24. Digital Signal Processing

Academic and Research Staff

Prof. A.V. Oppenheim, Prof. A.B. Baggeroer, Prof. J.S. Lim, Prof. B.R. Musicus, Dr. D.R. Mook, Dr. J.C. Lee

Graduate Students

T.E. Bordley, P. Chan, S.R. Curtis, D.S. Deadrick, W.P. Dove, F.U. Dowla, M. Feder, D.W. Griffin, W.A. Harrison, W. Huang, D. Izraelevitz, P. Kenny, D.M. Martinez, E.E. Milios, C. Myers, T.N. Pappas, M.D. Richard, H. Sekiguchi, R. Sundaram, P.L. Van Hove, M.S. Wengrovitz, A. Zakhor

24.1 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 Lincoln Laboratory, and with the 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 which 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 since 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 it runs on. 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 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. Recently, based on a general class of algorithms involving the estimation of a signal after its short-time Fourier transform has been modified, we have developed a very high quality system for time-scale modification of speech and also a system for very reliable pitch detection. We are also exploring new techniques for speech enhancement using adaptive noise cancelling when multiple microphones are available. Speech processing is also the basis for two projects on knowledge-based signal processing. One combines signal processing with the concepts of artificial intelligence and expert systems in the dement of a knowledge-based pitch detector, and the second the development of a system for knowledge-based speech modeling and enhancement.

In image processing, we are pursuing a number of projects on restoration, enhancement, and coding. One project is restoration of images degraded by additive noise, multiplicative noise, and convolutional noise. Our current work in this project involves development of a new approach to adaptive image restoration based on a cascade of one-dimensional adaptive filters. This approach, when applied to some existing image restoration systems, not only reduces the number of computations involved but improves the system performance. Another project 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 studying involves the transformation of an image into a one-bit intensity image by adaptive thresholding and then coding the one-bit intensity image by run-length coding. In addition, we are currently working on algorithm development for detection and identification of artificial objects in infrared radar images. In this project, we hope to incorporate both signal processing and artificial intelligence techniques, and hope to exploit both the intensity and range map information.

In the area of geophysical signal processing, there are a variety of ongoing and new projects. Work on determining characteristics of the ocean bottom by inverting a measured acoustic field is continuing, and associated with this, we continue to explore algorithms for evaluating Hankel transforms, and are investigating new field migration algorithms. Also, there are a number of projects related to a series of major experiments in the eastern Arctic in 1980 and 1982, and to a series of experiments in the marginal ice zone of the Arctic in 1983 and 1984. These projects include examining the properties of several velocity function inversion techniques for multi-channel ocean-bottom interaction data and extending the parabolic wave equation approximation for modeling underwater acoustics. Also of potential importance to geophysical signal processing is our work on knowledge-based signal processing with distributed sensor nets, being carried out in collaboration with Lincoln Laboratory.

There are also 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. This work has also been extended to reconstruction of multidimensional signals from one bit of phase, and, exploiting duality zero-crossing and threshold crossing information. 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. We have also obtained significant results and continue to explore new algorithms for high-resolution multidimensional spectral estimation. Recently, we have also 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.

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.

24.2 Improved Paraxial Methods for Modeling Underwater Acoustic Propagation

U.S. Navy - Office of Naval Research (Contracts N00014-81-K-0742 and N00014-77-C-0266)
Arthur B. Baggeroer, Thomas E. Bordley

Current paraxial techniques for modeling wave propagation in the ocean perform poorly when bottom interaction is important, due to their inability to handle wide-angle propagation. Such methods solve a simplified form of the wave equation, obtained by splitting the field into transmitted and reflected components, and then neglecting the reflected field. In spite of their poor performance, the fundamental assumption underlying these techniques remains true: the energy transmitted forward greatly dominates the energy reflected back; so modeling the transmitted field alone should suffice. The concern of this work is to develop paraxial methods less restricted in angle, yet retaining the numerical stability and efficiency which have made this class of solutions so valuable. The approach taken is to generalize the notion of a split to allow variation in range and depth, and to define a measure of the consistency of the split in term of the observed field alone. The algorithm studied generates the field iteratively using the current estimate of the field to generate the split for the next step by requiring that the new split be consistent with the observed field. The effectiveness of this approach is examined in the context of modeling sound propagation in the Arctic Marginal Ice Zone.

