Combining Approximate Front End Signal Processing with Selective Reprocessing in Auditory Perception

Frank Klassner **Victor Lesser** **Hamid Nawab**[†]

Computer Science Department
University of Massachusetts
Amherst, MA 01003
{klassner lesser}@cs.umass.edu

[†]**ECE Department**
Boston University
Boston, MA 02215
hamid@bu.edu

Abstract

When dealing with signals from complex environments, where multiple time-dependent signal signatures can interfere with each other in stochastically unpredictable ways, traditional perceptual systems tend to fall back on a strategy of always performing finely-detailed, costly analysis of the signal with a comprehensive front end set of signal processing algorithms (SPAs), whether or not the current scenario requires the extra detail. Approximate SPAs (ASPAs) – algorithms whose processing time can be limited in order to trade off precision in their outputs for reduced execution time – can play a role in producing adaptive, less-costly front ends, but their outputs tend to require context-dependent analysis for use as evidence in interpretation. This paper examines the IPUS (Integrated Processing and Understanding of Signals) architecture's ability to serve as a support framework for applying ASPAs in interpretation problems. Specifically, our work shows that it is feasible to include an approximate version of the Short-Time Fourier Transform in an IPUS-based sound-understanding testbed.

Introduction[1]

Since the early 1980's, many perceptual architectures have incorporated the basic design shown in Figure 1. This design scheme produces systems with a numeric-oriented front end that is logically separated from a symbolic-oriented interpretation component. The signal processing algorithms (SPAs) in the front end are permitted only one pass over the incoming signal, and the interpretation component is designed with the assumption that the front end's output is always an "adequate" decomposition of the signal. The development of this scheme can be attributed to several factors, including the influence of Marr's reconstructionist school of thought in computer vision (Marr, 1982) and early psychophysical research on human perception which ignored the role of expectations in human interpretation of visual and auditory signals. Both influences led to the view that symbolic interpretation follows and depends upon signal decomposition by the front end through inversion of the physical processes that led to the original signal (Draper, 1993).[2]

Since this design paradigm emphasizes one pass over input data, there is a tendency to build perceptual systems with fixed front ends that are expensive because they must provide detail for the most ambiguous interpretation cases even when that detail is unnecessary. Dorken (Dorken *et al.*, 1992) showed that, as the complexity of an acoustic environment increases (e.g. greater signal signature interference, more signal sources emitting simultaneously, etc.), classic interpretation systems can require front ends with a combinatorially explosive number of fixed SPAs with multiple parameter settings to avoid ambiguous signal-to-symbol mappings, with consequent front end processing time costs. Tsotsos (Tsotsos 1989) also demonstrated this property in the visual domain, through a proof of the NP-completeness of the machine-vision interpretation problem.

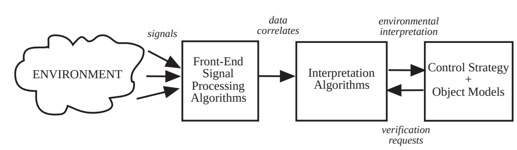


Figure 1: *Classic Signal Interpretation Architecture.*

The *Integrated Processing and Understanding of Signals* (IPUS) architecture (Klassner, 1996; Lesser *et al.*, 1995) was originally developed as a domain-independent framework for structuring feedback between a blackboard-based perceptual system's front end and interpretation components. By representing (1) reasons for uncertainty in interpretation hypotheses, (2) theoretical relationships among input signal appearance, control parameter values, and SPA outputs'

[2] This is *not* the only view within the machine perception community. There is an alternative view (Ellis, 1996; Draper, 1993; Strat 1991; Kohl *et al.*, 1987) advocating feedback between front end and interpretation components. This paper's research is complementary with this view.

appearance, (3) possibilities where SPAs were applied with parameter values inappropriate to the combination of signal objects in the environment, and (4) the processing contexts under which SPA outputs and interpretations are created, IPUS implements perception as the integration of search in a front-end-SPA space with search in an interpretation space. Uncertainty in one space's current state triggers search in the other space for an explanation and resolution of the ambiguity or discrepancy. Front-end-SPA search involves the (re)application of SPAs with control parameters chosen on the basis of the theory behind their operation, and is performed to find data that eliminates or reduces uncertainty (e.g. missing support for expectations) in the signal's current interpretation.

