Machine Learning and Multiscale Methods in the Identification of Bivalve Larvae

Sanjay Tiwari
Biology Department
Woods Hole Oceanographic Institution
Woods Hole, MA 02543

Scott Gallager
Biology Department
Woods Hole Oceanographic Institution
Woods Hole, MA 02543

Abstract

This paper describes a novel application of support vector machines and multiscale texture and color invariants to a problem in biological oceanography: the identification of 6 species of bivalve larvae. Our data consists of polarized color images of scallop and other bivalve larvae (between 2 and 17 days old) collected from the ocean by a shipboard optical imaging system of our design. Larvae of scallops, clams, and oysters are small (100 microns) with few distinguishing features when observed under standard light microscopy. However, the use of polarized light with a full wave retardation plate produces a vivid color, bi-refringence pattern. The patterns display very subtle differences between species, often not discernable to human observers. We show that a soft-margin support vector machine with Gaussian RBF kernel is a good discriminator on a feature set extracted from Gabor wavelet transforms and color distribution angles of each image. By constraining the Gabor center frequencies to be low, the resulting system can attain classification accuracy in excess of 90% for vertically oriented images, and in excess of 80% for randomly oriented images.

1 Introduction

Identification of bivalve larvae such as clams, oysters, and scallops is required to effectively manage these commercially important shellfish resources. For years researchers have searched for distinguishing, species-specific features for identifying larvae. Although scanning electron microscopy (SEM) is regarded as the most reliable approach for feature identification, the technique is very laborious requiring days to process just a few larvae. The optical approach to larval identification described here lends itself to non-destructive, rapid, flow-through optical systems where hundreds of larvae could be identified in seconds. Our imaging system has been put into use on research vessels, where this approach facilitates our ability to map the abundance and distribution of bivalve larvae in the world ocean and allows us to make rapid predictions of where the ocean currents will carry the larvae before settlement on the bottom.

1.1 Biological background

Many bivalve mollusks produce planktonic larvae which progress through a series of developmental stages before settling to the bottom and taking up a benthic existence. The larval shell is mineralized with aragonitic crystals which nucleate and grow on an organic matrix. Crystal optical orientation is controlled by protein organization within the matrix. Under polarized light, the shell exhibits light extinction where the optic axes of the two polarizing filters are normal to one another demonstrating that crystal orientation is arranged radially around the location of nucleation.

With the addition of a full wave retardation plate, distinct colored patterns are produced as the polarized light travels through the birefringent crystals, combines and cancels. Each color represents the degree of axial rotation of the crystal at that location. Because the protein compliment of the shell matrix is unique to given species, crystal axial rotation and, therefore, the color pattern produced under polarized light is also unique to that species. However, species-specificity in the color pattern is relatively subtle thereby requiring sophisticated statistical approaches to quantify differences.

1.2 The Approach

We characterize here bi-refringent spectral patterns for the larvae of six species of bivalves: *Argopecten irradians* (Bay Scallop) (Figure 1), *Crassostrea Viginica* (American Oyster) (Figure 2), *Mercenaria mercenaria* (Hard-shelled clam) (Figure 3), *Mya arenaria* (Sand gaper) (Figure 4), *Placopecten magellanicus* (Sea Scallop) (Figure 5), and *Spisula solidissima* (Atlantic surf Clam) (Figure 6). The inputs for our pattern recognition system are polarized color images of the six species. Since the problem of *species* recognition is complicated by the variation *within* a species,

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

IEEE COMPUTER SOCIETY

a good classifier has to have enough flexibility to accommodate this intra-species variation, while still working efficiently to capture inter-species variation. Our images are so alike that trained marine biologists cannot identify them with greater than 50% accuracy. In fact, since we cannot independently identify larvae from the ocean with accuracy, specimens for the training set must be raised in monocultures in the lab. Thus for us, an automated identification system is not an aid (like in some bio-medical applications), but an essential need. The performance of the classifier also depends on how accurately the feature sets capture species-related features. We show that a *soft margin* Support Vector Machine applied to feature sets constructed from *Gabor* transforms and *color angles* of the images, performs this task with classification accuracy of over 90% when the images are presented in a vertical orientation (as a result of preprocessing) and over 80% when the images are presented in random orientations. In the remainder of the paper, after reviewing related work in section 1.3, we describe Support Vector Machines (SVMs) in section 2. The texture features, which are constructed from Gabor wavelet transforms of different components of a color image, are described in section 3 and the rotation-invariant features derived from these follow in section 4. The color features, which use the elegant construction of the color distribution angles, are described in section 5. Finally, in section 6, we evaluate different SVM classifiers and different feature sets in the task of discriminating the larval images by species.

