MULTI-ASPECT TARGET IDENTIFICATION WITH WAVE-BASED MATCHING PURSUITS AND CONTINUOUS HIDDEN MARKOV MODELS

Paul Runkle and Lawrence Carin

Department of Electrical and Computer Engineering
Duke University
Durham, NC, USA
lcarin@ee.duke.edu

ABSTRCT

A wave-based matching-pursuits algorithm is used to parse multi-aspect time-domain backscattering data into its underlying wavefront-resonance constituents, or features. Consequently, the N multi-aspect waveforms under test are mapped into N feature vectors, $\boldsymbol{y}_n$. Target identification is effected by fusing these N vectors in a maximum-likelihood sense, which we show, under reasonable assumptions, can be implemented via a hidden Markov model (HMM). Algorithm performance is assessed by considering measured acoustic scattering data from five similar submerged elastic targets.

I. INTRODUCTION

We are interested in the identification of a concealed or distant target, assuming that the putative target has been detected and therefore that its nominal location is known. However, the target itself and its orientation are unknown, and to be determined. We assume that N scattered waveforms are measured, at N different target-sensor orientations. After performing matching-pursuits feature parsing on each of the N scattered waveforms, we effect a mapping of the N scattered waveforms to N feature vectors $\boldsymbol{y}_n$, where $\boldsymbol{y}_n$ represents the feature parameters for the nth scattered waveform. Maximum-likelihood target identification is effected by choosing that target T_i for which $p(\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_N \mid T_i) \geq p(\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_N \mid T_k) \; \forall \; T_k$. Note that the problem is treated statistically even if the data under test is noise-free, since the absolute target-sensor orientation is hidden and modeled as a stochastic parameter.

Over the last several decades, there has been a significant effort to develop modeling algorithms that faithfully predict the fields scattered from general targets [1], and the insight accrued from such studies can be exploited in the aforementioned feature parsing. When a pulse of electromagnetic or acoustic energy impinges a target, there are initially diffractions from localized scattering centers on the target surface. Such features are termed wavefronts, and are characterized by localized support in time and wide support in frequency, with the extent of each dictated by the incident-pulse bandwidth. After the initial scatterings from localized scattering centers, wave energy reverberates between scattering centers [1] or may circumnavigate the target [1], with each such reverberation or circumnavigation shedding energy, which is received at the sensor. This resonant portion of the scattered signal occurs after the initial diffractions, at what is termed "late time", and such features are characterized by localized support in frequency and extended support in time, with the relative extent of each dictated by the Q of the resonance [1]. To represent this underlying physics compactly, we have developed a wavefront-resonance matching-pursuits dictionary [1], which we exploit here to generate the feature vectors $\boldsymbol{y}_n$.

In addition to exploiting the underlying wave physics in construction of the wave-based matching-pursuits dictionary, we utilize such in the evaluation of the probabilities $p(\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_N \mid T_k)$. In particular, it is well known that wave scattering from most targets is characterized by angular sectors over which the angle-dependent scattered fields are slowly varying [1]. Each such sector is here termed a "state", and the number of states characteristic of a given target is dictated by the target complexity and sensor bandwidth. Since a given state is characterized by particular underlying physics, one can define a state-dependent probability of observing a given parameter vector $\boldsymbol{y}$.

Note that, as modeled, the N scattered waveforms sample N discrete states characteristic of the target under interrogation, where some states may be sampled more than once, and others not at all, depending on the (hidden) target-sensor orientation. If we assume that the probability of transitioning from one state to the next is dictated only by the current state occupied, then the state sequence $\boldsymbol{S}$ is a Markov process [2], and the probabilities $p_S(\boldsymbol{S} \mid T_k)$ can be evaluated by a Markov model. More properly, we utilize a *hidden* Markov model (HMM) [3], since the underlying states are hidden and the only observable is the sequence of parameter vectors $\{\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_N\}$.

