ARMA MODELING IN MULTICHANNEL FILTER BANKS WITH APPLICATIONS TO RADAR SIGNAL ANALYSIS

Kie B. Eom

Department of Electrical Engineering
and Computer Science
The George Washington University
Washington, DC 20052

ABSTRACT

In this paper, we consider the classification of radar signals by using stochastic models at different scales. The signal at a different scale is modeled by a hierarchical Autoregressive Moving Average (ARMA) model, and the features at coarse scales are extracted from the model without performing expensive filtering operation. The hierarchical modeling can increase the accuracy of radar signal classification by exploiting features at different scales. For radar signal classification, model parameters at five different scales obtained by hierarchical modeling are used as features. A minimum distance classifier is implemented, and is tested on real aperture radar signals.

1. INTRODUCTION

Recently, there have been studies on AR models in scale space with emphasis on model representation, parameter estimation and prediction when a signal is changed from fine to coarse (aggregation) and from coarse to fine scale (disaggregation) [1]. The hierarchical models based on multiscale AR models are potentially useful for improving the performance of radar signal processing algorithms in terms of accuracy and robustness.

In this paper, we consider the classification of radar signals by using features extracted from stochastic models at different scales. A simple approach to finding features at a coarse scale is to fit a model at a coarse scale after aggregation. The aggregation, also called as decimation filtering, is the process of changing scales from fine to coarse by filtering followed by an m:1 down sampling. After aggregation, the number of samples is reduced by m, the aggregation factor. In general, the analysis of signals at multiple scales requires aggregation over different scales, and it requires more computations than when features are extracted at one scale. The approach presented in this paper is based on the fact that the model at a coarse scale can be obtained from the model at a finer scale, if the signal follows an ARMA model. Therefore, the hierarchical modeling approach is computationally efficient because it does not perform expensive decimation filtering.

The support of the Advanced Research Projects Agency (ARPA order No. A369) and the Air Force Office of Scientific Research (Contract F49620-93-1-0576) is greatly appreciated.

When a signal following an ARMA model is aggregated, the aggregated data also follows an ARMA model. The AR polynomial of the aggregated model can be uniquely obtained from the AR polynomial of the model before aggregation. The moving average polynomial of the aggregated model is obtained from the correlation structure. Model disaggregation is a process which identifies models at finer scales from a model at a coarse scale. We generalize the model disaggregation approach in [6] to nonuniform filtering. The moving average parameters of the disaggregated model are estimated from the disaggregated correlations.

The radar signal classification algorithm presented in this paper is based on model aggregation. For each radar signal, an AR model of 30th order is fitted and parameters are estimated at the finest scale. Then the roots of the AR polynomial are sorted in order of magnitude, and two pairs of complex roots closest to the unit circle are selected as features at that scale. In coarser scales, the roots of the AR polynomial are computed by the the model aggregation algorithm. Two pairs of complex roots closest to the unit circle are selected at each scale. The features are computed over five different scales (m=1,2,4,8, and 16). A minimum distance classifier with Eucledian distance is implemented to classify radar signals. The multiscale classifier is applied to classify radar returns from 35 different targets. The results show that the multiscale classifier performs better than the classifier using features from a single scale.

2. HIERARCHICAL STOCHASTIC MODELS

In this section, we consider the stochastic modeling of signals at the output of decimation filter in Figure 1, assuming that the input process is ARMA(p,q). In an analysis filter bank, the input signal is decomposed by a bank of decimation filter in Figure 1. Suppose that the signal {x(i), i=1,...,N} at the finest scale follows an ARMA(p,q) model.

$$x(i) = \sum_{j=1}^{p} a_j x(i-j) + \sum_{j=0}^{q} b_j w(j), i = 1, ..., N \quad (1)$$

where $\{w(i)\}$ is a zero mean white noise sequence with variance σ_w^2, and a_j's and b_j's are real constants.

