

# Wavelet-SOM に基づいた類似画像検索システムの設計・実装と性能評価

呉 君 錫<sup>†</sup> 金子 邦彦<sup>††</sup>  
 牧之内 顕文<sup>††</sup> Sang-Hyun Bae<sup>†††</sup>

本論文では、画像データの特徴ベクトルを類似性により分類し、画像データベースの類似検索に用いる手法について述べる。各画像データに対して、ウェーブレット変換の多重解像度解析を用いて、色とテクスチャに対する特徴ベクトルを取り出す。このように得られた画像データの特徴ベクトルから、自己組織化特徴マップを用いて、画像の特徴ベクトルの類似性に基づいた位相特徴マップを生成する。画像の類似検索を行うときには、質問画像の特徴ベクトルの最近傍ノードのみを用いて検索を行うので、検索時間が短い。つまり、検索時には画像の特徴ベクトルを用いない。また、ノードベクトルが画像の特徴ベクトルとの位相関係を保つ性質を利用して部分画像による類似検索を試みる。本論文では、位相特徴マップをデータベースに格納して、例示画像による類似検索法と部分画像による類似検索法について適合率、再現率、一致度を用いて検索結果の評価を行う。

## Design, Implementation and Performance Evaluation of Wavelet-SOM Based Similar Image Retrieval System

KUN-SEOK OH,<sup>†</sup> KUNIHICO KANEKO,<sup>††</sup> AKIFUMI MAKINOUCHI<sup>††</sup>  
 and SANG-HYUN BAE<sup>†††</sup>

The paper describes a new method to extract and cluster image features for effective still image database. The features concerning color and texture are extracted using the multiresolution analysis. Contrast to traditional image databases where feature vectors extracted from stored images are stored and are used to match the feature vector of the input image, we use the Self-Organizing Maps neural network for clustering stored images and generate topological feature maps with node vectors represented similarity between feature vectors. No feature vectors is stored in the databases, since similar retrieval is performed between node vectors and feature vectors. And the node vectors are characterized by preserving topological relationship between the feature vectors. A prototype image database is developed and we experiments on the method of retrieval by example and subspace for image data. The paper reports on the architecture and experimental results.

### 1. はじめに

マルチメディアコンピューティング時代の到来により、デジタル画像に対する需要はますます増加している。一方、ハードウェア性能の向上、記憶媒体の大容量

量化、WWWの普及とDTP(DeskTop Publishing)の急激な増加とともに、大規模な画像データベースの構築が可能となり、その利用分野も広がる一方である。また、インターネットや多様なメディアを介して、様々な画像データベースにアクセスし、検索や、ブラウジングを行うことも可能となった。以上の背景から、色・テクスチャ・形状特徴のような、画像固有の情報に基づいた実用的で効果的な検索手法の研究がその重要性を増している。

既存の画像データベースシステムは、画像と関連付けたキーワードやテキストに基づいた画像検索をサポートしている。キーワードは、高レベルでの画像内容を記述する手法<sup>27)</sup>であり、一種の記号情報であるから、コンピュータによる情報検索に適した表現である。

<sup>†</sup> 九州大学大学院システム情報科学研究科

Department of Intelligent Systems, Graduate School of Information Science and Electrical Engineering, Kyushu University

<sup>††</sup> 九州大学大学院システム情報科学研究科

Department of Intelligent Systems, Graduate School of Information Science and Electrical Engineering, Kyushu University

<sup>†††</sup> 韓国朝鮮大学校自然科学大学数学電算統計学部

Department of Computer Science & Statistics, College of Natural Science, Chosun University, Korea

しかも人間にとっても直感的に理解しやすい有効な表現でもある。しかし、このキーワードによる画像検索は次のような点で容易ではない<sup>(21), (27)</sup>。まず、複雑なテクスチャパターンのような画像内容を、キーワードという人間の言葉で記述するのはしばしば難しい。第2に、各々の画像データのキーワード付けを手作業で行うには時間と労力が必要となる。第3に、同じ画像データに対しても、各々のユーザは独自の観点で画像を扱うことが一般的であるから、ある画像のキーワードを完全に記述しつくすことはできない。最後に、画像の特徴や中身を表すのにふさわしいキーワードが記述されていたとしても、ユーザが同じ意味を持つ別のキーワードを使って検索を試みることがありうる。キーワードによる表現に対して、色・テクスチャ・形状のような、低レベルの画像固有の特徴量で画像内容を表現する手法<sup>(27)</sup>が研究されてきた。特徴量は、画像の識別、すなわち、与えられた画像がどの類に属するかを決定するのに役立つ。このような低レベルの特徴量は計算機により自動的に取り出すことができるから、画像内容の記述の労力はキーワードよりはるかに少ない。

一般に、画像の特徴量をもとに画像検索を行う技術は、内容に基づく画像検索 (Content-based Image Retrieval: CBIR) と呼ばれ、多くの研究成果が報告されている<sup>(1), (8), (11), (12)</sup>。これは、互いに意味内容的に似た画像データからは、互いに近い特徴量が得られるという考え方によっている。このアプローチの有効性・実用性はすでに多くのシステムで実証されている<sup>(22)~(24)</sup>。

内容に基づく画像検索手法では、ユーザが検索したい画像の内容を表現する検索キーを指定して、大量の画像の中から所望の類似画像を検索する。検索結果は検索キーの類似画像である。類似検索は、検索キーとして指定された画像の特徴ベクトルと、データベース内の画像の特徴ベクトルのマッチングにより行われる。特徴ベクトルとは、特徴量の集まりを1つのベクトルとして表現したものであり、一般に高次元データである。したがって、大規模な画像データベースになると、特徴ベクトルだけで大量の記憶領域を占めることになる。また、与えられた特徴ベクトルに類似する特徴ベクトルを探し出す時間も課題である。このような理由から、画像の特徴ベクトルの集合から類似性を見出し、分類することで、類似検索に必要な記憶領域の削減と、検索処理の高速化が可能であると考えた。当然ながら、この分類は自動で行わなければならない。

本研究では、ウェーブレット変換<sup>(13), (14)</sup>で得られた

画像の特徴ベクトルを、その類似性により自動分類するために、Kohonenの自己組織化特徴マップ(以下、SOMという)<sup>(2), (6)</sup>を用いる。SOMの学習は複数の特徴ベクトルから成るデータ集合を2次元上にマッピングし、特徴ベクトルとの位相関係を保った2次元の特徴マップ(位相特徴マップ)を生成する。この位相特徴マップの各ノード(マップ層のユニット)に生成されるベクトルのことを通常、コードブックベクトル<sup>(2)</sup>というが、本論文では、各ノードが持つベクトルという意味でノードベクトルと呼ぶことにする。ここで、位相関係とは、特徴ベクトルの各要素が、画像の各画素の位置における性質を表現しているのと同様に、SOMのノードベクトルもその性質を保っていることを意味する。以上のように、SOMを用いて、特徴ベクトル間の類似性を見出し、互いに類似した特徴ベクトルどうしをマップの近傍ノードにクラスタリングすることが可能である<sup>(3)</sup>。

