Wavelet Based Recognition Using Model Theory for Feature Selection

Zbigniew Korona and **Mieczyslaw M. Kokar**

Northeastern University, 334 Snell Building
Boston, Massachusetts 02115
zbigniew@meceng.coe.neu.edu
kokar@meceng.coe.neu.edu

April 10, 1997

Abstract

An increase in accuracy and reduction in computational complexity of the common wavelet-based target recognition techniques can be achieved by using interpretable features for recognition. In this work, the Best Discrimination Basis Algorithm (BDBA) is applied to select the most discriminant complete orthonormal wavelet basis for recognition purposes. The BDBA uses a relative entropy criterion as a discriminant measure. Then, interpretable features are selected from the most discriminant basis by utilizing symbolic knowledge about the domain. The domain theory that contains this symbolic knowledge is implemented in a backpropagation neural network. The output of the backpropagation neural network gives a final recognition decision. The results of our simulations show that the recognition accuracy of the proposed Automatic Feature Based Recognition System (AFBRS) is better than the recognition accuracy of a system that performs recognition using the Most Discriminant Wavelet Coefficients (MDWC).

***Keywords*— backpropagation neural network, best basis, entropy, feature-based recognition, model theory, target recognition, wavelet packet decomposition**

1 Introduction

One of the obvious advantages of feature-based approach to recognition is the reduction of dimensionality of the input data space by extracting certain features that distinguish input signals. A variety of approaches are known for feature extraction. Commonly used techniques include the Fourier transform, moment feature space, the Hough transform, Wigner distribution feature space, orthogonal polynomials, and Gabor functions [9]. Recently, the most popular technique is the wavelet transform, replacing the Fourier transform in many applications (for example in fingerprint identification used by FBI).

The reason for the popularity of the wavelet transform as a feature extractor is the fact that the important target features are often expressed by combination of edges, spikes and transients in a measurement signal. Therefore, these features are characterized by local information in both the time and frequency domain. With FFT, signals can be analyzed locally in the frequency domain but not in the time domain. On the other hand, the wavelet transform enables to extract features in both time and frequency domains simultaneously [10].

In this work, we describe a feature-based methodology for target recognition which we call Automatic Feature Based Recognition System (AFBRS). The goal of this system is to increase accuracy of the commonly used wavelet-based target recognition techniques by incorporating symbolic knowledge about the domain. We utilize the following four main techniques within the AFBRS framework: (i) Discrete Wavelet Packet Decomposition (DWPD) as a tool for time-frequency signal decomposition, (ii) Best Discrimination Basis Algorithm (BDBA) as a tool for most discriminant basis selection, (iii) model theory as a framework for feature selection and feature extraction, and (iv) backpropagation neural network as a tool for classification.

The BDBA is used to select a complete orthonormal basis that maximizes a discriminant measure (statistical distance among classes) for a given reference database of target signatures. We use a relative entropy as a discriminant measure. The selected basis is the most discriminant basis in relation to the relative entropy criterion. The BDBA is based on the best basis selection technique developed by Coifman and Wickerhauser [2] and extended by Saito [10] to classification problems. For the purpose of this work, the elements (wavelet coefficients) of the most discriminant basis are considered as features.

The framework of model theory [1] is applied to reduce the number of features used in recognition by selecting only those features that are interpretable in terms of symbolic target features. The examples of symbolic target features are height, length, width, shape, color, and reflectance of the target. Such symbolic features are well understood and usually provide sufficient discriminant power for recognition. The framework of model theory consists of formal languages, models, and theories. Formal languages describe the world and the sensing process. Models represent sensor data, operations on data, and relations among the data. Theories represent symbolic knowledge about the world (domain) and about the sensors. Such a consistent representational structure helps us to maintain the correctness of data interpretation [8].

Features extracted from a measurement signal using the BDBA and model theory are used as inputs to a backpropagation neural network. The output of the backpropagation neural network gives a final recognition decision.

In this work, we show that the knowledge of the domain theory allows us to select the features (out of many possible features in the most discriminant basis) which result in an increase in the recognition accuracy of the AFBRS over a recognition system using Most Discriminant Wavelet Coefficients (MDWC). The MDWC are determined by applying a relative entropy criterion to a reference database of target signatures; the MDWC maximize the relative entropy criterion. The problem of distinguishing rectangles from triangles is used as an example of a recognition problem.

