MEIS

MICROELECTRONIC & INFORMATION SCIENCES CENTER

INSTITUTE OF TECHNOLOGY
UNIVERSITY OF MINNESOTA

227 Lind Hall / 207 Church Street S.E.
Minneapolis, Minnesota 55455
612/376-9122

THE IMPORTANCE AND DESIGN OF THE STUDENT-MODEL DATA STRUCTURE

Microelectronic and Information Sciences Center

Technical Report #31
Revised, August 1985

R. Doshi
Department of Computer Science
University of Minnesota

B. Koski
Department of Electrical Engineering
University of Minnesota

R. Hu
Department of Electrical Engineering
University of Minnesota

H. Wechsler
Department of Electrical Engineering
University of Minnesota

M. Wick
Department of Computer Science
University of Minnesota

MEIS technical reports are published in order to provide early access to research results intended for publication. Manuscripts are reviewed and selected for MEIS technical reports by the MEIS Executive Committee.

THE IMPORTANCE AND DESIGN OF THE STUDENT-MODEL DATA STRUCTURE

MEIS Technical Report #31

R. Doshi
B. Koski
R. Hu
H. Wechsler
M. Wick

The research reported in this paper was supported in part by the Microelectronic and Information Sciences Center.

The Importance and Design of the Student-Model Data Structure

Rajkumar Doshi, Bruce Koski, Dick Hu, Harry Wechsler, Mike Wick

ABSTRACT

The data structures used by an Intelligent Computer Aided Instruction system for the modelling the student are used for various purposes like Explanation and Tutoring. The development and design of the primitives that these data structures are made up of, is very crucial to the success of the ICAI system. We have developed these primitives for the domain of teaching Programming Languages. We will also show the power of reasoning and student modelling these data structures can support.

1. Introduction

Computer Aided Instruction (CAI) was started in the early 1960's. The main purpose was to interact with the student. Early programs were more or less page turners. In the 1970's, Intelligent Computer Aided Instruction (ICAI) systems were developed, where course materials were prepared independently of the teaching procedures so that different advice could be generated differently for each student [Carbonell 1970].

Today, research is focussed to make ICAI sensitive to the student's strength, weakness and preferred style of learning. ICAI should offer a reactive learning environment, where the student's interest and mis-understandings must actively guide the tutoring and explanations of the system. ICAI should be capable of pointing out the real source of difficulty; not just say right or wrong.

After reviewing previous attempts at trying to satisfy the criteria of ICAI, our research shows that it is very important to analyse the domain and determine the primitive operations and various inter-relationships among them. The next step is to design appropriate data structure(s) to be able to capture the strengths and weaknesses of the student. The success of the system depends on the quality of the data structure(s).

Note that the data structure primitives described in this paper cannot support user modelling i.e. personal traits of the students. The work presented in this paper is conceptual. We have not implemented it on a computer. However, we have manually simulated the working of our ideas.

1.1. Guide to the reader

In section #2, we will describe previous ICAI systems and their limitations in the area of Student Modelling (SM). In section #3, we will discuss our approach to the problem of student modelling. In section #4, we will discuss some of the very difficult issues in student modelling; also, we will illustrate our ideas with a very simple example.

2. Review of Previous Research Efforts

This section will serve to put our research into the proper context. The relevant systems that we will discuss are:

1. Spade-0
2. WEST
3. ACE
4. GUIDON
5. MACSYMA Advisor
6. PIP

For each of the above we will discuss the following:

- the domain,
- the input,
- the output,
- the process which takes the input to produce the output.

In this review, we have sometimes changed some of the original author's terminology; this helps bring the ideas into proper perspective and also standardize the subject material to reduce the amount of ambiguity.

2.1. SPADE-0

2.1.1. Domain: SPADE-0 [Miller 1983] is an environment designed to help students acquire general programming skills such as Top down design and isolation of programming bugs. The domain is programming to create Graphics. SPADE-0 is a menu driven system and it guides the student through a hierarchical planning process and discourages the student from premature coding. SPADE-0 also guides the student through a hierarchical debugging process.

2.1.2. Input: It takes as input a transcript from an interactive programming session with a student. In this session, the student works through with the help of the system, the design and implementation of a single programming task. This may include such things as overall plans, debugging techniques, test runs, etc.

2.1.3. Output: It was really designed as an experimental tool for building tutorial programs. As such, the "output" of the system is directed more towards system refinement than towards any conclusive statements about the student.

2.1.4. Process: The basic process of SPADE-0 is an interactive programming session between SPADE-0 and the student. During this session, the system will use different modes (set of rules) to help guide the student through the various aspects of developing a program, eg. plan revision, debugging, etc. The student, following well defined student models, enters his/her plans for a solution to the problem through a hierarchy of solution descriptions (eg. Top-down refinement). Later, after test runs, the student may revisit this solution plan, and edit it in attempt to resolve bugs. The goal of this processs is to tutor the student in programming by examining his/her approach to solving a problem and offering helpful guidance along a solution path.

2.2. WEST

2.2.1. Domain: WEST [Burton, Brown 1983] is a coach for a student playing the boardgame "How the West was Won". The computer and the student are adverseries. The object is to be the first player to land exactly on the last space on the board.

When it is the student's turn to make a move, the computer gives the student three numbers; the student has to use an arithmetic expression to determine the number of spaces to move. This expression is made up by the student, using 4 arithmetic operators. The constraint is that no number or operator be used twice in the expression. One simple strategy is to add up these three numbers and move that many squares. But, the student has also to learn how to take advantage of special moves, like:

- SKIP TOWNS
- SHORTCUTS
- BUMPS

If you land on a town, you can advance to the next town. If you land on a shortcut, you can proceed from the other end of the shortcut. If you land on a space that your opponent is occupying, the opponent is bumped back two towns.

So, sometimes it is smart not to move as far as you can, but to move less so as to be able to visit a town on the way and or to take shortcuts.

2.2.2. Input: The model used in WEST keeps track of the differences between the student's (computer's) behaviour and the expert's behaviour. The input to the model is the student's move and a set of better moves that the student could have made. This set of better moves is created by some specialised modules in the WEST system. Associated with each of these moves is a set of skills which must be used in some manner to play that move.

2.2.3. Output: To make a move might take a combination of one or more basic pre-defined skills. If the system finds that the student has misused the skill (under-use, over-use, incorrect-use), WEST may decide to present an explanation of that skill together with a better move that illustrates the use of this skill. In this way the student can see the usefulness of the skill at a time when he will be most receptive to the idea presented - immediately after he has attempted a problem whose solution requires that skill.

2.2.4. Process: The important aspects of the domain are the skill and the concepts the student is expected to master. These skills and concepts are identified as a collection of Issues. There are three levels of Issues skills. The lowest level contains the basic mathematical skills the student is practicing. The second level contains skills needed to play this game such as when to move forwards or backwards, or when use the special skill (discussed above). The third level contains general game playing skills such as watching your opponent in order to learn from his moves.

Associated with each issue are two procedures, a Recognizer and an Evaluator. The Recognizer procedure watches the student's moves and determines which skills the student has used in making his/her moves. This information is used to construct the Student Model. Each Evaluator knows how to use the Student Model to tell if the student is weak in some particular issue. At any point in the game, the student's weaknesses can be determined by running all of the issue Evaluators.

2.3. ACE

2.3.1. Domain: ACE [Sleeman, Hendley 1979] is a system which Analyses Complex Explanations. The system observes the student's behaviour by analysing the explanations the student gives during his attempt to solve a problem. The domain involves analysing Nuclear Magnetic Resonance spectra, given the molecular formula and the spectrum. Listed below are some of the system's facilities:

[1] A response mode which tells the student whether or not a solution to a sub-problem is correct.

[2] an EXPLAIN mode to tell the student why the last assertion made by the student was wrong.

[3] a Follow Student's Reasoning mode which takes a student's explanation and comments on it's correctness, completeness and consistency.

2.3.2. Input: The input is the student's argument to solve a given problem. The system will parse these arguments presented by the student into a series of facts, assertions and deductions which it then presents to the matcher. The system also gives the matcher a set of correct solutions. The input to the system also includes a set of previous arguments about this problem, that the student has already mentioned before in the current session.

2.3.3. Output: The output is an analysis of the student's explanations. Basically, the system's analysis tells whether the student's argument is correct or incorrect. If the argument is incorrect, the systems breaks down the arguments into individual facts and deductions, and then points out whether these are valid, inconsistent or incomplete.

2.3.4. Process: The student selects a problem from an existing library and begins an interactive dialogue with the system. The system asks the student for an analysis and arguments. The student then enters the necessary statements. The system extracts the meaning which is matched against to a set of prefabricated correct solutions.

The matcher has three basic stages. The first stage is to simply try to match the current explanation with the correct pre-stored ones. If this fails, the system then uses the last assertion of a previous argument list in conjunction with the current argument and again tries the process of the first stage. This process is repeated until either a match is found or the previous argument list is empty. If no match was found, then the student's argument must be inconsistent and so the argument is subject to extensive analysis. ACE does not really store or update a Student Model. The system always starts from afresh.

2.4. GUIDON

2.4.1. Domain: GUIDON [Clancey 1983] attempts to teach a student the rules of MYCIN by taking the student through a case study of diagnosing a desease. The student attempts to identify the desease while GUIDON points out the wrong lines of reasoning and suggests approaches that the student has not considered.

2.4.2. Input: The input to the process are: a cummulitive Student Model, the rules that MYCIN thinks should have been used and the values that the student has concluded. Also, it considers the inherent complexity of the rules, the background of the student and whether GUIDON believes that the student knows how to achieve the sub-goals that appear in the rules.

2.4.3. Output: GUIDON does not really conclude anything about the student (level of performance, amount of improvement, tired, frustrated etc.). It continues directing the student and giving appropriate advice where needed until it believes that the student has some level of confidence and that the student is aware of and understands the releveant MYCIN rules.

2.4.4. Process: During a session with GUIDON, the student asks questions and draws hypothesis about the desease causing the problem. While the student is doing this, GUIDON is also running the MYCIN program on the same set of facts. Now, as the student moves through the problem solving process, GUIDON can be asked questions about how to proceed next. It can also inform the student when his/her line of reasoning is going astray. GUIDON uses T-rules (teaching rules) to look at the Student Model and decide on the best way to advise the student. GUIDON looks for oppurtunities to deepen and broaden the student's knowledge. It updates the Student Model, whenever a domain rule fires in MYCIN.

2.5. MACSYMA Advisor

2.5.1. Domain: The MACSYMA advisor is a consultant on the usage of MACSYMA. If a student misunderstands a MACSYMA command, he can turn to the advisor for help.

2.5.2. Input: The student gives to the advisor the goal and the MACSYMA commands he used in trying to achieve that goal. The advisor has it's own data base of mathematical facts, the student's knowledge about the commands, which commands he knows and which it believes he does not know and common errors made by most people.

2.5.3. Output: The output explains the MACSYMA's solutions in greater detail. It intends to clarify the user's misconception on a particular MACSYMA command which lead to a incorrect solution.

2.5.4. Process: The advisor analyses the sequence of MACSYMA commands used by the student. It identifies the method the student used to solve his problem. By comparing against the correct commands for the particular method, the advisor can discover the student's misconception about the commands. The advisor can verify it's assumptions or partial solutions by asking the student for more information.

2.6. PIP (Pascal Instructional Program)

2.6.1. Domain: PIP [Dehlawi 83] was designed to teach elementary Pascal Programming. Dehlawi claims that it is equivalent to a 2 unit introductory course in computer programming at the college level. The student receives a task. They solve the problem (using several tools provided by PIP, like hints about problem and info about Pascal syntax). PIP suggests corrections to errors giving references to pertinent sections in the PIP manual.

2.6.2. Input: The student writes up a tentative solution (code). The student then has to ask the operating system to 'run' his/her program. The student (not the system) is supposed to manually note the error messages. PIP does not do any solution verification. The student then edits the incorrect program till he/she finds the correct solution.

2.6.3. Process: Besides having a crude task selection algorithm, PIP does not do anything intelligent at all. When the student's program runs successfully, the student asks for a post task interview. PIP gives the student a model solution. The student (not PIP) has to assess how closely his/her solution resembles the model solution. After this self assessment (in which PIP plays no role) the student answers questions about the skills used in the task. The student (not PIP) has to manually indicate confidence in each skill (Dehlawi has not said what is the type and nature of the units used to express this confidence) or, the student must indicate the need for more practice.

2.6.4. Output: The student is either right or wrong; it's a binary decision. And, even this decision is left upto the student. PIP plays no role to fulfill any of the goals of ICAI systems.

2.7. Limitations of the above Systems

We will now discuss the limitations for each system. However, we do want to explicitly mention that we are critiquing their student modelling capabilities only; they were very good in other areas like, explanation, problem solving, etc.

SPADE-0

The student's competence in programming is never evaluated. Rather than grade the student, the system guides the student through the debugging process.

WEST

When a student makes a wrong move, WEST assigns blame more or less equally to all the skills; that is, blame will be given even to Issues that are in fact understood.

ACE

The major limitation is that the system does not retain any info about the student between sessions; that is, there is absolutely nothing like a history, or statistics about previous effort. The student's arguments are taken independently and no cross reference is made to attempt to explain or model the student's knowledge. It makes no effort to understand the student's abilities or shortcomings. It does not even care whether the student got it right on the first try or the ninety ninth try; just that sooner or later he/she gets it correct.

GUIDON

GUIDON is the most sophisticated of the sytems we have even considered for review. But, we can identify the following problems. The student model does not recognise student behaviour that can be explained by simple variations of the expert's rules. Also, GUIDON does not consider the difficulty of the problem being tackled.

MACSYMA Advisor

The system assumes that once the student has used a command correctly, he knows exactly all the instances where that command can and cannot be used. The system does not consider interactions among commands, forgetfulness, partial understanding and partial misconceptions.

PIP

The system does not really have a good and complete understanding of the programming domain. PIP itself has not been given the tools to understand the role of plans and goals in programming. PIP does not have a solution verifier. PIP assigns this task of solution verification to the student. And, very importantly, the student model has no capability (primitives) to model the student.

3. Our Approach: What we consider important

3.1. Basic Ideas

The Student Model is probably the single most important data structure used by the various processes viz. the Psychologist, the Tutor and the Explanation-strategy modules.

For ICAI in programming languages, the SM should have primitives to allow us to represent facts. The distinction between a fact and other interpretations derivable from these facts is crucial. Facts are things that the student has or has not done; all of the following are facts:

- A while loop has been used to solve the problem
- The increment used was +2

In contrast, the following is an interpretation:

- The student is forgetfull and slow.

The key idea is to analyse the mistakes made by students and also the associated reasons, in the domain of problem solving using programming languages, so that we can represent any and all facts (mistakes/errors), which then helps us to derive (using some procedures) any and all interpretations about the student's performance.

Basic Idea #1

To be able to represent any and all facts about what the strategies the student has used or not used, we need a data structure built from:

- Primitives and
- Relationships between these Primitives.

Basic Idea #2

To meet the requirements of ICAI, we need two distinct data structures:

- ☐ Problem Specific Student Model and
- ☐ General Student Model.

Problem Specific Student Model (PSSM)

This model is used solely to represent facts about what the student has done and has not done. There is a PSSM for every attempt the student makes to solve a problem. We fill the slots of the PSSM by facts, which are obtained by analysing his/her solution.

General Student Model (GSM)

This model represents the student's overall depth and breadth of knowledge about the usage a specific programming language (in our particular case it is PASCAL). The model should also store other relevant statistics related to the usage of the Pascal constructs. The GSM stores facts and interpretations based on these facts (and other performance statistics).

The PSSM (facts) represents the input to a process (rules) that will compute and store relevant information in the GSM . This process will also modify confidence of the appropriate primitives in the GSM, depending on the student's performance and history.

Figure #1 illustrates our very high-level view an ICAI-PL (ICAI-Programming Languages) system can work. We have not developed all the components. By using the figure, we will show the role of the PSSM and the GSM data structures, with respect to the ICAI process as a whole.

Lets assume the following: The ICAI system gives a student a problem from it's library. The system has a specification of the problem in some specialised task description language. The student then turns in his or her solution (code in PASCAL) back to the ICAI system. The system will then attempt to verify the student's solution by trying to see if his or her code satisfies the problem's specification. If not, then the system will try to identify non-syntactic bugs. We have neither built nor worked on such a system. However, it is not an unrealistic assumption. Currently, it is not a part of our system. To give an example, PROUST [Lewis, Johnson 84] is an existing system which does error diagnosis and can cope with significant variability exhibited by novices' programs.

Now, lets further assume that, after finding the bugs, the system would fill the appropriate slots in the PSSM (Note: PROUST does not do this; it was not meant to; we are trying to see what we can do if we had a system that can do this). Again, this modified PSSM will then represent what the student has done, how and what he has not done, in his/her last attempt at solving the problem (for an example, see section #3.2).

We use this PSSM (which represents his/her latest attempt at solving the problem) to see what new information we can generate about the student's knowledge in programming. We would also use the GSM (his or her previous history or record) in making a conclusion. So, the PSSM and the GSM represent the input as far as our system is concerned. (The PSSM data structure will be defined in section #3.2; the GSM data structure will be defined in section #3.3.)

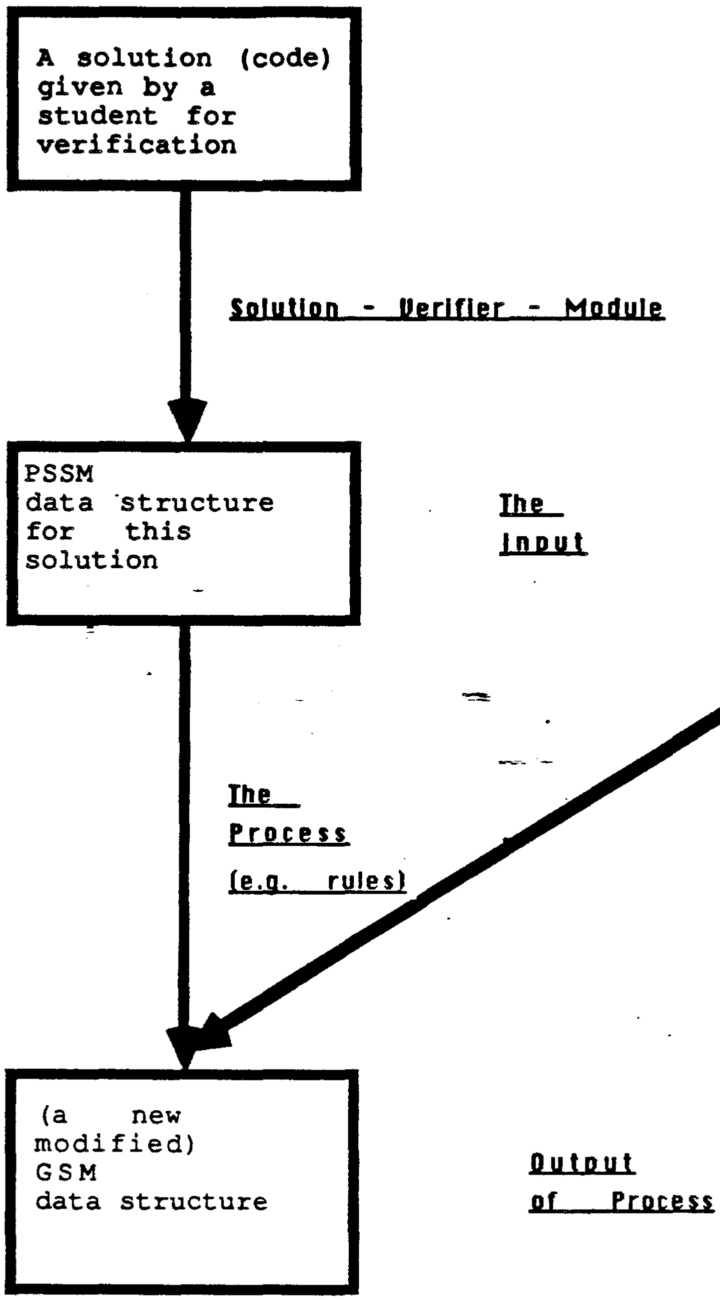


Figure #1 : Our View of the Student Modelling Process

There should then be some process (eg. a rule base) to find the possible cause(s) of the student's mistakes and/or misunderstandings. It may (or may not) also change it's belief in the student's overall competence. It may just modify it's belief of the student's competence in using some particular construct of PASCAL (what exactly this process will do will depend on many many factors viz. the level of the student, the complexity of the task being solved, how many similar tasks have been previously solved by the student, how much help has been given already, how long has he or she has been on the terminal, etc. etc.). We have done only little work on this process part of the problem. The work that we did do on this process part was done to convince ourselves, that, given PSMs and GSMs, such a process could indeed be feasible. The process part consisted of studying, understanding and evaluating the solutions written by novice programmers and then, we tried to see **how** we did this evaluation. Section #4 will discuss that in a little more detail.

Finally, the output of the process is a modified GSM.

3.2. PSSM : Development & Examples

We analysed the midquarter and final exams of many students to develop the components of this data structure. Our need was to be able to capture any and all mistakes that a student can make while writing algorithms. Before presenting the complete data structure, lets us trace the steps we took in building this model.

In the Figure #2 below, a student has proposed a solution to count from 0 to 10. The student does not have an initialization statement. We would like to capture this fact. Note that we have avoided saying that the student has **forgotten** to put the initialization statement.

```
Program count (output);
var
  i : integer;
begin
  while (i <= 10) do
  begin
    writeln(i);
    i := i+1;
  end;
end.
```

Figure #2: The Missing Initialization Problem

So, if our PSSM looked like as shown in Figure #3(a), then, after analysing the student's solution, Figure #3(b) shows the PSSM with it's slots filled.

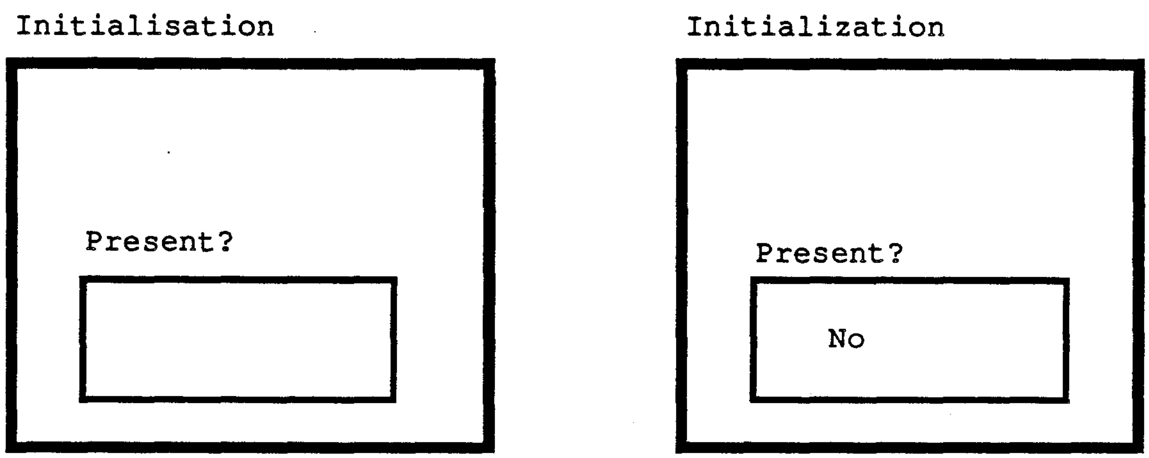


Figure #3: The initial and final PSSMs for Figure #2.

Lets take another small example as shown in Figure #4. The problem is to read in numbers, till we read in the number 100. The student has put an increment, which was not needed. We want to capture this; we will later on show how we interpret the student's mistake.

```
Program read_till_signal (input);
var
  i : integer;
begin
  readln(i);
  while (i <> 100) do
  begin
    readln(i);
    i := i + 1;
  end;
end.
```

Figure #4: The Increment Not Needed problem.

We have studied more complex examples. Basically, we took all the constructs and tried to violate them in many many ways. We then tried to see if our PSSM data structure allowed us to capture the mistakes. Whenever we could not represent a mistake, we developed appropriate primitives to then capture this mistake. As a concrete example, Figure #5 shows a complete PSSM for a while loop control structure.

Again, it is impossible to model a student or to explain or to teach a student, without knowing all the facts & their complete interdependencies. Also, there is always some inherent noise in the modelling process, because the ICAI systems dont have vision, audio, etc. And even if it did, we know from experience, that there will always be some uncertainity about the student's abilities. ICAI systems should not form opinions or judgements or interpretations about the student's performance and then purge the facts. Rather, the facts should be stored explicitly and interpretations should be derived when necessary.

3.3. The General Student Model: It's Architecture & Significance

As already mentioned before, the General Student Model (GSM) has many uses. It is used as a data structure to represent what the student has done correctly and what he or she has done incorrectly. The design of this structure should allow an ICAI system to reason about the true source of misconceptions and also, what some particular weaknesses can say about the familiarity with other similar problem solving skills. It is used to maintain various statistics. Also, it can maintain confidence measures with every node in the various levels of the GSM (explained below). Along with the confidence units, endorsements justifying some competence rating is also stored.

3.3.1. Architecture of the GSM

As mentioned in section #3, the GSM represents the input to the process of evaluating & understanding the student's performance and the GSM also represents the output of such a process. In developing an ICAI system, we think that before working the process part, the input and output should be very well thought out. Also, this specification will shape and give further insights in the process part.