24.3 Signal Reconstruction from Fourier Sign Information

U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
National Science Foundation (Grants ECS80-07102 and ECS84-07285)
Alan V. Oppenheim, Susan R. Curtis

In a variety of applications, it is necessary or desirable to recover a signal from the knowledge of the location of its zero crossings or from sign information. One such application occurs when a signal is clipped or otherwise distorted in such a way as to preserve the zero crossing or level crossing information, and it is desired to recover the original signal. Another application occurs in the theory of vision where studies have stressed the importance of edge detection as a means of classifying and identifying images but have not succeeded in developing a strong theoretical basis for this work. A third application occurs in some design problems where one might want to specify a filter response or antenna pattern in terms of zero crossing or null points (such as for interpolation) and derive the remainder of the response from these.

We have developed new theoretical results on reconstruction of multidimensional signals from sign information in either the Fourier or signal domain. Specifically, we have developed a set of conditions under which a broad class of multidimensional signals are uniquely specified by the sign of the real part of the Fourier transform. The problem of reconstruction from the zero crossings of the original signal is a dual to this problem since it involves recovering a signal from sign information in the signal domain rather than in the Fourier domain. We have also extended these results to permit reconstruction from crossings of an arbitrary threshold rather than just zero crossings.

In addition to developing conditions under which a signal is uniquely specified with zero crossings, we have also worked on developing practical algorithms for reconstructing signals from this information.

We have applied two algorithms previously used in other signal reconstruction problems to the problem of reconstruction from zero crossings. One of these algorithms involves solving a set of linear equations and the other involves an iterative procedure imposing different constraints in the space and frequency domains. We have successfully recovered a number of example images from zero crossings but more research is needed in this area.

We are also interested in finding new applications of our results. It is tempting to consider the application of these results to developing a low bit rate coding scheme which might involve transmitting the sign information alone or the zero crossing locations and recovering the original signal from these. In practice, however, recovery of the original signal is very sensitive to knowing the exact zero crossing locations and it is difficult to achieve good results even with very small errors in the zero crossing locations. We are currently exploring other applications of our results, including filter design and sampling theory.

24.4 Knowledge-Based Pitch Detection

Amoco Foundation Fellowship
National Science Foundation (Grants ECS80-07102 and ECS84-07285)
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
Sanders Associates, Inc.
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 are 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.

This project is developing 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, Englewood Cliffs, New Jersey 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.

24.5 Bearing Estimation of Wideband Signals by Multidimensional Spectral Analysis

U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)

National Science Foundation (Grant ECS80-07102)

Jae S. Lim, Farid U. Dowla

Directions of a stationary and homogeneous wavefield is conventionally determined by the frequency-wavenumber method. While the frequency-wavenumber method is useful in many applications, for bearing estimation of wideband planewaves of unknown temporal spectra and velocities, the computation of the frequency-wavenumber method may be unwarranted in view of the proposed spatial approach. The proposed approach is based on estimating the zero-delay covariance function, the time-space covariance function of the wavefield evaluated only at the zero temporal lag. Just as the cross-spectral covariance function determines a frequency-wavenumber spectrum, the zero-delay covariance function determines a zero-delay wavenumber spectrum, the spatial spectrum of the wavefield. The underlying theoretical basis of the proposed approach is that for a planar array the zero-delay wavenumber spectrum can be used to determine the bearings of the planewave sources. Unlike the frequency-wavenumber method which attempts to resolve sources at a single frequency, the proposed method attempts to resolve multiple sources at different frequencies in the spatial spectrum at once. Both the frequency-wavenumber and the proposed method utilize multidimensional spectral estimation algorithms to obtain the spatial spectrum. For a finite aperture array, the operation of a high-resolution wavenumber estimation method is important for separating the planewave sources in bearing. A new high-resolution spectral estimation method with useful computational property is introduced. This algorithm may be used either with the zero-delay covariance function or with the cross-spectral covariance function.

This work was completed in November.

24.6 The Estimate-Maximize(E-M) Algorithm and its Application to Signal Processing Problems

National Science Foundation (Grant ECS84-07285)
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
Alan V. Oppenheim, Meir Feder

In many signal processing problems one has to estimate unknown parameters. Usually this is done using the Maximum Likelihood (ML) criterion. Unfortunately, sometimes the parameters are connected to the observed signal in a complicated way and so maximizing the likelihood function is not an easy task.

