

No. **620**

感性表現ベースにおけるファジー検索の試み

by

松田紀之

March 1995

Institute of Socio-Economic Planning
Discussion Paper Series
No. 620

感性表現ベースにおけるファジー検索の試み

松田 紀之

March, 1995

本研究は'94年度筑波大学学内プロジェクトS”認知の工学的実現（サイバニクス）”の一環として行われたものある。

開発に当たって社会工学類 4年生の瀬野真氏（現在日本オラクル社）の協力を得た。

連絡先：筑波大学社会工学系

305 つくば市天王台1-1-1

tel: 0298-53-5170, 5374 fax: 0298-55-3849

email: mazda@shako.sk.tsukuba.ac.jp

感性表現ベースをにおけるファジー検索の試み

要 約

松田は画像と言語情報の補完的役割に注目しファッション解説特集を感性表現ベースとして構築してきた(e.g.,松田,1991). 特集の持つ情報展開の体系性や表現ベースの拡充性, 双方向的利用形態を考慮し, 基盤環境としてMacintoshのHyperCardを採用した.

HyperCardはオブジェクト指向による視覚的プログラミング環境を提供するもので, 音声を加えたマルチメディア情報基盤として利用価値が高い. しかし, 感性表現ベースの設計原理は同等の能力を有する基盤上に移植可能である.

感性表現ベースは, 蓄積された情報に対する利用者の自由な発想による検索要求に応えるインターフェイスを備えていなければならない. 本研究の目的は, 特集解説自体における情報の部分性と検索要求自体のあいまいさを処理することにある. 次の3項目を中心に試作を行った:

- ・記事クラスター毎のフレーム形式による情報記述
- ・ファジィ・シソーラスによるファジィ検索語群の生成
- ・検索項目の指定順序を反映した, 検索優先度の設定

排除的な要求や記事の体系については処理に不備な点が残されており今後の課題である.

感性表現ベースにおけるファジー検索の試み

1. 感性表現ベースと検索インターフェイス

広くデザインの世界では過去の類例から発想の契機を得ることが多いという。そこでは画像情報のみならず、言語による解説が重要な役割を果たしている。松田はこうした情報の補完的役割に注目し、資料として比較的体系だったファッション解説記事のデータベース化を進めてきた (e.g., 松田, 1991)。以降これを感性表現ベースと呼ぶ。ファッション解説記事の特色は、感性的なあるいは直感的な理解になじみやすいように情報が配置されている点である。

感性は理性と対立させられることが多いが、理性を欠くものとして反意語的に捉えていたのではその本質を見失うことになる。感性はむしろ、人の基本的な認知あるいは知識を指す概念とし、理性はその特殊な形態として把握すれば感性表現ベースの意義を理解しやすい。両者の特徴は次のように表現される：

