

## 感性の主因子に基づく画像・自然音アクセスに関する一検討

宝珍 輝尚 山田 恵一 都司 達夫

福井大学 工学部 情報・メディア工学科

〒 910-8507 福井市文京 3 丁目 9-1

hochin@pear.fuis.fukui-u.ac.jp

本論文では、感性の主因子を用いて感性語により多種のメディアデータの検索を行うシステムについて検討する。SD法を用いて画像、自然音に対する感性の主因子を求めたところ、画像に対しては、「評価性」、「力量性」、「活動性」、「壮快性」、「光輝性」という5つの主因子が求まり、自然音に対しては、「評価性」、「活動性」、「緊張性」、「力量性」という4つの主因子が求まった。これらの感性の主因子に対する評価尺度の算出方法を検討し、検索システムを構築した。また、おのおの感性の主因子を対応付けることにより画像と自然音の直接的な相互アクセスを可能とするシステムを実現した。

## On the Access of Pictures and Sounds Based on the Human Sensitivity Factors

Teruhisa HOCHIN Keiichi YAMADA Tatsuo TSUJI

Dept. of Information Science, Faculty of Eng., Fukui University

3-9-1, Bunkyo, Fukui-shi, Fukui 910-8507

hochin@pear.fuis.fukui-u.ac.jp

This paper revises the multimedia data retrieval system using sensitivity main factors. First, sensitivity main factors are obtained. The factors of evaluation, activity, potency, excitement, and brightness are obtained for pictures, and those of evaluation, tension, activity, and potency are obtained for sounds. The methods of deriving the feature values corresponding to these factors are presented. A multimedia data retrieval system with impression words has been constructed based on these factors and feature values. This system also enables users to access sounds directly from pictures, and vice versa, by adapting the sensitivity factors of pictures to those of sounds, and vice versa.

## 1 はじめに

近年、マルチメディアデータに対する検索の研究が盛んに行なわれている。スケッチ画像をキーとする類似画検索 [1]、内容を表現する一種のネットワークを用いる内容検索 [2, 3]、特徴量空間における所望の範囲を指定する検索 [4]、感性語を用いた検索 [4, 5, 6, 7] などの研究である。この中で、人間の感性に基づく検索ができる感性語を用いた検索は、利用者に理解しやすく利用しやすいという利点がある。しかし、感性語を用いた研究の多くではマルチメディアデータを扱うにあたり、そのメディア種別や対象に深く依存した特徴量を用いており [6, 7]、多種多様なメディアデータを統一的に扱うことは考慮されていない。また、良好な結果を得るためにはかなりの学習が必要となっている [4]。

このような問題を解決するために、筆者らは、様々なメディアデータを感性語により検索するための統一的な枠組みについて検討し、人間の感性は少数の支配的な因子（感性の主因子）によって記述できそうであるという心理学の研究成果 [8, 9, 10] を利用したマルチメディアデータ検索システムを構築してきた [12, 13]。本検索システムでは、感性語がいくつかの感性の主因子とのかかわりの程度（因子負荷量）によって表現できることを利用している。また、画像データと自然音データから人間の感性の主因子に相当する特徴量（評価尺度）を抽出し検索に利用している。しかし、このシステムでは、画像データに対応する自然音データを直接求めたり、自然音データに対応する画像データを直接求めることはできなかった。

そこで、本論文では、異種のマルチメディアデータ間の直接的なアクセスを可能とすることを目的として、画像データと自然音データ間の相互アクセスを可能とするための検討を行う。このために、まず、画像と自然音の感性の主因子を再考する。さらに、評価尺度の検討も行う。その上で、画像と自然音の感性の主因子の対応付けを検討する。

以下、2では画像と自然音に対する感性の主因子について述べ、3では評価尺度の導出について述べる。次に、4で、検索システムについて述べ、5で、検索例を示す。最後に、6でまとめを行う。

