

人間のあいまいな感性を反映する絵画検索システム

塩澤 秀和 西山 晴彦 松下 温

慶應義塾大学理工学部

本稿では、絵画検索システムについての提案を行い、実装したシステムについて述べる。まず人間の思考に合った画像表現の分析を行い、画像表現における曖昧性をどのように表せば、人間の画像検索に対する曖昧要求処理が我々の感覚通りに行われるかを考察する。そして、人間指向型の画像表現モデルとして、人間の記憶に残っているあいまいな部分を効果的に使うために、画像中のオブジェクトとその配置および、オブジェクトの色あいから感じられる印象語の2つの点に着目した。オブジェクトの配置と位置関係については、検索したい画像に含まれるオブジェクトの概形をフリーハンドで描く。また、それに対して色彩心理学に基づく印象語を付加するというインタフェースを用いた。

A Picture Retrieval System using Ambiguities of Human Feelings

Hidekazu SHIOZAWA Haruhiko NISHIYAMA Yutaka MATSUSHITA

Faculty of Science and Technology, Keio University

3-14-1 Hiyoshi Kohoku Yokohama 223, JAPAN

In this paper, we propose a method of retrieving pictures and describe a implementation which is based on this method. First, we analyze that the image expression suited for human considerations and explain how to express ambiguities of images in a human memory. Then, we propose the human oriented image expression model using spacial relationships of objects in a image and impression words related to colors of objects. We design a human interface of the system with the aim to express easily images in a human mind. You can retrieve pictures by drawing roughly outlines of objects and indicating impression words of the objects.

1 はじめに

我々は、人が画像を検索するとき、どのようなヒューマンインタフェースが望ましいものであるか、という観点から画像データベースの検索手法を研究してきた。人は画像を表現するとき、その略画を描いて説明することによって、視覚的、空間的な記憶を有効に表現しようとする。そこで、我々は「絵」を検索するには「絵」を描いて見せるのが最も自然な方法である、という立場から、レイアウト検索というものを提案してきた [1, 2]。

しかし、略画を描くだけでは、その雰囲気のようなものを十分に表すことができない。人間は画像を表現するとき、それから受ける「明るい」であるとか「どんよりしている」であるといった「印象」を、略画による表現に付加する [3]。絵だけでなく、言葉をもちいることで、人間はあいまいでおおざっぱなイメージを表そうとするのである。

我々は、画像およびその構成要素であるオブジェクトから受ける「印象による検索」というものを、レイアウト検索に加えるもう1本の柱として提案する。これによってユーザは、感性を活かした検索ができるようになるので、今までよりも、人間の感覚と表現に近いインタフェースが実現できるのである。

一方、レイアウト検索に関しても、オブジェクトを「意味をもった空間的広がり」としてとらえ、その形をユーザがかなり自由に入力できるように工夫した。さらに、オブジェクトの大きさ、色、種類(特徴)、印象語、レイアウトといったものを統一的に扱えるような手法を考え、実際に画像検索システムを実装した。

2 ヒューマンインタフェース

2.1 感性指向のインタフェース

ある画像を検索しようとするとき、ユーザは、ディスプレイ上の真っ白なキャンバスにオブジェクトの輪郭を線で囲い、そのオブジェクトの種類と属性を指定する。さらに、そのオブジェク

トによってどのような印象を受けるか言葉(印象語)を選択する。また、キャンバスには自由に色を塗ることができる。

このようなインタフェースを用意すれば、ユーザは「オブジェクトの種類」「オブジェクトの位置関係(レイアウト)」「色」「印象語」を指定して、目的の画像を自然な形で検索することができるようになる。

2.2 画像表現モデル

これを踏まえて、我々は、新しい画像表現モデルを提案する。このモデルは、画像をオブジェクトと画像から受ける印象の2つの要素からなると考えるのが特徴である。この画像表現モデルの概念図を図1に示す。