The Estimate-Maximize (E-M) algorithm,[1] is an iterative algorithm that converges to the ML solution. We assume that there is an extension of the observations for which the ML estimate of the parameters is simple. We will call that extension "complete data". At each iteration we first estimate the sufficient statistics of the complete data given the observed signal and the previous values of the parameters (Estimate E-Step) and then get the new values of the parameters by maximizing the likelihood function of the complete data where we substitute the estimated statistics instead of observed sufficient statistics (Maximize M-step).

Some examples of signal processing problems for which this idea is natural are:

- Parametric spectrum estimation from short data - the complete data will be a large data record.
- Signal parameters estimation when the signal is corrupted by noise - the complete data will be the signal alone and the noise alone.
- Array processing problems :

  (i) multiple sources problem - the complete data will be each source separately

  (ii) multipath problem - the complete data will be each path separately

  (iii) time delay estimation -- the complete data will be a large record for which the generalized cross correlator is optimal.

In this research we will try, on one hand, to find general properties of the E-M algorithm like convergence rate, adaptive versions, deterministic versions, etc., and on the other hand, to carry out the algorithm experimentally on some of the above problems and to establish this algorithm as a procedure for solving those signal processing problems.

References

1. A.P. Dempster *et al.*, "Maximum Likelihood from Incomplete Data via the EM Algorithm," Ann. of Royal Stat. Soc. (1977), pp.1-38.

24.7 Speech Synthesis from Short-Time Fourier Transform Magnitude Derived from Speech Model Parameters

National Science Foundation (Grants ECS80-07102 and ECS84-07285)

Sanders Associates, Inc.

U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)

Jae S. Lim, Daniel W. Griffin

Previous approaches to the problem of synthesis from estimated speech model parameters have primarily employed time-domain techniques. Most of these methods generate an excitation signal from the estimated pitch track. Then, this excitation signal is filtered with a time-varying filter obtained from estimates of the vocal tract response. This approach tends to produce poor quality "synthetic" sounding speech.

In this research, a new approach to speech synthesis from the same estimated speech model parameters is investigated. In this approach, a desired Modified Short-Time Fourier Transform Magnitude (MSTFTM) is derived from the estimated speech model parameters. The speech model parameters used in this approach are the pitch estimate, voiced/unvoiced decision, and the spectral envelope estimate. The desired MSTFTM is the product of a MSTFTM derived from the estimated excitation parameters and a MSTFTM derived from spectral envelope estimates. (The word "modified" is included in MSTFTM in these cases to emphasize that no signal exists, in general, with this MSTFTM.) Then, a signal with Short-Time Fourier Transform Magnitude (STFTM) close to this MSTFTM is estimated using the recently developed LSEE-MSTFTM algorithm.[1,2] Preliminary results indicate that this method is capable of synthesizing very high-quality speech, very close to the original speech.

This method has applications in a number of areas including speech coding and speech enhancement. In speech coding applications, the excitation parameters and spectral envelope can be coded separately to reduce transmission bandwidth. Then, these coded parameters are transmitted and then decoded and recombined at the receiver. In speech enhancement applications, the excitation parameters and the spectral envelope can be separated, processed separately, and then recombined.

References

1. D.W. Griffin and J.S. Lim, "Signal Estimation from Modified Short-Time Fourier Transform," Proceedings of the 1983 Int. Conf. on Acoustics, Speech and Signal Processing, Boston,

Massachusetts, April 14–16, 1983, pp.804–807.
2. D.W. Griffin and J.S. Lim, "Signal Estimation from Modified Short-Time Fourier Transform," IEEE Trans. Acoust., Speech, Signal Process., ASSP-32, 2, 236–243, April 1984.

24.8 Reconstruction of a Two-dimensional Signal from its Fourier Transform Magnitude

National Science Foundation (Grants ECS80-07102 and ECS84-07285)
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
Jae S. Lim, David Izraelevitz

In a wide range of applications, such as radio astronomy, crystallography, and holography, it is of great interest to reconstruct an unknown signal from measurements of the magnitude of its Fourier transform. Some progress has been made in the area in recent years. The uniqueness of the reconstruction has been proven for a large family of signals; specifically, all those two-dimensional signals with a non-factorable z-transform. In addition, several algorithms have been proposed for the reconstruction of the signal. Unfortunately, none of the proposed algorithms are guaranteed to converge to the true solution. It is encouraging to note, however, that much progress has been made lately in the related fields of reconstruction from various amounts of phase and magnitude information.

