

# Wavelet Based Image Retrieval Method

Kohei Arai

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

Cahya Rahmad

Electronic Engineering Department  
The State Polytechnics of Malang,  
East Java, Indonesia

**Abstract**—A novel method for retrieving image based on color and texture extraction is proposed for improving the accuracy. In this research, we develop a novel image retrieval method based on wavelet transformation to extract the local feature of an image, the local feature consist color feature and texture feature. Once an image taking into account, we transform it using wavelet transformation to four sub band frequency images. It consists of image with low frequency which most same with the source called approximation (LL), image containing high frequency called horizontal detail (LH), image containing high frequency called vertical detail (HL), and image containing horizontal and vertical detail (HH). In order to enhance the texture and strong edge, we combine the vertical and horizontal detail to be other matrix. The next step is we estimate the important point called significant point by threshold the high value. After the significant points have been extracted from image, the coordinate of significant points will be used for knowing the most important information from the image and convert into small regions. Based on these significant point coordinates, we extract the image texture and color locally. The experimental results demonstrate that our method on standard dataset are encouraging and outperform the other existing methods, improved around 11 %.

**Keywords**-component; Image retrieval; DWT; Wavelet; Local feature; Color; Texture.

## I. INTRODUCTION

Image retrieval has been used to seek an image over thousand database images. In the web based search engine, the image retrieval has been used for searching an image based on text input or image. Once an input taking into account, the method will search most related image to the input. The correlation between input and output has been defined by specific role. With expansion in the multimedia technologies and the Internet, CBIR has been an active research topic since the first 1990's. The concept of content based retrieval (CBR) in image start from the first 1980s and serious applications started in the first 1990s. Retrieval from databases with a large number of images has attracted considerable attention from the computer vision and pattern recognition society.

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 [1]. Unlike the traditional approach that using the keyword annotation as a method to search images, CBIR system performs retrieval based on the similarity feature vector of color, texture, shape and other image content.

Comparing to the traditional systems, the CBIR systems perform retrieval more objectiveness [2]. A very basic issue in designing a CBIR system is to select the most effective image features to represent image contents (3). Global features related to color or texture are commonly used to describe the image content in image retrieval. The problem using global features is this method cannot capture all parts of the image having different characteristics [4].

In order to capture specific parts of the image the local feature is used. The proposed method uses 2D Discrete wavelet transform with Haar base function, combined the two high sub-band frequency to make significant points and edge then estimate the important point called significant point by threshold the high value. After the significant points have been extracted from image, the coordinate of significant points will be used for knowing the most important information from the image and convert into small regions. Based on these significant point coordinates, and then extract the image texture and color texture locally.

## II. PROPOSED METHOD

### A. Wavelet Transformation

The wavelet representation gives information about the variations in the image at different scales. Discrete Wavelet Transform (DWT) represents an image as a sum of wavelet functions with different locations (shift) and scales [5]. Wavelet is the multi-resolution analysis of an image and it is proved that having the signal of both space and frequency domain [6]. Any decomposition of an 1D image into wavelet involves a pair of waveforms: the high frequency components are corresponding to the detailed parts of an image while the low frequency components are corresponding to the smooth parts of an image.

DWT for an image as a 2D signal can be derived from a 1D DWT, implement 1D DWT to every rows then implement 1D DWT to every column. 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.

The wavelet transform may be seen as a filter bank and illustrated as follow, 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.  $c[n]$  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  $c[n]$ .