Our GSM has 4 levels:

- Level #1: General Concepts level
- Level #2: Programming Concepts level
- Level #3: PASCAL Constructs level

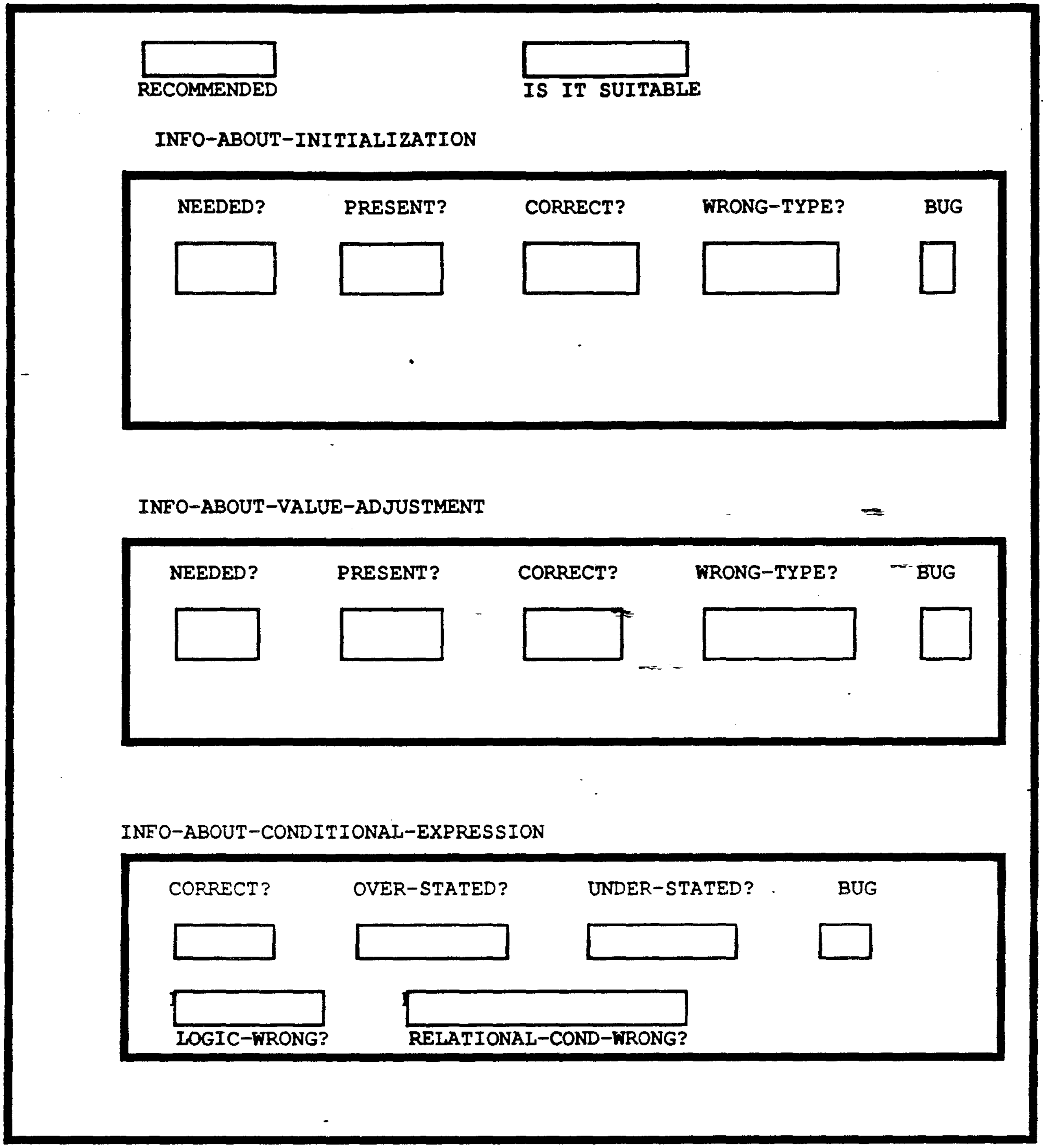


Figure #5 : The Primitives for a While-loop in the PSSM

- Level #4: Tasks level

In each level, we represent skills appropriate to the level in the form of entities or nodes. Skills within a level and between different levels are related via relationships. There are four kinds of relationships (not counting their inverses):

- Part-of
- Analogous
- Affects
- Immune

The GSM components can be related to each other in more than one way and so, there can be more than one relationship links between the nodes. The above mentioned four names of relationships represent three types of links; the **Part-of** link is mainly used when reasoning about a PASCAL construct and it's components. The **Analogous** link is mainly used when reasoning with Tasks level problems. Now, suppose a student repeatedly makes errors in the syntax of the while construct; this cannot tell us anything about his or her ability of the for construct. But, now suppose, the student makes repeated errors in specifying the conditional expression of the while construct; This could imply that (maybe), he or she needs to be tutored again on the concepts of branching, precedence rules of logical operators, etc. And, this also means that the ICAI system can also modify the confidence measure in the repeat construct. The **Affects** and the **Immune** links are used to give an ICAI system the power to decide when to propagate changes in confidence ratings and when not to do the same. These links (i.e. this kind of knowledge) is implicitly present in the rules (see section #4) used to process the PSSM to modify the GSM.

Level #1 - General Concepts level: This level represents problem solving skills at the level of a non-programmer. Some of these skills are very very basic, like for example, the concept about operations on integer numbers. Some other concepts like branching and iteration are less trivial, but these concepts are still very much a part of planning in daily activities.

Level #2 - Programming Constructs level: In this level, the skills relate to those needed in general high level programming languages like PASCAL or FORTRAN. So this level expresses skills that need to be learned by novice programmers, but these skills are language independent. This level itself has three sub levels, as shown in Figure #6 below.

Level #3 - PASCAL Constructs level: Here, the skills (the nodes) have the names of the actual PASCAL language constructs.

Level #4 - Tasks level: Here, we represent tasks that range from very trivial tasks (like declaring an integer) to complex tasks. Tasks can be related to each other or tasks can be related any underlying concepts (at the lower levels) by the following relationships (and their inverses):

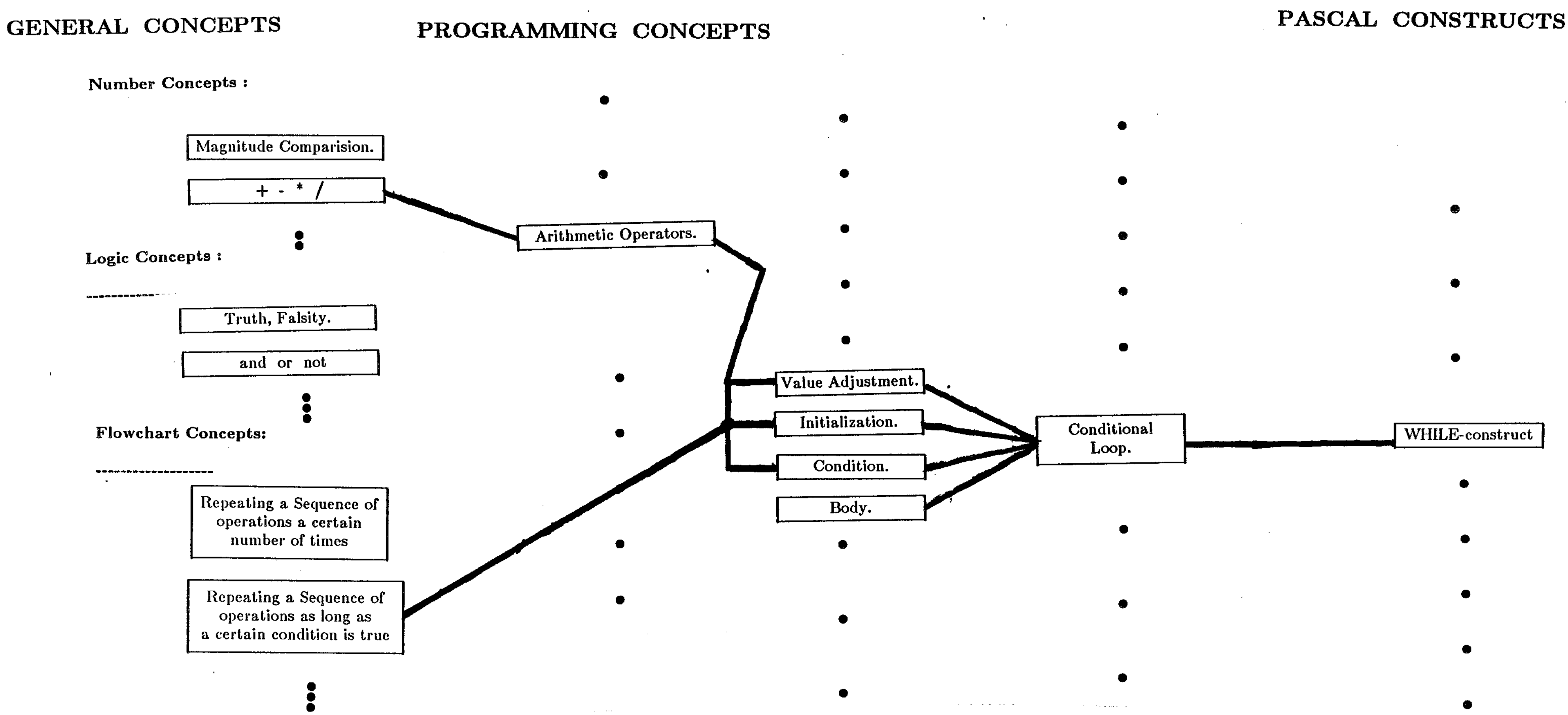


Figure #6 : A partial GSM data structure.

- pre-requisite-skill
- harder-skill
- analogous-skill
- based-on-the-concept-of

3.3.2. Usage & Power of the General Student Model

An ICAI system needs better primitives so that it can capture and then understand the student's real source of difficulty. Sometimes, the real difficulty could be due to biases developed by the student (especially, if the student is from a non computer science background) or due to mis-understandings about or incomplete knowledge of some basic problem solving skills (this refers to the General Concepts level and the Programming Languages Constructs level of the GSM). Also, many constructs depend on the complete understanding of it's various sub-constructs, like for example, the while loop is composed of:

(i) initialization of control variables,
(ii) testing of control variables & branching, if necessary,
(iii) performing some computation,
(iv) modifying the control variable,
(v) doing step #(ii) again.

Novice students repeatedly make errors in one or more of the above subconstructs. Each of these sub constructs would themselves be based upon other (one or more) general concepts. So, the GSM, with the current level of detail given so far, would facilitate (if also given appropriate procedures or rules to infer and manipulate the GSM) detecting underlying misconceptions of a student.

But, there should be more to the GSM than what we have described above; it will allow an ICAI system to just record, for example, that the student does or does not know the PASCAL while loop skill. Or, alternatively, it will allow to record some kind of numeric confidence rating, like for example "the student knows the while loop skill 0.53". Or, as another alternative, it give the ICAI system the ability to record qualitative measures, like for example "the student's knowledge about the while loop is not very good". Ofcourse, generating or inferring the abovementioned confidence measures (whether it is binary, numeric or qualitative) is another major problem in itself. We have done some work (basically, just enough to demonstrate & simulate our ideas of the PSSM & GSM data structures) on this problem; we will focus on this in section #4.

We had mentioned earlier that the GSM is a multi-purpose data structure. Suppose that a student makes a certain mistake where it is appropriate to remind him or her that, to enter a while loop, the control condition expression must be true. The GSM explained as above cannot facilitate this! So, we have to analyse the domain in more detail and augment this GSM.

While we were studying and evaluating student's solutions, we found ourselves saying things like:

- "the student forgot to initialize the while loop variable; he should have done that before entering the loop."
- "the student does not understand that a while loop would have been more appropriate here, because, the while loop allows variable number of iterations".

So, we realised that when an instructor teaches PASCAL, the students (should) learn the PASCAL constructs, their syntax, their components and also, equally important, the preconditions, the postconditions, powers and limitations for using each construct. Each problem solving skill can be considered to be a **resource** and so, for each we construct a Resource-Knowledge-Table as shown in Figure #7. We have the augmented the GSM with these new primitives.

4. Examples to the Power of Modelling

4.1. Manipulating the PSSM & GSM data structures

There has to be a process (eg. a rule base) that has (more) knowledge about the domain of programming and knowledge to use, maintain and manipulate the knowledge represented in the data structures. This process must typically deal with the following (and more):

- try to identify the real source of difficulty and try to diagnose the student's misconceptions.
- modify it's belief in the student's overall competence.
- modify it's belief in the student's competence in the usage of the various Pascal constructs (Level #4 in the GSM).
- (appropriately) propagate changes in belief to lower underlying levels (Level #3, Level #2, Level #1 in the GSM).
- maintain appropriate statistics (eg. the amount of help given for this problem, this session, last 10 days, etc; The frequency of help per problem solved; what he/she has used appropriately).
- try to judge or appraise the student.

As mentioned earlier, we dont have an exhaustive set of rules to do all of the

WHILE construct:

PRECONDITIONS:
Precondition #1: The loop variable(s) must be initialised.
Precondition #2: To enter the loop, the conditional must evaluate to true.
Precondition #3: Atleast one loop variable is needed.

POSTCONDITIONS:
Postcondition #1: The conditional will evaluate to false.

POWER:
Power #1: Allows iteration & looping.
Power #2: Allows variable number of repeated execution.
Power #3: Allows more than one termination criteria & permutations of conditional expressions using relational operators.
Power #4: The flow of control is taken care of automatically.

LIMITATIONS:
Limitation #1: Needs atleast one explicit loop variable.
Limitation #2: The loop variable must have been explicitly initialised.
Limitation #3: There must be some instructions to explicitly change the value of the loop variable (within the body).

Figure #7: Resource-Knowledge-Table for the While-loop construct.

abovementioned ideas. However, we did a fair deal of analysis of the students' solutions and we also analysed our own reasoning process to identify the issues and open problems in making a system that can reason like humans do. The development of such rules is a very complicated, intriguing and even amusing ordeal. The following are just some of the issues, that will illustrate our point:

- How do humans derive judgements about the student, such as, *ingenious, knowledgeable, confused, does not know, not intelligent, etc. Note* also that there are other kinds of judgements like, *sharp, slow* which depend on observations made over (long and short) periods of time. So, we have some more questions:
  - □ How do we formally define the above mentioned terms?
  - □ Judging someone's behaviour depends on what level of abstraction we look at the student's records. How many such levels do we humans use?
  - □ How far back in time do we search someone's records? The last attempt? The final exam? Last 5 weeks? Last 4 years? Also, why, how and when do we change these criteria dynamically?
- The above issues are compounded by the following:
  - □ We probably cannot itemise any and all misconceptions & mistakes the students can make, their permutations and their interactions.
  - □ We must also always consider the difficulty of the task the student is trying to solve.
  - □ We must also consider what the student has been taught and what concepts have not been taught.

It is crude to write *IF* .... *Then* .... rules for every instance of significant errors or observations. Our study is incomplete as of yet. There is definately some kind of structure and needs to be investigated rigourously and thoroughly to make it nearly decomposable. Nevertheless, we have developed many many rules; we will illustrate some of our ideas by using a simple example.

4.2. A Simple Example

Consider the problem of printing integers from 1 to 10. Figure #8 is an example solution by a novice programmer.

The code below has 2 bugs:

- The relational operator is wrong
- the 'counter' is never incremented.

The PSSM representing these facts is shown in Figure #9.

```
Program Print_1_to_10 (output);
var
  counter : integer;
begin
  counter := 1;
  while counter > 10 do
    writeln (counter);
  end;
end.
```

Figure #8: Buggy program to print from 1 to 10.

Now, we can use some rules to try to find out what the student understands and what he/she does not understand. Given below are some sample rules in English. Please refer to Figure #7 also. (In the following rules, the antecedent consists of queries to the PSSM-data-base; the consequent consists of queries to the GSM-data-base and also formulation of some initial hypotheses about the student's problems):

Rule #1: IF the relational operator is wrong
  but the initialization and scope of the
    conditional are correct
  THEN Check the history (the GSM).
    IF he/she did not make this mistake before,
      THEN it could be an oversight
      ELSE he/she probably misunderstands postcondition #1,

Rule #2: IF a statement to increment the loop variable is
    missing
  THEN Check the history (the GSM) for previous mistakes.
    IF he/she did not make this mistake before,
      THEN it could be an oversight
      ELSE he/she misunderstands limitation #3.

Lets assume hypothetically that the student's previous records allowed us to make the following judgements about the student's knowledge of the while loop construct:

- he/she misunderstands postcondition #1.
- he/she misunderstands limitation #3.

This information could further be used to build or modify the system's belief in the student's overall competence, by using more rules (in these rules, the antecedent consists of queries to a GSM-data-base; the consequent consists of changes to this GSM-data-base, modifications to the current set of beliefs

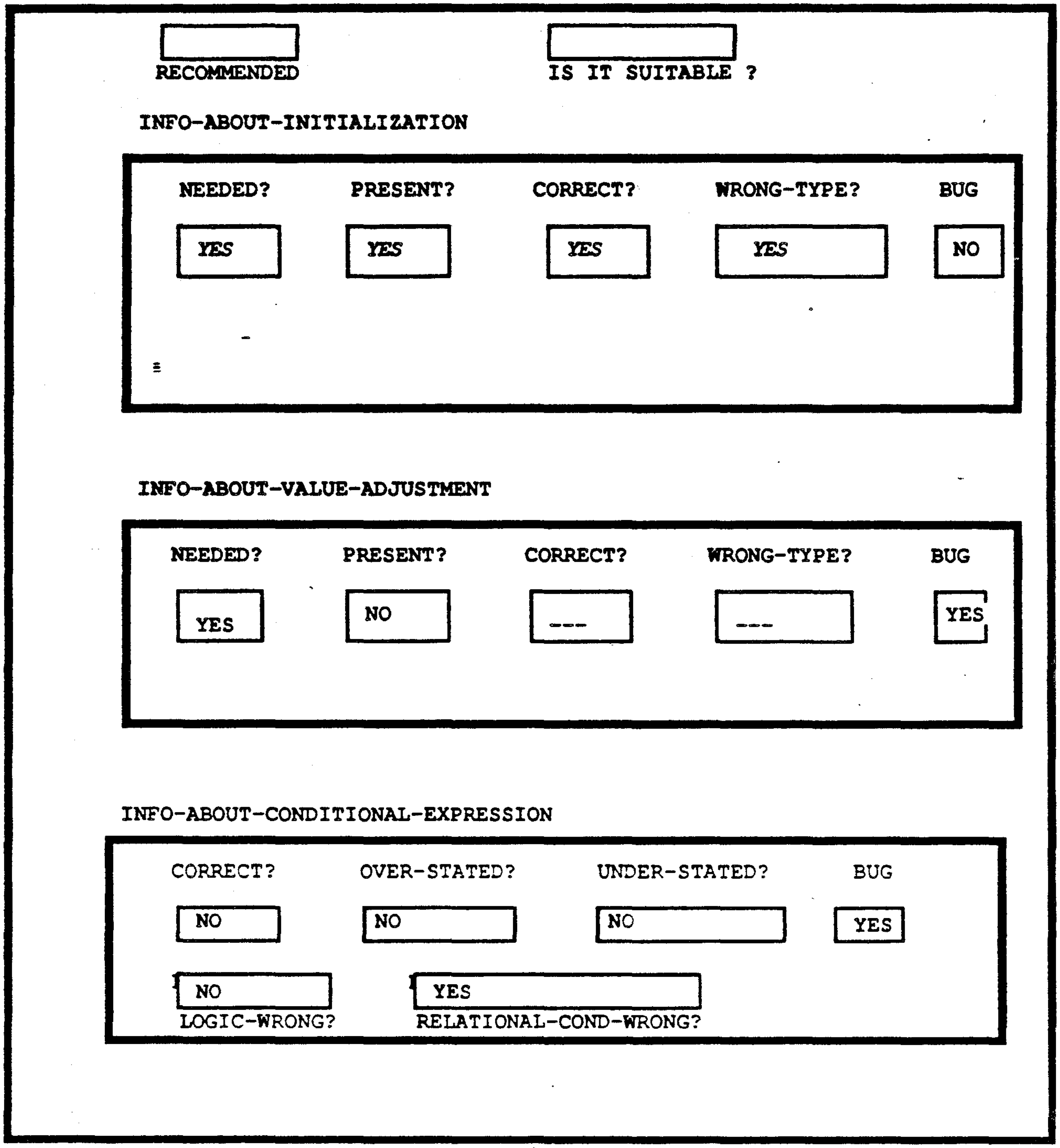


Figure #9 : The PSSM for Figure #8.

about his/her abilities and formulation of hypotheses about the student's abilities):

Rule #1: IF the student misunderstands postcondition #1
    and the task is basic
    and he/she has been taught similar examples
  THEN Record this mistake
    IF he/she has made this mistake frequently
      THEN decrement competence of the initialisation-node
        and the while-node
        and the basic concepts of truth/falsity &
          magnitude comparision.

Rule #2: IF the student misunderstands limitation #3
    and the task is basic
    and he/she has been taught similar examples
  THEN Record this mistake.
    IF he/she has made the mistake frequently,
      THEN decrement competence in the Value-Adjustment-node.

Again, note that, for various reasons (some of which have been mentioned in section #4.1), we feel that to encode the process in the form of such rules is very crude. The interactions among the big and small errors is very complicated; in fact if we use such a simple rule-based system, we would need many many 'demons' to reduce the complexity and make decisions at various levels of detail and abstraction. In fact, many times a 'small' error is more significant and telling than other so called 'big' errors. We have shown the above rules just to show what things have to be considered significant and what they mean when taken together; we do not want to advocate that the a rule representation should be used.

5. What more needs to be done

A lot of research about ICAI tutors and solution verification has been done, eg [Anderson 83], and the YALE Cognition and Programming Projects.

As far as student modelling ICAI-PL is concerned, we would like to continue with the following:

- A computer implementation of our ideas
- A rigorous and intensive study to understand and decompose the reasoning processes used by humans when they model novice programmers. We think that this might lead to some new ways for representing this modelling and data structure manipulation processes.

- We also intend to study how we grade the student in terms of measures; Note that the neasures can be either numeric or qualitative (or both). These issues were discussed in section #4.1. The study should throw more light on the range of units used any particular measure-scheme. Also, it should throw more light on how, when and by how much (the sensitivity) should these measures be modified.

Bibliography

(1) Anderson, T. R., **The Architecture of Cognition,** Harvard Press, 1983.

(2) Burton, R.R. and Brown, J.S., "An Investigation of Computer Coaching for Informal Learning Activities", D. Sleeman and J.S. Brown (editors), **Intelligent Tutoring Systems** , Academic Press, New York, 1982.

(3) Carbonell, J.R., "AI in CAI: An artificial intelligence approach to Computer Aided Instruction", **IEEE Transactions on Man-Machine Systems** , 11(4), 1970.

(4) Clancey, W., "Tutoring rules for guiding a case method dialogue", **Intelligent Tutoring Systems** , Academic Press, New York, 1982.

(5) Dehlawi, F., "Computer Assisted Instruction in Pascal Programming", Ph.D. Thesis, Stanford University, 1983.

(6) Johnson, Lewis and Soloway, E., "PROUST: Knowledge based Program Understanding", **YALE University, Cogntion and Programming Project, Report #7** ,1984.

(7) Genesereth, M.R., "An Automated Consultant for MACSYMA", **Fifth IJCAI,** 1977.

(8) Miller, M.L., "A Structured Planning and Debugging Environment for Elementary Programming", **Intelligent Tutoring Systems** , Academic Press, New York, 1982.

(9) Sleeman, D.H. and Hendley, R.J., "ACE: A system which analyses Complex Explanations", **Intelligent Tutoring Systems** , Academic Press, New York, 1982.

MEIS

MICROELECTRONIC & INFORMATION SCIENCES CENTER

INSTITUTE OF TECHNOLOGY
UNIVERSITY OF MINNESOTA

227 Lind Hall / 207 Church Street S.E.
Minneapolis, Minnesota 55455
612/376-9122

THE DEVELOPMENT OF SENSITIVITY TO KINETIC, BINOCULAR AND PICTORIAL DEPTH INFORMATION IN HUMAN INFANTS

Microelectronic and Information Sciences Center

Technical Report #32

Albert Yonas
Institute of Child Development
University of Minnesota

Carl E. Granrud
Institute of Child Development
University of Minnesota

MEIS technical reports are published in order to provide early access to research results intended for publication. Manuscripts are reviewed and selected for MEIS technical reports by the MEIS Executive Committee.

THE DEVELOPMENT OF
SENSITIVITY TO KINETIC, BINOCULAR
AND PICTORIAL DEPTH INFORMATION
IN HUMAN INFANTS

MEIS Technical Report #32

A. Yonas
C. Granrud

The research reported in this paper was supported in part by the Microelectronic and Information Sciences Center.

THE DEVELOPMENT OF SENSITIVITY TO KINETIC, BINOCULAR AND PICTORIAL DEPTH INFORMATION IN HUMAN INFANTS

Albert Yonas and Carl E. Granrud
Institute of Child Development
University of Minnesota

A version of this paper was presented at NATO Advanced Study Institute: Brain Mechanisms and Spatial Vision, Lyon, France, June 16-25, 1983. Published in D. Ingle, M. Jeannerod, and D. Lee (Eds.), Dordrecht, The Netherlands: Martinus Nijhoff Press, 1985.

