 伊理, 腰塚: 計算幾何学と地理情報処理, bit, 1986年9月号別冊, 共立出版
- 52) T.Takahashi, A.Hakata, N.Shima, Y.Kobayashi: Scene Description using Spatial Relationships Derived from Visual Information, SPIE Vol.1192 Intelligent Robots and Computer Vision VII: Algorithms and Techniques (1989) pp.702-708
- 53) 伯田, 高橋, 小林: 自然言語で指示された対象物の同定方法, 情報処理学会論文誌, Vol.31, No.12, pp.1735-1742 (1990)
- 54) 田中編: 知識工学, 朝倉書店(1985)
- 55) 石塚: 曖昧な知識の表現と利用, 情報処理, Vol.26, No.12. pp.1481-1486 (1985)
- 56) 石塚: Dempster & Shafer の確率理論, 電子通信学会誌, Vol.66, No.9,

pp.900-903 (1983)

- 57) 上野, 石塚: 知識の表現と利用, オーム社 (1987)
- 58) B. G. Buchanan, E. H. Shortliffe: RULE-BASED EXPERT SYSTEMS - THE MYCIN EXPERIMENTS OF THE STANFORD HEURISTIC PROGRAMMING PROJECT, Addison-Wesley Publishing Company (1984)
- 59) 島, 高橋, 小林: ヒューマンインタフェースにおける位置指示言語の評価実験, TV学会技術報告, Vol.12, No.58 (1988)
- 60) 国立国語研究所(編): 分類語彙表, 秀英出版 (1964)
- 61) 小林: プロダクションシステム, 情報処理, Vol.26, No.12, pp.1487-1496 (1985)
- 62) 高橋, 伯田, 小林: 2次元世界の位置関係作成とシーンの記述について, 人工知能学会全国大会(第3回) 9-1 (1989)
- 63) 高橋, 島, 岸野: 位置情報を手がかりとする画像検索方法, 情報処理学会論文誌, Vol.31, No.11, pp.1636-1643 (1990)
- 64) 西村, 高橋, 小林: 木構造モデルによる文書画像からの検索情報抽出, 電子情報通信学会研究会 PRU89-34 (1988)
- 65) ジェラルド・サルトン(著), 神保(訳): SMART情報検索システム, 企画センター (1974)
- 66) 八村: 身体運動情報の処理, 情報処理学会研究会人文科学とコンピュータ 6-3 (1990)
- 67) 黒川, 小野, 紙谷: 手の形態のコード化と意味的冗長性の除去-機械の非音声コミュニケーションに関する基礎研究-, 第4回ヒューマンインタフェースシンポジウム, 1988, pp.167-172
- 68) 河口: 数学ライブラリ32 多変量解析入門I, 森北出版社(1986)
- 69) 河口: 数学ライブラリ46 多変量解析入門II, 森北出版社(1985)
- 70) M. R. Anderberg: CLUSTER ANALYSIS FOR APPLICATIONS, ACADEMIC PRESS New York and London (1973)
- 71) T.Takahashi, F.Kishino: Hand Gesture Coding Based on Experiments Using A Hand Gesture Interface Device, acm/SIGCHI Bulletin, Vol.23, No.2, April 1991, pp.67-74

- 72) 高橋, 岸野: 手振り認識方法とその応用, 電子情報通信学会論文誌, vol. J73-D-II, No.12, pp.1985-1992 (1990)
- 73) A.Hakata, T.Takahashi, Y.Kobayashi: Object Identification by Language in a User Interface Using Language and Image Information, acm/SIGCHI Bulletin, Vol.21, No.3, January 1990, pp.33-36
- 74) T.Takahashi, Y.Kobayashi, A.Hakata, N.Shima; Unifying voice and hand indication of spatial layout, SPIE's 1989 Conference on Sensor Fusion II (1198-29) Nov 5-10 (1989)
- 75) 栄藤, 伴野, 小林: 指の方向検出を用いたヒューマンインタフェース, 第4回ヒューマンインタフェースシンポジウム, 1988, pp.173-176