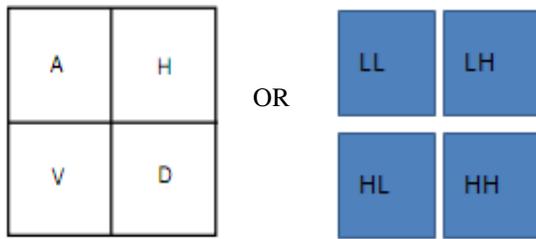


Figure 1. Level 1 of 2D DWT

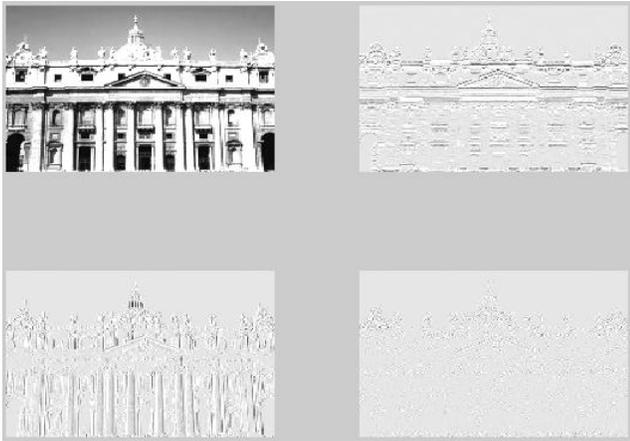


Figure 2. Example Level 1 of 2D DWT

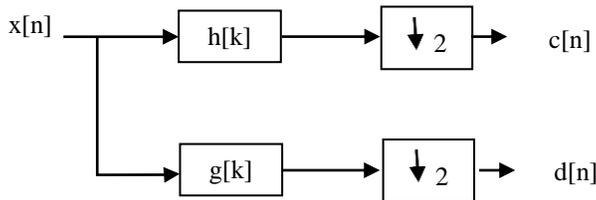


Figure 3. Two channel filter bank

For example, 1D Haar wavelet decomposition is expressed as follows, let  $x[n]$  be an input,  $x[n]= X_0, X_1, X_2, \dots, X_{N-1}$  which contains  $N$  elements. Then output will consist of  $N/2$  elements of averages over the input and is stored in  $c[n]$ . Also the other output contains  $N/2$  elements wavelet coefficients values and is stored in  $d[n]$ . The Haar equation to calculate an average  $AV_i$  (See Eq.1) and a wavelet coefficient  $WC_i$  (See Eq.2) from pair data odd and even element in the input data are:

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

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

where:

AV = Average  
WC = Wavelet coefficient

### B. Color and Texture

Texture contain 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 extracted from the original image is typical features for image retrievals [7]. The texture is characterized by the statistical distribution of the image intensity using energy of Gabor filter on  $7 \times 7$  pixels. Color is produced by spectrum 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 (See Eq.(3)) and the second order statistical moments (See Eq.(4)) HSV color space is similar to human perception color system so we used it to extract the color feature in the HSV color space on neighbor of significant points with size  $3 \times 3$  pixels.

The first order statistical moments is expressed as follows,

$$\mu = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N p(i, j) \quad (3)$$

where:

P = Pixel value  
MxN = Size of significant points and its neighbor.

The second order statistical moments is represented as follows,

$$\sigma = \sqrt{\frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (p(i, j) - \mu)^2} \quad (4)$$

where:

P = pixel value  
 $\mu$  = The first order statistical moments value  
MxN = Size of significant points and its neighbor

### C. Image Retrieval Algorithm

The proposed image retrieval algorithm is as follows,

1. Read Query image and Convert from RGB image to gray image and HSV image then Decomposition using wavelet transformation.
2. Make absolute for every Wavelet coefficients,  $WC_{new} = |WC_{old}|$ .
3. Combine Vertical Detail and Horizontal Detail,  $CVdHd(i, j) = \text{Max}(Vd(i, j), Hd(i, j))$ .
4. Choose significant points on  $CVdHd(i, j)$  by threshold the high value.
5. Choose points on HSV image and it neighbor ( $3 \times 3$  pixel) base on coordinate significant points on  $CVdHd(i, j)$  then Forming color feature vector by using The first order statistical moment and the second order statistical moment.
6. Forming texture feature vector by using Gabor transform on  $7 \times 7$  pixel neighbor of significant points and Implement min/max normalization on all feature vector with range  $[0, 1]$ .
7. Measure the distance between feature vector image query and feature vector image in the dataset by using

Euclidean distance then display image results with X top ranking from the dataset.

### III. EXPERIMENTS

The retrieval result is not a single image but a list of image ranked by their similarity. The similarity measure is computed by using Euclidean distance (See Eq.(5)) between feature representation of image query and feature representation of image in dataset. The feature representation is image feature refer to the characteristics which describe the contents of an image.

