

Image Retrieval Method Utilizing Texture Information Derived from Discrete Wavelet Transformation Together with Color Information

Kohei Arai¹

¹ Graduate School of Science and Engineering
Saga University
Saga City, Japan

Cahya Ragmad²

² State Polytechnics of Malang,
East Java, 65145 Indonesia

Abstract—Image retrieval method utilizing texture information which is derived from Discrete Wavelet Transformation: DWT together with color information is proposed. One of the specific features of the texture information extracted from portions of image is based on Dyadic wavelet transformation with forming texture feature vector by using energy derived from Gabor transform on 7 by 7 pixel neighbor of significant points. Using the Wang's dataset, the proposed method is evaluated with retrieval success rate (precision and recall) as well as Euclidian distance between the image in concern (Query image) and the other images in the database of interest and is compared to the other method. As the result through the experiments, it is found that the DWT derived texture information is significantly effective in comparison to the color information.

Keywords—Wavelet; DWT; local feature; Color; Texture; CBIR

I. INTRODUCTION

Henning Müller, et al. presented the paper which is entitled “Benchmarking image retrieval applications” [1]. In their paper, almost all the methods for image retrievals are tested and compared. One of the conclusions is Content Based Image Retrieval: CBIR is the most effective. Brahmi et al. mentioned the two drawbacks in the keyword annotation image retrieval. First, images are not always annotated and the manual annotation expensive also time consuming. Second, human annotation is not objective the same image may be annotated differently by different observers [2]. CBIR based image retrievals is based on the similarity of feature vector of color, texture, shape and other image content between a query image in concern and the images in a database of interest. Also, CBIR based image retrievals is objectively in comparison to the conventional image retrieval methods [3].

Although color and texture information is significant for image retrieval in particular, color and texture are different by portion by portion of whole image [4]. Therefore, local properties of color and texture information have to be used for image retrievals are commonly used to describe the image content in image.

There are so many methods for extraction of texture information from the local images. The proposed method uses 2D Discrete Wavelet Transform: DWT with Haar base

function, combined with the two high sub-band frequencies to make significant points and edge, for choosing any part of image by threshold the high coefficient value. Although, the image size of frequency components derived from the conventional DWT is decimated with the factor of 2 by 2, a Dyadic wavelet transformation maintains the image size of frequency component which represent texture information. Therefore, this paper proposes a method for texture information extraction based on the Dyadic wavelet transformation.

This paper is organized as follows: In the following section, Discrete Wavelet Transform: DWT is briefly described. Then, color and texture information is explained followed by the proposed method for image retrieval. After that, similarity and performance measure is discussed. Then the texture information extraction method with Dyadic wavelet transformation is proposed followed by experimental results. Finally, conclusion is given together with some discussions.

II. PROPOSED METHOD

A. Discrete Wavelet Transform

One of the specific features of the DWT is time and frequency analysis [5]. Namely, not only frequency component analysis, but also frequency component analysis at the time in concern is available for the DWT. Although the well-known Fourier analysis allows only frequency component analysis, the DWT makes frequency component analysis at the time in concern. The DWT represents an image as a sum of wavelet functions at different locations (shift) and scales [6].

The DWT is based on a base function (mother wavelet function). If the base function is bi-orthogonal function, then the original image before the DWT can be reconstructed through Inverse DWT: IDWT. Any decomposition of an 2D image into wavelet involves four sub-band elements representing LL (Approximation), HL (Vertical Detail), LH (Horizontal Detail), and HH (Detail), respectively as shown in Fig.1.

The DWT works as a filter bank and the functions of the filter bank are explained as follows,

On a one dimensional signal $x[n]$. $x[n]$ is input signal that contains high frequencies and low frequencies. $h[k]$ and $g[k]$ is channel filter bank involving sub-sampling as shown in Fig.2. $c[n]$ is called averages contains low frequencies signal. $d[n]$ is called wavelet coefficients contain high frequencies signal. $c[n]$ and $d[n]$ be sub sampled (decimated by 2: $\downarrow 2$) the next process for further decomposition is iterated on the low signal $c[n]$.

