論文

ファジィネスの高い蓄積データを 対象にした情報検索技法 ——絵画データベースを例として——†

磯本 征雄*¹ 野崎 浩成*¹ 吉根 勝美*² 永井 弘人*³

従来のファジィ・データベースは、検索条件と蓄積データの属性値の一致を基準に情報検索が行われてきた。しかし、データのファジィネスが大きい場合には、“一致度”を目安にした情報検索では、しばしば我々の日常感覚に合わない検索結果になる。属性値の曖昧さが大きい場合の情報検索の改善を図るために、筆者らは、蓄積データおよび検索条件のファジィな属性値をラベルとメンバーシップ関数で表記し、蓄積データが検索条件を満たす度合いを満足度として定義することで、新しい情報検索手法を定式化した。本論文では、絵画を実例として、満足度に基づく情報検索の有効性と実用性を議論する。

キーワード：美術作品、画像データベース、ファジィ論理、情報検索、データ構造、ファジィ数、ファジィ集合

1. はじめに

本論文では、極めてファジィネスの高い蓄積データを扱うデータベースを対象に、ファジィ情報検索の新技术を提案する。従来のデータベースでは、検索条件と蓄積データの属性値の一致度が情報検索の目安であった。この事情は、属性値をファジィ化したファジィ・データベース^{1,2,3,4,5,6,7)}でも同じである。しかし、蓄積データのファジィネスがさらに大きくなると“一致度”の基準はあいまいになり、可能性測度や必然性測度によるファジィ情報検索³⁾では、しばしば日常感覚に合わない検索結果になる。本論文で筆者らは、ファジィネスの大きなデータで起こる情報検索の困難を解決するために、蓄積データが検索条件の許容範囲内にあることを目安にした“満足度”による情報検索技法の定式化を試みる。

システム構成やデータ構造について、画像データベースの共通点を一般的に論ずることも可能である^{11,12)}。しかし、医用画像データ¹³⁾、地図データ¹⁴⁾、文献画像データ¹⁵⁾、金属材料画像データ¹⁶⁾の例に見るように、多くの場合に蓄積データの特徴に合わせた個別的議論が必要である。本論文の画像データベースでは、極めてファジィネスの大きな絵画を蓄積データとして^{8,9,10)}、データ属性の複雑・多様・曖昧さを一括してファジィネ

スで取り扱うことに特徴がある。

画像データベースにも、書誌情報のキーワードによる伝統的かつ基本的な情報検索を用いたものもある¹⁷⁾。一方、絵画にふさわしい情報検索法として、検索条件をラフスケッチで与え、これと絵画中の被描写体との形状一致を目安にした内容検索¹⁸⁾あるいは画像検索^{19,20,21)}の方法もあるが、これには画像理解という困難な課題が付随している²²⁾。別の方法として、画像の感性情報を名詞や印象語で表象し、これらと絵画の関連度を変量とした因子分析・多変量解析・回帰分析による風景画²³⁾や顔画像²⁴⁾の情報検索もある^{18,23,25,26)}。この中にはファジィ理論の応用^{22,24,25)}も見られるが、いずれの方法も、検索条件と蓄積データの“一致”や“類似”が基準であり、本論文における“満足度”による方法とは異なる。

本論文の情報検索の特徴は、データ属性をファジィ集合とそのメンバーシップ関数で表し、蓄積データが検索条件の許容範囲内にあることを目安にする満足度を導入したことである。検索条件のファジィ集合とその補集合を考えに入れて、蓄積データが検索条件に“包含される可能性”と“包含されない可能性”をもとに、前者を両者の和で除算した値を満足度とした。この方法は、メンバーシップ関数がシングルトンの場合に、従来のクリスプなキーワード検索と同等になる。本論文の方法は、非常にファジィな検索条件や蓄積データの数理モデルであり、美術鑑賞や美術評論など漠然とした内容の情報検索に適している。

本文第2章では、ファジィネスの大きなデータに関する取扱いの課題とその表記法を述べ、第3章でファジィ検索技法の定式化を議論する。第4章で、絵画を蓄積データとした実践的方策を論じ、本検索技法の実

† Information Retrieval Method for Highly Fuzzy Stored Data — A Case Study in Paintings Database —
Yukuo ISOMOTO, Hironari NOZAKI,
Katsumi YOSHINE, and Hirohito NAGAI

*¹ 名古屋市立大学 計算センター
Nagoya City University

*² 南山大学 経済学部
Faculty of Economics, Nanzan University

*³ 愛知県立港養護学校
Aichi Kenritu Minato-Yougogakko

用面での有効性を確認する。

2. ファジィ・データの取り扱い

本論文の画像データベースが対象にしている蓄積データは、図1に示す美術作品の絵画である。絵画データは、原画のイメージ情報の外に、表題、制作年、作家、所蔵美術館などの書誌情報、被描写体の场景情報、画像の雰囲気を示す感性情報からなる^{8,9,10)}(図2参照)。絵画では、医療用写真¹³⁾、設計図や指紋や地図¹⁴⁾、

