21. Digital Signal Processing Group

Academic and Research Staff

Prof. A.V. Oppenheim, Prof. A.B. Baggeroer, Prof. J.S. Lim, Prof. J.H. McClellan, Prof. B.R. Musicus

Graduate Students

A.J. Barabell, T.E. Bordley, P. Chan, S.R. Curtis, W.P. Dove, F.U. Dowla, G.L. Duckworth, C. Esmersoy, M.J. Glaser, D.W. Griffin, W.A. Harrison, D. Izraelevitz, T.H. Joo, A.L. Kurkjian, D.M. Martinez, E.E. Milios, D.R. Mook, C. Myers, S.H. Nawab, T.N. Pappas, M.D. Richard, J.-P. Schott, H. Sekiguchi, R. Sundaram, P.L. VanHove, D.L. Wang, M.S. Wengrovitz

21.1 Introduction

The Digital Signal Processing Group is carrying out research in the general area of digital signal processing. While a major part of our activities focus on the development of new algorithms, there is a strong conviction that theoretical developments must be closely tied to applications and to issues of implementation. The application areas which we deal with principally are speech, image and geophysical data processing. In addition to specific projects being carried out on campus, there is close interaction with Lincoln Laboratory and with the Woods Hole Oceanographic Institution.

In the area of speech processing, we have over the past several years worked on the development of systems for bandwidth compression of speech, parametric modeling of speech using pole–zero models, time–scale modification of speech, and enhancement of degraded speech. Recently we have obtained some important new results on time–scale modification of speech, growing out of a more general set of issues involving the estimation of a signal after its short–time Fourier transform has been modified. We are also exploring new techniques for speech enhancement using adaptive noise cancelling when multiple microphones are available.

There are also a number of projects related to image processing that we are currently pursuing. One project is restoration of images degraded by additive noise, multiplicative noise, and convolutional noise. Out of this project, we have developed a new image restoration system which is applicable to restoring images degraded by various different types of degradation. Our current work in this project involves development of new image restoration systems by exploiting additional available information such as the range map in infrared radar images. Another project is development of new image coding techniques by reducing quantization noise in PCM image coding or by reducing blocking effect in cosine transform image coding. Our approach to first decorrelate the quantization noise, and then reduce the quantization noise by a noise reduction system, led to a noticeable improvement in the performance of a simple PCM image coding system. We are currently working on the extension of these results to a more complex PCM image coding system. To reduce the blocking effect in cosine transform image coding, we have studied two approaches. In one approach, the coder is modified to account for the blocking effect and in the second approach, the coded image with blocking effect is processed to reduce the blocking effect. In both approaches, we have developed specific algorithms that significantly reduce the blocking effect in cosine transform coding. Another project that we are currently exploring is the development of a very low bit rate (below 50 kbits/sec) video-conferencing system. The specific approach we are currently studying is to model a human face, which is a regular feature in typical video-conferencing applications, with a set of parameters and then synthesize the image at the receiver from the coded parameters. This approach is analogous to modeling human speech for speech coding, which led to significant bit rate reduction without seriously degrading the speech intelligibility.

In the area of geophysical data processing, there are a variety of ongoing and new projects. During March–May 1980, we led a large acoustics and geophysics experiment, FRAM II, in the eastern Arctic. This was followed by an even more extensive program in March–May 1982, FRAM IV. Both of these experiments implemented an array of hydrophones and geophones with multichannel digital data recording. Work has been carried out on applying adaptive array processing to the measurement of the reverberation associated with the resulting acoustic signals, as well as the phase and group velocities of the seismic paths within the seabed and water column for refraction and bottom interaction studies. Work is also currently under way to examine the properties of several velocity function inversion techniques for multi-channel seismic ocean-bottom interaction data. The array data at FRAM II and FRAM IV are also being used to measure the scattering function of the channel at low frequencies and the directional spectra of the ambient noise in the Arctic. Associated with the acoustics experiment is a project aimed at extending the parabolic wave equation approximation for modeling underwater acoustics. The summers of 1983 and 1984 mark the beginning of a series of geophysical and acoustic experiments in the marginal ice zone of the Arctic. In these MIZEX experiments, large multichannel telemetered arrays will be used to receive acoustic signals propagated across ocean and crustal paths marking the transition between the open waters of the Atlantic, and the ice covered regions of the Arctic Ocean. The goals are to study the acoustic transmission properties of the laterally inhomogeneous and time varing water column, and to characterize the crustal velocity-depth function in this region.

Two additional projects related to geophysical signal processing in the context of ocean acoustics are the development of an algorithm for data processing to measure the acoustic reflection coefficient from the ocean bottom both for the deep water and shallow water cases. Out of this work has come a Hankel transform algorithm as well as a new method for generating synthetic data.

