Detection, Classification and Tracking in Distributed Sensor Networks

Dan Li, Kerry Wong, Yu Hen Hu and Akbar Sayeed
Department of Electrical and Computer Engineering
University of Wisconsin-Madison, USA
akbar, hu@engr.wisc.edu
kerryw, dli@cae.wisc.edu

Abstract – We outline a framework for collaborative signal processing in distributed sensor networks. The ideas are presented in the context of tracking multiple moving objects in a sensor field. The key steps involved in the tracking procedure include event detection, target classification, and estimation/prediction of target location. Algorithms for various tasks are discussed and some recent results on experiments with real data are reported. Directions for ongoing and future research are discussed.

Keywords: collaborative signal processing, sensor networks, detection, target localization, classification, tracking

1 Introduction

Networks of small, densely distributed wireless sensor nodes are being envisioned and developed for a variety of applications involving monitoring and manipulating of the physical world in a tetherless fashion. Typically, each individual node can sense in multiple modalities but has limited communication and computation capabilities. Many challenges must be overcome before the vision of sensor networks becomes a reality, including efficient methods for exchanging information between the nodes and collaborative signal processing between the nodes to gather useful information about the physical world.

This paper describes the key ideas behind the algorithms being developed at the University of Wisconsin (UW) for collaborative signal processing (CSP) in distributed sensor networks. We also describe the basic ideas on how the CSP algorithms interface with the networking/routing algorithms being developed by the UW-API team. We motivate the framework via the problem of detecting and tracking a single maneuvering target. This example illustrates the essential ideas behind the integration between UW-API and UW-CSP algorithms and also highlights the key aspects of detection and localization algorithms. We then build on these ideas to present our approach to tracking multiple targets that necessarily requires classification techniques. In the present form, all our algorithms are based on processing a single sensing modality, such as seismic or acoustic. Furthermore, current detection and classification algorithms are based on single node processing, whereas localization and tracking algorithms require collaboration between sensors.

2 Detection and Tracking Algorithm

2.1 Single Target

Figure 1 illustrates the basic idea of a region-based collaborative signal processing for detection and tracking of a single target. Under the assumption that a potential target may enter the monitored area via one of the four corners, four regions, A, B, C and D, are created by the UW-API protocols. Sensor nodes in each of the four regions are activated to detect potential targets. Some of the nodes in each region are designated manager nodes for coordinating the processing in that region.

Figure 1. A schematic illustrating detection and tracking of a single target using UW-API and UW-CSP algorithms.

In each region, every activated sensor node runs an energy detection algorithm whose output is sampled at an a priori fixed rate depending on the maximum velocity of the target. Suppose a target enters Region A. Tracking of the target consists of the following steps.
(a) Some of the nodes in Region A in the vicinity of the target detect the target. These nodes are the active nodes. The active nodes also yield closest point of arrival (CPA) information. The active nodes report their energy detector outputs to the manager nodes at \( N \) successive time instants (defining an observation period determined by the a priori information on target velocity and the size of the region).

(b) At each time instant, the manager nodes determine the location of the target (using an algorithm described in Section 3.2) from the energy detector outputs of the active sensors. Ideally, this location determination is at a finer resolution than the sensor spacing.

(c) The manager nodes use locations of the target at the \( N \) successive time instants to predict the location of the target at \( M (\leq N) \) future time instants.

(d) The predicted positions of the target are used by the UW-API protocols to create new regions that the target is likely to enter. This is illustrated in Figure 1 where the three dotted regions represent the regions that the target is likely to enter after region A. A subset of these regions is activated by the UW-API protocols for subsequent detection and tracking of the target.

(e) Once the target is detected in one of the new regions, the sensors in the original region (Region A in Figure 1) are put in dormant state to conserve energy.

Steps (a) – (e) are repeated for the new region and this forms the basis of detecting and tracking a single target.

