Proceedings of the 12th International Coral Reef Symposium, Cairns, Australia, 9-13 July 2012
5D Robots and machine learning techniques aiding coral reef science

Texture and Color Distribution-based Classification for Live Coral Detection

Joshua V. Stough[1], Lisa Greer[2], Matt Benson[2]

[1]Department of Computer Science
[2]Department of Geology, Washington and Lee University, Lexington, VA, USA
Corresponding author: stoughj@wlu.edu

Abstract. The endangered *Acropora cervicornis* coral has been considered a bellwether for coral reef habitat change. Given the recent discovery of a concentrated region of growth off Belize, our group is studying automatic abundance estimation of live *A. cervicornis* with the goal of future rapid assessment monitoring. In this paper we present a novel technique for the automatic segmentation of coral image sets. While others have had limited success applying machine learning techniques on color or texture-based features, our project presents several confounding factors in acquisition and in image content for which we must compensate. Our technique uses color features called quantile functions and SIFT texture features, and classifies local image regions as either live *A. cervicornis* or other image content using linear Support Vector Machines. We present promising results on a series of images for which we have manual segmentations to train with or test against. We also compare our results to those achieved using established raw color features. Our approach may not only greatly reduce the time cost of future abundance estimates of *A. cervicornis*, but also may be generalized to other coral vision problems.

Key words: *Acropora cervicornis*, Coral, Abundance estimation, Segmentation, Machine learning

Introduction

Acropora cervicornis (Staghorn coral) has been a prolific Caribbean framework reef builder throughout the Pleistocene (Jackson 1992). Yet this species has suffered a precipitous and well documented decline in the Caribbean-Atlantic Sector since the 1980's (Aronson and Precht 1997; Jackson 2001). It has been suggested that this collapse may be unprecedented for the recent geologic past (Aronson and Precht 2001; Greer et al. 2009). The species was added to the Endangered and Threatened Wildlife list (U.S. Department of the Interior) in 2007 and the once common coral is now rare on the vast majority of Caribbean reefs. Causes cited for the rapid decline of this (and other coral) species are virtually all tied directly or indirectly to human-induced environmental or climatic change (Hoegh-Guldberg 1999; Bellwood et al. 2004; Pandolfi et al. 2003) and white-band disease (Aronson and Precht 2001).

Most living *A. cervicornis* now exist in small patches and isolated colonies, and *A. cervicornis*-dominated 'reefs' are rare (Keck et al., 2005; Lirman et al. 2010). Our work suggests that an *A. cervicornis* refugium may exist off Ambergris Caye, Belize. In order to understand what environments/environmental factors promote survival and dominance of this species, skeletal framework and live tissue must be assessed in space and time. Given the apparent sensitivity of this coral to a number of possible stressors, the rapid rate of recent decline, and the rarity of reefs dominated by the species, rapid assessment of live tissue abundance and skeletal volume of this coral is needed at this potential refugium site.

Initial attempts at manual segmentation of live coral tissue from 62 calibrated m^2 quadrat photographs were successful but intensely laborious. It can take up to 4 hours to accurately digitize live tissue in each m^2 photograph of this morphologically complex species. Though point count methods are more efficient at low sample size, they may yield less than the desirable precision of 5%, given the heterogeneous and highly variable coverage in this dataset (Pante and Dustan 2012). Additionally, larger scale methods have shown limited accuracy (Hedley et al. 2004).

In order to assess the full extent of reef framework and live coral tissue, high precision rapid assessment is needed. In the following we describe our initial work toward automating this task through the use of machine learning.

Related Work

There has been extensive work in color and texture-based segmentation of biological images in general and coral habitat in particular (Cheng et al. 2001). Many successful methods fall under the term *supervised learning* and involve training some kind of classifier on local image features. When given a previously unlabeled feature, the classifier decides the label. These features are usually acquired using a training set of manually segmented or selected images. A classifier is fed these training features along with their class types and it determines some pattern for discriminating between the classes.

