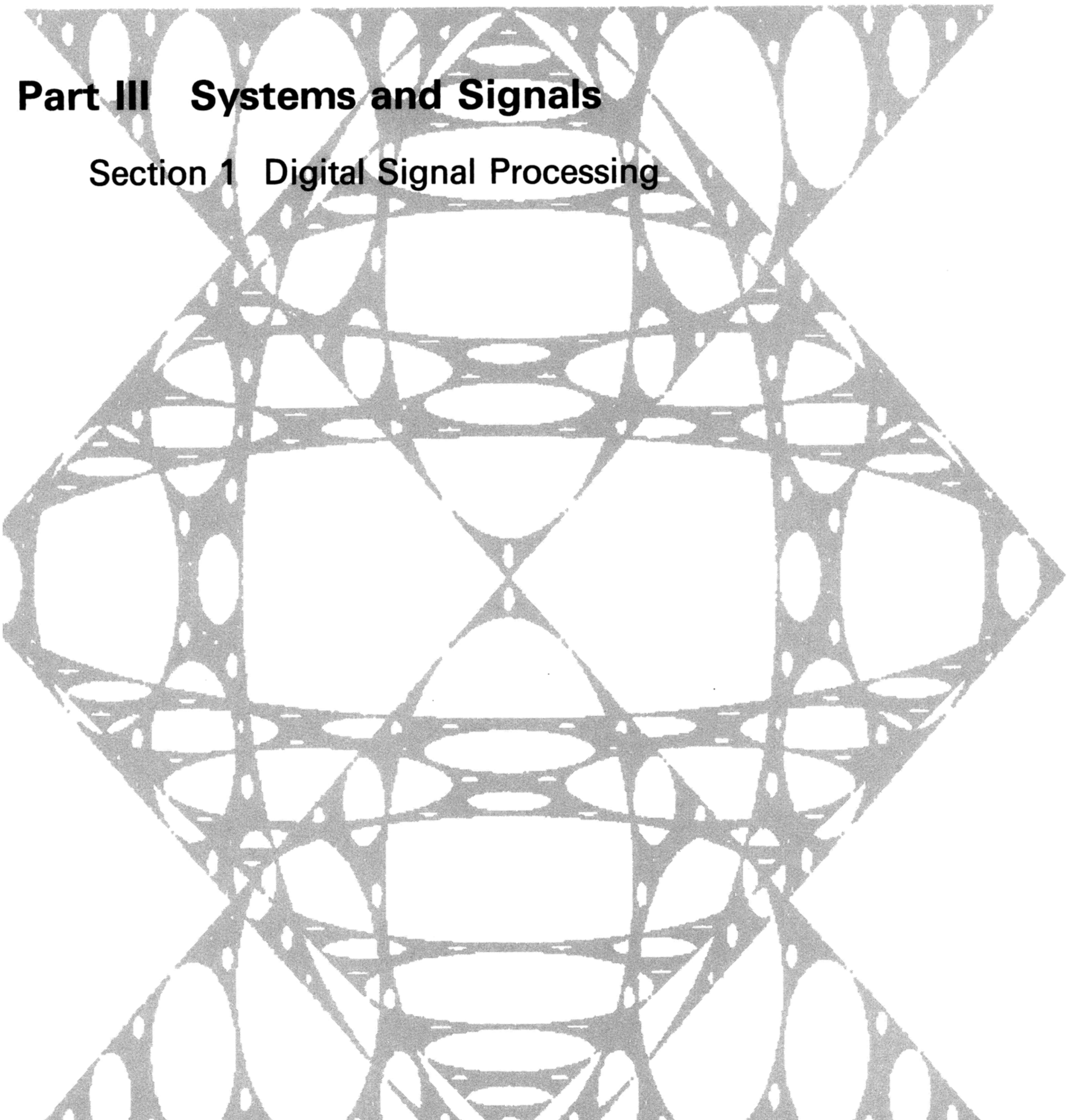
Part III Systems and Signals

Section 1 Digital Signal Processing



Section 1 Digital Signal Processing

Chapter 1. Digital Signal Processing

Academic and Research Staff

Professor Alan V. Oppenheim, Professor Jae S. Lim, Professor Bruce R. Musicus, Professor Arthur B. Baggeroer, Dr. Meir Feder, Giovanni Aliberti

Visiting Scientist

Dr. Ehud Weinstein[1]

Collaborating Scientist

Professor Ramesh Patil[2]

Graduate Students

Matthew M. Bace, Joseph E. Bondaryk, Michael S. Brandstein, Lawrence M. Candell, Daniel T. Cobra, Michele M. Covell, Dennis C.Y. Fogg, John C. Hardwick, Steven H. Isabelle, Jacek Jachner, Tae H. Joo, Rosalind W. Picard, G.N. Srinivasa Prasanna, James C. Preisig, Michael Richard, Gregory W. Wornell

Support Staff

Deborah A. Gage, Cynthia LeBlanc

Part-Time Assistants/Special Projects

Alfredo Donato, Hassan Gharavy

1.1 Introduction

The Digital Signal Processing Group is carrying out research in the general area of signal processing. In addition to specific projects being carried out on campus, there is close interaction with MIT Lincoln Laboratory and Woods Hole Oceanographic Institution. 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. The application areas that we currently are involved with are speech, image, video and geophysical signal processing. We also believe that algorithm development should be closely tied to issues of implementation because the efficiency of an algorithm depends not only on how many operations it requires, but also on how suitable it is for the computer architecture on which it runs. Also strongly affecting our research directions is the sense that while, historically, signal processing has principally emphasized numerical techniques, it will increasingly exploit a combination of numerical and symbolic processing, a direction that we refer to as knowledge-based signal processing.