In both the context of image processing and geophysical data processing we have obtained some significant results in the multi-dimensional high resolution spectral estimation problem. Specifically, we have developed new algorithms for maximum entropy power spectrum estimation which are computationally simple relative to previous approaches and applicable to both equally spaced and non-equally spaced data for both one-dimensional and two-dimensional signals. This algorithm has been applied to investigate the characteristics of multi-dimensional maximum entropy spectral estimates. In addition, we are investigating several approaches to improve the performance of the maximum likelihood method for spectral estimation.

Recently, we have proposed a new approach to the problem of estimating multiple signal and/or parameter unknowns using incomplete and noisy data. Our Minimum Cross-Entropy Method applies an information theoretic criterion to optimally estimate a separable probability density for the signal model. Not only does this new approach include all the various Maximum Likelihood and Maximum *A Posteriori* methods as degenerate cases, but it also directly leads to a simple iterative method of solution in which we alternate between estimating the various unknowns, one at a time. We are now exploring applications to statistical problems, iterative signal reconstruction, short-time analysis/synthesis, and noisy pole/zero estimation.

Another interesting area of research is the connection between signal processing algorithms and computer architectures. The "speed" of an algorithm depends not only on how many operations it requires, but also on how suitable it is for the computer architecture it runs on. With the advent of VLSI technology, it is now possible to build customized computer systems of astonishing complexity for very low cost. Exploiting this capability, however, requires designing algorithms which not only use few operations, but also have a high degree of regularity and parallelism, or can be easily pipelined. Directions we are exploring include systematic methods for designing multi-processor arrays for signal processing, isolating signal processing primitives for hardware implementation, and searching for algorithms for multi-dimensional processing which exhibit a high degree of parallelism.

There also are a number of projects directed at the development of new algorithms with broad potential applications. For some time we have had considerable interest in the broad question of signal reconstruction from partial information such as Fourier transform phase or magnitude. We have shown theoretically how under very mild conditions signals can be reconstructed from Fourier transform phase information alone. We have also developed a variety of theories and algorithms relating to signal reconstruction from Fourier transform magnitude and from partial short-time Fourier transform information. We are also exploring the application of some of these theoretical results to problems such as speech and image coding.

A recent and growing emphasis in our group is the combination of signal processing and artificial intelligence techniques. There are a variety of problems in signal analysis that can be approached either from the analytical viewpoint characteristic of signal processing or the symbolic viewpoint characteristic of knowledge-based systems and artificial intelligence. We believe there is considerable potential for combining these two viewpoints into what we refer to as knowledge-based signal processing. There are currently two projects under way directed at developing this approach in the context of specific signal processing problems. One attempts to exploit artificial intelligence concepts to develop a knowledge based pitch detector and the second, to explore knowledge-based signal processing in the context of signal enhancement. We also expect to couple our work on knowledge-based signal processing into a project at Lincoln Laboratory on distributed sensor nets.

21.2 Parabolic Wave Equation Modeling for Underwater Acoustics

U.S. Navy - Office of Naval Research (Contract N00014-77-C-0266)

Arthur B. Baggeroer, Thomas E. Bordley

In underwater acoustics, the parabolic (Schroedinger) wave equation is often used as an approximation to the hyperbolic wave equation when solutions are sought in regions with slowly varying inhomogeneities. Such inhomogeneities prevent separation-of-variables or other simple solution techniques from being employed. The rationale behind the parabolic approximation is that if the inhomogeneities are weak, then the component of the field which is due to reflections ought to be small. Thus, it should be legitimate to split the wave equation into a pair of coupled equations in the "transmitted" and "reflected" fields and then solve for the transmitted field with the reflected field set to zero.

This approximation has two significant effects. First, since the equation is reduced to first order in the direction of propagation or range, the original problem is reduced to a comparatively simple initial value problem for which the necessary initial conditions can be readily found. Second, the split fields are decoupled approximants to the true transmitted and reflected fields, possessing a parabolic rather than a circular dispersion relation. Therefore, both the shape and scaling of waves emanating from the source at large angles with respect to the nominal direction of propagation are wrong.

Such errors are of negligible importance in many marine problems. Inhomogeneities in the water column normally are significant only when large distances are involved. Over such ranges, the heavy attenuation in the bottom allows only those waves which possess a small grazing angle with respect to the bottom, or equivalently, which emanate at an angle close to horizontal with respect to the source, from contributing substantially to the solution. Ordinarily then, a good estimate of the field can be obtained, because those waves which contribute the most to the final solution are represented well while only those which contribute little are represented poorly.

When the signals propagating through the bottom are not small or are of interest in themselves, this approach is less satisfactory. Local variations in the acoustic field with range can be rapid due to the presence of distinct material layers in the earth. Thus, the approximation can be poor. In this research, our concern is to extend the parabolic approximation so that the fields in both the earth and the sea are modelled well.

