XXI. COGNITIVE INFORMATION PROCESSING*

Academic and Research Staff

Prof. M. Eden
Prof. J. Allen
Prof. B. A. Blesser
Prof. T. S. Huang
Prof. F. F. Lee
Prof. S. J. Mason
Prof. W. F. Schreiber
Prof. D. E. Troxel
Prof. I. T. Young
Dr. R. R. Archer
Dr. G. H. Granlund
Dr. E. G. Guttmann
Dr. L. P. A. Henckels
Dr. K. R. Ingham
Dr. J. I. Makhoul
Dr. D. M. Ozonoff
Dr. O. J. Tretiak
F. X. Carroll
Deborah A. Finkel
E. R. Jensen
J. W. Modestino
C. Nehring
D. E. Robinson
Sandra A. Sommers
J. S. Ventura

Graduate Students

B. S. Barbay
J. E. Bowie
B. E. Boyle
R. L. Brenner
J. R. Ellis, Jr.
I. S. Englander
A. M. Fakhr
A. E. Filip
W. B. Grossman
D. W. Hartman
M. Hubelbank
F. H. Ives
J. W. Klovstad
H. S. Magnuski
L. A. Marks
P. L. Miller
B. Minkow
O. R. Mitchell, Jr.
J. A. Myers
R. A. Piankian
G. Poonen
D. G. Sitler
J. R. Sloan
A. A. Smith
R. D. Solomon
M. Sporer
J. H. Stahler
W. W. Stallings
W. D. Stratton
H-m. D. Toong
K. P. Wacks
Y. D. Willems
K. H. Yung

RESEARCH OBJECTIVES AND SUMMARY OF RESEARCH

The Cognitive Information Processing Group works on problems related to the analysis, manipulation, and simulation of the sense data used in certain human perceptual and cognitive processes. Our major interest is increased understanding of visual processes, in particular, the ways by which humans can recognize and interpret pictorial information. Thus we do research on electronic methods for simplifying pictures, without altering significantly human subjective judgments of picture quality, on methods for altering pictures so as to enhance particular features and remove others, on the way humans perform sophisticated visual tasks, such as reading and writing, and on the way humans learn to recognize complicated patterns in microscopic images or photographs and categorize them without conscious effort.

These studies have a variety of applications, including computer input-output devices, communication systems, sensory replacement, and many others. The studies of language and picture processing have suggested ways to substitute other senses for sight, so that a blind person can "read" printed material in ways that are not too different from the ways sighted persons use such information sources. Image processing and pattern recognition studies are being applied to the classification of white cells in blood smears, to the detection of malignant cells in Papanicolau smears, to the diagnosis of blood dyscrasias by measurements on erythrocytes in smears, to the enhancement of x-ray images before radiological diagnosis.

*This work was supported principally by the National Institutes of Health (Grants 5 PO1 GM14940-05 and 3 PO1 GM15006-03S2), and in part by the Joint Services Electronics Programs (U.S. Army, U.S. Navy, and U.S. Air Force) under Contract DAAB07-71-C-0300), the World Health Organization (Grant R/00348), and a Grant from the Associated Press.

The mathematical methods, as well as the computer algorithms and technical processes developed within communication science, have proved relevant to a wide range of sensory and cognitive processes. Our research efforts have contributed toward the development of coherent theories of image processing and pattern recognition. We have found, for example, that research results, and even computer programs, relevant to speech analysis have a straightforward connection with research on the visualization of biological tissues. Thus our many discrete research activities share two unifying attributes: a common set of mathematical analytic techniques, and a core data-processing facility specifically created for research on two-dimensional signal processing and pattern recognition.

B. A. Blesser and D. M. Ozonoff have developed a procedure for discovering perceptual errors in the reading of medical x-rays when these have been made with and without techniques such as image enhancement.

O. J. Tretiak has developed a method for obtaining the internal structure of arbitrary objects through computer processing of multiple x-ray images, and this procedure is being applied to the reconstruction of skeletal bone revealed by radiation therapy.

I. T. Young has supervised the construction of a microscope with attached vidicon scanner so that large populations of blood cells may be studied in order to classify anatomical features and to study cell properties such as contact adhesiveness. Similarly, G. H. Granlund has been studying techniques for automatically classifying chromosomes by means of new staining techniques which make it possible to trace distinct contour patterns of the chromosomes when ultraviolet light projects on the differentially stained cells. J. E. Green has examined several thousand red blood cells from 6 normal subjects in order to obtain feature values needed by recognition algorithms for these cells. He has also begun to develop automatic methods for detecting malarial parasites in blood films. D. E. Robinson has applied digital processing to overcome some of the basic limitations of the pulse-echo technique used in the increasingly important area of ultrasonic techniques for medical diagnosis.

W. F. Schreiber has been working on a digital facsimile transmission system for newspaper use, with possible eventual application to x-ray transmission. The enhancement of real-life degraded images has been studied by T. S. Huang, and has resulted in techniques for estimating the characteristics of degrading systems, as well as a method for restoring motion-degraded images.

Several projects designed to develop aids for the handicapped have been completed this year. F. F. Lee has successfully applied tactile feedback of pitch information to allow deaf subjects to control their voices in an acceptable way, while D. E. Troxel has continued to develop techniques for TV scan acquisition of characters from printed text, as well as a model for the computer generation of handprinted characters. K. R. Ingham has implemented a time-shared computer system for the use of the blind which permits keyboard input and provides compressed digitally stored speech as output.

J. Allen has continued the development of a system for conversion of unrestricted English text to speech. He has used the Brown Corpus of 38,000 words obtained from a million-word random sample of text. These words were decomposed into 15,000 constituent morphs which were then merged with pronunciations and parts of speech from the Merriam Pocket Dictionary to obtain a complete lexicon suitable for speech synthesis.

Brief descriptions of the principal projects now under way, and some thoughts on plans for future projects follow.

1. Image Compression Techniques

The reduction of the number of bits required to transmit a visual image is a continuing goal. Coupled with such reduction is the attendant increase in subjective degradation of these images. We are studying visual noise perception in such pictures with controlled spatio-temporal spectra.

In the field of coherent optics, we have generated holograms of two-dimensional objects by computer, and are now attempting computer-generated holograms of three-dimensional objects.

We are now turning our attention to much higher quality images such as are encountered in x-rays and facsimile transmission. Results are expected to be somewhat different from those previously obtained, since it is well known that many aspects of image processing are strongly affected by resolution. We are interfacing a high-resolution facsimile transmitter and receiver with the PDP-9 computer, and plan to study various aspects of image quality and coding efficiency, such as optimum linear and nonlinear filtering, coding of both color and monochrome images by digital and analog means, and ways to transmit such images with a high degree of control over their properties.

2. Image Enhancement and Transformation

Up to the present time, most of our work in this area has been devoted to analyzing the mathematical and technical problems encountered in dealing with visual images degraded by a variety of physical processes. Potential applications for the techniques and processes under study include many situations in which visual images must be analyzed and interpreted. Initial applications have been to medical problems.

Although we have done some exploratory work on radiological images, relatively little biomedical work could be justified because some technical problems had to be solved first.

Based on work begun during the past year, increasing attention will be directed toward specific biomedical image studies in the future. Two such projects have been initiated.

a. Computation of x-ray sections from a series of standard films. This technique appears to be of particular importance in therapeutic radiology and our project is being carried on in collaboration with radiologists at the Peter Bent Brigham Hospital, in Boston.

b. Studies of the characteristic image degradations of x-ray camera systems and development of two-dimensional recursive digital filters so as to compensate for certain aberrations by computation.

There are some interesting possibilities that we expect to explore in the future.

a. Extension of the resolution of chromosomal images beyond the diffraction limit of the microscope by using the techniques of analytic continuation.

b. Use of computer processing to correct degradation of electron micrographs resulting from lens defects that induce linear, shift-variant aberrations.

c. Construction of an x-ray transmission system for remote diagnostic examination, where a very high-quality replica is required. (The project will include the design of a laser scanner, a receiver, and display unit with hard copy output and a coding scheme dependent on techniques of bandwidth compression.

3. Pattern Perception and Recognition

Because the group's work deals with visual or auditory signals that have a meaningful interpretation to human observers, we have always had certain relevant psychophysical or psychological projects. We have also carried on studies in the theory of pattern recognition. The applications of pattern-recognition techniques are varied, including new health care methods, oceanographic studies, and automated industrial quality control. At present, the following projects are being pursued.

a. The use of filtered random-noise visual patterns for the measurement of human spatial frequency response. Both psychophysical and objective measurements are made.

b. A study of the utility of a perceptual model for diagnostic radiology (in collaboration with the Department of Radiology, Peter Bent Brigham Hospital). This study is essential to any future work on x-ray image enhancement.

c. A study of the possibility of speech recognition based on the detection of higher order linguistic constructs (syllables and words rather than phonemes). Thus far, this scheme has worked successfully for the recognition of the spoken integers 1 to 999.

We expect these three projects to continue for some time. In particular, the application of image-processing analysis to radiology is expected to require several years of research effort. It is also quite possible that the studies of visual psychophysics may be applicable to the rapid and objective diagnosis of certain visual defects, such as astigmatism, myopia, and hyperopia.

Consideration of problems of biological image processing has given us several interesting ideas about perceptual processes that we intend to study in the next five-year period. They include the following.

a. The visual correlates of texture; that is, what physical signals are interpreted as textures by human observers. A related study concerns human judgment of shape factors.

b. A comparison of the utility of color values (a subjective judgment) with chromaticity as parameters for making discriminations by machine; for example, distinguishing cells by color of nucleus or cytoplasm.

4. Quantitative Microscopy

Major projects in this area will be outlined separately.

a. Quantitative study of cell morphology. We plan to extend our identification algorithms from the five principal types of leukocytes (neutrophil, lymphocyte, monocyte, eosinophil, and basophil) to the next five most common types: atypical lymphocyte, metamyelocyte, band cells, hypersegmented polyocytes, and nucleated red cells. As in previous work, the algorithms for identification of the cells will be developed to measure those properties of the cells observed in light microscopy, that is, nucleus and cytoplasm size, nucleus and cytoplasm color, nucleus shape, and possibly nucleus texture. These 10 cells account for the cells observed on 97% of all slides that we have examined.

We also plan an in-depth study of the five principle types of leukocytes. We propose to perform large population studies of these cell types to understand better their distributional properties and inherent biological variability. For example, we would like to study the distribution of cytoplasm and nucleus size and nucleus shape in a population of 10,000 neutrophils.

b. Study of leukocyte stickiness. Our recent studies have shown that the contact adhesiveness of leukocytes may be measured by preparing blood slides on a special centrifuge and then comparing the number of leukocyte clumps (2 or more leukocytes touching) with a number projected from a stochastic model. The excess of clumps over the projected number is a measure of the cell "stickiness." During the next eighteen months we plan a thorough clinical study, with the help of Dr. Robert Colman of the Hematology Unit of Massachusetts General Hospital to calibrate our measure and determine the nature of the difference between leukemic and nonleukemic cell adhesiveness. Also included, of necessity, will be a study of the effects of anticoagulating agents and preparation times.

c. Distributional property of cell hemoglobin. Using techniques that we have already developed for the quantitative analysis of red cell morphology, we plan to study the distribution of hemoglobin per cell per patient in a large cell population. At present, the only quantitative information at our disposal is the mean corpuscular hemoglobin (MCH) and the mean corpuscular hemoglobin concentration (MCHC), both measured as

the ratio of extremely global parameters. Our techniques yield the actual distribution of corpuscular hemoglobin so that more than just the mean value might be available. Such distributional information would be extremely valuable in evaluating the effects of therapy, for example, in anemias; or when the distribution itself is not unimodal, for example, in thalassemia.

d. Quantitative evaluation of blood stains. The use of Wright's stain for the preparation of peripheral blood specimens leaves much to be desired. We have been examining alternative staining preparations in the Romanowsky series in order to improve our morphological measurements in both consistency and cell discrimination. We are now continuing this work with the idea of incorporating in the staining procedure one of the enzyme stains reported by Dr. L. T. Yam of the Tufts New England Medical Center. We have consulted with Dr. Yam on this problem and it seems that a combination of a Romanowsky stain and a nonspecific esterase will provide an excellent preparation for cell differentiation.