本研究では、SOMの学習により得られた位相特徴マップをデータベースに格納し、例示画像による類似検索と部分画像による類似検索を行う手法を提案する。例示画像による類似検索では、例示画像の特徴ベクトルに最も近いノード(最近傍ノード)のみから類似画像を取り出すので、データベース内の個々の画像の特徴ベクトルを用いない。また、ノードベクトルが画像の特徴ベクトルとの位相関係を保つ性質を用いることで、画像の部分空間をキーとした部分画像による類似検索を試みる。

本論文では、自然画像をオブジェクト指向データベース管理システムであるJasmine<sup>(28)</sup>を用いて自然画像データベースプロトタイプの実装を行い、

- (1) 特徴ベクトルの類似性による自動分類
- (2) 例示画像による類似検索
- (3) 特徴量間のハイブリッド検索
- (4) 部分画像による類似検索

について述べ、検索結果の適合率、再現率、一致度の評価を行う。

本論文では、2章で関連研究について述べる。3章では特徴量抽出と類似画像の分類について述べ、4章では本プロトタイプシステムの実装について論じる。5章では実験とそれに対する考察を行い、6章でまとめる。

## 2. 関連研究

内容に基づく画像検索では、画像内容を表現する特徴量を用いる。特徴量の抽出には主に2つのアプローチがある。1つは、色ヒストグラム・テクスチャ・

形状情報といった画像の特徴量を抽出するのに、各々固有の画像処理技術を用いる。このようなシステムの例として、QBIC<sup>(9),10),22)</sup>、Photobook<sup>(24)</sup>、VisualSEEK<sup>(23)</sup>、ExSight<sup>(18)</sup>があげられる。この類のシステムは、いろいろな画像処理技術を駆使して個々の画像内容ごとに別々の特徴量抽出アルゴリズムを用いるので、検索システムが複雑になる傾向がある。これらの研究で主に用いられる特徴量抽出アルゴリズムとしては、色ヒストグラム、色レイアウト、テクスチャに対する統計的特徴量抽出、面積・円弧度・周囲長・モーメント特徴などによる形状特徴量の抽出などがあげられる。

もう1つのアプローチは、画像のウェーブレット係数で特徴量を表現する手法である。この手法を用いた多くの類似画像検索システムが実験用プロトタイプとして提案されている<sup>(20),21),25)~27)</sup>。ウェーブレット変換を用いることにより、画像の色・テクスチャ・形状・位置情報を1つの統一されたフレームワークで取り扱うことができるだけでなく、画像の並進、伸縮、回転に対して不変な画像検索にも有効である<sup>(25),26)</sup>。

一方、検索に有効な類似画像分類手法としてSOMが応用されている<sup>(4),5),19),30),31)</sup>。SOMの特徴として、マップ層内での相互結合がないのでハードウェアの負担が小さいこと、学習アルゴリズムが単純であること、外界からの入力に対して教師信号を必要としないで自己組織的に学習を行うこと、多次元の特徴ベクトルから類似性を自動的に見出し、近い特徴ベクトルを位相特徴マップ上の近くに配置し、特に類似性の高い特徴ベクトルは同一のノードに配置することがあげられる。SOMを用いた従来の研究では、類似画像を分類するために、離散コサイン変換による特徴ベクトル<sup>(4)</sup>、色平均やテクスチャ<sup>(30)</sup>、形状情報<sup>(5),19),30)</sup>、キーワード特徴ベクトル<sup>(4),19)</sup>を用いる。我々は従来から、ウェーブレット変換とSOMを組み合わせ、類似検索に有効な特徴量抽出と分類手法を提案してきた<sup>(31)</sup>。

本研究は関連研究と比較して、次のような特徴があげられる。(i)ウェーブレット変換を用いて特徴量を抽出する際、類似検索に有効な分解レベルを定める。ウェーブレット変換を用いた類似検索は多数報告されているが<sup>(20),21),25)~27)</sup>、類似検索に有効なウェーブレット変換の分解レベルを定めた研究はない。本研究では、分解レベルの変化が類似検索に与える影響を調べ、類似検索に有効な分解レベルを決定する手がかりとする。(ii)複数の特徴量を用いたハイブリッド検索と、特徴量ベクトルとノードベクトルとの位相関係に基づいた部分画像による類似検索を提案する。複数の

特徴量をサポートする類似検索はすでに報告されているが<sup>(9),18),23)</sup>、SOMを用いたハイブリッド検索や、特徴量ベクトルとノードベクトルとの位相関係に基づいた部分画像による類似検索は、我々が知る限りまだその研究例がない。(iii)少ないノード数でSOMのマップ層を構築することにより、学習時間の短縮をねらう。SOMを類似検索に用いる場合、学習時間と検索時間が問題となる。本研究では、SOMの学習により得られた位相特徴マップの各ノードのうち、画像が1つも分類されていない空きノードに着目し、その空きノードを少なくするような位相特徴マップを構築することにより、学習時間と検索時間の短縮を図る。

### 3. 画像の特徴量抽出と類似画像の分類

#### 3.1 システムの概要

図1に、今回実装した類似画像検索システムの構成図を示す。本システムは大きく分けて、特徴量抽出/分類プロセスと検索プロセスからなる。まず、特徴量抽出/分類プロセスでは、入力画像データに対してウェーブレット変換を行い、その結果から、色とテクスチャの特徴量を抽出し、特徴ベクトルを生成する。その特徴ベクトルを入力データとして、SOMの学習により、各特徴ベクトルの類似性を見つけ出し、位相特徴マップを生成する。位相特徴マップの各ノードはノードベクトルを持つ。データベース内には位相特徴マップそのものと、各ノードベクトルに最も近い特徴ベクトルを持つ画像IDのリスト(最整合画像リスト)が格納される。

検索プロセスでは、与えられた質問画像(例示画像または部分画像)から色とテクスチャの特徴ベクトルを求め、ノードベクトルとの距離が最小となる $k$  ( $k \geq 1$ )

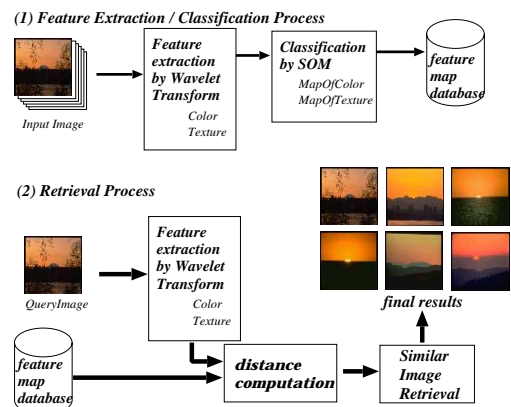


図1 類似画像検索システムの構成

Fig. 1 Overview of similar image retrieval system.