THE DEVELOPMENT OF SENSITIVITY TO KINETIC, BINOCULAR AND PICTORIAL DEPTH INFORMATION IN HUMAN INFANTS

Albert Yonas and Carl E. Granrud

Institute of Child Development
University of Minnesota

ABSTRACT

A program of research is described that explores the development of sensitivity to three classes of spatial information in human infants. The research suggests that sensitivity to kinetic, binocular and pictorial depth information develops in a fixed sequence. Some sensitivity to kinetic information may be present at birth or soon thereafter; sensitivity to binocular information appears between three and five months; and sensitivity to static monocular information appears between five and 7 months of age. These findings may direct research on the development of neural mechanisms that underlie the emergence of responsiveness to spatial information in human infants.

1. INTRODUCTION

Humans are a rather altricial species. Our spatial behavior develops slowly. We do not reach for objects and grasp them until four or five months after birth, and we do not crawl until seven or eight months. In fact, human infants do so little initially that it was thought for many years that the young infant was blind. The failure to discover the infant's perceptual abilities was due perhaps to pervasive empiricism and to the near-absence of empirical research studying infants. In recent years, however, there has been an explosion of research on infants' perception which has resulted in a tremendous increase in our knowledge of the infant's perceptual world. Our research has focused on the development of infants' perception of the three-dimensional spatial layout of objects and surfaces in the environment and of events occuring in three-dimensional space. The immediate goal of this research is to discover what information infants use in spatial perception and how infants' spatial perception changes throughout the first year of life. As a long range goal, we hope that understanding infants' perceptual development will aid in discovering the neural mechanisms that underlie spatial perception and its development.

A persistent problem encountered in studying spatial perception in human infants (or any non-verbal organism) is distinguishing responses to the meaning specified by the information in the proximal stimulus (i.e., the pattern of light reaching the eyes) from responses to the proximal stimulus itself. Simply demonstrating that infants detect a particular cue in the proximal stimulus, such as binocular disparity or motion parallax, is not equivalent to demonstrating that infants perceive the spatial layout of objects in the environment that is specified by that cue. This distinction between responses to proximal stimulus cues and the meaning specified by those cues must be kept in mind when one attempts to infer that the infant is detecting information about the three-dimensional layout of objects and surfaces in the distal world. We have attempted to disentangle responses to the distal spatial layout from responses to the proximal stimulus primarily by using spatially appropriate behaviors as indices of infants' depth perception. In many of the studies to be described in this chapter, reaching is used as a measure of spatial sensitivity.

By five months of age, human infants have a strong spontaneous tendency to reach for an object that is within reach, to grasp that object if possible, and to bring that object to the mouth for further exploration and, perhaps, ingestion. When an object is presented beyond the reach of the infant, the duration and number of reaches decrease. By 5 months of age (20-22 weeks), an infant, if given a choice between two objects at different distances, will consistently reach for the nearer object. While it is reasonable to infer that the perception of distance between infant and object or between two objects is controlling reaching, it is possible that the behavior is actually influenced by some non-spatial property of the display. One can infer more securely that spatial perception is taking place by designing control conditions which rule out alternate accounts of the effective variables. In many of the studies of sensitivity to pictorial depth information described in this chapter, a control condition is used in which binocular information is placed in conflict with pictorial cues. If preferential reaching is reduced when conflicting binocular information is present, the inference that reaching is controlled by the perception of relative distance is strengthened.

In addition to studying the direction or duration of reaching, one can analyze more precisely the trajectory and end points of reaches. If one finds that infants reach repeatedly to the precise location of, for example, a stereoscopically-specified virtual object, this would constitute convincing evidence of sensitivity to location in space.

Some behaviors, such as eye blinks, have none of the variability and spatial character of precise reaching, yet it is possible that blinks may also indicate sensitivity to spatial information. Again, the use of control conditions is necessary to clarify the precise stimulus event that is required for blinking to occur. If consistent blinking is evoked only by the particular spatio-temporal transformation that specifies that an object is going to collide with an infant, then it may be reasonable to infer that spatial perception is occurring. On the other hand, it may be the case that a reflex-like mechanism unrelated to other aspects of spatial perception may be responsible. If many different spatially appropriate responses are evoked by the stimulus information (i.e., if we have response convergence) and if many different types of stimulus information evoke the same response (i.e., if we have stimulus convergence) we feel most confident that spatial perception is responsible. While our ability to make such inferences with very young infants is limited, new methods may change this picture in the years ahead.

The research described in this chapter has explored the development of sensitivity to kinetic, binocular, and pictorial information for spatial layout. Within each class of information, we have isolated single depth cues and tested for emergence of responsiveness to each cue. This work has led us to two hypotheses about the development of spatial sensitivity. The first hypothesis is that the mechanisms that detect kinetic, binocular, and pictorial information function as three distinct systems. Second, sensitivity to these three types of information develops in a fixed sequence. Sensitivity to kinetic information may be present at birth or soon thereafter; sensitivity to binocular information appears between three and five months; and sensitivity to static pictorial information appears sometime after five months.

2. KINETIC INFORMATION

Theories of perception have often assumed that the visual system does not respond directly to motion in the optic array. Instead, the perception of motion was considered to be a higher-order process resulting from the integration of a series of static images (e.g., 1). Similarly, it has typically been assumed that distance and spatial layout are perceived primarily from depth cues in the static retinal image. In contrast, Gibson (2) and Johansson (3) have argued that in normal perception the static retinal image does not exist. The visual environment is rarely static. Objects and other organisms move, the observer is often in motion, the head moves, and the eyes move almost constantly. Moreover, as Johansson (3) points out, the eye, unlike a camera, has no shutter to freeze the continuously changing flow of light on the retina. The visual system must function to detect information in the continuous change in the distribution of light reaching the eye.

In this view, motion in the optic array does not make perception of spatial layout more complicated. Rather, motion carries specific information for spatial layout. For example, motion parallax can specify the relative distances of points in the environment (4) and the shapes of three-dimensional objects (5), accretion and deletion of texture specifies depth at an edge (6), and expansion patterns in the optic array provide information for direction of locomotion and time to impact with an object or surface (7).

Although it is now clear that motion-carried information plays an important role in adults' spatial perception (e.g., 3, 8), research on infant perception typically has viewed motion primarily as a variable that controls attention, rather than as a source of invariant information for the three-dimensional structure of the environment. However, research on infants' sensitivity to motion-carried information for spatial layout has now begun in earnest. We have recently conducted investigations of infants' sensitivity to kinetic information for impending collision, accretion and deletion of texture specifying depth at an edge, and motion parallax specifying the layout of a surface in three-dimensional space.

2.1 Kinetic Information for Impending Collision

James Gibson (9) first described the optical transformations that specify the approach of an object and imminent collision of an object with a viewer. He pointed out that when an object approaches a viewer, there is a centrifugal flow of texture in the optic array. In addition, accelerated magnification that fills the entire visual field specifies imminent collision. Schiff (10) investigated the responsiveness of various animal species to such expansion patterns. Newborn dark-reared chicks, for example, respond defensively to the expansion, but not contraction, of a projected image. In humans, consistent responsiveness to information for collision appears to develop over the first few months. Several studies have found that infants respond by blinking and head withdrawal when presented with optical expansion, and the frequency of these defensive responses occurring increases between 1 and 3 months of age. Pettersen, Yonas, and Fisch (11) found that the likelihood of a blink in response to an approaching object increased between the sixth and tenth weeks in full-term infants. Post-term six-week-olds responded more consistently than did full-term infants at the same age. Maturation thus appears to play an important role in the development of sensitivity to collision information. Sensitivity may also increase as infants learn the consequences associated with a symmetrically expanding contour filling the visual field.

While the findings from early studies suggest that infants may perceive the approach and withdrawal of objects from expansion and contraction in the optic array, it is possible that infants' responses in these studies may be more reflex-like than truly perceptual. We reasoned that if blinking and head withdrawal were actually defensive responses to the information for collision specified by optical expansion, infants should exhibit these responses only when presented with an 'explosive' expansion pattern specifying dangerous collision. Yonas, Pettersen, Lockman, and Eisenberg (12) varied the optical expansion to determine which characteristics elicited defensive blinking and head withdrawal in 3-month-olds (12-14 weeks). A rear-projected image expanded to fill either 30 degrees or 100 degrees of the visual field; the expansion pattern corresponded either to a small object on an 'explosive' collision course with the viewer or a larger object that appeared to stop short of the viewer. Only the explosive expansion which filled a 100 degree visual field produced reliable blinking and backward head movement. Responding in other approach conditions did not differ from that to the contracting image in the withdrawal conditions. As Yonas (13) has discussed, this pattern of results suggests that 3-month-olds' responses to optical expansion are not simply reflex responses to non-spatial characteristics of the proximal stimulus. Rather, the 3-month-old appears to respond to specific spatial characteristics of the display, suggesting the presence of kinetically-based depth perception by three months of age.

The extent of younger infants' sensitivity to kinetic information for depth has been explored in another series of experiments. Although 4-week-olds blink more frequently to optical expansion than to contraction (14), they do not appear to discriminate among various expansion conditions. Backward head movements do not distinguish the one-month-olds' responses to object approach and withdrawal. One-month-olds blink less frequently to an approaching object than do three-month-olds, but blink equally to an 'explosive' expansion pattern and a non-accelerating expansion. Thus, a conclusion that depth sensitivity is present, based only on blink rate differences between expansion and contraction conditions, may be unwarranted. By three months of age, the evidence for sensitivity to kinetic information for collision is more compelling.

It is possible, however, that three-month-olds' blinking and head withdrawal in response to kinetic information for impending collision is due to an isolated, reflex-like process that is not related to other mechanisms detecting spatial information. In contrast, mature reaching is goal-directed spatial behavior that cannot be accounted for by an isolated reflex-like process. We have recently used reaching as a measure in several studies of infants' sensitivity to kinetic information for spatial layout. As a result, these investigations are limited to older infants who will reach.

2.2 Accretion and Deletion of Texture

In addition to producing expansion patterns in the optic array, movement by the observer or by objects in the environment results in disruption of the pattern of visible texture projected to the observer's eyes. Since terrestrial environments are filled with opaque objects and surfaces, the visual world is divided into visible and occluded surfaces at each possible observation point. When the observer moves through the environment, or when an object moves relative to an observer, some surfaces are occluded and others are revealed (2). At the level of the proximal stimulus, movement produces accretion and deletion of visible texture. For example, when a single object moves through a cluttered environment, its leading edge deletes texture from the proximal stimulus while texture is added at its trailing edge. Gibson (2, 15) has argued that since this change in the proximal stimulus is lawfully determined by the spatial layout of objects and surfaces in the environment, accretion and deletion of texture in the proximal stimulus unambiguously specify an edge and one surface in front of another at that edge. Gibson's hypothesis has been supported by Kaplan (6) who found that for adults, accretion and deletion of texture provides effective information for depth at an edge.

In a recent study, Granrud, Yonas, Smith, Arterberry, Glicksman, and Sorknes (16) found that 7-month-old infants can also use this information to perceive the spatial layout of surfaces. Seven-month-old infants (26-30 weeks) viewed computer-generated random-dot displays in which accretion and deletion of texture provided the only information for contours or for depth at an edge (see Figure 1). When the display was in motion, accretion and deletion specified a foreground surface moving in front of and

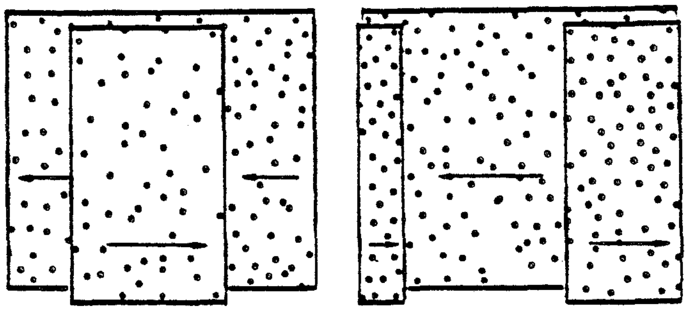


Figure 1. Schematic drawing of random dot display used for the study of sensitivity to accretion and deletion of texture. Two views are shown to depict change in the display over time. Interior lines indicate location of subjective depth contours.

occluding a moving background surface. When the display was motionless, it appeared to be a flat array of dots; the foreground and background were indistinguishable.

The infants reached significantly more often for the "foreground" than for the "background." A number of studies have shown that infants reach preferentially for the nearer of two surfaces when sufficient information for relative distance in available (17, 18). As no other information was available to distinguish the two regions, the reaching preference indicated that 7-month-olds are sensitive to accretion and deletion of texture as information for the spatial layout of surfaces.

Attempts to find evidence of 5-month-olds' sensitivity to accretion and deletion of texture were hampered initially by the low frequency of younger infants' reaching for the random-dot display used with 7-month-olds. Granrud and von Hofsten (19), therefore, tested 5-month-olds for sensitivity to accretion and deletion of texture produced either by a real object moving against a textured background or by an equally large aperture moving relative to the background. Less than 1 millimeter of differential depth separated the background from the object or aperture. Significantly more reaches were observed to the object than to the aperture. This was not the case in control conditions in which either the entire display was stationary or in which the object or aperture moved with the background, resulting in no accretion and deletion of texture. These results support the hypothesis that accretion and deletion of texture specifies spatial layout for infants as young as 5 months of age.

Granrud et al. (16) subsequently corroborated this finding. They presented 5-month-olds with a computer-generated random-dot display, in which accretion and deletion of texture again provided the only information for contours or for depth and specified two partially overlapping surfaces at different distances. The 5-month-olds showed a significant preference to reach for the apparently nearer region of the display. Again, since accretion and deletion was the only information that distinguished the two regions, this result indicated that 5-month-olds are sensitive to accretion and deletion of texture as information for the relative distances of two surfaces.

2.3 Motion Parallax

A third source of information carried by transformations in the optic array is motion parallax. As an observer's head or body moves, the retinal projections of objects move at differential velocities determined by the objects' distances. These relative motions in the optic array carry a considerable amount of information for the spatial layout of objects and surfaces in the environment. Although it has often been suggested that motion parallax is an important, or perhaps even primary, source of information used by infants for perceiving spatial layout (20, 21), until recently no empirical study had demonstrated infants' sensitivity to motion parallax. Rogers and Graham (4) developed a technique for isolating motion parallax from other sources of spatial information. They found, with adult subjects, that displays varying only the velocity of horizontally moving random dots can provide a compelling impression of three-dimensionality, not unlike that provided by stereopsis. In these experiments, each movement of the observer or of the display oscilloscope transformed the random-dot patterns to simulate the relative movement information produced by a three-dimensional surface.

Yonas, Smith, and Granrud (22) used Rogers and Graham's method, and moved an oscilloscope screen in front of 5-month-old infants. Two sinusoidal virtual surfaces were presented to each infant: a concave surface in which the horizontal central area of the screen was apparently beyond the reach of the infant, and a convex surface, in which the center of the screen appeared to be within reach. Thirty-second presentations of the two displays were alternated. Infants wore an eye patch to eliminate conflicting binocular information. Duration of reaching toward the screen in each condition was scored from a video recording. The convex display elicited reliably longer reaching than did the concave surface. This suggests that some sensitivity to motion parallax information for depth is present by 5 months of age. However, interpretation of these results should be tempered by the fact that we have, to date, been unable to replicate this finding.

2.4 Kinetic Information Conclusion

Our research suggests that 1-month-old infants may be sensitive to kinetic information specifying impending collision with an approaching object. Furthermore, a study by Gibson and Walker (23) indicates that 1-month-olds can perceive the rigidity or elasticity of objects from kinetic information in the optic array. Since there is now good evidence that stereopsis does not appear until about 3½ to 5 months (24), and that sensitivity to pictorial depth information may not appear until after 5 months (25, 26, 27, 28) the finding of sensitivity to kinetic information in 1- to 3-month-olds suggests that kinetic information is the first information to which infants respond in perceiving spatial layout. This developmental sequence is particularly interesting given the traditional view that motion perception is a higher-order process achieved through the integration of static images (1). If this view were correct, we might expect spatial perception based on kinetic information to be more complex and, therefore, a later achievement than perception based on pictorial information. The developmental primacy of sensitivity to kinetic information supports Gibson's (2) and Johansson's (3) theories that the visual system functions primarily to detect information for spatial layout from the continuously changing distribution of light in the optic array.

Finally, it should be noted that although infants are sensitive to kinetic information at a very early age, this sensitivity seems to undergo a good deal of development. The 1-month-old does not appear to be sensitive to all of the information for spatial layout carried by kinetic transformations in the optic array. The finding that 1-month-olds do not respond differently to optical expansion patterns specifying dangerous collision and gentle contact indicates that between 1 and 3 months infants become increasingly sensitive to specific variables in the optic flow. Future research will attempt to describe more precisely how infants' sensitivity to kinetic information changes during the first months of life.

3. BINOCULAR DEPTH INFORMATION

Kinetic information may be the only visual information for spatial layout to which infants respond during the first three months of life. However, infants' sensitivity to spatial information in the optic array undergoes rapid change during the first year, and there is now evidence that infants become sensitive to the spatial information provided by binocular vision between 3 and 5 months.

Binocularity provides information for spatial layout as a result of the eyes receiving slightly different views of the environment. There are two related sources of binocular information: (1) the convergence angle of the eyes in bifoveal fixation, and (2) the horizontal disparity between retinal images as the eyes fixate the same point in space. In studying infants, we can ask several different questions. For example, we might investigate the development of binocular sensitivity as a unified phenomenon, or we could study the development of sensitivity to binocular convergence or disparity in isolation. Finally, we might take a functional approach and ask whether, and at what point in development, infants can use binocular information in addition to monocular information to increase the veridicality of their spatial perception.

In a study that isolated binocular information for depth, Gordon and Yonas (29) asked whether infants' reaches would terminate at locations that varied systematically with the location of a stereoscopically-projected virtual object. The virtual object was presented using a stereoscopic shadow caster. Polarized filters were placed over the infant's eyes, and a real object was positioned between two point source lamps and a rear projection screen. In such a situation, an observer with binocular depth perception sees not a pair of shadows on the screen, but a single object at a precise location in space (specified by the convergence angle of the eyes and the disparity between the two images). Stereoscopic information specified a virtual object as within or beyond the reach of 5-month-old infants. An analysis of the infants' reaches revealed that the end points of the reaches did not vary reliably with the specified distance of the virtual object. However, binocular depth information did influence reaching. Infants reached more frequently and showed more prehensile activity when binocular information specified the virtual object as within reach (15 cm) than when it specified the object as beyond reach (30 cm). This study, therefore, revealed some binocular depth perception in 5½-month-olds (22-24 weeks).

Both convergence and disparity were available in the cast-shadow stereograms used by Gordon and Yonas (29). Several studies have now been conducted in an attempt to discover when sensitivity first appears to each of these sources of binocular information in isolation.

3.1 Convergence

Von Hofsten (30) investigated infants' ability to use convergence information in locating the position of a target object. He modified the convergence angle necessary for the infant's binocular fixation of the object and observed whether reaches were directed to the apparent location of the object. Infants from 4½ to 8 months wore 4-diopter displacing prisms which decreased convergence, making the object look farther away than it actually was; or wore 10-diopter displacing prisms which increased binocular convergence, making the object look closer than it actually was. Infants were tested under two lighting conditions: in the light, or in the dark with only the target object illuminated. Most reaches were directed to the location of the virtual object. Von Hofsten concluded that convergence is information for absolute depth by 4½ months. It could be argued, however, that the infants might have used binocular disparity to localize the virtual object, making use of the relative depth information provided by their hands and the virtual object. The possibility of this confound is strengthened by the fact that there was infrequent reaching when the infants were tested in the dark, a condition in which disparity cues relating the distance of hand and object were presumably absent. Thus, although infants can converge their eyes reliably to achieve bifoveal fixation by 2 months of age (31), we have no clear evidence to date that they can use convergence angle in isolation as distance information.

3.2 Binocular Disparity

Although a number of studies have investigated infants' stereopsis, it remains unknown when sensitivity to binocular disparity as information for spatial layout first appears. Atkinson and Braddick (32) found that two of four 2-month-olds showed a reliable preference for fixating a random-dot stereogram (33) which was paired with a non-stereoscopic random-dot pattern. In a second experiment, two infants (one of whom had shown a fixation preference in the first experiment) showed recovery from habituation of high-amplitude sucking to a disparity shift in a stereogram. These results suggest that disparity was detected in the stereogram. However, the small sample size and the finding that only one infant showed positive results in both experiments make these data difficult to interpret. A study conducted by Fox et al. (24) found more convincing evidence of infants' detection of binocular disparity. In Fox et al.'s study, 3½-month-olds tracked a moving virtual object specified by binocular disparity in a dynamic random-dot stereogram. Furthermore, tracking occurred only when the disparity was below the adult fusion threshold. In contrast, 2½-month-olds did not reliably track the virtual object. Held, Birch, and Gwiazda (34) found a similar developmental trend. They presented infants with a line stereogram paired with a similar zero-disparity stimulus in a forced-choice preferential looking procedure. The mean age at which the infants first showed a looking preference was 16 weeks.

These findings suggest that stereopsis first appears at about 3½ months of age. There is one difficulty with this conclusion, however. As we mentioned earlier, a persistent problem encountered when studying infant perception is disentangling responses to proximal stimulus cues from responses to the meaning specified by those cues. In studies of infants' stereopsis, detection of binocular disparity is not necessarily equivalent to perceiving the depth relations specified by disparity. In the studies by Atkinson and Braddick (32), Fox et al. (24), and Held et al. (34), the infants' fixation preferences and visual tracking could have been based on the detection of binocular disparity in the displays without the perception of depth. One might feel more secure in attributing stereoscopic depth perception to young infants if a spatially appropriate response such as reaching varied according to the spatial layout specified by binocular disparity. Granrud and Yonas (35) recented conducted a study using reaching as a measure to investigate infants' stereopsis. Five-month-olds viewed a pair of random-dot stereograms presented side by side. One stereogram contained crossed binocular disparity which specified a convex pyramid-shaped virtual object projecting outward from the picture plane toward the infant; in the other, an equal amount of uncrossed disparity specified a concave pyramid projecting away from the infant. The infants showed a significant preference to reach for the stereogram specifying a convex object. This finding provided evidence that 5-month-olds perceive the spatial layout specified by binocular disparity. This conclusion must be made cautiously, however. Few infants reached frequently for the stereograms and a number of infants did not reach at all. Moreover, there was a large amount of variance in the data from this experiment. Thus, we are now attempting to replicate this study in order to establish with greater confidence that 5-month-olds perceive depth from binocular disparity.

3.3 The Effects of Binocular Vision on Infants' Depth Perception

It is clear that infants by 5 months of age can perceive depth from binocular information (29, 35). Five-month-olds' sensitivity to binocular depth information, however, does not necessarily indicate that binocularity increases the veridicality of these infants' spatial perception. Jones and Lee (36) recently found that binocular vision enables adults to perceive the environment and control their actions in three-dimensional space more accurately than does monocular vision. Given the superiority of binocular depth perception in adults, we might expect binocular depth perception to be superior in infants as well. It is possible, however, that although stereopsis appears between 3 and 5 months, kinetic information, the first depth information to which infants respond, continues to be the most effective source of information for perceiving spatial layout through infancy.

Few studies to date have investigated the effects of binocularity on infants' depth perception. Early research showed that binocular information is not necessary for infants' spatial perception. Walk and Dodge (37) found that a congenitally monocular 10-month-old infant behaved similarly to normal binocular infants on a visual cliff. Walk (38) also found no difference in behavior on a visual cliff between normal binocular infants and normal infants with one eye patched. Of course, the finding that monocular vision is sufficient for discriminating the deep and shallow sides of a visual cliff does not demonstrate that binocular vision has no effect in infants' depth perception. More recently, Granrud, Yonas, and Pettersen (19) compared binocular with monocular depth perception in 5- and 7-month-olds, using reaching as the dependent measure. They presented infants with a small disc positioned within reach and a larger disc just beyond reach. The visual angles subtended by the discs were matched. The infants in both age groups showed a remarkably consistent reaching preference in the binocular condition, reaching for the nearer object on almost every trial. In contrast, in the monocular viewing condition, the 5- and 7-month-olds showed only slight tendencies to reach for the nearer object. These results indicated that binocular vision was superior to monocular vision for perceiving the objects' distances. Moreover, binocular vision seemed to result in highly veridical spatial perception while the infants' only slightly greater than chance performance in the monocular condition suggests that monocular vision resulted in only moderately veridical spatial perception. This raises an interesting question concerning the development of spatial perception: how veridical is spatial perception before the onset of stereopsis?

The best evidence currently available suggests that stereopsis first appears no earlier than about 3 to 4 months of age (24, 34) and perhaps as late as 5 months. Although 3-month-olds and younger infants can perceive spatial layout from kinetic information, the only moderate effectiveness of monocular information, including motion parallax, in Granrud et al.'s (17) study suggests that prior to the onset of stereopsis infants may be unable to achieve high levels of veridicality in spatial perception. The effect of the onset of sensitivity to binocular disparity on the 3- or 4-month-old's spatial perception remains unknown. Perhaps the development of stereopsis results in a sharp increase in the veridicality of spatial perception. On the other hand, if the advantage of binocular vision results primarily from binocular concordance (i.e., matched information in the light reaching the two eyes), as Gibson (2) and Jones and Lee (36) have suggested, and not from binocular disparity, binocular vision might enable infants to perceive the spatial layout of their environments more accurately than would monocular vision even before sensitivity to binocular disparity appears. We hope to test these hypotheses in the future using both reaching and transfer of habituation methodologies (see 39).