Because IPUS has the ability to selectively reprocess uncertain portions of a signal with specialized SPAs, the framework can potentially play an additional architectural role: supporting the use of *approximate processing* techniques to reduce the complexity of front ends by sacrificing precision in SPA output.

Approximate processing (Lesser *et al.*, 1988) refers to deliberate limitation of search processes in order to trade off certainty for reduced execution time. Approximate SPAs (ASPAs) are SPAs whose processing time can be limited in order to trade off their outputs' precision for reduced execution time. The availability of such SPAs permits formulation of perceptual control strategies that first use ASPAs to create a rough picture of the environment that is refined only where the front end outputs' interpretations are too uncertain. Refinement involves reprocessing these limited signal portions with SPAs that produce outputs having greater precision. These non-approximate and potentially highly specialized SPAs would be expensive if applied to the entire signal, but if applied only in restricted signal regions their costs become manageable.

This paper examines IPUS' feasibility as a framework for supporting ASPAs within the auditory scene analysis problem (Bregman 1990), which involves the segregation and identification of sounds in an acoustic signal. Specifically, our test application focuses on showing that an IPUS-based sound-understanding testbed (SUT) can use ASPAs to adaptively generate spectrograms that provide sufficient detail in the time-frequency domain for recognizing the sources responsible for generating the input signal.

Two key questions must be addressed in evaluating the suitability of IPUS (or, for that matter, any other framework) for ASPAs: **(1)** whether the framework provides enough structured support for the context-dependent nature of ASPAs' outputs, and **(2)** whether front-end ASPAs' time gains are overwhelmed by interpretation search and reprocessing due to increased uncertainty in the ASPAs' output. This paper considers the first question in **IPUS and ASPAs**, and the second question in **Performance Evaluation**. The paper ends with **Analysis and Conclusions**.

IPUS and ASPAs

The first subsection describes how IPUS supports principled, efficient, selective (re)application of SPAs. The next subsection shows why this benefits ASPAs, using the Quantized Short-Time Fourier Transform (QSTFT) (Nawab and Dorken, 1995) as a specific example. The last subsection gives an abstract trace of an IPUS-based system to unify the section's concepts.

IPUS Architecture

IPUS is instantiated by a domain's formal signal processing theory, and has four components for organizing and applying signal processing theory: discrepancy detection, discrepancy diagnosis, differential diagnosis, and signal reprocessing. (Lesser *et al.*, 1995) These components have the following functionality:

- detect discrepancies between data expectations and actual data observations,
- diagnose these discrepancies and ascribe reasons for observational uncertainty,
- determine reprocessing strategies for uncertain data and expected scenario changes, based on the results of the diagnosis, and
- determine differential diagnosis strategies to disambiguate data with several alternative interpretations.

The architecture follows an iterative process of "discrepancy detection, diagnosis, reprocessing" for converging on the appropriate SPAs and interpretations. Convergence is driven by the goal of eliminating or reducing various categories of interpretation uncertainty.

IPUS implements perception as the integration of search in a front-end-SPA space with search in an interpretation space. Uncertainty in the current state in one space triggers search for an explanation and resolution of the ambiguity or discrepancy in the other space. In general, the search process whose current state produces the lower uncertainty serves as the standard against which progress toward a complete interpretation or adequate front end is measured in the other. Within the interpretation search process "uncertainty" refers to the portion of the signal[3] explained by the current interpretation state and the strength of the negative (i.e. missing or incomplete) evidence against each hypothesis in the interpretation. Within the front-end search process "uncertainty" refers to the degree of inconsistency found among the results from SPAs whose outputs are supposed to be related according to their domain signal processing theory.