The astounding variability in classification methods comes from the variety of potential choices of both the image feature type and classifier. Marcos et al. (2005) used the combination of both texture features (as Local Binary Patterns) and color features (as Normalized Chromaticity and Hue-Saturation-Value) to train a neural network to classify living coral, dead coral, and sand on the Great Barrier Reef. Purser et al. (2009) used custom grating texture features and a self-organizing feature map to help differentiate between coral, sponge, and sand in cold-water coral habitats.

The work described in this paper most readily extends from Mehta et al. (2007). There, the authors used as a feature vector the concatenated raw pixel intensities within an N x N region, or *image patch*. Such features were computed at 100 epitomizing regions for corymbose, branching, and tabulate *Acropora* in a number of training images. Non-linear Support Vector Machines (SVMs) were used to provide a binary classification for each of the three coral morphology types in a one-against-all manner. The accuracy of classification varied among morphology types. SVMs can be useful classifiers in the context of highly variable coral imaging in that they are inherently stable in the presence of outliers.

Unlike the above work, our goal was to discriminate between live *A. cervicornis* and all other content, including other live branching and non-branching coral, dead *A. cervicornis*, algal cover, water, and hardground surfaces. Our project addresses several confounding factors in data acquisition and in image content that lead us to alternative choices from those in prior work.

Quadrat selection results in distortion that blurs the characteristic texture of *A. cervicornis* in parts of each image. Additionally, the image objects are viewed from many perspectives within and across images due to surface wave disturbance during photo acquisition. We have observed that under these conditions, the raw patch feature of Mehta et al. (2007), which assumes a pixel-scale correspondence across patches, results in little information gained about the collection of patches beyond their average color. To accommodate this variability and recover more subtle information about collections of patches, we used the regional intensity quantile function (QF) as a color feature. The QF captures the *distribution* of intensities within an image patch. Furthermore, these features are rotationally invariant and are amenable to linear statistical methods, such as the relatively efficient linear SVM that we use to determine our classifier.

In the following we describe QFs and their advantageous use in this context, and to a lesser

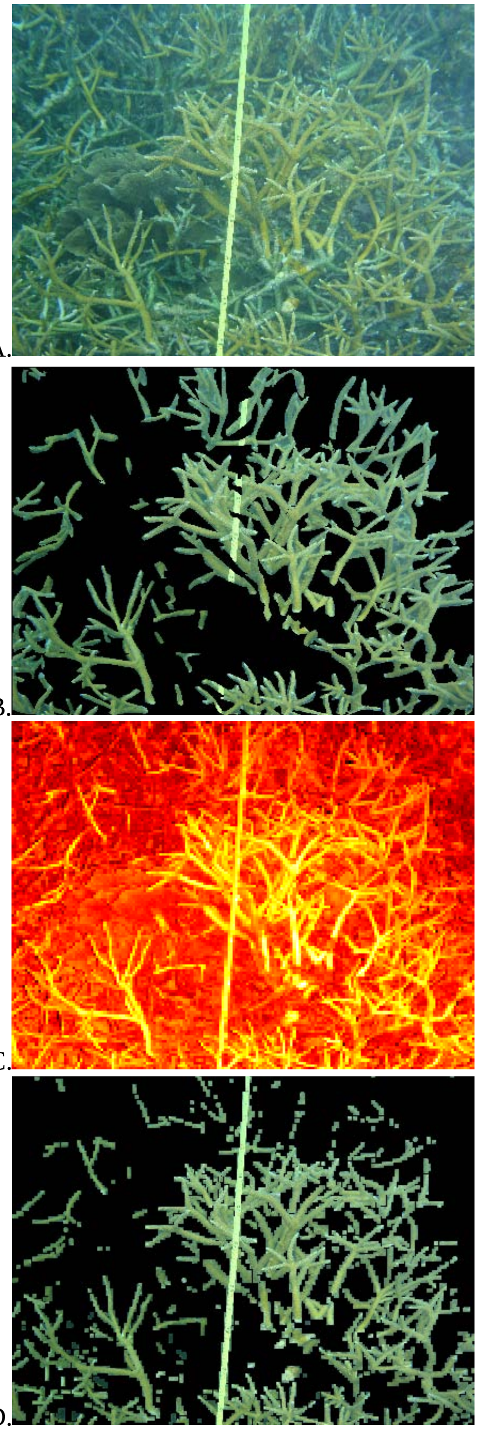
A.
B.
C.
D.


