

JPEG2000のビットプレーンを用いたテクスチャ画像の検索

野口達也*† 細谷 剛† 八木 秀樹†† 平澤 茂一†

† 早稲田大学大学院理工学部研究科経営システム工学専攻 〒169-8555 東京都新宿大久保3-4-1

†† 早稲田大学 メディアネットワークセンター 〒169-8050 東京都新宿区戸塚町1-104

E-mail: *†noguchi@hirasa.mgmt.waseda.ac.jp

あらまし 本論文ではテクスチャ画像を対象とした類似画像検索において、ビットプレーンの出現確率とオイラーベクトルを特徴量とした手法を提案する。JPEG2000ではビットプレーンはWavelet係数の値を2値展開し“1”の共起を同一ビット毎にまとめたものであり、復号の際に得ることができる。一方、オイラーベクトルは1枚のビットプレーンから生成されるため、JPEG2000との親和性が高い。そこで本論文では、オイラーベクトルを特徴量として追加したJPEG2000のビットプレーンを用いたテクスチャ画像の検索手法を提案し、Vis Tex データを用いた実験により検索精度が向上することを示す。

キーワード content-based image retrieval (CBIR), JPEG2000, Wavelet 変換, テクスチャ画像, ビットプレーン, オイラーベクトル。

Image Retrieval Using Bit-Plane for JPEG2000 Coded Texture Image

Tatsuya NOGUCHI*†, Gou HOSOYA†, Hideki YAGI††, and Shigeichi HIRASAWA†

† Department of Industrial and Management Systems Engineering, Major in Science and Engineering,
Waseda University, Okubo 3-4-1, Shinjuku-ku, Tokyo, 169-8555 Japan

†† Media Network Center, Waseda University, Totsuka-Machi 1-104, Shinjuku-ku, Tokyo, 169-8050 Japan.

E-mail: *†noguchi@hirasa.mgmt.waseda.ac.jp

Abstract We propose an image retrieval method that extracts features using probability and Euler vectors of wavelet coefficient's bit-plane. Each bit-plane is associated with probabilities represents the frequency of "1" occurrence, and bit-plane of wavelet coefficients can be extracted from JPEG2000 coded images. On the other hand, Euler vectors generated from a bit-plane, it has a great affinity for JPEG2000. In this paper, we propose the image retrieval method using bit-planes and the Euler vector of JPEG2000 coded texture image. We show by simulation result with using Vis Tex data set that retrieval precision of the proposed method is better than that of the conventional one.

Key words content-based image retrieval (CBIR), JPEG2000, Wavelet transform, texture image, bit-plane, Euler vector.

1. はじめに

画像の意味内容から画像を検索する技術 (CBIR: Content Based Image Retrieval) が盛んに研究されている [1]。CBIR の検索処理は特徴量抽出と類似度測定に分かれ、検索対象画像は自然画像とテクスチャ画像に大分される。

特徴量としては自然画像の場合、色、オブジェクト、回転不変量が用いられ [2]、テクスチャ画像では Wavelet 係数が多く用いられている。これは、離散 Wavelet 変換 (以下、DWT) で得られる係数がテクスチャの特徴をよ

く表す [3] ためである。そして、Wavelet 係数からより画像の特徴をつかむために、Wavelet 係数の確率分布の推定が広く研究されている。これには、サブバンド内の Wavelet 係数を一般化ガウス分布で近似する手法 [4] や、HMM と混合ガウス分布で近似する手法 [5] がある。しかしこれらの手法は特徴量を得るために画像を Wavelet 変換する必要があり、計算量が大きい。

検索に要する計算量だけでなく、画像の圧縮規格も考慮する必要がある。圧縮データから JPEG2000 の画像を検索する手法は多く提案されており、[6] では JPEG2000

