

Content-based Image Retrieval of Environmental Microorganisms Using Double-stage Optimisation-based Fusion

Yanling Zou ^{*}, Chen Li [†],
Kimiaki Shirahama [‡], Tao Jiang ^{*}, Marcin Grzegorzek [§]

Abstract

Environmental Microorganisms (EMs) are very tiny living beings which impact the entire biosphere by their environmental functions. Traditionally, a lot of manual efforts through morphological analysis using microscopes have been put on looking for EMs. However, these methods are time-consuming and laborious. To overcome this, we develop a *Content-based Image Retrieval* (CBIR) system for the EM image retrieval task within a *Double-stage Optimisation-based Fusion* framework. In the first stage, in order to effectively use the colour information of EM images, a *Multiple Colour Channel Fusion* (MCCR) method based on a *Particle Swarm Optimisation* (PSO) is developed to search for similar database images to a query image using local features. In the second stage, in order to enhance the retrieval performance of the first stage, a retrieval method based on *Immune Evolutionary Particle Swarm Optimisation - Shuffled Frog Leaping Algorithm* (IEPSO-SFLA) is devised to further combine global features. Finally, the experimental result shows that our double-stage fusion method obtains a mean average precision of 35.87% for 21 classes of EMs, which is superior to the existing methods.

Keywords: *Environmental Microorganisms, Content-based Image Retrieval, Microscopic Images, Double-stage Fusion, Multiple Colour Channel Fusion*

1 Introduction

Environmental Microorganisms (EMs) are present in every part of the biosphere (rivers, forests, mountains, etc.), playing critical roles in earth's biogeochemical cycles [1]. They are cost-effective agents for *in-situ* remediation of domestic, agricultural and industrial wastes and subsurface pollution in soils, sediments and marine environments. For example, *Vorticella* can digest organic pollutant in wastewater and improve the quality of fresh water. Traditionally, environmental microbiological researchers identify a new EM by checking references or asking experts. However, these traditional methods are time-consuming

^{*} Chengdu University of Information Technology(CUIT), Chengdu, China. Tao Jiang contributes the same as the contact author

[†] BMIE School, Northeastern University, Shenyang, China. Chen Li contributes the same as the first author

[‡] Pattern Recognition Group, University of Siegen, Siegen, Germany

[§] Department of Knowledge Engineering, University of Economics in Katowice, Poland

and inefficient. To improve the search effectiveness of EM information, especially for complex EM images, we develop a *Content-based Image Retrieval* (CBIR) system (EM-CBIR) [2, 3]. The system searches over a database to find EM images that are relevant to a query EM image. Through the CBIR method, features are automatically extracted from EM images to represent their semantic properties, such as colour, shape and texture, etc. Hence, EM-CBIR can help environmental researchers obtain useful EM information effectively and economically [4].

Since the contents of EM images are very complex to describe completely by a single feature, we combine different features in our EM-CBIR system to cover the colour, texture and shape information. In particular, our system is developed by addressing the following three issues: First, because the light colour and intensity of EM images are not stable, the colour feature extraction is affected. To manage this, we propose a *Multiple Colour Channel Fusion* (MCCF) approach to obtain more abundant colour information from different colour channels. Second, in different colour channels, we extract local features which are useful to describe various texture characteristics of EM images. Thirdly, global features are applied to describe the morphological properties of the EMs. Thus, these features show robust descriptive power in EM image retrieval tasks.

Putting the above three issues together, we develop a double-stage EM-CBIR system illustrated in Figure 1. Here, we aim to assist a user to quickly know what kind of EMs are included in an environmental sample. The first fusion stage is performed using the *MCCF* approach. This computes a similarity between a query and a database image by fusing their similarities in terms of local features on different colour channels. In the second stage, the resulting similarity is further combined with the similarity in terms of global features. Here, *Immune Evolutionary Particle Swarm Optimisation - Shuffled Frog Leaping Algorithm* (IEPSO-SFLA) is used to combine similarities by weighting them based on the correspondence accuracy [5], IEPSO-SFLA has the optimal searching ability and stability as well as the rapid convergence. Experimental results on real-world images show the effectiveness of the proposed system.

