

Joint Ship Classification and Learning

Katia Estabridis

Research and Intelligence Department

Naval Air Weapons Center

China Lake, CA 93555

email:katia.estabridis@navy.mil

Abstract—This paper proposes a ship recognition system that jointly classifies and learns from unlabeled data within a sparse representation framework. Compact dictionaries based on local descriptors serve as the basis for the classification system via l_1 minimization techniques. Previous research has demonstrated the advantages of exploiting sparsity within the recognition context. Creating a dictionary based on invariant descriptors provides robustness to changes in illumination and to affine transformations. Traditional approaches assume that the training data will span future test samples implying that the training set includes a complete object representation. Such data sets are difficult to obtain and in the end the system's performance is highly dependent on the training data sets. This framework implements a flexible learning approach where dictionaries can be augmented or updated with relevant data from unlabeled test samples.

KEY WORDS

ship recognition, joint classification and learning, l_1 minimization.

I. INTRODUCTION

Object recognition is an area of active research and of interest in a variety of applications. The research presented in this paper focuses on ship recognition from satellite imagery, while exploring scenarios where training or labeled data is limited. Traditional automatic target recognition systems assumed that the training data spans all future test data. In the era of 'Big Data' one may assume that large training data sets are available, but often, most of that data will be unlabeled. The proposed method minimizes the need for large training data sets by introducing a learning behavior into the recognition system. Thereby, enabling knowledge expansion by adding new information from the test image to a selected dictionary class (when appropriate). Joint classification and learning within a sparse framework is the essence of the proposed work. Compact dictionaries based on local descriptors serve as the basis for the classification system via l_1 minimization techniques. Creating dictionaries based on invariant descriptors provides robustness to changes in illumination and to affine transformations, while providing a flexible framework for updating or augmenting a dictionary class during the learning phase.

Meaningful learning happens when new information is acquired and linked to existing knowledge. Therefore, the goal is to engage learning in an near orthogonal-sense by acquiring only descriptors from the test image not present in the current dictionary class. Reverse Orthogonal Matching Pursuit (OMP),

[22], provides a method to select candidate descriptors from the test image for dictionary augmentation. This technique is applied throughout the training and during joint classification and learning, to obtain compact representations that are more computationally efficient. Local descriptors such as SIFT and SURF provide target representations that are robust to changes in illumination, scale, rotation and to some degree to out-of-plane rotations. However, these descriptors depend on finding consistent key-points. Images with poor contrast provide few key-points, which in turn provide few descriptors. On the other hand good quality images provide a large number of key-points and therefore lots of descriptors. Real world imagery will be of non-uniform image quality. This work addresses these challenges through adaptive filtering, to enhance the imagery to obtain more reliable and consistent key-points. Although this does not ensure the same number of descriptors per class, it provides better target representations. A ranking approach is implemented to accommodate the uneven number of images per class and the uneven number of descriptors per image. This is required since some classes will have far more descriptors than others and also have more within intra-class variability than others, thus requiring more training samples to obtain better target representations. The dimension mismatch among the descriptor classes frequently biases the answer towards the class with the most descriptors (because of the ship similarities and noise), if the answer is based solely on the entire energy present per class. Applying a ranking approach, at the image level, per class solves this problem.

The proposed work expands on the work developed by Estabridis, [1], for face recognition. There it was possible to only utilize a single image per class during training. The intra-class variability associated with faces is significantly less than that of ship classes. Additionally the training and test imagery utilized for face identification consisted of cropped images thereby limiting the presence of clutter. The raw ship imagery utilized in this effort is from different satellites with varying resolution and image quality that include clutter in the form of glint and wake.

The algorithm can be summarized as a four-step process composed of: glint removal, dictionary training (only once), recognition and learning.

Glint removal is a necessary procedure because, if not removed, it will generate unwanted key-points in the form of noise by creating spurious descriptors, while increasing the computational burden. Interest points are usually selected

for feature computation after the image has been processed through a filter bank. Glint is worse than wake since it offers no information about the ship itself. Wake can provide information about the ship and its the movement. A process based on non-parametric Bayes is employed to remove the unwanted glint.

Given a training image, features are computed and stored in a sub-dictionary as columns of a matrix. New training images are added via Reverse-OMP in order to minimize redundancy in the dictionary and reduce its size which increases computational efficiency. This process is employed in the training phase as well as during joint classification and learning.