図1 本論文で蓄積対象とするデータの例

学術資料写真¹⁵⁾、顔写真²⁴⁾などに比較してデータ属性のファジィネスが非常に大きい。情報検索についても、美術鑑賞とか美術評論などのファジィな表現に合わせる必要である。例えば、被描写体の配置は、“画面右上部”のように制約が大雑把で緩い。制作年には、“12世紀中期”のように、しばしばファジィ集合の得意とする表現が使われる。絵画の雰囲気を表す“明るい”や“厳しい”などの用語も、解釈の曖昧さは避けられない。これらデータ属性の複雑・多様・曖昧さを一括してファジィネスとしてファジィ集合のラベルとメンバーシップ関数で表すことで、ファジィなデータに対する情報検索の改善が可能になる^{8,9,10)}。以下に本章では、メンバーシップ関数の表記法について、情報検索時におけるデータ属性の振る舞いの違いを配慮して、数値的変量と言語的真理値に分けて議論する。

(a) 数値的変量：ファジィ数

年代、長さ、大きさなどを数値的変量とみなし、これらの可能性分布をメンバーシップ関数として次式で表記する(図3参照)。

図2 絵画のデータベースにおける蓄積データの種類

図3 ファジィ数の形状 (式 (1) 参照)

$$F(x; X_1, X_2, X_3, X_4) = \begin{cases} 0 & \text{for } x \leq X_1 \\ (x - X_1) / (X_2 - X_1) & \text{for } X_2 < x < X_2 \\ 1 & \text{for } X_2 \leq x \leq X_3 \\ (X_4 - x) / (X_4 - X_3) & \text{for } X_3 < x < X_4 \\ 0 & \text{for } X_4 \leq x \end{cases} \quad (1)$$

この形状は台型であり，“ほぼ X_2 から X_3 くらいの値らしい”ことを表す。特殊な場合として， $X_2 = X_3$ では三角型となり，“大体 X_2 くらい”を表す。極限值 $X_2 \rightarrow -\infty$ は“大体 X_3 以下らしい”を，極限值 $X_3 \rightarrow +\infty$ は“大体 X_2 以上らしい”を表す。

(b) 確信度および言語的真理値

絵画では，“朗らか”や“明るい気分”などの叙情的雰囲気の強弱の扱いが重要である^{23,26)}。蓄積データは，鑑賞者や評論家の確信度を値域 $[0,1]$ の変量として，式(1)のファジィ数で表す。一方，検索者の叙情的感性は，しばしば“やや明るい気分”や“とても朗らか”の様に言語的に表現される。そこで，システム内部では，“「明るい気分」は「やや真らしい」とか「「朗らか」は「とても真らしい」と読み換え，言語的真理値^{27,28)}(表1参照)を用いた蓄積データや検索条件の属性値の扱いをも許す。

(c) 存在可能性関数

絵画内の被描写体の配置データは，2次元平面 X - Y 座標系上の座標点 (x, y) を台集合とするメンバーシップ関数で表し，これらを存在可能性関数と呼ぶ。存在可能性関数は，式(1)を2次元に拡張した次式で定義する(図4参照)。

$$f(x, y; X_1, X_2, X_3, X_4, Y_1, Y_2, Y_3, Y_4) = F(x; X_1, X_2, X_3, X_4) \times F(y; Y_1, Y_2, Y_3, Y_4) \quad (2)$$

表1 言語的真理値のメンバーシップ関数

真理値	ファジィ集合	メンバーシップ関数
t_1	完全に真らしい	$\mu_{t1}(u) = \frac{1(u=1)}{0(u \neq 1)}$
t_2	とても真らしい	$\mu_{t2}(u) = u^2$
t_3	真らしい	$\mu_{t3}(u) = u$
t_4	やや真らしい	$\mu_{t4}(u) = u^{1/2}$
t_5	やや偽らしい	$\mu_{t5}(u) = (1-u)^{1/2}$
t_6	偽らしい	$\mu_{t6}(u) = (1-u)$
t_7	とても偽らしい	$\mu_{t7}(u) = (1-u)^2$
t_8	完全に偽らしい	$\mu_{t8}(u) = \frac{0(u \neq 0)}{1(u=0)}$
t_9	分からない	$\mu_{t9}(u) = 1$

図4 存在可能性関数の2次元表示

図4左は，式(2)で表される座標点 (x, y) 上の帰属度(メンバーシップ値)を示す。図4右は，左図の X - Y 座標への射影である。座標点 (X_1, Y_1) ， (X_4, Y_1) ， (X_4, Y_4) ， (X_1, Y_4) をこの順に結ぶ矩形外は存在可能性0の領域で， (X_2, Y_2) ， (X_3, Y_2) ， (X_3, Y_3) ， (X_2, Y_3) をこの順に結ぶ矩形内は存在可能性1の領域である。その中間域は，ファジィな存在可能性0~1の領域である。この定義により，存在可能性関数は，式(1)の数値的変量の場合と同じ数学的演算が可能になる。