21.3 Adaptive Image Restoration

U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
National Science Foundation (Grant ECS80-07102)
Jae S. Lim, Philip Chan

There have been many image restoration techniques developed with the assumptions of global statistical stationarity and an additive signal independent noise model. These assumptions are usually not valid for many applications. The aim of this thesis research is to develop an adaptive algorithm for restoring images corrupted by noise. Without the usual assumption of global stationarity, the algorithm adapts itself to the local statistics of the processed picture elements. The result is the simultaneous reduction of noise and preservation of edges.

The adaptive algorithm will first be developed for an additive noise model. The ultimate aim will be to develop an algorithm for the restoration of infrared radar images corrupted by speckle noise which is signal dependent and non-additive.

21.4 Signal Reconstruction from Partial Fourier Domain Information

U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
National Science Foundation (Grant ECS80-07102)
Bell Laboratories Fellowship
Alan V. Oppenheim, Jae S. Lim, Susan R. Curtis

In a variety of practical problems, only the magnitude or the phase of the Fourier Transform (FT) of a signal is available, and it is desired either to reconstruct the signal exactly or to synthesize a signal which retains many of the important characteristics of the original signal. Our work in this area involves two distinct problems. One problem is to identify those portions of the FT which contain most of the "intelligibility" information; the other is to develop conditions under which a signal can be exactly reconstructed from various portions of the FT.

On the intelligibility problem, past work has shown that a signal synthesized with the correct phase and unity or average magnitude maintains many of the important characteristics of the original signal, whereas a signal synthesized from the correct magnitude and zero or random phase does not. In addition, a signal synthesized from the "signed-magnitude" (magnitude and one bit of phase) is intelligible. Recently, we have found that a signal synthesized from one bit of phase alone is also intelligible. This one bit of phase is contained in both the signed-magnitude and the phase, and thus this result helps to explain the earlier results.

On the exact reconstruction problem, past work has shown that, under mild restrictions, a finite length signal can be exactly reconstructed from its FT phase (to within a scale factor) or from its signed-FT magnitude. In addition, almost all multi-dimensional signals with finite support are uniquely specified by their FT magnitude, although this is not true in the one-dimensional case. Recently, we have developed some new interpretations of these results and removed the finite-length restriction from the signed-magnitude result. Currently, we are developing conditions under which a signal can be exactly reconstructed from one bit of its FT phase.

Despite the potential applicability of these results to a wide range of problems, so far they have been mostly of theoretical value. In addition to solving some remaining theoretical questions, we plan to explore applications of these results.

21.5 Knowledge-Based Pitch Detection

Amoco Foundation Fellowship
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
National Science Foundation (Grant ECS80-07102)
Alan V. Oppenheim, Randall Davis, Webster P. Dove

Knowledge-based signal processing is an effort to design signal processing programs that go beyond purely numerical processing of the data and try to symbolically reason about the problem in order to better solve it. Problems appropriate for this area those whose model is either too complex to be solved directly with a numerical algorithm, and those for which the model is not well understood.

Pitch detection falls into this category both because the speech signal model is not well specified, and because the model for the generation of pitch is not fully understood.

The aim of this project is to develop a program called the **Pitch Detector's Assistant** (PDA) which will both serve to reduce the effort involved in generating hand edited pitch and provide a laboratory for studying and programming the knowledge that makes humans better pitch trackers than existing automatic algorithms.

Existing methods of semi-automatic pitch detection[1] require the user to make a voicing decision and select a pitch individually for every frame. The PDA program is intended to analyze as much of the utterance as it is sure of and then help the user with the remaining difficult portions. Thus we expect a dramatic reduction in the time from the current 30 minutes per second of speech analyzed.

Although there have been projects which combine signal processing and AI technology for particular problems such as speech understanding[2,3] and underwater acoustic signal recognition,[4] the actual signal processing present in these systems has only been used as a means for generating symbolic objects. These objects are then manipulated by the AI portions of the program until an interpretation of the data is complete. The symbols do not provide information to assist subsequent numerical processing, and thus the information flows one way from numeric to symbolic form. The pitch detection problem choice is motivated by the observation that these other problems are ones of recognition (i.e. signals in, symbols out) and naturally lead to solutions which process numerically first and symbolically later.

By choosing a problem which involves signal output we assure the use of numerical processing in later portions of the program. The creation and study of programs which emphasize the interaction between symbolic and numerical processing is the primary purpose of the knowledge-based signal processing effort at M.I.T.