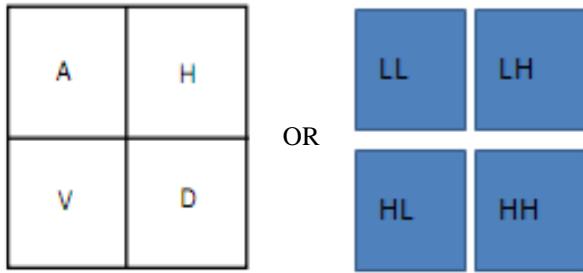


Fig. 1. Level 1 of 2D DWT

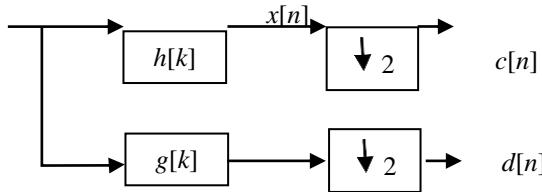


Fig. 2. Two channel filter bank

B. Example 1D Haar wavelet decomposition

Let the $x[n]$ as an input $x[n] = X_0, X_1, X_2, \dots, X_{N-1}$ contains N elements the output is $N/2$ elements of averages and is stored in $c[n]$. Also, it assumed to be containing $N/2$ elements wavelet coefficients values and is stored in $d[n]$. The Haar equation to calculate an average AV_i in equation (1) and a wavelet coefficient WC_i in equation (2) from pair data odd and even element in the input data are:

$$AV_i = \frac{X_i + X_{i+1}}{2} \quad (1)$$

$$WC_i = \frac{X_i - X_{i+1}}{2} \quad (2)$$

C. Wavelet Multi-Resolution Analysis

One dimensional wavelet transformation is expressed with the following equation.

$$F = C_n f \quad (3)$$

where F, f denotes wavelet frequency component and radar echo signal as a function of time. C_n denotes wavelet transformation matrix which is expressed as a bi-orthogonal function based on base functions. C_n can be determined with a reference to the appendix. Therefore,

$$C_n C_n^T = I \quad (4)$$

Then, f is converted to the followings,

$$F_I = (L_I, H_I),$$

$$F_2 = C_n L_I = (L_2, H_2),$$

$$F_3 = C_n L_2 = (L_3, H_3),$$

...

$$F_m = C_n L_m^{-1} = (L_m, H_m) \quad (5)$$

Also f is reconstructed as follows,

$$C_n^{-1} F_m = C_n^{-1} (L_m, H_m) = L_m^{-1}, \dots, C_n^{-1} F_2 = L_I, C_n^{-1} F_I = f \quad (6)$$

The suffix of 1 to m is called “level”. Level m implies that wavelet transformation is applied m times. Multi-resolution Analysis: MRA ensure that the original signal can be reconstructed with the wavelet coefficients or frequency components of the level m . The frequency components derived from MRA are corresponding to the level m . Therefore, MRA does work as a filter bank.

There are some base functions such as Haar, Daubechies¹, etc. In order to conduct a preliminary experiment, Haar base function is selected for showing an effectiveness of the DWT for texture information extractions.

D. Texture

Texture is pattern, tracery, crease patter, etc. with intensities and colors. Texture contains repeating pattern of local variations in image intensity also an area that can be perceived as being spatially homogeneous. Texture provides important characteristics for surface and object identification. Texture information can be extracted from the original image and are typical features that can be used for image retrievals [7]. The texture is characterized with the statistical distribution such as probability density functions, co-occurrence matrices, etc. of the image intensity. It also may extracted by using energy of Gabor filter on 7 by 7 pixel neighbor of significant points.

E. Color

Color is visual perceptual property corresponding in human to the category called red (600 nm of wavelength), green (550 nm), blue (450 nm), black, yellow, etc. Color is produced by spectra of light that absorbed or reflected then received by the human eye and processed by the human brain. To extract the color feature, the first order statistical moments and the second order statistical moments in the Hue, Saturation, and Intensity: HSI color space on neighbor of significant points with size 3 by 3 pixels is used. HSI color space is similar to human perception color system so we used it to extract the color feature.

The first order statistical moment is as follows,

$$\mu = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N p(i, j) \quad (7)$$

where:

¹ Daubechies base function is defined as $\{\alpha_k\}$ satisfying the following conditions,

$$\phi(x) = \sum_k \alpha_k \sqrt{2} \phi(2x - k)$$

$$\beta_k = (-1)^k \alpha_{1-k}$$

$$\varphi(x) = \sum_k \beta_k \sqrt{2} \phi(2x - k)$$

P = Pixel value

MN = Size of significant points and its neighbor.