Figure 1: Example result (top to bottom): A) typical reef quadrat image; B) manual trace of live *A. cervicornis*; C) visualization of the local color-based classifier score; D) final automatic segmentation. Note the manual result both includes and excludes the tape. The automatic result reflects a Dice coefficient of .745 (see Results).

extent we describe our use of the well-established SIFT texture feature and linear SVM classifier (Lowe 1999; Cortes and Vapnik 1995). We present results on our 62-image series using both the color and texture features computed at multiple scales (Fig. 1).

Materials and Methods

This section includes a detailed description of the construction of QFs, followed by brief reviews of SIFTs and linear SVMs. We conclude with our experimental setup.

Regional Intensity Quantile Functions

Broadhurst et al. (2006) describe an approach for probabilistically representing the appearance of an object region within an image in the context of medical image segmentation. The basic unit of appearance is the regional intensity quantile function (QF), derived from the single-channel intensity histogram within an image patch. Quantile functions were computed for each of the red, green, and blue color channels and the concatenated result was used as the patch feature provided to the SVM. Fig. 2 shows the average QFs for live *A. cervicornis* and anything else.

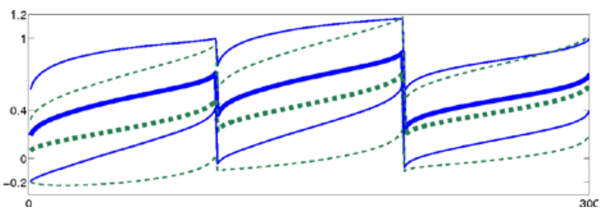


Figure 2: QF feature variability over a dense grid in an example image. Bold for live *A. cervicornis* and dotted for other, with mean (thick) and ±2σ (thin) curves, which can indicate intensities outside the possible [0,1]. Note that live tends to be brighter, while the significant overlap suggests the difficulty of classification here.

Quantile functions are a useful parameterization of one-dimensional distributions. For example, certain common changes in a distribution, such as mean shift and variance scaling, are represented as linear changes in the QF feature space. The QF space thus confers improved efficiency to the extent that these are common differences for corresponding distributions to have, as in different patches of *A. cervicornis* within an image.

If q and r are the continuous, one-dimensional intensity distributions in two regions between which we wish to measure the similarity. The Mallows, or Earth Mover's distance (Levina 2002) between q and r, with cumulative distribution functions Q and R, respectively, is defined as

$$M_p(q,r) = \left(\int_0^1 |Q^{-1}(t) - R^{-1}(t)|^p dt\right)^{1/p}.$$

An n-dimensional QF is then the discretized inverse cumulative distribution function on intensities in a region, i.e., $Q^{-1}(t)$ or $R^{-1}(t)$ in the above equation. Rather than provide the percentile given the variable value as the cumulative distribution does, the QF provides the variable value given the percentile score. Let these discretized quantile functions be denoted **q** or **r**. Coordinate j of **q** or **r** stores the average of the [(j−1)/n , j/n] quantile of the intensity distribution for that patch, i.e,

$$\mathbf{q}_j = \int_{(j-1)/n}^{j/n} Q^{-1}(t)dt$$

After discretization, the Mallows distance above corresponds (up to a scale factor) to the L_p vector norm between **q** and **r**.

Through quantile functions, patch-scale intensity distributions are understood as points in an n-dimensional Euclidean space in which distance corresponds to the M_2 metric, and mean and variance changes in intensities are linear.

