

# Comparative Study on Discrimination Methods for Identifying Dangerous Red Tide Species Based on Wavelet Utilized Classification Methods

Kohei Arai<sup>1</sup>

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

**Abstract**—Comparative study on discrimination methods for identifying dangerous red tide species based on wavelet utilized classification methods is conducted. Through experiments, it is found that classification performance with the proposed wavelet derived shape information extracted from the microscopic view of the phytoplankton is effective for identifying dangerous red tide species among the other red tide species rather than the other conventional texture, color information.

**Keywords**—hue feature; texture information; wavelet descriptor; red tide; phytoplankton identification

## I. INTRODUCTION

Image retrieval success rate (search hit ratio) is not good enough due to a poor visual signature or image feature followed by a poor similarity measure as well as clustering and classification performance. There is some information which can be extracted from images. That is (1) Halftone, color, and spectral information, (2) Spatial information including shape, size, texture, etc., and (3) Relational information such as relation between objects and the other objects included in images. The conventional image retrieval methods use the color information such as HSV<sup>1</sup>: Hue, Saturation and Value (Intensity), RGB: Red, Green, and Blue, etc. as the spectral information. Meanwhile texture information is also used in conventional image retrieval methods as the spatial information. On the other hand, Bachattarian [1], Euclidian<sup>2</sup>, Mahalanobis<sup>3</sup> distance measures [2] are well known as the similarity or distance measure. Not only hierarchical<sup>4</sup> and non-hierarchical clustering<sup>5</sup> as well as Bayesian rule of classification<sup>6</sup> and Maximum Likelihood classification<sup>7</sup>, but also Vector quantization<sup>8</sup>, Support vector machine<sup>9</sup>, etc. are proposed and used for image retrievals. Relational information such as the relations among image portions or segments, semantic information, knowledge based information, relational similarity to classify semantic relations [3] etc. are tried to use in image retrievals. Spatial and spectral information derived from the images in concern is applicable

image retrievals. There are some moment based spatial information extraction methods [4], [5], texture feature based spatial information extraction methods [6] and spectral information based image retrieval methods [7], [8], [9]. Furthermore, some attempts are made for image retrievals with wavelet descriptor as a spatial information extraction [9], [10]. In general, these conventional methods have not so good performance in terms of retrieval success rate.

All the spectral and spatial information are used in image retrieval except shape information. There are some trials to use shape information extracted from image using Fourier descriptor and the others. There are some definitions for Fourier descriptors. Zahn and Roskies proposed Z type descriptor [7] while Granlund proposed G type descriptor [11]. Z type descriptor is defined as the cumulative angle changes of the contour points from the starting point is expanded with Fourier series while G type descriptor defined as the length between the contour points from the start point of contour line in concern is expanded with Fourier series. Both of descriptors have the following problems:

- (1) It is hard to express local properties,
- (2) It cannot represent the shape of contour when the shape is not closed,

The results depend on the start point on the contour line in concern for tracking. On the other hand, Z type descriptor has another difficulty that the convergence speed is not fast so that it takes relatively large computational resources and the reproducibility of low frequency component is not good enough. Meanwhile, G type descriptor has another difficulty that Gibbs phenomenon [12] would occur at the end points of the closed curve of contour lines results in the end points cannot be preserved.

The shape descriptor proposed here is wavelet based descriptor not the Fourier type of descriptor. Therefore, the proposed wavelet based descriptor allows shape description through frequency-time analysis while Fourier based descriptor allows only frequency components representation of shape. There is some advantage for the wavelet based descriptor in shape information extraction rather than Fourier descriptor. Wavelet descriptor is proposed for best matching method to measure similarity between two feature vectors of the two shapes [9], [10]. This is impractical for higher dimensional feature matching. Therefore, wavelet descriptors are more suitable for model-based object recognition than

<sup>1</sup> [http://www.cs.rit.edu/~ncs/color/t\\_convert.html](http://www.cs.rit.edu/~ncs/color/t_convert.html)

<sup>2</sup> [http://en.wikipedia.org/wiki/Euclidean\\_distance](http://en.wikipedia.org/wiki/Euclidean_distance)

<sup>3</sup> [http://en.wikipedia.org/wiki/Mahalanobis\\_distance](http://en.wikipedia.org/wiki/Mahalanobis_distance)

<sup>4</sup> [http://en.wikipedia.org/wiki/Hierarchical\\_clustering](http://en.wikipedia.org/wiki/Hierarchical_clustering)

