

視覚心理空間を利用した画像検索法

藤田健二、古郡 了、畑 秀二
マツダ(株) 横浜研究所

本稿では「暖かい」、「近代的」といった形容詞をキーワードに用いて、形容詞から連想される画像を検索する方法について述べる。

一般に、人間の感覚を表現する形容詞と画像との対応関係は非常にあいまいなものとなる。そこでまず因子分析法を適用し、人間の視覚因子を抽出する。次に、形容詞を用いて視覚心理空間を構成し、この空間上で形容詞と画像とを対応づける特定化の方法を提案する。

この方法によって心理空間を介しての画像検索が可能となる。同一空間上に形容詞と画像を配置することによって対応関係の記述や検索処理を簡素化することができる。

A NEW REFERENCE METHOD TO SEARCH PICTURE IMAGE USING HUMAN PSYCHOLOGICAL SPACE BROUGHT BY VISION

Kenji FUJITA, Satoru FURUGORI, Syuji HATA
Technical Research Center(Yokohama), MAZDA Motor Corporation
2-5, Moriya-cho, Kanagawa-ku, Yokohama-shi, Kanagawa, Japan

In this paper, an 'image filing' system which has a new reference method to search picture is proposed. Adjectives such as 'warm' or 'modern' are used as keywords to search picture which may be imagined in human mind from the adjective keywords.

First, the correlations between the human feelings aroused by the picture and those by various adjectives are investigated by the panel-test. Then, using the factor analysis, a mathematical model of multi-dimensional space, which may be viewed as a human psychological space brought by vision, is formed so that all the picture and adjectives can be identified by the coordinates in the virtual space.

The advantages of this new filing system are as follows. Since the file indexes of the picture to be stored are simplified, the work load for the file maintenance is enormously decreased. Also the exploited mathematical scheme makes the searching process dynamic and flexible.

1. はじめに

今日のように、人々の感性や嗜好が多様化する中で、短期間に質の高い商品を開発することが求められる時代においては、コンピュータを活用することによって、人間の知的な活動や、創造的な活動を支援していくことが情報処理分野の重要な課題となってきている。

なかでも、商品コンセプトに基づいて新しい造形を作り上げていく工業デザイナーには、「ひらめき」や「インスピレーション」といった人間の感性に基づく高度な創造性が求められる。「何かを感じ、ひらめく」といった発想の過程をながめると、デザイナーは過去の体験や知識から形成された感性イメージ領域と「絵画」や「写真」といった具体的なイメージ領域とのやりとりを行うことが多い。

そこで、画像を用いて人間の感性領域に積極的に働きかけていくために、感性領域の心理表現語と具体的な画像との翻訳機能を実現することを試みる。そのためには、人間の視覚心理の構造を明らかにするとともに、視覚心理と画像とを対応づける、即ち画像を特定化する方法を検討する必要がある。またこれらの多量の画像を蓄積し、提供する画像データベースの機能が重要な役割を持つことになる。

画像をイメージ情報（例えば、“優しいイメージ”，“爽やかなイメージ”など）として捉えると連想されるキーは多様である。とりわけ、「暖かい」、「近代的」といった感覚属性と画像との対応関係は極めてあいまいなものとなる。このような情報を従来のキーワード法で管理しようとすると、予め画像個々に対して、あらゆる感覚属性を評価し、関連する属性データとしてそのすべてを登録しておく必要がある。これは一般に、多くの労力と時間を要する上、感覚属性と画像との関連度の記述やその処理機構を複雑にする。

そこで本稿では、人間の視覚心理を表現する空間（イメージ空間）を、形容詞を用いて定義し、この空間上であいまいさを許容した画像特定化の方法を提案する。画像を空間上に布置することは、感覚属性と画像との対応関係や関連度を暗に記述することであり、結果として特定化処理を簡単に行うことができる。