Each time an SPA is executed within IPUS, the hypotheses representing the execution's results are annotated with the name of the SPA and the control parameter values used in the execution. This annotation is the outputs' *parameter context.* In addition to

[3] In the SUT, percent of input signal energy accounted for by the current interpretation.

the parameter context, each SPA output is annotated with a *processing context*, or a data structure listing the SPA sequence that generated the hypothesis from the input signal.

Within IPUS, three sets of information (in addition to SPA code) used to define SPAs are important for supporting selective SPA (re)application. The first is a set of rules defining how individual SPA control parameters should be modified to eliminate or reduce various classes of distortions that could be manifested in the algorithm's outputs. The second key definition element is a list of "supercontext methods" that take as input a parameter-context and an optional "information category" label. These methods return context patterns indicating the range of values for each control-parameter in an SPA parameter-context that would permit the SPA to produce correlates having the same or greater detail in the specified "information category" as found in the specified parameter context. As a simple example, assume an SPA that selects local maxima from a spectrogram on the basis of whether their energy values are greater than a threshold control-parameter. The supercontext method for this SPA would, when supplied with a particular parameter context and the information category "peaks," return a pattern indicating that any execution of the SPA with a parameter context having threshold values below that of the given parameter context would provide at least the same number of peaks as were produced by the given parameter context. The third key SPA definition element is a mapping function that takes as input two parameter contexts and the output hypotheses produced from the first context (i.e. execution of an SPA), and returns a list of the hypotheses modified to reflect how they would appear had they been produced by the second context.

Together, processing contexts and SPAs' mapping functions enable IPUS-based systems to examine their reprocessing history for processing contexts that would provide SPA outputs that were at least as detailed as those required by a current reprocessing request. This process is called *context mapping*, and, along with IPUS' diagnostic support, will be shown by the next subsection to be important for efficient ASPA usage.

ASPA Illustration

The discrete Short-Time Fourier Transform (STFT) (Nawab and Quatieri, 1988)

$$X[n,k] \;=\; \sum_{m=-\infty}^{\infty} x[nL+M+m]w[m]e^{-j2\pi mk/N},$$

$$0 < k < N-1, 0 \leq M < L$$

is a common tool for representing the time-dependent frequency content of a discrete signal $x[n]$. Figure 2 shows how its matrix representation can be viewed as a picture of time-dependent frequency tracks indicating the presence of some signal-producing sources.

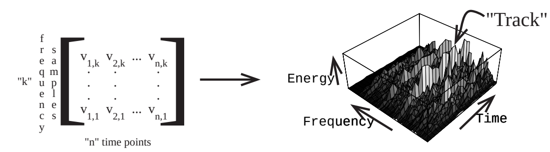


Figure 2: *Abstract "Tracks" in a STFT Spectrogram.*

An STFT instance has particular values for its parameters: analysis window length (number of signal points L analyzed at a time to produce a column in Figure 2's matrix), frequency-sampling rate (number of points N computed per column in Figure 2's matrix), and decimation interval (signal points M between consecutive analysis window positions). For values of N that are powers of 2, there is an efficient algorithm for the STFT based on the Fast Fourier Transform (FFT), which requires $O(n\log n)$ real multiplications and $O(n\log n)$ real additions. Conceptually, the algorithm computes a series of FFTs on successive blocks of L data points in the signal. Though commonly used, this version of the STFT has the drawback of computing values for *all* points within the spectrogram matrix, whether or not they are needed for interpretation.