2.2 Multiple Targets

Figure 1 illustrates detection and tracking of a single target. Tracking of multiple targets requires more sophisticated processing. If multiple targets are sufficiently separated in space or time, essentially the same procedure as described in Section 2.1 may be used: a different track is initiated and maintained for each target. Sufficient separation in time means that the energy detector output at a particular sensor exhibits distinguishable peaks corresponding to the CPAs of the two targets. Similarly, sufficient separation in space means that at a given instant the spatial target signatures exhibit distinguishable peaks corresponding to sensors that are closest to the targets at that instant.

The assumption of sufficient separation in space and/or time may be too restrictive in general. In such cases, classification algorithms are needed that operate on the temporal target signatures to identify and classify them. This necessarily requires a priori knowledge of typical signatures for different target classes. In this case, a time series segment is generated for each detected event at a node. Some form of temporal processing, such as an FFT, is performed and the transformed vector is fed to a bank of classifiers corresponding to different possible target classes. The outputs of the classifiers that detect the particular target are reported to the manager nodes as opposed to the energy detector outputs. Steps (a) – (e) in Section 2.2 are repeated for all the active classifier outputs to generate and maintain tracks for different classified targets. Some initial results on classification are described in Section 4.

3. Signal Processing Algorithms

3.1 Detection

Energy detection is the simplest form of detection that uses minimal a priori information about the target. The detector essentially computes a running average of the signal power over a window of pre-specified length. The output of the detector is sampled at a pre-specified rate. The window duration and sampling rate are determined by target characteristics, such as the expected duration of its signature in the particular sensing modality used. An event is detected when the detector output exceeds a threshold. Due to signal averaging, the noise component in the output of the detector is modeled as a Gaussian random variable whose mean and variance can be determined from the statistics of the background noise in the original signal. The threshold is dynamically adjusted according to the variance of detector output noise so that the detector maintains a constant false alarm rate (CFAR). If the detector output is below the current threshold, the sensor signal is assumed to consist of background noise only and these sensor measurements are used to update the threshold value.

The output parameters from the energy detector that are communicated to the manager nodes consist of: 1) the onset time when the detector output exceeds the threshold, the time of the maximum (CPA), and the offset time when the detector output again falls below the threshold; and 2) the detector output at CPA. For target localization purposes, the detector output at any required instant within the offset and onset times may also be communicated. For classification purposes, the sensor time series for some fixed duration around the CPA (and within the onset and offset times) may also be communicated to the manager nodes.

3.2 Target Localization

We have developed an algorithm for estimating target location at a particular time instant by using measurements from multiple (4 or more) sensors at that time instant. Such an energy-based algorithm is an attractive alternative to existing target localization methods for the following reasons:

(a) A key requirement for accurate target localization methods, such as those based on time-delay estimation [1], is accurate synchronization among different sensor nodes. This turns out to be quite expensive in a wireless packet data sensor network.

(b) Coherent localization methods, such as beamforming [3,4], also require additional assumptions, such as
plane wave (far field) approximation for the incoming wave. Such assumptions are violated in dense sensor networks.

c) Communication among sensor nodes should be minimized to conserve power. Exchange of time series data among sensor nodes, as required by some algorithms, consumes too much energy to be feasible.

Our energy-based target localization algorithm assumes an isotropic exponential attenuation for the target energy source:

\[ y_i(t) = s(t) / \|r(t) - r_i\|^{\alpha} \]  

(1)

where \( y_i(t) \) is the energy reading at the \( i \)th sensor, \( r(t) \) denotes the unknown coordinates of the source with respect to a fixed reference, \( r_i \) are the coordinates of \( i \)th sensor, \( s(t) \) is the unknown target signal energy, and \( \alpha \) is the decay exponent which is assumed to be known (or estimated via experiments). The algorithms first computes the ratios \( y_i(t)/y_j(t) \) for all pairs of sensors to eliminate the unknown factor \( s(t) \). The unknown target location \( r(t) \) is then estimated by solving a nonlinear least squares problem involving the ratios. A sample contour plot based on the algorithm is shown in Figure 2 below.

