

## 位置関係を利用した画像データベース検索システム

高橋友一      島 則之      岸野文郎  
ATR通信システム研究所

あらまし 「山の麓に川がある絵」といった画像の内容検索を、問い合わせに含まれている対象の位置関係を画像検索の手がかりにして行う画像検索システムSPADE(SPAatial relationships based feature image retrieval/DEscription system)について述べる。画像データベースの検索方法として、言語を用いる時、多義性、漠然性を持つ言語、画像間の共通の情報を検索キーとする必要がある。SPADEでは、画像認識、理解技術により抽出される対象の位置情報と位置関係のもつ漠然性を表現した位置関係の辞書を用いて、問い合わせ文に含まれる位置関係を満たす対象を含む画像の検索を行なっている。最後に、SPADEで文書画像データベースの内容検索を行った例を示す。

### An Image Database Retrieval System using Spatial Relationships

Tomoichi TAKAHASHI Noriyuki SHIMA Fumio Kishino

ATR Communication Systems Research Laboratories

Sanpeidani, Inuidani, Seika-cho, Soraku-gun, Kyoto 619-02, Japan

**Abstract** A method of retrieving feature images using spatial relationships between objects in an image is discussed. In pictorial information retrieval, we often desire to retrieve pictures having some relationships, for example, "find some pictures in which a river is running at the foot of a mountain". Our method infers automatically necessary relationships between objects in a feature image using lexicon of spatial relationships. This lexicon are consisted of procedures and rule representations which calculate basic relationships of objects and their inherent fuzziness. Our method are implemented SPAatial relationships based feature image retrieval/DEscription system(SPADE) and an example of document image database on SPADE is shown.

## 1. はじめに

画像データベースの検索方法として、利用者の問い合わせ内容を言葉やスケッチ画で表現した情報を検索キーとし、

- (1) 画像内容を反映した検索キーワードを、予め画像に付与し、キーワード検索で対応する画像を検索する。
- (2) 画像の構造を表現しているアブストラクト画像を検索キーとして、同じ構造を持つ画像を検索する。
- (3) 画像の画素ラン数等の特徴パラメータで、対象画像を分類する。
- (4) 定性的な言葉や感性を定式化、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て、検索する。

等の方法が提案されている。[1]

言語で画像を検索するには、画像情報、言語情報と多義性、曖昧さのあるメディア間の共通情報を手がかりにして、画像データベースを検索する必要がある。本論文では、「山の麓に湖がある画像」のように画像内の対象間の位置関係を画像検索の手がかりとする画像検索システム SPAD E (SPAtial relationships based feature image retrieval/DEscription system) について述べる。SPAD E では、(1) 画像認識、理解技術で特徴抽出される情報で画像を管理、(2) 問い合わせ文の解析から対象に関する情報と参照されている対象間の位置関係を抽出、(3) 問い合わせ文に含まれている位置関係を、対象の位置情報と位置に関する知識を用いて推論、等により画像の内容検索を実現している。

## 2. 位置関係を利用した画像検索

画像に含まれている対象間の位置関係を用いた検索システムとしては、画像内の領域を対象領域に含まれている領域の集合 (point-set) で表現し、point-set に対する演算を基に対象間の位置関係を求める PROBE [2]、領域間の相対位置関係を 2-D string 表現で表現し、検索する IIB [3] や、風景画に含まれる山の領域を画像処理技術で認識し、その認識された領域の形状、位置、大きさ、色といった属性情報で検索するシステム [4] がある。これらのシステムは、その目的とする位置関係の導出にあたって、目的対応の画像処理を行なっている。

SPAD E では、画像認識、理解技術により、画像に含まれる対象の位置情報、属性情報が自動抽出される事を前提している。画像認識、理解技術により抽出された対象の位置情報と属性情報から構成される図形画相当の情報が、画像データの検索キーとなる。「山の麓に川がある」といった内容検索に対しては、「山」や「川」が存在している条件の検索以外に、位置関係「麓」を反映した検索が必要になる。位置関係については、画像処理技術で位置関係を認識するのではなく、画像認識、理解技術により得られた対象位置と位置関係に関する知識から照合する。(図.1)