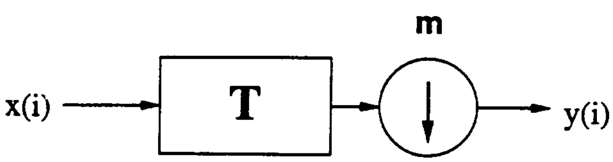

Figure 1. Decimation filter

Equation (1) can be rewritten as

$$A_p(z)x(i) = B_q(z)w(i), i = 1, ..., N \tag{2}$$

where

$$\begin{aligned} A_p(z) &= 1 - a_1 z^{-1} - a_2 z^{-2} - \cdots - a_p z^{-p}, \\ B_q(z) &= 1 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_q z^{-q} \end{aligned} \tag{3}$$

and z^{-1} is the delay operator, and we assume that the roots of $A_p(z)$ and $B_q(z)$ lie inside of the unit circle for stability and invertibility of the model.

Suppose that the signal at a coarse scale $\{y_m(i), i = 1, ..., N/m\}$ is obtained by applying an finite impulse response (FIR) filter of length L followed by an m:1 down sampling operation. This process is called the aggregation process.

$$y_m(i) = \sum_{j=0}^{L-1} h_j x(im - j) = H(z)x(im), \tag{4}$$

$$\text{where } H(z) = h_0 + h_1 z^{-1} + \cdots + h_{L-1} z^{-L+1} \tag{5}$$

In previous work [1], the aggregation by uniform filtering is considered. In this paper, we generalize previous results to arbitrary FIR filtering and show that $\{y_m(i)\}$ is an ARMA(p,q^*), where $q^* = [(p(m-1) + q + L - 1)/m]$, as summarized in Theorem 1.

Theorem 1: The aggregated data $\{y_m(i)\}$ defined (4) follows an ARMA(p, q^*) model in (6), where $q^* = [(p(m-1) + q + L - 1)/m]$.

$$C_p(z)y_m(i) = D_{q^*}(z)v(i), \tag{6}$$

where the AR polynomial $C_p(z)$ is

$$\begin{aligned} C_p(z) &= 1 - c_1 z^{-1} - c_2 z^{-2} - \cdots - c_p z^{-p} \\ &= (1 - r_1^m z^{-1})(1 - r_2^m z^{-1}) \cdots (1 - r_p^m z^{-1}) \end{aligned} \tag{7}$$

$r_1, ..., r_p$ are roots of $A_p(z)$, and the MA polynomial $D_{q^*}(z)$ is

$$D_{q^*}(z) = 1 + d_1 z^{-1} + d_2 z^{-2} - \cdots + d_{q^*} z^{-q^*} \tag{8}$$

Corollary 1: If the MA order of disaggregated model is less than or equal to AR order ($q \leq p$), then the MA order of aggregated model is also less than equal to AR order ($q^* \leq p$).

The parameters of MA polynomial $D_{q^*}(z)$ of $\{y_m(i)\}$ can be estimated by solving simultaneous nonlinear equations of correlations, and the correlation of $\{y_m(i)\}$ is computed in terms of the correlations of $\{x(i)\}$

$$\begin{aligned} R_{yy}(k) &= H(z)H(z^{-1})R_{xx}(mk) \\ &= (z^{-(m-1)} + g_{m-2} z^{-(m-2)} + \cdots + g_1 z^{-1} + g_0 \\ &\quad + g_1 z + \cdots + g_{m-2} z^{m-2} g_{m-1} z^{m-1})R_{xx}(mk), \end{aligned} \tag{9}$$

where g_i is the coefficient of z^i in the polynomial $H(z)H(z^{-1})$.

Since $R_{xx}(-k) = R_{xx}(k)$, (9) can be rewritten as follows for $L \geq 0$.

$$\begin{bmatrix} R_{yy}(0) \\ R_{yy}(1) \\ \cdot \\ \cdot \\ R_{yy}(L) \end{bmatrix} = G_m(L) \begin{bmatrix} R_{xx}(0) \\ R_{xx}(1) \\ \cdot \\ \cdot \\ R_{xx}(mL + m - 1) \end{bmatrix} \tag{10}$$

where $G_m(L)$ is the following Lx((L+1)m) matrix.

$$\begin{bmatrix} \mathbf{g^0} & \mathbf{0m} & \mathbf{0m} & \mathbf{0m} & \cdot & \cdot & \cdot & \mathbf{0m} \\ \mathbf{g^-} & \mathbf{g^+} & \mathbf{0m} & \mathbf{0m} & \cdot & \cdot & \cdot & \mathbf{0m} \\ \mathbf{0m} & \mathbf{g^-} & \mathbf{g^+} & \mathbf{0m} & \cdot & \cdot & \cdot & \mathbf{0m} \\ \cdot & & \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \mathbf{0m} & \mathbf{0m} & \cdot & \cdot & \cdot & \mathbf{0m} & \mathbf{g^-} & \mathbf{g^+} \end{bmatrix}. \tag{11}$$