<sup>5</sup> <http://www.daylight.com/meetings/mug96/barnard/E-MUG95.html>

<sup>6</sup> [http://en.wikipedia.org/wiki/Naive\\_Bayes\\_classifier](http://en.wikipedia.org/wiki/Naive_Bayes_classifier)

<sup>7</sup> [http://www.ccrs.nrcan.gc.ca/glossary/index\\_e.php?id=341](http://www.ccrs.nrcan.gc.ca/glossary/index_e.php?id=341)

<sup>8</sup> [http://en.wikipedia.org/wiki/Learning\\_Vector\\_Quantization](http://en.wikipedia.org/wiki/Learning_Vector_Quantization)

<sup>9</sup> [http://en.wikipedia.org/wiki/Support\\_vector\\_machine](http://en.wikipedia.org/wiki/Support_vector_machine)

data-driven shape retrieval, because for shape retrieval, which is usually conducted online, speed is essential.

Contour of the object extracted from the original image can be expressed with wavelet based descriptor. The proposed image retrieval method is based on the hue information and texture as well as the proposed wavelet described shape information of extracted objects to improve image retrieval success rate.

The following section describes the proposed image retrieval method followed by some experiments for reproducibility of the proposed wavelet descriptor in comparison to the conventional Fourier descriptor with several simple symmetrical and asymmetrical shapes. Then it is validated with the image database of phytoplankton [13].

## II. PROPOSED METHOD

### A. Research Background

There are a plenty of red tide species. Small portion of red tide species can be listed up in Figure 1.

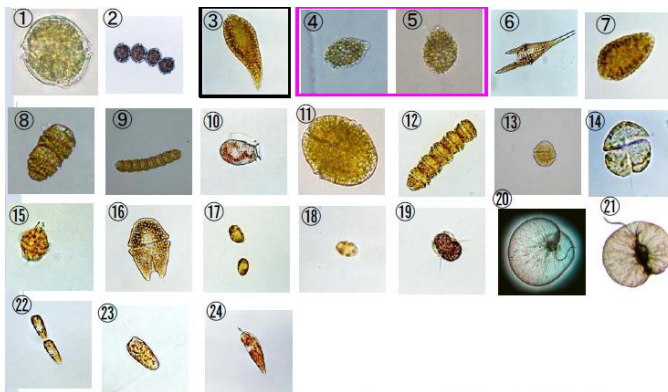


Figure 1 Photos of a portion of red tide species

These red tide species can be classified into three categories, (a) Caution level of species, (b) Warning level of species, and (c) Dangerous species. Fishes and shells take these dangerous red tide species. After that human habitats eat the fishes and shells. Then such persons get a bad situation and have an illness condition. Therefore, these red tide species are classified into dangerous species.



(a) Caution level species



(b) Warning level species



(c) Dangerous species

Figure 2 Three categories of red tide species

Identifying these dangerous red tide species is important. It, however, is not so easy to classify because these three categories of red tide species are quite resemble. Usually, the local fishery research institutes measure red tide from the research vessels with microscope. They used to count the number of red tide with microscope camera acquired imagery data on the ship. Then identify the red tide species in the same time quickly. Even though human perception capability is superior to that by machine learning based automatic classification, there are some mistakes. The purpose of the research is to improve classification performance by using considerable features which can be extracted from the microscopic imagery data.

### B. Process Flow of the Proposed Image Classification

Image classification method based on hue information [14] and wavelet description based shape information [15] as well as texture information of the objects extracted with dyadic wavelet transformation [16] is proposed. Object is assumed to be focused so that the frequency component in the object is relatively high in comparison to the other (background). Figure 3 shows the process flow of the proposed image classification method.

One of the image features of hue information (angle) is calculated for the entire image in the color image database. Dyadic wavelet transformation<sup>10</sup> is also applied to the images then texture information is extracted from the transformed resultant image. Based on the Dyadic wavelet transformation, HH<sup>11</sup> image of edge is extracted from the original image. Morphological operations<sup>12</sup>, opening and closing are then applied to the edge extracted images to remove inappropriate isolated pixels and undesirable image defects. After that the

<sup>10</sup>

[http://cas.enscm.fr/~chaplais/Wavetour\\_presentation/ondelettes%20dyadiques/Dyadic\\_Transform.html](http://cas.enscm.fr/~chaplais/Wavetour_presentation/ondelettes%20dyadiques/Dyadic_Transform.html)

<sup>11</sup> HH denotes high frequency component in horizontal direction and high frequency component in vertical direction

<sup>12</sup> [http://geol.hu/data/online\\_help/MorphologyFilters.html](http://geol.hu/data/online_help/MorphologyFilters.html)

resultant image is binarized with appropriate threshold then contour of the object is extracted. Then the Dyadic wavelet transformation is applied to the contour in order to extract shape information (Wavelet descriptor). After all, Euclidian distance between target image and the other candidate images in the color image database is calculated with extracted hue, texture and shape information then the closest image is retrieved.