1.3 Related Work

Support Vector Machines (SVMs) have a burgeoning literature: good introduction are [2], [19] and [8]. SVMs have been successfully used in various visual pattern recognition tasks, like object (specially, face) detection and recognition and image retrieval: In [15] and [12] a trainable system for object detection using SVMs is described. In [11] an application to face detection is described; in [18] a view-based object SVM recognition system is described; and in [1], color and luminescence information is exploited for classifications with SVMs.

Gabor functions have a long history of work starting in the 1940s. In [4], both this history and relations to areas of mathematics (especially Weyl-Heisenberg Frames and the Balian-Low Theorem) and applications are all explored. In Texture segmentation, Gabor analysis has been in use since the early 1980s: In [7], a polar, analytic form of a 2-D Gabor wavelet is developed and a multiresolution family of these wavelets is used to compute information-conserving *micro* and *macro* features. In [10], the use of Gabor wavelet features is proposed for content-based image retrieval. In [6], a rotation-invariant texture classification is described using Steerable Pyramids. In [16], another rotation-invariant texture classification method is described which is robust with respect to noisy conditions. Various comparative analyses between different multiscale texture segmentation schemes have been made by: [13] and [17]. In [3], "IrisCode" is constructed by demodulation of the iris pattern.

The color distribution angles are defined in [5] and shown to capture important low-order statistical information about color and edge distributions. The use of color-angle based features in multiscale representation of natural images has been further explored in [21]. The use of HSV (Hue, Saturation, and Value) space to split an image into chromatic and achromatic channels is explored in [9].

2 Support Vector Classification

Support Vector Machines (SVMs) create a function from a set of labeled training data. The function can be a classification function or a general regression function. SVMs are capable of learning in *high dimensional spaces* with small training set, which they accomplish by minimizing a bound on the empirical error and the complexity of the classifier at the same time [12]. For binary classification, given a *pattern space* of inputs, SVMs operate by finding a hypersurface in the space which attempts to split the positive examples from the negative ones. The SVM algorithm formulates the training problem as one that finds, among all possible separating hypersurfaces, the one that maximizes the distance between the closest elements of the two classes (negative and positive). We give a more precise description below.

2.1 Binary Classification

For a given pattern space, $\mathbf{X}$, suppose we are given labeled training

$$\mathcal{T} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{\ell}, \quad \mathbf{x}_i \in \mathbf{X}, \quad y_i \in \mathbf{Y} = \{-1, +1\},$$

Suppose $\mathbf{X} \subset \mathbb{R}^d$, with the inherited inner product $\langle ., . \rangle$. Then, among all hyperplanes $\mathbf{H}$ (defined by: $\langle \mathbf{w}, \mathbf{x} \rangle + b = 0$), separating the data, there exists a unique *optimal hyperplane*, distinguished by the maximum margin of separation between any training point and the hyperplane. It is the solution of:

$$\underset{\mathbf{w}\in\mathbf{H},\ b\in\mathbb{R}}{\text{minimize}}\ \tau(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|^2$$

(where $1/\|\mathbf{w}\|$ is the width of the margin) subject to the inequality constraints:

$$y_i(\langle \mathbf{w}, \mathbf{x}_i \rangle + b) \geq 1 \ \forall\ i = 1, \ldots m. \tag{1}$$

In practice, a separating hyperplane may not exist (for example, because of a high noise level). To allow for the

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

IEEE COMPUTER SOCIETY

possibility of examples violating (1), one introduces slack variables:

$$\xi_i \geq 0 \, \forall \, i = 1, \ldots, m. \tag{2}$$

which relax the constraints to:

$$y_i(\langle \mathbf{w}, \mathbf{x}_i \rangle + b) \geq 1 - \xi_i \, \forall \, i = 1, \ldots, m. \tag{3}$$