The initial dictionary for all c classes is a concatenation of all the sub-dictionaries or descriptors per class into $A = [A_1, A_2, \dots, A_c]$. The recognition problem is then formulated as a sparse representation where the linear equations, $y_i = Ax_i$ constitute an undetermined system of equations. The sparse solution, x_i , is obtained via l_1 minimization techniques. The final solution contains the matches for each input descriptor, y_i , from the test image in relation to all the classes present in the dictionary providing the identification of the test sample.

Utilizing SIFT as dictionary atoms eliminates the need to align test images with the training images providing resilience to viewing angle, scale and illumination changes. More details on the robustness of SIFT can be found in [2]. The SIFT key-point selection process was modified to obtain better coverage, and incorporated a filtering stage in order to enhance target features.

Learning involves altering the current knowledge of the dictionary thus it has a long-term impact on the system. The decision to augment the dictionary is based on a conservative ratio test measuring the sparseness of the solution.

One of the main contributions of this paper is the design of an adaptable and computationally efficient ship recognition algorithm capable of handling scenarios when very few labeled training samples are available. The algorithm learns from the unlabeled data thereby minimizing the need for large labeled training data sets. Moreover, it offers good performance in the presence of clutter.

II. RELATED WORK

Object recognition is an ongoing and active area of research. Recently, sparse representations and the theory of compressed sensing have provided innovative approaches for object recognition that exploit the sparsity of the solution. Regional descriptors like SIFT, SURF, HOG among others are proven tools utilized in computer vision ranging from object recognition, to image rectification, and registration algorithms. The proposed methodology in this paper combines both regional descriptors and sparse representation within a joint classification and learning framework. In the interest of brevity, the literature review is restricted to techniques that are similar to the proposed approach.

Wright, [3], formulated the face recognition problem as a sparse representation based on the raw images to create a fixed dictionary. This approach is successful as long as the training

and testing images are aligned, the training data spans the illumination space of future input images, and the training and test images are within the same scale. Subsequently Wagner, [4], achieved stable performance under a wide range of illuminations, misalignments and with tolerance to small amounts of pose variation and occlusion, but not including scale changes. This approach, requires several image samples per class for training. Good performance is reported on the Multi-PIE database, [5], within the tested scenarios.

Two approaches within the sparse framework using SIFT dictionaries are those of Zepeda [6] and Kang [7]. In [6], sparse coding is applied to the SIFT features to create a codebook for querying of large databases. Kang utilizes SIFT dictionaries, with a focus on security. Secure-SIFT features are first obtained from several training samples, followed by the K-SVD algorithm, [8], to condense the information into a fixed dictionary. Good results were reported on the CalTech 101 database, [9]. One must point out that algorithms tested on the CalTech 101 database do not necessarily address real world challenges as shown by Pinto, Cox, and DiCarlo, [10]. They developed an object recognition system that performed extremely well on the CalTech 101 database but it performed near chance level for a 2-category problem when scale and rotational changes were introduced.

Estabridis, [1], expanded on the SIFT + Sparse framework proposed in [6] and [7], by adding a feedback loop to introduce dictionary learning while classifying. The methods were specifically applied to face recognition and achieved very good classification rates while addressing the cases where labeled training data is limited. The techniques developed herein for face recognition are adapted to the ship recognition problem.

Other work that proposes SIFT for face analysis is presented in [11] and [12]. Krizaj, [11], proposes Grid-SIFT and points out that the SIFT descriptor is invariant to illumination changes but that the key-point detector is not. In Grid-SIFT, descriptors are calculated at pre-determined intervals, with reported superior performance when compared to regular SIFT and other methodologies. Parua in [12] showed the potential of utilizing a hierarchical feature face recognition system based on SIFT and Gabor features plus fiducial points. The system's performance is improved by computing descriptors around the interesting/fiducial points. The work presented here overcomes these challenges by introducing a filter bank prior to key-point selection.

Yang [13] proposed a supervised hierarchical sparse coding approach model based on local descriptors for image classification. The hierarchical model is constructed by max pooling over the sparse codes of local descriptors within a spatial pyramid. The algorithm requires several images for training. Good results are reported with respect to the Multi-PIE database for frontal views.