In (11), $\mathbf{g^0}$, $\mathbf{g^-}$, $\mathbf{g^+}$, and $\mathbf{0m}$ are row vectors of m elements defined by the following equations.

$$\mathbf{g^0} = [\ g_0 \quad 2g_1 \quad 2g_2 \quad \cdot \quad \cdot \quad 2g_{m-1}\ ] \tag{12}$$

$$\mathbf{g^-} = [\ 0 \quad g_{m-1} \quad g_{m-2} \quad \cdot \quad \cdot \quad \cdot \quad g_1\ ] \tag{13}$$

$$\mathbf{g^+} = [\ g_0 \quad g_1 \quad g_2 \quad \cdot \quad \cdot \quad \cdot \quad g_{m-1}\ ] \tag{14}$$

$$\mathbf{0m} = [\ 0 \quad 0 \quad 0 \quad \cdot \quad \cdot \quad \cdot \quad 0\ ] \tag{15}$$

For $k > q^*$, the correlations $R_{yy}(k)$ can be recursively computed as

$$R_{yy}(k) = \sum_{i=1}^{p} c_i R_{yy}(k - i). \tag{16}$$

3. MODEL DISAGGREGATION

Model disaggregation is the identification of finer scale models from a coarser scale model. Suppose that $\{y_m(i)\}$ is the time series obtained by aggregating $\{x(i)\}$ as shown in (4). The disaggregated sequence $\{x(i)\}$ is not observable, and is assumed to follows an ARMA model with unknown parameters. By Theorem 1, the order of AR polynomial does not change after aggregation. However, there are many-to-one relations between AR polynomials of disaggregated and aggregated models. We resolve this ambiguity in disaggregation of AR polynomial by selecting minimum phase poles.

Suppose that $s_1, ..., s_p$ are roots of AR polynomial $C_p(z)$ of the aggregated model. Then $C_p(z)$ can be factorized as shown in the following equation.

$$C_p(z) = \prod_{j=1}^{p} (1 - s_j z^{-1}) \tag{17}$$

The following disaggregated AR polynomial is obtained by choosing the minimum phase roots.

$$A_p(z) = \prod_{j=1}^{p}(1 - r_j z^{-1}), \tag{18}$$

where r_j is the minimum phase root which maps to s_j by m-th power operation.

Now, we need to find MA polynomial $B_q(z)$ of the disaggregated model. The MA polynomial $B_q(z)$ of the disaggregated model can be estimated from the correlation function of the disaggregated data $\{R_{xx}(k)\}$. However, the disaggregated data and its correlation function is not available, and we need to estimate $R_{xx}(k)$ from the correlation function of the aggregated data $\{R_{yy}(k)\}$. By Corollary 1, $q^* \leq p$ if $q \leq p$, and we assume that $q^* \leq p$ and $q \leq p$. Then we need to consider $R_{yy}(k)$ for $k \leq p$ because $R_{yy}(k)$ does not depend on MA parameters for $k > p$. Therefore, from (10),

$$\begin{bmatrix} R_{yy}(0) \\ R_{yy}(1) \\ \vdots \\ R_{yy}(p) \end{bmatrix} = G_m(p) \begin{bmatrix} R_{xx}(0) \\ R_{xx}(1) \\ \vdots \\ R_{xx}(mp+m-1) \end{bmatrix} \tag{19}$$

where $G_m(p)$ is defined in (11). However, $R_{xx}(k)$ for $k > p$ can be obtained in terms of $\{R_{xx}(k),\ k \leq p\}$ and the AR parameters using the following relation.

$$R_{xx}(k) = \sum_{i=1}^{p} a_i R_{xx}(k-i) \tag{20}$$

Therefore, we have

$$\begin{bmatrix} R_{xx}(0) \\ R_{xx}(1) \\ R_{xx}(2) \\ \vdots \\ R_{xx}(mp+L-1) \end{bmatrix} = \begin{bmatrix} I \\ ----- \\ \mathbf{f}_{p+1}^T \\ \vdots \\ \mathbf{f}_{mp+m-1}^T \end{bmatrix} \begin{bmatrix} R_{xx}(0) \\ R_{xx}(1) \\ R_{xx}(2) \\ \vdots \\ R_{xx}(p) \end{bmatrix}, \tag{21}$$

where I is the (p+1)x(p+1) identity matrix, and $\mathbf{f}_{p+1}, \cdots, \mathbf{f}_{mp+m-1}$ are p-vectors determined recursively in terms of AR parameters.