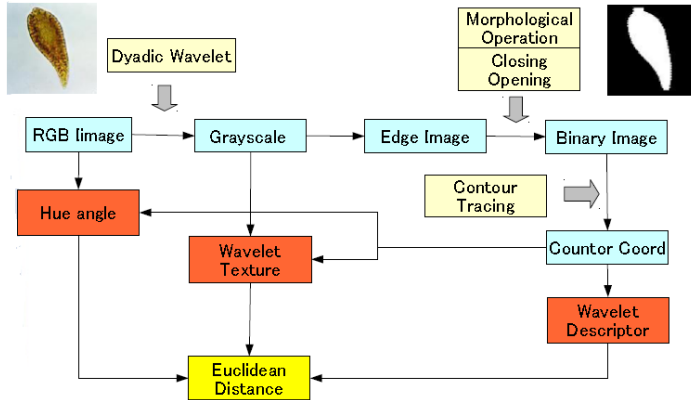


Figure 3 Process flow of the proposed image classification method.

### C. Dyadic wavelet transformation

Using dyadic wavelet, frequency component can be detected. Dyadic wavelet allows to separate frequency components keeping image size with that of original image. Dyadic wavelet is called as a binary wavelet and has high pass and low pass filter components,  $\{h[k], g[k]\}$  and reconstruction filter  $\{\underline{h}[k], \underline{g}[k]\}$ . Low and high frequency components,  $C_n$  and  $d_n$  are expressed as follows,

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

$$d_n [i]=\sum_k g [k] C_{n-1} [i + k2^{n-1}] \quad (2)$$

Then original image is also reconstructed with the low and high frequency components as follows,

$$C_{n-1} [i]=1/2 \sum_k \underline{h}[k] C_n[i-k2^{n-1}]+\sum_k \underline{g}[k] d_n [i-k2^{n-1}] \quad (3)$$

If a new parameter  $s[m]$  is employed, then lifting dyadic wavelet is defined as follows,

$$h^{new}[k]=h^{old} [k] \quad (4)$$

$$\underline{h}^{new}[k]=\underline{h}^{old} [k] + \sum_m s[-m] \underline{g}^{old} [k-m] \quad (5)$$

$$g^{new}[k]=g^{old} [k] - \sum_m s[m] h^{old} [k-m] \quad (6)$$

$$\underline{g}^{new}[k]=\underline{g}^{old} [k] \quad (7)$$

### D. Dyadic wavelet based descriptor (Shape information)

Image classification method with hue and texture information is conventional. In the proposed method, another feature, shape information is employed. Fourier descriptor is used, in general, to represent shape information. Although Fourier descriptor represents frequency component of the contour line, location information cannot be described. In other words, Fourier descriptor does support only frequency analysis, and does not support time-frequency component analysis.

Wavelet descriptor which is proposed by this paper supports a time-frequency component analysis so that not only

frequency component but also location of contour edge can be discussed [17].

Let  $u(i)$  be distance between a point in the closed object contour line and a certain point  $i$  on the line, then the closed object contour line can be represented as  $u(i), i=1,2,\dots,n, i=1$  corresponds to 0 degree while  $i=n$  corresponds to 360 degree, respectively.  $u(i)$  can be converted with dyadic wavelet transformation. Then the contour line can be represented with high frequency component of the dyadic wavelet transformed sequence as is shown in Figure 4. Then average of the high frequency component of pixel value is used for a feature of the image classification.

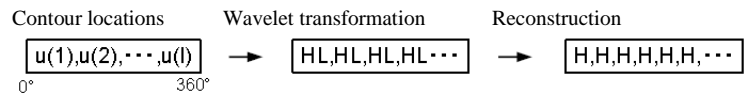


Figure 4 Dyadic wavelet descriptor for representation of the closed object contour lines.

