General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

*Frangione
X8135
242
T.o. Don Friedman)
702
B-11*

CLINICAL OPHTHALMIC ULTRASOUND IMPROVEMENTS
(FINAL REPORT)

J. B. GARRISON, PhD. and P. A. PIRO, M.D.

Applied Physics Laboratory/The Johns Hopkins University
and
Department of Ophthalmology, School of Medicine

August 1981

THE JOHNS HOPKINS UNIVERSITY ∎ APPLIED PHYSICS LABORATORY
Johns Hopkins Road, Laurel, Maryland   20707

CLINICAL OPHTHALMIC ULTRASOUND IMPROVEMENTS
(FINAL REPORT)

J. B. GARRISON, PhD. and P. A. PIRO, M.D.

Applied Physics Laboratory/The Johns Hopkins University
and
Department of Ophthalmology, School of Medicine

August 1981

TABLE OF CONTENTS

A. SUMMARY

The purpose of this contract was:

1. Demonstration of use of a switched array to reduce data collection time to a few milliseconds to avoid eye motion problems in the eye itself.

2. Selection and use of an eye tracking system.

3. Demonstration of all the key elements in the Experimental Clinical System (NASA Phase II system) by use of equipment being assembled on a private foundation grant for automatic tracking of heart wall motion with two-dimensional echocardiograph data.

These demonstrations are for the purpose of answering questions raised during the review of research proposals submitted to the National Eye Institute[1] and the National Aeronautics and Space Administration[2] for Clinical Ophthalmic Ultrasound Improvements.

The results presented here have been funded by the subject NASA contract ($26,000), by APL IR&D ($10,000), and Wilmer Institute ($5,000). This report covers the work supported by all these sources since the funds were applied to a single program. In addition, we are making use of equipment and techniques funded by the W. W. Smith Charitable Trust and the National Institutes of Health.

We believe that we have met the objectives of the NASA contract and have demonstrated that we have satisfactory solutions to the questions raised during the review of the initial proposals.

One of the major questions was the effect of eye motion during the data collection time. After we established the parameters of the switched array configuration we consulted with Professor D. H. Robinson of the Wilmer Institute to obtain his assessment of the effects of eye motion on the performance of the system as it is now configured. A copy of his letter to us is included in this report (next two pages). In addition to the specific question of eye motion he included two very helpful suggestions regarding "operational" use of the system.

---

1. "Grant Proposal for Clinical Ophthalmic Ultrasound Improvements," Principal Investigator: Philip A. Piro, dated November 1, 1978.
2. "Program Plan for Clinical Ophthalmic Ultrasound Improvements," dated July 2, 1979.

THE WILMER OPHTHALMOLOGICAL INSTITUTE
THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
THE JOHNS HOPKINS HOSPITAL

DEPARTMENT OF OPHTHALMOLOGY

LABORATORY OF VISUAL AND
OCULOMOTOR NEUROPHYSIOLOGY

Telephone· 301-955-3587

June 27, 1980

Please address reply care of

Rm. 355 Woods Research Bldg.
The Wilmer Institute
601 N. Broadway
Baltimore, Maryland 21205

John B. Garrison, Ph.D.
Building 1 E-216
Applied Physics Laboratory
Johns Hopkins Road
Laurel, Maryland 20810

Dear John:

This letter summarizes the points we discussed today which suggests that
eye movements should not interfere with the resolution of your proposed ultra-
sound scanner. You indicate that you seek a resolution for your ultrasound
localizing ability of 50μ. On this basis a displacement of objects in the eye
of 30μ will begin to interfere with the system's resolution. This amounts to a
displacement of objects at the radius of the eyeball (assumed to be 12mm) which
is equivalent to an angular rotation of 10 min arc. Between microsaccades the
eye drifts at a typical velocity of 12 min arc/sec. If your exposure time is
only 20 msec the eye would have only travelled through 0.24 min arc which is much
less than the amount that would interfere with your system's resolution. In
fact, your exposure time could go up to 100 to 200 msec without effecting the
resolution very much. The eye usually does not drift in one direction for that
long but turns around and drifts back with an average net displacement between
microsaccades of about 1.5 min arc, again well within your limit of 10 min arc.
Consequently, you could probably have an exposure time which lasted from one
microsaccade to the next, typically 1 sec, without degrading resolution. A
typical microsaccade is 10 min arc so that you could even tolerate a smaller
microsaccade occurring during the exposure time. During long fixations the eye
travels typically 6 min arc from mean eye position which means the total move-
ment would be less than 12 min arc most of the time, but peak drifts and micro-
saccades could be as large as 24 min arc. This suggests that you probably
would not be able to record continuously during normal human miniature eye
movements.

You have indicated to me that you have an optical device which can look
at the eye at a 60 cycle scan rate and determine movements of 1.0μ or 0.3 min
arc. Consequently, you can tell easily whether the eye has drifted by an
amount which will degrade resolution. It sounds to me as though you could use
this method to determine when the eye had been displaced by 10 min arc for any
reason due to slow drifts, microsaccades or translations of the eye due to head
movements. If your limit was exceeded, the exposure would be discarded and

2

you would take another picture. You have indicated that you would control for cardiac pulsation by timing your exposures during diastole and if the patients can maintain reasonable fixation it might even be possible to continue scanning the eye through several periods of diastole. Certainly eye movements will not be a problem for your 20 msec exposure time and so the only question is how long can the exposure time be before eye movements become a problem. You appear to have the means to detect such movements and answer this question.

Since microsaccades will probably be a limiting factor and because they occur in both eyes (although they are not always exactly conjugate) it seems to me that you could monitor fixation in the eye opposite to that being exposed to ultrasound. I would recommend that the head be held in something like a vacuum sandbag. The eye being monitored should be able to clearly see a fixation target since eye drift in the dark is larger than in the light by a factor of 5, and that the patient be asked to hold his or her breath and fixate a target spot during the period of the ultrasound exposure. With all these safeguards it seems to me that eye movements need not degrade the resolution of your ultrasound system.

I hope these comments are helpful to you.

Sincerely,

David A. Robinson
Professor

DAR:amc

3

B.   PROGRESS REPORT

1.   Background

We submitted a proposal to NIH in November 1978 on high resolution ophthalmic ultrasound, and subsequently found that there was interest at NASA in establishing a jointly sponsored program by NEI and NASA if both the medical and technical approaches, methods and objectives were found to be satisfactory. A proposal was therefore submitted to NASA. There was agreement among Wilmer Eye Institute, APL, and NASA that Study Section approval and NEI interest in funding were requirements for serious NASA consideration.

The study section was concerned about several aspects of the proposal and it was not approved. The primary areas of concern were:

a.   Effects of eye motion during the data collection time

b.   The overall complexity, cost, and time scale of the technical approach, and the performance that can be expected

c.   Clinical application of the techniques, if successful, because of cost.

Since this initial proposal, we have received funding from NASA, APL, and the Wilmer Eye Institute to investigate the areas which were of most concern to the study section when the initial proposal was reviewed.

In addition we received funding from the W. W. Smith Charitable Trust for correlation tracking of images on TV tape. As a result of this funding we have at APL a significant portion of the computing and display equipment required for the initial ophthalmic ultrasound investigations. This equipment will be in operation at APL this calendar year; with the exception of the deAnza Image Array Processor it will remain at APL, and will be available for use on the proposed program. The deAnza unit will be available for about a year.

2.   Progress and Pilot Demonstrations

a.   Effects of Eye Motion

The use of a switched array will reduce the data collection time to a few tens of milliseconds. We will monitor eye motion during the data collection using a correlation image tracking technique developed on another medical program at APL. Data collection will be synchronized with the EKG to take data during the late diastolic period of the arterial pressure pulse in the eye. We believe this method will eliminate the problems associated with eye motion.

This section describes the pilot demonstrations with a switched array, the correlation tracking technique, and a simulation program for the imaging technique.

4

Array and Data Recording Technique

Our initial proposal called for the use of a single transducer which was mechanically moved to generate the appropriate synthetic aperture. The data collection times would have been many seconds and the study section, quite correctly, was concerned about the effects of eye motion in this time interval. In our response to the study section request for additional information we said that we would prefer to use a switched array in order to reduce the data collection time to a few milliseconds and would investigate the feasibility of this approach as well as the possibility of using an eye tracking system to compensate, during the image computation, for the effects of eye motion.

We have established the feasibility of using a switched array, and the program is now based on this approach.

Figure 1 shows a block diagram of the data collection portion of the system as it is being implemented for collecting experimental data. The array was specially fabricated for the Applied Physics Laboratory by Aerotech. It is planar and has 49 elements arranged as a 7x7 matrix. The elements are on .3 centimeter centers. Each element is a square .15 centimeters on a side. The beamwidth of each element is about .48 radians and the elements operate at a frequency of 2.5 MHz in water.