The *soft margin classifier* is obtained by minimizing the objective function:

$$\tau(\mathbf{w}, \xi) = \frac{1}{2}\|\mathbf{w}\|^2 + C \sum_{i=1}^{m} \xi_i$$

subject to the constraints (2) and (3). Next, when the pattern space $\mathbf{X}$ is not a subset of $\mathbb{R}^d$, the SVM constructs a maximal margin linear classifier in a high dimensional feature space, $\Phi(\mathbf{X})$, for a map $\phi : \mathbf{X} - \rightarrow \mathbf{F}$, where $\mathbf{F}$ is a Hilbert Space with inner product $\langle ., . \rangle_{\mathbf{F}}$. This is equivalent to the existence of a positive semi-definite kernel function, $k(\mathbf{x}, \mathbf{x}')$, defined on $\mathbf{X} \times \mathbf{X}$, such that:

$$\langle \mathbf{\Phi}(\mathbf{x}), \mathbf{\Phi}(\mathbf{x}') \rangle_{\mathbf{F}} = k(\mathbf{x}, \mathbf{x}').$$

The function defining the *decision hypersurface* is given by:

$$f(\vec{x}) = \left\{ \sum_{i=1}^{\ell} \alpha_i y_i k(\mathbf{x}_i, \mathbf{x}) \right\} - b. \tag{4}$$

By eliminating the *primal* variables, $\mathbf{w}$ and b, one obtains the dual optimization problem for the soft margin classifier: find optimal coefficients, $\vec{\alpha}$, which maximize the functional,

$$W(\vec{\alpha}) = \sum_{i=1}^{\ell} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{\ell} y_i y_j \alpha_i \alpha_j k(\mathbf{x}_i, \mathbf{x}_j), \tag{5}$$

subject to the constraints:

$$0 \leq \alpha_i \leq C, \qquad i = 1, \ldots, \ell, \tag{6}$$

and

$$\sum_{i=1}^{\ell} \alpha_i y_i = 0. \tag{7}$$

C is a regularization parameter, controlling a compromise between maximizing the margin and minimizing the number of training set errors.

Let $\hat{\vec{\alpha}} = (\hat{\alpha_1}, \hat{\alpha_2}, \ldots, \hat{\alpha_\ell})$, maximize the objective function given by equation 5. In general only a limited number of Lagrange multipliers, $\hat{\vec{\alpha}}$, will have non-zero values. The decision surface therefore only depends on a small number of data points with non-zero $\hat{\alpha_i}$; these data points are called *support vectors*. For our classification problem, we utilize two kernel functions: The *Gaussian Radial Basis Function (RBF) Kernel* and the (nonhomogeneous) *Polynomial Kernel*. The Gaussian RBF kernel is defined by:

$$k(\mathbf{x}, \mathbf{x}') = e^{-\gamma ||\mathbf{x} - \mathbf{x}'||^2}.$$

and the polynomial kernel is defined by:

$$k(\mathbf{x}, \mathbf{x}') = (\langle \mathbf{x}, \mathbf{x}' \rangle + 1)^d.$$

2.2 Multi-class Classification and Training

The above theory is for binary classification problems. We extend this to multi-class classification in the following manner: Our multi-class scheme uses pairwise classification, i.e. a 6-class classifier is constructed from $k = 6 \times 5/2 = 15$ two-class classifiers. An input pattern is classified as belonging to the class receiving the largest number of votes.

There are various ways to train SVMs. We use a simple and fast method, the Sequential Minimal Optimization Algorithm (or SMO). Training an SVM requires the solution of a very large quadratic programming optimization problem. SMO breaks this large problem into a series of smallest possible quadratic programming problems i.e. optimizing a minimal subset of just two points at each iteration, which can be solved analytically. We refer to [14] for details.