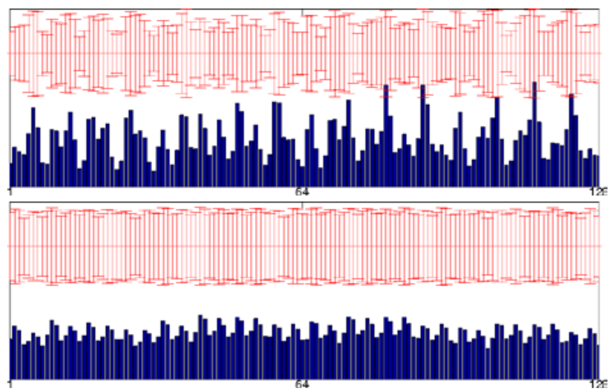


Figure 3: Average SIFT features over a dense grid in an example image (blue), with a visualization of each component's standard deviation (red), for live tissue (top) and other (bottom). Note that the live coral, as might be expected, tends to have higher contrast and more consistent structure than the class of everything else.

SIFT: Scale-Invariant Feature Transform

The SIFT texture feature is a representation of the distribution of gradients (or edge orientations) within some image patch (Lowe 1999). This feature has been extensively proven in object recognition tasks, where a particular object or class of object is viewed from varying orientations and scales across images (Mikolajczyk and Schmid 2005). With this in mind, we tested the ability of SIFT features to discriminate live *A. cervicornis* through a linear SVM classifier.

We used an open MATLAB implementation which decomposes each image patch into 16 equal parts, determining for each part an 8-dimensional histogram of the gradient energy in the 8 cardinal directions (Lazebnik et al. 2006). The resulting feature is thus 128 dimensions for each image patch (see Fig. 3).

Linear Support Vector Machines

Support Vector Machines (SVMs) have become a popular means of learning two-group classifiers since their introduction by Cortes and Vapnik (1995). The most basic algorithm learns a discriminating hyperplane in a mapping of a feature space for which there is a sample population, including explicitly labeled positive and negative cases. In the case where features in the unmapped (or linearly mapped) space are thought to be separable, linear SVM is an efficient algorithm for determining the decision boundary given by

$$\mathbf{w}' \cdot \mathbf{q} + b > 0,$$

where **w** and **q** are column vectors representing the hyperplane normal and feature descriptor, and *b* is the distance of the hyperplane to the origin in the space. Note that the classifier response is continuous and can be considered a confidence (Fig. 1). In validation, these confidences can be used along with the correct label for each patch to generate a precision-recall curve describing the strength of the classifier.

In our case, features were acquired from a dense grid of patches in a training image, where the labeling was inferred by comparing the patch support to a binary image of the manual trace (Fig. 4). QF and SIFT features were trained separately, as they exist in inherently different spaces. We used an open MATLAB implementation for linear SVM training (Yang et al. 2009).

Figure 4: Close-up of *A. cervicornis* with manual trace outline. Average basal branch thickness is ~1.5 cm)

Experimental Setup

Video line-transect and photographic m^2 quadrat survey data were collected in 2011. A total of sixty-two quadrat photos were digitized and scaled, and all live *A. cervicornis* tissue was manually segmented in map view (Fig. 4). We consider this sample representative of the larger local reef structure, though that belief will be tested in future surveys.

Due to highly variable white balance within images, results were improved by the additional preprocessing step of adaptive contrast enhancement, simply equalizing the local luminance across the image while maintaining color.

For each image we sampled a dense grid of non-overlapping patches, with different grids for different scale patches (we tested 8, 16, 24, and 32 pixel square patches). These image patches along with their corresponding labels from the manual segmentation were fed to a linear SVM. We thus obtained classifiers for each scale within each image separately.

We tested each image against the training SVMs given by the other images in leave-one-out tests. In practice it would not make sense to manually trace all but one image, but for this initial experiment, it is informative to have all possible results given the choice of training image (see Discussion).

Results

We report the average precision (AP), or area under the precision-recall curve, as well as the Dice Similarity Coefficient (DSC), common measures in the image segmentation literature. AP measures the strength of a classifier over all possible decision boundaries. The DSC is the ratio of the intersection and the average of the two classifiers (automatic and manual) given a particular decision boundary (here, an SVM response greater than zero).