3. 情報検索のファジィ論的方法

従来のクリスプなキーワード検索では，検索条件と蓄積データのキーワードが，“完全一致”，“前方一致”，“中間一致”，“後方一致”した時に情報が見つかったとされている。一方，被描写体の配置，叙情的雰囲気「暖かい」や「楽しい」，制作年「1455~1458年頃」などファジィネスの大きなデータでは，キーワードの“一致”の意味は不明瞭になり，従来の情報検索では議論できなくなる。図5は，ファジィネスの大きなデータについて，この事情を具体的に示す事例である。図5の太線は，検索条件「1430~1470年頃」である。図5左に，2つの蓄積データ「1440年頃(細線)」と「1455~1460年頃(破線)」を並列に図示した。図5右には，同様に2つの蓄積データ「1440年頃(細線)」と「15世紀頃(破線)」を並べて図示した。

ファジィ情報検索では，検索条件が蓄積データに一致する度合いを，“可能性測度”と“必然性測度”で評価する方法がとられてきた^{3,27,28,29,30,31,32,33,34)}。こ

図5 検索条件と蓄積データの積集合

の方法を図5左に適用すると、2つの蓄積データの可能性測度と必然性測度は共に1で、検索条件と蓄積データの一致度のよさを示す。同様に、図5右の“1440年頃”の可能性測度と必然性測度は共に1であり、完全に検索条件範囲内にあることを納得させる。一方、“15世紀頃”では可能性測度1で必然性測度0であるが、検索条件の範囲外と範囲内の比率が同程度であり、単純に可能性測度と必然性測度のみを見たのでは、“1440年頃”と“15世紀頃”を対比させるための十分な結果が得られない。そこで本論文では、検索条件が蓄積データを包含する度合いを検索の目安として顧慮した。

検索条件 C が蓄積データ D を包含する度合いには、 $I_1(D, C) = \|D \cap C\| / \|D\|$ の $I_1(D, C)$ で評価する方法がある³⁵⁾。この方法によると、検索条件「1430～1470年頃」に対して、図5左の蓄積データ「1440年頃(細線)」と「1455～1460年頃(破線)」および図5右の蓄積データ「1440年頃(細線)」は、共に包含度 $I_1(D, C)$ が1となり、完全に検索条件に包含される常識的な結果を得る。図5右の「15世紀頃(破線)」は、検索条件をかなり外れているが、包含度は約0.5という妥当な値になる。一方、図6の検索条件と蓄積データが全く同じ三角型ファジィ数「1450年頃」の例では、包含度 $I_1(D, C)$ は1となり、蓄積データは検索条件に完全に包含されることになる。しかも、検索条件と蓄積データが共にクリスプな「1450年」の場合でも、包含度 $I_1(D, C)$ はやはり1である。この結果に対して筆者らは、ファジィな属性とクリスプな属性を比較した場合、クリスプな属性に対してファジィな属性が検索条件を満足する度合いを少なく見積もるのが自然な評価であると考ええる。以下にこの考え方に沿って、筆者らは、検索条件が蓄積データを包含する可能性と包含しない可能性に基づいて、新たに満足度による評価方法を定式化する。

検索条件 C のファジィ属性をファジィ集合 a_c と台集合 $\{u\}$ 上のメンバーシップ関数 $\mu_c(u)$ とし、 C の否定のファジィ属性をファジィ集合 $\neg a_c$ (a_c の補集合) とメンバーシップ関数 $(1 - \mu_c(u))$ とする。 a_d が a_c に包含される可能性 p_{1cd} と、 a_d が a_c に包含されない可能性 p_{0cd} を定義して、 a_d が a_c の許容範囲に含まれる度合いを満足度 π_{cd} として次式で計算する。

$$\pi_{cd} = p_{1cd} / (p_{1cd} + p_{0cd}) \quad (3)$$

この式に従えば、図6の例における満足度は2/3であり、上記の議論に沿った結論を与える。

さて、数値的変量と真理値では、メンバーシップ関数に意味の違いがある。数値的変量のメンバーシップ

関数は、“ほぼ40～50メートルくらい”や“明治初期”や“群集”や“集落”や“中年”など数量に関わる用語をラベルとしてデータ属性値の可能性分布を示し、台集合の値域は任意である。一方、真理値のメンバーシップ関数は、真理値のファジィ化であり、検索条件や蓄積データの真偽の不確定ないしは確信度のあいまいさを示し、台集合の値域は $[0, 1]$ である。これら数値的変量と真理値の違いを満足度の計算方法にも反映させて、 p_{1cd} と p_{0cd} を以下の様に評価する。

