XXI. DIGITAL SIGNAL PROCESSING

Academic Research Staff

Prof. Alan V. Oppenheim
Prof. Arthur B. Baggeroer
Prof. James H. McClellan

Graduate Students

Bir Bhanu
Thomas E. Bordley
Patrick W. Bosshart
David S. K. Chan
Gregory L. Duckworth
Mark G. Hall
David B. Harris
Samuel Holtzman
Richard B. Kline
Stephen W. Lang
David C. LeDoux
Steven J. Leverette
Jae S. Lim
David R. Martinez
Thomas L. Marzetta
Syed H. Nawab
Robert W. Patterson
Michael R. Portnoff
Thomas F. Quatieri, Jr.
Antonio Ruiz
Kenneth B. Theriault
José M. Tribolet
Pirooz Vatan

1. HOMOMORPHIC SPEECH ANALYSIS-SYNTHESIS

U. S. Navy – Office of Naval Research (Contract N00014-75-C-0951)

Thomas F. Quatieri, Jr., Alan V. Oppenheim, Antonio Ruiz

a. Simulation of a Homomorphic Vocoder Based on Charged Coupled Device (CCD) Technology

We have completed simulations of various speech analysis-synthesis configurations based on both the conventional chirp-z-transform (CZT) realization of the discrete Fourier transform and the sliding CZT realization of the discrete sliding Fourier transform. These realizations are amenable to CCD technology and allow for real-time, low-cost implementation of the homomorphic vocoder.

A comparative study was performed, illustrating the tradeoffs between synthetic speech quality and implementational complexity for the two schemes.

b. Quality Improvement

Techniques for synthetic speech quality improvement were tested and evaluated in collaboration with studies of the Speech Group at Lincoln Laboratory. Issues such as interpolation, amplitude measurements, buzziness, hoarseness, and coding were explored. For male speech, formal listening tests indicate that for low bit rates (2400 to 3600 bps) the homomorphic system is comparable in quality to more established schemes such as LPC and the channel vocoder. For high bit rates (8000 to 9600 bps), the homomorphic system was judged to have the highest quality.

The female synthetic speech, on the other hand, unlike that of the male, tends in general to be degraded by a "hoarseness." We are investigating ways of rigorously characterizing this degradation and are exploring adaptive techniques for improvement.

We shall soon be completing a pitch-synchronous complex cepstral vocoder. This scheme, we hope, will yield an understanding of the effects of phase of the estimated vocal tract impulse response on synthetic speech quality.

2. ENHANCEMENT OF DEGRADED SPEECH

U. S. Navy – Office of Naval Research (Contract N00014-75-C-0951)

Jae S. Lim, Alan V. Oppenheim

Continuing our work on enhancing degraded speech, we have attempted to develop a complete analysis/synthesis system in which the synthesis parameters are estimated from noisy speech data. The particular analysis/synthesis system we have considered is based on an all-pole model of speech. Our approach has been to apply a Maximum A Posteriori (MAP) estimation procedure in estimating the coefficient vector of an all-pole system from noisy speech accounting for the presence of noise. In general, a MAP estimation procedure for noisy speech leads to solving a set of nonlinear equations. Two suboptimal procedures which require solving only sets of linear equations have, however, been developed. These methods have been applied to both synthetic and real speech data with white Gaussian background noise, and our preliminary listening test indicates that both systems are capable of significant noise reduction. We are now engaged in a more formal subjective test which is directed toward evaluating the two linear systems in terms of their performance in enhancing speech intelligibility and quality when the background noise is of various different spectra.

3. TIME-VARYING LINEAR PREDICTIVE CODING OF SPEECH SIGNALS

U. S. Navy – Office of Naval Research (Contract N00014-75-C-0951)

Mark G. Hall, Alan V. Oppenheim, Alan S. Willsky
[Prof. Willsky is Assistant Director of the Electronic Systems Laboratory, M. I. T.]

During this year we have completed a project involved with time-varying predictive coding of speech. The project involved a generalization of linear prediction using time-varying coefficients. By representing each time-varying coefficient, either in terms of a power series, or in terms of a Fourier series, a set of equations to determine the coefficients was obtained. These coefficients are reminiscent of those in the time-invariant case in that they are block-symmetric or block-Toeplitz. The basic problem can either be formulated in a covariance form or in a correlation form and the relative characteristics of these two approaches were explored. Through a study of a number of synthetic examples and examples using real data, we concluded that the covariance form with a power series representation for the coefficients was the most preferable and that this approach has the potential for representing a long nonstationary segment of speech with fewer total coefficients than would be required through the use of time-invariant LPC in which an analysis window is moved through the data.