理性：非冗長性、一貫性、厳密性、無矛盾性を指向する

感性：冗長性、非一貫性、あいまいさ・不明確さ、矛盾を許容する

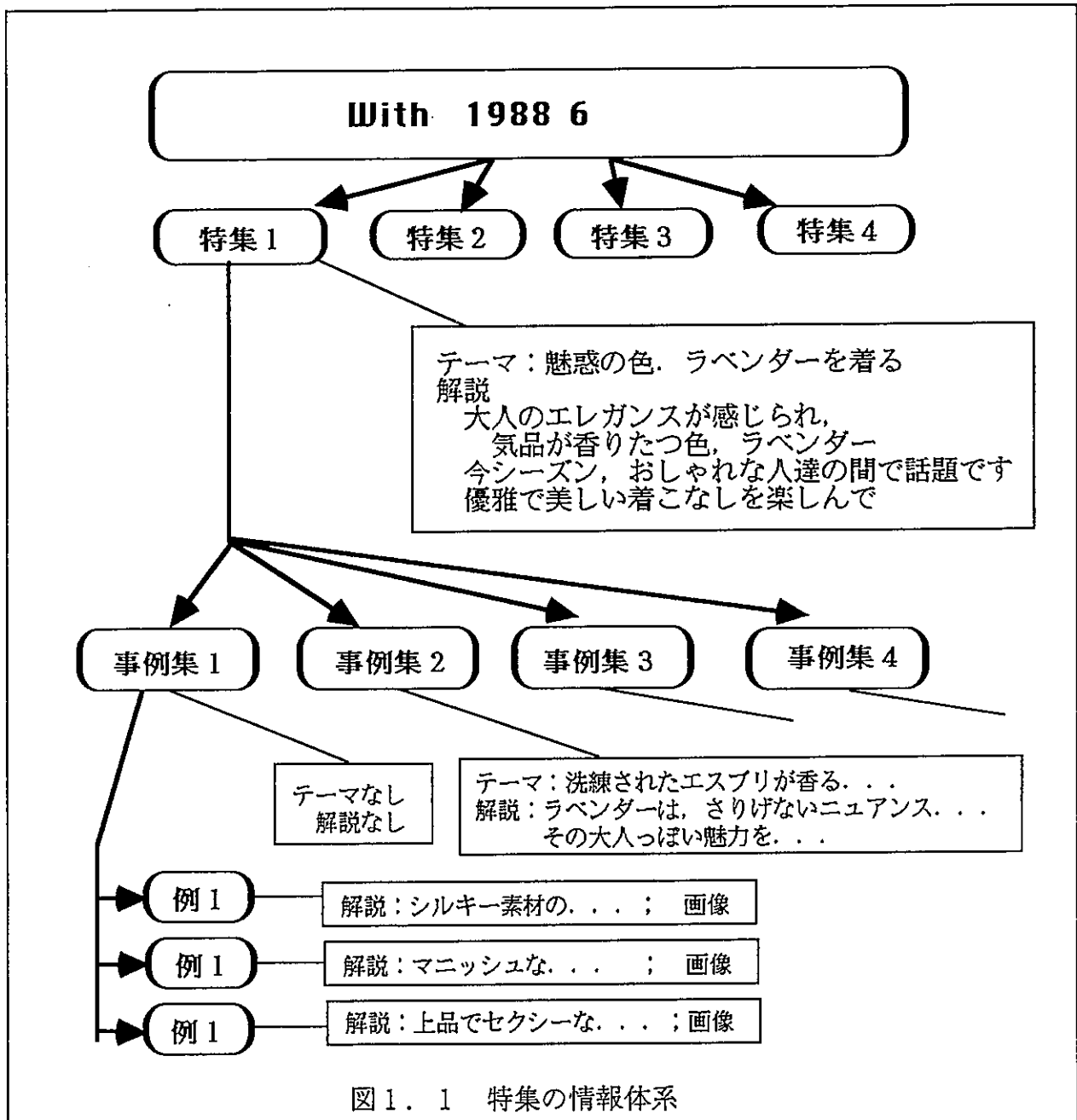
この章では感性表現ベースを概説し、実用化に向けた検索インターフェイスが備えるべき機能について検討する。

1-1. 感性表現ベース

感性表現ベースは、資料の系統性・継続性・事例の豊富さなどの要因を考慮し、主要月刊女性雑誌 (More, with, Classy) のファッション解説記事をHyperCard上に移植したものである。一般に月刊雑誌のファッション事例集では、感性情報は主として画像によって担われており、解説記事は稀にしか顧みられることのない従の役割しか持たないような印象を持たれている。しかし、画像の中で注視する焦点や、素材、着心地・着こなし、期待できる印象・TPOなどを手際よく伝達するためには、言葉による解説が必要である。一方、アイテムの形状とか模様や形状については、視覚による把握が優れている。ファッション事例集はこうした意味で、言語表現と画像が極めて効果的に相互補完的に組み合わせられた専門知識の表現とみなせる。

