Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science
Research Laboratory of Electronics
Room 36-615
Cambridge, MA   02139

AD-A169 166

DTIC
ELECTE
JUN 2 4 1986
S
D
D

Signal Processing and Interpretation

using

Multilevel Signal Abstractions

Evangelos E. Milios

Technical Report No. 516

June 1986

DTIC FILE COPY

ADA169166

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT
	Approved for public release; distribution
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
Research Laboratory of Electronics Massachusetts Institute of Technology		Office of Naval Research Mathematical and Information Scien. Div.

6c. ADDRESS (City, State and ZIP Code)	7b. ADDRESS (City, State and ZIP Code)
77 Massachusetts Avenue Cambridge, MA 02139	800 North Quincy Street Arlington, Virginia 22217

8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
Advanced Research Projects Agency		N00014-81-K-0742

8c. ADDRESS (City, State and ZIP Code)	10. SOURCE OF FUNDING NOS.			
1400 Wilson Boulevard Arlington, Virginia 22217	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
			NR 049-506	

11. TITLE (Include Security Classification)
Signal Processing and Interpretation...

12. PERSONAL AUTHOR(S)
Evangelos E. Milios

13a. TYPE OF REPORT	13b. TIME COVERED		14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Technical	FROM _____	TO _____	June 1986	221

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This thesis advances multilevel signal abstractions as a useful conceptual framework in signal processing. More specifically, multilevel abstractions for one-dimensional harmonic spectra are developed and demonstrated to be useful in combining heuristics with algorithmic techniques in a variety of signal processing problems associated with real helicopter data, such as adjusting spectral estimation parameters and tracking fundamental frequency and power of helicopter spectra. Furthermore, multilevel abstractions for two-dimensional wavenumber spectra are used as the basis for a diagnosis system that searches for inappropriate parameter settings in an array processing system for direction determination. The diagnosis system illustrates the use of multilevel signal abstractions in automating reasoning about complex signal processing systems.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS ☐	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
Kyra M. Hall RLE Contract Reports	(617) 253-2569	

DD FORM 1473, 83 APR   EDITION OF 1 JAN 73 IS OBSOLETE

Signal Processing and Interpretation
using
Multilevel Signal Abstractions

by
Evangelos E. Milios

Submitted to the Department of Electrical Engineering and Computer Science
on May 2, 1986 in partial fulfillment of the
requirements for the Degree of Doctor of Philosophy in
Electrical Engineering and Computer Science

ABSTRACT

This thesis advances multilevel signal abstractions as a useful conceptual framework in signal processing. More specifically, multilevel abstractions for one-dimensional harmonic spectra are developed and demonstrated to be useful in combining heuristics with algorithmic techniques in a variety of signal processing problems associated with real helicopter data, such as adjusting spectral estimation parameters and tracking fundamental frequency and power of helicopter spectra. Furthermore, multilevel abstractions for two-dimensional wavenumber spectra are used as the basis for a diagnosis system that searches for inappropriate parameter settings in an array processing system for direction determination. The diagnosis system illustrates the use of multilevel signal abstractions in automating reasoning about complex signal processing systems.

Thesis Supervisor: Dr. Alan V. Oppenheim

Title: Professor of Electrical Engineering

2

Στην μνημη του πατερα μου

To the memory of my father

Acknowledgements

Many individuals have contributed, directly or indirectly, to the research described in this thesis. I would like to gratefully acknowledge:

Alan Oppenheim, my thesis advisor, for his enthusiastic support of the project and his advice on several issues, both technical and non-technical, which has had a profound influence on me throughout the years.

Hamid Nawab for closely collaborating on this project since its early stages. Hamid played a key role both in the conception and maturation of this thesis. His enthusiastic attitude and his advice on various matters have been extremely helpful.

Randy Davis and Bruce Musicus, my thesis readers, for valuable criticism and feedback during the last and formative period of this thesis.

Richard Lacoss and Group 21 at Lincoln Laboratory, for generous moral and financial support over the years of this research, for providing the real helicopter data for this thesis and for making available their expertise with that data.

Webster Dove and Cory Myers for valuable interaction and exchange of ideas over a period of several years and for being always willing to share their Lisp machine ·expertise. In addition, for designing, implementing and making available the KBSP package, a sophisticated signal processing environment for the Lisp machine that supported much of the research in this thesis.

Cory Myers, Michele Covell, Jacek Jachner and Tae Joo for useful feedback on an early draft of this thesis.

The once "senior" DSPG members, especially Bruce Musicus, Greg Duckworth, and Tom Bordley for motivating me to "learn the computer", and thus unwittingly affecting later decisions in my graduate career.

All Digital Signal Processing Group members since the fall of 1981 for providing a stimulating, rewarding, friendly and cooperative social environment. Especially Thrasyvoulos Pappas, for being a special friend, and Mike Wengrovitz, for being a good friend and officemate for an unusually long period of five years.

My old-time friends and classmates, (Drs) Nicholas Leventis and Petros Maragos, who recently came to the Boston area, for making me feel left behind and thus prompting me to finish.

My wife, Evie Tastsoglou, for sharing with me the joys and frustrations of our graduate student lives.

Contents

List of Figures

11

13

14

Chapter 1

Introduction

1.1 Signal abstractions in signal processing

In this thesis, we explore the use of signal abstractions as an organizational principle for real-world signal processing systems in the context of frequency, power and direction tracking of helicopters based on acoustic microphone signals.

In the case of signal processing problems lacking satisfactory models for the signal and noise and involving low signal-to-noise ratios, such as helicopter tracking, signal processing systems with adequate performance on real data are usually developed by heuristic modifications of the algorithms derived from idealized models. The programming paradigm commonly used in expressing such heuristic modifications has been that of procedural abstraction, whereby the program is decomposed into several procedures, with each procedure accomplishing an identifiable task that can be used as a module in other procedures [Abelson & Sussman 1985]. Procedural abstraction alone is not always appropriate for expressing heuristic modifications, leading to programs that are difficult to understand and maintain, especially when such modifications are related not directly to the numeric signal, but to higher-level structures in the signal, such as peaks or harmonically-related sets of peaks.

16

As a step toward increasing the clarity and flexibility of signal processing programs that perform on real data, this thesis uses signal abstractions. A signal abstraction is a view of a signal as consisting of multiple levels, with higher levels of abstraction obtained by suppressing information (i.e. detail) in lower levels. As an example, consider a helicopter spectrum. At the lowest level of abstraction, it is a discrete function of power versus frequency. At a higher level, it consists of several well distinguished peaks. At an even higher level, it consists of two sets of harmonically related peaks, corresponding to the main and the tail rotor of the helicopter. A high level of abstraction supports a smaller set of inferences than a lower level of abstraction, but hopefully does so more elegantly, i.e. with simpler programs, and economically, i.e. with less computation.

In the thesis, we arrive at the suitability of signal abstractions as a useful concept in signal processing from two distinct routes. The first is an extension of data abstraction for suppressing implementation details in the computer representation of signals, as proposed by [Kopec1984, Dove1984]. The second is the observation that when human experts perform signal processing interactively, they tend to view signals not as sequences of numbers, but in terms of features, recognizing in them higher level entities, such as collections of peaks, which have more direct relevance to the phenomenon producing the signals.

Signal abstractions are not the same as data abstraction, because each serves to combat a different kind of complexity present in signal processing systems. Data abstraction applied to signals as in [Kopec1980, Kopec1984] is concerned with the complexity of computer representations of signals and with ensuring that signals can be specified in programs completely with a minimal amount of implementation overhead, such as memory allocation for the numeric samples of the signal or interface with the code for various signal processing operations. In contrast, the signal abstractions of this thesis are concerned with the complexity of the signal as an information carrier.

Furthermore, data abstraction is a design methodology. Design is centered around objects specified by their desired use. At first, we ignore implementation details and we concentrate on the use of an object. If use of the object is not directly implementable in terms of the underlying programming language, the object is decomposed into other objects, specified by their use, which, when combined, will implement the use of the original object. This cycle is repeated until objects can be directly implemented in terms of the underlying programming language.

Signal abstractions are a conceptual framework for thinking about signal processing systems. A major point of the thesis is that within this framework one may obtain solutions to signal processing problems which are expressed in the informal signal terms that people often use when interactively processing and interpreting real-world signals. Moreover, the same framework is useful towards designing programs that "reason" about existing complex signal processing systems by coupling mathematical models and heuristic rules [Apte & Weiss 1985]. These points are demonstrated through two different systems, both in the context of acoustic helicopter tracking.

The first system illustrates the power of signal abstractions in providing new solutions to existing problems that rely on signal correlation. We take a more general view of signal correlation and we expand both its mechanism and its scope, by treating it as a problem of matching signal abstractions. Signal matching is used for ordering signals according to heuristic criteria and for associating elements in different signals. The system we implemented performs tracking of fundamental frequency and power of acoustic helicopter signals from a single microphone signal using its short-time spectrum. Spectral matching is used as the basis for adjusting the parameters of a spectral estimation technique to best suit the acoustic data at hand. It is also used for identifying faulty channels in multichannel acoustic data. The helicopter pitch tracking system relies on explicit associations between higher levels of abstraction of acoustic spectra at different times. The problem of computing power and frequency tracks of helicopter signals was motivated by

18

the goal of localization of helicopters using signals from geographically separated sensors. Although the full problem is beyond the scope of this thesis, a preliminary study was carried out indicating the plausibility of the approach.

The second system is an example of the use of abstractions of two-dimensional spatial Fourier spectra (wavenumber spectra) in a diagnosis system that identifies maladjusted parameters of a complicated signal processing system performing direction determination of acoustic waves from multichannel microphone data. The diagnosis system takes as input abstract descriptions of the aircraft scenario and the corresponding output of the signal processing system and produces an explanation of their differences in terms of maladjusted signal processing parameters and/or propagation phenomena. The diagnosis system uses means-ends analysis [Newell & Simon 1963] to search the space of all possible explanations as directed by the differences between the input scenario and the output of the signal processing system, where both are expressed as wavenumber spectrum abstractions. The current diagnosis system assumes that an approximate description of the input scenario is available a priori, including information such as the azimuth of the helicopters, the distance of the helicopters from the sensor array, their approximate speed and whether they are approaching the array or are going away from it. In a real use of the diagnosis system, such input scenario information, which does not need to be very accurate, can be provided by a more robust but less accurate localization system, such as one based on the power and frequency tracks computed by the first system of this thesis.

The signal abstractions used in this thesis were defined to agree closely with the high level features, in terms of which human subjects think about signals in our acoustic signal processing application. Human subjects are people experienced in processing and interpreting a specific type of real data, in our case acoustic helicopter signals, who are being observed while interactively performing that task. This domain-specific approach suggests that the initial step in building a real-world signal processing system is to record the "thinking aloud" of human subjects, or

at least informally focus on the thinking of human subjects, both for defining appropriate signal abstractions and for extracting heuristic rules that humans use in problem solving.

In our two case studies, it turned out that the signal abstractions our human subject uses are heavily influenced by the theory underlying the phenomenon, the theory that is related to the signal formation process as well as the theory related to the signal processing involved. As a consequence, our signal abstractions are abstractions in the Fourier domain (spectra of one-dimensional acoustic signals and two-dimensional wavenumber spectra), since Fourier theory is the basis of both of our case studies.

In our helicopter signal tracking system, signal abstractions are computed from the low-level signal. Computation of abstractions of acoustic spectra from the corresponding numeric spectra raises issues related to extraction and classification of signal features at various levels of abstraction. Numeric thresholds involved in classification are set through a training procedure, which is facilitated by the rule implementation of the classification criteria.

In the diagnosis system, signal abstractions are externally provided or generated from other signal abstractions by mappings expressed in the Fourier domain. Previously proposed causal analysis techniques for diagnosis are based on the analysis of intermediate data states that have been saved either during system operation [Hudlicka & Lesser 1984] or have been regenerated through simulation [Davis1985]. Either approach is not very suitable for the diagnosis of our direction determination system or other complex signal processing systems, because the amount of intermediate data that has to be analyzed can become prohibitive.

1.2 Related previous studies

One problem, in which progress has been made through the use of concepts similar to signal abstraction, is stereopsis [Marr1982]. The two eyes form slightly different images of the world. The relative difference in the positions of objects in the two images is called disparity, and is caused by the differences in their distance from the viewing eyes. Our brains can measure this disparity and use it to determine the relative distances of objects from the viewer. The computational theory proposed by [Marr1982] proposes three steps in measuring stereo disparity: first, a particular location on a surface in the scene must be selected from one image, second, the same location in the other image must be identified, and third, the disparity between the two corresponding images must be measured.

The difficult problem here is to find matching locations in the two images. Marr proposes the use of image elements that correspond to things such as surface markings, shadows, discontinuities in surface orientation, and so forth. Higher level elements such as these allow physical constraints to be translated into matching constraints, which restrict the allowable ways of matching image elements. Three constraints are proposed: compatibility, uniqueness and continuity.

Marr goes on to suggest that several versions of an image, each smoothed by a lowpass filter of a different bandwidth, can be computed and used for stereo matching. Low-bandwidth versions have less matchable features, thereby easing the task of search for corresponding features between the two images, at the expense of reduced disparity resolution. High-bandwidth versions can then be used for finer disparity resolution. Furthermore, by restricting the number of possible disparities being looked for, he and his colleagues proposed algorithms for stereo matching with successful results on random dot stereograms (pairs of images composed of dots which are slightly displaced with respect to their counterpart in the other image of the pair. If viewed through an appropriate optical device, they give the impression of depth).

A related piece of work is scale-space filtering [Witkin1984]. The idea here is to describe an one-dimensional signal in terms of the zero-crossings of the second derivative of multiple versions of it, each obtained by low-pass filtering with filters of varying bandwidth. Scale-space is a two-dimensional space, one dimension being the same as that of the signal and the other being the varying bandwidth of the low-pass filter (scale). Zero crossings of the second derivative of low-pass filtered versions of the signal form contours in scale-space. These contours are claimed to "precisely localize large-scale events and effectively manage the ambiguity of descriptions at multiple scales". Multiple scales can be thought of as multiple levels of abstraction, because different degrees of smoothing incur different degrees of loss of detail. A problem with the scale-space philosophy is that important events are not always large-scale events. On the contrary, they get smoothed out at large smoothing scales. Furthermore, large smoothing scales introduce shifting and merging of events, thereby complicating the task of association of events at different scales that are due to the same underlying cause.

A third problem, which has been approached with a notion resembling signal abstractions, is correlation of well logs by matching curve elements [Vincent et al 1979]. Correlation of well logs from two nearby wells or from different azimuths of the same well is an important step towards the description of the subsurface of an oil field. Well log correlation attempts to recognize and match similar details on curves that are similar, but not identical. The output consists of dip as a function of depth. Dip is the angle of the line connecting similar details with respect to the horizontal (well logs are functions of depth and are drawn vertically). Although numeric correlation measures have been applied to this problem, they are deficient because they do not allow explicit expression of the causes of similarity of well logs and because only similar features of the same thickness can produce a high value for correlation.

In [Vincent et al 1979] correlation of well logs relies on matching of features of elements of well logs. Elements are of five possible types: peaks, troughs, spikes,

steps and levels. The first phase of the method is element extraction. Each curve is decomposed into a sequence of elements. Each element is assigned a pattern vector, which is a series of measurements further characterizing the shape of the element.

The second phase is correlation among similar elements. The basic criterion used is comparison of pattern vectors. Pattern vectors are likely to be better measures of similarity of elements in well logs than some uniform numeric correlation measure, because they concentrate on the elements' important features. Numeric correlation, on the other side, cannot easily differentiate between important and less important features.

A likeness coefficient is computed for each pair of elements from the two well logs being correlated. The larger elements (higher peaks and deeper troughs) are considered first and the pairs with the highest likeness are established as correct correspondences. Established correspondences are then checked as to whether they satisfy additional geologically derived criteria. Finally, trigonometric calculations on the established correspondences between elements are used to compute dips as a function of depth. Dip measurements thus obtained can be used to determine geological formations.

An attempted solution to a fourth problem, that of speaker separation of voiced speech [Parsons1976], can be interpreted as involving signal abstractions. Speaker separation involves suppression of a common type of interference in speech applications, which is caused by the speech of a competing speaker. The human brain is adept at clarifying such speech, but it relies heavily on binaural data. Although Parsons' system had a very heuristic flavor to it and was not phrased in the context of signal abstraction, it can be interpreted as an example of a system involving interference suppression by applying constraints to an appropriate abstract speech representation. The representation used was an abstraction of the short-time spectrum.

The spectrum of voiced two-talker speech has two sets of harmonics, which can

be identified and used to reconstruct the two speech signals from each of the two talkers separately by inverse transforming each signal's share of the spectrum. The major steps of the process are shown in Figure 1.1.

```
                      mapping
    peaks ─────────────────────────▶ peaks
      ▲                                │
      │                                │
  abstraction                      refinement
      │                                │
      │                                ▼
   spectrum                        spectrum
      ▲                                │
      │                                │
      │                                ▼
   speech                           speech
```
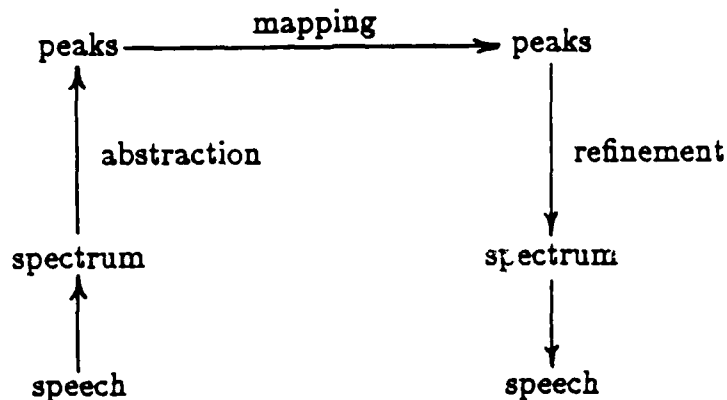
Figure 1.1: Major steps in voiced speech separation using the spectrum

The first step (abstraction) is to compute Fourier domain abstractions of the time waveform up to a level of abstraction, at which assumptions about the nature of the signal, such as continuity or consistency of certain characteristics, can be applied to achieve the goal of the processing (separation of two signals in this case). This level of abstraction is that of harmonically related sets of spectral peaks, one set per speaker. Several problems are encountered in computing such an abstraction. First, spectral peaks have finite width, both because pitch varies over time and because of the finite length of the segments used. This results in peaks from different speakers overlapping or merging into a single peak. Second, determining the pitch of each talker cannot always be done reliably, especially when the pitch periods are very close together.

The second step (mapping) is to use the assumption of continuity of pitch to derive Fourier domain abstractions of the separated signals from the Fourier domain abstraction of the input signal.

24

The final step (refinement) is to lower the level of abstraction of the resulting signals down to the time waveform level, i.e. compute the resulting low level signals by following the reverse procedure of the abstraction step.

Peaks are represented in terms of peak model parameters. A peak model is used which accounts for broadening of spectral peaks, which occurs due to windowing of the time signal with a hanning window, and for the phenomenon of pitch glide, namely the rate of change of pitch over time, which is assumed linear. The peak model is used to separate peaks that seem to overlap. Several heuristic criteria are used to test for overlap. Once overlap is suspected, peaks are separated by fitting a two-peak model to the composite peak. In order to determine pitch, use is made of the constraint that each talker's harmonics form a set of harmonically related peaks.

After having separated overlapping peaks and correctly grouped peaks into harmonic sets, speaker separation has in principle been accomplished. However, the objective is to produce two time series, each corresponding to a speaker. The final task is to refine the harmonic set descriptions of the two signals into time series representation. This problem was given a straightforward solution through the use of the assumed peak model. The four peak parameters, amplitude, phase, pitch glide and frequency are used to synthesize the peaks of the spectra of the windowed speech segments. After the spectra have been formed as the sum of their peaks, they are transformed back to the time domain to obtain (overlapping) segments of the output speech. Finally, the overlapping segments are added together to produce continuous speech output.

Unfortunately, the speaker separation system did not work very well in its limited domain of separation of two-speaker voiced speech. However, we feel that the above interpretation of the work as a sequence of signal abstraction, mapping and refinement may prove useful in other signal processing problems involving signal enhancement.

Finally, work in shape analysis of intensity profiles of images has led to a number of approaches for one-dimensional shape descriptions, which may be viewed as related to our notion of signal abstractions. [Ehrich & Foith 1976] and later [Sankar & Rosenfeld 1979] proposed hierarchical tree descriptions of waveforms based on dominance relations of peaks and valleys. Although elegant, such descriptions do not always yield a tree corresponding to the most natural description of a signal, or they are much too complicated to be useful in the case of noisy signals. The work by [Cheng & Lu 1985] on well log correlation by tree matching is a practical application of the relational tree approach of [Ehrich & Foith 1976], and has been claimed to work well even in cases with significant stretching or shrinking between well logs.

A different approach is linguistic analysis of waveforms [Lozano-Perez 1977, Pavlidis & Ali 1979]. In this approach, the waveform is segmented into elements, which are then viewed as a sequence of "words". Local shapes of interest are then detected by "parsing" this sequence of words using a different grammar for each shape. A difficulty associated with the linguistic approach is the need for smoothing to eliminate noise, which introduces the problem of how to select an appropriate smoothing scale that preserves the essential characteristics of the signal. Shape grammars can be embedded to form a hierarchy, as pointed out by [Anderson 1982], in which case the result can be viewed as describing waveforms at multiple levels of abstraction.

1.3 Overview of the thesis

The thesis begins in chapter 2 with an overview of a specific acoustic processing problem that was used as a catalyst for exploring signal abstractions. Multichannel waveforms are obtained from an acoustic microphone array monitoring the sounds emitted by various sources in the vicinity of the array. This array data can be processed either as individual one-dimensional waveforms or as multichannel data

for source detection and direction determination through array processing. Power and frequency characteristics of the signals from the sound sources may be computed by estimating the spectra of individual channels at different times.

In chapter 2, we review the characteristics of helicopter signals and the issues associated with estimating their spectra. Then we address the problem of localization of helicopters from power and fundamental frequency traces obtained from geographically separated microphones. This analysis motivates the material in chapters 3 and 4, which focus on the computation of power and fundamental frequency traces from real helicopter data. An alternative existing localization technique based on array processing is then briefly reviewed. According to this technique, the azimuth of an acoustic source with respect to two geographically separated microphone arrays is computed and the two azimuth measurements are used to localize the source. The sensitivity of the azimuth determination system to improper parameter adjustment motivated the research presented in chapters 5 and 6. Finally, protocol collection and analysis, as it was adapted for recording the problem solving activity involving interactive signal processing, is reviewed. Analysis of the collected protocols was used to help define appropriate signal abstractions for the problem.

In chapter 3, a set of signal abstractions, which we call an extended spectrum, is defined for acoustic spectra characterized by collections of harmonically related peaks. The extended spectrum is defined by three levels of abstraction, the numeric spectrum, the list of peaks of the spectrum and the list of harmonically related peak sets. Computation of the higher levels is based on the immediately lower levels and is achieved by algorithmic procedures, the results of which are pruned and/or modified by collections of heuristic criteria, implemented in rule form. Collections of parameters needed for these transitions are introduced as part of the extended spectrum. Such parameters may include thresholds, which can be adjusted by a training procedure, whereby the human specifies which of the rules of the system are applicable to particular training situations. A program then adjusts the related thresholds to ensure that the preconditions of such rules are satisfied by the training

27

situation at hand.

In the second part of chapter 3, examples of the use of the extended spectrum are presented in three problems involving spectral matching. The first problem is detection of faulty channels based on their spectra. The second problem is adjustment of the parameters of a specific spectral estimation technique by matching spectra obtained from different parameter setups. The third problem is the representation of the temporal evolution of acoustic spectra by forming explicit links between harmonically related sets of peaks over time.

Chapter 4 presents a system for helicopter pitch tracking, which is built on the ideas of chapter 3. Results of its performance on several real data scenarios of total duration more than 20 minutes are presented. Feedback is involved in the completion of gaps or extension of the linked sequences of harmonically related peak sets by focusing the search aided by the links already established. By concentrating on the explicit representation of the relationships between spectra at different times, pitch and power tracking come as a natural consequence.

Moreover, the generated pitch and power tracks are based on harmonic sets at consecutive times which are explicitly linked with each other to form chains. These linked harmonic sets are accessible for inspection, thereby facilitating the process of tracing down discrepancies between the results of the program and the visual experience of the user. The behavior of the system can then be improved by modifying existing heuristics or adding new ones for dealing with particular discrepancies as they are identified during testing of the system on real data.

Chapter 5 presents the signal abstractions and components and Chapter 6 the overall design of a diagnosis system for a complex signal processing system for determining the azimuth of helicopters from multichannel microphone signals. The faults that are being diagnosed are related to mismatch between parameters of the signal processing system and the scenario generating the input multichannel signals from which azimuth determination takes place possibly including propagation phe-

nomena that distort the acoustic waves collected by the microphones. The input to the diagnosis system consists of an approximate description of the scenario generating the multichannel signals, including azimuth, power and frequency content of the helicopter sounds, distance of the helicopters from the microphones and speed, and of a description of the actual output of the signal processing system. Both the scenario and the output of the signal processing system are described as wavenumber spectra, a convenient representation of direction and frequency information of acoustic plane waves impinging on a sensor array. The output of the diagnosis system is an explanation of the differences between the scenario and the output of the signal processing system. The explanation is in the form of a sequence of mappings of wavenumber spectra which introduce distortions to their input wavenumber spectra. The cumulative effect of those distortions causes the output of the signal processing system to be different from the acoustic source scenario, for example showing less sources than those that are actually present (a resolution problem) or azimuths different from the true ones (a direction shifting problem).

In Chapter 5 we introduce the notion of mappings of wavenumber spectrum abstractions as a mechanism for representing the operation of the direction determination system based on multichannel array processing and the associated wave propagation effects. Mappings are not necessarily associated with a single component of the signal processing system, but are related to effects that might be attributed to more than one component. For example, a resolution problem may be attributed to improper array aperture, proximity of the azimuths of two helicopters and overlap of their temporal spectra. In the same chapter we address matching of wavenumber spectrum abstractions to identify their differences, ordering multiple differences in order of importance, and associating a difference with a set of mappings that may be able to explain it.

Chapter 6 presents the architecture of the diagnosis system, which can be viewed as a search through the space of all possible sequences of mappings for a sequence that explains the differences between the scenario and the output of the signal pro-

cessing system. The diagnosis system hypothesizes a mapping that explains some of the overall differences, thereby recursively reducing the top-level problem into one or more subproblems of explaining the remaining differences. Verification of the hypothesized explanations takes place, which causes explanations to be modified or rejected, if the verification fails. Diagnosis is initially attempted at the highest level of abstraction, while verification takes place at the lowest possible level. If a verifiable explanation cannot be found at the highest abstraction level, the search is repeated at a lower abstraction level. An example of the operation of the implemented diagnosis system is shown to illustrate the main points of the work.

Chapter 2

An acoustic signal processing and interpretation problem and the associated knowledge

To explore the issues associated with signal abstractions in signal processing systems we selected as a catalyst a particular problem involving acoustic signals, for which there are people with experience in processing and interpretation of the signals. The problem is to build a system that takes as input a set of signals from microphone sensors. The sensors are placed on the ground and monitor sounds in their environment. The goal of the system is to produce interpretations for the acoustic sources (such as aircraft) present in the area within the acoustic range of the sensors. Source interpretations consist of detection (i.e. whether there are sources present), identification and localization.

In this chapter we review knowledge about the relationships between source characteristics and their manifestation in the acoustic signals. Such knowledge is derived from the theory of signal formation and propagation, from experiments with real data and from formal observation of experienced persons while they interac-

31

tively process and interpret real signals. Later chapters of this thesis build upon parts of this knowledge and focus on the various ways it has been represented in the signal processing programs of the thesis.

Sections 1 and 2 of this chapter provide background material for helicopter detection and localization using the spectra of one-dimensional microphone signals. Source localization based on power and frequency traces from multiple microphones is examined in section 2 and provides the motivation for the material in chapters 3 and 4, which focus on the extraction of such traces from real helicopter data. Section 3 briefly reviews an existing alternative localization technique based on azimuth measurements obtained by array processing of multichannel data. This azimuth measurement system served as the starting point for the system presented in Chapters 5 and 6, which identifies incorrect parameter settings in test runs of the azimuth measurement system. Although currently the two localization systems are independent, we envision an integrated system in which robust, but not very accurate, localization from power and frequency traces will assist the system for identifying incorrect parameter settings to ensure proper operation of the azimuth measurement system. The latter system is capable of providing accurate azimuth measurements, but the correctness of its results are sensitive to improper parameter settings.

2.1 Characteristics of acoustic waveforms from a single sensor

The main causes of helicopter sound are aerodynamic in nature and they are a result of the fluctuations and motions of the distributed pressures on the blades. These pressures are in turn due to the rotor blade/wake interactions. Since the wake is assumed to be periodic, the resulting sound is periodic and hence its spectrum is characterized by a fundamental frequency and its harmonics.

There are two such harmonically related peak sets present in helicopter sound, due to the main rotor and tail rotor. The tail rotor fundamental is usually several times higher than the main rotor fundamental, and by design not a multiple of it. Typically, the main rotor fundamental ranges between 8 and 20 Hz, with harmonics up to about 100 Hz. The tail rotor fundamental ranges between 70 and 140 Hz, with harmonics up to about 800 Hz. The experimental acoustic data used in this thesis was sampled every 0.5 ms, corresponding to a sampling frequency of 2048 Hz.

The pressure disturbances at the blades of a flying helicopter act as generators of acoustic waves, which can be considered spherical to within a reasonable approximation.

In Figure 2.1 we show a typical scenario from a real experiment with two helicopters travelling with approximately the same speed along the same straight line path and with a distance between the two helicopter equal to about 45 seconds multiplied by the helicopter speed. Figure 2.2 shows parts of the acoustic waveform collected by the sensor in Figure 2.1 and the corresponding spectra. Part (a) of Figure 2.2 shows the time waveform over 100s (the sampling rate being 2048 Hz). The amplitude of the acoustic waveform is highest at the closest point of approach (CPA) of the first helicopter. The CPA of the second helicopter is outside our data window. Under the assumption of a point source and preservation of energy of the acoustic wave, the power (amplitude squared) of the acoustic waveform is inversely proportional to the square of the distance between the sensor and the source.

Parts (b),(c) and (d) of Figure 2.2 show portions of the acoustic waveform at a small scale (1s), the corresponding spectrum over the frequency range 0-1024 Hz and the same spectrum expanded and rescaled over the low frequency range 3-80 Hz (first, second and third row respectively. For display purposes, the low-frequency spectrum is shown in linear scale, whereas the spectrum over 0-1024 Hz is shown in logarithmic scale). Part (b) is a section when the first helicopter is significantly before CPA, part (c) is a section very close to the CPA of the first helicopter, and

33

part (d) is a section before the CPA of the second helicopter and after the CPA of the first helicopter, at a time when the signal from the first helicopter has almost died away.

In (b) the time waveform is very noisy, i.e. has a low signal-to-noise ratio, and does not show a clear periodicity. The logarithmic spectrum shows a clear set of harmonics of about 94 Hz, which corresponds to the tail rotor harmonics of the first helicopter. The linear spectrum shows clearly the first two harmonics of 14 Hz, a strong peak at 35 Hz and smaller peaks at 44, 54 and 70 Hz.

In (c) the time waveform shows a clear periodicity of about 13 periods in a second, i.e. 13 Hz. The linear spectrum therefore, has a strong peak at 13 Hz and two strong harmonics at 26 and 39 Hz, corresponding to the main rotor. The logarithmic spectrum does not show a clear harmonic structure anymore in the frequency range above 80 Hz.

In (d) the time waveform is very noisy and does not show a clear periodicity. The logarithmic spectrum shows distinct harmonically related peaks, corresponding to a fundamental of about 95 Hz. Therefore, they belong to the tail rotor of the second helicopter, since the tail rotor of the first helicopter is now well below 90 Hz, due to Doppler shift. The linear spectrum shows harmonics of 12 Hz, due to the main rotor of the first helicopter, as well as harmonics of 15 Hz, due to the main rotor of the second helicopter.

Apart from the signal-to-noise ratio, we note that the frequencies of the spectral peaks change over time in the example of Figure 2.2. Frequency shift is due to the motion of the source relative to the propagation medium. According to the Doppler phenomenon, the frequency of a sinusoidal component of the sensor signal is equal to the frequency of the corresponding component of the acoustic source scaled by the factor $\frac{1}{1-v/c}$, where c is the speed of sound and v is the speed of the source along the direction of propagation, with $v > 0$ if and only if the direction of v coincides
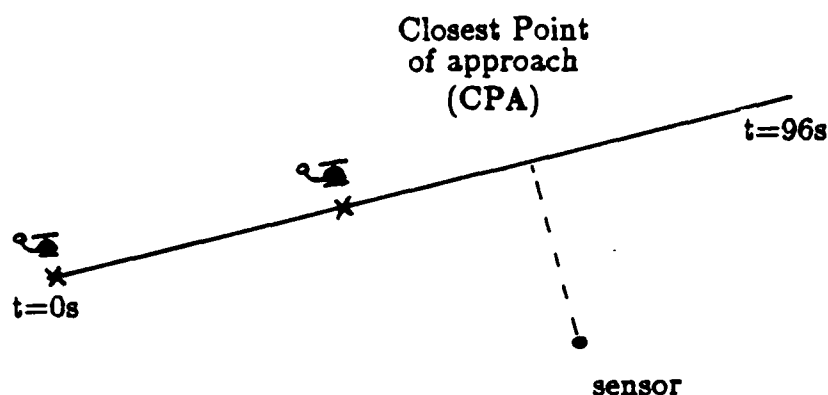
Figure 2.1: A straight source path and two helicopters one following the other. Associated time waveforms and spectra are shown in the next figure.

with the direction of propagation. Figure 2.3 shows the frequency at the sensor as a function of time for a straight source path and different source speeds and CPA distances from the sensor.

The above observations on the time waveforms and power spectra of helicopter signals suggest that detection of helicopters within the acoustic range of a sensor may be assisted by search for harmonically related peaks in the power spectrum. Furthermore, the relation between power of the spectral peaks and distance between the source and the sensor and the relation between the frequencies of the source and those of the sensor signal may help in source localization. These relations, however, are crude approximations, while the accurate modelling of helicopter signatures is an extremely difficult problem, because the total sound radiated by helicopters is a composite of rotor noise, engine noise and drive-system/gear box noise. Moreover, the rotor noise is classified into three major categories: rotational noise due to the time varying pressure distributions on the blade, broad-band noise due to turbulence phenomena at the rotor blades and blade slap. The relative importance of these categories in the far field depends on the rotor tip speeds, blade slap being the most
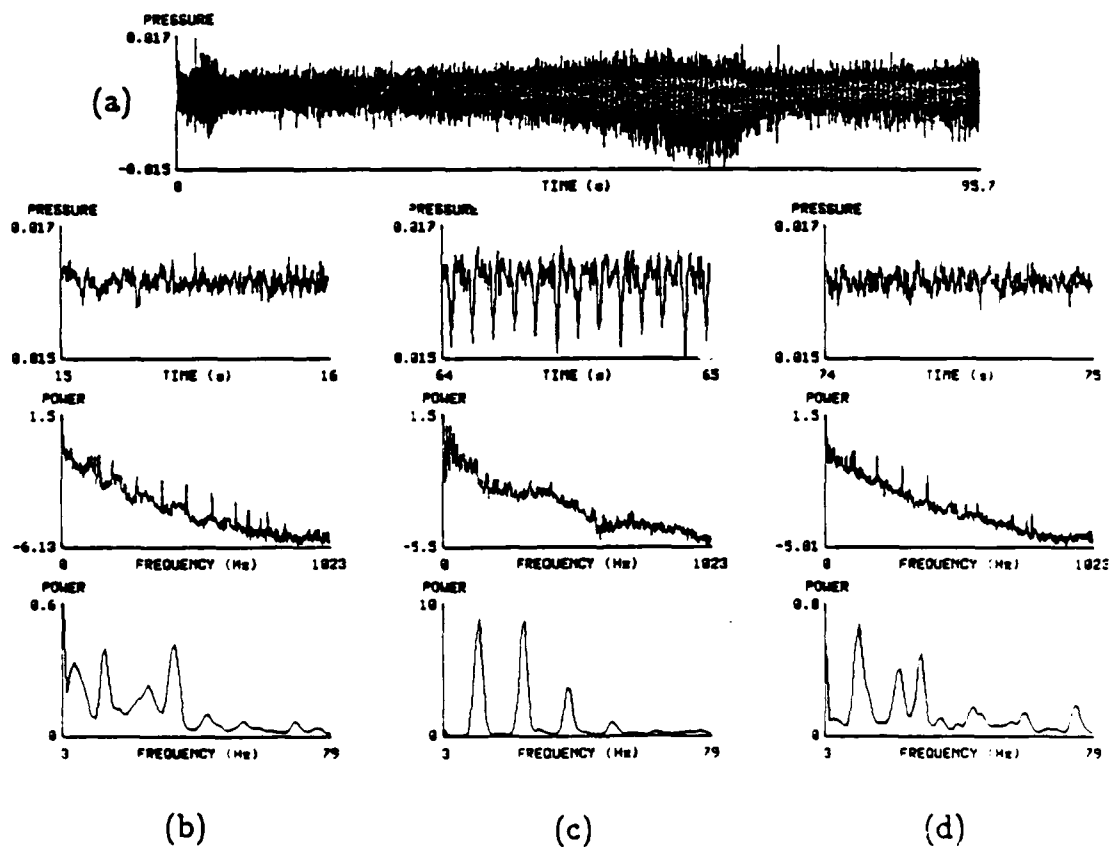
Figure 2.2: Time waveforms and power spectra of a two-helicopter signal and for a straight path shown in the previous figure.
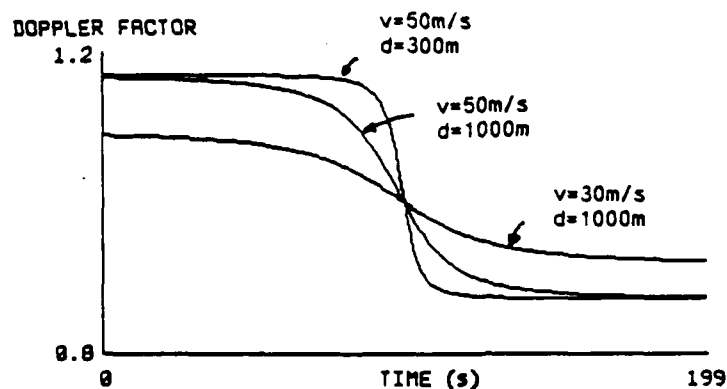
Figure 2.3: Sensor signal frequency as a function of time for a straight source path and several source speeds and CPA distances from the sensor.

prominent at low to moderate tip speeds [Hubbard & Harris 1981]. The complexity of helicopter sound is further compounded by atmospheric propagation effects, which depend on conditions such as atmospheric temperature, pressure, humidity, and wind speed and direction, as well as environmental noise from tree leaves and ground vehicles. In spite of these difficulties in accurately modelling helicopter data, in many cases humans are capable of interpreting helicopter signals by combining interactive signal processing and consistency checking of their hypotheses based on the available crude models.

Estimation of spectra of harmonic acoustic signals The initial step towards determining power and fundamental frequency traces of acoustic signals is estimation of the spectra of these signals. There is a wide variety of techniques for estimating the spectrum of a signal [Kay & Marple 1981]. Specific techniques also depend on the choice of one or more parameters. Selection of a spectral estimation technique together with values for the associated parameters involves various tradeoffs and therefore should be based on the characteristics of the signal. In the rest of this section we briefly review some of the tradeoffs associated with the method

37

used for estimating helicopter spectra, which consists of averaging periodograms of consecutive frames of the time signals.

Essential requirements for any spectral estimation technique are ability to resolve closely spaced peaks and ability to reduce noise in the spectrum [Oppenheim & Schafer 1975]. The variance of the spectral estimate is a measure of how much noise there is in the spectral estimate. In the case of real signals noise is due to additive noise sources (for example tree noise because of wind in acoustic helicopter data) as well as changes to the signal itself (such as Doppler shift, nonuniformity of the propagation path because of wind or variable atmospheric conditions along the path). For periodogram-based spectral estimation, the variance of averaged periodograms is inversely proportional to the number of periodograms for statistically independent periodograms. Ability to resolve closely spaced peaks increases with the length of data used to compute a single periodogram. Long data implies windowing with a long temporal window, and this in turn implies convolution with a narrow spectral window, thus providing high spectral resolution. In spectral estimation applied to real data various tradeoffs are present.

For fixed data length (and contiguous non-overlapping blocks), there is a tradeoff between low variance of the estimate (requiring many, therefore short blocks of data) and high resolution (requiring long, therefore few blocks of data).

For fixed window length and variable total data length a different tradeoff is present, if peak frequencies shift over time. Frequency shifting causes peaks from different periodograms to only partially overlap, therefore giving rise to a broad peak with insufficient accentuation when added together or averaged. This undesirable phenomenon limits the total data length that can be used for reducing noise. Consequently, the tradeoff in this case is between noise suppression and peak broadening accompanied by insufficient accentuation. This situation occurs due to the Doppler effect when the acoustic source is moving.

As an example of the above tradeoff, let us consider the case of a helicopter flying

38

in a straight line. The Doppler effect is most pronounced near CPA, thus limiting the length of time over which the signal can be considered stable. Near CPA is also where the signal-to-noise ratio is highest, because that is where the acoustic wave is strongest. Therefore, when the helicopter is far from the microphone and the signal-to-noise ratio is low, the signal is relatively stable, thus permitting averaging over longer pieces of data.

2.2 Source localization from power and fundamental frequency traces of sensor signals

In the previous section we raised the notion of using power and fundamental frequency traces from multiple geographically separated sensors to localize the corresponding acoustic sources. In this section we explore the idea in greater detail by assuming that power and fundamental frequency traces have already been computed from the data and we present geometrical arguments for source localization. This geometric analysis motivates the next two chapters, which examine the issues in extracting reliable power and fundamental frequency traces from real helicopter data.

In source localization from power and fundamental frequency traces we make several assumptions, which may not be exactly satisfied in practice, but may serve as a source of heuristics for guiding the signal processing and interpreting its results. A set of assumptions relates to acoustic wave propagation: (a) power of the sensor signal is inversely proportional to the square of the distance between the source and the sensor and (b) sensor frequencies are related to source frequencies by the factor $\frac{1}{1-v/c}$. Another set of assumptions applies to the source motion: source speed is assumed to be constant and the source path is assumed to be piecewise linear. Finally, we are concerned with only a single source, to avoid the problem of consistent association of harmonically related peak sets with sources over time.

A further issue is propagation delay of the acoustic waves. If the acoustic source is not moving, a sensor experiences constant delay, which is proportional to the sensor's distance from the source. For a moving acoustic source, the delay experienced by a sensor is proportional to its instantaneous distance from the source, thus introducing a time-varying warping of the time axis at the sensor and complicating the localization process.

In the following two subsections we address source localization in two special cases, in order to give the flavor of the issues involved. In the first case, we neglect propagation delay, but we allow the source to maneuver in a piecewise linear fashion with constant speed. In the second case, we take into account propagation delay, but we restrict the source to move in a straight line with constant speed.

2.2.1 Source localization assuming negligible propagation delay

Propagation delay can be neglected if the distances between sensors are small enough to imply negligible propagation time of sound between pairs of sensors.

As an initial step, we concentrated on whether power and fundamental frequency increase, decrease or remain constant over time and we developed a qualitative vocabulary for describing portions of the source path in their position relative to the sensor, as shown in Figure 2.4.

Figure 2.4 shows various cases involving the local shape of the acoustic source path in the neighborhood of a landmark point combined with the relative positions of the path and the sensor. A landmark point is a point on the source path, at which the tangent to the path is perpendicular to the line connecting the sensor and the landmark point (Figure 2.5). Landmark points are important because at those points the Doppler shift becomes momentarily zero, and the frequency of the
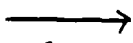
	Fund. Frequency	Power	
before FPA	increasing	decreasing	
after FPA	increasing	increasing	
before CPA	decreasing	increasing	
after CPA	decreasing	decreasing	
aligned before CPA	flat	increasing	
aligned after CPA	flat	decreasing	
far from sensor	flat	flat	

Figure 2.4: Correspondence between the sign of temporal derivative of power and fundamental frequency and the local character of the source path, assuming constant speed for the source.

sensor signal is the same as the frequency emitted by the acoustic source. Landmark points are either closest points of approach (CPA) or "farthest" points of approach (FPA). A landmark point is a closest point of approach if the center of curvature of the path at the landmark point is farther away from the path than the sensor is. An example is the case of a straight line path, where the center of curvature is at infinity. A landmark point is a farthest point of approach if the center of curvature of the path at the landmark point is closer to the path than the sensor is. This implies a small radius of curvature, or equivalently a highly curved source path, implying that the acoustic source is taking a turn or making a maneuver.
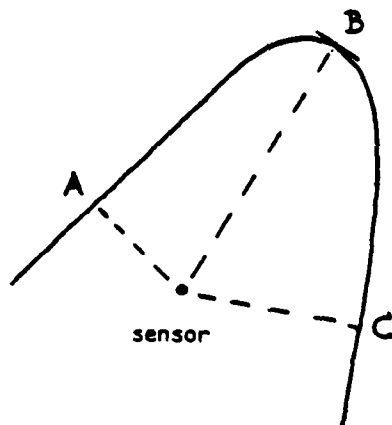
Figure 2.5: Landmark points of a source path. Points A and C are closest points and B is a farthest point of approach.

As an example, consider the case "before-CPA". In this case, the power of the sensor signal increases with time, because the source comes closer to the sensor. Frequency decreases because the speed of the source along the direction of propagation is positive and decreases, approaching the value zero at CPA.

42

The vocabulary shown in Figure 2.4, combined with the assumption of negligible propagation delays between sensors, helps in forming a qualitative characterization of the source path independent of the sensors. In principle, it is possible to parse a given pair of power and fundamental frequency traces and qualitatively characterize portions of the source path, as they are seen from a single sensor. Such parsing can be done at more than one sensor, leading to a qualitative characterization of the source path independent of the sensors by combining the results. Then we concentrate on the time intervals over which the source path is linear, and use power and fundamental frequency values to locate the linear source path, assuming that the problem is planar. This is the goal of the system that is described next.

Parsing power and fundamental frequency traces. The first step towards using the vocabulary shown in Figure 2.4 is to segment the power and fundamental frequency traces from each sensor (obtained using techniques such as the one presented in the next two chapters of the thesis) according to the sign of their derivative, which we approximate by the first difference of the traces. In the case of a piecewise linear source path, sharp changes in direction give rise to increasing frequency and a power minimum over the duration of the maneuvers in the traces from those sensors for which the maneuver portion of the path shows an FPA-type landmark point (see Figure 2.4). Therefore, segmenting the power and fundamental frequency traces and combining such information from multiple sensors allows us to determine the time intervals over which the source is maneuvering, as the intervals over which at least one sensor shows frequency increasing with time.

