

感性語句を含む自然言語文による 画像検索のための形状特徴空間の構築

原 田 将 治^{†*} 伊 東 幸 宏^{††} 中 谷 広 正^{††}

本研究は、感性語句を用いた自然言語文による画像検索システムの構築を目的とする。感性語句を用いた画像検索は、ユーザが画像内容についての専門的知識を持たない状況で、直感的印象に基づいて未知画像を検索する場合などに有効な方法である。本研究では、特に、比較表現を用いて対話的に検索を進める方式を考えている。感性語句を含む自然言語文を解釈して画像特徴と突き合わせるためには、両者の媒介となる場を設定する必要がある。これまでの研究により、色特徴について言及する感性表現を解釈するための媒介として、HSV空間を用いればよいことが分かっているが、それ以外の特徴、たとえば形状特徴について言及する文を解釈するための媒介となる場として、どのような条件を満たす場を考えればよいかは明らかでない。そこで本論文では、感性語句と画像特徴とをマッチングするための媒介となる場を満たすべき条件を明らかにし、椅子画像を題材として、形状についての感性語句を解釈するための特徴空間を構築する手法について述べる。また、実際に構築した椅子の形状特徴空間の妥当性を評価し、その特徴空間を利用して試作した画像検索システムの動作について述べる。

On Constructing Shape Feature Space for Interpreting Subjective Expressions

SHOUJI HARADA,^{†*} YUKIHIRO ITOH^{††} and HIROMASA NAKATANI^{††}

The purpose of our research is to study a methodology of building an image data retrieval system. The system can accept natural language sentences that contain subjective expressions. Image retrieval systems having an ability to handle subjective expressions are useful especially when users, who have no knowledge about contents of an image database, try to retrieve unknown images. We have proposed a method to retrieve images interactively by using comparative sentences. In order to interpret sentences including subjective expressions and match the result with each image, a space that binds subjective expressions and image features is necessary. HSV space can be used as such a space in the case that subjective expressions mention color of images. However, in the case that the expressions mention other features, an appropriate space has not been explored. In this paper, we discuss the conditions that the space should satisfy and we propose a method of constructing a space that satisfies those conditions. We show a process of constructing a shape feature space of "chairs," and evaluate the shape feature space. We also show an experimental image retrieval system that stores pictures of office chairs and accepts natural language inputs including subjective expressions.

1. はじめに

本論文では、画像データベースの日本語インタフェースの構築、特に、感性語句を含む日本語入力を解釈す

るために必要な特徴空間の構築方法について論じ、椅子画像を検索するための形状特徴空間の構築を行う。また、その形状特徴空間を利用して試作した画像検索システムについて述べる。

画像データベース検索システムとして、これまで、キーワードを画像ごとにあらかじめ付与しておき、そのキーワードを用いて検索要求を指定させるタイプ^{1)~3)}や、スケッチ画やシルエット画像を入力とし、画像解析結果の画像特徴値を比較して画像検索を行うタイプ^{4)~9)}などが提案されてきている。これらの方式はそれぞれ異なった状況で有効であり、いずれも画像

[†] 静岡大学大学院電子科学研究科
Graduate School of Electronic Science and Technology,
Shizuoka University

^{*} 現在、株式会社富士通研究所
Presently with Fujitsu Laboratories Ltd.

^{††} 静岡大学情報学部
Department of Information Science, Shizuoka University

データベースの活用には欠かせない検索方式である。

さらに、もう1つの検索方式として、「暖かい」、「地味な」といった感性語句を用いて、相当する印象を与える画像を検索するという方式も検討されている^{10)~13)}。この方式は、画像内容に関する専門知識を持たないユーザが未知の画像を検索しようとする場合に特に有効である。たとえば、絵画を検索したり、インテリア製品をカタログから選んだりする場合、絵画の様式やインテリア製品の仕様についての知識がないと客観的キーワードの指定は難しく、検索したい絵画やインテリアのイメージを具体的に持っていない限りスケッチ画を描くことは不可能である。このような場合でも、検索したい画像の直感的印象を指定することであれば可能な場合が少なくない。

我々はこの方式の有効性に着目し、感性語句を用いた自然言語文による画像データベース検索について検討を行ってきた^{14),15)}。この方式をとる場合、画像ごとにそれと対応すべき感性語句をあらかじめ与えておくことは、印象の主観性やそれに基づく判定の煩わしさ、判定結果の不安定さから、現実的ではない。感性語句を用いた検索条件指定を、画像から何らかの方法で抽出することが可能な画像特徴を用いた検索条件表現に変換したうえで、画像特徴値を用いて検索条件と各画像とを比較していくのが普通である。そのためには、感性語句による条件指定から画像特徴を用いた条件指定に変換する機構が不可欠である。我々は、色に関する条件を指定する感性語句を解釈して画像特徴と突き合せを行う方法を提案した¹⁴⁾。この方法では、色相・彩度・明度で表されるHSV空間上に、感性語句で指定される色の範囲を設定することによって、色特徴に関する感性語句と画像特徴との対応付けを行っている。ここで用いた枠組みは、基本的には色以外の特徴に対しても拡張可能であるが、色特徴以外の特徴に拡張しようとする、色特徴におけるHSV空間に相当する特徴空間が与えられなければならない。しかしながら、一般に、様々な形状を網羅的に整理しうる空間は知られておらず、また、感性語句の解釈のベースとなりうる特徴空間が満たすべき条件についても明らかにされていない。そこで、本論文では、感性語句の解釈のベースとなりうる特徴空間が満たすべき条件を整理し、さらに、そのような空間を構築する手法について論じる。本論文では、例題として、椅子の形状特徴を取り上げ、特徴空間の構築方法について検討し、さらに、その特徴空間を利用して構築した日本語インタフェースについて述べる。次章では、色特徴に関する感性語句を解釈して画像特徴と突き合せを行う仕組

