

2次元画像の位置関係記述におけるあいまい性の扱いの一検討\*

5R-9

島 則之 高橋 友一 †

ATR 通信システム研究所†

1 はじめに

画像と言語を統合化することで、ユーザにとって使い易い様々なインターフェースを構築することができる。例えば言葉により『家の前に人が立っていた図』と指定すると画像の内容でデータベースを検索することができたり[1]、また画像を目前にして『立っている人をもう少し右にずらせ』というような、画像を操作することも可能である。そのためにはユーザの言った言葉と画像を何らかの方法でマッピングする必要がある。本稿では2次元画像を対象に、位置に関する言葉を画像にマッピングするさいのあいまい性の取り扱いについて述べる。

2 画像情報

画像が表現している情報は大きく2種類に分けることができる。1つは、画像の表す現実世界の情報(実世界情報と呼ぶ)で、もう1つが画像自身が持つ属性の情報、例えば位置や形や大きさ等の情報(視覚情報と呼ぶ)である。我々は図1で示す様な2次元画像を対象に、視覚情報のうち位置に関する情報に注目して、その画像の内容を記述するシステムを構築中である[2]。図2は本システムが図1の画像の位置関係を記述した例を示している。

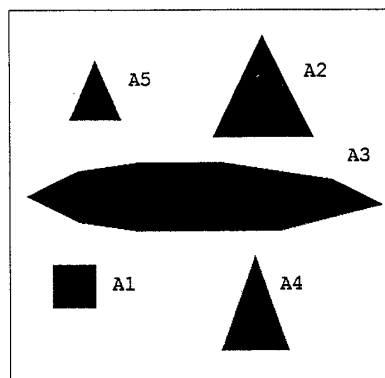


図 1: 対象とする 2 次元画像例

- A3(湖) は A2(山) の下にある. (0.89)
- A1(家) は左下にある. (0.80)
- A1(家) は下にある. (0.57)
- A1(家) は左にある. (0.54)

図 2: 2 次元画像位置関係記述例 (抜粋)

3 画像位置記述システム

2次元画像の各物体は凸多角形であると仮定し、凸多角形の各頂点座標が本システムに入力されるものとしている。この入力データから『物体Aは画像の真中にある』等の位置関係の記述を行なうために、表1のように処理するデータを4つの階層に分割している。レベル1は各物体固有の物理量(面積等)を表すデータで、レベル2のデータは複数物体間で一意に決定される物理量(距離等)を表している。レベル3はレベル2までの物理量をもとに『右』等の位置関係を表すデータである。レベル4はレベル3のデータを組み合わせて求められるデータ(『右上』等)としている。レベル1やレベル2のデータは入力データから一意的に求められるもので、本システムでもそれらのデータは手続き的に計算を行なって求めている。レベル3以上の位置関係は人間の主観を伴うあいまいな関係で、一意には求められないものである。従って、本システムでは評価関数を用いて、関連する物理量から位置関係を求めるようにルールにより定義している。

表 1: 位置データのレベル

分類	位置関係の情報等
物体データ	中心座標, 面積等
物体間データ	距離, 角度, 隣接関係等
基本的関係	上下左右, 遠近, 端等
応用的関係	右上, 右下, 右端等

4 マッピングにおける問題点

位置情報を位置関係の言葉へマッピングする際、次のような問題点をあげることができる。

- (1) レベル3以上の位置関係の計算機内での表現  
例えばある物体が『右』にあるかどうかの判断は人間の主観に依存しており、言葉の持つあいまい性を計算機で扱えるように数値化する必要がある。
- (2) 画像が表現していると思われる現実世界との対応  
位置の言葉に関しても、『手前』『奥』などという言葉は、画像を現実世界に反映させた時に使用される言葉であり、画像の表現している世界に依存するような言葉をどのように扱うかという問題がある。
- (3) 注目している物体以外による影響  
例えば注目している2物体間の位置関係は、他の注目外の物体の位置によって影響を受ける。

本稿では(1)と(3)の問題点について述べる。(2)の問題点は画像ごとに対象世界のデータベースをもたせることで対処できるものと考えられる。

\*A Study on Vagueness in Description on Spatial Relationship of 2-Dimensional Images