References

1. C.A. McGonegal, L.R. Rabiner, and A.E. Rosenberg, "A Semiautomatic Pitch Detector (SAPD)," IEEE Trans. ASSP 23, 570-574 (December 1975).
2. L. Erman, R. Hayes-Roth, V.R. Lesser, and D.R. Reddy, "The Hearsay-II Speech Understanding System: Integrating Knowledge to Resolve Uncertainty," Computing Surveys 12, 213-254 (June 1980).
3. B. Lowerre and R. Reddy, "The HARPY Speech Understanding System," in W. Lea (Ed.), Trends in Speech Recognition (Prentice-Hall, 1980) pp. 340-360.
4. H.P. Nii, E.A. Feigenbaum, J.J. Anton, and A.J. Rockmore, "Signal-to-Symbol Transformation: HASP/SIAP Case Study," AI Magazine 3, 23-35 (Spring 1982).

21.6 Multi-Dimensional High-Resolution Spectral Analysis and Improved Maximum Likelihood Method

U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
National Science Foundation (Grant ECS80-07102)
Jae S. Lim, Farid U. Dowla

Although MEM, MLM, and AR-modeling spectral estimation are high resolution spectral estimation algorithms, in multidimensional situations there are problems with each of these methods. The MEM algorithms are iterative and their applications to real world problems is prohibitive from numerical viewpoints. MLM does not quite achieve the resolution which one would like it to have. The AR-modeling has problems as the shape of the spectrum is distorted when the filter mask does not conform to certain symmetry.

We are developing a closed-form high resolution spectral estimation algorithm based on the concepts of MLM, MEM, and AR spectral estimation. We have found a useful relationship between the MLM and AR signal modeling in multidimensions. By exploiting this relationship and by studying the problem of array design in multidimensions for these algorithms, we propose to present an algorithm with computational properties like the MLM but whose resolution property is better than the MLM.

The performance of the algorithm is being evaluted on synthetic data and on real data in a multichannel radar tracking problem.

21.7 Processing and Inversion of Arctic Refraction Data

U.S. Navy - Office of Naval Research (Contract N00014-77-C-0266)
Arthur B. Baggeroer, Gregory L. Duckworth

Several seismic refraction and long-range propagation experiments were carried out during the *Fram 2* experiment in the Pole Abyssal Plain of the Eastern Arctic Ocean (May 1980). This work concerns the digital signal processing and inversion methods used to analyze and interpret these multichannel array data for crustal and oceanic velocity structure, and the results of these analyses. Examined are: i) methods for obtaining the partial plane wave decomposition of the data subject to the experimental limitations imposed by the Arctic environment, ii) methods for inverting these decompositions for oceanic and crustal structure, and iii) synthetic seismogram techniques using normal mode and WKBJ theory for use in inversion, algorithm testing, and model verification.

For the plane wave decomposition, conventional time and frequency domain array processing techniques for velocity analysis and slant stacking are contrasted to an optimal adaptive technique. In this method, the time varying statistics of the data are estimated and used to design the optimal time varying frequency-wavenumber filters used to filter the data to obtain the power or slant stack output. The criteria of phase velocity, group velocity (time of arrival), and amplitude estimator bias and stability are discussed. It is shown that the different methods often have complimentary advantages and disadvantages.

To recover the structure, several inversion methods for estimation of the ocean/crust velocity structure from the velocity analyzed refraction data are examined. In this work, an emphasis is placed on techniques which fully utilize the data obtained in this experiment. The incorporation of the directly estimated delay time, $\tau(p)$, and offset, $X(p)$, data as a function of horizontal slowness, p, is carried out. The judicious use of offset data in a velocity/depth migration inversion procedure constrains the resulting models more tightly, especially in the case of sparse offset sampling necessitated by the Arctic environment. The use of the WKBJ approximation in this inversion method allows more complete utilization of the data through inclusion of non-geometric arrivals and amplitude constraints. An iterative algorithm based around velocity-depth migration which checks for consistency between the τ and offset data using the WKBJ approximation is developed for both types of arrivals. In regions where the lateral homogeneity is not adequate for the migration technique, an approximate linear least-squares inversion procedure is used. This method utilizes all primary and multiple $\tau(p)$ and $X(p)$ information simultaneously in a tomographic technique to invert for laterally varying structures under the assumption that lateral variations are not excessive.

In addition to the refraction experiments, long-range (100–500 km) propagation data are used for the determination of water column and near-bottom sediment velocity structure. The use of the velocity spectral estimation techniques and sonogram analysis on the long range propagation data allows estimation of the modal dispersion curves and travel time structure of the arrivals. These dispersion curves can be inverted for water column sound speed structure, and the travel time data used for ocean bottom velocity structure determination. The use of both normal mode and WKBJ theory tightly constrains the velocity structure using single-shot data as long as the lateral homogeneity of the structure to be examined is sufficient.

The WKBJ synthetic seismogram technique, normal mode analysis and synthesis, and an approximate normal mode analysis/synthesis technique using WKBJ method to approximate the dispersion characteristics of an acoustic model have been examined. Their properties and utility in the inversion techniques are studied, and the results are used as test data for the velocity analysis/slant stack methods.