$$\mathbf{f}_{p+1} = \begin{bmatrix} 0 \\ a_p \\ a_{p-1} \\ \vdots \\ a_1 \end{bmatrix}, \text{ and } \mathbf{f}_{p+k} = \sum_{i=1}^{k-1} a_i \mathbf{f}_{p+k-i} + \begin{bmatrix} 0 \\ \vdots \\ a_p \\ \vdots \\ a_k \end{bmatrix}, \tag{22}$$

for k=2,...,(p+1)(m-1).

Define

$$F = G_m(p) \begin{bmatrix} I \\ ----- \\ \mathbf{f}_p^T \\ \vdots \\ \mathbf{f}_{mp+m-1}^T \end{bmatrix} \tag{23}$$

The matrix F is invertible under the assumption of no hidden periodicity [6]. Therefore from (19), the correlations of disaggregated data, $R_{xx}(k)$ are obtained from the correlations of aggregated data as

$$\begin{bmatrix} R_{xx}(0) \\ R_{xx}(1) \\ \vdots \\ R_{xx}(p) \end{bmatrix} = F^{-1} \begin{bmatrix} R_{yy}(0) \\ R_{yy}(1) \\ \vdots \\ R_{yy}(p) \end{bmatrix} \tag{24}$$

The correlations $R_{xx}(k)$ for $k > p$ are obtained using (20). The MA parameters of disaggregated model are estimated from the correlation function.

4. APPLICATION TO RADAR SIGNAL CLASSIFICATION

A signal at different scales gives features that are not observable from a single scale. The hierarchical approach discussed in previous sections enable us to find models at different scales without actual aggregation and resampling. Therefore, we can incorporate features from different scales in radar classification without adding much computational complexity by using a hierarchical model.

The hierarchical modeling approaches in multichannel filter banks are applied to radar signature classification. Figure 2 shows MMW RAR signatures of a T-72 tank and ZIL truck. As we can observe from Figure 2, radar returns from two targets are noisy but have distinct features. Even at a coarser scale, differences between two radar signatures can be observed, and it demonstrates the importance of coarser scale features in radar signal classification. Since features at different scales are not observable from a single scale, classifiers using multiscale features may improve the classification accuracy. However, filtering and resampling for obtaining coarser scale signals require more computation. Fortunately, the hierarchical modeling approach provides us a tool for computing features at coarser scales without adding significant computational burden.

For radar signal classification, we used spectral features extracted by fitting an AR(30) model to the radar signature. Spectral features are frequently used in signal classification applications, such as speech classification. Each radar return from a target has distinct spectral peaks caused by its target shapes. Therefore, the spectral envelope of a radar signature can characterize a target, and spectral peaks are well detected by the AR spectral estimation method. However, AR spectral estimation approach also generates spurious spectral peaks. The effects of spurious spectral peaks can be removed by using features at different scales. Figure 3 shows the estimated power spectra which are estimated by fitting AR(30) models to RAR signals in Figure 2. As we can observe in Figure 3, the spectral envelopes of two targets are quite different.

As discussed in Section 2, the signal at a coarser scale follows an ARMA model if the signal at a finer scale is an ARMA process. The model parameters at a coarser scale can be obtained from the model parameters at a finer scale. For extracting spectral features at a coarser scale, AR polynomial for the coarser scale model should be computed. The AR polynomials of the model at coarser scales (m=2,4,8,16) are obtained by Theorem 1. The roots of AR polynomials are good features for RAR signature classification, because spectral peaks are well detected by them. We choose two pairs of dominant complex poles at each scale. The dominant poles are selected by sorting roots of AR polynomial by magnitude. Therefore, we have four features corresponding to real and imaginary parts of two pairs of complex poles. Since a radar return consists of even and odd polarimetric signatures, features from both signatures are combined (eight features). A minimum distance classifier is implemented in the feature space using the Eucledian distance measure to classify the RAR signatures.