さらに、この方法に基づいて、デザイナー支援のためのイメージファイリングシステムを試作した。システムの概要及び機能についてもあわせて報告する。

2. 画像特定化のモデル

人間の感覚心理を比較的容易に表現できると考えられる形容詞からの画像検索を考える。この場合、問題となるのは、形容詞と画像との対応関係が複雑かつあいまいとなる点である。こうしたあいまい性の中には形容詞相互の意味のあいまい性と画像からいだけく視覚心理のあいまい性の両面を含んでおり、対応関係をより複雑にしているように思われる。従って、形容詞から画像を特定化するためには

(1) 形容詞間の相関を明確にするとともに、画像からいだけくあいまいな視覚心理を定量化する。

(2) 形容詞から画像を対応づける特定化モデルを決定する。

以上の課題を踏まえ、画像特定化のモデルを次のように考える。まず、形容詞そのものの持つ少数の潜在的因子を求める（例えば、“陽気な”、“明るい”、“活発な”などはある共通因子を含むと考えられる）。すなわち、相関を持つ形容詞をこれらの因子に縮約し、各形容詞を因子を用いて説明することを考える。さらに、これらの因子を座標軸とし、具体的な形容詞を配置させたイメージ空間を構成し、形容詞と画像とをつなぐ共通の土俵を作る。次に、これらの因子を用いて、画像の持つあいまいな視覚心理を定量化し、空間上にマッピングする。両者の位置関係を比較することによって、形容詞から画像を特定化することができる。と考える。

3. イメージ空間の定義と画像の特定化

3.1 因子モデルとイメージ空間

人間の視覚心理に関する因子を抽出するために、ここでは因子分析法を適用する。今、形容詞間の相関を m 個の因子で説明できると仮定すると、一般に因子モデルは (1) 式のように書くことができる。

$$[Z] = [A] [F] + [E] \quad (1)$$

$[Z]$: 観測値 $[A]$: 因子負荷量
 $[F]$: 因子得点 $[E]$: 独自因子
 p : 形容詞数 n : 画像数

観測値 z_{pn} は、画像 n について観測される形容詞 p の評価データである。つまり、画像を見てどう感じるかを形容詞毎に調査したデータである。因子負荷量 a_{pm} は、観測値 z_{pn} に m 番目の因子がどれだけ反映するかを表す。言いかえると、形容詞の各因子に対する重み係数である。また、因子得点 f_{mn} は m 番目の因子が画像 n にどの程度強くはたらくているかを示す値（スコア）である。

ここで、因子が互いに無相関と仮定すれば、因子負荷量 $[A]$ は次式を解くことにより推定できる。

$$[C] = [A] [A]^T + [D] \quad (2)$$

$[C]$: 形容詞間の相関行列
 $[A]^T$: $[A]$ の転置行列
 $[D]$: $[E]$ の分散行列

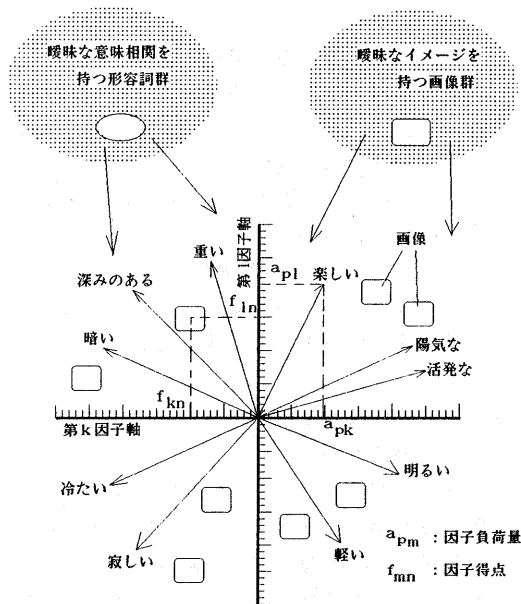


図1 イメージ空間の概念

次に、(2)式より定まる因子負荷量を用いて画像の因子得点を推定する。因子得点の推定法はいくつか知られているが、ここでは回帰法を適用する。

$$[F] = [A]^T [C]^{-1} [Z] \quad (3)$$

こうしてモデル(1)のもとで求まる視覚心理因子、及び形容詞の因子負荷量からイメージ空間を構成できる。さらに(3)式より得られる因子得点から画像を同一空間上にマッピングすることができる。ここで作られるイメージ空間の概念図を図1に示す。