21.8 Signal Estimation from Modified Short-Time Fourier Transform

U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
National Science Foundation (Grant ECS80-07102)
Jae S. Lim, Daniel W. Griffin

In several applications, including time-scale modification of speech and speech enhancement by spectral subtraction, the Short-Time Fourier Transform (STFT) of a signal is modified so that it no longer corresponds to the STFT of any sequence. In these cases, we would like to find a time domain sequence that has a STFT closest to the desired STFT. Several methods are being investigated for estimating a sequence whose STFT is closest to a given modified STFT or modified STFT magnitude. One of the methods developed for estimating a signal from modified STFT magnitude has produced excellent results when applied to time-scale modification of speech. Informal listening tests have judged the results of this method superior to Portnoff's time-scale modification method based on the phase vocoder.

21.9 Speech Enhancement Using Adaptive Noise Cancelling Algorithms

U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
National Science Foundation (Grant ECS80-07102)
Jae S. Lim, William A. Harrison

This research is directed towards evaluating the performance of some general adaptive noise cancelling algorithms in a distributed noise environment. General adaptive noise cancelling algorithms utilize one or more reference microphones that record a correlated version of noise that is additively corrupting a desired signal in the primary microphone. In practice, the desired signal is often recorded by the reference microphones. Under these conditions, the algorithm will attempt to cancel part or all of the desired signal. Under certain conditions, modifications in an adaptive noise cancelling algorithm such as the Widrow–Hoff least mean square algorithm, allows one to still reduce the background noise without severely distorting the desired signal. One application of this work is in reducing the background noise in a jet fighter pilot's speech. Test cases conducted with simulated data have shown some promise that the ANC algorithm can be used to improve the SNR of the pilot's speech.

21.10 Overspecified Normal Equations for Autoregressive Spectral Estimation

U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
National Science Foundation (Grant ECS80-07102)
Jae S. Lim, David Izraelevitz

There is a one–to–one relationship between a set of P normalized positive definite correlation estimates and the P predictor coefficients derived using autoregressive modeling. Several researchers have proposed the use of $M > P$ correlation estimates to provide a better Pth order model. Specifically, the normal equations are augmented to provide M linear equations between the correlation estimates and the predictor coefficients. Since the system of equations is now overspecified, a least squares solution is required.

In this work a study is presented of some of the properties of the method of overspecified normal equations as applied to the problem of spectral estimation. The main contribution of this research is the derivation of the relationships between the number of correlations used, the model order and the signal–to–noise ratio of the signal, to the characteristics of the resulting spectral estimate. The characteristics studied are the spectral height, bandwidth and area. The method is shown to be a spectral density estimator like the ME method, where spectral *areas* rather than spectral values should be interpreted as estimates of power.

The relationships derived point to the number of correlations used over the minimum, i.e. model order, as a signal-to-noise enhancer. The resulting spectrum is equivalent to the ME spectrum under higher signal-to-noise conditions. Another result is the requirement of a proportionality constant dependent on the number of correlations and the model order which is necessary for unbiased signal-to-noise measurements. This constant is not required however, for measurements of relative power within the same spectral estimate, as in the power ratio of two sinusoids in noise.

The second part of the research presents some empirical studies using computer simulations which verify the theoretical predictions and provide the region of validity of the analysis. Further experiments study the interfering effect of several closely spaced sinusoids. The method of overspecified normal equations is shown to be much more sensitive to this interference than the ME method. Finally, some further empirical studies are made of the resolution capabilities of the method. Using the data derived, an empirical model is derived which seems to agree to some extent with the data.

21.11 Spectral Analysis Methods for Non-Stationary Time Series

U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
National Science Foundation (Grant ECS80-07102)
Jae S. Lim, Dennis M. Martinez

Spectral analysis methods play an important role in the study of signals and systems. Using spectral techniques, a very broad class of signals (time series) can be described in a manner which provides insight into the signal properties. Furthermore, for a very important class of systems (linear time-invariant), the effect of the system on the signal is easily determined. Consequently in engineering practice, spectral methods are fundamental to most signal analysis methods and system design procedures.

Most of the currently used spectral representations in continuous time map a time series defined over the infinite time interval $-\infty \leq t \leq \infty$, into a functional defined over a particular contour in the complex plane $(\sigma, j\omega)$. For example, the Fourier transform maps the signal x(t) defined for $-\infty \leq t \leq \infty$ into the functional $X(j\omega)$ defined for $-\infty \leq \omega \leq \infty$. Similar mappings are in wide-spread use in discrete time. While such representations provide insight about the global properties of a time series, they do not provide insight into its local properties.