即ち、自然言語による問い合わせ「山と川がある画像」の実行は、

```
attribute(x, 山),
attribute(y, 川)
```

の 2 つの情報を含む画像を検索するのに対し、

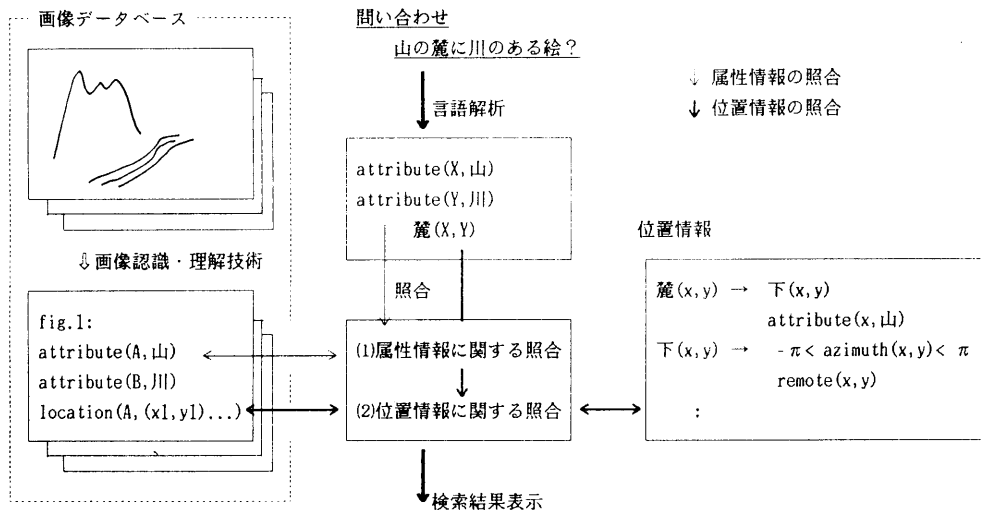


図. 1 SPAD E システムにおける照合処理

問い合わせ「山の麓に川がある画像」に対しては、上記の条件に加えて、「麓(x,y)」の条件を満たす対象x,yを含む画像を検索する。

### 3. 位置関係の定式化と照合

位置関係は、人間の主観を反映した関係で、その定義は、例えば「右」を広辞苑では「南を向いた時、西にあたる方」と他の関係を用いて説明してあるように、明確に規定する事は難しい。〔5〕〔6〕

例えば、AがBの右にあるならば、

1. Bの中心からみて右半面（Aの左端の頂点の角度が $-\pi/2$ から $\pi/2$ の範囲）にある。
2. AとBは、お互いに離れている。

と、基本物理量、物理的關係で記述される関係を満たしている。しかし、条件1,2.を満足しているからといって、「AがBの右にある絵」ではなく「AがBの上にある絵」と表現する方が適切な時もある。（図.2）

位置関係の漠然さを含め、対象の位置情報と属性情報で管理されている画像と問い合わせに含まれている位置関係との照合を行なう為に、位置関係を以下の2種類の方法で定義した。

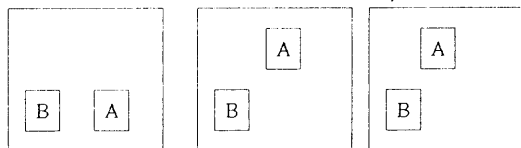
- (1) 言語によって表現している位置関係は異なる。例えば、日本語の「上」は英語の「on, over, above」に対応する。従って、位置関係の言葉毎に関係を定義するのではなく、対象の位置座標から、図形画における対象の位置関係を表現する。その為に、位置関係を、基本物理量、物理的關係、位置関係1、2の4レベルの階層の位置関係に整理し、その関係で言語の位置関係を定義した。（表.1）

表. 1 位置関係の4レベル

| レベル   | 内容                        | 位置関係の例                        |
|-------|---------------------------|-------------------------------|
| 基本物理量 | 対象の頂点座標から計算できる量。          | 中心座標、面積等 (11)                 |
| 物理的關係 | 位置関係を定義する時に必要な対象間の物理量や関係。 | 物理量：距離、角度<br>関係：包含、離れている等(17) |
| 位置関係1 | 基本的物理量、関係で定義される位置関係       | 上下左右、遠近、端等 (6/15)             |
| 位置関係2 | 位置関係1の合成で定義される位置関係        | 右上、右端、右隣等 (12/16)             |