Rainey [14] investigated several approaches similar to those of face recognition with a focus on the ship recognition problem. Their best results were obtained by utilizing hierarchical multiscale local binary patterns coupled with support vector machines. They achieved a recognition rate of 90% when train-

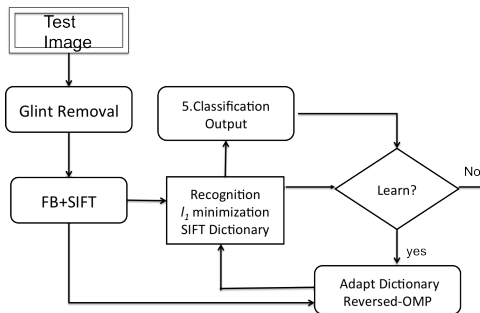


Fig. 1. System's Architecture

ing on 80% of the available pre-processed data. Furthermore the data needs to be pre-processed, (rotated, cropped, aligned and resized) to obtain such results. By comparison the data utilized in this work is mostly in its raw state except for the glint removal procedure.

Tan [15] presented a survey of promising algorithms for a face recognition system that utilized a single training sample per class. The algorithm based on Local Binary Patterns (LBP) developed in [16] is outlined as one of the best approaches. LBP is a regional operator that extracts textural information. Although algorithms based on LBP have been shown to be successful, as demonstrated in [4] and [14], the LBP approach requires both the training and testing samples to be aligned. Whereas the algorithm proposed in this paper does not require the training samples to be aligned with the test samples, since the regional SIFT descriptor is invariant to image translation, scaling, and in plane rotation.

III. RECOGNITION AND LEARNING

Formulating the recognition problem as a sparse representation offers a variety of advantages. First, a recognition problem should have a sparse solution when compared against an entire database since the object under test can only belong to one class. Second, powerful tools developed under compressed sensing by Candes, [17], can be exploited. And most importantly, dictionaries offer a flexible setting to add, new classes, new dictionary atoms to an existing class, and to replace existing dictionary atoms when needed. It is with these basic ideas that the underlying system has been developed. The goal is to design a flexible architecture that enables simultaneous classification and learning while addressing the cases when labeled data is limited. The algorithm has four basic steps and the system's architecture is depicted in Fig 1.

- (A) Glint Removal
- (B) Filter Bank, key-point selection
- (C) SIFT at key-points
- (D) Recognition
- (E-F) Learning

A. Glint Removal

Glint removal is a necessary step in this approach due to the feedback mechanism of the system's architecture. Glint is not associated with the ship in any way and it only adds noise

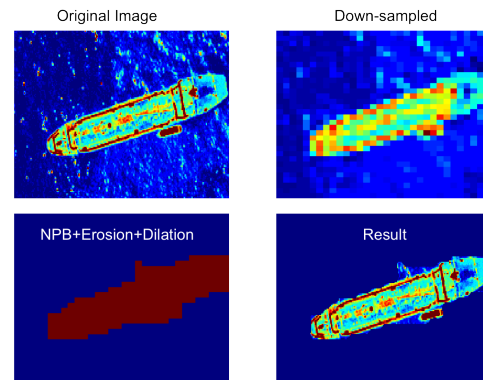


Fig. 2. Glint Removal, Upper left panel. Original image, Upper right panel. Downsampled image, Bottom left panel, Downsampled image processed by NPB + morphology methodology, Bottom right panel. Up-sampled solution after NPB + morphology

to the dictionary. If left untouched the proposed methodology while in learning mode could augment the dictionary with descriptors associated with glint. The glint removal process is based on non-parametric Bayes (NPB), by modeling the image as a Gaussian Mixture Model (GMM). The input to the GMM is the downsampled image followed by morphological post-processing to eliminate unwanted clusters. The GMM is described by Bishop in [18], Winn in [19], Yu in [20], and others, has enough level of complexity to not only model the data, but to learn its clustering structure. Fig 2. depicts a processed ship image.