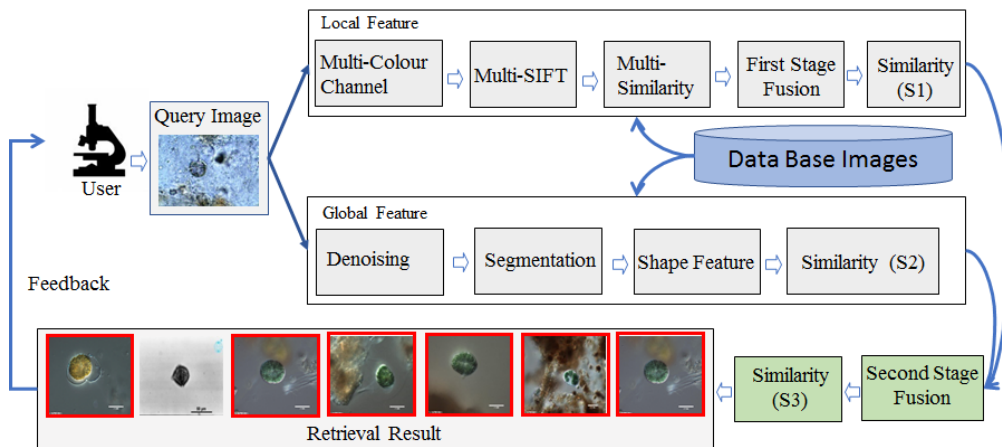


Figure 1: A work flow of our double-stage EM-CBIR system.

2 Related Work

EM image retrieval and classification: There are some existing works for EM image retrieval and classification. In [6, 7], a semi-automatic EM image classification approach is proposed. In this method, image segmentation, shape feature extraction and support vector machines are used to discriminate between regions of an interesting EM and the others. In [2], the characteristics of shapes in different segmented EM regions are represented by an edge-based feature descriptor and used for an image retrieval task.

All the aforementioned methods require image segmentation to extract shape features from EM images. Thus, their effectiveness depends on the accuracy of image segmentation, which is still a difficult task. In contrast, [8] introduces an EM classification approach that firstly uses Sparse Coding (SC) to extract useful local features by analysing a large number of image patches, and then employs Weakly Supervised Learning to perform the joint localisation and classification of an EM by finding a dense region of local features specific to this EM. However, because SC is computationally expensive, we particularly choose *Scale Invariant Feature Transform* (SIFT) features [9] as the local feature in our work. SIFT features are usually extracted from grey images, and ignores the useful colour information which finally hinders the ultimate performance. Hence, we adopt an approach to extract SIFT features from multiple colour channels.

Some works extract SIFT features by applying multiple colour channel strategy to obtain a larger amount of information. For example, [10] extracted SIFT features over three channels of HSV colour space, and [11] devises several SIFT feature extraction methods to effectively make use of colour spaces. While these methods only consider one colour space consisting of three colour channels, our method extracts SIFT features from much more colour channels (ten channels).

Fusion and optimisation: In [12], late fusion is used to enhance the classification and the clustering results in multi-view learning tasks, where features extracted from multiple camera views of a single object are effectively combined. There is another fusion approach, early fusion, which combines all features into a single high-dimensional vector. However, this causes the curse of dimensionality problem where important features may be weakened when it performs retrieval tasks based on this single vector. Late fusion exploits different features separately from each other, so it is more robust to the curse of dimensionality problem than early fusion. Thus, our double-stage fusion method uses late fusion [13], where similarity matrix in terms of local and global features are separately exploited and merged into one retrieval result.

There are several optimisation algorithms developed in the last years, such as *Particle Swarm Optimisation* (PSO) [14, 15], *Immune Evolutionary Algorithm* (IEA) [5] and *Shuffled Frog Leaping Algorithm* (SFLA) [16, 17]. These algorithms which have swarm intelligence are essentially random search algorithms. Algorithms described above can be used to solve practical problems, only in need of designing relative evaluation functions, independent of the strict mathematical feature and the structure characteristic for optimisation problems. Therefore, they have been widely used in many fields. Since we have ten local features for each EM image, the computation cost is expensive. Hence, we apply PSO in our first stage fusion work, which has fast convergence rate to guarantee the computational efficiency. IEP SO-SFLA integrates the global search idea of the PSO into SFLA, to simultaneously pursue the information of two optimal solutions in the sub-swarm and the whole-swarm, so as to search thoroughly near by the space gap of the worst solution. In addition, it integrates IEA into SFLA making immune evolutionary iterative computation

to the optimal solution in the whole-swarm, so as to fully use the information of optimal solution. This algorithm can not only be free from trapping into a local optimum, but also quickly converge close to the global optimal solution with a higher precision. Thus, IEPSO-SFLA is used to compute optimal weights for the second fusion stage.

