

# ウェーブレット分解係数の階層的相関関係を用いた テクスチャ類似画像検索

小早川 倫広<sup>†</sup> 星 守<sup>†</sup> 大森 匡<sup>†</sup>

本稿は、画像に対してウェーブレット変換を実行して得られるウェーブレット分解係数の階層的相関関係に着目したテクスチャ特徴量を定義した後、このテクスチャ特徴量を用いたテクスチャ画像類似検索を提案する。本方式の基本的アイデアは、テクスチャ特徴量としてウェーブレット分解係数の隣接する分解レベル間の階層的相関関係を用いることである。テクスチャ類似画像検索の性能評価のための画像データは、実際に市販されている 51 種類の生地を用意し、1 枚の生地からスキャナで解像度を 10 段階変化させ画像を読み取り、計 510 枚の生地データベースを構築した。評価実験の結果、提案するテクスチャ類似画像検索方式は画像の解像度に対して頑強なテクスチャ検索方式であることが分かった。

## Texture Image Retrieval Based on the Hierarchical Correlations of Wavelet Coefficients

MICHIHIRO KOBAYAKAWA,<sup>†</sup> MAMORU HOSHI<sup>†</sup> and TADASHI OHMORI<sup>†</sup>

In this paper we propose a robust texture image retrieval using hierarchical relations between the decomposed subimages obtained by wavelet transform. Key idea is to describe texture information in terms of the hierarchical correlations between the wavelet coefficients of the adjacent level. Firstly, we express the pyramidal structure of wavelet coefficients by associating the nodes of a complete quad tree with the wavelet coefficients. Secondly, we define a hierarchical dissimilarity vector between a parent node and his child, to express a hierarchical relation between them. Thirdly, to describe a relation among child nodes, we compute a covariance matrix of the dissimilarity vectors. We associate the covariance matrix with the parent node. We define the texture vector by the diagonal of elements of the covariance matrix. And then define the texture feature vector of level  $l$  by the pair of the mean and the standard deviation of texture vectors of level  $l$ . Finally, by applying the discriminant analysis to the set of the texture feature vectors, we make an effective index of the database. For retrieving similar images, we use the  $k$ -nearest neighbor search in the index space. The similarity between two images is defined by the Euclidean distance between the corresponding feature vectors of the images. To evaluate the performance of the retrieval, we made experiments on "Cloth Collections" consisting of 51 textile patterns with 10 different resolutions (image size is  $1024 \times 1024$  pixels). The experiments showed that the performance of retrieval is good and that the proposed method is robust with respect to resolution.

### 1. はじめに

マルチメディアデータを効率的に管理・運用するためには、マルチメディアデータを検索する機能が必須となる。従来のように、マルチメディアデータに対してキーワードを付加してデータベースを構築した場合、データベースの利用者はデータベースを検索するために的確なキーワードの入力を要求され、的確なキーワードを連想できない利用者は、検索を開始

することができないという問題があった。この問題の 1 つの解決手法として、マルチメディアデータそのものを問合せとして検索を行う、内容に基づく類似検索 (Content-Based Retrieval) が着目を浴びている<sup>2)</sup>。

たとえば、内容に基づく類似画像検索では、データベースの利用者は、写真やスケッチなどの画像を用いて画像データベースの検索を開始する。画像データベースは、画像データベースに登録されている画像の中から利用者の問合せ画像に類似した画像を検索し結果を表示する。

内容に基づく類似画像検索では、画像の内容を記述する特徴量の検討が重大な課題となる。一般に、画

<sup>†</sup> 電気通信大学大学院情報システム学研究科  
Graduate School of Information Systems, University of  
Electro-Communications

像から抽出される特徴量は高次元ベクトルとして表現されることから、検索効率を高めるための索引生成技術が不可欠となる。最近の画像検索の研究として、IBMのQBIC<sup>1)</sup>、電総研のTRADEMARK<sup>5)</sup>やARTMUSEUM<sup>6),8)</sup>、NTTのオブジェクトに基づく類似画像検索システム ExSight<sup>11)</sup> などがある。これらのシステムで、IBMのQBICだけが、さまざまな画像内容に対応する検索システムである。

我々は、さまざまな画像内容に対する特徴量抽出を、個々の画像内容ごとに別々の最適化された特徴量抽出アルゴリズムを用いるのではなく、ウェーブレット分解空間を基礎として、ウェーブレット分解空間から画像内容に対応した特徴量を考えることによって、画像内容を記述する特徴量を統一のとらえてシステムを構成する。この枠組みで我々は、写真、スケッチ、輪郭などから画像を検索する対話的類似検索システムをすでに開発した<sup>3),7)</sup>。

本稿は、画像のテクスチャに着目し、上記の枠組みの上にズームング(拡大・縮小)・平行移動に対してロバストなテクスチャ画像検索法を提案する。テクスチャ検索における回転に対する頑強性は、次の理由から要請しない: 回転したテクスチャは、元のものとは異なるテクスチャである。たとえば、縦縞・横縞・斜め縞のテクスチャは、同一のテクスチャであるというよりも異なるテクスチャであると見なすほうがよい。

2章では、画像に対するウェーブレット変換<sup>4)</sup>を簡単に説明する。3章では、画像に対してウェーブレット変換を実行して得られる分解空間を4分木で表現し、4分木で表される分解レベル間の階層関係に着目した、テクスチャ特徴量を提案する。4章では、画像の索引生成アルゴリズムと画像データベースへの検索アルゴリズムを示す。5章では、実際に生地をスキャナで読み取ってテクスチャ画像を作成し、このテクスチャ画像を用いて画像データベースを構築し、検索能力について評価実験を行う。最後に、本稿のまとめを行う。

## 2. 離散ウェーブレット変換

離散ウェーブレット変換は、離散コサイン変換やFFTなどと同じように画像や信号を空間周波数解析するための解析方法である。本章では、2次元離散ウェーブレット変換について簡単に説明する。

