

絵画からの画面構成の抽出と検索への応用

戸 嶋 朗† 八村 広三郎†

人が画像などの視覚情報を検索するときには、印象などの感性的情報に頼っている割合が大きいと考えられる。そこで、利用者が想起した印象を言葉（印象語）で表現し、システムがこれに適合する画像を選び出すという感性検索の方法が広く検討されている。すでに色彩情報を用いた感性検索の試みが行われ、比較的良好な結果が得られている。ここでは、その成果をふまえて、さらに構図に着目した感性検索を試みる。すなわち、画像から画面構成を表す情報を抽出し、画面構成に基づいた印象語検索を行う方法を報告する。画像処理によって画面構成を抽出する際には、背景と画面に描かれている対象物（オブジェクト）とを正確に分離する必要がある。この目的のために、安定した性能を発揮する領域分割法を開発して利用した。この結果、背景とオブジェクトの分離がうまく行われ、印象語検索に利用できる画面構成情報の抽出が可能であることが分かった。対象画像の画面構成の情報は、オブジェクト領域から得られる何種類かの特徴量の組合せにより、形式的に表現する。画面構成と関連すると思われる印象語（「賑やかな」、「安定感のある」など合計7語）について、このような、オブジェクト領域から得られる特徴量についての条件と、ここから求める適合度の評価式をあらかじめ設定しておく。検索時には、対象画像の特徴量から与えられた印象語に対応する適合度を求める。このような処理を約150枚の静物画の画像データすべてに対して行い、適合度の高いものを抽出、表示するという方法で、印象語による画像の検索を実現した。

Extraction of Compositional Features from Paintings and Application for Retrieval

AKIRA TOJIMA† and KOZABURO HACHIMURA†

When retrieving visual image data, we use impression-related information. Image retrieval by specifying impression words has been studied. Our previous work realized it by the use of color information, and it gave us fairly good results. As the next step, we have been working with a method of extracting compositional features from images and retrieval based on them. For the extraction of compositional features, more robust and reliable methods of region segmentation based on color are needed. We made some modification to the conventional methods. As a result, we could separate background regions and object regions and this made the extraction of information about composition possible. Compositional information consists of some of spatial image features. For each impression words, for example 'busy', 'stable', etc., ranges of image features have been predetermined. When retrieving image, image features extracted from object pictures are compared to predetermined features for impression words, and degree of conformity for these words is calculated. The pictures with highest conformity value are displayed as retrieval results. Experiment has been carried out on 150 still-life pictures, and it has proved the method is promising.

1. はじめに

多くの画像の中から必要とする画像を検索する機能は、画像データベースを実現するうえで、非常に重要である¹⁾。特に、画像を、付随する文字数値情報だけでなく、その内容に基づいて検索するという、いわゆる内容検索の研究^{2)~4)}がさかんに行われている。この画像の内容検索にはさまざまな手法が考えられてい

るが、その1つに、利用者が思いついた印象を言葉で表現し、システムがこれに適合する画像を選び出すという感性検索の方法がある。すでに、色彩情報を用いた感性検索の試みがなされており、画像の種類を限定した環境では、比較的良好な結果が得られている⁵⁾。本研究では、その成果をふまえたうえで、画像の構図に基づく印象語検索を試みる。

構図情報を用いた画像（絵画）の検索についての関連研究に、西山ら⁶⁾の研究がある。これでは、検索要求を領域ベースの略画で表現し、この領域間の相互関係をもとに、対象絵画とのマッチングにより検索する

† 立命館大学理工学部

Faculty of Engineering and Science, Ritsumeikan University

ものである。また、文献7)でも構図をもとにした画像検索を行っているが、ここでは、画像をいくつかのモデル化した構図テンプレートにあてはめて、最も良く表現するテンプレートでの特徴量を検索に用いるという方法をとっている。

本研究では、対象画像中のオブジェクト領域の分布状況をもとに、これと印象語との関連を求め、印象語での画像検索を行うという方法をとる。

いわゆる「構図」には定式化の難しい部分が存在するため、画像の構図についての情報を的確に抽出することは難しい。ここでは、対象物が画面内にどのように配置されているかを表す、「画面構成」⁸⁾を画像処理により抽出し、これを検索に利用することを考える。この画面構成は構図を表現する中心的な要素でもあり、また、構図の第一次近似ともいえる。

ここで検索対象とした画像は静物画を中心とする絵画である。絵画を選んだのは、絵画は一般にその制作過程において画面に作者の意図を反映させるため、観察者に印象を与えるような部分を強調して描かれるという性質があり、通常の画像と比べて印象を与える要素を抽出しやすいという考え方に基づいている。絵画の中でも、静物画を選んだのは、その画面構成が特に印象と強い関係を持つと考えられるためである。