3 Double-stage EM-CBIR Method

In this section, we present our double-stage fusion approach by sequentially describing the first fusion stage using local features, the global feature-based method, and the second fusion stage using IEPSO-SFLA optimisation.

First fusion stage: To increase the discrimination power of local features, we decompose an EM image into ten colour channels, as shown in Figure 2 [18, 19]. Then, we extract SIFT features on each channel. Especially, these are extracted from interest points detected by Harris corner detector [20]. In Figure 2, interest points detected on ten colour channels are depicted by yellow circles. As can see from this, different colour channels yield different interest points, from which we can extract SIFT features characterising various characteristics of the EM. Note that SIFT feature extraction does not need any pre-processing such as segmentation.

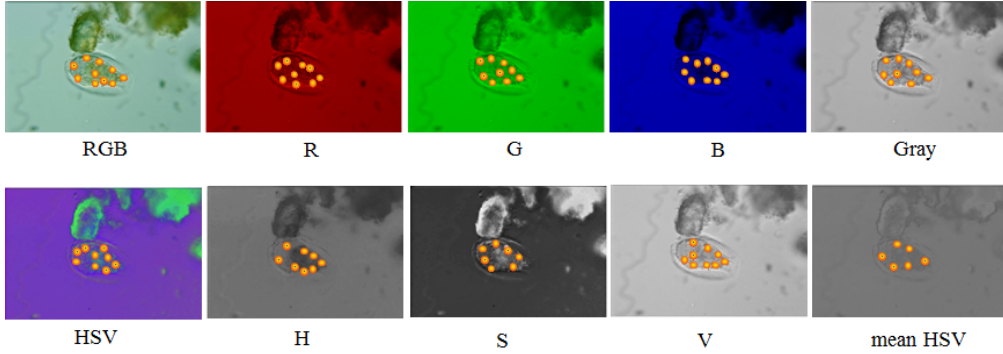


Figure 2: Examples of interest points detected on different colour channels.

For the j th channel ($1 \leq j \leq J(=10)$), we match SIFT features in a query image with the ones in a database image using random sample consensus (RANSAC) [21]. Their overall distance d_j is computed as the sum of Euclidean distances between matched SIFT features. To avoid any deviations from different value ranges, we normalise similarity for each channel so that the values range from 0 to 1. Afterwards, we define the similarity between the two images on the j th channel as $s_j = 1 - d_j$, and use MCCF to fuse similarities on J channels as follows:

$$S_1(J) = \frac{1}{J} \sum_{j=1}^J x_j s_j \quad , \quad (1)$$

where $\mathbf{x} = (x_1, \dots, x_J)^T$ is a weight vector in which x_j represents the weight for the similarity on the j th colour channel. Our late fusion method using PSO aims to compute the optimal weight vector $\hat{\mathbf{x}}$ that maximises the *mean of Average Precisions* (mAP) on $C = 21$ EM classes (see Figure 4). For each EM class, an *Average Precision* (AP) represents the average of precisions each of which is computed at the position where a relevant EM image is ranked. The value of an AP increases when relevant EM images are ranked at higher

positions. By taking the mean of such APs over C classes, the mAP expresses an overall retrieval performance using \mathbf{x} .