The second order statistical moment is as follows,

$$\sigma = \sqrt{\frac{1}{MxN} \sum_{i=1}^M \sum_{j=1}^N (p(i,j) - \mu)^2} \quad (8)$$

where:

P = pixel value

MN = Size of significant points and its neighbor

F. Propose Method for Image Retrieval

The procedure proposed here is as follows,

- 1) Read Query image and Convert from RGB image to gray image and HSI image.
- 2) Decomposition using wavelet transformation.
- 3) Make absolute for every wavelet coefficient,

$$WC_{new} = | WC_{old} |.$$
- 4) Combine Vertical Detail and Horizontal Detail,

$$CV_d H_d(i,j) = \text{Max}(V_d(i,j), H_d(i,j)).$$
- 5) Choose significant points on $CV_d H_d(i,j)$ by threshold the high value.
- 6) Choose points on HSI image and it neighbor (3 by 3 pixel) base on coordinate significant points on $CV_d H_d(i,j)$ then Forming color feature vector by using The first order statistical moment and the second order statistical moment.
- 7) Forming texture feature vector by using Gabor transform on 7 by 7 pixel neighbor of significant points.
- 8) Implement “MinMax” normalization on all feature vector with range [0 1].
- 9) Measure the distance between feature vector image query and feature vector image by using Euclidean distance.
- 10) Showing image with X top ranking from the dataset.

G. Similarity

The similarity is measured by using Euclidean distance between features of the query image in concern and the images in the dataset of interest. The retrieved result is not a single image but a list of image ranked with their similarity. The feature representation is image feature refer to the characteristics which are described with the contents of an image.

$FQ = (Q_1, Q_2, \dots, Q_n)$

$FD = (D_1, D_2, \dots, D_n)$

$$\text{dis}(FQ, FD) = \sqrt{\sum_{j=1}^n (Q_j - D_j)^2} \quad (9)$$

where:

FQ = Feature vector of query image.

FD = Feature vector of image in data set

n = The number element of feature vector

If the Euclidean distance between features of the query image in concern and the images in the dataset of interest is short, then it is to be considered as similar features. For example, a distance of 0 have meaning an exact match with

the query and 1 mean totally different. Base on the rank of that the measured similarity, the retrieval results are to be displayed.

H. Performance Measure

There is two-step approach to retrieve the relevant image from the dataset of interest. First, CBIR method extracts a feature vector for every portion of the image in concern and put the feature into the feature dataset. Second, the user provides a query image and a feature vector of the query image is extracted. Then the feature vector of query image is compared to the feature vectors of the portion of images in the database of interest for finding the feature vector which shows a shortest distance between two feature vectors.

The performance of CBIR system of the aforementioned method is evaluated for showing image with X top ranking from the dataset. It is a common way to evaluate the performance of the CBIR system based on the well-known precision and recall evaluations. Precision measures the retrieval accuracies. It is a ratio between the number of relevant images retrieved and the total number of images retrieved. Recall measures the ability of retrieving all relevant images in the dataset. It is a ratio between the number of relevant images retrieved and the whole relevant images in the dataset.

The performance of the CBIR system is calculated as follows,

$$\text{Precision} = \frac{NRRI}{XR} \quad (10)$$

where:

$NRRI$ = Number of relevant retrieved images

XR = X Top ranking of retrieved images

$$\text{Recall} = \frac{NRRI}{TR} \quad (11)$$

where:

TR = Total number of relevant images in dataset

III. EXPERIMENTS

A. Performance Measure

The Wang's dataset [8] is used for the experiments. The retrieved results are compared with the standard system SIMPLICITY, FIRM and also color salient points by using the same dataset [8] [9] [10]. The proposed method is developed with MATLAB tool for the evaluation of performance matrices.

Fig.3 shows the query image and relevant image as a result of the proposed method and system. The relevant image results for “Bus”, “Dinosaur”, “Elephant”, “Flower”, and “Horse” as a query image are #6, #10, #7, #8 and #7, respectively. It seems that “Horse” and “Elephant” are very similar due to the fact that the backgrounds of these images are resembled.

Query Image	Result									
352	335	786	366	396	923	360	928	978	393	310
499	493	439	495	449	489	482	424	442	436	466
527	997	553	561	570	130	565	574	514	869	584
688	659	622	660	629	613	612	621	334	236	648
793	791	765	796	755	34	763	805	711	162	718

Fig. 3. Example results for the bus, dinosaur, elephant, flower and horses query

In the Table 1, the comparison of the averaged precision results between the proposed method and other method show that the proposed method is superior to the other conventional methods for most of the cases. The proposed method achieved the improvements of 12%, 17%, 11%, in comparison to the Firm, Simplicity, and Color Salient Points, respectively.

