Lung Sound Recognition Using Model-Theory Based Feature Selection and Fusion

Zbigniew Korona and Mieczyslaw M. Kokar
Department of Electrical and Computer Engineering
Northeastern University
360 Huntington Avenue
Boston, MA 02115

Abstract

In this paper we describe the application of a new automatic signal recognition methodology to the recognition of lung sounds. Two main features of this methodology are: the use of symbolic knowledge about the signal sources (in addition to sensory inputs), and fusion of multisensor information. We show steps involved in the design of an automatic multisensor recognition algorithm that extracts symbolic features and utilizes symbolic knowledge for recognition. We also compare the performance of the resulting algorithm with both a single-sensor system and with a system that selects features using an entropy-based criterion. To evaluate the methodology we used both normal lung sounds (*bronchial* and *vesicular*) as well as adventitious sounds (*rhonchi*, *wheezes* and *crackles*). Our experiments show that the recognition accuracy can be improved through the use of symbolic knowledge about the signals, and that our methodology is feasible for this type of application.

Keywords: lung sounds, recognition, sensor fusion, formal methods, wavelets

1 Introduction

A stethoscope is one of the simplest basic diagnostic tools of a physician for assessing the health of a respiratory system [40]. Although not everyone agrees as to what extent a stethoscope should be used and relied upon, the sounds radiated by diseased lungs are regarded by most clinicians as information of important diagnostic value. While even the most accurate lung sound analysis and recognition system alone is not sufficient for a successful diagnosis of lung diseases, a stethoscope can be a major component of a diagnostic system.

Although a stethoscope is convenient to use, it is subject to interference from external noise. The proper diagnosis also requires significant training and experience of the medical personnel. Thus, the stethoscope may be unreliable in noisy environments such as an ambulance, a busy emergency room, or a medical assistance airplane [32].

Diagnosis of lung diseases based on lung sounds is a complex task. One of the reasons is that many particular lung sounds (e.g. crackles, wheezes) occur in many types of diseases and some occur in both healthy and diseased lungs (e.g. bronchial). The other reason is that many features of lung sounds depend on various characteristics of the patient, such as weight, age, sex and physical condition, as well as on the environment, such as humidity, temperature, and time of day. The difficulty of correct diagnosis increases with the increase in the level of noise. Therefore, it is important to develop an automatic diagnostic system which would assist the physician in decision making. Such a system would practically emulate a diagnostician who is relieved from stress, tiredness, and other harsh conditions.

The automatic signal recognition process usually consists of acquiring signals originating from a source through one or more sensors, and then classifying these signals based upon some prior knowledge. The prior knowledge may include probabilities of occurrence of particular classes of signals, features of signals, or relations among the features. The classification decision rule can be implemented as a statistical classifier, a decision tree, a neural network, or using some other classification approach.

It is not feasible to perform signal recognition based upon raw sensor signals. Instead, an abstraction of a sensor signal must be created first, and then classification can be performed on the abstracted features. Some of the main questions of this abstraction process are: what are the "right" features, how to extract these features in a computationally efficient way, and how to construct a recognition rule based upon these features?

Robustness of an automatic recognition system can be improved by using more sensors that provide complementary or redundant information [22, 38]. In such a case, information from multiple sensors needs to be combined into one representation; this process is called *sensor data fusion*. In recent years, sensor data fusion has been a very active research field (cf. [2, 16, 17, 25, 37].) In spite of many attempts reported in the literature, there is still a lack of a unified methodology for developing fusion systems [24]. Our approach to fusion is presented in Sections 5.2 and 6.4.

Both single and multisensor recognition algorithms could perform better if they were able to incorporate more prior knowledge, and use it to select the right features. In most cases, the prior knowledge is either in the form of probabilities or heuristics. In many cases, symbolic knowledge about the sources and about the scenarios exists; this knowledge could be incorporated into the recognition process and then used to *reason* about sources, their features and scenarios. To achieve this goal, *symbolic features* need to be extracted from signals, and symbolic knowledge needs to be represented in a structure that guarantees *sound* reasoning. *Theorem provers* (cf. [39, 35, 34]) guarantee sound reasoning; they could be used for recognition, at least in the design of recognition systems. Unfortunately, while there are many algorithms for extracting quantitative features, there is a lack of a methodology for extracting symbolic features, as well as a lack of libraries of symbolic knowledge that could be used in such a reasoning process.

In this paper we present a methodology for model-based feature selection and fusion for multisensor recognition, and show the results of the application of this methodology to lung sound recognition. First, we describe our design of a lung sound recognition system. Then we compare the resulting system with both a single-sensor system and with a system that selects features using an entropy-based feature selection method described in [31].

We begin with a brief overview of the lung sound recognition problem (Section 2). In Section 3 we describe the scenario for our study. Then, in Section 4, we give an outline of the entropy-based feature selection method [31]. In Section 5 we briefly describe our methodology and an Automatic Multisensor Feature-based Recognition System (AMFRS). In Section 6 we describe the steps of our methodology as applied to the design of an AMFRS for this specific scenario. In Section 7 we show the performance results of our system and compare them to a single-sensor system and to a multisensor system that uses an entropy-based measure for feature selection. Finally, in Section 8 we present our conclusions and directions for future research.

2 Lung Sound Recognition: Background

2.1 Lung Sounds

In this section we show examples of lung sounds and common lung diseases in which these sounds are present.

Lung sounds, that can be heard through a stethoscope, can be classified into two groups [6]: normal breathing sounds that occur when no respiratory problems exist and adventitious (abnormal) sounds when a problem exists. The normal breath sounds are both inspiratory and expiratory. They occur when the air moves in and out of the chest during a regular breathing cycle. The normal sounds are classified according to the character of the sound and the location where this sound is heard. The amplitude and frequency of these sounds can vary from location to location and from person to person.

Disease	Description	Lung sounds
Atelectasis	Collapse of lung tissue. Caused by shallow breathing (avoiding pain).	bronchial, fine crackles
Pneumonia	Inflammation of the air cells. Sounds may be detected two days before X-ray detection.	fine crackles
Upper airway obstruction	Obstruction of larynx or trachea by foreign body, or diseased upper airways.	stridor
Chronic bronchitis	Inflammation of the bronchial tubes.	rhonchi, coarse crackles
Bronchial asthma	Spasm and constriction of bronchial passages.	wheezes

Table 1: Examples of diseases and corresponding lung sounds

There are two important types of normal respiratory sounds: *bronchial* and *vesicular*.

Abnormal sounds are those sounds that occur unexpectedly during a regular breathing cycle. They include normal sounds, when they appear at atypical locations, and adventitious sounds. For example, sounds of a bronchial or bronchovesicular nature are considered to be abnormal if they appear in locations where vesicular sounds should be heard. Considerable variability in the character of lung sounds among different individuals makes it sometimes difficult to be sure of the "abnormality" of a sound. There are two major types of adventitious sounds: continuous and discontinuous. Among continuous adventitious sounds, the most well known are *rhonchi*, *wheezes* and *stridor*. Wheezes are continuous sounds, longer than 250msec, high pitched, with dominant frequency of 400Hz or more (a hissing sound). Rhonchi are also longer than 250msec in duration, low pitched, with dominant frequency of 200Hz or less (a snoring sound). *Stridor* is similar to wheezes. Both are high-pitched, but stridor can often be distinguished by its tendency to be accentuated in inspiration.

The most common discontinuous adventitious sounds are *crackles* and *squawks*. Crackles are discrete, nonmusical sounds, usually shorter than 10msec. They are a sequence of short interrupted sounds with a wide spectrum of frequencies, between 200 and 2000 Hz. The characteristics, number and timing of crackles vary in different diseases [21]. *Fine crackles* are characterized by the initial deflection width of 0.9ms and two cycle duration of 6ms, while these characteristics for *coarse crackles* are 1.25ms and 9.5ms, respectively. *Squawks* are a combination of wheezes and crackles; they may start with fine crackles and then sound like short inspiratory wheezes.

The sounds of diseased lungs are regarded by most of clinicians as having a relatively high diagnostic value. In Table 1 we list a number of examples [26, 6] of acute and chronic lung diseases and associated lung sounds.

2.2 Lung Sound Analysis and Recognition Methods

Over past fifteen years, much research has been carried on microcomputer-based respiratory and cardiac sound analysis systems. However, while cardiac sound analysis systems are already commercially available and used in clinical practice to detect such anomalies as malfunctioning heart valves [10], similar respiratory sound analysis systems are not widely used. Lung sounds are highly nonstationary stochastic signals due to changing air flow rate and lung volumes during a respiration cycle. This makes the analysis of lung sounds difficult.

There have been many different attempts to analyze lung sounds in the frequency domain and fewer to analyze them in the time domain. Frequency domain analysis has been usually performed using the fast Fourier transform (FFT). However, no standard FFT routine has been applied. The FFT's used have differed in sample frequency, the number of FFT points and the window function [3, 27, 28, 30]. The simple spectrum-describing features have been derived from the FFT analysis [12, 28]. Examples of such features are: the frequency of maximum power, the mean frequency of power spectrum, and the maximum frequency of the signal. These features were tested using simple statistical tests such as paired t-tests [27]. There were few attempts to use FFT components directly as features. The principal component analysis was used to deal with a huge number of such features [3, 36]. Much research was also devoted to the autoregressive (AR) modeling of lung sounds [33]. The first such attempt is described in [6], where power spectral density is estimated by means of linear prediction.

Although most of the lung sound recognition research was performed in the frequency domain, some work was also done in the time domain. In [4] asthmatic wheezing was categorized by the time spent wheezing divided by respiratory cycle duration. In [18] time-domain analysis of tracheal sounds has been applied to apnoea detection.

The standard frequency analysis, such as the FFT or autoregression spectral estimation techniques, do not provide insight into the timing of the appearance and disappearance of various frequency components [11]. For example, fast transient signals like crackles, occurring either at fixed or random times in the breathing cycle, are not easily detected in the frequency domain. On the other hand, by analyzing signals in the time domain, important frequency information is lost. Therefore, there is a need for time-frequency tools which enable the analysis of lung signals both in the time and frequency domains.