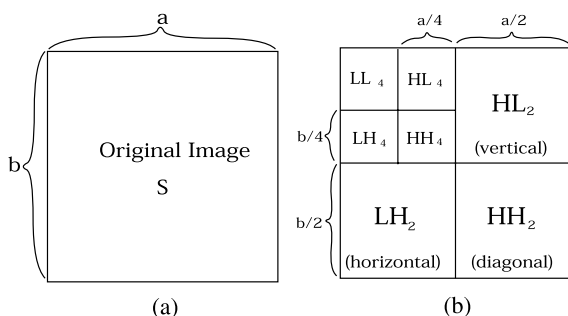


図2 多重解像度解析による分解画像の階層構造表現

Fig. 2 Hierarchical representation of subimage by multiresolution analysis.

個の最近傍ノードから類似画像を検索する。

### 3.2 ウェーブレット変換による特徴量抽出

(1) ウェーブレット変換 画像データをウェーブレット変換することによって画像データを階層化された分解画像に分割する。この画像データの階層化の理論は多重解像度解析として Mallat によって考案された<sup>13)</sup>。画像をウェーブレット変換して得られるウェーブレット分解係数空間から、色・テクスチャ・形状情報などの様々な画像内容に対応した特徴量をシステム側で効率良く取り出すことができる<sup>21),26)</sup>。

画像のウェーブレット変換は、次のような手順で行われる。まず、水平方向に対して、フィルタ処理を行い2つの帯域に分解し、2:1 ダウンサンプリングを行う。同じ処理を垂直方向に対して行い、4つの帯域成分が得られる。4つの帯域成分は、原画像の縮小画像を表する1つの低域成分 ( $LL_2$ ) と、差分画像を表す3つの高域成分から構成される。高域成分は、原画像の水平方向 ( $LH_2$ )、垂直方向 ( $HL_2$ )、斜め方向 ( $HH_2$ ) であり、画像の詳細情報を表現する。以上の処理を低域成分  $LL_i$  ( $i = 2, 4, 8, \dots$ ) に対して再帰的に行う。図2は原画像  $S(1 : a, 1 : b)$  に対して、ウェーブレット変換による2レベルの多重解像度解析を行ったときの、分解画像の階層関係を示している。

本論文では、各画像に対する色とテクスチャの特徴量を得るために、Haar ウェーブレット変換を用いる<sup>15)</sup>。Haar ウェーブレット変換は、画像処理に適用するとブロック歪みなどが生じる欠点があるが<sup>21)</sup>、画像の類似検索の結果にはそれほど影響を与えない<sup>25),26)</sup>。反面、Haar ウェーブレット変換は高速な画像の分解アルゴリズムであり、実装が簡単である利点がある。

(2) ウェーブレット分解レベルと類似検索との関係 画像のウェーブレット変換を行う際、有効かつ妥当な分解レベルは決まっていない<sup>35)</sup>。したがって、個々のアプリケーションごとにユーザが分解レベルを指定

するのが普通である。

この節では、類似検索における興味深い問題として、分解レベルの変化が類似検索に与える影響を調べ、類似検索に有効な分解レベルを決定する手がかりとする。ここでは、色特徴量についてのみ考える。

$128 \times 128$  画素を持つ  $S(1 : 128, 1 : 128)$  の自然画像を、分解レベルを4から7まで変化させ、ウェーブレット変換を行う。各々の分解レベルごとに得られる低域ウェーブレット係数から色特徴量を得て、これら4つを比較する。画像のサイズが  $128 \times 128 (= 2^7 \times 2^7)$  であるから、分解レベル7までウェーブレット変換が可能である。画像の色空間は  $YIQ$  表色系 (NTSC方式  $\bar{Y}$ ) を使う。特徴ベクトルは、レベル4のときは192次元、レベル5のときは48次元、レベル6のときは12次元、レベル7のときは3次元のベクトルである。特徴ベクトルの生成に関する詳細は、後述の(3)を参照されたい。

分解レベルと類似検索の関係を評価するために、多数の自然画像を用意して、すべての画像の色特徴ベクトルと、質問画像として選んだ画像の色特徴ベクトルとのマッチングを行い、最も近い10個の画像を検索結果  $R_i$  ( $i = 4, 5, 6, 7$ ) として得て、これらの比較を行う。マッチングはユークリッド距離で行う。比較のための尺度として一致度を用いる。ここでの一致度とは、検索結果として得られた類似画像  $R_i$  が、他のレベルの検索結果  $R_j$  ( $j = 4, 5, 6, 7$ ) にどれだけ含まれるかを示す値で、次のように定義する。

$$\text{一致度} = \frac{R_i \cap R_j}{R_i}$$

ここで、 $i = j$  の場合、一致度は1となる。評価実験では、200枚の自然画像を用いる。その詳細は後述の5章に説明している。質問画像は全画像の中からランダムに抽出した20枚である。

表1に、分解レベルの変化にともなう一致度の変化を示している。実験結果により、以下のことが明らかになった。

- 分解レベルが高くなるほど一致度は低くなる
- 分解レベルの影響を受ける度合いは、レベル7, 6, 5, 4の順でレベル7が最も高い
- ウェーブレット分解レベルは、 $128 \times 128$  画素の画像データの場合、分解レベル4, 5が妥当である

(3) 特徴ベクトルの生成 Haar ウェーブレット変換を用いて、原画像  $S(1 : 128, 1 : 128)$  から、色とテクスチャの特徴ベクトルを生成する手順を示す。

(Step1) 色空間の変換: 画像の特徴量抽出では、 $YIQ$  表色系を用いた。一般的に、カラー画像の

表1 分解レベルの変化にともなう一致度の変化

Table 1 Variation of correspondence degree with increment of decomposition level.

	$i = 4$	$i = 5$	$i = 6$	$i = 7$
$j = 4$	1			
$j = 5$	0.975	1		
$j = 6$	0.685	0.73	1	
$j = 7$	0.495	0.51	0.61	1

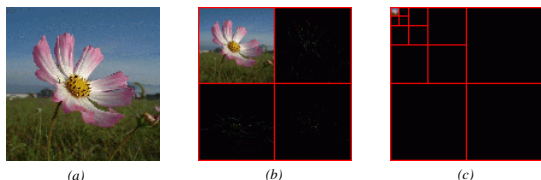


図3 画像のウェーブレット分解例。(a)原画像,(b)1レベル分解,(c)4レベル分解

Fig. 3 The example of image decomposition using wavelet transform. (a) Original image, (b) 1-level decomposition, (c) 4-level decomposition.