台集合 $\{u\}$ が数値的変量の場合

数値 u のメンバーシップ関数は u の許容度や可能性分布を示す。したがって、検索条件と蓄積データの論理積も、蓄積データが検索条件の部分集合である許容度または可能性を示しており、 u の値域全域で同等にメンバーシップ値に比例した重みで扱う必要がある。そこで、蓄積データが検索条件に含まれる可能性 p_{1cd} は、台集合 $\{u\}$ 上の総和をとって次式で計算する(図7左参照)。

$$p_{1cd} = \int_{u_{\min}}^{u_{\max}} \{\mu_c(u) \wedge \mu_d(u)\} du. \quad (4-a)$$

同様に、蓄積データが検索条件に含まれない可能性 p_{0cd} を次式で計算する(図7右参照)。

$$p_{0cd} = \int_{u_{\min}}^{u_{\max}} \{(1 - \mu_c(u)) \wedge \mu_d(u)\} du. \quad (4-b)$$

積分区間 $[u_{\min}, u_{\max}]$ は、メンバーシップ関数が $0 < \mu_c(u)$ 及び $0 < \mu_d(u)$ を満たす台集合 $\{u\}$ の領域の上限と下限である。ここで文献³⁵⁾に対比すると、 p_{1cd} は $\|D \cap C\|$ に対応し、 $\mu_c(u)$ の積分値が $\|D\|$ に対応する。

台集合 $\{u\}$ が真理値の場合

真理値 u のメンバーシップ関数は、 u に関する確信度を示す。蓄積データと検索条件の論理積は、真理値 u において蓄積データが検索条件を満たす確信度を与えており、値域 $[0, 1]$ 内の最大値を指標にするのが一般

図6 検索条件と蓄積データの属性値が共に三角型ファジィ数で完全に重なり合った状態を示す。

—— 検索条件と蓄積データのメンバーシップ関数
 検索条件の補集合のメンバーシップ関数

的な考え方である。そこで、蓄積データが検索条件を満たす可能性 p_{1CD} を次式で計算する(図8左参照)。

$$p_{1CD} = \bigvee_u \{ \mu_C(u) \wedge \mu_D(u) \}. \quad (5-a)$$

同様に、蓄積データが検索条件に含まれない可能性 p_{0CD} を次式で計算する(図8右参照)。

$$p_{0CD} = \bigvee_u \{ (1 - \mu_C(u)) \wedge \mu_D(u) \}. \quad (5-b)$$

ただし、上式の p_{1CD} と p_{0CD} は、検索条件と蓄積データのファジィ集合の一致度を測る目安である可能性測度 p_{1CD} と必然性測度 $(1 - p_{0CD})$ に対応付けられる^{3,29,30)}。

式(4)~(5)で計算される p_{1CD} と p_{0CD} の和は1になるとは限らないが、式(3)のファジィ満足度 π_{CD} では、領域 $[0,1]$ に規格化されている。 π_{CD} は、メンバーシップ関数の帰属度が1未満の領域で、前述の $I_1(D,C)$ とは異なる振る舞いをする。 π_{CD} の特徴は、付録で述べることにして、本文での議論を割愛する。なお、存在可能性関数に以上の議論を適用する場合は、数値的変量 u を位置座標 (x,y) と読み換えて計算する。

ところで、メンバーシップ関数 $\mu_C(u)$ と $\mu_D(u)$ が共にシングルトンの場合には、

$$(p_{1CD}=1 \text{ and } p_{0CD}=0)$$

または

$$(p_{1CD}=0 \text{ and } p_{0CD}=1)$$

となり、式(3)は通常のキーワード検索における次の関係に帰着する。

$$\text{if } a_C \cap a_D \neq \phi \text{ then } \pi_{CD}=1, \quad (6-a)$$

$$\text{if } \neg a_C \cap a_D \neq \phi \text{ then } \pi_{CD}=0. \quad (6-b)$$

クリスプな属性 a_C と a_D について、 $a_C \cap a_D \neq \phi$ と $\neg a_C \cap a_D \neq \phi$ は排他的であり、いずれか一方のみが成り立つ。

4. 実現の方策

本ファジィ・データベースの特徴は、データ属性の記述にファジィ集合のラベルとそのメンバーシップ関数を含め、式(3)の満足度に基づいて情報検索を行うことにある。これに伴い、条件検索の手順を、ファジィ集合名をキーワードとするクリスプ検索と、式(3)で定義される満足度に基づいて序列化するファジィ検索の2段階に分ける。以下に本章では、データベースの具体的な実現の方策について述べる。

4.1 蓄積データの格納

本画像データベースには、いわゆるキーワードの外に、数量の可能性分布、意味の強弱、または真偽の確信度を示すファジィ集合とそのメンバーシップ関数が

図7 検索条件およびその補集合と蓄積データの関係 (メンバーシップ関数が式(1)で表される場合)