1次元ウェーブレット変換の拡張である2次元ウェーブレット変換は、 $x$ 軸と $y$ 軸に関するスケール関数 $\phi(2^{-l}x - m)$ 、 $\phi(2^{-l}y - n)$ とマザーウェーブレット $\psi(2^{-l}x - m)$ 、 $\psi(2^{-l}y - n)$ の組合せ

$$\phi_{l,mn}(x, y) \equiv \phi(2^{-l}x - m)\phi(2^{-l}y - n), \quad (1)$$

$$\psi_{l,mn}^1(x, y) \equiv \psi(2^{-l}x - m)\phi(2^{-l}y - n), \quad (2)$$

$$\psi_{l,mn}^2(x, y) \equiv \phi(2^{-l}x - m)\psi(2^{-l}y - n), \quad (3)$$

$$\psi_{l,mn}^3(x, y) \equiv \psi(2^{-l}x - m)\psi(2^{-l}y - n), \quad (4)$$

によって変換が実現される(ただし、 $m, n$ は整数)。ここで、 $\phi_{l,mn}$ 、 $\psi_{l,mn}^1$ 、 $\psi_{l,mn}^2$ 、 $\psi_{l,mn}^3$ の張る空間をレベル $l$ の分解空間と呼び、 $V_l$ 、 $W_l^1$ 、 $W_l^2$ 、 $W_l^3$ と表す。このとき、レベル $l-1$ の分解空間 $V_{l-1}$ とレベル $l$ の分解空間 $V_l$ 、 $W_l^1$ 、 $W_l^2$ 、 $W_l^3$ の間には、 $V_{l-1} = V_l \oplus W_l^1 \oplus W_l^2 \oplus W_l^3$  (5) が成り立つ( $\oplus$ は直和)。

画像は、関数 $f(x, y)$ によって表現でき、ウェーブレット基底関数 $\phi_{0,mn}(x, y)$ によって次のように表される。

$$f(x, y) = \sum_m \sum_n a_{0,mn} \phi_{0,mn}(x, y), \quad \in V_0, \quad (6)$$

ただし、係数 $a_{0,mn}$ は式(7)によって与えられる。すなわち、分解係数 $a_{0,mn}$ は、関数 $f(x, y)$ と基底関数 $\phi_{0,mn}(x, y)$ の内積

$$\begin{aligned} a_{0,mn} &= \langle f, \phi_{0,mn} \rangle \\ &\equiv \int \int_{-\infty}^{+\infty} f(x, y) \phi_{0,mn}(x, y) dx dy \end{aligned} \quad (7)$$

である。

ここで、空間 $V_0$ への射影成分の値である分解係数 $a_{0,mn}$ は、画像を $2^L \times 2^L$ 画素で表現するときの $(m, n)$ 画素の画素値であることに注意されたい。すなわち、 $2^L \times 2^L$ 画素からなる画像は行列 $I \equiv (a_{0,mn})_{m=0, \dots, 2^L-1; n=0, \dots, 2^L-1}$ で表現できる。以後、本稿では画像を行列 $I$ を用いて表す。

ここで、ウェーブレット変換による原画像 $I \equiv (a_{0,mn}) \in V_0$ の分解について眺めてみる。原画像 $I$ に対してウェーブレット変換を実行することにより、原画像 $I$ は分解レベル1の近似画像(縮小画像) $A_1$ と差分画像 $D_1^1$ 、 $D_1^2$ 、 $D_1^3$ に分解される。ただし、分解レベル1の各分解画像は、それぞれ次式から計算される。

$$A_1 \equiv (a_{1,mn}) = (\langle f, \phi_{1,mn} \rangle), \quad A_1 \in V_1, \\ m, n = 0, \dots, 2^{L-1} - 1,$$

$$D_1^k \equiv (d_{1,mn}^k) = (\langle f, \psi_{1,mn}^k \rangle), \quad D_1^k \in W_1^k, \\ m, n = 0, \dots, 2^{L-1} - 1; \quad k = 1, 2, 3.$$

近似画像に対して再帰的にウェーブレット変換を実行することにより、レベル $L$ の分解空間まで分解する。このようにして、近似画像群 $\{A_1, A_2, \dots, A_L\}$

と差分画像群  $\{D_1^k, D_2^k, \dots, D_L^k\}_{(k=1,2,3)}$  を得る．ここで，近似画像群  $\{A_1, A_2, \dots, A_L\}$  と3つの差分画像群  $\{D_1^k, D_2^k, \dots, D_L^k\}_{(k=1,2,3)}$  はそれぞれ，レベル1の近似画像  $A_1$  と差分画像  $D_1^k$  を底辺としたピラミッド構造をとることに注意されたい．

### 3. テクスチャ特徴量

レベルが高くなるに従って，分解係数のサイズが1/4ずつ縮小するピラミッド構造は，4分木でも表現できる．図1に示すように，レベル  $l+1$  の中間ノード  $n_{l+1,ij}$  は，左から順にレベル  $l$  の4つの子ノード  $n_{l,2i-2j}, n_{l,2i+1-2j}, n_{l,2i-2j+1}, n_{l,2i+1-2j+1}$  を持つ高さ  $L$  の4分木  $QT$  を考る．以後，4分木の各ノード  $n_{l,ij}$  にさまざまな付加情報  $v_{l,ij}$  を割り当てる．ノード  $n_{l,ij}$  の値  $v_{l,ij}$  は， $\text{Val}(n_{l,ij}; QT)$  または，簡単に  $\text{Val}(n_{l,ij})$  で表す．