Section 2 describes the Discrete Wavelet Packet Decomposition (DWPD). Section 3 introduces the Best Discrimination Basis Algorithm (BDBA). Section 4 presents a framework of model theory. Section 5 conatins a conceptual description of the proposed recognition methodology (ABFRS). Section 6 describes a feature selection procedure for a recognition problem of distinguishing rectangles from triangles. The simulation results are shown in Section 7.

2 Discrete Wavelet Packet Decomposition

A discrete signal can be decomposed into multiresolution frequency channels by using Discrete Packet Wavelet Decomposition (DWPD). For a given signal $s^j(n) \in l^2(\mathbf{Z})$ at resolution level j, a lower resolution low frequency signal can be derived by using a lowpass filter with an impulse response $h(n)$ and then by subsampling the output of the lowpass filter by two [5],

$$w_1^{j-1}(n) = \sum_m h(2n-m)s^j(m). \tag{1}$$

Similarly, a lower resolution high frequency signal can be derived by using a highpass filter with an impulse response $g(n)$ and then by subsampling the output of highpass filtering by two [5],

$$w_0^{j-1}(n) = \sum_m g(2n-m)s^j(m). \tag{2}$$

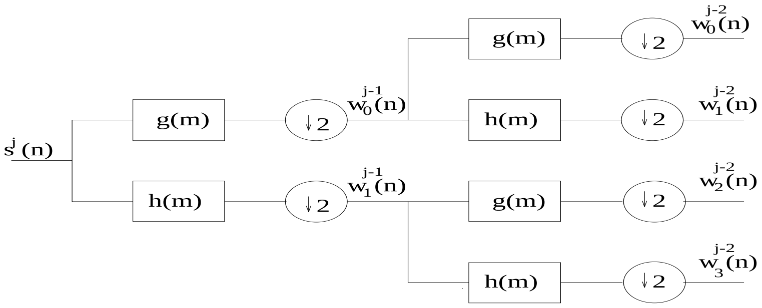

Figure 1: Discrete Wavelet Packet Decomposition

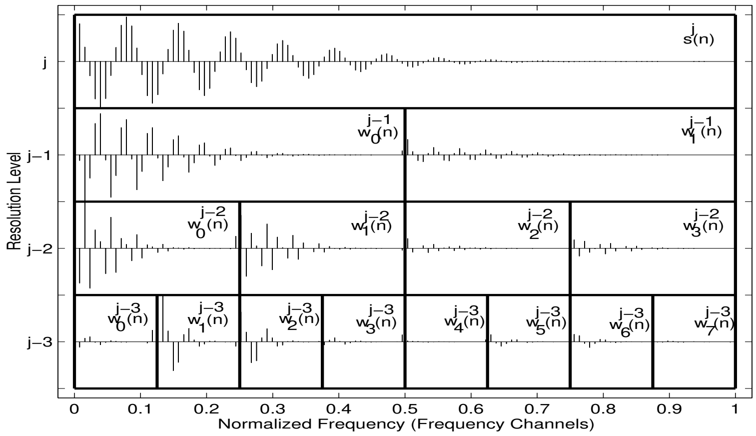

Figure 2: DWPD of Exponentially Modulated Signal

The original signal $s^j(n)$ can be recovered from the two lower resolution signals $w_1^{j-1}(n)$ and $w_0^{j-1}(n)$. However, filters $h(n)$ and $g(n)$ must meet some constraints in order to guarantee perfect reconstruction: the regularity constraint [4] and the orthonormality of the filter impulse responses. In such a case eqs. 1 and 2 can be considered as a decomposition of the signal onto an orthonormal basis, where $w_1^{j-1}(n)$ and $w_0^{j-1}(n)$ are vectors of wavelet coefficients. The original signal $s^j(n)$ can be reconstructed by

$$s^j(n) = \sum_m h(2m-n) w_1^{j-1}(m) + \sum_m g(2m-n) w_0^{j-1}(m). \tag{3}$$

Using the DWPD both outputs from the lowpass filter $(w_1^{j-1}(n))$ and highpass filter $(w_0^{j-1}(n))$ can be further uniformly decomposed into low and high frequency signals [11] as shown in Fig. 1. The output of any subsequent lowpass filter is given by

$$w_{2b}^{j-l-1}(n) = \sum_m h(2n-m) w_b^{j-l}(m). \tag{4}$$

where $b = 0, \ldots, 2^l - 1$ for every resolution level $j-l, l = 1, \ldots, L$. The output of any subsequent highpass filter is given by

$$w_{2b+1}^{j-l-1}(n) = \sum_m g(2n-m) w_b^{j-l}(m). \tag{5}$$

where $b = 0, \ldots, 2^l - 1$ for every resolution level $j-l, l = 1, \ldots, L$.