4. DIGITAL SEISMIC SIGNAL PROCESSING

U. S. Navy – Office of Naval Research (Contract N00014-75-C-0951)
National Science Foundation (Grant ENG76-24117)

David B. Harris

A seismic surveying technique called wave equation migration is being investigated for possible applications of two-dimensional digital signal processing algorithms. The key component of the migration algorithm is a difference equation approximation to the wave equation. This difference equation is used to extrapolate a wave field recorded on the boundary of a region backwards into the region. An ideal transfer function for the two-dimensional difference equation can be derived from the wave equation. Currently, methods are being sought to approximate this transfer function which is all-pass with a specified phase.

5. TIME SCALE MODIFICATIONS OF SPEECH SIGNALS

Michael R. Portnoff

The objective of our research in this area is to modify a speech signal in such a manner that the resulting signal is perceived as identical to the original except for its rate of articulation. In particular, we seek to preserve such qualities as naturalness, intelligibility, and speaker-dependent features, while avoiding the introduction of such objectionable artifacts as "glitches," "burbles," and reverberation often present in vocoded speech.

We have developed and demonstrated a high-quality system for time-scale compression and expansion of speech based on short-time Fourier analysis.[1,2] This system is capable of compressing speech by ratios as large as 3:1 and expanding speech by arbitrarily large ratios. Furthermore, the performance of this system does not appear to be sensitive to the presence of broadband noise in the speech source material.

References

1. M. R. Portnoff, "Implementation of the Digital Phase Vocoder Using the Fast Fourier Transform," IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. ASSP-24, No. 3, pp. 245-248, June 1976.

2. M. R. Portnoff, "A Mathematical Framework for Time-Scale Modification of Speech Signals," Ninety-third Meeting, Acoustical Society of America, State College, Pennsylvania, June 7-10, 1977. (Abstract in J. Acoust. Soc. Am., Vol. 61, Suppl. No. 1, Spring 1977, p. S68.)

6. ALGORITHMS FOR HIGHLY PARALLEL COMPUTER STRUCTURES

National Aeronautics and Space Administration (Grant NSG-5157)

James H. McClellan, David C. LeDoux

In order to process photographic images from Earth observation satellites rapidly, computers with a very high degree of parallel processing capability are being designed. One such machine[1] would be composed of a minimum of 16, 384 very simple processors arranged in a 128 × 128 array. Each processing element contains a bit-serial adder with a small amount of memory and is able to transfer data, one bit at a time, to any of its four nearest neighbors. A central control unit issues an instruction to the entire array and the individual processors decide whether or not to execute it, depending on a mask bit in each processor. In use, each processor operates on one pixel in a sampled photo and the parallel processing allows great time savings over more traditional serial computers.

A frequent operation in satellite image processing is that of image registration which involves computing the cross-correlation between two images of the same scene. Correlations may be computed indirectly and sometimes more efficiently by using two-dimensional transforms such as the Fourier transform or the Fermat number transform (FNT). The computation of the FNT does not involve multiplications or complex numbers, and is well suited to the simplicity of the small processors. We have investigated the use of the FNT to perform correlation on a highly parallel computer developed by NASA. Programs have been writen and timing estimates have been obtained. The FNT can be more efficient than direct computation of the correlation (on this machine), depending on the sizes of the images used.

We are now investigating the applicability of the parallel computer structure to several new image restoration algorithms, particularly those which attempt to correct for the effects of the Earth's turbulent atmosphere.

References

1. L. W. Fung, MPPC: A Massively Parallel Processing Computer, NASA Goddard Space Flight Center, September 1976.

7. WINOGRAD FOURIER TRANSFORM ALGORITHM (WFTA) IMPLEMENTATION

National Aeronautics and Space Administration (Grant NSG-5157)

Syed H. Nawab

This research is concerned with some implementation issues for the Winograd Fourier Transform Algorithm. Although this DFT computation algorithm has fewer multiplies than the FFT, speed comparisons between these algorithms requires a further investigation of computation structures used in the implementations.

Attention has been focused on two-address general register machines and computation structures with parallelism. For the register machines, a register allocation study has been carried out to determine the relationship between algorithm speed and execution times for the machine instructions. In the parallelism study, existing FFT structures have been generalized to yield WFTA parallel structures. Speed comparisons for these parallel structures are now being made.