2. 画像の画面構成と印象

2.1 画像から得られる印象

人が画像から受ける印象は、画像中の対象物、色彩、構図、動勢のような要素が関係していると考えられる。絵画では、これに加えてタッチ、マチエール(テクスチャ)なども印象を与える要素となる。現在の画像処理、パターン認識の技術では、絵画から対象物を特定したり、動勢などを抽出したりすることは困難であるが、色彩分布や、画面構成についての情報はある程度数値的な表現ができ、その抽出が可能である。

2.2 画面構成と印象

文献9)によると、絵画の構図とは、「対象が画面に構成されたときに抽出された骨組みと、その骨組みの上に肉付けされた、形態と色彩の関係」であると表現されている。一般の画像においても構図の意味はこれに準ずるものと思われる。このような意味を持つ構図という概念は、定式化が難しい要素を含み、画像処理によってあらゆるパターンを抽出することは困難である。しかし、画面の中でのオブジェクトの配置という観点での画面構成の情報を抽出することは必ずしも不可能ではない。この画面構成を抽出することで、画像から人間が得る印象の一部を抽出することができる

考えられる。

画像を見たときに人が受ける印象のうち、特に構図に関連するものを取りあげ、これらを表す印象語と対応する画面構成(すなわち、オブジェクトの配置についての情報)を考えると、以下のようなものをあげることができる。

「安定感のある」：三角形(ピラミッド形)状にオブジェクトが配置されており、画面の上部から下部にかけて徐々にオブジェクト領域の割合が増していく画面構成。

「不安定な」：「安定感のある」画面構成とは対照的に、画面の上部から下部にかけて徐々にオブジェクト領域の割合が減っていく画面構成。

「整然とした」：水平線上に3個以上のオブジェクトが一定の間隔で配置されている画面構成。

「寂しい」：小さなオブジェクトが画面内に点在する画面構成。オブジェクトの数が少ない、もしくは背景領域が広いほど、より寂しい画面である印象を与える。

「賑やかな」：オブジェクトが画面内に多数点在している画面構成。オブジェクトの数が多く、もしくはオブジェクト領域の面積が画面に対して大きいほど、より賑やかな印象を与える。

「大胆な」：大きな面積を持つ単一のオブジェクトが画面の中央に存在する画面構成。

「ごちゃまじりとした」：複数の小さなオブジェクトが画面の中央に寄り集まって存在する画面構成。

絵画に対する印象語に関しては、SD法を用いた調査の結果を文献5)において報告している。この場合は、構図だけに限定せずに、色彩使用なども含めて、それぞれの絵画の全体的な印象について評価したものである。この調査の結果では、絵画に対する印象の要素となる因子は、「好き」、「嫌い」などの「評価性の因子」、「動的な」、「静的な」などの「活動性の因子」、「強い」、「弱い」などの「力量性の因子」にまとめられることが分かった。この結果を勘案し、今回は、特に静物画の印象に強く関連していると思われる「活動性」と「力量性」の因子に着目し、これに該当すると思われる一般的な印象語を、さらに文献10)、11)などを参考にして選択した。

このような印象語と画面構成の対応関係を前提とすると、各画像から画像処理によりオブジェクト領域を抽出し、これから得られる情報をもとに、絵画の持っている印象を推測することができる。

2.3 画面構成情報の抽出手順

以上で述べたように、本研究では画像中のオブジェクトの配置状況を「画面構成」として考えているので、

このような画面構成情報の抽出処理は次のような手順によって行うことができる。

- (1) 色彩情報に基づく画像の領域分割を行う。
- (2) 領域分割後の画像について、各領域の属性をもとに、各領域を背景の領域とオブジェクトの領域に分離して抽出する。
- (3) オブジェクトを構成する領域の個数、位置、分布状況などのパラメータを抽出し、これらを「画面構成情報」として求める。

3. 色彩に基づく領域分割

本研究では、領域分割の処理に、クラスタリングによる初期領域分割と領域統合とを組み合わせた手法を用いている。画面構成の抽出のためには、背景領域とオブジェクト領域の分離を行う必要がある。このことを念頭におき、ここでは、クラスタリングと領域統合のそれぞれに新たな手法を導入した。

3.1 逐次クラスタ生成法によるクラスタリング

クラスタリングの代表的な手法に K 平均法がある。しかし、この手法では、初期クラスタ中心の選択方法に問題がある。すなわち、たとえば、画像中で一定の画素間隔で初期クラスタ中心を選択する方法では、類似した画素値を持つ初期クラスタ中心が選ばれる可能性があり、このため、クラスタリングの結果、本来は 1 つの領域として抽出されるべき画像の背景領域がいくつかのクラスタに分割されることがある。