TABLE I. AVERAGE PRECISION RESULTS

Category	Firm	Simplicity	Color salient points	Proposed method
Bus	0.60	0.36	0.52	0.68
Dinosaur	0.95	0.95	0.95	0.94
Elephant	0.25	0.38	0.40	0.60
Flower	0.65	0.42	0.60	0.75
Horses	0.65	0.72	0.70	0.71
Average	0.62	0.57	0.63	0.74

B. Dyadic Wavelet Utilized Method

Another example of image database which includes 24 of species of plankton is shown in Fig. 4. In the figure, there are 24 of color images. Process flow of the proposed image retrievals is shown in Fig. 5. In this case, Dyadic wavelet transformation is used for texture information extraction. Although the conventional DWT makes a down-sampling

(from 1 to 1/4), the input and output image size are same for the Dyadic wavelet transformation. Therefore, much texture information can be extracted with Dyadic wavelet transformation rather than the conventional DWT.



Fig. 4. Another example of image database which includes 24 of species of plankton

Dyadic wavelet transformation can be expressed as follows,

$$C_n[i] = \sum_k h[k] C_{n-1}[i + k2^{n-1}] \quad (12)$$

$$d_n[i] = \sum_k g[k] C_{n-1}[i + k2^{n-1}] \quad (13)$$

where C_n d_n denotes low and high frequency components, respectively. Inverse Dyadic wavelet transformation is represented as follows,

$$C_{n-1}[i] = \frac{1}{2} \left(\sum_k h[k] C_n[i - k2^{n-1}] + \sum_k g[k] d_n[i - k2^{n-1}] \right)$$

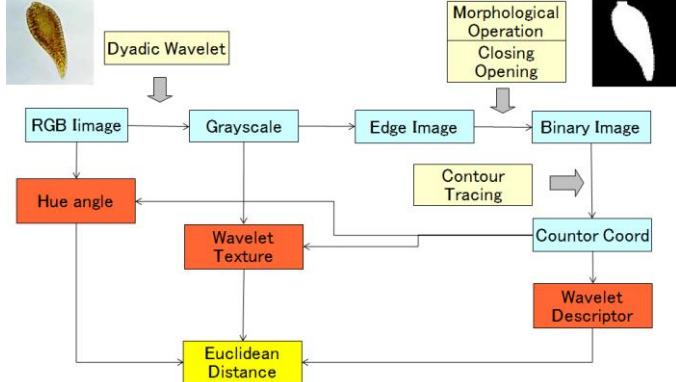


Fig. 5. Process flow of the proposed image retrievals

On the other hand, color information is represented with the following Hue angle, Saturation and Intensity: HSI.

Hue angle:

$$H = \begin{cases} 60 \frac{G-B}{MAX-MIN} + 0 & (if MAX = R) \\ 60 \frac{B-R}{MAX-MIN} + 120 & (if MAX = G) \\ 60 \frac{R-G}{MAX-MIN} + 240 & (if MAX = B) \end{cases} \quad (14)$$

Saturation:

$$S = MAX - MIN \quad (15)$$

Intensity:

$$I=MAX \quad (16)$$

If the image number 3 in the Fig.4 is query image, then the Euclidian distance between the query image and the image of interest calculated. Table 2 shows the distance for the cases of which only texture information is taken into account ($D(t)$) and only hue information is used ($D(h)$). The difference is quite obvious that the distance between the query image and the other image of interest for using texture derived from Dyadic wavelet transformation is much longer than that for hue information. Also, it is quite obvious that the extracted texture information with Dyadic wavelet transformation is greater than that with the conventional DWT due to the fact that much detail frequency components are maintained for Dyadic wavelet transformation rather than that for the conventional DWT. Therefore, it may concluded that the followings,

- 1) Texture information is greater than color information,
- 2) Texture information can be extracted with Dyadic wavelet transformation much effectively than the conventional DWT
- 3) Forming texture feature vector by using energy derived from Gabor transform on 7 by 7 pixel neighbor of significant points is also effective for texture information extractions.

TABLE II. EUCLIDIAN DISTANCE BETWEEN THE QUERY IMAGE AND THE IMAGE OF INTEREST

Plankton	$D(t)$	$D(h)$
a.catenella1cell	0.751	0.428
a.catenella4cell	0.143	2.842
c.antiqua	0.0	0.0
c.antiqua2	1.561	1.096
c.antiqua3	1.926	0.053
c.furca	6.563	0.710
c.marina	1.802	0.060
c.polykrikoides2cell	1.915	0.256
c.polykrikoides8cell	0.560	0.121
d.fortii	6.611	0.506
g.catenatum1cell	1.069	0.013
g.catenatum5cell	3.543	0.073
g.instriatum	0.490	0.468
g.mikimotoi	0.785	1.461
g.polygramma	2.124	0.115
g.sanguineum	4.948	0.211
h.akashiwo	1.295	0.186
h.circularisquama	1.455	0.022
m.rubrum	2.396	0.576
n.scintillans4	1.037	1.3676
n.scintillans5	2.764	0.6447
p.dentatum	3.882	0.401
p.dentatum2	4.732	0.114
p.signoides	3.476	0.132