An example of the DWPD of an exponentially modulated signal is shown in Fig. 2. The signal consists of 128 samples and is decomposed into three resolution levels of wavelet coefficients. The resolution level

$j-1$ consists of $2^{j-(j-1)} = 2$ frequency channels, the resolution level $j-2$ consists of $2^{j-(j-2)} = 4$ frequency channels, and the resolution level $j-3$ consists of $2^{j-(j-3)} = 8$ frequency channels.

3 Best Discrimination Basis Algorithm

Best Discrimination Basis Algorithm (BDBA) is based on a best basis selection technique developed by Coifman and Wickerhauser [2] for signal compression and extended by Saito [10] to signal classification. The goal of the BDBA is to select a complete orthonormal wavelet basis that maximizes a discriminant measure (statistical distance among classes) for a given reference database of target signatures.

One of the common discriminant measures is Kullback-Leibler Distance which is also known as relative entropy. In the case of two classes represented by two normalized energy distributions E_1 and E_2, relative entropy is defined as:

$$I(E_1, E_2) = \sum_{n=1}^{N} E_1(n) \log(E_1(n)/E_2(n)), \tag{6}$$

where $\sum E_1(n) = \sum E_2(n) = 1$ and $n = 1, \cdots, N$. The relative entropy measures how much the distributions of E_1 and E_2 differ. We use the symmetric version of the relative entropy which is a symmetric quantity:

$$D(E_1, E_2) = I(E_1, E_2) + I(E_2, E_1). \tag{7}$$

The important property of the relative entropy measure is that it is additive:

$$D(E_1, E_2) = \sum_{n=1}^{N} D(E_1(n), E_2(n)). \tag{8}$$

This property is used for the multiple-class classification, i.e., for the classification into $j = 1, \cdots, K_c$ classes. In such a case the relative entropy is defined as:

$$D(\{E_j\}_{j=1}^{K_c}) = \sum_{m=1}^{K_c-1} \sum_{n=m+1}^{K_c} D(E_m, E_n). \tag{9}$$

In this work, the reference database is transformed into the wavelet domain using the DWPD as defined in Section 2. This results in a decomposition of each reference signal s_i with respect to a hierarchy of subbases (frequency channels). The normalized energy distribution E_j for each class $c_j \in C$ of targets is determined by a time-frequency energy map [10]

$$E_j = \sum_{i=i_j}^{i_j+N_j-1} W_i^{\,2} / \sum_{j=i_j}^{i_j+N_j-1} \| s_i \|^2 \,. \tag{10}$$

In this equation $i = i_j, \cdots, N_j - 1$, indexes reference signals belonging to the same class $c_j \in C$ in the reference database, and E_j is a matrix whose elements are computed by taking a sum of the squares of wavelet coefficients at each position in the complete wavelet packet decomposition matrix W_i for all reference signals belonging to the same class $c_j \in C$ and by normalizing this sum by the total energy of the reference signals belonging to a given class $c_j \in C$.

The most discriminant subbases $bb(l, b)$, where $l = 0, \ldots, L$ is the wavelet packet decomposition level and $b = 0, \ldots, 2^l - 1$ is the b-th subbasis on a given decomposition level l, are determined by the following recursive procedure [10, 2]:

1. Set $bb(L, b) = w(L, b)$ and $d(L, b) = D(\{E_j\}_{j=1}^{K_c})$ for $b = 0, \cdots, 2^L - 1$, where $w(L, b)$ is the b-th wavelet packet decomposition subbasis b at the coarsest decomposition level L.