### E. Texture Information

Also texture information is useful for discrimination. Texture information can be derived from dyadic wavelet transformation. Texture information is defined as high frequency component of pixel value derived from dyadic wavelet transformation. Dyadic wavelet transformation is applied to the 2x2 pixels defined in Figure 5. Pixel value of the pixel in the object is replaced to the high frequency component detected with dyadic wavelet. Thus image which represents texture information of the detected object image is generated [18].

### F. Hue angle

Thus contour of the object is detected. Then Red, Green, and Blue: RGB of the original object image can be transformed to Hue, Saturation, and Intensity: HSV information. Hue information in unit of radian, in particular, is useful for discrimination of the target image classifications of phytoplankton images. HSV, on the other hand, is expressed in Figure 6 (Color coordinate system).

RGB to HSV conversion is also be expressed as follows,

$$V = \max(R, G, B)$$

$$S = (V - X)/V \text{ where } X = \min(R, G, B)$$

$$R = V: H = (\pi/3) (b - g)$$

$$G = V: H = (\pi/3) (2 + r - b)$$

$$B = V: H = (\pi/3) (4 + g - r)$$

where  $r = (V - R)/(V - X), g = (V - G)/(V - X), b = (V - B)/(V - X), H$  ranges from 0 to 360,  $S$  ranges from 0 to 1,  $V$  ranges from 0 to 1, HSV representation and  $R, G, B$  also range from 0 to 1. These three features, hue,  $H$ , texture,  $xx$  and shape information,  $yy$  composes three dimensional feature space results in measurement of Euclidian distance between a query image and the images in previously created image database. Using the distance, a query image can be retrieved from the image in the database. Thus image classifications can be done with hue and texture information as well as shape information derived from dyadic wavelet descriptor.

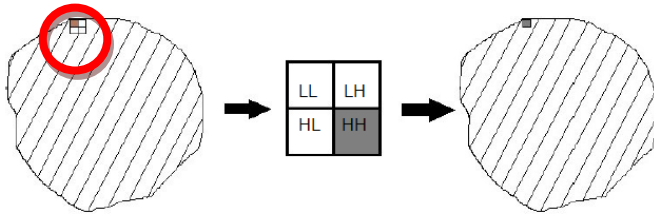


Figure 5. Detected object and 2x2 of matrix in the object to detect texture information with 2x2 of dyadic wavelet transformation.

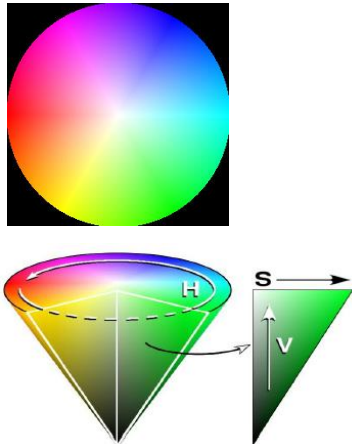


Figure 6 Hue angle and Saturation as well as Value (Intensity).

G. Alternative shape information (Fourier Descriptor)

The location coordinate is expressed in the complex plane representation for the G type of Fourier descriptor, that is,

$$Z_s = X_s + iY_s \tag{11}$$

Then space or time domain locations can be transformed with eth following equation,

$$F_i = \frac{1}{S} \sum_{s=0}^{S-1} Z_s \exp\left(\frac{-2\pi ks}{S} i\right) \tag{12}$$

It can be inversely transformed with the following equation,

$$Z_s = \sum_{i=0}^{S-1} F_i \exp\left(\frac{2\pi ks}{S} i\right) \tag{13}$$

If these transformation and inverse transformation is perfect, then the original shapes are completely reproduced. The reproducibility for the shapes of circle, triangle, square, and trapezium (asymmetric shape) of the proposed wavelet descriptor is better than that of the conventional Fourier descriptor as shown in Figure 7.

In the comparison, the original image is binarized and the contour is extracted. Then shape information is extracted with both Fourier descriptor (G-type) and Dyadic wavelet descriptor. After that, image is reconstructed with the extracted shape information then compares the reconstructed images with two descriptors, Fourier and Dyadic wavelet descriptors. The difference between the reconstructed contours and original image is shown in Table 1. Thus it is found that the reproducibility of Dyadic wavelet descriptor is better than the conventional Fourier descriptor.