In describing stationary time series, global properties are often sufficient, but this is not the case for non-stationary time series such as are encountered in speech analysis. Consequently a variety of spectral representations for non-stationary time series have been developed in the past. In speech analysis, the most common non-stationary spectral representation in use is the short-time Fourier transform. Time dependent spectral representations such as the short-time Fourier transform often provide considerably more insight into the properties of non-stationary time series than the Fourier transform itself.

The purpose of the research being conducted is twofold; (1) to review and compare presently defined non-stationary spectral representation methods; and (2) to investigate new techniques for spectral analysis of non-stationary time series.

21.12 Speech Coding Using the Phase of the Long-Time LPC Residual Signal

National Science Foundation (Grant ECS80-07102)
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
Alan V. Oppenheim, Evangelos E. Milios

The potential of using the long-time phase of the Fourier transform of the LPC residual signal in speech synthesis has been investigated. The modified residual obtained by using the correct phase and constant magnitude leads to speech that is close to the original, but slightly hoarse. However, if the correct magnitude over a low-frequency band of width about a fifth of the total spectrum and smoothed magnitude in the rest of the spectrum is used, and combined with the correct phase, the hoarseness disappears. Because the phase has a uniform distribution of values between $-\pi$ and π, uniform quantization and binary encoding of the values is appropriate. The low-frequency magnitude has a bell-shaped distribution of values; thus, Huffman coding can be used to advantage. The possibility of using the previous scheme in coding as well as the perceptual importance of the various parts of the long-time Fourier transform of speech and the LPC residual are topics of further research.

21.13 The Numerical Synthesis and Inversion of Acoustic Fields Using the Hankel Transform with Application to the Estimation of the Plane Wave Reflection Coefficient of the Ocean Bottom

U.S. Navy - Office of Naval Research (Contracts N00014-77-C-0196 and N00014-81-K-0742)
National Science Foundation (Grant ECS80-07102)
Alan V. Oppenheim, George V. Frisk,[11] Douglas R. Mook

---

[11] Woods Hole Oceanographic Institution

The plane wave reflection coefficient is an important geometry independent means of specifying the acoustic response of a horizontally stratified ocean bottom. Its determination is an integral step in the inversion of acoustic field measurements to obtain the parameters of the bottom and it is used to characterize an environment for purposes of acoustic imaging. This research, completed this year, studied both the generation of synthetic pressure fields to estimate the plane wave reflection coefficient and the inversion of measured pressure fields to estimate the plane wave reflection coefficient. These are related through the Sommerfeld integral, which has the form of a Hankel transform. The Hankel transform was extensively studied in this work and both its theoretical properties and numerical implementation were considered. A fast, NlogN, numerical algorithm was developed for both the Hankel and Abel transforms. The results of the theoretical and numerical studies on the Hankel transform have broad application beyond ocean acoustics. When these results were applied to the generation of synthetic data, the result was hybrid numerical-analytical algorithms which generate extremely accurate synthetic fields without sacrificing computation speed. These algorithms are capable of accurately incorporating the effects of trapped modes guided by slow speed layers in the bottom. The general results for the Hankel transform were also applied to study the inversion of measured pressure field data for the plane wave reflection coefficient. Practical issues associated with the inversion procedure were addressed, including the removal of the source field, sampling the effect of field measurements over a finite range, and uncontrolled variations in source-height. A phase unwrapping and associated interpolation scheme was developed to handle improperly spaced data.

A preliminary inversion of real pressure field data was performed. In parallel, an inversion of a synthetically generated field for similar bottom parameters was also performed and the results of processing the real and synthetic data compared. The estimate for the depth-dependent Green's function obtained from the real data was found to share many features with the depth-dependent Green's function estimated from the synthetic data, suggesting that the total inversion to obtain the plane wave reflection coefficient will soon be possible. Errors in the present estimate of the plane wave reflection coefficient were associated with uncontrolled variations.

21.14 Optimal Signal Reconstruction and ARMA Model Identification Given Noisy and Incomplete Observation Data

National Science Foundation (Grant ECS80-07102)
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
Jae S. Lim, Bruce R. Musicus

There are a large number of signal processing applications in which signals and/or model parameters must be estimated from noisy or incomplete observation data. For example, we may need to reconstruct a finite length signal from noisy measurements of its phase of magnitude. We may want to filter, interpolate and/or extrapolate a finite interval of stationary data, and possibly fit a pole/zero model to the data at the same time. We may be given grouped, truncated, censored, or heavily quantized data and need to estimate the parameters of the probability distribution from which they were drawn. Conventional Bayesian estimation methods, when applied to these problems, would typically require a numerically unpleasant multi-dimensional optimization of a multi-dimensional integral. We have developed a new approach to the problem of estimating multiple unknowns given noisy and incomplete data, however, which leads to a particularly simple and elegant estimation algorithm. We start with a probabilistic model describing the signals and parameters, together with a set of constraints which the signals and parameters are known to satisfy. We then optimally fit a separable probability density to this given model by minimizing a cross-entropy criterion. Surprisingly, all Maximum Likelihood and Maximum *A Posteriori* algorithms for this problem can be treated as degenerate cases of this single approach. Furthermore, it is easy to derive a simple iterative algorithm for calculating the optimal estimates. Each step of this algorithm treats each unknown one at a time, iterating between well-known filtering, interpolation, and/or extrapolation steps to estimate each signal, followed by well-known parameter estimation steps for each parameter. Convergence is guaranteed under mild conditions. Furthermore, this approach has been shown to work well in a variety of problems ranging from statistics to pole/zero estimation.

