CLASS-BASED IDENTIFICATION OF UNDERWATER TARGETS USING HIDDEN MARKOV MODELS

Nilanjan Dasgupta, Paul Runkle, and Lawrence Carin

Department of Electrical and Computer Engineering
Duke University
Durham, NC, 27708

ABSTRACT

It has been demonstrated that hidden Markov models (HMMs) provide an effective architecture for classification of distinct targets from multiple target-sensor orientations. In this paper, we present a methodology for designing *class-based* HMMs that are well suited to the identification of targets with common physical attributes. This approach provides a means to form associations between existing target classes and data from targets never observed in training. After performing a wavefront-resonance matching-pursuits feature extraction, we present an information theoretic tree-based state-parsing algorithm to define the HMM state structure for each target class. In training, class association is determined by minimizing the statistical divergence between the target under consideration and each existing class, with a new class defined when the target is poorly matched to each existing class. The class-based HMMs are trained with data from the members of its corresponding class, and tested on previously unobserved data. Results are presented for simulated acoustic scattering data.

1. INTRODUCTION

Hidden Markov models provide a natural architecture for target detection and classification using multi-aspect sequential data, when the target identity and orientation are unknown [1]. In our previous work, we have defined the HMM states to correspond to angular sectors of a target over which the scattering physics is relatively invariant, with the state transition probabilities derived from the relative target-sensor motion. When sensing a sequence of scattered waveforms from multiple target-sensor orientations, the target is effectively sampled from a sequence of states, which we model as a Markov process. Since the orientation of the target is generally unknown, the sequence of sampled states is also unknown (hidden), and is statistically deduced (via an HMM) from the sequence of observed features.

While it has been demonstrated that such an approach is successful for discriminating distinct targets in a variety of sensing environments [1,2], here we address the issue of designing HMMs to identify target *classes*, where members of a given class may share common physical attributes such as size or physical composition. Such a detection strategy may be incorporated into a hierarchal identification framework, where a putative target is first detected, then assigned to a class, and finally (with sufficient information) identified as a particular target within that class.

The class-based model design is summarized in Fig. 1. Data from an unknown target is submitted to a matching-pursuits [3] feature extractor, with the features subsequently vector quantized (VQ) [4] using a nearest neighbor mapping in the feature space. A class is statistically defined according to its VQ codebook, the angular sectors that define its state boundaries, and the probability mass function of the VQ codes as a function of state. If data from a target in the training set is not well matched to any existing class, its features are used to design a new class. The class state partitions are determined by successively maximizing the statistical divergence of the feature distributions between adjacent regions of the target using a tree-based algorithm as described in Section 2. In Section 3, we discuss the details of the class-association decision followed by a description of the HMM architecture and training process in Section 4. The classification results on simulated acoustic scattering data are presented in Section 5.

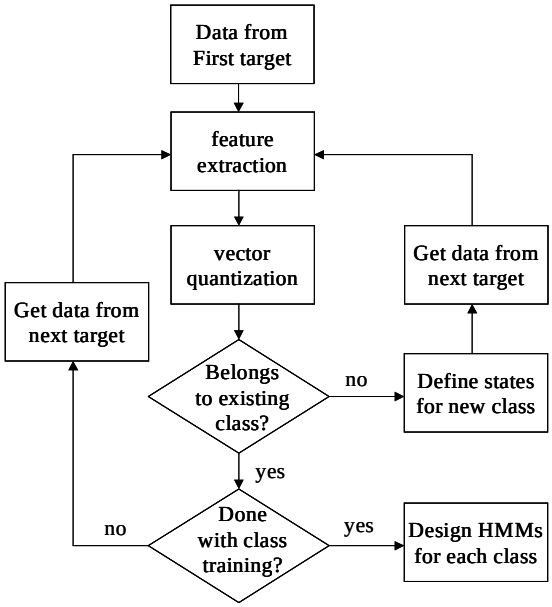

Fig. 1. Flowchart of class association prior to HMM design.