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

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

$$dis(FQ, FD) = \sqrt{\sum_{j=1}^n (Q_j - D_j)^2} \quad (5)$$

where :

- FQ = Feature vector of query image.
- FD = Feature vector of image in data set
- n = Number element of feature vector

If the distance between feature representation of image query and feature representation of image in dataset small then it to be considered as similar.

The performance of the CBIR system is calculated by showing image a number of top ranking from the dataset. We used precision and recall to evaluate the performance of the CBIR system. Precision measures the retrieval accuracy; it is ratio between the number of relevant images retrieved and the total number of images retrieved (See Eq.(6)). Recall measures the ability of retrieving all relevant images in the dataset. It is ratio between the number of relevant images retrieved and the whole relevant images in the dataset (See Eq.(7)).

The performance of the CBIR system is calculated by using equation (6) and equation (7).

$$\text{Precision} = \frac{NRRI}{XR} \quad (6)$$

where :

- NRRI = Number of relevant retrieved images
- XR = X Top ranking of retrieved images

$$\text{Recall} = \frac{NRRI}{TR} \quad (7)$$

where:

- NRRI = Number of relevant retrieved images
- TR = Total number of relevant images in dataset

We used The Wang's dataset [8] To evaluate the effectiveness of our approach and compared with the standard system SIMPLICITY, FIRM and also Color salient points by using the same dataset [8],[9],[10].

Figure 4, 5, and 6 shows query image and the retrieved image. Query image number is shown in the query image while retrieved image number is also indicated in the retrieved image. Figure 4 shows relatively good retrieval accuracy while Figure 5 and 6 for flower and horses shows strange retrieved results. Image number 236 and 334 for flower and image number 34 for horses are not correct.

The comparison average precision results between the proposed method and other method is showed in the Table 1 and the result by using the proposed method is improve. The improvement compared other method are 17%, 12%, 11%, respectively.

Image retrieval experiments with five query images including the aforementioned three query images are conducted. Figure 7 shows the query image and relevance image as results of this system. The relevance image results for bus, dinosaur, elephant, flower, and horse as a query image are 6, 10, 7, 8 and 7, respectively.

Table 2 shows the comparison average precision results between the proposed method and other methods. It shows that ability of system to retrieve relevance image from image in dataset is improve. The improvement compared other method are 12%, 17%, 11%, respectively. The graphic comparison of average precision can be seen in Figure 8.

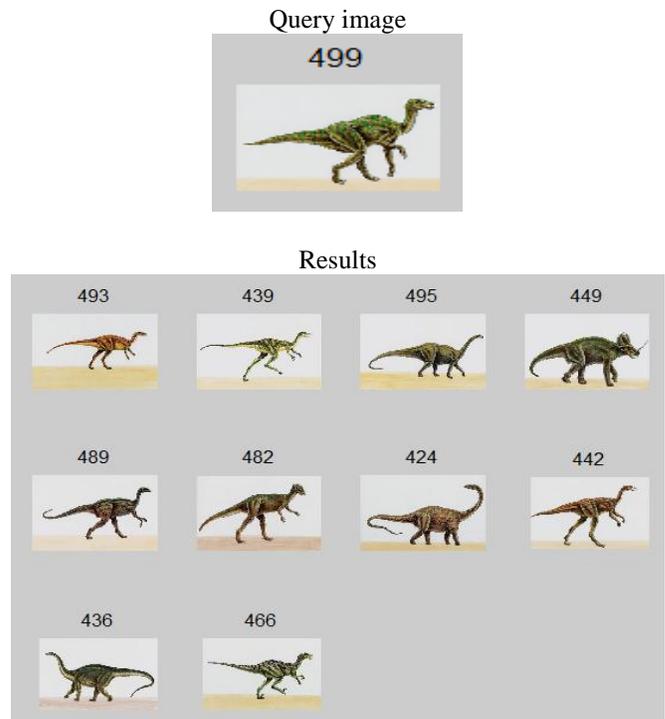
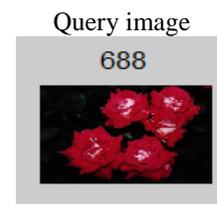