IV. CONCLUSION

Image retrieval method utilizing texture information which is derived from Discrete Wavelet Transformation: DWT together with color information is proposed. One of the specific features of the texture information extracted from portions of image is based on Dyadic wavelet transformation with forming texture feature vector by using energy derived from Gabor transform on 7 by 7 pixel neighbor of significant points. Using the Wang's dataset, the proposed method is evaluated with retrieval success rate (precision and recall) as well as Euclidian distance between the image in concern (Query image) and the other images in the database of interest and is compared to the other method. As the result through the experiments, it is found that the DWT derived texture information is significantly effective in comparison to the color information.

Through the experiments with the Wang's image dataset, it is found that the proposed method is superior to the other conventional methods for most of the cases. The proposed method achieved the improvements of 12%, 17%, 11%, in comparison to the Firm, Simplicity, and Color Salient Points, respectively. Also, it is found that texture information is greater than color information. Furthermore, it is found that texture information can be extracted with Dyadic wavelet transformation much effectively than the conventional DWT. Moreover, it may concluded that the forming texture feature vector by using energy derived from Gabor transform on 7 by 7 pixel neighbor of significant points is also effective for texture information extractions.

Further investigations are required for utilization of relational information, shape information, and so on. Also, it is desirable to conduct a research on image retrieval based on deep learning.

REFERENCES

- [1] Henning M'uller, Antoine Geissbuhler, et al. "Benchmarking Image Retrieval Applications". San Francisco, CA : In The Seventh International Conference, September 2004. pp. 334–337.
- [2] Kherfi, M., Brahmi, D. and Ziou D. "Combining Visual Features with Semantics for a More Effective Image Retrieval". : in ICPR '04, vol. 2, 2004. pp. 961–964.
- [3] H. Yu, M. Li, H.-J. Zhang and Feng, J. "Color texture moments for content-based image retrieval". : In International Conference on Image Processing, 2002. pp. 24–28.
- [4] N. Sebe, Q. Tian, E. Loupias, M. Lew and T. Huang. "Evaluation of salient point techniques". s.l. : In International Conference, 2004.
- [5] Stephane Mallet. "Wavelets for a Vision". : Proceeding to the IEEE, Vol. 84, 1996. pp. 604-685.
- [6] I.Daubechies. "Ten lecturer on wavelet". Philadelphia, PA:Sosity for Industrial and Applied Mathematics Analysis, vol. 23, Nov. 1992. pp. 1544–1576.
- [7] Kohei Arai and Y.Yamada. "Image retrieval method based on hue information and wavelet description based shape information as well as texture information of the objects extracted with dyadic wavelet transformation". s.l. : 11th Asian Symposium on Visualization, NIIGATA, JAPAN, 2011.
- [8] <http://wang.ist.psu.edu>.
- [9] Hiremath P.S and Jagadeesh Pujari. "Content Based Image Retrieval using Color Boosted Salient Points and Shape features of an image". : International Journal of Image Processing (IJIP) Vol.2, Issue 1, January–February 2008. pp. 10-17.
- [10] J. Li, J.Z. Wang, and G. Wiederhold. "IRM: Integrated Region Matching for Image Retrieval". : in Proc. of the 8th ACM Int. Conf. on Multimedia, Oct. 2000. pp. 147-156.

AUTHORS PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 and also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Commission-A of ICSU/COSPAR since 2008. He received Science and Engineering Award of the year 2014 from the minister of the ministry of Science Education of Japan and also received the Best Paper Award of the year 2012 of IJACSA from Science and Information Organization: SAI. In 2016, he also received Vikram Sarabhai Medal of ICSU/COSPAR and also received 20 awards. He wrote 34 books and published 520 journal papers. He is Editor-in-Chief of International Journal of Advanced Computer Science and Applications as well as International Journal of Intelligent Systems and Applications. <http://teagis.ip.is.saga-u.ac.jp/>

Cahya Rahmad, He received BS from Brawijaya University indonesia in 1998 and MS degrees from Informatics engineering at Tenth of November Institute of Technology indonesia in 2005. he is a lecturer in The State Polytechnic of Malang Since 2005 also a doctoral student at Saga University Japan Since 2010. His interest research are image processing, data mining and patterns recognition.