In the area of knowledge-based signal processing, there are currently two research projects. One involves the concept of symbolic correlation, which is concerned with the problem of signal matching using multiple levels of description. This idea is being investigated in the context of vector coding of speech signals. Symbolic correlation will

[1] Tel Aviv University, Faculty of Engineering, Tel Aviv, Israel.

[2] MIT Laboratory for Computer Science.

entail the use of both symbolic and numeric information to efficiently match a speech signal with stored code vectors. The second project in this area deals with the representation and manipulation of knowledge and expressions in the context of signal processing. This work examines issues such as the representation of knowledge, derivation of new knowledge from that which is given, and strategies for controlling the use of this knowledge.

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 speech modeling, time-scale modification of speech and enhancement of degraded speech. A new model-based speech analysis/synthesis system has been developed and shown to be capable of high-quality speech production. It is currently being used in several mid- and low-rate speech coding systems. A coding system based on this model has achieved a bit rate of 4.8 kbps while still maintaining high speech quality. Current work seeks to reduce the bit rate to 2.4 kbps. A different model has been proposed with dual excitation and filter structure to improve speech processing performance. Research on adaptive noise cancellation techniques in a multiple microphone environment has led to an approach based on maximum likelihood estimation that has shown substantial improvements over previous techniques.

A number of interesting image processing projects are currently being conducted. One such project is examining texture regions in images. The goal of this research is to develop a model which is capable of synthesizing a continuum of structural and stochastic textures. Another project is focusing on a new technique of motion estimation with improved performance at motion boundaries. Fuzzy segmentation information is used to iteratively refine the estimates of the motion field. A third project seeks to develop a receiver-compatible scheme to reduce the effects of channel imperfections on the received television picture. Adaptive modulation is being investigated in order to make more efficient use of the currently underutilized bandwidth and dynamic range of the NTSC signal. Several other image processing projects have recently been completed. These include the estimation of coronary artery boundaries in angiograms, motion compensation for moving pictures, and magnitude only reconstruction of images.

In the area of geophysical signal processing, one research project involves the transformation of side-scan sonar data. In practice this data is corrupted by a number of factors related to the underwater environment. In this project, the goal is to explore digital signal processing techniques for extracting the topographic information from the actual sonographs. Concepts under study include the removal of distortions caused by towfish instability and reconstruction based on multiple sonographs taken from different angles. A second project involves near-field beamforming for underwater acoustics.

There are also a number of projects directed toward 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. This work has also been extended to the reconstruction of multidimensional signals from one bit of phase and exploiting duality, zero crossing and threshold crossing information. Reconstruction from distorted zero crossings has been studied as well as reconstruction from multiple threshold crossings which has been shown to be a better-conditioned problem than reconstruction using only a single crossing.

We are also examining the problem of narrowband signal detection in wideband noise, in particular in the presence of relative motion between transmitter and receiver. New algorithms are applied to the detection of gravitational waves generated by the orbiting motion of binary star systems. Another area of research is the estimation of the angle of arrival of multiple sources to an antenna array which is in motion. Algorithms are pursued which are computationally efficient for general antenna geometries, with application in radar and sonar systems.

Research also continues on the relationship between digital signal processing and stochastic estimation. We are exploring algorithms which can be applied to statistical problems, iterative signal reconstruction, short-time analysis/synthesis, and parameter estimation. The theory of functional approximation using generalized polynomials is being studied for application to windowing of arrays for near-field beamforming, windowing of time series with non-uniform samples, windowing of non-uniform and multidimensional arrays, and mask design for x-ray lithography.

A new area of interest is chaotic dynamical systems. Present work is aimed at applying non-linear dynamical system theory and fractal process theory to digital signal processing scenarios. Self-similar or fractal signal models are being investigated and novel spread spectrum techniques in the use of iterated cellular autonoma for modeling signals. The chaotic dynamics at work in quantization and overflow problems are also being explored.

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. The 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 multidimensional processing that exhibit a high degree of parallelism. We are also investigating highly-parallel computer architectures for signal understanding, in which a mixture of intensive computation and symbolic reasoning must be executed in an integrated environment. Another research direction involves computer-assisted design of signal processing chips. The approach taken is to use design specifications that include multiple, interacting performance specifications and generate designs at a multitude of performance levels to present possible trade-offs among the designs.