The Quantized Short-Time Fourier Transform (QSTFT) (Nawab and Dorken, 1995) was developed for approximating a signal's STFT using an order of magnitude fewer additions than the FFT-based STFT and *no* multiplications. This performance is achieved through first quantizing each L-point block in the signal to the set (-1,0,1), backward-differencing the quantized signal, and then applying an O(n^2) implementation of the Fourier Transform (i.e. the summation of the x[]w[]e^j terms) to each block. When compared to STFTs with N's up to 256, evaluation of the basic QSTFT for the entire frequency-domain spectrum takes fewer additions, and no multiplications, due to the increased number of zeros in the modified signal. A band-limited QSTFT that computes only a limited region surrounding the estimated highest-energy frequency generally requires fewer mathematical operations than the complete-spectrogram STFT for N's up to 1024.

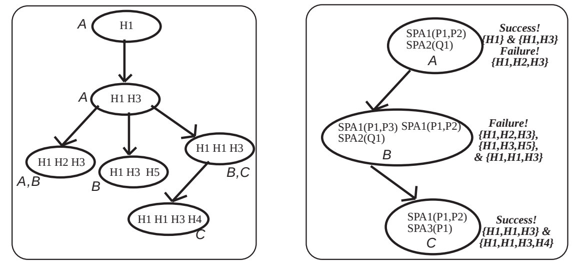
Interpretation of an STFT's (or, any other SPA's) output depends on the appropriateness of the algorithm's parameter values to the current scenario. It can be shown through analysis of Fourier theory that a fixed STFT with a long analysis windows will provide fine frequency resolution for scenarios containing sources with time-invariant frequency tracks, but at the cost of poor time resolution for sources with time-varying components. Conversely, a fixed STFT with short window lengths will provide fine time resolution for scenarios containing sources with time-varying components such as chirps or reverberatory decays, but at the cost of poor frequency resolution for sources with close frequency components. The band-limited QSTFT (and ASPAs in general) suffers from these same context-dependent issues, but its approximate nature introduces even more uncertainty issues. For example, the termination of a track in a QSTFT spectrogram can result not only from the actual signal cessation of a source or STFT-related resolution distortions, but also from the appearance of new sources with frequency content that shifts the estimated maximum-energy frequency so that the computed spectrogram region does not include the track.

Thus, as signal sources change and interact in complex scenarios, perceptual systems must be carefully designed to ensure that these changes can be interpreted at "face value" and are not the result of a distortion introduced by SPA parameter values no longer appropriate to the current context. This issue becomes particularly important for ASPAs because of the greater number of assumptions involved in their design (e.g. frequency region with maximum energy remains stable, for QSTFT). The IPUS framework's diagnostic component and SPA theory representation address this problem by supporting the focused (re)application of front-end SPAs in regions of the signal where SPAs' signal processing theory indicates the possibility of distortions from inappropriate SPA parameter values. These (re)applied SPAs are executed with control parameters chosen to provide more precise data that will confirm or disconfirm the distortion hypothesis. In addition, the IPUS reprocessing component's ability to use context-mapping (e.g. using earlier reprocessing results from a non-approximate STFT instance or an expensive QSTFT instance with very high frequency resolution to eliminate search for newly hypothesized sources with frequency tracks that theoretically should have been present in the earlier results) helps to preserve ASPAs' front-end time gains.

Summary

We conclude with an abstract trace of an IPUS-system execution that unifies the concepts in this section. Figure 3 shows a hypothetical interleaving of progress in the two IPUS search spaces.

The system behavior can be summarized as follows. Initially, the interpretation system uses front end $\mathcal{A}$ with ASPA SPA1 to collect evidence, and hypothesizes that one perceptual object of type H1 is present. Attempting to account for more signal energy, the system then explores the interpretation state $\{\mathbf{H1},\ \mathbf{H3}\}$ and finds that $\mathcal{A}$'s SPA sequence has also already produced evidence to support the interpretation. When attempting to explain the remaining signal energy, the system finds that an additional single object of either type H1, H2, or H5 could be hypothesized. Choosing H2 first (i.e. state $\{\mathbf{H1,H2,H3}\}$), the system's discrepancy detection finds that A does not provide evidence for the H2 instance. When discrepancy diagnosis