B. Filter Bank, Key-point Selection

The SIFT key-point detector searches for stable interest points that are repeatable and can be found again at similar locations in other images. Interest points are sensitive to image content and to image contrast. Some images will generate no key-points due to their poor contrast or lack of texture, making the image classification impossible. Variations of the SIFT key-point approach have been proposed in order to overcome such limitations. Dense or Grid SIFT is an approach that samples a regular grid resulting in good coverage of the entire image. On the downside Dense SIFT cannot reach the same level of repeatability, of the original SIFT approach, unless the image is densely sampled including sampling at various scales, resulting in an insurmountable number descriptors. Tuytelaars, [21], proposed a dense interest point approach that is a hybrid methodology of the SIFT and Dense SIFT key-point detector approach that obtains better coverage around regions of interest in the image. Key-points on the water that surrounds the vessel are of no value to the recognition system and represent a source of noise which unnecessarily increases the computational burden, with a potential decline in performance. For this reason a filter bank prior to applying the key-point detector, is applied as a pre-processing step to enhance the structures of interest within the ship imagery resulting in better coverage of the ship. Afterward, the key-point selection threshold is decreased automatically until at

least ' N ' key-points are obtained. This procedure increases the number of descriptors obtained per image specially when contrast is poor. The SIFT descriptors are computed only after the desired ' N ' points are obtained.

C. SIFT Dictionaries

SIFT descriptors for each image class are computed and stored as columns of the dictionary. SIFT was introduced by Lowe and it can be summarized in four major steps: scale space extrema, key-point localization, orientation assignment and, descriptor representation. The SIFT descriptor representation is a vector which is invariant to scale, translation, in-plane rotations, and partially invariant to illumination and geometric distortions. More details on SIFT can be found in [2].

D. Recognition

The recognition problem is a two-step process formulated within the sparse representation framework. First the algorithm finds a sparse representation for each SIFT descriptor obtained from the test image. Second it finds the class with the maximum number of descriptor matches. The set of N descriptors obtained from the test image is represented as $Y = AX$. Each column in Y is $y_i \in R^{M \times 1}$, $i = 1, 2, \dots, N$ and each column in X is the sparse solution $x_i \in R^{K \times 1}$, $i = 1, 2, \dots, N$ for each of the descriptors obtained from the test image by solving the constrained linear system, (1). The solution to each descriptor is a sparse linear combination of the dictionary elements in $A \in R^{M \times K}$ ($M \ll K$).

$$\min_x \|x_i\|_1 \quad \text{subject to} \quad y_i = Ax_i \quad (1)$$

The solution for the l_1 minimization problem in (1) can be approximated with Orthogonal Matching Pursuit (OMP), a greedy algorithm. Tropp in [22] showed that OMP can yield equivalent results to those of the Basis Pursuit algorithm (gradient type technique) which has been shown to be robust to noise. OMP is an attractive choice for its speed and ease of implementation. Computational efficiency is important since this approach requires solving for the sparse solution N times, which can be done in parallel since each sparse solution is independent of the others. The resultant matrix $X = [x_{11}, x_{12}, \dots, x_{CN_c}]$ containing the solution for all the test input descriptors is used to count the matches obtained for each image class. A ranking approach is employed to avoid biasing the results, since some of the classes have far more images than others. The highest ranked images from each class are selected during the recognition process to identify the ship in the test image. Finally, the class with the maximum energy is then selected for identification and potential learning.

E. Unsupervised Learning from Unlabeled Data

Learning can benefit any kind of system as long as the learning is conducted in a meaningful way. Learning by definition means acquiring new information or modifying existing knowledge.

In this framework, the goal is to improve the system's performance by incorporating a learning behavior into the algorithm. The proposed approach takes advantage of unlabeled

data to expand its current knowledge. There are two questions that must be answered during the learning process. The first one is to decide *when to learn* and the second one relates to *how to learn*. The *when to learn* is currently solved as a ratio test and the *how to learn* is solved via Reversed-OMP.

F. Learning Decision Process

The tested scenarios include 4-class and 6-class problems making a ratio test an appropriate approach. The decision process applies a conservative ratio test to the ranked cumulative sparse solution after solving (1). The ratio test takes into account the top 2 classes with the most energy. Learning is enable if the ratio test if $q_r/q_{r-1} < t$.