In the last several years, the wavelet transform started to be used as a tool for time-frequency analysis and recognition of lung sounds [1, 13, 15]. The transient and random nature of lung sounds [29] make them especially suitable for the wavelet processing. Decomposition of a nonstationary lung signal into several frequency channels allows to perform signal analysis on separate resolution levels. By doing this the amount of nonstationarity is reduced. In [1], the Malvar's wavelet transform is used to perform a segmentation of infant's lung sounds. In [13], a wavelet-based crackle detector is proposed. This detector applies nonlinear operators to the frequency channels carrying the most information in order to increase the signal-to-noise ratio (SNR). Then, AR parameters are extracted from each frequency channel and used to make a recognition decision. The AR modeling approach assumes that a system formed by lungs and the chest wall can be modeled as an all-pole filter. The limitations of the AR modeling are discussed in [20, 14]. In [14], higher-order statistics was used for autoregressive modeling of lung sounds. The main advantage of the higher-order statistics component was robustness to symmetric noise. The use of the wavelet transform to the separation of discontinuous adventitious lung sounds from vesicular sounds was reported in [15]. The main thrust in [15] was to distinguish "explosive peaks" associated with crackles and squawks from noise, that can be associated with any signal.

While in this paper we also use the wavelet transform, our approach is significantly different from all the approaches discussed above. The main difference is that we assume that we have some symbolic description of particular types of lung sounds. Consequently, our approach is both data and knowledge driven, while all the approaches discussed above are essentially data driven.

3 Lung Sound Recognition Scenario

In our lung sound recognition experiments, we used a tape recording of various lung sounds, from various patients, prepared and described by L. H. Murphy from Pulmonary Service Lemuel Shattuck and Faulkner Hospitals [26]. The characteristics of the patients, such as sex, age, weight, or any special conditions, were not known.

Since the goal was to evaluate our approach to the recognition of lung sounds, we created two types of data: one collection of sounds to serve as a reference data set (examples, or *signatures*) and another collection to serve as a test data set. Additionally, since one of our goals was to show an improvement in the robustness of recognition due to fusion, we generated two data sets of signatures. The main idea of fusion is to combine information from multiple sensors. Fusion is needed even when the same sensor is used multiple times (e.g., in auscultation, the same stethoscope is used to acquire lung sounds at different spots of the patient's chest). In our scenario, we simulated such a situation by creating two data sets of signatures, DL_r and DL_i, by picking different breathing cycles for the two data sets. We refer to these data sets as *reference databases* for $Sensor_r$ and $Sensor_i$.

DL_r and DL_i contained four types of lung sounds, one normal sound, *bronchial*, and three adventitious sounds: *fine crackles*, *rhonchi* and *wheezes*. Each signature covered one breathing cycle. Figure 1 shows five examples of lung sound signatures for each of the four classes of lung sounds in DL_r.

Additionally, we created two *test database*, one for $Sensor_r$ and one for $Sensor_i$. They were generated by first picking specific signals (breathing cycles) from the tape and then adding noise to the signals. For each class of signals we selected five signals (different than the signals selected for the reference databases) and then added zero-mean Gaussian

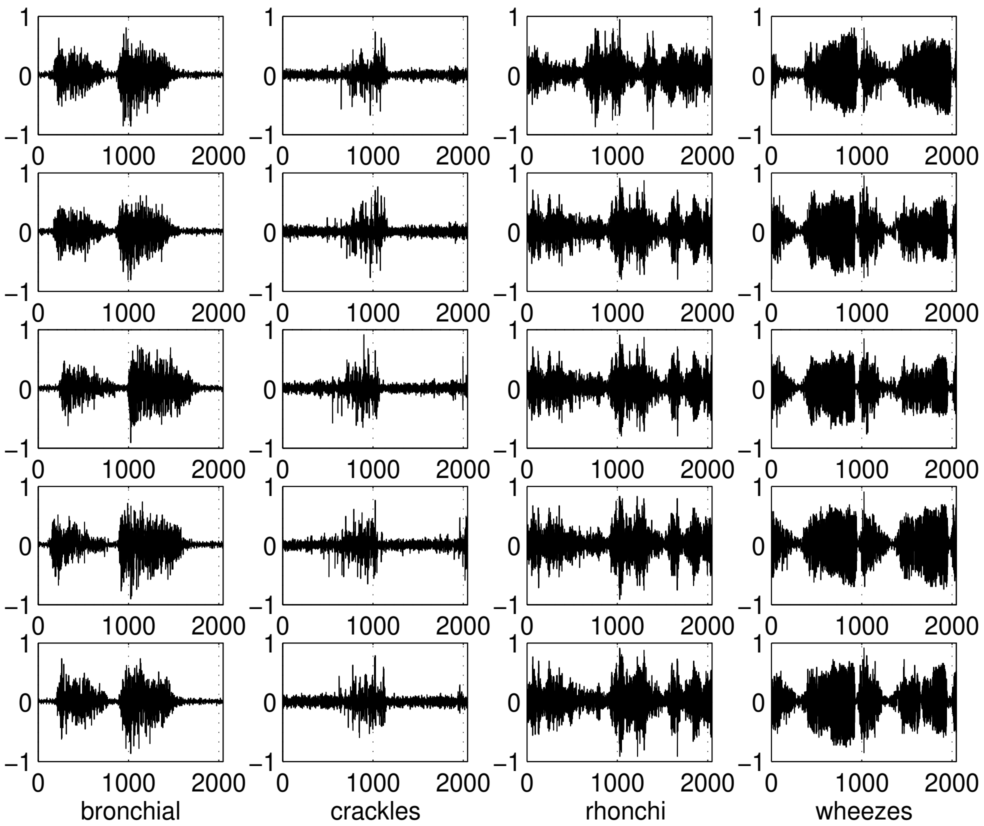

Figure 1: Five Signatures for Each of the Four Classes of Lung Sounds in DL_r (normalized amplitude vs. time [msec])

noise on eleven levels of standard deviation (see Section 7).

4 Entropy-Based Feature Selection

In this section we describe entropy-based feature selection. We use this method to select an initial set of features.

A signal $s(n)$ can be represented on various decomposition levels $j = 1, 2, \ldots$, as described below. For a given signal $s^j(n)$, at the decomposition level j, a lower resolution low frequency signal $s^{j+1}(n)$ and a lower resolution high frequency signal $w^{j+1}(n)$ are derived by using a lowpass filter with an impulse response $h(n)$, and a highpass filter with an impulse response $g(n)$, respectively, and subsampling the output of the low pass filter by two:

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

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

If the filters $h(n)$ and $g(n)$ meet the regularity and the orthonormality constraints [8], the original signal $s^j(n)$ can be reconstructed from the two lower resolution signals $s^{j+1}(n)$ and $w^{j+1}(n)$. In such a case, Eqs. 1 and 2 can be considered as a decomposition of the signal onto an orthonormal basis $\{g_m, h_m\}$, $m \in \mathbf{Z}$, where particular elements of the sequences $s^{j+1}(n)$ and $w^{j+1}(n)$ are coefficients associated with particular basis functions. The coefficients $w^j(n)$ are called *wavelet coefficients*. In the Discrete Wavelet Packet Decomposition (DWPD) scheme [7, 19], outputs of both filters are further decomposed into low and high frequency signals. Consider signals $s(n) \in R^N$, where $N = 2^J$, $J \in \mathbf{Z}$. Such signals can be decomposed into decomposition levels $j = 0, 1, \ldots, J$. Denote the original signal as $w^0_{b=0}(n) \in R^N$, where b indexes *frequency bands*. The discrete wavelet packet decomposition then can be represented by the following two equations:

$$w^{j+1}_{2b}(n) = \sum_m h(2n - m) w^j_b(m), \tag{3}$$

$$w^{j+1}_{2b+1}(n) = \sum_m g(2n - m) w^j_b(m), \tag{4}$$

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

Discrete Wavelet Packet Decomposition (DWPD) can be viewed as recursive decomposition of a vector space into two mutually orthogonal subspaces. Therefore, the original signal $s(n)$ is represented by the coefficients in the two subbases at level 1, in four subbases at level 2, and so on. The two subbases for the frequency bands $2b$ and $2b+1$ at level $j+1$ represent the same vector subspace as one base in the frequency band b at level j. Therefore the related bases at consecutive levels are not independent. In this way, the DWPD scheme generates a number of dependent subbases out of which a base for the whole space can be chosen in many different ways.

Best Discriminant Basis Algorithm (BDBA) [7] selects a complete orthonormal basis that is "best" for representing a signal with respect to some "information cost" measure, e.g., the *Shannon entropy* of the coordinate vector. To find a basis in which a class of signatures is best discriminated from all other classes of signatures in terms of "statistical distance" among classes, a measure based on the *relative entropy* can be used [31]. For any two nonnegative sequences s_1 and s_2 with $\sum s_1(n) = \sum s_2(n) = 1$, such a measure is represented as:

$$I(s_1, s_2) = \sum_{n=1}^{N} s_1(n) \log(s_1(n)/s_2(n)), \tag{5}$$

where N is the number of elements in each sequence. The relative entropy measures how much the distributions of s_1 and s_2 differ. For two equal sequences the relative entropy is zero, for different sequences it is greater than zero. The symmetric version of the relative entropy is defined as

$$D(s_1, s_2) = I(s_1, s_2) + I(s_2, s_1). \tag{6}$$

In the case of multiple sequences $j = 1, \cdots, K$, the relative entropy is defined as the sum of all pairwise combinations of sequences

$$D(s_1, \ldots, s_K) = \sum_{m=1}^{K-1} \sum_{n=m+1}^{K} D(s_m, s_n). \tag{7}$$

In order to measure the discriminant power of a signature database, first a *time-frequency energy map* $E_k(j,b,n)$ is computed for each class k of signatures in the database [31],

$$E_k(j,b,n) = \frac{\sum_{s=1}^{N_k} \left(W_s(j,b,n)\right)^2}{\sum_{s=1}^{N_k} \parallel W_s(.,.,.) \parallel^2}, \tag{8}$$

$$\parallel W_s(.,.,.) \parallel = \sqrt{\sum_j \sum_b \sum_n \left(W_s(j,b,n)\right)^2} \tag{9}$$

where $W_s(j,b,n)$ is a complete wavelet packet decomposition matrix of wavelet coefficients of a signature s, j is the decomposition level, b is the frequency band, and n is the sample number. The subsets (sequences) $E_k(j,b,.)$ of this matrix are used for computing the values of the relative entropy measure $D(s_1,\ldots,s_K)$.