注1) 括弧内の数字は各レベルで定義した関係数。

注2) 位置関係は（空間的／相対的）な位置関係。



AはBの右にある。

図. 2 位置関係の漠然性

AはBの上にある。

- (2) 画像の表現している世界固有の位置関係の言葉、例えば、風景画における麓等、は対象の属性情報と位置関係で定義する。

例) 位置関係「麓(x,y)」:-

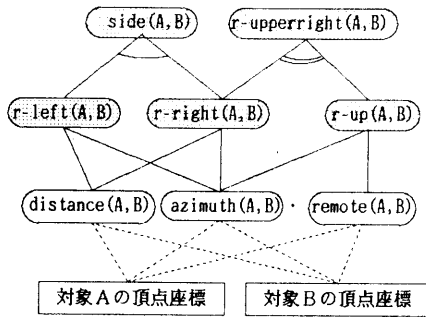
位置関係「下(x,y)」:-

属性情報 - attribute( x,山)-

基本物理量、物理的關係は対象の位置情報（頂点座標）が与えられると計算可能な量、関係である。対象間の位置関係は、人間の知覚の漠然さを伴っている。この位置関係の漠然性を反映し、問い合わせ「AがBの右にある絵」に対し適切な画像を選択するには、位置関係の漠然さを表現する事は必要である。

画像情報と言語情報に含まれている対象間の位置関係を照合する辞書は、基本物理量、物理的關係は、対象の外接多角形の頂点座標をもとに手続きで記述され、位置関係は基本物理量、物理的關係および他の位置関係をもとに、ルール表現で記述されている。（図.3）

図. 4 に、「右」の表現する範囲を定義した例と、「右」のもつ漠然性も同じく対象の中心間の角度、距離を用い、角度が0°に近い程、距離がある程度離れている程、右らしいと定義した例を示す。〔7〕



☐ 物理量、関係    ☐ 漠然性をもつ位置関係  
 : 手続きによる表現  
 | ルールによる表現  
 / or関係    \ comb関係

図.3 位置関係のレベル

「右上」の様に「右」と「上」の関係の合成で定義される位置関係2に含まれる関係の漠然さは、関係の定義に用いられている関係「右」、「上」の漠然さを反映している。「右」の様に発見的な関数で、その位置関係の漠然性を評価するのではなく、And, Or, Comb関係等の関係の定義方法に応じ、MYCINの確信度と同じ方法で評価する。[8]

#### 4. 位置関係を検索キーとする照合手順

検索要求で指示された対象間の位置関係—位置関係名(対象1、対象2)—と、データベースに含まれている対象の頂点座標を、位置関係の定義に基づき照合する手順を以下に示す。

- (手順1) 検索要求に含まれている山とか川の様な属性情報をもつ対象を含んでいる画像を候補画像とする。
- (手順2) 検索要求に含まれている対象の位置関係を位置関係の定義情報に従い、位置関係1、物理的関係あるいは属性情報の積の形の集合に展開し、関係を満足している画像をさらに選択する。
- (手順3) 各位置関係に対して定義されている関係の漠然さの値から、画像毎に指定された位置関係の充足度を評価する。

手順2において、「麓」のように画像の表現世界の情報を含んでいる位置関係は、その属性情報も含めた形で展開される。(図.5)

```

(defrel s-right-definition
  (definition
    (rel 3 s-right (?A))
    <=
    (and (rel 2 distance1 (Frame ?A) (?Distance))
         (rel 2 azimuth1 (Frame ?A) (?Azimuth1))
         (rel 2 azimuth3 (Frame ?A) (?Azimuth3))
         (test (and (<= ?Azimuth3 (#L/ pi 2))
                    (<= (#L/ pi -2) ?Azimuth3))))))
    (cf-calc(* (+ (* -3 ?Distance (- ?Distance 1)) (#L/ 4))
              (- 1 ( #L/ (abs ?Azimuth1) (/pi2))))))

(defrel s-uperright-definition-1 "A is right on up"
  (definition
    (rel 4 s-uperright (?A))
    <=
    (rel 3 sright(?A)))

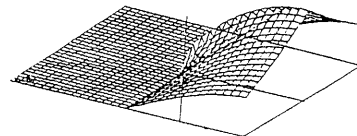
```

```

(defrel s-uperright-definition-2 "A is right on up"
  (definition
    (rel 4 s-uperright (?A))
    <=
    (rel 3 stop (?A)))