み¹⁴⁾について総括し、感性語句解釈のベースとなる特徴空間が満たすべき性質を整理する。また、そのような条件を満たす特徴空間構築の基本的な方法を提案する。3章では、椅子の形状を題材にして、形状特徴空間構築のプロセスと結果を示す。4章では、形状特徴空間を用いて試作した画像検索システムの動作例を示す。また、5章では提案する手法を椅子以外の対象に適用する場合の問題点を整理する。

2. 基礎的考察

2.1 特徴空間が満たすべき条件

2.1.1 従来の方法の総括

感性語句を用いたインタフェースでは、検索結果を見ながら対話的に要求を明確化しながら検索を進められることが重要である。なぜなら、感性語句の使用法は個人に依存するところが相当程度あり、計算機がすべてのユーザを満足させるような解釈を与えることは困難であるため、要求の仕方を修正しながら対話的に検索を進められるようにすることが全体の検索効率を向上させると考えられるからである。検索結果を見ながらの対話的検索を考える場合、感性語句が比較級で用いられることが多い。そこで、感性語句を用いた日本語文、特に比較表現の使用を認め、対話的に検索が行えるインタフェースの開発を進めてきている。以下、その概要を整理する。

感性語句、つまり、言語を用いて画像の検索を行う場合、言語情報を画像情報に直接対応付けることができないため、なんらかの媒介となる場が必要となる。我々は色特徴に言及する言語表現を画像データと対応付ける媒介の場としてHSV空間を用い、以下のような方法で対応付けを行った¹⁴⁾。

- (1) 感性語句の意味表現を、感性的属性概念を用いて定義する。たとえば、「柔らかい」は「標準以上の柔らかさを内包している」と定義する。
- (2) 感性的属性概念ごとに、そのような印象を標準以上に与えうる色の範囲をHSV空間上に定義する。
- (3) さらに、そのような印象を最も強く与える色が特定できる場合には、そのような点をHSV空間上にマークする(代表点と呼ぶ)。「鮮やかな」などのように最も強く印象を与える色があまり色相に依存せず、トーン属性(明度、彩度)だけに依存する傾向が強い場合には、トーン属性値をマークする。その際、代表点は空間上で線を構成する。
- (4) 各画像から検出される色情報からHSVに相当

する値を算出しておく。

- (5) 感性語句による条件指定が行われた場合、感性的属性概念に対応する HSV 空間上での領域内に含まれる色を持つ画像を検索する。
- (6) 感性語句を程度副詞が修飾する場合には、程度副詞の度合いに応じて HSV 空間上での領域に対して拡大/縮小などの操作を施す。
- (7) 比較表現中に感性語句が用いられた場合(たとえば、「もっと柔らかい色の椅子」など)比較対象となる画像(現在提示している候補画像)の色に対応する HSV 空間上の点(現在点と呼ぶ)から、感性的属性概念に対応する代表点(代表点が線を構成している場合には、同一色相の代表点)に近づけたあたりに適当な領域を定めて検索範囲とする。どのくらい近づけるかは現在点と代表点との距離、感性語句を修飾する程度副詞情報から定める。検索範囲の大きさは現在点と代表点との距離に応じて定める。

我々は、以上の方針を定めるにあたり、HSV 空間が人間の心理的印象に基づいて構成された空間であることから、いくつか提案されている表色系から HSV 空間を採用した。人間の心理的印象に基づいて構成された特徴空間では、印象が近い色は空間上でも近くに存在することが期待できるので、ある感性的特徴概念ごとにそのような印象を与える色の範囲をまとめた領域として定義することが可能である。実際、我々も簡単な心理実験を通して、同一の印象を与える色が HSV 空間で閉じた領域を構成することを確認している。また、この空間では印象の違いが空間上の距離として表されることになるため、たとえば「この椅子より柔らかい色の椅子」といった比較表現に対して、上述のように、最も「柔らかさ」を感じさせる位置へ距離を近づけて検索範囲を設定すればよい。さらに、HSV 空間上での位置は、画像から抽出可能な RGB などの物理的画像特徴から算出することが可能であるため、画像として与えられた情報を特徴空間上に位置づけることは自動的に実行できる。以上の理由から HSV 空間が言語情報と画像情報との媒介の場として適切なものであると考えられる。

2.1.2 特徴空間が満たすべき条件

以上の考察から、感性語句を解釈して言語情報と画像情報とを突き合わせるためには、感性語句で言及する特徴に応じた特徴空間が必要であると考えられる。この特徴空間は、以下の条件を満たしていなければならない。

A) 人間の感性を反映し、同様な印象を与えうる特徴

の値が近傍に位置すること

- B) 連続的な値をとり、印象の違いを距離として求められること
 - C) 物理的特徴量から空間上の位置が求められること
- A) により、感性的属性概念で表される印象を与える範囲をまとめた領域として定義できる。また、B) によって、比較表現に対し、比較対象となった特徴点から感性的属性概念で表される印象を最も強く与える特徴点との印象の差異、つまり、距離を縮めた領域を検索範囲として用いることが可能となる。さらに、C) によって、画像解析によって自動的に求められる物理的特徴量を用いて、各画像の特徴空間上での位置を自動的に算出できることが保証される。

色特徴に関してはすべての色を定義した空間(表色系)が存在し、その中で HSV 空間が、前項で述べたように上述の条件を満たし、我々の手法に適した特徴空間であった。しかし、色以外の特徴について、すべての特徴量を網羅し、かつ、上述の条件を満たすような特徴空間は、一般には存在しない。たとえば、形状に関していえば、すべての3次元物体の形状を表現しうる形状特徴空間を考えることが容易ではない。また、それが可能であったとしても、形状が与える印象は、その形状を持つものに大きく左右されてしまう。たとえば、「どっしりした」という印象を与える形状は、「どっしりした椅子」と「どっしりしたテレビ」とでは大きく異なる。したがって、具体的画像を検索するための言語解釈に必要な特徴空間は、対象の「もの」ごとに個別に定義する必要がある。そこで、次節では、対象物として椅子を取り上げ、上述の3条件を満たす個別形状特徴空間の基本的構成方法について述べる。