The BDBA algorithm [31, 7] starts with the highest level of decomposition, i.e., with the leaves of the tree. It then prunes the tree by replacing two frequency bands $2b$ and $2b+1$ at the level $j+1$ with one frequency band b at level j, whenever this substitution gives more discriminant power to the representation, as measured by the discriminant measure D, i.e., whenever

$$\begin{aligned} & D(E_1(j,b,.),\ldots,E_{N_k}(j,b,.)) \geq \\ D(E_1(j+1,2b,.),\ldots,E_{N_k}(j+1,2b,.)) &+ D(E_1(j+1,2b+1,.),\ldots,E_{N_k}(j+1,2b+1,.)). \end{aligned} \tag{10}$$

The result of this algorithm is a complete orthonormal basis consisting of selected subbases (frequency bands) for each decomposition level j. The pairs $((j,b,n), W_s(j,b,n))$ constitute *features* (where the second elements are coefficients of the signal in the best basis and first elements are their locations in the DWPD). The total number of component features is equal to the number N of samples of the signal. In order to reduce the computational cost, only $k_f < N$ features are used for recognition. In our approach, k_f is the number of constants in a source theory (see Section 6.2). In the entropy-based approach (cf. [31]), the k_f elements of the most discriminant basis that maximize the value of the relative entropy measure are used; they are called Most Discriminant Wavelet Coefficients (MDWC).

In this paper we compare the performance of an algorithm that uses MDWC as features versus our algorithm, which selects features out of the complete best basis using symbolic information (source theories). Our algorithm is described in the following section.

5 Model-Theory Based Feature Selection and Fusion

5.1 System Overview

Our system, called Automatic Multisensor Feature-based Recognition System (AMFRS), consists of four main processing blocks: DWPD, *Feature Selection*, *Feature Fusion* and *Classification*. The DWPD transforms signals into the wavelet domain according to the algorithm described in Section 4. The output of this algorithm is a set (matrix) of wavelet coefficients associated with particular locations in the input signal. *Feature Selection* picks k_f of the wavelet coefficients, which are then used as *symbolic features*. In this process the *interpretation function* (see Section 6.2) associates particular constants in source theories with these selected features. *Feature Fusion* combines features from both sensors into one set of fused features; they become elements of a fused model. Here again, the interpretation function associates constants of the fused theory with the particular features from both sensors. The fused theory and model are constructed during the design phase of the AMFRS (see Section 5.2). Finally, *Classification* (a backpropagation neural network) implements *soft model checking*, i.e., checking whether the model relations among the features, as defined by the corresponding theories (see Section 6.2), are preserved. Specifically, this block decides whether the features extracted from the signal fulfill one of the relations defined in the source theory.

5.2 System Design

The design process of an AMFRS consists of four steps: DWPD/BDBA, *Model Construction*, *Model Fusion* and *System Implementation*. The inputs to the design process are two reference databases of signatures, DL_r and DL_i, described in Section 3, and two source theories. First, the DWPD is used to transform the database signatures from the time domain into the time/frequency (wavelet) domain. Then the Best Discriminant Basis Algorithm (BDBA) is applied to select the *most discriminant basis* (features) for each reference database. In the *Model Construction* step, the designer constructs *models* for the given source theories (see Section 6.2). In this step, each constant, function and relation symbol of the theories are assigned values in the domain; this assignment constitutes an *interpretation function*. Constants are assigned locations in the DWPD. Function and relation symbols are assigned procedural interpretation (how to find the value of a function and how to check relations among the objects of the domain). In our case we had only one function for each sensor (f_r and f_i); they were assigned the functions W_{sr} and W_{si}, respectively, i.e., the functions that compute wavelet coefficients (see Section 4) for particular sensors. Pairs of interpretations of constants and the values of the function on them are interpretations of *symbolic features*.

The *Model Fusion* operation combines two source theories and their models into one fused theory and its model [23]. In this operation, constants, functions and relations from the two theories are combined into one theory. And so are their interpretations. Fusion is a *formal system operator* that has multiple models and theories as inputs and a single theory and its model as output [23]. This process involves three operators: *expansion, reduction* and *union*, as described below [5]. Since there is no known theory that embeds these two theories (no physical law is known for fusing data from these two sensors), the designer, using the operators described below, constructs a fusion operator based on symbolic knowledge about the domain and the relation between the symbolic knowledge and the interpretable features.

Reduction Operator: A language L^r is a reduction of the language L if the language L can be written as

$$L = L^r \cup X, \tag{11}$$

where X is the set of symbols not included in L^r. A theory (subtheory) T^r for the language L^r is formed as a reduction of the theory T for the language L by removing formulas from the theory T which are not legal sentences of the language L^r (i.e., those sentences that contain symbols of X). A model M^r for the language L^r is formed as a reduction of the model $M =< A, I >$ for the language L by restricting the interpretation function $I = I^r \cup I_x$ on $L = L^r \cup X$ to I^r

$$M^r =< A, I^r > . \tag{12}$$

The reduction operator preserves the theorems of the original formal system, provided that they are not reduced by the operator.

Expansion Operator: A language L^e is an expansion of the language L if the language L can be written as

$$L^e = L \cup X, \tag{13}$$

where X is the set of symbols not included in L. A theory T^e for the language L^e is formed as an expansion of the theory T for the language L by adding a set of new axioms of the language L^e to the theory T. A model M^e for the language $L^e = L \cup X$ is formed as an expansion of the model $M =< A, I >$ for the language L by giving appropriate interpretation I_x to symbols in X

$$M^e =< A, I \cup I_x > . \tag{14}$$

The expansion operator preserves all the theorems of the original theory in the expanded formal system. The expansion operator is not unique. In our application we used a special form of this operator in which we utilize some special properties of our recognition problem. Since our intent was to replace some of the constants, functions and relations with new ones, we used the expansion operator to introduce new symbols into the original language. These new symbols were interpreted using the interpretation of the original symbols, and then the original symbols were removed by the (following this step) reduction operator. We used the following two operations for deriving interpretations for new symbols.

1. (Relation restriction) Given an n-placed relation $R \subset A^n$ in the model M, we can expand this model with a n^e-placed ($n^e < n$) relation $R^e \subset (A')^{n^e}$, where $A' \subset A$. R^e is called a *restriction* of R and is denoted as $R^e = R\mid_{(A')^{n^e}}$. This operation is a combination of projecting the relation R onto selected axes A^{n_e} and, at the same time, restricting its domain to the subset $A' \subset A$.

2. (Product of relations) Given a n_1-placed relation $R_1(x_1, \cdots, x_{n_1})$ and a n_2-placed relation $R_2(y_1, \cdots, y_{n_2})$ in the model M, we can expand this model with a new n-placed relation $R^e(z_1, \cdots, z_n)$, where $n = n_1 + n_2$, derived as a Cartesian product of the relations $R_1(x_1, \cdots, x_{n_1})$ and $R_2(y_1, \cdots, y_{n_2})$. Hence, $R^e(z_1, \cdots, z_n) = R^e(x_1, \cdots, x_{n_1}, y_1, \cdots, y_{n_2}) = R_1(x_1, \cdots, x_{n_1}) \times R_2(y_1, \cdots, y_{n_2})$.

In the same manner, we construct a new function in the expanded model M^e using the following two procedures:

1. (Function domain restriction) Given a function $f: A \rightarrow A$ in the model M, we can expand this model with a function $f^e: A' \rightarrow A$, $(A' \subset A)$, where $f^e = f\mid_{A'}$ is a function whose domain has been restricted from A to $A' \subset A$. The function $f^e: A' \rightarrow A$ is called a *restriction* of f;

2. (Union of functions) Given a function $f_1: A' \rightarrow A$, $(A' \subset A)$, and a (complementary) function $f_2: (A \backslash A') \rightarrow A$ in the model M, we can expand this model with a new function $f^e: A \rightarrow A$ derived as the union of the functions f_1 and f_2. Therefore, $f^e = f_1 \cup f_2$.

Union Operator: This operator generates a language L as a union of the languages L_1 and L_2

$$L = L_1 \cup L_2 \tag{15}$$

and a theory T for the language L, as the union of the theory T_1 for the language L_1 and the theory T_2 for the language L_2

$$T = T_1 \cup T_2. \tag{16}$$

To define a union of two models we expand our notation by including explicitly constants, relations and functions. The union $M =< A; R, f, X; I >$ of two models $M_1 =< A_1; R_1, f_1, X_1; I_1 >$ and $M_2 =< A_2; R_2, f_2, x_2; I_2 >$ is defined as

$$M = M_1 \cup M_2 =< A_1 \cup A_2; R_1 \cup R_2, f_1 \cup f_2, X_1 \cup X_2; I_1 \cup I_2 > . \tag{17}$$

This operator does not guarantee that the resulting structure is a model of the union of two theories; this property is a proof obligation and needs to be checked with each specific case of the application of this operator.

The three models developed in this design process are then used to build the AMFRS in the *System Implementation* step. In this step, among others, a neural-network based classifier is built using the knowledge of the fused model, and trained using the reference databases.