図8 検索条件およびその補集合と蓄積データの関係 (メンバーシップ関数が式(1)および言語的真理値で表される場合)

表2 画像データの表形式による記述

コード番号	被描写体番号	項目名	データ値	属性値	メンバーシップ関数			
					X1	X2	X3	X4
01	0	表題	〇〇〇〇					
01	0	作者	△△△△					
01	0	年代		昭和40年代	1963	1965	1975	1977
01	1	名称	山					
01	1	位置		左上				
01	1	標高	高い	とても真らしい	0.6	0.8	0.9	1.0
01	2	名称	川					
01	2	位置		右下				
01	2	全長		約50km	45	50	50	55
01	3	名称	家					
01	3	位置		左下				
01	0	叙情	楽しい	約0.7	.6	.7	.7	.8
01	0	叙情	暖かい	やや真らしい	0.3	0.4	0.5	0.6

格納される。表2は、蓄積データのデータ構造を示す。表2において、コード番号は箇々の絵画の識別番号である。被描写体番号は、画像内の箇々の被描写体ごと

に付けた番号で、同じ被描写体の属性には同じ被描写体番号が付けられる。項目名は、そのレコードのデータ属性名を示す。データ値は、蓄積データの内容である。その右欄に、ファジィ集合のラベルを属性値に、さらにその右欄にメンバーシップ関数のパラメータ X_1, X_2, X_3, X_4 を与える。ただし、被描写体番号 0 は、個々の被描写体に関わりを持たないデータであることを示す。

4.2 クリस्प検索

クリस्प検索は、本データベースの情報検索第1段階であり、ラベル(つまりファジィ集合名)をキーワードとする従来のキーワード検索によって、検索条件と同じ属性を持つ蓄積データを選び出す。絵画の場合には、被描写体の名称、作品の題名、制作年、作家、所有者、絵画の雰囲気“力強さ”、“懐かしさ”、“優しさ”などがキーワードとなり、シソーラスによる同義語・類似語・関連語の参照も行う。クリस्प検索は、次節に述べる第2段階ファジィ検索の満足度計算に際して、検索条件と蓄積データの同じ属性のファジィ・データを正しく矛盾なく対応させるための前処理でもある。また、クリस्प検索で得られた蓄積データは、式(3)で計算される満足度がゼロの場合でも、ソーティングによる順序付けの優先度は低くなるが、検索条件に合ったデータとして扱う。

4.3 ファジィ検索

ファジィ検索は、前節のクリस्प検索に続く第2段階であり、次の処理をする。

- ① 式(3)～(5)により満足度を計算する。
- ② 満足度に基づいてソートする。
- ③ 満足度の高い順に提示する。

検索者は、こうしてソート順に提示された画像を見ながら、最終的に自分の欲している絵画を決める。

ここで、ファジィ検索条件は、被描写体の存在可能性関数、ファジィ数、言語的真理値である。存在可能性関数は、ディスプレイ画面を1枚のキャンバスとみなし、各々の被描写体の名称ごとに、マウスで内枠矩型と外枠矩型を指定する。ファジィ数の場合は、ディスプレイ画面上の座標に式(1)の X_1, X_2, X_3, X_4 をこの順序でマウスを使って指定する。言語的真理値は、アイコン・メニューから適切な用語を選ぶ。

検索条件が複数個ある場合には、各条件 C に重視度 w_C を付け、重視度と満足度をもとに、総合満足度 T_D を次式で計算する。

$$T_D = \sum_{c=1}^{C_{\max}} w_C \times \pi_{CD} \quad (7)$$

ここで、 $C_{\max}$ は、検索条件の総数である。蓄積データ D の総合満足度 T_D を第1ソートキーにして、以下重視度 w_C の大きい順に満足度 π_{CD} を第2, 3, ... キーとして降順にソートし、これを検索結果とする。ソートされた原画データは、この順序にディスプレイ画面上に表示され、目当ての検索画像であるか否かを検索者自身が決める。画面表示の際には、逐次前進するだけでなく、後退や飛び越し等の表示順序の操作も可能である。図9は、存在可能性関数とそれによって検索さ

図9 検索条件の存在可能性関数とその検索結果

れた絵画の上位3件までを示した実例である。

本画像データベースのひな型では、ウィンドウズの稼働するパソコンが使用された。実用性の観点から、場景情報、書誌情報、感性情報による情報検索を効率よく処理するために、汎用データベース管理システム“ACCESS”を導入した。蓄積データは、中学校から高等学校にわたる美術教育の教科書から抜き出した約300件の画像データを蓄積している。