デジタル表現には  $RGB$  色空間が用いられる。しかし、 $RGB$  色空間は、 $R$ 、 $G$ 、 $B$  各属性の明るさのみで画像を表現するため、3 属性間の相関が強く、各属性を独立に扱えない。一方、 $YIQ$  表色系は、輝度情報 ( $Y$ ) と、色彩情報 ( $IQ$ ) に分けて画像情報を表す。結果として、特徴ベクトルは、輝度情報部分と色彩情報部分の組合せになる。類似検索における  $YIQ$  表色系の優位性は Jacobs ら<sup>26)</sup> により実験結果が報告されている。本論文では、各画像データを  $RGB$  色空間から、 $YIQ$  表色系に変換し、特徴量抽出に用いる。

(Step2) 画像のウェーブレット変換:  $YIQ$  表色系に変換された画像  $S_{C_i}(1:128, 1:128)$  に対して、4 レベルの Haar ウェーブレット変換を行う。ここで、 $C_i$  ( $i = 1, 2, 3$ ) は  $YIQ$  表色系の各チャンネルを表す。ウェーブレット変換結果、1つの低域成分と12個の高域成分で表現されるウェーブレット係数行列  $EC_i(1:128, 1:128)$  が得られる。図3は、画像の輝度成分を表す  $Y$  チャンネルに対するウェーブレット変換結果を示している。図3(c)で、左上の  $8 \times 8$  部分行列  $EC_i(1:8, 1:8)$  は最も低域な成分であり、画像全体に対する平均的な色成分を表す。また、この低域成分と隣り合う3つの部分行列 ( $EC_i(1:8, 9:16)$ ,  $EC_i(9:16, 1:8)$ ,  $EC_i(9:16, 9:16)$ ) は画像のより詳細な情報を表す。

(Step3) 特徴ベクトルの生成: (Step2) のウェーブレット変換により得られた部分行列の最も低域な

成分である  $8 \times 8$  行列  $EC_i(1:8, 1:8)$  から、 $192 (= 64 \times 3)$  次元の色特徴ベクトル  $FV_{col}$  を得る。また、テクスチャ特徴ベクトルとしては、高域成分の水平方向の部分行列  $EC_i(1:8, 9:16)$  から  $192$  次元の  $FV_{txtr}$  を得る。なお、今回、特徴量抽出に高域成分の垂直方向と斜め方向を用いていないが、今後それらを考慮した特徴量を検討する予定である。

### 3.3 SOMによる類似画像の分類

類似画像の分類は、ウェーブレット変換から得られた色の特徴ベクトル  $FV_{col}$  とテクスチャの特徴ベクトル  $FV_{txtr}$  を用いて、SOMの学習により行う。

SOMは2層構造を持つニューラルネットワークである。第1層は  $n$  次元の入力層であり、そのノード数  $n$  は特徴ベクトルの次元数から  $n = 192$  である。第2層は、複数のノードが2次元に配列されているマップ層(または競合層)で構成され、学習前にランダムな重み値が与えられ初期化される。

類似画像分類は、以下に示す位相特徴マップ生成過程と最整合画像リスト生成過程からなる。

(1) 位相特徴マップ生成過程 位相特徴マップを生成するために、学習に必要なパラメータ(学習率、近傍半径、マップ層のサイズ、近傍関数、学習回数)を与え、SOMの学習を行う。位相特徴マップの生成過程は次のとおりである。まず、入力として与えられた特徴ベクトルに最も近い重みベクトルを持つノードをマップ層から選ぶ。次に、そのノードと近傍領域内のノードを特徴ベクトルに近づけるように更新する。この処理を指定された学習回数だけ繰り返す。学習の結果、マップ層の各ノードに生成されるベクトルのことをノードベクトルといい、

$$CBV_i = [cv_{i1}, cv_{i2}, \dots, cv_{ij}, \dots, cv_{im}]^T$$

で表す。ここで、 $i$  ( $1 \leq i \leq k$ ) はマップ層のノード番号を示す。 $m$  は入力層のノード数(画像の特徴ベクトルの次元数)を、 $k$  はマップ層のノード数を表す。このように得られた位相特徴マップは、特徴ベクトル相互の類似関係が2次元上にマッピングされたものである。

(2) 最整合画像リスト生成過程 最整合画像リスト生成過程では、位相特徴マップを用いて、特徴ベクトルとノードベクトル間の距離計算を行い、最も近いノードを求め類似画像を分類する。特徴ベクトルとノードベクトル間の類似度は、ユークリッド距離によって計算した。最近傍ノード  $BMN_i$  は、

$$BMN_i = \min_i \{ \|FV - CBV_i\| \}$$

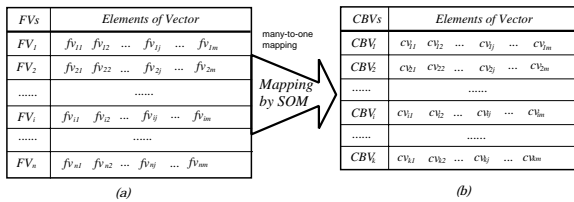


図4 特徴ベクトルとノードベクトルとの関係。(a) ウェーブレット変換により生成された特徴ベクトル,(b) SOMにより生成されたノードベクトル

Fig. 4 The relationship between feature vector and node vector. (a) Feature vector generated by wavelet transform, (b) node vector generated by SOM.

を用いて求める。ここで、 $FV$  は  $FV_{col}$  または  $FV_{textr}$  を表す。特徴ベクトルとノードベクトルとの関係を図4に示す。この2つのベクトルは、その類似性により多対一の対応関係にある。

以上の処理をデータベース内のすべての画像について、色とテクスチャの特徴ベクトルを用いて行うことで、色とテクスチャの位相特徴マップと最整合画像リストが生成される。

#### 4. システムの実装

本章では、類似画像検索システムの実装について述べる。実装は、富士通と Computer Associates 社が共同開発したオブジェクト指向データベース管理システム Jasmine<sup>28)</sup> 上で行った。Jasmine は、オブジェクトデータベースアクセス機能と、インターネットを経由して Jasmine サーバにアクセスする Web ベースのアプリケーション構築機能を持っている。オブジェクトデータベースアクセス機能は ODQL (Object Database Query Language) により、データの定義、操作、手続きの記述を行うための機能である。Web ベースのアプリケーション構築機能には、Web とデータベース間接続サーバ機能を持つ WebLink という機能があり、HTML の拡張タグを用いて HTML ページ内に ODQL 文を記述することで、WWW ブラウザを用いたデータベース内のデータに対する閲覧、更新および検索処理ができる。

類似画像検索システムは画像データベース部と検索用ユーザインタフェース部を実装した。

(1) 画像データベース部 画像データベース部は、Jasmine のオブジェクトデータベースアクセス機能を用いて実装を行った。図5のシステムの構成要素に必要なすべてのクラスとメソッドを、埋め込み ODQL (ODQL 文がホスト言語 C++ に埋め込まれた形式で記述) の機能を使って実装した。画像データベース部