2. Determine the most discriminant subbasis $bb(l,b)$, for $l = L-1, \cdots, 0$ and $b = 0, \cdots, 2^l - 1$ using recursively the following test:
   Set $d(l,b) = D(\{E_j\}_{j=1}^{K_c})$.
   If $d(l,b) \geq d(l+1, 2b) + d(l+1, 2b+1)$,
   then $bb(l,b) = w(l,b)$
   else $bb(l,b) = bb(l+1,2b) \oplus bb(l+1,2b+1)$ and $d(l,b) = d(l+1,2b) + d(l+1,2b+1)$.

The most discriminant subbases determined by utilizing the above procedure form a complete orthonormal most discriminant basis. The total number of components (wavelet coefficients) in the complete basis is equal to the number of samples of the measurement signal N. In order to reduce the computational cost, usually only $k_f < N$ features are used for recognition. One of the methods proposed in [10] for selection of k_f features is to use Most Discriminant Wavelet Coefficients (MDWC) determined by using a relative entropy measure. In this work, we propose to select k_f features by using a framework of model theory. We show that features selected in this way result in better recognition accuracy.

4 Model Theory

A natural requirement for each recognition system is that the interpretation of the data be "correct". Intuitively, this means that the targets identified by a recognition system really exist in the world, that these targets have the features as identified by a recognition system, that the relations recognized by a recognition system really exist in the world, and that the whole interpretation does not violate the constraints that the world is known to obey, e.g., the laws of physics. To maintain the correctness of interpretations, an Automatic Feature Based Recognition System (AFBRS) proposed in this work uses the framework of model theory [1, 8].

Model theory is a branch of mathematical logic which deals with the relation between a formal language and its interpretations, or models. The formal language is used to state the truthfulness of the models. This is achieved by having a basic truth definition, which specifies for each pair consisting of a sentence and a model the truth value "true" or "false". This truth definition is the bridge connecting the formal language with its interpretation by means of models [1].

According to model theory, a first-order *formal language* L is a collection of symbols. These symbols consist of relation symbols ($P_i,\ i = 0, \ldots, u$), function symbols ($F_j,\ j = 0, \ldots, v$) , and constant symbols ($c_k,\ k = 0, \ldots, z$). Therefore, the language L is represented as

$$L = \{P_0, \ldots, P_u,\ F_0, \ldots, F_v,\ c_0, \ldots, c_z\}. \tag{11}$$

Each relation symbol P_i is an n-placed relation, where $n \geq 1$. Similarly, each function symbol F_j is an m-placed function, where $m \geq 1$.

A *model* for language L is a pair $< A, I >$, where A is the *universe* (domain) of the model and I is the *interpretation* function. The interpretation function maps the symbols of the language L to appropriate relations, functions and constants in the universe A. Therefore, each n-placed symbol P_i from the language L corresponds to an n-placed relation $R_i \subset A^n$ on A, each m-placed F_j corresponds to an m-placed function $G_j : A^m \rightarrow A$ and each constant symbol c_k corresponds to a constant $x_k \in A$.

In this work, the framework of model theory is used to reduce the number of features actively participating in the recognition process from N (number of components of the complete basis) to $k_f < N$. A formal language L describes the world and the sensing process. A model M represents sensor data, operations on data, and relations among the data. They consist of domains, operators, and relations among the elements of the domains. A theory T represents symbolic knowledge about the world and about the sensors. A theory is *consistent* if it has a model. In other words, a consistent theory is such a theory that all its sentences hold in some model of this theory (under some interpretations of its nonlogical symbols). Such a consistent representational structure helps us to maintain the correctness of data interpretation [8].

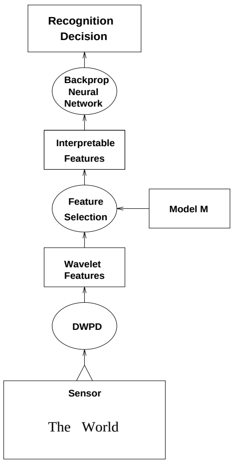

Figure 3: Automatic Feature Based Recognition System (AFBRS)

5 AFBRS Methodology

Fig. 3 shows a conceptual view of the Automatic Feature Based Recognition System (AFBRS).

The meaning of each block in the AFBRS is the following:

1. *The World* is a domain of the recognition problem with *Sensor* providing information about the domain to the AFBRS.
2. *Feature Extraction* extracts features from measurement data using the BDBA and a model theory framework.
3. *Model Checking* tests whether the measurement data represent a model M of a given theory T of the domain.
4. *Reasoning* uses the theory T to arrive with a final recognition decision.