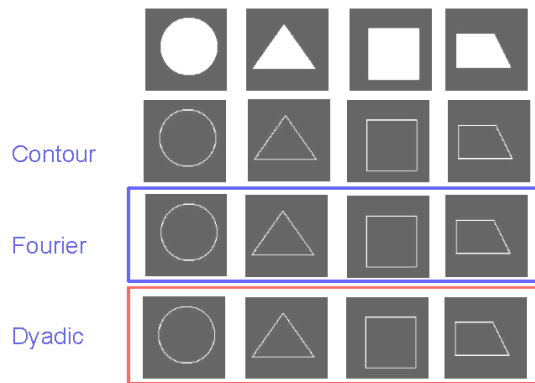


Figure 7 Comparison of reproducibility of the shapes between Fourier and Dyadic wavelet descriptor.

TABLE I. COMPARISON OF THE DIFFERENCE BETWEEN THE ORIGINAL AND RECONSTRUCTED CONTOURS WITH FOURIER AND DYADIC WAVELET DESCRIPTORS.

Shape	Fourier	Dyadic
circle	0.4121	0.1809
triangle	0.5391	0.1280
square	0.3689	0.1101
trapezium	0.4660	0.1929

This method can be expanded to 3D object. Once 3D object image is acquired through scanning in roll/pitch/yaw directions with the appropriate step angle, and then contour lines of the acquired 2D images are extracted. After that the 3D object shape complexity is represented with the wavelet descriptor as a resultant image which includes series of the high frequency components derived from dyadic wavelet transformation as shown in Figure 4 something like the following sequence,

- (1) HH..H(for roll angle 0), HH..H(for roll angle 10),..., HH..H(for roll angle 350),
- (2) HH..H(for pitch angle 0), HH..H(for pitch angle 10),..., HH..H(for pitch angle 350),
- (3) HH..H(for yaw angle 0), HH..H(for yaw angle 10),..., HH..H(for yaw angle 350).

It is an image and is totally visual. This image represents 3D object shape complexity as an index. Also this index is shift invariant and rotation invariant. Namely, the index is not changed even if 3D object is translated and rotated.

III. EXPERIMENTS

A. Euclidian Distance

One of the measures for classification performance evaluation is Euclidian distance among the classes in concern. Shorter Euclid distance implies a poor classification performance while longer distance means a good performance. If the texture and hue information are used for classification, then Table 2 of Euclid distance is calculated while those of the case of utilizing wavelet descriptor and texture information is shown in Table 3. In Table 2 and 3, Euclid distance between class #5 of *Chattnella\_Antiqua* and the others are listed because primary red tide species at this moment is *Chattnella\_Antiqua*.

TABLE II. CALCULATED EUCLID DISTANCE WHEN TEXTURE AND HUE INFORMATION ARE USED AS FEATURES

	file name	Euclid Distance (t, h)
1	a.catenella1cell	0.0342
2	a.catenella4cell	0.7316
3	c.antiqua	0.0024
4	c.antiqua2	0.0602
5	c.antiqua3	0.0000
6	c.furca	0.0798
7	c.marina	0.0074
8	c.polykrikoides2cell	0.0035
9	c.polykrikoides8cell	0.0148
10	d.fortii	0.0817
11	g.catenatum1cell	0.0061
12	g.catenatum5cell	0.0012
13	g.instriatum	0.0470
14	g.mikimotoi	0.0448
15	g.polygramma	0.0214
16	g.sanguineum	0.0356
17	h.akashiwo	0.0260
18	h.circularisquama	0.0148
19	m.rubrum	0.0044
20	n.scintillans4	0.2059
21	n.scintillans5	0.7316
22	p.dentatum	0.0196
23	p.dentatum2	0.0358
24	p.signoides	0.0271

TABLE III. CALCULATED EUCLID DISTANCE WHEN TEXTURE AND WAVELET DESCRIPTOR INFORMATION ARE USED AS FEATURES

	Euclid Distance (w, t)		
	daubechies1	dyadic	fourier
1	0.0044	0.0263	0.0234
2	0.7619	0.6435	0.5593
3	0.0034	0.0082	0.0565
4	0.0104	0.0886	0.0211
5	0.0000	0.0000	0.0000
6	0.0071	0.0048	0.0719
7	0.0026	0.0281	0.0061
8	0.0136	0.0329	0.0323
9	0.0259	0.0436	0.0279
10	0.0098	0.0679	0.0104
11	0.0032	0.0730	0.0186
12	0.0144	0.0419	0.0385
13	0.0021	0.0488	0.0192
14	0.0114	0.0405	0.4726
15	0.0158	0.0305	0.0142
16	0.0053	0.0319	0.0189
17	0.0148	0.0136	0.0063

18	0.0016	0.1090	0.0287
19	0.0164	0.0294	0.3365
20	0.0147	0.1654	0.0651
21	0.7619	0.6435	0.5593
22	0.0087	0.0119	0.0411
23	0.0260	0.0025	0.0407
24	0.0086	0.0356	0.0650

Also, Euclid distance between *Chattnella Antiqua* and the other species are calculated and shown in Table 4 for the case of utilizing all these three features of wavelet descriptor, texture, and hue information together.