の圧縮、復号過程から得られるコンテキスト情報の出現確率をビットプレーンごとに求め特徴量としている。また、[7]ではDWTにより得られたサブバンド間のビットプレーンに子孫関係を仮定しそれをまとめ、ヒストグラムを特徴量としている。どの研究もビットプレーンを利用しており、また、特徴ベクトルの距離計算にはマンハッタン距離を利用している。[8]では、圧縮データから得られるサブバンド毎のビットプレーンにおけるビット“1”の出現確率を特徴量としており、計算量が少ない。しかし、この手法は出現確率のみをビットの特徴量としているので、ビットの位置関係や散らばり具合は考慮していない。そのため検索精度は高くなく、この特徴量では画像の特徴を完全に抽出してるとは言い難い。

そこで、本研究ではビットの位置関係や散らばり具合を表すオイラーベクトル[9]を特徴量とし追加し、検索精度の向上を目指す。

2. 準備

2.1 JPEG2000 符号化方式

JPEG2000の符号化方式は、入力された画像を量子化、DWTしEBCOT符号化により符号化する。入力画像は1回のDWTにより4つのサブバンドに分割される、つまり、 L 回DWTにより $3L + 1$ つのサブバンド $1HL, 1LH, 1HH, 2HL, \dots, L \times HL, L \times LH, L \times HH$ に分割される。EBCOT符号化は、ビットプレーン単位で符号化を行いエントロピー符号化は算術符号により行なう。JPEG2000の符号化と逆符号化の流れを図1に示す。

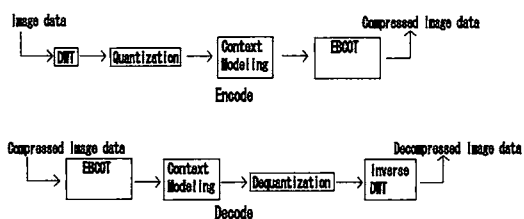


図1 JPEG2000の符号化と逆符号化

また、EBCOT符号化では、ビットプレーン中の値をその周囲の値と比較することで3つのパスに分類し、かつコンテキストを決定している。3つのパスは、注目係数とその周辺係数の関係により significance pass, magnitude pass, clean-up pass に分類される[10]。

それぞれの割当ては、significance pass においては有意である係数が周辺にあるが自身は有意でない係数に、magnitude pass においては有意である係数に、clean-up pass においてはそれら以外に割当てられる。このpassの情報と周囲の係数の状態(有意かどうか)をもとにさらに、ビットのコンテキストを決定する。

2.2 オイラーベクトルの画像検索への適用

オイラーベクトルは、いくつかのオイラー数をベクトルでまとめたものであり、これを画像の特徴量に適用したものに手法[9]がある。ここでは手法[9]に紹介されている画像のビットプレーンからオイラーベクトルを抽出する方法を示す。

この場合のオイラーベクトルは、ビットプレーンから画像を作った場合に、ビット1の集合数とそうでない部分の集合数の差を表し、オイラーベクトルの要素であるオイラー数は、画像サイズを幅 M 、高さ N 、そのビットプレーンから作られる画像を $I_k, (k = 0, 1, \dots, K - 1)$ 、 I_k 中の n 行目($n = 1, 2, \dots, N$)を I_{k_n} とすれば、以下の式により定義される。

$$E(I_k) = \sum_{n=1}^N R_n(I_k) - \sum_{n=1}^{N-1} O_n(I_k) \quad (1)$$

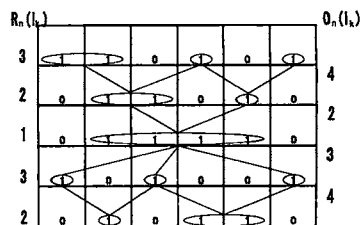
オイラー数を求める為の $R_n(I_k)$ 、 $O_n(I_k)$ は下のアルゴリズムにより決定される。また、オイラー数の計算例を図2に示す。

[$R_n(I_k)$ の計算]

1. I_k のビットの n 行目($n = 1, 2, \dots, N$)で注目するビットが1であった場合、右のビットが1かどうかみる。
2. 右のビットが1であった場合、更にその右もみる。
3. 右のビットが1でなくなったら、 $R_n(I_k)$ に1を加算する(連続する1をランと呼ぶ)。
4. n 行目の右端まで1~3を繰り返す。