5. まとめ

本画像データベースの特色は、データの複雑さ・曖昧さ・多様さを、ファジィ集合とそのメンバーシップ関数の形状に集約した点にある。これに伴い、ファジィ検索条件及びその補集合を共に評価し、検索条件を許容限界とみなしたファジィネスの大きなデータに有効な情報検索技法を定式化した。しかも、メンバーシップ関数がシングルトンになる極限で、通常の文献情報データベースのキーワード検索に一致する論理的連続性も保証されている。

本論文の主題は、ファジィ集合とメンバーシップ関数による情報検索技法の定式化であった。絵画のようなファジィネスの大きなデータを蓄積する際には、データを大雑把に記述できても厳密な表現はほとんど不可能であり、その点において本論文の手法によるデータベース化が有力な手段となる。もちろん実用面では、メンバーシップ関数の形状決定が重要であり、利用状況を観察しながらメンバーシップ関数を汎化する学習機能も必要である。さらに式(7)の定式化では、ファジィ測度論^{29,30)}による数理モデル化の考察も残されている。そうした課題は、本論文の論点外としてここでの議論を割愛し、今後の課題として別の論文で述べることにした。

最後に、本研究は、美術教材のデータベース化を目的とした研究課題として、文部省科学研究費補助金(試験研究(B)(1)課題番号06558021)を得て進められたことを感謝を込めて付記する。

参 考 文 献

- 1) 深海, 馬野, 水本, 田中: Fuzzy データベース検索操作言語について, 信学技報, Vol.AL78-85(1979)
- 2) 宮本, 小西, 三宅: ファジィ命題索引にもとづく文書検索と画像検索, 日本ファジィ学会誌, Vol.3, 1, pp.98-107(1991)
- 3) 浅居喜代治編著: ファジィデータベース, ファジィ情報処理入門, オーム社, (1993)
- 4) Amihai MOTRO: VAGUE: A User Interface to Relational Databases that Permits Vague Queries, ACM Transactions on Office Information Systems, Vol.6, No.3, pp.188-214 (July, 1988)
- 5) 馬野, 深海: ファジィ・データベースの現状について, 日本ファジィ学会誌, Vol.3, No.1, pp.2-14(1991)
- 6) 宮本, 三宅: ファジィ情報検索について, 日本ファジィ学会誌, Vol.3, No.1, pp.15-26(1991)
- 7) 藤代, 白井, 池辺, 國井: 個別化されたファジィデータ管理のためのグラフ指向スキーマ設計, 日本ファジィ学会誌, Vol.3, No.1, pp.86-97(1991)
- 8) 中谷, 吉根, 磯本, 石井: 画像の被描写体を対象とするファジィ論的検索技法, アドバンスド・データベースシステム・シンポジウム '92論文集, pp.17-25(1992)
- 9) 磯本, 吉根: 画像データベースのためのデータ構造と情報検索技法, 日本ファジィ学会第10回ファジィシステムシンポジウム, pp.627-630(1994)
- 10) Y.Isomoto, K.Yoshine: Data Structure and Retrieval Method of Scenic Image Database Based on Fuzzy Set Theory, PROCEEDINGS of the International Joint Conference of Fuzzy Systems and the Second International Fuzzy Engineering Symposium, FUZZ-IEEE/IFES'95 Vol.II, pp.794-756(1995).
- 11) 岡崎: 画像データベース概論, 情報処理, Vol.33, No.5, pp.448-456(1992)
- 12) 野口: 画像データベースにおけるデータ表現・管理, 情報処理, Vol.33, No.5, pp.457-465(1992)
- 13) 西原: 医用画像データベースの現状と将来, 情報処理, Vol.33, No.5, pp.478-485(1992)
- 14) 東明: 地図情報データベースシステム, 情報処理, Vol.33, No.5, pp.486-496(1992)
- 15) 飯沢: 文献画像データベースシステム, 情報処理, Vol.33, No.5, pp.497-504(1992)
- 16) 稲田, 小林: 金属材料画像データベース, 情報処理, Vol.33, No.5, pp.505-511(1992)
- 17) 菊川, 川淵: 映像資料情報データベースシステムの開発, 信学技報, Vol.ET87, 4, pp.43-48(1987)
- 18) 加藤, 栗田: 画像の内容検索—電子美術館への応用—, 情報処理, Vol.33, 5, pp.466-477(1992)
- 19) 浦谷, 柴田, 野口, 相沢: 静止画像検索システム FORKS の試作, 情処学論, Vol.28, 7, pp.758-767(1987)
- 20) 平田, 原: 概略画像を用いた画像検索, 信学技報, Vol.DE92-2(1992)
- 21) 岡崎, 美濃, 板井: 風景画像の認識とその認識対象属性による画像検索—風景画像中の山を例として—, 信学技報, Vol.PRU86-118(1987)
- 22) 増井 他: ファジィ理論を用いた風景画の画像理解, 日本ファジィ学会誌, 第6巻5号, pp.1000-1011(1995)
- 23) 加藤, 栗田, 坂倉: フルカラー絵画データベース