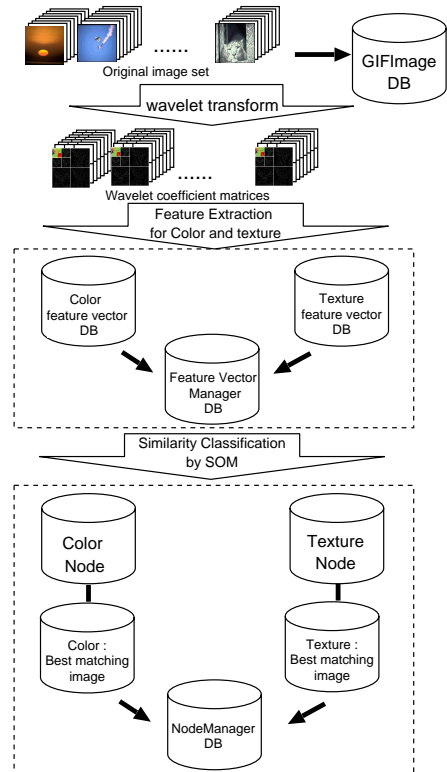


図5 画像データベースの構造

Fig. 5 Image database architecture.

では、3章で述べた特徴量抽出と類似画像分類手法を埋め込み ODQL により実装した。

(2) ユーザインタフェース部 ユーザインタフェース部は、Web ベースのアプリケーション構築機能を用いて実装を行った。WWW ブラウザから画像データベースへアクセスするために、類似検索を行うテンプレートを記述する。ここで、テンプレートとは、データベース内のデータを操作するために、WebLink アプリケーション側が提供する埋め込み WebLink タグを持つ HTML ページである。ユーザインタフェース部では、例示画像による検索機能、部分空間による検索機能の2つの検索機能を持つテンプレートを実装した。

#### 5. 実験と考察

類似検索の検索精度を評価するために、色・テクスチャによる類似検索に対して適合率、再現率、一致度を測定した。画像データは、スウェーデン大学で公開されている画像コレクション (SUNET<sup>9)</sup>) からダウンロードした。SUNET の画像コレクションは種類別にディレクトリ構造になっている。今回、実験に用い

た画像は、SUNET から 10 種類 (動物, 植物, 飛行物など) の画像を人間の目で 200 枚選んだものである。

SOM の入力層のノード数は, 3 章で説明したとおり, 192 である。マップ層のノード数は空きノードが少なくなるように注意深く決める必要がある。その理由は, 検索時に, 質問画像の特徴ベクトルとノードベクトル間にユークリッド距離による類似度を計算するとき, 空きノードが多いと無駄な計算時間がかかるからである。データベース内の各画像はマップ層の各ノードに 1 対 1 対応になるようにマッピングされるのが理想である<sup>30)</sup>。しかし, データベース内では互いに類似した画像が含まれているため, 一般的に 1 ノードには複数の画像がマッピングされる。その結果, マップ層のノードには, 画像が 1 つも分類されていない空きノード (empty node) が生じる。

我々は, マップ層の大きさを全画像のデータ数とほぼ同じ大きさ 256 (16 × 16) から, 200 (20 × 10), 196 (14 × 14), 100 (10 × 10), 25 (5 × 5) のように徐々に小さくしながら, 位相特徴マップを生成してみた。その位相特徴マップを比較し, 次の 3 つの結果が得られた。(i) マップの大きさが小さいときは粗い分類が, 大きいときは細かい分類が行われる。(ii) マップの大きさにかかわらず, お互いに類似性の高い画像どうしは同一のノードに分類される。(iii) マップが大きくなるほど空きノードが多くなる。たとえば, 我々の実験では 200 枚の画像に対して, マップの大きさが 14 × 14 のとき空きノードの数は 99 個で, 全ノードの 50% に達した。

以上の結果に基づいて, 今回の実験ではマップ層のノード数を, 類似画像の分類は多少粗いものの, 空きノードが 0 となったマップの大きさ 5 × 5 = 25 として実験を行うことにした。このように, 少ないノード数でマップを構築することにより, (i) 学習時間が短縮される, (ii) 検索時に質問画像の特徴ベクトルとノードベクトル間の距離の計算が少なくなるという効果が得られた。

データベース内のすべての画像に対して, 色とテクスチャの特徴ベクトルを求め, SOM の学習を行った。マップ層の各ノードが持つ重みベクトルは, ランダム値で初期化された 192 次元のベクトルである。学習回数は  $T = 10,000$  とし, 学習時の近傍集合の初期値  $\sigma_0 = 3$ , 学習率の初期値  $\alpha_0 = 0.5$  とした。近傍集合と学習率は, 学習が行われる間に更新され続ける。SOM の学習による位相特徴マップは, SUN Microsystems 社の Ultra 1 (OS: Solaris2.5, memory: 128 MB) 上で, 色特徴ベクトルに対しては 12.5 秒,

テクスチャ特徴ベクトルに対しては 12.4 秒で生成された。また, 色特徴ベクトルは 371 KB, テクスチャの特徴ベクトルは 367 KB であったが, マップ層のノード数を 5 × 5 と設定したとき, SOM の学習により生成されたノードベクトルのサイズは, 色特徴ベクトルは 42 KB (11.3%), テクスチャの特徴ベクトルは 49 KB (13.4%) で表現された (括弧内は, 元の特徴ベクトルとの比率)。

学習の後, 位相特徴マップとその最整合画像リストをデータベースに格納し, 次の検索手法によって検索実験を行った。ここで, 検索対象は, データベース内の全画像である。

(1) 例示画像による類似検索法 この検索手法は与えられた質問画像 (例示画像) の特徴ベクトルとノードベクトル間のユークリッド距離を計算し,  $k$  個の最近傍ノードから類似画像を順次取り出す手法である。検索には色とテクスチャの特徴量を用いた。図 6 には, 色特徴量による類似検索の結果を示す。検索時間 (質問画像を与え, ブラウザに検索結果を表示するまでの時間) は 3 秒から 5 秒であった。検索結果を評価するために, 2 つの尺度を用いる。まず, 検索精度を定量的に評価するために, 適合率 (Precision ratio) と再現率 (Recall ratio) を測定した。また, SOM の類似画像分類の精度を評価するために一致度を用いた。