3 The Texture Feature Set

The texture feature set is constructed using the *Gabor* wavelet transforms of the image calculated for a number of scales and orientations. We now describe this in greater detail: A two-dimensional Gabor elementary function, $g(x, y)$, is given by:

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-1/2\left(x^2/\sigma_x^2 + y^2/\sigma_y^2\right) + 2\pi i\omega x\right]$$

Its Fourier transform $G(u, v)$ is given by:

$$G(u, v) = \exp\left[-1/2\left((u-\omega)^2/\sigma_u^2 + v^2/\sigma_v^2\right)\right]$$

where $\sigma_u = 1/2\pi\sigma_x$ and $\sigma_v = 1/2\pi\sigma_y$. Gabor functions form a complete but nonorthogonal basis of $L^2(\mathbb{R})$, the space of square-integrable functions. *Gabor wavelets*, a class of self-similar functions, are obtained by appropriate dilations and rotations of the *mother wavelet* $g(x, y)$ through the generating function:

$$g_{mn}(x, y) = a^{-m} g(x', y'), \;\; a > 1, \;\; m, n, \text{ integer}$$

$$x' = a^{-m}(x\cos\theta + y\sin\theta) \text{ and } \; y' = a^{-m}(-x\sin\theta + y\cos\theta)$$

where $\theta = n\pi/K$ and K is the total number of orientations and a^{-m} the scale factor. Since Gabor wavelets are non-orthogonal, the redundancy is reduced by ensuring that the half-peak magnitude support of the filter responses

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

IEEE COMPUTER SOCIETY

in the frequency distribution touch each other (in a non-overlapping fashion). This condition gives an explicit formula for the filter parameters in terms of the highest and lowest center frequencies, and the total number of scales and orientations (see [10]), in the manner described below.

Let U_f and L_f denote the highest and lowest center frequencies. For scale parameter σ, the lowest center frequency, L_f is defined to be $\frac{1}{4\sigma}$ cycles per pixel. This value is chosen so that the Gaussian smoothing window size covers at least one cycle of the lowest spatial frequency. In our application the highest center frequency, U_f is set to 0.045 cycles per pixel. Let K be the number of orientations and S be the number of scales in the multiresolution decomposition, then the optimal filter parameters σ_u and σ_v (and thus σ_x and σ_y) are given by:

$$a = (U_f/L_f)^{-\frac{1}{S-1}}, \quad \sigma_u = \frac{(a-1)U_f}{(a+1)\sqrt{2\log 2}}$$

$$\sigma_v = \tan\left(\frac{\pi}{2k}\right)\left(U_f - 2\log\left(\frac{\sigma_u^2}{U_f}\right)\right)\left(m - \frac{m^2\sigma_u^2}{U_f^2}\right)^{-1/2},$$

where, $m = 2\log 2$. Next, given a gray scale image $I(x, y)$ its Gabor wavelet transform is then defined to be:

$$W_{mn}(x, y) = \int I(x_1, y_1) g_{mn}(x - x_1, y - y_1) dx_1 dy_1$$

The mean μ_{mn} and the standard deviation σ_{mn} of the magnitude of the transform coefficients are used to represent the image.

$$\mu_{mn} = \int\int |W_{mn}(xy)| dxdy \text{ and}$$

$$\sigma_{mn} = \sqrt{\int\int (|W_{mn}(x, y)| - \mu_{mn})^2 dxdy}$$

For a color image, one can define similar quantities in both RGB space and HSV (Hue, Saturation, and Value) space. In other words, one can define quantities: μ_{mn}^R, μ_{mn}^G, μ_{mn}^B, σ_{mn}^R, σ_{mn}^G, and σ_{mn}^B for each of the three RGB color bands and similarly in HSV (Hue, Saturation, and Value) space. We will refer to the first set as RGB Gabor Features and the second as HSV Gabor Features. In the experiments, we found four scales $S = 4$ and eight orientations $K = 8$, to be optimal, resulting in ($32 \times 2 = 64$ features for each color channel) a 192 dimensional feature space.

4. Rotation Invariance

Let S be the number of scales and K the number of orientations. Let μ_{mn} and σ_{mn} for $m = 1, \ldots, S$ and for $n = 1, \ldots, K$, be the features defined above. We compute the magnitudes of the Discrete Fourier Transforms of these features to make them rotation invariant. In other words, define rotation-invariant features, μ_{mn}^{ROT} for $m = 1, \ldots, S$ and for $n = 1, \ldots, L+1$ where, $L = [\frac{K}{2}]$ = the greatest integer less than $K/2$, as follows:

$$\mu_{mn}^{\mathrm{ROT}} = \left| \sum_{k=1}^{K} \mu_{mj} \exp(-\frac{i2\pi kn}{K}) \right|$$

Since *circular translation* or rotation does not affect the magnitude of the Fourier Transform (but only the phase), these features are rotation-invariant. Similarly one defines, $\sigma_{mn}^{\mathrm{ROT}}$. As above, one defines these features for each of the RGB (or HSV) color channels. The accuracy of these features is dependent on the number of orientations, K. The higher the K, the better the feature set. In section 6, we describe results obtained by chosing $K = 90$, (which makes $L = 46$). Then for $S = 4$, we get $4 \times 46 = 184$ values each for μ_{mn} and σ_{mn} (for each color channel), producing a $184 \times 2 \times 3 = 1104$ dimensional feature space.