ファッション解説記事は、各号毎にTPOに応じた幾つかの特集から構成されている

(図1.1参照)。特集はさらに中心となるアイテムや、素材又は色、デザイン、演出方法等のキーにより複数の具体例の集まり (以下クラスターと呼ぶ) に分けられている。一



つのクラスターは、通常見開き 1 ページに収まるように編集されていて見渡しによる把握を容易にしている。

感性表現ベースの画面構成はこのクラスター単位である（図 1. 2 参照）。クラスターを総括する特集全体のテーマとその解説はクラスターごとに添えられている。情報の節約の観点からは冗長ではあるが、簡単にクラスター単位での電子的切抜きが行える利点を重視した。

画像は基本的に見開きページ毎にスキャンし、画面に呈示されているクラスターの発行情報（雑誌名・発行年号・クラスター番号）とファイル名でリンクが取れるようフォルダー

系列を設定している。クラスター画像を特集毎に集成したQuickTime仕様のファイルも用意されているので、実際に雑誌をくるような感覚での閲覧が可能である。個々の具体例の画像も利用可能である。

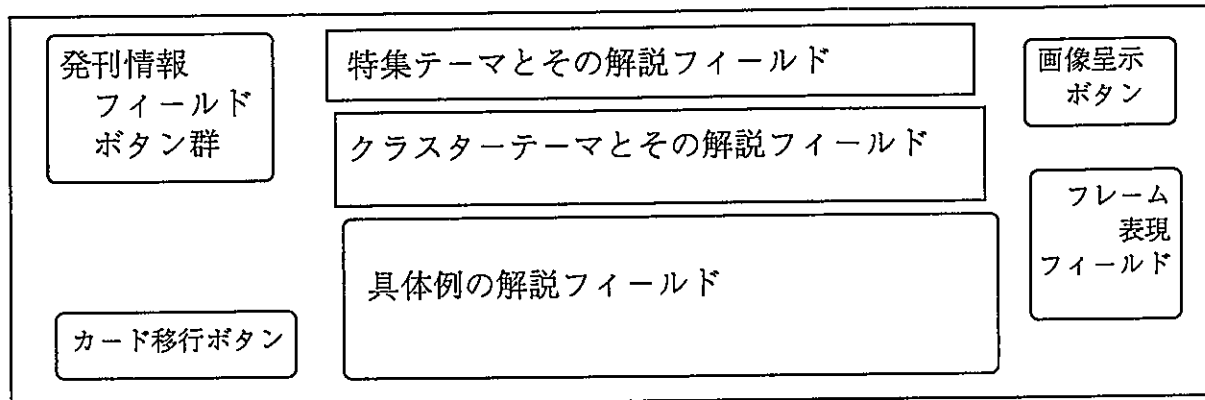


図 1. 2 感性表現ベースの画面構成

このように画像と連携した記事閲覧機能や、クラスターあるいは雑誌の移行機能を利用すれば、感性表現ベース自体でかなり現実的な情報が取得できるよう設計されている。ただし、特定の語句やその結合による検索機能を備えなければ発想支援の道具とはならない。本研究の目的は、感性的理解を妨げないインターフェイスのもとで、この機能を提供することにある。

1-2. 感性的検索要求

デザイナーが発想を具体化する契機として感性表現ベースを利用することを想定すると、検索要求は

- 青いスーツでさわやかな印象を与えるもの
- フレアースカートのきこなし
- 麻を主体にしたアイテムの印象

といった形で記述できるであろう。具体的な語句はさまざまであるが、大きく分類すればアイテム、色、素材、着こなし・印象

に分けられる。これがシステム側のテンプレートとなるが、この例にあるように一定の順序でユーザがすべての項目を明示できることは稀であろう。従って、項目の欠損をエラーとして扱わないインターフェイスが求められる。また、項目はユーザが任意の順序で選択できることが望ましい。この選択順序は、情報獲得の優先度を反映するものとして解釈することができる。