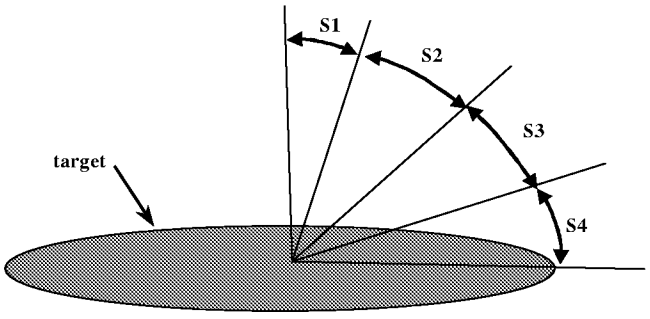

Figure 1. Schematic of a state decomposition for a generic target. Four states are depicted here.

II. Wave-Based Matching Pursuits

Matching pursuits is an algorithm developed by Mallat and Zhang [4] for decomposition of a sampled waveform $\boldsymbol{f}=[f_1, f_2, ..., f_K]^T$ in terms of a prescribed set of normalized vectors $\boldsymbol{e} \in D$, where

D is termed a "dictionary". The dictionary can take an arbitrary form, and therefore can be tailored to the features that underlie the waveform(s) under study. We perform inner products, $\langle \boldsymbol{f}|\boldsymbol{e}\rangle = \boldsymbol{f}^H\boldsymbol{e}$, between $\boldsymbol{f}$ and the sampled dictionary elements $\boldsymbol{e}$, and select that dictionary element $\boldsymbol{e}_1$ for which $|\langle \boldsymbol{f}|\boldsymbol{e}_1\rangle| \geq |\langle \boldsymbol{f}|\boldsymbol{e}\rangle|$, $\forall\ \boldsymbol{e}\in D$. A remainder $\boldsymbol{R}_1 = \boldsymbol{f} - \langle \boldsymbol{f}|\boldsymbol{e}_1\rangle \boldsymbol{e}_1$ is defined. The process is repeated on the remainder $\boldsymbol{R}_1$, and after I iterations I dictionary elements are extracted. As discussed above, we apply a wavefront-resonance dictionary [5], the details of which will be further elucidated in the talk.

III. Continuous-HMM Target Identification

A. States and state-transition probabilities

After applying matching pursuits on scattered waveforms from N different target-sensor orientations, each K-dimensional waveform, $\boldsymbol{f}_n$, is mapped into an FI-dimensional parameter vector $\boldsymbol{y}_n$, where F represents the number of features characteristic of each dictionary element (*e.g.*, time shift, oscillation frequency, etc.). The requisite number of matching-pursuits iterations, I, is dependent on the scattering complexity and on dictionary compactness; for the data considered in Sec. IV, we have found three matching-pursuits iterations sufficient.

As discussed in the Introduction, underlying scattering physics dictates that general targets have scattered fields $\boldsymbol{f}$ (and hence $\boldsymbol{y}$) that vary strongly with target-sensor orientation, the degree of variability depending on the detailed target geometry and the sensor wavelength. However, one can generally define angular sectors, representative of target-sensor orientations, over which the physics and the scattered fields $\boldsymbol{f}$ are slowly varying. We define such angular sectors as "states". The N signals $\boldsymbol{f}_n$ are sample waveforms from N states, and in practice particular states are often sampled multiple times, depending on the target orientation and sensor motion. The actual states sampled are "hidden" since the target is distant or concealed.