3.2 画像の特定化

視覚心理が m 個の因子によって説明できるとすれば、(2)式から求まる因子負荷量は形容詞と各因子との相関係数そのものを表わすことになる。一方、因子得点はある因子が各画像に対してどの程度強く働いているかを示す量であり、因子負荷量とは意味や扱いが異なる。従って、直接両者を比較演算することはできないが、各因子との関係を表わしている点については共通しており、因子を介して対応付けることができると考える。

空間の持つ意味を考慮し特定化の方針を次のように決めた。

- (1) まず、因子負荷量を用いて形容詞から因子を対応づけ、次に因子得点を用いて因子から画像を対応づけることを考える。即ち、形容詞と各因子の相関を見た場合、相関の高い因子であれば、その因子について因子得点の高い画像を特定化の対象とし、相関の低い因子であれば得点の低い画像を対象とする。
- (2) 個人あるいはクラス（性別、年齢別など）によって視覚心理はばらつく。そこで特定化においては、一つの画像に絞りにくく、類似の視覚心理を与える複数の画像を取り出す。
- (3) 原点近傍に位置づけられる画像は因子得点が低く、特徴のない画像と考えられる。従って、原点からより遠くに位置付けられている画像を優先して特定化する。

以上の方針に従って、今回は図2に示すようなくさび状の領域を特定化の対象とした。特定化は次の手順で行なう。

- (1) 空間における形容詞の位置ベクトル \vec{P}_i と画像の位置ベクトル \vec{Q}_j とのなす角度を逐次求める。
- (2) 角度が許容範囲 ϵ 内であれば、類似の視覚心理を与える画像として特定化する。
- (3) 画像が複数特定化された場合、各画像に対して原点からの距離 d_j を逐次求め、遠い画像から順位づける。

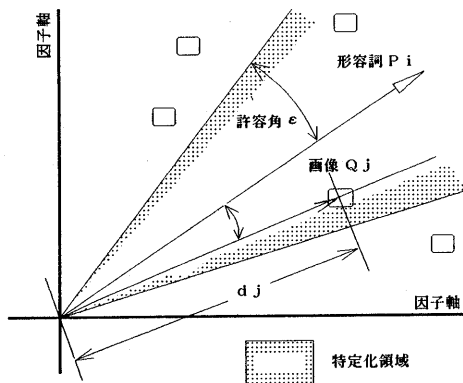


図2 画像の特定化

3.3 画像の追加蓄積の簡略法

画像を追加蓄積する場合、3.1で示した因子モデルを適用し、推定式(2),(3)から因子得点を求めることができるが、次のような問題が予想される。

- (1) 観測値 $[Z]$ を得るためには、索引となるすべての形容詞について画像を評価する必要がある。一般にこの作業は非常に多くの労力と時間を要する。
- (2) 得られた観測値に推定式(2)(3)式を適用すると、再度、因子や因子負荷量を求めることになり、効率が良くない。
- (3) 追加画像の質(画像の持つ視覚心理情報)に偏りがある場合、先に求めた因子や因子負荷量と値が大きすぎる恐れがある。この差が著しいと画像のマッピングができなくなる。

こうした問題を解決するために、ここでは次のような簡略法を考える。

まずイメージ空間作成時に得られる形容詞の相関行列 $[C]$ 及び因子負荷量 $[A]$ を追加画像の因子得点推定に流用する。本来、画像を評価する実験条件が一致していれば相関行列や相関行列から求まる因子負荷量はほぼ同様なものが得られるはずである。 $[C]$ 、 $[A]$ を流用することは先に得られるイメージ空間をそのまま利用することであるから再度、因子や因子負荷量を求める必要がなくなる。同時に因子得点を求める上で、追加画像の質や種類の偏りに影響されることはない。