4. PICTORIAL INFORMATION

The third general category of optical information for spatial layout comprises the pictorial "cues" for depth. Interest in the pictorial cues initially arose in response to the technical problem of portraying depth in a painting. These cues were first described in detail by Leonardo da Vinci (40). When psychologists began to explore the stimulus variables that account for spatial perception, they assumed that the same cues which evoke the apparent three-dimensionality in a flat painting might play a role in our perception of the real world's spatial layout (e.g., 41). There is now considerable evidence that adults can perceive distance from a number of pictorial cues such as texture gradients (42), relative size (43, 44, 45, 46), familiar size (47, 48, 49, 50), interposition (51, 52, 53) and shading (54). Until recently, the role of pictorial information in infants' spatial perception had not been studied. However, we have now investigated infants' sensitivity to several types of pictorial depth information. Our first studies explored infants' sensitivity to Ames' trapezoidal window illusion (55) which combines several pictorial cues. Subsequent studies have tested infants for sensitivity to the cues of familiar size, relative size, interposition, and shading, isolated from other sources of depth information.

4.1 Trapezoidal Window Studies

Yonas, Cleaves, and Pettersen (26) investigated infants' responsiveness to the illusion of the Ames trapezoidal window (see Figure 2A & 2B). The display was created by photographing a rectangular window rotated 45 degrees about its vertical axis, and cutting out the internal spaces from the photograph. Adults viewing this display monocularly experience a powerful illusion of a slanted rectangular window. The pictorial cues of relative size, linear perspective, angular perspective, and shading combine to create this illusion. The expectation was that infants who were responsive to the depth information in the display would reach to the apparently nearer side of the window. A preliminary study determined that infants' reaching was sufficiently accurate to differentiate the distances of the sides of a similar real slanted window. In the first experiment, 7-month-olds were shown either a trapezoidal window display (Figure 2B) or a control display with sides of different sizes but without differential distance information (Figure 2C). Both displays were viewed monocularly. Infants reached to the larger (apparently nearer) side of the trapezoidal window twice as often as to the smaller side. Reaching to the size-control display was equally frequent to both sides. A second experiment compared reaching with monocular and binocular viewing of the trapezoidal window. Binocular viewing significantly reduced the tendency to reach for the pictorially nearer side of the display. Thus, it appears that pictorial depth accounted for the preferential reaching and that 7-month-olds perceive the trapezoidal window as a slanted rectangular window, as do adults. Finally, 5-month-old infants were shown (monocularly) either a real rectangular window rotated 45 degrees or the trapezoidal window, presented in the fronto-parallel plane. These infants reached consistently to the nearer side of the real window, but reached equally to either side of the trapezoidal window. This result suggests that pictorial information is ineffective at 5 months, or that conflicting information (accommodation and motion parallax) override its effects.

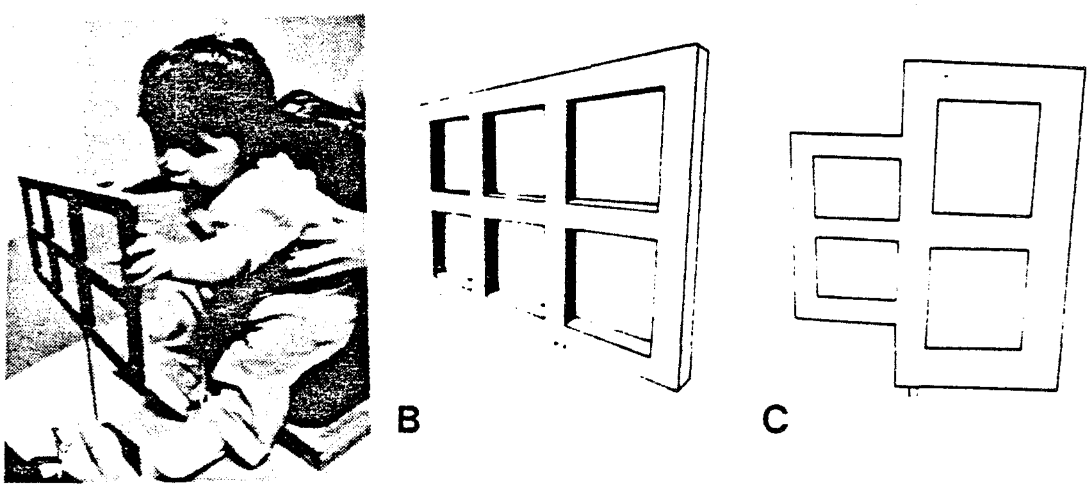


Figure 2. Infant reaching for trapezoidal window. B. Line drawing of trapezoidal window oriented perpendicular to line of sight. C. Line drawing of control display.

A later study by Kaufmann, Maland, and Yonas (56) further investigated infants' responsiveness to the pictorial information in the trapezoidal window. In their experiment, the window was presented with the smaller side rotated toward the viewer. When viewing this display monocularly, adults report that the orientation of the display becomes ambiguous, as pictorial and other cues are put in conflict. Kaufmann et al. tested 7-month-olds, whose reaching suggested a similar experience; fewer reaches were directed to the near side of the rotated trapezoidal window than to the near side of a real rotated object. Five-month-olds, however, reached equally to the near side of the trapezoidal window and the real window. This behavior, which appears to be uninfluenced by the pictorial depth information in the trapezoidal window, suggests that sensitivity to such information develops between 5 and 7 months of age.

4.2 Familiar Size

After establishing that infants can respond to pictorial information for the spatial layout of a surface (the trapezoidal window), we began a series of studies designed to investigate infants' sensitivity to various pictorial cues isolated from all other spatial information. The first of these studies examined infants' sensitivity to familiar size.

The depth cue of familiar size is based on the lawful relationship between an object's distance from an observer and the visual angle subtended by the object at the observer's station point: as distance increases visual angle decreases. However, since visual angle is determined by both object size and object distance, visual angle alone carries no information for size or distance. If an object's size were known, however, visual angle could be used as information for distance. Despite numerous controversies concerning the effectiveness of familiar size as a depth cue, there is now convincing evidence that adults and infants can use familiar size to judge the distances of familiar objects.

Recently Yonas, Pettersen, and Granrud (28) investigated 5- and 7-month-olds' sensitivity to familiar size as a cue for distance. Each age group was subdivided into three treatment groups (an experimental group and two control groups). The experimental group viewed monocularly (and one at a time) larger than life-size and smaller than life-size photographs of adult female faces, a familiar class of objects (see figure 3). One control group viewed these same photographs binocularly. The second control group viewed monocularly large and small checkerboard ovals (a class of objects with no specific familiar size) that were equal to the faces in size.

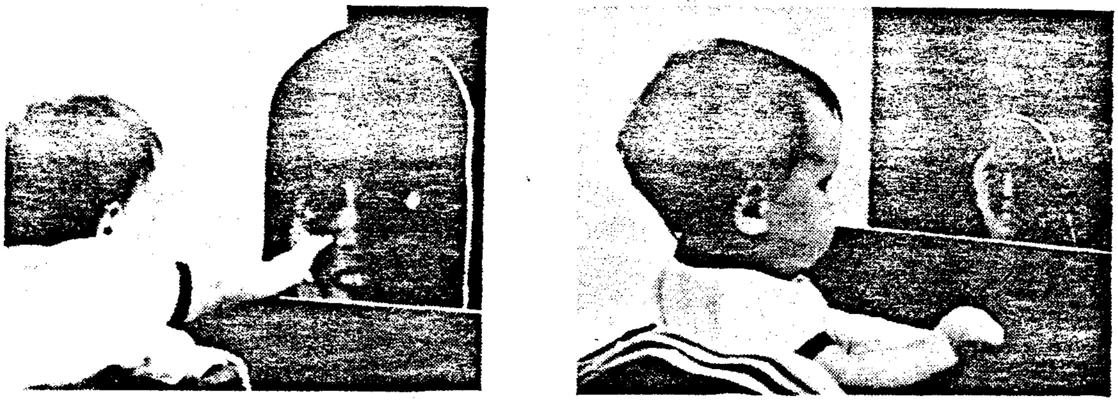
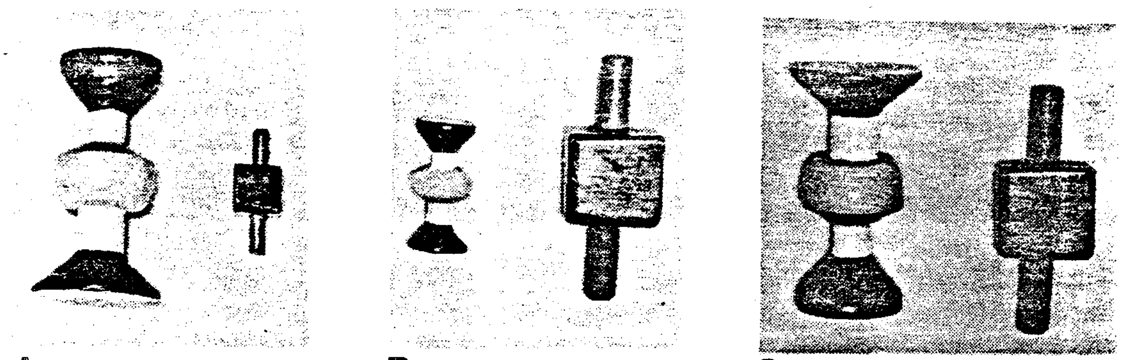


Figure 3. Infant viewing face displays used in familiar size study.

The results indicated that perceived distance was determined by familiar size for the 7-month-olds. In the experimental condition, the 7-month-olds reached significantly more for the large faces than for the small faces, and showed no significant differences in their reaching for the large and small faces viewed binocularly or for the large and small checkerboards. Results from the control conditions rule out the possibility that the preference to reach for the large faces in the experimental condition was due to a preference to reach for large objects without regard to apparent distance. Thus, in the experimental (i.e., monocular) condition, the infants apparently perceived the large faces as within reach and the small faces as beyond reach, on the basis of the familiar size of faces. This conclusion is further supported by Yonas et al.'s (28) findings for a group of adults who viewed these same displays monocularly. The mean distance judgments were 8.0 inches (22.6 cm) for the large faces, a distance within reach for most 5- and 7-month-olds; and 20.6 inches (52.3 cm) for the small faces, a distance well beyond these infants' reach. In the binocular control condition, the lack of a reaching preference indicated that the infants detected binocular information which overrode familiar size and specified the faces' actual distances. The 5-month-olds' results were equivocal. Although the duration-of-reaching data were similar for the two age groups, the experimental effects for the 5-month-olds were smaller and their reaching performances did not differ significantly between conditions. While these results suggest that 5-month-olds may use familiar size as distance infor-



A B C

Figure 4. Objects used in familiar size study. A and B. Pairs of objects seen by infants during familiarization. C. Pair of objects seen by infants during test phase.

mation, they provide no clear evidence of 5-month-olds' sensitivity.

A recently completed study by Granrud, Haake, and Yonas (57) further investigated infants' ability to use familiar size in distance perception, using objects with which the infants had only brief experience. There were two phases in the experiment: a familiarization phase and a test phase. During the familiarization phase the infant played with a pair of objects, either pair A or pair B pictured in Figure 4, for ten minutes. During the test phase, a pair of objects identical to those seen in the familiarization phase, but now equal in size (pair C in Figure 4), was presented to the infant at a fixed distance.

It was predicted that if the infants remembered the sizes of the two objects from the familiarization phase and could use these familiar sizes to perceive the objects' distances, the objects in the test phase should be perceived as being at different distances. Since the two objects subtended equal visual angles, the object with the smaller known size (the object that resembled the smaller object from the familiarization pair) should be perceived as nearer. Reaching was used as the dependent measure. Under monocular viewing conditions, a reaching preference for the object in pair C which resembled the smaller object in the familiarization pair might indicate sensitivity to familiar size as information for relative distance. With binocular depth cues available to specify that the objects were actually equidistant, however, reaches should be equally distributed between the objects.

The results from 7-month-olds confirmed the predictions. In the test phase of the experiment, 7-month-olds viewing the objects monocularly showed a significant preference to reach for the object resembling the smaller object in the familiarization phase. Seven-month-olds in the binocular condition reached equally to the two objects in the test phase. These results indicated that in the monocular condition the 7-month-olds could use the objects' familiar sizes acquired during the familiarization phase to extract information for distance from the objects' visual angles in the test phase. The lack of a reaching preference in the binocular condition rules out interpretations of the results not based on depth perception. It is interesting that familiar size information operates as a source of depth information after such brief experience with novel objects. This finding refutes Helmholtz's (41) hypothesis that distance perception based on familiar size requires such extensive experience that children are insensitive to this cue. It also conflicts with Brunswik's (58) hypothesis that perceptual learning requires "slow stamping in" (p. 241) by experience. This finding further refutes the notion that the familiar size effect is, for adults, a result of a cognitive strategy (59). It appears more likely that the effect, rather than reflecting a conscious decision, is an example of a powerful direct impact of memory on perception.

Unlike 7-month-olds, 5-month-olds in this study gave no evidence of sensitivity to familiar size. These infants reached equally for the two objects in the test phase of the experiment. We cannot conclude from this result that 5-month-olds are insensitive to familiar size. It is possible, of course, that future studies might reveal sensitivity in these infants. However, if 5-month-olds are insensitive to familiar size it is probably not due to an inability to encode and store information about objects in memory since results from a number of studies suggest that 5-month-olds have impressive memory capabilities (e.g., 60). Rather, insensitivity to familiar size would most likely be due to an inability to use memory in perception. Granrud et al's (57) findings, therefore, suggest the hypothesis that while 7-month-olds' perception is not completely determined by the information present in the visual input, insofar as memory can be used to perceive distance from visual angle, 5-month-olds' perception may be determined directly by the visual input. The ability to use memory in

Figure 5. Infants reaching for pairs of objects in relative size study.

perception might first appear between 5 and 7 months of age.

4.3 Relative Size

The 7-month-old's visual system apparently can use knowledge of an object's physical size to constrain the interpretation of its angular size and recover information for the object's distance. When two or more objects are viewed, the objects' relative angular sizes can be used to achieve additional constraints on the interpretation of visual angle. When two objects that are equal in size are at different distances from an observer, the nearer object subtends a larger visual angle than the farther. If the two objects can be assumed to be approximately equal in distal size, the relationship between the objects' visual angles can be a source of information for the objects' relative distances. This source of information, or depth cue, is called relative size.

Although there is substantial evidence that adults can employ relative size information to perceive relative distance (e.g., 43, 44, 45, 46, 61), no research on the development of this ability in infants had been carried out until recently. Yonas, Granrud, and Pettersen (27) explored sensitivity to relative size in infants from 5 to 7 months of age. In these experiments, infants viewed two different-sized objects of various shapes (squares, circles, and triangles) presented side by side suspended in front of a dark vertical surface (see figure 5). The two objects were always at equal distances from the infant. Seven-month-old infants were presented with pairs of discs and triangles, both monocularly and binocularly. They showed a significantly greater tendency to reach for the larger object in the monocular condition than in the binocular condition.

This same pattern of reaching preferences was observed when a group of 5½-month-olds (22 to 24 weeks) was presented with a pair of discs. They reached for the larger disc on more trials with monocular presentation than on binocular trials. Five-month-olds' (20 to 22 weeks) behavior in the monocular and binocular conditions did not differ. In both conditions, they reached more frequently to the larger object.

The results indicate that for 5½- and 7-month-old infants the relationship between angular sizes of objects is an effective source of information for those objects' relative distances. Results from the binocular viewing conditions eliminate the possibility that infants preferred to reach for the larger objects without regard to their apparent distances. If this were the case, infants should have reached more frequently for the larger objects under both binocular and monocular viewing conditions, but no such preference was observed in the binocular condition. Thus, the most plausible explanation for the results is that infants perceived the larger objects as closer than the smaller objects in the monocular conditions, and detected stereoscopic information which specified the objects as equidistant in the binocular conditions. In contrast, the 5-month-old infants did not provide evidence of sensitivity to relative size.

4.4 Interposition

Another recent study of infants' sensitivity to pictorial depth information investigated 5- and 7-month-olds' perception of interposition (25). In the first experiment, infants viewed monocularly the displays pictured in Figure 6A and 6B. Figure 6A is the interposition display. Adults tend to see this figure as consisting of three overlapping diamonds, rather than as one diamond and two adjacent arrowheads. The control display, shown in Figure 6B, contains no information for overlap. It was predicted that if the infants perceived the interposition display as consisting of overlapping surfaces, they would reach preferentially for the apparently nearer side of the display.

The 7-month-olds showed a significant preference to reach for the apparently nearer part of the interposition display. Five-month-olds also reached preferentially for the apparently nearer part of the interposition display. However, this preference did not differ significantly from chance. Neither group of infants showed a reaching preference when viewing the control display.

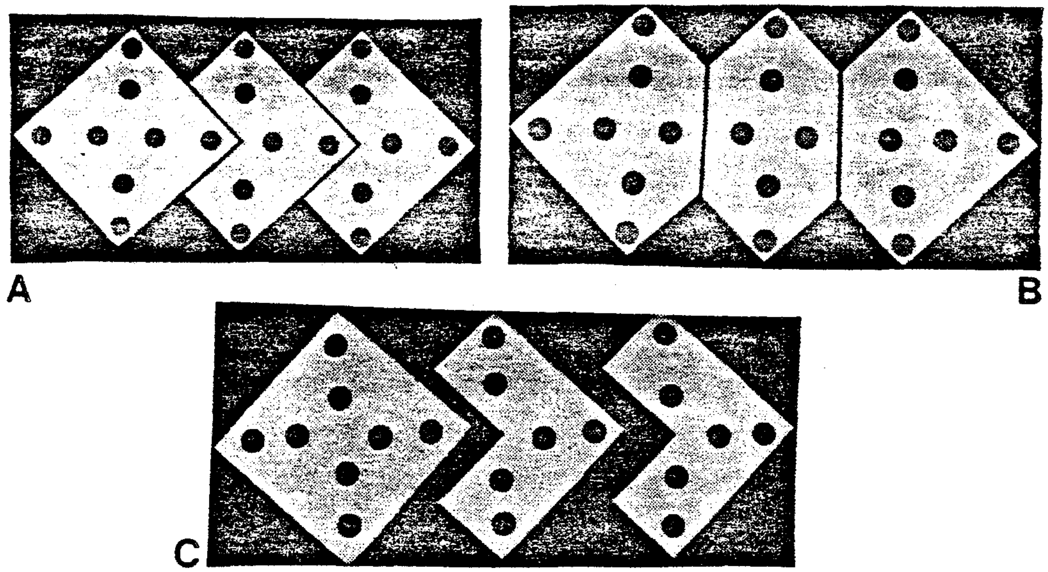


Figure 6. Displays used in interposition study. A. Stimulates three overlapping surfaces. B and C. Control displays.

These results suggest that by 7 months of age infants might perceive pictorial interposition. It is possible, however, that the infants' reaching preferences were due to a preference to reach for the larger, diamond-shaped, section in the interposition display rather than to the perceived relative distances of the display's sections. A second experiment was conducted to control for this possibility. Five- and 7-month-olds were presented with the interposition display (Figure 6A) and a control display (Figure 6C). The two displays share many proximal stimulus features. Since the sections of the control display were separate, however, this display contained no information for overlap. If the infants' reaching is determined by interposition, that is, by the apparent relative distances of the display sections, they should prefer to reach for the pictorially nearer side of the interposition display but not for the diamond shape in the control display. On the other hand, if the reaching preference observed in the first experiment can be accounted for by non-spatial proximal stimulus variables such as the sizes and shapes of the display sections, the infants should show equal reaching preferences in the two conditions.

The 7-month-olds again showed a significant preference to reach for the apparently nearer part of the interposition display, and although they also showed a slight preference to reach for the diamond-shaped section in the control display, their reaching preference was significantly greater in the interposition condition than in the control condition. This result indicated that 7-month-olds are sensitive to pictorial interposition. Since the

7-month-olds' reaching preference when viewing the interposition display cannot be accounted for by non-spatial proximal stimulus variables such as the size and shape of the display sections, it seems that the interposition display was perceived as consisting of overlapping surfaces, and reaches were for the apparently nearer side. The 5-month-olds also showed a tendency to reach for pictorially nearer side of the interposition display. However, this tendency did not differ significantly either from chance or from their tendency to reach for the diamond shape in the control condition. Therefore, even if the 5-month-olds' tendency to reach for the pictorially nearer side of the interposition display represented an actual preference, this could be accounted for entirely in terms of non-spatial proximal stimulus variables such as size and shape. Thus, these data provide no evidence that 5-month-olds perceive pictorial interposition.

4.5 Shape from Shading

We recently demonstrated infants' sensitivity to one final pictorial depth cue: shading. Shading provides information for the three-dimensional shapes of objects because of the lawful relationship that exists between the orientation of a surface and the luminance level in the retinal image projected by that surface. A surface that is oriented orthogonally to the source of illumination will be highly illuminated and will reflect a large amount of light. A surface that is oblique or parallel to the incident light will be less strongly illuminated and will reflect less light. As a result, for surfaces with uniform reflectance, smoothly varying gradients of luminance (i.e., shading) in the proximal stimulus correspond to changes in surface orientation relative to the source of illumination (when illumination is constant across a scene). This correspondence between distal surface orientation and shading in the proximal stimulus can provide a perceiver with information for the three-dimensional shapes of objects (2, 42). Figure 7 illustrates shading information for shape. Adults usually see this as a photograph of a surface containing a convexity (the region more brightly illuminated on top) and a concavity (the region more brightly illuminated on the bottom). When the photograph is inverted, the depth reverses. This demonstrates that shading is the only information for shape in the photograph. It also indicates that the visual system employs an "assumption", or constraint that light comes from above.

In the first study of the development of sensitivity to

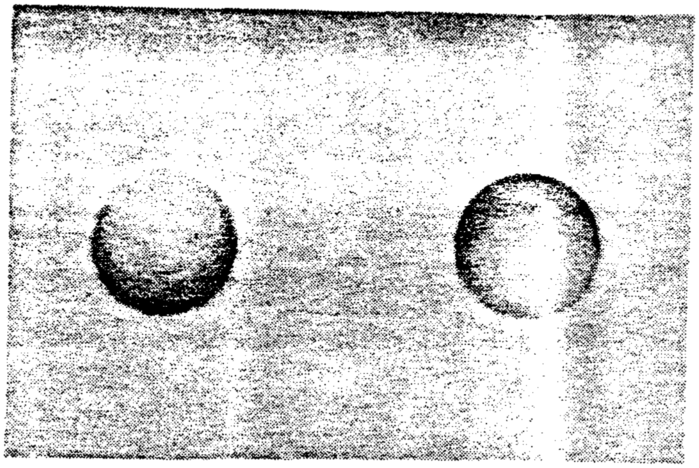


Figure 7. Convexity and concavity depicted by shading.

shading, Benson and Yonas (62) presented 3-year-olds with a surface containing a convexity and a concavity and trained them to point to the convexity. They then showed the children a photograph of the surface containing only shading information for shape (see Figure 7). The children exhibited transfer of training to the photograph, pointing to the "convexity", the area which is brighter on top, on 86% of the trials. In a similar study, Granrud, Yonas, and Opland (63) found that 7-month-olds perceive shape from shading. There were two conditions in the experiment. In one condition, 7-month-olds viewed a surface containing a convexity and a concavity. In the other, they viewed an enlargement of the photograph in Figure 7. With monocular vision, the infants reached preferentially for the convexity when viewing both the actual surface and the photograph. With binocular vision, the infants reached preferentially for the convexity in the actual surface but reached equally for the two areas in the photograph.

The results from the real surface condition indicated that 7-month-olds reach preferentially for convexities. The infants' preference to reach for the "convexity" in the monocularly viewed photograph indicated that they perceived this region to be an actual convexity. This result demonstrated these infants' perception of shape from shading. This conclusion is supported by the finding of no reaching preference when the infants viewed the photograph binocularly which indicated that the infants' reaching was based on perceived depth and not on non-spatial proximal stimulus variables, such as a preference to reach for the region in the photograph that was more highly illuminated on top. Five-month-olds are currently being tested in this experiment.

Infants' perception of shape from shading is particularly interesting because it illustrates the sorts of constraints, or "decoding principles" (64), that might operate in the visual system. Information for an object's shape is confounded in retinal image luminance since a given luminance value could result from an infinite number of combinations of three variables: surface orientation, surface reflectance, and illumination intensity. In order to recover information for shape from luminance gradients (i.e., shading), the visual system apparently constrains the possible interpretations of the proximal stimulus through a process analogous to making assumptions about the scene that is being perceived. If there is no evidence to the contrary, the visual system seems to operate on the "assumptions", or constraints, that illumination comes from above, that illumination is uniform across the scene, and that reflectance is uniform across a surface (65, 66). Without these constraints, shading would be ambiguous and there would be no basis for perceiving a convexity and a concavity in Figure 7. That infants perceive a convexity and a concavity when viewing this photograph suggests that the visual system incorporates these constraints very early in life.