Figure 4 Example results for the dinosaur as query image



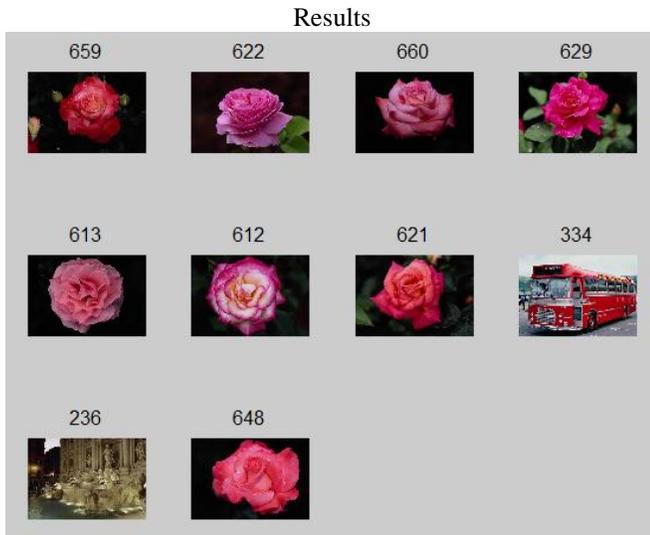


Figure 5. Example results for the flower as query image

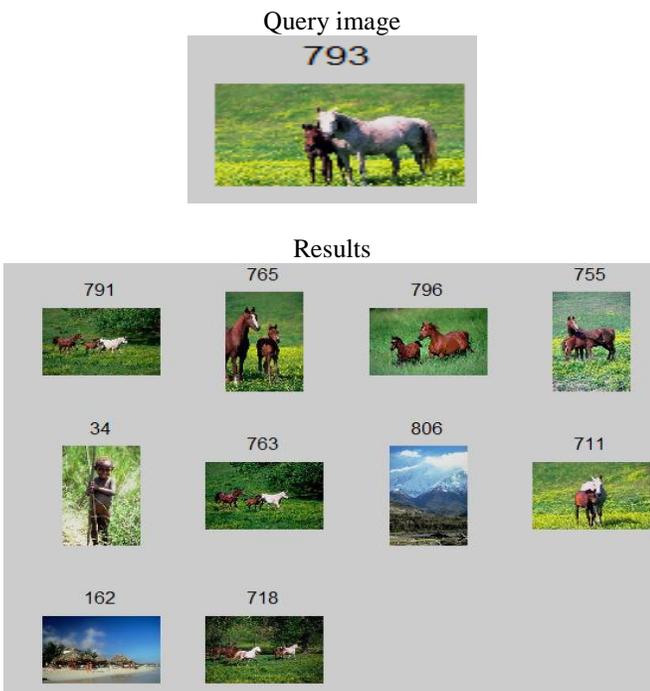


Figure 6. Example results for the horses as query image

TABLE 1. Comparison average precision results between the proposed method with other method

Method	Average Precision results
Simplicity	57%
Firm	62%
Color salient points	63%
Proposed method	74%

Query image number 2 shows almost same image retrieval accuracy and is more than 0.9 so that it is easy to retrieve this image. Meanwhile query image number 3 shows much poor

image retrieval accuracy below 40% for the conventional methods with Firm, Simplicity, and Color salient points of gradient vector while the proposed method shows relatively high accuracy.

TABLE 2. 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

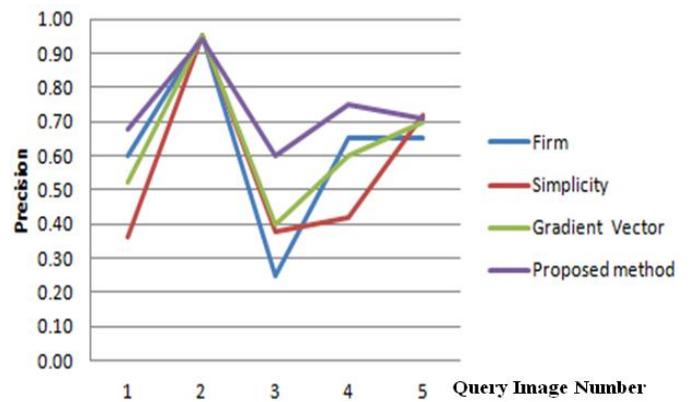


Figure 8. Comparison of average precision

Such this image need time-frequency components of image feature for image retrievals. The difference between the image retrieval accuracy of the proposed method and the conventional methods is around 20%, significant difference. On the other hand, both of spatial and color features are required for image retrievals of the query image number 5. The image retrieval accuracy of the conventional method with Simplicity and Color salient points of gradient vector is almost same as that of the proposed method so that these features work for image retrievals for this image.