[$O_n(I_k)$ の計算]

1. I_k のビットの n 行目($n = 1, 2, \dots, N - 1$)のランの画素それぞれの下3画素をみる。
2. 下3画素に $n+1$ 行目のランがあれば、 $O_n(I_k)$ に1を加算する。
3. 1,2を n 行目の全てのランで繰り返す。



$$E(I_k) = \sum R_n(I_k) - \sum O_n(I_k) = 11 - 13 = -2$$

図2 オイラー数の計算例

3. 従来手法 [8]

3.1 特徴量抽出

特徴量には、サブバンド内の係数のビットプレーンの出現確率を用いる。

ある画像 I (サイズは $M \times N$) を L 回 DWT すると $1HH, 1LH, 1HL, 2HH, \dots, L \times HH, L \times LH, L \times HL, L \times LL$ と $3L + 1$ 個のサブバンドに分解される。それぞれのサブバンドの Wavelet 係数は、レベル l ($l = 1, 2, \dots, L$) のサイズ $\frac{M}{2^l} \times \frac{N}{2^l}$ のサブバンド内の Wavelet 係数の位置を m ($m = 1, 2, \dots, \frac{M}{2^l}$), n ($n = 1, 2, \dots, \frac{N}{2^l}$) とすれば、 $w_{HHl,m,n}, w_{LHl,m,n}, w_{HLl,m,n}$ と表わされる。

また、Wavelet 係数 HHl,m,n の K 次ビットプレーンは、

$$\mathbf{b}_{HHl,m,n} = (b_{HHl,m,n,0}, \dots, b_{HHl,m,n,K-1})$$

と表される。同様に $w_{LHl,m,n}, w_{HLl,m,n}$ の K 次ビットプレーンは、

$$\mathbf{b}_{LHl,m,n} = (b_{LHl,m,n,0}, \dots, b_{LHl,m,n,K-1})$$

$$\mathbf{b}_{HLl,m,n} = (b_{HLl,m,n,0}, \dots, b_{HLl,m,n,K-1})$$

と表せる。但し $b_{HHl,m,n,K-1}, b_{LHl,m,n,K-1}, b_{HLl,m,n,K-1} \in \{0, 1\}$ である。ここで、あるサブバンド内のビットプレーンで 1 が出る出現確率をビットプレーンの次数ごとに定義すれば、各サブバンド内での特定ビットの出現確率は下式のように定義できる。

$$p_{HHl,k} = \sum_{m=1}^{\frac{M}{2^l}} \sum_{n=1}^{\frac{N}{2^l}} \frac{2^{2l} b_{HHl,m,n,k}}{MN} \quad (2)$$

$$p_{LHl,k} = \sum_{m=1}^{\frac{M}{2^l}} \sum_{n=1}^{\frac{N}{2^l}} \frac{2^{2l} b_{LHl,m,n,k}}{MN} \quad (3)$$

$$p_{HLl,k} = \sum_{m=1}^{\frac{M}{2^l}} \sum_{n=1}^{\frac{N}{2^l}} \frac{2^{2l} b_{HLl,m,n,k}}{MN} \quad (4)$$

よって、画像内全てのサブバンドでは

$$\mathbf{p}_{HHl} = (p_{HHl,0}, \dots, p_{HHl,K-1}) \quad (5)$$

$$\mathbf{p}_{LHl} = (p_{LHl,0}, \dots, p_{LHl,K-1}) \quad (6)$$

$$\mathbf{p}_{HLl} = (p_{HLl,0}, \dots, p_{HLl,K-1}) \quad (7)$$

と表せる。これを次式のようなサブバンド毎に 1 枚の画像の特徴量 I_{fea} として、検索を行う。

$$I_{fea} = (\mathbf{p}_{HH1}, \mathbf{p}_{LH1}, \mathbf{p}_{HL1}, \dots, \mathbf{p}_{HLL}) \quad (8)$$