1.2 Receiver-Compatible Adaptive Modulation for Television

Sponsors

National Science Foundation Fellowship
National Science Foundation
(Grants ECS 84-07285
and MIP 87-14969)
U.S. Navy - Office of Naval Research
(Contract N00014-81-K-0742)

Project Staff

Professor Jae S. Lim, Matthew M. Bace

There have been numerous proposals for methods to improve upon the quality of the current NTSC television picture. Most of these proposals have concentrated on methods for increasing either the spatial or the temporal resolution of the television picture. While these proposals promise significant improvements in picture quality, the fact remains that until an effective scheme to combat channel noise has been introduced, these improvements will never be fully realized. Degradations such as random noise ("snow"), echo, and intersymbol interference (channel crosstalk) are still the greatest barrier to high-quality television.

This research will attempt to develop a receiver-compatible scheme to reduce the effects of channel imperfections on the received television picture. In particular, the method of adaptive modulation will be employed in an attempt to make more efficient use of the currently underutilized bandwidth and dynamic range of the NTSC signal. By concentrating more power in the higher spatial frequencies and using digital modulation to send additional information in the vertical and horizontal blanking periods, the existing television signal can be altered so that it is more robust in the presence of high-frequency disturbances. Furthermore, it is possible to adjust the parameters of this scheme so that the modulated signal may be received intelligibly even on a standard receiver (although an improved receiver will be required to realize the full benefits of adaptive modulation).

Before a conclusion can be reached as to which adaptive modulation scheme is optimal, many details must be considered. Among the parameters which may be varied are: the control over the adaptation and compression factors, the form of the input low-pass filters, the interpolation scheme to be used at both the transmitter and the receiver, and the encoding of the digital data. These parameters will be adjusted to optimize the performance of the modulation scheme with respect to two fundamental performance criteria: the degree to which the channel degradations are removed when the signal is received on an improved receiver and the degree to which the signal is distorted when received on a standard receiver.

1.3 Reconstruction Of Nonlinearly Distorted Images From Zero Crossings

Sponsors

National Science Foundation
(Grant ECS 84-07285)
U.S. Navy - Office of Naval Research
(Contract N00014-81-K-0742)

Project Staff

Professor Alan V. Oppenheim, Joseph E. Bondaryk

It has been shown theoretically that band-limited, multidimensional signals can be specified to within a constant factor by the information contained in the locations of their zero crossings. It has been shown experimentally that two-dimensional, band-limited signals can be reconstructed to within a constant factor from zero crossing information alone. The two-dimensional signals used were derived from images and their zero crossings corresponded to the threshold crossings of the images.

In this research, the problem considered is that of two-dimensional, bandlimited signals which have been affected by memoryless, nonlinear distortions. It is shown that such distortions retain the information required by the above theory, if they contain a monotonic region. Therefore, reconstruction to within a scale factor of an original, band-limited image from the threshold crossing information of a distorted image is made possible. The zero crossing coordinates of the two-dimensional signal derived from a distorted image are substituted into a Fourier Series representation of the original signal to form a set of homogeneous, linear equations with the Fourier coefficients of the original signal as unknowns.

The least squares solution to this set of equations is used to find the Fourier coefficients of the original signal, which are inverse Fourier transformed to recover the original image. By comparison of the distorted and reconstructed images, the nature of the distortion can be described. Some of the numerical problems associated with the reconstruction algorithm are also considered.

The reconstruction process is particularly stable for images which have bandedge Fourier components of high magnitude. It is shown that reconstruction from the zero crossing information of these images is similar to reconstruction from halftones, binary-valued images used to represent images which contain a continuous range of tones. This theory is finally extended to include reconstruction of bandlimited images from the zero crossings of distorted halftones.

This research was completed in May 1988.

1.4 Development of a 2.4 Kbps Speech Vocoder

Sponsors

National Science Foundation Fellowship
National Science Foundation
(Grant MIP 87-14969)
U.S. Navy - Office of Naval Research
(Contract N00014-81-K-0742)

Project Staff

Professor Jae S. Lim, Michael S. Brandstein

The recently developed Multi-Band Excitation Speech Model has been shown to accurately reproduce a wide range of speech signals without many of the limitations inherent in existing speech model based systems.[3] The robustness of this model makes it particularly applicable to low bit rate, high-quality speech vocoders. Griffin and Lim[4] first described a 9.6 Kbps speech coder based on this model. Later work resulted in a 4.8 Kbps speech coding system.[5]

Both of these systems have been shown to be capable of high-quality speech reproduction in both low and high SNR conditions.

The purpose of this research is to explore methods of using the new speech model at the 2.4 Kbps rate. Results indicate that a substantial amount of redundancy exists between the model parameters. Current research is focused on exploiting these redundancies so as to quantize these parameters more efficiently. Attempts are also underway to simplify the existing model without significant reduction in speech quality.

1.5 Correction of Geometric Distortions in Side-Scan Sonographs Caused by Sonar Motion Instabilities

Sponsors

Scholarship from the Federative Republic of Brazil
National Science Foundation (Grants ECS 84-07285 and MIP 87-14969)
U.S. Navy - Office of Naval Research (Contract N0001-81-K-0742)

Project Staff

Professor Alan V. Oppenheim, Daniel T. Cobra

Since its introduction in the early sixties, side-scan sonar has proved to be a very important tool for underwater exploration, and in particular for marine geology. Its applications include surveying the seafloor, search and location of objects on the bottom of the sea, and prospection of mineral deposits.