ここで，近似係数  $a_{l,ij}$  と3つの差分係数  $d_{l,ij}^k$  の組をウェーブレット分解係数ベクトル

$$c_{l,ij} = (a_{l,ij}, d_{l,ij}^1, d_{l,ij}^2, d_{l,ij}^3)^T, \quad (8)$$

$$l = 1, \dots, L; \quad i, j = 0, \dots, 2^{L-l} - 1,$$

として定義する．ここで， $^T$  は，転置を表す．

4分木の各ノード  $n_{l,ij}$  に，ウェーブレット分解係数ベクトル  $c_{l,ij}$  を割り当てたものを，Wavelet Tree (WT) と呼ぶことにする．この4分木 WT の各ノード  $n_{l,ij}$  の値は，

$$\text{Val}(n_{l,ij}; WT) = c_{l,ij}, \quad (9)$$

$$l = 1, \dots, L; \quad i, j = 0, \dots, 2^{L-l} - 1,$$

である．

ウェーブレット変換は，分解レベルをあげる(変換

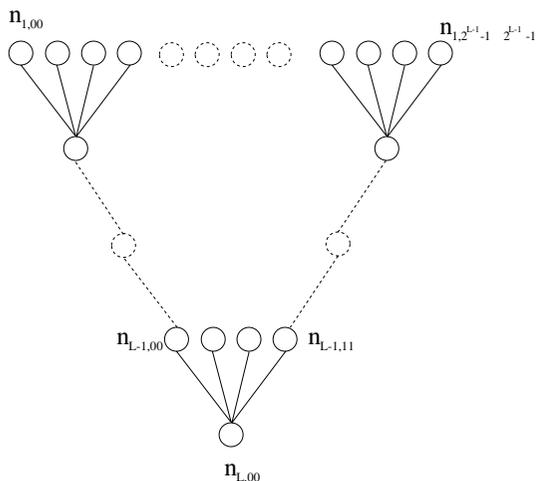


図1 高さ  $L$  の4分木構造

Fig. 1 The structure of "Quad Tree" of height  $L$ .

領域を拡大する)ことにより，より低い周波数成分を求めている．このことから，WTの隣接するレベルの親ノード  $n_{l+1,mn}$  とその子  $n_{l,ij}$  の間には，周波数成分の変動が存在する．ここで，WTの親ノード  $n_{l+1,mn}$  とその子  $n_{l,ij}$  の間の不一致度を階層偏差ベクトル

$$g_{l,ij} \equiv (g_{l,ij}^0, g_{l,ij}^1, g_{l,ij}^2, g_{l,ij}^3)^T = c_{l,ij} - c_{l+1,mn}$$

$$= \begin{pmatrix} a_{l,ij} - a_{l+1,mn} \\ d_{l,ij}^1 - d_{l+1,mn}^1 \\ d_{l,ij}^2 - d_{l+1,mn}^2 \\ d_{l,ij}^3 - d_{l+1,mn}^3 \end{pmatrix}, \quad (10)$$

$$l = 1, \dots, L-1; \quad i, j = 0, \dots, 2^{L-l} - 1,$$

$$m = \lfloor \frac{i}{2} \rfloor, n = \lfloor \frac{j}{2} \rfloor,$$

で定義する．親ノード  $n_{l+1,mn}$  の子ノード  $n_{l,2m-2n}, n_{l,2m+1-2n}, n_{l,2m-2n+1}, n_{l,2m+1-2n+1}$  間の相関関係を共分散行列

$$S_{l+1,mn}$$

$$= \begin{pmatrix} s_{l+1,mn}^{00} & s_{l+1,mn}^{01} & s_{l+1,mn}^{02} & s_{l+1,mn}^{03} \\ s_{l+1,mn}^{10} & s_{l+1,mn}^{11} & s_{l+1,mn}^{12} & s_{l+1,mn}^{13} \\ s_{l+1,mn}^{20} & s_{l+1,mn}^{21} & s_{l+1,mn}^{22} & s_{l+1,mn}^{23} \\ s_{l+1,mn}^{30} & s_{l+1,mn}^{31} & s_{l+1,mn}^{32} & s_{l+1,mn}^{33} \end{pmatrix},$$

$$l = 1, \dots, L-1; \quad m, n = 0, \dots, 2^{L-(l-1)} - 1,$$

で定義する．ただし，行列  $S_{l+1,mn}$  の要素  $s_{l+1,mn}^{\alpha\beta}$  の値は，第  $\alpha$  子と第  $\beta$  子の階層偏差ベクトルの内積である．

行列  $S$  の対角要素は，親ノードとその子ノード間の2次の変動スペクトルの総和を表現している．この2次の変動スペクトルの総和を用いてテクスチャを考える．すなわち，行列  $S$  の対角要素を取り出し，これらをテクスチャベクトル

$$f_{l,mn} \equiv (f_{l,mn}^0, f_{l,mn}^1, f_{l,mn}^2, f_{l,mn}^3)^T$$

$$= (s_{l,mn}^{00}, s_{l,mn}^{11}, s_{l,mn}^{22}, s_{l,mn}^{33})^T, \quad (11)$$

$$l = 2, \dots, L; \quad m, n = 0, \dots, 2^{L-l} - 1,$$

と定義して，テクスチャ類似検索の索引生成に用いる．

4分木の各ノード  $n_{l,mn}$  にテクスチャベクトル  $f_{l,mn}$  を割り当てた木を“Texture Feature Tree (TFT)”と呼ぶ．このTFTが，我々のテクスチャ解析の基本的な4分木になる．

レベル  $l$  のテクスチャベクトルを用いて，レベル  $l$  のテクスチャ特徴ベクトル  $F_l$  を求める．テクスチャ特徴ベクトル算出アルゴリズムを次に示す．