3.2 類似度計算

式 (8) から求めた I_{fea} を質問画像 I_{que} , 非検索画像 I_{db} ごとに作成し、それらのマンハッタン距離 $L_1(I_{que}, I_{db})$ もしくはユークリッド距離 $L_2(I_{que}, I_{db})$ で類似度を計算し、各画像の検索、順位付けを行なう。

$$L_1(I_{que}, I_{db}) = \sum_{l=1}^L \left\{ \left| \mathbf{p}_{queHHl} - \mathbf{p}_{dbHHl} \right| + \left| \mathbf{p}_{queLHl} - \mathbf{p}_{dbLHl} \right| + \left| \mathbf{p}_{queHLl} - \mathbf{p}_{dbHLl} \right| \right\} \quad (9)$$

4. 提案手法

4.1 従来手法 [8] の問題点

従来手法 [8] はサブバンドごとのビットの出現確率を計算することで画像の特徴量を抽出しているが、画像の形状やビットの散らばり方の情報は含まれない。そこで、それらの特徴を表現する事ができるオイラーベクトル [9] を特徴量として新たに追加することを考える。

この際、ビットプレーンの出現確率が 0 から 1 の値をとるのに対し、オイラーベクトルはそのような定義域は取らない為、両者のマンハッタン距離を同じ比重で比較する事ができない。そこで、オイラーベクトルもビットプレーンの出現確率と同等に扱えるように 0 から 1 の値になるように正規化を行う。

4.2 提案手法

ある画像 I (サイズは $M \times N$) を L 回 DWT して得られる $1HH, 1LH, 1HL, 2HH, \dots, L \times HH, L \times LH, L \times HL, L \times LL$ の $3L + 1$ 個のサブバンドの Wavelet 係数 $w_{HHl,m,n}, w_{HLl,m,n}, w_{LHl,m,n}$ に対し、2.2 節と同様の方法によりオイラーベクトル

$$\mathbf{e}_{HHl} = (e_{HHl,0}, \dots, e_{HHl,K-1})$$

$$\mathbf{e}_{LHl} = (e_{LHl,0}, \dots, e_{LHl,K-1})$$

$$\mathbf{e}_{HLl} = (e_{HLl,0}, \dots, e_{HLl,K-1})$$

がえられる。ここで、

$$\|\mathbf{p}\| = \sum_{l=1}^L \{ |\mathbf{p}_{HHl}| + |\mathbf{p}_{HLl}| + |\mathbf{p}_{LHl}| \} \quad (10)$$

$$\|\mathbf{e}\| = \sum_{l=1}^L \{ |\mathbf{e}_{HHl}| + |\mathbf{e}_{HLl}| + |\mathbf{e}_{LHl}| \} \quad (11)$$

とし、各サブバンドのオイラーベクトルを以下のように正規化をする。

$$\mathbf{e}'_{HHl} = \left(\frac{e_{HHl,0} \|\mathbf{p}\|}{\|\mathbf{e}\|}, \dots, \frac{e_{HHl,K-1} \|\mathbf{p}\|}{\|\mathbf{e}\|} \right)$$

$$e'_{HL_i} = \left(\frac{e_{HL_{i,0}} \|p\|}{\|e\|}, \dots, \frac{e_{HL_{i,K-1}} \|p\|}{\|e\|} \right)$$

$$e'_{LH_i} = \left(\frac{e_{LH_{i,0}} \|p\|}{\|e\|}, \dots, \frac{e_{LH_{i,K-1}} \|p\|}{\|e\|} \right)$$