Assume that a given target can be represented by M contiguous states, with (in two dimensions) consecutive angular support of φ_1 , φ_2 , ..., φ_M (see Fig. 1). If we assume that the target orientation is uniformly distributed, then the probability that the first sampled waveform $\boldsymbol{y}_1$ will be in state m is

$$\pi_m = \varphi_m \ / \sum_{i=1}^{M} \varphi_i \qquad (1)$$

Moreover, if we assume that the sensor moves in one direction relative to the putative target center (the target has been detected, and is now being identified) with angular sampling $\Delta\varphi$, with $\Delta\varphi < \varphi_m$ $\forall \varphi_m$, then the probability of transitioning from state m to state l, represented by $a_{m,l}$, is given by

$$a_{m,m} = 1/2 \ \ if\ \ \Delta\varphi \leq \varphi_m \leq 2\Delta\varphi \quad ; \ a_{m,m} = \frac{\varphi_m - \Delta\varphi}{\varphi_m} \ \ if\ \ \varphi_m \geq 2\Delta\varphi$$
$$a_{m,m+1} = a_{m,m-1} = \frac{\Delta\varphi}{2\varphi_m} \qquad ; \ a_{m,l} = 0 \ \ if\ \ |m-l| > 1 \qquad (2)$$

In deriving (2), although the sensor moves in a fixed direction, we assume the target orientation is unknown and therefore that the *absolute* motion relative to the target is unknown. We see from (2) that the state transitions are modeled as a Markov process, which simplifies implementation of the subsequent identification algorithm. Rigorously speaking, the Markov model is an approximation. For example, if state m has width $\Delta\varphi \leq \varphi_m \leq 2\Delta\varphi$, and waveforms are sampled from state m on two consecutive observations, then on the next observation $a_{mm}=0$. It is possible to alter the Markov model to account for time-dependent state-transition probabilities, but, as discussed in Sec. IV, we have found the approximation in (2) to be appropriate for the cases investigated thus far. The expressions in (2) define a tri-diagonal state-transition matrix $\boldsymbol{A}$.

B. Continuous HMM

Maximum-likelihood target identification is effected by selecting that target T_i for which $p(\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_N \mid T_i) \geq p(\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_N \mid T_k)\ \forall\ T_k$. Recall that the N measured waveforms, and hence $\boldsymbol{y}_n$, sample N *hidden* states of the target under interrogation. We therefore evaluate $p(\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_N \mid T_k)$ by considering all possible M^N state sequences, with each weighted by its probability of occurrence. In particular,

$$p(\boldsymbol{y}_1,\boldsymbol{y}_2,...,\boldsymbol{y}_N|T_k) = \sum_{m}^{M}\sum_{n}^{M}\cdots\sum_{l}^{M}\sum_{q}^{M}[\pi_m a_{m,n}\cdots a_{l,q}]P$$
$$P = [p(\boldsymbol{y}_1|S_m,T_k)p(\boldsymbol{y}_2|S_n,T_k)\cdots p(\boldsymbol{y}_{N-1}|S_l,T_k)p(\boldsymbol{y}_N|S_q,T_k)] \qquad (3)$$

where we sum the indices m, n, l and q over all M states for target T_k. In (3) we have utilized the Markov state-transition model from Sec. IIIA.

The form in (3) is characteristic of a hidden Markov model (HMM) [3], and the requisite multiple sums can be evaluated efficiently via the well-known forward-backward algorithm [3]. Alternatively, recently HMM researchers have employed the Viterbi algorithm [6].

In (1) and (2) we have defined the initial estimates for π and $\boldsymbol{A}$; it remains to describe how the continuous distribution $p(\boldsymbol{y}_n \mid S_m, T_k)$ is generated, it defined as the probability of extracting the feature vector $\boldsymbol{y}_n$ when the target-sensor orientation is in state S_m of target T_k. We consider training data from state S_m and target T_k, from which we determine the distribution of the associated matching-pursuits parameter vectors $\boldsymbol{y}$ in the FI-dimensional parameter space. Using the K-means algorithm [7] in conjunction with the Mahalanobis distance [8] metric, we generate L discrete FI-dimensional vectors $\boldsymbol{v}_1$, ..., $\boldsymbol{v}_L$, which coarsely represent the distribution of the training data. Subsequently, The Mahalanobis distance [8] is then used to cluster the training data, with C_l representing the cluster associated with $\boldsymbol{v}_l$. We represent C_l with a Gaussian distribution $g_l(\boldsymbol{y}_n \mid S_m, T_k)$, with mean $E(\boldsymbol{y}_n) = \boldsymbol{v}_l$, $\boldsymbol{y}_n \in C_l$. Moreover, for simplicity, we approximate the FI parameters as uncorrelated. Finally, we have the mixture distribution

$$p(\boldsymbol{y}_n|S_m,T_k) = \sum_{l=1}^{L} w_l g_l(\boldsymbol{y}_n|S_m,T_k) \ , \qquad \sum_{l=1}^{L} w_l = 1 \qquad (4)$$