The information contained in reflected sound waves is used by side-scan sonar to produce an acoustic image, called a sonograph, which constitutes a composite representation of the topographic features and the relative reflectivity of the various materials on the seabed. Due to several factors, however, sonographs do not provide a precise depiction of the topology. The image can suffer from radiometric interferences such as those caused by dense particle suspension in the water, shoals of fish, or by ultrasonic waves generated by passing ships. Large-scale geometric distortions can be caused by deviations in the ship's trajectory from the ideal straight path. Small-scale distortions can be caused by motion instability of the towing body on which the transducers are mounted, due to underwater currents or to the ship's sway being communicated to the towed body. As a result, the interpretation of sonographs often requires extensive practice and can be a tedious and time-consuming task.

Radiometric distortions and large-scale geometric distortions have been successfully corrected through standard image-processing techniques. Our goal is to develop techniques to detect and correct small-scale geometric distortions due to sonar motion instabilities by digital post-processing of sonographs. This is a problem which hitherto had not been successfully addressed in the literature. This project is being conducted under MIT's joint program with the Woods Hole Oceanographic Institution, with the cooperation of the U.S. Geological Survey.

---

3 D.W. Griffin and J.S. Lim, "A New Model-Based Speech Analysis/Synthesis System," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Tampa, Florida, March 26-29, 1985, pp. 513-516.

4 D.W. Griffin and J.S. Lim, "A High Quality 9.6 kbps Speech Coding System," IEEE International Conference on Acoustics, Speech, and Signal Processing, Tokyo, Japan, April 8-11, 1986.

5 J.C. Hardwick, *A 4.8 Kbps Multi-Band Excitation Speech Coder*. S.M. thesis, MIT, 1988.

1.6 Representation and Manipulation of Signal Processing Knowledge and Expressions

Sponsors

National Science Foundation
(Grants ECS 84-07285
and MIP 87-14969)
U.S. Navy - Office of Naval Research
(Contract N00014-81-K-0742)

Project Staff

Professor Alan V. Oppenheim, Michele M. Covell

The phrase "signal processing" is used to refer to both "symbolic" and "numeric" manipulation of signals. "Symbolic" signal processing manipulates the signal description as opposed to the signal values with which "numeric" signal processing is primarily concerned. Efforts have been made to create computer environments for both types of signal processing.[6] Some issues that arise as a result of this work concern uniform representation of knowledge, derivation of new knowledge from that which is given, and strategies for controlling the use of this knowledge. This research is concerned with these areas and how they apply to digital signal processing.

Representations that have been used in symbolic signal processing[7] have been largely distinct from those used in numeric signal processing.[8] The types of representations used are further separated by the control structures that the numeric and symbolic information commonly assume, the distinction essentially being the same as the distinction between Algol-like languages and logic programming languages. This dichotomy results from the differing amounts of available knowledge about appropriate approaches to the problems being addressed. By separating the control structure from application knowledge, this dichotomy can be avoided.

Strategies for controlling when knowledge about a signal is used should be provided and new strategies should be definable, since these control structures provide additional information about the problem space, namely, approaches that are expected to be profitable. Control strategies can also be used to outline new approaches to a problem which would not be considered by simple trigger-activated reasoning.

Finally, the ability to derive new knowledge from that which is given is desirable. This ability would allow the amount of information initially provided by the user to be minimized. The environment could increase its data base with new conclusions and their sufficient pre-conditions. Two immediate advantages of providing the environment with this ability are the reduction in the programming requirements and the possible "personalization" of the data base. A reduction in programming requirements is available since information that is derivable from given information need not be explicitly encoded. Commonly, this type of information is provided to improve the performance of the derivation process. Secondly, since the environment would add information to the data set according to conclusions prompted by the user's queries, the data set would expand in those areas which the user had actively explored.

---

[6] C. Myers, *Signal Representation for Symbolic and Numeric Processing*. Ph.D. diss., Dept. of Electr. Eng. and Comp. Sci., MIT, 1986; W. Dove, *Knowledge-Based Pitch Detection*. Ph.D. diss., Dept. of Electr. Eng. and Comp. Sci., MIT, 1986.

[7] Ibid.; E. Milios, *Signal Processing and Interpretation Using Multilevel Signal Abstractions*. Ph.D. diss., Dept. of Electr. Eng. and Comp. Sci., MIT, 1986.

[8] C. Myers, *Signal Representation for Symbolic and Numeric Processing*. Ph.D. diss., Dept. of Electr. Eng. and Comp. Sci., MIT, 1986; G. Kopec, *The Representation of Discrete-Time Signals and Systems in Programs*. Ph.D. diss., Dept. of Electr. Eng. and Comp. Sci., MIT, 1980.

1.7 Iterative Algorithms for Parameter Estimation from Incomplete Data and Their Applications to Signal Processing

Sponsors

National Science Foundation
(Grants ECS 84-07285
and MIP 87-14969)
U.S. Navy - Office of Naval Research
(Contract N00014-81-K-0742)

Project Staff

Professor Alan V. Oppenheim, Dr. Meir Feder

Many signal processing problems may be posed as statistical parameter estimation problems. A desired solution for the statistical problem is obtained by maximizing the Likelihood (ML), the A-Posteriori probability (MAP) or some other criterion, depending on the a-priori knowledge. However, in many practical situations the original signal processing problem may generate a complicated optimization problem, e.g., when the observed signals are noisy and "incomplete."