2.2 特徴空間の構成方法

はじめに、我々は SD 法¹⁶⁾を用いて心理的イメージ空間を測定することにした。SD 法では以下のようにして測定される。まず、何人かの被験者にいくつかの形容詞対による尺度、たとえば、明るい-暗いといった尺度を用いて対象を評価させる。次に因子分析を用いてそれぞれの形容詞尺度を説明しうる少数の潜在的な因子を求めることでイメージ空間を構成する。ここで求められるイメージ空間は人間の印象を反映し、同様な印象を与えるものどうしがイメージ空間上でも近くに位置し、空間上の値の差が印象の違いを表しているため、必要となる特徴空間の条件 A), B) を満たしている。

しかし、求められたイメージ空間はそれぞれの対象の座標を客観的、自動的に求める方法はなく、未知の対象に対しては形容詞尺度を評価しなければ空間上の

座標を求めることができず、条件 C) を満たしていない。そこで、我々はさらに重回帰分析を用いた画像特徴量からイメージ空間上のそれぞれの軸、つまり、因子を求める回帰式を求めることにした。求められた回帰式を用いて物理的特徴量からイメージ空間上の位置を算出することで条件 C) も満たす形状特徴空間を構成する。

2.3 類似研究との比較

感性語句と画像情報とを対応付けるための媒介となるべき場としての空間の構築を試みた研究として加藤ら^{11)~13)}の正準相関分析を用いた手法の研究があげられる。正準相関分析とは多対多の変数間の関係を分析し、それらに対応付けるための少数次元の空間を求める手法である。加藤らの検索手法では、希望する画像の印象を n 個の形容詞の中から選択するという方式で行われ、形容詞が選択された/されないと二値を各形容詞に対して割り付けた n 次元ベクトルと画像特徴値で構成されたベクトルとを対応付けるための空間を正準相関分析により求めている。

この方法は、形式的手順で空間を定義できるため応用範囲が広く、有効な方法である。たとえば、感性語句と音声特徴との対応付けなどにも適用可能である。しかしながら、この方法では、検索された結果を見ながらそれと比較した際の印象の違いを次の入力に反映することが困難であるため、比較表現を用いながら対話的に検索を進めていくことができない。なぜなら、正準相関分析により構成された空間からは、入力となる n 次元ベクトルの間の心理的距離や方向を求められず、条件 B) を満たせないからである。たとえば、あるベクトル p 、その任意の 1 つの要素だけを反転させたベクトル p' 、それと別の 1 つの要素だけを反転させたベクトル p'' を考える。このとき、 p で表される印象と p' で表される印象との違いと、 p で表される印象と p'' で表される印象との違いとで、どちらの方が違いが大きいかは、空間上に反映されていないため求められない。また、ある形容詞 W が選択されているベクトルに対応する印象を与える画像に対して、もっと W という印象が強い画像を検索したい場合に、どのようなベクトルを指定すればよいかは不明である。したがって、感性語句を含んだ比較文を用いて対話的に検索を進めようとする目的には適さない。すなわち、この方法は、異なった目的に基づいて感性語と画像特徴との媒介空間を規定しようとするものであり、上述の 3 つの条件を意識した場合には適用できない。

3. 形状特徴空間の構成

3.1 SD 法によるイメージ空間の測定

SD 法によりイメージを測定するためには、形容詞尺度ならびに評価対象を選出しなければならない。形容詞尺度は網羅的な収集を行ったうえで選出した。まず、EDR 概念辞書¹⁷⁾の中から椅子の形状を修飾する概念を抽出し、859 個の概念を選択した。さらにそれらを類似性に基づいて分類し、58 種類のグループに分類した。各グループを代表する概念見出しは以下のとおりである。

四角い、粗末だ、美しい、大きい、突起する、質素さ、古風だ、複雑さ、かわいい、不統一だ、鈍だ、精巧だ、優雅だ、平らだ、派手さ、素敵だ、豪華だ、硬い、重い、凹み、ごてごてする、面白い、軽い、陰気だ、ゆったりする、今風だ、荒々しさ、自然的だ、すっきり、新鮮だ、古い、平凡だ、凸凹だ、丸い、鋭い、静寂さ、単純さ、雑だ、小さめだ、地味だ、柔らかい、重々しい、安定、特殊だ、強い、太さ、細長だ、きちんと、窮屈だ、つまらない、醜い、統一的だ、不自然だ、不安定、野性的だ、不整だ、弱々しさ、不気味だ

なお、ここでは、検索すべき椅子の形状に影響を及ぼしうるものを幅広く抽出した。そのため、「四角い」「丸い」のように比較的客観的に形状に言及する語彙から、「好き」「嫌い」のようにきわめて主観的な語彙まで多岐にわたる。主観性の高い語彙に関しても、SD 法により、たとえば、統計的に多くの人が「好む」形、「嫌う」形を評価することができる。本研究では、できるだけ広範囲な言語表現に対応できるようにすることをねらって、主観性の高い語彙も含めて実験を行ったが、特定の応用に限定する場合には、あまりに主観的で、形状に対する影響の個人によるばらつきが大きそうな語彙は対象からはずしてもよい。このグループを基に表 1 の左側に示す 20 対の形容詞尺度を決定した。

評価対象となる画像は 1 冊のカタログに含まれる事務・会議用の椅子の中から選出した。カタログには全部で 294 種類の椅子が掲載されており、脚の形や背の形を基に分類し、63 グループにタイプ分けが行われた。その中から約 2 倍の 124 種類の椅子を選択した。2 倍の種類を選択したのは後に行う重回帰分析を評価するためのデータを得るためである。また、対象画像は色特徴の影響が出ないように白黒の 2 値画像で行った。

次に被験者に、それぞれの画像に対し形容詞尺度を

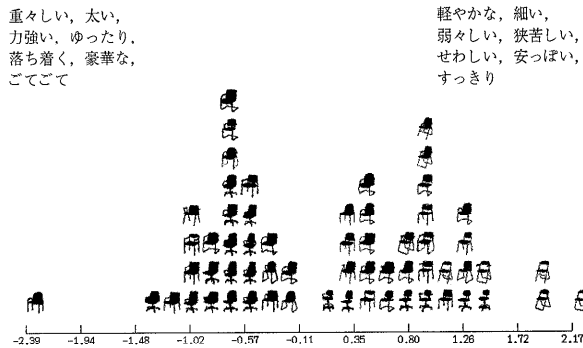


図2 第1因子の因子スコア

Fig. 2 Factor scores of the first factor.