## [ テクスチャ特徴ベクトル算出アルゴリズム ]

- F-1 画像  $I$  に対してウェーブレット変換を実行し、ウェーブレット分解係数を得る。ここで、近似係数  $a_{l,ij}$  と 3 つの差分係数  $d_{l,ij}^k$  の組をウェーブレット分解係数ベクトル  $c_{l,ij}$  とする。
- F-2 隣接するレベル間のノードの階層偏差ベクトル  $g_{l,ij}$  を算出する。
- F-3 階層偏差ベクトル  $g_{l,ij}$  を用いて、レベル  $l+1$  の行列  $S_{l+1,mn}$  を求める。ここで、レベル  $l+1$  の行列  $S_{l+1,mn}$  の対角成分を、レベル  $l+1$  のテクスチャベクトル  $f_{l+1,mn}$  とする。
- F-4 レベル  $l$  のテクスチャベクトルの平均ベクトル  $\mu_l = (\mu_l^0, \mu_l^1, \mu_l^2, \mu_l^3)^T$  と標準偏差ベクトル  $\sigma_l = (\sigma_l^0, \sigma_l^1, \sigma_l^2, \sigma_l^3)^T$  :

$$\mu_l^k = \frac{\sum_{i=0}^{2^{L-l}-1} \sum_{j=0}^{2^{L-l}-1} f_{l,mn}^k}{2^{L-l} \times 2^{L-l}},$$

$$\sigma_l^k = \sqrt{\frac{\sum_{i=0}^{2^{L-l}-1} \sum_{j=0}^{2^{L-l}-1} (f_{l,mn}^k - \mu_l^k)^2}{2^{L-l} \times 2^{L-l}}},$$

$$k = 0, 1, 2, 3. \quad (12)$$

を計算する。

- F-5 レベル  $l$  の平均ベクトル  $\mu_l$  と標準偏差ベクトル  $\sigma_l$  の組をレベル  $l$  のテクスチャ特徴ベクトル  $F_l \equiv (\mu_l, \sigma_l)$ , (13) と定義する。

テクスチャ特徴ベクトル算出アルゴリズムにより求めたテクスチャ特徴ベクトル  $F$  を用いてテクスチャ類似検索を行う。

## 4. 索引生成方法と検索方法

本章では、 $N$  枚の画像からなる画像データベースの索引生成方法と検索方法を説明する。

## 4.1 索引生成アルゴリズム

画像データベース中の画像は、テクスチャ特徴ベクトル算出アルゴリズムにより特徴ベクトルが抽出される。さらに、より効率的な索引空間を生成するために判別分析が実行され、画像データに対する索引が生成される。判別分析は、多変量解析の手法の 1 つで、クラス内分散とクラス間分散の比を最小にする線形写像である(詳しくは、文献 9), 10) を参照されたい)。  $N$  枚の画像からなる画像データベースの索引生成アルゴリズムを下記に示す。

## [ 索引生成アルゴリズム ]

- I-1: 次のことを  $N$  枚の画像に対して行う: 画像データベース中の  $n$  番目の画像  $I_n$  に対してテク

スチャ特徴ベクトルアルゴリズムを実行し、レベル  $l$  のテクスチャ特徴ベクトル  $F_{n,l}$  を算出する。ここで、テクスチャ特徴ベクトル  $F_{n,l}$  は、クラス  $n$  に属するものとする。

## I-2: テクスチャ特徴ベクトルの集合

$\{F_{n,j}\}_{n=1,\dots,N;l=2,\dots,L}$  に対してクラス数を  $N$  として判別分析を実行する。得られた線形写像によって、 $n$  番目の画像  $I_n$  のレベル  $l$  のテクスチャ特徴ベクトル  $F_{n,l}$  は、判別空間上の点  $\hat{F}_{n,l}$  に写像される(判別分析で得られた判別空間を画像データベースの索引空間と呼ぶ)。

- I-3: 索引空間に写像された  $n$  番目の画像  $I_n$  のテクスチャ特徴ベクトル  $\hat{F}_{n,l}$  の重心  $i_n$  を求める(これを  $n$  番目の画像の索引ベクトルとし、画像の索引として検索に用いる)。

## 4.2 検索アルゴリズム

類似画像検索には、 $k$ -近傍探索を用いる。システムは、問合せ画像  $I_q$  から、データベース登録画像と同様にテクスチャ特徴ベクトルを算出し、索引空間において検索アルゴリズムに従って検索を実行する。問合せ画像  $I_q$  が入力され、検索の結果の画像が求められるまでの手順を下記に示す。

## [ 検索アルゴリズム ]

- R-1: 問合せ画像  $I_q$  に対してテクスチャ特徴ベクトル算出アルゴリズムを実行し、問合せ画像  $I_q$  のテクスチャ特徴ベクトル  $F_{q,l(l=2,\dots,L)}$  を算出する。
- R-2: 問合せ画像  $I_q$  のテクスチャ特徴ベクトル  $F_{q,l}$  を得られた線形写像により索引空間上に写像する。ここで、写像された問合せ画像  $I_q$  のテクスチャ特徴ベクトル  $\hat{F}_{q,l}$  の重心ベクトル  $i_q$  (問合せベクトルと呼ぶ)を求める。
- R-3: 索引空間中の画像  $I_n$  の索引ベクトル  $i_n$  と問合せ画像  $I_q$  の問合せベクトル  $i_q$  とのユークリッド距離

$$d_n = \|i_n - i_q\|_2, \quad (14)$$

を類似度として  $k$ -近傍探索を実行する。問合せベクトル  $i_q$  と距離が近い画像から順に検索結果として、データベース利用者に表示する。

## 5. 評価実験

## 5.1 生地コレクション

評価実験を行うため、生地画像を用いたテクスチャ画像データベースを構築した。用いた生地画像は、51 種類の生地をスキャナで 10 段階 (125 dpi, 150 dpi, 200 dpi, 250 dpi, 300 dpi, 350 dpi, 400 dpi, 450 dpi,



(a)

(b)

図 2 検索結果 . (a) 141 番の画像から 150 番の画像は、問合せ画像の同種画像、(b) 201 番の画像から 210 番の画像は、問合せ画像の同種画像

Fig. 2 Results of a retrieval: (a) The images with the number from 141 to 150 are brother images of the query image, (b) The images with the number from 201 to 210 are brother images of the query image.