An iterative framework for maximizing the likelihood, the EM algorithm, is widely used in statistics. In the EM algorithm, the observations are considered "incomplete" and the algorithm iterates between estimating the sufficient statistics of the "complete data" given the observations and a current estimate of the parameters (the E step), and maximizing the likelihood of the complete data, using the estimated sufficient statistics (the M step). When this algorithm is applied to signal processing problems, it yields, in many cases, an intuitively appealing processing scheme.

In the first part of this research, we investigate and extend the EM framework. By changing the "complete data" in each step of the algorithm, we can achieve algorithms with better convergence properties. In addition, we suggest EM type algorithms to optimize other (non ML) criteria. We also develop sequential and adaptive versions of the EM algorithm.

In the second part of this research, we examine some applications of this extended framework of algorithms. In particular, we consider

- Parameter estimation of composite signals, i.e., signals that can be represented as a decomposition of simpler signals. Examples include:
  - Multiple source location (or bearing) estimation
  - Multipath or multi-echo time delay estimation
- Noise cancellation in a multiple microphone environment (speech enhancement)
- Signal reconstruction from partial information (e.g., Fourier transform magnitude).

The EM-type algorithms suggested for solving the above "real" problems provide new and promising procedures, and they thus establish the EM framework as an important tool to be used by a signal processing algorithm designer.

Portions of this work were supported in part by the MIT – Woods Hole Oceanographic Institution Joint Program. This research concluded in August 1988.

1.8 Assisting Design Given Multiple Performance Criteria

Project Staff

Professor Bruce A. Musicus, Professor Ramesh Patil, Dennis C.Y. Fogg

In this thesis, we explore ways for computers to assist designers of signal processing chips. Our system's novelty is that the design specifications include multiple, interacting performance specifications in addition to the traditional functional specification. The performance factors that can affect a design include: throughput rate, latency, chip area, power usage, technology, and pin count. The ability to explicitly consider tradeoffs between various performance factors changes both the designed object and the process by which the object is designed. The system's output is a set of possible designs. We limit the system's expertise to a niche of signal processing where computation is not extremely regular.

The design process is divided into two phases. The first transforms the functional specification into various architectures described at the register transfer level. The second phase chooses implementations for the various functional units. We have developed two techniques for the second phase. One does a heuristically limited enumeration of possible combinations. The other formulates the task as a linear programming problem to find optimal assignments. Both methods generate designs at a multitude of performance levels thereby giving the user a sense of the possible tradeoffs among designs.

1.9 A Dual Excitation Speech Model

Sponsors

National Science Foundation
(Grants ECS 84-07285
and MIP 87-14969)
U.S. Navy - Office of Naval Research
(Contract N00014-81-K-0742)
U.S. Air Force - Electronic Systems Division
(Contract F19628-85-K-0028)

Project Staff

Professor Jae S. Lim, John C. Hardwick

One class of speech analysis/synthesis systems (vocoders) which has been extensively studied and used in practice is based on an underlying model of speech. Even though traditional vocoders have been quite successful in synthesizing intelligible speech, they have not been successful in synthesizing high-quality speech. The MultiBand Excitation (MBE) speech model, introduced by Griffin, improves the quality of vocoder speech through the use of a series of frequency dependent voiced/unvoiced decisions. The MBE speech model, however, still results in a loss of quality as compared to the original speech. This degradation is caused in part by the voiced/unvoiced decision process. A large number of frequency regions contains a substantial amount of both voiced and unvoiced energy. If a region of this type is declared voiced, then a tonal or hollow quality is added to the synthesized speech. Similarly, if the region is declared unvoiced, then additional noise occurs in the synthesized speech. As the signal-to-noise ratio decreases, the classification of speech as either voiced or unvoiced becomes more difficult and, consequently, the degradation is increased.

A new speech model has been proposed in response to the aforementioned problems. This model is referred to as the Dual Excitation (DE) speech model, due to its dual excitation and filter structure. The DE speech model is a generalization of most previous speech models, and with the proper selection of the model parameters it reduces to either the MBE speech model or to a variety of more traditional speech models.

Current research is examining the use of this speech model for speech enhancement, time scale modification and bandwidth compression. Additional areas of study include further refinements to the model and improvements in the estimation algorithms.

1.10 Motion Estimation from Image Sequences

Sponsors

AT&T Bell Laboratories
Doctoral Support Program
National Science Foundation
(Grants ECS 84-07285
and MIP 87-14969)
U.S. Navy - Office of Naval Research
(Contract N00014-81-K-0742)

Project Staff

Professor Jae S. Lim, Steven H. Isabelle

Most biological systems possess some type of mechanism for sensing motion in the world around them and use this information to assist them in many tasks. For example, visual motion information is used for object tracking, feedback for physical movement, and forming inferences about the surrounding environment. Research in the area of motion estimation from image sequences is aimed at equipping computers with some of these capabilities. Many motion estimation techniques rely on simple models for object motion to form local estimates of motion directly from image luminance at many points in an image, then combine these local estimates in accordance with same global model of the motion field. These algorithms tend to perform poorly in regions of motion field discontinuities such as foreground/background boundaries.

