Technical Report 628

An Investigation of the Feasibility for Implementing an Advanced Terrain Representation System

John F. Patterson, Dennis M. Buede, and Robert N. Kraft
Decisions and Designs, Inc.

Nancy B. Mitchell
Army Research Institute

Battlefield Information Systems Technical Area
Systems Research Laboratory

U. S. Army
Research Institute for the Behavioral and Social Sciences

March 1984

U. S. ARMY RESEARCH INSTITUTE

FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the

Deputy Chief of Staff for Personnel

Research accomplished under contract to
Department of the Army

Decisions and Designs, Inc.

Technical review by

Randall M. Chambers
Beatrice J. Farr

NOTICES

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report 628	2. GOVT ACCESSION NO.: AD-A148 235
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): AN INVESTIGATION OF THE FEASIBILITY FOR IMPLEMENTING AN ADVANCED TERRAIN REPRESENTATION SYSTEM	5. TYPE OF REPORT & PERIOD COVERED: Technical Report, Period ending October 1982
	6. PERFORMING ORG. REPORT NUMBER: DDI/PR 82-10-331
7. AUTHOR(s): John F. Patterson, Dennis M. Buede, and Robert N. Kraft (Decisions and Designs, Inc.); Nancy B. Mitchell (ARI)	8. CONTRACT OR GRANT NUMBER(s): MDA903-81-C-0568
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Decisions and Designs, Inc., Suite 600, 8400 Westpark Drive, P.O. Box 907, McLean, VA 22101	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 2Q162717A790
11. CONTROLLING OFFICE NAME AND ADDRESS: U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333	12. REPORT DATE: March 1984
	13. NUMBER OF PAGES: 86
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): --	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: --

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

--

18. SUPPLEMENTARY NOTES

Contracting officer's representative was Nancy B. Mitchell. A related report is ARI Research Note 84-68, Psychological Research on Advanced Terrain Representation: Formatting the Visual Material.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Military training	Videodiscs	Surrogate travel
Training devices	Terrain	Computer-generated imagery
Simulation	Tanks (combat vehicles)	Terrain boards

use single quotes

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report describes a feasibility study for developing a videodisc-based computer system for training small unit armor tactics. This system will generalize the techniques of "surrogate travel" ~~in order~~ to provide an advanced terrain representation (ATR) system permitting free travel within a simulated environment. The purpose of ATR is to provide a ground-level perspective of a simulated battle. Using pictures taken from many points and perspectives on the terrain, the system displays the appropriate pictures in a fashion simulating the movement of the user's tank. (Continued)

DD FORM 1473, 1 JAN 73 — EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

 SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ARI Technical Report TR 628
20. (Continued)

The results of this investigation indicate that present technology is inadequate to allow for the use of computer-generated imagery (CGI) overlays if the visual base is photographs of real terrain; that a terrain board may be appropriate to CGI use; and that a computer-generated image base, while lacking some other desirable properties, provides the greatest certainty that overlays can be made available. Originator-supplied keywords include: Computer-generated imagery, and Training devices.

Technical Report 628

An Investigation of the Feasibility for Implementing an Advanced Terrain Representation System

John F. Patterson, Dennis M. Buede, and Robert N. Kraft
Decisions and Designs, Inc.

Nancy B. Mitchell
Army Research Institute

Submitted by
Franklin L. Moses, Chief
Battlefield Information Systems Technical Area

Approved as technically adequate
and submitted for publication by
Jerrold M. Levine, Director
Systems Research Laboratory

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
5001 Eisenhower Avenue, Alexandria, Virginia 22333

Office, Deputy Chief of Staff for Personnel
Department of the Army

March 1984

Army Project Number
2Q162717A790

Human Performance
Effectiveness and Simulation

ARI Research Reports and Technical Reports are intended for sponsors of R&D tasks and for other research and military agencies. Any findings ready for implementation at the time of publication are presented in the last part of the Brief. Upon completion of a major phase of the task, formal recommendations for official action normally are conveyed to appropriate military agencies by briefing or Disposition Form.

FOREWORD

The Battlefield Information Systems Technical Area of the Army Research Institute (ARI) has supported the Advanced Terrain Representation (ATR) systems development research presented in this report. The research on the feasibility of implementing the ATR has been done in the interest of producing simulated terrain travel which can be used to enhance both product development and training where navigation over terrain is required.

The information presented in this report provides a basis for evaluating the efficacy of producing a terrain travel simulation and offers, in addition, prospects for future considerations as technological development proceeds. Based on this guideline, ARI has proceeded with the development of a terrain travel simulation.

Edgar M Johnson

EDGAR M. JOHNSON
Technical Director

AN INVESTIGATION OF THE FEASIBILITY FOR IMPLEMENTING AN ADVANCED TERRAIN REPRESENTATION SYSTEM

EXECUTIVE SUMMARY

Background

Battle simulations are important training techniques which have been developed largely for the purpose of tactical and leadership training. Several types of battle simulations, such as CAMMS, Dunn Kemp, and BATTLE, are currently used for training at several levels of command and for somewhat different purposes, but all battle simulations have one or more of the following characteristics in common.

- The simulation is manpower intensive, requiring a well-trained controller who derives minimum benefit from the exercise.
- The information sent by the controller is less than fully realistic because the entire playing surface is not in full view.
- The simulation cannot be played in real-time.
- The most significant missing element is tactical surprise.

This report describes the technical considerations for the development of a battle simulation which would provide a ground-level view, be played in real-time, incorporate tactical surprise, and eliminate the need for a controller. This prospective simulation is based on a generalization of the "surrogate travel" technology which, by definition, produces an interactive system utilizing a videodisc, a microcomputer, and a CRT. In its current form, it allows the user to control his movement through a display on a CRT in four directions: forward, backward, right, and left. The generalization proposed for the present system would provide an opportunity to travel freely, in more than four directions, over open terrain, in a manner similar to the movement of a tank. The purpose of this particular simulation system is the training of small armored units in tactics, leadership, and land navigation. This report contains the results of research which was done by Decisions and Designs, Inc. (DDI), in cooperation with the Army Research Institute (ARI), to determine the feasibility of implementing the simulation now called an Advanced Terrain Representation (ATR).

Method

The production of a videodisc, which is the base for the visual display shown on the CRT, requires pictures that are taken from many grid centers and from many directions of view. These pictures are then displayed in a fashion simulating the movement of the user's tank.

If the user is to have any appropriate ground-level view of the terrain and sufficient interactivity and realism to permit him to maneuver his vehicle in the direction called for by the situation represented on the display, then specific objectives must be met. They are as follows:

- a simulated area on the order of 2km by 4km;
- a sense of continuity and travel coherence must be maintained as the user moves from grid center to grid center;
- the feeling that one is free to travel in virtually any direction;
- sufficient detail and texture to support visually guided behavior; and
- the ability to overlay dynamically changing symbols, such as those of friendly or opposing forces.

In order to achieve these goals, two closely related but distinct information sources are needed, image data and intervisibility data. The image data are generated by the procedures used to create the ground-level pictures for each grid, then placed on the videodisc, and later presented to the user. Intervisibility data refer to information which is extracted, stored, and used to control the overlay of dynamic imagery. The intervisibility data must include information concerning the extent to which objects within the full field of view will occlude or obscure an object which is overlayed.