Selection of the desired element is controlled by means of a series of relays that are part of a selector assembly. Physically, the array is mounted directly onto the bottom of the selector assembly which is fastened to a traverse mechanism. The selector assembly with the sensor attached is shown as Figure 2. A control processor, which consists of a commercially available microprocessor unit with modifications, is utilized to control the relays within the selector assembly and also for generating appropriately timed trigger pulses to fire the pulse generator within an Aerotech Ultrasonic transducer analyzer (Model UTA-3) to excite the selected element within the array. The electrical signal from the array, due to reflection from the object under investigation, is gated and amplified within the transducer analyzer and fed into a deAnza image array processor.

The image processor is a commercial unit for digitizing television images. It is capable of digitizing at about a 13.4 MHz sample rate. The system will be implemented so that when a selected element within the array is excited, and the return signal gated by the transducer analyzer, it will arrive at the image processor at the beginning of a selected horizontal line. The data will be digitized at the 13.4 MHz rate and stored within the image processor memory. Upon completion of storing one line, the microprocessor will select another element in the array and initiate the process to store data from it. This will continue until data has been collected from all selected transmit/receive element combinations.

The processor is interfaced to a PDP 11/34 minicomputer that controls the operation of the image processor and will read the data from the memory and write it on a magnetic tape for transfer to the Laboratory's 3033 computer for initial imaging computation.

Initial tests of the array indicate that it is peforming to specification. The sensor is configured as shown in Figure 3. The 49
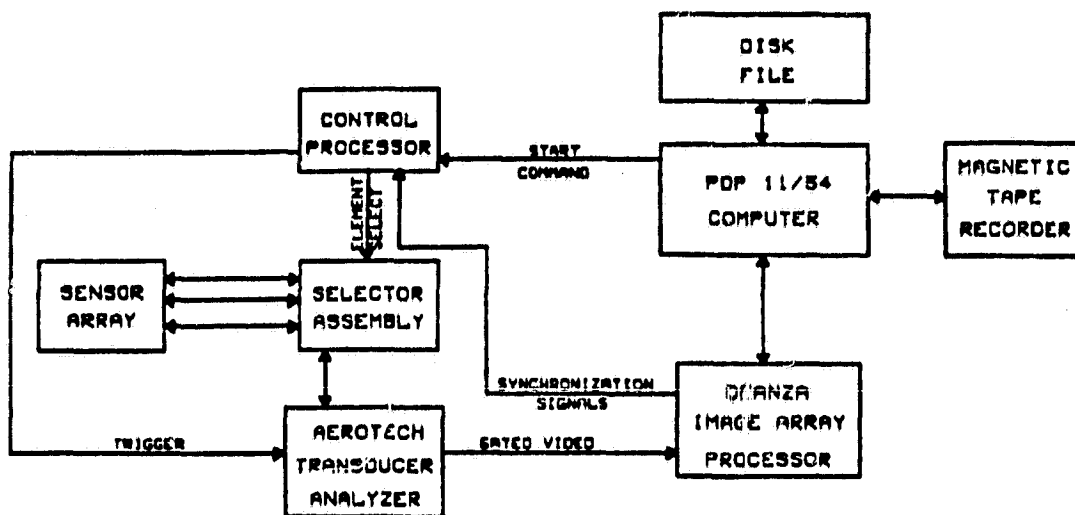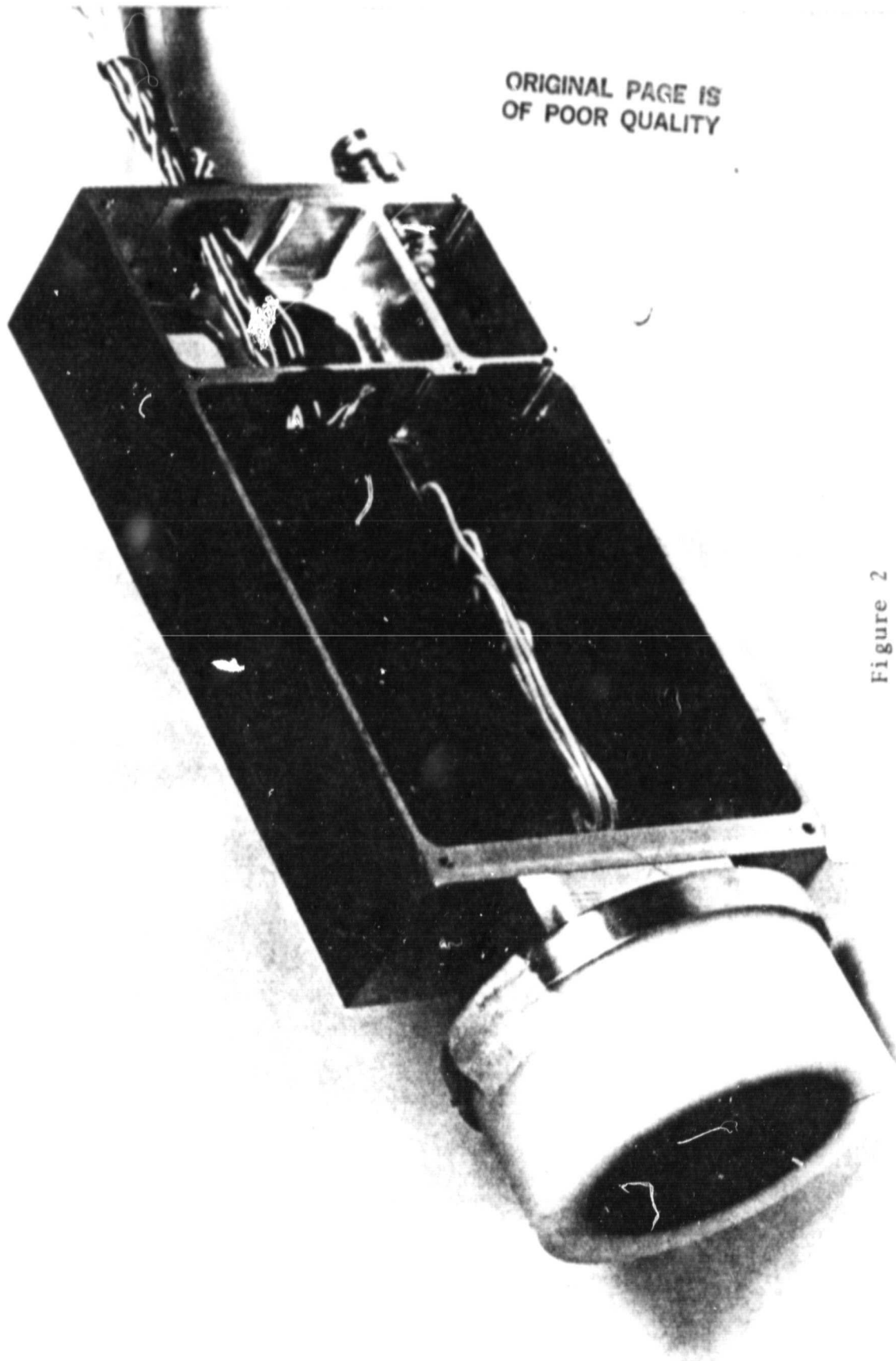
DATA COLLECTION SYSTEM
BLOCK DIAGRAM

Figure 1

Figure 2

Array and Selector Assembly

7

elements are electrically connected into a row-column addressing system. Also shown in the drawing are 14 relays that are required to select the desired element. These relays are controlled by logic within the control processor assembly and are under software control.

The relays utilized within the selector assembly were procured after considerable investigation. The transducer analyzer excites the elements by means of a narrow 300 volt pulse. It was difficult to find relays that could withstand this voltage without arcing over and were capable of operating relatively fast. The typical operation time for the relays is 1 millisecond. This limits us to a rate of about 1000 operations a second. As a result, to read out all 49 sensors will take something under 100 ms. When solid state switches are developed, this time will decrease significantly with the data collection time being controlled principally by the propagation time of the signals within the water. Development of the switches is made difficult by the need for them to provide isolation to the high voltage fast rise time pulse while still presenting a minimal loss for the low level return signal.

In implementing the system, even with the relays, some difficulty has been encountered due to the electrical capacity between the relay contacts and the coil passing the transmit pulse through the control processor and causing it to temporarily malfunction. To solve this problem, an additional assembly has been placed between the selector assembly and the controller processor using electro-optical isolators to provide a high level of isolation between the control processor and the signals from the transducer analyzer.

Another function of the control processor is to provide a trigger signal with the proper timing relationship to the synchronization signals generated by the image array processor. As part of this task, the control processor assigns a horizontal line within the image field to hold the data from each element of the sensor array. The control processor, after properly positioning the relays, generates a trigger pulse to the transducer analyzer at the appropriate time so that the return signal will arrive at the image processor at the correct time to be stored in the desired line. Figure 4 is a simplified schematic of this circuitry.