Figure 2. An example of energy based collaborative target localization using real data obtained from SITEX00 experiment. Dots indicate sensor locations. The center of the contour plot indicates estimated target location. Dotted circles indicate possible target locations based on a pair of the sensor energy readings among the four pairs of sensors (triangles) whose readings are used for calculation.

3.3 Target Tracking

Given the trajectory of target locations in the past, it is possible to fit the data samples into a dynamic model and therefore predict the future target locations. Tracking is a complicated problem when multiple targets present. For a single target moving at constant speed, tracking can be accomplished by fitting the data into a linear or polynomial model using a least squares fit.

4 Neural Network Based Classification

As mentioned earlier, classification algorithms are needed in general for tracking multiple targets. Classification algorithms operate on time-series data associated with each detected event. However, the variability in signatures poses a significant challenge in efficient classifier design. In general, some a priori knowledge of the statistical characteristics of signatures for different target classes is required. In this section, we report results from our initial investigations in which we applied neural networks and nearest-neighbor classifiers to SITEX00 data to generate some benchmark numbers.

4.1 Spectral Signatures

The choice of feature vectors on which the classifiers operate is important. Our initial experimentation has shown
that spectral characteristics of target signatures vary significantly between different target classes and hence may be fruitful for classification. However, the presence of Doppler shifts due to motion of targets makes the use of spectral characteristics a little problematic. In Section 5, we discuss some ongoing work aimed at accounting for such effects.

To investigate the utility of spectral signatures, we applied the k-means clustering algorithm to feature vectors of individual classes to extract typical templates in each class.

Multiple temporal signature vectors of length 128 were extracted from the time series for each detected event by using 128-length overlapping segments with a 64-point overlap between them. After subtracting the mean, FFT to the 128-length vectors is computed and only the first 64 samples of the FFT spectrum are kept (since real signals). The magnitude FFT vectors were used as the feature vectors. The L2 norm is used as a distance measure between two feature vectors. Since different classes have different number of feature vectors, we extracted one typical template for approximately every twenty feature vectors. The extracted templates for LAV (wheeled vehicle) and Tank (tracked vehicle) time series are plotted in Figure 3. It is evident that these two vehicles exhibit different dominant spectral content.

### 4.2 Classifier Structure

The structure of the neural network classifier is shown in Figure 4. The input $x$ is a length 128 vector corresponding to the raw data collected from the acoustic or seismic sensors. (The input vector corresponds to half a second of the time series sampled at 256 samples per second). The Fourier transform of the input vector:

$$\tilde{x} = |\text{DFT}(x)|$$

is computed and only the first half (64 points) of the coefficients are kept in $\tilde{x}$. Finally, each DFT vector is normalized

$$\hat{x} = \frac{\tilde{x}}{||\tilde{x}||}$$

The neural network classifier was trained and tested using data from the SITEX00 repository. Suppose there are $M$ different target types to be classified. The training of the neural network associated with the $j$th target type, proceeds as follows. When the input training vector corresponds to the $j$th target, the desired output is set to one. On the other hand, the desired output is set to zero for training vectors corresponding to other targets. This training procedure is applied to all the neural networks corresponding to the different target classes. The well-known error back propagation algorithm [6] is applied for updating the weights of the hidden neurons during the training phase.

Table 1 summarizes the results of applying the neural network classifier to the SITEX00 data in the form of a confusion matrix. The matrix shows the performance of the classifier as well as the total number of signature vectors and fed to the individual classifiers for different classes.
available for each target class. The matrix was generated via a 3-way cross validation procedure. The available event signatures were divided into three groups. For each run, two groups were used for training and the remaining one for assessing the performance of the classifier. The performance is averaged over the three possible runs. Table 2 lists the probability of false alarm and the probability of correct detection computed from Table 1. However, it should be noted that the probabilities are simply a coarse way of interpreting the data in Table 1 in a different way — their accuracy is limited due to the limited data. As a comparison, Table 3 lists the probabilities of false alarm and correct detection using the k-nearest neighbor algorithm which chooses the target type with the smallest $L^2$ distance between the event signature and the templates for different target classes (Table 3).