- 適合率と再現率による評価 適合率と再現率は次のように定義した。すなわち, 全データベースからの適合画像を  $I_{relevant}$ , 質問画像に対する検索結果を  $I_{retrieved}$  とし, 適合率と再現率はそれぞれ,

$$\text{適合率} = \frac{I_{retrieved} \cap I_{relevant}}{I_{retrieved}}$$

$$\text{再現率} = \frac{I_{retrieved} \cap I_{relevant}}{I_{relevant}}$$

で表す。質問画像はランダムに 20 枚を選んだ。適合画像は, 各質問画像に対して 2 人の被験者の判断により前もって全データベースの中から選んでおく。質問画像に対して, 5 個の最近傍ノードから, 距離が最小となる画像 10 枚を順次取り出し, 検索結果とする。この検索結果の適合率と再現率を図 7 に示す。実験結果では, 色による検索に比べて, テクスチャによる検索の方が精度が落ちる。つまり, テクスチャによる検索の方が, データベース内から被験者が選んだ適合画像と検索結果との差が大きかった。その理由は, 被験者が全画像データから適合画像を割り当てる際, 色よりテクスチャの方が被験者の個人差に影響を受けるからであると思われる。

- 一致度による評価 もう 1 つの検索結果に対する

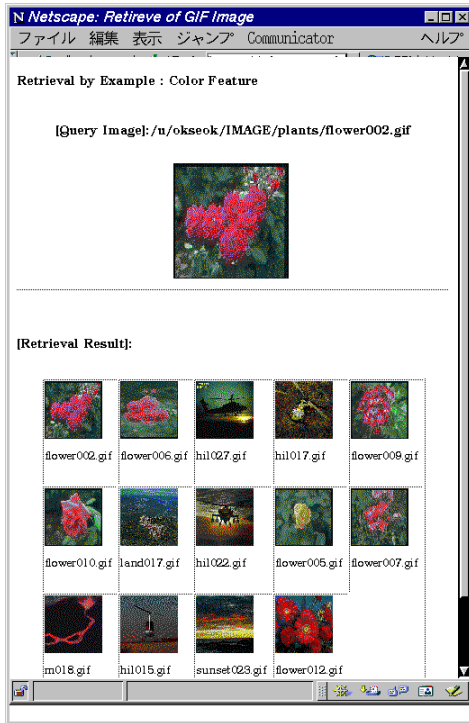


図 6 色特徴による類似画像検索

Fig. 6 Similar image retrieval by color feature.

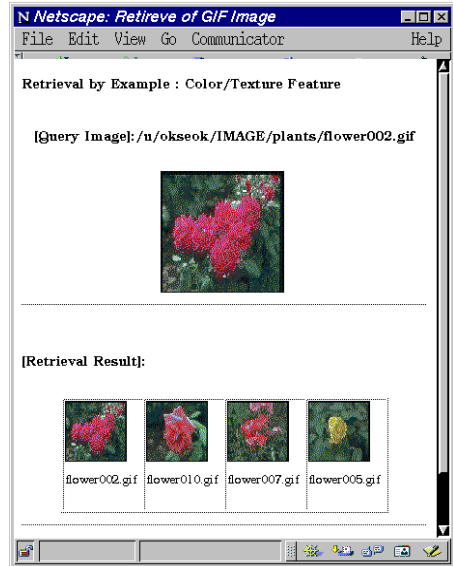


図 8 色とテクスチャによるハイブリッド検索

Fig. 8 Hybrid retrieval by color and texture feature.

分類の精度を表すものである．検索結果から求めた一致度は、色による検索では 89%，テクスチャによる検索では 79% が得られた．この一致度の評価により、SOM による類似画像の分類が有効であることが明らかになった．

(2) 特徴量間のハイブリッド検索 本システムでは、図 5 で示すように、色とテクスチャの特徴ベクトルに対して各々の位相特徴マップが構築されている．これら複数の分類マップを組み合わせたハイブリッド検索機能を実現した．

ハイブリッド検索では、与えられた質問画像に対して、色による検索、テクスチャによる検索を行い、その 2 つの積を求め検索結果とする．図 8 に、ハイブリッド検索による検索結果を示す．

適合率と再現率は図 7 に示す．ハイブリッド検索では、色やテクスチャの検索結果から質問画像に最も近い画像に絞られるので、色やテクスチャのみでの検索精度より良い結果が得られた．一方、一致度は 95% の高い値が得られた．

この適合率・再現率および一致度による検索結果の評価により、ハイブリッド検索の方がより精度の良い検索結果が得られることが明らかになった．

(3) 部分画像による類似検索法 SOM の学習により生成されたノードベクトルは画像の特徴ベクトルと位相関係を保持している．色特徴ベクトルの場合、原画像  $S(1:128, 1:128)$  に対して 4 レベルのウェーブレット変換を行い、最も低域成分である  $8 \times 8$  画素を

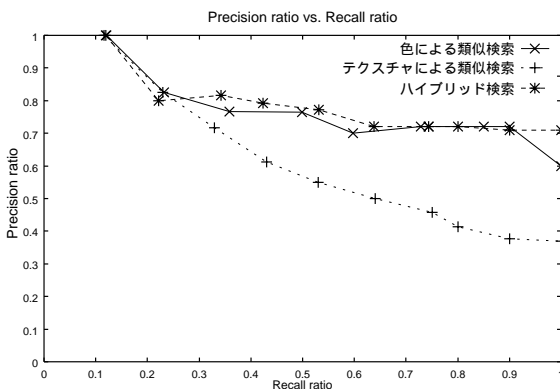


図 7 検索効果

Fig. 7 Retrieval effectiveness.

評価尺度として、3.2 節 (2) と同様に一致度を用いる．ここでは、画像データ全体に対する検索（全件検索）結果を  $R_{fv}$  として、次のように定義する．

$$\text{一致度} = \frac{R_{fv} \cap I_{\text{retrieved}}}{R_{fv}}$$

ここで、 $R_{fv}$  は 3.2 節 (2) と同様に、データベース内の全画像からユークリッド距離による全件検索を行い、質問画像に最も近い画像を上位 10 枚取り出したものである．この一致度は SOM の類似画像



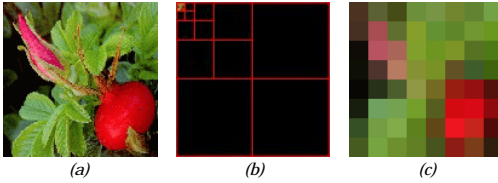


図9 ウェーブレット変換による原画像と低解像度画像との関係。(a) 原画像,(b)分解レベル4のウェーブレット変換,(c)低解像度  $8 \times 8$  画像

Fig. 9 The relationship between original image and low-pass resolution image using wavelet transform. (a) Original image, (b) 4 level wavelet transform, (c) enlarged lower-resolution  $8 \times 8$  subimage.

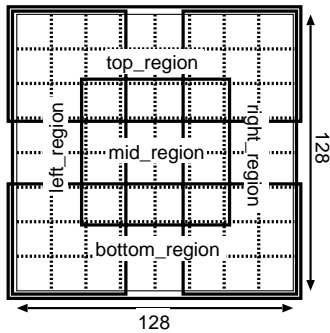


図10 類似検索のための画像の部分空間表現

Fig. 10 Subspace representation for similar image retrieval.