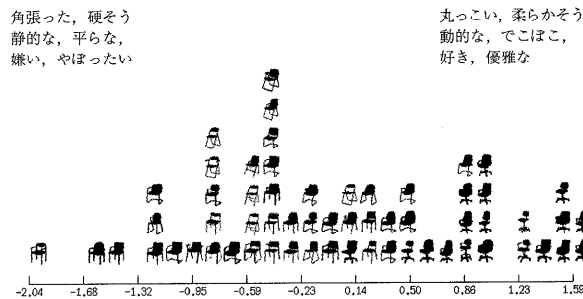


図3 第4因子の因子スコア

Fig. 3 Factor scores of the fourth factor.

表1 因子分析結果

Table 1 Results of the factor analysis.

SD スケール	因子			
	第1	第2	第3	第4
重々しい-軽やかな	0.88	-0.06	-0.20	-0.17
太い-細い	0.85	0.00	-0.21	-0.20
力強い-弱々しい	0.85	-0.04	-0.21	-0.15
ゆったり-狭苦しい	0.76	0.14	0.25	-0.17
落ち着く-せわしい	0.72	0.19	0.29	-0.20
安っぽい-豪華な	-0.71	0.27	0.11	0.42
ごてごて-すっきり	0.56	-0.42	-0.35	-0.38
平凡な-珍しい	-0.17	0.81	-0.17	-0.01
整った-いびつな	0.19	0.74	0.24	-0.07
新鮮な-古風な	-0.24	-0.70	0.26	-0.08
地味な-派手な	-0.10	0.70	0.07	0.11
不統一な-まとまり	-0.08	-0.68	-0.31	0.08
複雑な-単純な	0.35	-0.61	-0.35	-0.41
美しい-醜い	-0.08	0.05	0.70	-0.05
丸っこい-角張った	0.07	-0.19	0.05	-0.71
硬そう-柔らかそう	-0.49	0.08	0.10	0.66
静的な-動的な	-0.28	0.04	0.53	0.63
平らな-でこぼこ	-0.33	0.50	0.27	0.51
嫌い-好き	-0.16	-0.23	-0.31	0.50
優雅な-やぼったい	0.30	0.11	0.08	-0.42
因子寄与度	4.95	3.69	1.75	2.72
因子寄与率 (%)	37.8	28.2	13.3	20.7

用いて評価してもらう。我々は被験者として学生 67 人に対しアンケートを実施した。アンケートは Web ブラウザ上で対象画像の 2 値画像を見せ、それに対し形容詞尺度に評価値をチェック (マウスでクリック) するといった簡単な操作で入力が行えるアプリケーションを作成し、それを用いて行った。図 1 (p. 2365 参照) にその外観を示す。中央に評価対象となる画像が表示され、ユーザは左右に 10 個ずつ提示されている計 20 個の SD スケールを用いて評価を行う。スケールはラジオボタン形式になっており、マウスでクリックすることで図のような丸い印を付けることができるようになっている。また、対象画像の順番はランダムでかつ、似たような形のものが連続しないようにした。

得られたアンケート結果から因子分析を行い¹⁸⁾、表 1 のような結果を得た。結果から椅子の印象は互いに独立した 4 つの因子に依存していることが分かる。これらの結果から各画像の因子スコア、つまり、その因子軸上の値を算出しプロットした結果のうち第 1 因子、第 4 因子をそれぞれ図 2、図 3 に示す。左右に示されたそれぞれの SD スケールは一番因子負荷量の高かった因子、つまり、相関の高い因子に対し付与されており、因子負荷量の高いものから順に上から並べられて



(A) 椅子面積 (B) 凸閉包面積 (C) 細線化画像

図4 画像例

Fig. 4 Example images.

いる。たとえば、図2では、その画像が左側にあるほど、「重い」・「太い」・「力強い」・「ゆったり」・「落ち着く」・「豪華な」・「ごてごて」といった印象が強くなり、逆に右側にあるほど、「軽やかな」・「細い」・「弱々しい」・「狭苦しい」・「せわしい」・「安っぽい」・「すっきり」といった印象が強くなる。それぞれのSDスケールがなるべく1つの因子だけに相関が高くなるように因子軸の回転が行われているが、いくつかは2つ以上の因子と相関の高いものがあり、たとえば、「安っぽい ⇔ 豪華な」は第1因子だけでなく第4因子とも相関が高く、第1因子の図2において左方向にあり、かつ、第4因子の図3において右方向にあるものが、「豪華な」という印象が強い。

全体的かつ直感的な印象では、比較的客観性の高そうな語彙に関しては、ある因子と相関の高い形容詞対として、似たような語感のものが集められているように見受けられる。たとえば、「重々しい ⇔ 軽やかな」と「太い ⇔ 細い」などは第1因子と相関の強いグループに、「丸っこい ⇔ 角張った」「柔らかそう ⇔ 硬そう」などは第4因子と相関の強いグループに入っている。「好き ⇔ 嫌い」のような主観的な側面の強い語彙に関しては、椅子に限定した場合に他のどんな尺度が「好き ⇔ 嫌い」という評価に関与する傾向にあるかを示しているとみることができる。

以上のように、椅子の形状の心理的な印象は、4つの軸から構成されるイメージ空間で整理できることが分かった。しかし、各々の画像に対する空間上の座標たる因子スコアの値はSDスケールによる評価値がなければ求めることができず、また、その評価は人手によるものであり手間がかかる。そのため、計算機により自動的に値を抽出するために画像解析により得られる特徴と空間を結び付ける必要がある。本研究では前述のとおり重回帰分析により行うことにした。詳細は引続き次の節で述べる。

3.2 重回帰分析による物理的特徴量との関係付け
画像特徴量からSD法で得られたイメージ空間への

表2 重回帰分析に用いた特徴量

Table 2 Image features for multiple regression analysis.