A. INTERPRETATION SPACE B. FRONT-END SPACE

Figure 3: *Figure 3A shows a system's progress within its interpretation space. Labels outside each state indicate the front end(s) being used to find support evidence for the state's interpretation (set of object instances Hn). Figure 3B shows the front ends explored by the system. Labels outside each state indicate the interpretation expected for the front end's outputs and whether the outputs actually supported it.*

shows that SPA1's current parameters do not provide enough detail for the interpretation, the system's reprocessing component generates and applies front end $\mathcal{B}$'s SPAs to selected regions of the signal, and this time finds negative evidence for H2's instance, causing it to abandon interpretation state $\{\mathbf{H1,H2,H3}\}$ and proceed to explore $\{\mathbf{H1,H3,H5}\}$. The SPAs in front end $\mathcal{B}$ fortuitously served as a supercontext of a front-end that provided, via context mapping, indisputable negative evidence for any instance of H5, leading the system to explore interpretation state $\{\mathbf{H1,H1,H3}\}$. This time discrepancy detection shows that $\mathcal{B}$ did not provide positive evidence for the second instance of type H1; however, the system's discrepancy diagnosis component finds that $\mathcal{B}$ was inappropriate to the interpretation. The system's reprocessing component then uses SPA theoretic constraints to determine that front end $\mathcal{C}$, by using non-approximate SPA3 in a limited portion of the signal, should be appropriate for supporting or disproving the existence of the extra H1. According to Figure 3B the outputs produced by the new front end do in fact support the second H1 hypothesis, and ultimately support the creation of a final H4 hypothesis. The final interpretation state is $\{\mathbf{H1,H1,H3,H4}\}$, and because it accounts for enough signal energy, interpretation search stops.

Performance Evaluation

The first subsection describes the environment in which IPUS SUTs are evaluated, the next section describes the three SUT versions evaluated, and the final section lists the experimental results.

Evaluation Domain

For the evaluation experiments presented in this paper the SUT was supplied with a sound-model library containing 40 common sound-sources. The sounds were specifically selected to provide a reasonably complex subset of the acoustic behaviors (e.g. impulsive, harmonic, periodic, chirping) and sound interactions (e.g. masking, start/end time blurring, overlapping frequency content) that can arise in random real-world auditory scenarios. (Klassner, 1996) provides a full catalogue of the library sounds and their models. The sounds' durations ranged from 0.2 to 30.0 seconds.

As an indication of the potential for interactions among sounds randomly selected from the library and placed in scenarios with random start times, it should be noted that the expected frequency range of each narrowband track (e.g. $\leq$ 100 Hz wide) of each library sound overlaps a track of at least one other sound. For each track of a given sound, on average another 4.2 sounds have overlapping frequency content. Note that the greater the number of overlapping tracks there can be in a spectral region, the greater the amount of interpretation search that must be done to determine (1) whether in fact overlapping tracks are present in a scenario, and (2) which subset of the tracks that could be in the region of overlap are actually present.

Experiment Design

The goal of the experiments in this paper is to indicate indicate how well ASPA- and non-ASPA-based IPUS SUTs handle complex environments, and with what types of costs for front-end search and interpretation search. For this paper, "complex environments" are those having three or more sounds that share some frequency content occur at overlapping time periods.

The following 5-step method was used to generate 15 complex acoustic scenarios. First, four sounds were randomly selected from the SUT library. Second, a random instance of each sound was selected (each sound had at least 5 instances, sampled at 16 KHz). Third, start-times for each instance were randomly selected with uniform distribution within a 7-second base timeframe. Fourth, a 5-second window was randomly chosen within the base timeframe such that all four sounds were included for at least their length or 1 second, whichever was shorter. When start times precluded such a window, steps 3 and 4 were repeated until this criterion was met. Fifth, each scenario was scaled so that all had the same average power. Since some sound-creation events, such as footsteps and phone ring sequences, are really composed of several instances, the average number of instances in each scenario is 7.3 rather than four, as might have been expected.