To further explore these ideas, we built a simulation system, which allows manual input of the location of the sensors and samples of the source path, and computes the power and frequency traces for each sensor based on the idealized models for power and frequency. We also built a system that first performs segmentation of the frequency traces from several sensors arranged uniformly along a straight line, then identifies the maneuvering intervals as those over which the frequency in at

43

least one trace has positive slope, and finally uses geometrical calculations on the power traces over the straight-line intervals (computed as the complement of the maneuvering intervals) to localize the straight line portions of the source path. The maneuvering portions of the source path are filled in by using spline interpolation as an initial approximation. Therefore, because of our simplifying assumption of a piecewise linear source path, there is no need in our system to use combined information from both power and frequency traces to characterize portions of the source path, as shown in Figure 2.4. The overall structure of the localization system, with its input obtained from the simulation system, is shown in Figure 2.6.
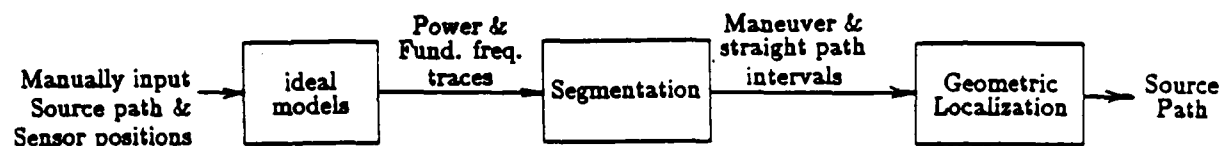
Figure 2.6: Simulation system for source localization based on power and frequency traces from multiple sensors, with manual input of the source path and sensor positions.

Segmentation of power and frequency traces is not a straightforward task due to the presence of noise. In our simulation system, noise is due to the sampling of the source path (and linear interpolation between the samples) and the inaccuracy of the manual input (e.g. a straight line path is as straight as the human eye and hand can make it). In a system operating with real data, additional noise results from more complicated signal generation and propagation effects and from measurement errors in computing power and frequency from power spectral estimates of the sensor

signals.

We perform segmentation in four steps (Figure 2.7). The first step finds all time intervals over which the slope of the power and frequency traces is large or small, marking them as sloped or flat respectively. The second step splits all sloped intervals into subintervals over which the corresponding traces are monotonic and marks them with the label UP or DOWN. In the third step, we merge short intervals with their neighbors, if appropriate. Figure 2.8 shows the rules used for merging and sketches of examples of their preconditions. The merging rules apply on existing intervals and replace them by merged intervals. The rules apply repeatedly until no rule applies any more. In the fourth and last step, we combine time intervals obtained from power and frequency traces of multiple sensors to compute the time intervals over which the source is maneuvering, using the assumption of negligible propagation delay between sensors. Their complement consists of the time intervals, over which the source path is a single straight line segment.
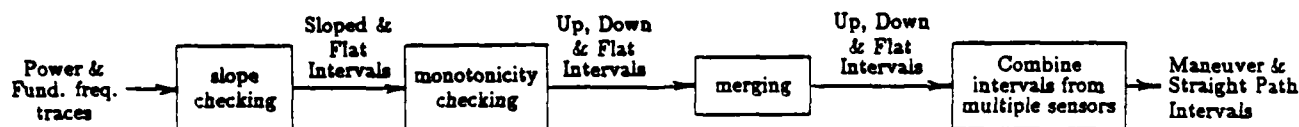
Figure 2.7: The overall segmentation process

Geometrical considerations for localization of straight source path segments from power and frequency traces   With sufficient sensor coverage of

45

```
If:    - the middle interval is flat, i.e. it has small slope,
  and - the left and right intervals have the same slope sign
        (both increasing or both decreasing with time)
  and     - the middle interval is very short
        or - its small slope has the same sign as that of the
              left and right intervals,
Then:  merge the three intervals into one with the same slope
       sign.



If:    - the left and right intervals are flat
  and  - the middle interval is short
  and  - it has a moderate maximum slope,
Then:  merge the three intervals into a single flat
       interval.



If:    - the left and right intervals are flat
  and  - the middle interval is sloped
  and  - the signs of the  small slopes of both the left and
         right intervals are the same with that of the middle
         interval,
Then:  merge the three intervals into one with the same slope
       sign.

If:    - two adjacent intervals have the same slope sign
Then:  merge them into a single interval with the same slope
       sign.
```

Figure 2.8: Rules for merging monotonic intervals of either power or fundamental frequency traces. Each rule (except the last one) operates on a set of three consecutive intervals. Intervals are characterized by the sign of the average slope which is a symbolic characterization UP, DOWN or FLAT. Some rules go to a lower level of detail and have tests that rely on the maximum, minimum and average slope of a trace over an interval. The merging process has two phases. In phase 1, the second and third rule are applied. In phase 2, all rules are applied to the intervals resulting from phase 1.

the area under monitoring, and with sufficiently long straight source path segments, we can ensure that there will be at least two sensors exhibiting a closest point of approach for each straight source path segment. To achieve localization of such a segment, we concentrate on such a pair of sensors, as shown in Figure 2.9.

Using the assumption of negligible propagation delay between the two sensors, we can measure various ratios of distances of the source from the two sensors at various points in time by finding the corresponding power ratios at that time. Figure 2.9 shows the geometry for a generalized triangulation method for finding the exact location of the straight source path segment to within a 180-degree ambiguity and the ambiguity as to whether the path crosses or does not cross the segment defined by the two sensors. If more than two sensors display a CPA for the source path segment in question, we can use more than one pair to reduce the ambiguity and increase the accuracy of the solution.

In Figure 2.9, we assume that the distance $(A_1 A_2)$ between the sensors A_1 and A_2 is known and that the ratios $\mu = \frac{(A_1 C_1)}{(A_1 C_2)}$ and $\nu = \frac{(A_2 C_2)}{(A_2 C_1)}$ can be measured from ratios of powers of the power traces from the two sensors at the CPA times. Note that the ratios are related to powers of only one trace, therefore there is no need for the two sensors to be calibrated (i.e. to give the same signal power for the same acoustic wave). From the Figure we see that $\mu = \sin \phi$ and $\nu = \sin \theta$. Furthermore,

$$\frac{\tan \phi}{\tan \theta} = \frac{(C_1 A_1)}{(C_2 A_2)} \qquad (2.1)$$

implying that

$$\lambda \equiv \frac{(C_1 A_1)}{(C_2 A_2)} = \frac{\mu}{\nu} \frac{\sqrt{1 - \nu^2}}{\sqrt{1 - \mu^2}} \qquad (2.2)$$

Therefore,

$$\lambda = \frac{(C_2 H)}{(C_2 A_2)} \Rightarrow \frac{(C_2 H)}{(A_2 H)} = \frac{\lambda}{1 - \lambda} = \tan \phi \qquad (2.3)$$

and

$$\tan \psi = \frac{(A_2 H)}{(A_1 H)} = \frac{1 - \lambda}{\lambda} \tan \phi \qquad (2.4)$$

and finally all distances necessary to specify the position of C_1 and C_2 can be

47

computed as follows:

$$(A_2 H) = (A_1 A_2) \sin \psi \qquad (2.5)$$

$$(C_1 C_2) = (A_1 A_2) \cos \psi = (A_1 H) \qquad (2.6)$$

$$(C_1 A_1) = (C_1 C_2) \tan \phi \qquad (2.7)$$

$$(C_2 A_2) = (C_1 C_2) \tan \theta \qquad (2.8)$$
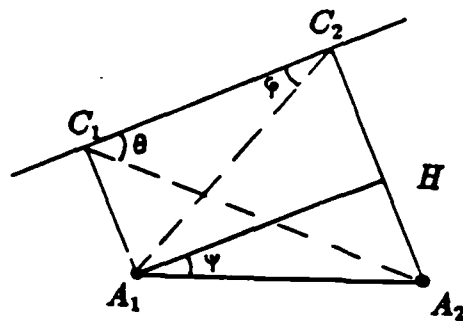
Figure 2.9: Geometry of generalized triangulation method for localizing a straight source path segment. Points A_1 and A_2 are the sensors, while points C_1 and C_2 are the corresponding closest points of approach.

If the source path segment happens to be almost orthogonal to the sensor segment, the computation of Figure 2.9 becomes very sensitive numerically. In this case, a different approach can be used (shown in Figure 2.10), which does not suffer from numeric sensitivity, but requires calibration of the two sensors, because λ cannot be calculated from μ and ν, which are both very close to 1, and therefore must be computed directly from its definition as $\frac{(C_1 A_1)}{(C_2 A_2)}$. A fundamentally ill-conditioned case, for which no robust special method has been found, occurs when the source path segment is very far from the two sensors (compared to the distance between them), as shown in Figure 2.11. In this case λ, μ and ν are all very close to 1.

48
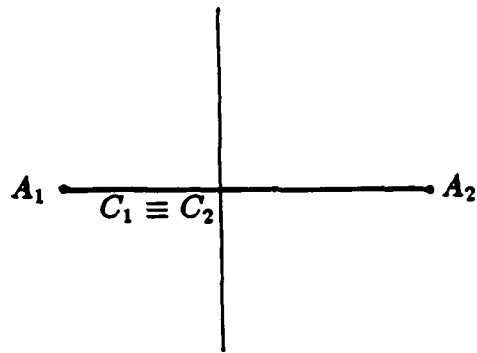
Figure 2.10: Geometry of specialized localization method in a case when the default method of the previous figure is numerically sensitive.
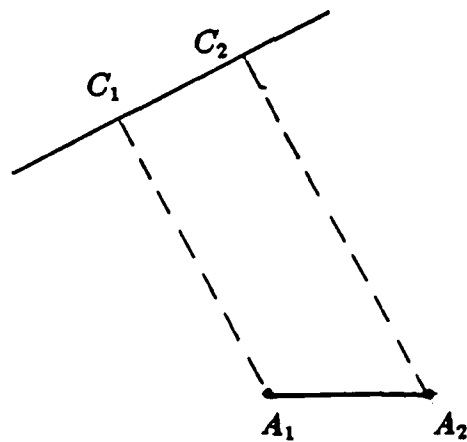
Figure 2.11: Fundamentally ill-conditioned geometry.

Examples of the simulation system   We now present examples of the operation of the simulation system previously described. The user provides the position of the sensors and samples of the source path through interactive graphics and the mouse device. The system computes power and frequency traces for each sensor using the ideal models of propagation, namely that $P \sim \frac{1}{d^2}$ and that $f = f_0 \frac{1}{1-v/c}$, where P is the power of the sensor signal, d is the distance between the source and the sensor, f is the frequency of the sensor signal, f_0 is the frequency of the signal emitted by the acoustic source, v is the magnitude of the projection of the source velocity on the line connecting the source and the sensor. For each sample of the source path, power and frequency are computed from the above formulas. For computing v, it is assumed that the direction of the source velocity is the same as that of the line connecting the current path sample with the next one.

Figure 2.12 shows ten sensors numbered 1 to 10 arranged in a straight line and several points of a single source path (small dark circles and a triangle denoting the first point) entered by the user. Figure 2.13 shows the corresponding power and frequency traces for sensors 1, 5 and 8. Based on knowledge of the location of the sensors and the traces of figure 2.13, two solutions are obtained by the simulation system for the source path, shown by continuous lines in Figure 2.12. The circle on one end of the line denotes the beginning of the path. Figure 2.14 shows a more complicated example including a maneuver portion, while the corresponding power and frequency traces for sensors 1, 5 and 10 are shown in Figure 2.15.

The previous examples of the simulation system indicate that in the case when propagation delay can be neglected or is uniform across sensors, approximate source localization is possible using power and frequency traces from multiple sensors, provided we can extract reliable measurements of power and frequency from the time waveforms.

The geometric solutions explored in this section are suboptimal in the ideal case
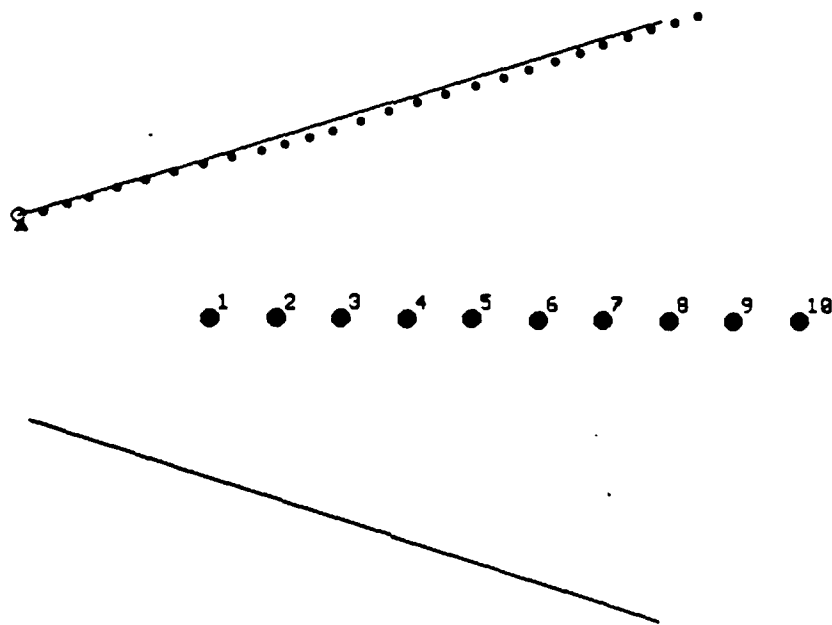
Figure 2.12: A localization example, where the source path is a straight line. Numbered thick dots are the sensor locations. Thin dots are points of the source path. Continuous line is the program's estimate of the source path from power and frequency traces of all sensors. The system finds two solutions, which are symmetric with respect to the line of the sensors. The lower line segment is the second solution.
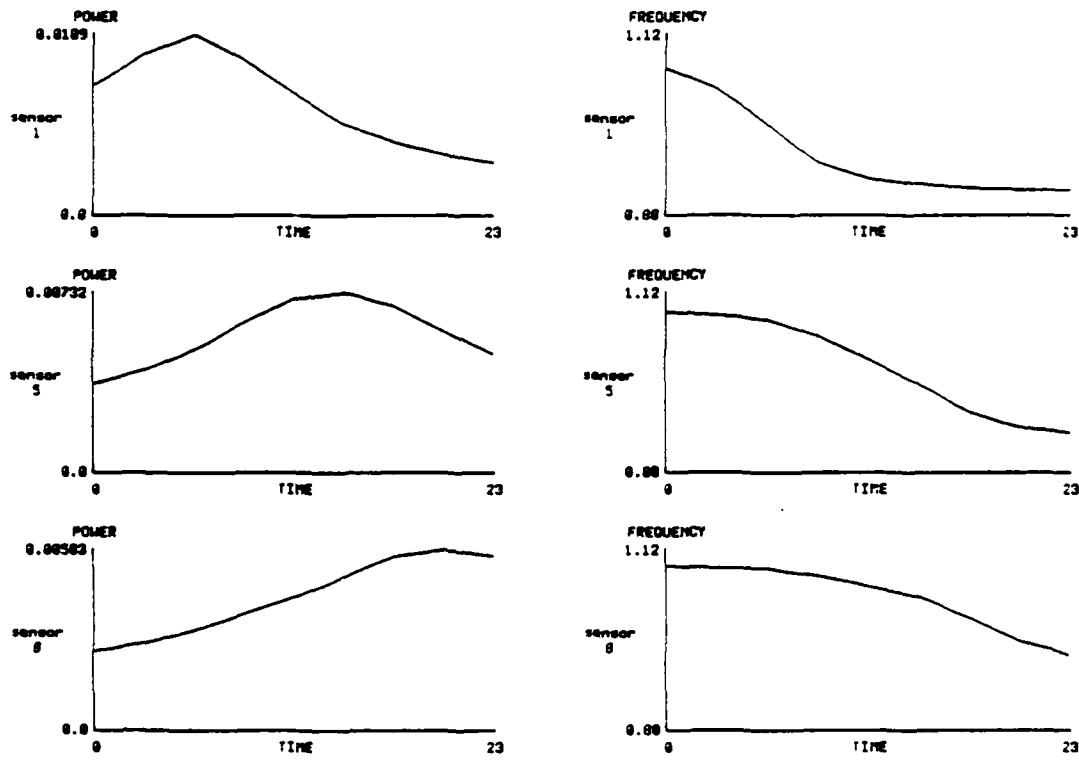
Figure 2.13: Power and frequency traces from sensors 1,5 and 8 corresponding to the example, where the source path is a straight line
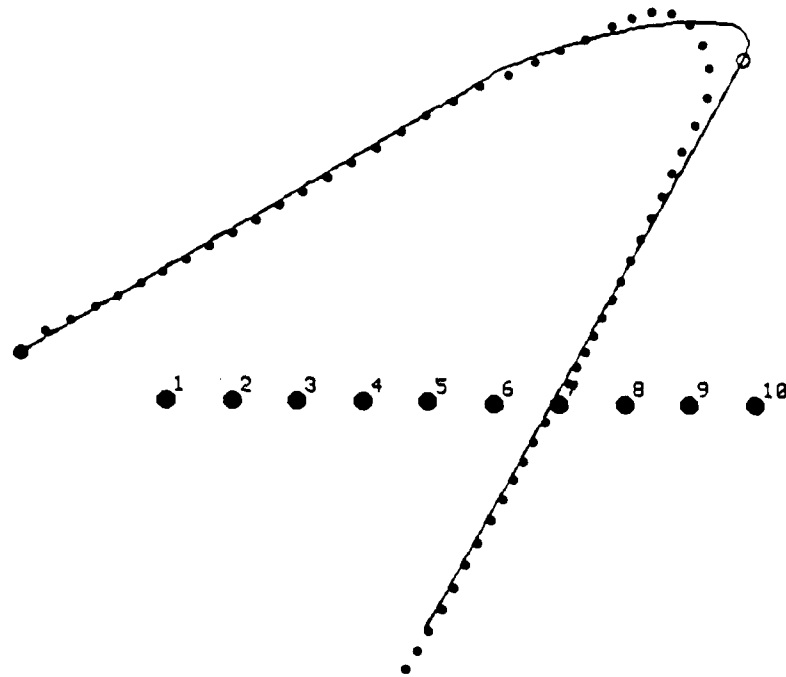
Figure 2.14: A localization example, for which the source path includes a maneuver.

when sufficiently good measurements for power and frequency traces can be obtained from the raw acoustic sensor signals. An "optimal" solution for the source path can be formulated, for example in the least-square sense, if the localization problem is viewed as a problem of estimating the position of the source at a finite number of time points, given measurements of power and frequency at several sensors at the same time points. Although such an approach is plausible, it is likely that it would have problems in practice, because it is not possible to make reliable measurements of power and frequency at all times. In the next two chapters we will see that power measurements from real data tend to be accurate only in the vicinity of the CPA, where the signal is strong. Therefore, a localization technique that has some potential on real helicopter data would need to be data-driven in the times when the signal is strong, but goal-driven when the signal is weak, with the goals set by the correspondences shown in Figure 2.4. A limited form of goal-driven processing for extending and completing frequency tracks will be demonstrated in
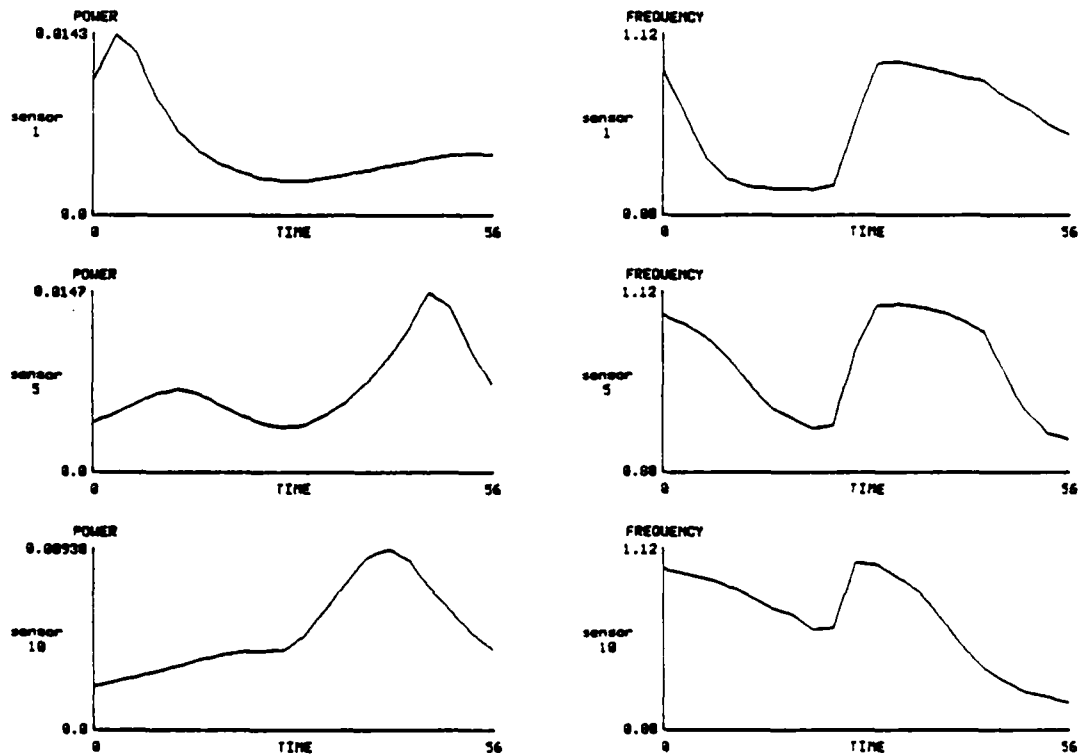
Figure 2.15: Power and frequency traces from sensors 1,5 and 10 corresponding to the example of the previous figure, for which the source path includes a maneuver. The lower line segment is the second solution for the first straight segment of the source path.

2.2.2 Localization from power and frequency with significant propagation delay.

In this section we relax our assumption of negligible propagation delay, but we restrict the acoustic source path to a single straight line and we assume unknown but constant speed for the acoustic source. The assumption of significant propagation delay is more realistic in the case of helicopter signals and geographically separated sensors.

In this case, the elapsed time between the emission of the sound by the source and the reception by the microphone is neither negligible nor the same for all microphones. Instead, each sensor experiences a time-varying time warping, which is different from time warping in other sensors.

As a first step towards localizing the straight source path from sensor measurements, assume that T_1 and T_2 are the times of maximum power in the two sensors A_1 and A_2 of Figure 2.16. With finite propagation delay, these times are different from the times τ_1 and τ_2 the source was at the corresponding closest points of approach. The relation between τ_i and T_i, $i = 1, 2$ is given by:

$$T_i = \tau_i + \frac{(A_i C_i)}{c} \tag{2.9}$$

where $\frac{(A_i C_i)}{c}$ is equal to the propagation delay of the acoustic wave from the source to the sensor, c being the speed of sound. Therefore,

$$T_2 - T_1 = \tau_2 - \tau_1 + \frac{(A_2 H)}{c} \tag{2.10}$$

$$= \frac{(C_1 C_2)}{v} + \frac{(A_2 H)}{c} \tag{2.11}$$

$$= (A_1 A_2)(\frac{\cos \theta_1}{v} + \frac{\sin \theta_1}{c}) \tag{2.12}$$

and similarly for the pair of sensors A_1 and A_3 of Figure 2.17,

$$T_3 - T_1 = (A_1 A_3)(\frac{\cos\theta_2}{v} + \frac{\sin\theta_2}{c}) \qquad (2.13)$$

In the above equations, $T_2 - T_1$ and $T_3 - T_1$ are measured from the power traces of the two sensors and $(A_1 A_2)$ and $(A_1 A_3)$ are known distances between sensors, therefore the unknowns are θ_1, θ_2 and v, the speed of the acoustic source. To solve these trigonometric equations for θ and v, we note that the difference $\theta_1 - \theta_2$ is also known and equal to some angle α. To proceed, we define

$$p_1 = c\frac{T_2 - T_1}{(A_1 A_2)} \qquad (2.14)$$

$$p_2 = c\frac{T_3 - T_1}{(A_1 A_3)}, \qquad (2.15)$$

we eliminate v/c from equations 2.12 and 2.13, and we substitute θ_2 by $\theta_1 - \alpha$. After some algebra, we obtain a single equation with one unknown, θ_1:

$$\frac{p_2 - p_1 \cos\alpha}{p_1 \sin\alpha} \cos\theta_1 - \sin\theta_1 = -\frac{1}{p_1} \qquad (2.16)$$

This is a standard equation, with solutions given by

$$\theta_1 = \xi - \sin^{-1}(\frac{-\cos\xi}{p_1}) \qquad (2.17)$$

where ξ is an angle satisfying

$$tan\xi = \frac{p_2 - p_1 \cos\alpha}{p_1 \sin\alpha} \qquad (2.18)$$

Equation 2.16 in general has two solutions for θ_1 in the interval $[0, 2\pi]$. Once a value for θ_1 is determined, $(A_2 C_2) - (A_1 C_1) = (A_1 A_2)\sin\theta_1$. The ratio $\frac{(A_2 C_2)}{(A_1 C_1)}$ can be measured from the ratio of the corresponding powers in the two power traces (assuming calibrated sensors). The result is two linear equations with two unknowns $(A_2 C_2)$ and $(A_1 C_1)$, which can be solved to compute the exact location of the source path.

Issues of numerical sensitivity of the above localization method still remain to be explored. A simulation system similar to the one implemented for the case of negligible propagation delay would be appropriate for that purpose.

56
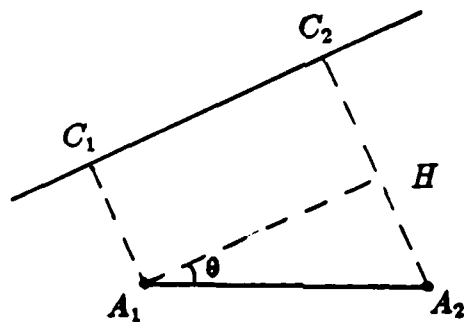
Figure 2.16: Geometry for source localization in the case of significant propagation delay.
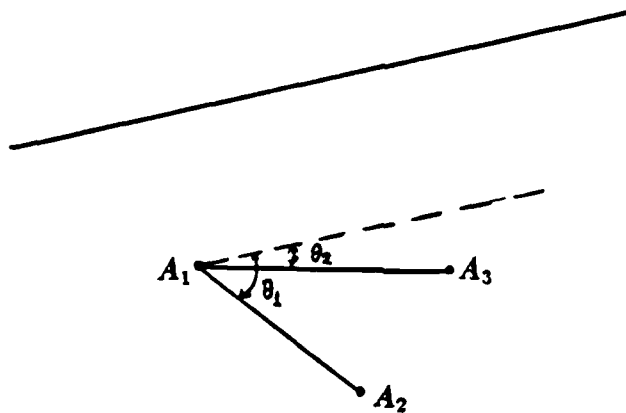
Figure 2.17: Three non-collinear sensors allow computation of v

2.3 Source localization using array processing

A different class of methods for source localization relies on the assumption that the acoustic wave is approximately a plane wave far from the acoustic source. An array of sensors with spatial extent much smaller than the distance between the array and the acoustic sources can then be used to determine the direction of arrival of the acoustic waves. Figure 2.18 shows a typical microphone array, which was used for collecting the data for the experiments of this thesis. The directions of arrival of an acoustic wave at two geographically separated sensor arrays are then used to find the position of the source.
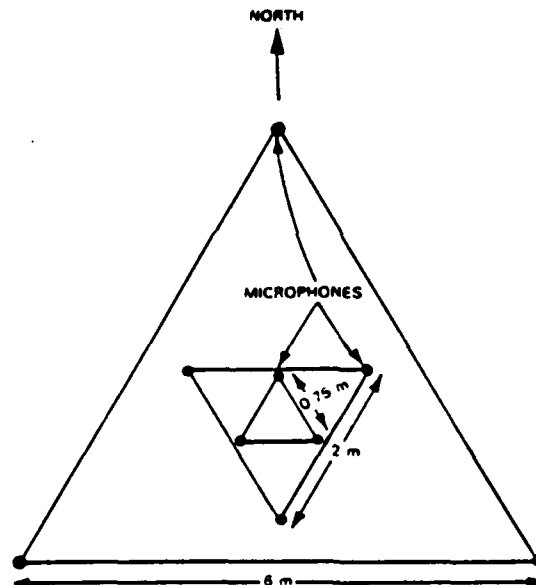
Figure 2.18: Typical microphone array, used for collecting the data for the experiments of this thesis

One particular system for direction determination, proposed and implemented by [Nawab et al 1985], is based on the two-dimensional spectrum of the spatial covariance of the sensor signals, called the wavenumber spectrum. This system is capable of providing accurate azimuth measurements, which are then used for source localization by combining azimuth measurements from geographically separated locations. Experience with the system has shown that its performance is sensitive

to inappropriate parameter settings, and this fact motivated the work presented in Chapters 5 and 6 of this thesis towards a system for identifying inappropriate parameter settings in test runs of the azimuth measurement system. In test runs, the source scenario, i.e. the locations, speeds and characteristics of the acoustic sources are approximately known a priori. This section provides a brief overview of the azimuth measurement system described in more detail in [Nawab et al 1985].

In the wavenumber spectrum, acoustic sources manifest themselves as radial ridges at angles equal to the angles of arrival of the corresponding acoustic waves. Figure 2.19 shows a typical wavenumber spectrum and Figure 2.20 shows the stages of the direction determination algorithm and their major parameters.
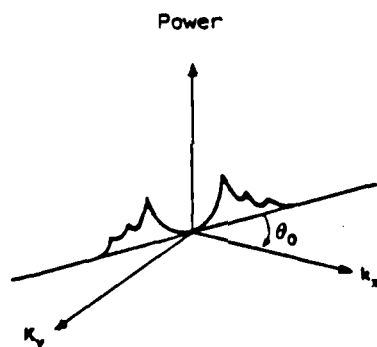
Figure 2.19: A typical wavenumber spectrum, in which an acoustic source is manifested as a radial ridge. The shape of the ridge is a frequency-scaled version of the temporal spectrum of the acoustic signal (from [Nawab et al 1985]).

In practical situations, not all microphones are operational at any given time, due to hardware failures in the microphones or the cables from the microphones to the signal processing unit or other kinds of interference (e.g. fluctuating power supply voltage). Inoperative microphones generate data that violates the assumptions of the array processing algorithms and, therefore, causes serious deterioration in

Microphone Signals → **Analog Filter** → **Digitization** → **Digital Filter** → **Covariance Estimator** → **Spectrum Estimator** → Direction Spectra

(rate/sensor)
(gain-adj/sensor)

(epsilon)
(number-sensors)
(analysis-interval)
(block-length)
(block-overlap)
(cov-method)
(sensor-positions)

(gain/sensor)
(min-cutoff/sensor)
(max-cutoff/sensor)

(low-freq/sensor)
(high-freq/sensor)
(output-rate/sensor)
(window-length/sensor)
(window-shape/sensor)

(r-min)
(r-max)
(az-min)
(az-max)
(az-rate)
(r-rate)
(int-min)
(int-max)
(int-rate)
(spec-method)

Figure 2.20: Stages of the direction determination algorithm and their major parameters

their performance. An alternative is to exclude inoperative microphones from the array processing, in effect reducing the number of sensors used. This leads to a more graceful degradation of the performance compared to that incurred by inclusion of faulty microphone data.

The signal processing for both spectral analysis of microphone signals and direction determination of acoustic sources involves several parameters. Proper setting of those parameters depends on the particular data being analyzed, and no parameter setting exists that is appropriate for all types of input data. For this reason, human signal processors typically analyze data interactively, and they arrive at an appropriate parameter setup by visually inspecting intermediate results and adapting the parameters involved.

The following subsections provide a more detailed overview of the concept of wavenumber spectrum and its application to direction determination.

2.3.1 Definition and properties of wavenumber spectra

In the far field of the (approximately point) acoustic source, the wave $p(\mathbf{r}, t)$ (where p is pressure, $\mathbf{r}$ is position and t is time), can be assumed to be a plane wave. Microphones located in the far field sample the plane wave at specific values of the spatial variables. We can define the *wavenumber spectrum* as:

$$P_{zd}(\mathbf{k}) = \int_{-\infty}^{\infty} P(\mathbf{k}, w) dw \qquad (2.19)$$

where $P(\mathbf{k}, w)$ is the frequency-wavenumber spectrum, defined as the Fourier transform of the space-time covariance function of $p(\mathbf{r}, t)$. The wavenumber spectrum can be shown to be equal to the spatial Fourier transform of the space-time covariance function for time $= 0$ (zero-delay covariance).

Under the assumption that the wave is plane:

$$p(\mathbf{r}, t) = p(t - \alpha \cdot \mathbf{r}). \qquad (2.20)$$

It can be shown that in the latter case the wavenumber spectrum of the acoustic wave $p(t - \alpha_\mathbf{P} \cdot \mathbf{r})$, where $\alpha_\mathbf{P}$ is the projection of the vector α onto the sensor plane, is a distribution in wavenumber space confined to an axis at azimuthal angle θ, where θ is the bearing of α. Furthermore, the distribution on that axis is proportional to the temporal spectrum of $p(t)$, $P(\omega)$, with the frequency axis compressed by the magnitude of $\alpha_\mathbf{P}$, i.e. it is proportional to $P(\omega/|\alpha_\mathbf{P}|)$.

In the presence of several acoustic sources, which generate acoustic waves impinging upon the microphone array from different directions, the wavenumber spectrum consists of several radial distributions, one per acoustic wave, at azimuth angles corresponding to the bearings of the wave vectors, each distribution compressed by the cosine of the elevation angle of the corresponding wave vector.

2.3.2 A direction determination algorithm based on the wavenumber spectrum

The basic concept behind the direction determination algorithm presented in [Nawab et al 1985] is to compute the wavenumber spectrum as the spatial Fourier transform of the zero-delay covariance of the microphone signals, which are viewed as spatial samples of the acoustic wave at the microphone locations.

Several issues are associated with the direction determination algorithm in its practical form. The zero-delay covariance function is both severely undersampled and spatially windowed. Both of these are consequences of the fact that we can afford only a few sensors in a limited spatial configuration. The consequence of windowing in the space domain results in smearing of the ideal wavenumber spectrum, because windowing in the space domain is equivalent to convolution with the

Fourier transform of the window in the wavenumber domain. Undersampling in the space domain results in spatial aliasing in the wavenumber domain. Aliasing can be reduced by designing the array so that the smallest distance between sensors is shorter than half the smallest wavelength of the anticipated plane waves. Spectral smearing due to spatial windowing can be reduced by spreading out the available sensors to increase the aperture of the array. We see that with a fixed number of sensors these two requirements are contradictory.

With a finite number of sensors we can estimate spatial samples of the zero-delay covariance. Using such samples to estimate the wavenumber spectrum is a problem in two-dimensional spectral estimation. High resolution methods can be used instead of the two-dimensional extension of the periodogram method, thereby improving the resolution of the wavenumber spectrum.

Another problem is the fact that a 180-degree ambiguity in direction is introduced when the acoustic wave is real valued. In this case, its temporal spectrum is even, therefore it extends to both sides of the origin in the wavenumber spectrum, as shown in Figure 2.19. To eliminate this ambiguity, the sensor signals are processed so as to suppress the negative sideband of their temporal spectrum, resulting in complex-valued signals, termed the "complex analytic representation" (CAR) of the original real-valued signals.

Thus the overall computation consists of the following sequential steps, shown in Figure 2.20, together with the system parameters associated with each step:

1. An analog bandpass filter is applied to the input microphone signals. Parameters are the gain of each sensor and the minimum and maximum cutoff frequency of the bandpass filter.

2. The filtered microphone signals are sampled. Parameters are the sampling rate and the gain adjustment to compensate for the different power levels of the resulting discrete signals.

63

3. A digital filter is applied to the real-valued discrete signals. This filter suppresses the negative frequencies of its input, therefore producing a complex-valued output, the complex analytic representation (CAR) of the original signals, and it may also act as a bandpass filter. Parameters are: the output sampling rate (the output is decimated with respect to the input), the window length and shape used for computing the DFT of the input and the low and high frequency of the filter acting as a bandpass filter.

4. The zero-delay covariance samples are computed. Parameters needed are: The sensor positions, block length and overlap, the total analysis interval, the particular method used for covariance estimation (temporal or spatial averaging) and the parameter epsilon added to the diagonal elements of the covariance matrix to ensure its stability.

5. The wavenumber spectrum is estimated. Parameters are: the region of the wavenumber space to be sampled, specified by a minimum and a maximum radius and a minimum and maximum azimuth, the sampling rate for both azimuth and radius, and the spectral estimation method (Bartlett or MLM).

Finally, ridges are extracted from the wavenumber spectrum obtained from the above method, by integrating along each radial direction and finding local maxima in the resulting function of radial power versus azimuth.

For typical parameter values, the dominant computations are the digital filter computation and the operations for adequately sampling the wavenumber space, the latter being about three times the former.

2.4 Protocol collection and analysis

As we mentioned earlier, developing accurate models for the generation and propagation of helicopter sound is a very difficult task. Therefore, it is reasonable to pursue design of a system that performs adequately on real helicopter data through the combination of the approximate theory underlying generation and propagation of helicopter signals, experimental results from a variety of real data and a study of the human problem solving activity. The previous sections of this chapter reviewed theories underlying acoustic signal formation and propagation. In this section we discuss some of the issues associated with identifying human skills by observing experienced practitioners as they process and interpret real helicopter data.

Protocol collection is a general methodology for knowledge acquisition in the design of knowledge-based systems [Hayes-Roth et al 1983]. It is a systematic record of the "thinking aloud" of a skilled practitioner during problem solving. The protocol transcript is later analyzed and formalized in order to understand the nature of the problem solving and to design the program architecture and the appropriate representations.

Protocol collection is a difficult procedure, because it involves interactions among a human subject (the "expert"), a domain problem, and the interviewer. A first issue is the choice of appropriate domain problems, that are constrained enough to be covered during a protocol session and at the same time rich enough to illuminate the knowledge representations and inferences in the domain. In the case of medical diagnosis, interviews with physicians are usually held on the basis of a detailed printed description of a patient, resulting in a verbatim transcript of the thinking process of the physician [Kuipers & Kassirer 1984]. In the case of signal processing, the activity is more like probing a piece of data with a set of available tools and less like diagnostic reasoning based on a detailed description of the problem. To limit the scope of the problem solving in our protocol sessions, we limited the amount of data and the set of tools made available to the subject.

A second problem is whether to allow the subject to introspect. A recent study [Kuipers & Kassirer 1984] indicates that considering the subject's introspective accounts may be misleading, because "a subject has no privileged knowledge of the factors that influence his behavior". So what kinds of information can be reliably extracted from a protocol? The same study concludes that "a subject's statement of what is currently in his focus of attention is unlikely to be in error". The implication for protocol collection is that the subject should not be encouraged to introspect and should only report the information and intentions that are within his current focus of attention. In signal processing, this is facilitated by concentrating on the information derived from the signals, and on the choice and use of tools to probe the signals.

Previously proposed computer-aided knowledge acquisition methods support domain experts in providing new knowledge and debugging existing knowledge in a large rule-based system [Davis & Lenat 1982]. Such methods are, however, limited to operating within a predetermined knowledge representation. We are interested in studying the experts for determining the representations for signals and knowledge related to signals, even before attempting to capture large quantities of signal processing knowledge. After determining appropriate signal and knowledge representations, we could use the previous aids to knowledge acquisition.

2.4.1 Analysis of the protocol transcript

Appendix A presents a complete protocol transcript carried out in the spring of 1984. The subject was a person experienced in interactive processing and interpretation of helicopter acoustic waveforms. The subject was provided with a particular piece of data and facilities for graphically displaying the waveform with different time scales and for computing and displaying its spectrum using averaged periodograms. The subject was asked to think aloud regarding both his current interpretation of the data and the information he visually extracted from the signal

graphs. The subject's verbalizations were tape recorded and saved along with all signal graphs displayed during the protocol session.

Analysis of the verbatim transcript takes place in the following stages:

1. Identify the signal entities referred to and their relationships in a single spectrum.

2. Identify relationships of interest between entities in spectra at different times and from different channels.

3. Identify processing plans and their relation to the formation of tentative interpretation hypotheses.

The protocol transcript shows that the subject's reasoning about spectra is centered around spectral peaks and their harmonic relations (S20 - S57 in appendix A). Missing harmonics (S32, S52) and extraneous peaks, i.e. visually prominent peaks that do not seem to belong to previously identified harmonically related peak sets are also of interest. Furthermore, plateaus (flat areas next to peaks) are an indication of possible hidden peaks.

Relationships of interest between spectra from different channels are similarity of peak patterns and noise levels, determined primarily by the visual prominence of the peaks in the spectra (S37, S44).

Relationships of interest between spectra from different times are evolution of peak patterns and noise and peak levels (S44, S46, S54, S56, S57). In comparing spectra from different times, we expect changes in the power levels of the peaks, their frequencies and the noise level of the spectra, and we want to characterize such changes. In comparing spectra from different channels (not appearing in appendix A), we expect the spectra to be the same, and any serious dissimilarities of one channel with the rest of the channels indicate a faulty channel.

Regarding processing plans, protocols indicate that the expert alternated be-

tween two modes: exploration and confirmation. In exploration mode, he was looking for general characteristics of interest, such as harmonically related peaks. In confirmation mode, he was looking for specific characteristics, such as existence of a peak or peak features in particular narrow frequency ranges.

Strategic planning of signal processing is guided by the heuristic of finding a consistent hypothesis about the acoustic sources present in the signals, which is supported as fully as possible by the existing data. Consistency requires constant testing for similarity of peak patterns in all available channels. Consistency also requires establishing continuity of peak patterns over time and a plausible explanation of their evolution.

* An analyzed protocol transcript is obtained from a verbatim transcript by separating signal descriptions provided by the subject, hypotheses about the source motion made by the subject, and the plans formed by the subject for probing including the purpose of the probing. Figure 2.21 shows part of an analyzed protocol.

2.4.2 Signal processing and its relation to source hypotheses.

From the protocols, we identified two kinds of signal processing actually taking place during a problem solving session: The first kind is signal processing that is actually performed by the expert on the signals. For example, periodogram-based spectral estimation may be used to change the signal description from the time domain into the Fourier domain, thus making it easier to identify high frequency harmonics that are often not apparent in the time waveform. The second kind is human visual extraction of signal information from waveforms and spectra. For example, a human locates the "loud" regions in a time waveform or identifies the harmonics in a spectrum by visual inspection. The human identifies higher level entities in the signals, e.g. peaks and harmonic sets and reasons on the basis of

68
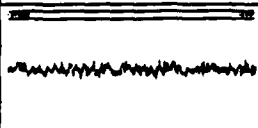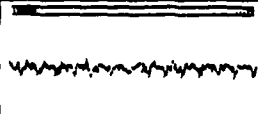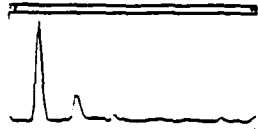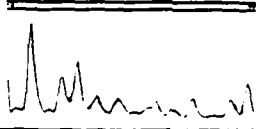
SIGNAL	SIGNAL DESCRIPTION	SOURCE HYPOTHESIS	PLAN
	Signal has no apparent periodicity	No source or source far away	Compute spectrum with averaging over 16 blocks
	Spectrum has clear harmonic peaks at 210, 315 and 527 Hz	Source present but far from the sensor. Source tail rotor fundamental at 105 Hz	Check signal 30 secs later
	Signal has a periodicity of 16 Hz	Source is close. Source main rotor fundamental is 16 Hz	Compute spectrum with averaging over 8 blocks looking for the main rotor harmonics
	3 strong harmonics of 16 Hz, a strong harmonic at 91 Hz	Source main rotor fundamental at 16 Hz confirmed	Compute spectrum 10 secs later with averaging over 8 blocks looking for Doppler shift
	3 strong harmonics of 14 Hz	Doppler shift detected. Source is close to CPA	....

Figure 2.21: Part of an analyzed protocol. Protocol should be read by row, left to right. The plan indicates computation and/or display of a signal, shown as the signal of the next row. The human operator extracts a relevant signal description, from which he formulates a source hypothesis.

such entities at multiple levels of detail.

Protocol analysis also gives indications concerning the interaction between the above two kinds of signal processing. Determining the nature of such interaction is important for designing a computer program that performs the task of the human operator. Protocols indicate that the human operator interleaves the two kinds of processing, using the second kind to plan the first kind. Human visual extraction of signal information, such as harmonically related peaks in a spectrum or periodic structure in a time signal, does not necessarily involve direct association with real world entities.

Forming source hypotheses does involve associating signal descriptions with the presence and motion of acoustic sources, such as helicopters. Planning the next signal processing operation attempts to confirm the anticipated characteristics in the signal, based on the source hypothesis. Therefore planning sets up a context for the derivation of relevant signal descriptions, by limiting even more the signal characteristics of interest in the result of the particular signal processing operation.

Chapter 3

Abstraction and Signal Matching

In this chapter, we focus on the initial step towards helicopter localization from one-dimensional spectral information, namely the extraction of reliable power and frequency traces from real helicopter data. As we saw in chapter 2, accurately modelling the generation and atmospheric propagation of helicopter sound is a very difficult problem. In spite of the lack of accurate models, human experts are capable of extracting a significant amount of information out of real helicopter data, by a process involving interactive signal processing and reasoning about various signal entities, such as harmonically related sets of spectral peaks, using highly idealized models as a guide. Human experts tend to view helicopter spectra not as mere functions of power versus frequency, but as collections of such higher level signal entities.

To represent this complex view explicitly, we introduce the notion of the extended spectrum. The Extended Spectrum is a representation of multiple levels of detail associated with a spectrum as a single conceptual (and hence representational) entity, including all numerical parameters necessary to derive it completely from the time waveform.

As we saw in the section of the previous chapter on protocols, an important

component of the processing and interpretation of helicopter signals is spectral matching.

One instance of spectral matching is the identification of faulty sensors: power spectra from sensors close to each other ideally are the same, because an acoustic plane wave results in signals differing only in phase. In real data, spectra from different sensors differ because of noise. Whereas differences in the noise peaks are expected, differences in prominent features, such as extra peaks or extra harmonically related peak sets, present in only one of a number of sensors indicate that the sensor may have a fault. A program for identifying a faulty channel has been implemented that relies on extraneous or missing peaks in only one of a group of spectra from different sensors.

Another instance of spectral matching is related to adjustment of spectral estimation parameters to a particular piece of r·al helicopter data. In chapter 2 we saw some of the tradeoffs involved in estimating the spectra of helicopter signals. Adjustment of the spectral estimation parameters in practice is guided by the observation that helicopter spectra have sharp peaks, which are harmonically related, and correspond to the main and tail rotor of the helicopter. Therefore, the parameters that optimize the "peakiness" of the main and tail rotor harmonics in the resulting spectrum are, in a sense, the most appropriate for the task of tracking the power and fundamental frequency of the helicopter. An iterative program has been implemented that performs matching of spectra obtained from the time waveform with different parameter settings to find the direction of change of the parameters that increase the "peakiness" of the resulting spectrum. Peakiness is defined at multiple levels of abstraction using the notion of the extended spectrum.

A third instance of spectral matching is involved in forming power and fundamental frequency traces as needed for the localization techniques presented in chapter 2. In this case, matching takes place between spectra obtained at consecutive times, and associates harmonically related peak sets with neighboring frequencies.

Pairwise associations lead to the formation of sequences of such sets. Power and frequency traces are then extracted from these sequences. We have implemented a program for this task, described in chapter 4, together with the results of its application to a variety of real helicopter data. Our experience indicates that a sequential approach for source localization, whereby we first extract power and frequency traces and then use the geometric localization techniques and qualitative associations between power and frequency and local path shapes, is not always successful because harmonically related peak sets are often hidden in noise. Our iterative approach uses a combination of data-driven processing, which consists of forming portions of harmonic set sequences, at time intervals when such harmonic sets are prominent, and goal-driven processing, which consists of searching for harmonic sets with specific fundamental frequencies in order to complete or extend already formed sequences. This is similar to the notion of "islands of certainty" in artificial intelligence [Barr & Feigenbaum 1981, Erman et al 1980].