```

(a) 定義ルール



(b) 漠然さの評価関数の形

図.4 位置関係「右」の定義例

#### (1) 山の麓に川のある絵。

```

→attribute(x,山)
attribute(y,川)
麓(x,y)
麓(x,y)
→attribute(x,山)
down(x,y)
down(x,y)
→TRUE(- π < azimuth(x,y) < π)
remote(x,y)

```

```

⇒attribute(x,山)
attribute(y,川)
TRUE(- π < azimuth(x,y) < π)
remote(x,y)

```

#### (2) 家の麓に川のある絵。

```

⇒attribute(x,家)
attribute(y,川)
attribute(x,山)
TRUE(- π < azimuth(x,y) < π)
remote(x,y)

```

図.5 問い合わせの展開形

従って、SPADEでは、問い合わせ「家の麓に川のある絵」の様に、意味のおかしい問い合わせは、言語解析ではなく、一つの対象が「家」と「川」の属性を持つことになり、検索文の意味解析ではなく、検索時に該当画像なしで終了する。

##### 5. SPADE上の文書画像データベースの例

対象画像として、66枚の科学論文の第1頁を対象としSPADEシステムを用いて実現した文書画像データベースの例を示す。一般に、画像に含まれる対象の位置情報、属性情報の自動抽出は難しい。このデータベースの対象とする文書画像に含まれる対象の位置情報、タイトル、著者名等の属性情報は、論文レイアウトの知識を用い、自動抽出できる。[9]

図.6に、文書画像の原画像、領域抽出された画像と自動抽出された対象の位置情報と属性情報のリストを示す。

文献書誌情報による検索の他に、「タイトルの横に著者名のあった文献」といった文書レイアウトの問い合わせは、

```
attribute( X, title)
attribute( Y, author)
side(X,Y)
```

の条件を満たす画像を検索する事になる。

「2段組の文献」の検索といった対象世界固有の「2段組」といった意味は、「麓」と同じく属性情報を含む形で、

```
attribute( X, title)
attribute( Y, title)
side(X,Y)
```

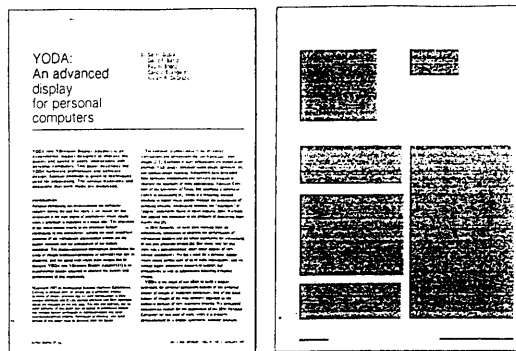
と展開される。

図.7にSPADEの検索画面の構成を示す。

右上のウィンドウが検索要求ウィンドウで「タイトルの横に著者名がある文献」のような検索文を入力する。左下のウィンドウが検索結果表示ウィンドウで、関係充足条件を満たした3件の画像番号が関係充足度とともに表示される。左上のウィンドウが対象の図形の表示ウィンドウで検索した結果を表示する。図.7では、3件のうちの2件の文書画像の1/8縮小画像が表示されている。

##### 6. まとめ

画像データベースの検索として、画像内の対象間の位置関係を検索キーとした内容検索について述べた。本方法は、画像処理技術とは独立



入力画像

領域抽出後の図形画

```
(dat POSITION FRAME ((0.0 0.0)
(0.70 0.0) (0.70 1.0) (0.0 1.0)) 1)
(dat POSITION B1 ((0.07 0.66)...
:
(dat POSITION B9 ((0.07 0.09)
(0.36 0.09) (0.36 0.20) (0.07 0.20)) 1)
(dat ATTRIBUTE B1 TITLE 1)
(dat ATTRIBUTE B2 AUTHOR 1)
(dat ATTRIBUTE B3 PAGE 1)
(dat ATTRIBUTE B4 TEXT 1)
(dat ATTRIBUTE B5 NONE 1)
(dat ATTRIBUTE B6 NAME 1)
(dat ATTRIBUTE B6 VOLUME 1)
(dat ATTRIBUTE B6 NUMBER 1)
(dat ATTRIBUTE B6 DATE 1)
(dat ATTRIBUTE B7 ABSTRACT 1)
(dat ATTRIBUTE B8 TEXT 1)
(dat ATTRIBUTE B9 FOOTNOTE 1)
```