In this research we are considering several new approaches to the problem of signal reconstruction from Fourier transform magnitude. Our purpose is to incorporate recent research results to develop an algorithm for reconstruction from Fourier transform magnitude which will be guaranteed to yield a solution as long as the underlying signal satisfies the conditions for uniqueness.

24.9 Motion Compensated Noise Reduction for Motion Video Scenes

Advanced Television Research Program
National Science Foundation (Grants ECS80-07102 and ECS84-07285)
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
Jae S. Lim, Dennis M. Martinez

Signals which correspond to motion video scenes possess a great deal of spatial and temporal redundancy. Many image processing algorithms exploit the spatial redundancy in one form or another. However, when dealing with 3-D signals generated from motion video scenes, there is also a significant amount of temporal redundancy among adjacent image frames. In fact, often the only changes in the scene from one frame to the next are those introduced by motion within the scene. By properly identifying the relevant motion trajectories of the objects within a scene, motion information can be used in various image processing problems.

This research is concerned with two fundamental issues. The first issue involves investigating techniques for performing motion estimation given a set of noisy image frames. Basically, the problem is to estimate the displacement field in those regions of the scene where motion is present. The displacement field is a 2-D vector field, which is the function of the two spatial coordinates and the temporal coordinate, that describes the motion trajectories. The second issue is to develop algorithms which can incorporate motion information for the purpose of noise reduction. Filters which incorporate motion information exploit the temporal redundancy of these signals, and should theoretically permit greater signal-to-noise improvements than are possible using only spatial information.

24.10 Knowledge-Based Harmonic Source Detection and Direction Determination

National Science Foundation (Grants ECS80-07102 and ECS84-07285)
Sanders Associates, Inc.
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
Alan V. Oppenheim, Evangelos E. Milios

Architectural and design issues in knowledge-based signal processing are examined in the context of acoustic waveform processing for harmonic source detection and direction determination. The system under construction takes as input a set of noisy acoustic waveforms and produces hypotheses about the existence and direction of harmonic sound sources present in the waveforms. The approach combines signal processing and symbolic information processing using techniques derived from the domain of artificial intelligence and knowledge-based systems. The system uses signal processing operators as probing tools and has the capability of interpreting the results of their application and of constructing concise symbolic descriptions of waveform features. A unique aspect of the system is that it fully integrates signal processing and symbolic processing, as opposed to previous systems, which used signal processing as a simple front end. In the system under construction, the problem of planning the application of a variety of signal processing operators is posed and it is handled by applying knowledge about the operators based on past experience. The mechanism for formalizing the relevant knowledge is that of *protocol analysis*, a technique by which a person with experience in processing acoustic data is given particular problems to solve and all his actions, as well as the related reasoning, are recorded and analyzed. Particular rules are extracted from this analysis, which also dictates the overall system architecture. The system is being implemented in a rule language implemented in LISP and uses a highly modular signal processing package also in LISP.

The work is being done in collaboration with the Distributed Sensor Networks program at M.I.T. Lincoln Laboratory.

24.11 Knowledge-Based Speech Analysis and Enhancement

Amoco Foundation Fellowship
National Science Foundation (Grants ECS80-07102 and ECS84-07285)
Sanders Associates, Inc.
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
Alan V. Oppenheim, Cory Myers

The problems of speech modeling and speech enhancement for speech degraded by noise are ones which have been of great interest in the signal processing community. As part of our efforts to explore methods by which signal processing and symbolic processing techniques can be made to interact, we are building a system for symbolic reasoning about speech analysis and enhancement. For analysis, the system produces a time-varying parametric model of the speech signal where, not only do the parameters change over time, but the models used may also change over time. For example, the system may perform all-pole modeling of the signal but may change the number of poles which are to be used in different portions of the signal. The system attempts to develop a parametric model of a speech signal given both the speech signal (or a noisy version of it) and a symbolic description of how the signal was produced. The symbolic description of the signal includes information about the speaker, such as gender and age, and information about the recording environment, such as noise characterization, sampling rate and speech bandwidth. The system is also given a symbolic description of the content of the signal in the form of a time-aligned phonetic transcription. The output from the analysis portion of the system is a time-varying parametric representation of the speech signal (excluding excitation information).