## 2 感性の主因子

### 2.1 SD 法

心理学の研究では、形容詞が表す感性的な性質はいくつかの基本的な因子（感性の主因子）によって表現されることが明らかとなっている。これは、Semantic Differential (SD) 法という手法を用いて解析した結果得られるものである。SD 法とは心理学者 C. E. Osgood が考案した手法で、複数の反対の意味を持つ形容詞の対を尺度とし、その間をいくつかの段階に分けてある対象物を被験者に評価させるものである [8, 9, 10]。これらの形容詞対がそれぞれ単一次元の特性を抽出すると仮定し、因子分析を適用すると、次元の少ない、より簡潔なものとすることができる。

### 2.2 画像における感性の主因子

検索に用いる感性の主因子を得るために、画像に対する感性的印象を測定する。測定方法は、男子大学生 12 名に対して 40 個の画像を用いて SD 法により、2, 1, 0, -1, -2 の 5 段階で評価してもらう。用いた形容詞対は表 1 に示した 16 個である。画像は、風景や植物などの自然画である。

SD 法により得られたデータを各画像、各形容詞対ごとに平均を求め、因子分析を行なった。1 より大きい固有値を持つ因子を採用することにすると、因子数は 5 となった。因子数を 5 として因子分析した結果を表 1 に示す。

第一因子は『明るい-暗い』『大胆な-繊細な』などの因子負荷量が多いことから『壮快性』の因子とした。

第二因子は『暖かい-冷たい』などの因子負荷量が多いことから『活動性』の因子とした。

第三因子は『重い-軽い』などの因子負荷量が多いことから『力量性』の因子とした。

第四因子は『うれしい-悲しい』などの因子負荷量が多いことから『評価性』の因子とした。

第五因子は『自然な-不自然な』『潤いのある-渴いた』などの因子負荷量が多いことから『光輝性』の因子とした。

表 1: 画像の感性語の因子分析結果

形容詞対		第一因子	第二因子	第三因子	第四因子	第五因子
美しい	— 醜い	0.225	-0.139	-0.634	-0.186	-0.259
自然な	— 不自然な	0.462	0.282	0.238	-0.150	0.997
新鮮な	— 古くさい	0.678	-0.154	-0.069	0.288	-0.103
澄んだ	— 濁った	-0.484	-0.441	-0.490	0.478	-0.218
潤いのある	— 渴いた	-0.086	0.294	-0.197	0.029	-0.802
鋭い	— 鈍い	0.035	0.339	-0.007	0.156	0.782
暖かい	— 冷たい	-0.094	0.934	-0.023	-0.059	0.157
明るい	— 暗い	0.930	-0.193	-0.380	0.420	0.285
単純な	— 複雑な	0.174	-0.117	0.032	0.217	-0.053
大胆な	— 繊細な	0.923	-0.007	0.014	-0.002	0.024
動的な	— 静的な	0.367	-0.202	0.123	0.332	0.180
緊張した	— ゆったりした	-0.176	0.518	0.481	0.305	-0.094
大きい	— 小さい	-0.360	0.823	0.059	0.157	-0.556
重い	— 軽い	0.249	-0.088	1.016	-0.136	0.371
強い	— 弱い	0.330	-0.009	0.061	0.039	0.379
うれしい	— 悲しい	0.572	0.193	-0.184	0.670	-0.203

## 2.3 自然音における感性の主因子

男子大学生 12 名に対して 40 個の自然音を用いて SD 法により 5 段階で評価をしてもらう。用いた形容詞対は表 2 に示した 16 個である。

画像のときと同様に、SD 法により得られたデータを各自然音、各形容詞対ごとに平均を求め、因子分析を行なった。固有値が 1 より大きい因子の数は 4 であったので、自然音検索に用いる因子数は 4 とした。因子数を 4 として因子分解した結果を表 2 に示す。