Thirty-five (35) radar returns are classified by the multiscale classifier based on hierarchical modeling. Each radar return consists of even bounce (LL) and odd bounce (LR) polarimetric signatures, and each return has 128 channels. Since each channel in radar return is from stationary targets, multiple channels are averaged to reduce noise in radar signature. From each radar return, a training signal is obtained by averaging outputs from the first 64 channels. The test signal is obtained by averaging the remaining 64 channels. Then model parameters are estimated by fitting an AR(30) model to each signal to extract spectral features. At each scale, 8 features corresponding to the dominant spectral peaks are extracted using the hierarchical modeling approach. Radar returns are classified using features from five different scales (m=1,2,4,8,16), In our experiments with 35 RAR returns, 86 percent of the targets are correctly classified.

Acknowlegement

The author wishes to thank Professor R. Chellappa for valuable discussions and support for this research.

5. REFERENCES

[1] K.-B. Eom and R. Chellappa, "Hierarchical statistical modeling for speech compression," *Proc. 1994 International Conference on Acoustics, Speech, and Signal Processing*, April, 1994.

[2] S. G. Mallat, "A theory for multiresolution signal decomposition: The Wavelet representation," *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol. 11, pp. 674-693.

[3] L. M. Novak, "A Comparison of 1-D and 2-D algorithms for radar target classification," *Proc. IEEE Int'l Conf. on Systems Engineering*, August, 1991.

[4] O. Rioul, "A discrete time multiresolution theory," *IEEE Trans. on Signal Processing*, vol. 41, pp. 2591-2606, 1993.

[5] P. P. Vaidyanathan, "Multirate digital filters, polyphase networks, and applications: A tutorial," *Proc. of IEEE*, vol. 78, pp. 56-93, 1993.

[6] W.W.S. Wei and D.O. Stram, "Disaggregation of time series models," *J. R. Statist. Soc. B.* vol. 52, pp. 453-467, 1990.

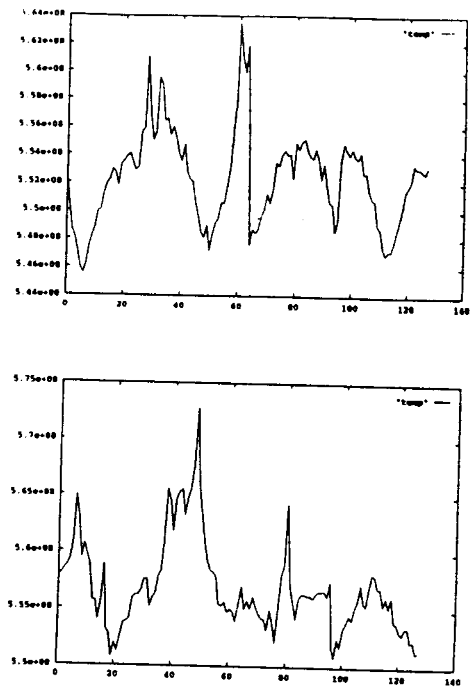

Figure 2. MMW radar returns (even bounce) from T-72 tank and ZIL truck.

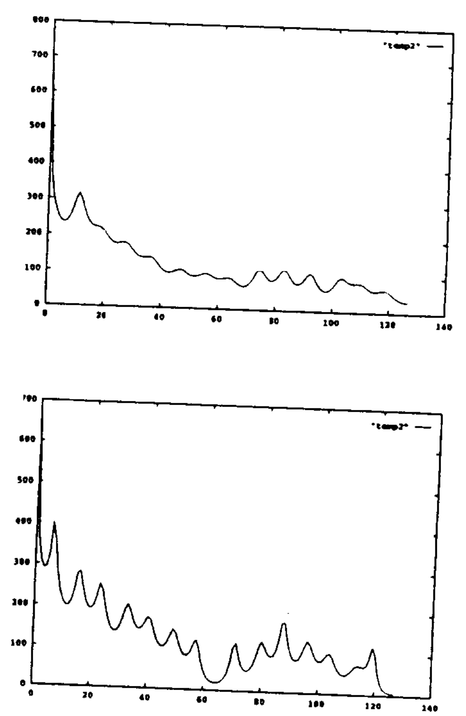

Figure 3. Power spectra estimated by fitting AR(30) models to MMW radar returns in Figure 2.