5. Sensory Aids

Our activities here center around the development of a reading machine for the blind, with major emphasis on text-to-speech conversion. We have several independent projects.

a. Developing hardware and algorithms for more reliable and more flexible character recognition.

b. Constructing a dictionary for generating a large set of words, developing letter-to-sound rules, building a parser, and devising rules to establish the acoustic correlates of speech from the written text alone.

c. Developing better, more intelligible, pleasant synthetic speech (in conjunction with the Speech Communication Group of the Research Laboratory of Electronics).

d. Developing an inexpensive minicomputer-based Braille translation facility.

During the coming years, we anticipate that the direction of this work will change in the following ways.

a. More blind subjects will be involved in test situations designed to evaluate the acceptability and usefulness of the reading machine and related devices to a blind user.

b. The linguistic aspects of the program will be extended to research in speech recognition, to speech training for deaf subjects, and to studies of the reading-writing process, e.g., design of systems of potential value in the training of dyslexics.

c. Another possibility is the development of sensory aids directed toward improving professional skills of the handicapped. (The system that K. R. Ingham has developed in our group for training blind programmers is an example.)

d. A more recent objective is to develop devices and training methods that will enable those who are born deaf or with severe hearing loss to learn to speak so that they can be more easily understood. It is our belief that a continuous, prolonged usage of a suitable tactile feedback apparatus should be introduced to a deaf child for use throughout the language-acquisition phase of his life. Preferably this apparatus can be used both for monitoring important parameters of his own speech, and extracting the same parameters of the speech of others.

We have constructed pitch-to-tactile display equipment and performed preliminary study on the use of such equipment in static pitch control of the deaf. Work is under way on a study of dynamic pitch control.

6. Text-to-Speech Conversion

In the area of text-to-speech conversion, we report several new results. In order to compute the sound symbols (phonemes) corresponding to orthographic words, it is

necessary to have both a morph lexicon with a scheme for decomposing words into morphs, and a set of letter-to-sound rules for converting words which cannot be treated lexically to phoneme strings. We have completed the construction of a 13,000 morph lexicon from the 40,000 word Brown Corpus, and have made several improvements to the word-decomposition program. A complete letter-to-sound algorithm has also been developed, so that low-frequency English words can be converted directly to phoneme strings without lexical look-up. Further work will result in a smaller dictionary as more bound morphemes are recognized, so that redundant words can be removed from the lexicon.

In order to synthesize speech properly, an accurate vocal-tract model must be controlled with appropriate parameters. Segmental controls have now been chosen so that high-quality speech at the phonemic level is possible. At present, emphasis is being placed on prosodic parameters that reflect stress, intonation, and pauses. Vowel durations as a correlate of stress have been studied, and consonant cluster durations are now under investigation. We are also studying fundamental frequency contours that reflect stress and intonation properly, as well as the relation of pauses to the syntactic structure of the sentence.

We have nearly completed the design for a new all-digital speech synthesizer. This vocal-tract model will be built from TTL logic, and enable us to achieve reliable yet flexible synthesis. Present estimates show that a complete sample-point iteration of the model should take only 30 ms, which is well within the 100-kHz sampling rate. Since the model is effectively a high-speed special-purpose computer, it can be reprogrammed (by a new read-only memory) to simulate other vocal-tract models. Thus, while it accommodates the current model satisfactorily, there is ample flexibility, both in instructions and time, to incorporate the results of future refinements in vocal-tract models.

In the related field of speech recognition, several projects are under way. In one study we have employed the fundamental frequency and energy envelope waveforms to locate syllable boundaries, so that phonemic rules can be more effectively employed. In another study we are developing local parsing strategies centered around stressed content words which can then be expanded to predict poorly articulated function words by means of a grammar network. Finally, a dictionary is under development that will permit nearest-neighbor look-up so that the effects of wrong, too few, and too many phonemes can be corrected automatically.

M. Eden, J. Allen, B. A. Blesser, T. S. Huang,
F. F. Lee, S. J. Mason, D. M. Ozonoff,
W. F. Schreiber, O. J. Tretiak,
D. E. Troxel, I. T. Young

A. SCANNING WITH A MOVING MIRROR

Recently, in attempting to make a laser beam scan a straight line by means of a rotating mirror, we discovered or rediscovered a law of geometrical optics: when a ray is reflected from a mirror, the intersection of the reflected ray and a plane is generally a conic section, as shown in Fig. XXI-1. There are at least two special cases, shown in Figs. XXI-2 and XXI-3, which do produce exactly straight-line scans. These arrangements have been used in existing scanners, but each has certain drawbacks. The most convenient optical arrangement for many cases is that of Fig. XXI-4, where the incident beam makes a small angle with respect to that plane that is the locus of the reflected rays produced as the mirror turns.

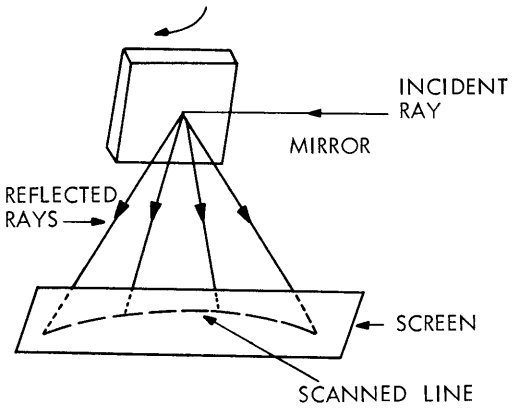

Fig. XXI-1.

General case of ray reflected from mirror.

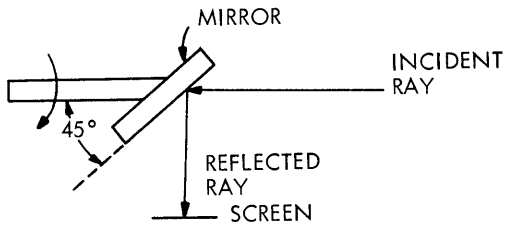

Fig. XXI-2.

Mirror, screen, and scanned line perpendicular to paper. Axis of mirror rotation is in the plane of the paper, parallel to incident ray.

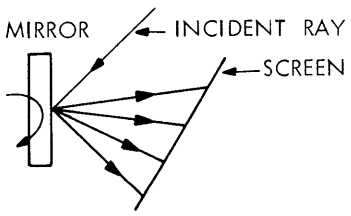

Fig. XXI-3.

Mirror, screen, and axis of rotation are perpendicular to paper. Incident and reflected rays and scanned line are in the plane of the paper.

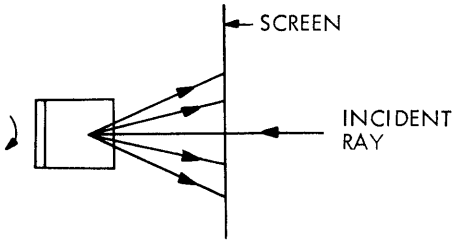

Fig. XXI-4.

Screen perpendicular to paper. Reflected rays are in the plane of the paper. Incident ray and mirror are tilted.

We proposed to straighten the hyperbola produced by the last arrangement, by introducing a countereffect that results from tilting the mirror on its axis. Some crude experiments quickly demonstrated that this was possible, at least approximately. For the range of angles used, we found that the condition for straightness was that the angle between the mirror plane and its axis of rotation be equal to the angle between the incident and reflected rays.

After this experimental result, two independent analytical solutions of the problem were found. We present here the more general solution, which has been numerically verified.

We have constructed several scanners on this principle, using total scan angles of more than 45°, and mirror tilts up to 15°. In all cases, the scanned line can be straightened, by adjustment of the mirror tilt, so that we have been unable to measure the error.

1. Experimental Arrangement and Problem

The scanner arrangement is shown in Fig. XXI-5. Both the mirror and the screen are perpendicular to the plane of the paper. The angle of oscillation for this position of the mirror is $\theta = 0$.

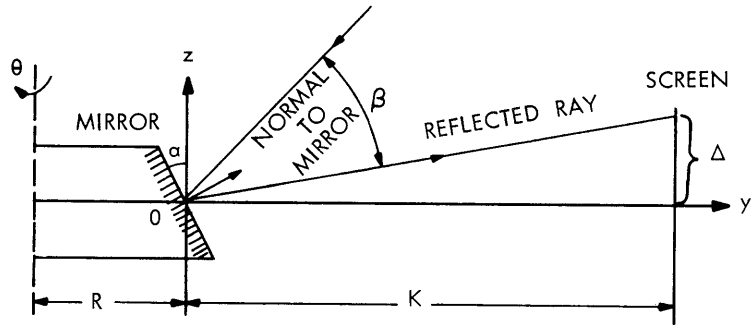

Fig. XXI-5.
Experimental arrangement.

Let Δ be the z coordinate of the point of intersection of the reflected ray and the screen. We are interested in finding values of α and β for which Δ is almost a constant independent of θ.

2. Principal Result

$$\Delta = \frac{\left[K - \frac{\cos\alpha(1-\cos\theta)}{\cos\alpha\cos\theta + \sin\alpha\tan\beta} R\right]\left[-\sin 2\alpha\cos\beta\cos\theta + \sin\beta\cos 2\alpha\right]}{\left[\cos\beta(1-2\cos^2\alpha\cos^2\theta) - \sin 2\alpha\sin\beta\cos\theta\right]} + \frac{\tan\beta\cos\alpha(1-\cos\theta)}{\cos\alpha\cos\theta + \sin\alpha\tan\beta} R. \quad (1)$$

Note that when $\theta = 0$,

$$\Delta\Big|_{\theta=0} = K \tan (2\alpha-\beta). \tag{2}$$

3. Special Case

When $R = 0$, or $R \ll K$, and α, β are small,

$$\Delta = -K\beta \frac{1 - 2(\alpha/\beta) \cos \theta}{1 - 2 \cos^2 \theta}. \tag{3}$$

If we make the assumption that the angle of oscillation, θ, is small, then

$$\Delta = \text{constant}, \qquad \text{for } \frac{\alpha}{\beta} = \frac{2}{3}. \tag{4}$$

This condition is shown in Fig. XXI-6.