Improved performance could be achieved if the frames of the image sequence could be segmented into regions containing similar motion. As posed, this is a circular proposition, since an image frame must be segmented based on motion to estimate motion. This circularity may be resolved by using a technique of iterative refinement which, starting with a crude estimate of the motion field, produces a crude segmentation which is then used to refine the motion estimate.

This research is aimed at investigating techniques of image motion estimation with improved performance at motion boundaries by incorporating fuzzy segmentation information. As a by-product, the segmentation produced is itself useful for applications such as object tracking and object detection.

1.11 MultiLevel Signal Matching for Vector Quantization

Sponsors

Canada, Bell Northern Research Scholarship
Canada, Fonds pour la Formation de Chercheurs et l'Aide a la Recherche Postgraduate Fellowship
National Science Foundation (Grants ECS 84-07285 and MIP 87-14969)
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)

Project Staff

Professor Alan V. Oppenheim, Jacek Jachner

This thesis investigates the use of multilevel signal representations, or hierarchies of signal representations with increasing levels of abstraction of detail, to perform efficient signal matching for Vector Quantization in a speech coder. The signal representations range from the numeric sequence that completely characterizes a signal to high-level representations that abstract detail and group signals into broad classes.

The use of abstraction of detail in signal matching has been proposed in the context of a helicopter sound signature detection problem.[9] The current work focuses on low-bit-rate speech coding, for which recently proposed approaches, such as the Code Excited Linear Predictor (CELP) structure,[10] rely on a Vector Quantizer to code the residual signal after linear prediction stages. The residual is matched with one codeword signal out of a codebook of such signals, such that an error criterion between input and codeword signals is minimized. In the CELP coder, the error criterion is a time-varying linearly weighted mean-square error. The signal matching operations required for Vector Quantization (VQ) represent the major source of complexity in the CELP structure, and limit the practical size, hence performance, of VQ codebooks.

This work seeks to reduce matching complexity by using multilevel signal representations simplified by abstraction of detail. The computational benefit of such signal representations is twofold. First, the matching operations on simplified or partial signal representations, termed partial errors, are substantially simpler to perform than applying the error criterion to the complete signal representation. Next, the simplified signal rep-

---

9 E. Milios, *Signal Processing and Interpretation Using Multilevel Signal Abstractions*. Ph.D. diss., Dept. of Electr. Eng. and Comp. Sci., MIT, 1986.

10 M. Schroeder and B. Atal, "Code-Excited Linear Prediction (CELP): High-Quality Speech at Very Low Bit Rates," International Conference on Acoustics, Speech and Signal Processing, 1985.

resentations partition the codebook into equivalence classes of codewords that share a partial signal representation value. The partial error is the same for all codewords in a class and needs to be computed only once.

When seeking the optimal codeword that minimizes the error criterion, the partial errors provide sufficient information to eliminate some codewords from further consideration. Multiple signal representations with progressively greater representation detail are used for matching, and a branch-and-bound procedure[11] prunes the set of eligible codewords.

In situations where a non-optimal codeword is still acceptable, the "best" codeword is selected using only the simple partial errors. This yields significant reductions in complexity. The cost in performance is minimized by the adaptive selection of the partial representations used to match each input signal.

The formulation of suitable multilevel representations and the efficient evaluation of partial error functions for VQ in the CELP coder are the main directions of current work. Structured codebooks which simplify the adaptive multilevel matching procedure are being investigated. Performance simulations are conducted using the SPLICE[12] signal processing environment on LISP machines.

1.12 Application of Nonparametric Statistics in Narrowband Signal Detection

Sponsors

National Science Foundation
(Grants ECS 84-07285
and MIP 87-14969)
Sanders Associates, Inc.
U.S. Navy - Office of Naval Research
(Contract N00014-81-K-0742)

Project Staff

Professor Alan V. Oppenheim, Tae H. Joo

The problem of detecting a narrowband signal in wideband noise occurs in many different contexts. This research is motivated by the problem of detecting a periodic gravitational wave signal which is generated by the orbiting motion of binary stars, for example. Because of the relative motion between the gravitational wave source and the receiver, the received data are modeled as a nonlinear, frequency-modulated signal in additive noise.

When the received data are processed as multidimensional data, the periodic gravitational wave detection problem becomes the problem of detecting a sinusoidal signal in wideband noise using multiple observations. Using the time-frequency relationship of the signal specified by the frequency-modulation structure, the spectra can combined to detect the signal. A critical component of this detection problem is the requirement that the sinusoidal signal must exist in all channels. Because local disturbances can exist in real-world channels, the detector should decide that the received data contain a gravitational wave only when the sinusoidal signal exists in all channels.

If the sinusoidal signal exists in all channels, the optimal detector computes the average of the periodograms then threshold-tests the average. However, a shortcoming in using the average as the detection statistic is that only the total energy is tested. Hence, a strong signal which is present in one channel can exceed the threshold to result in a false alarm. To rectify this incorrect decision, the existence of the signal in all channels must be explicitly incorporated into a detection statistic. Using nonparametric statistics, a new detection statistic is developed which measures the coherence between channels to enforce the signal existence requirement. This detection statistic will be tested using

---

[11] E. Lawler and D. Wood, "Branch-and-Bound Methods: A Survey," *Oper. Res.*, 14(4):699-719, 1966.

