

## Enhancing Red Tide Image Recognition using Semantic Feature and Rotation of Algae Image Angle

Sun Park<sup>1</sup>, Min A Jeong<sup>2</sup>, Yeong Ju Kim<sup>3</sup>, Jingwan Park<sup>4</sup>, Seong Ro Lee<sup>5</sup>

<sup>1</sup> Institute Research of Information Science and Engineering,  
Mokpo National University, South Korea

<sup>2,3,4</sup> Department of Computer Engineering, Mokpo National University, South Korea

<sup>5</sup> Department of Information Electronics Engineering, Mokpo National University,  
South Korea

<sup>1,2,4,5</sup>{sunpark, majung, jgpark, srlee}@[mokpo.ac.kr](mailto:mokpo.ac.kr), <sup>3</sup>[Xfile-7@daum.net](mailto:Xfile-7@daum.net)

**Abstract.** This paper proposes the red tide algae image recognition method using rotation of image angle and semantic feature based on NMF (nonnegative matrix factorization). The experimental results demonstrate that the proposed method achieves better performance than other red tide recognition methods.

**Keywords:** red tide algae, image recognition, NMF (nonnegative matrix factorization), rotation of image angle, semantic feature.

### 1 Introduction

Since the red tide of South Korea has a broad occurrence range and spreads quickly, the direct detection of red tide blooms in real time for mitigation activity is difficult. Seven of the 63 species of harmful red tide algae damage fisheries in coastal areas of South Korea. The HABs have inflicted massive mortality on fin fish and shellfish, damaging the economies of fisheries for almost every year from 1993 to 2012. Particularly, there were huge harmful algal blooms of *Cochlodinium polykrikoides* in 1995, resulting in about 95 million dollars in fishery damage [1].

There have been many studies on red tide because the damage to sea farming by red tide blooms of harmful algae occurs every year. The research on red tide events of harmful algae in the coastal area of South Korea has mostly focused chemical properties and investigation of biological causes and is generally used for mitigation activity [2]. However, research on a variety of automatic methods for monitoring and recognizing red tide images of harmful algae on the coast of South Korea is still insufficient [1].

The studies of red tide recognition mostly use classification techniques, such as Bayesian classifiers [3], SVM[3], SVDD (support vector data description) [4], and FCM (fuzzy c-means) clustering [5]. These methods [3, 4, 5] are difficult to classify red tide algae images except training algae images. In addition, they do not consider matching point between training images and query image, since the training images

have different forms regarding various species of red tide algae (i.e., there are blooming red tides with relation to 63 species of algae in the coastal area of South Korea). In our previous study, we proposed the red tide image recognition method using roundness and entropy [6]. This method can recognize various species of red tide algae image, however the method is affected by image angle.

In order to resolve the above limitations of few training algae images and without matching point, this paper proposes a new red tide image recognition method using rotation of image angle based on semantic features of NMF. The proposed method has following advantages. First, the method uses rotation of algae images angle to construct image vector set with respect to preprocess for enhancing the accuracy of image recognition. Second, training of the method uses the semantic features of 63 species of red tide algae image by means of NMF which can recognize various species of red tide algae images. Final, our method uses shape roundness and entropy of red tide images which improves the recognition accuracy of algae images to apply matching point of the images for the images recognition.

## 2 Related Works

Recently, studies of red tide recognition mostly use classification techniques as follows. Jiang et al. proposed the red tide algae classification method using Bayesian classifier and SVM (support vector machine). Their method removes unknown and polluted algae images using Bayesian classifier. Red tide algae images are classified using SVM [3]. However, there is no good performance to remove unknown and polluted images. Besides, they use only 8 species of red tide algae for training images classification. Jiang et al. proposed red tide algae recognition method using two step classifiers. Their method uses first step classifier in which SVDD (support vector data description) removes unknown and polluted algae images. Algae images are classified into the predefined class label using second step SVM classifier. Also this method uses only 8 species of red tide algae for learning image classification [4]. Lili et al. use SVM and FCM (fuzzy c-means) clustering method for classifying red tide images. Their method uses SVM to identify class of test sample. The method identifies the misclassified data using probability based method, and then the misclassified data is reclassified using FCM. Their method also uses 5 species of red tide algae to train image classification [5].