初期クラスタ中心を決める方法はいくつか提案されており、これらを利用することも考えられるが、ここでは、本研究での画面構成の抽出に必要な、背景とオブジェクトの分離を念頭においた新たなアルゴリズムを考案して利用した。

単純な K 平均法では、あらかじめ K 個の初期クラスタ中心を決めておき、この初期クラスタ中心を用いてクラスタリングを行うが、クラスタリングを行う前段階で適切な K 個のクラスタ中心を求めることは難しい。

本研究で用いた方法では、クラスタ中心には、その時点で、まだどのクラスタにも分類されていない画素の中から最も出現頻度の高い画素値を選ぶようにする。そして、クラスタ中心を 1 つ決定するごとに、このクラスタ中心との色差が閾値以内の画素をこのクラスタに属するものとする。この方法によって、先に述べた単純な K 平均法と異なり、画素値の類似したクラスタ中心が複数選ばれるということは起こらない。クラスタが 1 つずつ形成されていくことから、この方法を「逐次クラスタ生成法」と呼ぶ。

なお、この逐次クラスタ生成法では、前処理として、256 の代表色への減色処理¹²⁾と原画像を白黒濃淡化した画像からのエッジ点(エッジ画像から閾値処理によって抽出した点をエッジ点と呼ぶ)の抽出を行う。クラスタ中心には、どのクラスタにも属していない画素の中で最も多く含まれる色を選ぶ処理を行うが、少数の代表色をあらかじめ決めておくと、この処理が効果的に行える。また、エッジ点の抽出を行っておくのは、濃淡値が急激に変化する領域からクラスタ中心を選ばないようにするためである。

逐次クラスタ生成法による、クラスタリングのアルゴリズムを以下に示しておく。

- (1) 画像中のエッジ点を除いた画素の中から、最も出現頻度の高い色 1 色を求め、これを初期クラスタ中心とする。
- (2) 現在のクラスタ中心との色差が閾値以下の画素をこのクラスタに属するとする。すでに他のクラスタに属している画素でも、色差がこれまでに求めている他のクラスタ中心との色差より小さい場合には、このクラスタに属するように変更する。
- (3) クラスタに属する画素値の平均を求め、クラスタ中心の値との間に閾値以上の差がある場合には、この平均値を新たなクラスタ中心とし、(2)に戻る。それ以外の場合は(4)に進む。
- (4) まだどのクラスタにも属していない画素のうちから、最も出現頻度の高い色を求め、次のクラスタ中心として、(2)からの処理を繰り返す。クラスタ数があらかじめ設定した値に達したら(5)に進む。
- (5) この段階で、どのクラスタにも属していない画素は、クラスタ中心のうち最も色差の小さいクラスタに属するものとし、クラスタリングを終了する。

3.2 エッジ情報を利用した領域統合

クラスタリングによる初期領域分割の結果に対して領域統合を行う。本研究での領域統合は、微小領域の統合、領域ラベリング、領域統合の一連の処理により成り立っている。この領域統合を経て画像全体の領域分割が完了する。

領域統合の一般的な手法では、統合の判断に用いる特徴量としては、領域間の色差を使うものが多いが、これでは統合が不完全なものになりやすい。たとえば、1 つのオブジェクトであっても、陰の部分が存在する場合は、かなりの色差が生じるためにうまく領域統合が行えない。また、背景に、グラデーションや色ムラ

があると、背景を1つの領域として統合することができない。

このような問題に対処するため、領域統合に際して、色差だけでなく濃淡値のエッジ情報も考慮する。背景に比較的強い色ムラやグラデーションがあっても、濃淡画像上にエッジは現れないことが多い。したがって、互いに隣接する領域間に多少の色差があっても、領域の境界上にエッジ点が存在しない場合は、これらの領域を統合する。

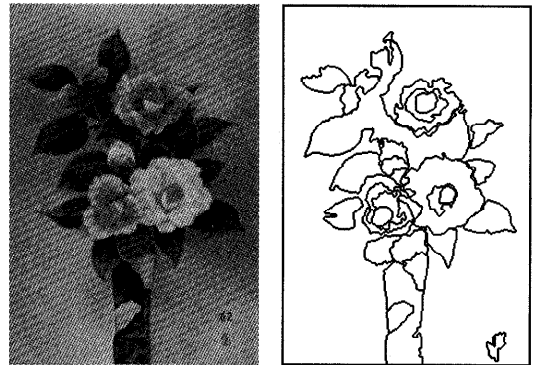
領域統合の際に、統合される領域に属する画素における色の平均値として統合後の領域の色を算出すると問題が生じる場合がある。すなわち、本来は1つの背景領域に属すると考えられる領域がオブジェクトによって複数に分断されている場合、これらの領域間でその色がわずかに異なったものとなり、このため、これらの領域を同じ背景領域として判定することが困難になる。したがって、ここでは、背景の領域候補のクラスタに属している領域の色は、領域の色から求めるのではなく、クラスタリングの結果のクラスタ中心の色を設定するようにしている。