Results

The results of this study demonstrate that the development of an ATR system is feasible and can meet these goals. The major conclusions are:

1. The tradeoff between increasing terrain coverage and decreasing the minimum travel speed is critical. Increases in coverage require larger distances between the grid centers from which the terrain pictures are taken. However, decreases in the minimum travel speed can only be achieved by decreasing these distances

between grid centers. The minimum time between the presentation of images to the user is also a factor in this tradeoff.

2. In order to achieve an acceptable minimum presentation time, the grid pictures must be allocated to multiple videodiscs in square block patterns. That is, if four videodiscs were used, two videodiscs would store pictures from alternate grids on the odd rows of the grid pattern and the other two videodiscs would cover the even rows in the same fashion. Our recommendation is to use nine videodiscs in a three by three block pattern.

3. A panoramic (360^{o}) image provides a continuous pivoting (scanning) capability. However it is not achievable at this time if terrain boards are used as the source material for the imagery. In addition, a frame buffer or special effects generator must be included in the system to take advantage of a panoramic image. Therefore, sixteen or twenty-four equally spaced, discreet images should be photographed from each grid and used to simulate pivots. The images should be laid out sequentially on the videodiscs to enable the "step" mode of the videodisc player to be utilized.

4. The overlay of dynamic symbols (e.g., tanks) on the terrain image requires a range image in registration with the terrain image. The only alternative at this time for creating such a range image is to build an object model of the terrain. The ATR system requirement to achieve this symbolic overlay is a frame buffer with four frames; one each for the terrain image, range image, symbol(s), and composite image.

Three major uncertainties remain. They can be resolved only by implementing a prototype system and through experimentation:

1. The responsiveness of the ATR system design to the user's open field travel demands,

2. The ability of terrain that is computer-generated to provide the needed stimuli for simulating travel, and

3. The accuracy with which dynamic symbols can be overlayed on noncomputer-generated terrain.

AN INVESTIGATION OF THE FEASIBILITY FOR IMPLEMENTING AN ADVANCED TERRAIN REPRESENTATION SYSTEM

CONTENTS

PRECEDING PAGE BLANK

CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

AN INVESTIGATION OF THE FEASIBILITY FOR IMPLEMENTING AN ADVANCED TERRAIN REPRESENTATION SYSTEM

1.0 INTRODUCTION

One of the major goals of the research presented in this report is the development of a low cost alternative to field training exercises and engagement simulations. The recent development of videodisc-based surrogate travel provides the technological basis for the development of such a battle simulation.

Surrogate travel uses photographs that have been stored on a videodisc to provide a realistic simulation of ground-level travel. Ground-level scenes are photographed at ten-foot intervals along the roads of a town, and then replayed at the request of the user. The scenes can be sped up, reversed, and stopped. Of even greater significance, choice points (such as right and left turns onto side streets) can be selected, thereby providing a freedom of travel that is unavailable on film or videotape. Although surrogate travel was developed to represent travel through urban streets, the concept can be generalized to represent travel through open terrain. By using this technique of terrain representation, a battle simulation, which presents the player with a ground perspective on the battle, can be implemented. The investigation of the feasibility of developing an Advanced Terrain Representation (ATR) is the subject of this report.

1.1 <u>The General Concept of an Advanced Terrain Representation</u>

The concept of an ATR is the extension of the surrogate travel technology to free travel over open terrain. Since a tracked vehicle is not limited to travel over roads, the visual representation of off-road travel is essential to a battle simulation aimed at training the tactics of combined arms. The addition of computer-generated dynamic overlays representing friendly and opposing forces would create a realistic training technique.

Figure 1-1 represents one means of segmenting terrain for image capture. Photographs taken from each of these grid centers are represented by the walls surrounding the grid. When these photographs are stored on a videodisc, they can then be displayed on a CRT, and visual travel is produced by moving from grid center to grid center. The variable presentation capability using the random access mode of the videodisc player (as opposed to the deterministic, sequential presentation of film) allows the user to control ıe direction of travel; by doing so, the user can respond to a situation by maneuvering in the appropriate direction. The addition of dynamic overlays can create the tactical situation to which users can respond. The objective of this project is the utilization of this approach to provide a realistic battle simulation for the purpose of training small unit armor tactics.

1.2 Report Organization

The four sections of this report include the project goals, this section; a summary of the supporting psychological research (Kraft, Patterson, and Mitchell, 1982); a discussion of pivot construction, detailed imagery, and dynamic overlays, Section 3.0; and Conclusions and Recommendations, Section 4.0.

Goals one through four of the project (large coverage, coherent travel, travel in any direction, and 360° pivots) are chiefly governed by the procedure for acquiring the image data and placing it on the videodisc. Travel coherence, travel

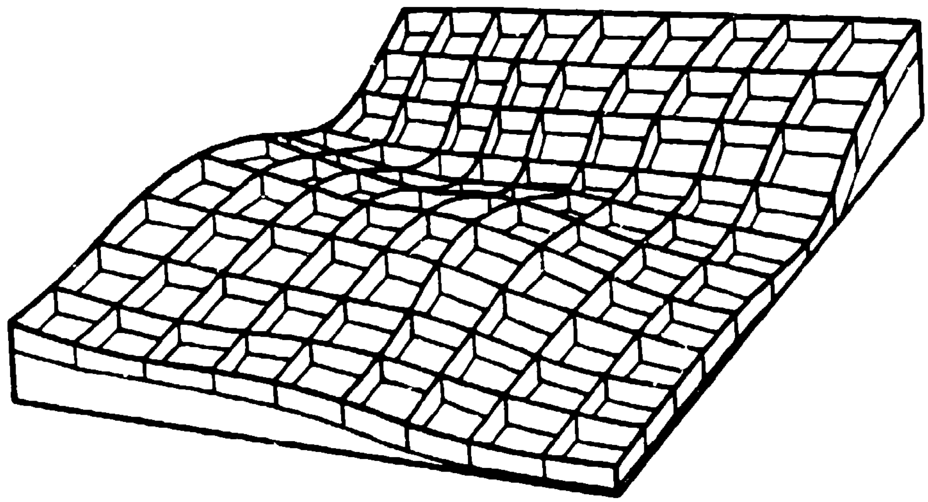

Figure 1-1
EXAMPLE OF ATR GRID

freedom, and 360° pivots are largely a matter of the granularity of the grid and the number of different directions of view photographed for each grid location. In this case, granularity is defined as the distance between grid centers. More pictures will provide the user a richer travel experience, but will result in less coverage per videodisc. The companion report to this document (Kraft, et al., 1982) provides the results of empirical research on the effects of expanding or reducing granularity relative to these issues. A summary of this research in provided in Section 2.0. This section also contains a definition of the system parameters as well as a discussion of the tradeoffs between these parameters that will define the extent to which these four goals can be achieved. Design alternatives for producing a travel system are also presented in Section 2.0.

The last three goals (360° pivots, detailed imagery, and dynamic overlays) have implications for aspects of the system other than the source to image mapping. Image detail is a matter of the richness of the visual source, regardless of the granularity with which it is filmed. The possibilities for using two sources other than the real world (terrain boards and computer-generated imagery) are discussed in Section 3.0. Dynamic overlays require a particular approach to extracting intervisibility information from the image source. The technological implications for achieving this information are discussed in Section 3.0 and the system design issues are presented in a later portion of the same section. Finally, there are specific technological options available for achieving 360° pivots, in addition to a large number of discrete views represented by photographs. These options were outlined in the research plan (Patterson, Kraft, and Buede, 1982) and are described in Section 3.0.