TABLE IV. CALCULATED EUCLID DISTANCE WHEN TEXTURE WAVELET DESCRIPTOR AND HUE INFORMATION ARE USED AS FEATURES

	Euclid Distance (w, t, h)		
	daubechies1	dyadic	fourier
1.0000	0.0345	0.0431	0.0414
2.0000	0.9294	0.8352	0.7722
3.0000	0.0042	0.0086	0.0565
4.0000	0.0611	0.1072	0.0638
5.0000	0.0000	0.0000	0.0000
6.0000	0.0800	0.0799	0.1074
7.0000	0.0078	0.0291	0.0096
8.0000	0.0141	0.0330	0.0325
9.0000	0.0298	0.0460	0.0316
10.0000	0.0823	0.1062	0.0823
11.0000	0.0069	0.0732	0.0195
12.0000	0.0144	0.0419	0.0385
13.0000	0.0471	0.0678	0.0508
14.0000	0.0462	0.0604	0.4747
15.0000	0.0266	0.0373	0.0257
16.0000	0.0360	0.0478	0.0403
17.0000	0.0299	0.0293	0.0267
18.0000	0.0149	0.1099	0.0323
19.0000	0.0170	0.0297	0.3365
20.0000	0.2064	0.2641	0.2159
21.0000	0.9294	0.8352	0.7722
22.0000	0.0214	0.0229	0.0455
23.0000	0.0442	0.0359	0.0542
24.0000	0.0284	0.0448	0.0704

#### IV. CONCLUSION

Comparative study on discrimination methods for identifying dangerous red tide species based on wavelet utilized classification methods is conducted. Through experiments, it is found that classification performance with the proposed wavelet derived shape information extracted from the microscopic view of the phytoplankton is effective for identifying dangerous red tide species among the other red tide species rather than the other conventional texture, color information

ACKNOWLEDGMENT

The author would like to thank Mr. Yuji Yamada for his effort to conduct the experiments.

REFERENCES

- [1] Duda R.O., P.E. Hart, and D.G. Stork, (2001), Pattern Classification, (Second Edition), John Wiley & Sons Inc.
- [2] Arai K. (1996), Fundamental theory for image processing, Gakujutsu-Tosho Shuppan Publishing Co., Ltd.
- [3] Séaghdha, D.O., Ann Copestake, (2009), Using lexical and relational similarity to classify semantic relations, Computational Linguistics 621-629.
- [4] Teh C.H. and R. T. Chin,(1988), On image analysis by the methods of moments, IEEE Trans. On Pattern Analysis and Machine Intelligence, 10, 4, 496-513
- [5] Taubin G. and D. B. Cooper,(1991), Recognition and Positioning of Rigid Objects Using Algebraic Moment Invariants, SPIE Conf. On Geometric Methods in Computer Vision, 1570, 175-186.
- [6] Niblack W.,(1993), The QBIC Project: Querying Images By Content Using Color, Texture and Shape, SPIE Conf. On Storage and Retrieval for Image and Video Databases, 1908, 173-187.
- [7] Zahn C.T., and Ralph Z. Roskies. (1972), Fourier Descriptors for Plane closed Curves. IEEE Trans. On Computer,c-21(3):269-281.
- [8] Huang C.L. and D.H. Huang,(1998), A Content-based image retrieval system. Image and Vision Computing, 16:149-163.
- [9] Yang H.S., S.U. Lee, K M. Lee., (1998), Recognition of 2D Object Contours Using Starting-Point-Independent Wavelet Coefficient Matching. Journal of Visual Communication and Image Representation, 9, 2, 171-181.
- [10] Tieng Q.M. and W. W. Boles, (1997), Recognition of 2D Object Contours Using the Wavelet Transform Zero-Crossing Representation, IEEE Trans. on PAMI 19, 8, 1997.
- [11] Grandlund H., (1972), Fourier preprocessing for hand print character recognition, IEEE Trans. on Computers, 621, 195-201.
- [12] Gibbs, J. W., (1899), "Fourier Series". Nature 59, 200 and 606.
- [13] Arai K. and Yasunori Terayama (2010), Polarized radiance from red tide, Proceedings of the SPIE Asia Pacific Remote Sensing, AE10-AE101-14.
- [14] Arai K. et al. (1991), Takagi and Shimoda ed., Image Analysis Handbook, Tokyo Daigaku Shuppan-kai publishing.
- [15] Arai K. (1998), Methods for Image Processing and Analysis of Earth Observation Satellite Imagery Data, Morikita Shuppan Publishing Co., Ltd.
- [16] Arai K. and L. Jameson (2001), Earth observation satellite data analysis based on wavelet analysis, Morikita-Shuppan Publishing Co., Ltd.
- [17] Arai K. (2002), Java based Earth observation satellite imagery data processing and analysis, Morikita-Shuppan Publishing Co., Ltd.
- [18] Arai, K., (2011), Visualization of 3D object shape complexity with wavelet descriptor and its application to image retrievals, Journal of Visualization, 15, 155-166.
- [19] Datta, Ritendra; Dhiraj Joshi, Jia Li, James Z. Wang (2008). "Image Retrieval: Ideas, Influences, and Trends of the New Age". *ACM Computing Surveys* 40 (2): 1–60
- [20] Prasad, B E; A Gupta, H-M Toong, S.E. Madnick (1987). "A microcomputer-based image database management system". *IEEE Transactions on Industrial Electronics* IE-34 (1): 83–8.