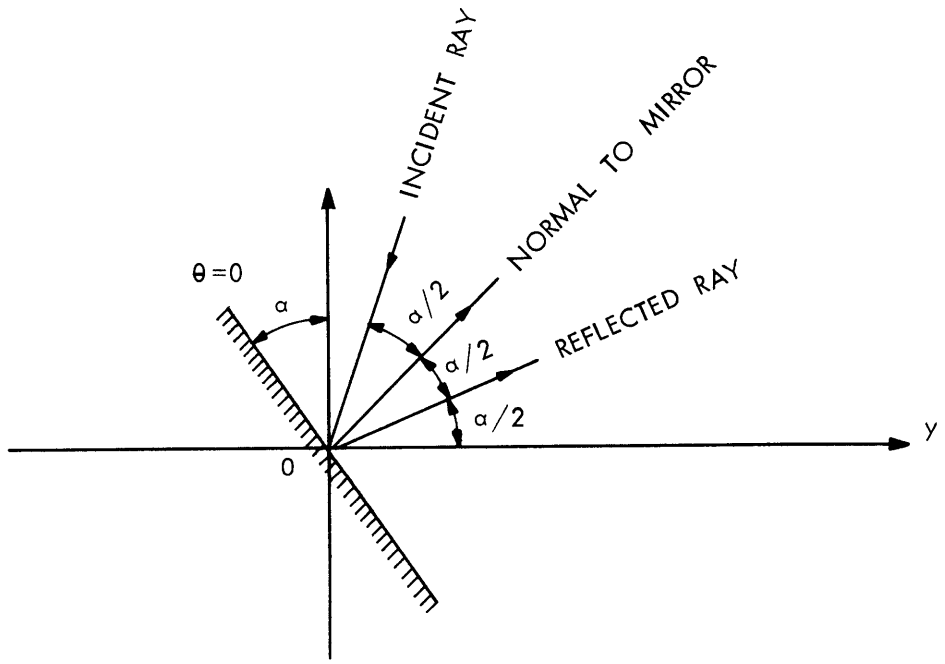

Fig. XXI-6. Special case.

4. Computer Results

Since the angle of oscillation, θ, may not be very small in practice, a computer program was written to calculate

$$\delta \equiv (\Delta - \Delta|_{\theta=0})/K, \tag{5}$$

where Δ is given by Eq. 3, for various values of α/β, especially around $\frac{\alpha}{\beta} = \frac{2}{3}$. We found that for $\beta = 10°$, the value $\frac{\alpha}{\beta} = 0.67$ gives the smallest deviation, δ.

For K = 10", the range $\frac{\alpha}{\beta} = .66\text{-}.68$ insures that $(\Delta - \Delta|_{\theta=0})$ is smaller than .003", for a maximum angle of oscillation $\theta_{max} = 15°$. In terms of α, the best value is $\alpha = 6.7°$. Tolerance in α is $\pm.1°$.

T. S. Huang, O. J. Tretiak, W. F. Schreiber

B. A NOISE MODEL FOR VIDEO PREAMPLIFIERS

In electro-optical communication systems, we often must design low-noise amplifiers for the output of a photosensor such as a silicon photodiode. In this report we develop a simple noise model of a FET-input preamplifier that has been shown experimentally to give reasonable results.

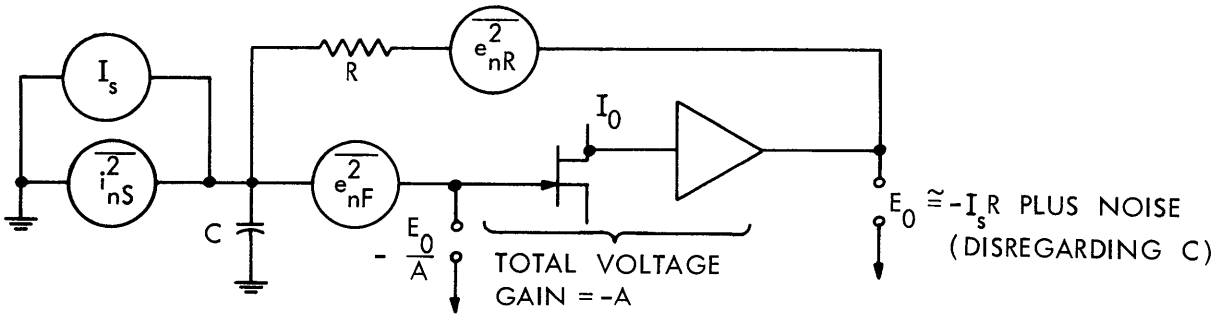

Fig. XXI-7. Model for a video preamplifier, showing the principal noise sources.

We make the following assumptions.

1. The signal is a current source, whose noise component $\overline{i_{nS}^2}$ is that of shot noise only.

2. The feedback resistor is a perfect resistance, whose noise, $\overline{e_{nR}^2}$ is Johnson noise only.

3. The amplifier is flat in the band of interest, with zero Z_{out}, infinite Z_{in}, whose noise is equivalent to the shot noise of a current, I_o. Input device transconductance is g_m, and $A \gg 1$.

We first convert each noise source into an equivalent voltage or current, and then calculate its contribution to output noise.

1. Shot Noise of the Signal

$$\overline{i_{nS}^2} = 2eI_sF = \int_0^F G\,df$$

$G = 2eI_s$ (units of current^2/cycle).

Capital letters refer to complex amplitudes of sine waves analyzing the circuit configuration shown in Fig. XXI-8.

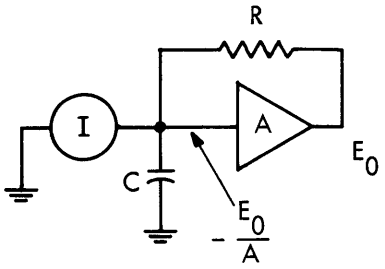

Fig. XXI-8.
Circuit configuration for analyzing the effect of the shot noise of the signal.

$$\frac{E_o + \frac{E_o}{A}}{R} + I = I_c - \frac{E_o}{A} = \frac{I_c}{sC}; \qquad I_c = -\frac{E_o sC}{A}$$

$$\frac{E_o + E_o/A}{R} + I = -\frac{E_o sC}{A}.$$

$$E_o = -\frac{IR}{1 + \frac{1}{A} + \frac{sCR}{A}} \simeq -\frac{IR}{1 + \frac{sRC}{A}}.$$

These relations hold true for any signal or noise current. On a current^2 spectral basis, for a noise current $G = 2eI_s$, where G_o is the output spectrum, and G_i is the input spectrum,

$$G_o = \frac{G_i R^2}{1 + \frac{\omega^2 R^2 C^2}{A^2}} = \frac{2eI_s R^2}{1 + \frac{\omega^2 R^2 C^2}{A^2}}.$$

2. Johnson Noise of the Feedback (Load) Resistor

$$\overline{e_{nR}^2} = 4KTRF$$

$G = 4KTR.$

Using the previous conventions, and referring to the arrangement of Fig. XXI-9, we find that for a noise component E_n,

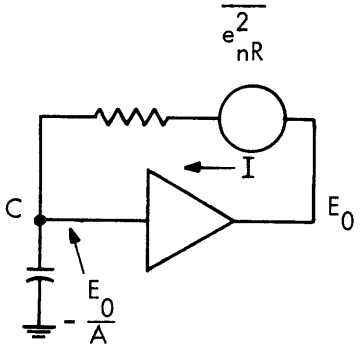

Fig. XXI-9.
Configuration for analyzing the effect of Johnson noise of the feedback resistor.

$$E_n + IR = E_o + \frac{E_o}{A} - \frac{E_o}{A} = \frac{I}{sC}, \qquad I = -\frac{E_o sC}{A}.$$

$$E_o = \frac{E_n}{1 + \frac{1}{A} + \frac{sRC}{A}} \cong \frac{E_n}{1 + \frac{sRC}{A}}.$$

On a spectral basis,

$$G_o = \frac{G_m}{1 + \frac{\omega^2 R^2 C^2}{A^2}} = \frac{4KTR}{1 + \frac{\omega^2 R^2 C^2}{A^2}}.$$

3. Shot Noise of the Input Stage

The noise current in the drain circuit of the first stage is $i_n^2 = 2eI_oF$, where I_o is the drain current. This is equivalent to an input noise voltage

$$\overline{e_{nF}^2} = \frac{2eI_oF}{g_m^2},$$

and on a spectral basis

$$G = \frac{2eI_o}{g_m^2}.$$

In Fig. XXI-10, for any noise component E,

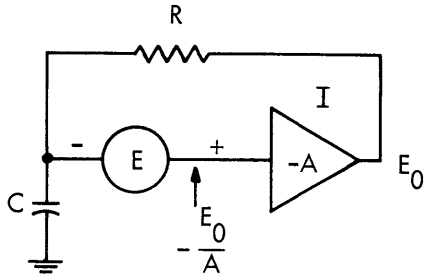

Fig. XXI-10.
Configuration for analyzing the effect of amplifier noise.

$$\frac{E_o + \frac{E_o}{A} + E}{R} = I - \frac{E_o}{A} - E = \frac{I}{sC}, \qquad I = \frac{-sCE_o}{A} - sCE.$$

$$E_o = -E\,\frac{1 + sCR}{1 + \frac{1}{A} + \frac{sCR}{A}} = -E\,\frac{1 + sCR}{1 + \frac{sCR}{A}}.$$

On a spectral basis,

$$G_o = \frac{2eI_o}{g_m^2}\,\frac{1 + \omega^2R^2C^2}{1 + \frac{\omega^2R^2C^2}{A^2}}.$$

If we add these three noises, we get a complete spectrum of output noise,

$$G_{n,o} = \frac{2eI_sR^2}{1 + \frac{\omega^2R^2C^2}{A^2}} + \frac{4KTR}{1 + \frac{\omega^2R^2C^2}{A^2}} + \frac{2eI_o}{g_m^2}\,\frac{1 + \omega^2R^2C^2}{\frac{\omega^2R^2C^2}{A^2}}. \tag{1}$$

All 3 components start decreasing at $\omega = A/RC$, but the last component has a zero also, so it starts rising at $\omega = 1/RC$ and levels off at $\omega = A/RC$. Since the signal starts decreasing at $\omega = A/RC$, the narrow-band SNR is flat until $\omega = 1/RC$, and then decreases continuously for all higher frequencies, the amplifier noise eventually predominating. If we define Operating Spot Noise Figure as input SNR/output SNR (all in volts^2 or current^2 units), we have

$$\text{OSNF} = \frac{\text{SNR input}}{\text{SNR output}} = \frac{S_{in}}{N_{in}} \cdot \frac{N_{out}}{S_{out}}.$$

If we assume that "input noise" is exclusively signal shot noise, we have

$$\text{OSNF} = 1 + \frac{2KT}{e} \cdot \frac{1}{I_sR} + \frac{I_o/I_s}{g_m^2R^2}\,(1 + \omega^2R^2C^2). \tag{2}$$

This result is displayed graphically in Fig. XXI-11, where the output noise spectral density, $i_n^2/\Delta f$, is plotted as a function of frequency for the various noise sources.