G. Learning with Reversed-OMP

Meaningful learning happens when new information is acquired and linked to existing knowledge. Therefore the goal is to acquire only new descriptors from the test image not present in the selected dictionary class. Reverse-OMP is the proposed technique to conduct the actual learning process in order to learn as orthogonally as possible to the current set of atoms in a selected dictionary class. The descriptor dictionary is an overcomplete basis meaning that the atoms in the current dictionary are already correlated at some level, thus the term near-orthogonal. These ideas are similar to those presented in compressed sensing theory where an ideal basis is an orthonormal basis which will contain more information than an overcomplete basis of the same size. Reverse-OMP provides a way to find the least correlated descriptors from the test image to those of the selected dictionary class $A_c = [a_{c1}, a_{c2} \dots, a_{cN_c}]$. This is opposite to solving for (1) where the sparse solution was the desired one. In this case although solving for $A_c = Y_j X_j$, the desired solution is \tilde{X}_j , with $Y_j \in R^{M \times P}$ being the set of descriptors obtained from the test image, and \tilde{X}_j being the complement of the sparse solution of (2) such that $Y_j = Y_j(X_j) \cup Y_j(\tilde{X}_j)$. The updated dictionary after learning from a new input image j is $A_{cj} = [A_{c(j-1)} Y(\tilde{X}_j)]$.

$$\min_x \|x_{jd}\|_0 \quad \text{s.t.} \quad Y_j x_{jd} = a_{cd}, d = 1, 2, \dots, N_c \quad (2)$$

Learning is also evoked during the training phase in order to minimize the redundancy and size of the dictionary. Basically the same process as described above is executed but without the learning decision since the ship identity is known. Dictionary augmentation is conducted while the number of atoms of the dictionary class is under a predetermined size, $atoms(A_c) < S_c$. This is done to keep the number of descriptors in the dictionary manageable. At each learning stage a maximum of n new atoms can be added to a dictionary class. A fading memory scheme is also employed to replace dictionary atoms that are not being used during the classification process. The conjecture is that if an atom's history shows no usage, then it must be a non-descriptive atom.

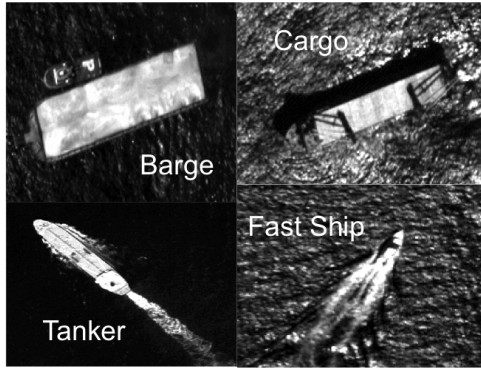


Fig. 3. Sample of Ship Classes: Barge, Cargo, Tanker, Fast-ship

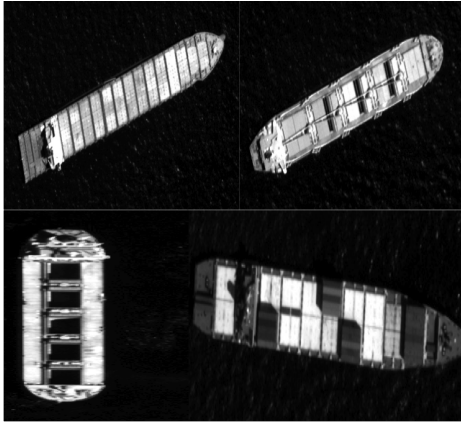


Fig. 4. Cargo Class

IV. RESULTS AND CONCLUSIONS

The database utilized for testing the algorithm consists of satellite imagery provided by SPAWAR-Pacific and it includes some of the same data used in [14]. Two problems were studied in this work: 4-class and the 6-class problems. The 4-class problem consists of barge, cargo, container and tanker. The 6-class problem includes the 4-class problem with the addition of cloud and fast-ship classes. Each class contains 200 images, of which a random set is set aside for training as listed in Tables I and II. Samples of the aforementioned classes are shown in Fig 3. The data set includes a mix of orientations, sizes and resolutions. The ship recognition problem has more intra-class variability when compared to the face recognition problem. Ideally one would like to initiate the training dictionary with a single image, but the intra-class variability associated with some of the classes, specially the cargo class makes it impossible (see Fig 4). Although the system learns from new unlabeled data it cannot learn from other sub-classes if they differ too much from the training data. Therefore, for ship recognition it is necessary to increase the number of samples during the training phase.