To find $\hat{\mathbf{x}}$, PSO considers a swarm of M particles (candidate weight vectors) $X = \{\mathbf{x}_m\}_{m=1}^M$. It iteratively updates the m th particle \mathbf{x}_m as $\mathbf{x}_m^t = \mathbf{x}_m^{t-1} + \mathbf{v}_m^t$ where t is the iteration index and \mathbf{v}_m^t is the velocity computed as follows:

$$\mathbf{v}_m^t = \psi \mathbf{v}_m^{t-1} + c_1 r_1 (l(\mathbf{x}_m)^t - \mathbf{x}_m^{t-1}) + c_2 r_2 (g(X)^t - \mathbf{x}_m^{t-1}), \quad (2)$$

where c_1 and c_2 are learning factors in the range of $[0, 2]$, r_1 and r_2 are random numbers between 0 and 1, and ψ is an inertial weight parameter that progressively decreases along iterations [22]. In addition, $l(\mathbf{x}_m)^t$ represents the best weight vector found by the m th particle \mathbf{x}_m until the t th iteration (i.e., this maximises mAP), and $g(X)^t$ is the best weight vector found by the whole set of M particles X . This way, inspired by the social behaviour of a swarm, PSO attempts to find $\hat{\mathbf{x}}$ by updating \mathbf{x}_m based on the local best solution $l(\mathbf{x}_m)^t$ and the global best one $g(X)^t$.

Global feature-based method: To ensure the accuracy of removing impurities in microscopic images, we use a semi-automatic segmentation method which purifies an EM region boundary that is roughly specified by a user [6]. Figure. 3 shows an overview of this method that combines conventional manual segmentation utilities with a novel automatic approach. First, as depicted in Figure. 3 (a), we use a cursor operation to draw a rectangle that approximates the region of an EM. Sobel edge detector is then applied to search the contour of this EM as shown in (b). Subsequently, morphological operations are used to fill holes and smooth the contour as in (c). Finally, (d) presents the final step where the foreground depicting the EM is selected by a single click to get the segmentation result in (e).

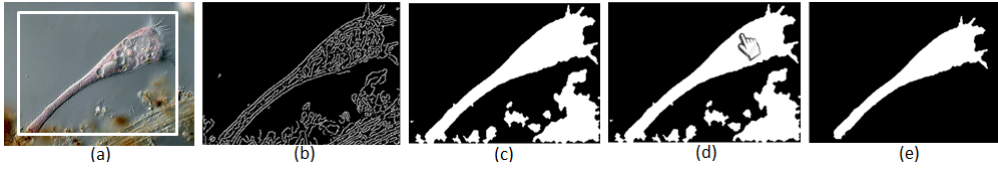


Figure 3: An illustration of our semi-automatic segmentation method.

To discriminate structural properties of each segmented EM region, we use *Internal Structure Histogram* (ISH) that is a contour-based shape feature and invariant to rotation and colour changes [2]. ISH is extracted by equidistantly distributing sample points on the contour of the EM region. Then, we create a histogram representing the distribution of angles, each of which is defined by a combination of three sample points. Finally, we obtain the similarity (S_2) by computing the similarity between each pair of EM regions as the Euclidean Distance between their global feature vectors.

Second fusion stage: For a query and a database image, we perform IEPSO-SFLA to combine their local feature-based similarity S_1 computed at the first fusion stage and their global feature-based similarity S_2 computed above:

$$S_3(I) = \frac{1}{I} \sum_{i=1}^I w_i S_i, \quad (3)$$

where $I = 2$ and $\mathbf{w} = (w_1, w_2)^T$ is a weight vector to fuse S_1 and S_2 . Using IEPSO-SFLA, we aim to find the optimal weight vector $\hat{\mathbf{w}}$ leading to the maximum mAP.

Before describing IEPSO-SFLA, SFLA is a biologically-inspired optimisation method that considers P ‘frogs’, each of which \mathbf{w}_p ($1 \leq p \leq P$) corresponds to a candidate weight vector. SFLA partitions frogs into α groups consisting of β frogs (i.e., $P = \alpha \times \beta$), and attempts to find the optimal $\hat{\mathbf{w}}$ by performing the local search in each group and shuffling search results over groups. However, SFLA is an unconstrained optimisation method, so it has problems of slow and premature convergence. Thus, IEPSO-SFLA is proposed by integrating the global search ideas of IEA and PSO into SFLA, so as to quickly converge to the optimal or a nearly optimal weight vector.