8. IMPLEMENTATION OF MULTIDIMENSIONAL DISCRETE SYSTEMS FOR SIGNAL PROCESSING

Joint Services Electronics Program (Contract DAAB07-76-C-1400)

David S. K. Chan, James H. McClellan

This research seeks to develop understanding of the implementation of multidimensional discrete systems by studying the properties of these systems and developing techniques for the synthesis of efficient implementations. Such understanding will lead to more systematic methods of analysis and synthesis of these systems and better utilization of modern resources for multidimensional signal processing.

Our research has established a framework for studying the implementation of a very general class of multidimensional discrete systems, called oriented discrete systems,[1] which is much broader than the class of first-quadrant causal systems that have generally been considered. It constitutes the practical, rather than analytical, generalization of one-dimensional causal systems to multidimensions. We have shown that, by broadening our perspective to include this class of systems, realizations having smaller dimension and requiring less overall memory than under first-quadrant causality can be obtained in certain cases.[1]

Our research has also dealt with structures for the realization of linear shift-invariant multidimensional systems, not only in the first-quadrant causal case,[2,3] but also in the more general case. Using a framework we developed for the characterization of realization structures,[4,5] we have extended the technique of continuously equivalent analog networks to the minimization of coefficient sensitivity and roundoff noise in multidimensional digital networks under structure transformation. The refinement of these techniques is now under way.

Finally, we have looked into block implementations of multidimensional systems using fast convolution algorithms, and implementation using simple interconnections of multiple processors.

References

1. D. S. K. Chan, "Multidimensional Oriented Discrete Systems – Characterization and Realization," Eleventh Annual Asilomar Conference on Circuits, Systems, and Computers, Monterey, California, November 7-9, 1977.

2. J. H. McClellan and D. S. K. Chan, "A 2-D FIR Filter Structure Derived from the Chebyshev Recursion," IEEE Trans. on Circuits and Systems, Vol. CAS-24, No. 7, pp. 372-378, July 1977.

3. D. S. K. Chan, "A Simple Derivation of Minimal and Near-Minimal Realization of 2-D Transfer Functions," to appear in Proc. IEEE (Correspondence).

4. D. S. K. Chan, "A Novel Framework for the Description of Realization Structures for 1-D and 2-D Digital Filters," EASCON '76 Record, pp. 157-A-157-H.

5. D. S. K. Chan, "A Unified Framework for the Realization Problem in Linear Digital Filtering," Record of the 1977 International Conference on Acoustics, Speech, and Signal Processing, Hartford, Connecticut, May 9-11, 1977, pp. 543-547.

9. ALGORITHM FOR COMPUTATION OF THE ACOUSTIC PLANE-WAVE REFLECTION COEFFICIENT OF THE OCEAN BOTTOM

U.S. Navy – Office of Naval Research (Contract N00014-77-C-0196), Woods Hole Oceanographic Institution

David R. Martinez, Arthur B. Baggeroer, Alan V. Oppenheim, George V. Frisk

[George V. Frisk is with the Woods Hole Oceanographic Institution.]

For a spherical acoustic wave incident on the ocean bottom, the reflected pressure field and the plane-wave reflection coefficient are related through a two-dimensional spatial-wave number Fourier transform.

$$R(k_x, k_y) = -\frac{i\beta\, e^{-i\beta_1(z+z_o)}}{2\pi} \iint_{-\infty}^{\infty} \Phi_n(x, y, z)\, e^{-i[k_x x + k_y y]}\, dxdy.$$

The objective of this project is to calculate the plane-wave reflection coefficient R(k) from samples of the reflected pressure field. An algorithm based on the "projection-slice" theorem[1] associated with the two-dimensional Fourier transform has been implemented on an HP 2100 at Woods Hole Oceanographic Institution. We have investigated the effectiveness of the algorithm for the perfectly reflecting hard bottom $R(k) = 1$, and it has yielded accurate results.

References

1. R. M. Mersereau and A. V. Oppenheim, "Digital Reconstruction of Multidimensional Signals from Their Projections," Proc. IEEE 62, 1319-1338 (1974).

10. ACCURACY BOUNDS FOR NORMAL INCIDENCE ACOUSTIC STRUCTURE ESTIMATION

U.S. Navy – Office of Naval Research (Contract N00014-75-C-0852), Department of Ocean Engineering, M. I. T.

Arthur B. Baggeroer, Kenneth B. Theriault

Kenneth B. Theriault submitted a doctoral thesis in August, 1977 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Electrical Engineering and Computer Science, M. I. T., and the Woods Hole Oceanographic Institution. The thesis analyzed how accurately one can estimate acoustic structure with signals at normal incidence in the presence of noise. Several specific models and acoustic structures were analyzed which indicated the effects of reverberation and source energy distribution. Results indicate that very high signal-to-noise ratios (>40 dB) are required to obtain accurate results.