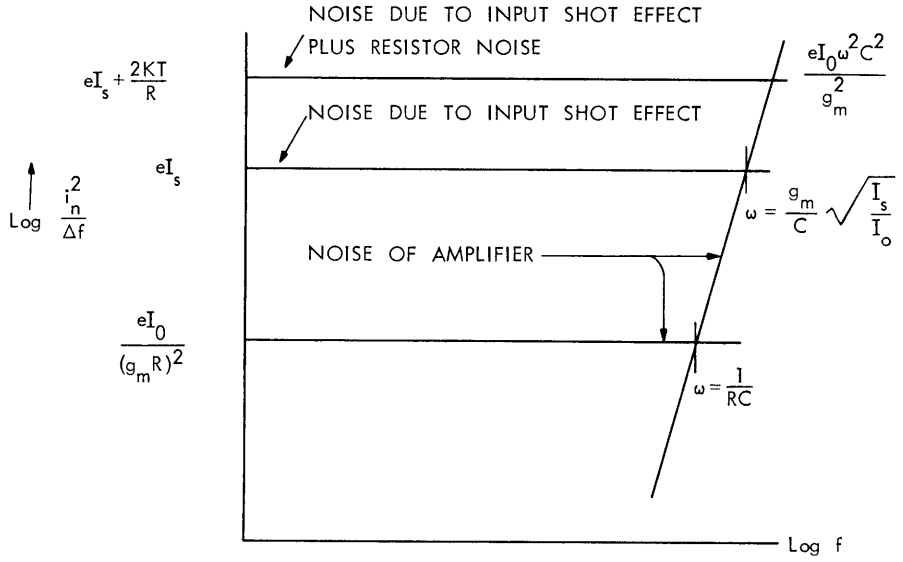

Fig. XXI-11. Graphical display of the output noise spectrum from separate effects.

We are also interested in finding the total noise output. For frequencies below A/RC we can ignore the frequency-dependent term in the denominator of (1) and obtain

$$\overline{e_{n_o}^2} = \int_0^F df\, G = 2eI_sR^2F + 4KTRF + \frac{2eI_o}{g_m^2}\left(F + \frac{(2\pi RC)^2F^3}{3}\right)$$

or

$$\overline{e_{n_o}^2} = 2eFR^2\left[I_s + \frac{2KT}{eR} + \frac{I_o}{(gmR)^2}\right] + \frac{2eI_oF^3}{3g_m^2}(2\pi RC)^2. \qquad (3)$$

Note that the signal

$$e_{s_o}^2 = I_s^2R^2$$

so that the SNR contribution of the last term, which is likely to be dominant at high frequencies, is independent of R.

W. F. Schreiber

C. GENERATION OF MINIMUM-PHASE PICTURES

1. Minimum Phase in Two Dimensions

The motivation for this research stems primarily from curiosity, since no one, until now, has had any idea of what a minimum-phase (MP) picture would look like. The plan was to take an original picture and from it generate another picture whose autocorrelation function (or equivalently, the magnitude of its Fourier transform) is the same as the first, but whose phase is minimum when compared with the phase angle of all pictures having the same autocorrelation function. We are also interested in possible applications in the field of picture bandwidth compression because, for MP signals, the phase angle can be determined from the logarithm of the magnitude of the Fourier transform. We reasoned that if the MP picture closely resembles the original picture, the number of bits needed to transmit the picture could immediately be halved. Unfortunately, MP pictures do not in the least resemble the original image. In fact, that may be the most important finding of this research.

Before discussing the definition of two-dimensional MP functions, we shall review the definitions of MP functions in one dimension. Since all experiments were performed on a digital computer, all of the equations are written in terms of discrete functions and their Z transforms.

There are essentially 3 interchangeable definitions of one-dimensional MP functions, of which the first is probably most basic.

a. Definition 1

A minimum-phase function is that function chosen from the set of all functions having the same autocorrelation function (i.e., the same magnitude of the Fourier transform) which has the minimum phase for all ω.

b. Definition 2

The second definition, stated in terms of the pole-zero pattern of the Z transform of f(n), is

f(n) is minimum-phase if all of its poles and zeros lie inside the unit circle.

c. Definition 3

The third definition, given in terms of the complex cepstrum of f(n), is

f(n) is minimum-phase if its complex cepstrum, $\hat{f}(n)$, is causal, i.e., $\hat{f}(n) = 0$, $n < 0$.

Extending these definitions to two dimensions is not straightforward. While the basic concept of minimum phase given by Definition 1 may be carried over directly to two dimensions, it is not a practical means of testing a function for minimum phase. The

idea of pole-zero patterns (Definition 2) is not useful in two dimensions because the two-dimensional Z transform of f(n, m) no longer has discrete poles and zeros, but rather contains contours of poles and zeros in the $Z_1(Z_2)$ plane, with $Z_2(Z_1)$ treated as a parameter.

Also, the straightforward extension of Definition 3 turns out to be too restrictive, since in two dimensions a function can be minimum-phase even though its complex cepstrum, $\hat{f}(n, m)$, is nonzero for negative values of n or m (but not for both). This can be most easily understood by using the concept of real part sufficiency, which will be discussed after the "working definition" of two-dimensional minimum-phase functions used in this report.

d. Definition 4

A two-dimensional function, f(n, m), is minimum-phase if its complex cepstrum, $\hat{f}(n, m)$, is nonzero only over either of the shaded regions shown in Fig. XXI-12.

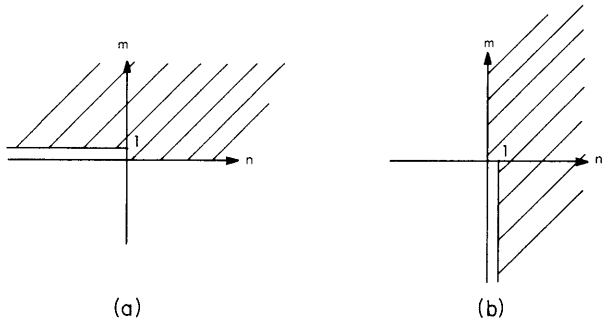

Fig. XXI-12. Allowable nonzero regions for the complex cepstrum of a minimum-phase function.

Another way of saying this is that $\hat{f}(n, m)$ must exhibit real part sufficiency, and its region of support must not include the third quadrant. Notice that this definition includes the case wherein $\hat{f}(n, m)$ is strictly causal.

2. Real Part Sufficiency

A function f(n, m) exhibits real part sufficiency (RPS) if F(u, v) can be determined from Re [F(u, v)], or equivalently, f(n, m) can be determined from Ev [f(n, m)]. We begin by defining F(u, v), the two-dimensional Fourier transform of f(n, m).

$$F(u, v) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} f(n, m) \exp(-jun) \exp(-jvm) \tag{1}$$

$$\text{Re}\,[F(u, v)] = \frac{1}{2}\,[F(u, v) + F^{*}(u, v)]. \tag{2}$$

Taking the inverse transform of the right side of (2), and assuming f(n, m) real, we have

$$\frac{1}{2}\,[f(n, m) + f(-n, -m)] = \text{Ev}\,[f(n, m)]. \tag{3}$$

Now, the question is, Over what region of the n-m plane can f(n, m) be nonzero so that f(n, m) can be recovered from Ev [f(n, m)]? From our experience with one-dimensional functions, we might begin by requiring that f(n, m) be causal. The shaded

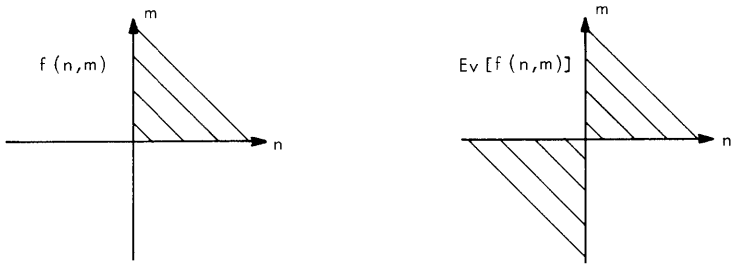

Fig. XXI-13. Nonzero region of a strictly causal function in two dimensions and its even part.

regions of Fig. XXI-13 indicate regions over which the function is nonzero. Clearly, f(n, m) = Ev [f(n, m)] · u(n, m), where

$$u(n, m) = \begin{cases} 2 & n, m > 0 \\ 1 & n, m = 0 \\ 0 & \text{elsewhere} \end{cases} \tag{4}$$

The first quadrant, however, is only a subregion of the allowable nonzero regions. Consider the shaded areas of Fig. XXI-12. If f(n, m) is nonzero over either of these areas, it may be recovered from Ev [f(n, m)] by multiplying the even part by the appropriate function. For Fig. XXI-12b,

$$f(n, m) = \text{Ev}\,[f(n, m)] \cdot u(n, m),$$

$$u(n, m) = \begin{cases} 2 & n \geq 0 \quad m > 0 \\ 2 & n \geq 1 \quad m \leq 0 \\ 1 & n, m = 0 \\ 0 & \text{elsewhere} \end{cases} \tag{5}$$

For Fig. XXI-12a, u(n, m) becomes

$$u(n, m) = \begin{cases} 2 & m \geq 0 \quad n > 0 \\ 2 & m \geq 1 \quad n \leq 0 \\ 1 & n, m = 0 \\ 0 & \text{elsewhere} \end{cases} \tag{6}$$

Notice that the search for allowable nonzero regions excludes the third quadrant. This is because our discussion of real part sufficiency is geared to the finding of minimum-phase functions. It is quite possible for a complex cepstrum $\hat{f}(n, m)$ to exhibit real part sufficiency and be nonzero in the third quadrant. Then, by definition, f(n, m) is a maximum-phase function.

3. Generating the Minimum-Phase Picture

The creation of the MP function may be viewed as a filtering operation where the original signal is passed through an appropriate all-pass filter. This filter has the effect of reflecting to the inside of the unit circle all poles and zeros that lie outside the circle. Extending this idea to two dimensions presents once again the conceptual problem of poles and zeros in two dimensions.

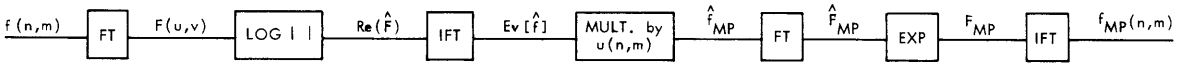

Fig. XXI-14. Diagram for generating a minimum-phase picture.

A much neater approach is to make use of the RPS requirement on the complex cepstrum. The calculation proceeds in seven steps (Fig. XXI-14).

(a) Calculate the Fourier transform of f(n, m) as in Eq. 1.

(b) Calculate $\log |F(u, v)| = \text{Re}\,[\hat{F}(u, v)]$.

(c) Calculate the inverse Fourier transform of $\text{Re}\,[\hat{F}(u, v)]$. This is the $\text{Ev}\,[\hat{f}(n, m)]$.

(d) Multiply $\text{Ev}\,[\hat{f}(n, m)]$ by u(n, m) given in Eq. 5 or Eq. 6. This gives a complex cepstrum corresponding to the minimum-phase version of f(n, m).

(e) Take the Fourier transform of $\hat{f}_{MP}(n, m)$. It is important to realize that $\text{Re}\,[\hat{F}_{MP}(u, v)] = \text{Re}\,[\hat{F}(u, v)]$ and, therefore, $|\hat{F}(u, v)| = |F_{MP}(u, v)|$.

(f) Take the complex exponential of $\hat{F}_{MP}(u, v)$ to obtain $F_{MP}(u, v)$.

(g) Take the inverse Fourier transform of $F_{MP}(u, v)$ to yield $f_{MP}(n, m)$.

Since all of the calculations were done on a computer, a discrete Fourier transform (DFT) was used. This requires a slight modification of the function u(n, m), as well as

of allowable nonzero regions. To see why this is so, let us begin by defining the two-dimensional DFT as

$$F(k,l) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f(n,m) \exp\left(-j \frac{2\pi}{N} nk\right) \exp\left(-j \frac{2\pi}{M} ml\right).$$

Then

$$\text{Re}\,[F(k,l)] = \frac{1}{2}\,[F(k,l) + F^{*}(k,l)].$$

Taking the inverse DFT, we have

$$\frac{1}{2}\,[f(n,m) + f(N-n, M-m)] = \text{Ev}_P[f(n,m)],$$

where the arguments are computed modulo (N, M).