3.3 領域分割の処理例

図1に領域分割の処理例を示す。原画像を同図(a)に、クラスタリングと領域統合の処理を行った結果を同図(b)に示す。(b)では、領域の境界のみを表示している。また、図1(c)は、図1(a)の原画像を白黒濃淡値に変換した画像からエッジ点を抽出したものである。なお、クラスタリングはクラスタ数を8として行った。原画像の図1(a)では、背景部分にグラデーションがあるので、従来の領域統合の手法では背景を1つの領域に統合することは難しいが、本手法では、エッジ点の存在を判定に利用したことにより、背景を1つの領域として統合することができた。

このほか、約150枚の画像について処理を行った結果、大部分の画像で背景抽出の処理は良好に行えることが分かった。一部の画像については、オブジェクトによって背景領域がいくつかの部分に分割されるため、背景を1つのクラスタにはまとめることができない場合もあったが、これらは次章で述べる背景とオブジェクトの分離抽出処理においては、大きな問題にはならなかった。

以上の領域分割の処理には、350×300画素の画像の場合で2分前後(SPEC int95: 4.8, SPEC fp95: 5.4の処理性能を持つワークステーションでの処理時間)の時間を要した。



(a) 原画像 (b) 領域分割の結果 (領域境界のみ表示)
(a) Original image (b) Result of region segmentation



(c) エッジ画像
(c) Edge image

図1 領域分割の処理例
Fig. 1 Region segmentation.

4. 背景とオブジェクトの抽出

以上のようにして、領域分割が完了した画像から背景領域とオブジェクト領域を分離して抽出する。

4.1 背景の抽出

前述した領域分割の処理では、エッジ点を考慮する統合処理によって、緩やかな色彩/濃淡変化をする背景領域の部分が、できるだけ過分割されないような工夫を行っているため、領域分割処理後の画像において、最も出現頻度の高い色を持つ領域を背景領域として考えることができる。また、領域の周囲のある割合以上が画像の4辺と接している領域も背景領域として見なすことができる。

したがって、以下のような、領域の性質を利用して背景領域の抽出を行うことができる。

- 画像中で大きな面積を占める。
- 画像の4辺に接することが多い。

4.2 オブジェクトの抽出

オブジェクトの領域は、領域分割後の画像から背景



図 2 オブジェクト領域の抽出結果
Fig. 2 Extraction of object regions.

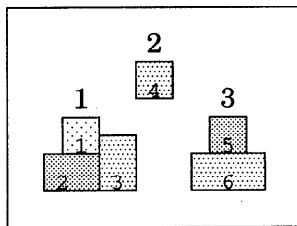


図 3 ミクロオブジェクトとマクロオブジェクト
Fig. 3 Micro objects and macro objects.

領域を除去して残った領域である。図 1 (b) からオブジェクト領域を抽出した結果を図 2 に示す (ここでは背景は白で表している)。領域統合とオブジェクトの抽出が良好に行われていることが分かる。

オブジェクト領域については、以下のように、2 つのレベルで考える。

まず、領域分割と背景領域の分離抽出処理で得られたすべての領域を個々にオブジェクト領域として見なし、これらをミクロオブジェクトと呼ぶ。また、隣接関係で結ばれたミクロオブジェクト領域の集まりを 1 つのオブジェクト領域としてグルーピングし、これをマクロオブジェクトと呼ぶ。図 3 にミクロオブジェクトとマクロオブジェクトの概念図を示す。図 3 の例では、ミクロオブジェクトの数は 6 個、マクロオブジェクトの数は 3 個となる。

ミクロオブジェクトからは、画像全体がどのような色彩領域で構成されているかを詳細に知ることができ、マクロオブジェクトからは、描かれている対象物全体の、画面内における大局的な様子を得ることができる。これらのミクロ/マクロのそれぞれのオブジェクト領域両方のデータを複合的に利用することにより、絵画の印象に関する基本的な情報を得ることができる。