6 AMFRS Design for Lung Sound Recognition

6.1 Entropy-Based Features

The first step in our feature selection methodology is to transform each of the lung sound signatures in the DL_r and DL_i reference databases into the wavelet domain by utilizing the DWPD. For this, we use the Daubechies_6 compactly supported wavelets with extreme phase and highest number of vanishing moments compatible with their support width [9]. Figure 2 shows the result of the DWPD for four signatures corresponding to four classes of lung sounds. This figure shows wavelet coefficients for four decomposition levels. For each decomposition level, a number of frequency bands are shown: one band for level zero (top), two for one, four for two and eight bands for three (bottom).

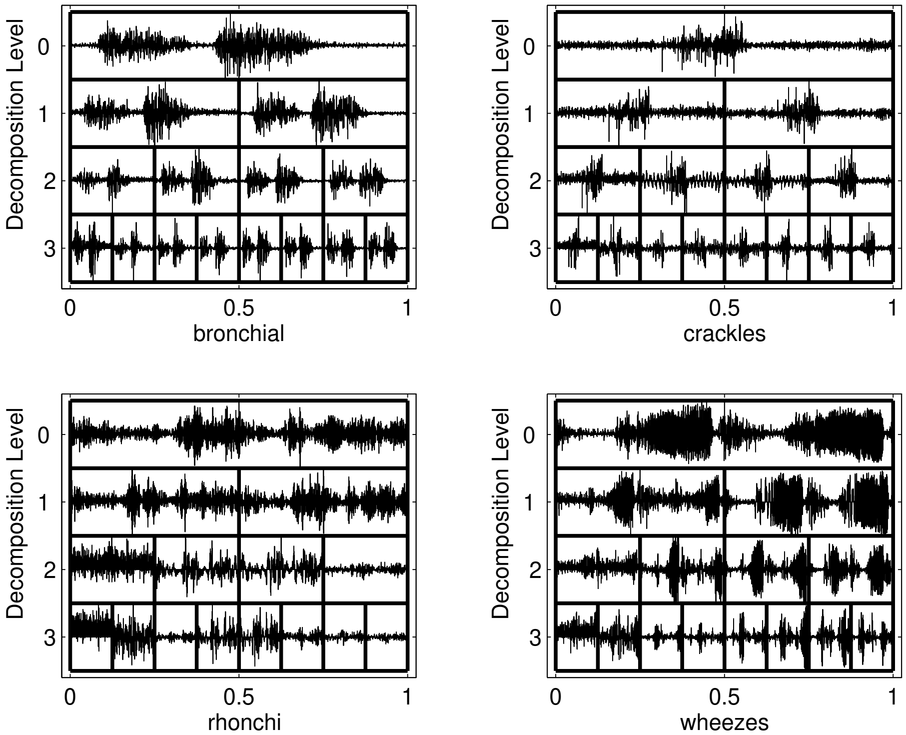

Figure 2: The DWPD of Four Lung Sound Signatures in the DL_r Reference Database

The best known wavelets are the Haar wavelets. The Haar wavelets graphs are made from functions that are piecewise constant with values of either 0 or 1. The approximation of most signals with Haar wavelets is typically worse than with some more complex wavelets; it requires more decomposition levels since many pieces are needed to represent any sloping line. For that reason, we used Daubechies_6 wavelets represented by six numbers. Daubechies wavelets represented by eight or more numbers could work even better but they would require more computational effort.

In the next step, a complete orthonormal Most Discriminant Basis (MDB) 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 signature. Figure 3 shows the discriminant value (as determined by the relative entropy measure) of each component in the MDB for the DL_r reference database. In this case the MDB consists of eight subbases on the third level of the DWPD decomposition. Notice that many elements in the MDB have the discriminant value equal to zero. In a similar manner, we developed an MDB for the reference database DL_i associated with $Sensor_i$.

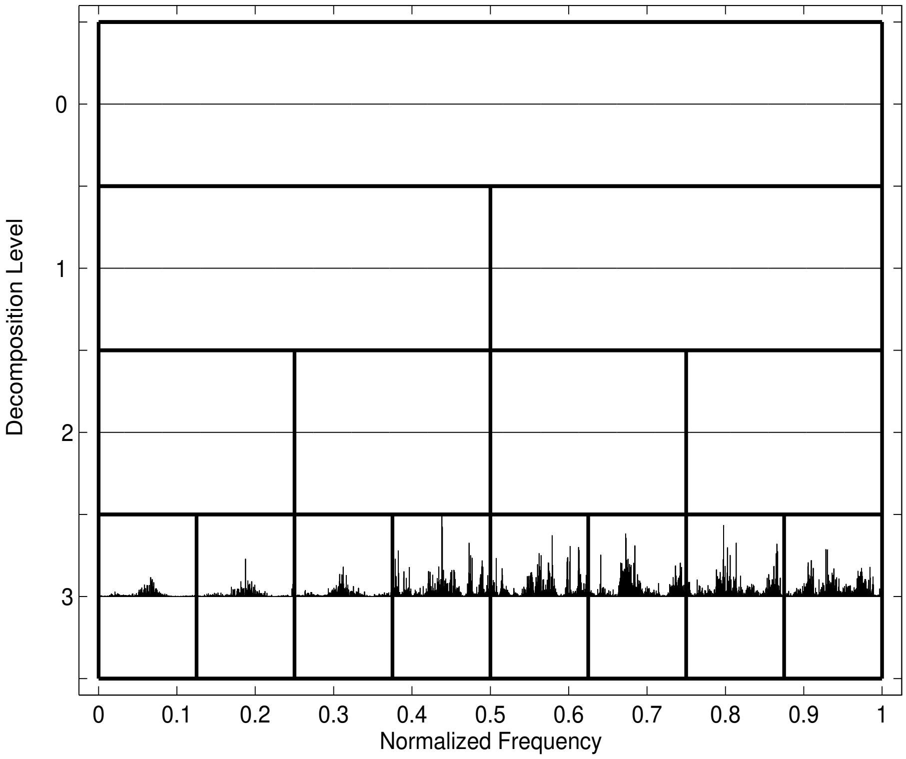

Figure 3: Most Discriminant Basis (MDB) for the DL_r Reference Database

6.2 Symbolic Features

In the following steps, according to our design methodology, we use the symbolic knowledge (source theories, which are presumed to be known at the design time) to select features out of the MDB. Below, we describe the languages, the theories and the models that were used in our experiments. An explanation of the meaning of the features is given in Section 6.3.

The languages L_r, L_i for $Sensor_r$ and $Sensor_i$, respectively, are:

$$L_r = \{bron_r, crac_r, rhon_r, wheez_r, f_r, +, 0, 1, 2, 3, 4, 6, C_{r_0},, \ldots, C_{r_9}\}, \tag{18}$$

$$L_i = \{bron_i, crac_i, rhon_i, wheez_i, f_i, +, 0, 1, 2, 5, 7, C_{i_0}, \ldots, C_{i_9}\}, \tag{19}$$

with the intended interpretation

- $bron_r, crac_r, rhon_r, wheez_r,\ bron_i, crac_i, rhon_i, wheez_i$ are 10-placed relation symbols (classes of lung sounds to be recognized using the $Sensor_r$ and $Sensor_i$ data, respectively),

- $f_r,\ f_i$ are 1-placed function symbols (functions that map feature indices in the DWPD of the $Sensor_r$ and $Sensor_i$ data, respectively, into feature values),

- $+$ is a 2-placed function symbol,

- $0, 1, 2, 3, 4, 5, 6, 7$ are constant symbols,

- $C_{r_0}, \ldots, C_{r_9},\ C_{i_0}, \ldots, C_{i_9}$ are constant symbols (feature indices).

The theories T_r, T_i for $Sensor_r$ and $Sensor_i$, respectively, consist of the following formulas.

Theory T_r

$$\begin{aligned}
&bron_r(C_{r_0}, \ldots, C_{r_9}) \Longleftrightarrow \\
&f(C_{r_0}) \leq f(C_{r_2}) \leq f(C_{r_1}) \ \wedge\ f(C_{r_0}) + 2 \leq f(C_{r_1}) \ \wedge \\
&f(C_{r_3}) = f(C_{r_4}) = f(C_{r_5}) = f(C_{r_6}) = f(C_{r_7}) = f(C_{r_8}) = f(C_{r_9}) = 0,
\end{aligned} \tag{20}$$

$$\begin{aligned}
&crac_r(C_{r_0}, \ldots, C_{r_9}) \Longleftrightarrow \\
&f(C_{r_2}) \leq f(C_{r_1}) \leq f(C_{r_0}) \ \wedge\ f(C_{r_2}) + 1 \leq f(C_{r_0}) \leq f(C_{r_2}) + 2 \ \wedge \\
&f(C_{r_5}) \leq f(C_{r_3}) \leq f(C_{r_4}) \leq f(C_{r_6}) \ \wedge\ f(C_{r_5}) + 2 \leq f(C_{r_6}) \leq f(C_{r_5}) + 4 \ \wedge \\
&f(C_{r_9}) \leq f(C_{r_8}) \leq f(C_{r_7}) \ \wedge\ f(C_{r_9}) + 3 \leq f(C_{r_7}),
\end{aligned} \tag{21}$$

$$\begin{aligned}
&rhon_r(C_{r_0}, \ldots, C_{r_9}) \Longleftrightarrow \\
&f(C_{r_1}) \leq f(C_{r_2}) \leq f(C_{r_0}) \ \wedge\ f(C_{r_1}) + 6 \leq f(C_{r_0}) \ \wedge \\
&f(C_{r_3}) \leq f(C_{r_6}) \leq f(C_{r_5}) \leq f(C_{r_4}) \ \wedge \\
&f(C_{r_4}) \leq f(C_{r_3}) + 2 \ \wedge\ f(C_{r_7}) = f(C_{r_8}) = f(C_{r_9}) = 0,
\end{aligned} \tag{22}$$

$$\begin{aligned}
&wheez_r(C_{r_0}, \ldots, C_{r_9}) \Longleftrightarrow \\
&f(C_{r_0}) = f(C_{r_1}) = f(C_{r_2}) = 0 \ \wedge\ f(C_{r_6}) \leq f(C_{r_4}) \leq f(C_{r_3}) \leq f(C_{r_5}) \ \wedge \\
&f(C_{r_6}) + 6 \leq f(C_{r_5}) \ \wedge\ f(C_{r_9}) \leq f(C_{r_8}) \leq f(C_{r_7}) \ \wedge\ f(C_{r_7}) \leq f(C_{r_9}) + 2,
\end{aligned} \tag{23}$$

$$C_{r_0} \leq C_{r_1} \leq C_{r_2} \leq C_{r_3} \leq C_{r_4} \leq C_{r_5} \leq C_{r_6} \leq C_{r_7} \leq C_{r_8} \leq C_{r_9}. \tag{24}$$