In order to recover f(n, m) from $\text{Ev}_P\,[f(n,m)]$, multiply the even part by $u_P(n,m)$.

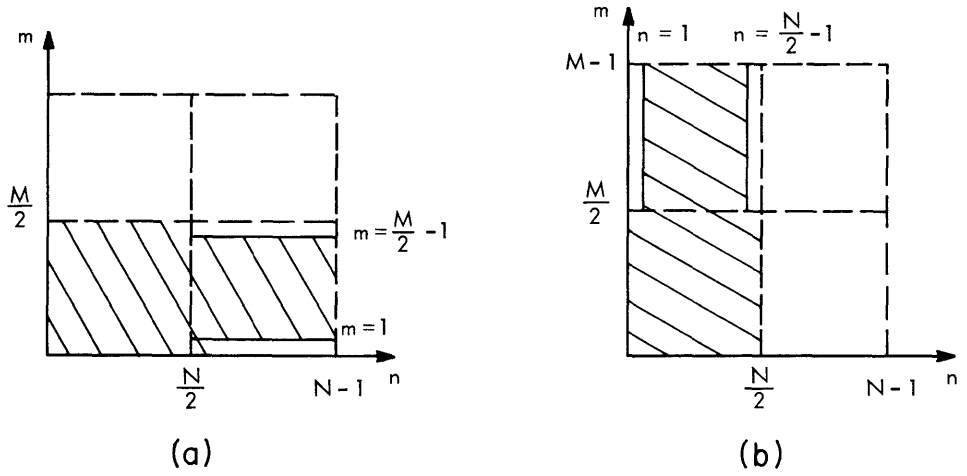

Fig. XXI-15. Allowable nonzero regions for the complex cepstrum of a minimum-phase function using the discrete Fourier transform (DFT).

The allowable nonzero extent of f(n, m) is shown in Fig. XXI-15. Thus the form of $u_P(n,m)$ is

$$u_P(n,m) = \begin{cases} 1 & \text{at } (0,0),\ \left(0, \frac{M}{2}\right),\ \left(\frac{N}{2}, 0\right),\ \left(\frac{N}{2}, \frac{M}{2}\right) \\ 2 & \text{elsewhere in shaded region} \\ 0 & \text{elsewhere} \end{cases}$$

The block diagram for the actual computation is shown in Fig. XXI-16.

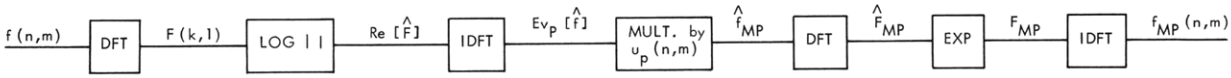

Fig. XXI-16. Diagram for generating a minimum-phase picture using DFT.

4. Examples of Minimum-Phase Pictures

Examples of minimum-phase pictures are shown in Fig. XXI-17. The format of each example is as follows.

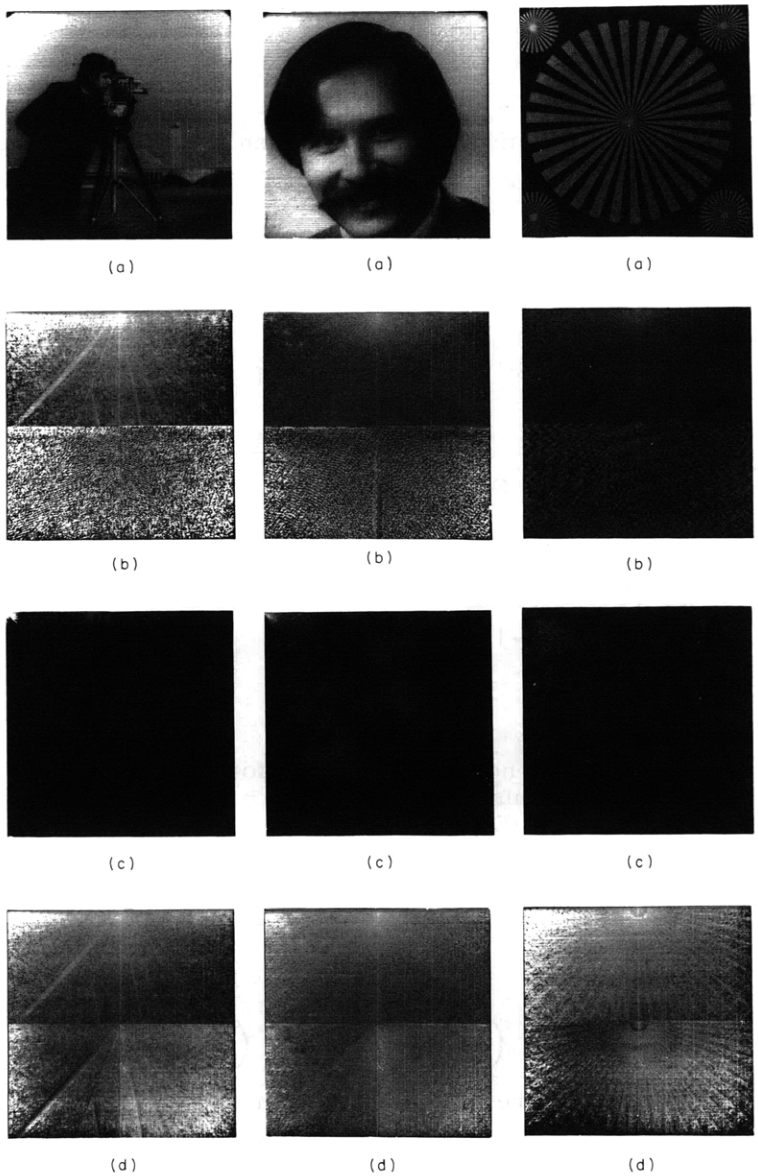

Fig. XXI-17. Minimum-phase pictures.

(a) Original picture, 256 × 256

(b) Upper: logarithm of the magnitude of the FFT of the picture. Lower: phase angle of the FFT

(c) Minimum-phase picture

(d) Upper: logarithm of the magnitude of the FFT of the minimum-phase picture. Lower: phase angle of the minimum-phase picture.

Note that only one-half of the log magnitude and phase angle are shown because of the complex conjugate symmetry of the Fourier transform of real input data.

5. Comments and Conclusions

It does not seem possible to make conclusive statements about MP pictures, but the pictures generated in this experiment have some characteristics in common.

(a) As expected, most of the energy is concentrated near the origin.

(b) The structure of the phase angle appears similar to that of the log magnitude.

(c) Most of the high frequencies in the MP picture are concentrated near the origin.

(d) Aside from the broad statements (a-c), it seems impossible to predict what the MP reconstruction of a given picture would look like.

A. E. Filip

D. ULTRASONIC SYSTEMS FOR MEDICAL DIAGNOSTIC VISUALIZATION

Ultrasonic waves have been used to obtain diagnostic information in many ways during the last twenty years. The first attempts to visualize the ventricles in the brain[1] by direct through transmission were unsuccessful because of diffraction and refraction by the skull. Later, a pulse-echo system similar in principle to sonar and radar was used[2, 3] and developments of these techniques are now widely used in routine clinical diagnosis.[4]

Various one-dimensional A-mode techniques are used for detecting shifts in the midline structures of the brain and measuring the diameter of the fetal head, as well as for locating foreign bodies and tumors and measuring the depth of the eye. When one of the echoes is moving, M-mode is used to obtain a curve of echo displacement with time from which the velocity of movement of the echo-producing structure may be calculated. This method is used mainly in studying the heart, but is also used in research on the physiology of headache.

In a suitable scanning, or B-mode, system, the ultrasonic echoes may be displayed at the correct geometric position of the echo-producing structure to yield an echogram that resembles an anatomical cross section. Such a display has been found useful in diagnostic studies of the abdomen, uterus, kidneys, liver, spleen, aorta, eye, and breast.

Recently, acoustic holography in its several forms and synthetic-aperture side-looking radar have offered promising new vistas for diagnostic ultrasonic application. The somewhat similar field of seismic exploration for oil has made great advances in the last decade. In this report we shall outline the limitations in ultrasonic visualization as it is performed at present, and, in the light of these limitations, examine some improvements suggested by procedures used in similar techniques.

1. Properties of Tissues

The acoustic properties of the parts of the body fit readily into three groups: soft tissue, bone and cartilage, and air-containing tissue.[5] In general, air-containing tissue, such as lung and the gastro-intestinal tract, can be regarded as a barrier to ultrasound. The attenuation is very high and the reflection coefficient at the boundary of any air-containing area is large. Most of the body is composed of soft tissue. Included in this group are muscle, fat, blood, and the material of most organs, such as kidney, liver, and so forth. The velocity of propagation and characteristic impedance of all materials in the group are similar to those of water, that is, 1.5×10^3 ms^{-1} and 1.5×10^6 kg m^{-2} s^{-1}, respectively. The total variation of the properties of other soft tissues around these figures is of the order of 8% and 20%, respectively. The attenuation of tissue varies more widely, and at 1 MHz ranges from 0.07 Np/cm (for fat) to 1.2 Np/cm (for liver). The attenuation of most soft tissues is proportional to frequency; one of the exceptions is liver, which has constant attenuation.

Bone has a higher velocity and an impedance of from 3.3 to 4×10^3 ms^{-1} and from 6 to 8×10^6 kg m^{-2} s^{-1}. The reflection coefficient between bone and soft tissue is very much greater than that between two soft tissues. When the sound propagates into bone the increased absorption and the effects of refraction and mode conversion at the boundary severely distort and attenuate the beam.

The exact identity of echo-producing interfaces in the body is not known with certainty. Nevertheless, from clinical experience during the past 15 years, it is obvious that the boundaries of a structure or organ give much greater echoes than the internal substance of the structure. Internal echoes from fluid-filled structures, such as cysts and the bladder, are of quite low amplitude; somewhat larger echoes are obtained from within solid structures such as muscle and solid tumor.

The echo pattern obtained from within the body is extremely variable with only small changes in the transducer beam axis direction. This is due in part to specular reflection which causes inclined interfaces to deflect the reflected ultrasonic energy away from the line of sight where it is lost to the transducer.

2. Frequency of Operation

The frequency of operation of diagnostic equipment is fixed by the required

resolution, and by the available power, attenuation, and signal-to-noise ratio requirements. In general, a higher frequency gives better resolution but the absorption is larger, thereby limiting the penetration. In practice, a frequency around 2 MHz is used in the abdomen and up to 30 MHz has been used in the eye. These frequencies correspond to a range of wavelengths of 0.75-0.05 mm.

At present, the highest frequency consistent with the desired tissue penetration is used. In some newer techniques, when data are to be acquired on a computer, it may be desirable to limit the frequency of operation by the resolution and acquisition requirements, rather than by the absorption.

3. Present Pulse-Echo Techniques

In pulse-echo techniques we use a single transducer (or two adjacent transducers) to measure the distance to an echo-producing interface by measurement of the travel time of the returned echo. The transducer beam axis and the intensity-modulated trace are moved in synchronism and all intensified spots are stored as they appear, either on film or with a storage tube. This (B-mode) display is a cross section of echo-producing interfaces and resembles an anatomical cross section. The section includes all lines of sight of the transducer. That is, the transducer views from the side and the resulting echogram is of a section viewed from above (Fig. XXI-18). This is exactly equivalent to radar and sonar displays.