IEPSO-SFLA iteratively tries to refine three different frogs to better ones in the following way (the goodness of a frog is measured by the mAP obtained based on its weight vector). Frog $\mathbf{w}_{worst,old,basic}^k$ is the worst frog in the k th group and updated to $\mathbf{w}_{worst,new,basic}^k$ by Equation (4). Here, this update is based on the basic SFLA algorithm, so the index *basic* is used for $\mathbf{w}_{worst,old,basic}^k$ and $\mathbf{w}_{worst,new,basic}^k$. The second frog $\mathbf{w}_{best,old}^g$ is the best frog in the whole set of P frogs, and refined into $\mathbf{w}_{best,new}^g$. The last frog $\mathbf{w}_{worst,old}^k$ is the worst frog in the k th group, and Equation (6) is used to change it to $\mathbf{w}_{worst,new}^k$.

$$\mathbf{w}_{worst,new,basic}^k = \mathbf{w}_{worst,old}^k + r_1(\mathbf{w}_{best}^k - \mathbf{w}_{worst,old,basic}^k), \quad (4)$$

$$\mathbf{w}_{best,new}^g = \mathbf{w}_{best,old}^g + N(0, 1)\sigma_o e^{-\frac{At}{T}}, \quad (5)$$

$$\mathbf{w}_{worst,new}^k = \mathbf{w}_{worst,old}^k + r_1(\mathbf{w}_{best}^k - \mathbf{w}_{worst,old}^k) + r_2(\mathbf{w}_{best,new}^g - \mathbf{w}_{worst,old}^k), \quad (6)$$

where r_1 and r_2 are random numbers between 0 and 1. \mathbf{w}_{best}^k is the best frog in the k th group, $N(0, 1)$ is the random number sampled from the standard normal distribution, σ_o is the standard deviation corresponding to initial weight, A is the dynamic adjustment coefficient of the standard deviation (usually in the range of $[1, 10]$), t is the iteration index and T is the total number of iterations. By repeating the refinement of these three frogs, IEPSO-SFLA raises the overall goodness level of frogs to find the optimal $\hat{\mathbf{w}}$.

4 Experimental Results

4.1 Experiment Setting

We perform experiments on a real EM dataset (EMDS) consisting of 21 classes of EMs $\{\omega_1, \dots, \omega_{21}\}$ as shown in Figure 4. Each class is represented by 20 microscopic images. In our experiments, we use each EM image as a query image once and all the remaining images for testing. First, we evaluate the PSO-based MCCF method using local features. The EM-CBIR system extracts SIFT features from ten different colour channels (RGB, R, G, B, I, HSV, H, S, V and mean HSV). Then, we launch late fusion based on PSO for weighting similarities on 10 colour channels. Some tests are conducted using various swarm sizes in the range of $[50, 1200]$. We chose the best solution obtained with the swarm size of 50. Learning factors are set as $c_1 = c_2 = 2$. Second, we evaluate the IEPSO-SFLA based retrieval method using global features is combined for improving the retrieval performance of the first stage. According to preliminary experiments, we initialise SFLA partitions frogs P into $\alpha = 20$ groups consisting of $\beta = 42$ frogs, P is 840 (42×20). Normally, the dynamic adjustment coefficient of the standard deviation is initialised in the range of $[1, 10]$. Here, we fix A to be 1, and set $T = 10$ for the total number of iterations. Then, tests are conducted. Finally, we evaluate the performance of our system in terms of AP and mAP for all 21 classes.

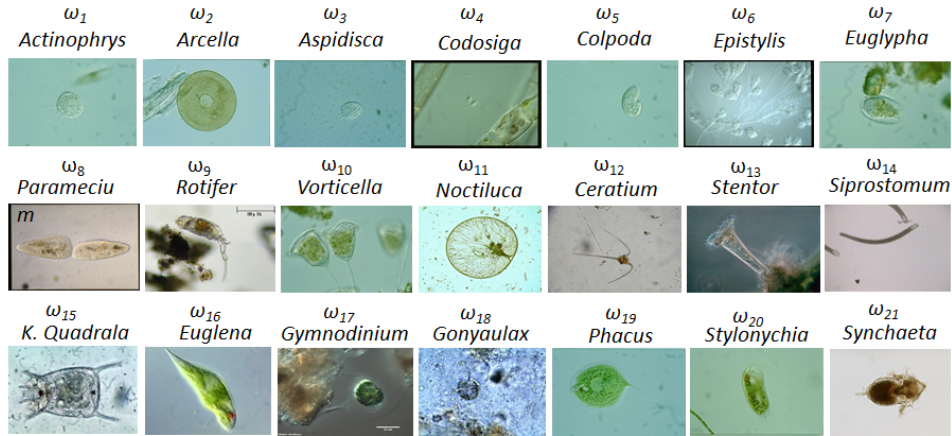


Figure 4: EM categories of EMDS.