The current system as shown above only allows reception from the element which was used for transmission. We have an approach that will allow transmission on any element and reception on any element. The required number of relays are already provided in the selector assembly as shown in Figure 2, but the required isolation amplifiers have not been installed. This will be accomplished following successful operation of the basic system.

We had expected to have some results with simple targets by the time of this report. We were unable to do so primarily because of hardware delivery schedules. In the past year or two hardware delivery times have skyrocketed because of severely curtailed inventories resulting from the inflationary economic environment.

Eye Tracking

Our initial proposal required an eye tracking system. With the change to a switched array and data collection times of a few milliseconds it is not necessary to track eye motions for compensating
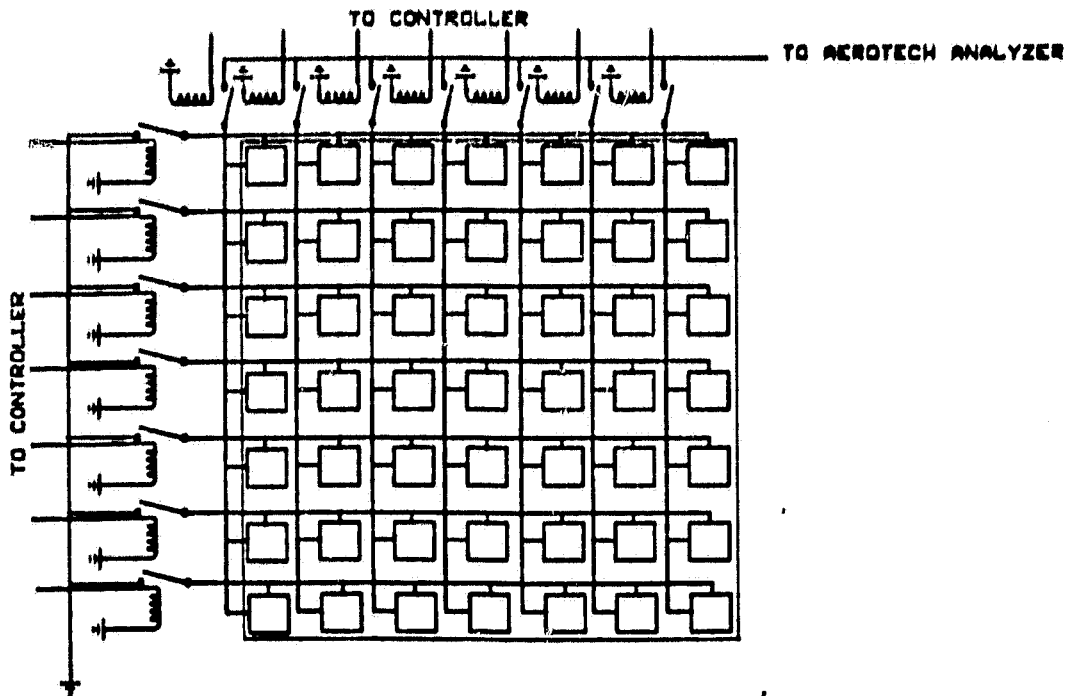
SENSOR ARRAY CONFIGURATION
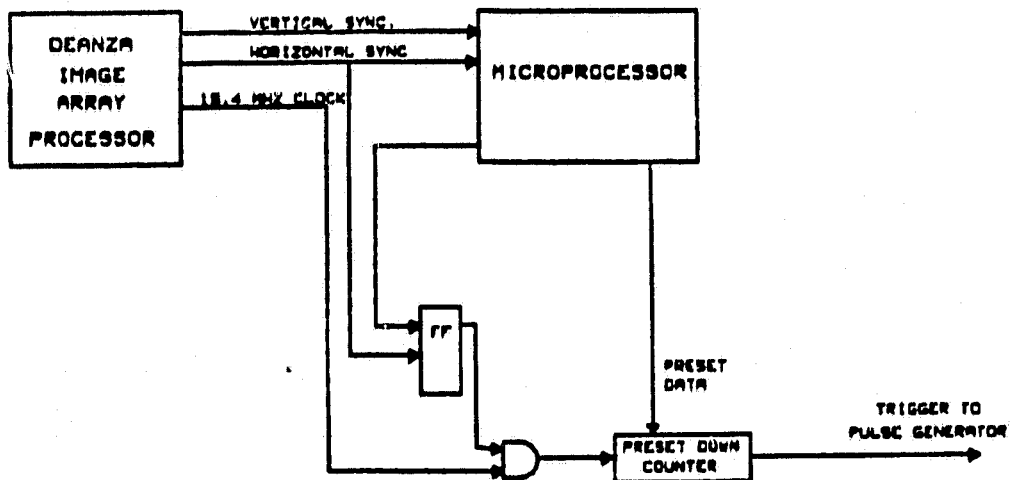
Figure 3

TIMING CONTROL

Figure 4

in the image formation process. Our general approach is to monitor eye motion during the data collection interval and to accept, or reject, this block of data if eye motion exceeded a threshold which could be expected to degrade the resulting image.

Thus, although we believe this eliminates the need for using the results of a tracking system in the calculation of the images, we still require a precision method of measuring small eye motions. Since submitting our initial proposal, we have made a reasonably complete survey of existing eye tracking systems and did not find any inexpensive systems which meet our requirements.

Since our initial proposals we have, on another program, developed a correlation tracking method to follow the movement of elements of a time series of pictures from one picture to the next. The source of the pictures is immaterial. One or more "control points" on the initial picture is selected for tracking through the sequence.

The translation of each individual control point is followed from picture to picture. A small square centered on an individual control point is selected from the current image field. This defines the local feature is called "feature window". A somewhat larger square, centered on the same picture coordinates, from the next image defines the "search window". The small square is moved within the search area, comparing the two images until the position of maximum correlation is found. The peak of the correlation function is located to a fraction of a pixel by quadratic interpolation. This defines the new position of the control point in the new field and, thus, measures its motion. This field becomes the current field, and it then defines the local features for each control point. We move to the third field in the same manner, and so on.

This process is similar to so-called "map matching" techniques, except that in our applications the "map" is changing (hence, de-correlating) with time. The solution is to make the latest image field define the local features for each point. In this way we are faced with the smallest possible de-correlation of the image as we search for the translation that has taken place.

In a program from NHLBI (Grant #NL24357)we have a requirement to accurately and rapidly track the motion of the order of 100 1 mm diameter spherical markers implanted in a canine heart. The dog is x-rayed with a standard 35 mm biplane system at up to 90 frames per second. Earlier this year we conducted a pilot study to determine if the correlation tracking system would perform this function and to obtain a measure of the absolute accuracy of the sytsem.

A test was made of the accuracy of the correlation tracking method in which we compared the positions of the images of the markers as measured on a digitizing table with those determined by the correlation method. This gave an RMS difference between the methods of 0.05 ball diameters; that is 1/20 of the diameter of the image of the balls on the film.

In this case the field of view of the x-ray system is the order of 10 cm at the dogs heart. The markers are 1 mm in diameter. Thus, in terms of the location accuracy on the film, they were located to .005 cm

in a field covering 10 cm. This is one part in 2000 of the field of view.

We plan to track the motion of the limbus by using this technique. The use of the limbus for measuring right-left position of the eye is common. In the majority of eye tracking systems it cannot be used for up-down motion because the boundaries are covered by the eyelid. In our application, the eyelids can easily be held so that it can be used for both up-down and right-left tracking.

The limbus is approximately one cm in diameter and has well-defined edges. We assume that the image of the limbus will be approximately centered in the picture and its height will be one half of the total picture height. The field of view in this case is 2 cm by 2 cm; in the case of the multiple markers it was 10 cm by 10 cm. Thus, the 1 mm markers would scale to a diameter of 0.2 mm in this case. For each such small area taken around the limbus we can expect a tracking accuracy of 1/20 of its diameter=0.01 mm. For the limbus we have a well-defined edge; the accuracy will probably be poor along the circumference, but it will be very good normal to the circumference.

We can use much larger areas for correlation, the limbus moves more or less as a rigid body and the optical image quality will be much higher than the x-ray images, we are confident that the tracking accuracy will be a few microns.

Standard TV systems operate with a frame rate of 30 per second with each frame consisting of two interlaced fields. We will have, therefore, at a minimum, a data rate of 30 per second. It is relatively easy to use field data thus providing a data rate of 60 per second. We plan to record continuously, beginning a few frame times ahead of each data collection interval and continuing for a few frame times after the data collection interval. We expect this to be adequate.

The use of this method provides us with an accurate method of measuring eye motion without significant equipment costs to this program. We will need only a simple lens system for immersion into the saline bath used with the echo system.*

Simulation

We have in operation a computer simulation of the switched array system. This is a relatively complete simulation for the case of N point targets, of variable strength and location, with a three-dimensional ultrasound array, and in a uniform lossless medium.