を得る。これらの修正されたオイラーベクトルとビットプレーンの出現確率ベクトルの連携をし、

$$P'_{HH_i} = \{P_{HH_i}, e'_{HH_i}\} \quad (12)$$

$$P'_{LH_i} = \{P_{LH_i}, e'_{LH_i}\} \quad (13)$$

$$P'_{HL_i} = \{P_{HL_i}, e'_{HL_i}\} \quad (14)$$

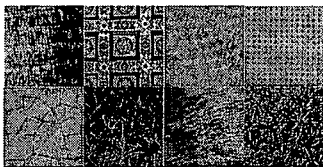
これを1枚の画像の特徴量 I_{fea} として、マンハッタン距離により検索を行う。

$$I_{fea} = (P'_{HH_1}, P'_{LH_1}, P'_{HL_1}, \dots, P'_{HL_L}) \quad (15)$$

5. 評価実験

5.1 評価用データセット

手法の有効性を確認する為に、Vis Tex データセット [11] を用いた。Vis Tex データセットは 512×512 のサイズのテクスチャ画像であり、今回はそのうちの Bark 0, 6, 8, 9, Brick 1, 4, 5, Buildings 9, Fabric 0, 4, 7, 9, 11, 14, 15, 17, 18, Flowers 5, Food 0, 5, 8, Grass 1, Leaves 8, 10, 11, 12, 16, Metal 0, 2, Misc 2, Sand 0, Stone 1, 4, Terrain 10, Tile 1, 4, 7, Water 5, Wood 1, 2 を用いた (図 3 に使用画像の一例を示す。)。1枚の画像をサイズ 128×128 の16枚に分割し計640枚とし、質問画像と被検索画像が分割もと同じ画像である場合を正解とした。



Brick0, Tile4, Fabric0, Tile7
Brick4, Leaves18, Fabric4, Food8

図 3 Vis Tex の画像例

5.2 パラメータと評価指標

DWT は Daubechies 9-7 フィルターを用いて $L = 3$ レベルまで行い、ビットプレーンは $K = 8$ ビットとした。比較する手法は、BitPlane の従来手法 (BitPlane)、オイラーベクトル (Euler)、正規化したオイラーベクトルと BitPlane を一つのベクトルにしたもの (Euler+BitPlane) の3つとする。評価指標には再現率、適合率と F 値を用い、再現率、適合率は、検索された正解画像数が1~15

枚の場合の再現率、適合率を、平均 F 値は平均再現率、平均適合率から算出した。 P を適合率、 R を再現率、 F を F 値とすれば各評価指標は以下の定義式で表せる。

$$P = \frac{N_r}{N_n} \quad (16)$$

$$R = \frac{N_r}{N_c} \quad (17)$$

$$F = \frac{2PR}{P+R} \quad (18)$$

但し、 N_r は検索された正解画像数、 N_n は正解と判定された画像数、 N_c はデータセットの全正解画像数を表す。

5.3 実験結果

上記の条件で実験を行ない、図 4 にそれぞれの手法の適合率、再現率のグラフを示す。表 1 に平均適合率、平均 F 値を示す。また、再現率が 0.33 の場合と 0.73 の場合に適合率が各手法間で向上、低下した画像の枚数を表 2, 3 に示す。表 2, 3 は、データセット中の画像に対して、行項目の適合率が列項目の適合率よりも向上、低下した画像の枚数を示す。

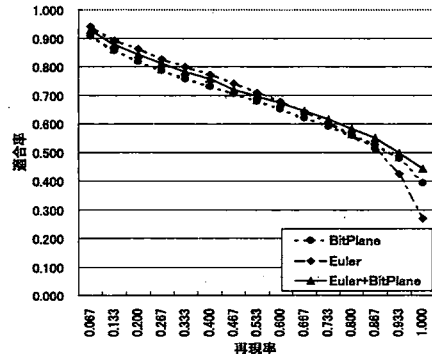


図 4 実験結果

表 1 平均適合率と平均 F 値

	BitPlane	Euler	Euler + BitPlane
平均適合率	0.671	0.683	0.694
平均 F 値	0.594	0.599	0.603

表 2 適合率の変化した画像数 (再現率 0.333)

	BitPlane		Euler	
	向上	低下	向上	低下
Euler	196	122	—	—
Euler + BitPlane	175	121	116	138

表 3 適合率の変化した画像数 (再現率 0.73)