X ₁	偏平度。椅子の高さを h 、横幅を w とすると $w \geq H$ の場合は $1 - (h/w)$ 、 $w < h$ の場合は $(w/h) - 1$ 。
X ₂	凸閉包面積に対する椅子の面積の割合
X ₃	左右からの閉包面積に対するその上部の凸部分面積
X ₄	左右から閉包しさらに上下からの閉包した面積に対するその上部の凸部分面積
X ₅	輪郭線の長さ(周閉長)に対する輪郭線の平均直線長
X ₆	複雑度(周閉長の2乗/面積)
X ₇	椅子の高さに対する椅子面積部のヒストグラムの標準的な1本脚の椅子画像との差
X ₈	同上ヒストグラムの標準的な2本脚の椅子画像との差
X ₉	同上ヒストグラムの標準的な折り畳みの椅子画像との差
X ₁₀	同上ヒストグラムの標準的な4本脚の画像との差
X ₁₁	椅子の高さに対する重心の高さ
X ₁₂	上下からの閉包した部分の面積に対する椅子の面積
X ₁₃	左右からの閉包した部分の面積に対する椅子の面積
X ₁₄	凸閉包の円形度
X ₁₅	椅子の高さに対する細線化画像の線分の長さ
X ₁₆	細線化画像の分岐点数
X ₁₇	細線化画像の連結数に対する横連結の割合
X ₁₈	細線化画像の連結数に対する縦連結の割合
X ₁₉	細線化画像の連結数に対する斜め連結の割合
X ₂₀	水平方向投影長
X ₂₁	垂直方向投影長
X ₂₂	細線化画像の端点数
X ₂₃	脚の数
X ₂₄	折り畳み
X ₂₅	肘掛け
X ₂₆	1本脚
X ₂₇	2本脚
X ₂₈	4本脚

軸に対する重回帰分析を行う。そのために必要な画像特徴の定義を行った¹⁵⁾。ここでは、各因子軸上に実際の画像を配置し、視察によりイメージ空間や印象に対して影響を及ぼしそうな特徴をできる限り広く想定し、それに相当する画像特徴を数多く定義した¹⁹⁾。たとえば、凸閉包面積(図4B)に対する椅子面積(図4A)の割合、細線化画像(図4C)の線分の長さ、重心といった画像解析から得られる特徴22種類を定義した。また、客観的基準に基づき形状特徴の印象に影響すると思われる肘掛けの有無や脚の数といった記号データも用いることにした。これらのデータは、たとえば、肘掛けがある椅子は1、ない椅子は0といったように数値化した。用いた画像特徴を表2示す。以上の特徴量を説明変数と見なし、3.1節で用いた124個の椅子データの半数の62データを用いて重回帰分析を行い、式(1)のような回帰式を得た。なお、重回帰分析にあたっては、因子スコアも画像特徴量も正規化した値を用いている。

$$F_1 = -2.5X_1 - 7.7X_2 - 1.6X_{10} - 6.9X_{14} - 0.1X_{22} + 0.6X_{24} - 0.7X_{28} + 10.0$$

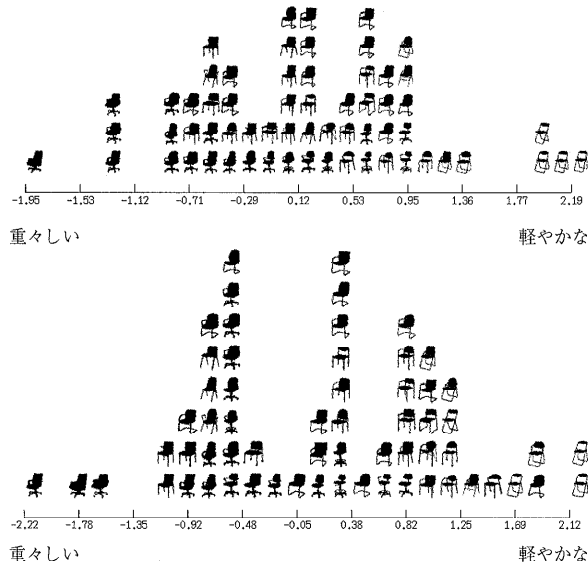


図5 第1因子として算出された軸(上)と因子軸(下)

Fig. 5 Calculated values (top) and factor scores (bottom).

$$\begin{aligned}
 F_2 &= -8.0X_4 - 3.1X_7 - 9.0X_{14} - 35.8X_{19} \\
 &\quad + 10.3 \\
 F_3 &= -4.2X_4 + 2.2X_8 - 0.5X_{25} + 0.7X_{26} \\
 &\quad - 0.7X_{27} + 0.3 \\
 F_4 &= -1.4X_5 - 1.8X_{10} - 6.3X_{14} - 0.2X_{23} \\
 &\quad + 0.7X_{26} + 6.8 \quad (1)
 \end{aligned}$$

以上のようにして、上述の3条件を満たす椅子の形状特徴空間を構築した。

3.3 形状空間の評価

以上の方法で構築した形状特徴空間がどの程度心理的イメージ空間を反映しているか確認するため、重回帰分析に用いなかった残りの62個のデータを用いて評価を行った。それらのデータは前述のとおりSD法の因子分析において、重回帰分析に用いたデータと合わせて測定されているため、それぞれの因子の因子スコアは算出されている(図5下)。それらの画像に対し先の重回帰分析により求められた回帰式を用いて画像特徴量から値を算出した。回帰式 F_1 による値プロットしたものを図5の上部に示す。それぞれの因子において因子スコアと回帰式から算出した値をプロットした結果、ならびに、その相関係数を図6、図7、図8、図9に示す。これらの結果から第1因子・第3因子・第4因子に関しては高い相関が得られており、それぞれの因子を反映した特徴軸が定義できているといえる。第2因子はその他の因子に比べやや低い相関となっている。これは、第2因子が「平凡な-珍しい」といった印象を反映する軸であり、形状特徴の印象というよ