第一因子は『大胆な—繊細な』『重い—軽い』『強い—弱い』などの因子負荷量が大きいことから『力量性』の因子とした。

第二因子は『新鮮な—古くさい』などの因子負荷量が大きいことから『評価性』の因子とした。

第三因子は『動的な—静的な』『潤いのある—渴いた』などの因子負荷量が大きいことから『活動性』の因子とした。

第四因子は『緊張した—ゆったりした』などの因子負荷量が大きいことから『緊張性』の因子とした。

## 3 評価尺度

画像・自然音に対する感性の主因子に対応する特徴量(評価尺度)の抽出方法について述べる。

### 3.1 画像の評価尺度

#### 3.1.1 壮快性

「明るい—暗い」と「大胆な—繊細な」が同程度の重みをもっているため、両方を考慮することとする。

明暗は、明度そのもので良いであろう。次は、大胆さである。ここでは、画像に含まれる最も多い色相と、次に多い色相が相反するもの(例:赤 ←→ 青)であれば大胆であり、逆にその 2 つの色相が同系色であれば鈍い画像であると考えた。そこで色相配列の最大値と次に大きい値の距離を評価尺度とすることとした。距離の大きいものが「大胆」を表し、小さいものが「繊細」を表すものとする。

以上の明るさと大胆さを 0 から 1 に正規化し、それらが構成する二次元平面の原点からの距離を壮快性の評価尺度とする。

表 2: 自然音の感性語の因子分析結果

形容詞対		第一因子	第二因子	第三因子	第四因子
美しい	— 醜い	0.215	0.317	0.178	0.328
自然な	— 不自然な	-0.300	0.185	0.557	-0.752
新鮮な	— 古くさい	-0.022	1.518	0.102	-0.291
澄んだ	— 濁った	0.422	0.485	-0.043	0.211
潤いのある	— 渴いた	0.255	0.357	0.723	0.292
堅い	— 柔らかい	0.126	-0.067	-0.292	0.124
暖かい	— 冷たい	0.190	0.367	0.423	-0.303
明るい	— 暗い	-0.101	0.082	-0.317	0.244
単純な	— 複雑な	0.441	0.018	0.364	0.015
大胆な	— 繊細な	0.950	-0.078	0.237	-0.971
動的な	— 静的な	0.054	-0.002	-0.752	-0.774
緊張した	— ゆったりした	0.008	0.076	-0.042	-1.170
大きい	— 小さい	0.103	0.186	-0.106	-0.153
重い	— 軽い	0.730	-0.089	0.209	-0.101
強い	— 弱い	0.555	0.153	-0.286	0.239
うれしい	— 悲しい	-0.231	0.431	0.342	-0.022

### 3.1.2 活動性

文献 [9] によると、「暖かい」を表す色は赤色、「冷たい」を表す色は青色としている。ここでは、画像の代表的な色相を導出し、その色相の赤・青の含んでいる度合いを活動性の評価尺度とした。

ここで、色相は角度 ( $0^\circ \sim 360^\circ$ ) によって表現されているので、赤の色相値 ( $113^\circ$ ) を最大、青の色相値 ( $353^\circ$ ) を最小となるように次式により評価尺度  $Mia$  を求める。

$$Mia = 2hue + 14 \quad (0 \leq hue \leq 113) \quad (1)$$

$$Mia = 353 - hue \quad (113 \leq hue \leq 353) \quad (2)$$

$$Mia = 2hue - 706 \quad (353 \leq hue \leq 360) \quad (3)$$

その評価尺度値  $Mia$  が小さいものは、青を多く含んでいると考え、評価は「冷たい」とする。 $Mia$  が大きいものは、赤を多く含んでいると考えて、評価は「暖かい」とする。

### 3.1.3 力量性

彩度と明度の2軸からなる平面で表されているトーン平面では、彩度軸側には主に色の濃いものが分布し、明度軸側には色の薄いものが分布している。

ここでは、まず画像の彩度・明度の平均値を求め、彩度と明度で構成された平面上の点とする。そして、原点を中心に原点とその点を結ぶ直線と明度軸との成す角度を力量性の評価尺度とした。その評価尺度値の大きいものが「重い」を表し、小さいものが「軽い」を表すとした。

### 3.1.4 評価性

明度と彩度の2軸からなるトーン平面の原点に近い部分に暗く地味な色彩が多く、遠い部分に明るく派手な色彩が多いため、明度と彩度の平均の座標と原点までの距離を評価尺度とした。距離が大きいものが「うれしい」を表し、小さいものが「悲しい」を表すものとする。