対象の位置情報と属性情報

図. 6 文献画像DBのデータ

した位置関係の定義辞書を用い、問い合わせに含まれている対象間の位置関係情報と画像の管理情報として認識された対象の位置座標・属性情報を関係つけている所に特徴がある。

自然言語により、画像の内容記述を手がかりに画像データベースを検索する時、画像の認識、理解技術以外の今後の課題として、

- ・「車の前に人がいる絵」と「車の後ろに人がいる絵」と人と車の位置関係が同じでも、車の向きにより異なる絵を検索する時。
- ・同じ絵でも、発話者の視点により、異なる問い合わせで参照される時。

等、自然言語からの必要な検索キーの作成と画像情報との照合に用いる位置情報の表現抽出が挙げられる。

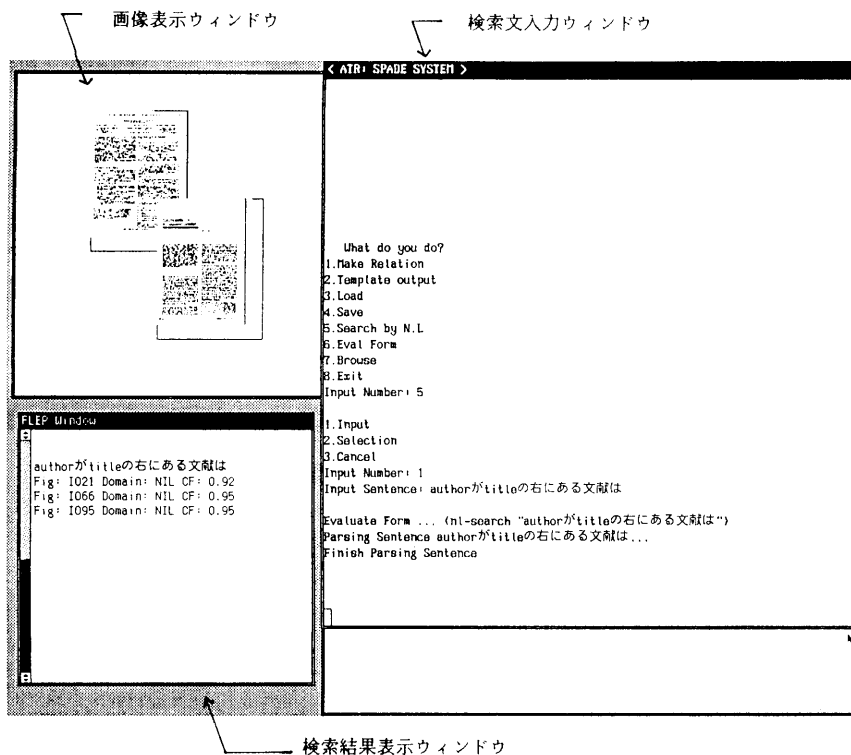


図. 7 SPADE検索画面

#### 参考文献

- (1) 坂内: 画像検索技術、電子情報通信学会 Vol.71 No.9 pp.911-914 (1988)
- (2) J.A.Orenstein, F.A.Manola: PROBE Spatial Data Modeling and Query Processing in an Image Database Application IEEE Tran. on Softw. Eng. Vol. SE-14, No.5, pp.611-629(1988)
- (3) S.K.Chang, C.W.Yan, D.C.Dimitroff, T.Arndt: An Intelligent Image Database System IEEE Tran. on Softw. Eng. Vol. SE-14, No.5, pp.681-688 (1988)
- (4) 岡崎、美濃、坂井: 風景画像の認識とその認識対象属性による画像検索、信学技報 IE86-117 (1986)
- (5) A.Herskovits: Language and spatial cognition, Cambridge University press(1986)
- (6) 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)
- (7) 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)
- (8) 田中編: 知識工学 朝倉書店(1985)
- (9) 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、PRU89-34 (1988)