Another major focus of our research is the development of new algorithms for solving almost-oeplitz sets of linear equations. Equations of this type arise quite commonly in the filtering and signal reconstruction problems we have been investigating, and efficient solution methods are especially critical because the problems must be solved repetitively as the algorithms iterate. Levinson-style recursions have been most commonly discussed in the literature. Fast Choleski methods for solving the same problems can be derived as "mirror images" of the Levinson-style recursion. Using both forward and backward recursions to minimize storage, these fast Choleski methods are particularly valuable for solving the band diagonal, almost-Toeplitz problems that arise when filtering noisy ARMA signal data. These algorithms also strongly resemble "inside out" Euclidian polynomial algorithms, an observation which directly leads to new "doubling algorithms" for solving almost-Toeplitz problems in only $0(N\log^2 N)$ operations. Connections between Euclidian polynomial algorithms and Levinson recursion algorithms are also being explored in other contexts as well; in particular, we have developed a new Euclidian algorithm with "partial pivoting" for evaluating inverse z-transforms and inverse Laplace transforms of rational polynomials.

21.15 The Use of Speech Knowledge in Speech Enhancement

Schlumberger-Doll Research Center Fellowship
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
National Science Foundation (Grant ECS80-07102)
Alan V. Oppenheim, Cory Myers

The problem of speech enhancement from speech corrupted by noise is one which has been of great interest in the signal processing community. Techniques based on estimation theory and based on speech analysis and resynthesis models have been used to differing degrees of success. Such techniques generally make use of only a small amount of the knowledge that is available about the speech process. We propose to build a system for the enhancement of noisy speech which will be able to use more and different types of speech knowledge. This system will attempt to mix various processing tools for signal enhancement with several artificial intelligence tools for symbolic reasoning. Through close cooperation among these tools it is hoped that a speech enhancement system which utilizes large amounts of speech knowledge can be realized.

The major issues being studied in this project are the representation and use of signal processing knowledge. Current work involves the acquisition and representation of acoustic–phonetic knowledge that can be used for speech enhancement. Future work will involve acquisition and representation of knowledge about perception of noisy speech and the behavior of various signal processing algorithms. Mechanisms for reasoning about the choice of signal processing algorithms in different contexts will also be developed. Emphasis will be paid to those situations in which outside symbolic information, e.g., speaker sex, phonetic transcription, etc., is available.

This work is part of a more general project in knowledge–based signal processing, i.e., the attempt to combine techniques from both signal processing and artificial intelligence in a close, cooperative manner.

21.16 Estimation of the Degree of Coronary Stenosis Using Digital Image Processing Techniques

U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
National Science Foundation (Grant ECS80-07102)
Jae S. Lim, Thrasyvoulos N. Pappas

The aim of this research is the development of an algorithm for evaluating the degree of coronary artery stenosis from coronary cine–angiograms. A cine–angiogram is a sequence of x–ray pictures of the coronary arteries in which a contrast agent has been injected via a catheter. The precise measurement of the stenosis of the coronary arteries is important in the treatment of patients with ischemic heart disease.

The first step will be the determination of the percentage diameter reduction from a single frame of the cine–angiogram. This will require the detection of the boundaries of the coronary arteries and the analysis of the variation of their diameter. Preprocessing of the image for noise reduction will hopefully lead to improved boundary detection.

Subsequent steps will involve the analysis of multiple frames of the cine-films, and the investigation of densitometric procedures, which use the brightness information within the artery to obtain estimates of the cross-sectional percentage area reduction.

21.17 Automatic Target Detection in Aerial Reconnaissance Photographs

U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
National Science Foundation (Grant ECS80-07102)
Jae S. Lim, Michael D. Richard

The detecting of small anomalous regions in images has aroused much interest in such areas as optical aerial reconnaissance, radar analysis, terrain classification, and medical diagnosis through imagery. A recently developed algorithm[1] has proven highly successful in detecting small objects or targets in natural terrain such as trees, grass, and fields of aerial photographs. The algorithm uses a significance test to distinguish each image pixel as either background or non-background (i.e., target). Specifically, the background is assumed to be characterized by a nonstationary Gaussian random process. The algorithm further represents the background by a two dimensional (2-D) autoregressive model. The resulting significance test is expressed as the error residuals of 2-D linear prediction.