取り出したものである。この特徴ベクトルの各要素は、原画像の各チャンネルの  $16 \times 16$  画素の平均である。このことを図9で示している。色とテクスチャの特徴ベクトルから、SOMの学習により得られたノードベクトルの各要素は、その特徴ベクトルの各要素が持つ色とテクスチャの代表値である。このような位相関係から、ユーザは検索パラメータとして、画像の部分空間を指定することで、画像の特定の部分に対する類似検索が可能となる。

部分空間は、図10のように、画像に写っている内容として、左右上下に広がる背景(または対象)と中央の対象物(または背景)を考え、5つの領域を定義した。図10の中の太い点線の部分は、原画像  $S(1:128, 1:128)$  を簡単に表現するために、 $8 \times 8$  に分割して表したものである。検索時には、問合せのパラメータとして、部分空間の領域と特徴ベクトルを与える。ノードベクトルから部分空間の特徴ベクトルに該当する要素を取り出して類似度を計算し、類似検索を行う。図11に色特徴量を用いて、与えられた質問画像に対する右の部分空間(right\_region)を指定したときの検索結果を示す。

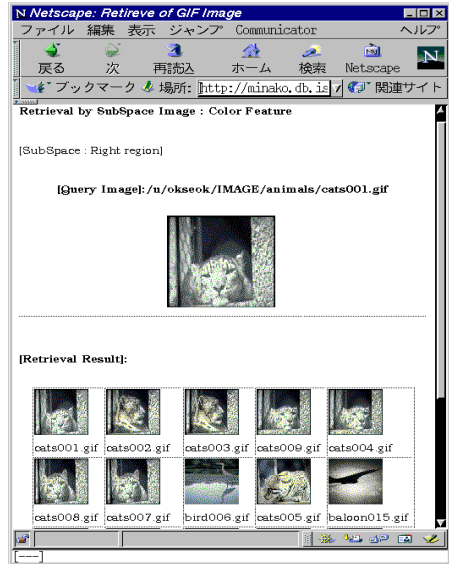


図11 部分画像による類似画像検索(色特徴ベクトルを使用)  
Fig. 11 Similar image retrieval by subimage using color feature.

部分画像による類似検索では与えられた質問画像に対して必ずしも自分が分類されたノードとマッチされない場合が生じる。その理由は、SOMの学習時にノードベクトルは特徴ベクトルに近似するように更新されるため、特定の部分空間に対してマッチするノードと、全空間に対してマッチするノードが違うからである。このように、ノードベクトルを用いた部分画像検索では試みとしては多様な検索法への発展性があるものの、部分空間の取扱いについてオブジェクトの位置情報のような特徴量を考慮した、さらなる議論が必要である。

## 6. まとめ

本研究では、自然画像に対しウェーブレット変換の多重解像度解析を行い、画像の色とテクスチャの特徴ベクトルを抽出した。ウェーブレット変換による画像の特徴ベクトルの抽出において、検索に有効なウェーブレット分解レベルを定めた。類似画像の分類には、ウェーブレット変換から得られた特徴ベクトルを入力データとし、SOMによる学習を行った。そのとき、検索に有効なマップを構築するために、空きノードに着目して、マップの大きさを定めた。その結果、学習時間や、検索時に質問画像の特徴ベクトルとノードベクトル間の距離の計算時間が少なくなった。しかし、マップの大小が類似検索の結果に影響を与える度合いについては、今回の実験に基づいてさらに研究を進めていきたい。

SOMの学習により得られた位相特徴マップをデータベースに格納することにより、例示画像による類似検索と特徴量間のハイブリッド検索が可能であることを実験で示した。そして、画像の特徴ベクトルとノードベクトル間の位相関係に基づいた部分画像による類似検索は、SOMの特徴を生かした類似検索の試みであった。

検索結果の評価尺度として、適合率・再現率に加え、一致度による評価を行った。本研究で定義した一致度がより客観的な評価尺度として使えるようにさらなる検討を行いたい。

今回、画像データの特徴量として色とテクスチャのみで行ったが、それ以外の特徴量(形状・位置関係など)について検討を行い、これらの要素を類似画像分類に考慮することで検索効率の向上を図ることができると思われる。また、テスト画像を増やし、SOMの特徴を十分に生かした類似画像の検索手法についてさらに検討を行う予定である。

謝辞 本研究の一部は、文部省科学研究費補助金(課題番号：10308012)の援助を受けている。

#### 参 考 文 献

- 1) 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).
- 2) Kohonen, T.: *Self-Organizing Maps*, Series in Information Sciences, second edition, Vol.30, Springer-Verlag, Berlin (1997).
- 3) 馬場則夫, 小島史男, 小沢誠一: ニューラルネットの基礎と応用, 共立出版 (1994).
- 4) 波多野賢治, 亀井俊之, 田中克己: 多段階自己組織化マップによるビデオ映像記述支援と類似シーン検索, 情報処理学会論文誌, Vol.39, No.4, pp.933-942 (1998).
- 5) 遠藤勝博, 上野雅浩, 田辺隆也, 山本 学: 自己組織化マップを用いた静止画像のクラスタリング方法, SOM研究会 (2000).
- 6) Kohonen, T., Hynninen, J. and Laaksonen, J.: SOM\_PAK: The Self-Organizing Map Program Package, Technical Report A31, Helsinki University of Technology, Laboratory of Computer and Information Science (1996).
- 7) Russ, J.C.: *The Image Processing Handbook*, CRC Press, Boca Raton (1995).
- 8) Gudivada, V.N. and Raghavan, V.V. (Eds.): Content-based Image Retrieval System, *IEEE Computer*, Vol.28, No.9, pp.18-22 (1995).
- 9) Flickner, M., et al.: Query by Image and Video Content: The QBIC System, *IEEE Computer*, Vol.28, No.9, pp.23-32 (1995).
- 10) Faloutsos, C., et al.: Efficient and Effective Query by Image Content, *J. Intell. Inform. Syst.*, Vol.3, pp.231-262 (1994).
- 11) Eakins, J.P. and Graham, M.E.: Content-based image retrieval, *Report to the JISC Technology Applications Programme* (1999). <http://www.unn.ac.uk/iidr/research/cbir/>.
- 12) Rui, Y., Huang, T. and Chang S-F.: Image Retrieval: Current Techniques, Promising Directions, and Open Issues, *J. Visual Communication and Image Representation (JVCIR)*, Vol.10, No.1, pp.39-62 (1999).
- 13) Mallat, S.G.: Multifrequency channel decompositions of images and wavelet models, *IEEE Trans., Acoust., Speech and Signal Proc.*, Vol.37, No.12, pp.2091-2110 (1989).
- 14) Chui, C.K.: *An introduction to Wavelets*, Academic Press Inc., Florida (1992).
- 15) Stollnitz, E.J. DeRose, T.D. and Salesin, D.H.: Wavelets for computer graphics: A primer, part 1., *IEEE Computer Graphics and Applications*, Vol.15, No.3, pp.76-84 (1995).
- 16) Stollnitz, E.J. DeRose, T.D. and Salesin, D.H.: Wavelets for computer graphics: A primer, part 2., *IEEE Computer Graphics and Applications*, Vol.15, No.4, pp.75-85 (1995).
- 17) 下垣弘行, 加藤俊一, 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, Vol.J72-DII, No.4, pp.535-544 (1989).
- 18) 串間知彦, 赤間浩樹, 紺谷精一, 木本晴夫, 山室雅司: オブジェクトに基づく高速画像検索システム: ExSight, 情報処理学会論文誌, Vol.40, No.2, pp.732-741 (1999).
- 19) 上原祐介, 遠藤 進, 指田直毅, 増本大器, 長田茂美, 棚橋純一: MIRACLES: マルチメディア情報のパノラミック検索システム—Web 検索への応用, 電子情報通信学会技術研究報告(データ工学), Vol.100, No.31, pp.17-24 (2000).
- 20) 小早川倫広, 星 守, 大森 匡: ウェーブレット変換を用いた対話的類似画像検索と民族資料データベースへの適用, 情報処理学会論文誌, Vol.40, No.3, pp.899-911 (1999).
- 21) Wang, J.Z., Wiederhold, G., Firschein, O. and Wei, S-X.: Content-based Image Indexing and Searching Using Daubechies' Wavelets, *Intl. J. Digital Libraries (IJODL)*, Vol.1, No.4, pp.311-328, Springer-Verlag (1998). <http://www-db.stanford.edu/~zwang/project/imsearch/IJODL97/>.
- 22) Niblack, W., et al.: The QBIC project: Query Image by content using color, texture and shape, *SPIE Storage and Retrieval for Image and Video Databases*, San Joes, pp.173-187