Section 4.0 of this report presents our conclusions and recommendations.

2.0 SIMULATING OPEN FIELD TRAVEL

The critical parameters for an ATR system are described in this section which also includes a summary of the research results of Kraft, Patterson, and Mitchell (1982), and a discussion of the major tradeoffs between the various system parameters. A glossary of parameter abbreviations can be found in Section 2.3. The section is concluded with a discussion of the details of the ATR travel concept, videodisc organization, travel components, and hardware requirements.

2.1 Issues in Videodisc Generation for ATR

As part of the ATR research, an empirical investigation (Kraft, Patterson, and Mitchell, 1982) was conducted to help design a videodisc-based system for Advanced Terrain Representation (ATR). Ideally, the system would present a complete and veridical representation of a natural tactical environment. However, because of the storage constraints inherent in videodisc technology, the amount of information which can be presented is limited. Psychological research was conducted to help define the bounds for a compelling, pedogogically effective system within the storage constraints. The research was designed to examine tactically motivated perceptual issues related to the format of the visual material. The primary issue was concern with the most efficient way to represent a large piece of terrain in a perceptually coherent and tactically informative fashion. The purpose of this research was to establish the bounds of perceptual acceptability for use in guiding subsequent technological development.

Angle of View, Jump Distance, Number of Viewing, and Travel Directions were format variables examined empirically because they represent the basic building blocks for constructing a videodisc-based representation of open-field terrain.

Angle of View refers to the angular size of the photograph used to create the visual display for the system and is directly related to the focal length of the camera lens. Angle of View typically ranges from 30° and below (telephoto) to 45° ("normal") up to 90° (wide angle) and beyond. The wider the angle, the greater the amount of information contained in the photograph, and the more the objects in the photograph appear to be stretched out.

Jump Distance refers to the maximum jump size that can be allowed between grid centers, and Number of Viewing Directions and Number of Travel Directions are both determined by the number of photographs taken from each grid center.

The proper parameters for each of these format variables was determined by four experiments which are summarized below. In Experiment 1 distance perception was examined as a function of viewing angle. Findings showed that distance perception along the depth plane was significantly affected by viewing angle; the wider the angle, the greater the perceived distance. A viewing angle of 90° appears to be the widest distortion-free angle, closely approximating distance perception in the real world. The results of Experiment 1 also demonstrated that the perceived distance between any two objects depicted in the scene was unaffected by viewing angle.

The perception of hills was examined in Experiment 2. The results demonstrated that viewing angle significantly affected steepness perception, and further, that visual travel over terrain interacted with perceived steepness. A 90° viewing angle created no more distortion or variability than any of the other viewing angles being tested. This experiment also replicated the major finding of the first experiment, i.e., widening the Angle of View increases estimates of distance. In addition, Experiment 2 results demonstrated that height perception can be affected by viewing angle: when viewers were visually on a hill, viewing angle significantly affected height perception; when the hill was viewed from a distance, perceived height remained constant across the different viewing angles.

The two format variables of concern in Experiment 3 were: Jump Distance and Number of Viewing Directions in two types of terrain. In lightly wooded terrain, visual travel remained coherent up to Jump Distances of 20m, but began to fall apart at 30m. At a Jump Distances of 40m, travel coherence was unacceptably poor. These results suggested that the maximum allowable jump size for lightly wooded terrain should be 25m. Pivots remained coherent with successive angular displacements of 15°, but become increasingly incoherent with displacement of 22.5° and 30°. These results suggested that the maximum allowable angular displacement in lightly wooded terrain should be 15° and thus, 24 viewing directions would be needed to specify a complete 360° pivot.

In open terrain, visual travel remained coherent up to Jump Distances of 35m. At 55m, travel coherency broke down, but at 75m, a high degree of travel coherence returned. It appears that the greater the Jump Distances, the greater the likelihood that a significant landmark will disappear between successive views or that the general terrain characteristics will change from from one view to the next. It was still possible to maintain coherent linear travel with a Jump Distance of 75m, given a highly homogeneous stretch of terrain. Since a potential loss of coherence can occur at 75m, however, the maximum allowable Jump Distance in open terrain should be set at 50m. Pivot coherence in open terrain was consistently high across all the levels of angular displacement that were tested. A 30° angular displacement was as coherent as a 7.5° angular displacement. Thus, the maximum allowable angular displacement in open terrain should be set at 30°, indicating that 12 viewing directions would be needed to specify a complete 360° pivot.

The objective of Experiment 4 was the determination of the Minimum Number of Travel Directions needed to provide the ATR user with a sense of free travel. Subjects viewed film sequences representing linear travel along a path which was oblique to the desired direction of travel. The results suggest that for a 90° viewing angle, subjects will become uneasy about going astray when the angular discrepancy between the desired and actual direction of travel is approximately 15°, indicating an upper limit of 24 travel directions. Subjects will experience a strong need to correct their travel path when the angular discrepancy reaches approximately 22.5°, indicating a lower limit of 16 travel directions. These results apply to both lightly wooded and open terrain. For a more complete description of this research, see Kraft, Patterson, and Mitchell, 1982.

It should be noted that a nonempirical investigation was conducted concerning image resolution. The results of this investigation indicate that a resolution of 256 x 256 pixels would provide sufficient detail and texture to support visually guided behavior for the ATR visual display.

Table 2-1 summarizes the empirical findings for formatting the ATR videodisc imagery.

FORMAT VARIABLE	PARAMETER VALUE
VIEWING ANGLE	90°
MAXIMUM ALLOWABLE JUMP DISTANCE	
- Lightly Wooded Terrain	25m
- Open Terrain	50m
MINIMUM NUMBER OF VIEWS	
- Lightly Wooded Terrain	24
- Open Terrain	12
MINIMUM NUMBER OF TRAVEL DIRECTIONS	16 or 24

Table 2-1

FORMAT RECOMMENDATIONS BASED ON EMPIRICAL RESEARCH

2.2 ATR System Parameters

The foundation of the ATR system is the terrain imagery which will be captured from the center of each grid that has been laid upon the terrain. The baseline assumptions for the ATR system are as follows:

(1) The terrain grid will consist of homogeneous squares.

(2) One terrain image (a given view direction from a given grid) will be stored on one videodisc frame.

(3) Only the "random access" and "step" modes of the videodisc player will be used. Random access of the videodisc frames takes one to three seconds for current optical videodisc systems, and this mode will be used for moving the user from one grid to another. The "step" mode can access adjacent videodisc frames in one tenth of a second or more (as controlled by the system) and will be used to present adjacent view directions to the user.

This system assumes the minimum availability of specific hardware (see Table 2-6 and Figure 2-10). More sophisticated designs are presented in Section 2.6.