さらに、画像評価の負担を軽減するために、因子や形容詞相互の関係を利用することによって評価に使用する形容詞を少数の形容詞に絞りこむ。推定式(3)は、形容詞に関する線形式にあてはめることによって得点を推定しようとする方法である。従って、各々の因子を代表する形容詞を選んでおけば大きな情報落ちが避けられ、簡便な近似解として利用できると思われる。以下に、追加画像の因子得点推定式を示す。

$$[F] = [A^*]^T [C^*]^{-1} [Z] \quad (4)$$

$m \times n$ $m \times p^*$ $p^* \times p^*$ $p^* \times n$

$[Z]$: 追加画像の観測値

$[F]$: 追加画像の因子得点

$[A^*]$: $[A]$ から抜きだした絞りこんだ形容詞の因子負荷量

$[C^*]$: $[C]$ から抜きだした絞りこんだ形容詞の相関行列

p^* : 絞りこんだ形容詞数

z の正規化に使用する平均値、標準偏差は空間作成時に求めた値を使用する。

4. 官能評価実験と結果

4.1 実験方法

イメージ空間作成のための実験と多量の画像を追加するための実験と2回に分けて行った。

(1) イメージ空間作成のための実験

首都圏在住の20代から40代の男女150名に対して50枚の画像によるSD尺官能評価実験を行った。日常頻繁に用いられる180個の形容詞対(例えば、"豪華な-質素な"といった対語)

に対して各々7段階で評価してもらった。画像は自然・風景、インテリア・建築など6ジャンルのもを用意した。また評価パネルは150名から無策意に抽出し、1人当たり18形容詞対、50画像を評価してもらい、結果として1画像1形容詞対当たり15名の平均がとれるように計画した。

実際の実験にあたっては、画像をスライドにし、18形容詞対、50画像を2時間で評価してもらった。画像の提示順序や形容詞の順序を入換えることにより順序効果をなくすように配慮した。

(2) 画像追加のための実験

追加する画像を500枚に増やし、逆に形容詞を因子を代表する18形容詞対に絞りこんで(1)と同様な実験を行った。この場合も1画像1形容詞対当たり15名の平均がとれるようにした。形容詞の絞りこみに際しては、まず、実験(1)で得られた観測値を因子分析し、各因子と特に相関の高い形容詞を選択した。

4. 2 因子分析結果

(1) 因子と因子負荷量

まず、実験(1)の観測値を分析し、因子数mを決定することによって因子モデル(1)を同定した。因子数mを決定する方法として、様々な情報量基準が提案されているが、相関行列の固有値を用いるもの、寄与率(共通性)¹⁶⁾を用いるものが一般的である。ここでは各因子の寄与率を求めた。寄与率とは、180形容詞対による観測値zの全分散のうち、各因子によって説明される割合を表わし、各因子負荷量の自乗和から求まる。表1に計算結果を示す。これを見ると第6因子と第7因子の間の落差が大きい。そこで今回はm=6を採用することにした。6因子の寄与率を合計すると78.5%となり、人間の視覚心理の大半を説明できると考えた。

次に、各形容詞の因子負荷量を求めるために推定式(2)を主因子法を使って解いた。また各因子の意味づけを容易にするためにバリマックス法による座標変換を行った。この結果得られた因子負荷量の大きさ及び符号に着目し因子を意味づけると、それぞれ画像の「豊かさ」「暖かさ」「新しさ」「動き」「簡素さ」「幻想性」を表わしていると解釈した。各因子と特に相関の高かった形容

表1 各因子の寄与率(%)

因子	1	2	3	4	5	6	7	...
寄与率	20.3	17.9	12.0	11.5	11.4	5.4	2.0	...
累計	20.3	38.2	50.2	61.7	73.1	78.5	80.5	...

$$[\text{寄与率} = \sum_i a_{im}^2 / \sigma \quad \sigma: \text{観測値}[Z] \text{の全分散}]$$

表2 因子の意味づけ

因子の意味づけ	相関の高い形容詞の例	
	正の相関	負の相関
1 豊かさ	立派な 豪華な...	庶民的な 質素な...
2 暖かさ	温厚な 家庭的な...	機械的な 冷たい...
3 新しさ	近代的な 進歩的な..	古典的な 伝統的な..
4 動き	行動的な 興奮した..	静かな 落ち着いた..
5 簡素さ	淡泊な 単純な...	濃厚な 華美な...
6 幻想性	幻想的な 芸術的な..	現実的な 実用的な..