There are three types of information available to the AFBRS:

- the target signatures $s_i \in S$, $i = 1, \cdots, K$, where S is a signal space;
- classes $c_j \in C$, $j = 1, \cdots, K_c$, of targets to be recognized, where C is a target class space;
- symbolic features $b_j \in B$ for each class $c_j \in C$ of targets, where B is a symbolic feature space.

The goal of the AFBRS is to increase the recognition accuracy by using all three types of information in an interpretable way. The *recognition accuracy* is defined as

$$recognition\ accuracy(\%) = \frac{number\ of\ correct\ recognition\ decisions}{total\ number\ of\ recognition\ decisions} * 100 \tag{12}$$

where the *total number of recognition decisions* is equal to the number of target signatures in the reference database. Similarly, we define *misclassification rate* as

$$misclassification\ rate(\%) = \frac{number\ of\ incorrect\ recognition\ decisions}{total\ number\ of\ recognition\ decisions} * 100 = 100 - recognition\ accuracy \tag{13}$$

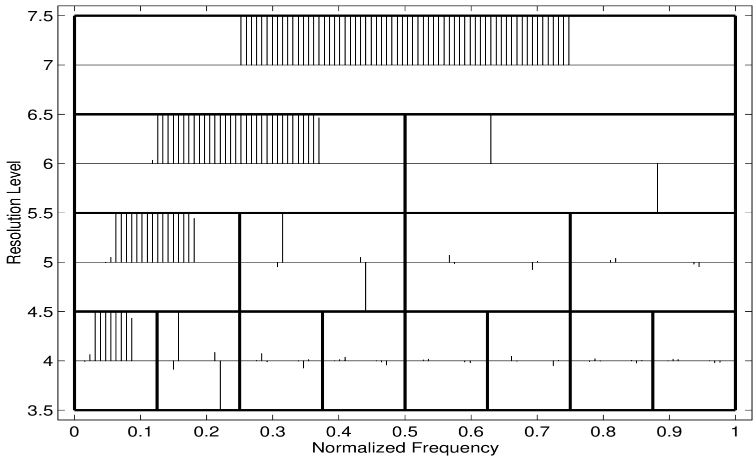

Figure 4: DWPD of a Rectangle Range Signal

In this work, we investigate a method for extracting features $f_i \in F$ of the signal $s_i \in S$ represented in the wavelet domain, such that these features are interpretable in terms of the symbolic features $b_j \in B$ for every target class $c_j \in C$. We define a recognition problem as a mapping $S \rightarrow C$. This mapping is composed of three mappings: (i) $S \rightarrow W$, (ii) $W \rightarrow F$, and (iii) $F \rightarrow C$. The mapping $S \rightarrow W$ is determined by the Discrete Wavelet Packet Decomposition (DWPD). The DWPD transforms a sensor signal $s_i \in S$ into a time/frequency (wavelet) representation $W_i \in W$. The mapping $W \rightarrow F$ is determined by the Best Discrimination Basis Algorithm (BDBA) and a model theory framework. The BDBA selects the most discriminant basis. The model checker selects features from the most discriminant basis. Conceptually, a mapping $F \rightarrow C$ is determined by a domain theory. The domain theory classifies targets based on the extracted features. The problem with a domain theory as a classifier is its sensitivity to noise. To make the ABFRS robust in noisy conditions, we use a backpropagation neural network to implement the mapping $F \rightarrow C$.

6 Feature Selection for Target Recognition

As an example of the target recognition problem, we consider a problem of distinguishing triangles from rectangles using a 1-D range sensor. This scenario is described in more detail in [7]. Each of the triangle/rectangle (target) signatures in the reference database is transformed into the wavelet domain using the DWPD with the Daubechies_6 compactly supported wavelets with extreme phase and highest number of vanishing moments compatible with their support width [3]. Figs. 4 and 5 show examples of the DWPD performed on a range signal corresponding to a rectangle and to a triangle respectively.