2. STATE-PARSING

While it is recognized that man-made targets often exhibit scattering that is a strong function of aspect, it is not necessarily clear how to *a priori* partition the target into contiguous states, each representing distinct scattering phenomenology. Here we describe a feature-based state parser used to define the HMM structure for a particular target class. We implicitly assume that the scattering phenomenology is sufficiently represented by the features themselves.

When a target, *T*, in the training set is not sufficiently matched (elucidated in Section 3) with any existing target class $\mathbf{C}_l$, $l = 1\ldots L$, the target features $\mathbf{Y} = \{\mathbf{y}_\theta\}$, are used to define a new class $\mathbf{C}_{L+1}$, with θ corresponding to aspect. By default, data from the first target in the training set is used to define the first class. After pre-whitening the features, a *K*-means algorithm employing a Euclidean distortion metric is applied to the data, where the number of discrete codes, *K*, is adapted to achieve a pre-defined minimum distortion. Therefore the VQ operation

$$VQ: \mathbf{Y} \rightarrow \mathbf{v},\ \mathbf{v} = \{v_\theta : v_\theta \in 1,..., K\} \tag{1}$$

is simply a mapping from feature space onto the integers 1,…,*K*. While VQ represents a discrete nonparametric representation of the feature distribution, other statistical models (such as continuous Gaussian mixtures) may be employed. We desire that the target be partitioned into states exhibiting distinct scattering behavior, manifested by the interstate dissimilarity between elements of **v**. A natural measure of the dissimilarity between two probability mass functions $p(x)$ and $q(x)$ is the information-theoretic Kullback-Liebler distance, or divergence, defined by

$$D(p \,\|\, q) = \sum_{x=X} p(x) \log \frac{p(x)}{q(x)} \tag{2}$$

In this context, given quantized data **v,** we initially seek to partition the target (Fig. 2a) into two contiguous sectors S_L and S_H along the aspect $\theta_{L|H}$ that maximizes the state-conditioned divergence of the codes

$$\theta_{L|H} = \max_\theta D(p(v \mid S_L) \,\|\, p(v \mid S_H)) \tag{3}$$

The sectors S_L and S_H may each be further partitioned given sufficient divergence between the quantized feature distributions in the subsectors. This iterative tree-based process is illustrated in Fig. 2b, where the sector S_H is partitioned into subsectors S_{HL} and S_{HH}, where $S_H = S_{HL} \cup S_{HH}$. In this example, there is no partition within S_L to achieve sufficient divergence to warrant further subdivision. When the hierarchical partitioning process is completed, the contiguous sectors may be arranged in order from bottom to top of the final nodes in the tree. For this example, the states of the target are represented by the adjacent partitions S_L, S_{HLL}, S_{HLH}, and S_{HH}, with each exhibiting sufficiently distinct feature distributions. Such an approach is similar to that employed in CART [5]. After completing the tree-based state-parsing, for convenience we re-label the states as a function of increasing aspect, $S_1,\ldots,S_M$, with the states spanning angular intervals of $\psi_1,\ldots,\psi_M$, respectively.

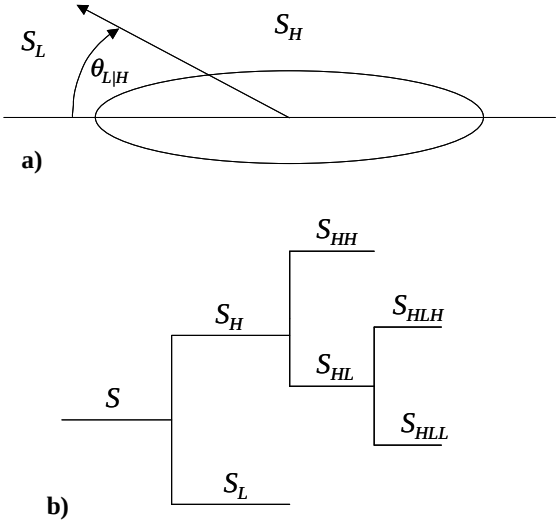

Fig. 2. Tree-based partitioning of target states. a) The target is initially divided into two angular sectors that have distinct feature distributions as defined by maximizing the divergence. b) Each sector is subsequently subdivided if there is a partition within the sector to achieve sufficient divergence, ideally corresponding to distinct scattering physics within the partitioned sector. The final set of sectors are the contiguous target states.