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 1974 to 1978 also was with National Space Development Agency of Japan (current JAXA) from 1979 to 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 was appointed professor at Department of Information Science, Saga University in 1990. He was appointed 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 also

appointed councilor of Saga University from 2002 and 2003 followed by an executive councilor of the Remote Sensing Society of Japan for 2003 to 2005. He is an adjunct professor of University of Arizona, USA since 1998. He also was appointed vice chairman of the Commission "A" of ICSU/COSPAR in 2008. He wrote 30 books and published 332 journal papers

APPENDIX

Table 1, 2, and 3 shows the Euclid distance between *Chatmella\_Antiqua* and the other species for the case that the support length of base function of wavelet is two. Meanwhile, Table 4, 5, and 6 shows the Euclid distance between *Chatmella\_Antiqua* and the other species for the case that the support length of base function of wavelet is four.

	file name	Euclid Distance (t, h)
1	a.catenella1cell	0.0342
2	a.catenella4cell	0.7316
3	c.antiqua	0.0024
4	c.antiqua2	0.0602
5	c.antiqua3	0.0000
6	c.furca	0.0798
7	c.marina	0.0074
8	c.polykrikoides2cell	0.0035
9	c.polykrikoides8cell	0.0148
10	d.fortii	0.0817
11	g.catenatum1cell	0.0061
12	g.catenatum5cell	0.0012
13	g.instriatum	0.0470
14	g.mikimotoi	0.0448
15	g.polygramma	0.0214
16	g.sanguineum	0.0356
17	h.akashiwo	0.0260
18	h.circularisquama	0.0148
19	m.rubrum	0.0044
20	n.scintillans4	0.2059
21	n.scintillans5	0.7316
22	p.dentatum	0.0196
23	p.dentatum2	0.0358
24	p.signoides	0.0271

	Euclid Distance (w, t)		
	daubechies1	dyadic	fourier
1	0.0044	0.0103	0.0234
2	0.7619	0.6027	0.5593
3	0.0034	0.0141	0.0565
4	0.0104	0.0775	0.0211
5	0.0000	0.0000	0.0000
6	0.0071	0.0064	0.0719
7	0.0026	0.0146	0.0061
8	0.0136	0.0264	0.0323
9	0.0259	0.0303	0.0279
10	0.0098	0.0510	0.0104
11	0.0032	0.0508	0.0186
12	0.0144	0.0407	0.0385

13	0.0021	0.0287	0.0192
14	0.0114	0.0345	0.4726
15	0.0158	0.0297	0.0142
16	0.0053	0.0185	0.0189
17	0.0148	0.0041	0.0063
18	0.0016	0.0892	0.0287
19	0.0164	0.0169	0.3365
20	0.0147	0.1305	0.0651
21	0.7619	0.6027	0.5593
22	0.0087	0.0250	0.0411
23	0.0260	0.0250	0.0407
24	0.0086	0.0455	0.0650