A complete orthonormal most discriminant basis is selected using the BDBA. The total number of components (wavelet coefficients) in this basis is equal to 128, i.e., it is equal to the number of samples in the target signature. Fig. 6 shows a complete orthonormal most discriminant basis for a given reference database and the discriminant value of each component obtained by using a relative entropy measure. In order to reduce the computational effort, we apply the framework of model theory to select $k_f < 128$ wavelet coefficients (features) from the complete basis, which are subsequently used to train a backpropagation neural network. After careful analysis of target signatures in the reference database and the most discriminant basis (shown in Fig. 6), we specify a language, a model and a theory for the reference database.

We selected nine features shown in Fig. 7. The first three features correspond to the rising edge of a rectangle, the next three features correspond to the vertex of a triangle, and the last three features correspond to the falling edge of the rectangle. These features are different from the nine features selected as MDWC,

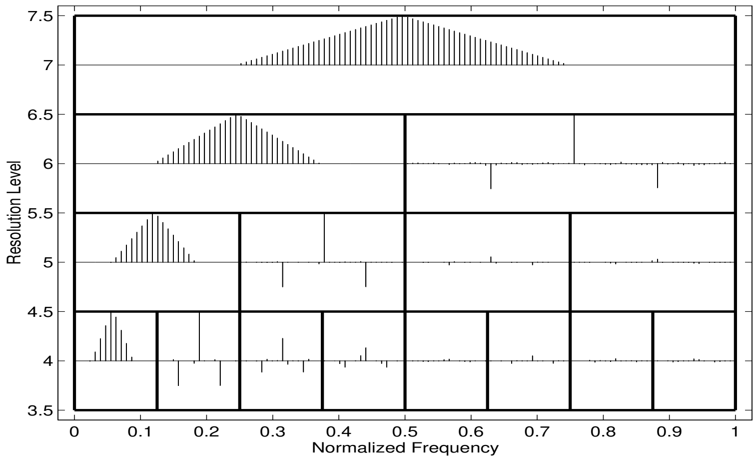

Figure 5: DWPD of a Triangle Range Signal

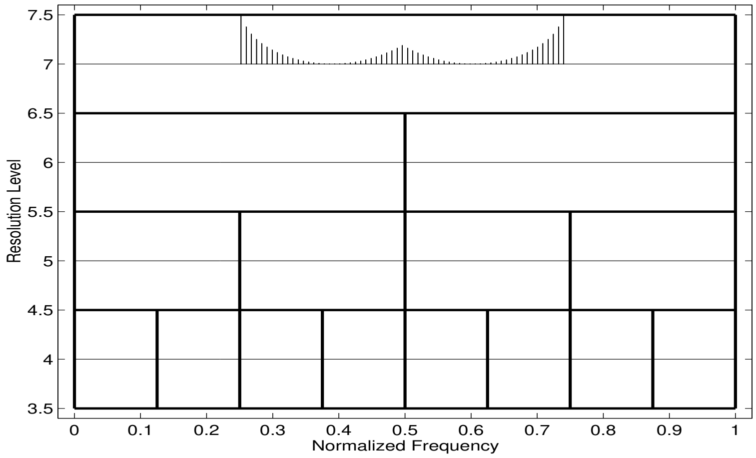

Figure 6: Most Discriminant Basis (MDB)

shown in Fig. 8.

Below, we describe an appropriate language, a model, and a theory for the range sensor.

- **Language** L_r

  A language L_r is the following:

$$L_r = \{\leq, f_r, C_{r_0}, C_{r_1}, \ldots, C_{r_7}, C_{r_8}, rectangle, triangle\}, \tag{14}$$

  where
  $\leq$ - is a 2-placed relation symbol (an order relation)
  f_r - is a 1-placed function symbol (a function that maps features' indices into features' values)
  $C_{r_0}, C_{r_1}, \ldots, C_{r_7}, C_{r_8}$ - are constant symbols (features' indices)
  $rectangle$, $triangle$ - are constant symbols (objects to be recognized)