Table 1. Confusion matrix for the neural network classifier.

<table>
<thead>
<tr>
<th>Det True</th>
<th>AAV</th>
<th>DW</th>
<th>5-ton</th>
<th>Hmv</th>
<th>LAV</th>
<th>POV</th>
<th>Tank</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAV</td>
<td>135</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>146</td>
</tr>
<tr>
<td>DW</td>
<td>5</td>
<td>147</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>164</td>
</tr>
<tr>
<td>5-ton</td>
<td>1</td>
<td>1</td>
<td>72</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>76</td>
</tr>
<tr>
<td>Hmv</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>LAV</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>184</td>
<td>20</td>
<td>0</td>
<td>210</td>
</tr>
<tr>
<td>POV</td>
<td>15</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>63</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>Tank</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>85</td>
<td>0</td>
<td>88</td>
</tr>
</tbody>
</table>

Table 2. Probability of detection ($P_D$) and probability of false alarm ($P_{FA}$) for the neural network classifier.

<table>
<thead>
<tr>
<th></th>
<th>$P_D$</th>
<th>$P_{FA}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAV</td>
<td>0.9246</td>
<td>0.0414</td>
</tr>
<tr>
<td>DW</td>
<td>0.8963</td>
<td>0.0164</td>
</tr>
<tr>
<td>5-TON</td>
<td>0.9474</td>
<td>0.0086</td>
</tr>
<tr>
<td>HMMV</td>
<td>0.8333</td>
<td>0.0013</td>
</tr>
<tr>
<td>LAV</td>
<td>0.8762</td>
<td>0.0160</td>
</tr>
<tr>
<td>POV</td>
<td>0.7500</td>
<td>0.0420</td>
</tr>
<tr>
<td>TANK</td>
<td>0.9660</td>
<td>0.0029</td>
</tr>
</tbody>
</table>

Table 3. Probability of detection ($P_D$) and probability of false alarm ($P_{FA}$) for the k-nearest neighbor classifier.

<table>
<thead>
<tr>
<th></th>
<th>$P_D$</th>
<th>$P_{FA}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAV</td>
<td>0.8797</td>
<td>0.0170</td>
</tr>
<tr>
<td>DW</td>
<td>0.8344</td>
<td>0.0844</td>
</tr>
<tr>
<td>5-TON</td>
<td>0.1667</td>
<td>0.0014</td>
</tr>
<tr>
<td>HMMV</td>
<td>0.5974</td>
<td>0.0156</td>
</tr>
<tr>
<td>LAV</td>
<td>0.7260</td>
<td>0.0170</td>
</tr>
<tr>
<td>POV</td>
<td>0.8394</td>
<td>0.0778</td>
</tr>
<tr>
<td>TANK</td>
<td>0.9195</td>
<td>0.0221</td>
</tr>
</tbody>
</table>

5 Ongoing and Future Research

As mentioned earlier, classification is critical for tracking multiple targets and, as we have seen, spectral characteristics of target signatures are useful in this context. However, the neural network classifier does not explicitly take into account the variations in spectral signatures due to related factors, such as Doppler effects. We are currently developing a framework for multiple target classification that explicitly takes Doppler effects into account. In this section, we briefly describe the rationale behind our approach.

An implicit assumption in the training of the neural network classifiers is that statistical characteristics of the target signatures do not change with time. In other words, the signatures are modeled as realizations of stationary processes. This assumption, however, does not really hold for real data. In particular, Doppler effects in acoustic and seismic signals are significant due to relatively low speed of sound in air and ground and make the data nonstationary.

5.1 Effect of Doppler on Spectral Signatures

To appreciate the significance of Doppler effects, consider the setup in Figure 6.

Figure 6. A simple geometry for a moving source to illustrate Doppler effects.