Theory T_i

$$\begin{aligned}
&bron_i(C_{i_0}, \ldots, C_{i_9}) \Longleftrightarrow \\
&f(C_{i_0}) \leq f(C_{i_2}) \leq f(C_{i_3}) \leq f(C_{i_1}) \ \wedge\ f(C_{i_0}) + 5 \leq f(C_{i_1}) \ \wedge \\
&f(C_{i_4}) = f(C_{i_5}) = 0 \leq f(C_{i_6}) \ \wedge\ f(C_{i_6}) \leq f(C_{i_4}) + 1 \ \wedge \\
&f(C_{i_7}) = f(C_{i_8}) = f(C_{i_9}) = 0,
\end{aligned} \tag{25}$$

$$
\begin{aligned}
crac_i(C_{i_0},\ldots,C_{i_9}) &\Longleftrightarrow \\
f(C_{i_0}) = f(C_{i_1}) \le f(C_{i_2}) = f(C_{i_3}) \ \wedge\ f(C_{i_3}) \le f(C_{i_0}) + 2 &\ \wedge \\
f(C_{i_5}) \le f(C_{i_4}) = f(C_{i_6}) \ \wedge\ f(C_{i_5}) \le f(C_{i_4}) + 2 &\ \wedge \\
f(C_{i_9}) \le f(C_{i_7}) \le f(C_{i_8}) \ \wedge\ f(C_{i_9}) + 2 \le f(C_{i_8}), & \qquad (26)
\end{aligned}
$$

$$
\begin{aligned}
rhon_i(C_{i_0},\ldots,C_{i_9}) &\Longleftrightarrow \\
f(C_{i_3}) \le f(C_{i_2}) \le f(C_{i_0}) \le f(C_{i_1}) \ \wedge\ f(C_{i_3}) + 7 \le f(C_{i_1}) &\ \wedge \\
f(C_{i_4}) \le f(C_{i_5}) \le f(C_{i_6}) \ \wedge\ f(C_{i_4}) + 7 \le f(C_{i_6}) &\ \wedge \\
f(C_{i_7}) = f(C_{i_8}) = f(C_{i_9}) = 0, & \qquad (27)
\end{aligned}
$$

$$
\begin{aligned}
wheez_i(C_{i_0},\ldots,C_{i_9}) &\Longleftrightarrow \\
f(C_{i_1}) \le f(C_{i_2}) \le f(C_{i_0}) \le f(C_{i_3}) \ \wedge\ f(C_{i_3}) \le f(C_{i_1}) + 2 &\ \wedge \\
f(C_{i_6}) \le f(C_{i_5}) \le f(C_{i_4}) \ \wedge\ f(C_{i_4}) \le f(C_{i_6}) + 2 &\ \wedge \\
f(C_{i_7}) \le f(C_{i_9}) \le f(C_{i_8}) \ \wedge\ f(C_{i_7}) + 2 \le f(C_{i_8}), & \qquad (28)
\end{aligned}
$$

$$
C_{i_0} \le C_{i_1} \le C_{i_2} \le C_{i_3} \le C_{i_4} \le C_{i_5} \le C_{i_6} \le C_{i_7} \le C_{i_8} \le C_{i_9}. \qquad (29)
$$

As we can see, the theories state that all the constants are ordered (Eqs. 24,29), and they define four relations (see also Section 6.3). The recognition problem is to decide whether and which one of these relations is fulfilled in the world. For this purpose we need to connect the theories to the world, i.e., construct *models* for the languages L_r and L_i. The models M_r and M_i of the languages L_r and L_i, respectively, are defined as:

$$
M_r =< A; bron_r, crac_r, rhon_r, wheez_r, W_{sr}, +, 0, 1, \ldots, 9; I_r >, \qquad (30)
$$

$$
M_i =< A; bron_i, crac_i, rhon_i, wheez_i, W_{si}, +, 0, 1, \ldots, 9; I_i >, \qquad (31)
$$

where $A = \{0, \ldots, 9\}$ is a universe of the models (in our case, these numbers are indices of the nine wavelet coefficients selected out of the complete DWPD of a given signal), and I_r, I_i are interpretation functions that map symbols of the languages L_r, L_i to appropriate relations, functions and constants in the universe A. I_r and I_i assign relations $bron_r, crac_r, rhon_r, wheez_r$, $bron_i, crac_i, rhon_i, wheez_i \subset A^{10}$ to the symbols (we use the same symbols to represent relations) in the languages L_r, L_i. Similarly, they assign functions $W_{sr}, W_{si} : A \to A$, in the model to the symbols f_r, f_i in the languages L_r, L_i, respectively (they represent the values of the ten wavelet coefficients). For simplicity, the function W_s was normalized to the same range as the indices of the features, and rounded so that its values are within the interval of integers from 0 to 9. And finally, the interpretation functions assign constants $0, \ldots, 9$ in the models to the constant symbols $C_{r_0}, \ldots, C_{r_9}, C_{i_0}, \ldots, C_{i_9}$ in the languages. The special constants $0, \ldots, 7$ in the models are assigned (the same) symbols in the languages.

6.3 Theory Construction

Although the construction of source theories is not addressed in this paper, we give an intuitive explanation on how we constructed these theories and what is the meaning of particular features.

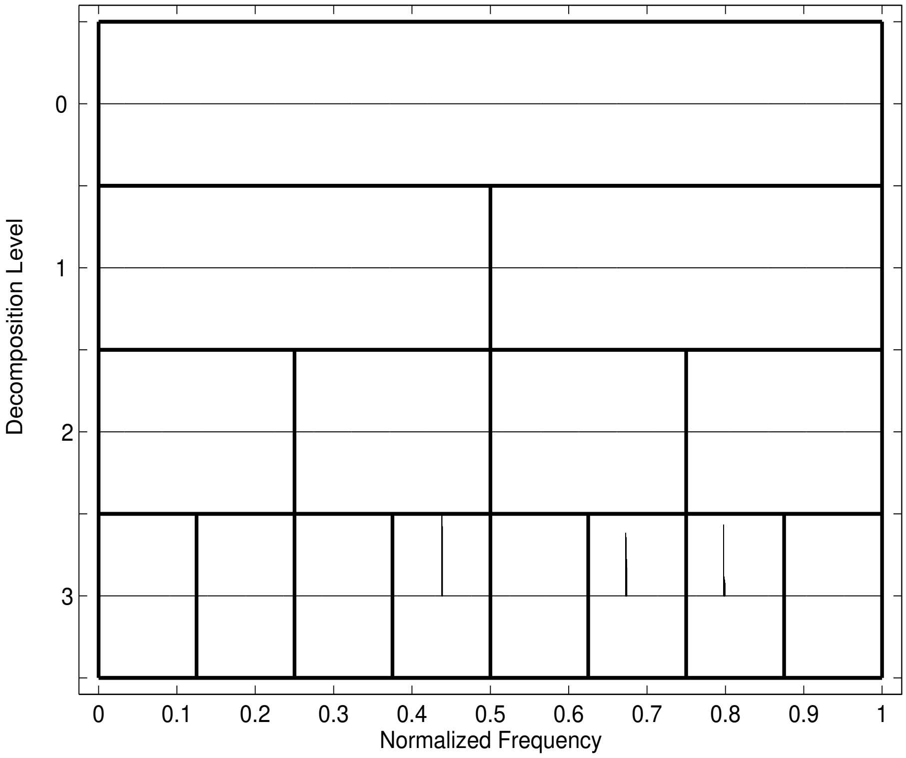

Figure 4: Features Selected From the MDB for the DL_r Database Using the Domain Theory

The first step in selecting these features was to visually analyze the Most Discriminant Wavelet Coefficients (MDWC) and to identify the patterns in the DWPD of the lung sounds by which particular classes of sounds can be discriminated. Since particular patterns are localized in time, we were able to focus on each pattern and review all the signatures in the database with the intent of finding such discriminating patterns. Based on this inspection, we selected ten features concentrated in three MDB areas (see Figure 4). The first three neighboring features are in the fourth subbasis (frequency channel) at the third level of the DWPD, the next four neighboring features are in the sixth subbasis at the third level of the DWPD, and the last three neighboring features are in the seventh subbasis at the third level of the DWPD. In Figure 5) we show the areas of concentration of features (for $Sensor_r$) with vertical bars.

All ten features, depending on their value, carry information about the type of symbolic lung sound features in the signal. In particular, the values of the first three features (the second vertical solid line in Figure 5) indicate the existence of the shape of main transients characteristic for *bronchial*, *crackles*, *rhonchi* or *wheezes*. The values of the last seven features (the first vertical solid line in Figure 5) indicate the existence of the shape of main transients characteristic for *crackles* or *wheezes*, or the existence of a transition between inspiration and expiration cycle characteristic for *bronchial* or *rhonchi*.