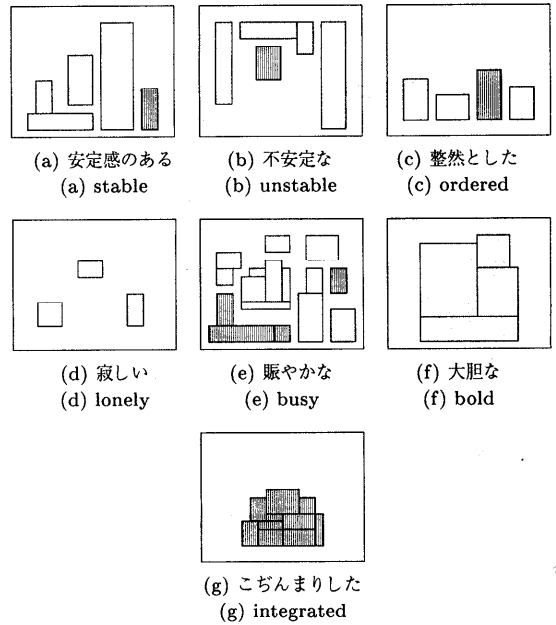


図 4 印象語と画面構成のモデル図
Fig. 4 Models of composition and an impression-related word.

5. 画面構成情報と印象語の関連付け

以上のようにして抽出した、ミクロ/マクロオブジェクト領域のデータから画面構成情報を導き、これと、画像の印象との関連付けを行う。

画面構成情報と印象との関連付けを行うために、ここでは、2 章で述べたような 7 種類の基本的な画面構成に基づく印象語に関して、図 4 に示すような画面構成のモデルを考える (図 4 においては、便宜上、オブジェクト領域を長方形で表現しているが、個々のオブジェクト領域の形状に特に意味はない)。すなわち、2 章でも述べたように、構図に基づいた絵画 (特に静物画) の印象を表現するのに基本的なものと考えられる、「安定感のある」、「不安定な」、「整然とした」、「寂しい」、「賑やかな」、「大胆な」、そして「ごちゃまじりした」の 7 種類の印象語と画面構成のモデルを採用している。

このほかにも、印象語と画面構成のモデルを設定することはできると考えられる。また、画面構成と印象との関連は、もちろんここであげたような単純なモデルだけで完全に表現できるものではないが、これらは画面構成と印象との最も基本的な関連付けとして考えている。

ミクロ、マクロのオブジェクト領域から求める特徴量には次のようなものを考える。

- マクロオブジェクトの個数 (特徴量 LB_N)
- ミクロオブジェクトの個数 (特徴量 SB_N)
- 各マクロオブジェクトの面積 (特徴量 LB_A)
- 各ミクロオブジェクトの面積 (特徴量 SB_A)
- 各マクロオブジェクトの重心の座標 (特徴量 LB_C)
- 各マクロオブジェクトの最上端, 最下端の座標 (特徴量 LB_Tp/Bt)
- マクロオブジェクト領域の集中度 (特徴量 Ig_R)
- 全画素に対するオブジェクト領域の割合 (特徴量 Ob_R)

マクロオブジェクト領域の集中度を表す特徴量 Ig_R は, 面積が最大のマクロオブジェクトの重心を中心とし, オブジェクト領域と面積が等しい円を描いたとき, オブジェクト領域を構成する全画素数に対する円内に存在するオブジェクト領域を構成する画素の数, として定義している.

これらの特徴量はそれぞれ単独ではあまり意味をなさないが, 複数の特徴量を複合的にとらえると画面構成の情報として利用することができる.

オブジェクト領域から得られるこれらの特徴量を用いて, 図4の画面構成のそれぞれのモデルに対して, 前提条件と評価式を設定する. これらの前提条件と評価式は, 与えられた画像の画面構成と印象語との対応関係を判断するための「適合度」と呼ぶ値を計算するのに利用する. これによって対象画像がどの印象語と関連があるかを知ることができ, 印象語を指定しての画像の検索を行うことができる.

以下に, 図4(d)における, 印象語「寂しい」に対応する画面構成のモデルの場合を例にとり, 適合度を算出する方法について述べる.

まず, 「寂しい」という印象に対応する画面構成のための条件 (前提条件) は, マクロオブジェクトの個数 (特徴量 LB_N) とミクロオブジェクトの個数 (特徴量 SB_N) の値が極端に大きくないことである. このことから, ここでは, 特徴量 $LB_N \leq 30$, 特徴量 $SB_N \leq 10$ を, 印象語「寂しい」の前提条件として設定している. この前提条件を満たしている場合に限り, ひきつづき適合度の計算を行う.

「寂しい」という印象は, オブジェクトの数が少ないほど, また背景領域の面積が広いほど強く感じると考えられる. このことから, 適合度の計算に必要な特徴量として次の3種類のものを利用する.