3. CLASS ASSOCIATION

Given that a class is characterized by its state partitions and the distribution of its codes within each state, we now establish two criteria sufficient for class membership. Since each class is representative of targets with distinct scattering behavior, a unique codebook is designed for each class, to obtain codes representative of the class within a specified minimum mean distortion. Features from a target under test for class membership are quantized using the codebook $\mathbf{Q}_l$ associated with class $\boldsymbol{C}_l$, with the codebook containing K_l codes. The number of class-dependent codes generally varies with the complexity of the features over the sampled target space, which in turn is a function of the geometric complexity of the target, the target's physical composition, and sensor bandwidth.

The first criterion for class membership considers the global fit between the target features **Y** and the class codebook $\mathbf{Q}_l$. If the average quantization distortion between the continuous features and the nearest-neighbor mapping is excessive, it is evident that the codebook does not sufficiently represent the target, which is therefore declared a nonmember of class $\boldsymbol{C}_l$. If the target data passes the global distortion test, the second criterion compares the statistics of **Y** and $\boldsymbol{C}_l$ at the state level. The data under test is partitioned according to the M_l angular sectors defined by the states $\mathbf{S}_l = \{S_{l,1},..., S_{l,M_l}\}$ associated with class $\mathbf{C}_l$. The total state-by-state divergence between the target and class is a natural measure of the local statistical mismatch, which is given by

$$d(\mathbf{Y}, \mathbf{C}_l) = \sum_{m=1}^{M_l} D(p(v_{\mathbf{Y}} \mid S_{l,m}) \| p(v_{\mathbf{C}_l} \mid S_{l,m})) \qquad (4)$$

where $p(v_{\mathbf{Y}} \mid S_{l,m})$ is the probability mass function of the quantized features $v_{\mathbf{Y}}$ from the target in the m^{th} state of class $\boldsymbol{C}_l$, and $p(v_{\mathbf{C}_l} \mid S_{l,m})$ is the probability mass function of the class $\boldsymbol{C}_l$ in its m^{th} state. When (4) is sufficiently small, the data is declared to be in $\boldsymbol{C}_l$. In the event that the features, **Y**, pass the global distortion test for more than one class, then the class that minimizes (4) is declared to best represent the target under test.

Examples of two targets that may pass the global distortion test, but fail the state-divergence test in (4) are a cylindrical shell and a spherical shell. Given a class designed on the distribution of the features from the cylindrical shell, the scattering from the sphere may be very similar to scattering from the cylinder at certain orientations. Therefore, there is likely to be a code in the codebook for the cylindrical class that is well matched to the features observed from the sphere (which are identical at all orientations). Thus the *distortion* for representing the sphere features may be low, but the state-dependent *distribution* of the spherical features may be quite disparate from the set of all cylindrical features via the cylindrical class codebook.

4. HMM DESIGN

The class-based HMMs are trained on the target data associated with their respective class. Prior to training the HMMs, a master codebook, **Q**, is constructed by merging the training features from every member of every class and re-applying the *K*-means algorithm. This ensures that the master codebook provides complete yet efficient sampling of the features from all trained classes. The existence of the master codebook requires that any data submitted for classification is vector quantized only once, rather than necessitating a separate VQ for each of the classes. The master codebook also serves to regulate the interclass likelihoods assigned to a particular observation sequence under test.