The number of directions of travel permitted within the grid, Number of Travel Directions (d), and the maximum jump distance permitted within the grid, Maximum Jump (j), are intimately related. Figure 2-1a depicts four travel directions (two perpendicular streets) and eight travel directions (every 45°) in the upper left corner. The case for sixteen travel directions is shown in the middle of Figure 2-1a and for twenty-four directions is depicted in Figure 2-1b. Both of

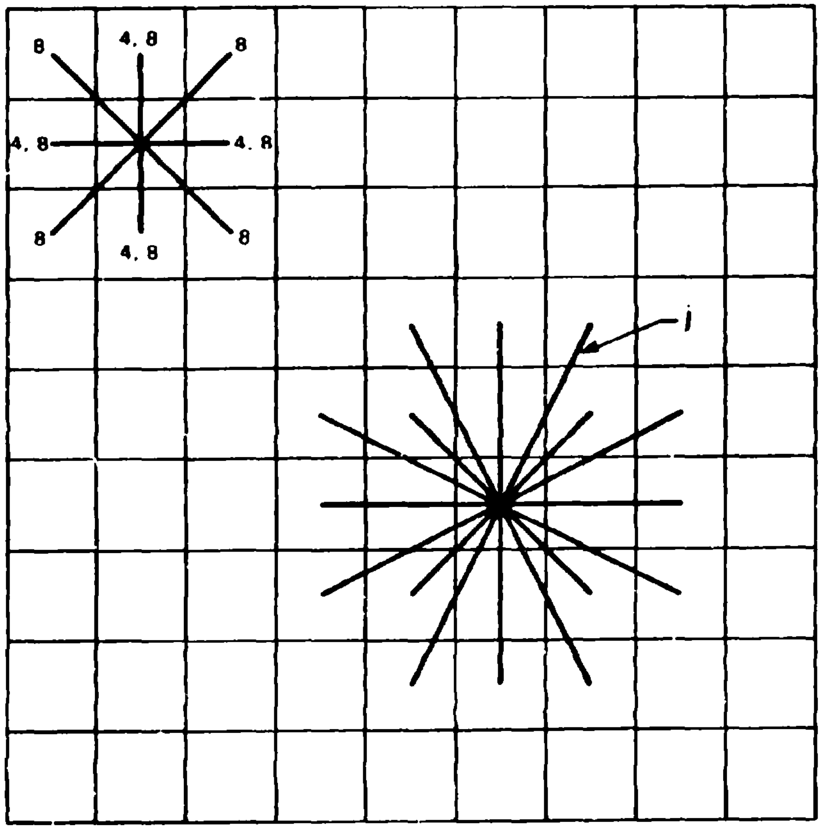

Figure 2-1a
4,8, and 16 TRAVEL DIRECTIONS

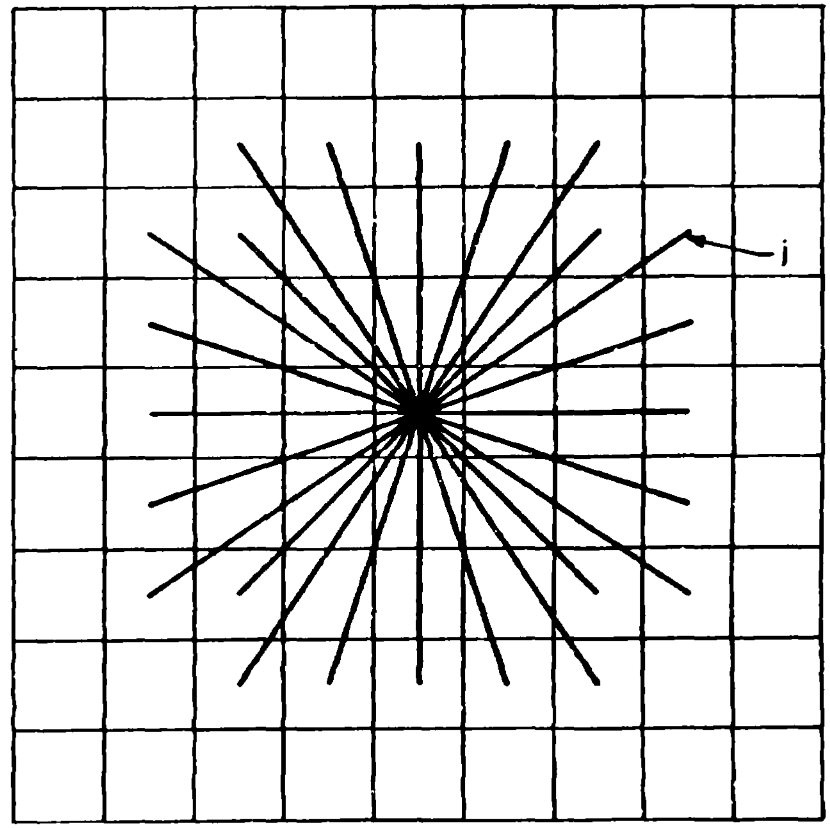

Figure 2-1b
24 TRAVEL DIRECTIONS

these cases were defined to achieve equal angular displacements between travel directions. In each case, the Maximum Jump, j, is labeled. Table 2-2 shows the Maximum Jump distance in terms of g for each of the examplary travel directions depicted in Figures 2-1a, b. The last column is the range of j's. In every situation, the minimum jump is g, corresponding to a north/south or east/west translation.

The Minimum Presentation Time (t) is a function of the ATR system's ability to access the videodisc file in anticipation of the grid to which the user's travel direction and speed are likely to take him. The system's access time is a function of hardware (e.g., the use of a frame buffer or not) and the number of non-redundant parallel videodiscs and players. These system features will be discussed in Sections 2.4, 2.5, and 2.6, but Table 2-3 presents our best estimate of t (in seconds) as a function of frame buffer capability and number of parallel videodiscs. (Note that a frame buffer will be required to refresh the terrain image if there is only one videodisc and player.)

Maximum Presentation Time (T), Table 2-4, must be determined by user acceptability. Based upon pilot tests this year, T may increase for very slow speeds (0-5 mph) since there will be little change over time. It may also be possible to use a frame buffer or special effects generator with continuous zoom capability to simulate travel during the presentation of a single terrain image. Although zoom and travel are not the same perceptually, for a period of one to two seconds they may be similar enough to convey travel and extend T. Additional discussion of this hardware option will be presented later in Section 2.0.

Minimum Speed (s) is the Maximum Jump (j) divided by Maximum Presentation Time (T): s = j/T. Since j is a function

Number of Travel Directions - d	Maximum Jump - j	j Range
4	g	----
8	1.41 g	1.0 - 1.41 g
16	2.16 g	1.41 - 2.16 g
24	3.46 g	2.82 - 3.46 g

Table 2-2

RELATIONSHIP BETWEEN NUMBER OF TRAVEL DIRECTIONS AND MAXIMUM JUMP

Number of Videodiscs/Players	Frame Buffer No	Frame Buffer Yes
1	--	1.0
4	1.0	0.5
9	0.5	0.33
16	0.33	0.25
25	0.25	0.20

Table 2-3

MINIMUM PRESENTATION TIME (SECONDS)

(Based upon a one-second random access search by the videodisc player)

System Parameter	Symbol	Definitions
Grid Size	g	distance between grid centers (meters)
Coverage	c	area of terrain covered by the grid (km^2)
Number of Travel Directions	d	number of directions of travel permitted within the grid
Maximum Jump	j	maximum jump distance permitted within the grid (meters)
Minimum Presentation Time	t	minimum time for presentation of terrain imagery (seconds)
Maximum Presentation Time	T	maximum time for presentation of terrain imagery (seconds)
Minimum Speed	s	minimum travel speed (greater than 0) permitted by the system (mph)
Maximum Speed	S	maximum travel speed permitted by the system (mph)
Number of View Directions	v	number of directions of viewing from a given grid location
Videodisc Frames	f	number of videodisc frames available for storage of terrain imagery

Table 2-4

SYSTEM PARAMETERS

of the particular travel direction chosen within each of the cases of 8, 16, 24 Travel Directions (Table 2-2), s also varies with d. For planning purposes s should be determined by the largest j in the range of each d. If T can be increased beyond one second, s can be decreased.