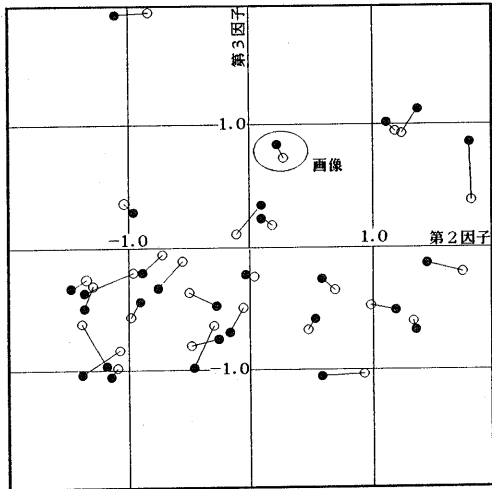
詞を表2に示す。

(2) 因子得点

予め、形容詞絞りこみに伴う推定値の変動を確認するために実験(1)で得られた50画像の観測値について18形容詞対と180形容詞対によるそれぞれの推定値を比較した(即ち、両者とも評価パネル、画像など同一条件で得られた観測値を使用している)。結果の一部を図3に示す。画像によってその変動は異なるが、相対変動は10-20%程度であり全体的に大きくなれば生じていない。従って、実験(2)で得られた観測値を使って500画像の因子得点を推定し空間上に配置させた。

(3) 個人による心理のずれ

一般にクラスタ(年齢、性別、職業別など)や個人によって言葉(形容詞)の定義や解釈あるいは感じ方、感度に違いがある。そこで、今回の実験では年齢、性別等の偏りをできるだけなくすことによって平均的な人間の心理が分析できるように配慮した。こうした平均値は多くの場合、大勢



● 180形容詞対による因子得点
○ 18形容詞対による因子得点

図3 形容詞絞り込みに伴う因子得点の変動
(第2-第3因子平面)

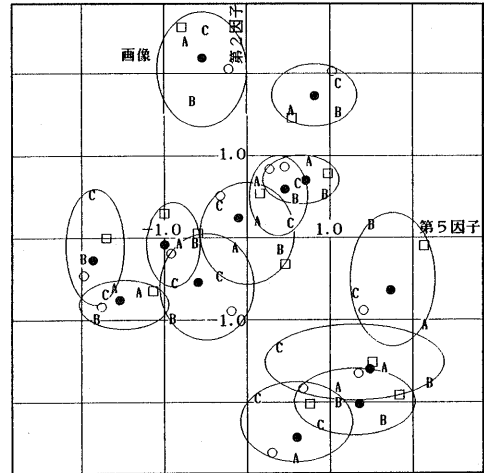
の感覚に近いと考えられるが、クラスタあるいは個人のばらつきが大きくなると平均値から離れる人間が増えてくる。そこで今回の実験では個人差の生じやすい形容詞(“好き”、“きれい”、“ナウい”、“ダサイ”など)は省いて評価を行った。更に、3.2で述べたように、平均値を中心に、ある許容範囲 ϵ を設け類似の視覚心理を与える複数の画像を取り出し、できるだけばらつきを許容するようにした。図4はクラスタ別に求めた因子得点の例である。こうした実際のばらつきに基づいて ϵ の値を決定した。

5. イメージファイリングシステムへの応用

以上述べてきたイメージ空間を利用した画像の特定化法を応用し、静止画検索システム“イメージファイリングシステム”を試作した。イメージ空間は先に得られた6つの因子を使って6次元空間として表現した。以下システムの概要、機能について紹介する。

5.1 システムの概要

商品のプランニングやデザイン活動の過程においては、商品コンセプトの創案、カラーニュートン作成等によるコンセプトのイメージ表現、あるいは新しい造形のヒントを得るために日頃から多くの写真、