In the next step we identified relationships among particular features. In our case, we used binary equality and inequality relations. These relations are captured by the

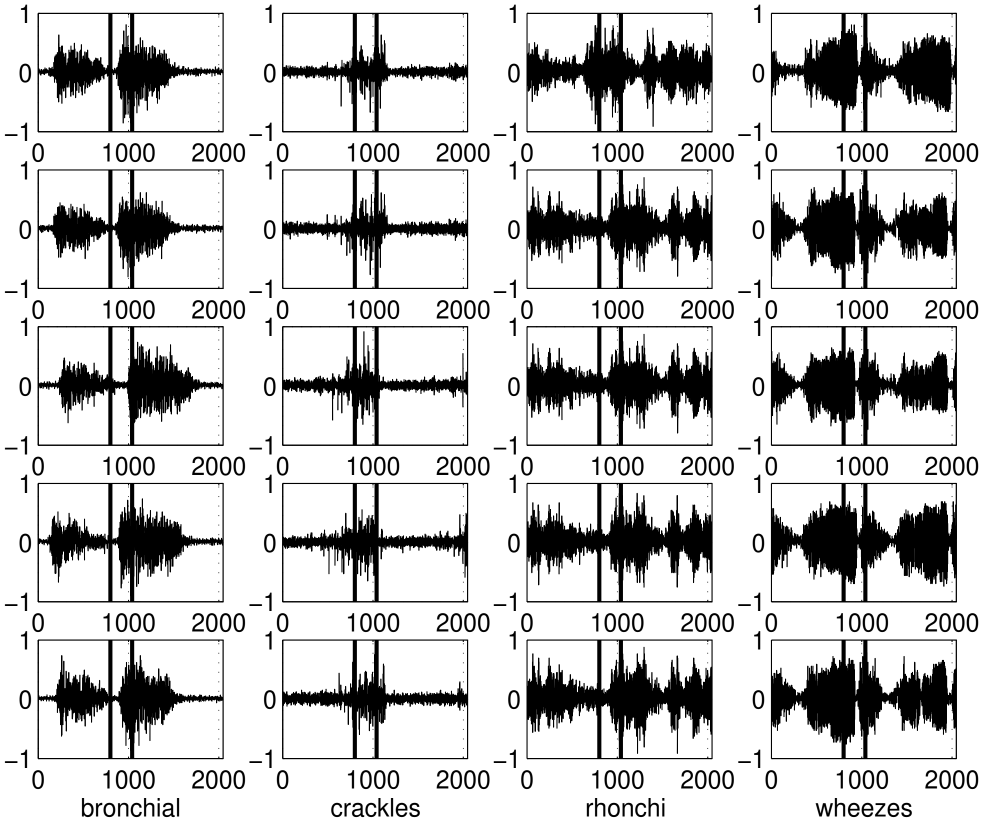

Figure 5: Areas of Concentration of Selected Symbolic Features for DL_r (normalized amplitude vs. time [msec])

formulas of theories, and then used by the model checker to decide whether a particular signal is an element of a particular class. Note, that the values of single features are not used for recognition.

For comparison, in Figure 6, we show which of the features were selected by the entropy-based measure as Most Discriminant Wavelet Coefficients (MDWC), assuming that $k_f = 10$. We can clearly see that the main difference is that our features are selected in groups (neighbors), while MDWC features happened to be rather far apart.

6.4 Model Fusion

Lung sound recognition based upon checking models may generate different decisions for each sensor. In order to arrive with only one decision, we need to have a one combined model that would incorporate the features from each of the sensors. We develop such a model in the step of *Model Construction*; we call it the *fused model*. In the process of *Model Construction* we combine the languages L_r, L_i into one fused language L_f, the theories T_r, T_i into one fused theory T_f, and the models M_r, M_i into one fused model M_f consistently, so that M_f is a model for T_f. Since we do not have a theory that

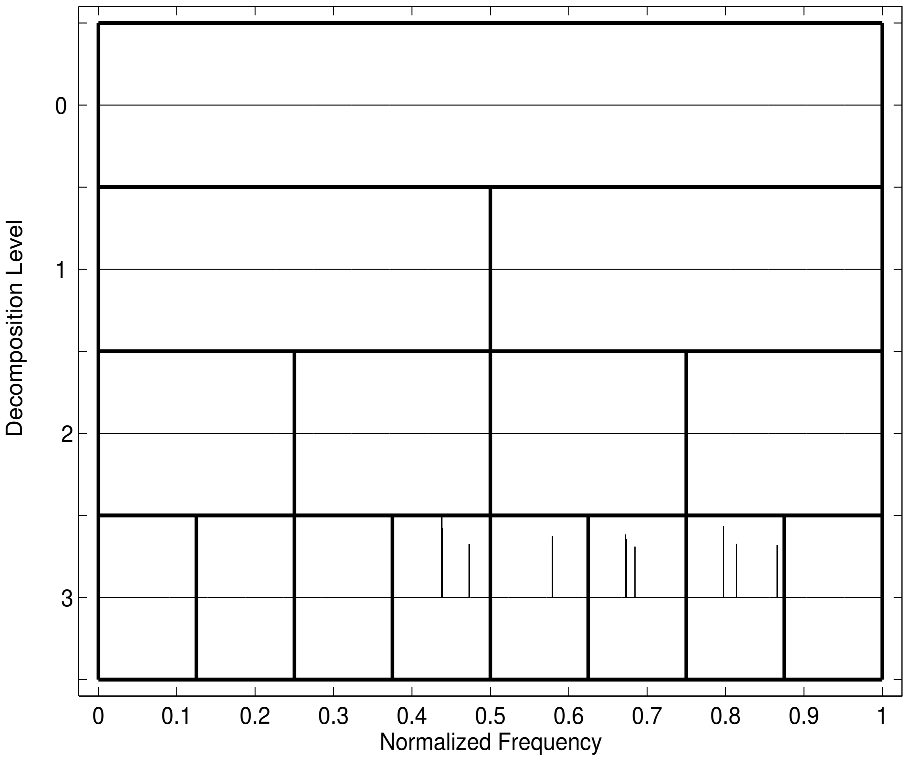

Figure 6: Most Discriminant Wavelet Coefficients (MDWC features) Selected from the MDB for the DL_r Database

would embed both T_r and T_i, we need to develop a *fusion operator* that accomplishes the above described result. In this process we use the operators of *reduction, expansion* and *union* described in Section 5.2. Below, we describe how we arrived with the resulting fused language, fused theory, and fused model for the lung sound recognition problem described in this paper.

The strategy for constructing a fused system is based upon the understanding that as a net result we want to have a language that is roughly of the same complexity as the initial languages L_r and L_i. To achieve this we need to remove some of the constants (preferably the least significant for recognition) from each of the languages. Another element of the strategy is the fact that the ultimate step in the fusion process will be the application of the union operator to the languages, models and theories. To avoid inconsistency problems with the union of functions, we want to restrict the domains of the initial functions so that after this operation the sets on which the two functions (i.e., the restricted f_r and the restricted f_i) are defined are disjoint.

In this particular case, we first remove the constants $C_{r_3}, \ldots, C_{r_7}$ and $C_{i_3}, \ldots, C_{i_7}$ from the languages L_r and L_i. This results in new (restricted) relations and functions. We need new names for these relations and functions:

$$X_r = \{bron'_r, crac'_r, rhon'_r, wheez'_r, f'_r\}, \tag{32}$$

$$X_i = \{bron_i', crac_i', rhon_i', wheez_i', f_i'\}. \tag{33}$$

First, we apply the expansion operator (see Section 5.2), then the reduction operator. The result is two new languages

$$L_r^{er} = \{bron_r', crac_r', rhon_r', wheez_r', f_r', +, 0, 1, 2, 6, C_{r_0}, C_{r_1}, C_{r_2}, C_{r_8}, C_{r_9}\}. \tag{34}$$

$$L_i^{er} = \{bron_i', crac_i', rhon_i', wheez_i', f_i', +, 0, 2, 5, C_{i_0}, C_{i_1}, C_{i_2}, C_{i_8}, C_{i_9}\}. \tag{35}$$

The new symbols are interpreted as restrictions of the relations and of the function, e.g.,

$$\begin{aligned} bron1_r' &= bron_r \cap (A \cap \{0,1,2,8,9\})^3, \\ f_r' &= f_r \mid_{A\cap\{0,1,2,8,9\}} . \end{aligned} \tag{36}$$

(Note that (36) interprets f_r' in terms of W_{sr}, since it uses f_r, which was interpreted in terms of W_{sr} in (30).) In parallel to this operation on the language, we restrict the theories T_r and T_i by removing from their formulas the six constants and terms that depend on these constants; we denote the resulting theories as T_r', T_i'.

Our next goal in this strategy is to combine these two languages so that as the final result we have one symbol for each class and one function symbol. The formulas of the theory should be derived as a combination of the formulas from appropriate initial theories T_r', T_i'. Similarly, the interpretation should be derived from the interpretations of these theories. We achieve this goal in three steps. First, we apply the union operator to the two languages, theories and their models. Then we apply the expansion operator. And finally, we apply the reduction operator.

The union operator is just an intermediate step that collects the pairs of languages, theories and models in three separate collections, corresponding to a new language, theory and model. The model domain remains the same, i.e., A. The expansion operator adds new symbols: $bron, crac, rhon, wheez, C_0, \ldots, C_9$. By applying the product of relations operator, these new symbols are interpreted in the following way.

$$\begin{aligned} C_0 = C_{r_0}, C_1 = C_{r_1}, C_2 = C_{r_2}, C_6 = C_{r_8}, C_7 = C_{r_9}, \\ C_3 = C_{i_0}, C_4 = C_{i_1}, C_5 = C_{i_2}, C_8 = C_{i_8}, C_9 = C_{i_9} \\ f = f_r \cup f_i', f:\ A \to A, \end{aligned} \tag{37}$$

$$\begin{aligned} bron = bron_r \times bron_i', crac = crac_r \times crac_i', \\ rhon = rhon_r \times rhon_i', wheez = wheez_r \times wheez_i', \\ bron, crac, rhon, wheez \subset A^{10}. \end{aligned} \tag{38}$$

And finally, the reduction operator removes all the unnecessary elements from the union. As a result we end up with the following fused language L_f.

$$L_f = \{bron, crac, rhon, wheez, f, +, 0, 1, 5, 6, C_0, \ldots, C_9\}. \tag{39}$$

The feature-level fused theory T_f consists of the following five formulas.