This research will explore several new areas to either develop a superior detection algorithm or to significantly improve the existing one. First, the issue of target modelling will be explored. The current algorithm models only the background and treats targets simply as anomalies in the background. The question arises as to how a suitable model for targets can be incorporated in a detection algorithm. Second, methods to detect and to fully determine the boundaries of larger objects will be considered. The present algorithm can detect only small point objects representing statistical irregularities in the background random process. The issue of detecting larger targets poses significant questions regarding object detection, image segmentation, and boundary extraction. Additional research in these two areas should improve the somewhat favorable results obtained by using linear predictive techniques to detect anomalous regions in images.

Reference

1. T.F. Quatieri, "Object Detection by Two-Dimensional Linear Prediction," M.I.T. Lincoln Laboratory Technical Report 632, January 1983.

21.18 Enhancement of Helium-Degraded Speech

U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
National Science Foundation (Grant ECS80-07102)
Toshiba Company Fellowship
Jae S. Lim, Hiroshi Sekiguchi

Considerable progress in the theory of signal processing by using the Short-Time Fourier Transform (STFT) has been made in the last several years. Particularly, new algorithms have been developed to reconstruct a signal from its modified STFT or from its modified STFT magnitude.

Among various applications of these algorithms, the enhancement of helium-degraded speech is fairly interesting. The problem of helium-degraded speech lies in the poor intelligibility of the speech uttered in high-pressured helium-oxygen atmosphere by deep sea divers. This speech is heavily distorted in the frequency domain primarily due to differences of acoustic properties in the high-pressured helium-oxygen gas. Moreover, it is considerably degraded by additional noise from life-support systems.

Some work has been done to improve the quality of helium speech. However, satisfactory results have not been achieved so far, partially because powerful methods like STFT have not been applied to accomplish a non-linear modification in the frequency domain.

This research intends to propose a new algorithm for enhancement of helium speech by using STFT techniques. The application of STFT on the problem is primarily motivated by intention to easily modify distorted frequency spectrum of helium speech in a non-linear way so that the modified version is closer to the frequency spectrum of the non-degraded original speech. Furthermore the STFT technique provides the possibility of a noise cancellation method with high performance.

21.19 Facial Parameterization for Low Bit Rate Video Conferencing

U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
National Science Foundation (Grant ECS80-07102)
Vinton Hayes Fellowship
Jae S. Lim, Ramakrishnan Sundaram

The idea of developing an algorithm for the construction and reconstruction of a human image based on edge information for purposes of video conferencing is being studied. The ultimate goal is to achieve a drastic reduction in the amount of picture information required to be sent over a channel, thus reducing bandwidth without much loss in quality. The algorithm will exploit the connectivity of the subimages of the human face, thus isolating each for independent edge detection.

Different edge detection schemes like the discrete Laplacian, Roberts' cross gradient, and transform methods are being studied and employed to get most effective detectors. Currently, work is being done on a continuous tone–gray level vertical face without glasses, beard, and moustaches. Those cases are likely to be looked into later.

21.20 Bottom Profile Determination in a Shallow Ocean

Hertz Foundation Fellowship
National Science Foundation (Grant ECS80-07102)
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
Alan V. Oppenheim, George V. Frisk,[12] Michael Wengrovitz

The problem of inverting reflected pressure field data from an acoustic point source in a shallow ocean to determine information about the bottom is being studied. Although there has been recent progress in this area, only the case of a deep ocean has been considered. The deep water case assumes that the reflected field consists only of components which are reflected from the bottom. In the shallow water case, it is necessary to consider the multiple reflections which occur from both the ocean bottom and the ocean surface. This requires the use of algorithms which dereverberate the reflected field data prior to use in an inversion scheme.

There are similarities between this problem and three other problems which have been studied. Iterative maximum likelihood dereverberation schemes have been proposed for use in the study of acoustic well–logging data from a borehole. It is hoped that some of these algorithms may be applicable to dereverberating reflected plane wave data as well as to the originally proposed cylindrical wave data. Also, in a separate study, a very fast algorithm to perform the Hankel transform was proposed. The Hankel transform relates the pressure field to the bottom reflection coefficient function in a deep ocean. This algorithm and other results which deal with removing singularities from the reflection coefficient function may also be applicable. Finally, there is a close relationship between principles involved in acoustic tomography and this problem. Recent progress in the development of inverse scattering algorithms for use in medical tomography may also be applicable to determining ocean bottom profile data from dereverberated scattered field data. This research will explore these and other new methods of inverting shallow ocean data to determine the characteristics of the underlying bottom.

---

[12] Woods Hole Oceanographic Institution