● 平均値 A 20代
□ 男性 B 30代
○ 女性 C 40代

図4 クラスタ別因子得点の比較
(第2-第5因子平面)

絵画、ポスター等が集められ、活用されている。本システムはこれらの画像を画像データベース上で一括管理するものである。商品の感覚的なコンセプトの仮説を試行錯誤で発想するための支援ツールとして活用していくことを目的としている。また、頭に描いたイメージを具体的な画像に置きかえ、相手に伝達するためのツールとしても活用することができる。

そこで本システムを設計する上で、以下の点について留意した。

- (1) 所望のイメージ画像を形容詞を用いて高速に取り出せること
- (2) 原画像の持つ視覚心理情報をできるだけ損なうことなく高精細な画像が蓄積、提示できること
- (3) プランナーやデザイナーにとってできるだけ容易な操作性を持つこと

5.2 システム構成

図5にシステム構成を示す。デザイン作業等、現場での利用、及び複数の利用者を想定し、形容詞情報、画像属性情報は全て、センターのホストコンピュータ上で一括管理する。またこれらの情報検索、管理用ソフトの他、端末装置からの検索

要求を受け付ける通信制御用ソフトを持つ。一方、ホストコンピュータとつながれる端末装置上にも、これらの情報やメッセージを送受信する通信ソフトの他、利用者とのマンマシーンインタフェースを容易にする画面制御用ソフト、画像出力コマンドの送信と光ディスク装置の状態を監視する画像出力制御用ソフトを開発し備えた。また、原画像の画質が大きく劣化せず、画像の追加蓄積が容易に行える追記型光ディスクを採用した。さらに、臨場感ある画像が提示できる大画面プロジェクタ^[3]を設け、光ディスク制御装置を介して端末装置に結合した。

5.3 操作性

本システムではキーワードの指定や検索、出力等のコマンド指定は全て端末画面上のメニューやアイコンをマウスで指示することによって行うことができる。図6にメニュー画面の様子を示す。システムには、334種類の形容詞がキーワードとして登録されているが、これらを相関の高い形容詞どうし7グループに分類し、各グループ毎にイメージを表わすガイド語を付けた。ガイド語を指定することによって相関の高い形容詞がメニュー表示され、入力したい形容詞を容易に選択することができる。

5.4 諸機能

本システムは、図7に示すような機能構成をなす。以下主な機能の概要を述べる。

(1) 形容詞の複合入力機能

2つの形容詞の因子負荷量ベクトルの内積は、形容詞間の相関係数そのものを表わす。つまり空間上での2つの形容詞の位置ベクトルのなす角度は、相関度を意味している。この位置関係を利用し、本システムでは形容詞の複合入力を受け付ける。それぞれの因子負荷量の平均値から画像を特定化する。

(2) 画像検索機能

本システムでは、「風景」、「建築物」といった画像対象自身の持つ属性（物理的属性）からと「暖かい」「淡い」といった視覚心理としての属性（感覚属性）からの2通りの検索形態をとる。両キーワードを併用することによって、利用者は、