感性的検索要求の顕著な特徴として、語句の曖昧さあるいは不明確さがあげられる。例

例えば、「青い」と指定した場合でも、必ずしも青いアイテムの事例だけを求めているのではなく、近接した青緑やラベンダーの事例も暗に想定されていると考えるのが自然であろう。操作的に言い換えれば、そのような関連語句をもつ事例の要求との一致度を、最大ではなくとも高いものとして処理するのが妥当である。

色合いや印象・きこなしの語句間のネットワークが、もっぱら個人的な経験に依存しているのに比べ、アイテムの体系は業界専門家の間で比較的一致している（図1. 3参照）。アイテムに固有の部分（袖、襟など）に対する名称についても同様である。しかし、いずれのタイプの語句関連も、関連度行列の値をうまく決めることで実用的な対応を見込める。後者に対しては、用語辞典スタックを作成し関連度行列のデフォルト値を提供する。

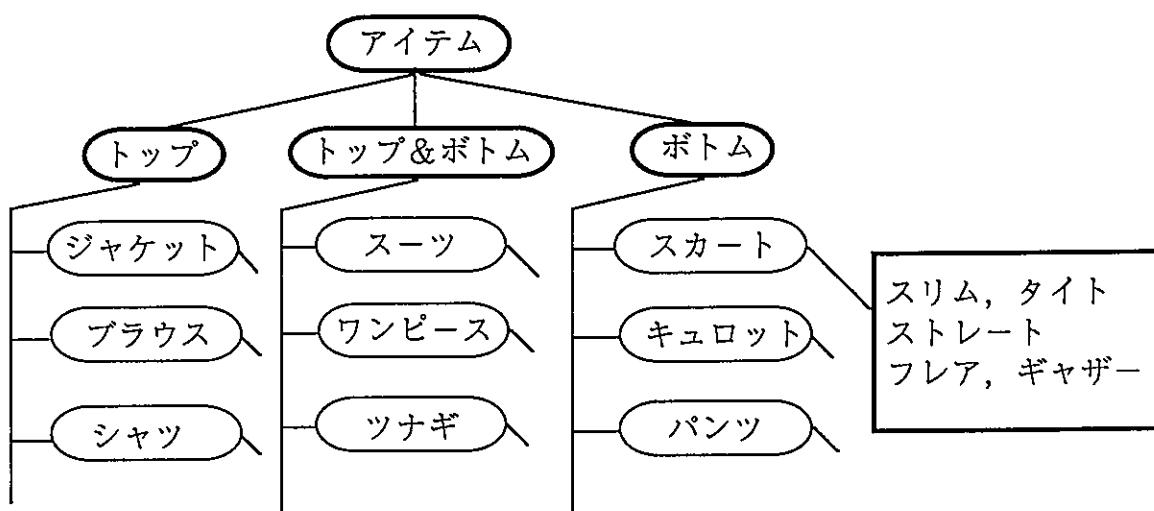


図 1. 3 ファッションアイテムの体系

こうした要件に応えるための基本方針は次の通りである：

- ・ユーザが指定した語句を関連行列に基き拡張する
- ・指定順序を優先順序として重みを算出する
- ・ユーザの希望する拡張範囲内で上の作業を繰り返す
- ・確定した検索語句を有する記事を見出す
- ・要求との一致度に従い検索結果を呈示する

2 検索システムの仕様

2-1. フレーム形式による解説再表現

感性表現ベースは先述したように、記事クラスターを一つの単位として主テーマおよび具体例群の情報をまとめている。省略形あるいは比喩が多用されている原文に検索をかけることは極めて困難である。したがって、本研究では記事クラスター毎に、解説対象をアイテム・色/模様・素材・印象/着こなしのスロットを持つフレームとして記述した。ただし、各スロットが常にデータを持っているとは限らない。これは、資料である雑誌記事の制約であり、感性的情報伝達の不完全さに起因するものである。