The conditions are specified by a 29-element vector permitting a wide choice of parameters. This simulation is written in APL and runs on the APL central computer from a remote terminal. This permits trial of various parameters and the tabulation of results in a few minutes of elapsed time. Its primary use to date has to quantitate a variety of choices of system parameters by computing the spatial impulse response of the system under ideal conditions.

---

* See letter from Professor D. A. Robinson (pages 2 and 3 of this report).

b. Overall Complexity, Cost, and Time Scale of the Technical Approach, and the Performance Which can be Expected

The study section was quite correct in concluding that the proposed program is difficult, and as for any research program, it is not easy to guarantee a successful outcome.

We believe that we have made very significant improvements in the definition and explanation of the program and that pilot studies have gone well. The availability of additional equipment at APL eliminates the need for purchase of several major items, and we have a well-defined step by step program that can be carried out and which will provide valuable data.

c. Clinical Application of the Techniques, If Successful, Because of Cost

We failed to make it clear in our initial proposal that the NASA High Speed Recorder was to be used only at APL.

Our estimate of the cost of a full system after completion of the development program is approximately $250,000. This puts it in the same price range as CT (computerized tomography) systems which have widespread clinical use.

C. SWITCHED ARRAY SYSTEM DESCRIPTION/ANALYSIS

1. Introduction

We are proposing the use of digital synthetic aperture techniques to obtain high resolution ultrasound images of the eye and orbit.

An understanding of the <u>principles</u> of synthetic techniques is essential in order to judge the methods we are proposing. This section presents the explanation of the necessary principles for our application.

The most common application of synthetic aperture processing involves the use of a system in an aircraft or satellite to form high resolution images of the ground. In this case the motion of the radar provides a look at a particular object on the ground from different directions as the radar passes overhead. In this application the fact that the source is moving from place to place at a relatively high speed provides the synthetic aperture. The bulk of the analyses and explanations of such systems are, quite naturally, based on the changing doppler frequency as the radar passes overhead. This approach does not, unless the reader is already thoroughly familiar with such systems, provide an easy understanding of the techniques we are employing, and tends to obscure the fact that it is <u>not</u> the speed of the source that is important; it is the <u>position</u> of the source at each transmission that is important. In our case, using an array of elements with electronic switching, it is simpler to explain <u>without</u> involving the speed of the source in the explanation.

2. Discussion of General Principles, Techniques, and Approach

Summary

The objective is to obtain high resolution in all three-dimensions. We want resolution of the order of the wavelength of the ultrasound signal. This permits the use of the lowest frequency for a given resolution thus minimizing the effects of attenuation which increases rapidly with frequency.

The general approach is to "pay" for improvements in performance by the use of high speed digital recording and processing. High speed is a relative term which changes with the state of the art, and needs to be quantified.

At the present time commercially available units developed for TV applications have digital memories which hold $512 \times 512 = 262,144$ 8-bit words and typically can read these words in or out at a rate of 10^7 words per second, in a predetermined (hard wired) way. More bits and larger memories, usually in multiples of the basic TV size are readily available.

This is not "state of the art", which usually implies special design and development but is closer to production equipment. The memory size for recording the data in general corresponds to the number of resolution cells produced after "synthetic aperture" processing of this recorded data.

"State of the art" A/D converters and shift registers are not difficult to obtain and use up to at least 50 MHz which is adequate for the highest ultrasound frequencies used for the eye.

This means that the approach we are investigating requires essentially no component development other than the array itself. In this area we have been pleasantly surprised by the ready availability and low cost of the experimental 49 element array which operates at 2.5 MHz. It appears that by the time we want a 10 MHz array it will be readily available.

The array requirements for synthetic aperture systems are, in general, simpler than for "on line" systems. A given synthetic aperture array can be used to obtain many different types of beam forming by changing the software. This is not the case for hard wired special purpose arrays.

Amplitude weighting on the aperture can be varied to optimize it for the case at hand, since all the data is stored and can be reprocessed with different weightings.

There is every reason to believe that digital equipment will continue to decrease in cost for a given performance.

Electronic switching for our application is very similar to that required for current two-dimensional phased array ultrasound systems and is not expected to be a major problem.

We are primarily interested in developing a useful system by using existing components in a new way; not in developing new components.

3. Synthetic Aperture Techniques with a Switched Array

The following discussion is quantitative but no attempt is made to do precise calculations. The primary intent is to explain the methods and provide a basis for understanding the effects of changing various parameters.

Figure 5 shows the geometry.

The geometry is three-dimensional. There are N individual transducers distributed over a sizeable solid angle (e.g., 30° half angle cone) as seen from the center of the volume to be examined at high resolution. This volume is assumed to be spherical. Its position within the eye is changed by moving the array with respect to the eye. We will use a simple manually-adjusted fixture to hold the array.

Each element is assumed to be focussed at the center of V and to have a beam width and depth of focus equal to or greater than the diameter of V.

Each element is able to deliver a reasonably (e.g., ± 2 dB) uniform amount of power passing through V, and be able to receive scattered signals originating anywhere within the volume, again with only small variations in gain over V.

14

TRANSMIT/RECEIVE
ARRAY ELEMENTS

Ro

ELEMENT BEAMWIDTH

SPHERICAL VOLUME
TO BE IMAGED WITH
HIGH RESOLUTION
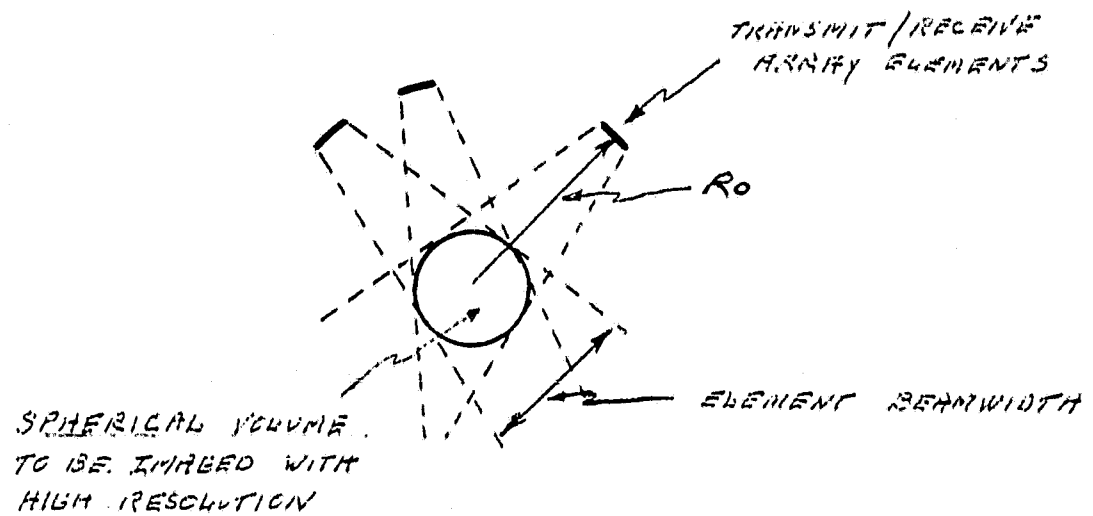
ARRAY CONFIGURATION

FIGURE 5

Note that increasing the distance (Ro) from the center of V to the array permits the elements to become larger to maintain the same beamwidth. If this is done then the same fraction of total transmitted power (and scattered power) is still received at each element.

Since, in general, the transmitted power capability of an element increases with its area, the performance of the system from a signal to noise ratio standpoint can be improved by placing larger elements further away.

In cases where Ro is several (e.g., 5) times the diameter of V, the change in distance between closest and most distant part of V can be neglected for present purposes. However, in the extreme case of the elements on the surface of V, range effects are far from negligible.

In this discussion we assume a resolution of the order of one wavelength in each direction. Rates of and quantities of data collected are discussed in terms of a standard 512x512x8 bit digital TV memory which operates at or about a 10 megaword per second rate. This does not imply the use of one memory as a limit; it is just a convenient size unit of complexity in discussing general principles. This means that in order of magnitude we can collect 262,000 8 bit samples per scan, and again in order of magnitude, process these to obtain 262,000 resolution cells in the volume V. To avoid unnecessary complications in arithmetic we assume that this will provide a high resolution volume which is 64λ in diameter ($64^3 = 512^2$), and we ignore the difference between a spherical and cubical high resolution volume. At 2.5 MHz, $\lambda = 0.6$ mm and V would be about 38 mm in diameter; at 10 MHz, it would be about 9.5 mm in diameter.

