



Classification of coral reef images from underwater video using neural networks

Ma. Shiela Angeli C. Marcos, Maricor N. Soriano and Caesar A. Saloma

National Institute of Physics, University of the Philippines
1101 Diliman, Quezon City, Philippines
csaloma@nip.upd.edu.ph

Abstract: We use a feedforward backpropagation neural network to classify close-up images of coral reef components into three benthic categories: living coral, dead coral and sand. We have achieved a success rate of 86.5% (false positive = 6.7%) for test images that were not in the training set which is high considering that corals occur in an immense variety of appearance. Color and texture features derived from video stills of coral reef transects from the Great Barrier Reef were used as inputs to the network. We also developed a rule-based decision tree classifier according to how marine scientists classify corals from texture and color, and obtained a lower recognition rate of 79.7% for the same set of images.

©2005 Optical Society of America

OCIS Codes: (100.5010) Pattern recognition and feature extraction; (100.2000) Digital image processing; (330.1690) Color

References and Links

1. R. Kenchington and B. Hudson, *Coral Reef Management Handbook* (UNESCO 1984).
2. S. J. Purkis, "A 'reef-up' approach to classifying coral habitats from IKONOS imagery," *IEEE Trans. Geosci. Remote Sens.* **43**, 1375-1390 (2005).
3. E. J. Hochberg and M. J. Atkinson, "Spectral discrimination of coral reef benthic communities," *Coral Reefs* **19**, 164-171 (2000).
4. E. P. Green, P. J. Mumby, A. J. Edwards, and C. D. Clark, "A review of remote sensing for the assessment and management of tropical coastal resources," *Coastal Management* **24**, 1-40 (1996).
5. A. Uychiaco, P. Alino, and M. Atrigenio, "Video and Other Monitoring Techniques for Coral Reef Communities," in the *Proceedings of the 3rd ASEAN Science and Technology Week Conference, Marine Science: Living Coastal Resources*, L. M. Chou and C. R. Wilkinson, ed. (National University of Singapore and National Science and Technology Board, Singapore, 1992) pp. 35-40.
6. T. Carleton, J. and Done, "Qualitative video sampling of coral reef benthos: large-scale application," *Coral Reefs* **14**, 35 – 46 (1995).
7. A. Olmos and E. Trucco, "Detecting man-made objects in unconstrained subsea videos," in the *Proceedings of the British Machine Video Conference 2002*, P. L. Rosin and A. D. Marshall, ed. (Cardiff, U. K., 2002) pp. 517-526.
8. M. J. Rendas, M. Thomson and S. Rolfes, "Maerl mapping with an underwater vehicle using vision," presented in the International Symposium on Environmental Software Systems, Semmering, Austria, 27-30 May 2003.
9. V. Di Gesu, F. Isgro, D. Tegolo and E. Trucco, "Finding essential features for tracking starfish in video sequence," in the *Proceedings of the 12th IEEE International Conference on Image Analysis and Processing* **0**, M. Ferretti and M. Grazia Albanesi, ed. (Mantova, Italy, 2003) pp. 504-509.
10. M. Soriano, S. Marcos, C. Saloma, M. Quibilan and P. Aliño, "Image classification of coral reef components from underwater color video," in the *Proceedings of the MTS/IEEE OCEANS Conference* (Honolulu, Hawaii, 2001) pp. 1008-1013.
11. R. Duda, P. Hart and D. Stork, *Pattern Classification 2nd Edition* (John Wiley and Sons Inc., Canada, 2001).
12. A. M. Darwish and A. Jain, "A Rule Based Approach for Visual Pattern Inspection," *IEEE Trans. Pattern. Anal. Mach. Intell.* **10** (1), 56-68 (1988).
13. S. Katsuragawa, K. Doi, H. MacMahon, L. Monnier-Cholley, T. Ishida, and T. Kobayashi, "Classification of normal and abnormal lungs with interstitial diseases by rule-based method and artificial neural networks," *Journal of Digital Imaging* **10**(3), 108-14 (1997).
14. M. Swain and D. Ballard, "Color indexing," *International J. Computer Vision* **7**, 11-32 (1991).
15. M. Soriano, L. Garcia, and C. A. Saloma, "Fluorescent image classification by major color histograms and a neural network," *Opt. Express* **8**, 271-277 (2001), <http://www.opticsexpress.org/abstract.cfm?URI=OPEX-8-5-271>.
16. I. Pitas, *Digital Image Processing Algorithms* (Prentice Hall, London 1993).