5 The Color Feature Set

The color features are constructed from the *color distribution angles* and the *color edge distribution angles* of [5], which are defined as follows: Let the color image I be represented as an $M \times 3$ matrix, $I = [I_1, I_2, I_3]$, where M is the total number of pixels in the image and I_1, I_2, I_3, the three color channels. Let N be the column normalization function. In other words,

$$N(I) = ID_N$$

where the ith diagonal entry of the 3×3 diagonal matrix D_N is equal to the reciprocal of the length of the ith column of I. For the RGB color space, the color distribution angles $\phi_{\mathrm{RGB}}(I) = (\phi_{12}, \phi_{23}, \phi_{31}) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}$, are defined by:

$$\phi_{ij}(I) = cos^{-1}(\langle N(I)_i, N(I)_j \rangle) \tag{8}$$

where, $N(I)_i$ denotes the ith column of $N(I)$. By construction, these angles are rotation-invariant.

The color edge map ([5]) is defined to be the image convolved with the Laplacian of Gaussian, $\nabla^2 G * I$, in which the usual two-dimensional filter is replicated for each of the three color channels, and where, as before, $\nabla^2 G * I$ is represented as an $M \times 3$ matrix. The color edge angles are defined to be:

$$\phi_{\mathrm{RGB}}(\nabla^2 G * I) = (\phi_{12}(\nabla^2 G * I), \phi_{23}(\nabla^2 G * I), \phi_{31}(\nabla^2 G * I)). \tag{9}$$

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

IEEE COMPUTER SOCIETY

The color edge angles encode second order moment information about the color edge distribution and, by construction, are rotation-invariant. In addition to the color angles, we also use three more (rotation-invariant) numbers $\Sigma_{\mathrm{RGB}}(I) = (\Sigma_{12}(I), \Sigma_{23}(I), \Sigma_{31}(I))$ defined as follows:

$$\Sigma_{ij}(I) = \sigma(I_i . * I_j) \tag{10}$$

where $.*$ denoted term by term multiplication of vectors and σ is the standard deviation. For the HSV (Hue, Saturation and Value) color space, one can similarly define 9 invariants: $\phi_{\mathrm{HSV}}(I)$, $\phi_{\mathrm{HSV}}(\nabla^2 G * I)$ and $\Sigma_{\mathrm{HSV}}(I)$.

We should add that a color histogram (used for example in [1]) does not work very well on our images. In addition to much increased computation time, the histogram features never resulted in a classification accuracy of over 70%.

6 Implementation

Recall that the *Leave-One-Out (LOO) Cross-Validation* is obtained by training a classifier k times (where k is the sample size of the feature vectors), each time leaving out one of the feature vectors from training and using only that omitted vector for classification. The various LOO cross-validation error rates for different sets of features are described below. The feature sets were first calculated for 60 vertically oriented images (see first image in Figures 1– 6) of each of six bivalve species described in Section 1.2

Although a number of choices were available for the number of scales, S, and orientations, K, and the settings for the highest center frequency, U_f, and lowest center frequency, L_f, we found that $S = 4$, $K = 8$, $U_f = 0.045$ cycles/pixel and $L_f = 0.0065$ cycles/pixel were the best settings; we found that if we chose the center frequencies of our Gabor wavelets to be low, we obtained better features for classifying all six species. The table 1 describes various choices of L_f, U_f, S, and K, and the resulting LOO percentage errors. (The SVM used the Gaussian RBF Kernel with $\gamma = 5$ and the regularization parameter $C = 100$.) The parameter, γ, we use is fairly large; in other words the resulting model is sensitive to small changes in inputs (*i.e.* a very non-linear mapping).