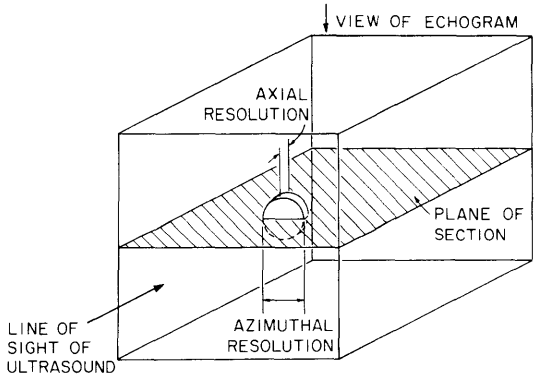

Fig. XXI-18.
Formation of image in pulse-echo system.

The resolution cell of a pulse-echo system has two components. The axial resolution is dependent on the time response of the transducer and, in general, can be made as small as necessary. The azimuthal resolution is fixed by diffraction considerations, and in practical systems is the limiting factor in resolution.

a. Axial Resolution

The axial resolution is set by the transducer material and construction details, and by electrical and mechanical loading. In general, a compromise must be made between

axial resolution and transducer sensitivity. It is quite practicable to produce a transducer such that the echo off a plane interface takes only three cycles of the resonant frequency, without an appreciable sacrifice in sensitivity. To achieve a shorter pulse, a reduction in sensitivity occurs. At 2 MHz, this pulse length corresponds to 1.5 μs and a range increment of 1.1 mm.

b. Azimuthal Resolution

The ultimate diffraction-limited resolution in azimuth at the focus for the Rayleigh criterion is given by[6]

$$\delta = \frac{.61\lambda F}{a},$$

where F is the focal length of the transducer, a the transducer radius, and δ the beam radius at the focus. With a focal length of 20 cm and transducer radius of 4 cm at 2 MHz, the resolution, δ, is 2.2 mm. This is at sharp focus, however, and the resolution falls off rapidly at distances away from the focus.

The distribution of intensity along the beam axis near the focus is given by

$$I = \left(\frac{\sin u}{u}\right)^2 \cdot I_o,$$

where $u = \frac{\pi}{2\lambda}\left(\frac{a}{F}\right)^2 \cdot z$, and z is the distance from the focus.

The beam diameter d at a distance z is related[7] to the intensity at that distance (I), the focal intensity (I_o), and the beam diameter at the focus (d_o) by $I_o \cdot d_o^2 = I \cdot d^2$, whence

$$d = d_o \sqrt{I_o/I} = d_o \frac{u}{\sin u}.$$

For the example already quoted,

$$u \approx \frac{\pi}{4} \cdot z.$$

For the beamwidth to double,

$$\frac{d}{d_o} = 2 = \frac{u}{\sin u},$$

whence $u \approx 2.0$, and $z = 2.5$ cm. That is, the depth of focus for a doubling of the beamwidth to 4.4 mm is ±2.5 cm.

A more practical approach to azimuthal resolution is to use weak focusing.[7] In this case the ultimate resolution is not achieved but a reasonable value is obtained over a

much larger penetration range. Typically a beamwidth of 1.2 cm may be obtained over a range from 10 cm to 30 cm, by using a transducer of diameter 3.5 cm, a radius of curvature 2.5 cm, and 2-MHz frequency.[8] This may be optimized for different ranges and transducer radii and the beamwidth can be reduced by an increase in frequency.

c. Sensitivity to Specular Reflection

It has been the experience of all investigators in this field that reflections from tissue are specular, and that there is only a small amount of scattered energy. The visible effects of this are that the received echo pattern varies considerably with small changes in transducer beam direction, and in commercially available manually scanned B-mode systems the appearance of the echogram can be altered considerably by changing the scanning pattern. Interfaces inclined to the beam are not displayed at all or displayed as very weak echoes. This difficulty can be overcome to some extent by compound scanning, by which the transducer line of sight is moved around the part to be examined with a rocking action so that it intersects each structure in the section plane at right angles. Nevertheless, structures inclined to the section plane may be missed completely on the display.

One of the effects of specular reflection is that the dynamic range of echoes to be processed must be very large to retain the small scattered component reflected from inclined interfaces. This, in turn, accentuates artifacts attributable to beamwidth, beam sidelobes, and multiple reflections within the examined structure.[9]

d. Data Acquisition Rate

In B-mode cross-sectional echoscopy the data are acquired on a plane-by-plane basis. For each plane it has been found desirable to perform a compound scan, with an oscillatory sector scan of included $\angle 40°$ and approximately 7 such scans in a slower arc sweep of 160° around the patient. It is difficult mechanically to obtain a scan cycle in less than 1.5 s, which means that there is a minimum time for scanning a single plane of the order of 10 s. The cyclic scan motion of included $\angle 40°$ and 1.5 s corresponds to an angular velocity of approximately 1 rad/s. With a maximum range of 45 cm in a water-coupling echoscope (penetration of 22 cm) the time taken for the scan motion to move the beam laterally 1 resolution unit of 2.5 mm is 5 ms, which fixes the maximum pulse repetition period to keep the returned echoes from appearing as separate dots. The travel time of an ultrasonic pulse returning from 45 cm is 600 μs. Thus the minimum pulse repetition period to avoid range ambiguities is approximately 1 ms which allows some time for reverberation in the coupling tank to decay between pulses, or time for use in data processing. Practical systems use a 400-1000 Hz pulse repetition rate, or 2.5-1 ms repetition period.

The time for each section is 10 s and, for a complete examination comprising

20 planes, the ideal examination time, including pauses between sections for recording purposes, is ~10 min. In practice, ~20 min is required, excluding the time needed for patient positioning and removal from the equipment. This is considerably more time than is required for diagnostic x-ray but less than that for most forms of radioisotope scans.

4. Present Acoustic Holography Techniques

In the holographic technique, the amplitude and phase of the incident information-bearing wave are recorded at a plane.[10] The recording contains all information in the incident sound wave. This may be converted to an optical image by using optical holographic reconstruction techniques. The difference in wavelength between ultrasound and light causes a demagnification of the image which is different in the azimuthal and axial directions, thereby precluding direct visual observation of the three-dimensional image. The information is presented in the form of section planes normal to the ultrasound beam. Alternatively, the phase and amplitude information can be used as data in a computer program which effectively propagates the wave backward from the

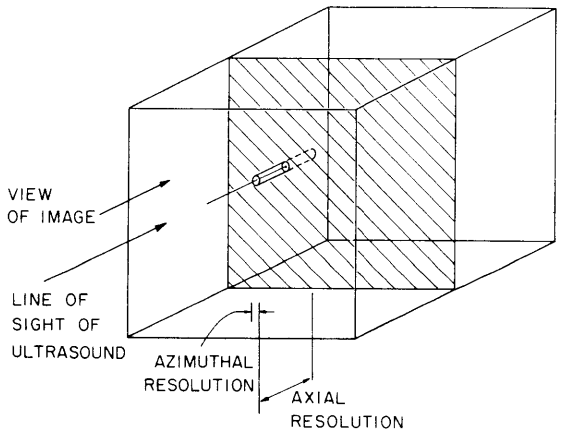

Fig. XXI-19.
Formation of image in holographic system.

receiving plane to a plane within the examined structure, again giving a planar image. The direction in which the image is viewed corresponds to the direction of propagation of the information-bearing wave (Fig. XXI-19).

a. Resolution

The axial and azimuthal resolutions are both limited by diffraction effects, and the resolution of practical systems approaches this limit. The azimuthal resolution is set by the Rayleigh limit and is better than the axial resolution which is a focusing or parallactic effect. The azimuthal resolution is given[6] by $0.61\lambda D/a$, and the axial resolution

by $2\lambda(D/a)^2$. For a distance (D) of 20 cm, 10-cm radius of aperture (a), and wavelength (λ) of 0.75 mm at 2 MHz, the azimuthal and axial resolutions are 0.9 mm and 6 mm, respectively.

b. Detection Systems

In present prototype systems one of two methods is used for detecting the acoustic wave field.[11] One method uses the deformation of an interface by the sound beam, which is sensed by means of a coherent optical technique. This system provides an optical image of a section plane in real time, because of the fast parallel processing available with optical data techniques. The system was found to require an acoustic lens to form an in-focus hologram at the hologram plane interface. The aperture is ~10 cm, and the object is ~20 cm from the lens. The operating frequency is 3 MHz, which gives theoretical azimuthal and axial resolutions of 1.2 mm and 1.6 cm, respectively.

In an alternative detection system either a scanned piezoelectric transducer or an array of transducers provides an electrical signal. This signal may be converted into an optical hologram by an oscilloscope scanning technique or used directly as data for a hologram reconstruction program in a digital computer. The scanning process can take up to 10 min for a typical hologram. The optical reconstruction technique requires only the time necessary for film development. The computer reconstruction time has been quoted as around half an hour.[12]

c. Practical Results and Limitations

The prototype units have been used in the field of nondestructive testing and have attained azimuthal resolutions approaching the theoretical diffraction limit with two-dimensional test objects. With a three-dimensional object the out-of-focus images obscure the planar in-focus output image. The images that are only slightly out of focus give rise to interference fringes which have a serious effect on image quality. Thus the early clinical trials of the interface-deformation optical reconstruction system for visualization of the breast are somewhat disappointing, because of the speckle effect of this coherent interference fringe noise. Earlier experiments on excised tissue and thin sections gave more encouraging results,[13] since their thickness was comparable with the axial resolution or depth of field and there were no out-of-focus contributions to give rise to interference fringes.

5. Related Techniques

a. Synthetic-Aperture Techniques

The synthetic-aperture technique used in side-looking sonar and radar is a combination of pulse-echo and holographic techniques.[14] It retains the axial resolution of pulse echo but uses the large synthetic aperture to obtain good resolution in azimuth along the direction of motion of the transducer. The "cross-track" resolution is not affected by

the technique and remains limited by the actual aperture (Fig. XXI-20). This method relies on the scattering properties of the reflection points to return echoes to the transducer at many points along the length of the aperture. The echoes are detected with a

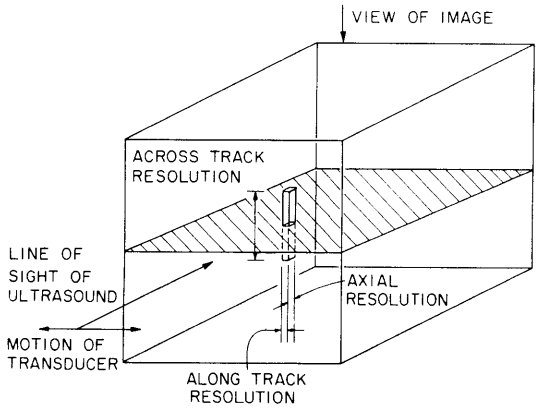

Fig. XXI-20.
Formation of image in synthetic-aperture system.

synchronous detector and stored on a film strip as individual brightness modulated A-mode displays. An elegant optical process is used to combine the information from many returns to a constant range and varying angles to make the synthetic aperture. When the aperture becomes an appreciable portion of the range, the contributions from a single point do not lie on a line of constant range, but on a hyperbolic curve whose curvature varies with distance (Fig. XXI-21). This makes the optical processing more

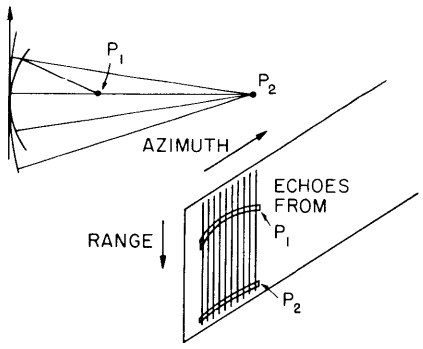

Fig. XXI-21.
Range curvature effect in synthetic-aperture technique.

difficult. It is desirable to use an aperture that is comparable to the distance to allow detection of specularly reflected echoes from inclined interfaces.