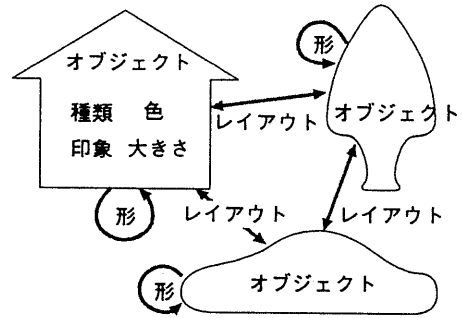


図1: 人間の思考に合った画像表現

まず、画像を「人間」や「動物」など、意味がひとまとまりとなる領域に分ける。そして、そのひとまとまりを1つのオブジェクトと考え、それがどういう種類であるかというカテゴリ、さらに細かい属性(アトリビュート)をつけ、意味ラベルとする。そしてそのオブジェクトの大きさ、色を見る。

さらに人間のあいまいな感性を反映するために、「寒い」であるとか「女性的な」であるとかの、画像から受けるイメージを印象語としてオブジェクトに付加する。

このモデルの特徴は、「こんな感じの絵だった」というあいまいな記憶による検索だけでな

く、「こういう感じの絵はあるか?」という創造的な検索の、全く違った観点からの検索要求を許容する点である。

2.3 印象におよぼす色彩の効果

人間が画像に対して抱く印象語を調査した結果、色あいや色づかいといった、色から感じるものが、非常に多いことがわかった [3, ?]。この点を踏まえて、印象語をオブジェクトの色あいから計算によって判断することにした。

その評価には、ファジーメンバシップ関数をもちいることで、あいまいな感覚をくみ取ることができるようにしている。なお、この印象語を色あいから求めるという方法によって、データベース登録に関しても、登録者は特別な知識や、熟練を必要としなくなり、データベースの変更にも柔軟に適応できるようにもなった。

3 データ表現

3.1 カテゴリとアトリビュート

オブジェクトの意味を大雑把な種類であるカテゴリと、そのカテゴリごとの特徴を表すためのアトリビュートに分類する。カテゴリおよびアトリビュートは整数で表され、値 0 は常に「不明、わからない、どちらでもよい」を表すものとしてもちいられる。

3.2 色の表現

画像データベースを構成する絵画や、写真といった画像データの多数の色をそのままあつかうためには、少なくとも 1 色 (1 ピクセル) につき 8 ビットは必要になってしまい、データベースが膨大になってしまう。そこで、人間のおぼろげな「記憶」を頼りにするデータベース検索システムを構成する上では、それほど多くの色をあつかう必要性はないであろうと判断し、絵画や写真といった画像は、登録される時にディザ処理をして、64 色に減色を行っている。

3.3 位置の表現

オブジェクトが画像の中でどの位置にあるか、従来の画像データベースシステムの多くでは、これを、重心などの代表点を用いたり、外接長方形を用いたりして表現していた [?]。しかし、このような方法では、オブジェクトの大きさや、複雑なオブジェクトの形状を記憶することができない [4]。

そこで、このデータベースシステムでは画像内でオブジェクトの占める位置を保存するために、画像に格子点を考え、各格子点にそのオブジェクトがかかっているかどうかを 1 ビットずつのフラグで表すことにした。つまり、オブジェクトの位置の保存には、従来のような重心や矩形ではなく、シルエットを粗いビットマップとして持たせることで、広がりや、形に加え、ある程度複雑な位置関係を表せるようにした。

3.4 位置関係の表現

今までの画像データベースでは、オブジェクトの位置関係を表すために、オブジェクトの重心など、代表点をもちいた位置関係か、外接長方形をもちいた位置関係の表現が、おもなものであった。しかし、これでは前述したように、オブジェクトの間の多少複雑な (囲まれているといった) 位置関係をあらわすものとしては、十分でない。

そこで、オブジェクトの位置を点の集合として表したことを利用して、オブジェクトの位置関係も、それぞれのオブジェクトを構成する格子点どうしの位置関係の集合として表すことにした。

オブジェクトの位置関係は、2つのオブジェクトを構成するそれぞれの点から選んだすべての組み合わせの位置関係を 8 方向で表して、8 方向それぞれの割合で表している。このため比較には、この配列の相関係数を用いている。

オブジェクトには、一般に多くの格子点が含まれていると考えられるので、個々の格子点については、わずか 8 方向だけであっても、全体としては、かなり微妙な位置関係も表現することができる。