For each class, a so-called geometric HMM is designed, which we briefly describe here, with the full details of the architecture provided in [1]. The HMM for the class $\boldsymbol{C}_l$ is $\boldsymbol{\Gamma}_l = \{\mathbf{A}_l, \mathbf{B}_l, \boldsymbol{\pi}_l, \mathbf{Q}_l, \mathbf{S}_l\}$, where $\mathbf{A}_l$ is the M_l x M_l matrix of state transition probabilities, $\mathbf{B}_l$ is a K_l x M_l matrix formed by the state-conditioned probability mass function $p(v_{\mathbf{C}_l} \mid S_{l,m})$ in each of the columns, and $\boldsymbol{\pi}$ is an M_l vector composed of the probability of sampling each of the states on the initial observation of the multi-aspect sequence. Typically, such an HMM is designed for each distinct target of interest. The novel approach presented here is that the HMMs are designed by incorporating the feature statistics across all class members (with multiple targets in the same class). Such an approach is valid since the state parsing is identical across class members. Note that the HMM explicitly incorporates sequential information, while the class-association described in Sections 2 and 3 only utilizes the distribution of features from *single* observations.

A unique feature of the geometric HMM is that **A** and $\boldsymbol{\pi}$ may be directly estimated from the knowledge of the relative target sensor motion, with the structure of **A** generally constrained to disallow transitions other than to adjacent states. Such a constraint follows from our concept of state as a contiguous angular sector of the target, and leads to a tri-diagonal structure in **A**.

5. CLASSIFICATION RESULTS

We present results for four target classes using simulated acoustic scattering data generated using a finite-element axis symmetric model for the targets. The targets under consideration were submerged steel ellipsoidal shells with fixed density of 7800 kg/m^3, and Poisson ratio of 0.3. The major axis length was 20 cm for classes $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$, and 60 cm for classes $\boldsymbol{C}_3$ and $\boldsymbol{C}_4$, with all classes having minor axis length of 11 cm. In this paper, we consider a class to constitute a set of targets that have *similar*, but not identical, physical attributes. We therefore varied the within-class Young's modulus over a uniform distribution of width 0.5 x 10^{11} N/m^2 (non-overlapping between classes), with the class means specified in Table 1. Furthermore, the shell thickness for all classes was sampled from a uniform distribution with mean 1 cm, and width 0.1 cm. As defined, these classes exhibited sufficient within-class variability, while still providing a challenging interclass identification problem as illustrated in Fig. 3. Free-field scattered responses were generated for each target with angular resolution of 1 degree over 90 degrees of the target (since the targets exhibited mirror symmetry across both major and minor axes).

	Major (cm)	Minor (cm)	Y. Mod N/m^2
C_1	20	11	2.0×10^{11}
C_2	20	11	2.5×10^{11}
C_3	60	11	2.0×10^{11}
C_4	60	11	2.5×10^{11}

Table 1. Description of ellipsoidal target classes. Each class was characterized by a uniform distribution of thickness and Young's modulus as described in the text, while the external dimensions remained fixed.

A matching-pursuits algorithm [3] is employed for feature extraction, with a dictionary D_γ indexed by the parameters $\gamma_n = \{\alpha_n, \omega_n, \tau_n, \phi_n\}$, where the dictionary elements are defined parametrically as

$$e_{\gamma_n} = \beta_{\gamma_n} \cos[\omega_n (t - \tau_n) + \phi_n] e^{-\alpha_n (t - t_n)} U(t - \tau_n) \qquad (6)$$

where $U(t)$ is the Heaviside step function ($U(t)$=0 for t<0 and $U(t)$=1 for t>0), and β_{γ_n} is a normalization constant. Note that this dictionary is capable of modeling both wavefronts (small temporal support, characterized by large damping α_n) and resonances (large temporal support, characterized by small damping α_n). The timing τ_n between consecutive extracted dictionary elements yields the aforementioned time delays. Given a scattered waveform $f(t)$, matching pursuits iteratively seeks to find the dictionary element, e_γ, that maximizes the projection energy $|\langle R_n, e_\gamma \rangle|^2$ with the residual R_n. Therefore, after N iterations the partial reconstruction of f is given by

$$R_0 = \sum_{n=1}^{N} \left\langle R_{n-1}, e_{\gamma_n} \right\rangle e_{\gamma_n} + R_N, \quad R_0 \equiv f(t) \qquad (7)$$