Our method achieved an overall AP of 72.1% using the color QFs as features, versus 71.4% using the raw color feature described in Mehta et al. (2007). Further, color QFs led to an average improvement in AP of 2.2% in 34 of 62 cases, with an average decline of 1.1% in the others. Thus, color QFs were better more often and better by more, though the raw color features were similarly successful overall. Also of note is that color QFs outperformed raw color features in 19 of the 25 difficult cases of less than average AP.

The overall DSC was 63.8%, with the classifier exceeding 60% DSC in 29 of 62 images. These results are consistent with validation literature in other fields in the case of long thin structures with a large boundary relative to area (Gerig et al. 2001). Fig. 5 shows examples achieving both low and high DSC. The abundance estimate over the entire study was 26.7%, versus 32.3% with the manual raters.

Results using the SIFT texture features were extremely noisy and appear to be of little value, at least given our training scheme. The SIFT features led to an average DSC of 17.1 %.

Discussion

In this paper we presented results of a supervised learning system for the segmentation of live *A. cervicornis* in images with strong confounding factors. We achieved encouraging success using intensity quantile functions on color channels, which lead to improved results over raw color features.

In continuing on this front, it will be difficult to qualify our results without some estimates of inter-rater variability. Anecdotal evidence suggests there is some difference in the manual tracings of different

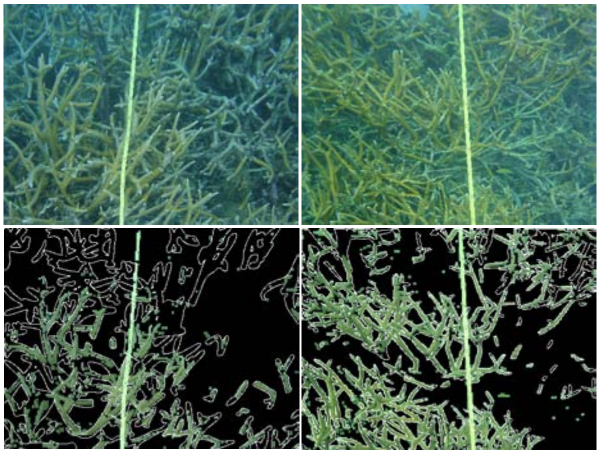


Figure 5: Example results with traces. Left column: example image on which the classifier achieved 52% DSC. Right: 76%.

raters over such a collection of high-resolution photographs. The goal of any automatic classifier can only be to differ from human raters no more than human raters differ from each other.

There are a number of directions to explore in potentially improving the segmentation quality and thus live tissue abundance estimates. While the SIFT features appear less informative than the QF-based features, there are published methods of more effectively combining classifiers of varying accuracy (Rohlfing et al. 2004). We will also consider alternative machine learning techniques to apply to our features for comparison against linear SVM.

Another goal is to determine the appropriate training image to use as the statistical atlas for any test case. Simple measures, such as average color similarity, do not correlate with the outcome AP. We must find what constitutes an effective training set.

Lastly, ongoing work is focused on the pre-process phase where we hope to correct the effect of the measuring tape and of errors in the quadrat selection that affect both our results and computed abundance estimates. In fact, it should be possible to perform the presented classification prior to quadrat selection. The longer term goal is to better automate rapid assessment monitoring efforts post acquisition. We envision a system with all the various post-acquisition steps integrated, from tract selection in photography to interactive trace correction using an adaptive classifier based on the presented color quantiles and potentially texture features.

Ultimately, we hope the methods generated here and in refinement will provide a crucial speed and efficiency for *A. cervicornis* coral monitoring efforts in the few remaining documented refugia.

Acknowledgement

Funding from the endowed Lenfest and RE Lee research funds, Washington and Lee University. We are grateful for field and laboratory support from H. Allen Curran and William Sullivan.