## 3 Non-negative Matrix Factorization

This section reviews NMF theory with algorithm and describes the advantage of semantic features of NMF in Example 1. In this paper, we define the matrix notation as follows: Let  $x_{*j}$  be  $j$ 'th column vector of matrix  $X$ ,  $x_{i*}$  be  $i$ 'th row vector, and  $x_{ij}$  be the element of  $i$ 'th row and  $j$ 'th column. NMF is to decompose a given  $m \times n$  matrix  $A$  into a non-negative semantic feature matrix  $W$  and a non-negative semantic variable matrix  $H$  as shown in Equation (1) [7].

$$A \approx WH \quad (1)$$

where  $W$  is a  $m \times r$  non-negative matrix and  $H$  is a  $r \times n$  non-negative matrix. Usually  $r$  is chosen to be smaller than  $m$  or  $n$ , so that the total sizes of  $W$  and  $H$  are smaller than that of the original matrix  $A$ .

The objective function is used minimizing the Euclidean distance between each column of  $A$  and its' approximation  $\tilde{A} = WH$ , which was proposed by Lee and Seung [7]. As an objective function, the Frobenius norm is used:

$$\sum_{i=1}^m \sum_{j=1}^n \|A_{ij} - \sum_{l=1}^r W_{il} H_{lj}\|^2 = \|A - WH\|_F^2 \quad (2)$$

Updating  $W$  and  $H$  is kept until  $\|A - WH\|_F$  converges under the predefined threshold or exceeds the number of repetition. The update rules are as follows:

$$H_{lj} = \frac{(W^T A)_{lj}}{\sum_{i=1}^m W_{il} A_{ij}} \quad W_{il} = \frac{A_{ij} H_{lj}}{\sum_{j=1}^n A_{ij} H_{lj}} \quad (3)$$

## 4 Proposed Method

This paper proposes a red tide algae image recognition method using the semantic feature of images and the preprocessing algae image. The proposed method consists of three phases: preprocessing, candidate image extraction, and recognizing red tide algae image.

### 4.1 Preprocessing

The different red tide algae images are recognized by different angle of same query image. In order to solve the above limitations, we rotate the algae image angle. The preprocessing phase consists of three steps: rotation of image angle step, fixing image size step, and normalization step. The phase is described as follows. First, rotation of image angle step rotates algae image angle according to based angel. Second, fixing image size step converts different algae image pixel size into same algae image pixel size (i.g., 120 by 100 pixel matrix). In addition, the center of image is almost reconciled with the center of image vector. Final, the normalization step normalizes all red tide algae image by values of mean and variance of image vectors in order to decrease an error occurring of image background and light.

### 4.2 Candidate image extraction

The phase consists of the calculating semantic feature of image step and the calculating semantic distance step. In the first step, the semantic feature of image is calculated using the NMF algorithm of Equation (1). We modified Turk's method [8] by using NMF for training and candidate algae image set. The training and candidate

algae image set are constructed by the modified Turk's method [8] based on the image semantic feature of NMF. In the second step, the candidate images are extracted having minimum value of Euclidean distance between a query image and the trained image set.

### 4.3 Recognizing red tide image

Our above NMF based method or algae image recognition methods of related works have problems in recognizing different image according to viewpoint of query images. Because there is no matching point between query image and candidate image set. In order to resolve the problem of without matching point, this paper proposes the roundness and entropy based algae image recognition method to enhance the quality of image recognition. The algae images can be recognized to compare roundness of shape between query image and candidate images (i.e. training image set).

In this paper, the roundness scale goes from 0 to 1. (i.e., the roundness of circle shape and horizontal shape is 1 and 0 respectively.). Roundness  $R()$  is defined as follows [9].

$$R(I) = \frac{4 \times \pi \times \text{area}}{\text{perimeter}^2} \quad (4)$$

where  $I$  is  $120 \times 100$  image matrix,  $\pi$  is pi that is the ratio of the circumference of a circle to its diameter (i.e., pi is 3.1416), area is area of image object, perimeter is a path that surrounds an area of image.