This chapter is divided into two parts. The first part defines the extended spectrum, a multilevel representation appropriate for spectra characterized by harmonically related peak sets, and addresses the issues involved in its computation from the numeric spectrum. The second part presents the first two systems outlined previously, i.e. the system for faulty channel identification and the system for adjusting the parameters of spectral estimation to maximize the peakiness of the resulting spectrum. Both systems rely on the notion of the extended spectrum. The second part ends with an overview of the issues associated with matching of spectra at different times, leading to the discussion of the system for extracting power and frequency traces from real helicopter data, fully described in chapter 4.

In all these problems, we view signal matching as a framework, because the designer of the matching program has choices along several dimensions which determine exactly how matching is performed. A single set of matching operations may not be general enough for most signal correlation applications. Instead of providing a universal set of matching operations, we structure matching around

multiple levels of abstraction, and we allow the program designer to attach specialized matching procedures for the various levels of abstraction [Bobrow & Winograd 1977].

3.1 The Extended Spectrum: a multilevel representation for harmonic spectra

The extended spectrum is defined at three levels of increasing abstraction:

- the spectrum level (power as a function of frequency),

- the peak level (a collection of peaks, with each peak characterized by its center frequency, its maximum power, several other features and a symbolic characterization of its prominence derived from those features),

- the harmonic peak set level (a collection of sets of harmonically related peaks, with each set characterized by a few basic features, plus a list of extraneous peaks, which are prominent peaks not belonging to any harmonic set).

In addition, the extended spectrum includes three sets of parameters, corresponding to the following three stages of the computation of the extended spectrum from the time waveform (Figure 3.1).

The first stage is the computation of an estimate of the spectrum from a given time waveform. The main issue here is the interaction of the requirements for good resolution, good noise reduction and lack of infinite data to work with, with tradeoffs involved as discussed in previous chapters.

The second stage consists of detection and symbolic characterization of spectral peaks. Peaks are detected as local maxima of the spectrum, i.e. as (discrete) frequencies at which the spectrum is stronger than both of the neighboring frequencies. We treat peak characterization as a classification problem, and several approaches are possible within this framework. Characterization of peaks is based on empirical criteria, hence our implementation needs to be flexible and to allow easy experimentation. We used rules as the programming technique to capture such empirical

List of
Harmonic Sets

↑

| Grouping into
Harmonic Sets |

- approx. multiple threshold
- fund. frequency range
- rules for fund. freq. hypotheses
- rules for harm. set pruning

↑

Peak
List

| Peak
Detection &
Characterization |

- heuristic rules
- thresholds

↑

Power
Spectrum

| Spectral
Estimation |

- time
- block length
- number of blocks
- block overlap

↑

Time Waveform
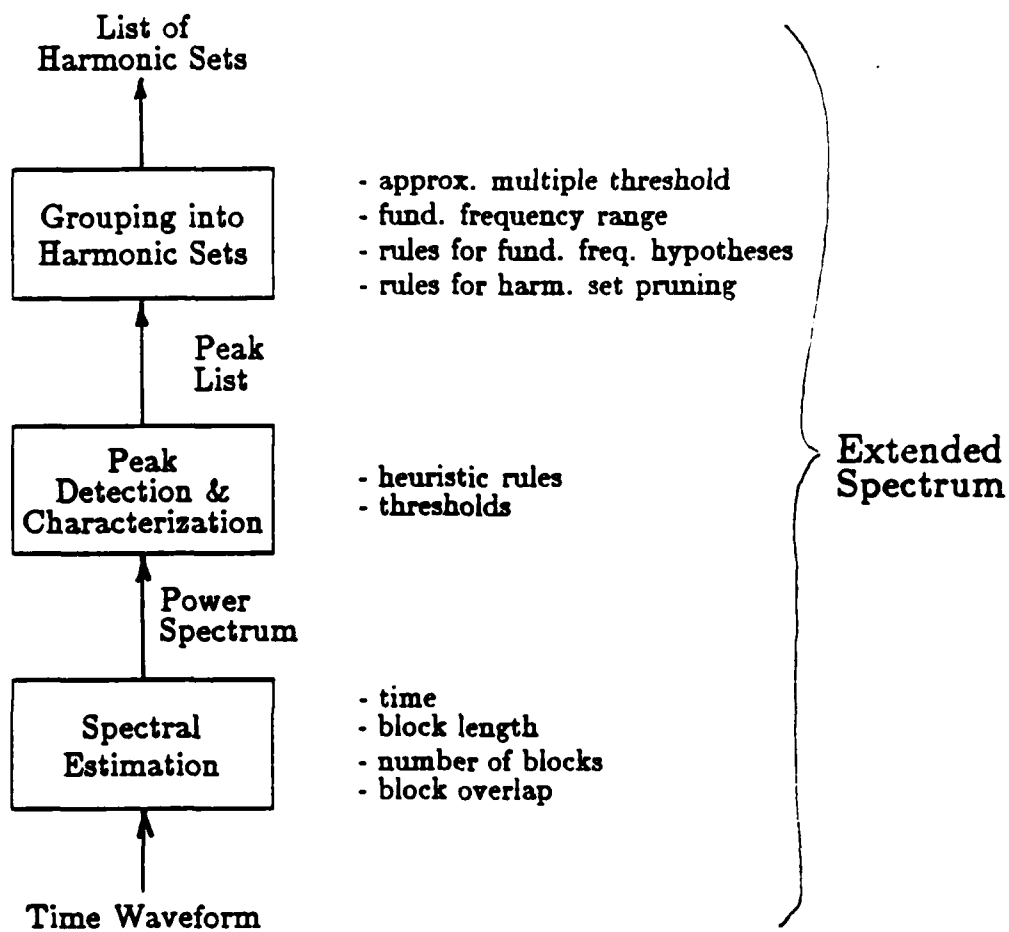
Extended
Spectrum

Figure 3.1: Stages of computation of the extended spectrum from the time waveform

criteria and we found it to be a convenient representation, because it was possible to use the same rules for two distinct purposes, classification and training based on examples, as will be explained later in this section.

The final stage is grouping of peaks into harmonic sets. The main problem here is that harmonic relations are not exact, but approximate. Grouping is achieved through the use of an approximate multiple criterion, selection of candidate fundamental frequencies and testing of peaks against the hypothesized fundamentals. Finally, harmonic sets thus formed are classified as potentially relevant or irrelevant to an acoustic source. Heuristic criteria are applied here, with the relevance thresholds for acceptance set low, because ultimately relevance of a harmonic set depends on its continuous presence over time as dictated by the scenario.

In this section we provide the components of the extended spectrum and then details of the above computations, with each component description preceding the details of its computation. Then we provide the representation of the extended spectrum as an abstract data type and point out how this extends the signal view introduced by the closure model of [Kopec1980]. Finally, we discuss some of the issues associated with our rule implementation of embedded classification.

3.1.1   Peak detection and symbolic characterization

Peaks correspond to local maxima of the spectrum, therefore they are located at zero crossings of the first difference of the spectrum. In general, we consider as a peak the whole section of the spectrum extending between two valleys and containing the local maximum.

Various pieces of information about a peak (features) can provide an incomplete characterization of a peak, but nevertheless sufficient for certain purposes. In Figure 3.2 we show the features we used in the extended spectrum. The last item in this figure refers to a set of characterizations associated with the peak according

77

to multiple classification criteria, which capture various cases of the peak being visually prominent, i.e. "standing out" in the spectrum. Prominent peaks are of interest because they are more likely to be due to acoustic sources. The classification criteria we used are shown in Figure 3.3. They were developed by experimenting with real helicopter data to produce results consistent with the experimenter's visual experience.
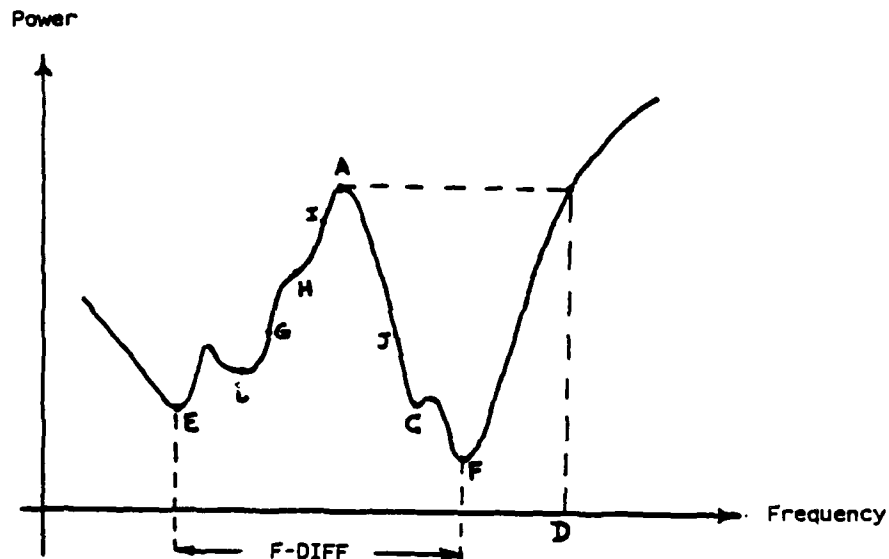
- Frequency and power of the peak at the location of maximum power (point A).

- Frequency and power of the enclosing valleys (points B and C).

- The closest frequency to the peak frequency, at which power exceeds the peak power (point D). It determines the size of the neighborhood, over which the peak is a local maximum.

- Frequency, power and frequency difference F-DIFF of the two enclosing valleys, the powers of which are below a certain percentage of the peak power (points E and F). Helps characterize peaks with rippled sides, therefore displaying many shallow maxima and minima.

- Frequency and power of all inflection points between the two enclosing valleys (points G, I and J). Helps identify plateaus.

- Symbolic characterization of the peak concerning its visual prominence in the spectrum.

Figure 3.2: Peak features used in the extended spectrum
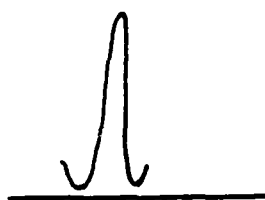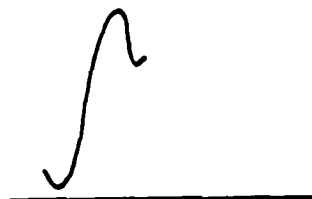
Large average   Prominent   Prominent but
prominence   Unbalanced   maybe rippled

- Large average prominence: the ratio of the power of the peak to the average of the power of two enclosing valleys is high.

- Prominent unbalanced: the ratio of the power of the peak to the minimum power of the two enclosing valleys is high while the ratio of the power to the average of the power of the two enclosing valleys is low.

- Strong: Power of the peak is within a certain fraction of the power of the strongest peak in the spectral region under consideration. This criterion is especially useful when we are examining short regions of the spectrum. Over long regions it tends not to be satisfied for most peaks because of the rapid decay of the spectrum with increasing frequency.

- Locally strong: Peak is a local maximum over a fairly large neighborhood.

- Prominent but maybe rippled: The enclosing valleys with power below a certain fraction of the power of the peak exist and are fairly close to each other. Such peaks appear to be prominent, but they may have rippled sides.

Figure 3.3: Classification criteria corresponding to symbolic characterizations of peaks

Rule implementation of the classification criteria To allow for easy experimentation, the classification criteria were implemented in the rule language YAPS [Allen1983]. The format of the rules is as follows:

```
Rule:
  parameter passing patterns
  data passing patterns
  test criteria
-->
  assign characterization to peaks.
```

A typical rule is shown in Figure 3.4. The lines after LET and before IF bind local variables (in lowercase) to patterns in the fact base. "peak" is the peak object being characterized, whereas the rest of the local variables correspond to thresholds used by the rule. The lines after IF and before THEN are tests defined on the local variables. The tests consist of Lisp functions that access various pieces of information in the peak object and test whether they belong or not to appropriate regions bounded by thresholds. The format of the tests is:

```
<a peak characteristic> BELONGS-IN <region bounded by thresholds>
or
<a peak characteristic> DOESNT-BELONG-IN <region bounded by thresholds>)
```

In our classification we accept overlapping classes (i.e. it is possible for a peak to have multiple symbolic characterizations). Overlapping classes result from multiple classification criteria applied independently. Thus the problem of classification with overlapping classes can be viewed as consisting of multiple binary classification problems solved independently (i.e. a given peak is classified in many ways and the complete characterization consists of a list of the positive characterizations).

Use of multiple classification criteria applied independently makes the system modular and allows experimentation with various collections of classification criteria

80

```
RULE PROMINENT-BUT-MAYBE-RIPPLED:

LET PEAK BE peak
    PROMINENCE-REGION BE pr
    PROMINENT-BUT-MAYBE-RIPPLED-REGION BE prrippled
    NARROW-WIDTH-REGION BE wr

IF
    PEAK-HORIZONTAL-DISTANCE BELONGS-IN wr, AND
    PEAK-PROMINENCE DOESNT-BELONG-IN pr, AND
    PEAK-GEN-PROMINENCE BELONGS-IN prrippled

THEN

    INSERT ''PROMINENT-BUT-MAYBE-RIPPLED'' INTO PEAK-CHARACTERIZATION
```

Figure 3.4: An example of a classification rule

just by "plugging" them in and out. In addition, it allows a training system to be built in addition to the classification system which is independent of the specific criteria used. The view of the rule base as useful in both classification and training is shown in Figure 3.5. The rules capture generic perceptions of what the various characterizations are, without reference to specific numerical values. For example, one kind of prominent peak is the one whose peak height is much larger than the average height of the enclosing valleys. At the rule level, what larger means is not specified. This is the role of the threshold level: the meaning of "larger" is application dependent and it is encoded in the form of one or more thresholds.

The classification and training programs are independent of both the rules and the thresholds. The classification program uses the rules and thresholds to produce peak characterizations. The training program uses the rules, peak characterizations provided by the user and the numeric spectrum to adjust the thresholds. This dual use of rules is achieved by constraining their form.
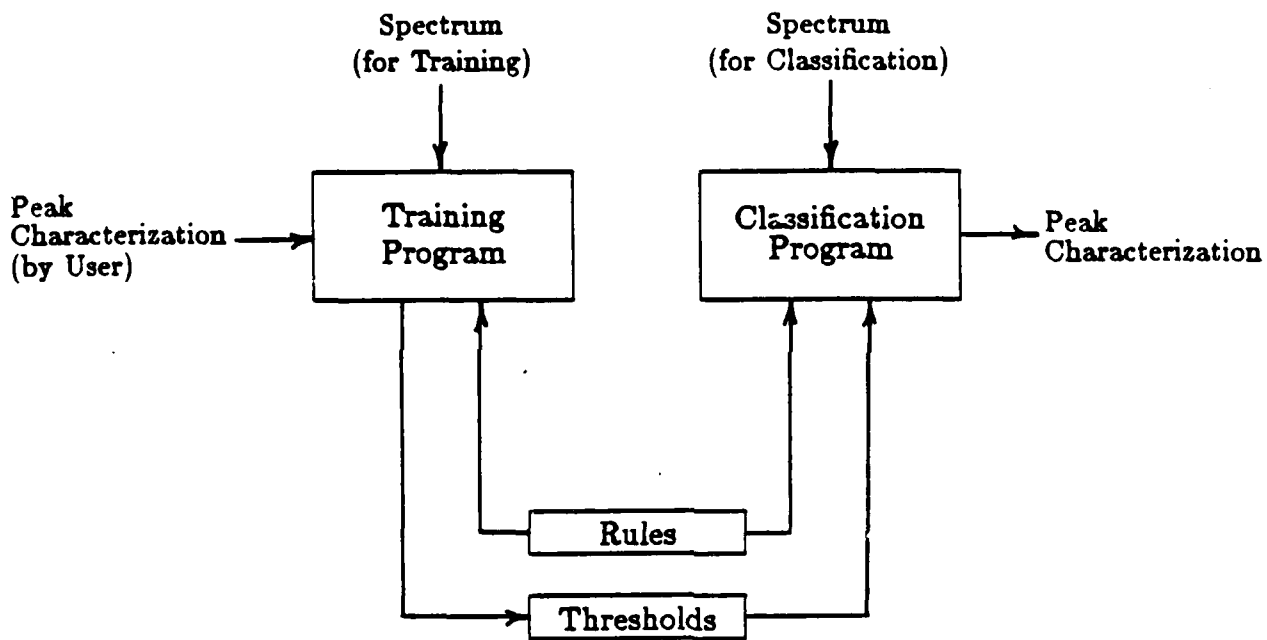
Figure 3.5: The two uses of the rule base: classification and training

The training system uses the left side of the applicable rules (as specified by the human trainer) and forces satisfaction of the test criteria by modifying the parameters that are used in the tests (thresholds). This is achieved by first matching the patterns in the left-hand side of the applicable rules and then considering the test clauses one by one. The peak characteristic is computed and the clause is forced into satisfaction by enlarging (or shrinking) the thresholding region. The above training procedure was actually used in setting the thresholds of the criteria shown in Figure 3.3.

The classification of each peak is later used in other classification problems at higher levels of abstraction. For example, the list of classes determines whether a peak is considered when hypothesizing fundamental frequencies or whether to classify a peak not belonging to any harmonic set as extraneous, and therefore worth considering further, or to ignore it. Thus a binary classification problem at higher levels of abstraction (for example whether a peak not in a harmonic set is considered important or not) is decomposed into several binary classification problems at the peak feature level (i.e. whether the peak is visually prominent according to any of the criteria for prominence). Such decomposition adds to the modularity of the binary classification problem at the higher abstraction level and makes explicit its solution in terms of binary classification problems at the peak level.

3.1.2 Grouping of peaks into harmonic sets

Harmonically related peak sets potentially have a direct correspondence to harmonic signal sources in the physical world. In the acoustic source localization problem discussed in chapter 2, features of the harmonic sets, such as the total power of their peaks and their fundamental frequency, depend on the type of the acoustic source, its distance from the sensor and its relative velocity with respect to the sensor. Hence harmonic sets are of immediate interest and relevance to the problem of localization and tracking of acoustic sources.

83

Several problems are present in grouping spectral peaks into harmonic sets. Peaks belonging to harmonic sets are not always exact multiples of the fundamental frequency. This is due to frequency modulation (because of the Doppler effect) and also because the exact frequencies of the harmonics may lie between two successive samples of the uniformly spaced discrete frequencies of the spectral estimate. Thus a notion of approximate multiple must be used.

Another problem is octave errors. For example, peaks at frequencies 24 Hz and 60 Hz can belong to a harmonic set with fundamental 3, 6, or 12 Hz. This problem is more acute with small fundamental frequencies because the spectral resolution over a frequency range of size equal to the fundamental frequency is small. The problem of octave errors is further compounded if approximate multiples are used. Finally, near overlap of harmonic sets (for example two harmonic sets with peaks at 8, 16, 24 Hz and 9, 18, 27 Hz respectively) may cause the almost overlapping peaks to get merged to a single peak with a frequency shifted with respect to the true frequencies of both harmonics (Figure 3.6).
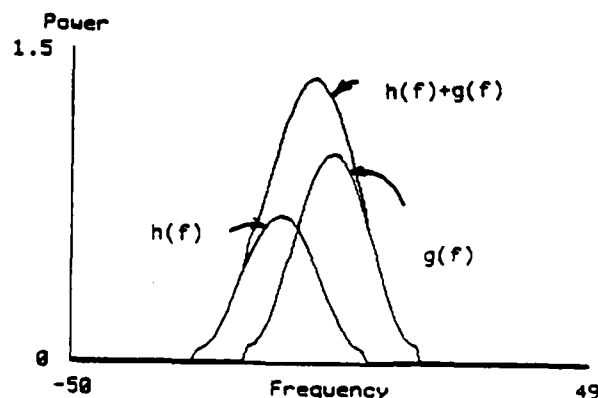
Figure 3.6: Harmonic set overlap can cause peak shifting.

The approach we followed for forming sets of harmonically related peaks consists of the following basic steps:

84

1. Find the candidate fundamental frequencies.

2. For each candidate fundamental frequency, collect peaks whose frequencies are approximate multiples of that fundamental. The dividend is considered an approximate multiple of the divisor if the remainder is a small fraction of the divisor.

3. Prune out or merge thus formed harmonic sets, if appropriate.

4. Find all prominent peaks that do not belong to any harmonic set after the pruning process and see whether they fit in one. If they do, insert them into the harmonic set.

Within this basic framework, we added several heuristic criteria for making the system more robust in processing real helicopter data. New heuristic criteria are derived while debugging discrepancies between the implemented criteria and the visual experience of the user/programmer of the system. We will next describe in more detail both the heuristic criteria we derived as they are embedded in sequential processing.

Finding candidate fundamental frequencies In finding candidate fundamental frequencies, we are only interested in fundamentals in a prespecified fundamental frequency range. Furthermore, we only consider peaks that belong to a prespecified frequency range of interest as candidates for higher harmonics. Both of these ranges are determined by the kinds of acoustic sources we are looking for. The criteria are as follows:

For each peak in the prespecified frequency range of interest (which is the cover of the fundamental frequency range and its second and third harmonics), we extract a candidate fundamental frequency:

- equal to the peak frequency (if the peak belongs to the fundamental frequency range)

- equal to the peak frequency divided by 2 (if the peak is prominent and its frequency divided by 2 lies in the fundamental frequency range). This means that the peak is a candidate second harmonic.

- Same as above, but considering the peak as a candidate third harmonic.

The above rules reflect the empirical fact that fundamental frequency peaks may be weak or missing altogether, and that if the fundamental is missing and neither the second not third harmonic is prominent, then most probably the harmonic set is a spurious one and not worth considering.

Grouping of peaks into harmonic sets After collecting the candidate fundamental frequencies, we proceed as follows: for each candidate fundamental frequency and for each peak in the frequency range of interest, we test whether the peak is an approximate multiple of the fundamental. If it is, we add it to the set of peaks that are possible harmonics of the fundamental.

Candidate fundamental frequencies are not necessarily integers. Hence in our approximate multiple test, we require that the quotient of the peak frequency and the fundamental be within a certain range of an integer (reasonable values for the slack are 5 to 10 %). For example, 61.3 can be considered a multiple of 12.1 of order 5, and the slack is equal to $|61.3/12.1 - 5| = 0.066$ or 6.6%.

It is possible for the same peak to appear in more than one collection of harmonics (since it is possible for the same peak to lie at the intersection of two harmonic sets, for example peak at 30 Hz is at the intersection of two harmonic sets with fundamentals equal to 6 and 15 Hz).

For each such collection of peaks formed around each candidate fundamental frequency, we form an instance of the harmonic set abstract data type. The harmonic set data type is composed of the following features of the harmonic set: fundamental frequency, the order of the harmonics that are present, whether the fundamental peak is missing, whether the fundamental peak is prominent, the number of promi-

nent harmonics, the average and total prominence of the peaks of the harmonic set, the average and total deviation of the peaks from being exact harmonics of the fundamental and a pointer to the harmonic peak objects, which contain more detailed information about the corresponding peaks.

Upon formation of a harmonic set instance, all lists of peaks of the same harmonic order are sorted by increasing slack, the fundamental frequency is being reestimated based on the peaks that are "best" harmonics of the original fundamental for each harmonic order and the harmonic set is sorted once more. We currently reestimate the fundamental frequency based on the frequencies of all harmonics. Alternately, we could reestimate the fundamental frequency based only on selected peaks of the harmonic set, for example the first four harmonics or the prominent harmonics only. It is still to be examined whether such selectivity improves performance and how heuristics of such type can be conveniently programmed. Furthermore, the process of reestimating fundamental frequency and sorting can go on until no change occurs to the harmonic set. We arbitrarily limited this to one iteration.

Harmonic set pruning Finally, all thus formed harmonic sets are pruned using a set of heuristics. These heuristics specify which of the harmonic sets are acceptable as possibly related to an acoustic source. The heuristics are specific to the kind of signals we are processing. However, the framework in which they are embedded is quite general and allows the user to experiment with different heuristics appropriate for other kinds of data.

Figure 3.7 shows all the harmonic pruning heuristics currently used. Each heuristic is designed to prune out a different kind of "false" harmonic sets that may arise as artifacts of the peak grouping technique described previously. Heuristic 1 is concerned with harmonic sets that may consist purely of noise peaks. Heuristics 2 and 3 are concerned with harmonic sets with sparse or too few peaks. Heuristic 4 is concerned with harmonic sets with similar fundamentals which should be merged

into a single set. Heuristics 5 and 6 prune out sets that differ from other sets only by a number of noise peaks. Finally, heuristics 7 and 8 deal with two types of "octave" errors. Octave errors arise when a fundamental frequency which is double or half the correct one is found, either in addition to or instead of the correct one. Octave errors are a known problem in pitch detection of voiced speech [Parsons1976].

1. If no peak in a harmonic set has large average prominence, the set is pruned out.

2. If a harmonic set consists of a single peak, it is pruned out.

3. If both the second and third harmonic of a harmonic set are missing, and its fundamental is not a prominent peak, the set is pruned out.

4. If two harmonic sets have very similar fundamentals, they are tentatively combined into a single harmonic set with fundamental frequency equal to the average of the two fundamentals. If the resulting set has at least as many peaks as each of the original sets or it has at least 3 peaks, it replaces both of the original sets. If not, the original set with the fewest peaks is pruned out, if it has at least 2 fewer peaks than the other set and at most 2 peaks in total.

5. If a harmonic set is a subset of another harmonic set and if its average prominence is much smaller than that of the latter set, the former harmonic set is pruned out.

6. If a harmonic set shares all its prominent peaks with another harmonic set, the former harmonic set is pruned out.

7. If the fundamental frequency of a harmonic set is equal to half the fundamental frequency of another set, and the former set shares all its prominent peaks except at most one with the latter, and the fundamental of the latter set is a prominent peak, whereas the fundamental of the former set is not, the former set, i.e. the set with the lowest fundamental frequency, is pruned out.

8. If the fundamental frequency of a harmonic set is equal to half the fundamental frequency of another set, and the latter set shares all its prominent peaks except at most one with the former set, and the fundamental of the former set is a prominent peak, whereas the fundamental of the latter set is not, the latter set, i.e. the set with the highest fundamental frequency, is pruned out.

Figure 3.7: Harmonic set pruning heuristics

The above heuristics apply equally well to both low- and high-frequency spectra. Low frequency spectra extend over a fairly narrow range of frequencies, 5-100 Hz, over which the decay of the spectrum with increasing frequencies is usually not very

rapid. The total number of peaks is in the tens and the spectrum is rather noisy with intense peak merging and splitting due to noise. High frequency spectra extend over a broad range of frequencies, 70-800 Hz, over which the decay of the spectrum is significant. The total number of peaks is in the hundreds and the interesting peaks are significantly more prominent compared to the low frequency spectra.

Examples of the extended spectrum representation We now present some examples of the extended spectrum representation of spectra of real helicopter data. We only show signal-related information of the extended spectrum. In addition to that, the extended spectrum includes parameter objects, which contain all the necessary numerical parameters and rules necessary to compute peak and harmonic set abstractions from the numeric spectrum.

In the first example we have a single helicopter far before its closest point of approach, and the extended spectrum is focused on the tail rotor frequencies, i.e. a frequency range of 80-1000 Hz. Part (a) of Figure 3.8 shows the spectrum of the signal in logarithmic scale. Part (b) shows the extended spectrum representation for part (a) in tabular form.

At the peak level, we show the frequency, power, prominence, generalized-prominence, horizontal-distance and characterizations of each peak, which is potentially relevant to an acoustic source. Prominence is equal to the ratio of the peak power to the average power of the two enclosing valleys. Generalized prominence is equal to the ratio of the peak power to the average power of the two enclosing valleys, whose powers are below a certain percentage of the peak power. Horizontal-distance is the frequency difference of these two valleys. If these valleys do not both exist for a peak, the corresponding entries are marked NIL.

For each harmonic set, which is potentially relevant to an acoustic source, we show its features as described previously. The harmonic set with fundamental frequency equal to 94.8 Hz is due to the tail rotor of the helicopter. It shows a full

Figure 3.8: Extended spectrum for single helicopter far before CPA and tail rotor frequencies

set of harmonics up to 9th order and it has 4 prominent harmonics (i.e. harmonics with the characterization *LARGE-AVG-PROMINENCE*). The harmonic set with fundamental frequency 80.7 Hz is missing several harmonics and its only prominent harmonic is shared with the harmonic set at 94.8 Hz. It is doubtful that the latter harmonic set is due to a helicopter source. No extraneous peaks, namely relevant peaks not belonging to a harmonic set and not only *LOCALLY-STRONG*, are present.

The second example (Figure 3.9) corresponds to the main rotor frequency range of the first example. At the peak level all peaks are considered relevant, even those with no characterization at all. This is not as costly as in the tail rotor frequency range, because there are significantly fewer peaks. Only one harmonic set is found with fundamental frequency equal to 14.7 Hz, corresponding to the main rotor of the helicopter. A prominent peak at 53 Hz is unaccounted for, so it is an extraneous peak, together with the strong peak at 6 Hz.

The third example (Figure 3.10) is a signal from two helicopters following the same path. The first helicopter is near CPA and the second before CPA. The extended spectrum shows two harmonic sets with fundamental frequencies at 91.9 and 94.8 Hz. The harmonic set at 91.9 Hz is from the tail rotor of the helicopter near CPA, and the one at 94.8 Hz from the tail rotor of the other helicopter.

The final example (Figure 3.11) is the extended spectrum of the main rotor frequency range of the previous example. It shows a harmonic set at 14.4 Hz, due to the main rotor of the helicopter before CPA. The main rotor sound of the helicopter near CPA is not there any longer. The harmonic set at 9.6 Hz is not easily attributable to an acoustic source because it shares all its prominent harmonics with the previous harmonic set. It is plausible that the main rotor harmonics of the helicopter near CPA are so weak and so close to the harmonics of the helicopter before CPA that they are drowned by them, and therefore are not visible.

*********** PROMINENT-HARM-SETS **************

Fund-freq =   14.7
Peaks at frequencies   (15 29 45 58)
fund-peak-present?   YES | fund-peak-prominent?   YES
Number of prominent harmonics   2 | Number of harmonics   4
avg-slack =   0.0413 | total-slack =   0.1653
avg-prom =   5.3 | total-prom =   21.3

**************** Extraneous peaks ******

freq   power | prom |gen-prom| hor-dist| characterization
-------|-------|-------|--------|---------|-------------------
  6   28411.7 | 2.5 | NIL | NIL | (IN-FUND-FREQ-RANGE STRONG)
 53   6272.0 | 2.9 | 4.6 | 14 | (LARGE-AVG-PROMINENCE SPLIT-PROMINENT)

Harmonic
Sets
&
Extraneous
Peaks

freq   power | prom |gen-prom| hor-dist| characterization
-------|-------|-------|--------|---------|-------------------
  6   28411.7 | 2.5 | NIL | NIL | (IN-FUND-FREQ-RANGE STRONG)
 16   38176.9 | 5.1 | NIL | NIL | (LARGE-AVG-PROMINENCE IN-FUND-FREQ-RANGE
                                           LOCALLY-STRONG STRONG)
 21   5344.9 | 1.4 | NIL | NIL | NIL
 25   4444.3 | 1.2 | NIL | NIL | NIL
 29   30847.5 | 12.6 | 12.6 | 10 | (LARGE-AVG-PROMINENCE STRONG)
 40   3706.5 | 2.7 | NIL | NIL | NIL
 45   4678.6 | 2.1 | NIL | NIL | NIL
 48   3168.4 | 1.1 | NIL | NIL | NIL
 53   6272.0 | 2.9 | 4.6 | 14 | (LARGE-AVG-PROMINENCE SPLIT-PROMINENT)
 58   1753.2 | 1.6 | NIL | NIL | NIL
 62   982.7 | 1.1 | NIL | NIL | NIL
 66   1832.0 | 1.3 | NIL | NIL | NIL
 72   1481.3 | 1.3 | NIL | NIL | NIL

Peaks

Linear spectrum
38200.9

0.0

3   Frequency (Hz)   78

Numeric
Spectrum

Figure 3.9: Extended spectrum for single helicopter far before CPA, main rotor frequency range.

92

```
*********** PROMINENT-HARM-SETS **************

Fund-freq =     91.9
Peaks at frequencies   (92 184 276 368 460 551 642 826 911)
fund-peak-present?           YES       | fund-peak-prominent?    YES
Number of prominent harmonics   6       | Number of harmonics     9
avg-slack =                    0.0146   | total-slack =       0.1311
avg-prom =                     19.6     | total-prom =        176.8

Fund-freq =     94.8
Peaks at frequencies   (92 191 286 382 477 572)
fund-peak-present?           YES       | fund-peak-prominent?    YES
Number of prominent harmonics   5       | Number of harmonics     6
avg-slack =                    0.0245   | total-slack =       0.1471
avg-prom =                     29.8     | total-prom =        176.6

**************** Extraneous peaks ******

freq·  power   | prom |gen-prom| hor-dist| characterization
-----|---------|------|--------|---------|----------------------------
 118   4063.6  |  5.3 |   7.1  |    15   | (PROMINENT-BUT-MAYBE-RIPPLED)
 403     48.0  |  5.4 |   7.1  |    13   | (PROMINENT-BUT-MAYBE-RIPPLED)
 587     41.6  |  2.8 |   7.0  |    18   | (LOCALLY-STRONG PROMINENT-BUT-MAYBE-RIPPLED)
```
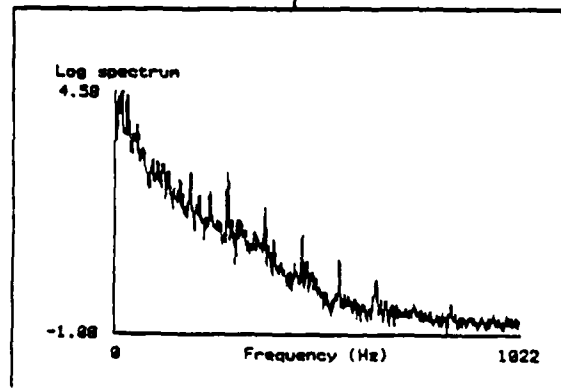
Harmonic Sets & Extraneous Peaks

```
freq    power   | prom |gen-prom| hor-dist| characterization
-----|----------|------|--------|---------|----------------------------
 88    4010.0   |  1.8 |  NIL   |   NIL   | (IN-FUND-FREQ-RANGE)
 92  183546.3   | 51.5 |  51.5  |    11   | (LARGE-AVG-PROMINENCE IN-FUND-FREQ-RANGE
                                            LOCALLY-STRONG)
 100   3487.7   |  1.3 |  NIL   |   NIL   | (IN-FUND-FREQ-RANGE)
 104   2578.0   |  1.9 |  NIL   |   NIL   | (IN-FUND-FREQ-RANGE)
 118   4063.6   |  5.3 |   7.1  |    15   | (PROMINENT-BUT-MAYBE-RIPPLED)
 184  19896.6   |  9.4 |   9.4  |    18   | (PROMINENT-UNBALANCED
                                            PROMINENT-BUT-MAYBE-RIPPLED)
 191  17219.3   |  9.6 |   9.6  |     9   | (LOCALLY-STRONG PROMINENT-BUT-MAYBE-RIPPLED)
 276   2513.4   | 23.0 |  23.0  |    11   | (LARGE-AVG-PROMINENCE LOCALLY-STRONG)
 286   1116.4   | 12.8 |  12.8  |    13   | (LARGE-AVG-PROMINENCE)
 368    329.8   | 19.0 |  19.0  |    15   | (LARGE-AVG-PROMINENCE)
 382   1436.3   | 81.2 |  81.2  |    18   | (LARGE-AVG-PROMINENCE LOCALLY-STRONG)
 403     48.0   |  5.4 |   7.1  |    13   | (PROMINENT-BUT-MAYBE-RIPPLED)
 460    146.5   | 22.9 |  22.9  |    21   | (LARGE-AVG-PROMINENCE LOCALLY-STRONG)
 477     77.4   | 18.0 |  18.0  |    12   | (LARGE-AVG-PROMINENCE)
 551    106.2   | 28.4 |  28.4  |    18   | (LARGE-AVG-PROMINENCE LOCALLY-STRONG)
 572     37.8   |  5.5 |  11.8  |    16   | (LARGE-AVG-PROMINENCE SPLIT-PROMINENT
                                            PROMINENT-BUT-MAYBE-RIPPLED)
 587     41.6   |  2.8 |   7.0  |    18   | (LOCALLY-STRONG PROMINENT-BUT-MAYBE-RIPPLED)
 642    163.9   | 17.9 |  17.9  |    18   | (LARGE-AVG-PROMINENCE LOCALLY-STRONG)
 698     18.6   |  3.3 |  NIL   |   NIL   | (LOCALLY-STRONG)
 826     18.6   |  2.4 |   7.8  |    44   | (LOCALLY-STRONG)
 890      8.8   |  3.6 |  NIL   |   NIL   | (LOCALLY-STRONG)
 911      3.1   |  4.4 |   8.1  |    24   | (PROMINENT-BUT-MAYBE-RIPPLED)
```

Peaks

```
Log spectrum
5.26




-9.747
     0         Frequency (Hz)          1022
```

Numeric Spectrum

Figure 3.10: Extended spectrum for a two-helicopter signal, tail rotor frequency range.

```
********** PROMINENT-HARM-SETS **************

Fund-freq =      9.6
Peaks at frequencies   (29 38 58)
fund-peak-present?              NO      | fund-peak-prominent?    NO
Number of prominent harmonics    2      | Number of harmonics      3
avg-slack =                 0.0328      | total-slack =        0.0983
avg-prom =                     6.6      | total-prom =          19.9

Fund-freq =     14.4
Peaks at frequencies   (14 29 44 58)
fund-peak-present?             YES      | fund-peak-prominent?   YES
Number of prominent harmonics    4      | Number of harmonics      4
avg-slack =                 0.0348      | total-slack =        0.1391
avg-prom =                    16.7      | total-prom =          66.9

*************** Extraneous peaks ******

freq   power  |  prom |gen-prom| hor-dist| characterization
--------------|-------|--------|---------|-------------------
NIL
```
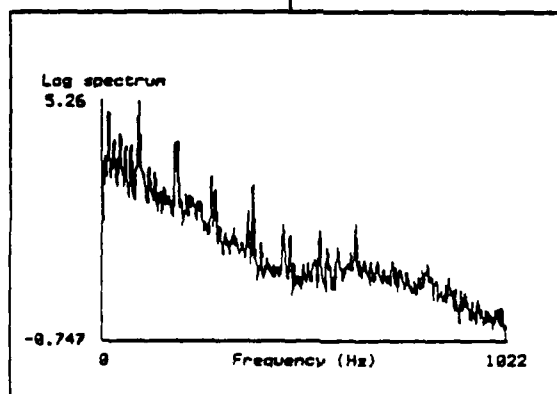
Harmonic
Sets
&
Extraneous
Peaks


```
freq   power    | prom |gen-prom| hor-dist| characterization
----------------|------|--------|---------|-------------------
  4  5083.1   |   1.3 |  NIL  |  NIL  | NIL
  7  7073.2   |   2.3 |  NIL  |  NIL  | (IN-FUND-FREQ-RANGE)
 14 94380.3   |  39.8 |  39.8 |   10  | (LARGE-AVG-PROMINENCE IN-FUND-FREQ-RANGE
                                        LOCALLY-STRONG STRONG)
 23  4320.3   |   1.4 |  NIL  |  NIL  | NIL
 29 18742.6   |   7.1 |   7.1 |   10  | (LARGE-AVG-PROMINENCE)
 38  3814.3   |   1.8 |  NIL  |  NIL  | NIL
 44 26734.6   |   8.9 |   8.9 |    8  | (LARGE-AVG-PROMINENCE LOCALLY-STRONG)
 50  3948.1   |   1.7 |  NIL  |  NIL  | NIL
 58 13844.3   |  10.9 |  10.9 |    9  | (LARGE-AVG-PROMINENCE)
 63  1204.8   |   1.3 |  NIL  |  NIL  | NIL
```

Peaks
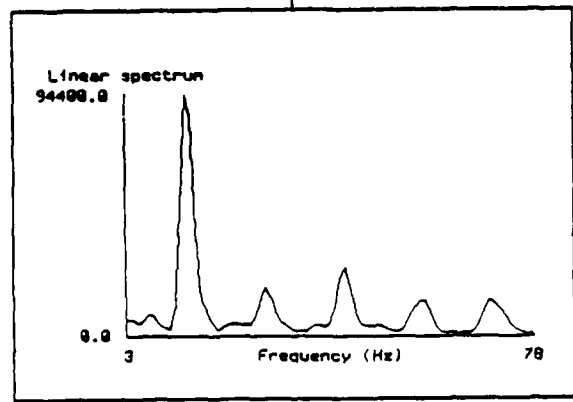
Numeric
Spectrum

Figure 3.11: Extended spectrum for a two-helicopter signal, main rotor frequency range.

3.1.3 The Extended Spectrum as a Data Abstraction

The extended spectrum is implemented as an abstract data type [Liskov & Zilles 1974, Kopec1980]. An abstract data type separates the use from the implementation of a data object, implemented in terms of more primitive objects: The set of allowed operations on instances of the abstract data type together with their usage specifications is the external view of the data type, whereas the actual details of how the object is represented in terms of more primitive objects is the internal view of the data type. Kopec was the first to advocate representation of signals as abstract data types, and used the term "closure model" for signals to indicate the representation of a specific signal as a function of several arguments, one of which is the independent variable plus several values for the function arguments (excluding the independent variable). Applying the function to a specific value of the independent variable and the values for the rest of its arguments generates the corresponding value of the signal.

In the case of the extended spectrum, in addition to separating use from implementation of a signal, we need to accommodate various degrees of suppression of detail (spectrum, spectral peaks, harmonically related peak sets), or levels of abstraction. This requires extension of the function arguments of the closure model into several parameter sets that are necessary for computing the transitions between successive levels of detail. We thus have three different viewpoints towards a signal, the first two inherited from the view presented in [Kopec1980]:

1. *Internal representation*, which is to be separated from the use of the signal as seen from the outside.

2. *Use of the signal.* Specifies operations that an outside user may perform on the signal. This includes fetching certain values associated with the signal at different levels of detail and fetching the parameter object for generating the most detailed description (i.e. estimating the numeric spectrum from the

95
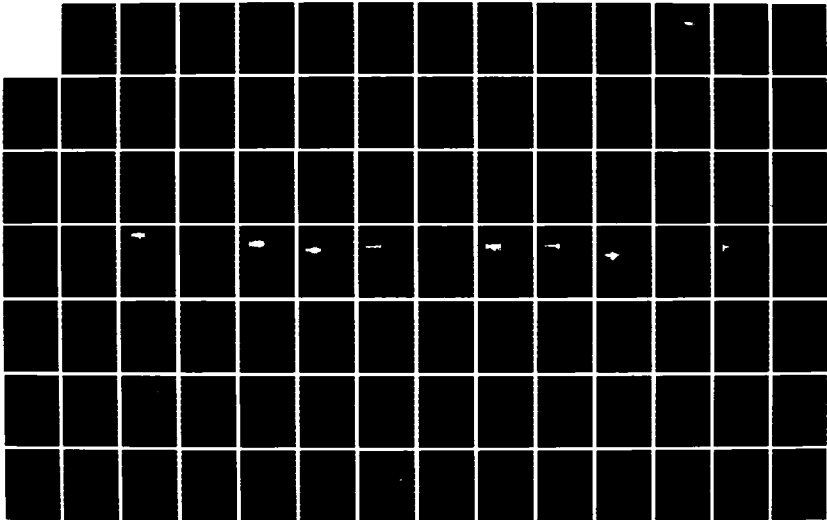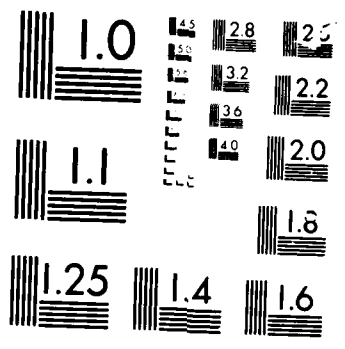
MICROCOPY

time waveform).

3. *Transition between successive levels of detail.* In contrast with the view of [Kopec1980], our signal representation is redundant, in the sense that if we have the most detailed level of description (i.e. the time signal), we can compute the spectrum and all its higher levels of description (i.e. peaks and harmonically related peak sets) from it. The algorithms associated with the transition between successive levels of detail have their own parameter objects.

The abstract data type we have defined for representing the extended spectrum holds the pieces of information shown in Figure 3.12. Two basic kinds of information are present, signal information at multiple levels of abstraction and parameter objects for the spectrum estimation and the transitions from the lower to the higher levels of abstraction (parameter objects are shown indented in the figure). Strictly speaking, the time signal and the parameter objects are a complete specification of the extended spectrum.

The spectral estimation parameter object includes the parameters used to estimate the spectrum from the time sequence. These parameters are:

- Starting time of the first block to be averaged.

- Individual block length.

- Degree of overlap between successive blocks.

- Total number of blocks.

The peak classification parameter object includes the classification thresholds for the various categories of peaks, as well as a selection of applicable empirical classification criteria. The latter need to be specified because not all criteria are applicable to all situations. For example, a different selection of criteria is applicable when we are classifying main-rotor harmonics and tail-rotor harmonics.

96

The harmonic set sorting parameter object, includes all parameters for determining the list of harmonic peak sets and the list of extraneous peaks from the list of characterized peak objects. These parameters are: the maximum slack used by the approximate multiple test, maximum distanced between two fundamental frequencies to be considered the same for purposes of merging two harmonic sets with similar fundamental frequencies, the fundamental frequency range of interest and the selection of rules for pruning harmonic sets.

Whether the signal information part of the extended spectrum is remembered ("memoized", according to the terminology of [Abelson & Sussman 1985]) after the first time it is computed for immediate access when it is needed later, is a matter of efficiency considerations, or of the desired tradeoff between memory space and computation time. Our implementation on the Lisp machine was not constrained by memory space, therefore we chose to memoize all signal information. In an operational system with both memory and time constraints, a more careful study is necessary of exactly what computation should be delayed. As we mentioned above, the issue we raise here is similar to the notion of "memoization" [Abelson & Sussman 1985], according to which the results of function calls in a programming language are saved to avoid recomputing the function, if it is called again with the same functional arguments.

The computations discussed in detail in the previous sections are defined as operations on the extended spectrum. They are summarized in Figure 3.13. These operations cannot be strictly classified under the internal or external view of the extended spectrum object. If the object is viewed as immutable, then they can be considered part of the internal view of the object. However, if we want to provide the user of the object with the ability to modify some of the parameter objects and *compute the modified extended spectrum*, we must provide him with access to the operations of Figure 3.13.
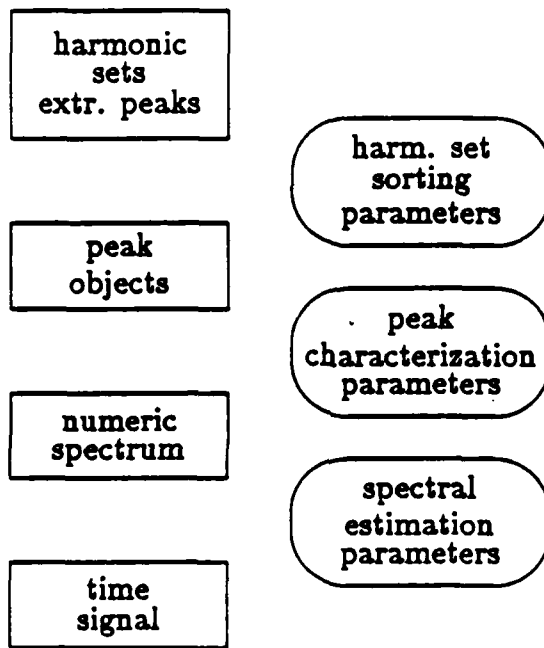
Figure 3.12: The extended spectrum data type

- Create an extended spectrum. This involves setting up an instance of the extended spectrum data type with the parameter objects filled in.