- XART MUSEUM—色彩感と略画による画像対話—,
情報処理学会研究会報告, IE88-118, pp.31-38(1988)
- 24) 中山 他: ファジィ理論を用いた印象言語による顔画像検索システム, 日本ファジィ学会誌, 第6巻5号, pp. 974-983, (1994)
- 25) 前田, 石飛: ファジィデータベース検索のための最適モデルを用いた自然言語インターフェース, 日本ファジィ学会誌, Vol.3, No.1, pp.108-119(1991)
- 26) 田邊, 大谷, 岸本: 主観評価に基づく形状類似尺度を用いた類似画像検索法, NTTR&D, Vol.39, No.5, pp. 697-707(1990)
- 27) A.Kaufmann, M.M.Gupta (田中英夫監訳, 松岡浩訳): ファジィ数学モデル, オーム社(1992)
- 28) 本多, 大里: ファジィ工学入門, 海文堂出版(1989)
- 29) 菅野, 室伏: ファジィ測度論入門 [I], 日本ファジィ学会誌, Vol.2, No.2, pp.174-181(1990)
- 30) 菅野, 室伏: ファジィ測度, 講座ファジィ第3巻, 日刊工業新聞社刊(1993)
- 31) H. Prade and C. Testemale: Generalizing Database Relational Algebra for the Treatment of Incomplete or Uncertain Information and Vague Queries, Information Sciences, Vol.34, pp.115-143(1984)
- 32) H. Prade and C. Testemale: Representation of Soft Constraints and Fuzzy Attribute Values by Means

of Possibility Distributions in Databases, in Analysis of Fuzzy Information II, (ed. by J. C. Bezdek), CRC Press (1986)

- 33) H. Prade and C. Testemale: Application of Possibility and Necessity Measures to Documentary Information Retrieval, in Lecture Note in Computer Science No.286, Springer-Verlag(1987)
- 34) H. Prade: Lipski's Approach to Incomplete Information Data Bases Restated and Generalized in the Setting of Zadeh's Possibility Theory, Information Systems, Vol.9, No.1, pp.27-42(1984)
- 35) Didier Dubois and Henri Prade: Fuzzy Sets and Systems: Theory and Applications, Mathematics in Science and Engineering Volume 144, Academic Press, pp.22-24(1980)

(1994年11月10日 受付)

(1995年6月1日 再受付)

[問い合わせ先]

〒467 名古屋市瑞穂区瑞穂町山の畑1
名古屋市立大学計算センター
磯本 征雄
TEL : 052-872-5790
FAX : 052-872-5781

著者紹介

磯本 征雄 (いそもと ゆくお)

名古屋市立大学計算センター

1966年 岡山大学理学部物理学学科卒業。
1970年 大阪大学大学院博士課程単位終了中退。同年 同大大型計算センター助手, 同講師を経て, 1984年より名古屋市立大学計算センター教授。理学博士。データベース, 人工知能, 教育工学, ファジィ理論等の分野で研究に従事。情報処理教育, 情報リテラシー教育を担当。

吉根 勝美 (よしね かつみ)

南山大学 経済学部

1986年 名古屋工業大学情報工学科卒業。1988年 同大学院博士前期課程修了。1990年 同後期課程中退。同年 名古屋市立大学計算センター助手。1995年より南山大学経済学部講師。ファジィ理論に基づくデータ処理及び解析, 教育工学分野の研究に従事。

野崎 浩成 (のざき ひろなり)

名古屋市立大学計算センター

1991年 愛知教育大学教育学部情報科学科卒業。1995年 東京工業大学総合理工学研究科修士課程修了。同年より名古屋市立大学計算センター助手。教育工学分野の研究に従事。

永井 弘人 (ながい ひろひと)

愛知県立港養護学校

1981年愛知教育大学教育学部美術教室卒業。同年 愛知県立名古屋養護学校勤務。1989年上越教育大学大学院障害児教育コース修士課程修了。同年より愛知県立港養護学校勤務。日本特殊教育学会, 教育工学学会, 教育システム情報学会の各会員。

付 録:

数値的変量の満足度に対する

検索条件と蓄積データへのファジィネスの影響

第3章式(3)の満足度 π_{cd} は、検索条件を可能性分布の許容度の上限として、蓄積データがこの許容範囲内にある度合いを示す。式(3)中に可能性測度(p_{1cd})や必然性測度($1-p_{0cd}$)に相当する変量も現れるが、 π_{cd} の満足度の定義はこれらとは異なる特性を持つ。これらの理解のために、検索条件と蓄積データのファジィネスの相対的な大きさの違いの満足度への影響について、式(4)に基づいて幾つかの具体例を考察する。