ユーザの検索要求は、これらのスロットに対応した語彙群から予約語を選択することで具体化する。予約語は原則的に書換え可能であるが、後述する用語間のファジー関係などを指定しなければならない。こうした指定を視覚的に支援する機能を現在検討中である。語彙群に分割することにより同音異義語の問題をかなり回避ものと期待される。これはカタカナ記述の表現が多いファッション記事において重要な課題である。例えば、色を指すカラーと襟を指すカラーは、それぞれ色のインスタンス、アイテム（の部分）のインスタンスとして区別される。また、どの語彙群に要求指定がなされたかによって検索すべきスロットが確定するので検索の速度向上にもつながる。

2-2. ファジィ・シソーラスの設定

ユーザの検索要求には不明確(vague)さや曖昧(ambiguous)さがつきものである。これは逆に、検索結果に対して関連語を含む検索結果に対する許容度が高いということにもなる。たとえば、「青」の指定に「ブルー」を含む結果が返されても的中感は損なわれないであろう。これは翻訳語間での同義語とみなせる。やや的中感は低くなるものの「青緑」や「緑」についてもどうようである。場合によっては、むしろユーザが明確にしえなかった潜在的な要求にかなうものと歓迎されることもあるだろう。しかし、「青」に「赤」を返すことは好ましくないと考えられる。ファジー・シソーラスにはこうした同義語・類義語に加え、ボトム⇔スカート⇔タイトスカートといったクラス関係が記載される。いずれの場合も、関連語との関係はファジー関連度として[0,1]の区間の数値で示す。一般にこうした関連は非対称なので検索語を行に、関連語を列にもつ関連度行列として表わす。この行列を管理するツールもシステムに用意されている。本研究では実用性を検証するために、約200語の予約語を設定した。

2-3. 検索優先度の設定

論理学での選言 ($A \cup B$) では結合された項同士は対等である。しかし、ファッション事例を求めるときには、「特定の事例がなければ代替例でもかまわない」というように優先順序を伴うことが多い。「ブラウスで見つからなければ、シャツで該当するものを」と

いうのは、同じカテゴリ内での優先順位として扱える。カテゴリ間の優先順位は、代替とは異なり注視対象に由来するものである。「清楚な印象を与えるブラウスで、色は青系統」という要求の場合、印象が主眼でありアイテム、色はやや従と見なせる。本研究では、潜在的な優先順序が検索語句を指定する順序に反映されるとの前提にもとづいて、カテゴリウエイト、カテゴリ内項目ウエイト[0,1]を既定値から算出し、ユーザの意向で自由に変更できるようにした。

2-4. 検索の手順

1) シソーラスに用意されている語句、ユーザが指定する語句をそれぞれ

$$W_T = \{w_{T1}, w_{T2}, w_{T3}, \dots\}$$

$$W_U = \{w_{U1}, w_{U2}, w_{U3}, \dots\}$$

で表わす。当然ながら後者は前者の部分集合である, i.e., $W_U \subseteq W_T$. ユーザ指定語句から検索語句集合 W_F を次の再帰的手続きで生成する。

$$\text{Step 1: } W_F = W_U$$

$$\text{Step 2: } A_\alpha = \{w_{sj} \mid \mu_{w_{Uj}}(w_{Ti}, w_{Tj}) \geq \alpha\} ; w_{Ti}, w_{Tj} \in W_F$$

$$\text{Step 3: } W_F = \text{union}(W_F, A_\alpha)$$

ただし、 A_α は α -カットと呼ばれる集合で次の式で定義される：

$$A_\alpha = \{x \mid \mu_A(x) \geq \alpha\} \quad x \in X, \alpha \in [0, 1]$$

すなわち、 α を適当な水準に設定することにより、集合 X の要素を適宜選択することができる。上の手続きの Step 2, 3 を繰り返すことにより W_F が拡充される。本システムでは、ユーザがこの繰返しの深さを変更できる（既定値は1）。