- Find peaks. This involves finding the peaks, creating a peak object for each, and finally filling in the appropriate slot with the ordered peak list.

- Classify peaks. This involves classifying the peaks in the peak list and filling in the classification slot of each peak object.

- Find fundamental frequencies. This involves going through all the peaks and applying to each the criteria for possibly extracting one or more fundamental frequency estimates out of each.

- Create tentative harmonic sets. This involves going through the fundamental frequency list and associate with each fundamental frequency all peaks that are its approximate multiples.

- Prune tentative harmonic sets. This involves going through all harmonic sets and extract those that seem to have strong evidence for being visually prominent.

- Find extraneous peaks. This involves going through all peaks and based on their classification and (non)membership of at least one pruned harmonic set classifying them as extraneous.

Figure 3.13: Operations defined on the extended spectrum data type

3.1.4 Perspective on the approach

Rule implementation of classification

In our rule implementation (built in the rule language YAPS [Allen1983]), a rule corresponds to a single independently applied binary classification problem. The left-hand side of the rule is used to bind the various rule parameters, while the test part is used to test for the conditions of the classification problem. If the test succeeds, then the action side of the rule adds the corresponding class to the current peak object (by side effect). If the test fails, nothing happens.

Rules that apply different criteria to the same task, e.g. peak classification, are grouped together into a rule base. Activation of a rule base is done via a function call: the function first inserts into the database the appropriate facts concerning the peak object being classified and the various threshold parameters, and then it activates the rule base. Rules are triggered by the existence of facts that match their left-hand side. After all the rules have fired, the rule base operation returns. Its goal has been accomplished by altering the peak objects by side effect.

A rule captures the intuitive notion of the corresponding binary classification problem, without specifying values for the thresholds involved. Such separation is useful, because the threshold values may vary with the kind of data at hand (for example low or high frequency spectra), but the same intuition may apply in many situations, therefore allowing the use of the same rules in more than one situation.

Embedded classification

In our system for computing extended spectrum representations, we perform classification at the low abstraction levels (e.g. peaks), and use its results when generating descriptions at higher abstraction levels. For example, whether a peak is being used as a source of candidate fundamental frequencies depends on its classification. We

thus embed (and solve) classification subproblems into an otherwise straightforward grouping procedure, for the purpose of reducing the number of items to be grouped (i.e. reducing the search space). The whole process of describing a harmonic spectrum at several levels of abstraction can be viewed as a process of search in the power set of all peaks for harmonic peak sets. By using symbolic peak characterization, candidate fundamental frequency selection, and harmonic set pruning we are able to cut down the size of the search space (which is enormous, considering that in typical spectra we have of the order of a few hundred peaks, giving rise to power sets of peaks (i.e. the sets of all possible subsets of peaks) of size of the order of 2^{100}).

Other uses of the rule approach

The rule approach was found convenient in splitting and merging hypothesized harmonic sets in harmonic set pruning. Rules operate differently in this case, because they act on a database of harmonic sets by constantly modifying it (by triggering on one, two or more sets in the database and replacing them by one or more modified sets, as appropriate) until no rule fires. This is how the rules shown in Figure 3.7 operate. The full implications of this approach are not well understood yet (issues of infinite looping, being able to have an indeterminate number of sets in the rules without explosion in the number of links in the discrimination net). Better understanding of these issues may lead to the design of appropriate control structures for rule based systems specifically for signal processing (see [Dove1986]).

101

3.2 Applications of matching of harmonic spectra

In the previous sections of this chapter we introduced the extended spectrum as a representation at multiple levels of abstraction for harmonic spectra. We are now going to illustrate the usefulness of the extended spectrum in three problems involving matching of helicopter data.

3.2.1 Correlation of harmonic spectra for faulty channel detection

In this section, we apply the notion of extended spectrum, by treating identification and characterization of faulty channels as a problem of matching their spectra in the appropriate representation.

In theory, when a plane wave is sensed by an array of closely spaced microphones (i.e. a few meters from each other), the only difference between signals from different microphones is a time lag, which affects only the phase of the spectrum of the signals. In practice, however, power spectra from different microphones are not the same. As an example, Figure 3.14 shows the time waveforms and power spectra of three channels from a typical helicopter data recording. Frequencies of signal peaks are only approximately equal in different channels, with frequency differences up to 1 or 2 Hz between corresponding peaks. Peaks due to noise are not coherent across sensors. Signal peaks do not always have similar amplitude patterns in different channels, i.e. different harmonics may be the strongest in different channels.

Furthermore, interference may be present in one or more channels due to mechanical or electrical problems, as in channel 6 of Figure 3.14, which has a strong 60-Hz interference in the form of strong harmonics of 60 Hz. Figure 3.15 shows the list of prominent peaks of the spectra shown in Figure 3.14. For the purposes of faulty channel detection, prominent peaks are those with one or more characterizations in

the set { LARGE-AVG-PROMINENCE, PROMINENT-BUT-MAYBE-RIPPLED, LOCALLY-STRONG }, as they were defined earlier in this chapter.
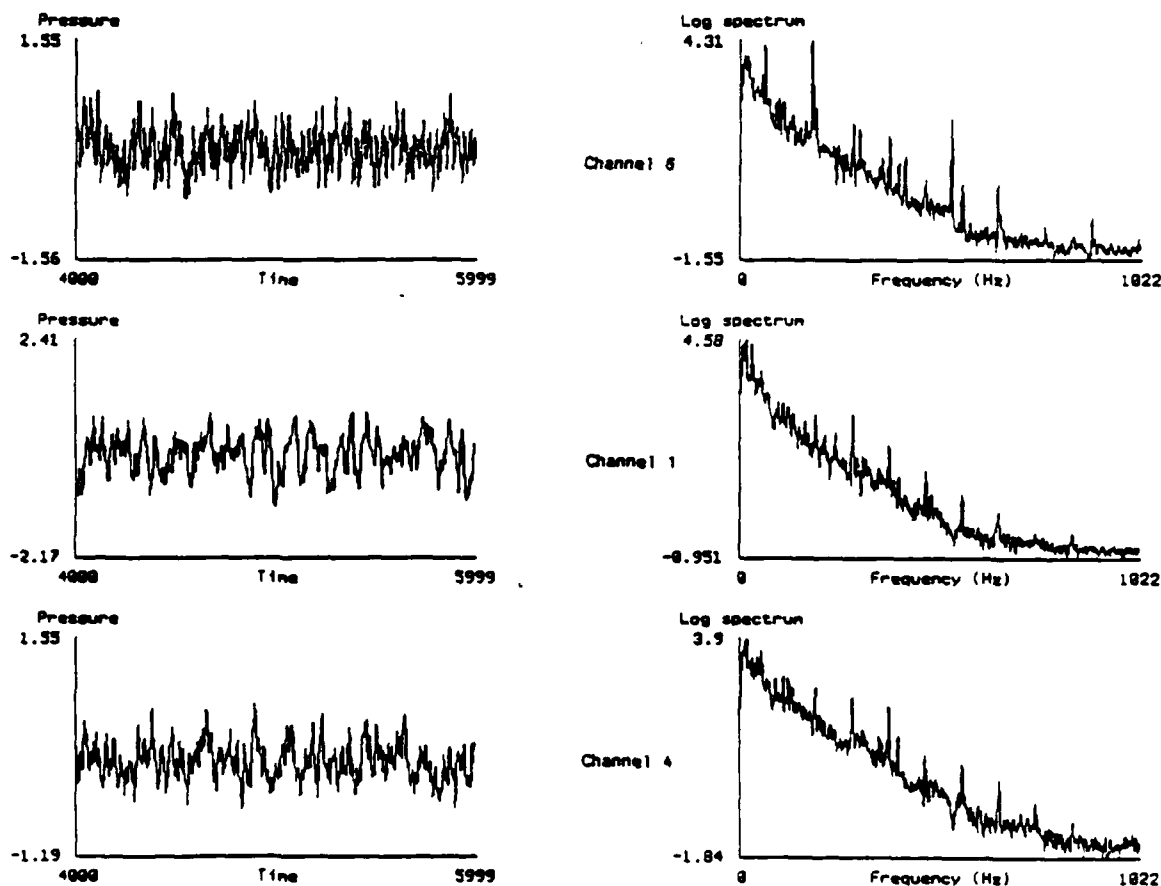
Figure 3.14: Time waveforms and spectra of channels 6,1 and 4, with channel 6 having several strong peaks not present in any other channel due to a 60-Hz interference

Inspection of the time waveforms can detect and characterize certain faults, such as microphone saturation, consisting of a signal with a high constant value, as shown in Figure 3.16. Other faults cannot always be visually detected from the time waveform, as in the case of a channel without signal, i.e. consisting of electrical noise only (Figure 3.17), when the signal is weak and does not give rise to a clearly visible periodic pattern in the waveform. Finally, there are faults that need more elaborate reasoning for their detection and characterization, such as the

103

Channel 6

freq	power	prom	gen-prom	hor-dist	characterization
81	138.1	1.6	NIL	NIL	(IN-FUND-FREQ-RANGE)
88	473.5	5.6	5.6	8	(IN-FUND-FREQ-RANGE PROMINENT-BUT-MAYBE-RIPPLED)
95	589.8	7.4	7.4	6	(IN-FUND-FREQ-RANGE PROMINENT-BUT-MAYBE-RIPPLED)
103	392.1	1.8	NIL	NIL	(IN-FUND-FREQ-RANGE)
106	444.0	2.0	6.5	13	(IN-FUND-FREQ-RANGE PROMINENT-BUT-MAYBE-RIPPLED)
180	20626.1	316.6	316.6	12	(LARGE-AVG-PROMINENCE LOCALLY-STRONG)
247	32.4	4.0	10.8	20	(LARGE-AVG-PROMINENCE SPLIT-PROMINENT PROMINENT-BUT-MAYBE-RIPPLED)
284	132.2	30.4	30.4	12	(LARGE-AVG-PROMINENCE LOCALLY-STRONG)
300	85.6	12.3	12.3	6	(LARGE-AVG-PROMINENCE)
360	14.5	4.1	5.8	22	(PROMINENT-BUT-MAYBE-RIPPLED)
379	55.3	34.1	34.1	9	(LARGE-AVG-PROMINENCE LOCALLY-STRONG)
403	11.0	7.7	7.7	7	(PROMINENT-BUT-MAYBE-RIPPLED)
420	16.2	25.1	25.1	18	(LARGE-AVG-PROMINENCE)
474	3.9	4.1	6.2	12	(LOCALLY-STRONG PROMINENT-BUT-MAYBE-RIPPLED)
540	160.8	476.2	476.2	15	(LARGE-AVG-PROMINENCE LOCALLY-STRONG)
569	2.8	22.5	22.5	12	(LARGE-AVG-PROMINENCE)
660	2.8	16.0	16.0	7	(LARGE-AVG-PROMINENCE LOCALLY-STRONG)
719	0.2	2.2	NIL	NIL	(LOCALLY-STRONG)
780	0.2	2.7	5.9	123	(LOCALLY-STRONG)
853	0.1	1.6	NIL	NIL	(LOCALLY-STRONG)
900	0.4	7.5	7.5	7	(LOCALLY-STRONG PROMINENT-BUT-MAYBE-RIPPLED)
967	0.1	2.0	NIL	NIL	(LOCALLY-STRONG)

Channel 1

freq	power	prom	gen-prom	hor-dist	characterization
81	364.4	1.8	NIL	NIL	(IN-FUND-FREQ-RANGE)
89	660.0	2.1	NIL	NIL	(IN-FUND-FREQ-RANGE)
95	1013.5	2.8	NIL	NIL	(IN-FUND-FREQ-RANGE)
99	316.3	1.0	NIL	NIL	(IN-FUND-FREQ-RANGE)
103	458.9	1.2	NIL	NIL	(IN-FUND-FREQ-RANGE)
107	894.8	2.4	NIL	NIL	(IN-FUND-FREQ-RANGE)
190	498.7	9.0	9.0	14	(LOCALLY-STRONG PROMINENT-BUT-MAYBE-RIPPLED)
240	176.2	4.3	5.8	14	(LOCALLY-STRONG PROMINENT-BUT-MAYBE-RIPPLED)
284	522.9	27.3	27.3	8	(LARGE-AVG-PROMINENCE LOCALLY-STRONG)
379	80.7	16.3	16.3	13	(LARGE-AVG-PROMINENCE LOCALLY-STRONG)
403	13.1	5.7	6.4	17	(PROMINENT-BUT-MAYBE-RIPPLED)
474	17.4	14.8	14.8	9	(LARGE-AVG-PROMINENCE LOCALLY-STRONG)
569	4.7	10.8	10.8	9	(LARGE-AVG-PROMINENCE LOCALLY-STRONG)
663	1.8	4.3	7.8	21	(LOCALLY-STRONG PROMINENT-BUT-MAYBE-RIPPLED)
759	0.4	1.5	NIL	NIL	(LOCALLY-STRONG)
853	0.4	2.2	NIL	NIL	(LOCALLY-STRONG)
960	0.2	1.7	NIL	NIL	(LOCALLY-STRONG)

============== Relevant peaks

Channel 4

freq	power	prom	gen-prom	hor-dist	characterization
88	718.0	5.4	NIL	NIL	(IN-FUND-FREQ-RANGE)
94	305.6	1.9	NIL	NIL	(IN-FUND-FREQ-RANGE)
107	809.7	7.1	7.1	11	(IN-FUND-FREQ-RANGE LOCALLY-STRONG PROMINENT-BUT-MAYBE-RIPPLED)
120	746.4	5.7	5.7	10	(PROMINENT-BUT-MAYBE-RIPPLED)
190	441.7	14.6	14.6	12	(LARGE-AVG-PROMINENCE LOCALLY-STRONG)
284	227.2	13.9	13.9	8	(LARGE-AVG-PROMINENCE LOCALLY-STRONG)
353	22.0	5.0	6.8	18	(PROMINENT-BUT-MAYBE-RIPPLED)
379	133.6	29.7	29.7	9	(LARGE-AVG-PROMINENCE LOCALLY-STRONG)
403	22.9	12.4	12.4	7	(LARGE-AVG-PROMINENCE)
474	7.3	18.4	18.4	13	(LARGE-AVG-PROMINENCE LOCALLY-STRONG)
569	4.0	9.8	9.8	10	(LOCALLY-STRONG PROMINENT-BUT-MAYBE-RIPPLED)
665	1.4	22.5	22.5	16	(LARGE-AVG-PROMINENCE LOCALLY-STRONG)
750	0.4	5.4	6.0	18	(LOCALLY-STRONG PROMINENT-BUT-MAYBE-RIPPLED)
853	0.1	4.2	7.5	34	(LOCALLY-STRONG)
912	0.1	2.8	NIL	NIL	(LOCALLY-STRONG)

Figure 3.15: Lists of prominent peaks of spectra of channels 6 (the faulty one), 1 and 4 shown in the previous figure.

60 Hz interference of Figure 3.14, which is best done by inspection of the power spectra. In our protocols, the human subject identified the fault by matching high level spectral characteristics of different channels and by isolating the channel that significantly departed from the majority of the channels.
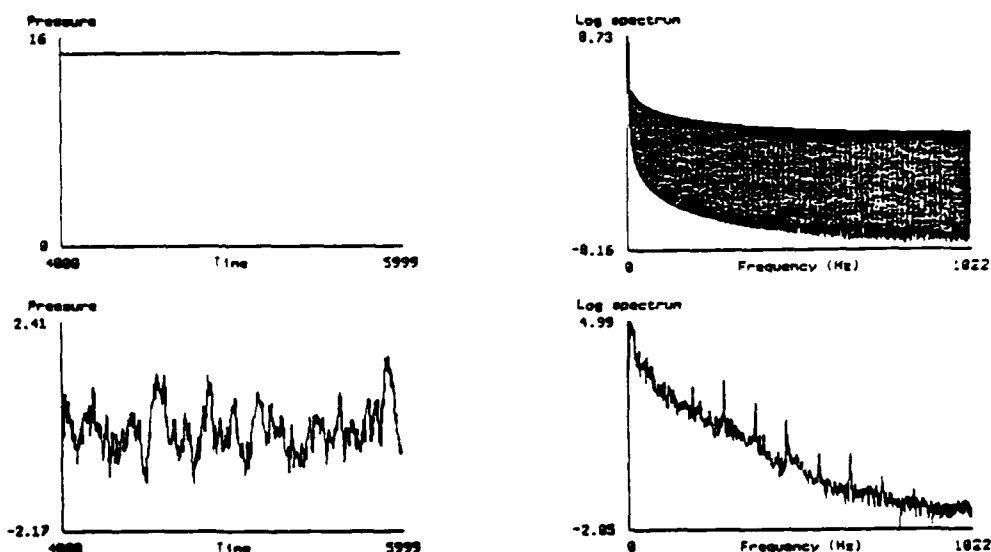

Figure 3.16: Saturated channel 0 and channel 1 at the same time. Note the vertical scale of the two plots. Vertical scale is [0 15] for channel 0 and [-2.2 2.4] for channel 1.

To illustrate the fact that the conditions describing "difficult" channel faults, such as a 60 Hz interference, can be expressed in terms of higher levels of abstraction, and therefore detected at those levels, we implemented a program for detecting faulty channels based on their different peak structure. Figure 3.18 shows a block diagram of the program.

The program operates on the basis of prominent peaks, i.e. peaks that visually "stand out". The peak characterizations obtained through the methods described earlier in this chapter are used to determine whether a peak is prominent.
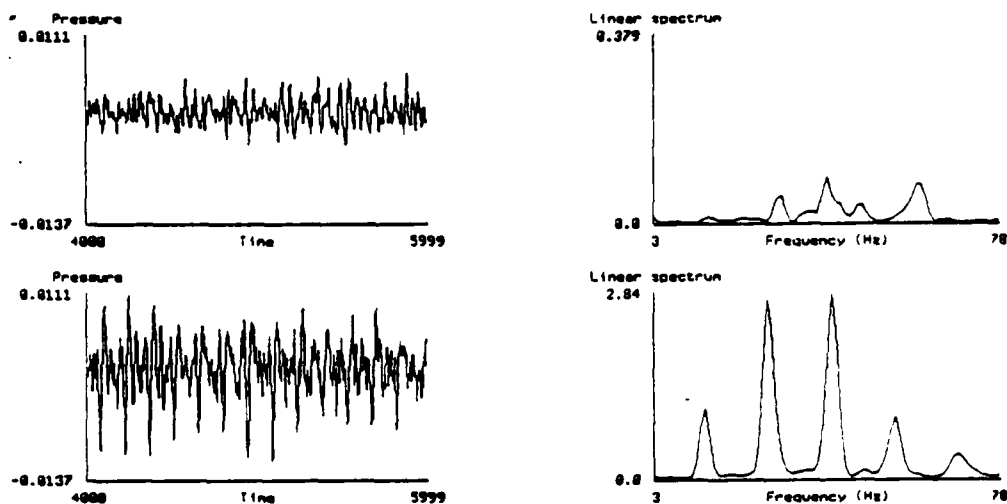
Operation of the program proceeds as follows.

Figure 3.17: Channel with a very weak signal, compared to another channel at the same time with a strong signal. Both time-waveforms (on the left) and log spectra (on the right) are shown.
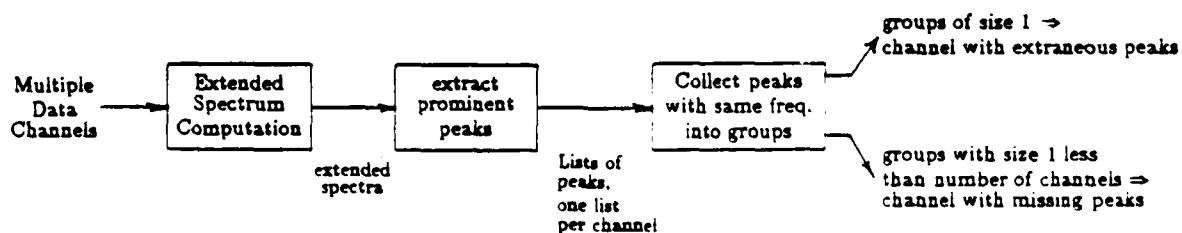

Figure 3.18: Block diagram of the program for faulty channel detection based on peak structure

A list of prominent peaks for each channel is formed based on the channel's extended spectrum. Then prominent peaks from all channels are collected into groups, where peaks in the same group have the "same" frequency (i.e. frequencies that differ very little. The largest permissible difference depends on the type of spectrum being analyzed. We defined it to be 3 Hz in the case of tail rotor frequency ranges). These groups collect together peaks from several channels that are apparently due to the same physical cause.

Missing and extraneous peaks identified from groups of size one or one less than the total number of channels are the starting point for detecting faulty channels, under the assumption of a single faulty channel. Moreover, the peak information used to detect a faulty channel can also be used to characterize the fault in terms of its spectral characteristics, although this was not implemented in this thesis.

As an example, consider the case of a single helicopter source and eight channels of data. The list of prominent peaks for each channel in this case consists of the list of prominent helicopter harmonics in the corresponding channel. The groups defined above consist of harmonics of the same order from different channels. Ideally, all groups will have eight members, one from each channel. If in a single channel the helicopter signal is very weak or missing altogether or the channel has some kind of broadband interference, one or more of the groups will have seven members. If a single channel has a strong harmonic interference, several groups will show up with only one member, corresponding to one harmonic of the interference.

The result of the program is now shown on the eight channels of real helicopter data, one of which (channel 6) has a strong 60 Hz interference. Part of the data was shown in Figure 3.14, including channel 6 and two other channels (1 and 4) for comparison.

Figure 3.19 shows the groups of peaks of different channels with the same frequency formed from the lists of prominent peaks of channels 1-8. Each peak is represented by an X at the intersection of its frequency and the channel it belongs

107

to. Only groups of size 1 or 7 are shown in the figure, corresponding to peaks present or missing in only one channel respectively. Note that channel 6 has eight peaks not present in any other channel and a peak at 189 Hz missing. No other channel has more than one peak not present in any other channel.

FREQ	1	2	3	4	5	6	7	8
88						X		
120				X				
162					X			
180						X		
189	X	X	X	X	X		X	X
213								X
247						X		
300						X		
360						X		
420						X		
540						X		
660						X		

Figure 3.19: Groups of peaks of different channels with the same frequency in table form. Each peak is denoted by an X at the intersection of its frequency and the channel it belongs to.

Figure 3.20 shows the channels with missing peaks (a) or peaks present only in those channels (b). Part (b) of this figure in addition shows the characterizations of each peak present in only one channel. It is worth noting that six out of the eight peaks present only in channel 6 have the characterization LARGE-AVG-PROMINENCE. Furthermore, seven of those peaks are harmonics of 60 Hz, indicating the presence of a strong 60 Hz interference.

```
***** Single missing peaks ******

Channel    6 is missing a peak at    189 Hz


***** Single present peaks ******

Channel    4  has an extraneous peak at    120 Hz
    with characterization (PROMINENT-BUT-MAYBE-RIPPLED)

Channel    5  has an extraneous peak at    162 Hz
    with characterization (PROMINENT-BUT-MAYBE-RIPPLED)

Channel    6  has an extraneous peak at     88 Hz
    with characterization (IN-FUND-FREQ-RANGE PROMINENT-BUT-MAYBE-RIPPLED)
Channel    6  has an extraneous peak at    180 Hz
    with characterization (LARGE-AVG-PROMINENCE LOCALLY-STRONG)
Channel    6  has an extraneous peak at    247 Hz
    with characterization (LARGE-AVG-PROMINENCE SPLIT-PROMINENT PROMINENT-BUT-MAYBE-RIPPLED)
Channel    6  has an extraneous peak at    300 Hz
    with characterization (LARGE-AVG-PROMINENCE)
Channel    6  has an extraneous peak at    360 Hz
    with characterization (PROMINENT-BUT-MAYBE-RIPPLED)
Channel    6  has an extraneous peak at    420 Hz
    with characterization (LARGE-AVG-PROMINENCE)
Channel    6  has an extraneous peak at    540 Hz
    with characterization (LARGE-AVG-PROMINENCE LOCALLY-STRONG)
Channel    6  has an extraneous peak at    660 Hz
    with characterization (LARGE-AVG-PROMINENCE LOCALLY-STRONG)

Channel    8  has an extraneous peak at    213 Hz
    with characterization (PROMINENT-BUT-MAYBE-RIPPLED)
```

Figure 3.20: Missing peaks and peaks present in only one channel with their characterizations, as produced by the implemented program.

The above experiment illustrates that some channel faults can be conveniently expressed, and therefore detected and characterized, at higher levels of abstraction, in this case peaks. Both the appropriate levels of abstraction and the conditions for the existence of various channel faults can be derived by observing human subjects perform the task of faulty channel detection using interactive signal processing.

3.2.2 Adjustment of spectral estimation parameters via matching of spectra

In the previous section we examined matching of spectral abstractions corresponding to spectra, which are ideally identical, for the purpose of identifying faults and interferences in particular channels. In this section, we perform matching of spectra obtained from the same time waveform but through different spectral estimation parameter setups. Spectra obtained through different parameter setups can be thought of as multiple levels of abstraction in the scale-space sense [Witkin1984], by viewing different amounts of periodogram averaging as different scales.

The purpose of matching here is to order spectra according to a specified criterion, thereby finding the "optimal" parameter setup according to that criterion. The criteria of interest are those that human subjects use based on a priori knowledge of general characteristics of the desired spectra.

One such criterion is "peakiness" of a harmonically related set of peaks. Relative peakiness of two versions of a harmonic set that is present in two spectra is determined by matching the two versions at multiple levels of abstraction, as shown in the diagram of Figure 3.21.

Initially, a decision is sought at the harmonic set level. If there is a clear answer at that level indicating that one harmonic set is "peakier" than the other, this is the final answer. If at the harmonic set level no clear decision can be made, matching
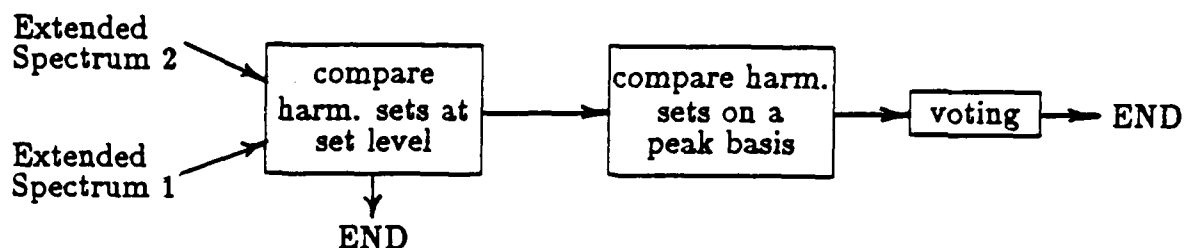
110

Figure 3.21: Block diagram of the determination of relative peakiness of two harmonic sets by matching at multiple levels of abstraction.

of the corresponding peaks is performed and voting takes place. If most peaks of one harmonic set are consistently more peaky than the corresponding peaks of the other, then again a final answer is possible in favor of the former harmonic set. If no clear decision can be made at this level, then the final answer is that the two harmonic sets have similar peakiness.

The program outlined in Figure 3.21 includes three sets of heuristics, implemented as separate YAPS rule bases.

The first rule base implements comparison of harmonic sets directly based on their features. Each rule implements a condition on features of the two harmonic sets that indicates that one harmonic set is more peaky than the other. The overall strategy is that in order to make a clear decision, two such conditions need to be satisfied, showing that set A is more peaky than set B, and no condition indicating that set B is more peaky than set A is satisfied. Figure 3.2.2 shows the list of heuristic conditions in the present status of the program. This list is not intended as the "best" collection of heuristics for the purpose, but as indicative of the nature of heuristics that are possible within the extended spectrum framework.

Set A is more peaky than set B if:

- A has a fundamental peak and B has none.

- A has a prominent fundamental peak and B has not.

- A has a lot more prominent harmonics than B.

- A has much higher average prominence than B.

- A has much higher total prominence than B.

Figure 3.22: Heuristic rules indicating that harmonic set A is more peaky than set B. At least two of them must be satisfied, and none must be satisfied in the opposite direction, in order to make a clear decision concerning whether set A is more peaky than set B.

The second rule base implements comparison of the features of two peaks for deciding which one is more peaky. Currently, the only basis for that decision is the prominence of the peaks, as the ratio of the peak power to the average power of the two enclosing valleys. The associated single rule declares that one peak is clearly peakier than the other if the ratio of the prominence of the two peaks is sufficiently high. The possible results of this rule base are that one peak is MORE, LESS, EQUALLY peaky as the other peak, or that a peak has NO-MATCH in the other harmonic set.

The third rule base implements the voting mechanism for deciding about the relative peakiness of two harmonic sets based on a peak-by-peak comparison. Figure 3.23 shows the rules currently in the program. Again, this list is intended to be indicative of the nature of the heuristics that can be conveniently expressed within the extended spectrum framework.

As an illustration of this approach, we present two examples of adjustment of the parameters of the spectral estimation by periodogram averaging. A data segment of length N is split into K sections of length M each, with 50 percent overlap, i.e. $N = (K+1) * (M/2)$. Each section of length M is multiplied by a

112

Set A is peakier than set B if:

- The number of peaks that are peakier in A is higher by at least 2 than the number of peaks that are peakier in B, and there is no more than one peak in B with no match in A.

- At least two peaks in A have no match in B, and no more than one peak in B has no match in A.

Figure 3.23: Heuristic rules implementing the voting mechanism for deciding about the relative peakiness of two harmonic sets based on a peak-by-peak comparison.

Hamming window and its periodogram is computed. Finally, the periodograms of all K sections are averaged to compute the power spectral estimate of the given data segment of length N. M is fixed to 0.5 s, thus fixing the resolution of the resulting power spectral estimate to 2 Hz. Hence, the parameter to be adjusted is the length of the data segment N.

In the first example, the signal characteristics are changing slowly, and at the same time the signal-to-noise ratio is low, corresponding to a helicopter flying in a straight path far before its closest point of approach. In this case, a large N, or equivalently a large K, is desirable for good noise suppression. On the other hand, signal frequencies change slowly, therefore a very large N is needed to introduce significant peak broadening due to frequency shifting with time. In this case, the "optimal" N is the largest N for which peak broadening is not significant. Figure 3.24 shows the resulting spectrum for several values of K. In comparing these spectra, we observe that the peaks due to the helicopter source become more distinct as K increases. The spectrum with K=44 is much stronger evidence of the presence of a helicopter source than the spectrum with K=2, because more harmonics of the helicopter source show up and they are much more visually prominent.

In the second example, the signal characteristics are changing rapidly, and the signal-to-noise ratio is high, corresponding to a helicopter flying in a straight path
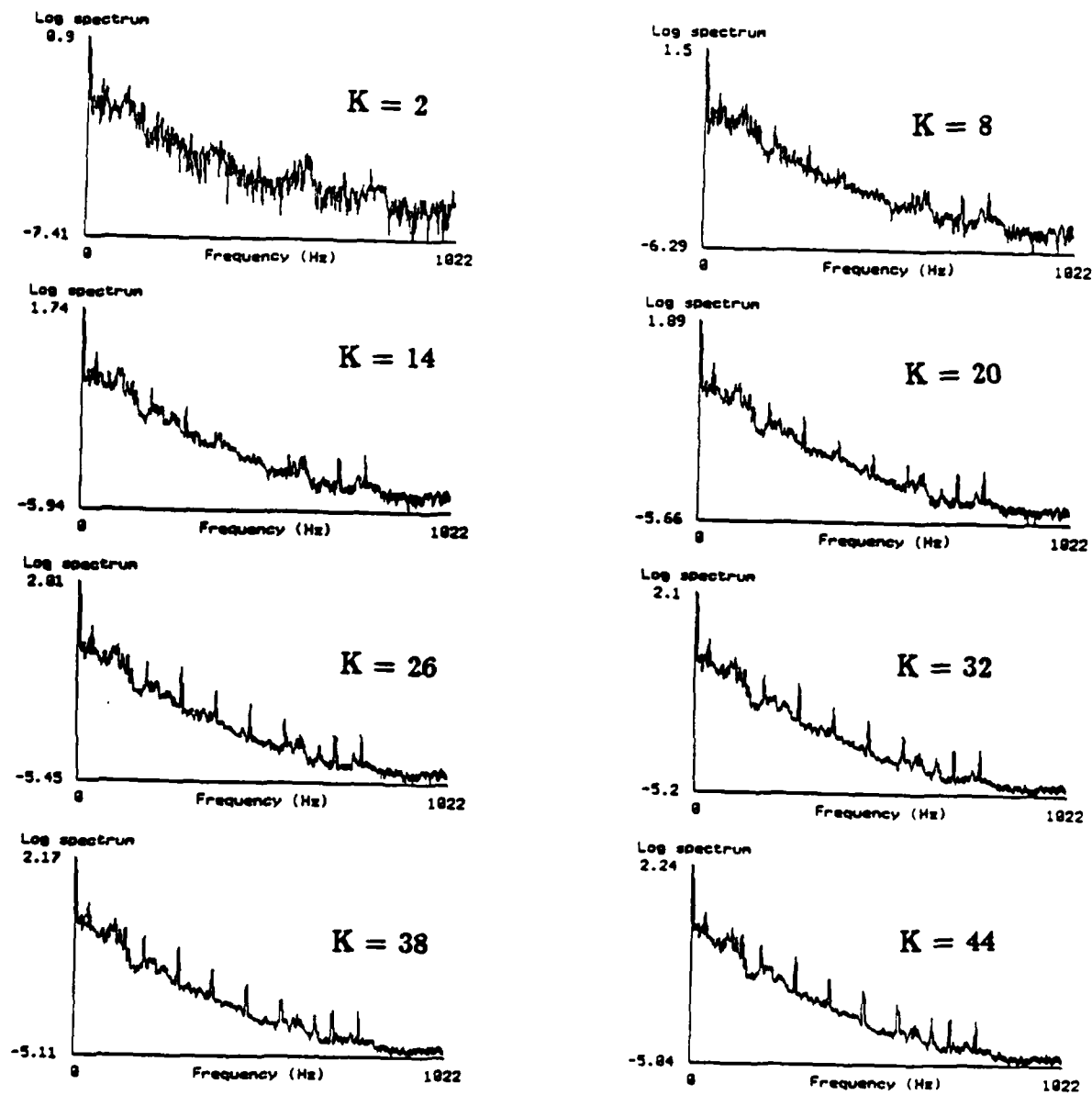
Figure 3.24: Acoustic spectra for a helicopter far before CPA and eight values for the number of periodograms averaged, 2,8,14,20,26,32,38,44.

near its closest point of approach, followed by another helicopter, which is still before its CPA. In this case a small N, or equivalently a small number of periodograms averaged, is needed to avoid smearing due to rapid change of signal frequencies. Since, however, the signal-to-noise ratio is high, sufficient noise suppression is obtained even with a small K. Figure 3.25 shows the resulting spectrum for several values of K. We notice substantial deterioration and merging in the peak patterns, especially the second and fourth harmonics of the two helicopters, as K increases.
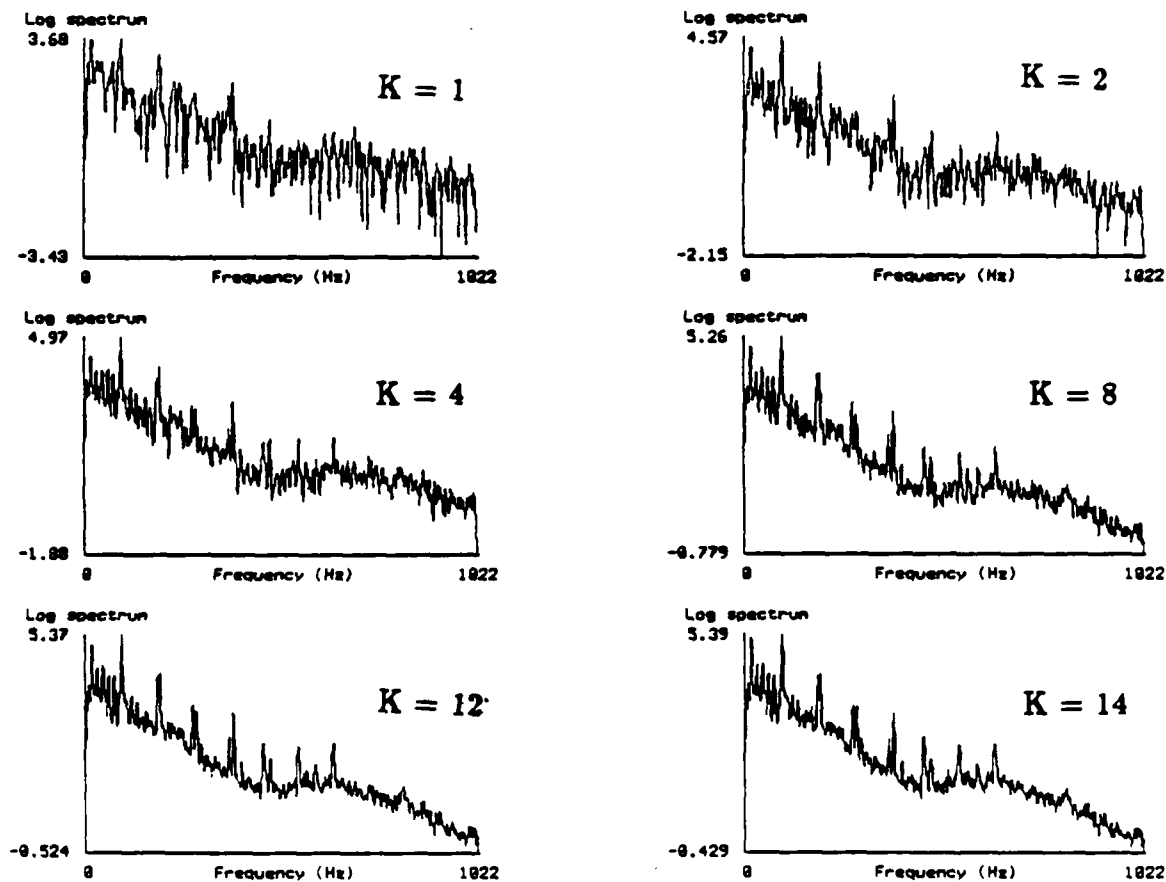

Figure 3.25: Acoustic spectra for two helicopters, one at CPA and the other before CPA, and six values for the number of periodograms averaged, 1,2,4,8,12,14

The results of the program for adjusting the parameters of the spectral estimation method based on the peakiness of the computed spectra on the above examples are shown next. The program inputs are:

1. a time waveform,

115

2. a "default" value for the single parameter to be adjusted, which is the number of periodograms averaged (block length and overlap are fixed), and

3. a value indicating the step change in the number of periodograms during the search process.

The program first computes the extended spectra for three values of the number of periodograms: the default minus the step, the default, and the default plus the step. Then it prompts the user to select a harmonic set from the spectrum obtained with the default value on which to focus the search. Then it applies the method outlined above for comparing harmonic sets based on their peakiness to two pairs of harmonic sets, consisting of the versions of the selected harmonic set in the default spectrum combined with each of the other two spectra.

If the middle harmonic set (obtained with the default number of periodograms) turns out to be the most peaky of the three, the search is finished, and the resulting number of periodograms is the default. Otherwise, one of two other values for the number of periodograms is selected as the default and the search proceeds recursively.

Figure 3.26 shows the table indicating the direction in which the search proceeds in all possible outcomes of the two comparisons of the default harmonic set with each of the other two. Entries whose outcome is MIDDLE indicate that the search is finished, those whose outcome is FIRST cause the new default value to be the current default minus the step, and those whose outcome is LAST cause the new default value to be the current plus the step. ERROR indicates a condition that is not expected, namely the default harmonic set to be the worst of the three. The search can be viewed as a search for a local maximum (of the peakiness of the harmonic set as a function of the number of periodograms averaged) by searching for a point where the derivative is negative in either direction, where the value of the derivative is determined qualitatively.

```
Number of periodograms:
A: default minus step
B: default
C: default plus step
```

A is _ peaky than B	B is _ peaky than C	Peakiest of A,B,C
LESS	MORE	B
EQUALLY	MORE	B
MORE	MORE	A
LESS	EQUALLY	B
EQUALLY	EQUALLY	B
MORE	EQUALLY	A
LESS	LESS	C
EQUALLY	LESS	C
MORE	LESS	ERROR

Figure 3.26: Table of outcome of the comparisons between harmonic set at the default number of periodograms and the same harmonic sets obtained with the default number plus or minus the search step.

Figures 3.27 and 3.28 show the trace of the program in two cases and different parameters.

In (a), the method is applied to the data where the helicopter is far before CPA. The default value is 8 periodograms averaged and the step is 6 periodograms. The focus is on the harmonic set with fundamental closest to 95 Hz, according to the value the user provided to the prompt of the program. The program successively adjusts the number of periodograms to 14, 20, 26, 32 and 38, which is the final answer. In all iterations except the last, the largest of the three values tried gave the "best" harmonic set in terms of peakiness. The last iteration showed the middle value to be the "best". In (b), the method is applied to the same data, with the only difference being the step, now equal to 4. In this case, the middle and right values give rise to SAME peakiness and the program stops. This happens because the program was set to require a substantial difference in peakiness to declare that one set is more peaky than another. This problem can be fixed by performing more elaborate reasoning to decide whether the program stops in such a case.

In Figure 3.28, the method is applied to two-helicopter data, the first helicopter being at CPA and the second before CPA and following the first. In all cases, the default is 8 periodograms. In (a), the focus is on a harmonic set of fundamental around 95 Hz, and the step is 4 periodograms. In this case, the number of periodograms goes down to 1 and then stops, providing the value 1 as the final result. In (b) the focus is on a harmonic set of fundamental around 92 Hz, and the step is 4 periodograms. In this case, the default of 8 periodograms is a local maximum of the peakiness of the selected harmonic set, and this is the final result. (c) and (d) are the same as (b) and (a), except that the step is 6 periodograms. In this case the final result is 1 periodogram.

The last example raises the issue of the robustness and consistency of the method. Different steps give different answers. The matter here is that this par-

```
(adjust-spec-est-params far-before-cpa-data 65 8 6)

Select a harmonic set to focus on by number

   Harmonic set 1 has fund-freq 94.82412
   Harmonic set 2 has fund-freq 80.72143

 Number to focus on?
1

Number of periodograms is   8 and direction is  ↓
Number of periodograms is  14 and direction is  ↓
Number of periodograms is  20 and direction is  ↓
Number of periodograms is  26 and direction is  ↓
Number of periodograms is  32 and direction is  ↓
Number of periodograms is  38 and direction is  ·

*****************

(adjust-spec-est-params far-before-cpa-data 65 8 4)

   Harmonic set 1 has fund-freq 94.82412
   Harmonic set 2 has fund-freq 80.72143

 Number to focus on?
1

Number of periodograms is   8 and direction is  ·
```

Figure 3.27: Trace of the program in data where helicopter is far before CPA. Default value is 8 periodograms and step is 6 in (a) and 4 in (b). In both cases, focus is on a harmonic set with fundamental around 95 Hz. Down-arrows indicate the direction of increase of peakiness in terms of the number of periodograms averaged.

```
(adjust-spec-est-params near-cpa-data 65 8 4)

Select a harmonic set to focus on by number

    Harmonic set 1 has fund-freq 94.844444
    Harmonic set 2 has fund-freq 91.89067

 Number to focus on?
1

Number of periodograms is   8 and direction is  ↓
Number of periodograms is   4 and direction is  ↓
Number of periodograms is   1 and direction is  •


****************

(adjust-spec-est-params near-cpa-data 65 8 4)

Select a harmonic set to focus on by number

    Harmonic set 1 has fund-freq 94.844444
    Harmonic set 2 has fund-freq 91.89067

 Number to focus on?
2

Number of periodograms is   8 and direction is  •


****************

(adjust-spec-est-params near-cpa-data 65 8 6)

Select a harmonic set to focus on by number

    Harmo·ic set 1 has fund-freq 94.844444
    Harmonic set 2 has fund-freq 91.89067

 Number to focus on?
2

Number of periodograms is   8 and direction is  ↓
Number of periodograms is   2 and direction is  ↓
Number of periodograms is   1 and direction is  •


****************

(adjust-spec-est-params near-cpa-data 65 8 6)

Select a harmonic set to focus on by number

    Harmonic set 1 has fund-freq 94.844444
    Harmonic set 2 has fund-freq 91.89067

 Number to focus on?
1

Number of periodograms is   8 and direction is  ↓
Number of periodograms is   2 and direction is  ↓
Number of periodograms is   1 and direction is  •
```

(a)

(b)

(c)

(d)

Figure 3.28: Trace of the program applied to two-helicopter data, the first helicopter being at CPA and the second before CPA and following the first.

ticular piece of data is problematic for a human also. There is heavy interaction and overlap between the two harmonic sets, resulting in the fact that peakiness of harmonic sets does not display a clear maximum, but seems to have erratic behavior even for a human.

As a final remark, we emphasize the voting principle used in the above program. In comparing harmonic sets based on their features, we used several criteria, with each criterion casting a vote, and a decision was made if there were enough votes supporting it. In comparing harmonic sets based on peak comparisons, each peak comparison casts a vote, and a small number of heuristic rules were used to interpret the result. Although voting was found to be adequate for our application, it should be noted that more elaborate schemes can be used instead, which employ confidence factors and means for combining them [Dove1986]. A careful consideration of the tradeoffs between simple and elaborate schemes must precede system design in other signal processing and interpretation applications.

3.2.3 Explicit representation of the temporal evolution of spectra by matching of spectra at different times

A third important application regards the matching of spectra at different times. This section is a high-level exposition of the ideas in a system for helicopter pitch and power tracking described more fully in the next chapter.

As we saw in chapter 2, the nature of the differences between spectra depends on the paths followed by the various acoustic sources that are present. In order to be able to apply the geometric arguments presented there, we must come up with power and frequency traces for each individual acoustic source present in the data. Obtaining appropriate spectral estimation parameters is only the first step. Harmonic sets in spectra at different times must further be associated with each other, before power and frequency traces can be extracted.

Two issues surface here.

The first is an issue of forming associations of harmonic sets over time. To do so, one has to make a hypothesis about possible acoustic source movements, which will then guide the search for associations. For example, if sources move always in straight lines, frequency always decreases due to Doppler shift. This is a strong constraint in associating harmonic sets over time: associations in which the frequency of the later harmonic set is higher than that of the earlier by more than a noise margin are not acceptable. The straight line hypothesis has other implications, too: a maximum of the power of the harmonic set occurs simultaneously with a steep decrease in frequency. The question here is in what form can such constraints be incorporated into a system for extracting power and frequency traces?

The second issue is that of representing associations between harmonic sets at different times. Harmonic sets are represented explicitly in the extended spectrum, therefore one way of representing the associations is by explicit links between harmonic sets of extended spectra at different times. Figure 3.29 shows extended spectra at three different times and the explicit two-way links between instances of a harmonic set. Such links allow tracing of the features of the harmonic set over time by following the chain forward or backward in time, thereby facilitating the computation of power and frequency traces.

Links between harmonic sets can also facilitate the application of constraints deriving from hypotheses about the nature of source movements: As the links are developed incrementally, one can examine the links built so far for conditions such as those mentioned previously. For example, if we develop links in a scenario, in which the acoustic sources move in straight lines, in a left to right manner (i.e. starting at the earlier times), by following the links backward in time we can find out if we are at a steep Doppler shift in frequency. We can then take this information into account in forming the next link.