For this study, we performed three matching-pursuits iterations, with the feature vector for the measured waveform constituted by $\mathbf{y} = \{\alpha_1, \omega_1, \alpha_2, \omega_2, \tau_2\text{-}\tau_1, \alpha_3, \omega_3, \tau_3\text{-}\tau_2\}$. Seven targets were generated from each class, with two from each class entered into the training set, and the remaining five targets from each class reserved for testing. In the statistical class-training phase described in Sections 2 and 3, no *a priori* knowledge of the physical class of the training data is assumed, leaving the algorithm to design classes based on the characteristics of the features alone. For the simulated data under consideration, our statistical class trainer correctly associated every target from each of the predefined physical classes. This not only is indicative of the success of the class association algorithm, but also of the ability of the features to well represent the scattering physics. The targets in classes $\boldsymbol{C}_1$, $\boldsymbol{C}_2$, $\boldsymbol{C}_3$, and $\boldsymbol{C}_4$ were parsed into 3, 4, 4, and 5 states respectively (over 90-degrees), indicating the relative scattering complexity across classes.

Each class-based HMM, $\boldsymbol{\Gamma}_l$, was trained using the data from the two targets associated with its class. The parameters for the model were specified by the geometric HMM [1] and refined using the Baum-Welch algorithm. Data under test was sampled at 5-degree intervals, with 10 multi-aspect observations per sequence, with each observation corrupted by white Gaussian noise at an SNR of 15 dB. Classification results are presented in Table 2, where it is observed that class confusion is predominantly across targets of the same physical dimensions.

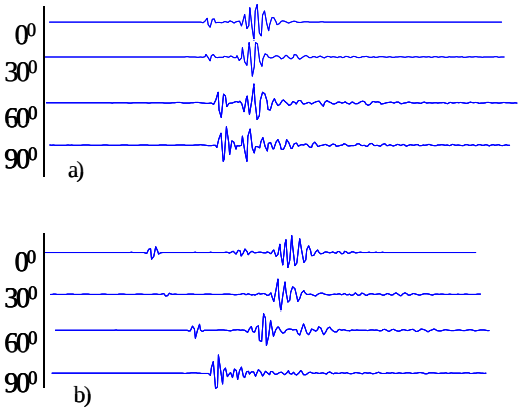

Fig. 3. – Sample waveforms of various aspects from classes a) $\boldsymbol{C}_1$ and b) $\boldsymbol{C}_3$. In the presence of moderate environmental noise, cross-class confusions may be likely given a single observation.

	C_1	C_2	C_3	C_4
C_1	0.91	0.06	0.03	0.0
C_2	0.02	0.97	0.01	0.0
C_3	0.02	0.03	0.91	0.04
C_4	0.0	0.0	0.18	0.82

Table 2 – Classification results of five testing targets from each class, with a total of 1800 testing sequences per class. The rows indicate the true target, while the columns indicate the outcome of the classifier.

6. CONCLUSIONS

We have presented a methodology for designing class-based hidden Markov models for the multi-aspect identification of targets sharing common physical attributes. The class associations for the training data may be known *a priori*, or as presented here, may be adaptively determined from the statistical characteristics of the features. The states for the HMMs were determined using a tree-based algorithm designed to maximize the interstate feature divergence within a given class. While we demonstrated the effectiveness of this approach to the classification of simulated acoustic scattering data, we believe it holds potential for a variety of multi-aspect sensing modalities.

REFERENCES

[1] P. Runkle *et al*., "Multi-aspect identification of submerged elastic targets via wave-based matching pursuits and hidden Markov models," *J. Acoustical. Soc. Am.*, vol. 106, pp. 605-616, Aug. 1999.

[2] P. Runkle, *et al*., "Multi-aspect target detection for SAR imagery using hidden Markov models," accepted for publication in *IEEE Trans. Geoscience and Remote Sensing.*

[3] S.G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Sig. Proc.*, Vol. 41, pp. 3397-3415, Dec. 1993.

[4] Y. Linde, A. Buzo, and R.M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Comm.*, Vol. 28, pp. 84-95, Jan. 1980.

[5] L. Breiman, *et al*., *Classification and Regression Trees*, Belmont, CA, Wadsworth International Group, 1984.