The complete Gabor feature set for vertically oriented images has length 192 (see section 3). With the above optimal settings for the center frequencies we experimented with various parameters for both the Gaussian RBF kernel, $k_{\mathrm{RBF}}(\mathbf{x}, \mathbf{x}') = e^{-\gamma||\mathbf{x}-\mathbf{x}'||^2}$, and the Polynomial kernel, $k_{POLY}(\mathbf{x}, \mathbf{x}') = (\langle \mathbf{x}, \mathbf{x}' \rangle + 1)^d$. In general, the RBF kernel performed a little better, but not by much. Table 2 gives the leave-one-out percentage error rates for various settings of γ, C, and d.

The 9 color angles of the image by themselves did not prove to be good features (see last two rows of Table 3),

Table 1: The L-O-Out % error rates for different values of the lowest center frequency, L_f (cycles/pixel); highest center frequency, U_f (cycles/pixel); number of scales, S; and number of orientations, K. Gaussian RBF Kernel with $\gamma = 5$ and $C = 100$.

All six species				
L_f	U_f	S	K	LOO % Error
0.0065	0.045	4	8	16.71
0.006	0.045	8	12	21.79
0.0065	0.045	5	8	24.36
0.005	0.1	4	6	33.33
0.005	0.3	4	6	52.14

Table 2: A Comparison of Gaussian RBF and Polynomial Kernels for different values of γ, C, and d. Feature set is RGB Gabor (no color angles)

Gaussian RBF Kernel			**Polynomial Kernel**	
$\vec{\gamma}$	**C**	**LOO % Error**	**d**	**LOO % Error**
2.5	100.0	15.32	3	16.99
3.0	100.0	14.2061	4.5	15.60
3.25	100.0	14.4847	5	15.04
4.0	100.0	15.0418	6	15.04
5.0	100.0	15.5989	7	15.88
7.0	100.0	15.8774	10	16.43

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 **IEEE**

IEEE COMPUTER SOCIETY

Table 3: The L-O-O % errors for different values of γ and C for the feature set of RGB Gabor Features and RGB Color Angles. Kernel = Gaussian RBF

Gabor and Color Angles		
γ	C	LOO % Error
3.0	100.0	15.04
4.0	100.0	13.93
5.0	100.0	13.09
6.0	100.0	13.93
Color Angles Only		
3.0	100.0	28.41
6.0	100.0	28.41

Table 4: The leave-one-out percentage errors for different values of γ and C for the feature set comprising of HSV Gabor Features and HSV Color Angles. The lowest error rate of 7.52% is for settings $\gamma = 2$ and $C = 70$. Kernel = Gaussian RBF

HSV Gabor and HSV Color Angles		
γ	C	LOO % Error
1	70.0	10.03
1	80.0	9.47
2	60	7.79
2	70	7.52
2	80	8.36
3	60	8.91
3	70	9.47
3	80	9.749

however when added to the Gabor features set (thereby creating a feature space of dimension 201), they improved the classification accuracy, as the Table 3 indicates. The table is for the RBF kernel (and Gabor settings of $S = 4$, $K = 8$, $U_f = 0.045$ cycles/pixel, and $L_f = 0.0065$ cycles/pixel).

We next considered both Gabor features and Color Angles in HSV space and this proved to be our best feature set for vertically oriented images, giving the lowest error rate of 7.52%. Table 5 describes the confusion matrix for these features. Finally, we calculated the rotation-invariant features (of dimension 1104, see section 4) in HSV space for images in random orientations. For the same choices, $\gamma = 2$ and $C = 70$, we obtained a L-O-O accuracy of 82.26%. Table 6 describes that confusion matrix.

7 Summary and Conclusion

We have shown how support vector machines, polarization microscopy, multiscale texture and color invariants can be

Table 5: The confusion matrix for classification of 6 species (60 images each), with each image in standard (vertical) orientation. Gaussian RBF Kernel, $\gamma = 2.0$, $C = 70$. $\mathbf{1}$ = *Argopecten*, $\mathbf{2}$ = *Crassostrea*, $\mathbf{3}$ = *Mercenaria*, $\mathbf{4}$ = *Mya*, $\mathbf{5}$ = *Placopecten*, and $\mathbf{6}$ = *Spisula*. The columns are the actual values and the rows the predicted values.