1. 特徴量 LB_N
2. 特徴量 SB_N
3. 特徴量 Ob_R

このように, 適合度の計算に用いる特徴量を複数考

える場合, それぞれの特徴量に重み (比率) をかけて加重和をとった形で適合度を計算する. 「寂しい」の場合の適合度 C は次のようにしている.

$$C = \frac{10 - LB_N}{10} \times w_1 + \frac{30 - SB_N}{30} \times w_2 + (1 - Ob_R) \times w_3$$

すなわち, 適合度の計算に必要な各特徴量を正規化したうえで, 適合度 C が $0 \leq C \leq 1$ の範囲に収まるように, 重み w_1, w_2, w_3 ($w_1 + w_2 + w_3 = 1.0$) で重み付けした加重和をとっている. なお, 前述の前提条件を満たさない場合には, 適合度は 0.0 とする. 「寂しい」の場合の重みのデフォルト値としては, $w_1=0.1, w_2=0.2, w_3=0.7$ と設定している. これは, オブジェクトの個数よりも, オブジェクトの面積の方が「寂しい」という印象に寄与する度合いが強いと考えられるからである. しかし, これらの重みの値については, 個人個人で適切な値が異なると考えられるので, この重み付けは後述するカスタマイズ機能で変更することができる.

他の印象語についても, 同様にして, 画像から得られるオブジェクト領域の特徴量によって前提条件と適合度の式を設定している.

6. 印象語を用いた静物画の検索

5章で述べたような手法により, 画面構成のモデルを仲介として, 与えられた画像と印象語との対応関係を導くことができるようになった. このことを利用して, 印象語を指定して, 画像データベース中からこれに合致するような画像を検索することができる. 以下では, 印象語による静物画の検索について述べる.

6.1 検索用データ

対象画像には, 絵画の中でも, 特に画面構成との関連性が強いと考えられる静物画を中心に用いた. 実験で用いた画像は約 150 枚で, その大部分 (約 9 割) は 16~18 世紀のヨーロッパ静物画 (油彩) であるが, 画面構成のバラエティをふやす目的で, 日本画など, その他の種類の絵画も含んでいる.

本研究では, さまざまなサイズ (画素数) の画像を対象としているが, ここで必要としている画面構成情報の抽出処理のためには, もとの画像を直接処理対象とする必要はない. したがって, 原画像とは別に, それぞれの画像を全画素数が 10 万画素程度になるように正規化したものを用意し, これを処理の対象としている. これらは, 画像データベースにおけるインデックス (索引) 画像としての働きをしていると考えられることもできる.

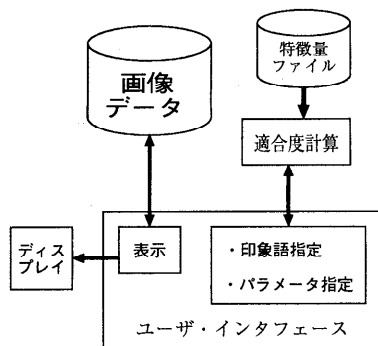


図5 画像検索システム

Fig. 5 Image retrieval system.

6.2 検索システム

画像検索システムの概略を図5に示す。

本システムでは、あらかじめ、すべての対象画像に対して領域分割およびオブジェクトの抽出を行っておき、5章で述べた特徴量をファイルに出力しておく。検索時にユーザの要求により適合度の計算方式を適宜変更できるようにするため、印象語との適合度は検索の実行時に求める。

通常は、あらかじめ設定してあるデフォルトの重みを用いて、適合度を計算するが、必要に応じてGUIにより各特徴量に対する重みを変更することもできる。これにより、印象語に関するユーザの個人差を反映させてカスタマイズできるようにしている。

6.3 検索例と評価

図6、図8、図9に画像の検索例を示す。図6は、印象語「整然とした」によって検索した結果で、最も適合度の高いものから3枚を表示している。この場合、特徴量LB_Bt(各マクロオブジェクトの最下端の座標)のみを適合度の計算に用いている。検索結果には、与えた印象語に合致したものが得られていると考えられる。図7は図6の各検索結果の画像に対応したオブジェクト領域の抽出結果を示したものである。

次に、図8は印象語「賑やかな」に対する検索例で、適合度の高い画像から3枚を表示している。この場合、特徴量に対する重みには、特徴量LB_N, 特徴量SB_N, 特徴量Ob_Rに対して、それぞれ、デフォルト値の0.1, 0.2, 0.7, を設定して用いている。さらに、図9も、印象語「賑やかな」で検索したものであるが、この場合は、それぞれの重みを、デフォルト値から、0.3, 0.4, 0.3に変更している。図8と図9から、特徴量LB_N, 特徴量SB_N(マクロオブジェクト、ミクロオブジェクトの個数を表す特徴量)に対する重みを大きくすることにより、オブジェクトの数が

表1 評価実験の結果

Table 1 Result of evaluational experiment.