The coefficients w_l are weighted to reflect the relative number of elements in cluster l.

C. Training of physics-based HMM

The statistic in (3) considers all possible M^N state sequences, weighted by the associated state-transition probabilities. The Viterbi algorithm yields the maximum-likelihood (ML) state sequence, defining the new likelihood function

$$p'(\mathbf{y_1},\mathbf{y_2},\cdots,\mathbf{y_N}|T_k) = \max_{\boldsymbol{S}} [\pi_m a_{m,n}\cdots a_{l,q}]P$$
$$P=[p(\mathbf{y_1}|S_m,T_k)p(\mathbf{y_2}|S_n,T_k)\cdots p(\mathbf{y_{N-1}}|S_l,T_k)p(\mathbf{y_N}|S_q,T_k)] \quad (5)$$

In addition to determining the ML state sequence, (5) can be used to refine the initial estimates for the vector $\pi=[\pi_1, \pi_2, ..., \pi_M]$ and the tri-diagonal $M\times M$ dimensional matrix $\boldsymbol{A}$, with elements *initially* defined in (1) and (2), respectively. In particular, recall that (1) and (2) are based on an initial angular state decomposition, inaccuracies in which will yield errors in π and $\boldsymbol{A}$, as well as errors in the probabilities $p(\mathbf{y}_n \mid S_m, T_k)$. Here we use the Viterbi algorithm as a tool to update the state decompositions as well as the probabilities π, $\boldsymbol{A}$, and $p(\mathbf{y}_n \mid S_m, T_k)$; (1) and (2) and the initial state decomposition can be viewed as starting points for subsequent optimization.

For a given set of data representative of target T_k, we use the Viterbi algorithm [6] to determine the ML state sequence for each set of N multi-aspect scattered waveforms available from the training data. For a fixed sensor rate of motion, each set of N multi-aspect scattered waveforms is representative of a different initial target-sensor orientation. By defining the ML state sequence for each set of N training waveforms, the Viterbi algorithm [6] associates each aspect-dependent scattered waveform with a particular state (not necessarily the same as the initial state decomposition). Therefore, in the training cycle, we update the states each scattered waveform is associated with. Moreover, after performing Viterbi ML estimation on all N-waveform training sequences, we have available the number of times state m was selected for waveform n=1, allowing a reestimation of π_m. Further, for each state m, we know the number of state transitions that occurred to state l, allowing a reestimation of $a_{m,l}$, permitting an updating of the state-transition-probability matrix $\boldsymbol{A}$. Finally, since the association between scattered waveforms and states has been updated via Viterbi estimation, we can update the probabilities $p(\mathbf{y}_n \mid S_m, T_k)$ as well. Therefore, after the first training cycle the Viterbi algorithm [6] updates the HMM parameters π, $\boldsymbol{A}$, and $p(\mathbf{y}_n \mid S_m, T_k)$. The process is then repeated, now performing Viterbi ML estimation on the refined HMM model. This again yields updated HMM parameters, and the process is iterated until the changes in $p(\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_N \mid T_k)$ are less than a prescribed threshold, at which point the HMM is declared optimized for the available training data.