17. T. Ojala, M. Pietikainen and T. Maenpää, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," IEEE Trans Pattern Analysis and Machine Intelligence **24** (7), 971-987 (2002).
18. S. Haykin, *Neural Networks: A Comprehensive Foundation 1st Edition* (Prentice Hall, New Jersey, USA, 1994).

1. Introduction

Conservation programs and monitoring surveys are conducted around the world to preserve coral reefs which play an important role in marine biodiversity and human livelihood. A basic way of monitoring reefs is to determine the population of different "benthos" or organisms in a reef area such as living or dead corals, algae, etc. [1]. It is utilized by marine scientists to deduce if a reef ecosystem is healthy, dying or at risk.

The current technology in reef surveys involves the use of satellite or airborne imagery of large spans of reefs, utilizing spectral analysis with typical image resolution of at best 0.5 – 1 meter per pixel. To correlate spectral image features to actual information such as living coral distribution, image processing, pattern recognition and water column correction are needed, as evidenced by numerous studies on this field [2, 3, 4]. However, to verify multi-spectral analyses, on-site inspection of the reef area, i.e. at a closer scale, is needed.

For on-site coral reef monitoring, video capture is the current protocol being adapted by marine scientists [1, 5, 6]. Manual counting of benthic populations are done in the laboratory. To obtain a statistical measure of the benthic distribution, reef objects underneath random points on the monitor screen are identified and counted while playing the reef video. The method is labor-intensive requiring a skilled eye and a substantial processing. In this study, we explore the possibility of developing a rapid and reliable computer automated system for classifying reef components from underwater video. We report high success rates in identifying coral reef images into three essential benthic groups (live corals, dead corals and sand/rubble) using only numerical descriptions of color and texture as feature inputs to a neural network (NN) classifier.

At present, we know of no other published works on image recognition of coral reefs at close range. There are only a few studies on benthic bottom classification using vision, some of which concentrate on identifying man-made objects in subsea videos [7], mapping maerl (coralline alga) habitats [8], and tracking starfish in video [9]. We emphasize that our study intends to classify benthic components which includes a variety of tropical coral images using underwater video only.

There are inherent difficulties in applying pattern recognition techniques to actual coral reef images. Corals occur in an immense variety of colors, textures, and shapes unlike faces or fingerprints which have a higher degree of similarity and well-defined features (e.g. face parts, minutiae). Corals are three-dimensional objects that can look differently at varying perspectives and scale. Marine scientists use color, texture and shape cues to classify coral reef images. In a previous work [10], we employed color and texture to classify six kinds of coral reef images (live corals, dead corals, dead corals with algae, abiotics, soft coral and other fauna). However, we obtained a poor average recognition rate (48%) because the identification task involved too many classes and the number of sample images were not equally distributed among the different classes. To improve performance, we now utilize color and texture to classify reef images according to three benthic groups: live corals, dead corals, and sand or rubble.

We use the NN as image classifier and compare our results against a rule-based classification scheme we call "two-step classifier" which are logic *if-then* rules set according to *a priori* knowledge of marine scientists on color-texture feature relationships descriptive of a certain benthos. Rule-based method is a type of decision-tree classifier that is widely used in segmentation, classification systems and database applications [11, 12, 13].

The paper proceeds as follows: Section 2 discusses the theoretical basis of our feature extraction and classifier details. Recognition performance of the neural network and two-step classifier is presented in Section 3 and Conclusions are discussed in Section 4.

2. Methodology

We consider three categories of coral reef components: live corals, dead corals, and sand or rubble. The three categories have significant amounts in our coral reef video ensemble. Color and texture features were extracted from the video data set that has been pre-classified by marine scientists from the Marine Science Institute of the University of the Philippines (<http://www.msi.upd.edu.ph>). The color feature should be invariant to changes in illumination, as underwater videos are prone to varying brightness. Also the texture feature should be rotation invariant, implying that texture description is preserved even with different perspective shots of the reef component. We note that scale is not a significant problem since video capture is maintained at a constant range from the reef surface (15 to 30 cm).