Maximum Speed (S) is the smallest jump in the range of Maximum Jumps for a given Number of Travel Directions (d) (as shown in Table 2-2) divided by t, Minimum Presentation Time. Figure 2-2 presents the ranges of Maximum Speed (S) and Minimum Speed (s) as a function of Grid Size for d equal to eight.

The Number of View Directions (v) is an independent parameter that directly impacts coverage. However, in order to preserve coherent travel, there should be a view direction for each travel direction so the traveler can look in the direction being traveled. This means that d is a lower limit for v.

There are approximately 54,000 frames on one side of a videodisc. Thus Videodisc Frames (f) equals 54,000, 216,000, 486,000, 864,000, and 1,350,000 for 1, 4, 9, 16 and 25 nonredundant videodiscs, respectively. The reason for concentrating on 1, 4, 9, 16, and 25 is to decrease Minimum Presentation Time, which will be explained in more detail in section 2.4.

Coverage (c) equals the Number of Videodisc Frames (f) times the area covered by one grid element (g^2), divided by the Number of View Directions per grid (v); then, $c = \frac{fg^2}{v}$. Or for a given coverage $f = \frac{cv}{g^2}$.

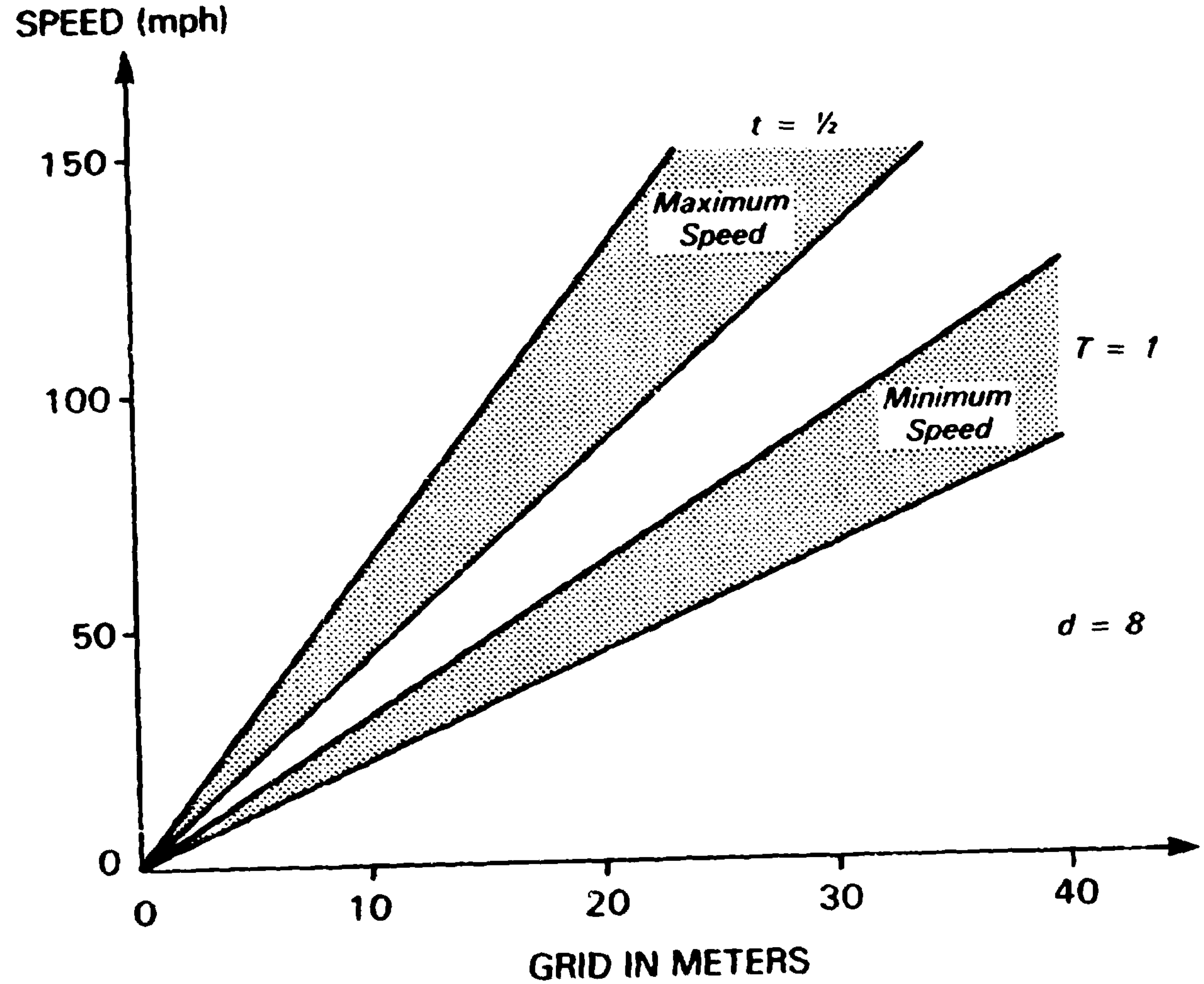

Figure 2-2
PLOT OF TRAVEL SPEED VS. GRID SIZE

2.3 ATR System Parameter Tradeoffs

In order to produce a reasonable dispersion between Minimum and Maximum Travel Speeds (s and S), there must be a dispersion between Minimum and Maximum Presentation Times (t and T):

$$s = \frac{j}{t}$$

$$S = \frac{j}{T}$$

where j = the Maximum Jump.

Referring back to Table 2-3 and our pilot studies which indicate that T should not be greater than one second, our conclusion is that a significant number of non-redundant videodiscs and players are needed to achieve a spread between T and t. See Table 2-4 for a glossary of terms.

Figure 2-2 shows the ranges of Maximum and Minimum Speeds, as a function of Grid Size (g), for the case of 8 Travel Directions. The ranges for S and s result from the ranges of Maximum Jump, depending upon the particular travel direction. The point of this figure is the high values for Minimum Speeds (s) when Grid Size (g) is greater than 10m. The range of Maximum Speeds (S) when Grid Size (g) equals 10m falls well within the reasonable limits for tank travel. Increasing the Number of Travel Directions to 16 and 24 as recommended empirically, exacerbates these problems because Maximum Jump increases even more. Table 2-5 presents the relationship between g and the empirical recommendations for j, as a function of Number of Travel Directions (d). In conclusion, g

	Number of Travel Directions d		
Maximum Jump Recommendations	8	16	24
Lightly Wooded: 25m	17.7m	11.6m	7.2m
Open: 50m	35.5m	23.2m	14.5m

Table 2-5

GRID SIZE FOR EMPIRICAL RECOMMENDATIONS
OF MAXIMUM JUMP

must be between 7 and 12m for lightly wooded terrain and 14-23m for open terrain.