The dataset as shown in Fig 3 includes clutter in the form of wake and glint. The glint removal process can only remove the noise associated with the glint but it cannot separate the

ship from the wake. In the six class problem the wake of the fast-ship actually helps classify the ship itself. Some images contain multiple ships and others include two classes where a cargo ship is being refueled by a tanker. Summary of the results are shown for learning on and learning off in tables I and II respectively. Three cases are addressed that include the 4-class aligned, which refers to the same data as that of 4-class raw but with the difference being that the ships are rotated and aligned. A 2% improvement is obtained by that technique. The raw cases, outlined in tables I and II, refer to data that is in its original state (no pre-processing). The biggest discrepancy between learning 'on' and learning 'off' is obtained when processing the 6-class problem. Here the difference is 5% improvement when learning is 'on' but not as much as expected. Tables III and IV show the confusion matrices for both cases, learning 'on' and 'off', for the 6-class raw utilizing only 13% of the training data. In this case 13% of the data reflects the class with the most training images corresponding to the cargo and tanker classes. The confusion matrices on both Table III and IV show in more detail where most of the errors occur. The cargo and tanker classes get confused with each other. Part of the explanation for the confusion is that about 15 images in the cargo data set include images of cargo ships being refueled by a tanker. Additionally some of the cargo classes include a crane located close to the middle of the ship that is getting confused with the characteristic pipeline down the middle of the ship associated with tankers. The current descriptor approach cannot differentiate between the two. Fig 5 highlights some of the mentioned issues. The data was not modified in order to stay compatible with the results shown in [14]. On the other hand the confusion matrices show that in some cases with very little training data it is possible to obtain classification rates $> 90\%$. The class fast-ship improves by 10% when learning is turned on. At the same time, the current approach induces negative learning when learning is turn on as it is the case of the tanker class shown in Tables III and IV. The conjecture is that when dictionary classes are augmented with new descriptors, some of them may have similarities with the tanker class resulting in a decline in performance. It must be pointed out also that all the results shown include clutter in the form of wake. Adding descriptors associated with the wake of the ships could be causing confusion during the classification process when wakes are similar. But on other cases such as that of the fast ship class, the wake is what is allowing the system to identify the ship. This will be investigated as part of future work.

The obtained results showed that it is possible to obtain very good classification rates even with a small training data set. The proposed learning methodology improves the overall classification rates, specially for the classes where the intra-class variability is less pronounced as it is the case of the fast ship class. Increasing the training samples for the fast ship class from 8 to 21 images resulted in a classification rate of 94%. Increasing the samples of the cargo and tanker class does not improve performance, indicating that a more complex descriptor is needed to discriminate between both classes. The

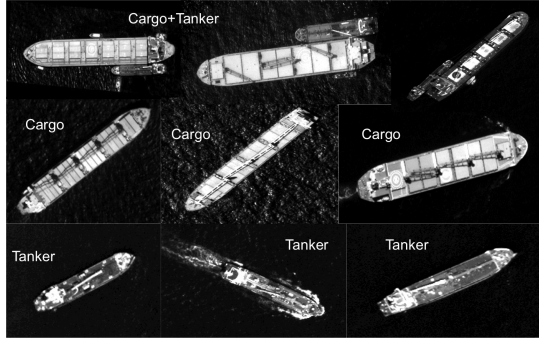


Fig. 5. Cargo-Tanker Examples

TABLE I
SUMMARY OF RESULTS, LEARNING ON

Trained Data %	4-Class Aligned	4-Class Raw	6-Class Raw
15%			82%
25%	86%	84%	84%

TABLE II
SUMMARY OF RESULTS, LEARNING OFF

Trained Data %	4-Class Aligned	4-Class Raw	6-Class Raw
15%			80%
25%	83%	82%	79%

TABLE III
6-CLASS CONFUSION MATRIX, LEARNING ON, TRAINED ON 13% OF THE DATA, C1: BARGE, C2: CARGO, C3: CLOUD, C4: CONTAINER, C5: FAST SHIP, C6: TANKER

	C1	C2	C3	C4	C5	C6	Total	#Trained
C1	149	13	0	5	3	9	0.83	21
C2	5	133	0	18	0	18	0.76	26
C3	2	5	169	0	0	3	0.94	21
C4	0	13	0	178	1	0	0.93	8
C5	16	12	0	1	143	20	0.74	8
C6	15	30	0	0	5	124	0.71	26
							0.82	

false alarm rate associated with the cloud class is very small and the probability of confusing a ship for cloud is almost non-existent, for the tested scenarios. Future work will include coupling geometric information with that of the feature vector and improving the descriptors and its coverage. As well as investigating the robustness of the algorithm when applied to partial views of ship imagery.