2.1 Color

To separate chromaticity information from brightness, we convert RGB image into a color space. Two color spaces were utilized to represent color information for the coral reef images: Normalized Chromaticity Coordinates (NCC) and Hue Saturation Value (HSV).

NCC [14] also known as normalized r - g color space, is a two-dimensional color space in which intensity information can be independent from the color. Color is represented by the proportion of red r , green g , and blue b values rather than the intensity of each. The corresponding r - g values are computed from the RGB space via the following expressions: $I = R + G + B$, $r = R/I$, and $g = G/I$, where I is the brightness or intensity value, and r and g are the chromaticity values. The chromaticity value $b = B/I$ can be derived from the constraint: $r + g + b = 1$. Mean r and g color histograms are computed for four major color groups: red-orange-yellow-brown color range, blue color, green color and white-gray. These major color groups were established through observations of the common hue ranges of benthic components [12]. The weighted value w corresponding to each major color is given by the dot product of the joint r - g color histogram C of the image and the mean histogram μ of the major color. For a sample i , we can compute w for a j^{th} major color as given by: $w_{ij} = C_i^T \mu_j$, where $i = 1, 2, \dots, N$ = number of samples, and $j = 1, 2, 3$ and 4 are bin numbers corresponding to major color groups [15]. This color feature will hereafter be referred to as “ r - g major color weights”.

The HSV color space has three attributes of color perception [16]. Hue H is measured as an angle between the actual color vector and a reference pure color vector (e.g. red). Saturation S is a measure of the percentage of white light added to a pure color. Value V corresponds to the brightness of the color. We compute the HSV components according to: $H = 0 + 60[(G - B)/(MAX - MIN)]$ if $R = MAX$, or $H = 120 + 60[(B - R)/(MAX - MIN)]$ if $G = MAX$, or $H = 240 + 60[(R - G)/(MAX - MIN)]$ if $B = MAX$, where: $S = (MAX - MIN)/MAX$, $V = MAX$, MAX is the maximum of the RGB values and MIN the minimum. We only use mean values of hue and saturation (mentioned in succeeding text as mean H-S) as a representation for the color feature.

2.2 Texture

Local Binary Patterns (LBP) is a texture descriptor that extracts statistics from spatial relationships of pixels. We chose LBP as a texture descriptor for coral images mainly because it is robust to brightness change and Gaussian blurring. It was also shown to be better at recognizing tilted three-dimensional textures than other texture paradigms [17].

Figure 1 illustrates the basic LBP operator. The gray-level values in a 3×3 pixel neighborhood containing eight pixels are thresholded to the value of the center pixel (see Fig. 1(a) and 1(b)). We assign 1 to neighboring pixel values greater than or equal to the center pixel value and 0 otherwise. The resulting binarized 3×3 pixel neighborhood is then multiplied with weights in powers of two (Fig. 1(c) and 1(d)). The sum of the neighborhood is unique and is assigned to the center pixel. The LBP operation is repeated for all pixels in the image frame to obtain the corresponding LBP image.

LBP8riu2 is a rotation invariant version of LBP, where the basic LBP operator is instead applied on a 3×3 circularly-symmetric neighbor set of eight pixels (see Fig. 2(a)). Only nine

LBP values are considered that exhibit the nine “uniform” patterns (see Fig. 2(b)) corresponding to the edges and spots of the image. The other rotation patterns that do not exhibit the nine uniform patterns are compressed into the 10th bin. The histogram of these values represents the texture feature of the image.

2.3 The Neural Network Classifier

NN is a pattern recognition model that does not need heuristics and recognition rules for classification [18]. It “learns” the rules of mapping by example. Although the NN normally requires a long training time to learn a specific task, it can find a complicated decision manifold that a linear pattern classifier cannot. We utilize the joint feature of color and texture as input to the network. Thus for each sample we have a 12 component vector for the joint mean H-S and LBP8riu2 histogram (hereafter referred to as HS-LBP) while we have 14 components for the joint r-g major color weights plus LBP8riu2 histogram (hereafter referred to as rg-LBP).

We employ a supervised three-layered feedforward, backpropagation NN using Matlab neural network toolbox. Weight updates are by gradient descent with a momentum term and an adaptive learning rate. The first two layers use the tangent and logarithmic sigmoid as activation functions, respectively. The NN is trained using the mean square error of 0.01 as the learning convergence criterion. Twelve (12) hidden units were utilized. Our choice of the number of hidden units falls under the rule of thumb in improving backpropagation as discussed by Duda and Hart in [11]. Momentum term was set to 0.9.