印象語	A	B	C	D	E	F	G	H	I	J	平均値
「安定感のある」	2	3	4	4	3	4	2	4	4	4	3.4
「不安定な」	2	3	2	3	2	2	1	4	2	3	2.4
「整然とした」	5	5	5	5	5	4	4	4	5	4	4.6
「寂しい」	4	4	2	4	2	3	4	4	4	3	3.4
「賑やかな」	5	4	5	5	5	5	4	5	4	5	4.7
「大胆な」	2	2	2	4	4	3	4	3	5	4	3.3
「ごちゃまじりした」	3	1	3	2	3	1	1	2	2	3	2.1
総合評価	4	3	3	4	3	4	3	4	4	4	3.6

多い画像が検索されるように変化したことが分かる。

さらに、本システムによる印象語検索の簡単な評価を行った。10人の被験者について、実際にシステムを利用して検索を行い、得られた検索結果の画像について、被験者が考えた印象との合致の度合を評価させた。すなわち、たとえば、「安定感のある」という印象語を用いて得られた検索結果の画像について、「まったく合致しない」の「1」から、「よく合致する」の「5」まで5段階で評点を得た。実験の結果は表1に示す。「A」から「J」は被験者を表している。

きわめて簡単な実験ではあるが、この結果から、印象語ごとに評価のばらつきはあるものの、全体的にはほぼ満足できると評価されていることが分かる。しかしながら、このうち、「不安定な」および「ごちゃまじりした」による結果については評価が低く、さらに検討することが必要と感じている。この原因として考えられることの1つには、対象として用いた静物画の集合の中に、これらの印象語にふさわしい絵画が多くは含まれていないということがある。実際、検索結果の画像の、印象語との適合度の値は、どちらの印象語の場合も、他の印象語の場合に比べて低く、検索結果だけを見た場合に被験者の評価が低くなるのは当然とも考えられる。いいかえれば、この場合は、必ずしも検索システムの評価ではなく、対象画像の集合についての評価を行っていると思えることもできる。

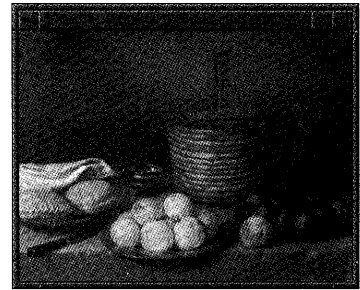
以上の検索例と実験で分かるように、本システムにより、入力した印象語に合致した絵画が検索できることが確かめられた。しかし、領域分割の結果が検索結果に影響を与え、我々の受ける印象と適合度の値との間に安定な対応関係が与えられない場合もある。領域分割の結果を向上させることも必要ではあるが、対象が絵画ということもあり、領域分割処理の確度を今以上に、さらに向上させることはあまり現実的ではない。オブジェクトからの特徴量の算出法を、ややラフな分割処理の結果に対してでも適用できるような、頑強なものに改良する方向で対応すべきであると考えている。



(a) 適合度：0.810
(a) Value of conformity: 0.810



(b) 適合度：0.800
(b) Value of conformity: 0.800

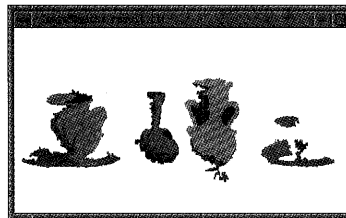


(c) 適合度：0.330
(c) Value of conformity: 0.330

図 6 印象語「整然とした」による検索例
Fig. 6 Retrieval with a word 'ordered'.



(a)



(b)



(c)

図 7 図 6 の各検索結果に対応するオブジェクト抽出結果
Fig. 7 Images of extracted objects corresponding to Fig. 6.



(a) 適合度：0.730
(a) Value of conformity: 0.730



(b) 適合度：0.722
(b) Value of conformity: 0.722



(c) 適合度：0.714
(c) Value of conformity: 0.714

図 8 印象語「賑やかな」による検索例
Fig. 8 Retrieval with a word 'busy'.

7. おわりに

本論文では絵画（静物画）から画面構成を抽出し、これを、印象語と対応付けて検索を行うシステムについて述べた。領域分割により、絵画の画像から背景領域とオブジェクト領域を分離して抽出し、オブジェクト領域から得られる特徴量を用いて印象語と画面構成情報の対応付けを行って、印象語による画像の検索が可能になった。

領域分割の手法については、特に背景領域とオブジェクト領域の分離抽出を念頭において、従来のものに対して改良を加えた。この結果、大部分の対象画像について良好な結果が得られるようになったが、なお一部の画像については若干の問題も残っている。今後さらに、領域分割手法の改良や、オブジェクト領域の特徴量の算出法について、より頑強な方法を検討する必要があると考えている。