The empirical research suggests that at least 16 and probably 24 view directions are needed to convey a continuity of scanning and traveling to the user. It is this variable together with Coverage (c) and Grid Size (g) that determine the number of Videodisc Frames (f) needed. Figure 2-3 shows the relationship between f and g when Coverage is set at our original goal of $8km^2$ (2 x 4) and Number of View Directions (v) is set at 16 and 24. This figure suggests that g must be kept above 17-20m to achieve nine or fewer videodiscs. Figure 2-4 plots f versus c for cases of g = 10, 15, 20 and 30m and v set at 16. This figure also demonstrates the need to keep g above 15m to achieve reasonable coverage with nine or less videodiscs.

One of the major issues that must be resolved during system development is the tradeoff between the need to increase Coverage and to decrease Minimum Speed, factors which pull Grid Size in different directions. Figure 2-5 shows that Grid Size (g) must be in the 5-10 meter range in order to achieve a Minimum Travel Speed (s) of 10-20 mph for 16 or 24 directions of view. One method for achieving such a tradeoff is the use of the concept of heterogeneous or adaptive grids. This concept would allow grid size to decrease when moving from open to lightly wooded terrain and would provide small grids for precise positioning and slow speeds. However, the use of such a concept would also introduce additional complexity into the system if smooth travel from one grid center to the next were to be maintained.

Even with the use of an adaptive grid concept, the size of a grid would need to average approximately 15m in order to confine the system to a reasonable number of videodiscs. A