122
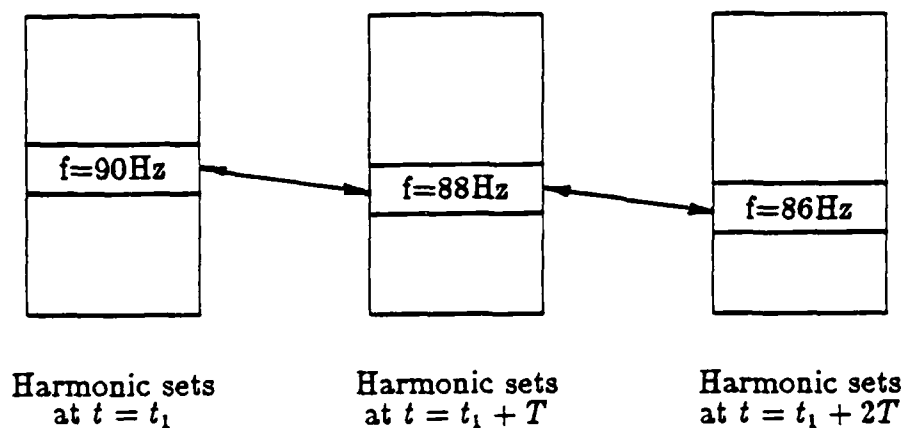
Figure 3.29: Two-way links between instances of a harmonic set in extended spectra at different times allow tracing of the features of the harmonic set over time.

A benefit from explicit representation of the association between harmonic sets at different times is that support for power and frequency traces built from such associations is readily accessible, in the form of the multiple levels of abstraction of the underlying extended spectra.

The above representation of the temporal evolution of harmonic sets is not constrained to correspond to extended spectra sampled uniformly in time. Although only uniform (but sparse) sampling was used in the system presented in the next chapter, both sparse and non-uniform sampling are important in our acoustic signal processing application because they are means for reducing the problem of high data rate combined with low information rate. Moreover, they are practiced by the human subject during protocol sessions. Protocols indicate that the human subject first samples spectra sparsely, then forms a qualitative hypothesis about the source path and finally samples spectra finer, in order to confirm the qualitative hypothesis. The system of the next chapter has the straight-line hypothesis about the source movement "hardwired" into the heuristic conditions for linking harmonic sets.

123

3.3 Summary

In this chapter we introduced the notion of extended spectrum, a multiple-level view of an acoustic spectrum, whose characteristics of interest are sets of harmonically related peaks. The computer representation of extended spectrum and its computation from the numeric spectrum were presented. For computation of the extended spectrum we embedded classification criteria into a sequential procedure and we argued that this approach enhances the robustness of the sequential procedure and provides an easily modifiable implementation, which can accommodate explicitly represented heuristic criteria.

Then we proceeded to demonstrate the usefulness of the extended spectrum by presenting systems based on it for solving three problems involving acoustic spectra.

The first system performs matching of spectra from different channels and identifies and characterizes faulty channels. As we saw in chapter 2, this system is useful in improving the performance of direction determination through array processing by excluding faulty channels.

The second system performs adjustment of the parameters of spectral estimation based on averaged periodograms, by maximizing the overall peakiness of the spectral estimate. The extended spectrum allowed us to express this visual criterion as multiple levels of matching of spectra obtained from the same time waveform but with different parameter settings. The need for such a system was seen in chapter 2 because no single parameter setting is appropriate for all scenarios encountered in helicopter signals.

The third system involves use of the extended spectrum to represent signals as a sequence of linked extended spectra obtained from non-overlapping sections of the time waveform. Links between extended spectra explicitly associate higher-level spectral entities, such as peaks and harmonic sets, over time. A system for helicopter pitch tracking based on these ideas will be presented in the next chapter.

124

Chapter 4

Helicopter Pitch and Power Tracking using Signal Matching

Pitch and power tracking of acoustic sources is the process by which we obtain traces of the fundamental frequency and power of helicopters from microphone signals, which in turn can be used for source localization as described in chapter 2. In the previous chapter we suggested that the notion of extended spectrum can be used for pitch and power tracking by forming explicit links between harmonic sets at different times. In this chapter we explore this idea further in the form of an implemented system for helicopter pitch and power tracking and we show experimental results of the operation of the system on several instances of real helicopter data.

4.1   Overview of the helicopter pitch tracking system

Figure 4.1 shows the overall architecture of the helicopter pitch tracking system. Excluding the feedback loops, the design philosophy is straightforward. First, spectral

estimation of short-time sections of the time waveform leads to a sequence of numeric spectra. Their peak and harmonic set structure is then explicitly represented by computing the corresponding extended spectra. Then links between harmonic sets at different times are built according to linking heuristics, which depend on the nature of anticipated source paths. Inspection of such links leads directly to pitch tracks.
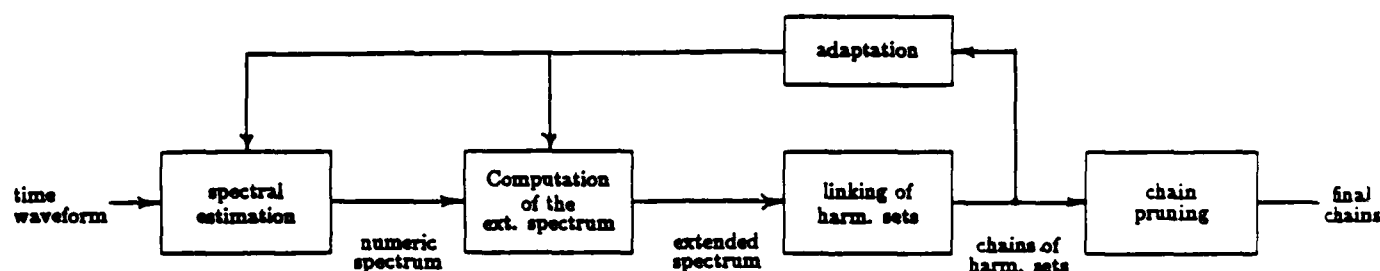
Figure 4.1: Architecture of the helicopter pitch and power tracking system.

The straightforward approach of linking harmonic sets at different times is augmented by heuristics in several places intended to improve robustness of the overall system.

In addition to the heuristics used to augment the straightforward approach, the performance of the system can be improved by incorporating feedback. Partially completed pitch tracks (as in Figure 4.2) arise when the associated harmonic sets are "weak" or non-existent in some spectra. Partially completed pitch tracks serve as starting points for a more focused search for harmonic sets that could close the

126

gaps. This is similar to the notion of "islands of certainty" in speech recognition, whereby recognition starts with the words that can be unambiguously recognized, which are then viewed as islands that grow to eventually cover the whole sentence.

FREQUENCY (Hz)

25

3

0.0   TIME (s)   46.5
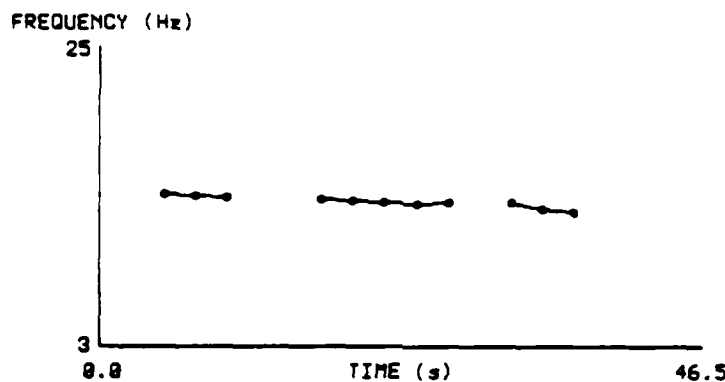
Figure 4.2: Partially completed pitch tracks. Harmonic sets are represented by circles. Lines connecting harmonic sets indicate that the sets are explicitly associated with each other and are part of a sequence of linked sets, called a chain.

Feedback can take two forms, as shown in Figure 4.1. In the simpler form, no new spectra are computed. Instead, harmonic set grouping is repeated in the existing extended spectra with a hypothesized fundamental frequency equal to that dictated by the partially completed pitch tracks and with the thresholds for accepting peaks set lower. In the more complicated form, the rate of change of the partially completed pitch tracks is taken into account and dictates an initial default value for the spectral estimation parameter (number of blocks), initiating an adaptation process for that parameter, as we described in the previous chapter. The same feedback philosophy is used to extend pitch tracks into times with very low signal-to-noise ratio.

4.1.1 Linking harmonic sets into chains

In forming chains of harmonic sets, i.e. sequences formed by harmonic sets linked together, as shown in Figure 4.2, we focus on pairs of extended spectra that are consecutive in time. For each pair of spectra, we exhaustively check all pairs of harmonic sets of the two spectra as to whether it is appropriate to link them. Currently, there is one linking rule, which links two spectra if the difference between the fundamental frequency of the first and the second sets lies within fixed bounds (Figure 4.3). The upper bound is the maximum allowable drop in frequency based on the acceptable Doppler shift assuming a straight line track. The lower bound is a negative number and corresponds to the maximum allowed increase in frequency. Under the straight line path assumption, such an increase can only be due to noise. Currently these two bounds are fixed, but in general they may be a function of the time distance between the two sets.
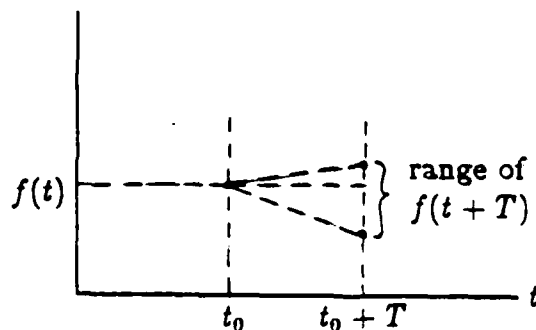
Figure 4.3: Upper and lower bound of the difference between frequencies of harmonic sets in successive times for linking.

Two possible kinds of extensions to the linking rule base are possible. One consists of more sophisticated criteria for linking two harmonic sets, based not only on their fundamental frequency, but other characteristics of the harmonic sets,

128

including their time difference. The second kind of extension consists of extending the locality of our criteria: currently whether two harmonic sets are linked depends only on their own characteristics. Extending the locality involves criteria that look at the links formed so far and take this information into account in deciding whether to link the harmonic sets under consideration.

To accommodate linking, the harmonic set data type is extended to include two more pieces of information, which are pointers to the previous and next harmonic sets in the chain. Furthermore, a pointer back to the parent extended spectrum is useful for access to information such as the time of the extended spectrum and its parameters (Figure 4.4).
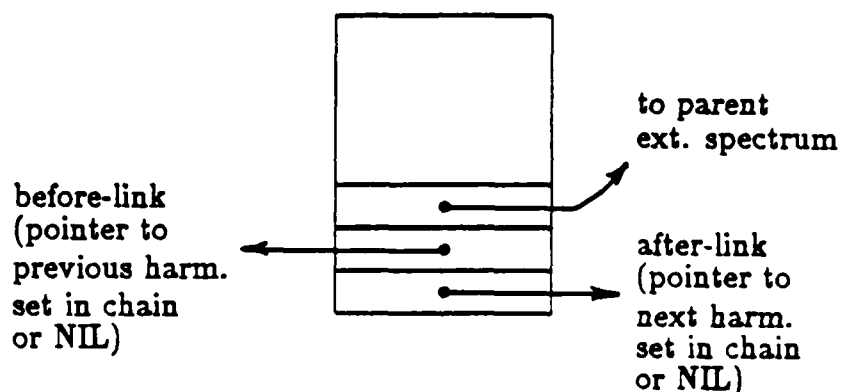
Figure 4.4: Additional slots in the harmonic set data type.

4.1.2 Harmonic chains as distinct objects and their linking

To be able to explicitly manipulate harmonic chains, we introduce the chain object (Figure 4.5). The chain object consists of pointers to the first and last harmonic set of the chain and pointers to the previous and next chain (chains can be linked together just as harmonic sets can). The chain object is basically a doubly linked list,

129

because harmonic sets are linked in two directions, forward and backward in time. Calling a chain an object emphasizes the view of the chain as a single independent conceptual entity, with its own features, which include the initial and final time, the initial and final fundamental frequency and the length of the chain.
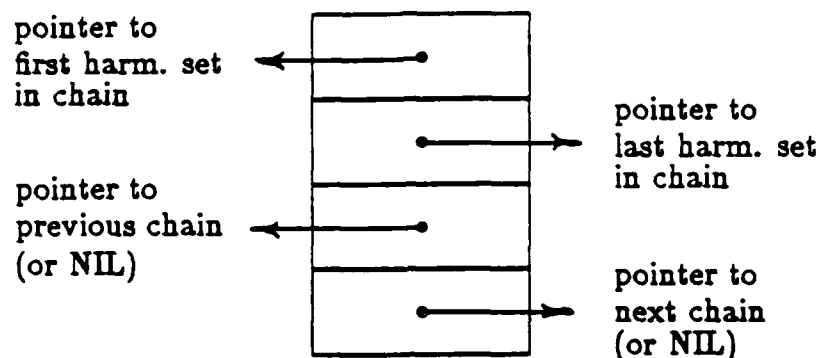
Figure 4.5: Chain object

Linking of chains is similar to linking of harmonic sets in terms of the procedures and issues involved. As in the case of harmonic sets, currently a single criterion is used based on the difference of the fundamental frequencies of the end set of the first chain and the start set of the second chain. The criterion specifies that two chains are to be linked if the start frequency of the second chain is within a range below or within a smaller range above the end frequency of the first chain. This criterion may be generalized into a rule base, including several criteria that link chains based on additional characteristics, such as rate of change of frequency or power along the chains to be linked.

4.1.3 Feedback mechanism for completion of pitch tracks

An interesting aspect of harmonic chains is their utility in implementing the feedback mechanisms mentioned in the beginning of this chapter. Feedback is initiated when pitch tracks are partially completed. Partially completed pitch tracks consist of multiple chains that could potentially link into a single one. Such chains are, as an initial step, linked together as described above. Then each chain link initiates a focused search for harmonic sets in the extended spectra over the duration of the link.

There are two reasons that harmonic sets may be missing from particular extended spectra, causing a gap in an initial pitch track. One is that the harmonic sets do exist, but are weak and get pruned by the harmonic set pruning criteria of the extended spectrum computation. A second reason is that the harmonic sets exist, but there is a "bug" in the extended spectrum computation, i.e. a case not covered by the heuristics used in this computation. A third reason is that the harmonic sets do not exist in the spectra along the gap for propagation reasons, such as wind or other atmospheric conditions.

To differentiate between the first two and the last case, a focused search for a harmonic set is conducted in the extended spectra along the chain link (Figure 4.6) with a hypothesis of a fundamental frequency as dictated by the frequencies of the linked chains (currently linear interpolation is used). All peaks that are approximate harmonics of the hypothesized fundamental frequency with a slack larger than the original slack used in the computation of the extended spectrum are collected. The harmonic set thus formed is tested against heuristic criteria for acceptance, similar to but weaker than the criteria used in pruning the harmonic sets in the original extended spectrum computation. If the harmonic set passes at least one of those criteria successfully, it becomes part of the corresponding extended spectrum and it is added to the chain.

131

FREQUENCY (Hz)

25

3
0.0   TIME (s)   46.5
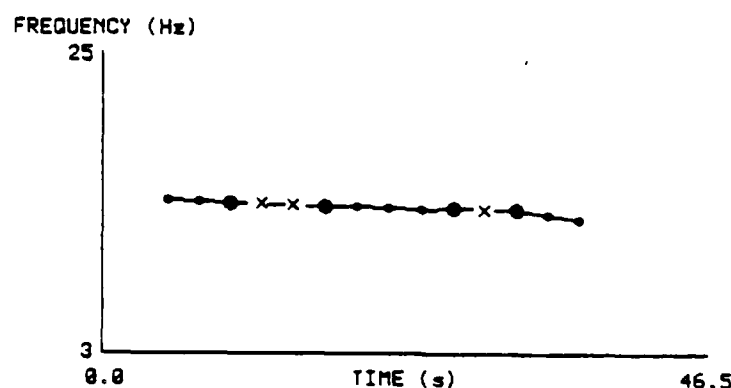
Figure 4.6: Completion of gaps in pitch tracks by focused search for harmonic sets in times along the gap. Newly found harmonic sets are shown as 'x's. Chain ends that are linked are denoted by larger circles than the normal harmonic sets. The chain links themselves are denoted by dotted lines.

The criteria for acceptance of the harmonic set formed after focused search based on a fundamental frequency hypothesis from the chain are shown in Figure 4.7. These tests are more lenient compared with the tests of the original computation of the extended spectrum and they do not attempt any merging or comparison of the harmonic set at hand with existing sets of the extended spectrum.

4.1.4 Feedback mechanism for extension of pitch tracks

A mechanism similar to that presented in the previous section is used for extending pitch tracks, or equivalently harmonic chains, beyond their ends. This is done by looking out of ends of existing chains that are not linked to another chain (Figure 4.8) and searching for harmonic sets with a fundamental frequency approximately equal to that of the harmonic set at the end of the chain. If thus formed harmonic sets pass one of the criteria of Figure 4.7, they are accepted as extension of the harmonic chain and inserted into the extended spectrum at the associated time.

A harmonic set is accepted if at least one of the following conditions is satisfied:

- It has at least two harmonics of order less than or equal to 3.

- At least its first or second harmonic have the characterization LARGE-AVG-PROMINENCE.

- It has at least three harmonics in total.

Figure 4.7: Criteria for acceptance of harmonic set formed after focused search based on a fundamental hypothesized by a chain.
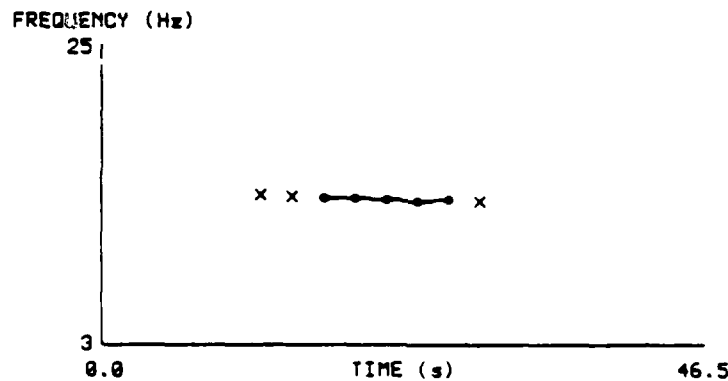
Figure 4.8: Extension of pitch tracks by focused search for harmonic sets beyond the ends of a chain. Newly found harmonic sets are shown by 'x's.

4.1.5 Interface for inspection of information supporting pitch tracks

Pitch and power tracks represented by links between harmonic sets are only a small part of a hierarchy of signal information explicitly represented in the underlying extended spectra. Ease of inspection of this information is desirable for several

reasons. If our system is used as an assistant, the human user needs to have easy access to this information to form a complete picture of the behavior of the signal tracking system. In developing the heuristics needed for robust operation of the system, there is a fair amount of experimentation involved, in which the deficiencies of the current heuristics of the system in specific scenarios are identified and rectified by the modification of existing or the insertion of new heuristics.

The main issue in inspecting the signal information contained in the extended spectra underlying pitch tracks is that there are more dimensions to this information than can be represented in a two-dimensional display, such as the pitch tracks shown for example in Figure 4.8. Therefore, inspection of extended spectra is more natural if performed interactively, with a concise output of the form shown in Figure 4.8 being only the initial output of the system and a starting point for a more detailed inspection. The human user then manually requests more detailed information on specific harmonic sets, displayed as circles in the concise pitch track output.

A variety of devices exist that permit input of the coordinates of a point of a two-dimensional display, such as a pitch track of the above form, into a program. Such devices include various types of joysticks, trackballs, touch-sensitive screens, and mouse devices [Norman & Draper 1986, Ch. 7].

To facilitate inspection of extended spectra underlying pitch tracks, an interactive facility was implemented based on a mouse device. Pointing to the harmonic sets shown as dots in the graphical representation of pitch tracks and actuating the mouse displays various pieces of information associated with the harmonic sets, depending on the type of actuation. The Symbolics Lisp machine has a mouse with three buttons, left (L), middle (M) and right (R), each of which can be pressed and quickly released ("clicked") once (1) or twice (2). Figure 4.9 shows the functionality we attached to the various mouse clicks.

A graphical representation like the one shown in Figure 4.8 essentially displays

- L-1: Display the harmonic set pointed at.

- M-1: Display extraneous peaks of extended spectrum containing the harmonic set pointed at.

- R-1: Display relevant peaks of the extended spectrum containing the harmonic set pointed at. For each peak show frequency, power, prominence, generalized prominence, horizontal distance and characterization.

- L-2: Display the log spectrum corresponding to the extended spectrum that contains the harmonic set pointed at. After displaying the log spectrum, the user has the ability to point to various points on the spectrum and read off their frequency and power by clicking the mouse.

- M-2: Display the linear spectrum over the low frequency range (3-80 Hz) corresponding to the extended spectrum that contains the harmonic set pointed at. Again the user can read points off the graph.

- R-2: Display all harmonic sets before pruning that are within a short distance (in frequency) from the harmonic set pointed at.

Figure 4.9: Mouse capabilities for displaying various kinds of information off the harmonic chain plots.

one characteristic (fundamental frequency or power) of a complex structure (the sequence of extended spectra linked through their harmonic sets) over time. Therefore it can be viewed either as a signal in the ordinary sense, i.e. a function of time, or as an indexing mechanism for accessing the underlying complex structure. An interface of the type presented here that uses a mouse-like device is a convenient way of using the graphical representation as an index to the underlying information.

4.2 Experimental results

In this section we present experimental results of the implemented frequency and power tracking program applied to real helicopter data. We are going to present a complete example including the results of all iterations of the feedback loop of the system shown in Figure 4.1 (without including adjustment of spectral estimation parameters) until it converges to a solution. The initial chains formed are conservative, in the sense that only prominent harmonic sets and strong set associations are taken into account. At each iteration, it is attempted to extend the chains formed so far and connect them with each other through focused search for less prominent harmonic sets that would serve that purpose. Harmonic sets identified after focused search are incorporated into the extended spectra and linked with other sets as appropriate. The system converges to a solution when focused search for less prominent harmonic sets does not yield any more sets. At that point, short chains are pruned out, and the final fundamental frequency and power traces are formed for each of the remaining chains. The power of a harmonic set is defined as the sum of the values of the power spectrum at the locations of the first 8 harmonics.

On each graph, the horizontal axis is time and the vertical axis is frequency or power. Each small filled circle corresponds to a harmonic set, whose x-coordinate is time and whose y-coordinate is its fundamental frequency or power. Solid straight line segments between small filled circles correspond to explicit links between har-

monic sets. Therefore chains are represented by sequences of small filled circles, which are connected with solid lines. Thick filled circles connected with dashed lines are links between chains. 'x' marks are harmonic sets identified after focused search, either to close gaps between chains or to extend pitch tracks.

The data shown in Figure 4.10 was obtained from a two-helicopter scenario, with both helicopters following the same straight line path and the second helicopter following the first by about 40 seconds. The closest point of approach of the first helicopter is around the 65th second of the data, whereas the CPA of the second helicopter is beyond the data window available. The sampling rate for extended spectra is one spectrum every 2.5 seconds, while the sampling rate of the data is 2048 Hz. Extended spectra correspond to numeric spectra computed by averaging 8 periodograms, each from a data window 0.5 seconds long and with 50 % overlap. Therefore the total data used for each spectrum is 2 seconds long.

In Figure 4.10, part (a) shows the envelope of the time waveform, which is a measure of the pressure of the acoustic wave at the microphone versus time. Part (b) shows the pitch tracks formed without any pruning of harmonic sets in individual extended spectra. In part (b) there are extraneous pitch tracks due to a variety of causes, such as octave errors and weak harmonic sets. Part (c) shows the initial pitch tracks formed by the pruned harmonic sets, but without any feedback yet. Part (d) shows the links between harmonic chains (with the corresponding ends shown as thick circles) and the harmonic sets identified through attempts to extend or close the gaps between chains. Part (e) shows the chains after the newly found harmonic sets have been incorporated into the extended spectra. Part (f) shows the result of another attempt to extend and link chains, starting with the chains of Part (e) this time. After that, no more harmonic sets can be found, leading to convergence of the program. Part (g) shows the final fundamental frequency traces and part (h) the corresponding power traces for the chains of part (g).

The results of parts (g) and (h) correlate very well with the helicopter scenario.

137
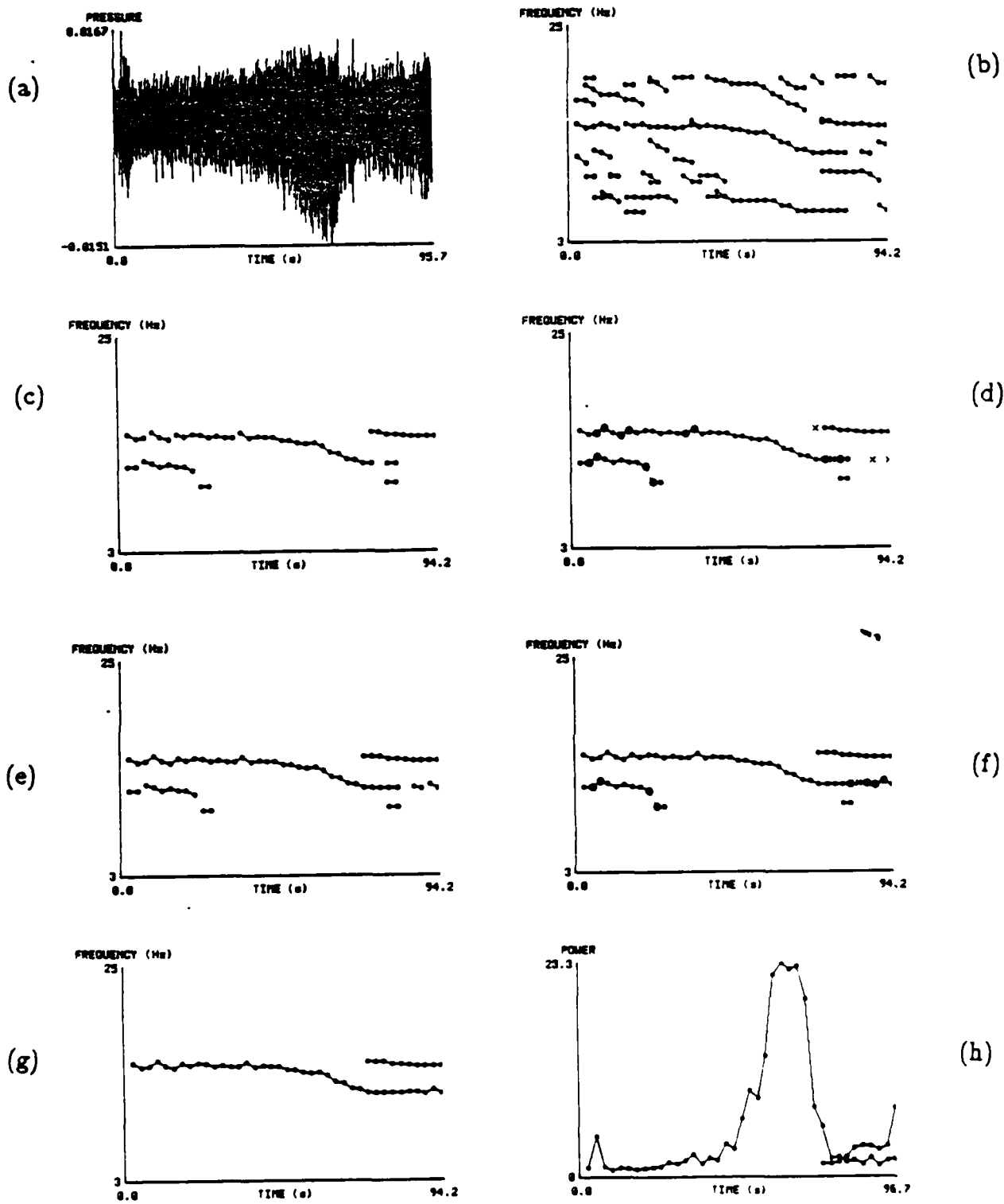
Figure 4.10: Fundamental frequency and power traces in the 3-25 Hz region from 95.7 seconds of real data from a two-helicopter scenario.

138

The longer chain in part (g) corresponds to the first helicopter, and the fundamental frequency shows a significant decrease around CPA, due to the Doppler shift, which is more intense around CPA, as we saw in Chapter 2. Moreover, the power trace in part (h) shows a clear maximum around CPA. The shorter chain in part (g) corresponds to the second helicopter, which has not yet reached CPA. The corresponding power trace shows a steady increase, but stops short of a maximum, because the CPA of the second helicopter is outside our data window.

Fundamental frequency and power traces computed from other scenarios of real helicopter data. We now present fundamental frequency and power traces from a variety of helicopter scenarios. We provide a brief description of the scenario and we show: (a) the envelope of the time waveform, (b) the chains obtained without any pruning of harmonic sets, (c) the chains obtained as the final output of the previous system shown in a frequency/time display and (d) the power traces of the final chains, or equivalently, the chains of part (c) shown in a power/time display.

The data shown in Figure 4.11 is from a single helicopter scenario. The helicopter is audible starting at 30s, and is at CPA at 235s. Weather was cloudy and calm, but there was heavy traffic on the neighboring roads, as well as heavy air traffic, including a single-engine plane takeoff and a jet flyby. The resulting chain, whose frequency is shown in part (c) and power in part (d) of Figure 4.11, correlates well with the scenario. Fundamental frequency shows the decrease at CPA due to Doppler shift, accompanied by a distinct maximum in power. Spectra in this example were sampled every 5 seconds and the whole scenario lasts 312 seconds.

Figure 4.12 shows a different two-helicopter experiment, with two helicopters following the same straight path at a difference of 11s, as concluded from the recorded CPA times. From the experiment records, the first helicopter is audible starting at t=30s, has a CPA at t=54s, and is no longer audible at t=60s. The second
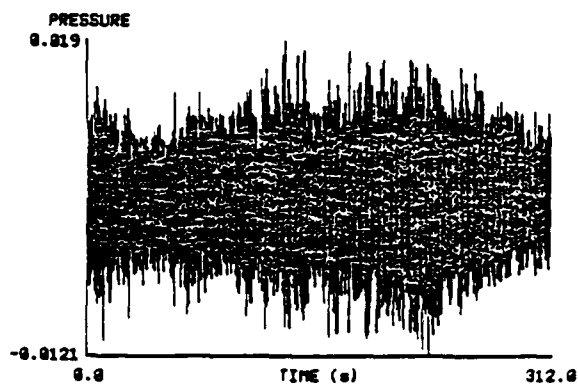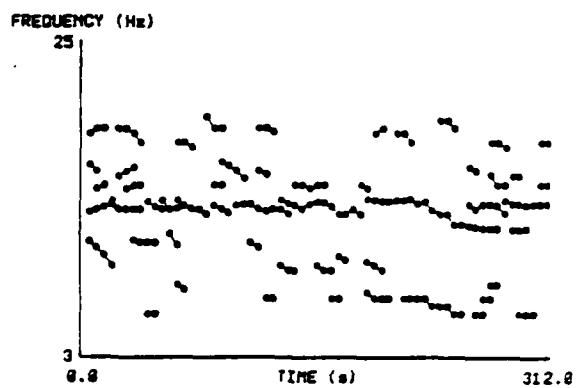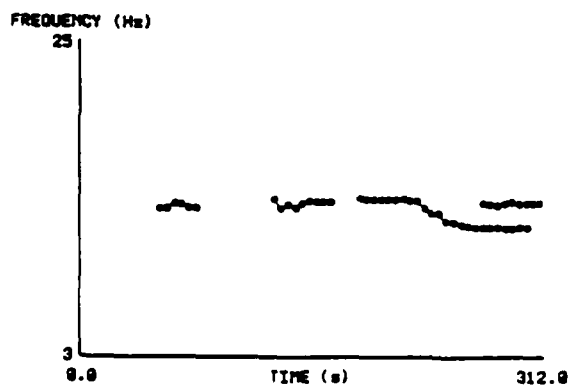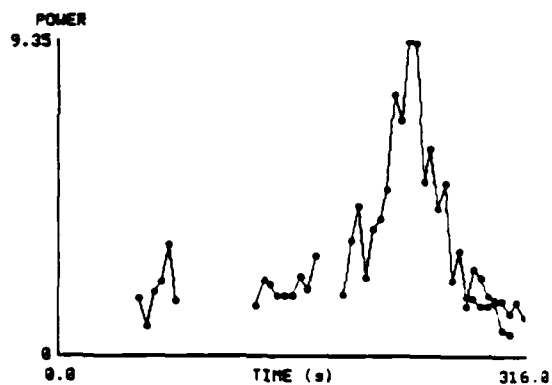
Figure 4.11: A single helicopter scenario. (a) the time waveform (b) chains without harmonic set pruning (c) final chains in frequency versus time (d) final chains in power versus time.

helicopter is audible starting at t=50s, has a CPA at t=65s and is no longer audible at t=70s. Loud truck traffic was recorded throughout the experiment, while wind speed was low (5mph) with cloudy weather. In spite of the weak signal and the loud noise sources, the resulting frequency and power traces show very good correlation with the recorded scenario.
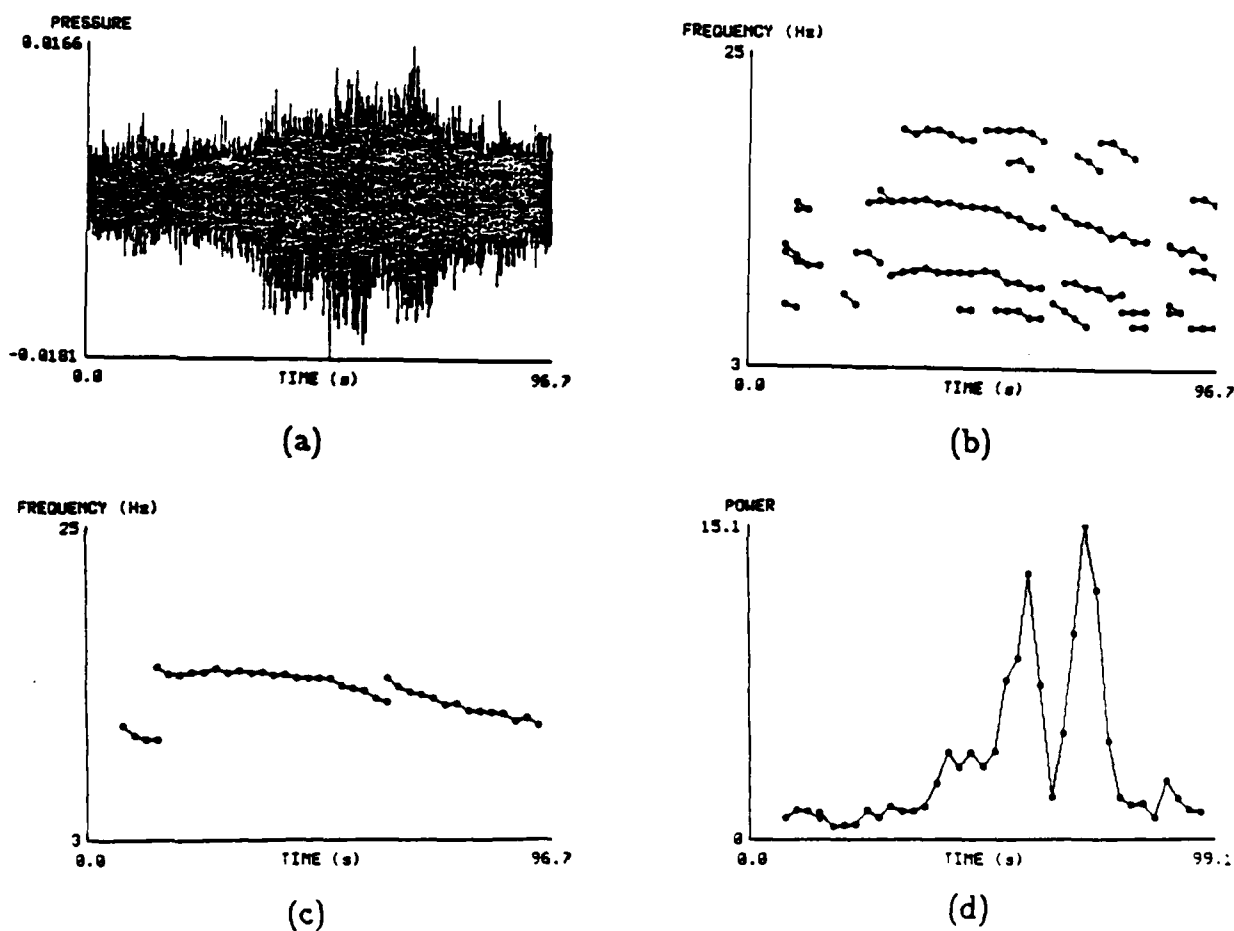
Figure 4.12: A two-helicopter scenario, with weak signals and the two helicopters 11s apart. (a) the time waveform (b) chains without harmonic set pruning (c) final chains in frequency versus time (d) final chains in power versus time.

Figure 4.13 shows data from a single-helicopter experiment but with a different type of helicopter. The data is 77s long and stops just after the CPA of the helicopter. Wind was at 6mph and no significant noise sources were recorded. Power and frequency traces are both in very good agreement with the scenario.
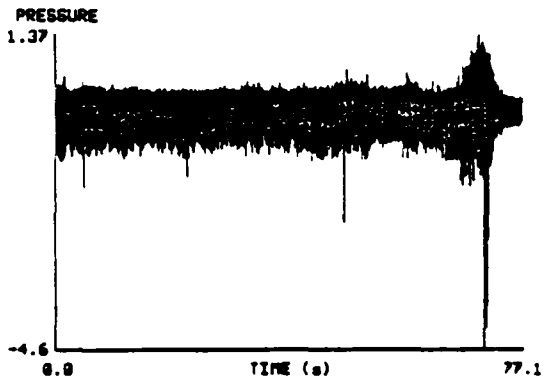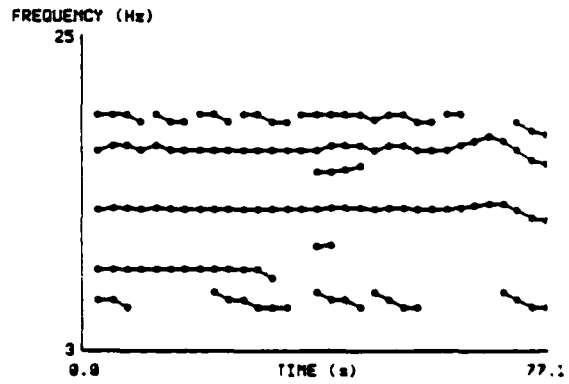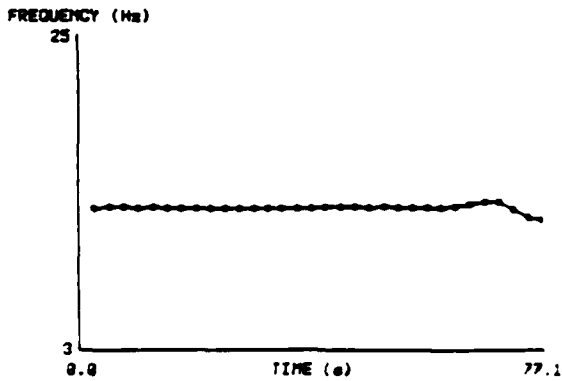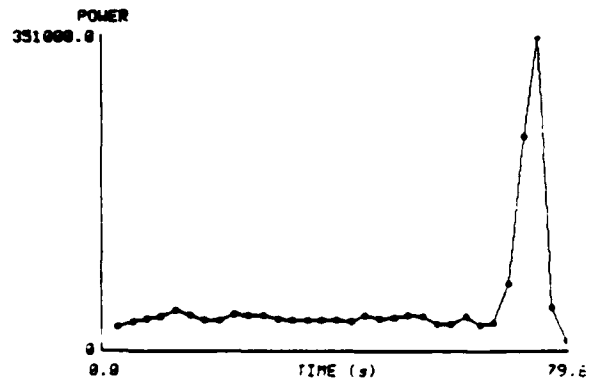
Figure 4.13: A single helicopter scenario with a different type of helicopter. (a) the time waveform (b) chains without harmonic set pruning (c) final chains in frequency versus time (d) final chains in power versus time.

Figure 4.14 shows data from another single-helicopter experiment with the same type of helicopter as Figure 4.13. Similarly, very good agreement of the results is observed with the recorded scenario.
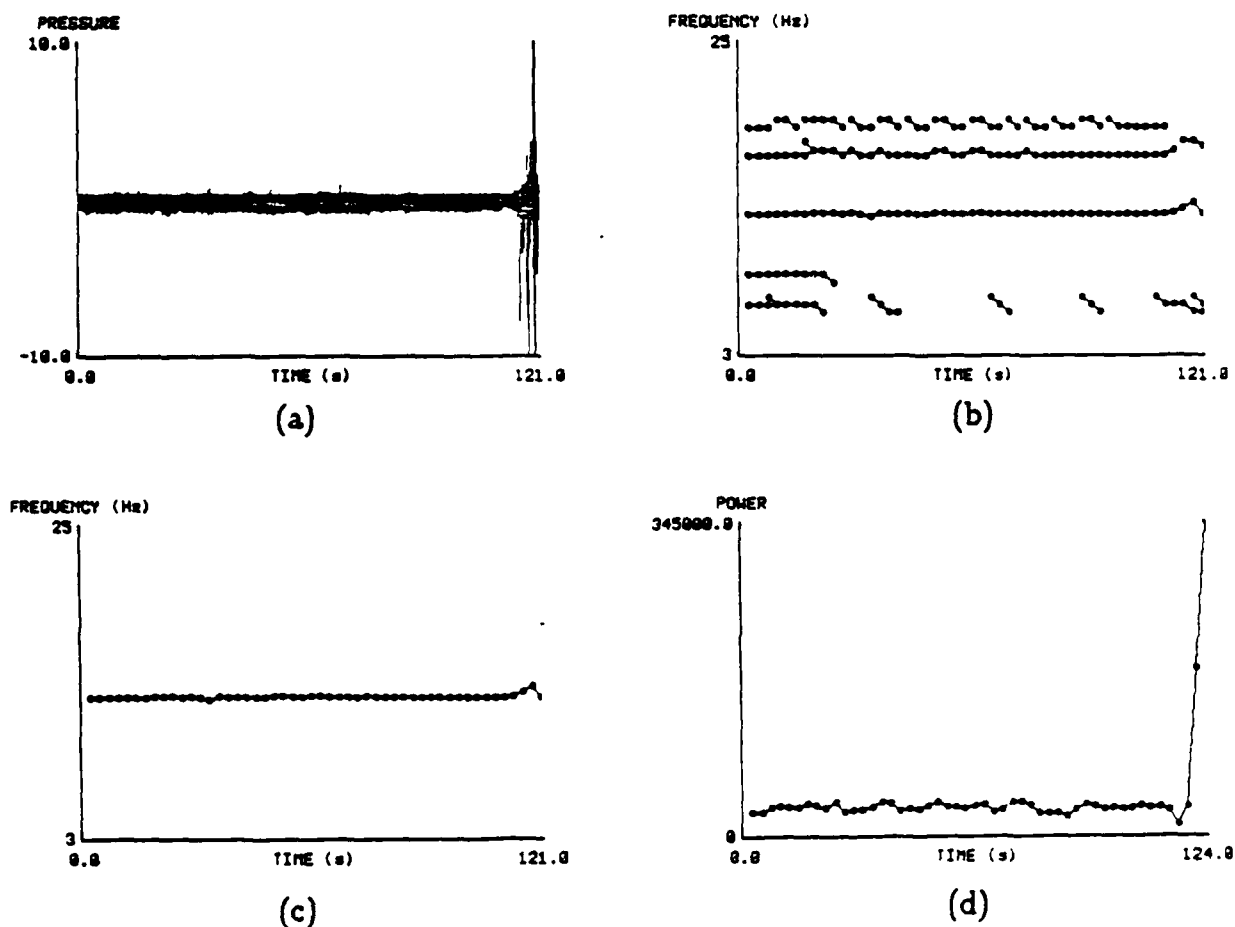
Figure 4.14: Another single helicopter scenario, with the same type of helicopter as in the previous figure and with the data window ending at CPA. (a) the time waveform (b) chains without harmonic set pruning (c) final chains in frequency versus time (d) final chains in power versus time.