For speech enhancement the system takes the parametric representation generated from the analysis section and combines it with classical speech synthesis techniques in order to enhance the signal when noise is present. The system combines three different techniques in order to produce an enhanced version of the noisy speech signal. In those regions of the signal where the system can be expected to do a reasonable job of speech modeling in spite of noise and where the measured parameters are reasonable in the context of the known symbolic information the system does resynthesis from the measured parameters. In those regions of the signal where the system fails completely to measure reasonable speech parameters the system uses a phoneme to speech synthesis system. In those regions where some measurements can be made reliably (but not all measurements can be made reliably) the system uses the good measurements to infer the unknown information. For example, the system may guess the third formant given the first and second formants and the identity of the phoneme.

The system is structured as an Expert System. This approach to system building emphasizes the knowledge used to solve a problem and tries to make this knowledge explicit in the system. Both phonetic knowledge and signal processing knowledge are represented in this way. The signal processing and symbolic processing interact closely, with the signal processing driving the symbolic processing in some cases and the signal processing being driven by the symbolic processing in other cases.

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

National Science Foundation (Grants ECS80-07102 and ECS84-07285)
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
Jae S. Lim, Thrasyvoulos N. Pappas

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

The techniques for quantifying the stenosis of the coronary arteries are similar to those used for stenosis determination of other arteries like the femoral and carotid. However, due to the location of the coronary arteries on the beating heart and their shape and size, their x–rays are very difficult to obtain and analyze. The performance of existing techniques is not satisfactory when they are applied to coronary arteries.

Coronary artery stenosis is determined as the maximum percent arterial narrowing within a specified length of the vessel. This requires the detection of the boundaries of the coronary arteries and the analysis of the variation of their diameter. Current algorithms find the boundaries of the artery as those points along a series of lines perpendicular to the vessel image where the film density change is maximum. We will consider a different approach. We will construct a parametric model of the film density along the lines perpendicular to the vessel image, and based on the observed density values we will estimate the parameters at each perpendicular line along the vessel. The parameters of our model will include the vessel diameter and centerpoint and will account for the structure of the background (e.g., variation in soft tissue thickness, presence of bones, the diaphragm, etc.) as well as the distortions introduced by the imaging system.

Since our estimation procedure will use more information about the blood vessels than just the maximum slope points of the film density at each perpendicular line, we expect it to have a better performance than the present methods. Preliminary results show that the estimation of the blood vessel boundaries based on such a parametric model is at least as good as that of the other methods.

The continuity of the vessel parameters as we move from line to line, which follows from the spatial continuity of the vessel, will also be incorporated into our model and is expected to significantly improve the accuracy of our estimation procedure.

We plan to test our algorithm both on real coronary angiograms and on x-rays of contrast medium filled cylindrical phantoms obtained over a wide range of radiographic conditions. The latter will provide information about the distortion introduced by the imaging system, and will enable us to compare our computer-derived measurements of the phantom diameters with their known values.

24.13 Low Bit Rate Video Conferencing

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

The attempt of this research is to achieve a drastic reduction in the channel capacity requirements for video transmissions. An algorithm based on local averaging and adaptive thresholding has been developed which reduces an 8-bit picture to a 1-bit binary picture without significant loss in quality. The resulting image is then smoothed by adaptive median filtering to remove random bit transitions. This is then suitable for run length coding. Compression factors of up to 60 are expected to result. Other forms of data encryption are also being examined. Currently the emphasis is only on intraframe redundancy reduction. The work is expected to lead to interframe analysis.

24.14 The Curvature Spherical Image, a New Representation of 3-D Objects

National Science Foundation (Grants ECS80-07102 and ECS84-07285)
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
Jae S. Lim, Patrick Van Hove

This research addresses the relationships between an object and the boundaries of its silhouettes, which will be referred to as contours, corresponding to various three-dimensional (3-D) orientations of the line of sight. For this purpose, special models of objects and silhouettes are considered. The property sphere of an object is defined on a unit sphere which is related to the object by a 3-D Gaussian mapping. Similarly, the property circle is related to the contour it represents by a 2-D Gaussian mapping. In earlier computer vision work, property spheres and circles have been used independently, and their applications have been limited to the representation of scalar fields of object properties.