4.6 The Development of Sensitivity to Pictorial Information

A consistent trend has emerged from our studies on infants' sensitivity to pictorial depth information. Yonas et al. (26) and Kaufmann et al. (56) found that 6- and 7-months-olds are sensitive to the spatial information in an Ames trapezoidal window, but failed to find such sensitivity in 5-month-olds. Yonas et al. (28) and Granrud et al. (57) demonstrated 7-month-olds' sensitivity to familiar size as depth information, Yonas et al. (27) found 7-month-olds to be sensitive to relative size as depth information and Granrud and Yonas (25) found sensitivity to pictorial interposition in 7-month-olds; none of these studies found evidence of sensitivity in 5-month-olds. It is unlikely that the failure to find sensitivity to pictorial depth information in 5-month-olds can be attributed to the inadequacy of reaching as the dependent measure, since Granrud et al. (17) have found that both 5- and 7-month-olds show highly consistent preferences to reach for the nearer of two objects when sufficient information for relative distance is available. Moreover, reaching has been used successfully as the dependent measure in studies demonstrating 5-month-olds' sensitivity to several types of depth information including accretion and deletion of texture (16, 19), binocular disparity (35), and a variety of combinations of cues (17, 18, 26, 29). Therefore, although we cannot be certain that 5-month-olds are insensitive to pictorial depth information, the current data suggest that while 5-month-olds are sensitive to kinetic and binocular information for perceiving spatial layout, sensitivity to pictorial information probably emerges between 5 and 7 months of age. This hypothesis is further supported by findings by Owsley (57) and Kellman (68) indicating that 4-month-olds may be unable to perceive objects' three-dimensional shapes from a sequence of static views.

Our consistent finding that the earliest evidence of sensitivity to pictorial information is observed between 5 and 7 months further suggests the hypothesis that infants may become sensitive to all pictorial cues simultaneously. If future research supports this hypothesis, it will conflict with Brunswik's (69) hypothesis that the acquisition of responsiveness to proximal stimulus cues is dependent on learning the correlation between a particular cue and the distal properties of the environment, since this hypothesis would predict that the appearance of sensitivity to different cues should be staggered, in a sequence corresponding to a hierarchy of cue validity. Moreover, this might suggest that a single mechanism (which begins to function after 5 months of age) exists in the visual system which is necessary for responsiveness to all pictorial cues.

5. CONCLUSION

An interesting pattern has emerged from our studies of infants' sensitivity to various types of spatial information. Kinetic information seems to the first depth information to which infants respond. The studies described in this chapter indicate that sensitivity to kinetic information for impending collision develops no later than 3 months and may be present as early as 1 month. We believe that binocular depth information is the second general class of spatial information to which infants become sensitive, with stereopsis developing between 3½ and 5 months of age. Sensitivity to pictorial information seems to appear last, developing after 5 months of age.

The staggered development of sensitivity to these three classes of information suggests that at least three distinct systems function in the detection of visual information for spatial layout: a kinetic depth system, a binocular depth system, and a pictorial depth system. On the other hand, the concurrent appearance of sensitivity to several different cues indicates that a common mechanism may be involved in detecting each cue. For example, our consistent finding of the first evidence of sensitivity to pictorial information between 5 and 7 months suggests that sensitivity to many or all pictorial depth cues depends on the development of a single mechanism. A single mechanism may also be involved in the detection of depth from kinetic information. At 5 months (20-22 weeks), when infants are unresponsive to pictorial information, they reach to the location of surfaces specified by the kinetic cues of accretion and deletion of texture. If detection of kinetic information is accomplished by a single mechanism, sensitivity to expansion information, accretion and deletion of texture, and motion parallax should appear simultaneously.

Of course, one should not forget that we can demonstrate the effectiveness of each depth cue in isolation. That is, each cue can function as a separate "module", to use Marr's (70) term. Marr's principle of modularity states that any large computational process should be divided into a set of specialized, nearly independent, subprocesses, or modules. He argued that if visual information processing were not organized in a modular way, incremental changes in design, presumably an essential requirement for evolution, would be unable to result in the improvement of one aspect of perception without degrading the performance of many others (see pp. 325-326). While the ability to respond to single depth cues in isolation suggests modularity of design in the infant's visual system, the concurrent development of sensitivity to many different pictorial cues raises the question of whether a common brain mechanism might underlie this sensitivity.

The gradual development of perception in human infants offers us an exceptionally fertile field for research investigating the mechanisms involved in spatial perception. Since the systems involved in detecting kinetic, binocular, and pictorial information seem to develop at different times in infancy, we can study infants' spatial perception when only one or two of the three systems are functioning. Thus, we can ask how veridical perception is when the infant can detect only kinetic information. In addition, we can investigate the contribution of binocularity to spatial perception by studying infants immediately before and after the onset of stereopsis. Similarly, by investigating infants' perception before and after the appearance of sensitivity to pictorial depth cues, we may be able to determine whether these cues are important in spatial perception or are only "secondary" cues as has often been suggested (2, 42, 71).

Infants also offer us a natural control group for studying perception with a minimal intrusion of cognitive factors. It has been argued, for example, that adult subjects' distance judgments in familiar size experiments may not be based on the immediate perception of distance specified by familiar size. Instead, subjects might infer the distance at which the familiar object would project a given visual angle (59, 72, 73). Yonas et al.'s (28) and Granrud et al.'s (57) findings that 7-month-olds respond to distance specified by familiar size, however, indicate that familiar size is an effective determinant of perceived distance. It is likely that the 7-month-olds' reaching was determined by the immediately perceived distances of the objects, and not by inferences or conscious estimates of the objects' distances. Similarly, our demonstrations of infants' sensitivity to the cues of relative size, interposition, and shading indicate that these cues are effective sources of information for depth perception. Results from these studies refute the hypothesis that pictorial cues are conventions invented by artists for portraying depth in a painting (74).

We should note that perception researchers have shown little interest in pictorial depth cues in recent years. This has perhaps been due to a shift in fashion or to the belief that pictorial cues are artists' conventions or are unimportant in depth perception. The research described in this chapter offers no support for these views. Furthermore, recent research in computer vision has now revealed the richness of information for spatial layout carried by pictorial cues (65, 66, 75, 76, 77; for reviews of this research see 78, 79). We hope that researchers who are attempting to understand spatial perception and the neural mechanisms underlying spatial vision will include in their work the full range of spatial information: kinetic, binocular, and pictorial cues. All three types of information are, in our view, equally interesting and worthy of study.

We also hope that the methods we have devised for studying spatial perception in human infants, particularly preferential reaching, will prove useful for those exploring spatial vision in other species. We believe that reaching is a powerful index of spatial sensitivity, and that it has been used by researchers less often than it deserves (for an extensive discussion of reaching as a measure of spatial perception, see Yonas & Granrud, 80).

Finally, the gradual development of perception in infancy may offer us an opportunity to discover the neural mechanisms underlying spatial perception. Concurrent development of a particular brain structure and sensitivity to a particular class of depth information might imply a relationship between the two. Thus, if we can link the study of developmental neurophysiology to behavioral studies of the development of responsiveness to visual information, it may be to the advantage of both fields. Progress in understanding the development of perception will undoubtedly be facilitated by progress in neurophysiology. It also seems realistic that understanding perceptual development in infancy will aid in discovering the brain mechanisms involved in spatial perception.

ACKNOWLEDGEMENTS

Preparation of this chapter was supported by National Institute of Child Health and Human Development grants HD-05027 and R01-HD-16924-01. This chapter was a collaborative effort; order of authorship does not reflect the relative contributions of the authors. The authors thank Isabel Smith and Martha Arterberry for their editorial assistance and Kaye O'Geay for typing the manuscript. Correspondence concerning this chapter should be addressed to Albert Yonas, Institute of Child Development, University of Minnesota, Minneapolis, MN, 55455, USA.

REFERENCES

1. Titchener, E. B. A Text-Book of Psychology (New York, Macmillan, 1910).
2. Gibson, J. J. The Senses Considered as Perceptual Systems (Boston, Houghton-Mifflin, 1966).
3. Johansson, G. Visual Event Perception. In R. Held, H. W. Leibowitz, & H. L. Teuber (Eds.), Handbook of Sensory Physiology: Perception (Berlin, Springer-Verlag, 1978).
4. Rogers, B., & Graham, M. Motion Parallax as an Independent Cue for Depth Perception. Perception 8(1979) 125-134.
5. Wallach, H., & O'Connell, D. N. The Kinetic Depth Effect. Journal of Experimental Psychology 45(1953) 205-217.
6. Kaplan, G. A. Kinetic Disruption of Optical Texture: The Perception of Depth at an Edge. Perception and Psychophysics 6(1969) 193-198.
7. Lee, D. N. The Optic Flow Field: The Foundation of Vision. Philosophical Transactions of the Royal Society of London-B 290(1980) Issue 1038, 169-179.
8. Johansson, G., von Hofsten, C., & Jansson, G. Event Perception. Annual Review of Psychology 31(1980) 27-63.
9. Gibson, J. J. Visually Controlled Locomotion and Visual Orientation in Animals. British Journal of Psychology 49(1958) 182-194.
10. Schiff, W. Perception of Impending Collision: A Study of Visually Directed Avoidant Behavior. Psychological Monographs 79(1965) Whole no. 604.
11. Pettersen, L., Yonas, A., & Fisch, R. O. The Development of Blinking in Response to Impending Collision in Preterm, Full-Term, and Postterm Infants. Infant Behavior and Development 3(1980) 155-165.
12. Yonas, A., Pettersen, L., Lockman, J. J., & Eisenberg, P. The

Perception of Impending Collision in 3-month-old Infants. Paper presented at the International Conference on Infant Studies, New Haven, Ct., 1980.

13. Yonas, A. Infants' Responses to Optical Information for Collision. In R. N. Aslin, J. Alberts, & M. Petersen (Eds.), Development of Perception: Psychobiological Perspectives. The Visual System, Vol. 2 (New York, Academic Press, 1981).
14. Yonas, A., Pettersen, L., & Lockman, J. J. Young Infants' Sensitivity to Optical Information for Collision. Canadian Journal of Psychology 33(1979) 268-276.
15. Gibson, J. J. The Ecological Approach to Visual Perception (Boston, Houghton-Mifflin, 1979).
16. Granrud, C. E., Yonas, A., Smith, I. M., Arterberry, M. E., Glicksman, M. L., & Sorknes, A. C. Infants' Sensitivity to Accretion and Deletion of Texture as Information for Depth at an Edge. Child Development, in press.
17. Granrud, C. E., Yonas, A., & Pettersen, L. A Comparison of Monocular and Binocular Depth Perception in 5- and 7-Month-Old Infants. Journal of Experimental Child Psychology, in press.
18. Yonas, A., Sorknes, A. C., & Smith, I. M. Infants' Sensitivity to Variation in Target Distance and Availability of Depth Cues. Paper presented at Society for Research in Child Development, Detroit, 1983.
19. Granrud, C. E., & von Hofsten, C. The Role of Accretion and Deletion of Texture in Infants' Perception of Objects. Unpublished manuscript, 1983.
20. Bower, T. G. R. The Visual World of Infants. Scientific American 215(1966) 80-92.
21. Gibson, E. J. Principles of Perceptual Learning and Development (New York, Appleton-Century-Crofts, 1969).
22. Yonas, A., Smith, I. M., & Granrud, C. E. Responsiveness of 5-month-old Infants to Motion Parallax Information for Depth. Unpublished manuscript, 1983.
23. Gibson, E. J., & Walker, A. S. Development of Knowledge of Visual-Tactual Affordances of Substance. Child Development, in press.
24. Fox, R., Aslin, R. N., Shea, S. L., & Dumais, S. T. Stereopsis in Human Infants. Science 207(1980) 323-324.
25. Granrud, C. E., & Yonas, A. Infants' Perception of Pictorially Specified Interposition. Journal of Experimental Child Psychology, in press.
26. Yonas, A., Cleaves, W., Pettersen, L. Development of Sensitivity to Pictorial Depth. Science 200(1978) 77-79.
27. Yonas, A., Granrud, C. E., & Pettersen, L. Infants' Sensitivity to Relative Size Information for Distance. Developmental Psychology, in press.
28. Yonas, A., Pettersen, L., & Granrud, C. E. Infants' Sensi-

tivity to Familiar Size as Information for Distance. Child Development 53(1982) 1285-1290.
29. Gordon, F. R., & Yonas, A. Sensitivity to Binocular Depth Information in Infants. Journal of Experimental Child Psychology 22(1976) 413-422.
30. Hofsten, C. von Binocular Convergence as a Determinant of Reaching Behavior in Infancy. Perception 6(1977) 139-144.
31. Aslin, R. N. Development of Binocular Fixation in Human Infants. Journal of Experimental Child Psychology 23(1977) 133-150.
32. Atkinson, J., & Braddick, O. Stereoscopic Discrimination in Infants. Perception 5(1976) 29-38.
33. Julesz, B. Foundations of Cyclopean Vision (Chicago, University of Chicago Press, 1971).
34. Held, R., Birch, E., & Gwiazda, J. Stereoacuity in Human Infants. Proceedings of the National Academy of Sciences, USA 77(1980) 5572-5574.
35. Granrud, C. E., & Yonas, A. Stereopsis in 5-Month-Old Infants. Unpublished manuscript, 1983.
36. Jones R. K., & Lee, D. N. Why Two Eyes Are Better Than One: The Two Views of Binocular Vision. Journal of Experimental Psychology: Human Perception and Performance 7(1981) 30-40.
37. Walk, R. D., & Dodge, S. H. Visual Depth Perception of a 10-Month-Old Monocular Human Infant. Science 137(1962) 529-530.
38. Walk, R. D. Monocular Compared to Binocular Depth Perception in Human Infants. Science 162(1968) 473-475.
39. Yonas, A., & Pick, H. L., Jr. An Approach to the Study of Infant Space Perception. In L. B. Cohen & P. Salapatek (Eds.), Infant perception: From Sensation to Cognition, Vol. II (New York, Academic Press, 1975).
40. Richter, J. P. The Notebooks of Leonardo da Vinci (New York, Dover, 1970).
41. Helmholtz, H. von A Teatise on Physiological Optics, Vol. 3. J. P. C. Southall (Ed. & Trans.) (New York, Dover, 1962) (originally published, 1925).
42. Gibson, J. J. Perception of the Visual World (Boston, Houghton-Mifflin, 1950).
43. Epstein, W., & Baratz, S. S. Relative Size in Isolation as a Stimulus for Relative Perceived Distance. Journal of Experimental Psychology 67(1964) 507-513.
44. Hochberg, J. E., & McAlister, E. Relative vs. Familiar Size in the Perception of Represented Depth. American Journal of Psychology 68(1955) 294-296.
45. Newman, C. V. Familiar and Relative Size Cues and Surface Texture as Determinants of Relative Distance Judgments. Journal of Experimental Psychology 96(1972) 37-42.
46. Ono, H. Apparent Distance as a Function of Familiar Size.

Journal of Experimental Psychology 79(1969) 109-115.
47. Epstein, W. The Influence of Assumed Size on Apparent Distance. American Journal of Psychology 76(1963) 257-265.
48. Epstein, W. Nonrelational Judgment of Size and Distance. American Journal of Psychology 78(1965) 120-123.
49. Fitzpatrick, V., Pasnak, R., & Tyer, Z. E. The Effect of Familiar Size at Familiar Distances. Perception 11(1982) 85-91.
50. Ittelson, W. H. Size as a Cue to Distance: Static Localization. American Journal of Psychology 64(1951) 54-67.
51. Buffart, H., Leuwenberg, E., & Restle, F. Coding Theory of Visual Pattern Completion. Journal of Experimental Psychology: Human Perception and Performance 7(1981) 241-274.
52. Chapanis, A., & McClearly, R. A. Interposition as a Cue for the Perception of Relative Distance. The Journal of General Psychology 48(1953) 113-132.
53. Dinnerstein, A., & Wertheimer, M. Some Determinants of Phenomenal Overlapping. American Journal of Psychology 70(1957) 21-37.
54. Yonas, A., Goldsmith, L. T., Hallstrom, J. L. The Development of Sensitivity to Information from Cast Shadows in Pictures. Perception 7(1978) 333-342.
55. Ames, A., Jr. Visual Perception and the Rotating Trapezoidal Window. Psychological Monographs 65(1951) Whole no. 324.
56. Kaufmann, R., Maland, J., & Yonas, A. Sensitivity of 5- and 7-month-old Infants to Pictorial Depth Information. Journal of Experimental Child Psychology 32(1981) 162-168.
57. Granrud, C. E., Haake, R. J., & Yonas, A. Infants' Depth Perception: A Familiar Size Effect Induced by Brief Experience. Unpublished manuscript, 1983.
58. Brunswik, E. In Defense of Probabilistic Functionalism: A Reply. Psychological Review 62(1955) 236-242.
59. Hochberg, C. B., & Hochberg, J. Familiar Size and Subception in Perceived Depth. Journal of Psychology 36(1953) 341-345.
60. Fagan, J. F. III Infants' Delayed Recognition Memory and Forgetting. Journal of Experimental Child Psychology 16(1973) 424-450.
61. Epstein, W., & Frankin, S. Some Conditions of the Effect of Relative Size on Perceived Relative Distance. American Journal of Psychology 78(1965) 406-470.
62. Benson, K. A., & Yonas, A. Development of Sensitivity to Static Pictorial Depth Information. Perception and Psychophysics 13(1973) pp. 361-366.
63. Granrud, C. E., Yonas, A., & Opland, E. Infants' Perception of Shape from Shading. Unpublished manuscript, 1983.
64. Johansson, G. On Theories for Visual Space Perception. Scandinavian Journal of Psychology 11(1970) 67-74.
65. Barrow, H. G., & Tenenbaum, J. M. Recovering Intrinsic Scene

Characteristics from Images. In A. R. Hanson & E. M. Riseman (Eds.), Computer Vision Systems (New York, Academic Press, 1978).

66. Horn, B. K. P. Obtaining Shape from Shading Information. In P. H. Winston (Ed.), The Psychology of Computer Vision (New York, McGraw-Hill, 1975).
67. Owsley, C. J. The Role of Motion in Infants' Perception of Solid shape. Perception, in press.
68. Kellman, P. Perception of Three-Dimensional Form in Infancy. Paper presented at the Society for Research in Child Development, Detroit, 1983.
69. Brunswik, E. Perception and the Representative Design of Psychological Experiments (Berkeley, University of California Press, 1956).
70. Marr, D. Vision (San Francisco, W. H. Freeman and Company, 1982).
71. Braunstein, M. L. Depth Perception Through Motion (New York, Academic Press, 1976).
72. Gogel, W. C. An Indirect Method of Measuring Perceived Distance from Familiar Size. Perception and Psychophysics 20(1976) 419-429.
73. Mershon, D. H., Gogel, W. C. Failure of Familiar Size to Determine a Metric for Visually Perceived Distance. Perception and Psychophysics 17(1975) 101-106.
74. Goodman, N. Languages of Art (Indianapolis, Ind., Bobbs-Merrill, 1968).
75. Clowes, M. B. On Seeing Things. Artificial Intelligence 2(1971) 79-112.
76. Huffman, D. A. Impossible Objects as Nonsense Sentences. In R. Meltzer & D. Michie (Eds.), Machine Intelligence, Vol. 6 (Edinburgh, The Edinburgh University Press, 1971).
77. Waltz, D. L. Generating Semantic Descriptions from Drawings of Scenes with Shadows (Tech. Rep. AI-TR-271). Cambridge, Mass., Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1972.
78. McArthur, D. J. Computer Vision and Perceptual Psychology. Psychological Bulletin 92(1982) 283-309.
79. Winston, P. H. Artificial Intelligence (Reading, Mass., Addison-Wesley, 1977).
80. Yonas, A., & Granrud, C. E. Reaching as a Measure of Infants' Spatial Perception. In G. Gottlieb and N. Krasnegor (Eds.), Measurement of Vision and Hearing During the First Year of Life: A Methodological Overview (Norwood, N.J., Ablex, in press).

MEIS

MICROELECTRONIC & INFORMATION SCIENCES CENTER

INSTITUTE OF TECHNOLOGY
UNIVERSITY OF MINNESOTA

227 Lind Hall / 207 Church Street S.E.
Minneapolis, Minnesota 55455
612/376-9122

APPLICATION OF ACOUSTIC MICROSCOPY TO THE EXAMINATION OF INTEGRATED CIRCUITS

Microelectronic and Information Sciences Center

Technical Report #33

William P. Robbins
Department of Electrical Engineering
University of Minnesota

Rolf K. Mueller
Department of Electrical Engineering
University of Minnesota

MEIS technical reports are published in order to provide early access to research results intended for publication. Manuscripts are reviewed and selected for MEIS technical reports by the MEIS Executive Committee.

APPLICATION OF ACOUSTIC MICROSCOPY TO THE EXAMINATION OF INTEGRATED CIRCUITS

MEIS Technical Report #33

W. P. Robbins
R. K. Mueller

The research reported in this paper was supported in part by the Microelectronic and Information Sciences Center.

Annual Report

May 1, 1983 to July 31, 1984

Application of Acoustic Microscopy to the Examination of Integrated Circuits

SRC Contract No. 83-01-024

Prepared By

William P. Robbins and Rolf K. Mueller
Dept. of Electrical Engineering
University of Minnesota
123 Church Street
Minneapolis, MN 55455

I. Introduction

A project to explore the application of acoustic microscopy to the examination of integrated circuits was initiated at the University of Minnesota under SRC sponsorship on May 1, 1983. The first year of the project, including a three month extension requested by SRC, ended on July 31, 1984. This document represents the first annual report for the project.

The interest in acoustic microscopy for the examination of integrated circuits lies in its potential for nondestructively visualizing subsurface or buried features which are not visible optically. Such features include buried metal layers, polycrystalline silicon layers, and buried oxide or nitride layers since these materials have acoustic properties which are significantly different from those of the host silicon lattice. But in addition such features as regions of different doping, localized regions of higher temperature, or regions of high electric field intensity, all within the semiconductor single crystal may also be capable of being imaged since they also involve some change in acoustic parameters compared to the host semiconductor crystal.

A major goal of the project is to obtain high quality acoustic images of the above mentioned features where high quality means not only a qualitatively crisp image but an image from which detailed quantitative information may be extracted. This goal requires the application of several of the imaging modalities available on the scanning laser acoustic microscope (SLAM) including transmission images (shadow projection), holography, surface acoustic wave images, and tomography.

These approaches require high precision measurements and extensive data processing to extract image information from the observed raw data. This in turn mandates a microscope system which is capable of acquiring analog data with high precision and then is capable of digitizing the data, formatting it, and storing it to disk for subsequent processing. Such a system did not exist at the start of this project. Consequently our efforts during this first year have centered on the development of the hardware and software to implement such a microscope.

In the later sections of this report we detail the progress that has been made in developing the microscope system during the past year. However we anticipate that acoustic microscopy is an unfamiliar subject to most readers, so the report begins with an overview of the field and describes the two types of acoustic microscopes currently being studied by the scientific community. We then compare the two types and address the reasons why a scanning laser acoustic microscope is most appropriate for the examination of integrated circuits. At this point we then report on the activities of the past year.

II. Description of Existing Acoustic Microscopes

A. Background

The concept of acoustic microscopy, that is the concept of a device which produces two dimensional mangified images of microscopic structures using sound waves as the probing raditation, was introduced in 1936 by Sokolof. To obtain the required high resolution, he proposed to use sound waves of gigahertz frequency. The technological limitations of his time, however, foiled his efforts. He succeeded to demonstrate the principle of his proposed system using much lower frequencies, sacrificing thereby the high resolution, and ending up with an acoustic camera rather than an acoustic microscope. Two decades later, however, his idea came to fruition. In the 1970's two rather different acoustic microscopy systems were developed independently and became the prototypes of all further development in acoustic microscopy.[1-5] This two systems are the scanning acoustic microscope (SAM) pioneered by C. F. Ouate, and the scanning laser acoustic microscope developed by Kessler and Korpel.

Both systems as their names imply are scanning systems. The acoustic response of an insonified (that is acoustically exited) object is observed locally at a well defined microscopically small spot on the object, and transformed into an electrical signal. The "detector" spot is then moved raster-fashion over an area of interest on the object. With the acoustic response signal keyed to the brightness or color response of a T.V. monitor, a two dimensional "acoustic" image is generated. The manner in which the objects are insonified however and the response detected differ significantly in the two systems.

B. The Scanning Acoustic Microscope.

The basic configuration for the Scanning Acoustic Microscope (SAM) is shown in Fig. 1. Two acoustic lenses are arranged confocally, one bringing sound waves generated by a piezoelectric transmitting transducer to a diffraction limited focal point on a thin object mounted on a transparent object carrier. The second lens recollimates the wave leaving the focal point on the other side of the object. The recollimated beam is intercepted by a piezoelectric receiving transducer. This transducer generates an electrical signal with an amplitude proportional to the local acoustic transmittivity of the object at the point of acoustic exitation. The focal point is then scanned raster fashion over the object to generate an image of the acoustic "transmittivity" of the object. An alternate form of this microscope with interesting applications in metallurgical research uses a single transducer/lens assembly operating as both transmitter and receiver to image the acoustic "reflectivity" of an object.