The source that emits energy at a frequency $f_0$ is moving at velocity $v$ parallel to the x-axis. The perpendicular distance between the source and the observer (sensor) is $d$. The sensor is located at a distance $x$ along the axis. A simple calculation shows that the frequency perceived by the sensor is related to source frequency as:

$$f = \frac{f_0}{1 - \frac{(v / v_s) \cos \alpha}{}}$$

where $\alpha$ is the angle between the x-axis and the line-of-sight between the source and the sensor. Figure 7 and Figure 8 plot the perceived frequency for different settings. The source frequency is $f_0 = 60$ Hz and the sensor is located at $x = 200$ m.
Figure 7 plots the perceived frequency as a function of source position for different $d$'s. It is evident that the perceived frequency is equal to the source frequency at the CPA and variation in the perceived frequency gets sharper for smaller $d$.

Figure 8 plots the perceived frequency as a function of source position for different source velocities. As expected, the Doppler shifts become larger at higher speeds. The important thing to note is that the changes in perceived frequency are significant at normal source speeds and thus must be taken into account for improved classifier performance.

Actual data from the SITEX00 experiments exhibits similar spectral trends. Figures 9 and 10 show the short-time Fourier transform (STFT) plots of a particular event in seismic and acoustic modalities, respectively. The variation in perceived frequency, similar to that in Figures 7 and 8, is evident. Furthermore, the seismic signature is shorter in time due to faster decay in the seismic modality.

5.2 Doppler-Based Composite Hypothesis Testing

In light of these observations, we are currently working on directly incorporating Doppler effects into classifier design. We illustrate the basic ideas in terms of a binary classification problem (e.g., whether it is a wheeled or a tracked target). In the presence of Doppler effects, we can model different target signatures as realizations of a wide-sense stationary (WSS) process modulated by time-varying instantaneous frequency (or Doppler profile). Furthermore, we assume that these processes have zero mean. Thus, the classification problem is analogous to a composite hypothesis-testing problem in the multivariate Gaussian model with zero means and different covariance matrices that are parameterized by the Doppler parameters. That is, under the $m^{th}$ hypothesis

$$H_m = x \sim N(0, R_n(\Omega_m)) \quad m=1,2$$
and the signature vector is a zero-mean Gaussian vector with covariance matrix $R_m(\Omega_m)$, where $\Omega_m = \begin{bmatrix} \alpha_m & v_m & d_m \end{bmatrix}$ is the parameter vector that characterizes the Doppler profile corresponding to Figure 6.

One way to deal with the Doppler parameters is to use an approach analogous to a generalized likelihood ratio test (GLRT): First, a maximum likelihood (ML) estimate of the Doppler parameters, $\hat{\Omega}_m$, is obtained for each hypothesis and then the likelihood ratio defined by the a posteriori probability density functions corresponding to these estimates, $f_m(x; \hat{\Omega}_m)$, is used for deciding between the two target classes. This is just one approach that illustrates how Doppler effect may be taken into account and by no means the best approach. In fact, ML estimates can be computationally quite expensive in general. We are currently pursuing several other directions as well.

6. Conclusion

In this paper we have presented the basic ideas behind a collaborative signal processing framework for tracking multiple targets in a distributed sensor network. The key components of the framework include event detection, estimation and prediction of target location and target classification. Most of the existing work is for tracking a single target and is based on a single sensing modality, such as acoustic or seismic. Tracking of multiple targets would in general require classification algorithms. Based on experimentation with real data, we have argued that differences in spectral signatures of different target classes may facilitate accurate classification. However, variations in spectral signatures due to Doppler effects make classification based on spectral signatures difficult. We have provided some preliminary results on neural network based classifiers that ignore Doppler effects and provide some benchmark numbers on performance. Finally, we briefly outlined an approach for classification that directly incorporates Doppler effects for potentially improved classification.

Acknowledgement

The work reported here is sponsored by the Defense Advanced Research Projects Agency (DARPA) and Air Force Research Laboratory, Air Force Material Command, USAF, under agreement number F30602-00-2-0555. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Defense Advanced Research Projects Agency (DARPA), the Air Force Research Laboratory, or the U.S. Government.

References