The synthetic-aperture processing technique does not require motion of a transducer but could be achieved equally well with a line of fixed transducers, when the result is effectively a one-dimensional hologram. The attraction of the synthetic aperture, from an ultrasonic point of view, is not in the increased resolution, as a physical transducer of sufficient size can be built, but in the elegant processing. This has the effect of

providing sharp focusing at all distances.

There seems to be no conceptual difficulty in extending the concept of two dimensions to provide improved azimuthal resolution in two orthogonal directions while obtaining the pulse-echo axial resolution. It could be achieved by writing the returns to a two-dimensional array at a fixed distance on a film. The in-focus return would be reconstructed in the normal way. A separate film would be required for each resolution increment in range. The effect of range curvature would then become two-dimensional, however, and it is not clear that an optical-processing scheme is realizable.

b. Beam Sharpening

Resolution in azimuth could be improved by an inverse-filtering process on the time series represented by the succeeding echoes from a given delay as the beam sweeps. This has been suggested for EEG recordings.[15] The beam pattern is of the form of $\sin x/x$ or $J_1(x)/x$, depending on the reflector. The spatial frequency transform is then a semicircle or a rectangle. Inverse- or deconvolutional-filtering techniques cannot extend the spatial bandwidth of such steep-sided spatial spectra without running into serious signal-to-noise ratio problems, and therefore no great improvement can be expected.

c. Beam Deconvolution

Beam deconvolution could be carried out by a decision theoretic approach.[16] It would appear that the problems of nonconstant "desired signal" would remain. A considerable amount of computation is also required on-line to perform the decision analysis. We considered it more fruitful to use such computing power in other ways. Similarly, such methods as analytic continuation[17] could be employed, but the cost in computing power could well be used to improve the input-data acquisition, rather than to process already acquired data.

d. Optical De-blurring

The coherent-optic and holographic techniques for de-blurring images have application in many classes of image degradation.[18] For such methods to succeed, however, the blur function or resolution cell, must be shift-invariant, at least over small regions. In an ultrasonic echogram the blur function for each echo depends on the direction from which the echo is received, and this information is not retained for later analysis. Therefore the blur function is not shift-invariant, even within small regions, and the method fails.

e. Seismic Processing

The literature on processing of seismic data is extensive. In seismic visualization, returns from a single source are received on a linear array. The positions of source

and receivers are then advanced and the procedure repeated. The source generates a spherical wave and the returned echoes are all specularly reflected highlights. Basically, processing includes calculation of the time offset required for each ray path from a reflecting point and search for correlation between traces on a field-point-by-field-point basis using a ray-path approach (Fig. XXI-22). The data records are gathered so that the echoes from a single reflection position obtained with a number of source and receiver locations are processed together. The time offsets caused by the geometry of the ray paths are applied (normal moveout), and then deviations of echoes from those expected

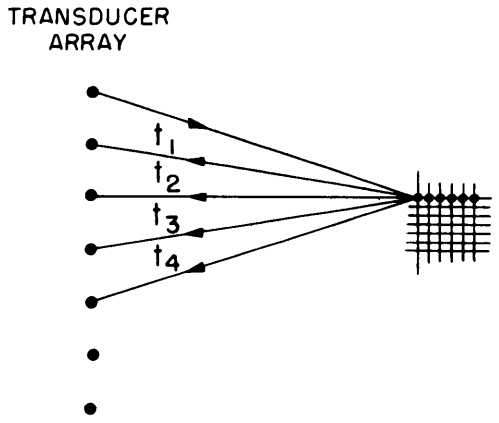

Fig. XXI-22.
Basis of seismic technique. The ray paths for each field point are analyzed consecutively.

for the ideal model are detected and interpreted as variations in the model. The emphasis in seismic processing is on enhancement of signal-to-noise ratio (filtering), eliminations of multiple reverberation artifact (de-reverberation), velocity measurement, and correction of display errors caused by dip of the reflecting layers (migration). To achieve these processes, analysis of a multiplicity of transmitters and receivers and coherent recording of data are required. The processing is simplified by the initial assumptions of horizontal layers and approximately vertical rays applied to most processes. These assumptions are not valid, and lead to a large increase in the complexity of equivalent processes in the ultrasound visualization case.

In seismic techniques consideration is not taken for any analysis of azimuthal resolution; flat layer interfaces are assumed, for which azimuthal resolution is not important.

It is very likely that the full potential of multiple-transmitter, multiple-receiver techniques evolved for seismic visualization will find application in ultrasound as fast on-line data-acquisition systems become more economical.

6. Improvement in Pulse-Echo Methods

Suggestions for improvements in pulse-echo methods normally center on two areas, improvement in azimuthal resolution and solution of the specular-reflection problem.

a. Strong Focusing

Improvement in azimuthal resolution can be achieved by employing strong focusing, using range gating to display only echoes close to the focus, and moving the transducer so that the operating region explores all of the part examined. The time to make a compound scan section is limited, at present, by mechanical considerations and is approximately 10 s. Increasing the amount of mechanical movement increases the scanning time, which is unacceptable as the present time is already too long.

b. Annular Array

To achieve sharp focusing over the entire range, electronic focusing of an annular array[19, 20] can be used to obtain a sharp focus. The results of analysis for a set of annular rings focused at 10 cm for various combinations of overall diameter and number of rings approach the diffraction limit.[20] For focusing on transmission, it is necessary to use appropriately delayed pulses to excite the annular rings. The delays, which may be obtained from flip-flop timing circuits, are varied for different focal distances. If 5 focal ranges are to be used, it would be necessary to transmit 5 differently focused pulses to cover all of the penetration range. At a 5-cm depth of field this would give a coverage of 25 cm, and the pulse repetition rate could be increased to 1 kHz. Thus the 5 pulses could be transmitted in 5 ms and not exceed the maximum pulse-duration time set by the mechanical scanning speed. The overall scanning time would not then be increased.

For focusing on reception, the received signals must be delayed. This may be achieved digitally or with analog delay lines. When focusing on reception, a dynamic focusing arrangement could be implemented either digitally or in analog form to achieve virtually perfect focus at each range in turn, with only a single transmitted pulse.

c. Specular-Reflection Improvement

One method suggested for overcoming the limitation of specular reflection is to use a wider transducer aperture.[8] The wider transducer aperture, by geometrical reasoning, allows reception of reflected energy from a more inclined surface. This technique requires a focus to retain azimuthal resolution. The techniques for achieving this have already been mentioned.

Another method is to use a higher sensitivity and strive to display echoes scattered from within the tissue structures.[21] Differences in echo texture are then used to indicate interfaces between structures. This has the advantage of giving some information about the tissue structures themselves but it tends to display artifact echoes arising from multiple reflections and beam irregularities. It is also likely to smear out echoes from strongly reflecting interfaces.

Alternatively, an attempt is made to obtain a specular highlight echo from each

interface in an extension of compound scanning. One method of achieving this is by omnidirectional scanning which requires movement of the transducer in such a way that the line of sight cuts the section plane obliquely at a number of angles and positions.[22] This technique greatly increases the scan time and, to a small extent, degrades the azimuthal resolution.

7. Wide-Aperture Annular Array

Techniques based on the seismic approach can be made quite simple or progressively more complex. In general, at each higher level of complexity the data rate is increased and scanning time is reduced at a cost of increased computation power requirement. With the approaches to be described, the advantages of the very high sensitivity technique to display scatter within tissue could be combined with highlighting of specular reflectors to provide a well-outlined structure with texture information.

The simplest form of wide-angle array achieves wide-angle coverage with few elements by using circular symmetry about the beam axis (Fig. XXI-23). The receiving array

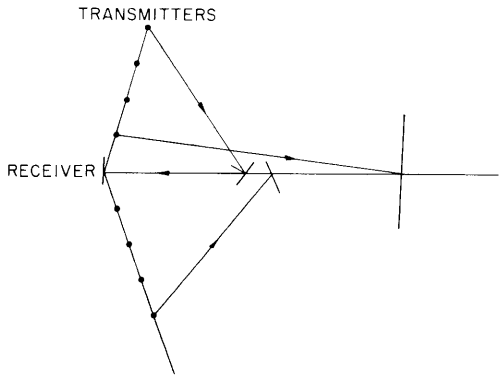

Fig. XXI-23.
Principle of wide-aperture array to achieve good resolution with an annular receiving array and sensitivity to inclined interfaces.

has as many as 8 close-spaced annular elements with sufficient aperture to achieve the desired diffraction-limited azimuthal resolution when sharply focused. The transmitting array has several widely spaced annular elements that are pulsed consecutively. They are spaced so that there is always a received echo of reasonable magnitude off an angled plane from one of the transmitting transducers. With this system echoes from inclined structures would be enhanced, achieving a reduction in compound scan and, therefore, allowing a faster scan. There would also be a greater detectability of interfaces inclined to the plane of scan to give more complete outlines. The axial resolution of a pulse-echo system would be preserved and the azimuthal resolution would approach that of a holographic system. For a receiving array of 10-cm diameter and a transmitting array of 4 annular rings with 5-cm spacing (40-cm OD), the resolution at a distance of 20 cm would be ~2 mm in each direction, and reflectors would be detected up to 30° of

inclination from the beam axis. As the sector in compound scanning is ±20°, the wide-aperture array would provide greater angle coverage with no need for mechanical scanning and the mechanical-scanning speed limitation would be removed. A simple arc scan could be used with the transducer array directed radially.

a. Transmission

With 4 outer rings and the central disk as transmitters, each line of sight requires 5 transmitted pulses for complete angle coverage. The minimum time between pulses to avoid range ambiguities is 1 ms as before, and the maximum time between lines of sight to avoid resolution of individual lines is 5 ms with a sector scan, or somewhat more when compound sector scanning is not employed. Thus the requirement for multiple transmitted pulses is not a factor in overall data-acquisition rate.

b. Receiver-Array Processing

The dynamic focusing of the receiver array can be implemented in a stepwise manner to yield discrete foci with a sufficient depth of field to maintain the required resolution at intermediate distances. It could also be achieved continuously. The signals from the individual annuli would be suitably delayed and added to form a focused beam. In the stepwise delay method, a set of delays would be permanently connected to each annular ring and the output for that focus brought to a terminal (Fig. XXI-24). The

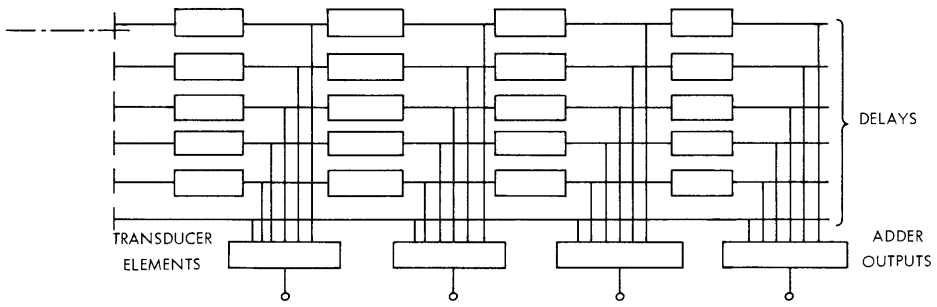

Fig. XXI-24. Receiver processing for dynamic focusing of the annular receiving array. For N transducer elements and M effective focal distances, (N-1) × M delays and M adders are required.

computer would select as input the terminal corresponding to the appropriate delay time. The best delays would be analog, with lumped-constant delay lines. The maximum delay required with a 5-cm aperture radius and 10-cm minimum distance is 7 μs. The delays for all discrete foci could be provided separately to give an output terminal for each beam. A multiplexer could then select the necessary output for the appropriate total

delay time. If digital delay is used, the sampling rate must equal the minimum unit of delay that is to be used, or ~25 ns.