On the other hand, the best resolution that can be expected, for an array that fills a hemisphere is $\lambda/2$ for a real aperture and $\lambda/4$ for a synthetic aperture. If we were to take these and apply them to the highest frequency in the spectrum of the transmitted pulse (which is about twice the center frequency), we would have a linear resolution about 8 times better and the diameters of V would drop to about 5 mm and 1 mm. If we need two samples per resolution cell in each direction, V would drop in diameter because of our assumption of memory size.

We point this out only to indicate the variation in numbers that can arise with apparently small changes in some of the assumptions.

Consider the case where we want to examine some volume, V, to measure reflections in this volume.

This "array" can be used in a number of different ways, but it has two primary functions.

1. To illuminate the volume in order to provide a scattered signal to analyze when received.

2. To receive this scattered signal and, using the prior knowledge of the transmitted signal, provide a measure of the reflecting properties of the medium.

There are several extreme methods of using such an array.

A'. The array can be focussed both on transmission and reception using all N elements, and by repeating the sequence of transmission and reception for each resolution cell in the volume obtain an echo picture of the volume. This "scanning" is required since in the transmit mode the array illuminates only a small portion of the total volume.

B'. Any one element can be used to illuminate the entire volume with one transmission and the received signals from all elements combined to provide information on all resolution cells.

C'. Method B can be repeated, using each element for transmission, in some sequence, and for each such transmission forming all beams. If the results for each resolution cell for each transmission are stored (without detection) these results can be combined in a second stage beam former.

Before proceeding, some general remarks about these three methods are in order.

It is clear that in cases A', B' and C' the same results will be obtianed if instead of using an "on-line" beam former as is implied, the received signals at each element are digitized and stored for later execution of the beam forming operation. Thus, the methods A', B' and C' can be "transformed" to the equivalent methods: A, B, C.

A. Focus the array, on transmit, on each resolution cell in turn, digitize and store the return signals at each element.

B. Transmit on one element, illuminating the entire volume, digitize and store the return signals at each element.

C. Repeat B for each transmit element, digitize and store the returns from each element for each transmission.

So far it is evident that it is immaterial from a performance standpoint whether the received signals are used immediately for beam forming or are stored for later beam forming. It is assumed that the storage is done by digitizing the return, at an adequate rate, and storing these words in any convenient fashion. This requires adequate attention to timing accuracy and dynamic range requirements. Both requirements are well within the state of the art.

The corresponding equivalence between using all elements simultaneously or sequentially on transmit raises a different set of questions, which we now consider.

Before proceeding it is necessary to emphasize that there are two important, qualitatively different, conditions which must be taken into account in image forming systems.

These are:

1) Signal to receiver noise limitations

2) Signal to clutter limitations.

17

Receiver noise is typically "white" noise and limits what can be seen unless the ratio of signal power to noise power is sufficiently high. In cases of an inadequate ratio, the performance can be improved by increasing transmitted power or reducing receiver noise. If this is not possible, then for a stationary assembly of targets and array elements, repetition of the full scanning sequence and integration of the results will improve performance. If this integration is coherent the signal to noise ratio increases directly as the number of repetitions; if it is incoherent (that is, following detection), it improves approximately as the square root of the number of repetitions.

Signal to clutter ratios refer to the fact that when a system is focussed on a given (small) volume element, signals coming from other points in the object will appear to come from the focal point because the focus can never be perfect. In this case increasing the signal to receiver noise ratio has no effect. It is this type of "noise" that poses the most difficult problems in imaging systems.

We now return to the consideration of simultaneous vs sequential illumination and compare cases A and B in preparation for the other more complex caess. This is done for both receiver noise and clutter.

In order to compare signal to receiver noise ratios in the different cases it is necessary to consider the power available among the different cases. We assume that a reasonable basis is that the peak power per element is one of the limitations; another is the power delivered per unit volume of tissue.

Tissue damage phenomena are not well understood and there are currently no clearly defined and agreed-upon criteria. It is clear that thermal heating can be a source of damage; there may be damage thresholds due to peak power and/or resonance phenomena. For present purposes we will take "average power" as one measure of relative safety, with peak powr as a second consideration. It is clear that the distinction between peak and average power requires consideration of the thermal time constant of the volume receiving the power.

To consider signal to receiver noise ratios it is sufficient to consider the problem of locating a single point reflector in the high resolution volume. We assume that there are N elements in the array, each capable of transmitting a peak power P for a length of time T, and assume that the time T is a small number of cycles of the resonant frequency of the transducer. With the elements distributed in a volume which subtends a substantial solid angle as seen from the volume V, the focal spot size is the order of one wavelength in diameter, like it or not. This can easily be seen from consideration of the limiting case of an array which subtends the full 4π solid angle. In this case if the array is focussed at a point, moving a quarter of a wavelength will cause the signals arriving from diametrically opposite directions to be 180 degress out of phase and therefore cancel.

The requirement for a relatively large solid angle subtended by the array essentially rules out the use of conventional dynamic focussing techniques. Such techniques can be used only on receive; they are not applicable in the transmit mode. They are useful primarily when high clutter rejection is not required, and accurate location or resolution of only a few targets is required.

18

When any one element is used for transmission, assuming no amplitude weighting to reduce side lobes, about half of the transmitted power will be in the main beam. Neglecting this factor, the power density at any point in the volume is (power transmitted)/(maximum cross sectional area of V). The transmitted power per element is P, and the maximum cross-sectional area of V is (total number of resolution cells$^{2/3}$xλ^2. For our "standard" case this is $(64)^2\lambda^2$.

Thus, for one transmit pulse from one element the energy density passing through any point in V is $(PxT)/(64)^2\lambda^2$.

If the scattering cross section of an object in a resolution cell is $\sigma(cm^2)$, the scattered energy will be σ times the energy density. The fraction of this, assumed isotropically scattered, energy collected by one receiving element is $(A/4\pi R_o^2)$ times the scattered energy. (Note that for the usual case the element size varies with Ro so that $A/4\pi R_o^2$ is constant.)

When N elements are used for transmission and they are focussed at a point the energy density is N^2 times as large as when a single element is used. (The total transmitted power is increased by N, and the focussing adds the second factor of N.)

The use of N elements on receive, assuming the beam former focusses them, provides an increase in received energy of N.

Thus, the signal energy in case A, for each cell, is

$$\frac{N^2 x(PxT)}{(64)^2\lambda^2} \times \sigma \times \frac{NxA}{4\pi R_O^2} \qquad .$$

For case B a single transmitted pulse illuminates the entire volume on each such pulse, the received energy is

$$\frac{PxT}{(64)^2\lambda^2} \times \sigma \times \frac{NxA}{4\pi R_O^2} \qquad .$$

Thus, to have the same signal to receiver noise ratio, case B must be repeated N^2 times with the result being integrated coherently.

It is now informative to consider the ratios of some characteristics of modes A and B when they are operated to produce the same signal to receiver noise ratios.

Ratio of number transmission cycles A/B = $(V/\lambda^3)/N^2$

Ratio of total energies delivered to V.   A/B = $(V/\lambda^3)/N$

Ratio of peak power density in V.   A/B = $N^2/2$.

We next examine the target to clutter performance of systems A and B.

In order to make the comparison in such a way that the same method can be used in considering case C, it is necessary to explain what happens. For this discussion we assume that the volume V is made of a random distribution of scattering elements and that the objective is to measure the scattering properties of each cell to some reasonable accuracy.

Figure 6 shows the geometry for the case of three elements. As always the elements are really in three-dimensions and V is a spherical volume. A planar section is shown for simplicity.

In A of Figure 6 we show the volumes which would contribute signals arriving at the same time as the signals from the focal spot if the elements were used one at a time for transmission and reception. These are marked 1,1 2,2 and 3,3 to indicate that transmission and reception occur at the same element. Each such volume is a curved sheet of thickness equal to the pulse length and, usually, with an area equal to the cross sectional area of the transmit (or receive) beams which we already know have an area of $(64)^2 x\lambda^2$. Thus with a 1λ pulse the volume is $(64)^2 x\lambda^3$ which is $(64)^2$ times the volume of the focal spot.

Thus for the sequential operation of elements 1, 2 and 3 and with reception only on the transmitting element each time we can add these three returns to be coherent at the focal spot.

Since the additions at the focal spot are coherent we gain a factor of 3^2 in the returned energy, thus getting $3^2\sigma$.

We now consider the other volumes which are not common to the focal spot.

For the returns at each element there are $(64)^2$ such returns. Since these elements contain random point scatters the energy will add as the number giving an energy of $(64)^2\sigma$. When these are added in the beam forming operation, we are adding three different sets of random samples, whereas at the focal spot we are adding three samples of the _same_ thing.

Thus the enegy from the focal spot is increased by N^2, and the clutter energy by N, giving a net gain of N.

B of Figure 6 shows the corresponding volumes for transmission on element 1 and reception on elements 2 and 3.