2) 上で確定した W_F の要素語句に対し、 q^a (カテゴリウエイト), q^b カテゴリ内ウエイトの積により重みづけを行う：

$$G = q^{a*} q^{b*} \mu(w_U, w_F)$$

この値の大きさの順（降順）に W_F の要素をソートしたリストを作り検索を実行する。

3) 検索要求と検索結果の一致度は、ヒットしたフレーム内での G の和で測る。Just-suggestiveness の原則により、上位 5 件を直ちに表示する。それ以降のものについてはユーザの選択による。

2. 3. 2 検索優先度付与の方法

検索優先度は、検索要求の指定順位に基づいて付与される。ユーザが選択する検索語はその順にしたがって検索優先度が高いであろうことが予想される。ユーザは設定された予

約語から検索語を選ぶ。このとき指定された順に各スロットと各キーワードに検索優先度が付与される。一つのスロットにおいて1番目に選択された検索語には1番高い優先度を、2番目に選択された検索語には2番目に高い優先度を与えるよう設定した。また、スロットも同様である。付与する優先度の値は、ユーザが設定できるように記録用フィールドを用意した。

表3. 1 各索引カードの索引語

フレーム	索引語構成	一致度	
		条件1	条件2
1	アイテム語句3	10.6	5.8
2	アイテム語句2, 色語句1	9.2	6.0
3	アイテム語句1, 色語句2	7.8	6.2
4	素材語句1, 色語句3	6.2	7.8
5	素材語句2, 色語句2	6.0	9.2
6	素材語句3, 色語句1	5.8	10.6
7	素材語句4	5.6	12.0

3. 評価実験

本システムに対する評価の脚掛かりとして、カテゴリーのウェイト(q^a)の影響をアイテム、素材に関して調べた。ただし、実験的に設定したフレームでは色に関するデータを表3. 1のように混在させた。評価は [アイテム, 色, 素材] カテゴリーのウェイトを

条件1 [1.0, 0.4, 0.2]

条件2 [0.2, 0.4, 1.0]

の2種類設定し行った。表3. 1に示した一致度から、本システムの検索規則がフレームのデータを良く反映していると言える。

4. 結びにかえて

感性表現ベースに対するユーザの曖昧さを伴う検索要求を補助する目的でファジーシソーラスを構成した。ファジー理論の α -カットを適用することで検索語句の拡張を実用性の範囲内に制限することが可能となった。しかし、優先度ウェイトについてはmin-max等のファジー演算を用いずに、検索順序の決定や検索結果の初期要求との一致度を算出し

た。従って、ファジー理論の有用性は現在のところ限定的でしかない。しかし、システムを拡充し、表現ベースの解説情報体系を反映したり、「…ではないもの」といった排除的要求を取入れたりするためには有効な基盤となることが期待される。

参考文献

- 宮本定明 (1993) ファジィ理論と情報検索モデル. 講座ファジィ 9 ファジィデータベースと情報検索. 日本工業新聞社
- 小川泰嗣 (1993) キーワード結合行列を用いたファジィ文書検索システム. 講座ファジィファジィデータベースと情報検索. 日本工業新聞社
- 野本弘平 (1993) 引用関係によるファジィ文献検索システム. 講座ファジィファジィデータベースと情報検索. 日本工業新聞社
- 水本雅晴 (1987) ファジィ理論とその応用. サイエンス社
- F.W.Lancaster, 中村幸男訳(1968) 情報検索システム. 紀伊国屋書店
- 寺野寿郎, 浅居喜代治, 菅野道夫 (1986) ファジィシステム入門. オーム社
- Parsaye, K., Chignell, M., Khoshafian, S., & Wong, H. 近谷英昭訳 (1992) 知的データベース—オブジェクト指向・演算・ハイパーメディア—. オーム社