500 dpi, 550 dpi) の解像度で読み取った画像であり、計 510 枚の画像 (サイズ  $1024 \times 1024$  画素) から構成されている。我々は、同一の生地からスキャナで読み取った 10 枚の画像を同種画像と呼ぶ。以下では、510 枚のテクスチャ画像を  $I_i$  と  $R_r$  で識別する。ここで  $I_i$  ( $i = 1, \dots, 51$ ) は、 $i$  番目の生地を表し、 $R_r$  ( $i = 1, \dots, 10$ ) は  $r$  番目の解像度であることを表す。すなわち  $i$  番目の生地  $I_i$  の解像度  $R_r$  の画像は  $(I_i, R_r)$  と表される。

$R_1$  のラベルが付加された画像は最も読み取り解像度が低く、テクスチャの全体的な情報を持っている。一方、 $R_{10}$  のラベルが付加された画像は、最も読み取り解像度が高く、細かい局所的なテクスチャ情報を持っている。

## 5.2 結 果

データベース中の各画像データは、テクスチャ特徴ベクトル算出アルゴリズムを実行し、画像のテクスチャ特徴ベクトルを求める。この画像は、Haar 基底を用いたウェーブレット変換によりレベル 10 まで分解されるが、解像度  $R_1$  の画像のレベル 4 からレベル

10 のテクスチャベクトルの値は 0 に近いので、これらのレベルのテクスチャ特徴ベクトルは使わず、レベル 2 と 3 のテクスチャ特徴ベクトル  $F_2$  と  $F_3$  を用いて索引生成アルゴリズムを適用し、索引空間を生成した。この操作より画像データベース中の画像 510 枚に対する索引ベクトルが得られる。

システムは、問合せ画像  $I_q$  が入力されると、問合せ画像に対してテクスチャ特徴ベクトル算出アルゴリズムを実行し、問合せ画像  $I_q$  のレベル 2 と 3 のテクスチャベクトル  $F_{q,2}$ ,  $F_{q,3}$  を計算し、索引空間に写像して問合せベクトル  $i_q$  を算出する。問合せベクトル  $i_q$  と索引ベクトル  $i_n$  との間のユークリッド距離を計算し、問合せベクトル  $i_q$  から距離の近いものから  $k$  枚の画像を表示する。この実験では、20 枚の検索結果を表示している。テクスチャ検索の結果を図 2 に示し、問合せ画像と検索結果として表示された画像間の類似度を 510 枚の問合せ画像の検索における最大類似度で正規化した値を表 1 に示す。

図 2(a) の検索では、10 枚中 10 枚 (すなわちすべて) の同種画像が検索できている。

表 1 問合せ画像と検索結果画像間の類似度. □ で囲まれた値は, 問合せ画像と同種画像との距離である.

Table 1 Similarity between the query image and the retrieved images. Similarity between the brother images and the query image are marked by □.

順位	検索実験 (a) に用いた問合せ画像	検索実験 (b) に用いた問合せ画像
1	0.000090	0.003316
2	0.000093	0.004549
3	0.000107	0.005679
4	0.000197	0.005918
5	0.000293	0.009337
6	0.000556	0.014542
7	0.000565	0.017887
8	0.000675	0.020711
9	0.000730	0.020725
10	0.000840	0.024562
11	0.000944	0.027806
12	0.000952	0.028290
13	0.000977	0.029242
14	0.001013	0.030576
15	0.001038	0.030748
16	0.001070	0.032551
17	0.001076	0.032981
18	0.001170	0.034074
19	0.001215	0.036475
20	0.001232	0.040528

また, 図 2 (b) の検索でも, 10 枚中 10 枚の同種画像が検索できている.

問合せ画像を入力し, 索引空間を生成, 近傍探索を行い 20 枚の画像番号を算出するまでに要する時間 (OS : Linux, CPU : PentiumII 300 MHz, Memory : 256 MB) は, 90 ms であった.

### 5.3 評価・考察

本節では, 51 種類の生地から構成される 510 枚の生地画像データベースに対して, 生地画像 510 枚を問合せ画像として検索を行い, 本提案方式の検索能力について検討する. まず, 問合せ画像によって同種画像がどれだけ検索できたかを表 2 のように整理する. ただし, 表の要素  $h_{\alpha\beta}^{\delta}$  は, 次のように定義される.

$$h_{\alpha\beta}^{\delta} = \begin{cases} 1 & \text{問合せ画像 } I_{\alpha}R_{\beta} \text{ の同種画像} \\ & I_{\alpha}R_{\delta} \text{ が検索された} \\ 0 & \text{それ以外,} \end{cases} \quad (15)$$

$(\delta = 1, \dots, 10, \alpha = 1, \dots, 51, \beta = 1, \dots, 10).$

行に関する合計  $h_{\alpha\beta}^{\bullet}$  は, 問合せ画像  $I_{\alpha}R_{\beta}$  に対する検索において同種画像を何枚検索できたかを示しており, 次式で表される.

表 2 問合せ画像  $I_{\alpha}R_{\beta}$  に対する検索結果.  $\alpha$ : 生地番号,  $\beta$ : 問合せ画像の解像度,  $\delta$ : 検索された画像の解像度

Table 2 Query-response pattern table:  $\alpha$ : cloth number,  $\beta$ : resolution of a query image,  $\delta$ : resolution of a retrieved image.