IV. Example Results

A. Submerged elastic targets considered

We consider measured acoustic backscattering data from five submerged elastic targets: (1) a cylindrical air-filled shell; (2) a duplicate shell, stiffened with equally spaced ring stiffeners; (3) a duplicate shell, stiffened with unequally spaced ring stiffeners and augmented by resiliently mounted, elastic internal rods; (4) a slightly larger, nearly periodically ribbed shell; and (5) a duplicate of shell 4, with a large number (~1000) of attached internal oscillators. Specifics on the measurement system and further details on the targets measured by the target facility section at the Naval Research Laboratory can be found in [9-11]. In particular, Targets 1 and 3 are described in [10], Target 4 in [9], and Targets 4 and 5 in [11]. These targets are geometrically similar, providing a challenging test to the HMM identification algorithm. Each of the targets was ensonified by an acoustic waveform with bandwidth from approximately 11-40 KHz, corresponding to relative target dimensions of $2.9 \le ka \le 10.4$, where k is the wavenumber and a is the average target radius. Finally, for each target, a five-state HMM was utilized.

B. Continuous vs. discrete HMMs

As an alternative to the continuous mixture density defined in (2), one can use a discrete probability function to describe the state-dependent distribution of the matching-pursuits parameter vector $\mathbf{y}_n$. This is done by defining J discrete elements $\mathbf{c}_j$ in the FI-dimensional parameter space, with these elements constituting a "codebook" $\boldsymbol{C}$ [7]. Each $\mathbf{y}_n$, extracted via matching pursuits for the nth scattered waveform $\boldsymbol{f}_n$, is mapped to its nearest neighbor in $\boldsymbol{C}$ (using the Mahalonobis distance [8]). In this manner, the continuous FI--dimensional space is effectively discretized into J regions, with b_{jm} representing the probability of realizing codebook element $\mathbf{c}_j$ in state m. The probabilities b_{jm} define a $J\times M$ dimensional matrix $\boldsymbol{B}$, which is "learned" during a training phase similar to that discussed in Sec. III.

We perform a comparison between the discrete and continuous HMM. With regard to the former, we utilize a 64-element codebook, and the vector π and matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are optimized via the Viterbi algorithm [6]. The codebook was generated via the K-means algorithm [7]. In Fig. 2 we show a comparison of discrete and continuous HMM performance, as a function of number of observations N, assuming 5° angular

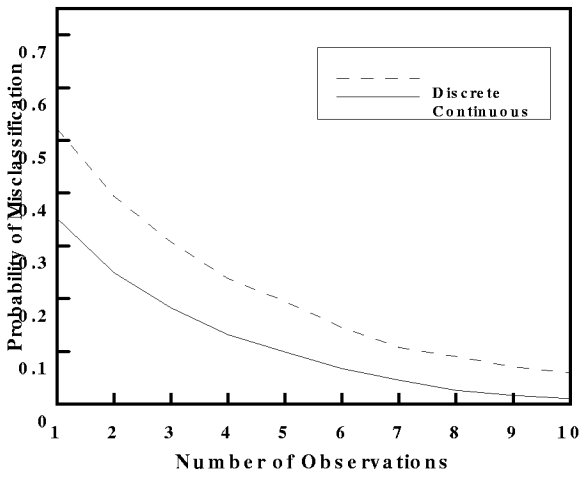

Figure 2. Average discrete and continuous HMM performance. Average SNR: 20 dB. sampling.

For both the continuous and discrete HMMs, each target required five states. The results in Fig. 2 are for noisy data, with the noise having a power spectral density (PSD) identical to the spectrum of the incident pulse, in an attempt to simulate additive clutter. For both the discrete and continuous HMMs, training was performed by using all 360 backscattered waveforms from each target (1° sampling was used in the measurements), and 15 noise realizations were considered for each scattered waveform. Since there are 181 unique observation sequences, we trained on 15×181 sequences. The testing was performed using all possible sequences of the 360 backscattered waveforms (for the N measurements characteristic of a given sequence), and 4 noise realizations, i.e., 4×360 test sequences. Since the noise realizations used for testing and training were different, the HMMs were tested and trained on distinct data. In Fig. 2, N=10 samples corresponds to a total of 45° of data (5° sampling), for which we see an average misclassification probability of less than 0.1, for both the discrete and continuous HMMs.