りはむしろ見慣れている度合いを表すような軸と思われるため、画像特徴との対応付けが難しかったためと考えられる。しかしながら、この因子の場合でも相関係数は0.48であり、我々の画像検索システムがもともと解釈の曖昧性を想定して設計されているものなので、十分使用に耐えられるものであると考えられる。

以上の結果から、重回帰分析により構成した形状特徴空間はSD法で得られた心理的イメージ空間を反映した空間であると考えられる。

4. 試作システムと動作例

構築した形状特徴空間を用いた検索手法の有効性を示すためにシステムの検索例を示す。インタフェース部はJava言語を用いてWebブラウザ上で実装しているため、実際にインターネット上の様々なユーザからの検索が可能となっている。また、画像データベース中には221枚の画像が蓄えられている。はじめにユーザは「落ち着いた感じの椅子を見せてください。」と入力したとする。すると図10のように5枚の候補画像が提示される。しかしながら、感性語句の解釈に曖昧性が存在するため、候補画像が必ずしもユーザの希望どおりの画像であるとは限らない。その場合には、ユーザは比較表現を用いて対話的に検索を行うことができる。たとえば、候補画像において、第1因子の「重々しい」・「落ち着いた」・「ごてごてした」といった印象が強すぎた場合、「2番の椅子よりすっきりした椅子が見たい。」といった入力が可能である(検索結果は図11)。

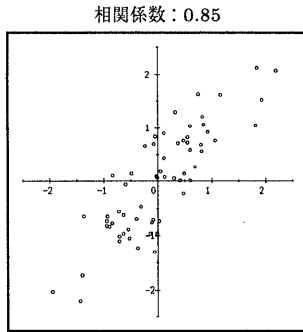


図6 第1因子の因子スコア（縦軸）と回帰式の値（横軸）のプロット

Fig. 6 Correlation at the first factor.

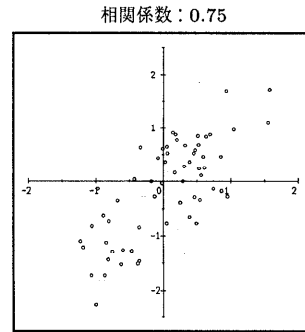


図8 第3因子の因子スコア（縦軸）と回帰式の値（横軸）のプロット

Fig. 8 Correlation at the third factor.

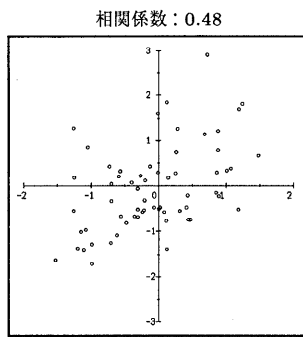


図7 第2因子の因子スコア（縦軸）と回帰式の値（横軸）のプロット

Fig. 7 Correlation at the second factor.

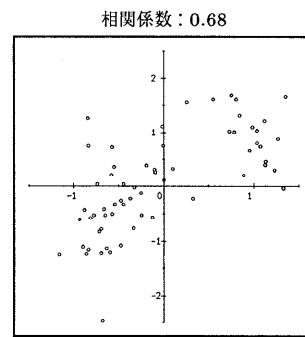


図9 第4因子の因子スコア（縦軸）と回帰式の値（横軸）のプロット

Fig. 9 Correlation at the fourth factor.

ここでもユーザの満足する画像が提示されず、さらに「10番より角張った椅子がありますか。」と入力されたとする。この場合、「角張った」印象は第4因子の値と相関があるため、第4因子軸上で「角張った」印象を与える値を持つ画像が検索されるが、さらに比較対象となった画像の第1因子の印象に近い値を持つことが条件として加えられる。そのため、比較対象となった画像の「すっきりした」印象を持ち、かつ、より「角張った」印象を与える候補画像が提示される（図12）。

以上のように、本システムでは感性語句と画像の対応付けは大まかに行ってはいるが、比較表現を用いて対話的に検索を進めることにより最終的には希望する画像を検索することが可能になっている。また、その比較表現を解釈するために構築した形状特徴空間が有効であることが確認できる。

5. 形状特徴空間構築手法の適用性

本章では、提案した形状特徴空間構築手法を実際の応用システムに適用する際に問題となることが予想される問題について考察する。まず、本手法を、「椅子」

以外の対象に適用する際の問題を考察し、次いで、本手法をコスト面から検討する。

5.1 椅子以外の対象への適用可能性

提案した手法を「椅子」以外の他の物体に適用する場合、以下の3点を考慮する必要がある。

(1) SD法によるイメージ空間の測定

SD法による心理的屬性空間の測定では、経験的に3～4次元空間で近似可能とされている²⁰⁾。ただし、本実験のように特定の具体的対象に対してSD法で評価した場合に、純粋に「形状」だけを評価しているか否かがやや曖昧である。

形状的な差異はあるものの、比較的単調な形状のバリエーションしか持たない対象に対して適用すると、形状以外のファクターが混入しやすいと予想される。したがって、本手法を適用するにあたっては、対象の形状のバリエーションをチェックし、ある程度豊富なバリエーションを有することを確認しておく必要がある。また、SD法による実験に際して、被験者にできるだけ「形状」のみを対象に評価するように教示しておくことも必要であろう。



図1 SD法のアンケートの実行例

Fig. 1 The interface of the questionnaire.