- **Model** M_r

  A model M_r is defined as

$$M_r =< A, I_r >, \tag{15}$$

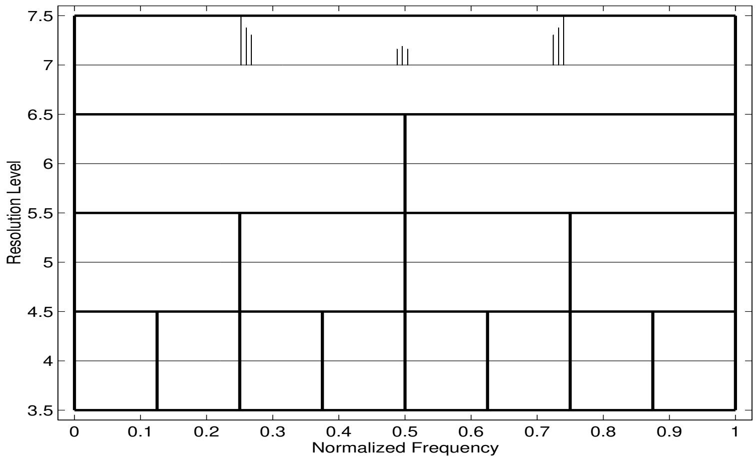

Figure 7: Features Selected Using the Domain Theory

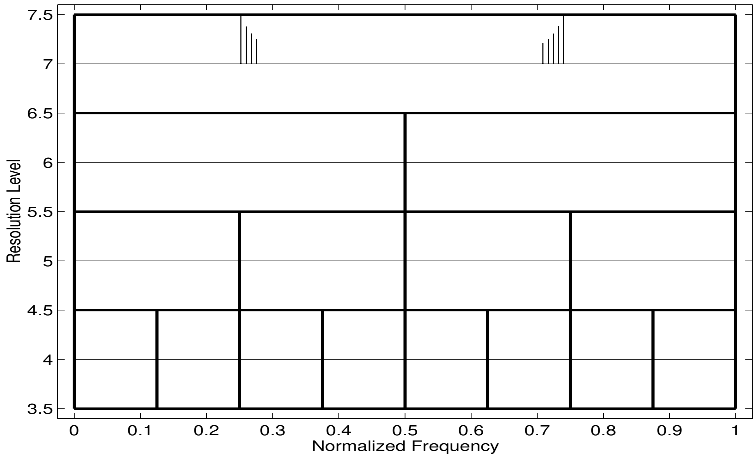

Figure 8: Features Selected as Most Discriminant Wavelet Coefficients (MDWC)

or as

$$M_r = < A, \leq, f_r, 0, 1, \ldots, 8 >, \tag{16}$$

where
A - is a universe of the model M_r, $A = \{0, \ldots, 8\}$
$\leq$ - is an order relation on A
f_r - is a function $A \rightarrow A$
I_r - is an interpretation function that maps symbols of the language L_r to appropriate relations, functions and constants in the universe A. It assigns an order relation on A in the model M_r to a symbol $\leq$ in the language L_r. We use $\leq$ to denote both the order relation symbol in L_r and the order relation in M_r. I_r also assigns a function $A \rightarrow A$ in the model M_r to a symbol f_r in the language L_r. We use f_r to denote both the function symbol in L_r and the function in M_r. And finally, I_r assigns constants $0, \ldots, 8$ in the model M_r to the constant symbols $C_{r_0}, \ldots, C_{r_8}$ in the language L_r.

- **Theory** T_r
  A theory T_r is the following:

$$rectangle \equiv \forall_{x,y \in \{C_{r_0}, \ldots, C_{r_8}\}} \; f_r(x) = f_r(y) \tag{17}$$

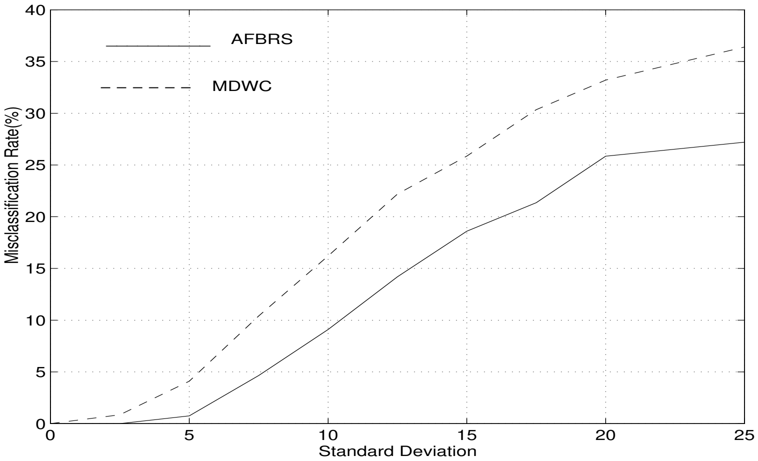

Figure 9: Misclassification Rates Using Model Theory (AFBRS) and Most Discriminant Wavelet Coefficients (MDWC) for Recognition