In a first stage, we have shown how the concepts of property spheres and circles can be usefully combined to relate the properties of an object and those of its contours. Specifically, it is proved that a slice through the property sphere of an object leads to the property circle of the contour corresponding to the line of sight perpendicular to that slice.

In a second stage, a new concept of object modeling has been introduced, where the property sphere is used as a domain for vector and tensor fields of object properties. In particular, a new representation of 3-D objects, referred to as the Curvature Spherical Image (CSI), maps the curvature tensor field of the surface of an object on its property sphere. The key advantage of this representation is that a slice through the CSI and a subsequent projection of the tensor field onto the slice produces the Curvature Circular Image (CCI) of the contour corresponding to the line of sight perpendicular to the slice. The CCI itself is a mapping of the radius of curvature scalar field of a contour on its property circle.

The study now progresses with attempts to use these new concepts in the matching of 2-D contour measurements with 3-D object models. An issue which will be addressed in the course of the research is the modelling of convex objects. The CSI, like the other spherical representations of 3-D objects, is defined for convex objects only. However, many of the concepts defined mathematically for convex objects are believed to produce good heuristics for non-convex objects.

This work is being done in collaboration with the Distributed Sensor Systems group at M.I.T. Lincoln Laboratory.

24.15 Shallow Water Acoustic Inversion

Hertz Foundation Fellowship
National Science Foundation (Grants ECS80-07102 and ECS84-07285)
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
Alan V. Oppenheim, George V. Frisk[28], Michael S. Wengrovitz

The problem of continuous-wave acoustic signal propagation in shallow water is being investigated. In this environment, components of the acoustic field alternately reflect off both the

[28]Woods Hole Oceanographic Institution

ocean surface and the ocean bottom. In effect, the water can be considered to be an acoustic waveguide bounded by the ocean surface and the underlying ocean bottom. Several aspects of this waveguide propagation problem are being studied. The first concerns the development of an accurate numerical model to predict the magnitude and phase of the acoustic field as a function of range from the source given the geoacoustic parameters of the water and the bottom. A technique has been developed which computes the field based on its decomposition into trapped (resonant) and continuum (non-resonant) components.

A second aspect being studied is the inverse problem. Here, features of the waveguide and the ocean bottom are desired and measurements of the field over an aperture are given. A parametric modeling technique has been proposed which reduces the effect of the inherent windowing and improves the estimate of the bottom reflection coefficient. Another inversion technique has also been developed which directly computes the sound velocity in the ocean bottom half-space given measurements of the acoustic field. This technique is based on accurately determining the branch point of a multivalued function.

The contribution of a branch point is an important component of the signal being processed in this problem. Current research is directed towards the more general digital signal processing problem of pole/zero/branch-point modeling.

24.16 Computing the Discrete Hartley Transform

Hertz Foundation Fellowship
National Science Foundation (Grants ECS80-07102 and ECS84-07285)
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
Alan V. Oppenheim, Avideh Zakhor

The discrete Hartley transform (DHT) resembles the discrete Fourier transform (DFT) but is free from two characteristics of the DFT that are sometimes computationally undesirable. The inverse discrete Hartley transform is identical to the direct transform and so it is not necessary to keep track of +i and –i versions as with the discrete Fourier transform. Also, the DHT has real rather than complex values and thus does not require provision for complex arithmetic or separately managed storage for real and imaginary parts. The DFT is directly obtainable from the DHT by a simple additive operation.

We have found an efficient way of computing the Hartley transform which enables us to compute the DFT with half as many multiplications as the FFT. Wang[1] and Bracewell[2] have also found algorithms for the fast Hartley transform. In this research, we will evaluate and compare these three algorithms in terms of their relative efficiencies. Statistical and deterministic error properties of these algorithms will also be investigated both theoretically and experimentally.

References

1. Z. Wang, "Fast Algorithms for the Discrete W Transform and for the Discrete Fourier Transform," IEEE Trans. Acoust. Speech. Signal Process., ASSP-32, 4, August 1984.
2. R.N. Bracewell, "The Fast Hartley Transform," Proc. IEEE, 72, 8, August 1984.