$$\begin{aligned}
&bron(C_0, \ldots, C_9) \Longleftrightarrow \\
&f(C_0) \leq f(C_2) \leq f(C_1) \ \wedge\ f(C_0) + 2 \leq f(C_1) \ \wedge\ f(C_3) \leq f(C_5) \leq f(C_4) \ \wedge \\
&f(C_3) + 5 \leq f(C_4) \ \wedge\ f(C_6) = f(C_7) = f(C_8) = f(C_9) = 0,
\end{aligned} \tag{40}$$

$$\begin{aligned}
&crac(C_0, \ldots, C_9) \Longleftrightarrow \\
&f(C_2) \leq f(C_1) \leq f(C_0) \ \wedge\ f(C_2) + 1 \leq f(C_0) \leq f(C_2) + 2 \ \wedge \\
&f(C_3) = f(C_4) \leq f(C_5) \ \wedge\ f(C_5) \leq f(C_3) + 2 \ \wedge \\
&f(C_7) \leq f(C_6) \ \wedge\ f(C_9) + 2 \leq f(C_8),
\end{aligned} \tag{41}$$

$$\begin{aligned}
&rhon(C_0, \ldots, C_9) \Longleftrightarrow \\
&f(C_1) \leq f(C_2) \leq f(C_0) \ \wedge\ f(C_1) + 6 \leq f(C_0) \ \wedge\ f(C_5) \leq f(C_3) \leq f(C_4) \ \wedge \\
&f(C_6) = f(C_7) = f(C_8) = f(C_9) = 0,
\end{aligned} \tag{42}$$

$$\begin{aligned}
&wheez(C_0, \ldots, C_9) \Longleftrightarrow \\
&f(C_0) = f(C_1) = f(C_2) = 0 \ \wedge\ f(C_4) \leq f(C_5) \leq f(C_3) \ \wedge\ f(C_7) \leq f(C_6) \ \wedge \\
&f(C_7) \leq f(C_6) \ \wedge\ f(C_9) \leq f(C_8),
\end{aligned} \tag{43}$$

$$C_0 \leq C_1 \leq C_2 \leq C_3 \leq C_4 \leq C_5 \leq C_6 \leq C_7 \leq C_8 \leq C_9. \tag{44}$$

The feature-level fused model M_f of the theory T_f is defined as

$$M_f =< A; bron, crac, rhon, wheez, W_{sf}, +, 0, 1, 2, 5, 6; I_f >, \tag{45}$$

where I_f is an interpretation function that maps symbols of the language L_f to appropriate relations, functions and constants in the universe A, as described above. The fused model M_f contains ten features created as a result of fusion of the five features from the DL_r database and the five features from the DL_i database. In particular, the first three features of this fused feature vector are the same as the first three features selected from DL_r; the next three fused features are the same as the first three features selected from DL_i; the next two fused features are the same as the last two features selected from DL_r; the last two fused features are the same as the last two features selected from DL_i. W_{sf} is the interpretation of the function symbol f. Since this function is a union of (subsets) of functions f_r and f_i, its values are determined by the values of these two functions.

7 Results of Experiments

To evaluate our approach to lung sound recognition, we present below the results of tests on the lung sound recognition scenario described in Section 3. The results of our experiments are summarized in Figures 7 and 8. The figures show misclassification rates for various levels of noise.

To evaluate our approach, we implemented three recognition systems: one unisensor and one multisensor, built according to our methodology, and additionally, a system that selects features using solely an entropy-based decision procedure (BDBA).

Since our main goal was to assess the quality of features selected according to our methodology, we decided to compare our features versus the BDBA selected features, due to the fact that the BDBA features are optimal with respect to the entropy-based classification measure, which is considered a good measure to evaluate the discriminative power of recognition systems. To make the comparison of features as fair as possible, we used the same signal classification algorithms as used by Saito [31], for both our features and the BDBA selected features. Also the number of features used was the same for both cases. Both the BDBA features and our model-theory based features were evaluated using the same test databases.

To evaluate the robustness of the system to noise, we added zero-mean Gaussian noise with standard deviation varied at 11 levels (with step 2.5) within the deviation range of 0 through 25, to each signal in the test databases. Each signal in the test databases was classified by the three recognition systems. The quality of recognition was measured in terms of the misclassification rate. The result was a misclassification, if the decision did not assign the correct class label to the signal. One data point in each of the figures was obtained by averaging the number of misclassifications over all the classification experiments for a given level of noise.

Figure 7 shows the resulting misclassification rates for the unisensor system based on our methodology (AUFRS) for different levels of noise. This figure also shows the misclassification rates of the recognition system using Most Discriminant Wavelet Coefficients (MDWC) as features. The misclassification results show that the AUFRS has a better recognition accuracy than the MDWC-based system.

Figure 8 shows the resulting misclassification rates for different levels of noise for the multisensor system (AMFRS). This figure also shows the misclassification rates of a multisensor recognition system using Most Discriminant Wavelet Coefficients (MDWC) as features, and of the two unisensor systems. The misclassification results show that for the lung sound recognition problem the AMFRS has a better recognition accuracy than the MDWC-based recognition system.

8 Conclusions

The main goal of this paper was to show a formal systematic approach to incorporating symbolic knowledge into recognition systems and to show that a system based on this methodology can be used for lung sound recognition. We addressed several problems related to the design of recognition systems that utilize multiple sensors and symbolic

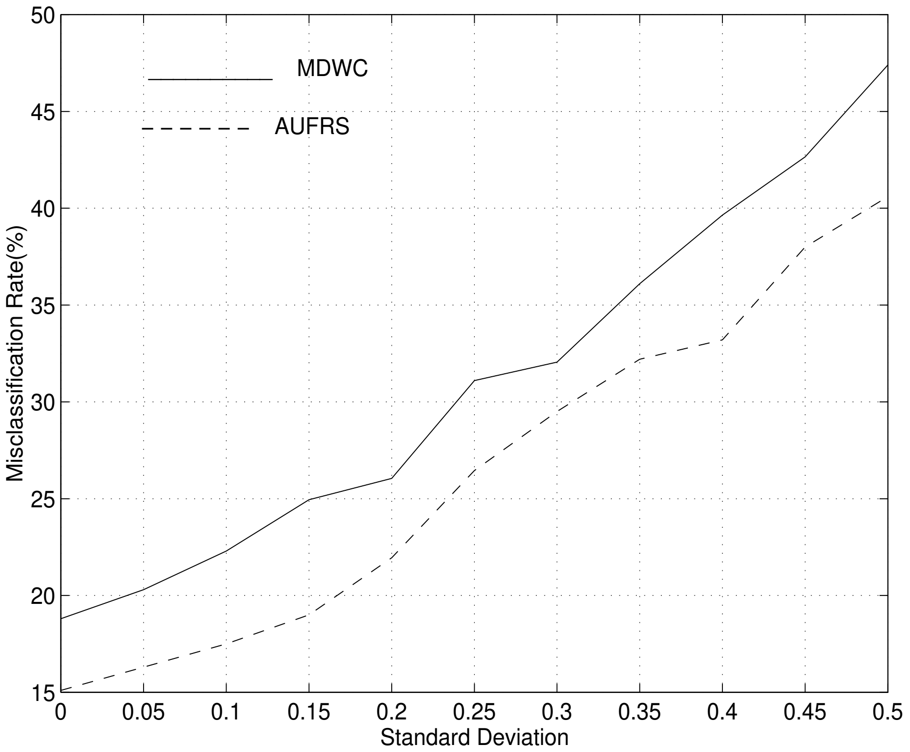

Figure 7: Misclassification Rates for the Unisensor Lung Sound Recognition Systems: Using Model Theory for Feature Selection (AUFRS) and Using Most Discriminant Wavelet Coefficients as Features (MDWC)

knowledge. Our main focus was to show how to select features out of sensor signals, given a collection of theories about signal sources. We assumed that such theories are known to the system designer. We used wavelets and entropy-based measures for feature extraction. Then we used feature selection guided by the models of particular source theories. Since we deal with multiple sensors, we also showed how, using source theories and their models, to combine features from different sensors.

Obstacles to incorporating symbolic knowledge into signal recognition systems include the lack of methods for extracting symbolic features from sensor signals and the lack of libraries of symbolic knowledge that could be used in the process of signal recognition. Additionally, there is a lack of a methodology for designing sensor fusion systems that incorporate both quantitative and symbolic knowledge. In this paper we showed the steps involved in the design of a multisensor recognition system that extracts symbolic features and utilizes symbolic knowledge for recognition. We also compared the performance of the resulting system with both a single-sensor system and with a system that selects features using an entropy-based feature selection method.

Our experiments showed that the entropy-based recognition system adapted well to the training set of signatures, but it did not have enough generalization power to perform equally well with the test data. The reason for that was that the features selected by the entropy-based criteria were too far apart within the time/frequency domain (see

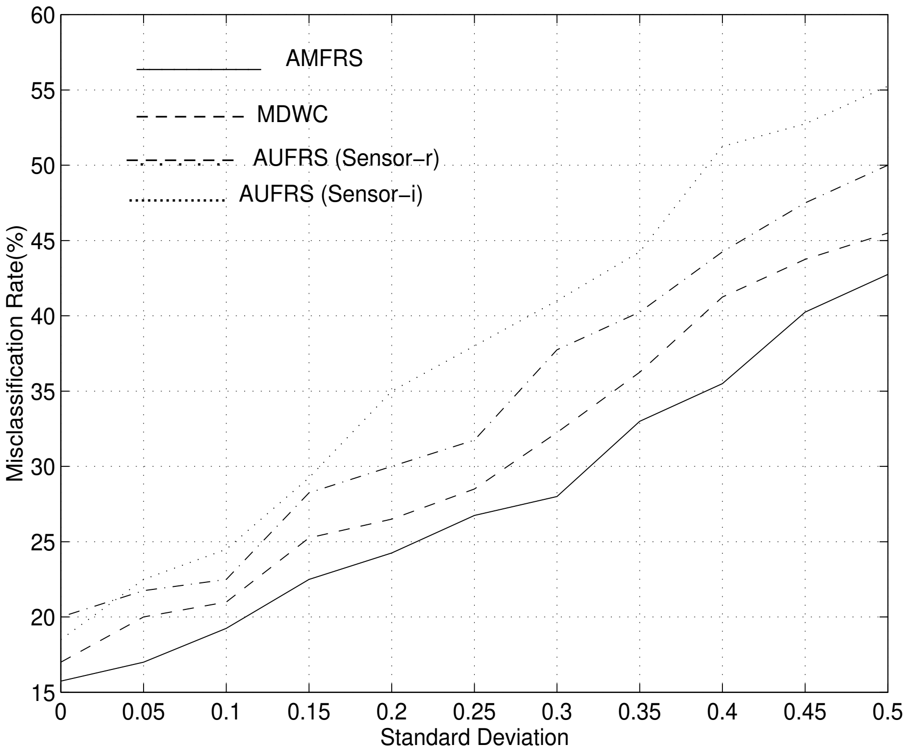

Figure 8: Misclassification Rates for the Multisensor Lung Sound Recognition Systems: Using Model Theory for Feature Selection and Fusion (AMFRS) and Using Most Discriminant Wavelet Coefficients as Features (MDWC)

Figure 6). Our system, on the other hand, due to selecting features that are grouped in two groups in the time/frequency domain, has a better generalization power and thus performs better on noisy signals.