Figure 4.15 comes from the same two-helicopter experiment as in Figure 4.10 but from a sensor several hundred meters away from the sensor of that figure along the path of the helicopter. There is a difference in real time between the two pieces of data, but they are similar in that both capture the data starting before the CPA of the first helicopter and ending just before the CPA of the second helicopter, where of course the CPAs are with respect to two different sensors. Again, except

for the double extension of the chain corresponding to the first helicopter before the 30th second of data, the result is in very good agreement with the scenario of the experiment.
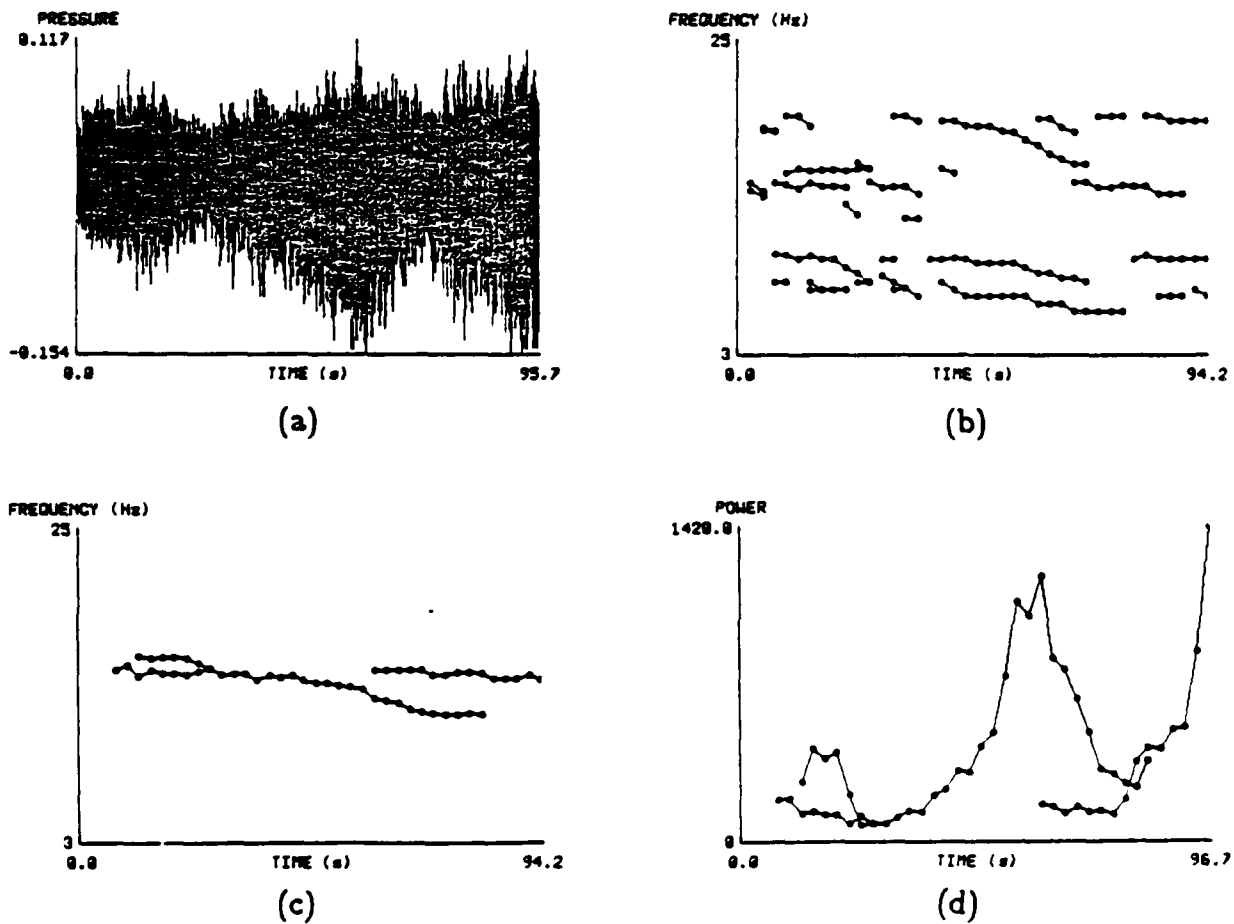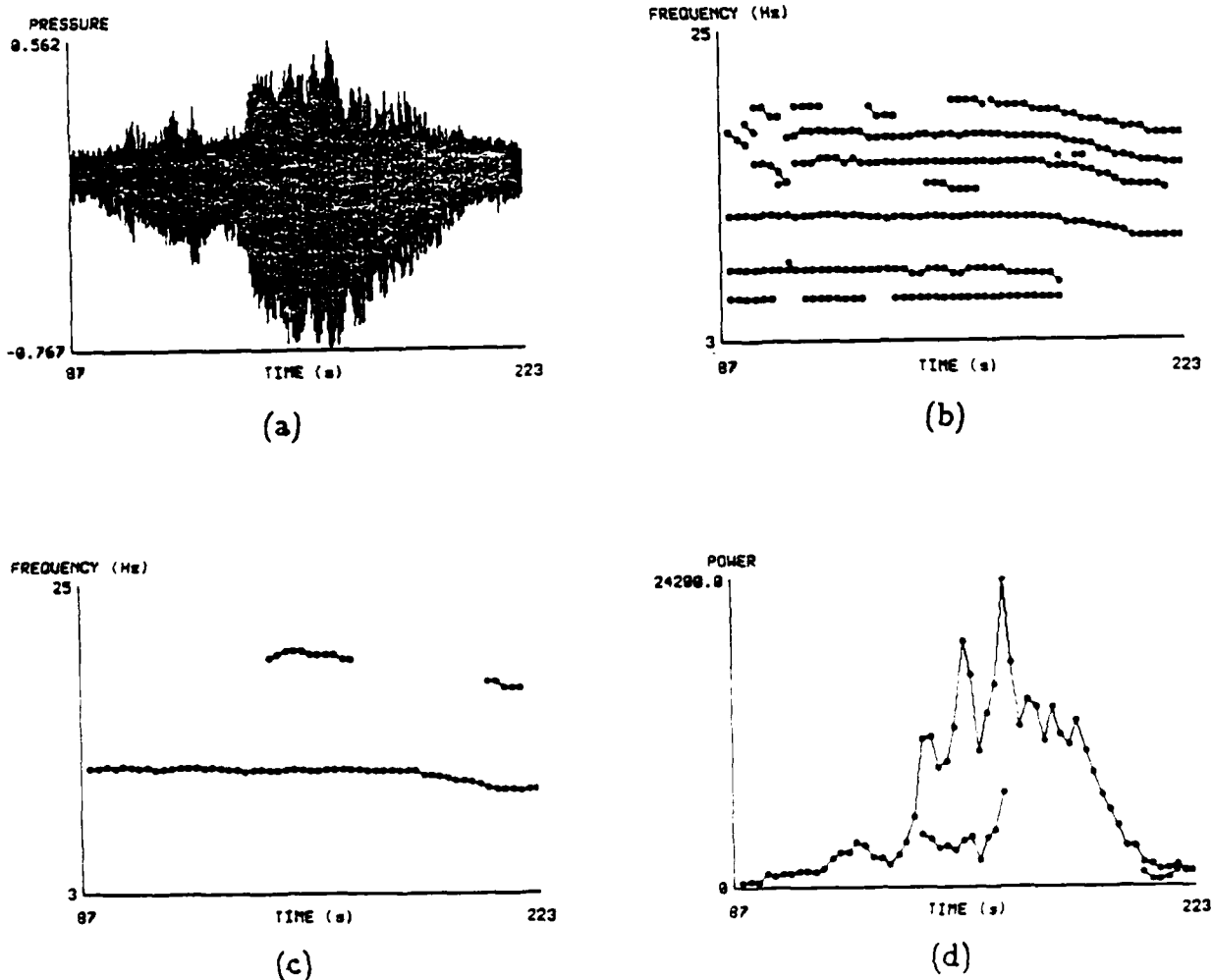

Figure 4.15: A two-helicopter scenario. (a) the time waveform (b) chains without harmonic set pruning (c) final chains in frequency versus time (d) final chains in power versus time.

Figures 4.16, 4.17 and 4.18 show portions of a single-helicopter scenario. Figure 4.16 shows the portion between 0s and 150s, in which the helicopter is audible at 45s for the first time, while it reaches CPA at 210s, which is outside this portion of the data. Again there is very good correlation of the output chain shown in parts (c) and (d) of Figure 4.16, in frequency versus time and power versus time display respectively. Fundamental frequency is fairly stable and power increases with time.
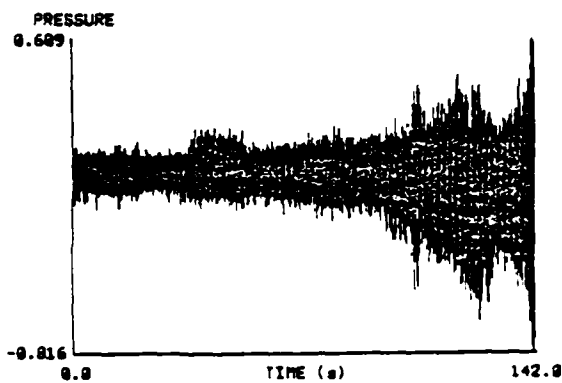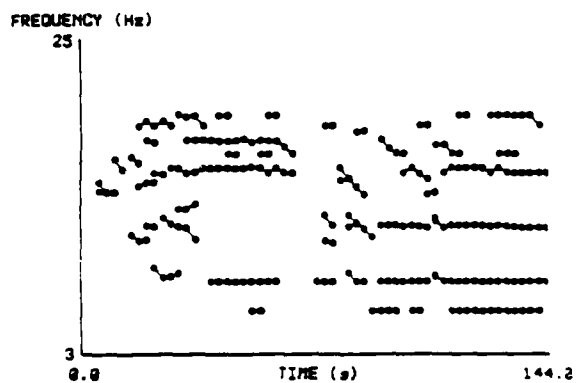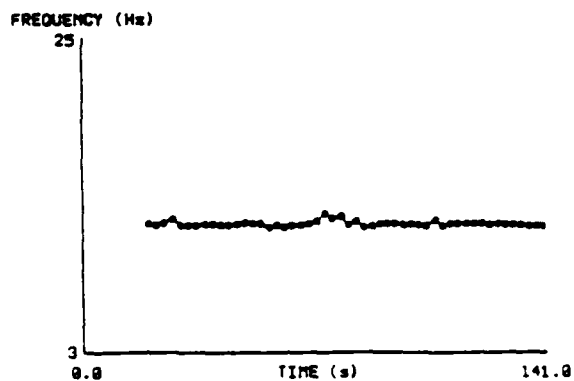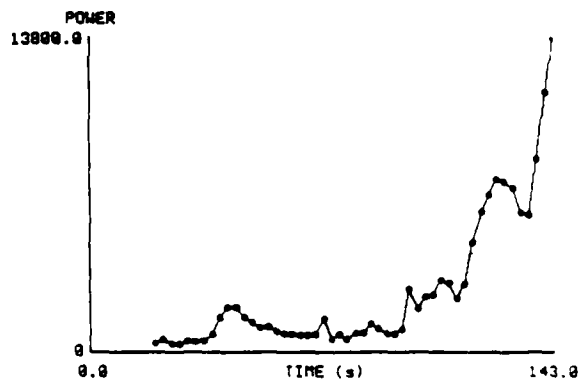
Figure 4.16: Another single helicopter scenario, portion before CPA. (a) the time waveform (b) chains without harmonic set pruning (c) final chains in frequency versus time (d) final chains in power versus time.

145

Figure 4.17 shows the portion of the data which includes the CPA of the helicopter. This portion overlaps with the last 55s of the data shown in Figure 4.16, starting at t=87s in the same time scale. Again the harmonic chain found shows good correlation with the scenario, having a strong Doppler shift in fundamental frequency around t=207s, at the same time when the CPA was recorded by the human observer. However, the power trace does not show a clear maximum, and the high-power portion is clearly well before the CPA time recorded by the human observer.

Figure 4.17: A single helicopter scenario, portion including CPA. (a) the time waveform (b) chains without harmonic set pruning (c) final chains in frequency versus time (d) final chains in power versus time.

To explore the above discrepancy between the computed power trace and the

146

recorded CPA time, we processed data from a different sensor very close to the one from which the data in Figure 4.17 originated, starting at t = 180s. Figure 4.18 shows the results of processing the latter portion of data. Power and fundamental frequency traces are in agreement with those of Figure 4.17, confirming that in this example the microphone signals show a power maximum before the recorded CPA time (t = 207s), while the fundamental frequency traces appear to agree with the recorded CPA time. This incompatibility cannot be readily explained from the available experiment records. Possible causes could be multipath propagation phenomena (reflections of acoustic waves off buildings or other structures), effects of strong winds or some change at the sound source itself (not steady speed or straight path). The experiment strongly suggests the complexity present in real helicopter data and the need to check the consistency of our data by comparing signals received at multiple geographically separated sensors.

Appendix B shows the complete results from the above experiments including the intermediate chains obtained during the program's iterations.

4.3 Discussion

In this chapter we presented a system for helicopter pitch and power tracking which is based on the concept of signal abstraction. Pitch and power tracking can be viewed as a natural consequence of the explicit representation of higher level signal entities consisting of peaks and harmonic sets, and of links between associated signal entities at different times in the form of chains of harmonic sets. Such explicit representation and retention of signal abstractions provides easily accessible support of the final pitch and power tracks, thus simplifying the task of explaining the behavior of the program and of correcting its behavior if it does not agree with the conceptions of the designer/user. An interface was implemented to facilitate access of the information underlying the pitch and power tracks.
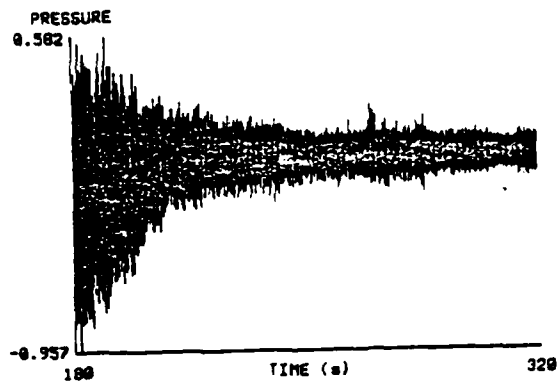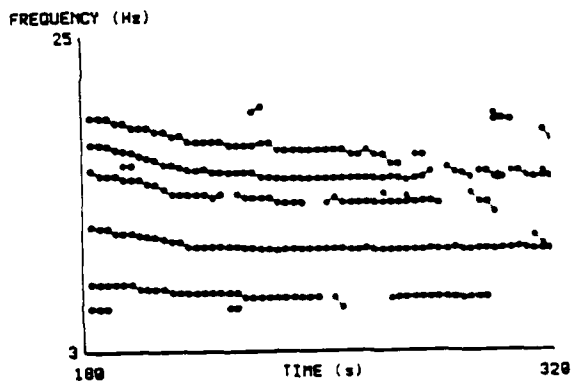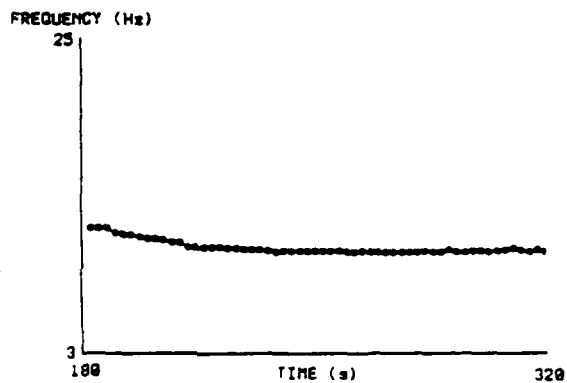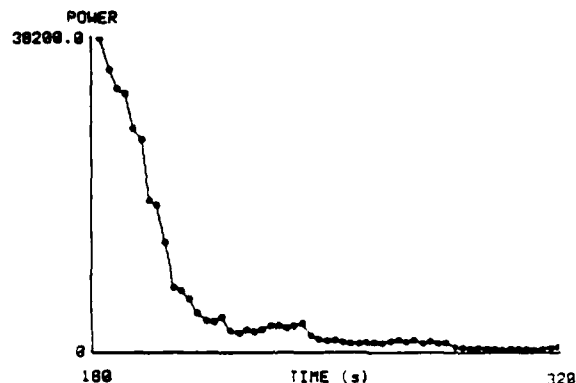
Figure 4.18: A single helicopter scenario, portion starting slightly before CPA. (a) the time waveform (b) chains without harmonic set pruning (c) final chains in frequency versus time (d) final chains in power versus time.

An important characteristic of the implemented program is locality of decision making. For example, the test whether to link two harmonic sets depends only on the two harmonic sets involved. When it is necessary to make non-local decisions at some level, we try to reduce them to local decisions at a higher level. For example, the decision whether to link two harmonic chains is handled at the harmonic chain level, where it is a local decision involving two chains, and not at the harmonic set level, where the decision is non-local.

Spectral abstractions explicitly represent higher level signal entities in the signals, therefore permitting incorporation of heuristics, that are more conveniently expressed in terms of the higher level signal entities than in terms of the numeric signal. Classification performed at lower abstraction levels helps to construct the higher levels of abstraction of a given numeric signal. Embedded classification is thus seen as an important aspect of the construction and use of signal abstractions.

Finally, both the system of this chapter and the parameter adjustment system for spectral estimation described in the previous chapter indicate that signal abstractions are not incompatible with traditional notions of signal processing, such as feedback. In fact, feedback is shown to enhance the performance of systems built on signal abstractions.

Chapter 5

Abstraction and Signal Mappings

In the previous chapters of this thesis we introduced the extended spectrum as a multilevel abstraction appropriate for harmonic spectra and we demonstrated its utility in a variety of signal processing tasks centered around matching of harmonic spectra. In this and the next chapter[1] we will present a multilevel abstraction for wavenumber spectra, which were defined in Chapter 2, and its use in a form of diagnostic reasoning about a complex signal processing system for direction finding, which is based on the wavenumber spectrum.

The task of the diagnosis system (Figure 5.1) is to identify the causes of mismatch between the known input scenario to the direction determination system, namely the position, speed and temporal signal characteristics of the acoustic sources present, and the output of the direction determination system, which consists of a wavenumber spectrum containing radial ridges at angles corresponding to the azimuth and with ridge shapes corresponding to the temporal spectra of the acoustic sources (defined in Chapter 2).

A key idea underlying the diagnosis system is the conceptualization of the direc-

---

[1]Chapters 5 and 6 are based on collaborative work with Hamid Nawab and Victor Lesser conducted at the MIT Lincoln Laboratory during the academic year 1984-1985.
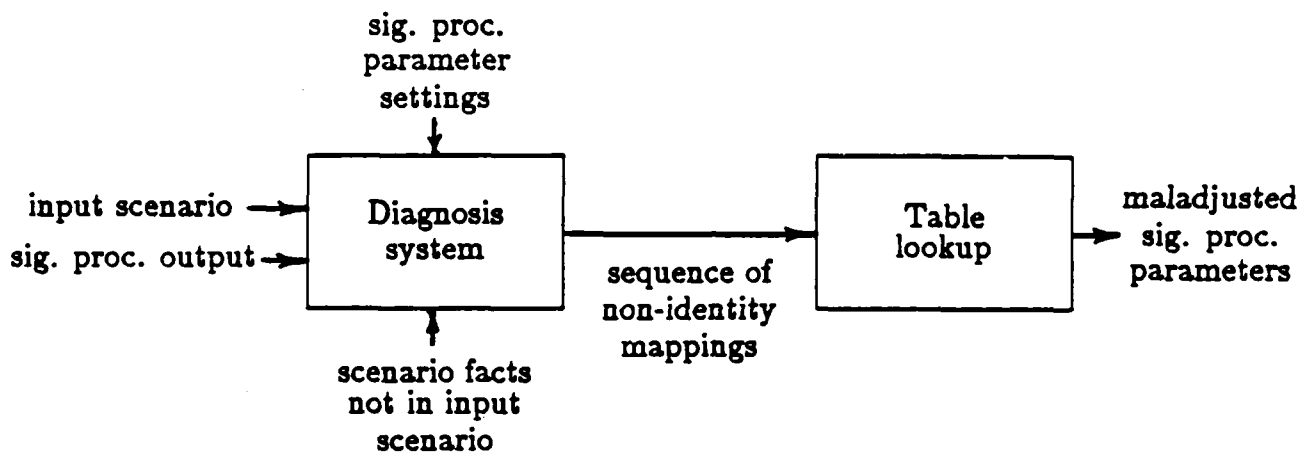
Figure 5.1: Block diagram of the diagnosis system for the direction determination system

tion determination system as a collection of mappings between wavenumberspectra, which act as identity mappings under ideal operation of the direction determination system. By taking this view and expressing the input scenario as a wavenumber spectrum (plus additional facts that cannot be represented in the wavenumberspectrum, such as speed of the acoustic sources and their distance from the microphone array), we can model both propagation phenomena and possible distortions introduced by the signal processing entirely in the Fourier (i.e. wavenumber spectrum) domain (Figure 5.2).
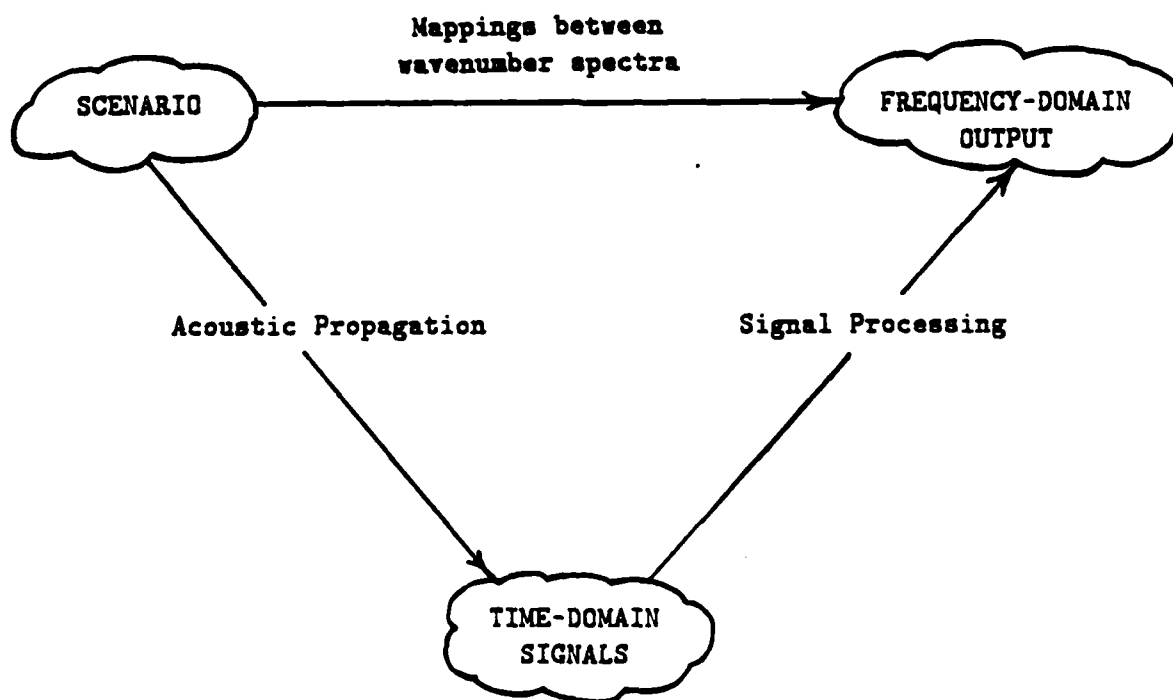
Figure 5.2: The effects of acoustic propagation and signal processing on the time-domain microphone signals can alternatively be conceptualized as mappings between wavenumber spectra

Diagnosis is viewed as a search through the space of all possible sequences of

mappings. Having several levels of abstraction for wavenumber spectra makes the search hierarchical, in the sense that it is performed starting at the higher abstraction levels and dropping the level of abstraction in case of failure. Search is followed by verification of its outcome. The principle used here is that verification takes place at the lowest possible abstraction level, because that is when all the detailed information about the diagnosis problem is being taken into account, whereas search takes place at the highest possible abstraction level in order to reduce the dimensionality of the search space. Verification is not as expensive as search, because it amounts to "execution" of a sequence of mappings and matching of the outcome of the execution against the output of the signal processing system.

It is not necessary for the mappings used to model the possible faults in the operation of the direction determination system to bear a direct relationship with specific components of the actual signal processing system (shown in Chapter 2). As an example, consider the mapping for modelling lack of resolution. Several factors can lead to lack of resolution, such as acoustic source directions (two acoustic sources very close in azimuth), temporal spectrum characteristics (two acoustic sources whose energy is concentrated in highly overlapping frequency bands) or inappropriate signal processing parameter settings (e.g. digital filter parameters).

By describing the input scenario in terms of the wavenumber spectrum, we can model both the atmospheric propagation phenomena and the signal processing transformations from the time-domain signals into the frequency-domain (i.e. wavenumber spectrum) output as mappings on wavenumber spectrum abstractions. Faults in either propagation phenomena or the signal processing transformations (in the form of inappropriate parameter settings) are associated with mappings that distort or modify their input wavenumber spectrum abstractions. The diagnosis task can be cast in terms of finding the one or more "offending" mappings that cause the discrepancy between the input scenario and the frequency domain output of the signal processing system.

Mappings are associated with system or scenario parameters. By explaining the system fault in terms of mappings, the fault is automatically associated with either system or scenario parameters. System parameters may be adjusted to modify the behavior of the system and reduce or possibly eliminate the fault. Associating a fault with scenario parameters helps recognize the limits of the system and avoid unnecessary processing. For example, if the distance between the aircraft and the sensor array is too large, it may not be possible to detect the aircraft.

One might ask the question why existing diagnosis techniques are not appropriate for complex signal processing systems that generate large amounts of complex intermediate data. The problem with existing techniques is that they are based on the analysis of intermediate data states that have either been saved during system operation (as in [Hudlicka & Lesser 1984]) or have been regenerated by simulation (as in [Davis1985]). In either case, the amount of intermediate data that has to be analyzed becomes prohibitive. Furthermore, signal processing data cannot be suitably abstracted for those techniques.

In our diagnosis system, analysis of intermediate data is avoided by effectively regenerating "intermediate" states from the system's input and output, all expressed as wavenumber spectrum abstractions. These "intermediate" states are not necessarily abstractions of actual states of the signal processing system, but they are states related to the alternative interpretation of the system operation in terms of the underlying Fourier theory (Figure 5.2).

In the following sections of this chapter we introduce the problem-specific concepts used in the design of the diagnosis system. The material is to a large extent specific to the direction determination system and was obtained through informal discussions with people experienced in the use of that system on real data. Although problem-specific, it is useful as an example of the ingredients needed to augment the diagnosis approach/architecture presented in Chapter 6 in order to produce a complete diagnosis system.

5.1 Wavenumber spectrum abstractions

In this section, we define the wavenumber spectrum abstractions that were used in the diagnosis system. The different levels of abstraction were designed to capture features of the wavenumber spectrum in varying degrees of importance for the diagnosis problem. At the same time, we introduce a specific diagnosis scenario, which we use as the context for illustrating the concepts of qualitative explanation and qualitative simulation, which rely on wavenumber spectrum abstractions and form the basis for the diagnosis system described in the next chapter.

5.1.1 Levels of abstraction of wavenumber spectra

The wavenumber spectrum is an explicit representation of important characteristics of the acoustic source scenario, such as the direction of arrival of acoustic waves at the sensor array and their frequency content. Therefore, it may be used to define useful abstractions for a conceptual description of the operation of the direction determination algorithm in terms of the underlying Fourier theory.

In diagnosing discrepancies between the input scenario and the output of the direction determination system, it is useful to represent the input scenario as a wavenumber spectrum, thereby having a uniform representation for reasoning in the Fourier domain, as shown in Figure 5.2. The wavenumber spectrum is described at multiple levels of abstraction, which helps reduce the search space in the diagnosis problem, by first seeking a solution at a high abstraction level, and by refining that solution at lower abstraction levels. Furthermore, wavenumber spectrum abstractions are described qualitatively, i.e. through ranges of values for the various features instead of values. Qualitative descriptions are adjusted to the accuracy with which the input scenario is known, by having broad ranges for low and narrow ranges for high accuracy.

The basic features of the wavenumber spectrum are shown in Figure 5.3. These features are organized in five different levels of abstraction, as shown in 5.4. Justification for the particular features selected and their organization in the particular levels of abstractions lies in informal study of the diagnostic reasoning of human subjects experienced in the behavior of the direction determination system. At the power and frequency levels, the shape of a radial ridge of the wavenumber spectrum is modelled by a rectangle, whose area is the power and whose radial limits are the minimum and maximum frequency $fmin$ and $fmax$. At the band level, we include the amplitude of the signal, in addition to its power and frequency. At the gaussian level, the signal is viewed as a gaussian truncated at both ends, specified by the frequencies at the two ends, $bmin$ and $bmax$.

- Direction (angle of radial line with a reference axis)

- Power (the total power of the temporal spectrum along the radial line)

- Minimum and maximum frequency (the limits of the frequency band over which the temporal spectrum has significant power)

- Amplitude (the maximum value of the temporal spectrum)

- Minimum and maximum band frequency (the limits of the frequency band over which the temporal spectrum has power at all. These frequencies are used to precisely account for overlap phenomena between signals).

Figure 5.3: Basic features of the wavenumber spectrum

We previously mentioned that an important aspect of the wavenumber spectrum abstractions used in the diagnosis system is that they are specified not by numeric values, but by ranges of values. This choice was made to account for the inaccuracy present in the specification of the input wavenumber spectra and measurement noise of the output wavenumber spectra. Single values are accommodated in the range system as ranges with identical ends. An arithmetic for ranges similar to that in [Abelson & Sussman 1985] was developed as a generalization of arithmetic for numbers.
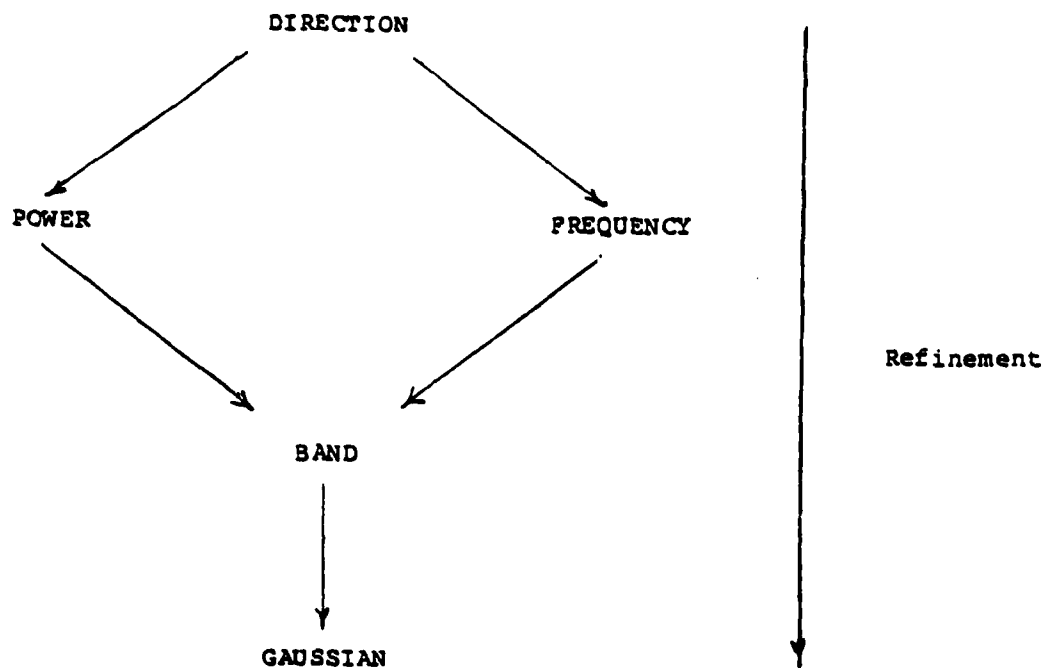
- Direction level (direction)

- Power level (direction and power)

- Frequency level (direction, minimum and maximum frequency)

- Band level (direction, minimum and maximum frequency, power and amplitude)

- Gaussian level (direction, minimum and maximum frequency, power, amplitude and minimum and maximum band frequency)

Figure 5.4: Levels of wavenumber spectrum abstractions.

157

Types of description relate to the nature of the wavenumber spectrum at different stages of its life from its generation by the physical sources to the output of the direction determination system. The types of description in the diagnosis system are shown in Figure 5.5. Types are useful in simplifying the search for an explanation of the system operation, because they place constraints on the construction of acceptable sequences of mappings in diagnosis. Particular mappings are applicable only to signals of specific types and have signals of specific types as outputs.

- Propagation (from the generation of the signal by the physical sources to it being sensed by the acoustic sensor)

- Continuous temporal (from the signal being sensed by the acoustic sensor to it being passed through analog to digital converter)

- Discrete temporal (from the signal being passed through the A/D converter until it being processed by the 2-D spectral estimator)

- Continuous spatial (Output of the 2-D spectral estimator - ideal)

- Discrete spatial (output of the 2-D spectral estimator - sampled)

Figure 5.5: Types of description of wavenumber spectrum abstractions.

For graphical computer display purposes, a wavenumber spectrum is represented as a circle (Figure 5.6). A radial ridge of the wavenumber spectrum is represented as a shaded sector centered at the mean of the direction range. The angular width of the sector corresponds to the width of the direction range. The radial extent of the sector (radial distance from the center corresponds to frequency) reflects the mean values of the $fmin$ and $fmax$ ranges. Finally, the amplitude range is shown as an interval label next to the corresponding sector.

158

A = [3, 5]
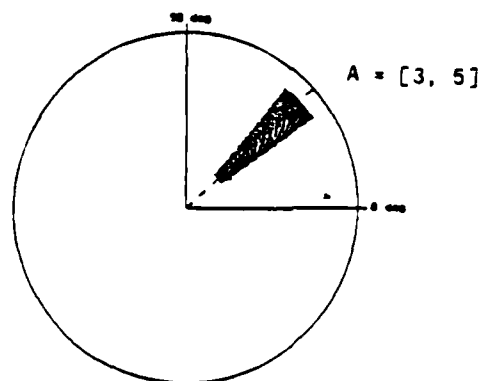
Figure 5.6: Computer-generated output from the diagnosis system.

5.1.2 A diagnosis scenario

In this subsection, we define an example of a diagnosis scenario that will be used to illustrate the concepts of this chapter. The same example will be solved by our diagnosis system in the next chapter.

The situation of the example is that of two aircraft flying in the area of a microphone array in such a manner that their directions with respect to the array are close to each other (Figure 5.7). A lack of resolution occurs in such a situation, in which the signal processing system detects just one acoustic source whose direction is between the two actual directions.

Let the two aircraft in our example be AIRCRAFT-1 and AIRCRAFT-2. The input wavenumber spectrum describes these aircraft qualitatively and is obtained by a priori knowledge of the scenario (see Figure 5.8). The output wavenumber spectrum is obtained as the output of the direction determination system. The scenario facts are also derived from knowledge of the scenario and they represent scenario knowledge that does not fit in the wavenumber spectrum description. System facts correspond to known parameter settings of the direction determination
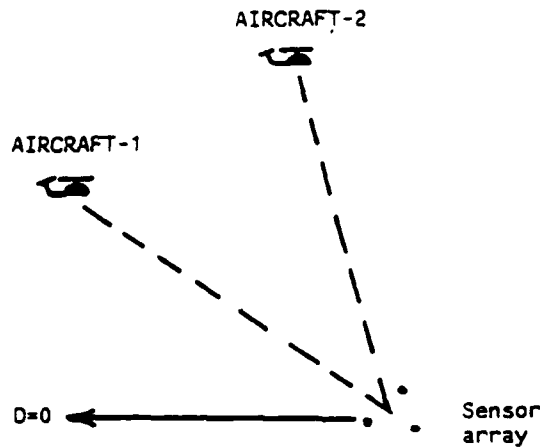
Figure 5.7: A diagnosis scenario.

system.

The diagnosis task consists of finding a sequence of mappings that explains the difference between the input and the output wavenumber spectra. The explanation should be "constructive", in the sense that the proposed sequence of mappings actually does map the input to something compatible with the output wavenumber spectrum at the lowest level of abstraction, i.e. the gaussian level. Compatibility means that there is overlap between the ranges of the corresponding features of the actual system output and of the output from the sequence mappings.

5.2 Mappings of wavenumber spectrum abstractions

The overall mapping from the frequency-domain description of the input scenario to the frequency-domain output is modelled in the Fourier domain as a sequence of "primitive" mappings. Each mapping is described in terms of how it transforms its input description. Such a transformation depends upon the values of parameters as-

Input:
  AIRCRAFT-1: direction in degrees D1 = [0 10]
              power in arbitrary units = [100 200]
              amplitude in arbitrary units = [7 8]
              fmin in Hz = [50 60]
              fmax in Hz = [90 120]

  AIRCRAFT-2: direction in degrees D2 = [35 50]
              power in arbitrary units = [100 200]
              amplitude in arbitrary units = [1 2]
              fmin in Hz = [0 30]
              fmax in Hz = [45 60]

Output:
  SIGNAL-3:   direction in degrees D3 = [20 20]
              power in arbitrary units = [100 200]
              amplitude in arbitrary units = [15 20]
              fmin in Hz = [0 20]
              fmax in Hz = [30 60]

Scenario facts:
    velocity of aircraft-1 = [150 150]
    direction of aircraft-1 increasing with time
    distance of aircraft-1 = [3 3]
    cos of the elevation of aircraft-1 = [0.5 0.5]

    velocity of aircraft-2 = [50 50]
    direction of aircraft-2 decreasing with time
    distance of aircraft-2 = [9 9]

System facts:
    cutoff frequency of the antialias filter = 200 Hz
    minimum frequency of the CAR filter = 0 Hz
    maximum frequency of the CAR filter = 100 Hz
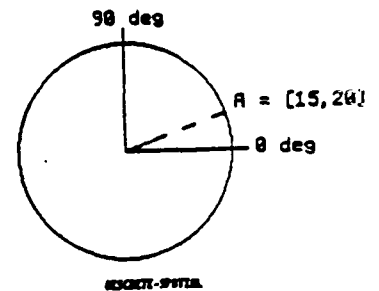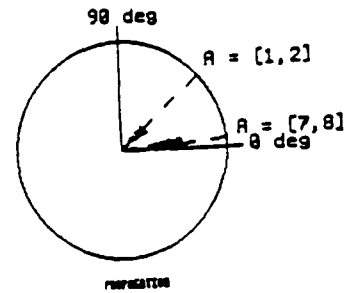
Figure 5.8: Wavenumber spectrum abstractions for the input and output of the direction determination system in the diagnosis scenario of the example.

161

sociated with the mapping. The mapping parameters have a direct correspondence with parameters of the direction determination system because they are derived from the same underlying theory. When the system is operating correctly, i.e. there are no discrepancies between the input scenario and the system output, each of the mappings is an identity mapping from its input to its output. In case of discrepancies, one or more of the mappings are non-identity mappings. The goal of diagnosis then becomes to identify a set of non-identity mappings that explain the discrepancies between input and output of the system.

The input wavenumber spectrum passes through several mappings before it becomes the wavenumber spectrum that is the output of the system. Such mappings may introduce distortions, giving rise to an output wavenumber spectrum that differs from the input one in several respects. Figure 5.9 shows examples of mappings and the associated distortions that arise because of wave propagation or signal manipulations in the direction determination system.

- Propagation: Power of temporal spectrum is attenuated according to the distance of the source from the sensor array. Frequencies are being scaled because of the Doppler phenomenon. Frequencies are also being scaled by the cosine of the elevation angle. Addition of atmospheric noise.

- Spatial sampling: Broadening by convolution with the transform of the spatial window. Spatial aliasing if smallest distance between sensors is too large.

- Temporal sampling: Temporal aliasing if analog antialiasing filters have failed.

- Complex analytic representation: It is computed on a block basis and this introduces broadening of the temporal spectrum by convolution with the transform of the temporal window.

- Temporal averaging of CAR samples for a good estimate of the zero-delay covariance: may introduce broadening of temporal peaks due to frequency modulation, if the peaks drift over the averaging interval.

Figure 5.9: Examples of mappings and the corresponding distortions.

Each mapping is specified at a particular level of abstraction by providing the description of the output signal as a function of the input signal at that abstraction

162

level and the scenario and system parameters. Figure 5.10 shows the information
associated with a mapping.

- Input signal types: these are the allowable types that the input signal may have.

- Output signal type: the type of the output signal.

- Differences affected: the kinds of differences between the input and the output signal
  that are explained by the mapping. (Note that we need to explicitly represent this
  piece of information for use when we are trying to explain the differences between an
  input and an output signal in the direction determination system. We only want to
  consider mappings that are relevant to the differences we are experiencing).

- Mapping parameters: Parameters of the direction determination system which can
  be changed to affect the outcome of the mapping.

- Input signal preconditions: Conditions that the input signal must satisfy in order for
  the mapping to be a non-identity mapping.

- Scenario preconditions: Conditions that the scenario must satisfy in order for the
  mapping to be applicable.

- Output signal postconditions: Conditions that provide a description as a function of
  the input signal description, the scenario facts and the system facts.

Figure 5.10: Ingredients of a mapping.

As an example, consider the equal-resolution mapping. (Figure 5.11). The input
wavenumber spectrum has two signals, each represented by a radial ridge, while the
output has only one signal. The two ridges are close to each other in direction and
therefore get merged into a single ridge because of lack of sufficient resolution. The
mapping is capable of explaining a resolution difference, occurring when the input
wavenumber spectrum has two signals and the output wavenumber spectrum only
one in some intermediate direction. Associated mapping parameter is the array
aperture at the direction and power level. At the frequency, band and gaussian
level, another parameter is relevant, *epsilon*, a small number added to the diagonal
of the covariance matrix to make the estimation of the spatial Fourier transform of
the sensor signals a stable procedure. The mapping parameters are not used by the
diagnosis system, but they will be useful when the system is extended to include

163

adjustment of the parameters of the signal processing system to fix the diagnosed faults (not implemented yet).

Input signal precondition at the direction level is that the difference of the direction ranges (according to the range arithmetic defined in Appendix B) intersects with the range [0, 100/array-aperture]. This means that the directions of the two signals must be close to each other, so that it makes sense to consider lack of resolution as a possible cause for the merging of the two signals into one. At the power level, we require in addition that the power of the two signals is non-zero. At the frequency level, we require in addition to the directions being close that the frequency ranges overlap (signals that occupy different frequency bands cannot be merged into one). Scenario preconditions are associated with required characteristics of the source movements (e.g. source velocity). There is none in the case of the equal-resolution mapping.

The output signal postconditions are the following. The direction of the output signal is equal to the cover of the ranges indicating the directions of the input signals. The power of the output signal is the range between 0 and the sum of the maximum powers of the two signals. The minimum and the maximum frequencies of the output signal equal the covers of the corresponding ranges of the input signals. Similar postconditions are valid for the amplitude and gaussian levels.

Figure 5.12 shows an example of the effect of the equal-resolution mapping in a situation with two aircraft that are close in direction and have identical frequencies. On the left is the wavenumber spectrum before the application of the equal-resolution mapping. On the right is the result after the application.

Figure 5.13 summarizes a number of mappings currently available in the diagnosis system. Each mapping in this figure has a complete representation similar to that shown in Figure 5.11 for the equal resolution mapping.

EQUAL RESOLUTION MAPPING

  INPUT SIGNAL TYPE: propagation, continuous-temporal,
                         discrete-temporal, continuous-spatial
  OUTPUT SIGNAL TYPE: continuous-spatial
  DIFFERENCES AFFECTED: resolution
  MAPPING PARAMETERS:
    DIRECTION   array-aperture
    POWER   array-aperture
    FREQUENCY   array-aperture, epsilon
    BAND   array-aperture, epsilon
    GAUSSIAN   array-aperture, epsilon

INPUT SIGNAL PRECONDITIONS: Per pair of input signals
  DIRECTION   Direction difference intersects [0, 100/array-aperture].
  POWER   Direction level preconditions and non-zero Power.
  FREQUENCY   Minimum and maximum frequencies intersect.
               Direction difference intersects range
                  [0,1000*epsilon)/(array-aperture*.0001*maximum-freq)].
  BAND   Power level preconditions.
               Frequency level pre-conditions.
               Amp in [0, inf]
  GAUSSIAN   Frequency level pre-conditions with but on the gaussian model.

SCENARIO PRECONDITIONS: none

OUTPUT SIGNAL POSTCONDITIONS: Per pair of input signals
  DIRECTION   Delete input signals.
               Create signal whose direction is the
                  cover of the two input directions.
  POWER   Direction level postconditions.
               Power of output signal is range
                  [maximum of lowest values of powers of input signals,
                   maximum of highest values of powers of input signals]
  FREQUENCY   Direction level postconditions.
               Minimum-freq of output equal to the cover of minimum
                  freqs of inputs
               Maximum-freq of output equal to the cover of maximum
                  freqs of inputs
  BAND   Frequency level postconditions.
               Power level postconditions.
               Amp of output signal: relation similar to that of power.
  GAUSSIAN   Band level postconditions with gaussian model.

Figure 5.11: The equal resolution mapping

165
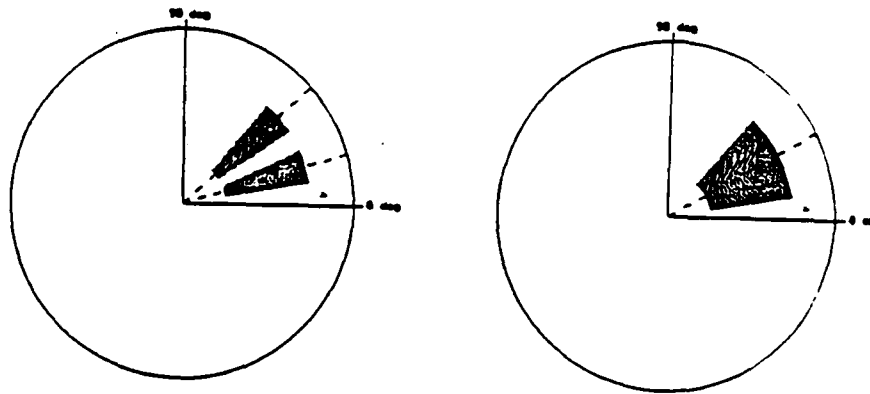
Figure 5.12: Example of the equal resolution mapping at the frequency level.

5.3 Differences between wavenumber spectra and their ordering

An important step toward diagnosis is finding and ordering differences between wavenumber spectra at multiple levels of abstraction. Differences can potentially be explained by the presence of non-identity mappings. Therefore, at a simplistic level, diagnosis may take place by finding the differences between the input and output wavenumber spectrum abstractions of the direction determination system, trying out all possible sequences of mappings that might explain these differences and finally choosing the simplest mapping sequence (by some criterion) that successfully explains the differences. In the next chapter we will see how this approach is complemented by a control strategy that helps reduce the search space for all possible combinations of mappings that explain a certain difference.

Multiple levels of abstraction play an important role, because differences at higher abstraction levels are considered first. If mappings for explaining those dif-

EQUAL RESOLUTION MAPPING: merges two signals that are close in direction and whose frequency distributions are the same. The effects of this mapping are dependent upon array aperture and the epsilon factor for the covariance matrix.

FAST VELOCITY MAPPING: shifts the direction of a signal because of fast aircraft velocity. The degree of direction shift depends upon the analysis-interval parameter of the signal processing system.

PRE-CPA DOPPLER MAPPING: shifts upward the frequencies of a signal due to an aircraft that is approaching a node with a significant velocity (compared to velocity of sound).

POST-CPA DOPPLER MAPPING: shifts downward the frequencies of a signal due to an aircraft that is receding from a node with a significant velocity (compared to velocity of sound).

ELEVATION COMPRESSION MAPPING: shifts the frequencies downward and the signal amplitude correspondingly upward. This occurs in accordance with the elevation of the aircraft with respect to the sensor array.

CAR FILTER MAPPING: causes frequency changes in the signal because of the filter performing the computation for complex analytic representations of the signals.

ANTIALIASING FILTER MAPPING: causes frequency changes in the signal because of the analog filter meant for avoiding temporal aliasing in the digitization process.

CAR-GHOSTING MAPPING: causes a ghost signal to appear at the direction which is 180 degrees away from the original signal. The phenomenon occurs if the original signal has considerable power in the lower frequencies. The effect of the mapping depends upon the block-length parameter of the signal-processing system.

SPATIAL ALIASING MAPPING: causes signals to wrap around at different directions due to the sparseness of the array. The effects of this mapping depend upon the minimum sensor separation in the acoustic array.

RADIAL-INTEGRATION MAPPING: causes frequency changes due to the frequency interval over which signal power is determined by the signal processing system.

RANGE SCALING MAPPING: causes amplitude changes due to the distance of the aircraft from the sensor array. The power scaling also depends upon the analysis interval and the block-length used by the signal processing system.

Figure 5.13: Summary of several mappings of the diagnosis system

ferences are found to successfully map the input wavenumber spectrum onto the output one at the lowest abstraction level, these are the final explanation. If such mappings fail to map the input onto the output at the lowest abstraction level, then differences at a lower abstraction level than the original level are considered and the process is repeated. This is an instance of the principle of searching at the highest possible abstraction level but verifying the result at the lowest possible abstraction level.

Figure 5.14 shows the differences currently used by the diagnosis system. Differences at lower abstraction levels have more detailed (and therefore stricter) preconditions than the same differences at higher abstraction levels, as for example the resolution difference. Other differences are not present at all at higher abstraction levels, as for example the frequency shifting difference.

In terms of implementation, the difference finding program is built as a rule base in YAPS [Allen1983]. Each rule implements one kind of difference of Figure 5.14 and it acts independently of other rules: if its preconditions are satisfied, it adds the name of the corresponding difference to the list of differences between the input and the output wavenumber spectra, which are both input to the difference finding program. The output of the difference finding program is the list of differences. Figure 5.15 shows the inputs and the outputs of the difference finding program. System facts are parameters of the direction determination system and scenario facts are scenario information, such as speed and distance of the acoustic sources from the sensor array.

As an example, consider the input and output for the example diagnosis scenario of Figure 5.7. The difference finding program in this case gives the following list of differences: AMPLITUDE-SCALING, GHOSTING, DIRECTION-SHIFTING and RESOLUTION.

In general, two wavenumber spectra have more than one difference. Since a

168

- Amplitude scaling: The output wavenumber spectrum has no signal with a direction range that overlaps with the direction range of a signal present in the input wavenumber spectrum. Applicable at all levels of abstraction.

- Amplitude scaling: The signal in the output wavenumber spectrum has a power range that is strictly less than the range of the corresponding signal in the input wavenumber spectrum. Applicable at the power level.

- Amplitude scaling: The signal in the output wavenumber spectrum has an amplitude range that is strictly less than the range of the corresponding signal in the input wavenumber spectrum. Applicable at the amplitude level.