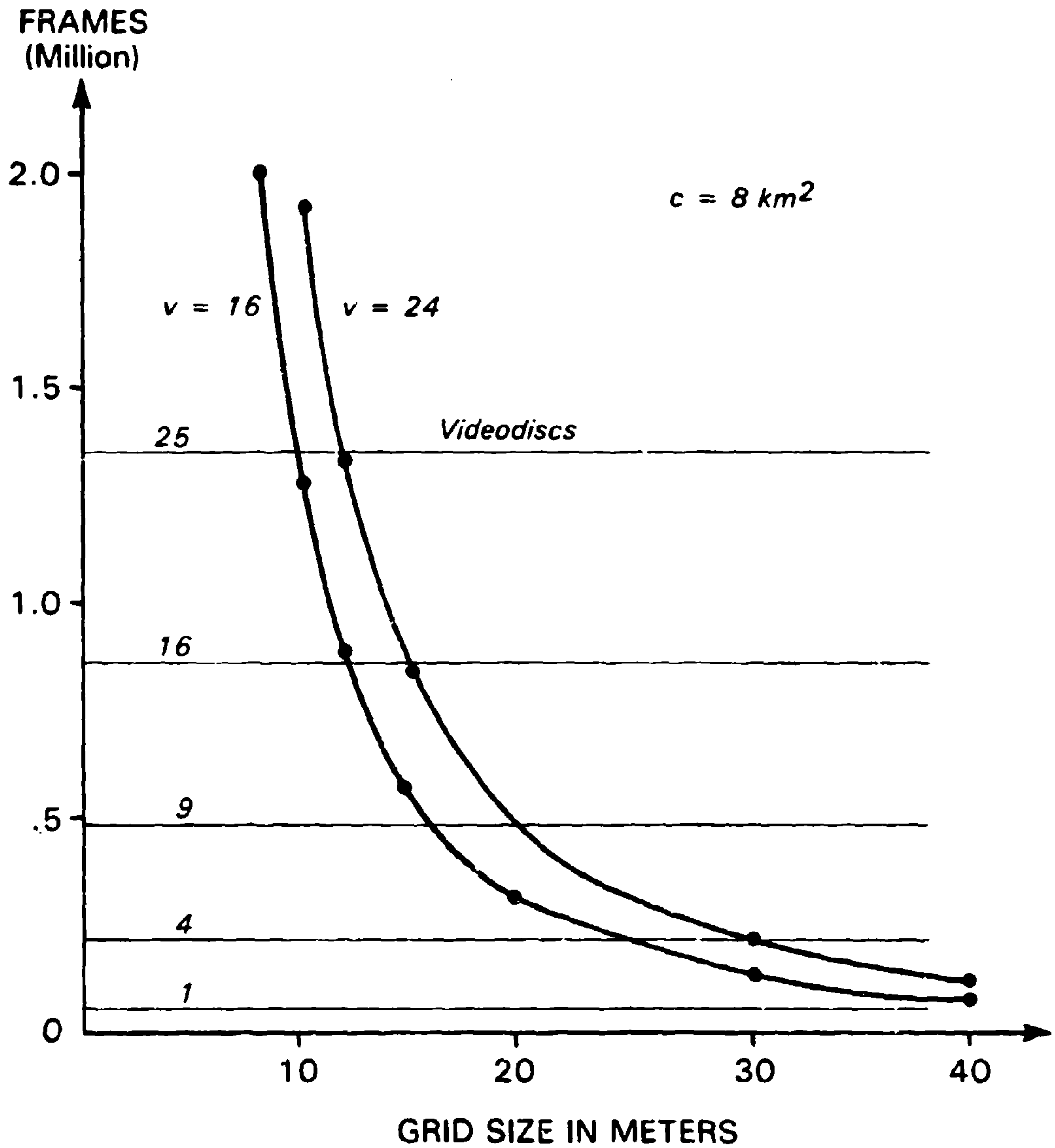

Figure 2-3
PLOT OF FRAMES VS GRID SIZE

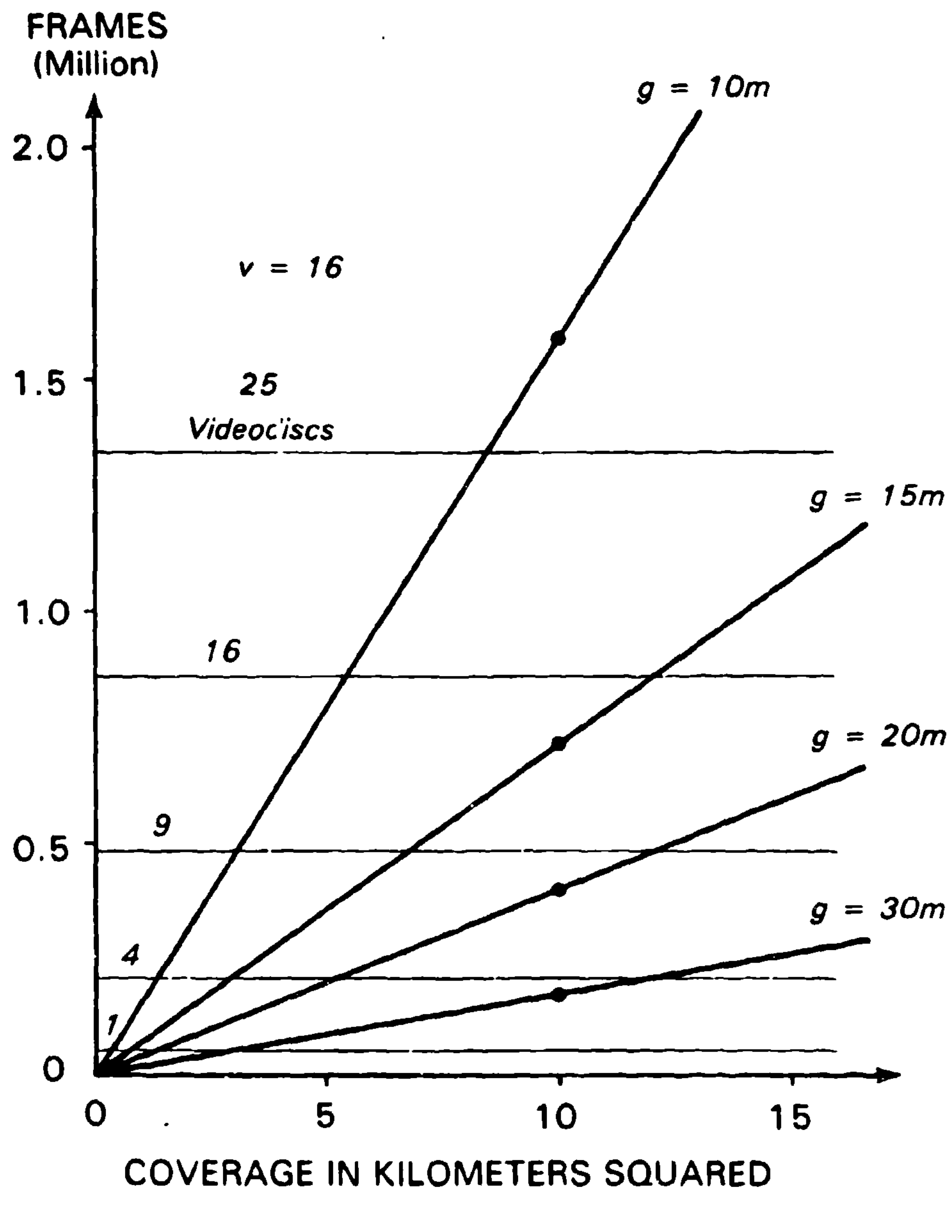

Figure 2-4
PLOT OF FRAMES VS COVERAGE

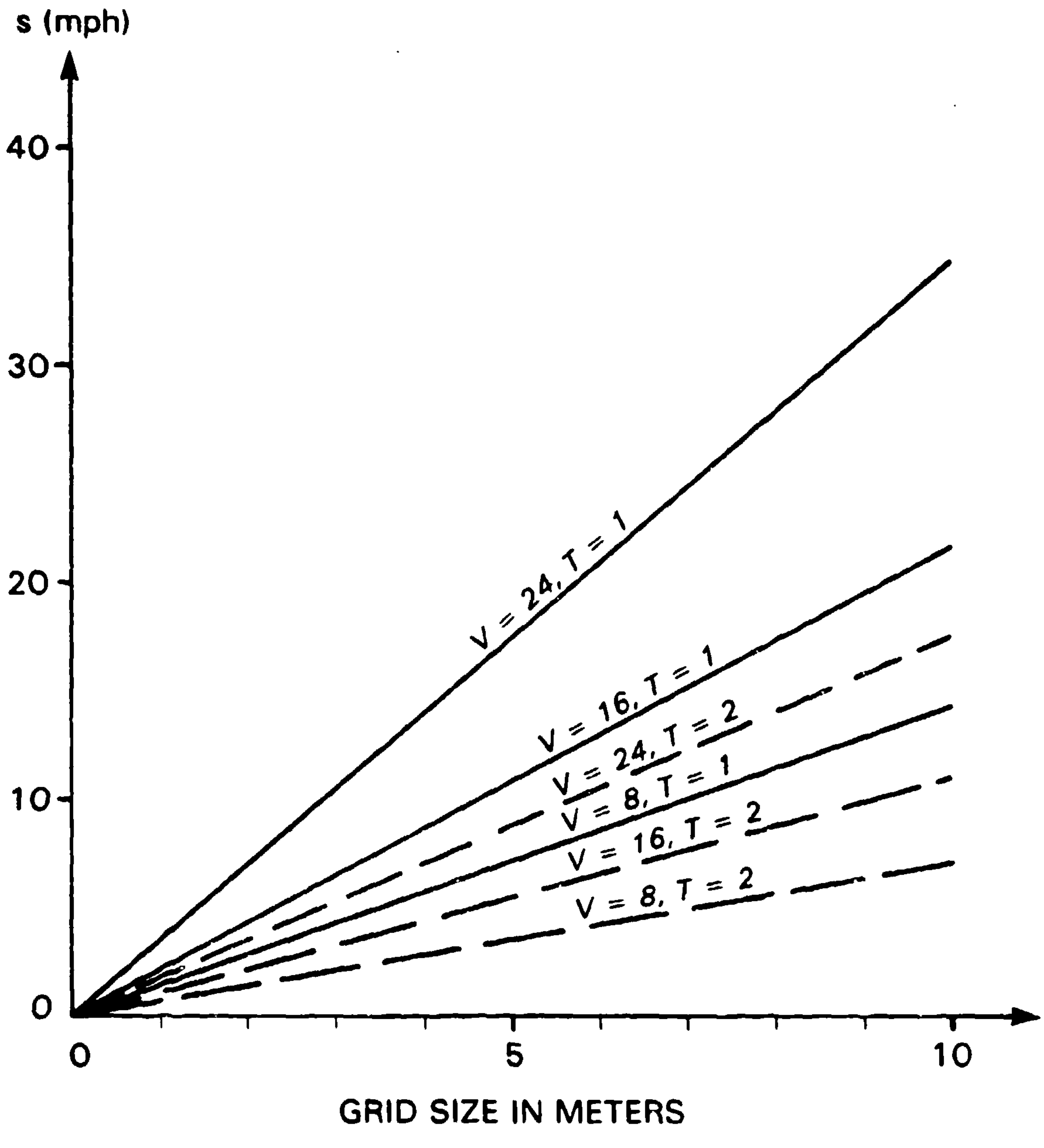

Figure 2-5
PLOT OF MINIMUM SPEED VS GRID SIZE

15m grid produces an average Minimum Speed above 35-45 mph and this is faster than a tracked vehicle will travel over rough terrain. By increasing the Maximum Presentation Time above one second per frame through the use of frame buffers or special effects generators, a satisfactory tradeoff between speed and coverage could be achieved. Both frame buffers and special effects generators can provide the continuous zooming capability needed for this purpose. Storage can also be reduced, and coverage increased, if more than one image is stored per videodisc frame. This possibility is discussed more fully in Section 2.4.2.

2.4 Videodisc Organization and Use

Currently, the state-of-the-art videodisc players require a search time of one to three seconds in the random access mode. However, faster search times would not negate the need to use parallel videodisc players, i.e., two identical discs on two parallel videodisc players, in order to prevent blank outs on the display. While one player is playing, the other player is searching. In addition to decreasing the Minimum Presentation Time (search time) between minimum (t) and maximum (T) search times, multiple videodiscs would also allow for incresing the amount of terrain coverage. (For the purpose of this report one second search times will be assumed.)

Videodisc players operate in three modes: random access mode mentioned above; step mode; and play mode. Play mode is used for applications such as commercial movies. The step mode is used for the alow motion presentation of images on adjacent frames. Since it is not possible to produce a videodisc for every possible travel sequence the user might take, play and step modes have no application to the simulation of movement. However, these modes may be useful for pivoting the user in place, that is providing the user the opportunity to turn his head to scan the environment. Section 2.5 addresses the use of the step mode for pivots.