†Noriyuki SHIMA Tomoichi TAKAHASHI

‡ATR Communication Systems Research Laboratories

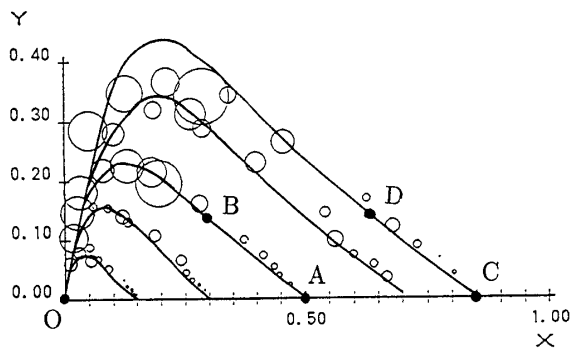


図 3: 『右』の主観評価実験結果

$$W(R, \theta) = (e^{-b\theta} + 1)(R - (1 - e^{-a\theta}))^3$$

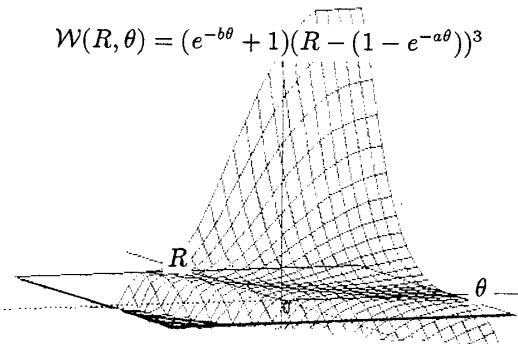


図 4: 『右』の度合の“差”を与える関数の概形

### 5 言葉の持つあいまい性の表現について

著者らは基本的な位置関係の言葉を定量化するための目安として、人間の定性的な言葉に対する主観評価実験を行なった[3]。図3にあるような4つの物体A,B,C,Dの基準物体Oに対する『右』の度合の比較をある被験者に行うと、物体Aと物体Bは同程度に右らしく、又物体Cと物体Dも同程度に右らしいという結果になった。物体Aと物体Cを比較してみると物体Aの方がCよりもはるかに右らしいという結果が得られた。しかしながら物体Bと物体Dを比較させると物体BとDはだいたい同じくらい右らしいという結果になった。

すなわち

- ある物体の基準物体に対する位置関係の度合を一意に決定するような関数は存在しない。
- ある物体の基準物体に対する位置関係の度合は他の物体の位置の状況で変化する。

ということが言える。

一般に『右』という言葉に関数として定義する場合、基準物体Oとある物体Xの2物体の位置関係から定義するのが普通であるが、その方法で物体AとB、物体CとDそして物体BとDが同じくらい右らしくなるようにその関数を定義すると、必然的に物体AとCも同程度の右らしさになってしまい、ここで述べたような人間の主観に対応できないのは明白である。

### 6 あいまい性を表現する関数

本稿では前述実験結果をもとに、基準物体Oとある物体Xとある物体Yの3つの物体の位置情報から物体Xと物体Yの位置関係言葉の度合の“差”を与える関数を定

表 2: 各2物体間の『右』の度合の差

	物体 A	物体 B	物体 C	物体 D
物体 A	0.000	-0.024	-0.476	-0.386
物体 B	0.024	0.000	0.000	-0.008
物体 C	0.476	0.000	0.000	-0.011
物体 D	0.386	0.008	0.011	0.000

(\*)i行j列の値はj列の物体がi行の物体にどの程度『右』の度合が高いかを示す。

義することを提案する。ここでは『右』の関数を例にとつて説明する。この『右』の関数では物体Xの物体Oに対する距離 $r_x$ と物体Yの物体Oに対する距離 $r_y$ の比  $R \equiv r_x/r_y$ と物体XのX軸に対する方位角 $\theta_x$ と物体YのX軸に対する方位角 $\theta_y$ の差  $\theta \equiv (\theta_x - \theta_y)/(\pi/2)$ の2つのパラメータで以下のように定義されている。

$$W(R, \theta) = (e^{-b\theta} + 1)(R - (1 - e^{-a\theta}))^3 \quad (1)$$

ここで係数aは右の度合の差を決定する、角度差と距離の2つの要因の重要度の比率に対応し、係数bは主観評価実験で得られた等高線のばらつき[3]に対応するものである。

図4にここで定義した『右』という言葉の関数の概形を示す。また表2には図3の4つの物体を対象にして、

(1)式で定義した関数から得られる各2物体間の右の度合の差を示している。この表から、物体AとB、物体CとD、物体BとDは同程度の『右』の度合で、物体AとCの『右』がかなり違うという主観評価実験の結果に沿うようになっていることが分かる。

### 7 おわりに

本稿では(1)の問題点に対しては、3物体の位置情報をもとに位置関係の度合の“差”に関数として定義することを提案した。ここで提案した手法での問題点(3)に対する手法として、心理学実験でよく使用される一対比較法の手法[4]をもちいることを現在検討中である。また今回提案した関数は物体のもつ大きさという要因を考慮に入れてない。物体のもつ大きさを考慮して位置関係を記述するシステムの実現も今後の課題である。

### 参考文献

[1] 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)

[2] 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”, 第3回人工知能学会全国大会, 9-1 (1989)

[3] 島, 高橋, 小林他: “ヒューマンインター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, pp.19-24 (1988)

[4] J.P. ギルホード著: “精神測定法”, 秋重義治 監訳, 培風館