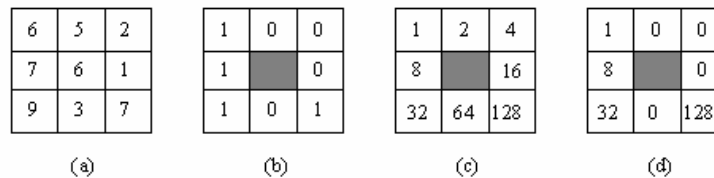


Fig. 1. The LBP operator. A 3x3 pixel neighborhood (a) is binarized into (b) by thresholding neighboring pixels with the center pixel. Weights in powers of two as in (c) is multiplied to (b) to obtain (d). The LBP value is computed as the sum of neighboring pixel values in (d), e.g. $1+8+32+128 = 169$. This value is assigned to the center pixel.

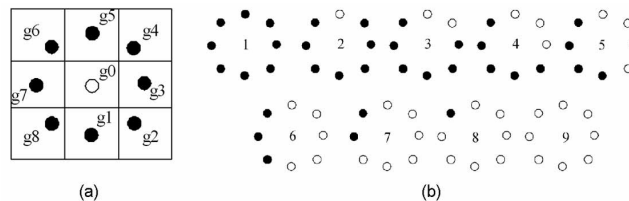


Fig. 2. Rotation invariant pixel neighborhood for LBP8riu2 in (a). Points g_6 , g_4 , g_8 and g_2 are obtained by bilinear interpolation. Binning is achieved through nine uniform patterns in (b) considered for the LBP8riu2 operator (black = 0, white = 1).

2.4 Two-Step Classifier

Live corals are dominantly colorful (ranging from reddish, brown, green, to blue hues) and have a regular or smooth texture. Dead corals are achromatic (gray or white). Sand on the other hand is achromatic and has a stochastic texture. Following such classification cues on the benthic groups, we proceed with a rule-based, decision-tree-type of classification based on two steps: first classifying whether the object is regularly, irregularly or smoothly textured and then determining if it is colorful or not (Fig. 3). *If-then* conditions are applied at each node. Using the r-g major color weights, colorfulness is identified by classifying an image to a major color group corresponding to $MAX(w_{ij})$. Linear Discriminant Analysis (LDA) is used to identify colorfulness when feature is mean H-S. Color permits us to resolve between live and dead coral. Classifying texture type is through a NN with similar architecture as discussed in

Section 2.3. The training set consists of images classified by the regularity, irregularity and smoothness of texture (Fig. 4(b)). We reiterate the difficulty of extracting unique features for living corals which occur in a variety of color and textures as seen in Fig. 4(b), first and third row.

2.5 Data Set

For the coral image data set, we used an underwater video of a coral reef transect from the Great Barrier Reef (Coral Navigator CD from the Australian Institute of Marine Science). Image blocks from 640×480 pixel-sized image frames are manually cut to comprise the training data set of 300 sub-images of live corals (149), dead corals (51) and sand (100). Figure 4(a) shows examples of the three categories. A total of 185 test images of varied block sizes, were obtained (live corals: 98 images, dead corals: 43 images, sand: 44 images).

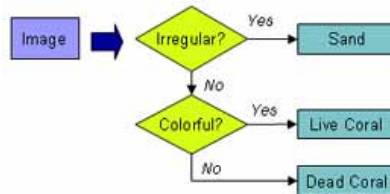


Fig. 3. Flowchart of two-step classification scheme for the coral reef images.



Fig. 4. (a) Image of a reef area from video. Shown are: **L** live coral, **D** dead coral and **S** sand/rock. (b) Examples of reef images with different textures. First row, images of regular texture; second row of irregular texture and third row of smooth texture. Images are of different sizes and cropped from image frames of a reef video.

3. Results and discussion

Figure 5 illustrates the confusion matrices of classification performance for the NN and two-step classifier. Clearly, the NN classifier performs better for coral component identification, showing a high recognition rate of 86.5 and 82.7% for features HS-LBP and rg-LBP, respectively. Average misclassification rate is 6.7 and 8.7%. The two-step classifier on the other hand, achieved a 79.7 and 79.3% success rate with misclassification rate of 10.1 and 10.3% for features HS-LBP and rg-LBP respectively.