### 3.1.5 光輝性

「自然な—不自然な」,「潤いのある—濁いた」,「鋭い—鈍い」といった形容詞対が関与している。ここでは、光沢のあるものが関係していると考え、彩度を評価尺度とすることとした。

## 3.2 自然音の評価尺度

### 3.2.1 力量性

音の最小値と最大値の差を評価尺度とする。差が大きいものが「大胆」,「重い」を表し、小さいものが「繊細」,「軽い」を表すものとする。従って、「評価性」の評価尺度  $M_{sp}$  を、音の最大値を  $v_{max}$ 、音の最小値を  $v_{min}$  としたとき、 $M_{sp} = v_{max} - v_{min}$  により求める。

### 3.2.2 評価性

高い音が「新鮮」な音であると考えた。そこで音高の平均を求め、この値を評価尺度とする。この値の大きいものが「新鮮」を表し、小さいものが「古くさい」を表すものとする。

### 3.2.3 活動性

テンポが速く、かつ、高い音が「動的」で「潤いのある」音であると考えた。そこで音高とテンポの平均の値からこの2軸の平面上の座標を求め、原点からの距離の値を評価尺度とする。この値の大きいものが「動的」,「潤いのある」を表し、小さいものが「静的」,「濁いた」を表すものとする。

### 3.2.4 緊張性

ここでは、音の周波数成分の片寄りが大きいほど緊張性が高いと考えた。そこで、パワースペクトル曲線の近似直線の上にある部分の平均により求まる直線よりも上方にある部分の数を評価尺度とする。ここでは、この部分の数が大きいものが「ゆったりした」を表し、少ないものが「緊張した」を表すものとする。

## 4 検索システム

### 4.1 感性語による画像・自然音検索

因子分析の結果得られた因子負荷量が表す感性語の重みと画像の5因子(自然音は4因子)の評価尺度とした特徴量によって検索を行なう。また、感性語の選択画面で、ユーザーが感性語の重み(1~10)も選択可能とした。これにより「非常に美しくやや暗い」など微妙なニュアンスのものも検索できるようになった。

使用する感性語は従来のもと同じ形容詞対を用いる。検索は以下に行なう。

1. 全ての画像・自然音データから特徴量を抽出する。
2. ユーザーは検索に用いる形容詞を1つ以上選択する。ただし、相反する形容詞(例:「明るい」と「暗い」)は同時に選択できない。
3. 選択した形容詞の重みを決定する。
4. 選択された形容詞の因子負荷量・重みと画像・自然音データから抽出した評価尺度の積をとり、検索対象となっているすべてのデータごとに和を求め評価値とする。
5. 評価値の大きいものから5個を検索結果として出力する。

手順4において、重みをかけることにより重み付けを行ない、微妙なニュアンスのデータの検索に対応している。

### 4.2 メディアデータによる他メディアデータ検索

感性語による検索結果に適した他メディアデータを検索では、感性語による検索結果の特徴量と他メディアデータの特徴量を用いて検索を行なう。検索手順を以下に示す。

1. 感性語による検索システムから画像・自然音を検索する。
2. 検索結果の特徴量と他メディアのデータの特徴量を比較し、5次元空間での距離をとり評

価値とする。ここで、活動性と評価性は、画像と自然音のそれぞれを対応させる。画像の光輝性は自然音の緊張性に対応すると考え、そのように対応させる。また、画像の壮快性は力量性に通じると考え、画像の力量性と壮快性を自然音の力量性に対応させて距離を算出する。

3. 評価値の小さいものから10個を検索結果として出力する。

## 5 検索例

各検索システムの検索例を以下に示す。

### 5.1 感性語による画像検索

感性語による検索例を示す。図1に検索画面を示す。ここでは、“自然で静的な”画像の検索を行なっている。また、それぞれの感性語の重みは、“自然な”が”8”、“静的な”が”3”である。検索結果を図2に示す。