$$triangle \equiv \forall_{x \in \{C_{r_0}, \ldots, C_{r_8}\}} \; f_r(x) \leq f_r(C_{r_4}) \tag{18}$$

$$C_{r_0} \leq C_{r_1} \leq C_{r_2} \leq C_{r_3} \leq C_{r_4} \leq C_{r_5} \leq C_{r_6} \leq C_{r_7} \leq C_{r_8} \tag{19}$$

7 Results of Experiments

We test the recognition accuracy of the AFBRS using a test database of triangle/rectangle (target) signatures. Each of the target signatures in the testing database is transformed into the wavelet domain using the DWPD with the Daubechies_6 compactly supported wavelets with extreme phase and highest number of vanishing moments compatible with their support width [3].

Fig. 9 shows the resulting misclassification rate for different levels of noise. This figure also shows the misclassification rate of a recognition system using the Most Discriminant Wavelet Coefficients (MDWC) as features.

8 Conclusions and Future Research

A recognition methodology (ABFRS) based on features selected by utilizing the Best Discrimination Basis Algorithm (BDBA) and a symbolic knowledge about the domain is proposed. The symbolic knowledge about the domain is incorporated into the ABFRS through a domain theory using a framework of model theory. A model of a domain theory represents a bridge between a symbolic knowledge about the domain and measurement data.

The simulation results show that the AFBRS has a better recognition accuracy than the MDWC based recognition system with respect to the test database. The MDWC based recognition system adapts well to a set of features characteristic to the reference database, but it does not have enough generalization power to perform equally well with the test database. Features used to train a classifier are concentrated in two time/frequency areas. We are able to increase generalization power of the recognition system by using features which are more spread across the time/frequency domain. To select such features, we utilize the knowledge of symbolic features specific to a given class.

In this work, we build the domain theory manually based on the analysis of the target signatures in the reference database and the most discriminant basis. In the future research we are going to incorporate learning capabilities into the ABFRS. We will also extend the ABFRS to perform recognition using information from multiple sensors.

9 Acknowledgments

This research is partially sponsored by the Advanced Research Projects Agency under Grant F49620-93-1-0490.

References

[1] C. C. Chang and H. J. Keisler. *Model Theory.* Amsterdam, New York, Oxford, Tokyo: North Holland, 1992.

[2] R. R. Coifman and M. V. Wickerhauser, "Entropy-Based Algorithms for Best Basis Selection," *IEEE Transactions on Information Theory,* vol.38, no.2, pp. 713–718, March 1992.

[3] I. Daubechies, *Ten Lectures on Wavelets.* Philadelphia, Pennsylvania: Society for Industrial and Applied Mathematics, 1992.

[4] I. Daubechies, "Orthonormal Bases of Compact Supported Wavelets," *Communications on Pure and Applied Mathematics,* vol. 91, pp. 909–996, 1988.

[5] L. Hong, "Multiresolutional Filtering Using Wavelet Transform," *IEEE Transactions on Aerospace and Electronic Systems,* vol. 29, no. 4, pp. 1244–1251, 1993.

[6] Z. Korona and M.M. Kokar, "Multiresolution Multisensor Target Identification," in *Proc. IEEE Int'l Conf. Multisensor Fusion and Integration for Intelligent Systems (MFI'94),* pp. 501–506, Las Vegas, Nevada, Oct. 1994.

[7] Z. Korona, "Model Theory Based Feature Selection for Multisensor Recognition," *A Dissertation,* Northeastern University, 1996.

[8] M.M. Kokar and J. Tomasik, *Towards a Formal Theory of Sensor/Data Fusion.* Progress Report, Northeastern University, May 1994.

[9] A.D. Kulkarni, *Artificial Neural Networks for Image Understanding.* New York: Van Nostrand Reinhold, 1994.

[10] N. Saito, "Local Feature Extraction and Its Applications Using a Library of Bases," *A Dissertation,* Yale University, Dec. 1994.

[11] M. V. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software.* Wellesley, Massachusetts: A.K. Peters, 1994.