3.5 印象語の表現

印象語データベースには、現在 50 語程度の印象語をいろいろと登録してみて、どのような印象語がこのデータベースシステムに適しているのか検討中であるが、この印象語データベースの形式を以下にしめす。

このデータベースは画像の中の個々のオブジェクトの「色あい」から受ける印象を検索要求として与えられるようなシステムをとっている。この色あいを判断する基準としては、人間の感覚により近いものとして、「色相」「明度」「彩度」の3つのパラメータをもちいることにした。

さらに、あいまいな検索をするために、色相・彩度・明度をそれぞれ、12段階・5段階・5段階に分けて、表すようにした。

データベースを作成するときには、それぞれの印象語とマッチするのがどのような色であるかを、具体的に色としてあたえる。これをデータベース上では、色相・彩度・明度のそれぞれについて段階化された数字として保存しておく。具体的な例を以下にしめす。

```
#051 清楚な 4 3
      = 11 3 4
      = 12 3 4
      = 7 2 5
      = 1 1 5
```

ここでまず初めの1行がインデックス行であり、印象語番号51番の「清楚な」という印象語が、登録されていることを示している。次にある2つの数は、それぞれ、この印象語に適應するルールの数と、パラメータとしての入力変数の数を示している。この場合パラメータの数は、色相・彩度・明度と固定されているので、常に3である。

2行目以降が、ルールを記録しておく実際のデータベース部分である。これは3つの入力変数、色相・彩度・明度が、それぞれどのような値をとるときにこの印象語が成り立つかどうかを表している。

このように、印象語を色によって求められるものに限定したため、画像データベースの登録

時に人間がいちいちその画像のオブジェクトに関する印象語を付加するという、作業をなくすことができるようになった。この制限のおかげで、自動的な登録と、客観的な評価が可能になり、画像データベースの容量の節約をもたらすようにもなった。また、画像の登録時に印象語が不必要となるため、新しくデータベースに印象語を付け加えた場合でも、その印象語の追加以前に登録した画像を、新しく追加した印象語で検索することが可能になった。

4 画像検索

4.1 画像どうしの比較

本システムでは、画像の検索の問題を、基本的には、ユーザの与える「検索要求画像」と、データベース上にある「データベース画像」との間の対等な比較問題としてとらえている。ただし、印象語をもちいた検索の部分だけは、例外的に、データベース画像が、検索要求画像の印象語リストにしめされた、印象語の条件を満たしているかどうかを評価するようになっている。この画像どうしの比較問題は、オブジェクトどうしの比較の問題としてあつかうことができる。

また、画像の一致度は、画像に含まれるすべてのオブジェクトの一致度を平均したものとしている。各オブジェクトの間の一致度は、0から1までの範囲の値として求められるので、画像の一致度も0から1までの範囲の値として求められることになる。

4.2 オブジェクトの対応関係の判断

与えられた検索要求画像と、データベース画像という2つの画像を比較するとき、検索要求画像のどのオブジェクトをデータベース画像のどのオブジェクトに対応させるべきであるか、ということが大きな問題となる。

本システムにおいては、この問題をオブジェクトの対応関係の、すべての組み合わせを調べることによって対処している。その際、オブジェ

クトのカテゴリまたはそのアトリビュートが異なるオブジェクトどうしを比較しようとした場合または、1つの画像の中で一致度が0.2以下であるオブジェクトの組み合わせが、全オブジェクトの約3分の1以上あるときには、その時点で画像どうしの比較をも打ち切ってしまう、画像どうしの一緻度を0とみなすようにしてしまっている

人間の記憶の確かさといった観点から考えた場合、これは妥当な選択であると思われる。なぜならこれは、オブジェクトの意味というものに対する人間の記憶というものがかなり正確であるということと、「わからない(どちらでもよい)」という選択肢を用意することによって、あいまい性に対応しているためである。

4.3 カテゴリとアトリビュートの比較

オブジェクトのカテゴリとアトリビュートによる比較では、2つのオブジェクトが与えられたとき、それらカテゴリとアトリビュートがともに一致する場合に1を返す。このときユーザがカテゴリまたはアトリビュートとして「わからない(どちらでもよい)」を与えた場合には、その項目は一致しているものとみなす。