11. DATA ADAPTIVE VELOCITY/DEPTH SPECTRA ESTIMATION IN SEISMIC WIDE-ANGLE REFLECTION ANALYSIS

National Science Foundation subcontract to Grant GX 41962 to Woods Hole Oceanographic Institution

Arthur B. Baggeroer, Steven J. Leverette

Steven J. Leverette submitted a doctoral thesis in August, 1977 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Electrical Engineering and Computer Science, M. I. T., awarded jointly with the Woods Hole Oceanographic Institution. His thesis dealt with the application of the Maximum Likelihood Method (MLM) of spectral analysis to estimating the velocity structure of the earth in wide-angle seismic reflection. The effects of the data windowing and the statistics of the estimator for a small number of observations were studied. The algorithm was applied to both simulated and field data, which demonstrated the superiority of the method over more conventional ones. Figure XXI-1 indicates a comparison of spectra estimated

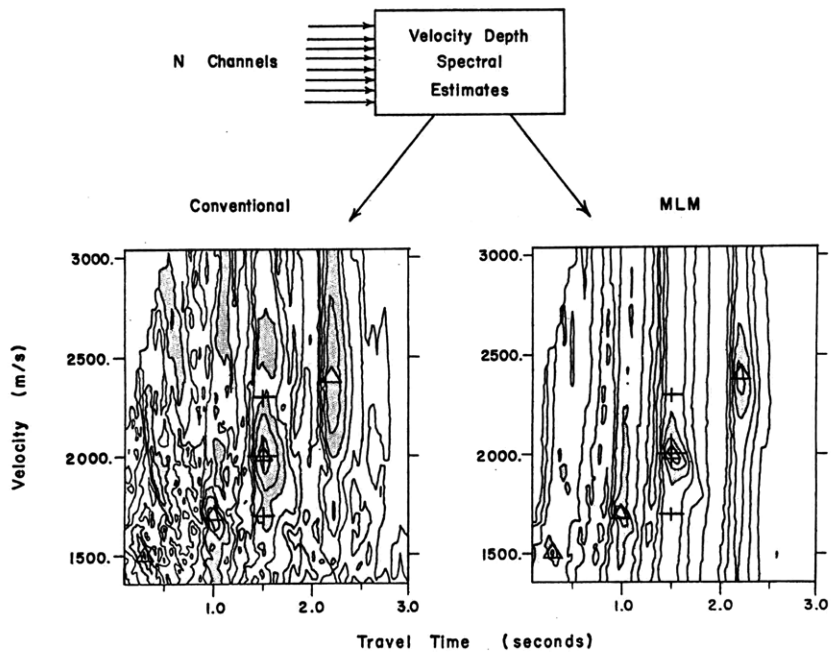

Fig. XXI-1. Comparison of velocity depth spectra for conventional and maximum-likelihood processing.

by using a conventional coherence criterion with spectra using the MLM algorithm.

12. A LINEAR PREDICTIVE APPROACH TO TWO-DIMENSIONAL SPECTRAL FACTORIZATION AND SPECTRAL ESTIMATION

Thomas L. Marzetta

Some recent results in the theory of time-series linear prediction have been extended to the two-dimensional case. In particular, an equivalence has been established among three separate domains: a class of two-dimensional analytic and positive-definite autocorrelation sequences, the class of two-dimensional minimum-phase prediction error filters and positive prediction error variances, and a class of two-dimensional reflection coefficient sequences and positive prediction error variances.

The two-dimensional reflection coefficient representation is the basis of a new approach to two-dimensional spectral factorization and autoregressive model fitting in which the prediction error filter is designed in the reflection coefficient domain. By constraining the reflection coefficient magnitudes to be less than one, the difficult minimum-phase requirement is automatically satisfied.

The remaining practical question concerns how to choose the reflection coefficients in an "optimal" way. For the spectral factorization problem a convenient (but generally suboptimal) method involves sequentially choosing the reflection coefficients to minimize the mean-square prediction error at each stage of the procedure. A program has been written to implement this algorithm and has been tested for two examples. In the first example, the spectrum was relatively smooth, and the algorithm performed well. In the second example, the spectrum was discontinuous, and the performance of the algorithm was less satisfactory. It is believed that the full potential of this new approach can only be realized by using a more sophisticated gradient algorithm for choosing the reflection coefficients.