(2) 重回帰分析で使用する画像特徴セットの設定

どのような画像特徴を基本セットとして重回帰分析にかけるべきかも、対象に依存するところが少なくない。たとえば、“顔”を対象にしようとする、全体や部品の形状だけでなく、部品の配置に関する情報をはじめ、“椅子”の実験で考慮していない画像特徴を導入する必要があると予想される。

我々は、“椅子”の画像特徴としてどのような特徴を基本セットとするかを決定するにあたって、以下のような手順を踏んだ。まず、抽出した軸ごとに、各軸に沿うように実際の形状を配置して、視察によってどのような特徴に変化があるかを考察し、イメージ空間での座標値決定に寄与しそうな画像特徴を列挙してみた。さらに、それらの画像特徴からイメージ空間での座標値を求める式を手で定義することを試み、ある程度の精度で可能であることを確認した¹⁵⁾。SD法で求めたイメージ空間が正確に“形状”に関する評価を表しているのであれば、視察によってある程度は重要な特徴を発見できるという感触を得ている。

実際に本手法を“椅子”以外の対象に適用する場合には、上述の予備実験のような方法を用いて、寄与が大きいような画像特徴を吟味しておく必要がある。

(3) 重回帰分析の適用可否のチェック

本手法では、物理的特徴とイメージ空間の対応付けを、重回帰分析を利用して行っている。これは、ある物理的特徴量がイメージ空間と密接な関係を持つ場合には、基本的に線形に近似できるタイプの関係になることを想定したものである。本論文で対象とした「椅子」の場合、設定した画像特徴がおおむねこの仮定を満たすものであることを上述の視察実験によって確認



図10 形状特徴による検索例1

Fig. 10 An example of the image retrieval.



図11 形状特徴による検索例2

Fig. 11 An example of the image retrieval.

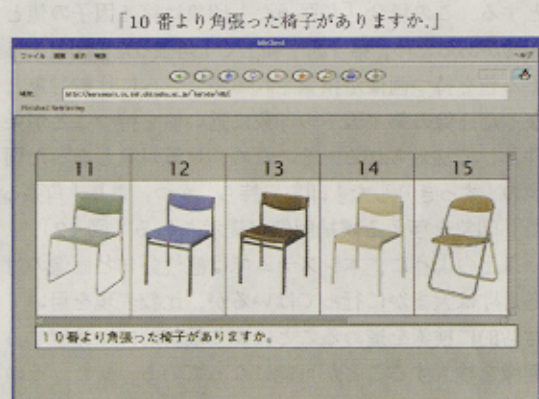


図12 形状特徴による検索例3

Fig. 12 An example of the image retrieval.

し、その結果を参考に重回帰分析を適用することとした。しかし、一般には、ある印象が強い場合と弱い場合にも高い値を示し、中程度の場合のみに反対に低い値を示すような、線形近似が困難そうな画像特徴

が存在する場合も想定できる。そのような特徴量が多い場合には、重回帰分析を適用することは困難である。したがって、画像特徴セットを定義するとともに、重回帰分析の適用可能性のチェックを、少なくとも視察レベルで行っておく必要がある。また、関係がありそうだが、その関係を線形に近似することが困難と判定できる画像特徴量に関しては、その画像特徴量を適当な関数で変換するなどして、線形近似可能なように特徴量の再定義を試みる必要がある。

提案した手法は、基本的には因子分析と重回帰分析という一般的手法を組み合わせているものなので、以上の3点を考慮しうる範囲であれば、“椅子”以外にも適用対象を広げることが可能であると考えている。この点に関しては、今後、実証的に確認していく必要がある。また、本手法が適用可能となる条件をさらに具体的に整理していく必要がある。

5.2 形状空間構築コスト

次に、形状特徴空間構築のためのコストについて検討する。新しい対象に対し、形状空間を構築するためには、以下の手順を踏むことになる。

- (1) 感性語句の整理とSD法で用いる形容詞対の決定
- (2) SD法で用いる対象物セットの決定
- (3) SD法によるイメージ空間の測定
- (4) 重回帰分析の対象とする画像特徴セットの決定
- (5) 重回帰分析による変換式の決定
- (6) 変換式の評価

このうち、(1)では、SD法の一般的手法や適用事例を解説した文献(20, 21)が参考になるし、各種の電子化辞書やコーパスを利用することで、ある程度のコスト削減をはかることができる。(2)のステップでは、イメージの測定に用いる対象物が形状バラエティを網羅している方が望ましいが、対象物が多すぎても被験者の負担が大きくなる。そのため、実験で用いる対象物を選定する必要がある、ある程度のコストが見込まれる。この工程に関しては、“椅子”のように対象物を網羅的に掲載しているカタログが存在すれば、それを参考に形状を何通りかに分類し、各分類から典型的な対象を選択するという方法が有効である。(3)のステップでは、複数の被験者を集めた実験が必要であり、最もコストがかかる可能性がある。我々は、被験者の負担をできるだけ低減できるようにWWW上で評価実験を行えるシステムを作成して実験を行った。しかし、実験にかかる被験者の負担を低減する方策について、さらに検討を重ねる必要がある。また、イメージ

空間測定に最低限必要な実験の規模についての心理学的な知見を考慮すべきである。

(4)のステップでは、前述のように、寄与の大きいような画像特徴を視察によって見当をつけておくことが有効である。しかしながら、変換式の精度を向上させるためには、このステップを、(5)、(6)のステップとあわせて繰り返して、徐々に修正を加えていく必要がある。そのため、ある程度のコストを見込まなければならない。これを軽減するためには、感性語彙の解釈に寄与の大きい画像特徴量に関する一般的検討が必要である。

本論文では、形状特徴空間構築の手法の提案に主題をおいて論じたが、実用システムへの応用を考える場合、さらに、コストの面からの評価と改善が必要である。

6. おわりに

感性語句を含む自然言語入力によって画像データベースを検索するシステムを構築するための特徴空間の構築について述べた。感性語句の解釈のベースとしての特徴空間が満たすべき条件を整理し、特に、椅子の形状特徴についての検討を行い、椅子形状特徴空間を構成し、形状特徴に関する自然言語入力文を受理して検索を行う実験システムを試作した。形状特徴空間はSD法を用いて心理的なイメージ空間を測定し、その空間に対し画像特徴量からの重回帰分析を用いて対応付けを行うという方法で構築した。重回帰分析に用いなかった別のデータを利用して、重回帰分析による対応付けによって求まるイメージ空間とSD法で測定したイメージ空間との比較を行い、両者の間に高い相関が得られることを確認した。このことから、構築した椅子形状特徴空間が必要な条件を満たすものであることが確認できた。また、この椅子形状特徴空間を利用して試作した画像検索システムの動作から、形状特徴空間を利用して比較文を解釈するインタフェースが画像検索に有効に機能しうることを確認した。今後、複数の被験者による使用実験を通してインタフェースの有効性を実際に評価していく必要がある。