[12] C. Meyers, *Signal Representation for Symbolic and Numeric Processing.* Ph.D. diss., Dept. of Electr. Eng. and Comp. Sci., MIT, 1986.

the gravitational wave data which are measured by the MIT 1.5 meter prototype interferometric gravitational wave detector.

1.13 Image Texture Modelling

Sponsors

National Science Foundation
(Grants ECS 84-07285
and MIP 87-14969)
U.S. Navy - Office of Naval Research
(Contract N00014-81-K-0742)

Project Staff

Professor Jae S. Lim, Rosalind W. Picard

Textured regions are the nemesis of many image processing algorithms. For example, algorithms for image segmentation or image compression often rely on assumptions of stationarity or high correlation of the image gray-level data. These assumptions generally fail in textured regions. In this research, we will look for a model for the textures which allows them to be synthesized with only a small number of parameters. Such a model would have applications for image coding, segmentation, and classification problems.

One desirable property of a texture model is an ability to synthesize syntactically regular "structural" textures, e.g., a tiled wall, as well as the more random "stochastic" textures, e.g., shrubbery. Most models assume only one of these cases. In this research, we examine the abilities of the Gibbs distribution (used in statistical mechanics to model intermolecular interactions on a lattice) to model both structural and stochastic textures. In particular, we will analyze the effects of different synthesis methods on the textures generated. The purpose of this analysis is to try to understand how structure can be incorporated into the random model. In addition to an analysis of the Gibbs distribution, we will explore some ideas for a new model.

The goal of this research is to develop a model which is capable of synthesizing a continuum of structural and stochastic textures. Although analysis is a major part of the modeling problem, the focus is on synthesis to see if one can characterize the texture completely.

1.14 Structure Driven Scheduling of DSP and Linear Algebra Problems

Sponsors

U.S. Navy - Office of Naval Research
(Contract N00014-81-K-0742)
OKI Semiconductor, Inc.
Sanders Associates, Inc.

Project Staff

Professor Bruce R. Musicus, G.N. Srinivasa Prasanna

In this thesis,[13] we explore issues related to the mapping of linear algebra and digital signal processing problems onto multiprocessors. The highly regular structure, high computation, and data-independent control of these problems makes them ideally suited for semi-automatic compilation onto multiprocessors. The large number of known, well-structured algorithms provides a ready knowledge base for the compiler.

The system to be developed, KALPANA (imagination), has knowledge of a variety of different algorithms, and maps (schedules) these different algorithms to the given architectures. For example, several good algorithms are known for matrix products. A problem involving a matrix product would exploit one of these. Solutions with optimal speed and efficiency are sought. Search complexity will be minimized by exploiting

[13] C. Myers, *Signal Representation for Symbolic and Numeric Processing.* Ph.D. diss., Dept. of Electr. Eng. and Comp. Sci., MIT, 1986.; G.N.S. Prasanna, *Structure Driven Scheduling of Linear Algebra and Digital Signal Processing Problems* Ph.D. diss. proposal, Dept. of Electr. Eng. and Comp. Sci., MIT, 1988.; M.M. Covell, *Representation and Manipulation of Signal Processing Knowledge and Expressions*, Ph.D. diss. proposal, Dept. of Electr. Eng. and Comp. Sci., MIT, 1987.

the problem structure, and choosing the simplest architectures — fully-connected multiprocessors. Low-level details like processor timing are also targeted to be handled. Note that the algorithm can change drastically as processor timing changes.

Exploiting the structure of the problem to derive a good algorithm and its implementation (Structure Driven Scheduling) is the common thread throughout this thesis. For example, an N by N matrix-matrix multiply can be broken up into N instances of matrix-vector multiplications. If schedules for matrix-vector multiplication are known, then we obtain an algorithm for the matrix product. If not, we can further split each matrix-vector multiply into N dot-products. If the schedules for dot-products on the specified architecture are known, then we obtain another algorithm for computing the matrix-matrix multiply. Note that the structure of the problem has enabled us to derive optimal solutions from the vast number of possibilities, without encountering NP-completeness or worse. Currently, the problem breakup is prespecified.

Note that the layout of the two matrices across the processors has to appropriate for a good schedule. One matrix has to be transposed relative to the other, to minimize communication. We have derived a general technique for optimal data layout for matrix expressions involving additions and multiplications. A dynamic programming approach to handle non-square matrices has also been explored.

The problems being targeted include matrix products and the Fourier Transform. Both are very important basic non-recursive numeric algorithms. Extensions to optimally-scheduled block diagrams containing these algorithms as basic operators are also being done.