The heart of the microscope is the acoustical lens. The principles involved in acoustic lens design are the same as those for optical lenses. Focussing is achieved by refraction at a curved interface between two media of differing wave velocity. Despite of this similarity it is much easier to design a diffraction limited acoustic lens for the SAM than a diffraction limited lens for an optical microscope. For one the function of the acoustic lens in a SAM is simply to focus a plane wave incident along its "optic" axis. This avoids all off axis aberrations such as astigmatism, coma, curvature of field etc. In other words the acoustic lens for the SAM can be optimized for on axis performance. In addition acoustic materials offer much higher relative indices of refraction than the best optical materials. This makes it possible to obtain diffraction limited acoustic lenses of large numerical aparture with one spherical

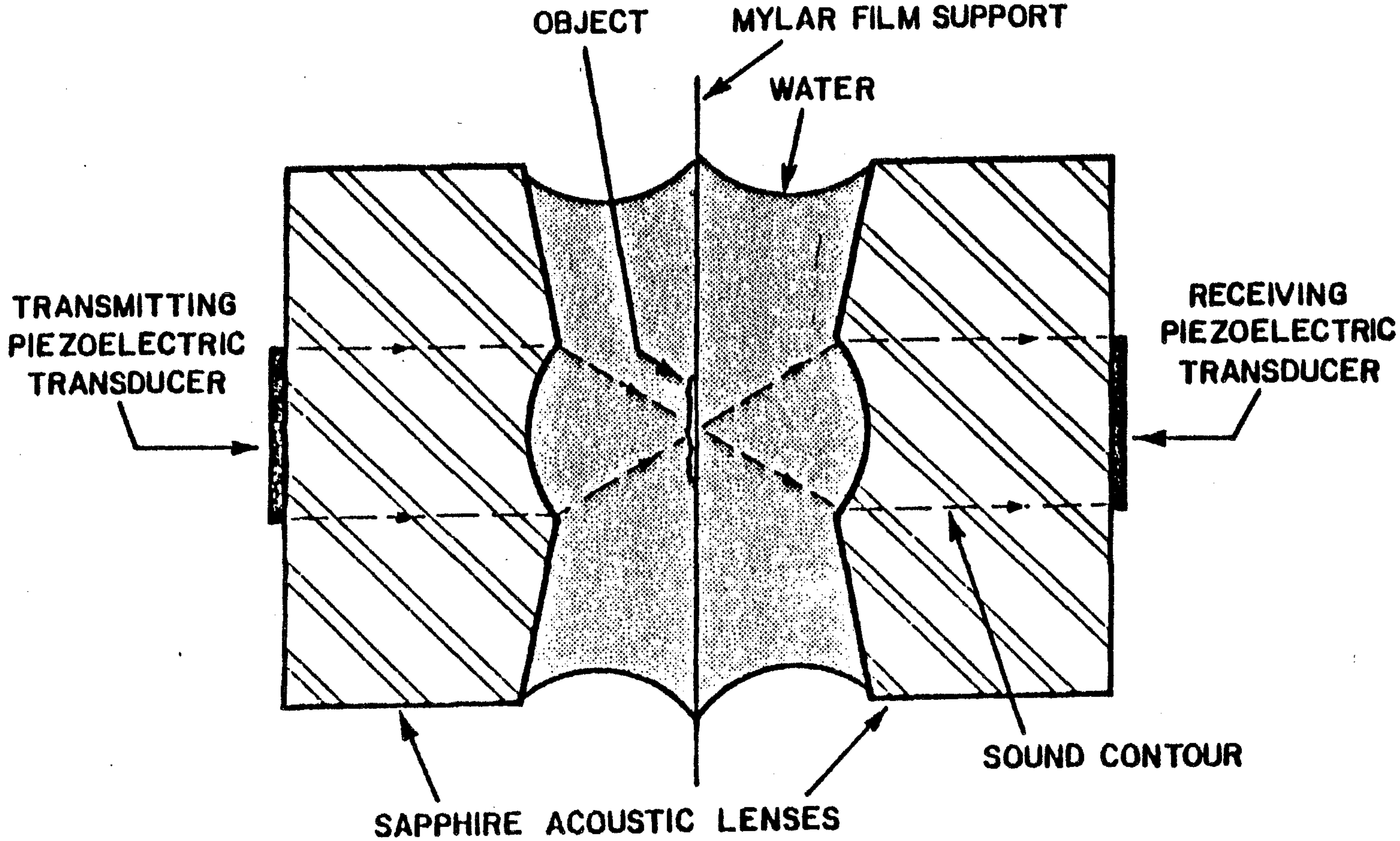


Fig. I Scanning Acoustic Microscope configuration.

interface. This advantage has its cost however. The large relative index of refraction implies large reflection losses at the refracting interface. For a saphire-water interface which constitutes the most commonly used lens in the SAM, the reflection coefficient is in excess of 90%. With two of these interfaces in the microscope, the transmitted sound intensity is less than 1% of the impinging intensity.

This large transmission loss can be overcome however by antireflection coatings. This is particularly important in the SAM, not only to avoid loss of contrast in the resulting imagery but more importantly to avoid loss in resolution, which in the SAM is primarily limited by signal attenuation. To appreciate this we observe that the resolution of the SAM is determined by the size of the focal spot. Since the lenses are diffraction limited, the spot size is directly related to the wavelength of sound in the transmission medium, i.e. in the liquid occupying the space between the two saphire lenses. (See Fig. 1) Practical design consideration impose a lower limit to the width of this liquid cell, and since in most liquids, the attenuation increases quadratically with frequency, the total attenuation in the acoustic pathway dictates the maximum frequency which can be usefully transmitted. This determines the minimum attainable wavelength and consequently the maximum resolution. For water as a coupling fluid and antireflection coating of the lenses, the resolution limit is around 0.5-1 micrometer which is about equal to the resolution limit of an optical microscope. Improvements in resolution are possible in principle through the use of coupling media with lower sound velocity or lower attenuations or both.

C. The Scanning Laser Acoustic Microscope.

The basic configuration of the SLAM is shown in Fig. 2. The SLAM is a transmission microscope. In contrast to the SAM, the entire object is (continuously) insonified. The transmitted sound field is observed over a detection plane which is defined by an optically reflective but (ideally) acoustically transparent layer imbedded within the acoustic transmission medium. As sound waves pass through this optically reflecting plane, it vibrates following the local particle motion in the medium. A focussed laser beam scanning over this mirror plane acts as a scanning acoustic detector. The reflected laser beam picks up the local mirror vibration as a shallow phase modulation. The mirror excursion for typically used acoustic intensities is of the order of .1 Angstrom units, which is about 10^4 times the minimum detectable excursion.

In order to derive an electrical signal proportional to the local acoustic field passing through the detector plane, the phase modulation of the laser beam has to be transformed into amplitude modulation which can then be converted into an electrical signal by a photo detector. The modulation conversion is accomplished by one of two techniques. The knife edge demodulator or the time delay interferometer.(6) The knife edge demodulator is a very simple optical device. It consists of an opaque screen which blocks off one half of the phase modulated beam before it enters the photo diode. It functions analogous to the well known "Schlieren method" which is widely used to obtain optical phase contrast images. The optical simplicity recommends this device but its simplicity brings with it some undesirable features. One is the modulation transfer function (MTF) of the knife edge demodulator shown in Fig. 3. The MTF is defined as the normalized amplitude of the high (acoustic) frequency portion of the photo detector output, which is obtained in response to a wavelike

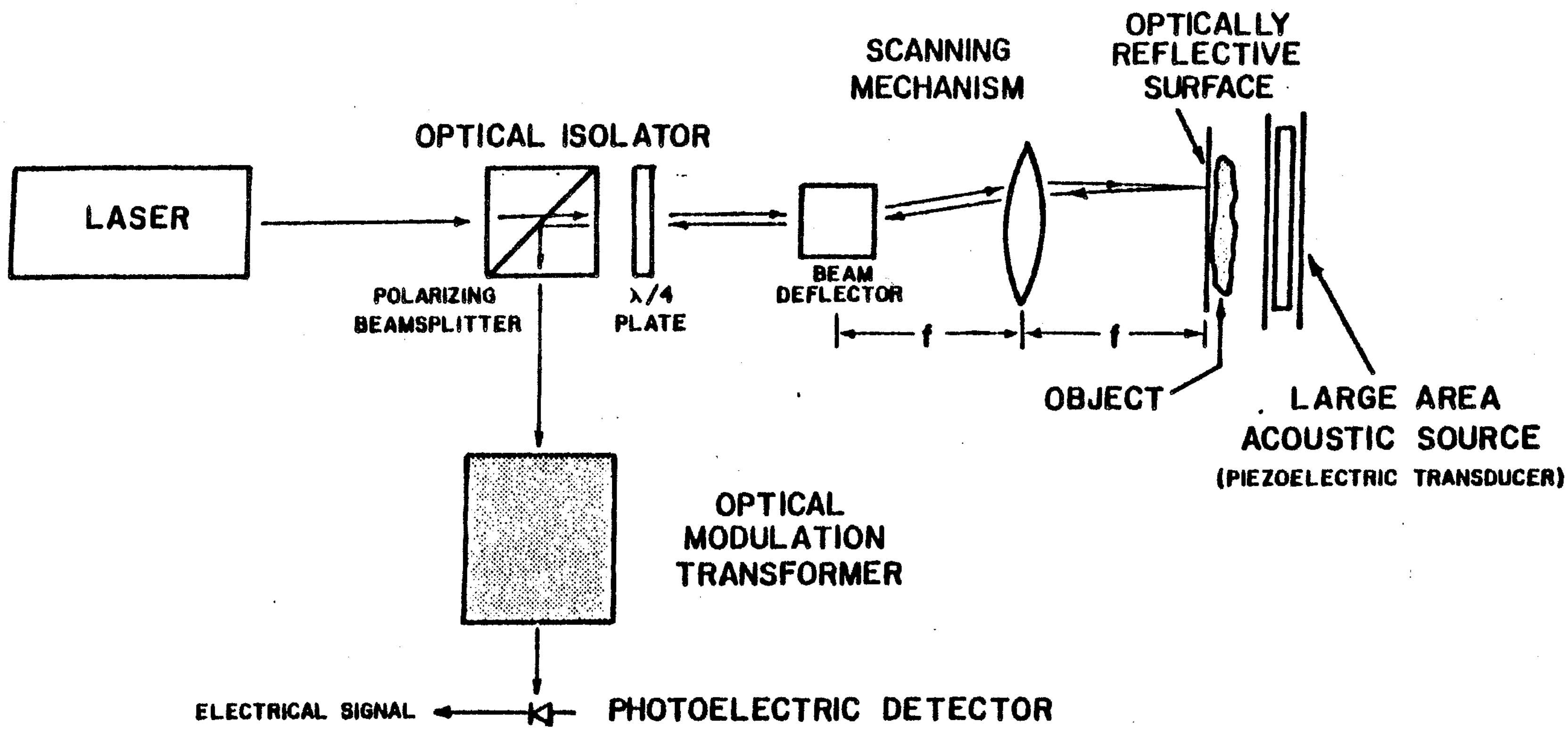


Fig. 2 Scanning Laser Acoustic Microscope configuration.

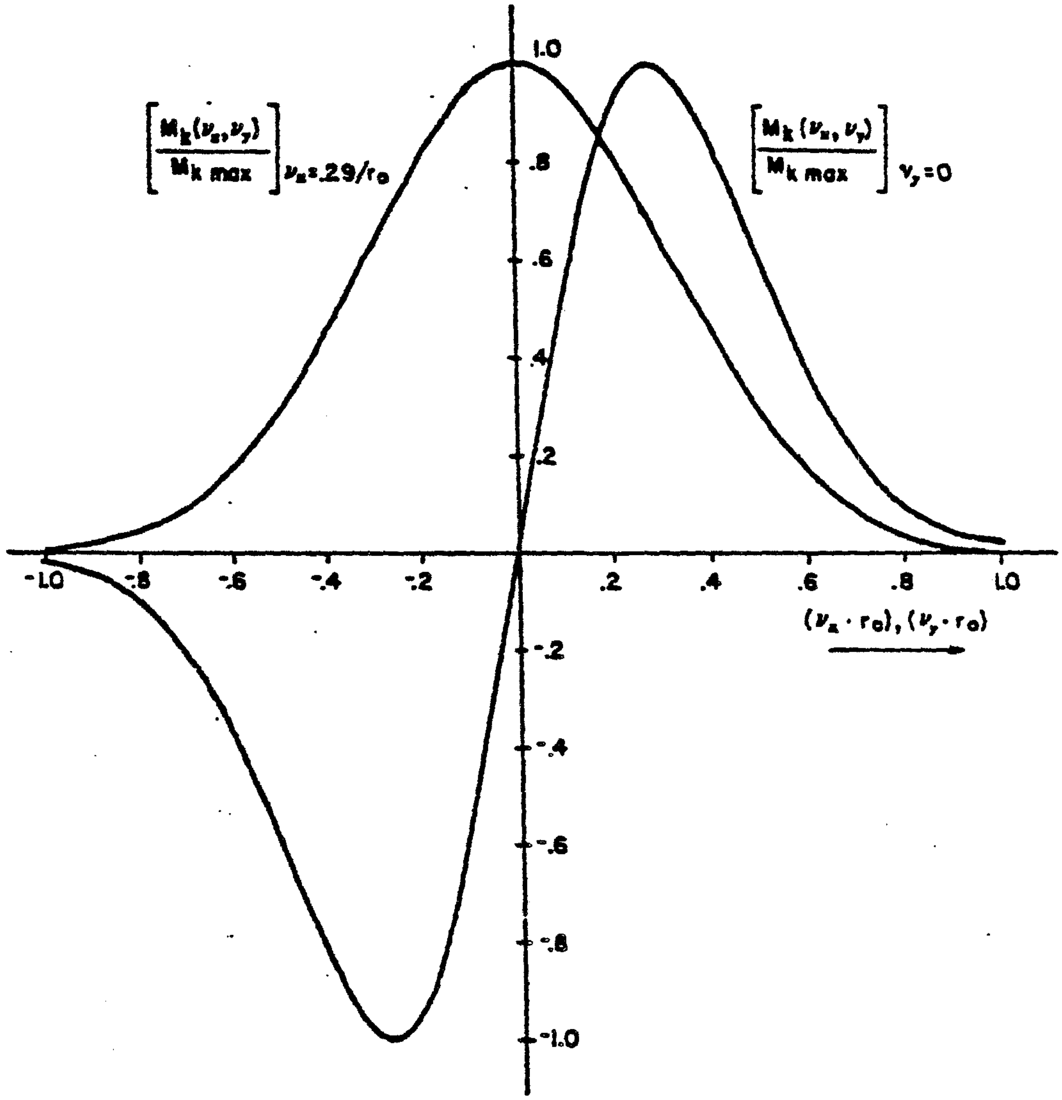


Figure 3. Modulation transfer function of the knife edge detector.

excitation of the mirror surface with constant amplitude but varying propagation vector $(\overline{k})$. The modulation transfer function for the knife edge is severely anisotropic. It is antisymetric for spatial frequency components k_x perpendicular to the knife edge and symetrical for spatial frequency components k_y parallel to the knife edge. Such a modulation transfer function gives rise to severe distortions in the image of the surface excursions if the photo detector output is directly keyed to the brightness of the T.V. monitor. It is possible by post detection filtering to compensate for the distortion introduced by the knife edge MTF but this introduces added complexity into the electronic system as the price to be paid for the simplicity in the optical system, and requires compromises in resolution and sensitivity of the system.

The time delay interferometer described in detail in Ref(6) does not have these difficulties. It has a well behaved isotropic MTF shown in Fig. 4 and about a 12dB higher efficiency in transforming phase modulation of a laser beam into amplitude modulation. With the time delay demodulator the photo detector output is directly proportional to the local mirror excursion, and therefore gives, if displayed on a T.V. monitor, the normal component of the acoustic particle motion in the detector plane of the microscope. The resolution of the SLAM is determined primarily by the scanning laser spot size and only to a lesser extent by the acoustic wave length.

Under ideal conditions the insonifying primary wave is a plane wave, and the thin essentially two dimensional transmission object is in direct contact with the optical reflector, which ideally does not interact with the acoustic field other than following the particle motions of the coupling medium.

In this idealized case the observed field amplitude is proportional to the amplitude transmission function of the object. In other words the conventional

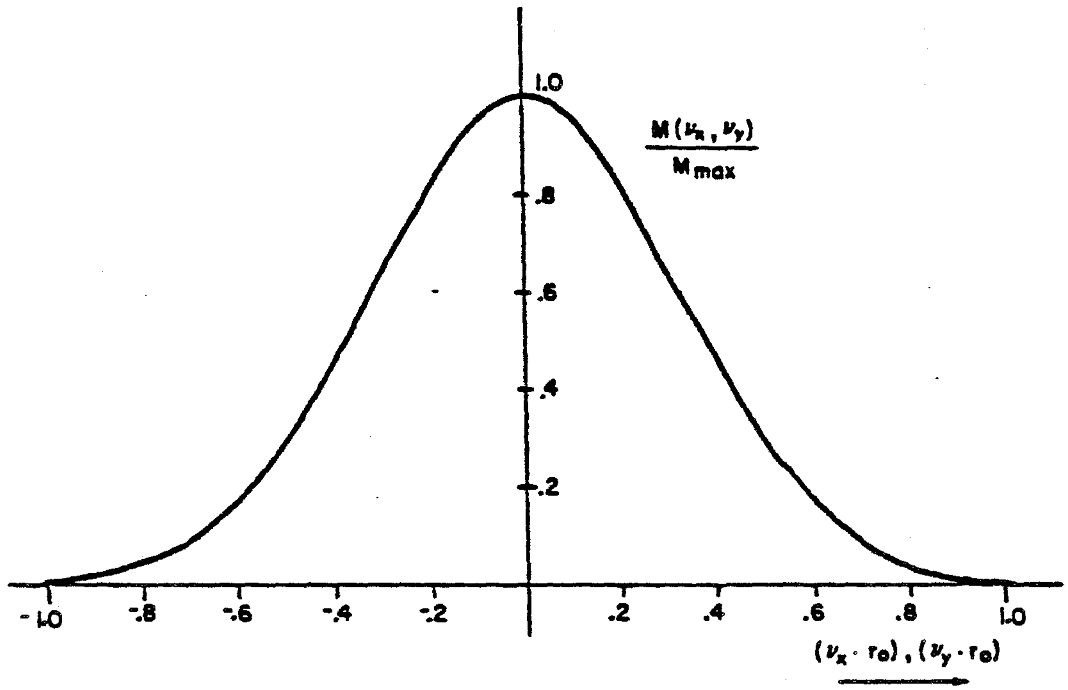


Figure 4. Modulation transfer function of the time delay interferometer.

SLAM generates the acoustic image by proximity focusing or shadow projection. The resolution with which the acoustic field (or the transmission function in this case) can be observed is determined only by the spot size of the probing laser beam. The wave length λ of the sound does not enter the picture because spatial frequencies in the transmission function which are larger than $1/\lambda$ generate evanescent waves which contribute as well as propagating waves to the motion of the mirror plane and are picked up by the scanning laser beam.

In the real world two things happen which spoil this ideal situation. First objects have a finite thickness. This leads to a blurring of the SLAM micrographs since proximity focusing becomes inadequate. As thickness and complexity of the object increases the two dimensional amplitude transmission function, which remains the observed quantity, becomes an increasingly complex functional of widely distributed object features, and direct display of the observed field amplitude at the mirror plane (the shadow projection image) eventually becomes unintelligible. A second problem which results in less than ideal performance in the SLAM is the fact that both thicker objects and physically realizable optically reflecting interfaces support surface waves. These surface waves contribute materially to the observed mirror excursion, again complicating the relationship between object features and observed acoustic exitation.

D. SAM and SLAM Compared.

The foregoing discussions show that the physical quantities observed and displayed in the two microscopes are quite different. In the SAM considering the reflection mode one observes the spatial distribution of the "reflection coefficient" at the water-object interface. Reflection coefficient in the sense used here means the ratio of the power incident on the focal spot at the object through the insonification cone of the transmitting lens to the power scattered back into the acceptance cone of the receiving lens. This reflection coefficient, besides its dependence on the cone angle, depends only on the topological characteristics of the small insonified area on the object and the elastic parameters of a thin surface layer underlying this area. Because of the geometry of the system, energy returning by multiple scattering into the acceptance cone of the receiving lens is negligibly small compared to the first order back scattered (reflected) energy. This implies that object features outside the immediate neighborhood of the insonified spot do not measurably affect the observed reflection coefficient. An object therefore with well defined topological and elastic features gives rise to a very crisp SAM micrograph, comparable in quality to an optical micrograph. This is a very useful feature for qualitative object evaluation. Quantitative interpretation is more difficult to obtain and is widely discussed in current literature.

In the SLAM on the other hand the observed quantity is the normal component of the acoustic "particle velocity" at the observation plane which contains complete information about the acoutic field traversing this plane. The local field amplitude (or intensity) however which is displayed in a conventional SLAM is not, as mentioned before, a good direct representation of topological or elastic object features. This would only be the case for very thin objects on a mirror interface which does not support surface waves. In general the observed field amplitude contains object information in an indirect encoded or transformed manner. To extract this information subsequent data processing is required. The advantage of the SLAM is that with (conceptually) minor modifications one can acquire complete field data in a digital format and use a digital computer to extract the desired object information. In contrast to the SAM which has a built in analog processor (the acoustic lenses), a wide variety of imagining modalities become available with a modified SLAM with digital data acquisition capability. These imagining modalities include proximity focusing, lens simulation, holography, tomography, and surface wave microscopy.

III. Progress in Current Reporting Period

We have during the present reporting period designed and built a modified SLAM with digital data aquisition capability. This system will be described in the following section. We have also developed algorithms for image reconstruction in a surface wave acoustic microscope which we felt should be the first imaging modality to be implemented with the new system. Results of exercising this algorithm on computer simulated data is given. Descriptions of test samples to be used to exercise the complete system, both hardware and software are presented.

A. The Modified SLAM.

In order to acquire a data base suitable for subsequent computer processing we have designed and built a new SLAM system which is equipped with special data acquisition features including a high precision optical and mechanical scanning system, which provides the required highly accurate raster scan. A control subsystem which puts all microscope operations under computer control and an analog and digital data acquisition system which supplies the data base for subsequent processing. A block diagram of the total system is shown in Fig. 5.

B. The Scanning and Control System

The scanning system is an integral part of the acousto-optic subsystem of the microscope which is shown schematically in Fig. 6. The light of a Spectro-Physics argon ion laser, model #164, after passing through an etalon and assorted spatial filters, enters the scanning system proper. The modulated beam which leaves the scanning system is then directed to one of the two available

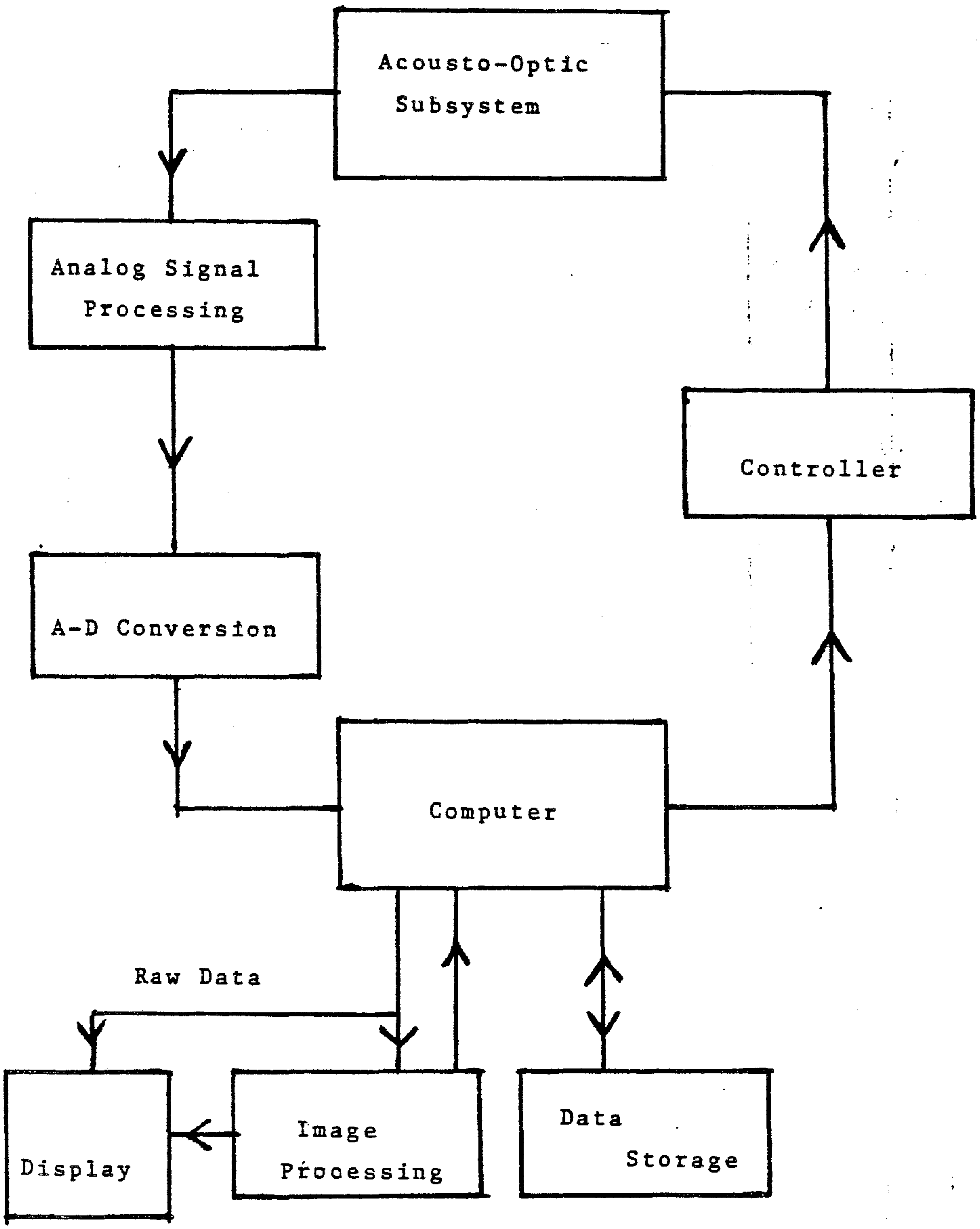


Figure 5. Block diagram of modified SLAM.