Two primary experiment runs were perfomed on the same 15 scenarios. The first (`LoRes`) used a version of the IPUS SUT whose default front end (i.e. the SPAs routinely applied in the first analysis of the signal) contained a non-approximate STFT, with $N = 512$, $L = 256$, and $M = 0$. It was directed to use non-approximate STFTs in limited signal regions during reprocessing. The second (`Appr`) was identical to the first, except that the default front end contained a band-limited QSTFT SPA with $N = 512$, $L = 256$, $M = 0$, and a 1000-Hz band radius. This second version was directed to use band-limited QSTFT's in restricted signal regions during reprocessing. A third reference experiment (`HiRes`) was performed on the 15 scenarios with a SUT having a default front end that contained a non-approximate STFT with $N = 2048$, $L = 1024$, and $M = 128$. All SUT versions modelled STFT-induced time- and frequency- resolution and QSTFT-induced windowing distortions in their diagnostic components, as well as distortions for other front-end SPAs applied to spectrogram results.

	Front End:	Appr	LoRes	HiRes
SPA Search Cost	Param Cntxt	13.70	7.45	3.20
	Total Ops	2.5e6	6.1e6	5.5e7
Interp. Search Cost	Total Hyps	16.48	14.10	8.14
	Answers	0.97	1.07	0.86
	Nonanswers	15.51	12.03	7.28
System Perform.	Hit Rate	0.56	0.61	0.60
	FAlarm Rate	0.42	0.43	0.39
	Track+ Rate	0.67	0.65	0.67
	Track- Rate	0.24	0.22	0.19

Table 1: *Experiment Results. SPA-search cost is averaged per scenario, while interpretation-search cost and system performance are averaged per scenario sound instances. Note that "Total Ops" includes both additions and multiplications performed during both initial front-end analysis and reprocessing.*

Results

Table 1 reports statistics for the SUT versions' performance in three evaluation categories:

SPA-search cost: the number of time-frequency-SPA reprocessing parameter contexts per scenario *(Param Cntxt)* and total number (first-pass + reprocessing) of spectrogram-based mathematical operations (additions and multiplications) per scenario *(Total Ops)* are reported.

Interpretation-search cost: the average number of answer hypotheses (both false alarm and hits), and the average number of considered but rejected sound-source hypotheses are reported.

System Performance: hit rate and false-alarm rate are reported, as well as the duration for which "hit" sound instances were tracked relative to the total amount of time for which all sound instances lasted *(Track+ Rate)* and the duration covered by all false alarm answer hypotheses relative to the total time covered by all answer hypotheses *(Track- Rate)*.

Analysis and Conclusions

With the understanding that the scope of the experiments is limited, two basic conclusions can be drawn from the experiment results:

1. the Appr and LoRes columns of Table 1 indicate that it is reasonable to consider using ASPAs in IPUS-based acoustic interpretation systems to trade off front-end complexity for moderate increases in interpretation search, and
2. the system performance rate in Table 1's HiRes column, when compared with that found in the Appr and LoRes columns, shows that while commitment to detailed initial front-end processing can save interpretation search, it is not guaranteed to outperform ASPA-based strategies, even in complex scenarios.

The first conclusion is based on the observation that the Appr SUT achieved system performance similar to the LoRes SUT, and took only 41% of the front-end mathematical operations that the LoRes SUT required. This is significant given that both systems have the same initial resolving power. Examination of the sounds missed by the Appr SUT shows that the hit rate difference is due solely to sounds whose entire frequency content fell outside the QSTFT's spectrogram window. All SUT versions included a simple threshold-based SPA that tracked time-domain signal energy relative to explained time-frequency energy, and triggered reprocessing when the ratio dropped below a threshold. We believe that a moderately more sophisticated, low-cost SPA could be designed to cause the Appr SUT to increase interpretation search and find the missed sounds. We also note that the Appr's increased number (15% greater) of considered interpretation hypotheses relative to the LoRes's did not adversely affect overall system time because the verification process for many of the extra sounds involved either context resuse or verification of only one narrowband track of a hypothesis that was then disbelieved.