These are the volumes for which the signal will arrive at these elements at the same time as the signal from the focal spot. In the case of transmission and reception on the same element these are defined by a sphere with origin at the element. In the other cases, e.g., 1, 2, they are ellipsoids with foci at elements 1, 2.

Thus, in the case of N elements, if we record on all elements when we transmit on one element we obtain $1/2(N+1)xN$ new samples. We do not get N^2 since transmitting on m and receiving on n, i.e., m,n the same as n,m.

This figure also shows the time delays which must be used in the beam forming calculations for modes A, B and C.

In mode A, the transmit time delays must be such as to cause the signals to arrive simultaneously at the focal spot. If these time delays are T1---TN on transmit, the same time delays are used in the receiver. One of these time delays can be taken at zero without loss of generality, making the set T1-T2--0--TN. The receive time delays will then be a set which is some constant K (representing the round trip time from element with zero time delay on transmit) greater than the set of
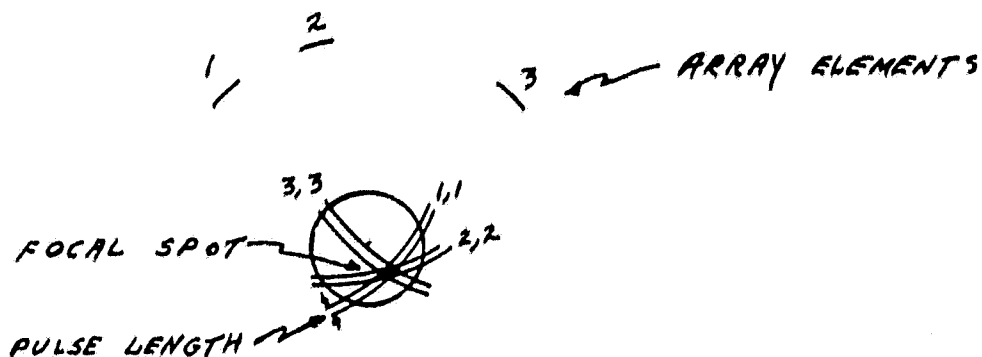
The spherical shells of thickness equal to the pulse length are the volumes which contribute clutter signals which add to the desired returns from the focal spot.

The notation m,n indicates transmission on element m and reception on element n.
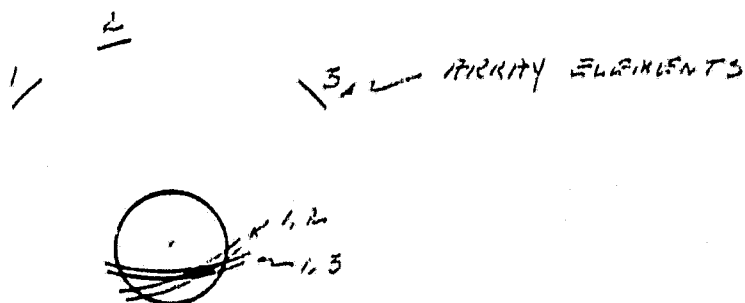
When m=n the shells are spherical.

Figure 6

Same as above except for the case of transmission on element 1 and reception on elements 2 and 3.

In this case the shells are ellipsoids with foci at the transmit and receive elements.

transmit delays. With this notation when element n is used for transmission the time delays on the reception are K plus the set

$$(Tn+T1), (Tn+T2) --- (Tn+Tn) --- (Tn+Tn)$$

and the overall calculations are not more difficult; there are just more of them.

Figure 6B also applies directly to case A, in that all combinations of transmit and receive are handled simultaneously on a single transmission.

All signals within the focal spots of the transmit and receive arrays are added coherently; those outside are assumed, in this model, to add incoherently.

There is a difference between modes A and C. In case A, the arriving signal in the focal spot can be written as

$$I = NxAx \sum_n e^{j\phi_n}$$

where A is the amplitude of the arriving signal from each transmitter which is assumed to be the same for each element. The beam formed signal from this focal spot will be

$$S = I \sigma \sum_o^N e^{j\phi_{\bar{n}}}$$

$$= NxAx\sigma x (\sum_n e^{j\phi_n}) \times (\sum_n e^{j\phi_n}) \quad .$$

That is, the arriving illuminating signals are summed in space. This sum is multiplied by σ and the received signal are summed. Thus, the product of the two beams is involved. The signals at the receiving elements are the terms in $\sum e^{j\phi_n}$, each multiplied by $(\sum e^{j\phi_n})$.

In the synthetic aperture case, where all the signals are individually recorded, the available signals are elements of the double sum

$$A \sigma \sum \sum e^{j\phi_n} e^{j\phi_n} \quad .$$

In the synthetic aperture system the focal spot is about 1/2 as large in each dimension because the phase shifts are additive.

We now consider the numbers for both systems to obtain a reasonable measure of scattering power per cell in the case where V is filled with a random collection of scatterers.

We found in a previous section that the signal from the focal spot increased directly as the number of samples for which the return from the focal spot added coherently, but the clutter added incoherently.

Starting with a single sample (i.e., transmit receive on one element we found the signal to clutter to be $1/(64)^2$.

If we use N elements we buy a factor of $1/2(N+1)xN$. Thus, if we want a signal to clutter ratio equal to 10, we have

$$\frac{1}{2}(N+1) \times N \geq (64)^2 \times 10 \quad .$$

This says that about 256 elements are required. The required number decreases linearly with the diameter of the high resolution volume. This comes about since we are assuming that the elements are designed to give a reasonable drop-off in power outside the beamwidth necessary to illuminate the volume V.

In the majority of medical applications the objective is to see well-defined structures embedded in a relatively uniform medium.

This is a considerably easier problem than the one of accurate local measurement of pure clutter which was the case we used in the discussion of clutter rejection.

For high resolution systems, the switched array synthetic aperture methods have many important advantages over more conventional systems.

At a given wavelength they provide higher resolution; for cases where attenuation limits the maximum frequency this is important.

Data collection times, and peak and average power delivered to the tissue being examined are orders of magnitude lower than for a "real" aperture system with the same resolution.

4. Instrumentation Considerations

For this case we assume a 49 element array of the type shown in Figure 5, and consider both immediate and near term instrumentation approaches.

We assume a frequency of 2.5 MHz and a sampling rate of 10 MHz. The resolution volume is taken to be 16λ diameter ($=16 \times .6 \approx 1$ cm). The round trip time through this 1 cm distance is $2/1.5 \times 10^5/\text{cm/sec} = 13$ microseconds at a sample rate of 10 MHz. This gives 130 digital words for each path through V. With a 49 element array the number of such paths is (25×49)=1225 and the total number of sampless is $130 \times 1225 \approx 160,000$. This number will fit "comfortably" in the 262,000 word deAnza memory. For the case of higher ultrasound frequencies the data is digitized at the appropriate rate, buffered and read into the memory at a 10 megaword persecond rate.

With our existing equipment which uses mechanical switching and direct digitizing by the deAnza system, we can obtain only one set of return data for each transmission, thus requiring two milliseconds for each set of returns. The total scan time will be about two seconds. This is satisfactory for all of the early investigations.

Without switching limitations it would require somewhat less than a full frame time of the deAnza memory (33 ms) to load all the data. If we take 5 kHz as a reasonable upper bound on the prf of the array in order to avoid second time around echoes, this by itself would make the minimum time 10 ms.

With electronic switching on the average there will be 24 receive elements that require recording for each transmission. Since these signals arrive at about the same time, the straightforward

solution is to use 25 A/D converters and buffers and transfer the contents to the deAnza memory prior to the next transmitted pulse.

It is premature to attempt to make a recommendation on the best configuration for a "final" system. This can be done only after we have better numbers from the initial year's work.

Indications are at this time that by gating with the EKG (to avoid the time around the peak of the arterial pressure pulse in the eye). It may be possible to operate over several seconds using from 1/3 to 1/2 of the total interpulse period. The results in this area will influence the design choices.

One approach we will certainly investigate is the use of one bit operation. In cases where the signal to noise ratio is low and where a large number of samples in which the noise is random from sample to sample are to be summed one bit A/D conversion prior to summing has many advantages. It can be shown that in such cases linear signal integration results and if the number of samples is large, the resultant signal to noise ratio and dynamic range are large.

If, for example, the signal to noise ratio is -10 dB prior to A/D conversion, and the output of the A/D converter is 0 or 1 (corresponding to the input level being less than 0 or greater than 0) the output signal to noise ratio after summing 2^N signals will be improved by 2^N. If $N=10$ ($2^N=1024$), the signal to noise ratio will be about +20 dB. The approach we are taking in this program has the characteristics which make one-bit operation attractive.