		$\delta$			
$I_{\alpha}$	$R_{\beta}$	1	...	10	total
1	1	$h_{1\ 1}^1$	...	$h_{1\ 1}^{10}$	$h_{1\ 1}^{\bullet}$
⋮	⋮	⋮	...	⋮	⋮
51	1	$h_{51\ 1}^1$	...	$h_{51\ 1}^{10}$	$h_{51\ 1}^{\bullet}$
subtotal	1	$h_{\bullet\ 1}^1$	...	$h_{\bullet\ 1}^{10}$	$h_{\bullet\ 1}^{\bullet}$
⋮	⋮	⋮	...	⋮	⋮
1	10	$h_{1\ 10}^1$	...	$h_{1\ 10}^{10}$	$h_{1\ 10}^{\bullet}$
⋮	⋮	⋮	...	⋮	⋮
51	10	$h_{51\ 10}^1$	...	$h_{51\ 10}^{10}$	$h_{51\ 10}^{\bullet}$
subtotal	10	$h_{\bullet\ 10}^1$	...	$h_{\bullet\ 10}^{10}$	$h_{\bullet\ 10}^{\bullet}$
total		$h_{\bullet\ \bullet}^1$	...	$h_{\bullet\ \bullet}^{10}$	$h_{\bullet\ \bullet}^{\bullet}$

$$h_{\alpha\beta}^{\bullet} \equiv \sum_{\delta=1}^{10} h_{\alpha\beta}^{\delta} \quad (16)$$

列に関する小計  $h_{\bullet\beta}^{\delta}$  は, 解像度  $R_{\beta}$  の問合せ画像の入力に対して, 解像度  $R_{\delta}$  の同種画像がどれだけ検索できたかを示しており, 次式で表される.

$$h_{\bullet\beta}^{\delta} \equiv \sum_{\alpha=1}^{51} h_{\alpha\beta}^{\delta} \quad (17)$$

表の小計  $h_{\bullet\beta}^{\bullet}$  は, 解像度  $R_{\beta}$  の問合せ画像によって何枚の同種画像が検索できたかを示しており, 次式で表される.

$$h_{\bullet\beta}^{\bullet} \equiv \sum_{\alpha=1}^{51} \sum_{\delta=1}^{10} h_{\alpha\beta}^{\delta} = \sum_{\alpha=1}^{51} h_{\alpha\beta}^{\bullet} = \sum_{\delta=1}^{10} h_{\bullet\beta}^{\delta} \quad (18)$$

列に関する合計  $h_{\bullet\bullet}^{\delta}$  は, 510 枚の問合せに対して解像度  $R_{\delta}$  の画像が何枚検索されたかを示しており, 次式で表される.

$$h_{\bullet\bullet}^{\delta} \equiv \sum_{\alpha=1}^{51} \sum_{\beta=1}^{10} h_{\alpha\beta}^{\delta} = \sum_{\beta=1}^{10} h_{\bullet\beta}^{\delta} \quad (19)$$

データベースに登録されている問合せ画像  $I_{\alpha}R_{\beta}$  の同種画像の総和 (ここでは 10) に対する検索された同種画像の総数の比を問合せ画像  $I_{\alpha}R_{\beta}$  に対するヒット率とし, 次式で定義する.

$$A_{\alpha\beta}^{\bullet} \equiv \frac{h_{\alpha\beta}^{\bullet}}{10} \quad (20)$$

ヒット率  $A_{\alpha\beta}^{\bullet}$  は, 個々の問合せ画像  $I_{\alpha}R_{\beta}$  に対する検索結果の良し悪しのみを示しているにすぎない. たとえば,  $A_{\alpha\beta}^{\bullet} = 1.0$  は, データベース中の同種画像

表3 ヒット率  $A_{\alpha\beta}^{\delta}$ . □で囲まれた値は、ヒット率 0.7 以上である.  
 Table 3 Hit ratios  $A_{\alpha\beta}^{\delta}$ . Hit ratios more than or equal to 0.70 are marked by □.

$\beta$	$\delta$										average $A_{\alpha\beta}^{\delta}$
	1	2	3	4	5	6	7	8	9	10	
1	0.80	0.67	0.43	0.33	0.25	0.14	0.08	0.06	0.08	0.09	0.29
2	0.82	0.76	0.63	0.47	0.31	0.19	0.16	0.09	0.08	0.06	0.36
3	0.96	1.00	0.80	0.67	0.53	0.33	0.29	0.19	0.16	0.14	0.51
4	0.88	0.98	0.96	0.94	0.76	0.65	0.45	0.31	0.25	0.18	0.64
5	0.73	0.90	1.00	1.00	1.00	0.94	0.73	0.55	0.39	0.25	0.75
6	0.41	0.71	1.00	1.00	1.00	1.00	0.96	0.82	0.59	0.47	0.80
7	0.18	0.49	0.88	1.00	1.00	1.00	1.00	0.98	0.90	0.63	0.81
8	0.12	0.29	0.76	1.00	1.00	1.00	1.00	1.00	0.96	0.82	0.80
9	0.06	0.18	0.55	0.96	1.00	1.00	1.00	1.00	1.00	0.94	0.77
10	0.04	0.14	0.39	0.75	0.94	1.00	1.00	1.00	1.00	1.00	0.73
$A_{\alpha\beta}^{\delta}$	0.50	0.61	0.74	0.81	0.78	0.73	0.67	0.60	0.54	0.46	0.64

$I_{\alpha}R_{\beta}$  をすべて検索できていることを意味する。したがって、以下では、提案方式の解像度に対する頑強性 (robustness) を次の 2 つの点から評価する。

- (1) 検索する際の有効な解像度はどの解像度であるかという視点
- (2) すべての解像度の画像をデータベースに登録するのではなくデータベースに登録する画像の解像度はどの解像度が有効かという視点

まず、問合せ画像の解像度と検索された同種画像の解像度との関係を表 3 に示す。この表の  $(\beta, \delta)$  要素  $A_{\alpha\beta}^{\delta}$  は、次式で定義する。

$$A_{\alpha\beta}^{\delta} \equiv \frac{h_{\alpha\beta}^{\delta}}{51} = \frac{\sum_{\alpha=1}^{51} h_{\alpha\beta}^{\delta}}{51} \quad (21)$$