4.2 Evaluation of Image Retrieval

In Figure 5 we compare APs and mAPs of ten colour channels and the first stage fusion using PSO on 21 classes of EM images. We can see that the first stage fusion significantly improves retrieval results. Among 18 of 21 classes, retrieval results are improved. In particular, it yields an impressive performance improvement for ω_2 , ω_4 , ω_5 and ω_6 , which have fewer key points only on a single colour channel. APs of the first stage fusion method for these classes amount to 44%, 35%, 35%, and 35%, respectively. By analysing the original EM images, we find that the performance gets worse for EMs which have high transparency, rough edge or more impurities, such as ω_{13} and ω_{17} in Figure 4. Table 1 presents the first stage fusion weights obtained by applying the PSO method on 10 colour channels. This shows that PSO works well for the RGB and the V channel, they are associated with the weights 0.72 and 0.86, respectively.

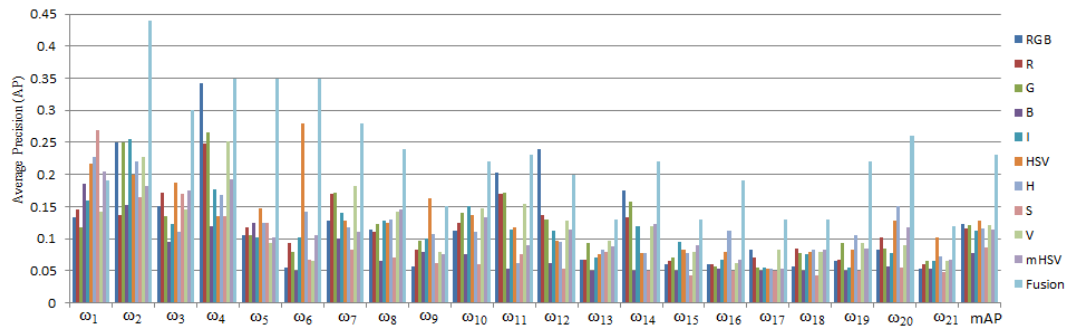


Figure 5: Evaluation of retrieval results using ten colour channels and the first stage fusion on 21 classes of EM images

Table 1: Fusion weights obtained by first stage fusion methods on ten colour channels. ‘I’ and ‘mHSV’ represent the channel Gray and mean HSV, respectively.

	RGB	R	G	B	I	HSV	H	S	V	mHSV
Weight	0.72	0.08	0.46	0.34	0.43	0.25	0.16	0.06	0.86	0.09

Figure 6 presents the comparison among retrieval results by the first stage fusion method

based on local features, results by the global feature-based method, and the ones obtained using the proposed double-stage optimisation-based method. The third method achieves the mAP of 35.87%, which obtains an improvement for ω_2 , ω_8 , ω_{11} , ω_{14} and ω_{17} . Compared to this, the mAP of the first stage fusion is 23%, the global feature method is 33.9%. Table 2 gives fusion weights obtained using the proposed double-stage optimisation-based method. This shows that these weights are specific to classes, and are not related to feature, as well as the similarity obtained by global feature-based method contributes more. Thus, the performance of our double-stage fusion method outperforms the other two methods.

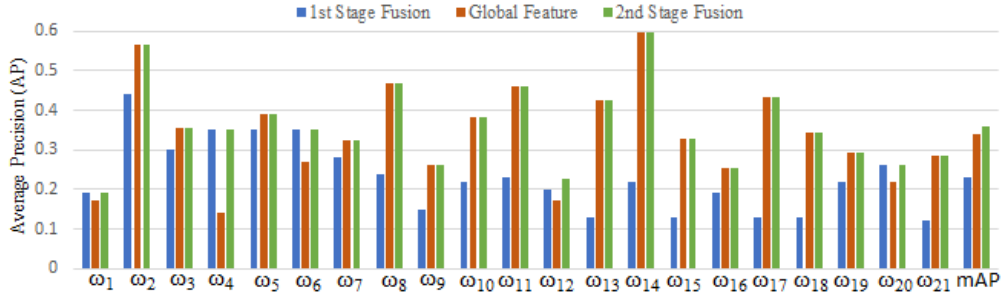


Figure 6: Comparison between the performances of the different fusion methods.