5. Image Computations

The basic requirements for image formation may be explained with respect to Figure 7. If we want to determine the contribution from point target A in XYZ image space, we first compute the two-way time delay TAR and then sample the signal received at R at the corresponding time following a transmission from T. We repeat this process for all desired transmit-receive pairs and sum the answers. For the single transmit-receive pair shown, it should be obvious that a target located anywhere on an ellipsoid, with focal points at r and t, passing through point A will produce a signal at the same delay, but that the ellipsoids (or spheres in the case of common transmit-receive elements) corresponding to other transmit-receive pairs will only intersect at point A within the image space. Thus, by sampling at the correct time following each transmission as defined by the transducer/image space geometry, a unique point is defined such that all returns from the target at this location add directly, or coherently.

We will, of course, want to repeat this process for many cells in the image space in order to form a complete image. Other than a desire to space contiguous cells at a separation not exceeding the system resolution, we can arbitrarily choose any coordinate system to define the cell locations. The choice of equally spaced cells in a rectilinear XYZ coordinate system will be used to further explain the image process, but is in no sense a requirement. The signals available for beam forming will be stored in memory in the deAnza system, each return signal occupying a horizontal scan line in the pseudo-TV format. The signals will have been digitized at a 13.3 MHz rate. If, for example, we want to form 10 equally spaced cells beginning at point A
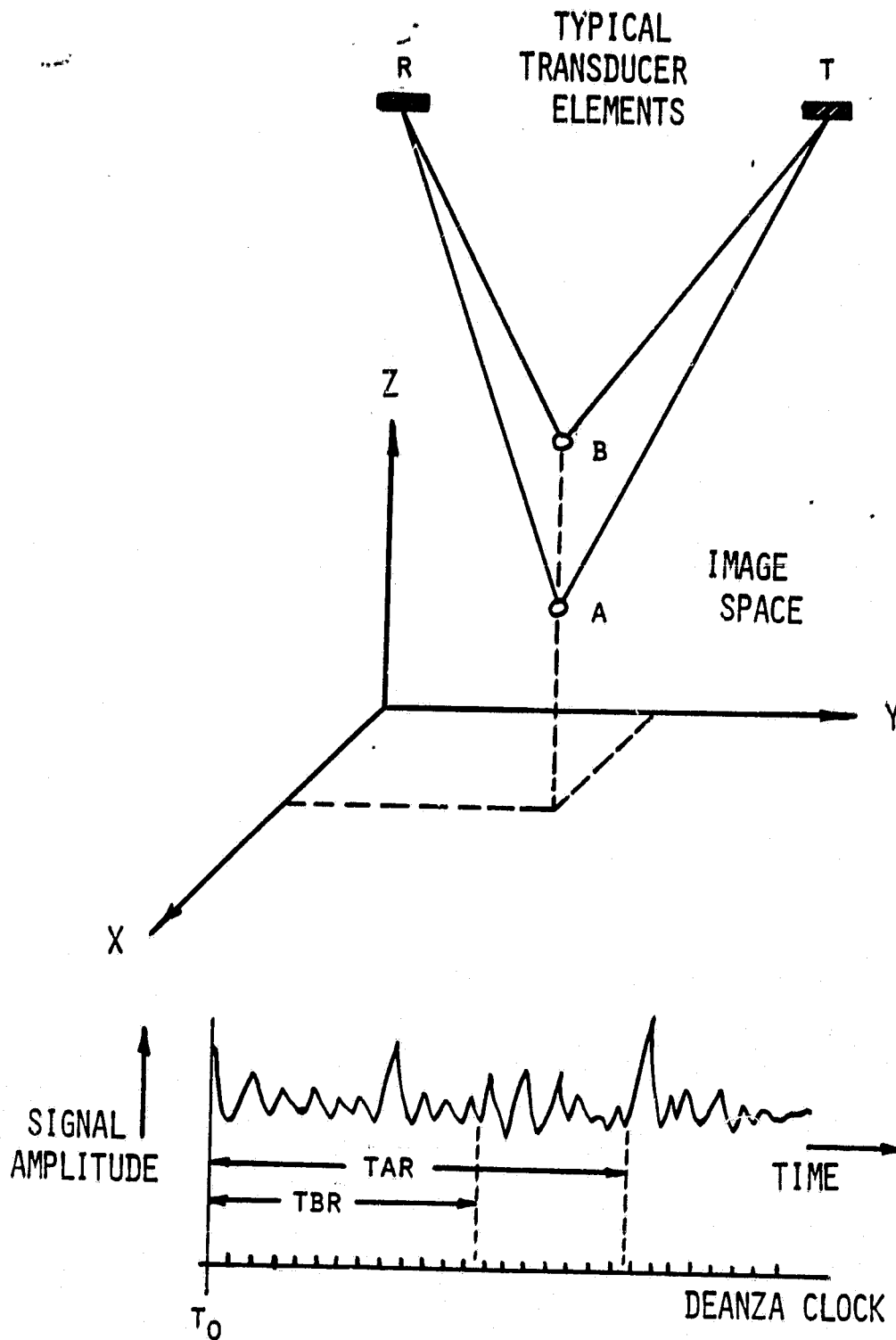
Figure 7

and ending at point B an immediate question arises: How do we assign signal values to the 10 cells when the available digital samples do not occur at the desired times? Even though the signal has been sampled at a high enough rate in the Nyquist sense, we desire to time align the various returns to the order of 1/16 wavelength prior to coherent summation. For the transducer now being used in the pilot study, the carrier frequency is 2.5 MHz and thus a minimum sample spacing of 400/16=25 ns is indicated. Thus, we would like to interpolate at least 3 additional samples between each sample in memory. Later, when a 10 MHz transducer is incorported into the system, a high-speed analog-to-digital converter and buffer will precede the deAnza memory to allow data to be input at a slowed down rate and thus require no furthur modification to the system timing.

The required interpolation could be accomplished digitally and consists, in exact form, of distributing the signal in each sample over adjacent samples with a $(\sin \pi t/T) \div (\pi t/T)$ weighting, function centered at the sample time where T is the sample spacing. It will be observed that this function causes no signal to be added to the original samples at intervals of T, but does specify the contribution at any intervening point from nearby samples. Each interpolated sample thus requires several multiplications and additions. We have concluded, subject to a more detailed study, that the interpolation process and the reclocking required to generate a fixed number of signal samples over a variable time span is best accomplished externally to the deAnza system. The latter requirement derives from the fact that the time separation for targets at points A and B in Figure 17 will differ in general for every transmit-receive transducer pair, as will the timing of the initial sample at B. Likewise, the timing for other cell columns along Z for different X-Y locations will have unique timing requirements. A further complication derives from the observation that the signal return for targets located at equal increments along the Z axis will not in general be separated by equal increments in range for a given transmit-receive pair. Thus, a processing system adaptable to a variety of transducer/image space geometries will need to accommodate a wide variation in individual range sample timing to insure that a sample may always be taken within $\lambda/16$ of the time computed from the geometry.

The proposed system is illustrated in Figure 8. Received signals from the array will be applied in analog form to the deAnza image processor. Signals from each transducer receiver will in general be stored on one horizontal scan line in a pseudo-TV format. The option exists for adding successive returns from the same transducer pair into the same memory location, achieving coherent integration. When Memory 1 is filled with the desired amount of raw data, which may consume several TV frames, data input to memory is inhibited and image formation commences.