The sampling rate required for the received signal depends on the subsequent processing. In seismic processing, a sampling rate of 2.5 kHz, or a sample each 4 ms, is used on signals of 125-Hz bandwidth. This corresponds to 10 times the Nyquist rate. In a 2-MHz ultrasound system the equivalent rate would be 40 MHz, or a sample each 25 ns.

If analog delays are used to form the focused beam, the data rate may be reduced substantially. For Nyquist sampling, a rate of 4 MHz is sufficient and, if the signal is envelope-detected, a rate of 500 kHz suffices. The slower data rate corresponds to a sample for each echogram picture point. It allows the storing of received echoes at high dynamic range for picture-processing techniques such as contrast enhancement, grey-scale manipulation, interpretive pattern recognition, and the manipulation of a set of pictures to form a three-dimensional matrix.[23]

The 4-MHz rate allows processes to be carried out on the input signal time series to achieve de-reverberation, filtering, and signal enhancement. To achieve other seismic processes, such as velocity measurement and migration, a much finer sampling interval is required. This could be achieved by direct sampling at the 40-MHz rate, or by interpolation of the 4-MHz sampled data. In the former case, considerable technological difficulties are confronted, while the latter introduces a large amount of computation.

c. Echo Processing

Since the total path length for a given reflector distance varies with the radius of transmitting annulus and reflector distance, a correction must be applied to obtain the true reflector distance from the observed delay time. This correction varies with distance of the interface and corresponds to normal moveout correction in seismic processing. If the received echo data for each transmitting annulus are stored separately, then velocity information could be inferred from a crosscorrelation study. Information on the angularity of the interface and its scattering properties could also be obtained which might prove useful in differential diagnosis.

For a less ambitious system, the contributions from the various transmitted pulses could simply be added to form a composite line of echo information that could then be stored as picture elements in the final echogram.

d. Implementation of a Practical System

A practical system using a simple version of the wide-aperture annular array can be implemented with the use of available hardware. The receiver delays would be fixed analog RLC delay lines with 4 focal beams available at 4 terminals. The appropriate

input would be selected by a multiplexer under computer control. The received echoes would be envelope-detected and digitized at a 500-kHz rate, or a sample each 2 μs. For a penetration range of 300 μs there would be 150 samples in a data line. The acquisition rate of a sample each 2 μs is well within the capabilities of present minicomputers.

After the data are acquired, unpacked from the data buffer, and added to the composite line buffer, an appropriate correction between time delay and interface distance is made. The corrections can be stored in memory in the form of indirect address offsets. By using a machine with multiple accumulators and two-level indirect addressing, the points can be packed in less than 10 μs each. Thus unpacking takes 1.5 ms. The pulse repetition period could be made 2 ms to allow a composite line to be scanned in 10 ms. After the composite line is obtained it is packed into a buffer which stores the picture. This operation again uses multiple accumulators and is achieved in ~10 μs per point. This exceeds, by a factor of two, the maximum period for a scanning rate of 1 rad/s.

After the scan is complete the data in the picture buffer are normalized with the number of times each cell is referenced and then are ready for output display.

8. Further Developments

A more general form of wide-angle array based on a rectangular or hexagonal grid rather than an annulus could be used. In this case, beam steering of a close-spaced receiving array could be employed to further reduce the amount of mechanical scanning. The processing requirements increase markedly, however. The essence of the approach is to use a small number of close-spaced receiving arrays of a size sufficiently large to obtain the required resolution. The coarse-grid transmitting array must be large enough to ensure specular reflection from a large proportion of inclined reflectors. Such a system can be made as simple or as complex as required to achieve the desired performance.

The ultimate requirement for such a system would be to provide a complete scan of the three-dimensional volume without mechanical scanning. The acquisition time would then be limited only by the data rate and storage capacity of the computer that is used to control the system. A volume of 25 × 20 × 20 cm suitable for displaying the contents of the pregnant uterus at a resolution of 2.5 mm would contain 640 K points. If each point is scanned from 5 directions and the average data rate is 25 μs per scanned point, the data could be acquired in approximately 80 s.

It is suggested that a wide-aperture array pulse-echo system has several advantages over a holographic system. It is relatively straightforward to display any section plane through the data matrix, rather than to be restricted to a set of parallel planes normal to the ultrasonic beam.[23] Structures not in the plane of section of interest do not interfere with the image. The properties of the medium do not degrade the image to the

same extent. Specular reflection does not produce highlighting, but is used to provide clean outlines of the anatomical structures. The effects of velocity variation, refraction, and so forth can be corrected if necessary during the formation of the data matrix on a line-by-line basis. There is no basic limitation on the processing methods, as there is with holography which requires linear shift-invariant operators. The limitations on the pulse-echo system are economic rather than technical.

9. Conclusion

The ideal azimuthal resolution of different visualization systems can be made comparable, but the axial resolution of the pulse-echo technique is usually better than that of holography. The main differences between pulse echo and holography are in the data-processing methods and the properties of tissue that is utilized. The holographic technique offers the advantages of the extremely high data-processing rates inherent in optical processing. To obtain three-dimensional information by either transmission or reflection methods, however, requires sufficiently strong scatterers at the interfaces to provide an information-bearing wave that apparently diverges from these points. Pulse-echo techniques are not limited by specular reflection, but the additional complexity to overcome its problems causes an increase in processing requirements.

Synthetic-aperture techniques provide an elegant method of achieving continuous focus in one direction which could be extended to two dimensions. Distortions arising from the range curvature effect, however, can add complexity in the one-dimensional case, and may preclude optical processing in the two-dimensional case.

Using an extension of the pulse-echo technique based on seismic exploration methods, a wide-angle array is proposed. In its simplest form, it is composed of an axially symmetric system that provides good resolution and sensitivity to inclined interfaces. It uses techniques and devices currently available to reduce the data at a rate comparable with present pulse-echo acquisition rates. The greater sensitivity to inclined echoes will allow faster scanning and a reduction in total examination time.

Further developments to remove the need for mechanical scanning are a straightforward extension of the simple axially symmetric system, although the increase in processing complexity is not small.

I would like to thank Dr. O. J. Tretiak for useful discussions during this work.

D. E. Robinson

References

1. K. S. Dussik, "Possibility of Using Mechanical High Frequency Vibration as a Diagnostic Aid," Z. Neurol. Psychiat. 174, 153-168 (1942).

2. J. J. Wild, "The Uses of Ultrasonic Pulses for the Measurement of Biological Tissues and the Detection of Tissue Density Changes," Surgery 27, 183 (February 1950).

3. D. H. Howry and W. R. Bliss, "Ultrasonic Visualization of Soft Tissue Structures of the Body," J. Lab. Clin. Med. 40, 579 (October 1952).

4. N. Lindgren, "Ultrasonics in Medicine," IEEE Spectrum 6, 48-57 (November 1969).

5. P. N. T. Wells, Physical Principles of Ultrasonic Diagnosis (Academic Press, Inc., London and New York, 1969), p. 5.

6. M. Born and E. Wolf, Principles of Optics (Pergamon Press, London and New York, 1970), pp. 435ff.

7. G. Kossoff, "Design of Narrow Beamwidth Transducers," J. Acoust. Soc. Am. 35, 905-912 (1963).

8. G. Kossoff, D. E. Robinson, and W. J. Garrett, "Ultrasonic Two-Dimensional Visualization Techniques," IEEE Trans. on Sonics and Ultrasonics, Vol. SU-12, No. 2, pp. 31-37, June 1965.

9. D. E. Robinson, G. Kossoff, and W. J. Garrett, "Artifacts in Ultrasonic Echoscopic Visualization," Ultrasonics, pp. 186-194, October 1967.

10. A. F. Metherell, H. M. A. El-Sum, J. J. Drekev, and L. Larmore, "Introduction to Acoustical Holography," J. Acoust. Soc. Am. 42, 733-742 (1967).

11. B. B. Brenden, "A Comparison of Acoustical Holographic Methods," in A. F. Metherell et al. (Eds.), Acoustical Holography (Plenum Press, New York, 1969), Chap. 4.

12. A. L. Boyer, P. M. Hirsch, J. A. Jordan, Jr., L. B. Lesem, and D. L. Van Rooy, "Reconstruction of Ultrasonic Images by Backward Propagation," in A. F. Metherell (Ed.), Acoustical Holography, Vol. 3 (Plenum Press, New York, 1971).

13. L. Weiss and E. D. Holyoke, "Detection of Tumors in Soft Tissue by Ultrasonic Holography," Surgery, Gynecology and Obstretics 128, 953-962 (1969).

14. L. J. Cutrona, E. N. Leith, L. J. Porcello, and W. E. Vivian, "On the Application of Coherent Optical Processing Techniques to Synthetic Aperture Radar," Proc. IEEE 54, 1026-1032 (1966).

15. D. G. Childers, R. S. Varga, and N. W. Perry, "Composite Signal Decomposition," IEEE Trans. on Audio and Electroacoustics, Vol. AV-18, No. 4, pp. 471-477 (December 1970).

16. T. Y. Young, "Epoch Detection – A Method for Resolving Overlapping Signals," Bell Syst. Tech. J. 44, 401-425 (1968).

17. J. L. Harris, "Information Extraction from Diffraction Limited Imagery," Pattern Recognition 2, 69-77 (May 1970).

18. J. L. Harris, "Potential and Limitations of Techniques for Processing Linear Motion-Degraded Imagery," in M. Nagel (Ed.), Evaluation of Motion Degraded Images, Proc. NASA/ERC Seminar, Cambridge, Mass., December 1968.

19. H. E. Melton, Jr., "Electronic Focal Scanning for Improved Resolution in Ultrasound Imaging," Ph.D. Thesis, Duke University, 1971.

20. M. Hubelbank and O. J. Tretiak, "Focused Ultrasonic Transducer Design," Quarterly Progress Report No. 98, Research Laboratory of Electronics, M.I.T., July 15, 1970, pp. 169-177.

21. G. Kossoff, "Improved Techniques in Ultrasonic Cross-sectional Echoscopy" (submitted for publication).

22. W. J. Fry, G. H. Leichner, D. Okayama, F. J. Fry, and Elizabeth K. Fry, "Ultrasonic Visualization System Employing New Scanning and Presentation Methods," J. Acoust. Soc. Am. 44, 1324-1338 (1968).

23. D. E. Robinson, "Display of Three-Dimensional Ultrasonic Data for Medical Diagnosis" (submitted for publication to J. Acoust. Soc. Am.).