REFERENCES

- [1] K. Estabridis, "Face recognition and learning via adaptive dictionaries," in *IEEE Proc. of Conf. on Technologies for Homeland Security*, 2012.
- [2] D. Lowe, "Distinctive image features form scale-invariant keypoints," *International journal of Computer Vision*, 2004.
- [3] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. on PAMI*, 2009.
- [4] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi, and Y. Ma, "Towards a practical face recognition system: Robust alignment and illumination by sparse representation," *IEEE Trans. on PAMI*, 2011.

TABLE IV
6-CLASS CONFUSION MATRIX, LEARNING ON, TRAINED ON 13% OF THE DATA, C1: BARGE, C2: CARGO, C3: CLOUD, C4: CONTAINER, C5: FAST SHIP, C6: TANKER

	C1	C2	C3	C4	C5	C6	Total	#Trained
C1	143	16	0	4	1	15	0.80	21
C2	7	127	0	20	0	21	0.73	26
C3	3	3	169	0	0	4	0.94	21
C4	0	11	0	177	0	4	0.92	8
C5	17	19	1	1	123	33	0.64	8
C6	10	29	0	0	1	135	0.78	26
							0.80	

- [5] R. Gross, I. Matthews, J. F. Cohn, T. Kanade, and S. Baker, "Multi-pie," in *IEEE Proc. of Conf. on Automatic Face and Gesture Recognition*, 2012.
- [6] J. Zepeda, E. Kijak, and C. Guillemot, "Sift-based local image description using sparse representations," *IEEE Trans. on MMSP*, 2009.
- [7] L. Kang, C. Hsu, H. Chen, and C. Lu, "Secure sift-based sparse representation for image copy detection and recognition," in *IEEE Int. Conf. on Multimedia and Expo*, 2010.
- [8] M. Aharon, M. Elad, and A. Bruckstein, "The k-svd: an algorithm for designing of overcomplete dictionaries for sparse representation," *IEEE Trans. on Signal Processing*, 2006.
- [9] F.-F. L., F. R., and P. P., "Learning generative visual models from few training examples: an incremental bayesian approach tested on 101 object categories," in *IEEE Proc. of CVPR*, 2004.
- [10] N. Pinto, D. D. Cox, and J. J. DiCarlo, "Why is real-world visual object recognition hard?" *PLoS Comput Biol*, vol. 4, 01 2008.
- [11] J. Krizaj, V. Struc, and N. Pavesic, "Adaptation of sift features for face recognition under varying illumination," in *IEEE Intl. Conf. on Information and Communication Tech., Elec. and Microelectronics*, 2011.
- [12] S. Parua, A. Das, D. Mazumdar, and S. Mitra, "Determination of feature hierarchy from gabor and sift features for face recognition," in *Proc. Conf. on Emerging Applications of Information Technology*, 2011.
- [13] J. Yang, K. Yu, and T. Huang, "Supervised translation-invariant sparse coding," in *IEEE Proc. of CVPR*, 2010.
- [14] K. Rainey, S. Parameswaran, J. Harguess, and J. Statny, "Vessel classification in overhead satellite imagery using learned dictionaries," in *Proc. of Defense, Sense and Security Conf. SPIE*, 2012.
- [15] X. Tan, S. Chen, Z. hua Zhou, and F. Zhang, "Face recognition from a single image per person : A survey," *Pattern Recognition*, 2006.
- [16] T. Ahonen, A. Hadid, and M. Pietikinen, "Face recognition with local binary patterns," in *Computer Vision, ECCV 2004 Proceedings, Lecture Notes in Computer Science*, 2004.
- [17] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. on Information Theory*, 2005.
- [18] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [19] J. M. Winn, "Variational message passing and its applications," Ph.D. dissertation, University of Cambridge, 2004.
- [20] S. Yu, "Advance probabilistic models for clustering and projection," Ph.D. dissertation, University of Munich, 2006.
- [21] T. Tuytelaars, "Dense interest points," in *IEEE Proc. of CVPR*, 2010.
- [22] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. on Information Theory*, 2007.