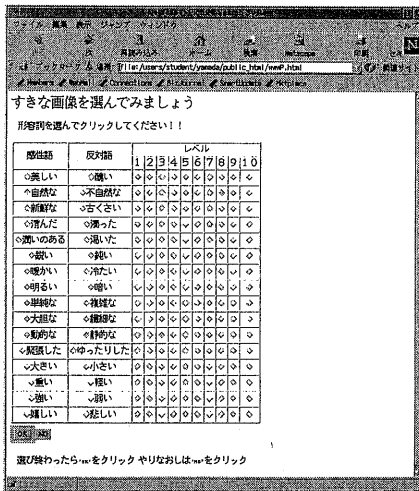


図1: 感性語による画像検索画面

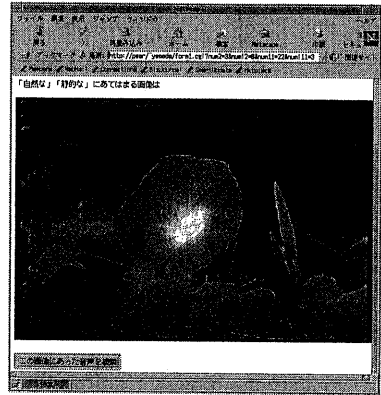


図2: 感性語による画像検索結果

### 5.2 感性語による自然音検索

感性語による検索例を示す。図3に検索画面を示す。ここでは、“小さく、澄んだ”自然音の検索を行

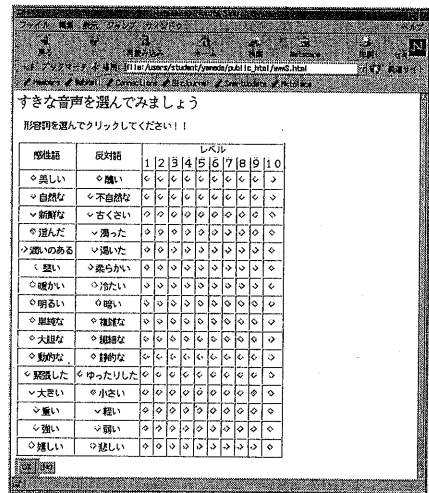


図3: 感性語による自然音検索画面

なっている。また、それぞれの感性語の重みは、“澄んだ”が”3”、“小さい”が”9”である。検索結果を図4に示す。ここでは、鳥の鳴き声、剣の交わる音、ガラスの割れる音、トンビの鳴き声などが求まっている。