- Direction shifting: There is a signal in the output wavenumber spectrum whose direction range does not overlap with the direction range of an input signal, and there is no other signal in the output wavenumber spectrum, whose direction range does overlap with that of the input signal. Applicable at all levels of abstraction.

- Resolution difference: There are two signals in the input wavenumber spectrum, whereas there is only one signal in the output wavenumber spectrum, whose direction range overlaps with the cover of the direction ranges of the two input signals. Applicable at the direction level.

- Resolution difference: There are two signals in the input wavenumber spectrum, whereas there is only one signal in the output wavenumber spectrum, whose direction range overlaps with the cover of the direction ranges of the two input signals, and whose power range either overlaps or is strictly less than the sum of the power ranges of the input signals. Applicable to all levels except direction.

- Ghosting: There is no signal in the input wavenumber spectrum with direction range overlap with a signal in the output wavenumber spectrum. Applicable at all levels.

- Frequency shifting: There is a signal in the output wavenumber spectrum with direction overlap with a signal in the input wavenumber spectrum, but the minimum frequency and maximum frequency ranges of the two signals do not both overlap. Applicable at the frequency, band and gaussian level.

Figure 5.14: Differences currently used by the diagnosis system

169
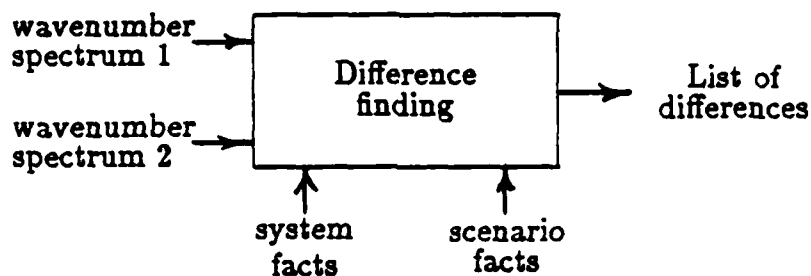
Figure 5.15: Inputs and outputs of the difference finding program.

mapping is associated with a single difference, we need to select one difference out of the list of differences, as the candidate mapping that will be verified whether it maps the input to the output wavenumber spectrum at the lowest abstraction level. The list of differences thus serves as a cueing mechanism in the search for the appropriate mapping.

In order to select one difference out of a list of differences, difference finding is followed by difference ordering. In our diagnosis system, ordering of the list of differences proceeds in two stages. In the first stage, the number of cases of PRESENT (signals in the input wavenumber spectrum with an associated signal in the output wavenumber spectrum), MISSING (signals in the input wavenumber spectrum without an associated signal in the output wavenumber spectrum) and UNASSO-CIATED (signals in the output wavenumber spectrum with no associated signal in the input wavenumber spectrum) are counted. The three counts, together with the scenario and system facts, are the only factors determining how the ordering of differences will be done.

Difference ordering is implemented as a rule base, with each rule making some change to the order of differences. The change consists of bringing to the top

or the bottom of the list particular differences, depending on the three counts of the previous paragraph or the scenario and system facts. Figure 5.16 shows two examples of rules in the difference ordering rule base, with their preconditions and actions. The first rule makes difference order changes based on the PRESENT, MISSING and UNASSOCIATED counts. The second rule makes an order change based on a scenario fact.

```
Rule: Unassociated-signals
----------------------------

IF: The number of missing plus the number of present signals is at least
two
    and there is at least one unassociated signal

THEN: bring GHOSTING to the top
      bring FREQUENCY-SHIFTING to the top
      bring AMPLITUDE-SCALING to the top
      bring DIRECTION-SHIFTING to the top
      bring RESOLUTION to the top


Rule: Cutoff-less-than-100-Hz
------------------------------------------

IF: The cutoff frequency of the antialiasing filter is less than 100 Hz

THEN: Swap the position of the differences FREQUENCY-SHIFTING and
AMPLITUDE-SCALING.
```

Figure 5.16: Examples of rules for ordering differences

For example, in the case of the input and output wavenumber spectra in the diagnosis scenario of Figure 5.7, the count is missing-direction = 2 and unassociated-direction = 1. Using the first rule of Figure 5.16, the ordered list of differences is RESOLUTION, DIRECTION-SHIFTING, AMPLITUDE-SCALING and GHOSTING.

5.4 Mapping selection to explain differences

The program for mapping selection returns an ordered list of mappings given a difference to be explained and the scenario and system facts. Figure 5.17 shows two examples of rules for mapping selection. In both examples, the difference to be explained is a FREQUENCY-SHIFTING difference. In the first example, the cutoff frequency of the antialiasing filter is less than 200 Hz, and in this case the antialiasing-filter mapping is the first one to be considered. In the second example, the cutoff frequency is greater than 200 Hz, in which case the elevation-compression mapping is a more likely explanation.

```
Rule: Frequency-shifting-antialiasing
-------------------------------------------------

IF: The difference is frequency-shifting and
    the cutoff frequency of the antialiasing filter is less than 200 Hz

THEN: the list of mappings in decreasing importance is:
    antialiasing-filter-mapping, car-filter-mapping,
    elevation-compression-mapping


Rule: Frequency-shifting-elevation-compression
-----------------------------------------------------
▼
IF: The difference is frequency-shifting and
    the cutoff frequency of the antialiasing filter is greater than 200 Hz

THEN: the list of mappings in decreasing importance is:
    elevation-compression-mapping, car-filter-mapping
```

Figure 5.17: Examples of rules for mapping selection.

5.5 Qualitative Explanation and Qualitative Verification

In a specific scenario, a mapping may act as an identity or non-identity mapping. An identity mapping does not distort its input signal. As an example, consider (ideal) low-pass filtering. If the highest frequency of the input signal is below the cutoff frequency of the filter, then filtering acts as an identity mapping, because the output signal is the same as the input signal. Otherwise, it acts as a non-identity mapping. We see that whether a mapping acts as an identity depends both on the characteristics of the mapping and on the input signal (i.e. the specific scenario).

The behavior of mappings, which arise because of signal processing operations in the direction determination system, can be modified by adjusting related system parameters. Mappings that arise because of wave propagation typically cannot be similarly affected, because the signal processing has no control over them. They operate on the *wavenumber spectrum before it is being picked up by the sensors* (propagation phenomena). Non-identity mappings, whose behavior can be controlled by parameter adjustment can be made to behave as identity mappings in a specific scenario by adjusting the corresponding system parameters.

Based on the above notion of mappings, we define *Qualitative Explanation* at a given abstraction level as the process of finding a minimal sequence of non-identity mappings that maps the given input wavenumberspectrum (derived from knowledge of the acoustic source scenario) onto a wavenumber spectrum that is compatible with the output wavenumber spectrum (computed by the direction determination system) at that abstraction level. Compatibility is defined as non-empty overlap of the ranges of the corresponding features. Qualitative Explanation is the "generate" step towards diagnosing the discrepancies between the input and the output of the signal processing system through a *generate-and-test* approach.

Ideally, the only difference between the input and output spectra is in their type

173

(Figure 5.5): the input spectrum is propagation type while the output spectrum is discrete spatial type. In this case, no non-identity mappings are present, and the explanation of the operation of the direction determination system is simple. In practice, the input signal is different from the output signal in some respects, because of the presence of non-identity mappings. The differences between the input and the output signal can serve as clues in identifying the non-identity mappings involved.

Determining a valid explanation is thus seen as a search through the space of all possible sequences of existing mappings. The search is guided by the differences between the input and output signal and proceeds by introducing mappings that partly explain these differences, thereby reducing the original problem into several subproblems that attempt to reduce the remaining differences (as in the General Problem Solver [Newell & Simon 1963]). The search first takes place at the highest level of abstraction (direction level) and drops abstraction level if no explanation can be found at the current level or if the explanation found does not yield an output compatible with the system output at the lowest abstraction level.

Checking for compatibility is the task of *Qualitative Verification*. In this task we have the input wavenumber spectrum and a sequence of mappings and we want to determine the wavenumber spectrum that would be computed by the direction determination system based on the underlying Fourier domain theory (i.e. without collecting the data or running it through the direction determination system) and confirm its compatibility (in the sense defined previously) with the wavenumber spectrum which is the actual output of the direction determination system. The task is carried out at the lowest abstraction level at which the input and output wavenumber spectra are specified. Qualitative verification is the "test" step in the "generate-and-test" diagnosis.

Figure 5.18 shows how the equal resolution mapping can help explain the difference between the input and output of the diagnosis scenario shown in Figure 5.7 in

174

page 160 at the direction level (i.e. ignoring all frequency and power information). The direction difference of the two signals in the input of the diagnosis scenario is too large for the equal-resolution mapping alone to explain the difference in number of signals between input and output. Another mapping, the fast velocity mapping, is capable of changing the direction of a signal, thus enabling the application of the equal-resolution mapping. By referring to the scenario facts, it turns out that the velocity of aircraft-1 is large, therefore the fast-velocity mapping is applicable to the input state, which changes the direction of the signal corresponding to aircraft-1.
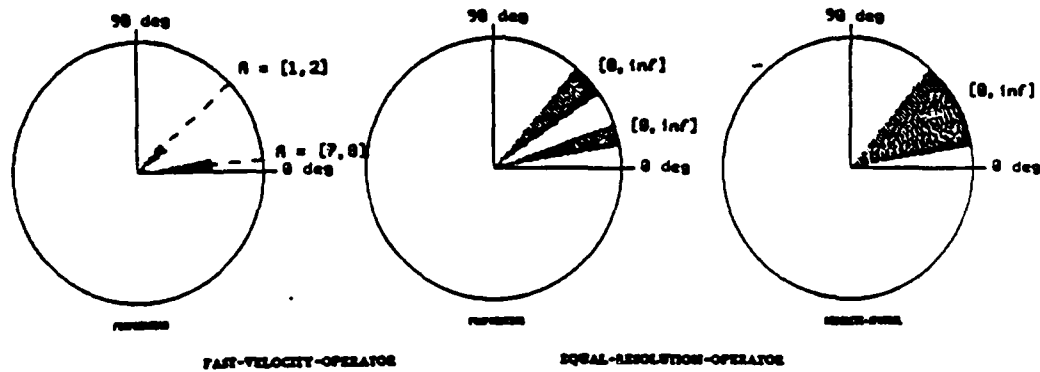
Figure 5.18: Fast-velocity and Equal-resolution mappings explain the difference at the direction level between input and output of the diagnosis scenario defined earlier

The fast velocity mapping followed by the equal resolution mapping gives an output that is compatible with the system output at the direction level. However, once we try to verify this explanation at the frequency level, we find that the frequencies of the two input signals do not overlap, therefore they violate a precondition of the equal resolution mapping at the frequency level. The equal-resolution mapping must be augmented by other mappings that will be found by search at levels below the direction level.

Chapter 6

Diagnosis of a Signal-Processing System using Signal Mappings

In the previous chapter we introduced the problem of fault diagnosis of a direction determination system and we defined wavenumber spectrum abstractions for expressing the intermediate states of the system in a Fourier domain model of its operation. In this chapter we present an implemented diagnosis system that uses the concepts and representations introduced so far in addition to a "means-ends" control strategy to search the space of all possible diagnosis outcomes. This control strategy extends the General Problem Solver framework [Newell & Simon 1963] to include qualitative descriptions of the states at multiple levels of abstraction.

6.1 Overview of the control strategy of the diagnosis system

In the diagnosis system, we assume that the user provides a qualitative description of the input scenario in the form of multiple levels of abstraction of a wavenumber spectrum for the situation. The diagnosis system uses the differences between the

input scenario and the output of the direction determination system to identify a sequence of mappings that explain these differences.

In the planning language, the diagnosis system can be viewed as performing a state-space search where the states are wavenumber spectrum abstractions and the operators are the various mappings as defined in the previous chapter. The initial state is derived through a straightforward computation from the user-specified description of the input scenario. The goal state is a description of the wavenumber spectrum at the output of the signal processing system.

The task of diagnosis can now be phrased as the following search problem: Identify a sequence of mappings such that:

1. the sequence maps the initial state into the goal state, and

2. no proper subsequence of the sequence maps the initial state into the goal state.

We talk about a sequence of mappings, because, as we will see later in this chapter, the order in which the mappings are applied to the input state is important. The requirement that there be no proper subsequence that can map the initial state into the goal state ensures that any mappings not necessary for explaining the discrepancy between the initial state and the goal state are excluded.

Another aspect of diagnosis is that there may exist more than one sequence of mappings that satisfies our search criterion. Unless intermediate signal states of the actual signal processing system are available for inspection, such multiple explanations cannot be disambiguated. Because in our signal processing system such intermediate states are not easily accessible (due to the large volume of data they constitute), we have designed our search strategy so that it finds the sequence with the smallest number of mappings. Equivalently, we apply the heuristic that "the simplest explanation is the most likely explanation for the cause of a system

fault".

The basic strategy (Figure 6.1) is to first construct a qualitative explanation at the highest abstraction level and then attempt to verify it. If the verification succeeds, the correct explanation has been found. If it fails, the current explanation is modified and fed back into the GPS explainer. If the explanation can be successfully readjusted, a new simulation is tried using the adjusted explanation. Otherwise, the level of abstraction of the analysis is dropped and a new explanation is sought at the new level of abstraction by the GPS explainer. If the GPS explainer fails to come up with an explanation at the current level of abstraction, the level is immediately dropped and a new explanation is sought.

The abstraction level of verification is chosen to be the lowest one at which a description of the input signal state is known. Verification makes use of the same mapping and state representation mechanisms and the same specific mappings as those in the GPS explainer. In particular, verification can be viewed as a degenerate case of the the GPS explainer at the detailed abstraction level. The difference is that in this case the verifier does not have to perform any search. Instead, it uses the mappings specified by the explanation to be verified.

There are two types of failure. In the first type, the pre-conditions of a mapping in the explanation are not satisfied by the state preceding the mapping. In the second type, the output of a mapping does not match the qualitative description anticipated for it in the original explanation. For both situations, the plan is readjusted by eliminating the failed mapping from the explanation and searching for a sequence of mappings to replace it.

In both types of failure, the sequence of mappings found by trying to adjust the original explanation can be at lower levels of abstraction. Adjustment of a failed explanation leads to the same kind of search as the original, where input and output are either intermediate states or pre-conditions of mappings, which, according to

178
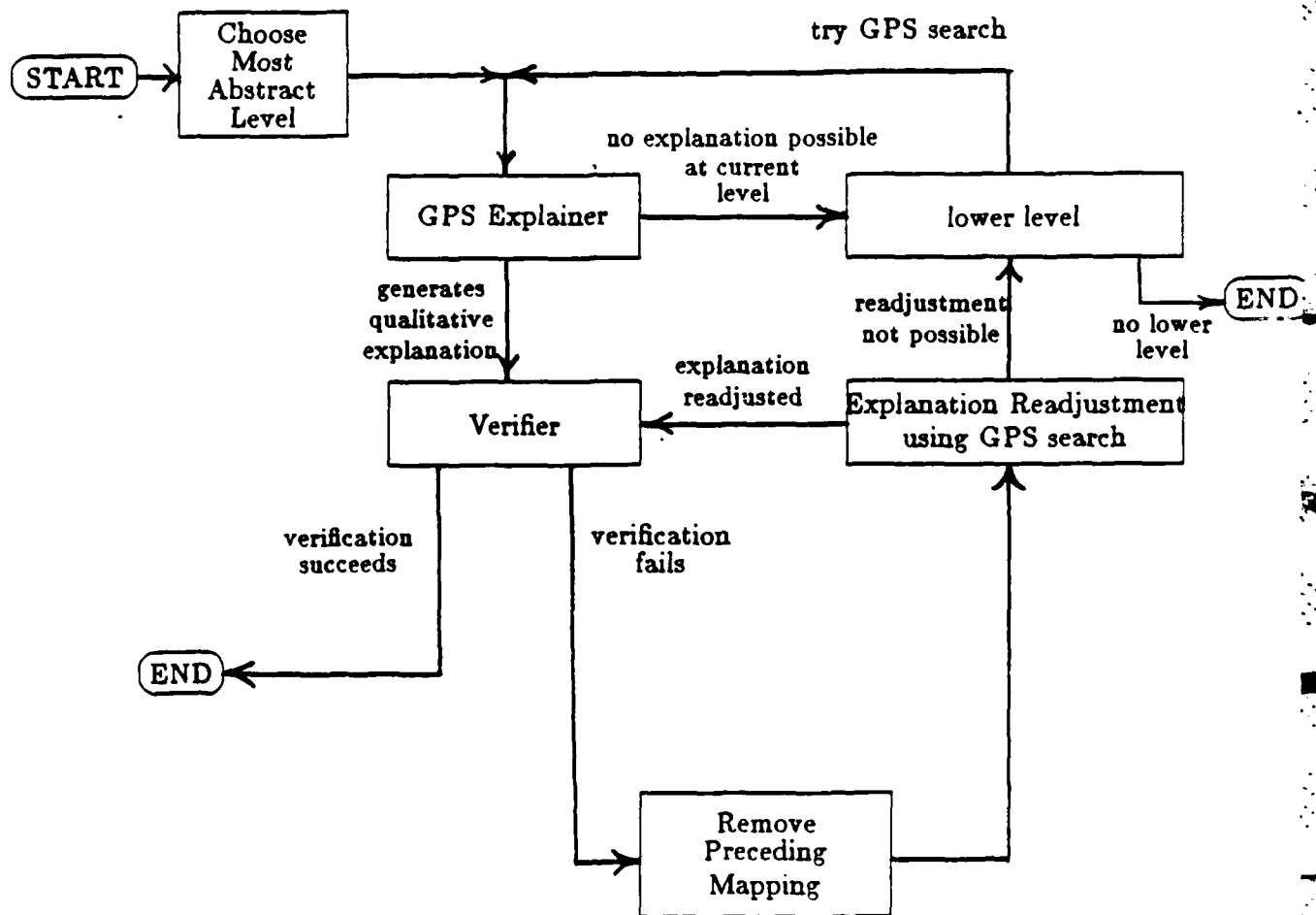
Figure 6.1: Diagnosis strategy by search for an explanation and verification.

our formulation, are also viewed as states.

6.2 GPS search for an explanation at a specific abstraction level.

In the previous section, we saw that a search for an explanation is a basic operation of the diagnosis system both in formulating an original explanation and adjusting an explanation that failed in its verification at a lower abstraction level. In this section we examine in greater detail the process of finding a sequence of mappings to explain the difference between an initial state S0 and a final state SF at a specific abstraction level I, where both S0 and SF are wavenumber spectrum abstractions.

Figure 6.2 is a diagram showing the GPS search. The first step is to find and order the differences between the qualitative descriptions of S0 and SF at abstraction level I. The most important difference is then selected, which leads to an ordered list of mappings that possibly explain it. The mapping Q closest to the top of the list and not already in the partially formed explanation is selected for inclusion in the explanation.

After the mapping Q has been selected, an attempt is made to apply it to the input state S0. Three possibilities exist at this point regarding the applicability of Q to S0, or equivalently, the difference between S0 and the preconditions of Q, pre(Q).

1. The difference between S0 and pre(Q) can potentially be explained. In this case, a new search problem is spawned, with input state S0 and output state pre(Q). If this subproblem succeeds in producing a subexplanation, Q is applied to the output obtained by applying the sequence of mappings composing that subexplanation to S0. The subexplanation becomes part of the final ex-
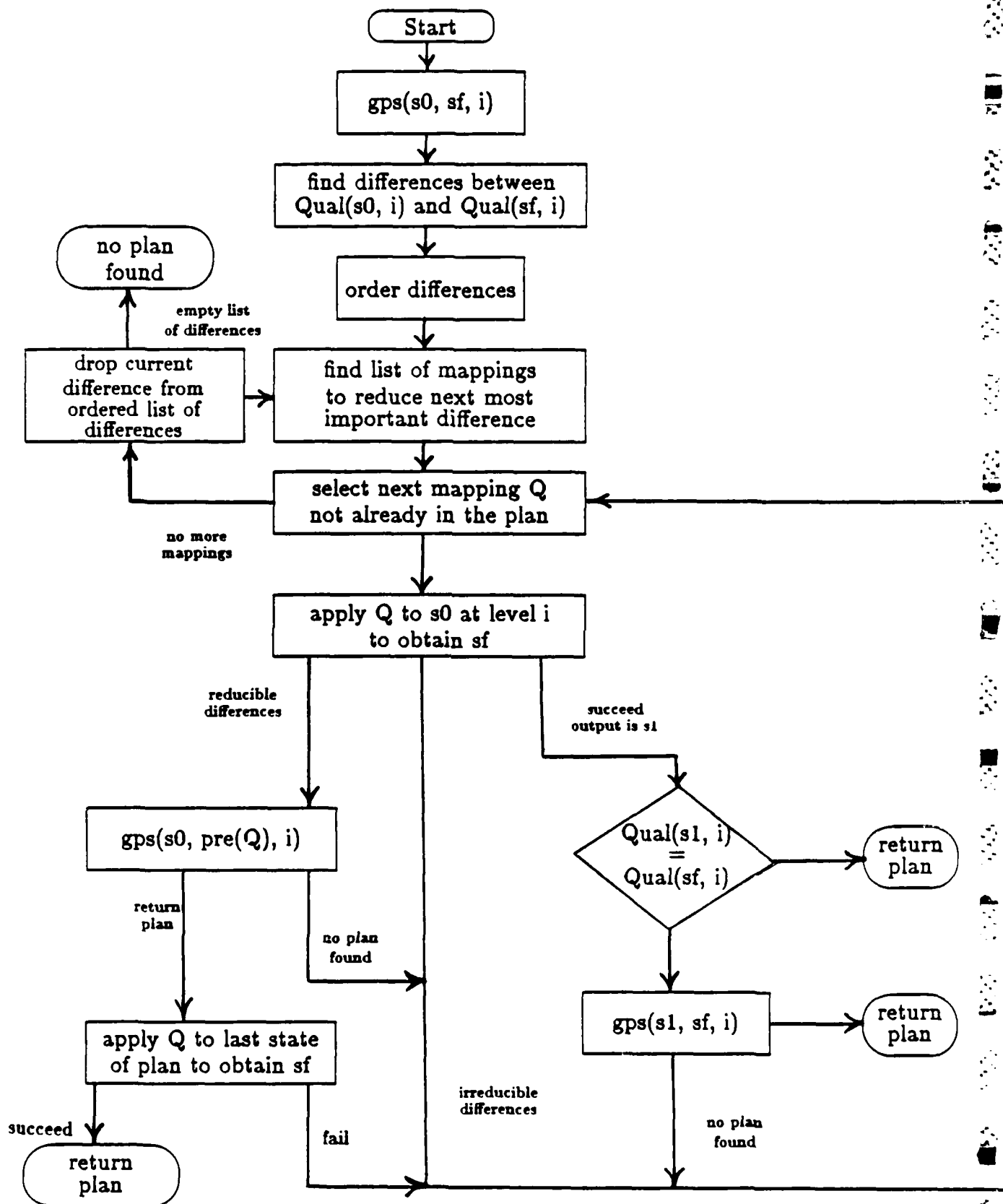
Figure 6.2: GPS Search for an explanation of the differences between S0 and SF at a specific level of abstraction I.

planation.

2. The difference between S0 and pre(Q) is irreducible. In this case, Q is discarded and a new mapping is tried from the list of mappings.

3. Q is successfully applied to S0 and the outcome is a state S1. Here we test whether S1 is compatible with SF at abstraction level I. If it is, the search is finished. If it is not, a new search problem is spawned, with input state S1 and output state SF at abstraction level I.

Control of the GPS search is accomplished through two important mechanisms.

First, no mapping is allowed to appear more than once in a particular explanation. This is in contrast to GPS search in other domains, such as algebraic simplification, where the same operator can be used many times. In such situations, the GPS search requires "depth heuristics" to avoid infinite loops. In our problem, the constraint that each mapping may appear only once in an explanation eliminates the need for depth heuristics.

Second, each mapping specifies the allowable types of the input and output states. Therefore, neighboring mappings in an explanation must satisfy their type constraints. This helps reduce the search space considerably. It should be noted, however, that mappings whose input and output states belong to the same single domain can appear in any order with respect to each other.

6.3   A diagnosis example

In the previous chapter we introduced a diagnosis scenario to illustrate the concepts of mapping, qualitative explanation and qualitative verification. In this section we present the results of the operation of the implemented diagnosis program on this scenario.

The scenario basically corresponds to a situation where lack of resolution on the part of the direction determination system results in an output wavenumber spectrum showing only one acoustic source, whereas it is known a priori that there are two aircraft in the area of the microphone array. However, the equal-resolution mapping alone is not sufficient to explain the lack of resolution in the above scenario, because of additional distortions present in both atmospheric propagation and processing in the direction determination system. It is due to these additional distortions that diagnosis in the particular scenario becomes nontrivial.

At the highest level of abstraction, the direction level, two differences between the initial search state S0 and the final search state SF (corresponding to the input and the output of the direction determination system respectively) are detected: **RESOLUTION** and **DIRECTION-SHIFT**. A direction-shift is a difference characterized by a direction in the initial state shifting to a different direction in the final state.

In the diagnosis system, the resolution difference is considered more important than direction-shift and the selected mapping is **EQUAL-RESOLUTION**. This mapping, as explained in the previous chapter, acts upon two signals whose directions are close to each other (direction level) and in addition have the same minimum and maximum frequencies (frequency level).

The equal-resolution mapping alone cannot explain the difference between S0 and SF at the direction level because one of its preconditions is that two sources be closer than 20 degrees in the initial state. The two sources in S0 are at least 25 degrees apart. This is the case of a reducible difference in Figure 6.2. In the search subproblem spawned, it is found that a mapping called **FAST-VELOCITY** may be able to explain the direction-shift difference now encountered. The fast-velocity mapping represents the effects of fast aircraft velocities on the output of the direction determination system. The scenario precondition of the fast-velocity mapping is that the speed of the aircraft be greater than 200 / analysis-interval m/s,

183

where *analysis – interval* is one of the parameters of the signal processing system. Indeed, aircraft-1 has a speed of 150 m/s, while the analysis interval is 4s. In the resulting state S1, aircraft-1 has a shifted direction in the range [10, 20], which does lie within 25 degrees of the direction of aircraft-2. Therefore the equal-resolution mapping applies and gives an output state with one signal in direction [10, 50]. This is compatible with the final state SF with direction 20 degrees. Thus the qualitative explanation at the direction level consists of the fast-velocity mapping followed by the equal-resolution mapping, as shown in Figure 6.3.
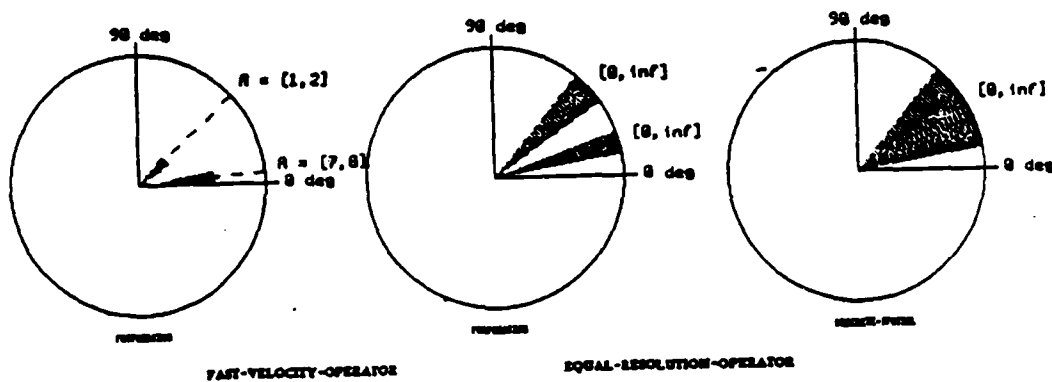
Figure 6.3: Qualitative explanation at the direction level.

The above plan is passed to the verification program at the gaussian level of abstraction. At this level it is found that the output state S1 produced by the fast-velocity mapping does not match the preconditions of the equal-resolution mapping because the two aircraft do not have overlapping minimum and maximum frequencies. This is a failure of the first type. According to Figure 6.1, the task now is to adjust the current explanation by producing a subexplanation that explains the difference between S1 and the preconditions of the equal-resolution mapping. Since the frequency-shift difference cannot be dealt with at the direction level, explanation must be searched for at the frequency level.

The ELEVATION-COMPRESSION mapping is selected as potentially capable of explaining a frequency-shift difference. The elevation-compression mapping represents the phenomenon of frequency scale compression for aircraft that are at high elevations with respect to the microphone array. The precondition of this mapping requires that an aircraft have a non-zero elevation with respect to the microphone array. In our example, the elevation of aircraft-1 is 45 degrees, therefore the elevation-compression mapping is applicable to state S1 and produces another state S2. The signal in state S2 corresponding to aircraft-1 has minimum and maximum frequency that satisfies the preconditions of the equal-resolution mapping. Therefore, a complete explanation at the frequency level consists of fast-velocity, elevation-compression and equal-resolution (Figure 6.4).
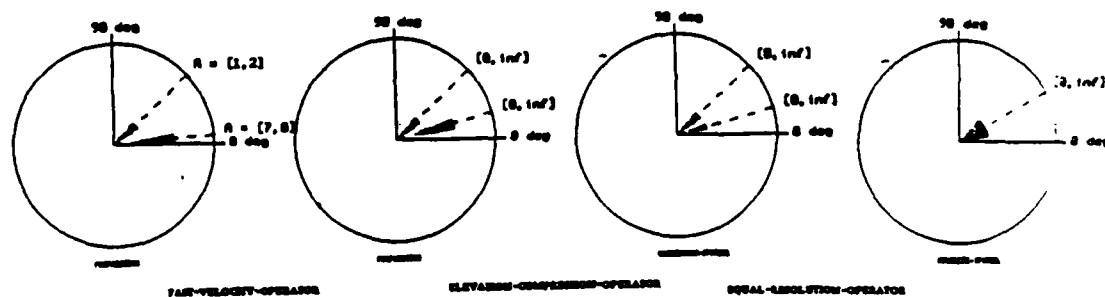
Figure 6.4: Qualitative explanation at the frequency level.

The adjusted explanation is again passed to the verification stage at the gaussian level of abstraction. It is found that the output state S2 of the elevation-compression mapping does not match the preconditions of the equal-resolution mapping. There is still a frequency-shifting problem. This is a failure of the second type, because the elevation-compression mapping did not achieve its intended goal. In accordance to Figure 6.1, we drop the elevation-compression mapping and seek an explanation to replace it. No mapping that can change the frequencies appropriately is found.

185

The abstraction level is then dropped to the band level, leading to the selection of the CAR-FILTER mapping. The CAR-filter mapping changes the amplitudes of the frequency components of a signal, making some of the components zero and reducing the amplitude of others. The CAR-filter mapping followed by the elevation-compression mapping succeeds in explaining the frequency difference. The final explanation consists of the following sequence of mappings: fast-velocity, CAR-filter, elevation-compression, and equal-resolution. This explanation is verified successfully at the gaussian level and is shown in Figure 6.5.
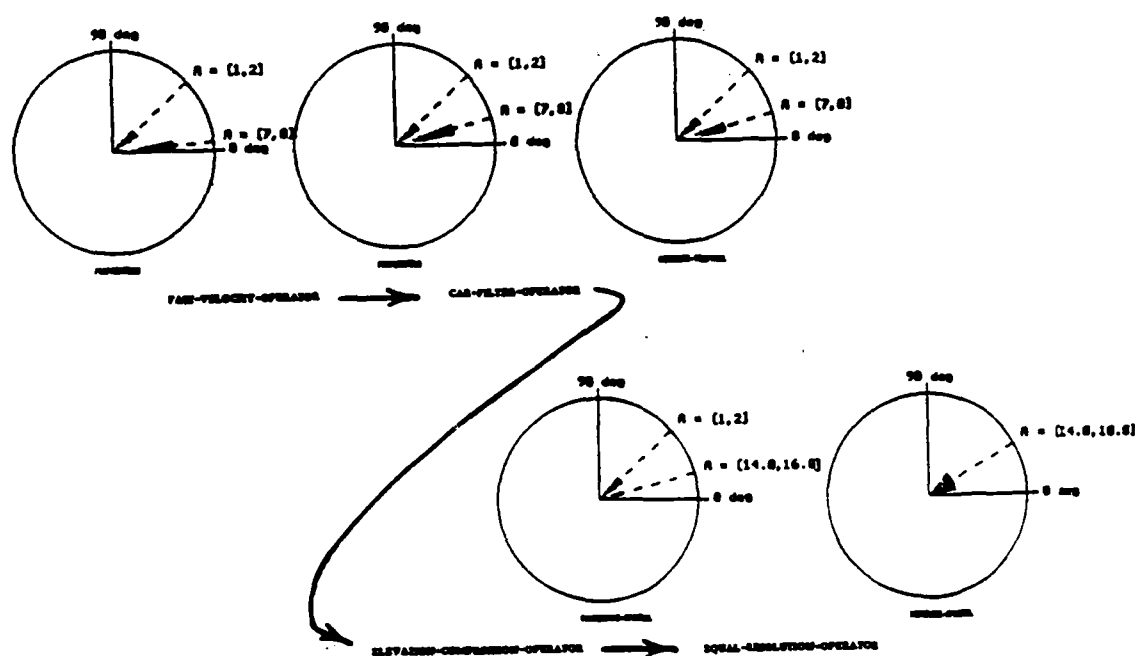
Figure 6.5: Qualitative explanation at the band level.

6.4 Discussion

In this chapter we presented a strategy for diagnosis of faults in systems that carry out signal transformations based on the underlying theory of an application domain. The strategy was described in the context of an automated diagnosis system

186

for a signal processing application. The approach was motivated by the following observations about the signal processing application:

- If we have some a priori knowledge regarding the correct answer to the problem being solved, diagnosis can be carried out using only Fourier domain representations for describing the signal processing. This is in contrast to the actual signal processing algorithms, many of which are designed to carry out operations on time-domain signals. Change of representation into the frequency domain is desirable because it facilitates use of conceptual abstractions available in the underlying Fourier theory of the signal processing system.

- We can use abstract models of the underlying Fourier theory for the signal processing to form a conceptual view of the system as a collection of interacting mappings that explain changes in the Fourier representation of the correct answer leading to the fault to be diagnosed.

- The incorrectly functioning signal processing system can be viewed as having one or more of its mappings acting as non-identity transformations. The diagnosis task then becomes a search for a sequence of such non-identity mappings.

- The number of possible mappings and the associated input and output states required to represent a particular diagnosis scenario in terms of the underlying Fourier theory is much smaller than the number of intermediate data states in the actual signal processing system.

Based on these observations, the diagnosis system has two major components, the qualitative representation of states and mappings between them and a search strategy.

The qualitative representation of states consists of multiple levels of abstraction, with descriptions in each level represented as numerical ranges of values. Matching two states is performed by comparing the numerical ranges of the corresponding values. Two states are considered "equal" if such ranges overlap. Range overlap implies that it is possible for the associated values to be the same. This is considered sufficient for the purposes of diagnosis, where a possible explanation is also an acceptable one.

187

A mapping between states is characterized by state preconditions, which must be satisfied by its input state, in the sense that there is overlap between the input state and the preconditions viewed as defining another state. The postconditions of the mapping specify an output state as a function of the input state and the system and scenario preconditions.

The overall search consists of two interleaved activities, explanation and verification. Explanation is formed via a GPS-type search at a specified abstraction level. Verification "executes" a formed explanation at the lowest possible abstraction level, with the purpose of checking whether the explanation is successful in transforming the initial into the final state.

Several mechanisms are used by the diagnosis system as an interface between the explanation and verification stages:

- *An Abstraction heuristic*, which considers explanations at a higher level of abstraction before explanations at lower abstraction levels, because plans generated at higher abstraction levels are generally simpler, in the sense that they include fewer mappings and fewer parameter settings.

- *A Recovery-from-Failure heuristic*, which first attempts to modify an explanation that failed in its verification through a search for a subexplanation to replace the unsuccessful mapping. The subexplanation may be at a lower abstraction level than the original explanation.

- *A Simplest Explanation heuristic*, which gives preference to explanations with the smallest number of mappings.

Although the implemented diagnosis system performs in the case of a signal processing system, the design of which is based on Fourier theory, the approach is more generally applicable to systems that have a formal theory underlying their design. This includes other types of signal processing systems, as well as systems outside the signal processing area. In all cases, sufficient information about the correct answer should be available to allow its representation in terms of the conceptual

188

models as dictated by the underlying theory. Such models are not just less detailed descriptions of the actual algorithms, but instead they are descriptions based on a representation mechanism quite different from the actual system under diagnosis.

Chapter 7

Summary and Future Directions

7.1 Summary

In this thesis, we advanced the concept of signal abstractions as an organizational principle in real-world signal processing systems.

We defined a set of signal abstractions for harmonic spectra, the extended spectrum, and we examined the issues involved in computing the extended spectrum corresponding to a given numeric spectrum. Such computation was shown to be made more robust by a combination of straightforward algorithmic procedures with heuristic criteria for classification and pruning. Parameters associated with this computation are packaged together into parameter objects that form a complete characterization of the extended spectrum.

The extended spectrum was shown to be useful in a variety of problems associated with harmonic spectra. Adjustment of spectral estimation parameters to maximize the peakiness of a particular harmonic set of the resulting spectrum was performed by matching of extended spectra at multiple levels of abstraction. The outcome of matching was a symbolic indication of the direction, in which a param-

eter change improves peakiness of the resulting spectrum.

Extended spectra allowed explicit representation of the evolution of spectra over time, leading to a helicopter pitch and power tracking system, which uses "islands of certainty", i.e. portions of tracks formed by visually prominent harmonic sets, as a guide in searching for less distinguished harmonic sets to complete and extend the tracks. The pitch and power tracking system employs feedback, whereby harmonic sets found by focused search become part of the current tracks, which are then further extended and completed at the next iteration of the feedback loop.

A second major application of signal abstractions in describing complex signal processing systems by sequences of mappings between signal abstractions derived from the underlying theory was presented in the context of a direction determination system. Mappings between wavenumber spectrum abstractions were used as the basis for a diagnosis system using a GPS-type search for producing and verifying explanations of discrepancies between the anticipated and actual output of the direction determination system. The explanations produced by the diagnosis system readily associate the discrepancies with parameter settings of the signal processing system, which can then be adjusted to reduce or eliminate the fault.

7.2   Future Directions

A direction that needs further exploration is that of generating highly adaptive signal processing systems through diagnosis and qualitative simulation. After diagnosis has indicated which parameter settings of a standard signal processing system are responsible for its fault, qualitative simulation of the underlying theory can possibly be used to automatically adjust the associated parameters. Qualitative models of the underlying theory may be used to represent the effects of parameter changes on the behavior of the corresponding system components [Bobrow 1985].
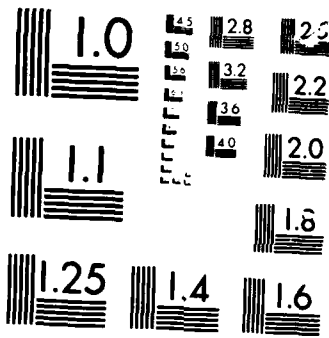
The approach followed in this thesis for formulating signal abstractions was that of observing human subjects perform signal processing in a narrow domain. In a standard signal processing course, however, prototypical signal abstractions are being taught, together with their properties, to enable students to develop an intuitive feeling for signal processing concepts. A possible research direction is the development of a formal theory of generic signal abstractions and the associated system models, along the lines of similar projects in the area of Artificial Intelligence [Bobrow 1985].

There is a need for the application of the approach of this thesis to other real-world signal processing problems with an aim towards systems that perform well and are robust in real tasks. Such systems will probably rely on a large body of domain-specific empirical knowledge combined with formal signal processing algorithms. They will also be "open-ended", in the sense that their performance will not be limited by the formal algorithms they use, but by the amount of empirical knowledge they incorporate. We feel that such signal processing systems can approximate the performance of experienced humans in interactive processing of real signal data.

Knowledge acquisition was performed by taking protocols of the problem-solving activity of human subjects and analyzing it. A combination of training techniques similar to those used in this thesis, of automatic knowledge acquisition tools as in [Davis & Lenat 1982] and of sophisticated signal processing workstations [Dove 1984] may lead to a more appropriate environment for knowledge acquisition. Furthermore, signal abstractions necessitate the existence of powerful user interfaces of the style used in this thesis for inspecting the wealth of information stored in the form of signal abstractions, for interpreting the results and for "debugging" the knowledge in the corresponding signal processing systems.

A more specific project could be undertaken along the lines of using diagnosis to adjust the parameters of the direction determination system without knowledge of the input scenario. Instead, as "reference input" one could use a scenario interpreta-

192

tion obtained using a different kind of processing, such as source localization from geographically separated microphones, or even a wavenumber spectrum obtained with a different parameter setup.

A tradeoff is present in signal processing between computation time and memory usage. With computer memory having been expensive, the trend has been towards signal processing algorithms that treat signals as streams of numbers that flow through the signal processing system, which has limited "information capacity". On the contrary, systems such as those presented in this thesis make intensive use of memory, because of memoization of multiple levels of signal abstractions. Recovery of unusable memory ("garbage collection") is a slow process, and may introduce unacceptable delays in real-time signal processing systems. However, it may be possible to facilitate recovery of unusable memory by introducing more structure in the allocation of memory for signal abstractions. This is an interesting future research direction.

A problem with highly nonlinear threshold-based systems of the sort presented in this thesis is their discontinuous behavior. By this we mean that a small change in the value of a parameter (threshold) may introduce a qualitative change in the behavior of the system. A better understanding of the issues involved in the quantization of information introduced by the existence of thresholds is needed.

Finally, assessing the performance of such systems is a difficult task. This is partly due to the lack of simplified models, such as white gaussian noise, that make performance evaluation mathematically tractable in traditional signal processing. Our frame of reference in this thesis has been the performance of humans in the same tasks. Although our systems have been designed to perform on real data, which makes it inappropriate to try to measure their performance on idealized signals, it may be possible to design somewhat more objective criteria than the performance of humans. As we develop experience with such systems, we may develop systematic procedures for analyzing the validity and usefulness of specific rules of the systems

and possibly automate them.

Appendix A

A raw protocol transcript

(E: Experimenter, S: Subject. The experimenter issues signal processing commands to a signal processing workstation with graphics capabilities. The subject orders the experimenter to issue signal processing commands c. his liking and views the resulting signals. The following is a transcript of the tape recording all the verbalizations of the subject and the experimenter. References are made to signal graphs inspected and commented by the subject. All such signal graphs are provided after the protocol transcript.)

E: You are provided with 15 seconds of acoustic data from 8 channels and your task is to inspect the data for presence of helicopters and infer as much as you can about their motion. I can compute and display the acoustic waveforms at any desired scale and spectra computed as averaged periodograms with variables being the starting time, the window size of a single periodogram, overlap between successive windows and the total number of periodograms to be averaged.

S1: Can I see the whole data from all the channels?

Plot W-1.

S2: OK. It looks like the general power level is fairly constant over time and similar for all channels except channel 1, which is significantly stronger.

S3: How many seconds is this? 20?

E: No, 15 s.

S4: I can see a consistent low of energy around the 5th second in all channels. Other than that, there are minor energy fluctuations here and there in every channel.

S5: Let me now look at the first 2 seconds of each of the channels.

Plot W-2.

S6: It looks like helicopter data. I will try to measure helicopter periodicity.

S7: Hm. Periodicity is hard to find. It is sort of there, but not very clearly visible in the waveform.

S8: It seems that the helicopter is either far before CPA or it is going away from the node. Signal-to-noise ratio is definitely no good in the first 2 seconds.

S9: Let me now see the last 2 seconds of each of the channels.

Plot W-3.

S10: Once again periodicity is not clear. What is the maximum amplitude (of the last 2 seconds of all channels)?

E: 2.6

S11: And what is the amplitude in the first 2 seconds?

E: 2.1

S12: This is an indication that signal energy is lower in the first 2 seconds, though a significant difference does not exist.

S13: Can I now see the 2 seconds of time waveform starting at the 4th second, where I previously observed a slight low in energy?

Plot W-4.

S14: The amplitude there is at the 2.2 level.

S15: Just for completeness, let us check the 2 seconds starting at the 8th second of data.

Plot W-5.

S16: As in all previous cases the signal-to-noise ratio is low and I cannot detect a nice periodicity.

S17: Amplitude is 2.3.

S18: Overall signal energy is increasing with time from 2.1 to 2.6. A plausible explanation is that we have a helicopter coming toward the node.

S19: Now I would like to take a look at spectral characteristics of the data. Let me see the spectrum of the first 2 seconds of channel 1 with

window size 0.5 s (1024 points), 50% overlap and a total of 8 blocks, in logarithmic scale.

Plot S-1-0.

S20: I see quite a number of distinct peaks in both the low and high frequency ranges.

S21: Let me find out the location of the tail rotor frequencies first.

S22: A strong fundamental is around 93 Hz. Another peak around 180 Hz and 270 Hz. Hm, there are rather two peaks at 180 and 270. That's interesting!

S23: Let me be more careful about these double peaks. There are peaks at 183 and 193 Hz, at 275 and 286 Hz, at 368 and 384 Hz, at 460 and 475 Hz.

S24: If we adopt the hypothesis that these peaks belong to two different helicopters, we have here two consistent sets of tail rotor harmonics with fundamentals around 92 and 96 Hz.

S25: So I think my current hypothesis is that there are two helicopters with tail rotor frequencies 92 and 96 Hz.

S26: Assuming that both helicopters are of the same type, the one with 96 Hz comes toward the sensors, while the one with 90 Hz is going away.

S27: Let me also add that the fact that energy increases with time is consistent with the fact that the incoming helicopter is dominant, while the outgoing helicopter is so far that it makes no significant contribution to the total energy.

S28: Let me now see the first 100 Hz of the previous spectrum in linear scale.

S29: What I hope to see here is further evidence of the two helicopters.

Plot S-1-1.

S30: Oh! I don't know why but I didn't really expect such a spectrum. A really dominant peak at 14 Hz. A number of smaller peaks at frequencies around 29, 44, 58 and 71 Hz. A large peak around 92 Hz.

S31: Let's see. The smaller peaks seem to be the second, third, fourth and fifth harmonic of 14 Hz.

S32: The large peak at 92 Hz must be the tail rotor fundamental of the outgoing helicopter. There is no peak at 96 Hz, so the tail rotor fundamental of the outgoing helicopter is dominant at the moment.

S33: Therefore, the peak at 14 Hz probably corresponds to the outgoing helicopter.

S34: By the way, what is the frequency resolution of this spectrum?

E: 1 Hz per bin.

S35: OK, it has pretty good resolution.

S36: Let me now see the same kind of log spectrum for the last 2 seconds of channel 1.

Plot S-1-2.

S37: I see a very noisy spectrum here.

S38: What's the energy level? (compares with spectrum of the first 2 seconds). It looks about the same as that of the first 2 seconds.

S39: Let's now look at the first 100 Hz of the spectrum in linear scale.

Plot S-1-3.

S40: We have a strong peak at around 94-95 Hz. Seems to be consistent with the tail rotor of the incoming helicopter.

S41: Still peaks at 14, 26, 29, 38-39 Hz. So a periodicity is present of the order of 13 Hz.

S42: I haven't been able to tell much from the linear spectrum. It has been rather confusing. I rely more on the tail rotor.

S43: Let me now look at a spectrum somewhere in between, say starting at the 4th second.

Plot S-2-0 (log), plot S-2-1 (linear).

S44: This log spectrum is not as clean as the first 2 seconds and not as dirty as the last 2 seconds. Peaks at 180 and 190 Hz are consistent with two helicopters.