C. Confusion matrix

The results in Fig.2 present the *average* misclassification probability, for the discrete and continuous HMMs, a misclassification defined as identifying the data as characteristic of target T_k when it was actually scattered from target T_i. Additional information is presented in the form of a confusion matrix, as presented in Tables 1 and 2, for the discrete and continuous HMMs, respectively. Using the target designations described in Sec. IVA, the confusion matrix considers data from target T_k, and quantifies the probability that it is classified as target T_i, for i and k from one to five (for the five targets considered here). Several observations can be made from these tables, which considered the same 20 dB average SNR data presented in Fig. 2, for the case of N=5 scattered waveforms (20° of data, with 5° sampling). First, as in Fig. 2, these tables demonstrate that the continuous HMM consistently outperforms its discrete counterpart. Moreover, for both the continuous and discrete HMM, data from target 4 had the highest probability of being properly classified. This is attributed to the Bloch wave [9] excited on the ribbed shell, which is characterized by an angle-dependent frequency that is clearly visible and distinct [9]. From Tables 1 and 2 we also note that targets 2 and 3 are often confused for one another. This is expected because, from Sec. IVA and [9-11], these targets are very similar geometrically.

True Target \ Classified Target	T1	T2	T3	T4	T5
T1	81.35	7.87	4.28	0.55	
T2	7.87	72.37	12.43	0.00	
T3	4.97	11.60	78.72	0.96	
T4	0.00	0.13	0.82	96.27	
T5					

Table 1. Confusion matrix for the discrete HMM, in percent.

True Target \ Classified Target	T1	T2	T3	T4	T5
T1	91.29	6.21	0.96	0.00	
T2	1.24	86.18	10.91	0.00	
T3	2.90	10.77	85.63	0.00	
T4	0.00	0.00	0.13	99.86	
T5					

Table 2. Confusion matrix for the continuous HMM, in percent.

V. Conclusions

Wave-based matching pursuits and a hidden Markov model (HMM) have been presented for the identification of concealed or distant targets. We have utilized the underlying wave physics to build a wavefront-resonance dictionary and as motivation for dividing the aspect-dependent scattered fields into "states", angular-dependent sectors over which the wave physics is slowly varying. The results from this investigation clearly demonstrate the superiority of the continuous HMM relative to its discrete counterpart.

References

[1] E. Heyman and L.B. Felsen, "A wavefront interpretation of the singularity expansion method," *IEEE Trans. Antennas Propag.*, vol. 33, pp. 706-718, 1985.

[2] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd edition, McGraw-Hill, 1991.

[3] L.R. Rabiner, "A tutorial on hidden Markov models," *Proc. IEEE*, Vol. 77, pp. 257-285, Feb. 1989.

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

[5] M. McClure and L.Carin, "Matching pursuits with a wave-based dictionary," *IEEE Trans. Signal Proc.*, vol. 45, pp. 2912-2927, Dec. 1997.

[6] A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inform. Theory*, vol. 13, pp. 260-269, April 1967.

[7] Y. Linde, et al., "An algorithm for vector quantizer design," *IEEE Trans. Comm.*, Vol. 28, pp. 84-95, Jan. 1980.

[8] L.L. Scharf, *Statistical Signal Processing, Detection, Estimation, and Time Series Analysis*, Addison-Wesley, 1990.

[9] B.H. Houston, et al., "Broadband acoustic scattering from a ribbed shell," *JASA*, vol. 98, pp. 2851-2861, Nov. 1995.

[10] J.E. Bondaryk et al., "Array processing for fluid-loaded, cylindrical shells," *JASA.*, vol. 97, pp. 1067-1074, Feb. 1995

[11] D.M. Photiadis, et al., "The effect of internal oscillators on the acoustic response of a submerged shell," *JASA.*, vol. 101, pp. 895-995, Feb. 1997.