7	c.marina	0.0074
8	c.polykrikoides2cell	0.0035
9	c.polykrikoides8cell	0.0148
10	d.fortii	0.0817
11	g.catenatum1cell	0.0061
12	g.catenatum5cell	0.0012
13	g.instriatum	0.0470
14	g.mikimotoi	0.0448
15	g.polygramma	0.0214
16	g.sanguineum	0.0356
17	h.akashiwo	0.0260
18	h.circularisquama	0.0148
19	m.rubrum	0.0044
20	n.scintillans4	0.2059
21	n.scintillans5	0.7316
22	p.dentatum	0.0196
23	p.dentatum2	0.0358
24	p.signoides	0.0271

Euclid Distance (w, t, h)			
	daubechies1	dyadic	fourier
1	0.0345	0.0357	0.0414
2	0.9294	0.8041	0.7722
3	0.0042	0.0143	0.0565
4	0.0611	0.0982	0.0638
5	0.0000	0.0000	0.0000
6	0.0800	0.0800	0.1074
7	0.0078	0.0163	0.0096
8	0.0141	0.0267	0.0325
9	0.0298	0.0337	0.0316
10	0.0823	0.0963	0.0823
11	0.0069	0.0511	0.0195
12	0.0144	0.0407	0.0385
13	0.0471	0.0551	0.0508
14	0.0462	0.0566	0.4747
15	0.0266	0.0366	0.0257
16	0.0360	0.0401	0.0403
17	0.0299	0.0263	0.0267
18	0.0149	0.0904	0.0323
19	0.0170	0.0174	0.3365
20	0.2064	0.2438	0.2159
21	0.9294	0.8041	0.7722
22	0.0214	0.0318	0.0455
23	0.0442	0.0437	0.0542
24	0.0284	0.0529	0.0704

Euclid Distance (w, t)			
	daubechies1	dyadic	fourier
1	0.0044	0.0097	0.0234
2	0.7619	0.5513	0.5593
3	0.0034	0.0161	0.0565
4	0.0104	0.0595	0.0211
5	0.0000	0.0000	0.0000
6	0.0071	0.0186	0.0719
7	0.0026	0.0029	0.0061
8	0.0136	0.0126	0.0323
9	0.0259	0.0126	0.0279
10	0.0098	0.0283	0.0104
11	0.0032	0.0222	0.0186
12	0.0144	0.0373	0.0385
13	0.0021	0.0007	0.0192
14	0.0114	0.0307	0.4726
15	0.0158	0.0268	0.0142
16	0.0053	0.0008	0.0189
17	0.0148	0.0114	0.0063
18	0.0016	0.0630	0.0287
19	0.0164	0.0088	0.3365
20	0.0147	0.0830	0.0651
21	0.7619	0.5513	0.5593
22	0.0087	0.0356	0.0411
23	0.0260	0.0293	0.0407
24	0.0086	0.0455	0.0650

Euclid Distance		
	file name	(t, h)
1	a.catenella1cell	0.0342
2	a.catenella4cell	0.7316
3	c.antiqua	0.0024
4	c.antiqua2	0.0602
5	c.antiqua3	0.0000
6	c.furca	0.0798

	Euclid Distance (w, t, h)		
	daubechies1	dyadic	fourier
1	0.0345	0.0355	0.0414
2	0.9294	0.7664	0.7722
3	0.0042	0.0163	0.0565
4	0.0611	0.0847	0.0638
5	0.0000	0.0000	0.0000
6	0.0800	0.0819	0.1074
7	0.0078	0.0079	0.0096
8	0.0141	0.0131	0.0325
9	0.0298	0.0194	0.0316
10	0.0823	0.0864	0.0823
11	0.0069	0.0230	0.0195
12	0.0144	0.0373	0.0385
13	0.0471	0.0470	0.0508
14	0.0462	0.0543	0.4747
15	0.0266	0.0343	0.0257
16	0.0360	0.0356	0.0403
17	0.0299	0.0283	0.0267
18	0.0149	0.0647	0.0323
19	0.0170	0.0098	0.3365
20	0.2064	0.2220	0.2159
21	0.9294	0.7664	0.7722
22	0.0214	0.0406	0.0455
23	0.0442	0.0463	0.0542
24	0.0284	0.0530	0.0704