4.4 大きさの比較

あたえられた2つのオブジェクトのうち小さいほうのオブジェクトの大きさを大きいほうのオブジェクトの大きさで、割ったもので計算している。おおざっぱな式であるが、オブジェクトの大きさを人間がどう感じるかについての考察が、まだ十分でないということと、ほかにもある多くのパラメータに比較して、オブジェクトの大きさというものの、画像検索に対する重要性がそれほどでもなく、実際のところこのような簡単な計算式で十分である。

4.5 色の比較

色相・彩度・明度の3つのパラメータは、それぞれ0から1までの一緻度として表現され、色の一緻度としては、これらの重みつき平均をも

ちいている。現在この重みとして、色相1、彩度1、明度3を用いているが、これは、色相環における補色間の距離(例えば緑と赤)を1とした場合、明度0の黒と明度1の白とでは、その3倍ぐらいの視覚的インパクトがあると考えたからである。つまり色相の一緻度を S_h 、彩度の一緻度を S_c 、明度の一緻度を S_v と表せば、あたえられた2色の一緻度は、

$$S_{color} = \frac{S_h + S_c + 3S_v}{1 + 1 + 3}$$

と式で表される。この3つのパラメータについては、どのように重み付けをするかまだ問題はあると思う。

4.6 方向関係の比較

あるオブジェクトから、ほかのオブジェクトを見たときの方向は8つの要素を持った整数配列として保存されている。

この方向配列を2つの画像について比較のするには、循環的関数に関する離散的相関係数を計算することとした。あるオブジェクトから見た他のオブジェクトに対する相関係数を、全オブジェクトに対して計算する。そして、その平均をとったものを方向に関する一緻度としている。

この方向関係は、そのオブジェクトそれ自身に対するものも持っているので、この方向に関する比較によって同時に、オブジェクトの形状も比較していることになる。

4.7 印象語の評価

本システムにおいては、この印象語の知識と判断のために、ファジー計算のうちもっとも基本的なものもちいている。

印象語データベースで用いられている段階化された色情報は、色相・彩度・明度の各パラメータについて、それぞれに対応するメンバシップ関数として表される。

ユーザからの検索要求として、ある印象語があたえられるとき、データベースシステムは、まずその印象語のルールを調べるために、印象

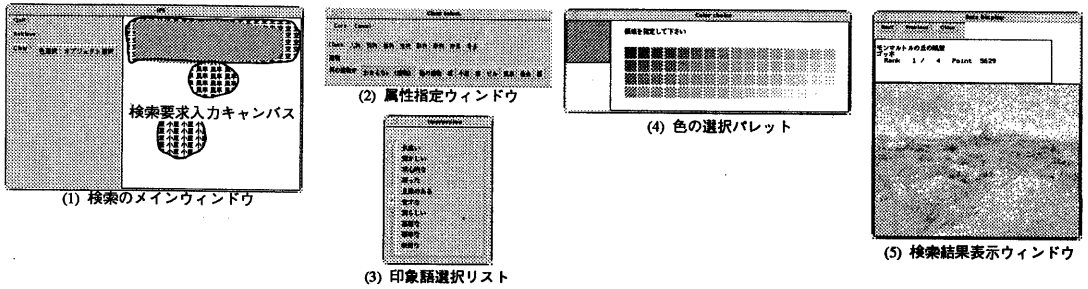


図 3: 画像データベースシステムの実装 (SparcStation + X-Window System)

語データベースを検索する。印象語データベースには、その印象語を満たすためには、色相・彩度・明度が、どの段階にあればよいか記憶されているルールがあるので、それから計算をする (図 2)。