The complexity of the algorithms was addressed in several ways. First, some of the algorithms are used off-line, i.e., they are used in the process of design rather than in the system operation. Second, the run-time recognition system uses only a small subset (k_f) features for recognition. The complexity of the DWPD algorithm, which is used on-line, is on the order $O(n)$, where n is the number of samples. And finally, the neural network classifier can be easily mapped to a parallel implementation.

This research can be extended in several directions. One of the questions that can be investigated is what is the sensitivity of the lung sound recognition process to the selection of different sets of features? In this paper we showed the impact of feature selection for two sets of features: one selected using an entropy-based criterion, and one selected using a theory of the signal source. Other selections, including random, could be analyzed.

We constructed theories manually, by observing relations among various features for all of the signals in the signature database. This process could be supported by the use of machine learning techniques. Also, a formal method tool (a theorem prover) could be used to help with checking the consistency of the derived theories.

Another direction for this research is to investigate domain-specific aspects, i.e., various features and theories characterizing lung sounds. The features of lung sounds depend on various characteristics of the patient, like weight, age, sex and the physical condition, as well as on the environment, like humidity, temperature, and time of day. Theories capturing these dependencies can be built. Once created, they can be easily utilized by our system. Also, the difficulty of recognition of different types of lung sounds may vary, and thus the dependence of misclassification rate on the type of the sound could be investigated separately for each sound type. Another aspect related to the lung sound recognition domain is the need to separate lung sounds from heart sounds and other signals that are acquired through a stethoscope in the process of auscultation. Although our data included, for instance heart sounds, the ability of this approach to deal with various kinds of noise (not only Gaussian) could be investigated further.

Acknowledgements

This research was partially sponsored by the Defense Advanced Research Projects Agency under Grant F49620-93-1-0490 and by University Research Engineers and Associates. The authors wish to thank Tom Callahan for his support and encouragement. The authors are grateful for many questions and suggestions received from the anonymous reviewers.

References

[1] E. Ademovic, G. Charbonneau, and J. C. Pesquel. Segmentation of infant respiratory sounds with Malvar's wavelets. In *Proc. 16th Int'l Conf. IEEE Engin. in Medicine and Biology Society*, pages 20a–21a, 1994.

[2] J. K. Aggarwal. *Multisensor Fusion for Computer Vision.* Springer-Verlag, 1993.

[3] S. W. Banham, R. B. Urquhart, J. E. S. Macleod, and F. Moran. Alteration in the low-frequency lung sounds in respiratory disorders associated with crackles. *Eur. J. Respirat. Dis.*, 65:58–63, 1984.

[4] R. P. Baughman and R. G. Louden. Quantization of wheezing in acute asthma. *Chest*, 5:718–722, 1984.

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

[6] A. Cohen and D. Landsberg. Analysis and automatic classification of breath sounds. *IEEE Tran. Biomed. Engin.*, BME-31, no. 9, 1984.

[7] R. Coifman, R. and M. V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory*, 38, no.2:713–718, 1992.

[8] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, XLI:909–996, 1988.

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

[10] L.-G. Durand, Guo. Z., H. N. Sabbah, and P. D. Stein. Comparison of spectral techniques for computer-assisted classification of spectra of heart sounds in patients with porcine bioprosthetic valves. *Medical & Biological Engineering & Computing*, 5:229–236, 1993.

[11] P. Forgacs. Crackles and wheezes. *The Lancet*, pages 203–205, 1967.

[12] N. Gavriely, Y. Palti, G. Alroy, and J. G. Grotberg. Measurement and theory of wheezing breath sounds. *Physiol. Respirat. Environ. Exercise Physiol.*, 57, no. 2:481–492, 1984.

[13] G. Guler, B. Sankur, and Y. Kahya. Multiresolution biological transient extraction applied to respiratory crackles. *Computers in Biology and Medicine*, 26:25–39, 1996.

[14] L. J. Hadjileontiadis and S. M. Panas. Autoregressive modeling of lung sounds using higher-order statistics: Estimation of source and transmission. In *1997 IEEE Signal Processing Workshop on Higher-Order Statistics*, pages 4–8, 1997.

[15] L. J. Hadjileontiadis and S. M. Panas. Separation of discontinuous adventitious sounds from vesicular sounds using a wavelet-based filter. *IEEE Transactions on Biomedical Engineering*, 44, no. 12:1269–1281, 1997.

[16] D. Hager, G. *Task-Directed Sensor Fusion and Planning: A Computational Approach.* Kluwer, 1990.

[17] D. L. Hall. *Mathematical Techniques in Multisensor Data Fusion.* Artech House, Boston - London, 1992.

[18] J. P. Heemels, H. H. Ross, P. Sipkema, and F. De Vries. Tracheal sound detector. *Med. & Biol. Eng. & Comput.*, 24:182–185, 1986.

[19] L. Hong. Multiresolutional filtering using wavelet transform. *IEEE Transactions on Aerospace and Electronic Systems*, 29(4):1244–1251, 1993.

[20] V. K. Iyer, P. A. Ramamoorthy, and Y. Ploysongsang. Autoregressive modeling of lung sounds: Characterization of source and transmission. *IEEE Tran. Biomedical Engineering*, 36, no. 11, 1989.

[21] T. Katila, P. Piirila, E. Paajanen, T. Rosqvist, and A. R. A. Sovijarvi. Original waveform of lung sound crackles: a case study of the effect of high-pass filtration. *J. Appl. Physiol.*, 71(6):2173–2177, 1991.

[22] L. A. Klein. *Sensor and Data Fusion Concepts and Applications.* SPIE, Bellingham, WA, 1993.

[23] M. M. Kokar and J. Tomasik. Towards a formal theory of sensor/data fusion. Technical Report COE-ECE-MMK-1/94 (available at http://www.coe.neu.edu/ kokar), Northeastern University, ECE, Boston, MA, 1994.

[24] P. J. Kolodzy. Multidimensional automatic target recognition system evaluation. *The Lincoln Laboratory Journal*, 6(1):117–146, 1993.

[25] R. C. Luo and M. G. Kay. Multisensor integration and fusion in intelligent systems. *IEEE Transactions on Systems, Man and Cybernetics*, 19-5:901–931, 1989.

[26] R. L. H. Murphy. A simplified introduction to lung sounds. Technical report, Lemuel Shattuck and Faulkner Hospitals, 170 Morton Street, Jamaica Plain, MA 02130, 1977.

[27] M. J. Musell. The need for standards in recording and analysing respiratory sounds. *Medical & Biological Engineering & Computing*, 30:129–139, 1992.

[28] M. J. Musell, Y. Nakazono, and Y. Miyamoto. Effect of air flow and flow transducer on tracheal breath sounds. *Medical & Biological Engineering & Computing*, 28:550–554, 1990.

[29] H. Pasterkamp, C. Carson, D. Daien, and Y. Oh. Digital respirosonography: New images of lung sounds. *Chest*, 96, no. 6:1405–1412, 1989.

[30] Y. Ploysongsang, V. K. Iyer, and P.A. Ramamoorthy. Inspiratory and expiratory vesicular breath sounds. *Respiration*, 57:313–317, 1990.

[31] N. Saito. *Local Feature Extraction and Its Applications Using a Library of Bases.* PhD thesis, Yale University, 1994.

[32] P. B. Samir, T. F. Callahan, and M. G. Callahan. An adaptive noise reduction stethoscope for auscultation in high noise environments. *The Journal of the Acoustical Society of America*, 103, no. 5:2483–2491, 1998.

[33] B. Sankur, Y. P. Kahya, E. C. Guler, and T. Engin. Comparison of AR-based algorithms for respiratory sounds classification. *Comput. Biol. Med.*, 24, No. 1:67–76, 1994.

[34] D. R. Smith. KIDS: a semi-automatic program development system. *IEEE Transactions on Software Engineering*, 16 (9):1024–1043, 1990.

[35] Y. V. Srinivas and R. Jullig. Specware^{TM}: Formal support for composing software. Technical Report KES.U.94.5, Kestrel Institute, 1994.

[36] R. B. Urquhart, J. McGhee, J. E. S. Macleod, S. W. Banham, and F. Moran. The diagnostic value of pulmonary sounds: a preliminary study by computer-aided analysis. *Comput. Biol. Med.*, 2:129–139, 1981.

[37] P. K. Varshney. *Distributed Detection and Data Fusion.* Springer-Verlag, 1996.

[38] E. Waltz and J. Llinas. *Multisensor Data Fusion.* Artech House, Norwood, MA, 1990.

[39] J. M. Wing. A specifier's introduction to formal methods. *IEEE Computer*, 9:8–24, 1990.

[40] G. M. Zenk. Stethoscopic detection of lung sounds in high noise environments. Master's thesis, Purdue University, 1994.