S45: The linear spectrum is very interesting. The two peaks at 90 and 95 Hz are still consistent with the two helicopter hypothesis.

S46: How does the strength of the outgoing helicopter change with time?

E: It has become smaller.

S47: Basically I cannot read too much from the lower end of the linear spectrum because of peak overlap. The peak at 14 Hz appears to be a combination of peaks from both helicopters, which add up to one peak because of resolution problems.

S48: At this time I think that the outgoing helicopter is still the dominant one.

S49: Let me now look at the linear spectrum starting at the 8th second.

Plot S-2-2.

S50: I see peaks at 86 and 95 Hz. Other peaks at 8 and 14 Hz.

S51: The strength of the 14 Hz peak is about the same as before.

S52: I would associate the 14 Hz peak with the incoming helicopter and the 8 Hz peak with the outgoing helicopter. The rest of the peaks are harmonics of 14 Hz. I don't see any harmonics of the 8 Hz peak.

S53: This spectrum seems to make things a little bit clearer. The peaks at 40 and 53 Hz seem to be harmonics of 13 Hz rather than 14 Hz.

S54: The 8 Hz peak was not visible at previous times because low frequencies are very noisy.

S55: Just to confirm all this, let's look at the spectrum starting at the 10th second.

Plot S-2-3.

S56: We see a peak at around 13-14 Hz, with the same magnitude as before and several of its harmonics.

S57: The tail rotor fundamental of the incoming helicopter at 95 Hz is there, while the fundamental of the outgoing helicopter is dying away.

S58: My conclusion from this data is that there is an outgoing helicopter, which was initially close. Its energy goes down very quickly. A second helicopter is coming toward the sensors but it has come nowhere near CPA bacause its fundamental is constant around 13 Hz. Most of my conclusion has been guided by the tail rotor frequencies and I was trying to confirm that with the main rotor frequencies. I feel I have found sufficient evidence for my conclusion.

W-1

W-2

(Pressure of Channel 1)
2.1
-3.2
24600   TIME in units of 0.5 ms   20590

(Pressure of Channel 2)
2.1
-3.2
24600   TIME in units of 0.5 ms   20590

(Pressure of Channel 3)
2.1
-3.2
24600   TIME in units of 0.5 ms   20590

(Pressure of Channel 4)
2.1
-3.2
24600   TIME in units of 0.5 ms   20590

(Pressure of Channel 5)
2.1
-3.2
24600   TIME in units of 0.5 ms   20590

(Pressure of Channel 6)
2.1
-3.2
24600   TIME in units of 0.5 ms   20590

(Pressure of Channel 7)
2.1
-3.2
24600   TIME in units of 0.5 ms   20590

(Pressure of Channel 8)
2.1
-3.2
24600   TIME in units of 0.5 ms   20590

W-3

(Pressure of Channel 1)
2.1

-3.2
8000   TIME in units of 0.5 ms   11998

(Pressure of Channel 2)
2.1

-3.2
8000   TIME in units of 0.5 ms   11998

(Pressure of Channel 3)
2.1

-3.2
8000   TIME in units of 0.5 ms   11998

(Pressure of Channel 4)
2.1

-3.2
8000   TIME in units of 0.5 ms   11998

(Pressure of Channel 5)
2.1

-3.2
8000   TIME in units of 0.5 ms   11998

(Pressure of Channel 6)
2.1

-3.2
8000   TIME in units of 0.5 ms   11998

(Pressure of Channel 7)
2.1

-3.2
8000   TIME in units of 0.5 ms   11998

(Pressure of Channel 8)
2.1

-3.2
8000   TIME in units of 0.5 ms   11998

W-4

W-5

S-1

POWER (log) 5.19 ... -0.514 ... FREQUENCY (Hz) 0 1022 — **0**

POWER (linear) 154000.0 ... 0.0 ... FREQUENCY (Hz) 3 98 — **1**

POWER (linear) 140000.0 ... 0.0 ... FREQUENCY (Hz) 3 98 — **2**

POWER (linear) 150000.0 ... 0.0 ... FREQUENCY (Hz) 3 98 — **3**

S-2

Appendix B

In this appendix we show the intermediate results from the iterations of the system for power and fundamental frequency tracking presented in Chapter 4 of the thesis. In Chapter 4 we only showed the power and fundamental frequency traces obtained after the final iteration of the system. The format of the figures is as follows:

- Part (a) shows the microphone signal (pressure as a function of time).

- Part (b) shows the harmonic chains obtained without pruning of the harmonic sets in a frequency versus time representation.

- Parts (c) and (e) show the harmonic chains resulting after 0 and 1 iteration of the system of Figure 4.1 in a frequency versus time representation.

- Parts (d) and (f) show the identified chain links (as thick filled circles connected with a dotted line segment) and the identified harmonic sets for confirming chain links or extending existing chains (shown as "x"'s) in a frequency versus time representation.

- Parts (g) and (h) show the final harmonic chains (including chain pruning) in a frequency versus time and power versus time representation respectively.

- The figure mentioned in the figure caption is the corresponding figure in Chapter 4.
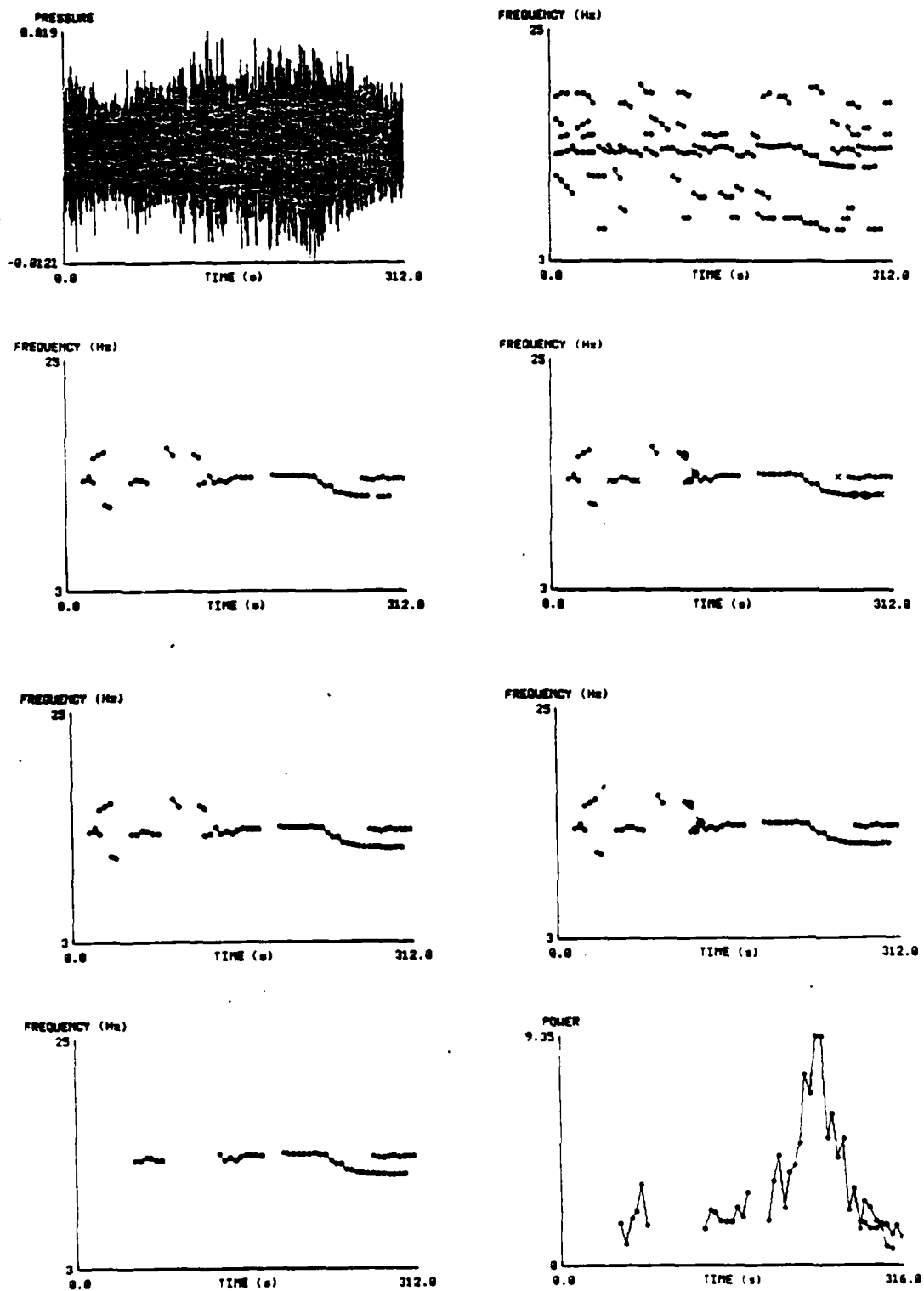
Figure B.1: A single helicopter scenario (Figure 4.11).
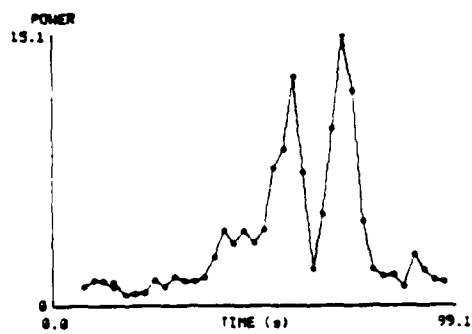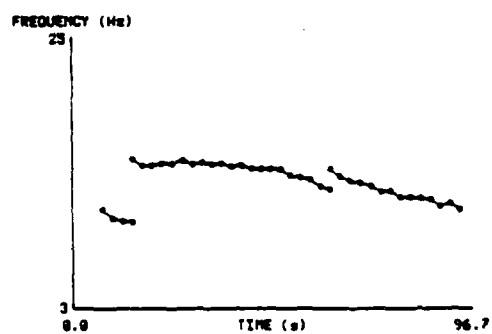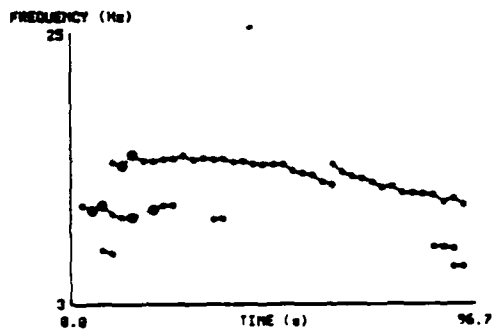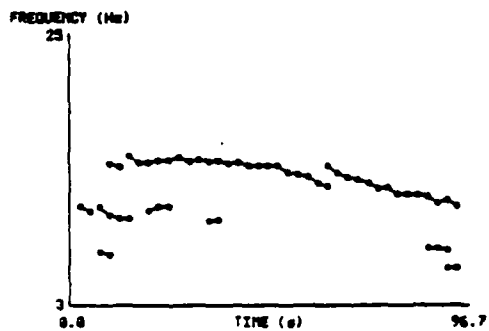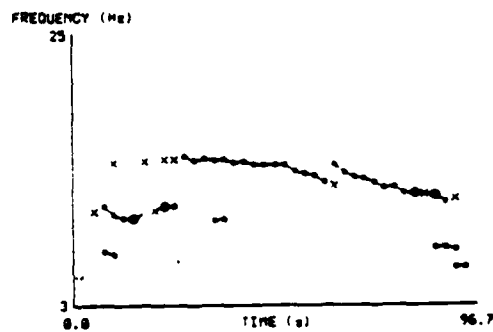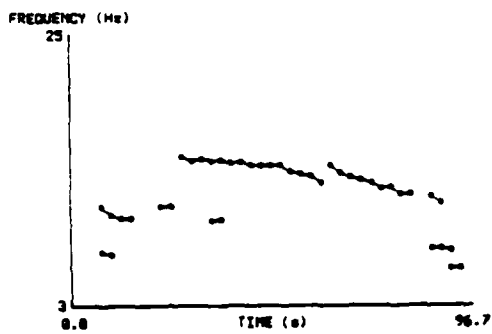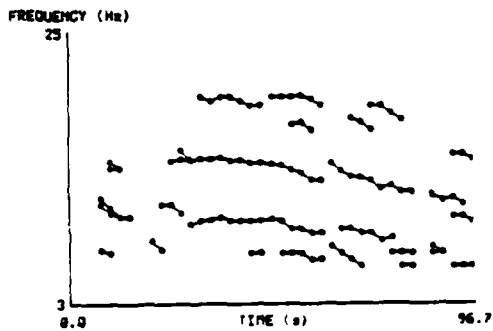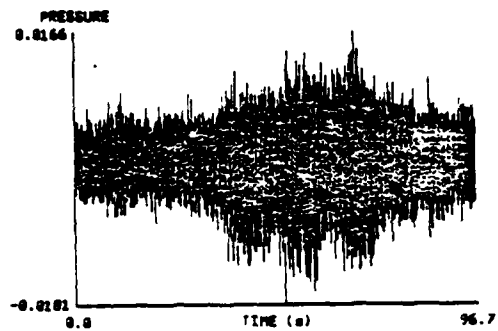
Figure B.2: A two-helicopter scenario, with weak signals and the two helicopters 11s apart (Figure 4.12).
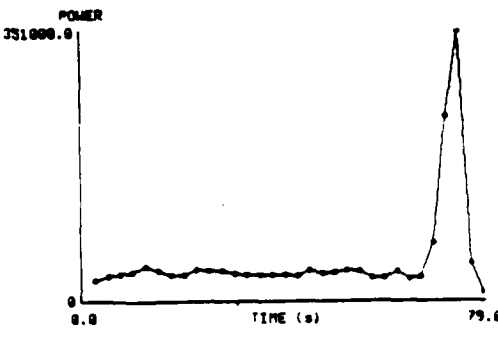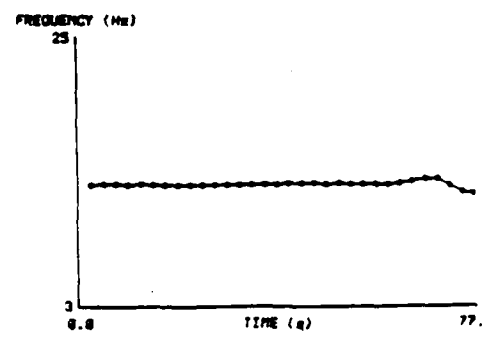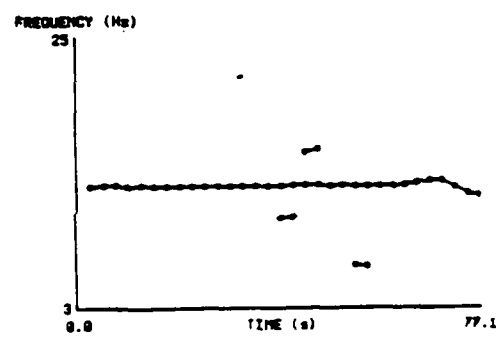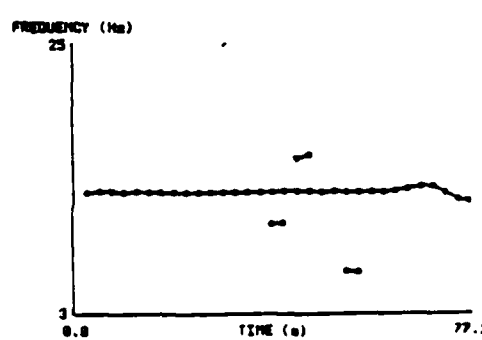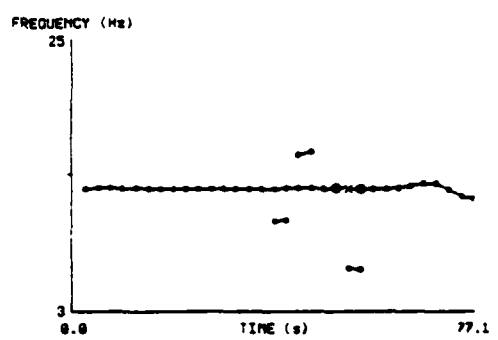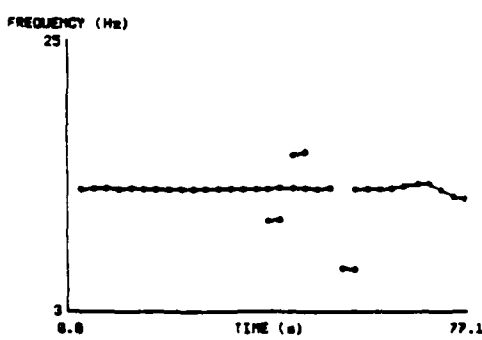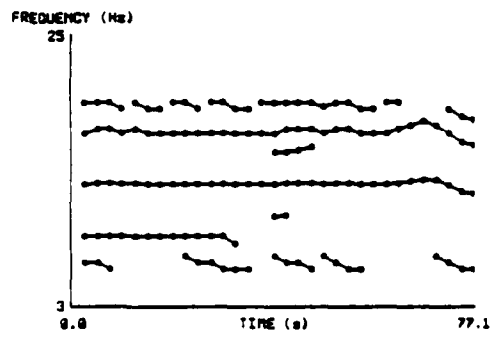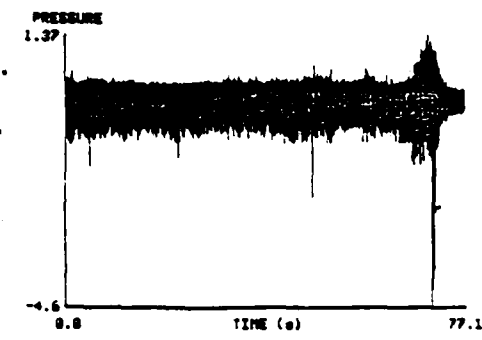
Figure B.3: A single helicopter scenario with a different type of helicopter. (Figure 4.13).
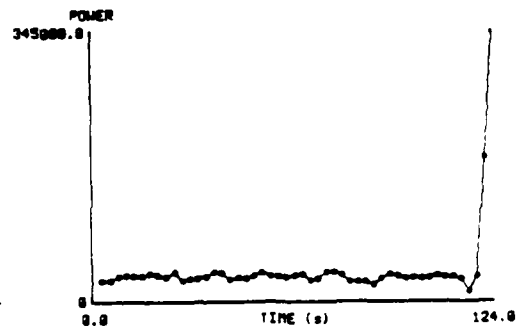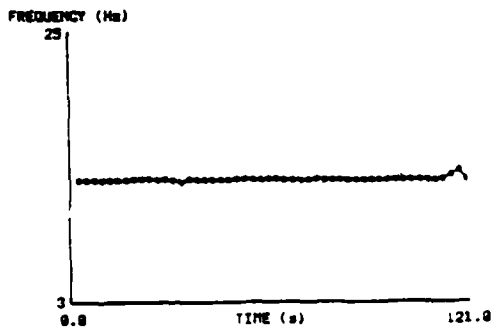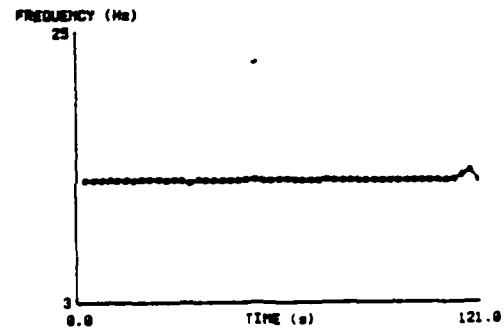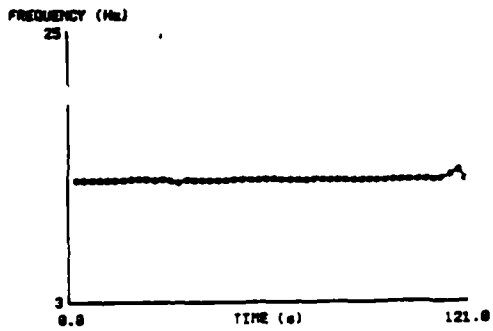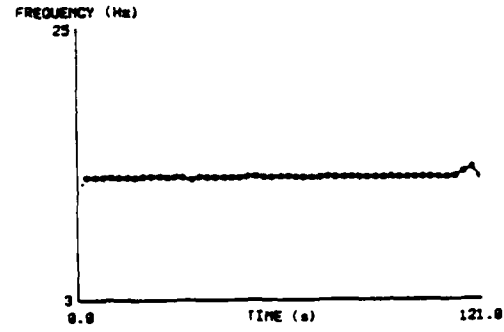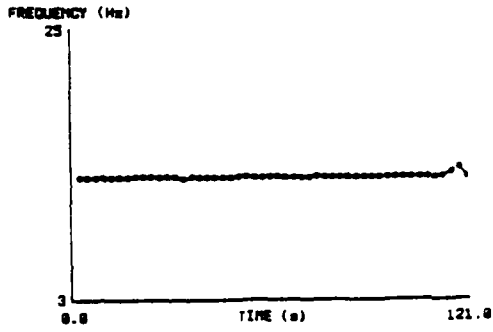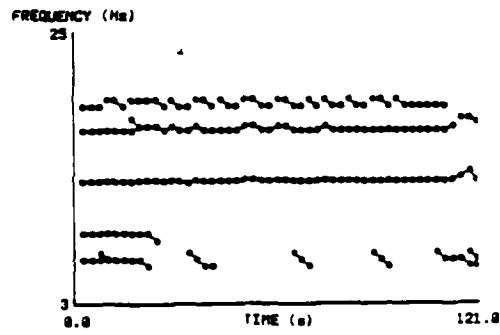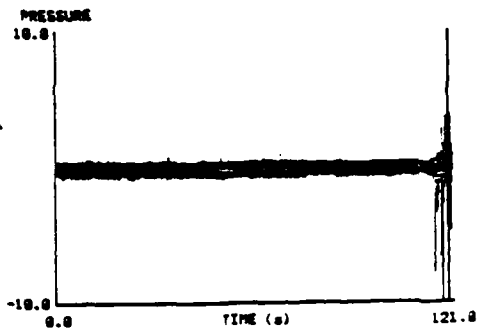
Figure B.4: A single helicopter scenario with the same type of helicopter as in the previous Figure and with the data window ending at CPA (Figure 4.14).
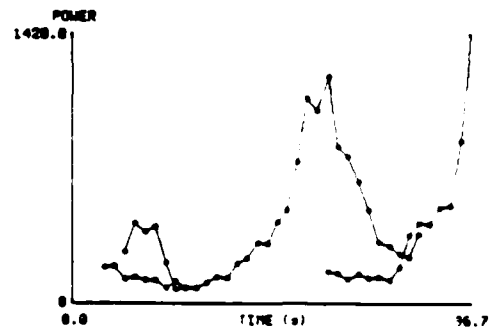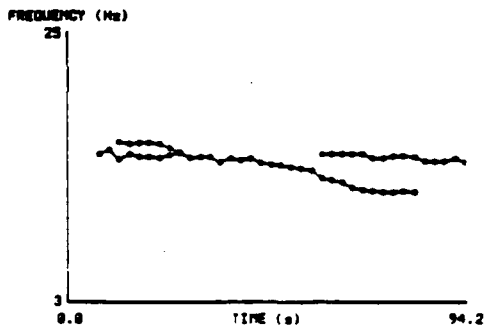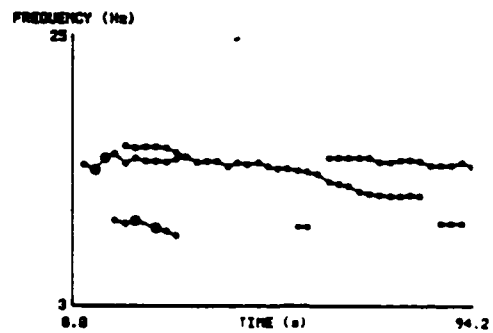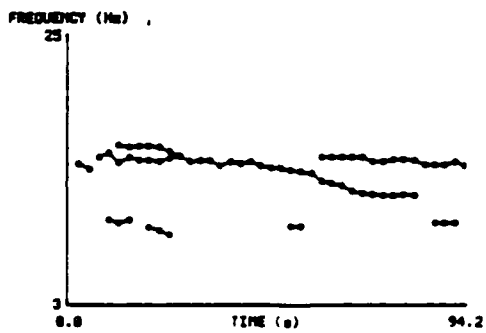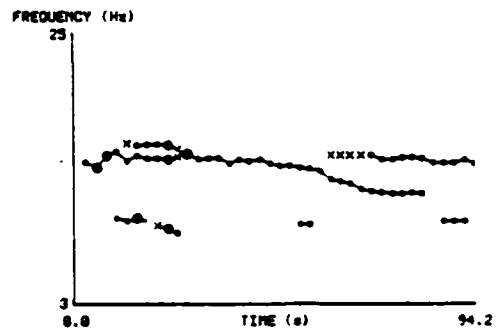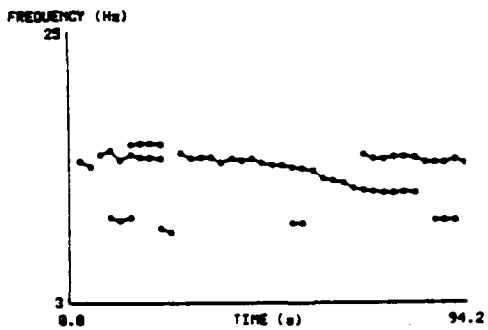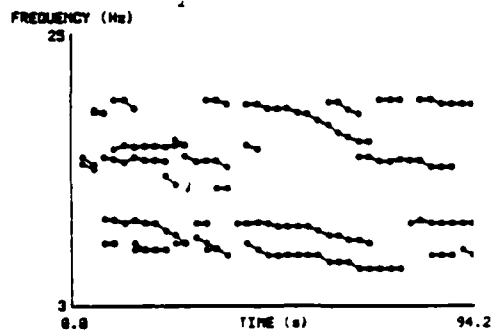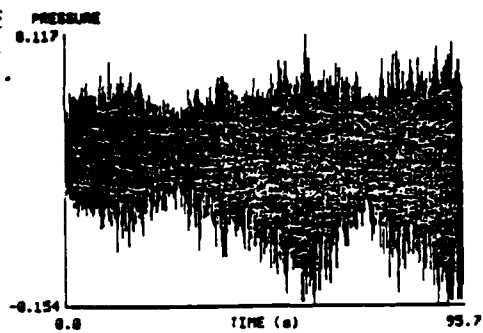
Figure B.5: A two-helicopter scenario (Figure 4.15).
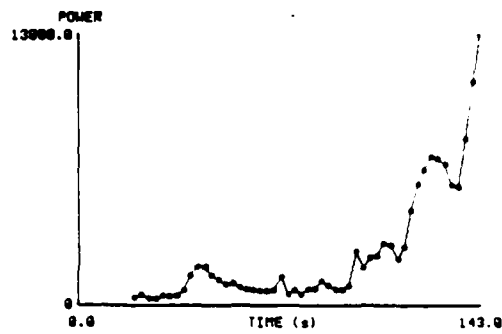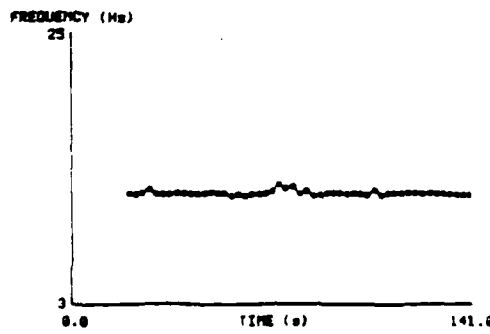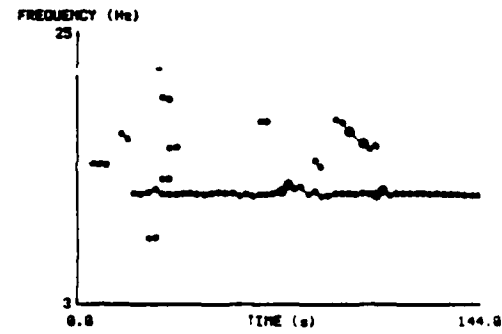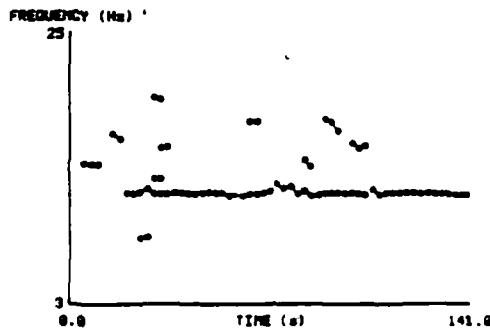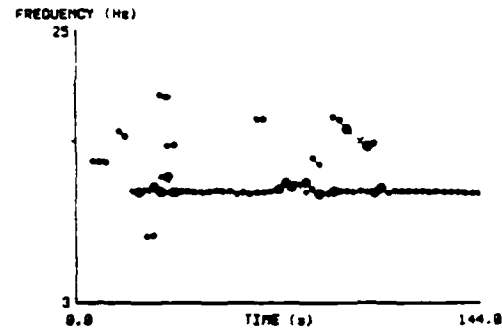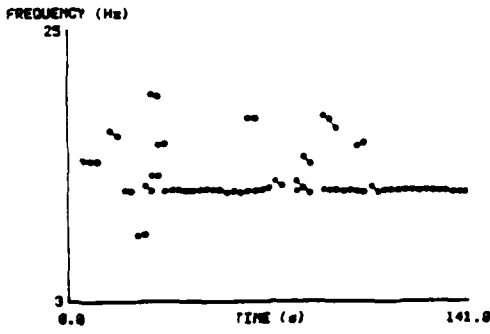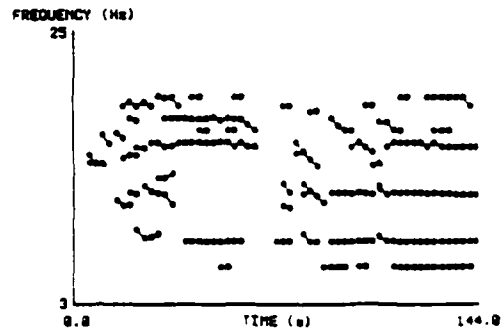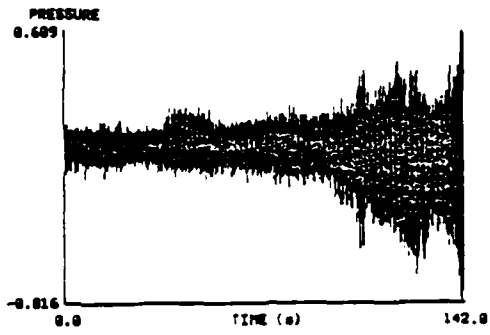
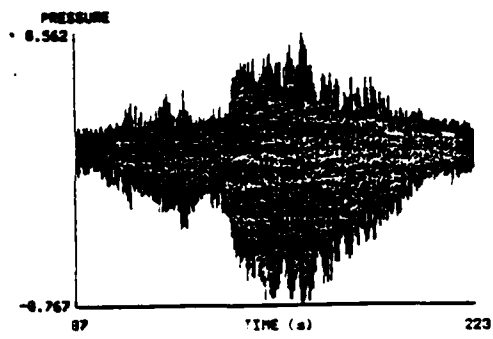Figure B.6: Another single helicopter scenario, portion before CPA (Figure 4.16).
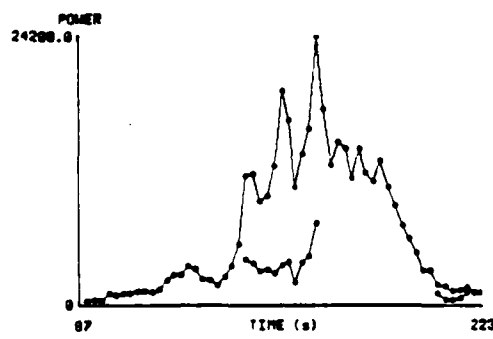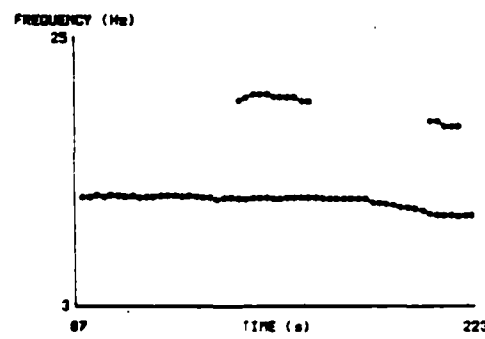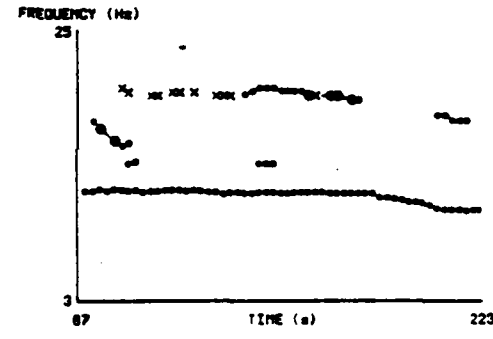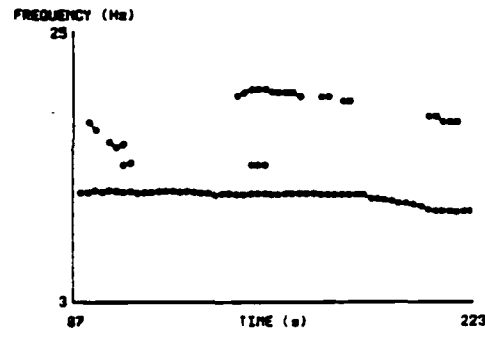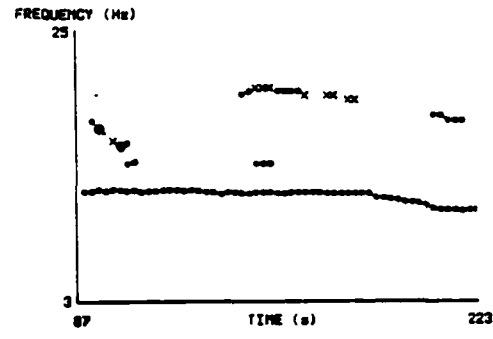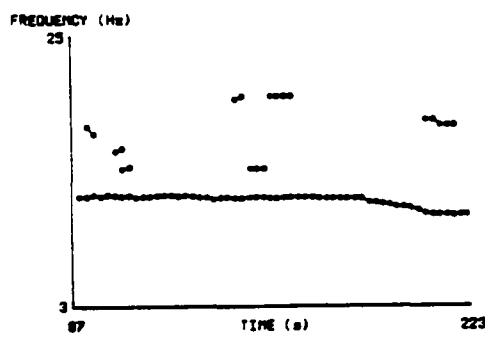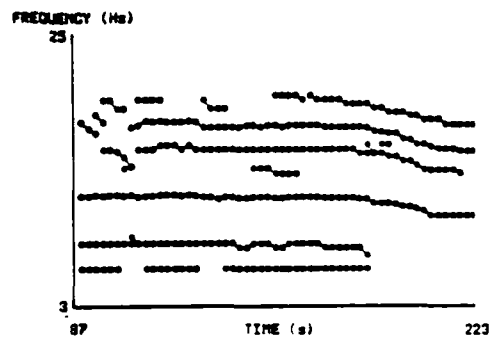
Figure B.7: A single helicopter scenario, portion including CPA (Figure 4.17).

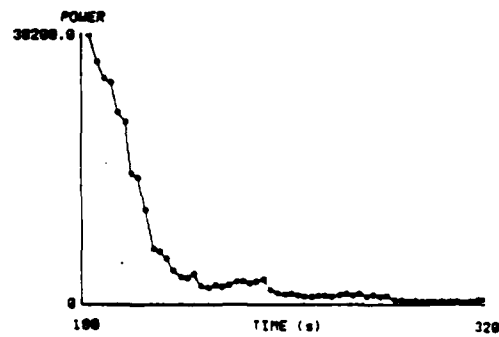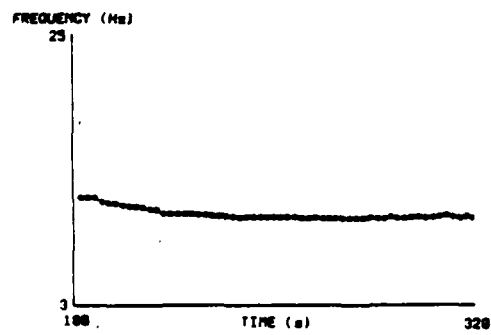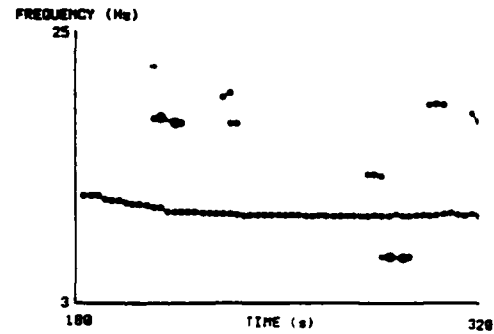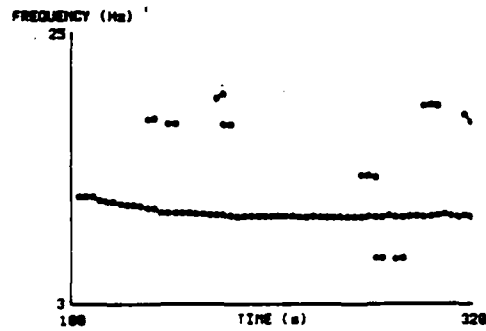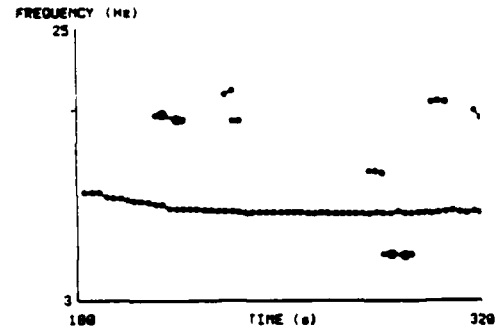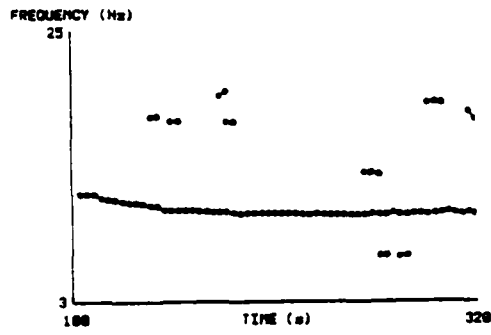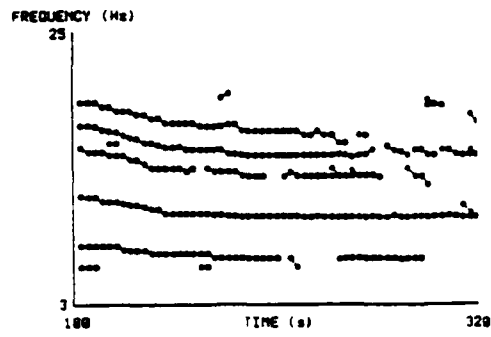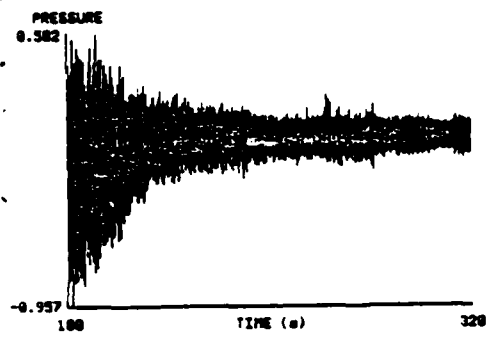Figure B.8: A single helicopter scenario, portion starting slightly before CPA (Figure 4.18).

215

References

[1] H. Abelson and G. Sussman. *Structure and Interpretation of Computer Programs*. MIT Press, Cambridge, Mass, 1985.

[2] E. Allen. YAPS: yet another production system. In *AAAI Proceedings*, pages 5–7, 1983.

[3] K. Anderson. Syntactic analysis of seismic waveforms using augmented transition network grammars. *Geoexploration*, 20:161–182, 1982.

[4] K. Anderson and J. Gaby. Dynamic waveform matching. *Information Sciences*, 31:221–242, 1983.

[5] C. Apte and S. Weiss. An approach to expert control of interactive software systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7:586–591, September 1985.

[6] A. Barr and E. Feigenbaum. *The Handbook of Artificial Intelligence*. W. Kaufmann, Los Altos, CA, 1981.

[7] D. Bobrow, editor. *Qualitative Reasoning about Physical Systems*. MIT Press, Cambridge, Mass, 1985.

[8] D. Bobrow and T. Winograd. An overview of KRL, a knowledge representation language. *Cognitive Science*, 1:3–46, 1977.

[9] Y. Cheng and S. Lu. Waveform correlation by tree matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7:299–305, May 1985.

[10] R. Davis. Diagnostic reasoning based on structure and behavior. In D. Bobrow, editor, *Qualitative Reasoning about Physical Systems*, pages 347–410, MIT Press, 1985.

[11] R. Davis. *Expert Systems: Where are we? and where do we go from here?* Technical Report AI Memo 665, MIT AI Laboratory, 1982.

[12] R. Davis and D. Lenat. *Knowledge-Based Systems in Artificial Intelligence (Part II)*. McGraw Hill, New York, 1982.

[13] W. Dove. *Knowledge-Based Pitch Detection*. PhD thesis, MIT, 1986.

[14] W. Dove, C. Myers, and E. Milios. *An Object-oriented Signal Processing Environment: the Knowledge-Based Signal Processing Package*. Technical Report 502, MIT Research Lab of Electronics, 1984.

[15] W. Dove, A. Oppenheim, R. Davis, and G. Kopec. Knowledge-based pitch detection. In *International Conference on Acoustics, Speech and Signal Processing*, Boston, MA, 1983.

[16] R. Ehrich and J. Foith. Represenation of random waveforms by relational trees. *IEEE Trans. Computers*, 25:725–736, July 1986.

[17] L. Erman, F. Hayes-Roth, V. Lesser, and R. Reddy. The Hearsay-II speech understanding system: integrating knowledge to resolve uncertainty. *Computing Surveys*, 12, June 1980.

[18] F. Hayes-Roth, D. Waterman, and D. Lenat. *Building Expert Systems*. Addison-Wesley, Reading, MA, 1983.

[19] J. Hubbard and W. Harris. Model helicopter rotor impulsive noise. *J. Sound Vib.*, 78(3):425–437, 1981.

[20] E. Hudlicka and V. Lesser. Meta-level control through fault detection and diagnosis. In *Proceedings of AAAI*, pages 153–161, 1984.

[21] S. Kay and S. Marple. Spectrum analysis - a modern perspective. *Proceedings of the IEEE*, 69:1380–1419, November 1981.

[22] L. Kinsler and A. Frey. *Fundamentals of Acoustics*. John Wiley and Sons, New York, 1962.

[23] G. Kopec. The integrated signal processing system isp. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 32-4:842–851, August 1984.

[24] G. Kopec. *The Representation of Discrete-Time Signals and Systems in Programs*. PhD thesis, MIT, 1980.

[25] B. Kuipers. The limits of qualitative simulation. In *IJCAI Proceedings*, Los Angeles, CA, 1985.

[26] B. Kuipers and J. Kassirer. Causal reasoning in medicine: analysis of a protocol. *Cognitive Science*, 8:363–385, 1984.

[27] R. Lacoss. *Distributed Sensor Networks Semiannual Technical Summary Report*. Technical Report, M.I.T. Lincoln Laboratory, September 1982.

[28] B. Liskov and S. Zilles. Programming with abstract data types. *ACM SIG-PLAN Notices*, 9(4):50–59, 1974.

[29] M. Lowson and J. Ollerhead. A theoretical study of helicopter rotor noise. *J. Sound Vib.*, 9:197–222, 1969.

[30] T. Lozano-Perez. Parsing intensity profiles. *Computer Graphics and Image Processing*, 6:43–60, 1977.

[31] D. Marr. *Vision*. W.H.Freeman and Co, San Francisco, 1982.

[32] D. Martinson, W. Menke, and P. Stoffa. An inverse approach to signal correlation. *Journal of Geophysical Research*, 87:4807–4818, June 1982.

[33] C. Myers. *Numeric and Symbolic Representation and Manipulation of Signals.* PhD thesis, MIT, 1986.

[34] C. Myers, A. Oppenheim, R. Davis, and W. Dove. Knowledge-based speech analysis and enhancement. In *International Conference on Acoustics, Speech and Signal Processing*, San Diego, CA, 1984.

[35] H. Nawab, V. Lesser, and E. Milios. Conceptual diagnosis of an algorithmic signal processing system. In *International Conference on Acoustics, Speech and Signal Processing*, Tokyo, 1986.

[36] H. Nawab, V. Lesser, and E. Milios. *Qualitative simulation in the diagnosis of a signal processing system.* Technical Report, M.I.T. Lincoln Laboratory, 1986.

[37] S. H. Nawab, F. Dowla, and R. Lacoss. Direction determination of wide-band signals. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 33:1114–1122. October 1985.

[38] A. Newell and H. Simon. GPS: a program that simulates human thought. In Feigenbaum and Feldman, editors, *Computers and Thought*, pages 279–293, McGraw Hill, 1963.

[39] P. Nii, E. Feigenbaum, J. Anton, and A. Rockomore. Signal-to-symbol transformation: HASP/SIAP case study. *The AI Magazine*, 23–35, Spring 1982.

[40] D. Norman and S. Draper. *User Centered System Design.* Lawrence Erlbaum Assoc., Hillsdale, NJ, 1986.

[41] A. V. Oppenheim and R. Schafer. *Digital Signal Processing.* Prentice Hall, Englewood Cliffs, New Jersey, 1975.

[42] T. Parsons. Separation of speech from interfering speech by means of harmonic selection. *Journal of the Acoustical Society of America*, 60:911–918, October 1976.

[43] T. Pavlidis and F. Ali. A hierarchical syntactic shape analyzer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1, January 1979.

[44] K. Prazdny. Waveform segmentation and description using edge-preserving smoothing. *Computer Vision, Graphics and Image Processing*, 23:327–333, 1983.

[45] E. Sacerdoti. Planning in a hierarchy of abstraction spaces. *Artificial Intelligence*, 5:115–135, 1974.

[46] P. Sankar and A. Rosenfeld. Hierarchical representation of waveforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1, January 1979.

[47] M. Schroeder. Period histogram and product spectrum: new methods for fundamental frequency measurements. *Journal of the Acoustical Society of America*, 43:829–834, 1968.

[48] P. Vincent, J. Gartner, and G.Attali. An approach to detailed dip determination using correlation by pattern recognition. *Journal of Petroleum Technology*, 232–240, February 1979.

[49] A. Witkin. Scale-space filtering: a new approach to multi-scale description. In *International Conference on Acoustics, Speech and Signal Processing*, page 39A.1.1, 1984.

DISTRIBUTION LIST

	DODAAD Code	
Director Defense Advanced Research Project Agency 1400 Wilson Boulevard Arlington, Virginia 22209 Attn: Program Management	HX1241	(1)
Head Mathematical Sciences Division Office of Naval Research 800 North Quincy Street Arlington, Virginia 22217	N00014	(1)
Administrative Contracting Officer E19-628 Massachusetts Institute of Technology Cambridge, Massachusetts 02139	N66017	(1)
Director Naval Research Laboratory Attn: Code 2627 Washington, D.C. 20375	N00173	(6)
Defense Technical Information Center Bldg 5, Cameron Station Alexandria, Virginia 22314	S47031	(12)
Dr. Judith Daly DARPA / TTO 1400 Wilson Boulevard Arlington, Virginia 22209		(1)

END

DTIC

7 — 86