このヒット率  $A_{\alpha\beta}^{\delta}$  は、解像度  $R_{\beta}$  の問合せ画像 (51 枚) に対して、解像度  $R_{\delta}$  の同種画像がどのくらい検索できたかを示す。

#### 問合せ画像の解像度

問合せ画像の解像度としてどの解像度を用いるかという視点から、検索能力を評価するために、表 3 を行方向で検証する。表 3 の  $\beta$  番目の行は、解像度  $R_{\beta}$  の問合せ画像の検索能力を示す。たとえば、5 番目の行は解像度  $R_5$  の問合せ画像によって検索したときの結果を示している。解像度  $R_5$  の問合せ画像は、解像度  $R_1$  から  $R_7$  の登録画像の 73% 以上が検索された。ここで、要求ヒット率を 0.7 とした場合、解像度  $R_5$  の問合せ画像は解像度  $R_1$  の登録画像から解像度  $R_7$  の登録画像に対して有効な検索が可能であることを示しており、解像度に対して幅広い有効範囲を持っていることが分かる。また、表 3 を全体的に眺めてみると、解像度  $R_{\beta}$  の問合せ画像が有効な解像度の範囲は、問合せ画像の解像度  $R_{\beta}$  が高くなるにつれ解像度の高い方に移動している。

実験に用いた生地データベースの場合、データベースの検索者は、解像度  $R_5$  の画像を問合せ画像として検索を実行すれば、解像度  $R_3$  から  $R_5$  の同種画像が必ず検索可能であることが分かる。同様に、5 番目の行から 10 番目の行を見ると、解像度  $R_5$  から解像度  $R_{10}$  の問合せ画像は、幅広い有効検索範囲 (7 段階) を持っていることが分かる。

表 3 の対角要素の値は、問合せ画像と同じ画像がどれくらいの割合で検索されたかを示している。解像度  $R_5$  以上の問い合わせ画像では、問合せ画像と同じデータベース中の画像は、100% 検索できたことを示している。

$A_{\alpha\beta}^{\delta}$  の 51 枚の生地に関する平均  $A_{\alpha\beta}^{\delta}$  は、解像度  $R_{\beta}$  の問合せ画像の平均的検索能力を示し、次式で表される。

$$A_{\alpha\beta}^{\delta} \equiv \frac{\sum_{\alpha=1}^{51} A_{\alpha\beta}^{\delta}}{51} = \frac{h_{\alpha\beta}^{\delta}}{51 \times 10} \quad (22)$$

右端の平均の列から、解像度  $R_5$  から  $R_{10}$  の問合せ画像群が検索に対して有効であるといえる。もし、検索者が解像度  $R_5$  から  $R_{10}$  の画像を問合せ画像として用いるならば、問合せ画像の同種画像を 73% 以上の割合で検索することができる。

表 3 の右端の列で最も高い値 0.81 を示している 7 行目 ( $A_{\alpha 7}^{\delta}$ ) は、解像度  $R_7$  の問合せ画像群は、同種画像を 81% の割合で検索できていることを示している。また、解像度  $R_4$  と解像度  $R_9$  の問合せ画像を用いることにより、同種画像を 88% の割合で検索できていることを示している。

#### データベース登録画像の解像度

次に、データベース中に登録する画像の解像度について検討するために、表 3 を列方向で検討する。表 3 の  $\delta$  列は、解像度  $R_{\delta}$  のデータベース登録画像の (平

均的)有効性を示している。画像データベース中の解像度  $R_4$  と  $R_5$  の画像群は、解像度  $R_4$  から  $R_{10}$  の問合せ画像に対して有効であり、最も検索有効範囲が広い(75%以上のヒット率)。

表3の列  $\delta$  の平均  $A_{\bullet\bullet}^{\delta}$  は、データベースに登録された解像度  $R_{\delta}$  の画像の有効性を示す。ここで、 $A_{\bullet\bullet}^{\delta}$  は

$$A_{\bullet\bullet}^{\delta} \equiv \frac{h_{\bullet\bullet}^{\delta}}{51 \times 10} \quad (23)$$

である。

解像度  $R_3$  から  $R_6$  のデータベース中の画像のヒット率  $A_{\bullet\bullet}^{\delta}$  は、いずれも0.7以上であり、データベースに登録する画像の解像度としては適している。 $A_{\bullet\bullet}^{\delta}$  の最大値は、解像度  $R_4$  のときの0.81であった。したがって、ある1つの解像度の画像のみでデータベース利用者を満足させるためには、画像データベース中に解像度  $R_4$  の画像を登録すればよい。また、解像度  $R_4$  から  $R_{10}$  の問合せ画像で問合せを行うときにヒット率をつねに75%以上におきたい場合、画像データベースには解像度  $R_4$  の画像のみでデータベースを構築すればよい。また、解像度  $R_1$  と解像度  $R_5$  の画像を登録することにより、すべての解像度の問合せ画像に対するヒット率をつねに80%以上に保持できる。

総合ヒット率  $A_{\bullet\bullet}$  は、510枚のヒット率の平均で、大まかな検索能力の指標として考えられる。ここで、総合ヒット率は

$$A_{\bullet\bullet} \equiv \frac{\sum_{\alpha=1}^{51} \sum_{\beta=1}^{10} A_{\alpha\beta}^{\bullet\bullet}}{51 \times 10} \quad (24)$$

で表される。今回の検索実験の結果では、総合的なヒット率は0.64であった。

本章では、問合せ画像の解像度と検索結果の関係、データベース画像の解像度と検索結果の関係に着目して検索実験の考察を行った。我々の提案したテクスチャ類似検索方法は、解像度が異なる問合せ画像でも、同種のテクスチャ画像を検索ができる。すなわち、解像度に関してロバスタなテクスチャ類似検索であることが分かった。