- (1993).
- 23) Smith, J.R. and Chang, S-F.: VisualSEEk: A Fully Automated Content-Based Image Query System, *Proc. ACM Intl. Conf. on Multimedia*, Boston, pp.87-98 (1996).
- 24) Pentland, A., Picard, R.W. and Schlaroff, S.: Photobook: Tools for Content-Based Manipulation of Image Databases, *SPIE Storage and Retrieval for Image and Video Databases II*, San Joes, pp.34-47 (1994).
- 25) Natsev, A., Rastogi, R. and Shim, K.: WALRUS: A Similarity Retrieval Algorithm for Image Databases, *ACM SIGMOD Intl. Conf. on Management of Data*, Philadelphia, PA, pp.395-406 (1999).
- 26) Jacobs, C.E., Finkelstein, A. and Salesin, D.H.: Fast multiresolution Image Querying, *Proc. SIGGRAPH95*, Los Angeles, California, pp.6-11 (1995).
- 27) Liang, K-C. and Kuo, C.-C.J.: WaveGuide: A Joint Wavelet-Based Image representation and Description System, *IEEE Trans. Image Processing*, Vol.8, No.11, pp.1619-1629 (1999).
- 28) Khoshafian, S., Dasananda, S. and Minassian, N.: *The Jasmine Object Database: Multimedia Application for the Web* (1998). コンピュータ・アソシエイツ(監修):日本語版,オーム社(1999).
- 29) Swedish University Network FTP server, <ftp://ftp.sunet.se/pub/pictures>.
- 30) Laaksonen, J., Koskela, M. and Oja, E.: Application of Tree Structured Self-Organizing Maps in Content-Based Image Retrieval, *Proc. ICANN '99*, Edinburgh, UK (Sep. 1999).
- 31) 呉 君錫, 金子邦彦, 牧之内顕文: Wavelet-SOM による画像の分類と検索手法, 第 10 回データ工学ワークショップ (DEWS '99) (1999).
- 32) Oh, K.S., Kaneko K. and Makinouchi, A.: Image Classification and Retrieval based on Wavelete-SOM, *Proc. 1999 Intl. Symp. on Database Applications in Non-Traditional Environments (DANTE '99)*, pp.164-167 (1999).
- 33) 呉 君錫, 金子邦彦, 牧之内顕文, 上野敦子: 自己組織化特徴マップに基づいた類似検索システムの設計・実装と性能評価, 電子情報通信学会技術研究報告 (データ工学), Vol.100, No.31, pp.9-16 (2000).
- 34) Koikkalainen, P.: Progress with the Tree-Structured Self-Organizing Map, *Proc. ECAI '94*, Cohn, A. (Ed.), pp.211-215, John Wiley & Sons (1994).
- 35) Prabhakar, S., Agrawal, D. and Abbadi, A.E.: Browsing and Placement of Multiresolution Images on Secondary Storage, *TRCS96-22*,

*Dept. of Computer Science Univ. of California*, (1996). [http://www.cs.ucsb.edu/T\\_Rs/Docs/](http://www.cs.ucsb.edu/T_Rs/Docs/).

(平成 12 年 6 月 20 日受付)

(平成 12 年 10 月 14 日採録)

(担当編集委員 田中 克己)



呉 君錫 (学生会員)

1993 年福岡工業大学工学部管理工学科卒業。1998 年同大学大学院工学研究科情報工学専攻修士課程修了。同年九州大学大学院システム情報科学研究科知能システム学専攻入学, 現在に至る。画像データベースの研究に従事。電子情報通信学会学生会員。



金子 邦彦 (正会員)

1990 年九州大学工学部情報工学科卒業。1995 年同大学大学院博士後期課程修了, 同助手を経て, 1999 年九州大学大学院システム情報科学研究院助教授。電子情報通信学会, ACM, IEEE Computer Society 各会員。



牧之内顕文 (正会員)

1967 年京都大学工学部電子工学科卒業。1970 年グルノーブル大学理学部応用数学科 Docteur-Ingénieur 取得。(株)富士通,(株)富士通研究所,九州大学工学部教授を経て,1996 年九州大学大学院システム情報科学研究科教授。現在,同大学院システム情報科学研究院教授。電子情報通信学会,ACM,IEEE Computer Society,人工知能学会各会員。



Sang-Hyun Bae

1982 年韓国朝鮮大学校電気工学科卒業。1988 年東京都立大学大学院博士課程修了。工学博士。1996 年韓国朝鮮大学校電子計算所長を経て,現在,同大学政策大学院情報通信政策学科主任教授。韓国インターネット情報学会理事。情報通信政策,人工知能,知能情報システムの研究に従事。韓国情報科学会,韓国インターネット情報学会,韓国海洋情報通信学会各会員。