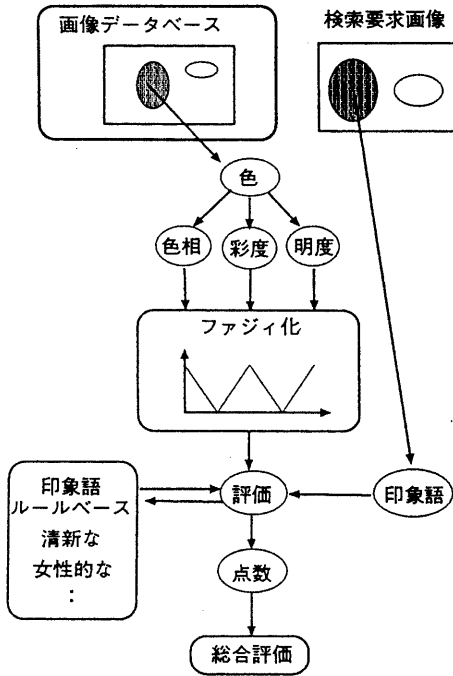


図 2: 印象語の評価の流れ

5 システムの実装

5.1 プラットフォーム

本システムは、Sun Microsystems の Sparc Station 上に X-Window System (Xview, Xlib) 上のアプリケーションとして実装した (図 3)。では以後、実際に我われが実装したシステムのインタフェースを述べてゆくことにしよう。

5.2 画像検索インタフェースの実装

システムの全体像を 図 3 にしめす。ここには、検索要求の例と、その時の検索結果の例が示されている。このシステムは検索要求を入力するためのキャンバスと、オブジェクトや領域の意味を入力するためのアトリビュート指定パネル、色を塗るためのパレット、そして検索結果を表しめするためのウィンドウの、大きく 4 つに分けられる。

メインウィンドウの「オブジェクト選択」ボタンを押すと、オブジェクトの意味を設定するためのウィンドウと印象語を選択できるウィンドウが表示される。ユーザは、ここでオブジェクトの意味(カテゴリとアトリビュート)と、印

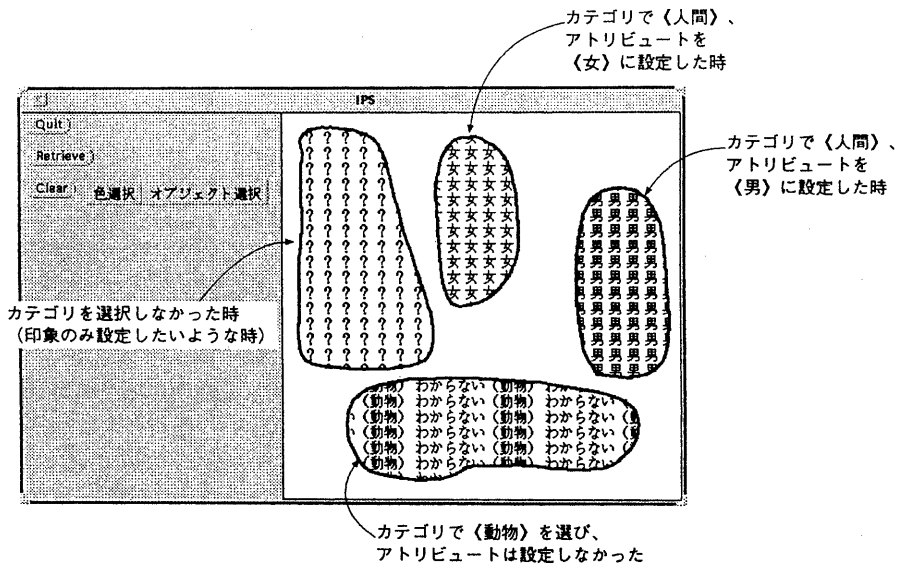


図 4: オブジェクト配置の例

象語を入力する。このとき、オブジェクトが何であったかを覚えていないときには「わからない」を選択をすればよい。

印象語は、図 3(3) にしめす印象語選択リストからこれと思う印象語を選択する。システムで使用している印象語は、色彩心理学の研究を参考にして [5, 6, 7]、色と関連のある印象語を我われが独自に選んだものである。

オブジェクトのアトリビュートを指定し終わったならば、検索要求入力キャンパスの上にオブジェクトの概形を、マウスで描く(図 4)。このように、オブジェクトの形は全く自由に入力できる。