Entropy of Equation (5) is given as follows [9, 10]:

$$\text{Entropy}(I_{st}) = - \sum_{i=1}^m p_i \log_2(p_i) \quad (5)$$

where  $I_{st}$  is  $120 \times 100$  semantic feature image,  $p$  is number of histogram of  $I_{st}$ .

## 4 Experiment

We used the training data of 3,150 algae images with respect to 63 species of red tide algae. Fifty images are included in the label of one species of red tide algae. We use 350 query images with respect to 7 species of most harmful algae on South Korea. Average accuracy is used to measure the recognition performance. We have conducted performance evaluation by testing proposed method and comparing it with 8 other recognition methods using same training and query images. We implemented 9 image recognition methods: PCA, NB, NN, SVM, B+SVM, SVDD+SVM, SVM+FCM, PCA+RE, and RNMF. The average accuracy of evaluation results of RNMF is approximately 17% higher than that of PCA, 14.6% higher than that of NB, 16.8% higher than that of NN, 12.99% higher than that of SVM, 10.5% higher than that of B+SVM, 8.9% higher than that of SVDD+SVM, 6.3% higher than that of SVM+FCM, 4.8% higher than that of PCA+RE.

## 5 Conclusion

In this paper, we proposed red tide image recognition method, and the architecture of a system to implement it. The proposed method uses semantic feature image of NMF from 63 species of red tide algae in coastal areas of South Korea for training algae images which can recognize various species of red tide algae images.

**Acknowledgement.** "This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the C-ITRC (Convergence Information Technology Research Center) support program (NIPA-2013-H0401-13-2006) supervised by the NIPA (National IT Industry Promotion Agency)", This work was supported by Priority Research Centers Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology(2009-0093828)

## Reference

1. Red Tide Information System, Korea National Fisheries Research & Development Institute, <http://portal.nfrdi.re.kr/redtide/index.jsp>, (2012)
2. Y. M. Kim, Y. G. Byun, Y. Huh, K. Y. Yu, Detection of *Cochlodinium Polykrikoides* Red Tide Using MODIS Level 2 Data in Coastal Waters, Journal of Korean Society of Civil Engineers, pp.535-540, (2007)
3. T. Jiang, C. Wang, B. Wang, J. Xie, N. Jiao, T. Luo, Real-time Red Tide Algae Classification Using Bayes Classifier and SVM, In proceeding of International Conference on Bioinformatics and Biomedical Engineering (ICBBE), pp.2888-2891, (2008)
4. T. Jiang, C. Wang, B. Wang, J. Xie, N. Jia, T. Luo, Real-time Red Tide Algae Recognition using SVM and SVDD, In proceeding of International Conference on Intelligent Computing and Intelligent System (ICIS), pp.602-606, (2010)
5. X. Lili, J. Tao, X. Jiozhen, Z. Shaoping, Red Tide Algae Classification Using SVM-SNP and Semi-supervised FCM, In proceeding of International Conference on Education Technology and Computer (ICETC), pp.389-392, (2010)
6. S. Park, Y. K. Ji, B. R. Cha, Red Tide Image Recognition using Roundness and Entropy based on Semantic Feature, Advanced Science and Technology Letters (ASTL) 04, pp. 61-66, (2012)
7. D. D. Lee, H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," Nature, vol. 401, pp.788-791, 1999.
8. M. Turk, A. Pentland, Eigenfaces for Recognition, Journal of cognitive Neuroscience, Vol. 3, No. 1, pp.71-86, (1991)
9. R. C. Gonzalez, R. E. Woods, S. L. Eddins, Digital Image Processing Using MATLAB, New Jersey, Prentice Hall, (2003)
10. J. Han, M. Kamber, "Data Mining Concepts and Techniques", Morgan Kaufmann, (2006)
11. Red tide information on Korea marine environmental information site, <http://portal.nfrdi.re.kr/redtide/index.jsp>, (2013)
12. Google Image Search, <http://images.google.com/imghp?hl=EN>, (2013)