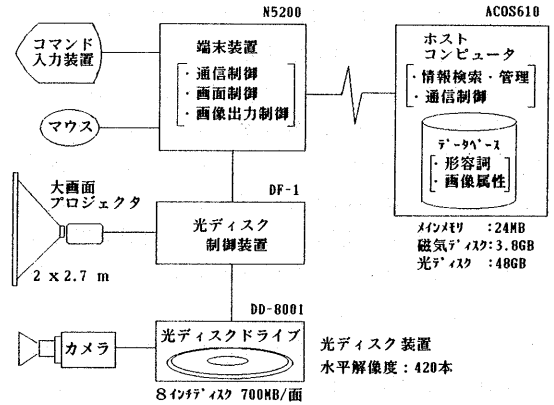


図5 システム構成

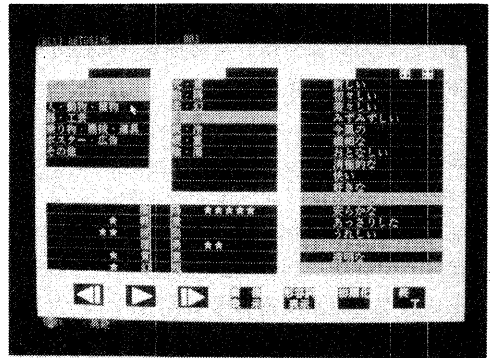


図6 メニュー画面

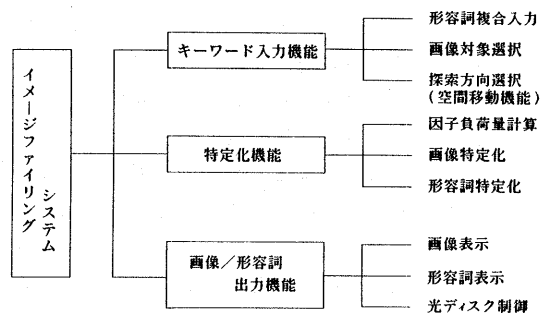


図7 機能構成

所望の画像を容易に引き出すことができる。

(3) イメージ空間移動機能

6軸の因子には、4.2で行ったように便宜上、「動一静」といった意味づけが施されている。各軸に向かって方向を定義することによって、利用者が自由に探索方向を操作できるようにした。移動後の点(新しい因子負荷量)は、空間上現在点の位置ベクトルを、指定軸方向に向かって、回転させることによって得られる。

さらに、空間上現在位置を端末グラフィック上に表示する機能を設けた。

(4) 形容詞検索機能

「画像対象」、「形容詞」、「探索方向」によって検索、表示される画像から、逆に形容詞への検索を可能にした。形容詞から画像への特定化と同様な方法を適用している。検索された画像が与える視覚心理の評価や新しく追加された画像の確認作業等に利用できる。

6. おわりに

「ひらめき」や「インスピレーション」といった人間の高度な創造性には、そのメカニズムが充分解明されていない部分が多い。従って、現状では、その発想過程を直接情報処理の過程に展開することは、極めて困難である。そこで今回、視覚心理を統計的に分析することによって人間の感性領域に積極的に働きかけていくことを試みた。具体的な方法として本稿では、画像を単なる「風景画」、「建築物」として捉えるのではなく、「優しいイメージ」、「爽やかなイメージ」といったイメージ情報として扱う新しいタイプの画像データベースの構築方法を提案し、その利用形態について述べた。以下、報告内容をまとめると、

- (1) イメージ情報を蓄積、管理する画像データベース技術と人間の視覚心理を分析する統計解析技術との融合を図ることによって、形容詞から画像への検索が可能であることを示した。
- (2) 複雑な相関を持つ形容詞からあいまいなイメージを持つ画像への検索システムにおいて、特定化モデルを提案することにより、形容詞と画像との対応関係の記述を簡素化し、容易に検索できることを示した。
- (3) 本研究の応用例として、デザイナー支援のた

めのイメージファイリングシステムの概要を報告した。

謝辞 本研究を進めるにあたり、多大な支援を頂いた山口真人課長をはじめとする日本電気(株)の方々、ならびに有意義な討論と助言を頂いたマツダ(株)横浜研究所、井上寛研究員に感謝します。

[参考文献]

- [1] 藤田,古郡,畑,麦谷:”デザイナー支援のためのイメージファイリングシステム”,第35回情報処理学会全国大会予稿集,(1987).
- [2] 古郡,藤田,畑:”6軸の視覚イメージ空間を用いた画像検索法について”,官能検査シンポジウム,(1987).
- [3] 井上,香川,畑,片岡:”視覚による人間感覚研究のための実験・解析システム”,第35回情報処理学会全国大会予稿集,(1987).
- [4] 小谷,畠中,坂下:”電気製品のカラー計画におけるイメージ分析”,官能検査シンポジウム(1986).
- [5] 小林:”カラーイメージ事典”,日本カラーデザイン研究所.
- [6] 田中,脇本:”多変量統計解析法”,現代数学社