Table 2: Fusion weights obtained by the proposed double-stage fusion on 21 EM classes. $W(S_1)$ and $W(S_2)$ represent weights for the local and global feature-based method, respectively.

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}	ω_{11}
$W(S_1)$	0.10	0.45	0.20	0.20	0.25	0.20	0.15	0.35	0.10	0.20	0.30
$W(S_2)$	0.90	0.55	0.80	0.80	0.75	0.80	0.85	0.65	0.90	0.80	0.70
	ω_{12}	ω_{13}	ω_{14}	ω_{15}	ω_{16}	ω_{17}	ω_{18}	ω_{19}	ω_{20}	ω_{21}	
$W(S_1)$	0.45	0.20	0.15	0.30	0.25	0.10	0.15	0.15	0.15	0.20	
$W(S_2)$	0.55	0.80	0.85	0.70	0.75	0.90	0.85	0.85	0.85	0.80	

In addition, we compare the proposed method to several commonly used similarity measurements, includes Manhattan similarity, Euclidean similarity and Cosine similarity in EM retrieval. They achieve a mAP of 28.2%, 35.87% and 18.2%, respectively. We also compare our method to three other existing contour-based shape features, namely Shape Signature (SS), Fourier Descriptor (FD), and Shape Context (ShC) in retrieval work. The SS yields a mAP of 31.4%, the FD achieves 28.8%, and the ShC obtains 28.7%. Through the retrieval results, we can conclude that our method has the highest mAP of 35.87%.

Figure 7 shows the examples of retrieval results visually. The first column shows the query images. From the second to the last columns, the database images are sorted by their re-weighted similarities from similar to dissimilar. The images in red boxes are the images relevant to a query. This convincing result verifies the suitability of the proposed method which can accurately find a specify EM from many images where various EMs appear.

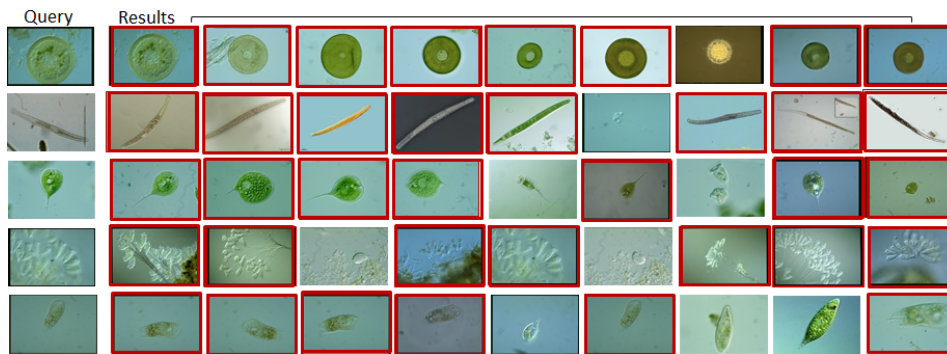


Figure 7: Examples of EM retrieval results using double-stage optimisation-based fusion.

5 Conclusion and Future Work

In this paper, we introduced an EM-CBIR method using double-stage optimisation-based fusion. Our method combines local feature-based and global feature-based approaches to enhance the final retrieval performance. In our experiments the double-stage optimisation-based fusion achieves a mAP of 35.87% and outperforms both the local feature and global feature based retrieval evaluation. This validates the effectiveness of the proposed approach. In our future work, we will adopt a deep convolutional neural network to improve the feature extraction performance of EM images [23].

Acknowledgments

Research activities leading to this work have been supported by China Scholarship Council (CSC), project (No. 2015GZ0197, 2015GZ0304) supported by Scientific Research Fund of SiChuan Provincial Science & Technology Department, project (No. 17ZB0080) supported by Scientific Research Fund of SiChuan Provincial Education Department.