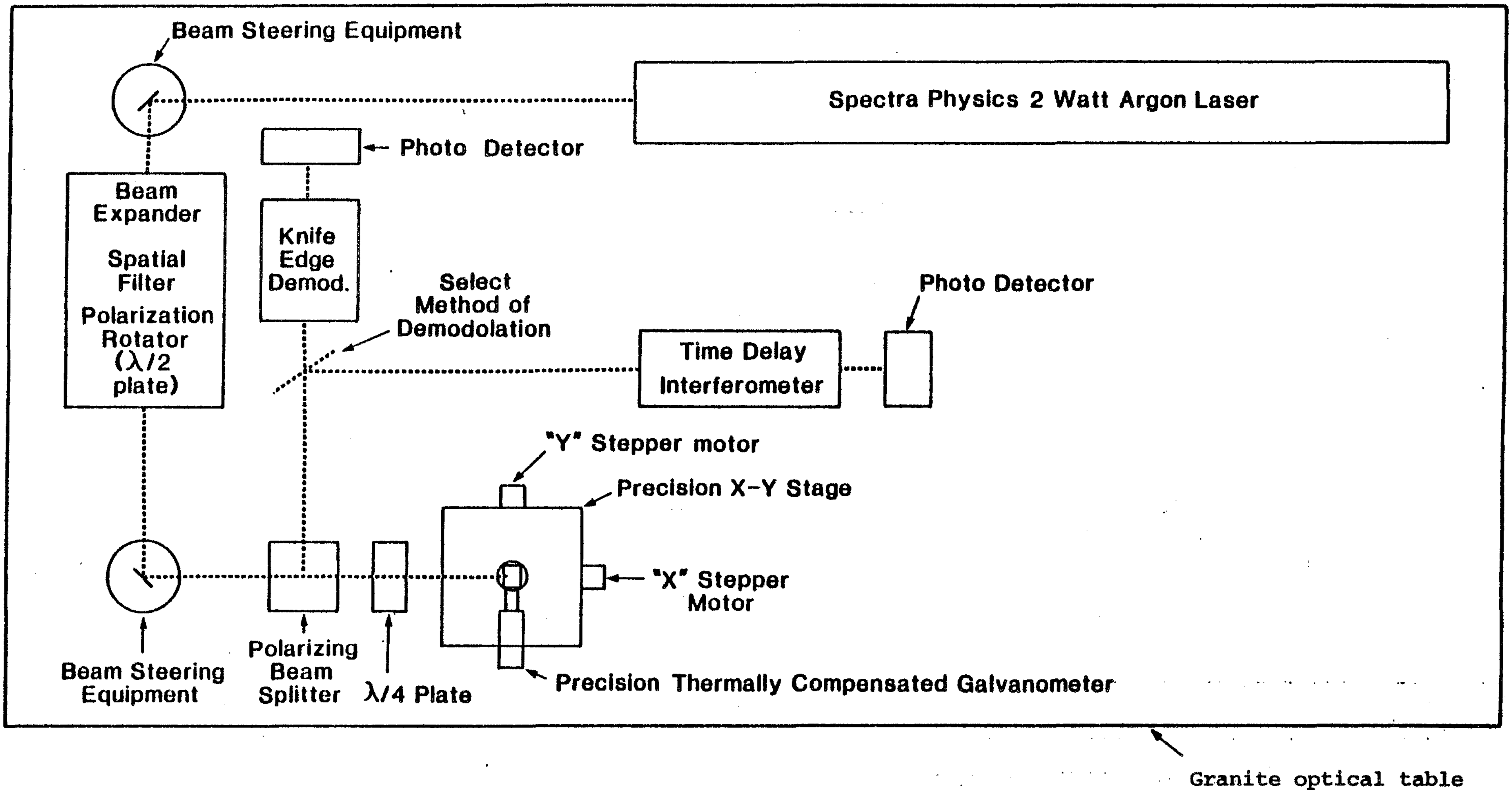


Figure 6. Diagram of acousto-optic sub-system of the modified SLAM.

detector options, the knife edge and the time delay interferometer. The scanning system, proper, consists of a scanning galvanometer and lens combination which scans the laser beam over one line of a raster pattern on the sample (or coverplate) surface. Let the direction of the line scan be the x-direction. A high precision Klinger translator stage, Model UT 100, provides the incremental scanning motion in the y-direction perpendicular to the scan line, by moving the acoustic stage underneath the scanning light beam. The translation stage is driven by a stepper motor. The incremental step size can be chosen as either 1.0 or 0.1 microns. The repeatability of the scan pattern is a fraction of the step size. The relative accuracy is 10^{-4} times the distance traveled by the stage. The translation stage has full x-y capability with maximum excursion of 5 cm. and a maximum slew rate of 2mm/sec. Both galvanometer and translation stages are computer controlled via a Z-80 based SLAM controller shown schematically in Fig. 7.

The Z-80 based SLAM controller (Fig. 7), built here, relieves the main computer, an IBM 9000, of the need to control every motion of the scanning optics and mechanics. The controller accepts macro commands from the 9000 to begin an experiment, pause, resume, adjust scanning speed, aspect ratio, frame location and magnification. A further important function of the controller is to upgrade the galvanometer performance. In order to achieve an accuracy and repeatability comparable to that of the translation stage for the position of the sampling points along the scan line, it is necessary to control and periodically up date the wave form which drives the galvanometer. This is accomplished by allowing for a calibration scan over a test object (located on the sample stage) between imaging scans. The (one dimensional) image of the test fixture acquired in this manner together with its known structure is used to calculate corrections to the

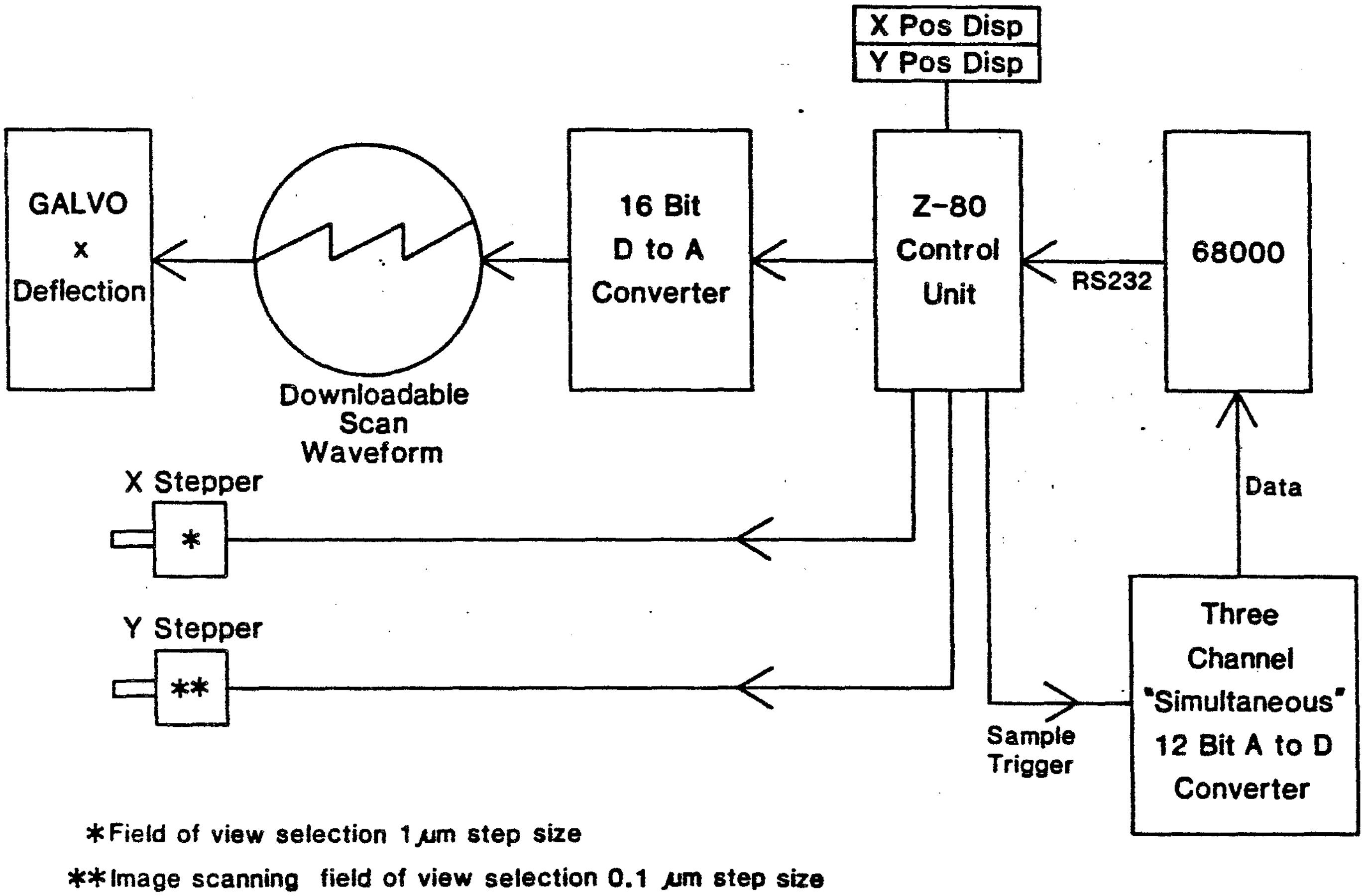


Figure 7. Microprocessor-based (Z-80) SLAM controller.

wave form which drives the galvanometer. These data are then down loaded from the computer to the controller which generates the corrected wave form. This procedure will assure equal accuracy in both scanning directions.

C. Drive and detection electronics

Our SLAM system operates at an acoustic frequency of about 100 MHz. The electronics which drives the acoustic stage and generates the analog signals which are proportional to the real and imaginary parts of the complex acoustic field amplitude at the observation plane is shown schetmatically in Fig. 8. The signals which drive the stage and the detectors are derived from two quartz controlled oscillators LO_1 at 83.5 MHz and LO_2 at 14.04 MHz. These two frequencies are mixed to produce a sum frequency of 97.54 MHz and difference frequency of 69.46 MHz. The sum frequency signal is amplified to drive the acoustic stage, and appears as the high frequency component in the output of the photo receiver, which contains the acoustic image information. Using the difference frequency of 59.96 MHz, the photoreceiver output is heterodyned down to an intermediate frequency (I.F.) of 28.08 MHz, twice the frequency of LO_2. This I.F. is in turn mixed down to base band by a quadrature detector. The dual outputs of the quadrature detector are proportional to the real and imaginary parts of the complex amplitude of the acoustic field at the locus of the scanning laser beam at the sample surface. The proportionality factor, as shown for example in Ref.(6) contains the intensity of the light intercepted by the photo detector. Ideally this light intensity is a constant, however in reality it contains the random laser intensity fluctuations which are, under optimal circumstances, about 5% of average intensity, and the optical reflectivity variations over the sample (or coverplate surface) encountered by the scanning

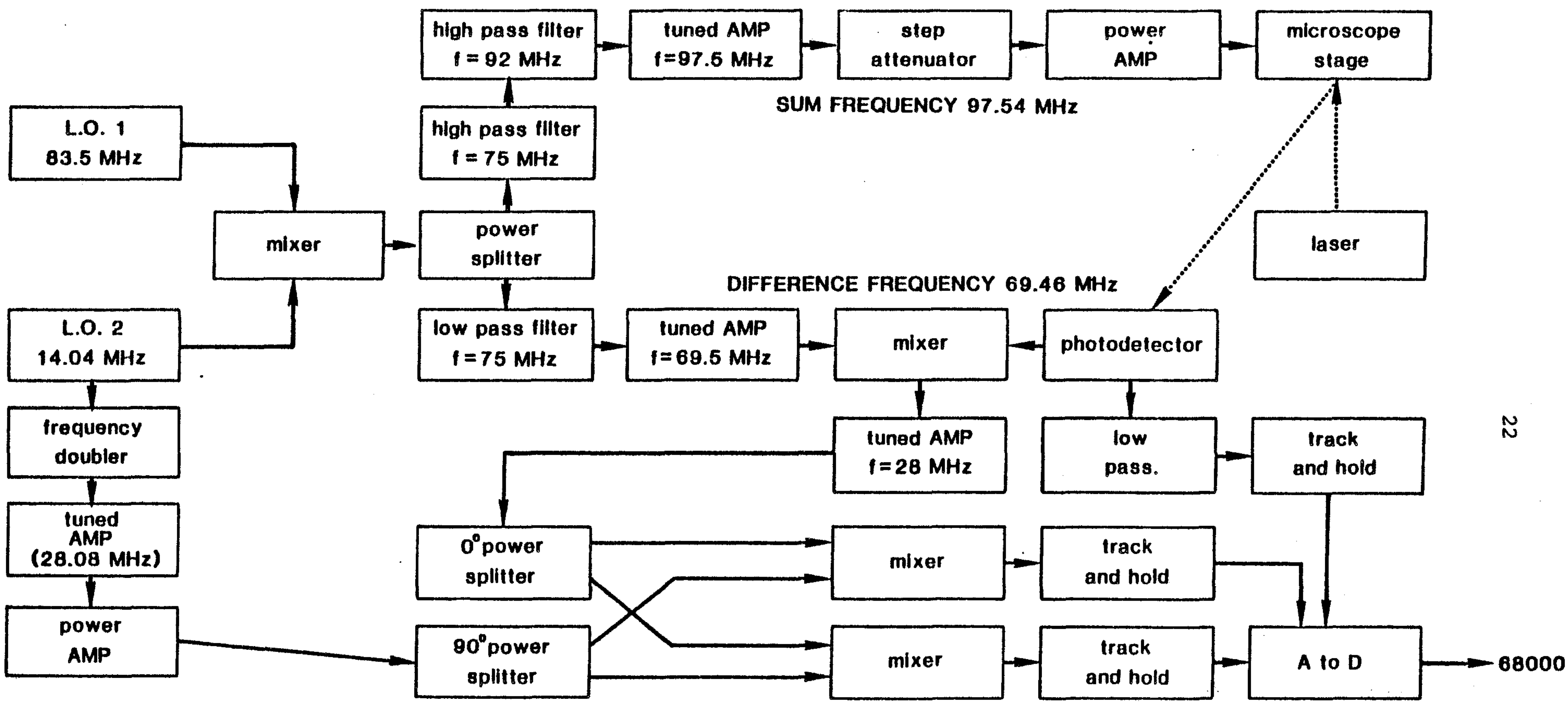


Figure 8. Diagram of the electronics for the acoustic input signal and the signal detection from the photodetector.

laser beam. These fluctuations are the largest single noise source in the image data and severely corrupt the acoustic data base. The "d.c." component of the photo diode output is proportional to this instantaneous light intensity and can be observed simultaneously with the high frequency output of the photodiode. This "d.c." component is sampled along with the acoustic signals for the same pixel and then software is used to divide out the intensity fluctuations so as to achieve the effect of a constant light intensity.

D. Digital Data Aquisition

As in any quadrature detection system, where the input signal is time varying, the two components (real and imaginary) of the detector output must be sampled as closely in time as possible. An added complication in our case is the need to simultaneously acquire a third piece of namely the optical intensity of the modulated laser beam. Failure to achieve effective simultaneity results in phase and magnitude errors in the data.

Effective simultaneity is achieved when the errors caused by the sample and hold time differential are less than other residual error sources. As a bench mark simulation studies have shown that objects with a relative velocity change of 10^{-4} can be imaged in the presence of the expected jitter in the raster-scan. The maximum relative velocity error which is introduced by the sample and hold time differential with the present data acquisition module, is on the order of 10^{-5} and therefore compatible with the scanning system accuracy.

A block diagram of the digital data aquisition module is shown in Fig. 9. Currently, samples are scanned at 16 lines per second (256 samples per line) or one frame of 256 lines every 16 seconds. Since image processing times with our present computer system is two orders of magnitude longer, there is no immediate

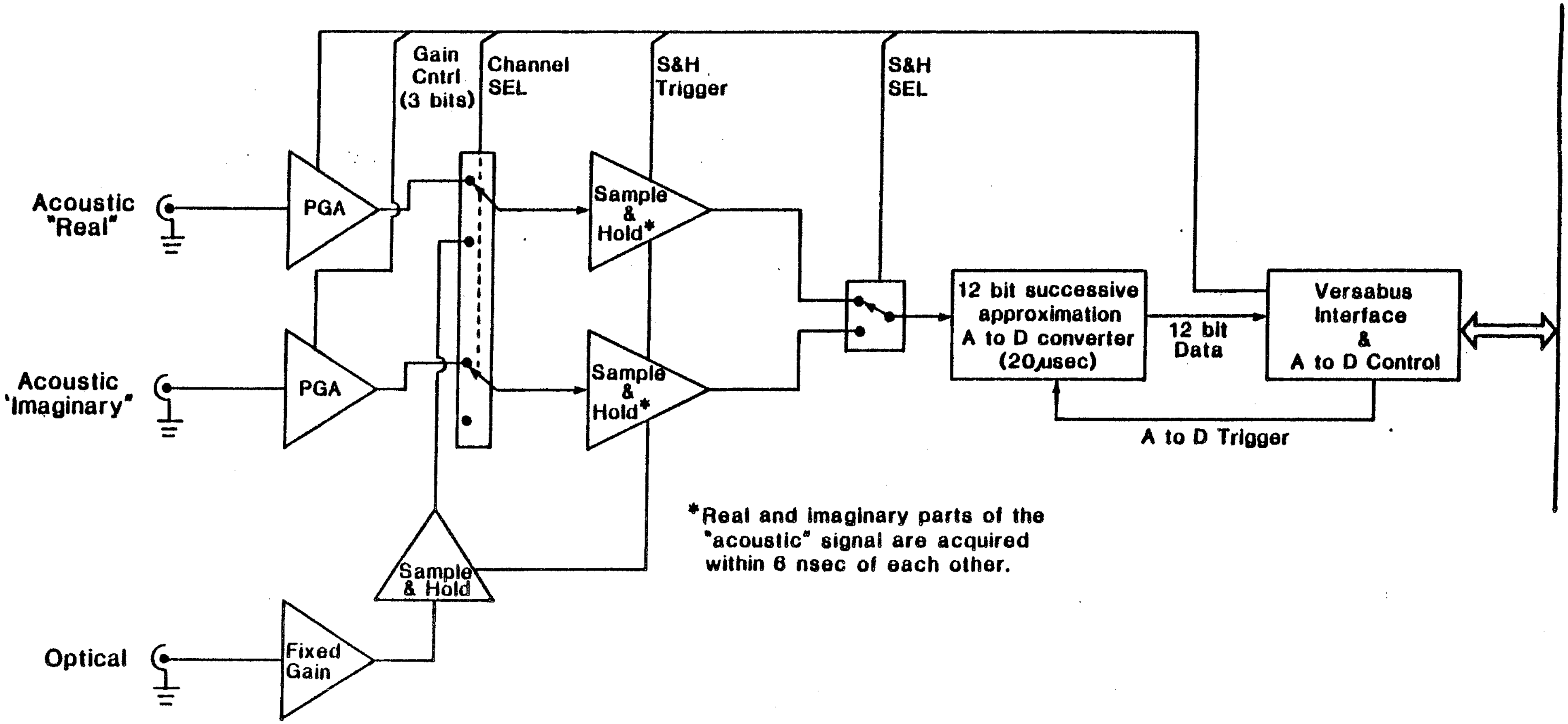


Figure 9. Diagram of the digital data acquisition system.

need to speed the image acquisition. Future use of currently available array processors may, however, reduce processing time to the point where efforts to speed the data acquisition become worthwhile.

E. The Surface Wave Microscope

We have found in previous studies of SLAM image formation (Ref. 6) that surface waves are generated in the mirror support e.g. the coverslip, a nylon support film, or the polished object surface. The occurance of these surface waves degrades the acoustic image in a conventional SLAM where one expects to obtain a proximity focussed image of the acoustic amplitude transfer function of the object. In actual fact however these surface waves carry object information which can be recovered by appropriate image data processing. This processing capability is now available in the modified SLAM discussed earlier. We have carried this thought one step further and introduced the concept of an acoustic surface microscope in which the probing radiation is a surface wave which traverses the object surface. The object surface at the same time represents the optical mirror in the modified SLAM. The desired object information is encoded in the complex amplitude distribution of this surface wave. For sufficiently small inhomogeneities in the object one can model the observed surface excursion $S(x,y)$ as a two dimensional wave propagating in an inhomogeneous material which affects in essence its propagation constant. If we let k_o correspond to the wave number of surface waves in the matrix material we can write for S:

$$(\nabla^2 + k_o^2)S = g(x,y) \cdot S. \qquad (1)$$

where $g(x,y)$ is a function directly related to the distribution of the inhomogenuities in a thin layer at the object surface. The function $g(x,y)$ in this notation is a complex function with the real part related to the local surface wave velocity, and the imaginary part related to the local surface wave attenuation due to absorption and conversion into bulk waves.

The objective of the surface wave microscope is to display the distribution of inhomogeneities in the object by imaging the function g(x,y). This is done by inverting equation (1) which is, as it stands, a prescription to evaluate S(x,y) given g(x,y). We however observe S(x,y) and want to reconstruct g(x,y). In the next section we describe several algorithms which do just that and show some results of exercising these algorithms on computer generated data.

F. Reconstruction Algorithms

Using Eq. (1) for finding g(x,y) essentially means that we must estimate the Laplacian of S. Two methods were examined for doing the estimation, a spatial domain approach and a frequency domain approach. The spatial domain approach uses difference equations as an approximation to differentiation. The frequency domain approach uses discrete Fourier transforms (DFTs).

A simulation of a cylindrical object was used to test and compare the two approaches. Good quality images were obtained with both methods. The spatial domain approach was faster but introduced a constant multiplicative error on g(x,y) which was dependent on the sampling rate. The frequency domain approach introduced Gibbs phenomena due to the use of DFTs. The results presented below are all obtained using the spatial domain method.

Figure (10) shows the simulated wave front for a $g(x,y)/k_0$ (difference in wave number inside the cylinder compared to the wave number outside the cylinder) of 10^{-3}. The reconstructed image using the spatial domain approach is shown in Fig. (11) and is seen to have good quality.