また、今回の試作システムは形状特徴空間の評価のために、形状に関するものとして入力を受理するようになっている。たとえば、「落ち着いた感じの椅子」に対しても、形状の側面だけから椅子を選択している。これを、先に検討した色特徴についての解釈系と組み合わせると色においても形状においても「落ち着いた感じの椅子」を選択できるようにしていく必要がある。さらに、値段や材質など、画像特徴以外のシンボルデータとあわせて自然言語で要求を入力できる検索システ

ムへと発展させていく必要がある。

参考文献

- 1) 清木 康, 金子昌史, 北川高嗣: 意味の数学モデルによる画像データベース検索方式とその学習機構, 電子情報通信学会論文誌 (D-II), Vol.J79-D-II, No.4, pp.509-519 (1996).
- 2) 芥子育雄, 池内 洋, 黒武者健一: 百科事典の知識に基づく画像の連想検索, 電子情報通信学会論文誌 (D-II), Vol.J79-D-II, No.4, pp.484-491 (1996).
- 3) 小野敦史, 天野督士, 斗谷充宏, 佐藤 隆, 坂内正夫: 状態遷移モデルとシーン記述言語による自動キーワード付与機能をもつ画像データベースとその評価, 電子情報通信学会論文誌 (D-II), Vol.J79-D-II, No.4, pp.476-483 (1996).
- 4) 加藤俊一, 下垣弘行, 藤村是明: 画像対話型商標・意匠データベース TRADEMARK, 電子情報通信学会論文誌 (D-II), Vol.J72-D-II, No.4, pp.535-544 (1989).
- 5) 栗田多喜夫, 下垣弘行, 加藤俊一: 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).
- 6) Jain, A.K. and Vailaya, A.: Image retrieval using color and shape, *Proc. 2nd Asian Conf. on Computer Vision*, Vol.II, pp.529-533 (1995).
- 7) 平田恭二, 原 良憲: 概略画像を用いた画像検索, 信学技法, DE 92-2, pp.9-16 (1992).
- 8) Lee, D., Barber, R., Niblack, W., Flickner, M., Hafner, J. and Petkovic, D.: Indexing for complex queries on a query-by-content image database, *Proc. 12th Int'l Conf. of Pattern Recognition*, pp.142-146 (1994).
- 9) 美濃導彦, 岡崎 洋, 坂井利之: 対象物の属性特徴による画像検索法, 情報処理学会論文誌, Vol.32, No.4, pp.513-522 (1991).
- 10) 中山万希志, 岩本 啓, 糊田寿夫: 印象語に基づく顔画像の検索, 人工知能学会研究会資料, SIG-HICG-9103-3, pp.17-24 (1991).
- 11) 加藤俊一, 栗田多喜夫: 画像の内容検索, 情報処理, Vol.33, No.5, pp.466-477 (1992).
- 12) 栗田多喜夫, 加藤俊一, 福田郁美, 坂倉あゆみ: 印象語による画像データベースの検索, 情報処理学会論文誌, Vol.33, No.11, pp.1373-1383 (1992).
- 13) Kato, T.: Database architecture for content-based image retrieval, *Proc. Int'l Society for Optical Engineering on Image Storage and Retrieval Systems*, Vol.1662, pp.112-123 (1992).
- 14) Harada, S., Itoh, Y. and Nakatani, H.: Interactive image retrieval by natural language, *Optical Engineering*, Vol.36, No.12, pp.3281-3287 (1997).
- 15) Harada, S., Sugiura, H., Nakatani, H. and Itoh, Y.: On Constructing Pictorial Feature Space For Image Retrieval, *IJCAI-95 Workshop note of Representation and Processing of Spatial Expressions*, pp.103-117 (1995).
- 16) 岩下豊彦: SD 法によるイメージの測定, 川島書店 (1983).
- 17) EDR 電子化辞書 (第2版), 日本電子化辞書研究所 (1995).
- 18) 田中 豊, 垂水共之, 脇本和昌: パソコン統計解析ハンドブック II 多変量解析編, 共立出版 (1984).
- 19) 長谷川純一, 奥水大和, 中山 晶, 横井茂樹: 画像処理の基本技法〈技法入門編〉, 技術評論社 (1986).
- 20) 日科技連官能検査委員会 (編): 新版官能検査ハンドブック, 日科技連 (1973).
- 21) 岩下豊彦: SD 法によるイメージの測定, 川島書店 (1983).

(平成 10 年 9 月 11 日受付)

(平成 11 年 2 月 8 日採録)



原田 将治 (正会員)

1993 年静岡大学工学部情報知識工学科卒業。1995 年同大学院修士課程修了。1998 年同大学院博士課程修了。同年 (株) 富士通研究所入社。博士 (工学)。ヒューマンインタフェース, 自然言語処理, 音声認識等に興味を持つ。



伊東 幸宏 (正会員)

1980 年早稲田大学理工学部電子通信学科卒業。1987 年同大学院博士後期課程修了。同年, 早稲田大学理工学部電子通信学科助手。1990 年静岡大学工学部情報知識工学科助教。現在, 同大学情報学部情報科学科助教授。工学博士。自然言語処理, ヒューマンインタフェース, 知的教育システム等に興味を持つ。電子情報通信学会, 人工知能学会, 言語処理学会, 教育システム情報学会, 日本認知科学会各会員。



中谷 広正 (正会員)

1974 年大阪大学基礎工学部情報工学科卒業。1976 年同大学院修士課程修了。現在, 静岡大学情報学部情報科学科教授。工学博士。画像処理, ヒューマンインタフェース等の研究に従事。電子情報通信学会, IEEE Computer Society, SPIE 各会員。