Various sequences of operations are possible in order to form an image from the basic data. The process involves assigning a signal value to each cell in the image space resulting from a properly timed signal sample for each transducer-pair. Since the deAnza system reads signals out as sequential time (range) samples, it is natural to select a contiguous set of image space cells for updating during each line scan such that a monotonic sequence of more or less uniformly spaced samples may be applied to the sequence of image space cells. An obvious choice is a set of cells aligned along the Z-axis (Figure 7) at a prescribed
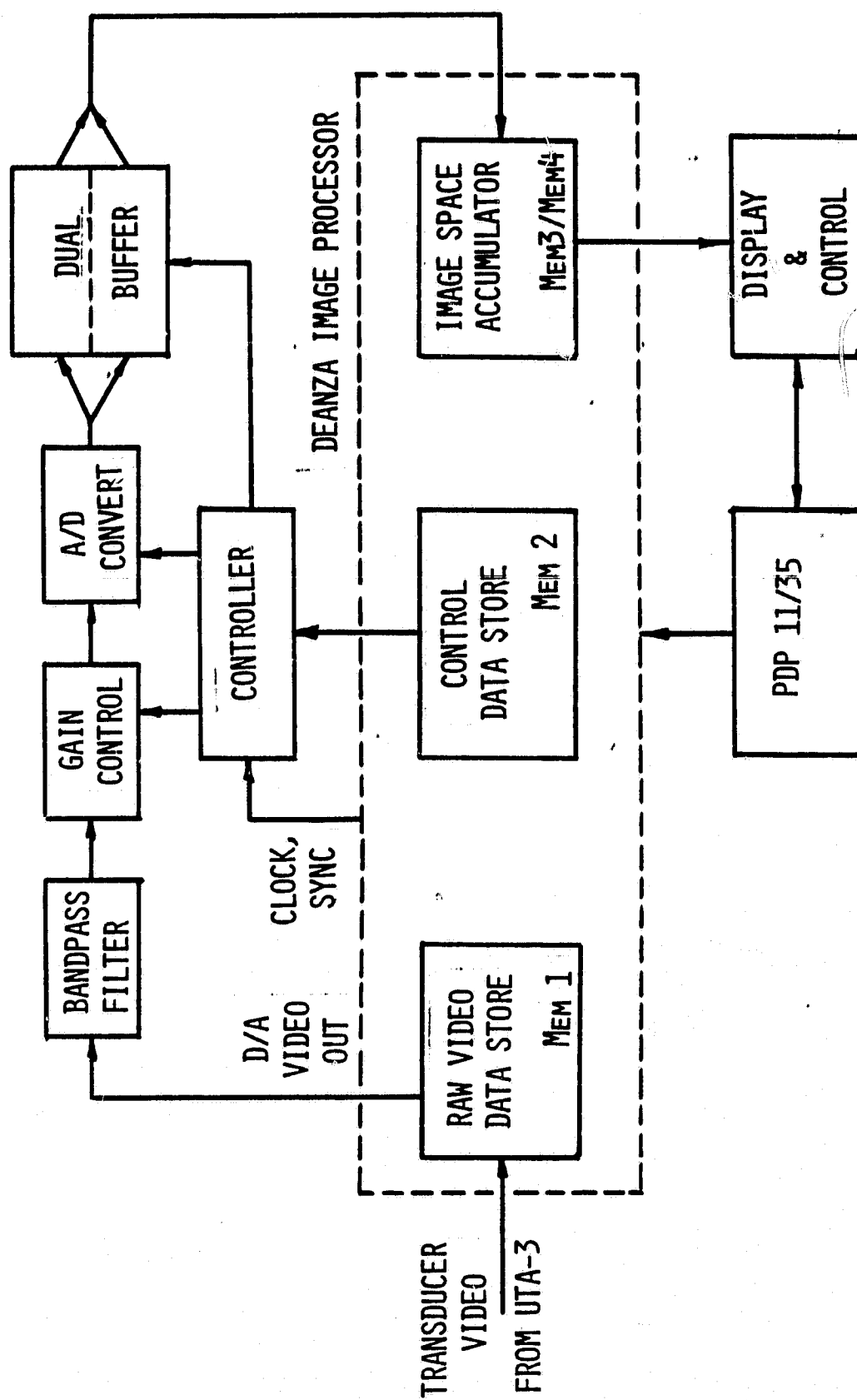
Figure 8

X-Y location, but this is by no means the only option. The image space will contain many locations in the X-Y plane and we need to read out the same line of data (for a particular transducer pair) many times. For example, if there are 32x32 cells in the X-Y plane, each transducer signal needs to be read out 1024 times. The data manipulation capabilities of the deAnza system will allow various techniques for minimizing the number of read cycles to form an image; for example, if only 128 range samples are taken, the returns for 4 transducer-pairs can be recorded on a single line for a corresponding reduction in processing time.

Raw data read out of memory #1 will be in analog form provided by a D/A converter within the deAnza system. After filtering and gain scaling the signal will be re-converted with controlled sample spacing and stored in a buffer memory. Pre-computed control data will be loaded into Memory 2. As each line is scanned a sequence of 8-bit words will be read out of the Control Data Store that will define the gain (5 bits) and sample spacing (3 bits) for each sampled point. Two words will be reserved at the beginning of each line to define the range to the initial sample. The A/D samples will be derived from a 40 MHz clock to allow the sample spacing to be varied in increments of 25 ns.

In the event that the number of cells in the X-Y plane exceeds 512, the Control Data Store will need to be refreshed during the image formation with a new set of pre-computed data. As samples are taken at the variable clock rate, they are entered into a double buffer. While one-half of the buffer is being loaded, the other half is read out at the 13.3 MHz clock rate of the deAnza system and stored in Memories 3 and 4 (combined to allow 16-bit samples per cell). Each active cell in the image space accumulator will receive a contribution from each transducer pair, the final result representing a summation of all inputs. The image space can accommodate 512x512 cells. This could consist of 64x64x64 cells in X-Y-Z space but we will have the capability to form other image space shapes such that the total number of cells does not exceed 262,144 cells. If we scan each line of the Raw Video Data Store 512 times to form a complete image, the time required will be 17 seconds. If fewer input scans are used, proportionally less time will be required.

Many options are available for displaying the data, the most obvious being as a sequence of planar slices. If we want to vary the orientation, this can be done by manipulating the data already in the image memory, but it may be more desirable to re-define the image cell coordinates based on a new orientation, and re-compute the timing required, and process the raw data again. It will also be desirable to vary the resolution of the system to expand the image space volume represented by the fixed number of cells. This may be done by a combination of transducer selection to control angular resolution along with the video filtering to control range resolution.

6. Increased Penetration

The system we are describing records the results of a scan in a digital memory. These memories have the capability of adding a new set of data to the existing set to permit picture averaging. This means that we can repeat a given scan N times and improve the signal to receiver noise ratio by a factor of N. This does not improve the signal to clutter ratio.

This is straightforward in the case of stationary targets. With eye motion effects it <u>appears</u> that by gating with the EKG to delete times when pulsatile motion is significant (for example, use 1/3 of the interpulse period) we can obtain several seconds worth of coherent integration, without significant degradation of resolution.

The use of a retrobulbar block will effectively eliminate eye motion, thus in cases where this is warranted from a medical standpoint it appears reasonable to integrate for periods up to the order of 30 seconds. This would permit the order of 1000 integrations with a gain in signal to noise ratio of 30 dB. This translates directly into additional penetration. If, for example, the attenuation is 20 dB/cm round trip, this would permit an additional 1-1/2 cm penetration.

It should also be noted, as pointed out in the general description of the technique, that the ability to produce better resolution at a given frequency obviously permits the use of a lower frequency for the same resolution. This also permits improved imaging at greater depths.

7. Array Design Considerations

As indicated in the preceding discussions, we want an array in which all the elements are pointed at the high resolution volume V, and have the appropriate beamwidth.

In our initial design we specified a 49 element 7x7 array with element on 3 mm spacing and element size of 1.5 by 1.5 mm. The "theoretical" beamwidth to the first null of an element 1.5 mm wide at a wavelength of .6 mm is approximately .6/1.5=.4 radians or a full beamwidth, null to null, of 0.8 radians. The actual beamwidth is 0.48 radians. This is not surprising because a significant area of transducer material around the electrodes is excited. This is nobody's fault; it was an experimental array built with mutual agreement to find out what problems might arise.

This means that we will operate at a greater range than planned in order to obtain adequte beam overlap.

We are initiating some experiments with a "lens" configuration to improve things in the next version. We want to tilt the beams from the individual element so the centers of these beams pass through a common point at a range of 5 to 10 cm from the face of the array. We believe this can be done by using a plate of standard lens material into which flat bottom holes are milled at different angles for each element. For the element at the center of the array no tilt is required; the tilt increases with distance from the center. This will permit the beams to be pointed in the way we want. We expect that we can also provide different beamwidths by going to holes with curved bottoms to control the beamwidth.

This method will not work over large angles, and for sufficiently large angles we plan to use several separate planar arrays mounted at appropriate angles to approximate a spherical surface.

8. Probable Applicability of CT Techniques

We have seen that the signals received at the individual receiving elements for each transmission consist of the sum of the returns from a volume shell within the medium, and that the signal from the focal spot is estimated by summing these signals so that the returns from the focal spot add coherently and the remainder of the returns add incoherently.

This corresponds to the "back projection" technique in computerized tomography. It differs in several respects; for the ultrasound case, the returns are bi-polar and the estimate from back projection improves more rapidly than for the x-ray case where only positive numbers are involved.

In the cases we have considered in the general discussion, we are assuming three-dimensional resolution; typical CT systems operate with a series of planar slices, which reduces the recording and processing problems. Newer CT systems are going to two-dimensional arrays of detectors to provide three-dimensional imaging.

An ultrasound system has the advantage of bi-polar signals; it has the disadvantage that ray paths are changed by the varying speed in an inhomogeneous medium, thus necessitating the computation of the corrected ray paths in the reconstruction process for the case of an arbitrary medium. An additional complicating factor for ultrasound is the specular reflections which occur for reflecting objects with sizes large compared to the wavelengths.

With the system we are proposing we can form an estimted image from the input data. We can take this estimated image and find the signals which would have been received if this estimated structure had been imaged. If the imaging system was perfect, the received signals would be identical from the actual and estimated structures. This will obviously not be the case; the difference in these two data sets is a measure of the "distortion" in the imaging process.