特徴量に対する重みを直接変更して行う検索のカス

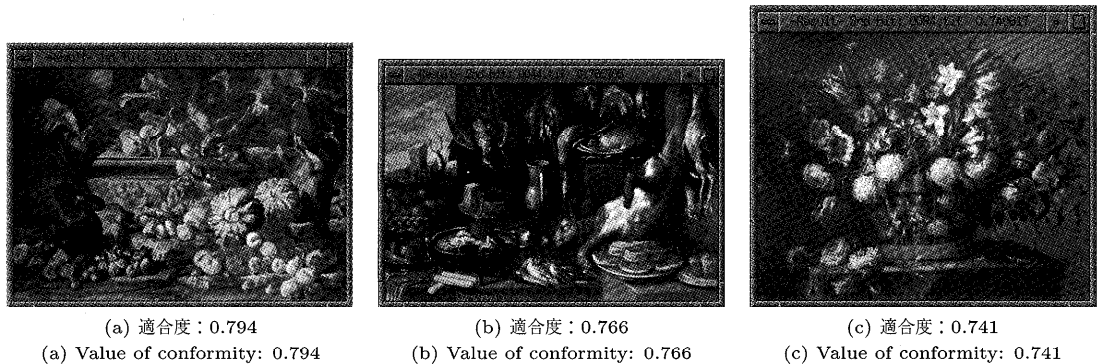


図9 特徴量の重みを変更した印象語「賑やかな」による検索例
 Fig. 9 Retrieval with a word 'busy', changing the weight value for image features.

タマイズ機能は、非常に基本的なものであるが、今後、画面構成のモデルに対する個人個人の印象を事前の予備実験により求め、この結果をシステムに反映するような仕組みを導入することによって、より自然にカスタマイズができるようにすることを考えている。

謝辞 本研究の一部は文部省科学研究費補助金(課題番号 07207116)の援助によって行われた。記して謝意を表す。

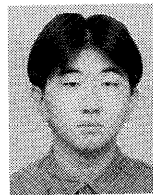
参 考 文 献

- 1) 岡崎彰夫: 画像データベース概論, 情報処理, Vol.33, No.5, pp.448-455 (1992).
- 2) Guivada, V.N. and Raghavan, V.V.: Content-Based Image Retrieval Systems, *IEEE Computer*, Vol.28, No.9, pp.18-22 (1995).
- 3) Flickner, M., et al.: Query by Image and Video Content: The QBIC System, *IEEE Computer*, Vol.28, No.9, pp.23-32 (1995).
- 4) 加藤俊一, 栗田多喜夫: 画像の内容検索—電子美術館への応用, 情報処理, Vol.33, No.5, pp.466-477 (1992).
- 5) 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, 情報処理学会研究報告, 人文科学とコンピュータ, Vol.27, No.6, pp.37-44 (1995).
- 6) 西山晴彦, 松下 温: 画像の構図を用いた絵画検索システム, 情報処理学会論文誌, Vol.37, No.1, pp.101-109 (1996).
- 7) 小林秀行, 大河内頼行, 太田俊二: 特徴量を統合し人の感性に近付けた画像検索システム, 電子情報通信学会技術報告, PRMU97-261, pp.75-80 (1998).
- 8) 下平幸宏, 八村広三郎: 絵画からの画面構成の抽出, 電気関係学会関西支部連合大会, G366 (1996).

- 9) 横山了平: 絵画の構図, 理工学社 (1993).
- 10) 小林重順: 造形構成の心理, ダヴィット社 (1978).
- 11) 鄭 絳宇, 阿部憲広: 画像の伝える感性情報の抽出と利用, 文部省科学研究費補助金重点領域研究研究成果報告書「感性情報処理の情報学・心理学的研究」, pp.165-170 (1995).
- 12) 渡辺孝志: カラー画像を 256 色で近似表示するための高速なアルゴリズム, 電子情報通信学会論文誌, Vol.J70-D, No.4, pp.720-726 (1987).

(平成 10 年 8 月 31 日受付)

(平成 11 年 1 月 8 日採録)



戸嶋 朗

1976 年生。1998 年立命館大学理工学部情報学科卒業。現在、同大学大学院理工学研究科博士課程前期課程情報システム学専攻に在学中。画像検索システムの研究に従事。



八村広三郎 (正会員)

1948 年生。1971 年京都大学工学部電気工学科第二学科卒業。1976 年同大学大学院工学研究科博士課程修了。工学博士。国立民族学博物館第五研究部助手、京都大学情報処理教育センター助教授、同大学工学部助教授を経て、1994 年より立命館大学理工学部情報学科教授。画像処理、コンピュータグラフィックスの教育・研究に従事。画像検索、感性情報処理、マルチメディアシステム等に興味を持つ。電子情報通信学会、画像電子学会、日本 ME 学会各会員。