	BitPlane		Euler	
	向上	低下	向上	低下
Euler	258	244	—	—
Euler + BitPlane	258	200	278	168

5.4 考 察

• 図4よりオイラーベクトルとビットプレーンを合せた手法は全ての再現率に対して、従来手法よりも高い適合率を示した。これにより、ビットプレーンのみよりもビットプレーンとオイラーベクトルを利用した方が高い検索精度を出すことが分かる。

• 再現率が低い場合はオイラーベクトルのみの手法 (Euler) が最も高い適合率を示しているが、表1より、平均適合率、平均 F 値が最も高かったのはオイラーベクトルとビットプレーンを合せた手法 (Euler+BitPlane) であった。これより、オイラーベクトルとビットプレーンを合せた手法は、平均的に最も良い手法であるといえる。

• 表2,3から、各手法間での適合率の向上した画像数は、低い再現率の場合は少なく、高い再現率の場合は多い。また、各手法間において適合率が向上した画像数と向上しない画像数の大小関係は、適合率の大小関係に等しい。

6. まとめと今後の課題

本論文では、JPEG2000の復号過程で得られるビットプレーンのオイラーベクトルを利用することでビットの出現確率を特徴量とする手法よりも高い性能を示した。また、オイラーベクトルを正規化しそれらを1つの特徴量ベクトルとした場合は、オイラーベクトルのみを特徴量にしたよりも、平均的に高い性能を示した。

今後の課題として、ビットプレーンの出現確率とオイラーベクトルをより有効な結果が得られるように組み合わせる方法、また、オイラーベクトルの正規化の方法、正規化したオイラーベクトルの統計的性質の解明と利用等も今後の課題である。

文 献

- [1] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Recognit.*, vol. 22, pp. 1349-1380, Dec. 2000.
- [2] Chiou-Shann Fuh, Shun-Wen Cho, and K. Essig, "Hierarchical Color Image Region Segmentation for Content-Based Image Retrieval System," *IEEE Trans. Image Processing*, vol. 9, pp. 156-162, Jan. 2000.
- [3] G. V. Wouwer, P. Scheunders, and D. V. Dyck, "Statistical texture characterization from discrete wavelet representations," *IEEE Trans. Image Processing*, vol. 8, pp. 592-598, Apr. 1999.
- [4] M. N. Do, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *IEEE Trans. Image Processing*, vol. 11, pp. 146-158, Feb. 2002.
- [5] M. N. Do and M. Vetterli, "Rotation invariant texture characterization and retrieval using steerable wavelet-domain hidden Markov models," *IEEE Trans. Multimedia*, vol. 4, pp. 517-527, Dec. 2002.
- [6] Ha-Joong Park and Ho-Youl Jung, "JPEG-2000 compressed image retrieval using partial entropy coding," *Lecture Notes in Computer Science*, vol. 4105, pp. 410-417, Sep. 2006.
- [7] J. Tang, W. Zhang, and C. Li, "An approach to compressed image retrieval based on JPEG2000 framework," *Lecture Notes in Computer Science*, vol. 3584, pp. 391-399, Aug. 2005.
- [8] M. H. Pi, C. S. Tong, S. K. Choy, and H. Zhang, "A fast and effective model for wavelet subband histograms and its application in texture image retrieval," *IEEE Trans. Image Processing*, vol. 15, pp. 3078-3088, Oct. 2006.
- [9] A. Bishnu, B. B. Bhattacharya, M. K. Kundu, C. A. Murthy, and T. Acharya, "Euler vector for search and retrieval of gray-tone images," *IEEE Trans. Systems, Man, and Cybernetic*, vol. 35, pp. 801-812, Aug. 2005.
- [10] Chung-Jr Lian, Kuan-Fu Chen, Hong-Hui Chen, and Liang-Gee Chen, "Analysis and architecture design of block-coding engine for EBCOT in JPEG 2000," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, pp. 219-230, Mar. 2003.
- [11] MIT Vision and Modeling Group. Vision Texture. [Online]. Available: <http://vismod.www.media.mit.edu>