Images in Vertical Orientation						
	1	**2**	**3**	**4**	**5**	**6**
1	54	0	0	3	3	0
2	2	58	0	0	0	0
3	0	0	57	0	0	1
4	6	0	0	53	1	0
5	4	0	0	1	55	0
6	0	0	4	0	0	56

Table 6: The confusion matrix for classification of 6 species (78 images each), with images in random orientations. Gaussian RBF Kernel, $\gamma = 2.0$, $C = 70$. $\mathbf{1}$ = *Argopecten*, $\mathbf{2}$ = *Crassostrea*, $\mathbf{3}$ = *Mercenaria*, $\mathbf{4}$ = *Mya*, $\mathbf{5}$ = *Placopecten*, and $\mathbf{6}$ = *Spisula*. The columns are the actual values and the rows the predicted values. Classification accuracy = 82.26%

Images in Random Orientation						
	1	**2**	**3**	**4**	**5**	**6**
1	56	0	0	10	4	0
2	0	76	0	1	0	0
3	0	0	70	0	2	8
4	14	1	0	59	13	0
5	5	0	0	7	55	1
6	0	0	7	0	1	89

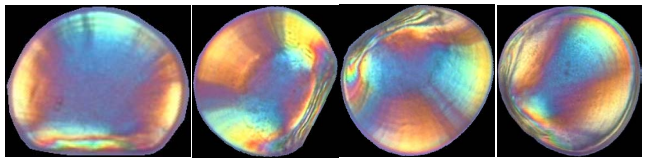

Figure 1: Argopecten (Bay Scallop) days 3, 5, 7 and 12. The first image in this set and below is in the standard (vertical) orientation; the remaining three are in random orientations. Note: all images have been resized to the same size in order that age (size) not influence the classification.

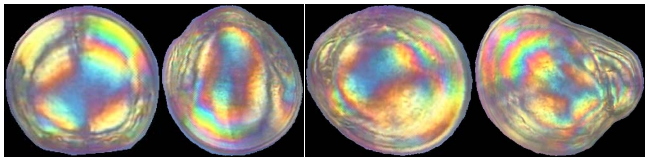

Figure 2: Crassostrea (American Oyster) days 6, 9, 12 and 16

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

IEEE COMPUTER SOCIETY

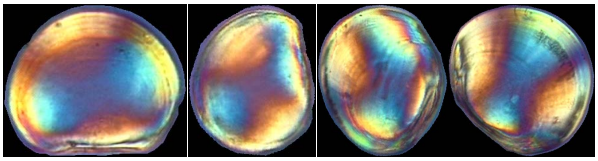

Figure 3: Mercenaria (Hard-shelled clam) days 6, 9, 10 and 10

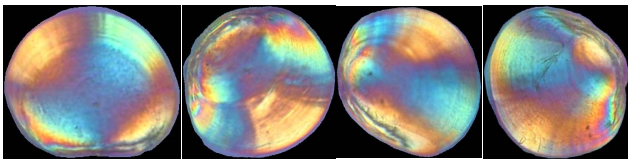

Figure 4: Mya (Sand gaper) days 2, 6, 6 and 7

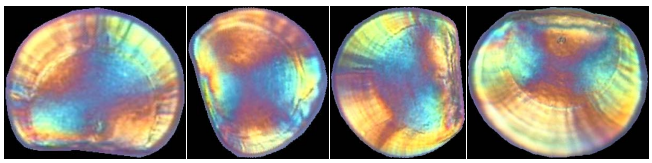

Figure 5: Placopecten (Sea scallop) days 4, 6, 11 and 21

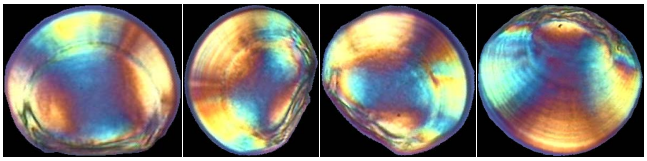

Figure 6: Spisula (Atlantic surf Clam) day 4, 10, 13 and 16

used to efficiently identify larvae of bivalve mollusks. We show further that our pattern recognition system can be optimized to achieve an accuracy of over 90% in the identification of vertically oriented images of six species of bivalve larvae, and an accuracy of over 80% in the identification of randomly oriented images. Rapid and accurate identification of larvae from the plankton will provide the basis for research into how larvae are transported by ocean currents as well as research into the population dynamics of commercially important shellfish stocks.