We note however that although NN is the better classifier, there are still several advantages of using the two-step classifier. First is the simplicity of its design which uses logical if-then rules as opposed to NN that requires many free parameters and can be mathematically complex. Another benefit is flexibility in introducing new features into the rules. These advantages are promising for the development of coral reef image databases.

Because color is an important aspect in this research, we emphasize that the proposed system will require retraining for video taken at a different depth. At greater depths, compensation for the bluish cast made by seawater should be performed through the use of either appropriate filters or underwater lighting during video recording.

4. Conclusion

We have performed computer-based image classification of three coral reef components (live corals, dead corals and sand/rubble) from their corresponding underwater video images. A recognition rate of 86.5% was achieved with a NN classifier. Input features of the NN classifier were based on color and texture which are attributes used by marine scientists in determining various benthic categories of a reef area. The color features were derived from the *r-g* chromaticity histogram and mean hue-saturation values. LBP8riu2 was used for the texture feature. Ours is the first report of classification of benthic bottom types at close-up range.

Our study provides a groundwork for developing rapid automated systems for benthic classification and mapping using a simple reef transect video. Trade-offs between improved recognition rate and NN training time are worth investigating with the use additional image features other than color and texture. A full system can classify more benthic groups including algae, soft corals, and other fauna which also play crucial roles in marine conservation efforts. To test robustness of a classifier system, experiments on other reef areas around the world are desirable. For sublevel classification, e.g. genera or species level, more research on suitable image features and classifier design is needed.

		Predicted		
		LIVING CORAL	DEAD CORAL	SAND
Actual	LIVING CORAL	83 (84.7)	9 (9.2)	6 (6.1)
	DEAD CORAL	3 (7.0)	39 (90.7)	1 (2.3)
	SAND	4 (9.1)	3 (6.8)	37 (84.1)
OVERALL SUCCESS RATE		86.5		

(a) Confusion Matrix for NN classifier
Color feature: mean H-S
Texture feature: LBP8riu2 histogram

		Predicted		
		LIVING CORAL	DEAD CORAL	SAND
Actual	LIVING CORAL	81 (82.7)	7 (7.1)	10 (10.2)
	DEAD CORAL	11 (25.6)	32 (74.4)	0 (0)
	SAND	1 (2.3)	3 (6.8)	40 (90.9)
OVERALL SUCCESS RATE		82.7		

(b) Confusion Matrix for NN classifier
Color feature: *r-g* major color weights
Texture feature: LBP8riu2 histogram

		Predicted		
		LIVING CORAL	DEAD CORAL	SAND
Actual	LIVING CORAL	63 (64.3)	26 (26.5)	9 (9.2)
	DEAD CORAL	0 (0)	41 (95.3)	2 (4.7)
	SAND	5 (11.4)	4 (9.1)	35 (79.5)
OVERALL SUCCESS RATE		79.7		

(c) Confusion Matrix for Two-step classifier
Color feature: mean H-S
Texture feature: LBP8riu2 histogram

		Predicted		
		LIVING CORAL	DEAD CORAL	SAND
Actual	LIVING CORAL	80 (81.6)	9 (9.2)	9 (9.2)
	DEAD CORAL	8 (18.6)	33 (76.7)	2 (4.7)
	SAND	6 (13.6)	3 (6.8)	35 (79.5)
OVERALL SUCCESS RATE		79.3		

(d) Confusion Matrix for Two-step classifier
Color feature: *r-g* major color weights
Texture feature: LBP8riu2 histogram

Fig. 5. Confusion matrices of classification performance for the NN and two-step classifier. Each block contains the number of samples classified and beside it is its percentage over the total number of samples in the actual class. The diagonals shown as shaded blocks, indicates successful per class recognition rates. A perfect classifier is perceived as having 100% recognition rate in the diagonal elements. Columns in off-diagonal elements indicate misclassification (false-positive) rates. (a) and (b) reveal results for the NN classifier while (c) presents results for the two-step classifier.

Acknowledgments

M. Quibilan and P. Aliño (UP Marine Science Institute) for assistance and advice.