① 検索条件のファジィネスが大きい場合

検索条件のファジィネスが大きくなると、蓄積データに対する制約が緩くなり、蓄積データが検索条件に含まれる可能性、つまり満足度も高くなる。図A1の例では、A、B、Cの3つの異なる蓄積データについて、ファジィネスの大きな検索条件との関連を見た。Aは、検索条件の周辺部でのみ重なりを持つが、検索条件およびその補集合との重なりがほぼ同程度で満足度も0.5程度となる。Bは、検索条件に完全に重なっているため、満足度は1になる。一方、Cは、検索条件と広い範囲で重なるが、補集合とも少し重なっており、その満足度は1よりやや小さくなる。

なお、検索条件のファジィネスが極端に大きくなる事例として考えられるのは、“UNKNOWN”の場合である。検索条件は全域で1となり、その補集合はゼロ(0)となって、蓄積データの満足度はすべて1となる。台集合全域で検索条件のファジィネスが1ということは、“制約条件の分からない状態”なので、すべての蓄積データがその条件を満たすと考えられるのも常識的処置といえよう。ただし、本画像データベースでは、第1段階のクリस्प検索で選ばれた蓄積データに限定して議論しており、この議論があくまでもファジィネスに関する事柄である点に注意を要する。

② 蓄積データのファジィネスが大きい場合

蓄積データのファジィネスが大きい場合には、図A2にも見るように、蓄積データは検索条件との重なりよりもその補集合との重なりが大きくなり、蓄積データが検索条件の許容範囲内である可能性つまり満足度も小さくなる。

蓄積データの属性があいまいになれば、検索条件を外れる部分も多くなり、その分だけ満足度も小さくなる。もし、図A2の状況において、蓄積データのファジィネスがさらに大きくなり、台集合の全域で“UNKNOWN”になったら、検索条件の補集合との重なりがますます大きくなり、満足度は限りなく0に近くな

る。

③ 蓄積データと検索条件のファジィネスがほぼ同程度の場合

蓄積データと検索条件のメンバーシップ関数がほぼ同じ場合には、図A3に見るように、蓄積データと検索条件が重なる部分によって満足度の大勢が決まり、両端の帰属度1未満のファジィな部分における補集合との重なり具合で満足度が1未満に押さえられる。式(4-a)における積分処理は、こうしたメンバーシップ関数の重なり程度に重みを付けて定量化したことによる。つまり、検索条件と蓄積データのファジィネスがほぼ同程度の場合には、満足度には、メンバーシップ関数の周辺部のファジィな部分の影響を見積もることになる。

ここでは、具体的事例に基づいて、定義式(3)の背景の一部を示した。③の例では蓄積データと検索条件のファジィネスがほぼ同程度な場合であり、文献³⁵⁾との

図A1 検索条件のファジィネスが大きい場合

—— 蓄積データのメンバーシップ関数
 —— 検索条件のメンバーシップ関数
 検索条件の補集合のメンバーシップ関数

図A2 蓄積データのファジィネスが大きい場合

—— 蓄積データのメンバーシップ関数
 —— 検索条件のメンバーシップ関数
 検索条件の補集合のメンバーシップ関数

図A3 検索条件と蓄積データのファジィネスがほぼ同程度の場合

—— 蓄積データのメンバーシップ関数
 —— 検索条件のメンバーシップ関数
 検索条件の補集合のメンバーシップ関数

対比で特色ある結果を得る。一方、①と②の議論では、可能性測度や必然性測度よりも、式(3)の満足度による評価の方が自然な結果を与えている。敢えて可能性測

度や必然性測度に言及するならば、満足度の同じ2つの蓄積データに対して、これらの違いをさらに詳細に知る助けにはなるであろう。

Information Retrieval Method for Highly Fuzzy Stored Data — A Case Study in Paintings Database —

by

Yukuo ISOMOTO, Hironari NOZAKI, Katsumi YOSHINE, and Hirohito NAGAI

Abstract :

Information retrieval of a conventional fuzzy database is executed on the base of equality of attribute values of stored data and the one of a retrieval condition. But, when the attribute values are very fuzzy, the information retrieval on the base of "equality" often gives us unreasonable results in comparison with our ordinary common sense. In order to overcome the difficulty of the information retrieval for very large fuzzy data, the authors express the fuzzy attribute value of stored data and a retrieval condition in a label and its membership function, and formulate the new method of fuzzy information retrieval by defining the satisfaction grade, that estimates how the attributes of stored data satisfy the one of a fuzzy retrieval condition. By example of paintings, the authors discuss usability and effectiveness of the information retrieval formulated here.

Keywords : Art objects, Image database, Fuzzy logic theory, Information retrieval, Data structure, Fuzzy number, Fuzzy set

Contact Address : **Yukuo ISOMOTO**

Nagoya City University

Yamanohata 1, Mizuho-cho, Mizuho-ku, Nagoya 465, JAPAN

TEL : 052-872-5790

FAX : 052-872-5781