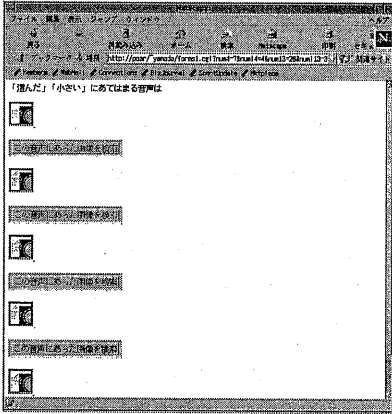


図 4: 感性語による自然音検索結果

### 5.3 メディアデータによる他メディアデータ検索

メディアデータによる他メディアデータ検索の検索例を示す。

図 2 で得られた朝顔の画像に対応する自然音を検索した結果の画面を、図 5 に示す。検索の結果、蝉

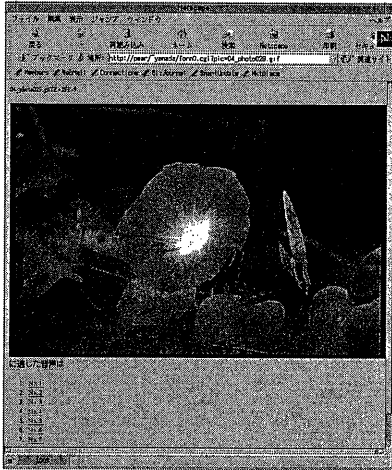


図 5: 画像による自然音検索結果

の声、せせらぎの音、蛙の鳴き声や栓を抜いた音が求まっている。

また、図 4 で得られた自然音の中で剣の交わる音に対応する画像を検索した結果の画面を、図 6 に示す。得られた画像の一つを、図 7 に示す。

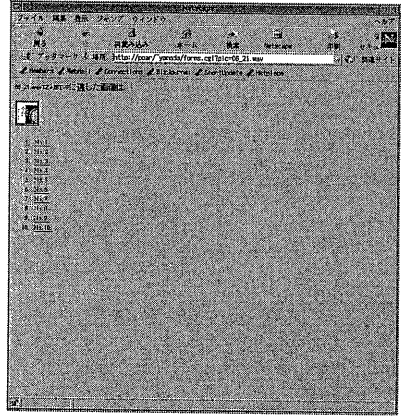


図 6: 自然音による画像検索結果画面

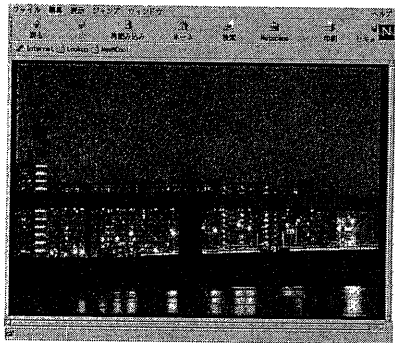


図 7: 自然音による画像検索結果

## 6 おわりに

感性の主因子を用いて多種のメディアデータの検索を行うシステムについて検討した。SD 法を用いて画像、自然音に対する感性の主因子を求めたところ、画像に対しては、「評価性」、「力量性」、「活動性」、「壮快性」、「光輝性」という 5 つの主因子が求まり、自然音に対しては、「評価性」、「活動性」、「緊張性」、

「力量性」という4つの主因子が求まった。これらの感性の主因子に対する評価尺度の算出方法を検討し、検索システムを構築した。また、おのおのの感性の主因子を対応付けることにより画像と自然音の直接的な相互アクセスを可能とするシステムを実現した。

今後は、実装したシステムの評価、動画や音楽への適用の検討、感性の個人差への対応などが課題である。

## 参考文献

- [1] 平田恭二, 原良憲: "概略画像を用いた画像検索", 信学技報, DE92-2 (1992-05).
- [2] 麻植 周, 前田 哲郎, 福田 慶郎, 上原 邦昭: "動画像データの内容記述とその構造化処理に関する検討", 情報処理学会データベースシステム研究会 DBS-108-4, pp. 25-32 (1996).
- [3] 宝珍輝尚, 新海敏之, 都司達夫: "対象世界のモデリングを用いた図形・画像データの内容検索", 情処学データベースシステム研報 104-17, 情処研報, Vol. 95, No. 65, pp. 129-136 (1995).
- [4] 栗田 多喜夫, 加藤 俊一, 福田 郁美, 板倉 あゆみ: "印象語による絵画データベースの検索", 情処論, Vol. 33, No. 11, pp.1373-1383 (1992).
- [5] 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探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).
- [6] 八村 広三郎, 英保 茂: "色彩分布と印象語に基づく絵画データの検索", 情処研報 CH-27, Vol. 95, No. 91, pp.37-44 (1995).
- [7] 丸山竜也, 勝木道哲, 柴田義孝: "イメージ語を用いたデザイン画像データベース検索システム", 情処学第45回全国大会 F-9, pp.4-87 - 4-88, (1993).
- [8] Snider, J. G. and Osgood, C. E. : "Semantic Differential Technique - A Sourcebook," Aldine Publishing Company (1969).
- [9] 金子隆芳: "色彩の心理学", 岩波書店, 岩波新書 134 (1990).
- [10] 井口征士 他: "感性情報処理", オーム社(1994).
- [11] 武者利光: "ゆらぎの世界", 講談社 (1980).
- [12] 近藤 央規, 宝珍 輝尚, 都司 達夫: "感性の主因子を用いたマルチメディアデータ検索システム", 電子情報通信学会 データ工学研究会 DE96-3, 信学技報 Vol. 96, No. 54, pp. 13-18 (1996-05).
- [13] 宝珍 輝尚, 山口 伸也, 都司 達夫: "感性の主因子を用いたマルチメディアデータ検索システムの改良", 電子情報通信学会 データ工学研究会 DE97-5, 信学技報 Vol. 97, No. 36, pp. 25-30 (1997-05).
- [14] 太田昭雄, 河原英介: "色彩と配色", グラフィック社 (1976).