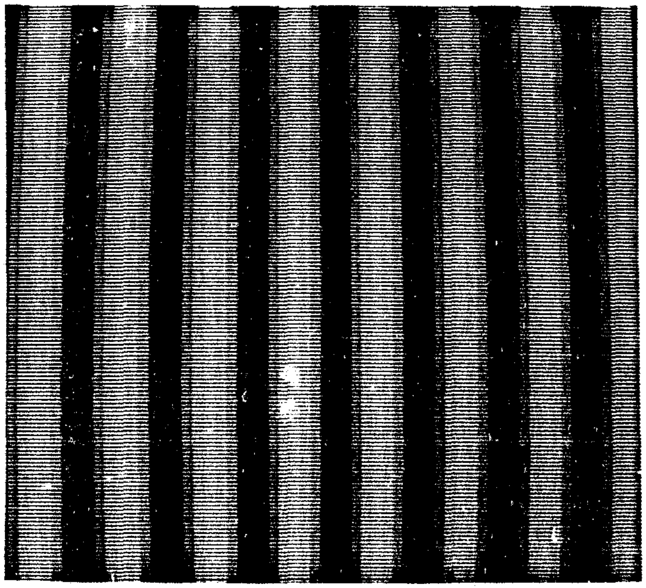


Figure 10 **Simulated acoustic wave field.**

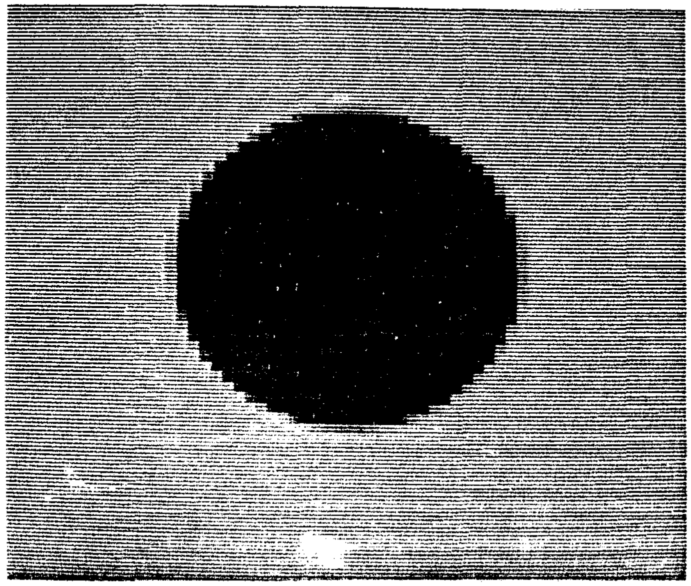


Figure 11. **Reconstructed image.**

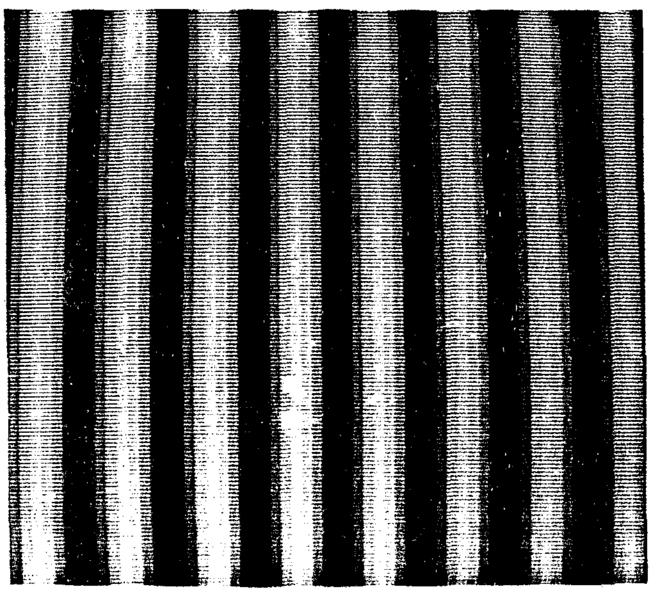


Figure 12. **Simulated acoustic wave field with noise.**

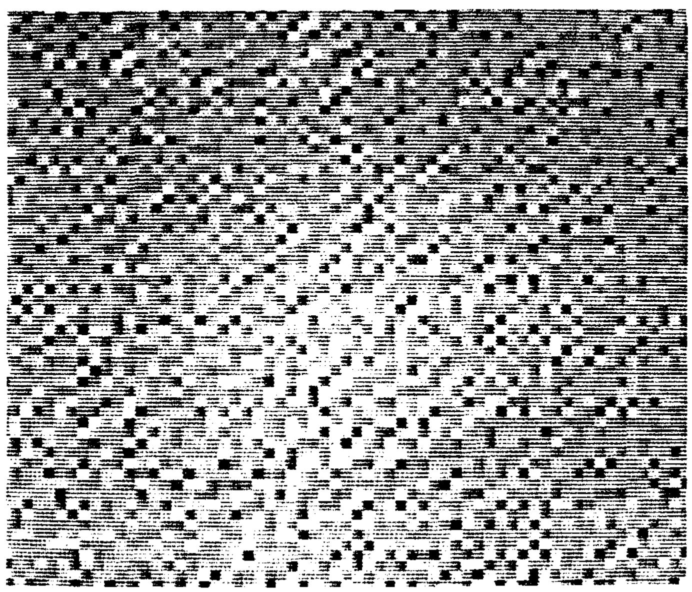


Figure 13. **Reconstructed image using noisy acoustic data.**

Real data will be corrupted by noise, with the best signal to noise ratio expected to be no more than 60 db. Thus to make the simulations more realistic, noise was added to the simulated field before the reconstruction was applied. A gaussian distribution noise with zero mean and a standard deviation of 10^{-3} was used. Fig. (12) shows the simulated wave front with noise added and it looks identical to Fig. (10) as one would expect since the signal to noise ratio is so large. Fig. (13) shows the reconstructed image which looks very noisy.

A noisy reconstruction is to be expected for two reasons. First the noise is about the same size as the expected difference between the wave numbers inside and outside the cylinder. Second, differentiating the measured field to obtain the Laplacian strengthens the high frequency noise by an estimated factor of five. A possible solution for this second difficulty is two-dimensional bandpass filtering. Fig. (14) shows the reconstructed image after such filtering. In this figure the cylinder can be clearly seen and the overall image quality has been improved. A cut through the object (plot of $g(x,1)/k_0$ versus one dimension through the cylinder), which is shown in Fig. (15) with plots before filtering and after filtering, clearly shows the improvement afforded by filtering.

Another environmental effect in the experimental measurements that complicates the reconstruction algorithm is multiple reflection of the wave field. Fig. (16) is the wavefront of a primary field with the addition of three reflections. Fig. (17) is the reconstructed image which shows good quality except at points where the fields have added up to nearly zero. Only at these points is any substantial error introduced.

In summary, a reconstruction algorithm has been developed that produces good quality images using simulated data which includes the effects of noise and

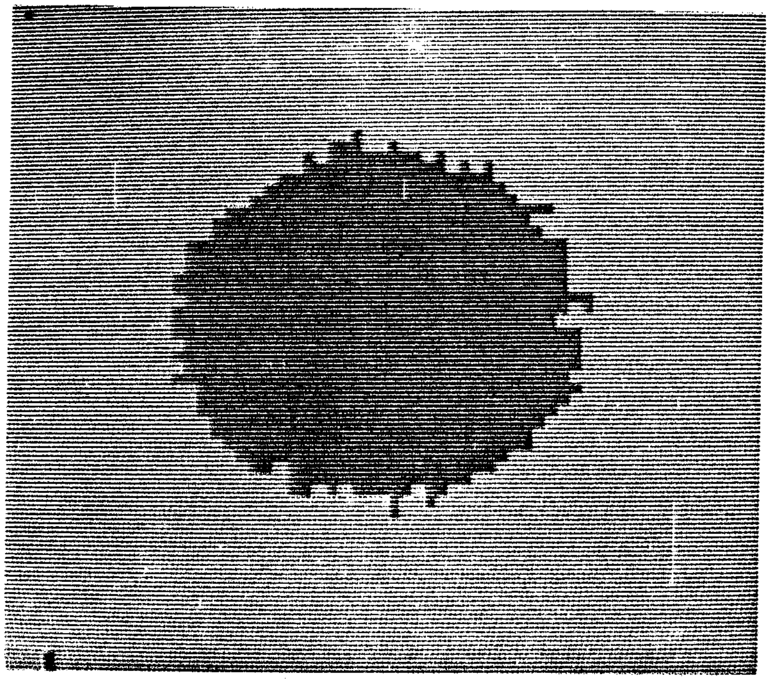


Figure 14. Image reconstructed from noisy acoustic data using bandpass filtering.

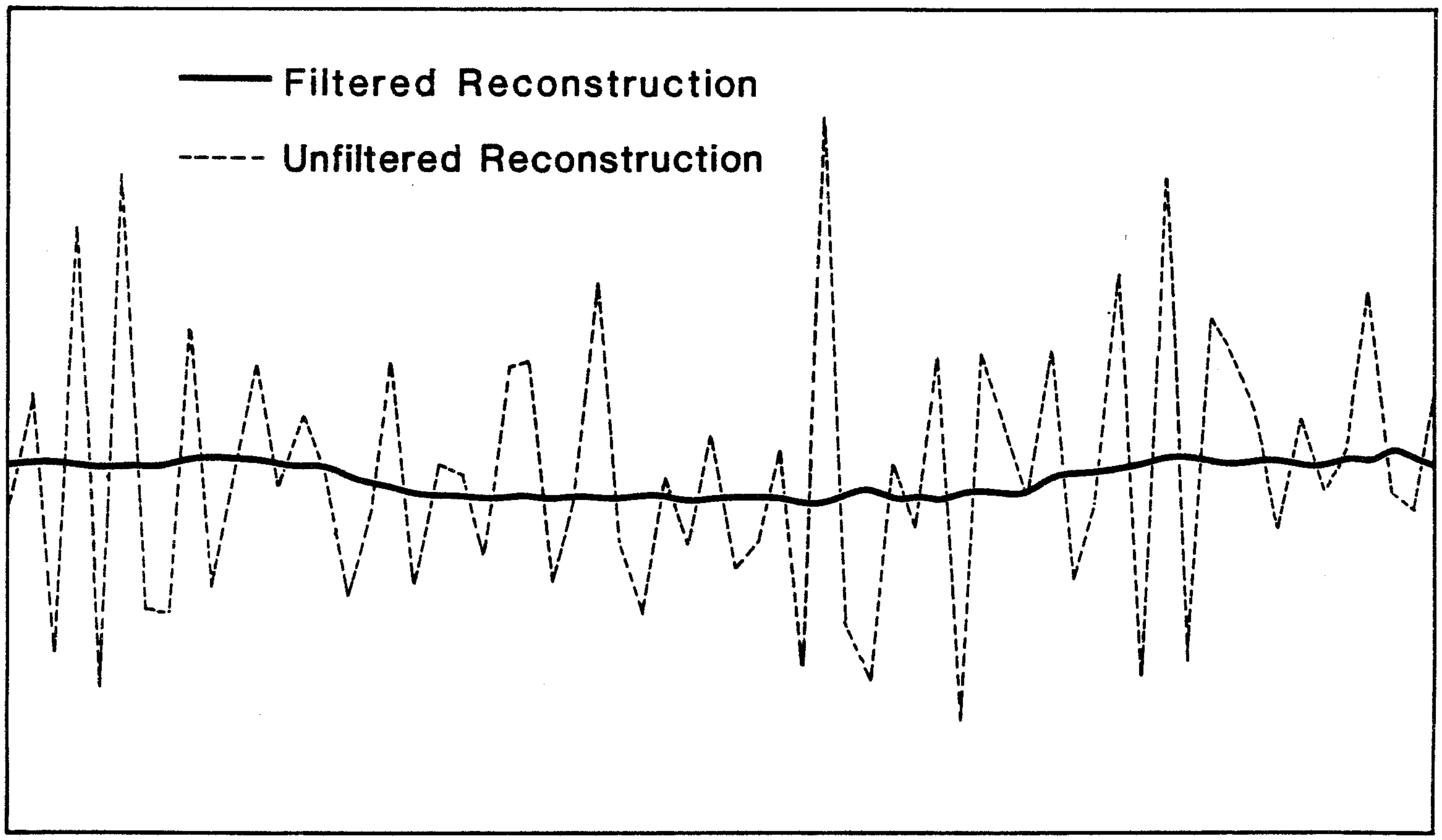


Figure 15. Horizontal cut through image showing wave vector magnitude position.

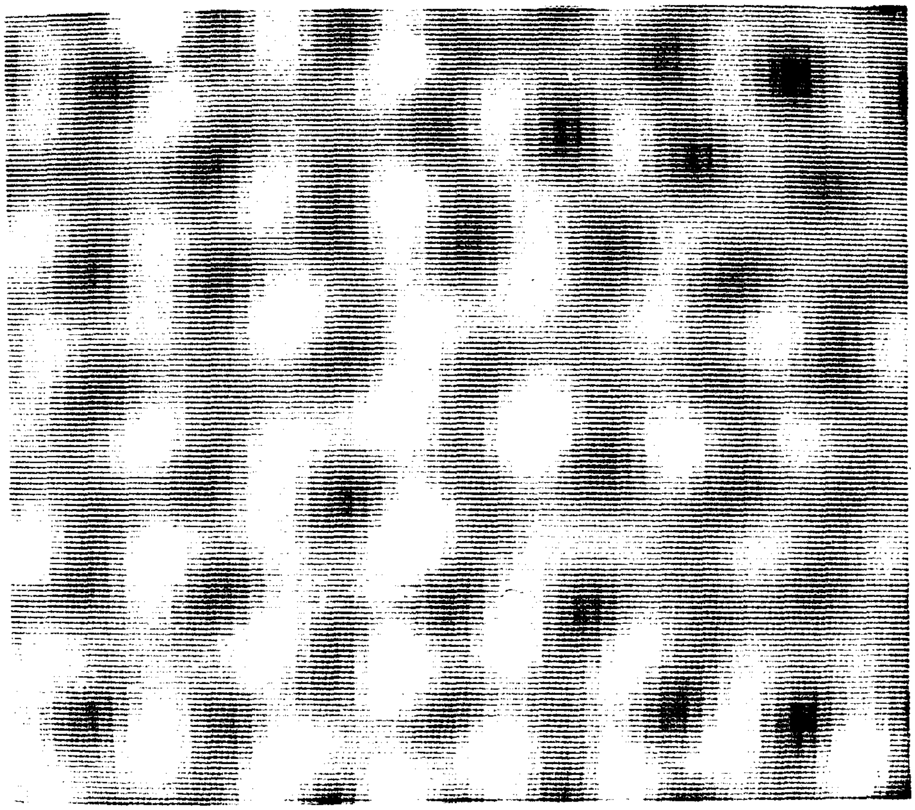


Figure 16. **Simulated acoustic field with multiple reflections.**

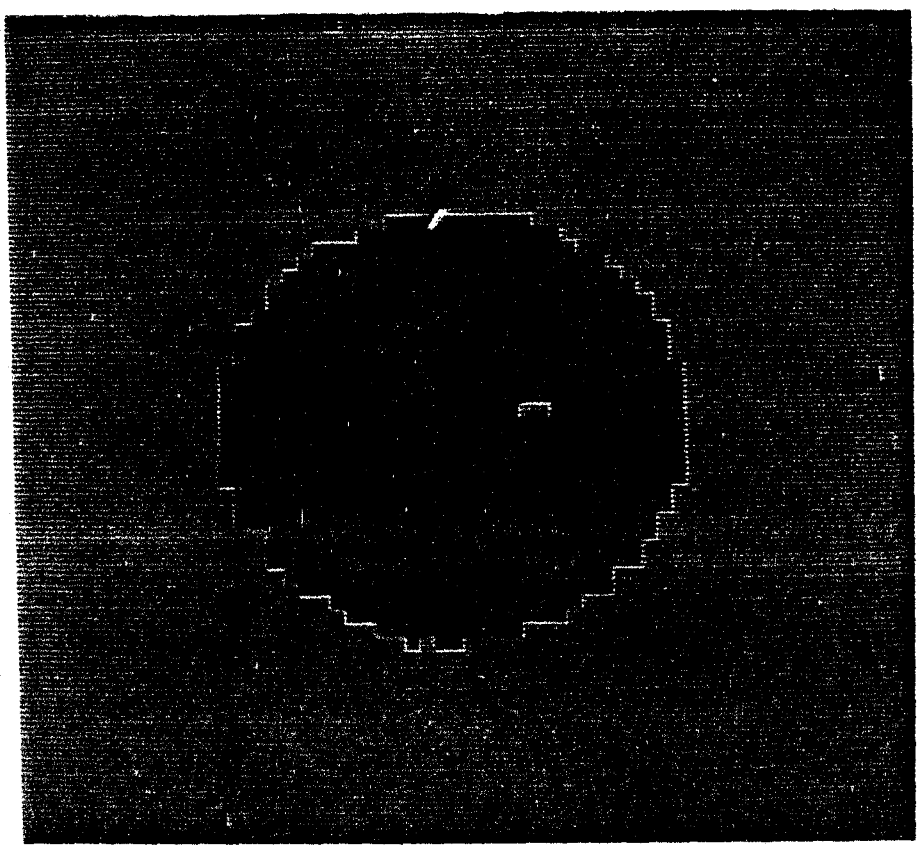


Figure 17. **Image reconstructed from data with multiple reflections.**

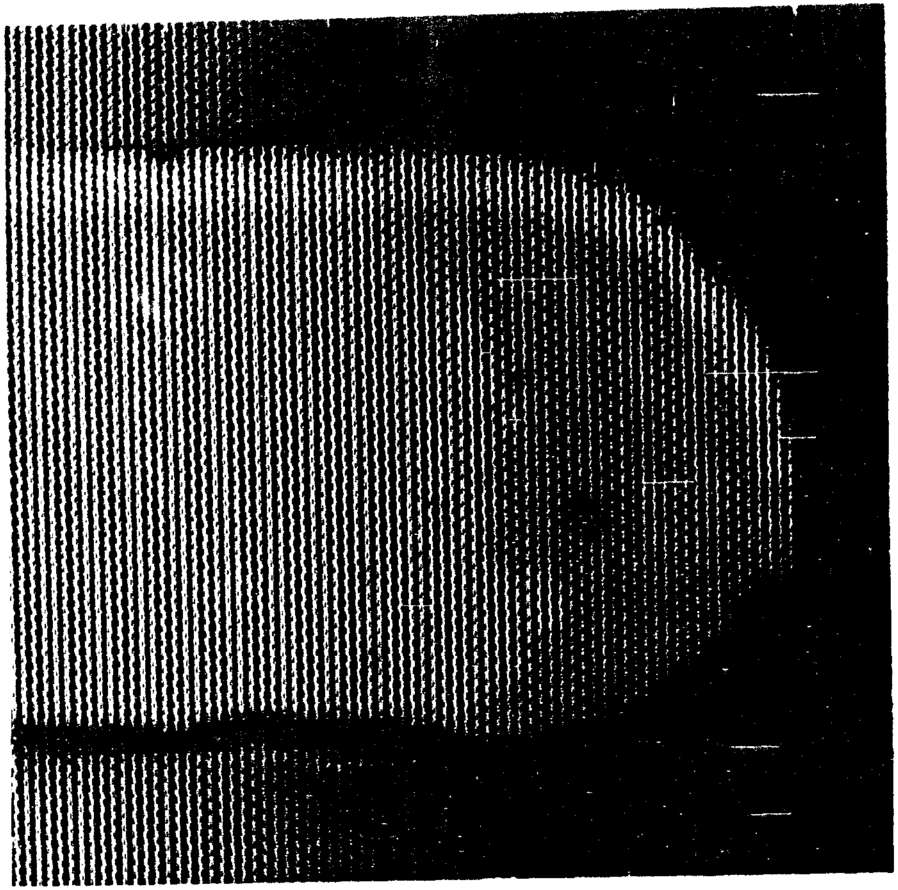
multiple reflections.

G. Test Samples

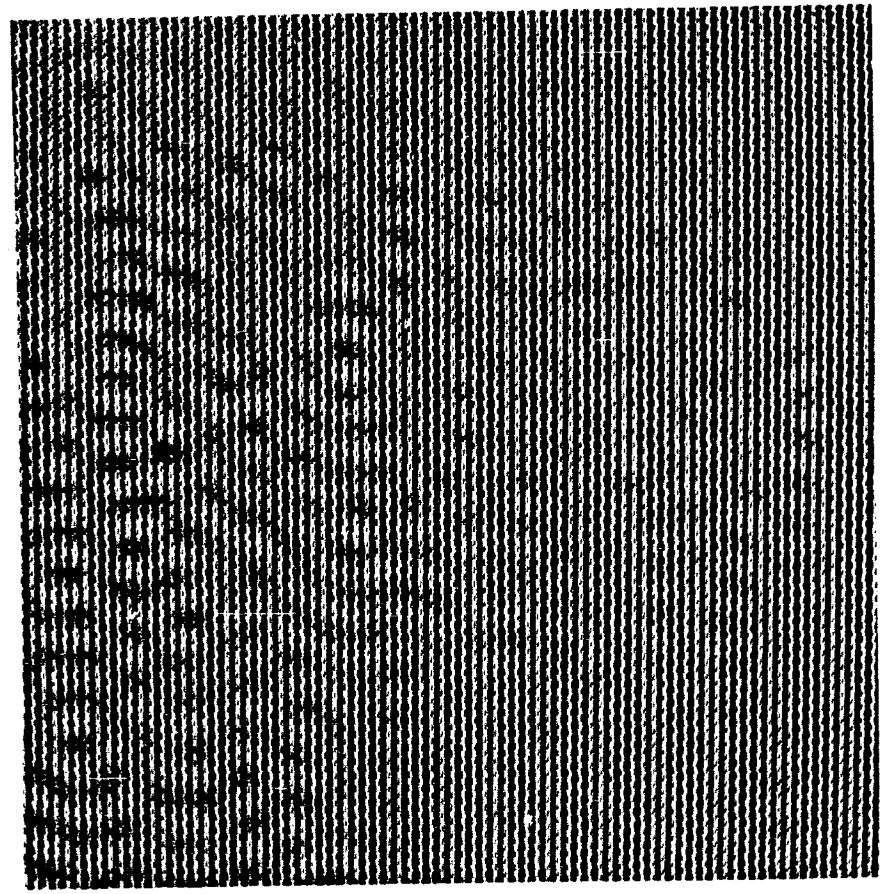
Realistic testing of the microscope requires exercising the hardware and software as a unit on test samples having "known" features. "Known" features means acoustic features not only of known lateral and vertical dimensions but also having known acoustic properties compared to the surrounding medium. Only if both the geometric and elastic properties of the test feature are known, can the operation of the microscope be completely characterized both in terms of spatial resolution and in terms of sensitivity to changes in acoustic properties.

Initial test features have been constructed on YZ lithium niobate surface acoustic wave delay lines for experimental convenience. The features have included simple geometric shapes (disks and long rectangular fingers) made from various thicknesses of thin copper films deposited on the delay line. Copper was chosen because it gives a relatively large acoustic impedance discontinuity (mass loading) per unit thickness of material. Part of the copper features are covered with a thin film of aluminum to eliminate impedance discontinuities due to piezoelectricity. Additional test samples with a greater variety of sizes and impedance changes will be designed and fabricated in the coming year to more fully characterize the microscope.

Initial operation of the microscope with the simple test features described in the preceeding paragraph has already been done. Fig. (18) shows the optical image obtained by the microscope of the copper fingers (about 400 microns in width). The acoustic field distribution of the same region is also shown in the same figure. There is no evidence of any feature in the acoustic field distribution and it also shows considerable amount of multiple reflections of



(a)



(b)

Figure 18. Images obtained from a test structure showing (a) an optical image and (b) acoustic wave field of same area.

the incident acoustic field. The reconstruction algorithms were applied to the raw acoustic data but the results were inconclusive because of incomplete cancellation of optical intensity variations in the acoustic signal which were due to the optical reflections from the test features. Nevertheless the results shown in Fig. (18) illustrate that most of the basic functions of the microscope including digital data acqusition and storage to disk are operational.

In addition, a series of silicon test samples composed of oxide features on a silicon wafer and buried beneath several tens of microns of polysilicon have been obtained from Hewlett-Packard. Dr. Hugh Grinolds, the industrial mentor for this project, was responsible for our obtaining these samples. They will be used in the future to more realistically exercise the microscope in its intended application.

IV. Plans for Second Year

The summary of the first year's activities given in the previous sections show that the diverse efforts, both in hardware and software, are beginning to merge as the microscope begins to operate as a complete system. The activities for the coming year will continue the ongoing efforts, seeing most of them to completion, and will emphasize the operation of the microscope as a complete system. This includes exercising it with test samples in order to begin to assess the ultimate capabilities of the microscope. A major milestone which we hope to reach early in the second year is the reconstruction of an acoustic image using data acquired from a test sample.

Specifically the tasks for the second year will be in three areas including (1) hardware design and fabrication, (2) development of reconstruction algorithms, and (3) evaluation of microscope capabilities. In the first area, most of the hardware activities will consist of completion of efforts already well underway which are concerned with the optics and data acquisition portions of the microscope. Only one new task will be undertaken during the coming year, the design and fabrication of a sample stage for holding samples and insonifying them with bulk waves or surface waves.

The development of reconstruction algorithms for bulk acoustic wave data will begin shortly. This is expected to be more difficult than the surface wave algorithm because of the necessity to account for surface wave generation by mode conversion from incident bulk waves. Additionally we will apply standard image enhancement techniques to the reconstructed images in order to determine if any improvement in image quality can be obtained. Once the major problems with the bulk wave reconstruction are in hand, development will start on the development of reconstruction algorithms for tomographic images.

The evaluation of the microscope's capabilities will be an ongoing effort during the year as appropriate test samples are fabricated and as the reconstruction algorithms are developed. Initial testing will be done with surface acoustic waves using SAW delay test samples described earlier. Non-piezoelectric samples particularly silicon test samples will be used when the sample stage mentioned above is completed. At this point it is hoped that both surface waves and bulk waves can be used. Test samples, both on SAW delay lines and on silicon wafers will be designed and fabricated for use in testing the microscope.

V. Conclusion

The discussion given in the preceeding sections of this report documents that the research is essentially on schedule towards the ultimate objective of project. The first year's efforts were centered on instrument development in order to produce an acoustic microscope which has never been designed or built before, namely a scanning laser acoustic microscope with digital control, data acquisition and storage, and the application of image reconstruction algorithms to raw experimental data. The need for and advantages of such an approach to acoustic microscopy are detailed in the text.

The major hardware tasks are nearing completion and during the coming year the qualitative emphasis of the project will shift towards assessment of microscope capabilities. This will entail operating the microscope with test samples of various kinds and from such trials we expect to pass a major milestone in the near future. That is the reconstruction of an acoustic image using data obtained from an actual sample rather than the simulations shown in this document. From such studies the potential of acoustic microscopy for the examination of integrated circuits should begin to become clarified.

References

1. S. Sokolov, U.S. Patent #2,164,125, 1939.
2. A. Korpel, L. W. Kessler, P. R. Palermo, Acoustic Microscope Operating at 100 MHz Nature, Vol. 232, p. 110, July 9, 1971.
3. R. A. Lemons, C. F. Quate, Acoustic Microscope, Appl. Phys. Letters, Vol. 24, p. 163, 1974.
4. L. W. Kessler, Acoustic Microscopy a review, J. Acoust. Soc. Amer., Vol. 55, p. 909, 1974.
5. C. F. Quate, et al., Acoustic Microscopy-A Review, Proc. IEEE, Vol. 67, p. 1092, 1979.
6. R. K. Mueller, R. L. Rylander, New Demodulation Scheme for Laser Scanned Acoustic Imaging Systems, JOSA, Vol. 69, p. 407, 1979.