References

[1] Chapelle, O., P. Haffner and V. Vapnik, 1999 "SVMs for histogram based image classification," *IEEE Trans. Neural Networks*

[2] Christianini, N. and J. Shawe-Taylor, 2000, *An Introduction to Support Vector Machines and other kernel-based learning methods*, Cambridge, U.K., Cambridge University Press.

[3] Daugman, J.G., 2001 "Statistical Richness of Visual Information: Update on Recognizing Persons by Iris Patterns," *Int. Journal of Computer Vision* **45**(1) : 25–38.

[4] Feichtinger, H.G. and T. Strohmer (eds.) 1998, *Gabor Analysis and Algorithms*, Birkhauser, Boston, pp 1-32.

[5] Finlayson, G.D., S.S. Chatterjee, and B.V. Funt, 1996, "Color Angular Indexing", *Proc. 4th Europ. Conf. Computer Vision*, vol 2, Cambridge, pp.16–27.

[6] Greenspan, H., S. Belongie, R. Goodman, and P. Perona, 1994 "Rotation invariant texture recognition using steerable pyramids" *Proc. IEEE Intern. Conf. Image Process.* Jerusalem, October 1994.

[7] Haley, G.M. and B.S. Manjunath, 1999 "Rotation invariant texture classification using a complex space-frequency model," *IEEE Trans. Image Process.* **8** : 256–269.

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 **IEEE**

IEEE COMPUTER SOCIETY

[8] Herbrich, R., 2002 *Learning Kernel Classifiers: Theory and Algorithms*, Cambridge, Massachusetts, The MIT Press.

[9] Lucchese, L, and S.K. Mitra, "Color Segmentation Based on Separate Anisotropic Diffusion of Chromatic and Achromatic Channels," to appear in *IEE Proceeding: Vision, Image, and Signal Processing*

[10] Manjunath, B.S. and W.Y. Ma, 1996 "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Machine Intell.* **18** : 837–842.

[11] Osuna, E., R. Freund and F. Girosi, 1997, "Training support vector machines: An application to face detection," *Proc. IEEE Conf. on Computer Vision and Patt. Recog.* pp 130–136.

[12] Papageorgiou, C. and T. Poggio, 2000, "A Trainable System for Object Detection," *Intern. Jour. Computer Vision* **38**(1) : 15–33.

[13] Pichler, O., A. Teuner, and B.J. Hosticka, 1996 "A comparison of texture feature extraction using adaptive Gabor filtering, pyramidal and tree structured wavelet transforms," *Pattern Recog.* **29** : 733–741.

[14] Platt, J.C., 1998 "Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines," *Microsoft Research Technical Report*, MSR-TR-98-14.

[15] Pontil, M. and A. Verri, 1998 "Object recognition with support vector machines," *IEEE Trans. Pattern Anal. Machine Intell.* **20** : 637–646.

[16] Porat, M. and Y.Y. Zeevi, 1989 "Localized texture processing in vision: analysis and synthesis in the Gaborian space," *IEEE Trans Biomed. Eng.* **36** : 115–129.

[17] Porter, R. and N. Canagarajah, 1997 "Robust rotation invariant texture classification: wavelets, Gabor filter, and GMRF based schemes," *IEE Proc. Vis. Image Signal Process.* **144** : 180–188.

[18] Roobaert, D. and M.M. Van Hulle, 1999, "View-based 3d object recognition with support vector machines." In Y.-H. Hu et al, editors, *Neural Networks for Signal Processing IX*, IEEE Press.

[19] Scholkopf, B. and A.J. Smola, 2002 *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, Cambridge, Massachusetts, The MIT Press.

[20] Swain, M.A. and D.H. Ballard, 1991 "Color Indexing," *Int. Jour. of Computer Vision*, **7**(11) : 11–22.

[21] Wanderly, J.F.C. and M.H. Fisher, 2001, "Multiscale Color Invariants Based on the Human Visual System" *IEEE Trans. Image Process.* **10**(11) : 1630–1638.

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 **IEEE**

IEEE COMPUTER SOCIETY