## 6. ま と め

我々は、すでにウェーブレット分解空間を用いた例示画像検索、輪郭画像検索、スケッチ検索の検索機能を実現している<sup>3),7)</sup>。本研究では、ウェーブレット変換に基づいた対話的類似画像検索システムにテクスチャ画像検索機能を追加した。

本稿では、テクスチャのズームング(拡大・縮小)、平行移動に対して頑強なテクスチャ画像検索を実現す

るために、ウェーブレット分解空間の階層関係に着目したテクスチャ類似画像検索を提案した。

基本のアイデアは、各レベルの分解係数のみでテクスチャが表現されるのではなく、各分解レベルの階層間にもテクスチャ情報が含まれるというものである。

分解係数の階層関係を表現するために、ウェーブレット分解係数を4分木で表現し、この4分木を用いてテクスチャ特徴量を定義した。テクスチャ特徴量は、各レベル間の階層関係と子ノードの間の相関関係を表した  $4 \times 4$  の共分散行列  $S$  で定義し、その対角要素をテクスチャベクトルとして用いた。

提案したテクスチャ類似検索方式の検索能力を評価するため、51枚の生地をスキャナで読み取り、生地データベースを作成した。ただし、1枚の生地から解像度を10段階変化させてスキャナで読み取り、計510枚の画像を作成し、生地データベースを構築した。スキャナの読み取り解像度を変化させることは、ズームングと平行移動を同時に行っていることに注意されたい。

510枚からなる生地データベースに対して、データベース登録画像すべてを問合せ画像として検索を実施し評価実験を行った。問合せ画像の解像度に関する視点とデータベース登録画像の解像度に関する視点の2つの視点から実験結果を考察した。

問合せ画像に適している解像度は、解像度  $R_7$  であった。解像度  $R_7$  の問合せ画像を用いて検索を実施すれば、データベース利用者は、同種の生地画像を81%で検索できた。

データベースの登録に最も適している解像度は  $R_4$  であった。データベースに解像度  $R_4$  の画像群を登録しておけば、データベース利用者は、目的の同種生地画像を81%で検索できた。

これらの実験結果から、提案したテクスチャ類似検索方式はさまざまな解像度の画像に対しても類似テクスチャ画像の検索が可能であった。すなわち、ズームング・平行移動に対して頑強なテクスチャ検索の手法であることが分かった。

本稿では、生地テクスチャを実験に用いたが、画像全体が同一なテクスチャである(複数のテクスチャ領域を持たない)ようなテクスチャ画像であれば生地以外のテクスチャ画像にも提案したテクスチャ類似検索方式は適用可能である。

## 参 考 文 献

- 1) Flickner, M., Sawhney, H., Niblack, W., Ashiley, J., Huang, Q., Dom, B., Gorkani, M.,

- Hafner, J., Lee, D., Petkovic, D., Steele, D. and Yanker, P.: Query by Image and Video Content: The QBIC System, *IEEE Computer*, Vol.28, No.9, pp.23-32 (1995).
- 2) Gudivada, V.N. and Reghavan, V.V.: Content-Based Image Retrieval Systems, *IEEE Computer*, Vol.28, No.9, pp.18-22 (1995).
- 3) Kobayakawa, M., Hoshi, M. and Ohmori, T.: Interactive image retrieval using wavelet transform, *Proc. SCI'99/ISAS'99*, Vol.6, pp.76-85 (1999).
- 4) Mallat, S.: *A Wavelet Tour of Signal Processing*, Academic Press (1998).
- 5) 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.J-72-D-II, No.4, pp.535-544 (1989).
- 6) 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).
- 7) 小早川倫広, 星 守, 大森 匡: ウェブレット変換を用いた対話的類似画像検索とその民俗資料データへの適用, 情報処理学会論文誌, Vol.40, No.3, pp.899-991 (1999).
- 8) 栗田多喜夫, 加藤俊一, 福田郁美, 板倉あゆみ: 印象語によるデータベースの検索, 情報処理学会論文誌, Vol.33, No.11, pp.1373-1383 (1992).
- 9) 奥野忠一, 久米 均, 芳賀敏郎, 吉澤 正: 多変量解析法, 日科技連 (1971).
- 10) 大津展之, 栗田多喜夫, 関田 巖: パターン認識—理論と応用, 朝倉書店 (1996).
- 11) 串間和彦, 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).

(平成 12 年 6 月 20 日受付)

(平成 12 年 9 月 27 日採録)

(担当編集委員 牧之内 顕文)



小早川倫広

1998 年 3 月電気通信大学大学院情報システム学研究科博士前期課程修了, 現在, 電気通信大学大学院情報システム学研究科博士後期課程在学中。データ工学, データ圧縮に関心を持つ。電気学会, 電子情報通信学会各会員。



星 守(正会員)

1970 年 3 月東京大学大学院工学系研究科修士課程修了, 同年 4 月電子技術総合研究所入所, その後千葉大学工学部を経て, 1992 年 4 月より電気通信大学大学院情報システム学研究科教授, 現在に至る。アルゴリズムとデータ構造(特に探索法のための), 多変量データ解析, 数値解析に関心がある。IEEE Computer Society, IEEE Information Theory Society, Association for Computing Machinery, Intern. Association for Statistical Computing, 電子情報通信学会, 情報理論とその応用学会, 日本行動計量学会, 日本分類学会各会員。



大森 匡(正会員)

1994 年 4 月から電気通信大学に助教授として勤務。1990 年東京大学大学院情報工学専攻博士課程修了, 工学博士。データ工学, 特に並列データベース処理技術に関心を持つ。電子情報通信学会, IEEE Computer Society, Association for Computing Machinery 各会員。