1.15 The Application of Function Approximation Using Generalized Polynomials to the Development of Optimum Windows

Sponsors

National Science Foundation Fellowship
National Science Foundation
(Grant MIP 87-14969)
U.S. Navy - Office of Naval Research
(Contract N00014-81-K-0742)

Project Staff

Professor Alan V. Oppenheim, James C. Preisig

The windowing of a time series of uniformly spaced samples to reduce spectral leakage, the spatial domain equivalent of windowing a uniform linear array to reduce leakage between far-field sources at different bearings, and the design of LTI FIR linear phase filters are problems whose optimum (in a min-max sense) solutions are well-understood and are based upon the approximation of functions on intervals (possibly disjoint) of the real line using trigonometric polynomials. The purpose of this work is to extend and apply the theory of function approximation using generalized polynomials to handle approximation on closed sets of multidimensional space using polynomial coefficients which may be real, complex or restricted to a finite set. Applications include windowing of arrays for near-field beamforming, windowing of time series with non-uniform samples, windowing of non-uniform and multidimensional arrays, and mask design for x-ray lithography.

1.16 Iterative Algorithms for Parameter Estimation from Incomplete Data and Their Applications to Signal Processing

Sponsors

Tel Aviv University, Department of Electronic Systems
Sanders Associates, Inc.
U.S. Navy - Office of Naval Research (Contract N00014-85-K-0272)[14]
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)

Project Staff

Dr. Ehud Weinstein

The Estimate-Maximize (EM) algorithm is an iterative method for finding the Maximum Likelihood (ML) or the Maximum-A-Posteriori (MAP) parameter estimates given incomplete data. The method has been found useful in a variety of signal processing problems including medical imaging, parameter estimation of superimposed signals, speech enhancement in multiple microphone environment, and signal reconstruction from spatial information.

We have developed an extension of the EM algorithm, termed the Cascade EM (CEM) algorithm, that is found useful in accelerating the rate of convergence of the algorithm, and in simplifying the computations involved.[15]

Using the EM and the CEM algorithms, we have considered the problem of simultaneous state estimation and parameter identification in linear dynamical continuous/discrete systems. The main result here is that one can use the Kalman smoothing equations for ML identification of the system parameters. We also develop a new method for calculating the log-likelihood gradient (score), the Hessian, and the Fisher's Information Matrix (FIM), that can be used for efficient implementation of gradient-search algorithms, and for assessing the mean square accuracy of the ML parameter estimates.[16]

We have also developed a computationally efficient algorithm for estimating the spatial and spectral parameters of multiple source signals using radar/sonar arrays. The most attractive feature of the proposed algorithm is that it decouples the spatial parameter optimization from the spectral optimization, leading to a significant reduction in the computations involved.[17]

1.17 Equalization (Identification) of Non-Minimum Phase Systems

Sponsors

Tel Aviv University, Department of Electronic Systems,
U.S. Navy - Office of Naval Research (Contract N00014-85-K-0272)[18]
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)

Project Staff

Dr. Ehud Weinstein[19]

The problem considered here is the following: we observe the output of a discrete time-invariant linear possibly non-minimum phase system H with input being a realization (sample function) from a stochastic process. We want to recover the input signal or equiv-

[14] Administered through the Woods Hole Oceanographic Institution.

[15] Ibid.

[16] Research developed with Mordechai Segal, Tel Aviv University.

[17] Research developed with Mordechai Segal and Bruce Musicus.

[18] Administered through the Woods Hole Oceanographic Institution.

[19] Research developed with Ofir Shalvi, Tel Aviv University

alently to identify the magnitude and phase of the inverse of H using a tap-delay line (Equalizer). This problem referred to as self-recovering or blind equalization is of prime interest in data communication, acoustical and geophysical signal processing.

We have developed necessary and sufficient conditions for equalization. Based on that, we propose several equalization criteria, and prove that their solution corresponds to the desired response. These criteria are universal in the sense that they do not impose any restrictions on the probability law of the input (unobserved) process. The resulting algorithms only involve the computation of a few moments of the system output, implying a simple tap update procedure.

1.18 Fractals and Chaos for Digital Signal Processing

Sponsors

Natural Sciences and Engineering
Research Council of Canada,
Science and Engineering Scholarship
National Science Foundation
(Grants ECS 84-07285
and MIP 87-14969)
U.S. Navy - Office of Naval Research
(Contract N00014-81-K-0742)

Project Staff

Professor Alan V. Oppenheim, Gregory W. Wornell

Generally, the objectives of this research are to apply non-linear dynamical system theory and fractal process theory to problems of interest to the signal processing community.

The first aspect of this research involves the use of self-similar or fractal signal models (e.g., 1/f) in solving signal processing problems. The ubiquity of self-similar phenomena and structures in nature suggests that these models may arise rather naturally in the efficient representation of information in various contexts. Novel spread spectrum techniques based on similar formulations are also being considered, as are techniques for modeling signals through the use of iterated cellular autonoma.

A second aspect of this research is aimed at exploiting the emerging theory of chaotic dynamical systems in problems pertaining to digital signal processing. Chaotic systems are characterized by exponential sensitivity to initial conditions, so that the long-term state evolution of such systems is essentially unpredictable. Consequently, chaotic dynamical systems are deterministic systems of arbitrarily low order capable of exhibiting highly complex and stochastic behavior. In some implementation scenarios, accumulator overflow effects have been shown to cause digital filters to exhibit such chaotic behavior. Present work is aimed at identifying and characterizing chaotic dynamics at work in other digital filtering scenarios involving quantization and overflow effects.