References

Aronson RB, and Precht WF (1997) Stasis, biological disturbance, and community structure of a Holocene coral reef. Paleobiology 23(3): 326-346

Bellwood DR, Hughes TP, Folke C, and Nyström M (2004) Confronting the coral reef crisis. Nature 429(6994): 827-833

Broadhurst RE, Stough J, Pizer S, and Chaney EL (2006) A statistical appearance model based on intensity quantile histograms. Proc Int Symp Biomedical Imag (ISBI)

Cheng HD, Jiang XH, and Li Wang J (2001) Color image segmentation: Advances and prospects. Patt Rec 34:2259-2281

Cortes C, Vapnik V (1995) Support-vector networks. Machine Learning. 20: 273-297

Greer L, Jackson J, Curran HA, Guilderson T, and Teneva L (2009) How vulnerable is Acropora cervicornis to environmental change? Lessons from the early to mid-Holocene. Geology 37: 263-266

Hedley JD, Mumby PJ, Joyce KE, and Phinn SR (2004) Spectral unmixing of coral reef benthos under ideal conditions. Coral Reefs 23: 60-73

Hoegh-Guldberg O (1999) Climate change, coral bleaching and the future of the world's coral reefs. Marine and Freshwater Research 50(8): 839-866

Jackson JBC (1992) Pleistocene perspectives on coral reef community structure. American Zoologist 32: 719-731.

Jackson, JBC (2001) What was natural in the coastal oceans. Proc. Nat Acad Sci 98(10): 5411-5418

Keck J, Houston RS, Purkis S, Riegl BM (2005) Unexpectedly high cover of Acropora cervicornis on offshore reefs off Roatan (Honduras). Coral Reefs 24:509

Lazebnik S, Schmid C, and Ponce J (2006) Beyond bags of features: spatial pyramid matching for recognizing natural scene categories. Proc Int Conf Comp Vis Pat Rec (CVPR).

Levina E (2002) Statistical Issues in Texture Analysis. PhD thesis, University of California at Berkeley

Lirman D, Bowden-Kerby A, Schopmeyer S, Huntingdon B, Thyberg T, Gough M, Gough T, Gough R, and Gough Y (2010) A window to the past: documenting the status of one of the last 'megapopulations' of the threatened Staghorn coral Acropora cervicornis in the Dominican Republic. Aq Conserv Mar Fresh Ecosyst. DOI: 10,1002/aqc. 1146.

Lowe DG (1999) Object recognition from local scale-invariant features. Proc Int Conf Comp Vis (ICCV) 2:1150-1157

Marcos MSAC, Soriano M, and Saloma C (2005) Classification of coral reef images from underwater video using neural networks. Optics Express Vol. 13, No. 22

Mathworks (2010) MATLAB version 7.10.0 (R2010a). Natick, MA

Mikolajczyk K, and Schmid C (2005) A performance evaluation of local descriptors. IEEE Trans. PAMI 27(10): 1615-1630

Mehta A, Ribeiro E, Gilner J, and van Woesik R (2007) Coral reef texture classification using support vector machines. Int Conf on Comp Vis Theory and Applications (VISAPP)

Pandolfi JM, Bradbury RH, Sala E, Hughes TP, Bjorndal KA, Cooke RG, McArdle D, McClenachan L, Newman MJH, Paredes G, Warner RR, Jackson JBC (2003) Global trajectories of the long-term decline of coral reef ecosystems. Science 301(5635): 955-958

Pante E, and Dustin P (2012) Getting to the point: Accuracy of point-counting methods in monitoring ecosystem change. J Mar Bio (Hindawi Press) Article ID 802875

Purser A, Bergmann M, Lundälv T, Ontrup J, and Nattkemper T (2009) Use of machine-learning algorithms for the automated detection of cold-water coral habitats: a pilot study. Mar Ecol Prog Ser 397: 241–251

Rohlfing T, Russakoff DB, and Maurer CR (2004) Performance-based classifier combination in atlas-based image segmentation using expectation-maximization and parameter estimation. IEEE TMI 23:8, p 983-994

Yang J, Yu K, Gong Y, Huang T (2009) Linear spatial pyramid matching using sparse coding for image classification. CVPR