References

- [1] I. L. Pepper, C. P. Gerba, and T. J. Gentry, *Environmental Microbiology*. San Diego, USA: Academic Press, 2014.
- [2] YL. Zou and C. Li and Z. Boukhers and K. Shirahama and T. Jiang and M. Grzegorzec, “Environmental microbiological content-based image retrieval system using internal structure histogram,” in *Proc. of CORES*, 2015, pp. 543–552.
- [3] YL. Zou and C. Li and K. Shirahama and T. Jiang and M. Grzegorzec, “Environmental microorganism image retrieval using multiple colour channels fusion and particle swarm optimisation,” in *Proc. of ICIP*, 2016, pp. 2475–2479.
- [4] C. Li, K. Wang, and N. Xu, “A Survey for the Applications of Content-based Microscopic Image Analysis in Microorganism Classification Domains,” *Artificial Intelligence Review*, pp. 1–70, 2017.

- [5] Z. Li, Z. Zhang, and C. Yu, “Shuffled frog leaping algorithm based on immune evolutionary particle swarm optimization,” *Journal of Computer Applications(in Chinese)*, vol. 31, no. 12, pp. 3288–3291, 2011.
- [6] C. Li, K. Shirahama, and M. Grzegorzec, “Application of content-based image analysis to environmental microorganism classification,” *Biocybernetics and Biomedical Engineering*, vol. 35, no. 1, pp. 10–21, 2015.
- [7] C. Li and K. Shirahama and M. Grzegorzec, “Environmental microbiology aided by content-based image analysis,” *Pattern Analysis and Applications*, vol. 19, no. 2, p. 531547, 2016.
- [8] C. Li, K. Shirahama, and M. Grzegorzec, “Environmental microorganism classification using sparse coding and weakly supervised learning,” in *Proc. of EMR@ICMR*, 2015, pp. 9–14.
- [9] D. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [10] A. Bosch, A. Zisserman, and X. Muoz, “Scene classification using a hybrid generative/discriminative approach,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 4, pp. 712–727, 2008.
- [11] K. van de Sande, T. Gevers, , and C. Snoek, “Evaluating colour descriptors for object and scene recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1582–1596, 2010.
- [12] J. Yu, M. Wang, and D. Tao, “Semisupervised multiview distance metric learning for cartoon synthesis,” *IEEE Transactions on Image Processing*, vol. 21, no. 11, pp. 4636–4648, 2012.
- [13] C. G. M. Snoek, M. Worring, and A. W. M. Smeulders, “Early versus late fusion in semantic video analysis,” in *Proc. of MM*, 2005, pp. 399–402.
- [14] J. Kennedy and R. C. Eberhart, “Particle swarm optimization,” in *Proc. of ICNN*, vol. 4, 1995, pp. 1942–1948.
- [15] R. C. Eberhart and Y. Shi, “Particle swarm optimization: developments, applications and resources,” in *Proc. of CEC*, vol. 1, 2001, pp. 81–86.
- [16] M. Eusuff and K. Lansey, “Optimization of water distribution network design using shuffled frog leaping algorithm,” *Journal of Water Resources Planning and Management*, vol. 129, no. 3, pp. 210–225, 2003.
- [17] E. Elbeltagi, T. Hegazy, and D. Grierson, “Comparison among five evolutionary-based optimization algorithms,” *Advanced Engineering Informatics*, vol. 19, pp. 43–53, 2005.
- [18] F. Perez and C. Koch, “Toward colour image segmentation in analog VLSI: algorithm and hardware,” *International Journal of Computer Vision*, vol. 12, no. 1, pp. 17–42, 1994.

- [19] H. Cheng, X. Jiang, A. Sun, and J. Wang, "Colour image segmentation: advances and prospects," *Pattern Recognition*, vol. 34, no. 12, pp. 2259–2281, 2001.
- [20] K. Mikolajczyk and C. Schmid, "Scale and affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 63–86, 2004.
- [21] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [22] M. G. Hinchey, R. Sterritt, and C. Rouff, "Swarms and swarm intelligence," *Computer*, vol. 40, no. 4, pp. 111–113, 2007.
- [23] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2016.