The second conclusion is based on the similar system performance rates between the Appr and LoRes systems as a group and the HiRes system. Note that Appr took only 5% of the total spectrogram-oriented operations required by the HiRes system. The fourfold frequency resolution gain afforded by the HiRes system's default front end did not significantly improve its performance over either of the lower-resolution systems. Since the experiment scenarios represent rather acoustically overloaded environments (i.e. several simultaneous sources having rapid changes in time-frequency characteristics), this result further encourages consideration of ASPAs for use in interpretation problems.

Even though the result of using ASPAs in our example application is promising, it is important to temper our hopes and conclude this paper with the understanding that different perceptual domains and applications can have widely differing balances between front-end and interpretation costs. For example, in some domains a reduction in overall front-end costs by as little as 10% could be very valuable if high-level interpretation costs are small, and there is not significant additional interpretation search caused by uncertainty introduced by approximate front-end processing. In other situations, however, where interpretation costs are extremely high, a reduction of front-end costs through use of ASPAs even by 80% may not be advantageous, if it engenders a slight increase in interpretation search.

Acknowledgment

This research was sponsored by the Department of the Navy, Office of the Chief of Naval Research, under contract number ONR #N00014-95-1-1198. The content of this paper does not necessarily reflect the position or policy of the government, and no official endorsement should be inferred.

References

Bregman, A., *Auditory Scene Analysis: The Perceptual Organization of Sound.* MIT Press. 1990.

Dorken, E., Nawab, S. H., and Lesser, V., "Extended model variety analysis for integrated processing and understanding of signals," *Proc. of 1992 IEEE Conf. on Acoustics, Speech and Signal Processing*, vol. V, pp. 73-76, San Francisco, 1992.

Draper, B., "Learning object recognition strategies," Ph.D. thesis, Computer Science Dept., Univ. of Massachusetts, Amherst, MA, 1993.

Ellis, D., "Prediction-driven computational auditory scene analysis," Ph.D. thesis, Electrical Engineering and Computer Science Dept., MIT, 1996.

Klassner, F., "Data reprocessing in signal understanding systems," Ph.D. thesis, Computer Science Dept., Univ. of Massachusetts, Amherst, MA, 1996.

Kohl, C., Hanson, A., and Reisman, E., "A goal-directed intermediate level executive for image interpretation," *IJCAI-87*, pp. 811–814, Milan, Aug. 1987.

Lesser, V., Pavlin, J., and Durfee, E., "Approximate processing in real-time problem solving," *AI Magazine*, vol. 9, no. 1, pp. 49–61, Spring 1988.

Lesser, V., Nawab, H., and Klassner, F., "IPUS: an architecture for the integrated processing and understanding of signals," *Artificial Intelligence*, vol. 77, no. 1, pp. 129–171, Aug. 1995.

Marr, D., *Vision.* San Francisco: W. H. Freeman & Co. 1982.

Nawab, H., and Dorken, E., "Quality versus efficiency tradeoffs in STFT computation," IEEE Trans. on Signal Processing, April 1995.

Nawab, H. and Quatieri, T., "Short-time fourier transform," *Advanced Topics in Signal Processing*, Prentice Hall, NJ, 1988.

Strat, T., "Natural object recognition," Ph.D. thesis, Computer Science Dept., Stanford, Aug. 1991.

Tsotsos, J., "The Complexity of Perceptual Search Tasks," IJCAI-89, pp. 1571–1577, Detroit, Aug 1989.