色を塗りたいときには、図 3(4) の色選択パレットから使いたい色を選んで、オブジェクトの配置と同様な手続きでキャンパス上に色を塗る。

あとは、「Retrieve」をマウスでクリックすればよい。候補となる画像が、評価点の高いものから順に表示される。

5.3 画像登録インターフェースの実装

画像登録は、画像に登場するオブジェクトをイメージスキャナで読みこみ、オブジェクトの輪郭をマウスで切り取ることでおこなう。これによって、オブジェクトの位置と形をほぼ忠実に再現できる。

切り取ったオブジェクトの意味(カテゴリ・アトリビュート)は、メニューから選択する。このメニューの内容とアイコンは、システムの実行時にファイルから読み込むようになっているので、画像登録の途中で指定したいオブジェクトのカテゴリやアトリビュートの数を増やすことができるようになっている。

6 アンケートによる評価

本システムを使用してもらった時の感想を数人にアンケート調査した。調査の方法は、ディスプレイ上にデータベースに登録されている画像のうちから 12 枚を同時に表示し、この中から印象に残った画像を検索してもらった。一人につき 2~3 枚を検索してもらった後に、以下

のようなアンケートを行った。このとき、ターゲットとする画像は表1に示すような順位で検索された。

表1: アンケートのための検索結果

順位	頻度(回)
1位	9
2~5位	2
6~10位	4
11~15位	1
出なかった	1

検索できなかったという結果が一回含まれているが、これは人物を配置した時に、本来「女性」と設定すべきところを「男性」と勘違いしてしまったために検索候補にあがらなかったのである。本システムでは、全ての項目に対して「わからない」というのを設けているので、はっきりと覚えていない限りこの項目を選んでもらうことで対応している。

「検索要求の表現は容易でしたか?」という質問に対しては、略画を描くことに関してはさほど難しくなく、自分の頭で思い描いたとおりにキャンバス上に表現できたので、満足しているようだった。

印象に関しては、もともとの「印象語」対「色ルール」の作成に使用した印象語が、絵画や色彩学などの専門家が使用する印象語であったり、日常的に使用する言葉ではないものであったのが問題点であった。

7 まとめ

まとめとしてまず、今後の課題をあげておきたい。

我われが選んだ印象語は同じような印象語がいくつかあるし、逆に付けようと思った印象語と似ている意味のものがないという欠点もあるので、まだ最適とはいえない。今後は、色だけではなく、形、レイアウト、オブジェクトのカテゴリなど、画像から求められるありとあらゆる

情報にもとづく印象語というものをサポートしてゆく必要があると思う。

本システムでは、オブジェクトの意味が、オブジェクトの色や印象語よりも重要度が高く検索に反映される。ユーザがどの何をどんなふうに見て覚えているかという指定のできる機構が必要であろう。

現在の検索システムは、検索時間がまだ十分に早いとは言えない。インデックスを充実させたり、画像をジャンル分けして、検索時間の短縮をはかる必要があろう。

最後になるが、本システムの検索方式は、従来の画像データベースシステムに比べて人間の感性という理想に近づいたものであることはまちがいない。あつかう画像の種類に対して適切なメニューの内容(オブジェクトの種類)を選べば、さまざまな分野で対応できるものと考えている。

参考文献

- [1] 西山 馨場 横山 松下. 略画作成による画像検索システムのインタフェース. 情報処理学会ヒューマンインタフェース研究会, 92-HI-44:pp133-140, 1992.
- [2] H.Nishiyama S.Kin T.Yokoyama Y.Matsushita. Image database reflecting human's sense. In *Proc. of 8th International Join Workshop on Computer Communication*, pages F4-2, 1993.
- [3] 千々岩. 色彩学. 福村出版, 1983.
- [4] J.A.Orenstein F.A.Manola. PROBE Spatial Data Modeling and Query Processing in an Image Database Application *IEEE Trans. on Software Engineering*, Vol.14, No.5, 1988.
- [5] 久野 フォルムス・色彩情報研究所. カラーインスピレーション. 講談社, 1992.
- [6] 香川 長谷川. 原色色彩語事典. 黎明書房, 1988.
- [7] 小林 道江. 応用色彩心理. 誠信書房, 1975.