Figure 9 shows the relation between image retrieval accuracy of the proposed method and those of the conventional methods with Firm, Simplicity, and Color salient points of gradient vector. In the figure, linear regressive equations are included with R-square values. The relation the image retrieval accuracy between the proposed method and the conventional method with Color salient points of gradient vector shows the highest R-square value of 0.909 followed by Firm and Simplicity. Therefore, the most significant feature for image retrievals is Color salient points of gradient vector followed by Firm and Simplicity for these retrieved images because the proposed method shows the highest image retrieval accuracy.

IV. CONCLUSION

In this research we proposed a method for image retrieval by using wavelet transformation. In order to enhance the texture and make strong edge, we combine the vertical and horizontal detail then estimate the important point called significant point by threshold the high value then by using it find the most important information from the image and convert it into small regions and extract the image texture and color locally. We proposed a method for image retrieval by using wavelet transformation. We combined the two high sub-band frequencies In order to make strong points and edge then detect the location of significant points. The experimental results demonstrate that our method on standard dataset is significantly improved around 11 %.

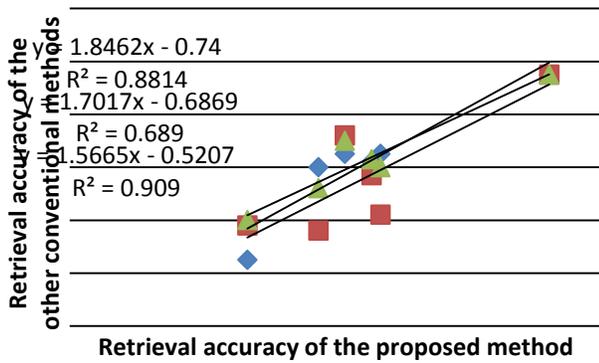


Figure 9 Relation between wavelet derived feature and the others

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 	806 	711 	162 	718 

Figure 7. Example results for the bus, dinosaur, elephant, flower and horses as a query, respectively.

The experimental results with the world widely used images for evaluation of image retrieval performance shows that the proposed method is superior to the other conventional method in terms of retrieving accuracy.

REFERENCES

[1] Kherfi, M., Brahmi, D. and Ziou D. "Combining Visual Features with Semantics for a More Effective Image Retrieval". ICPR '04, vol. 2, 2004. pp. 961-964.  
 [2] H. Yu, M. Li, H.-J. Zhang and Feng, J. "Color texture moments for content-based image retrieval". International Conference on Image Processing, 2002. pp. 24-28.

[3] Hui yu, Mingjing Li, Hong-Jiang Zhang, and Jufu Feng. "Color Texture Moments For Content Based Image Retrieval".  
 [4] N. Sebe, Q. Tian, E. Loupias, M. Lew and T. Huang. "Evaluation of salient point techniques". International Conference, 2004.  
 [5] I.Daubechies. "Ten lecturer on wavelet". Philadelphia, PA:Society for Industrial and Applied Mathematics Analysis, vol. 23, Nov. 1992. pp. 1544-1576.  
 [6] Stephane Mallet. "Wavelets for a Vision". Proceeding to the IEEE, Vol. 84, 1996. pp. 604-685.  
 [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". Proceedings of the 11th Asian Symposium on Visualization, NIIGATA, JAPAN, 2011.

- [8] <http://wang.ist.psu.edu>.
- [9] J. Li, J.Z. Wang, and G. Wiederhold. "IRM: Integrated Region Matching for Image Retrieval". Proc. of the 8th ACM Int. Conf. on Multimedia, Oct. 2000. pp. 147-156.
- [10] 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.

#### 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 wrote 30 books and published 307 journal papers

Cahaya Rahmad, He received BS from Brawijaya University Indonesia in 1998 and MS degrees from Informatics engineering at Tenth of November Institute of Technology Surabaya 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 researches are image processing, data mining and patterns recognition.