In order to decrease the random access search time of the videodisc player below one second, it is necessary to have multiple videodiscs operating in parallel. That is, one videodisc is displaying an image while others are searching for the next image. A frame buffer provides some redundancy since it can refresh the image it has "received" until the next image is received, thus freeing the videodisc player to begin searching for the next frame.

Section 2.4.1 describes how the grids that are laid upon the terrain are allocated to videodiscs to achieve the Coverage shown in Figures 2-3 and 2-4 with the search times that are consistent with the Maximum Presentation Times shown in Table 2-3. Section 2.4.2 describes how the decrease in image resolution from 512 x 512 pixels (picture elements) to 256 x 256 can be used to increase Coverage.

2.4.1 <u>Grid/videodisc patterns</u> - If the terrain imagery were to be stored on one videodisc, then a frame buffer would be required to refresh the image while the videodisc player is searching for the next image. The pattern in which the grids are stored on the single videodisc is not critical since a random access will be needed for each new image using the disc's directory. Two or three videodiscs and players can be used in series to increase the Coverage, but will have no systematic impact on Minimum Presentation Time.

The use of four videodiscs in the grid allocation pattern shown in Figure 2-6 would permit the videodisc players to search in parallel for every possible Travel Direction depicted in Figures 2-1a and b. As the pattern is defined every other grid on the odd-numbered rows is stored on either the

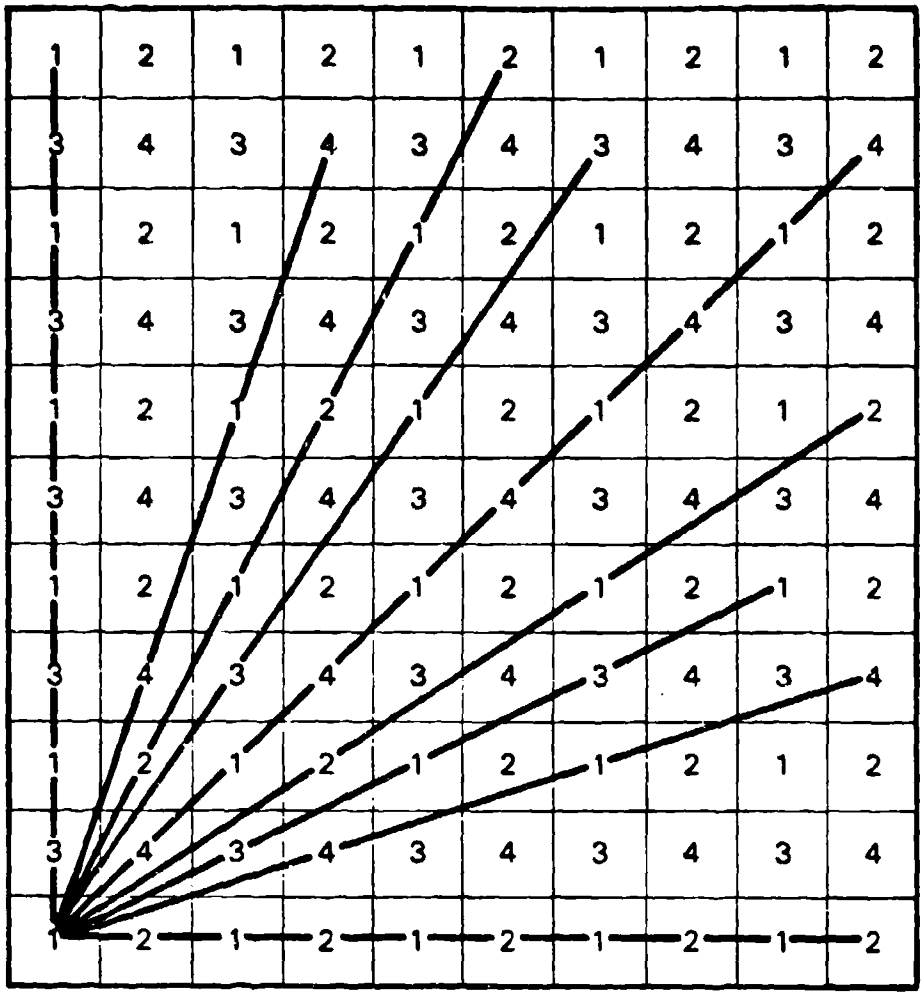

Figure 2-6
GRID ALLOCATION FOR FOUR VIDEODISCS/PLAYERS

first or second videodisc, while alternative grids for the even-numbered rows are stored on the third and fourth videodiscs. This pattern permits two videodiscs/players to operate in parallel for every possible travel direction shown in Figures 2-1a and 2-1b. The lines in Figure 2-6 demonstrate one quadrant of these travel directions. (The other three quadrants are symmetrical.) The breaks in each line indicate the stopping points for travel from the lower left corner and define the two videodiscs that would be operating in parallel for each Travel Direction. If there is no frame buffer in the system and the videodisc random access search time is one second, this arrangement permits a Minimum Presentation Time of one second. The addition of a frame buffer for image refreshing will decrease t to 0.5 seconds since two videodisc players are searching in parallel, each with a one second response. It is important to note that this pattern requires that each successive grid be used in the north/south, east/west, and 45° diagonal travel directions, thus reducing the Maximum Jump in these directions to g for N/S and E/W and 1.41g for the 45° diagonals.

2.4.2 <u>Other options for extending Coverage</u> - There are at least two additional ways of extending terrain coverage on a videodisc. Extended coverage can be accomplished in the following ways:

(1) by packing four images onto one videodisc frame, thereby decreasing the image resolution from 512 x 512 pixels to 256 x 256 pixels; and

(2) by using a transmissive videodisc player which can assess both sides of an optical videodisc without manual or mechanical assistance.

The packing concept is shown in Figure 2-7. The outer square represents a single videodisc frame of 512 x 512 pixels. Placing four images, or pictures, on one frame reduces the pixels per image to 256 x 256. Pilot research indicated that this amount of degradation would not seriously impact the user's perception. However, any further decrease in resolution would impact the user's perception significantly. Packing four images onto each videodisc frame increases coverage, everything else being equal, by a factor of four. The display of packed frames does, however, require additional computer logic and either a frame buffer or special effects generator to find and display the appropriate images of the four images available on each frame.

A comparison of Figures 2-8 and 2-3 reveals the advantages to be gained from packing. For the same eight square kilometer coverage, packing would require a half-million frames and a 9 videodisc player configuration to produce a 10 meter grid center coverage. Without packing, the same coverage would require two million videodisc frames and a 25 videodisc player configuration. The nine videodisc configuration without packing would require 17 to 20 meter grid centers. With packing the 9 videodisc player configuration would allow 10 meter grid centers and would place Grid Size within the desired range shown in Table 2-5.

The second option for extending Coverage is to use a transmissive videodisc player (e.g., Thomson CSF) that has a tunable laser capable of accessing information on both sides of a videodisc without requiring that the videodisc be turned over. Most videodisc players such as Pioneer, RCA, and DISCO-VISION do not currently provide this capability. Accessing both sides of the videodisc in this way nearly doubles the

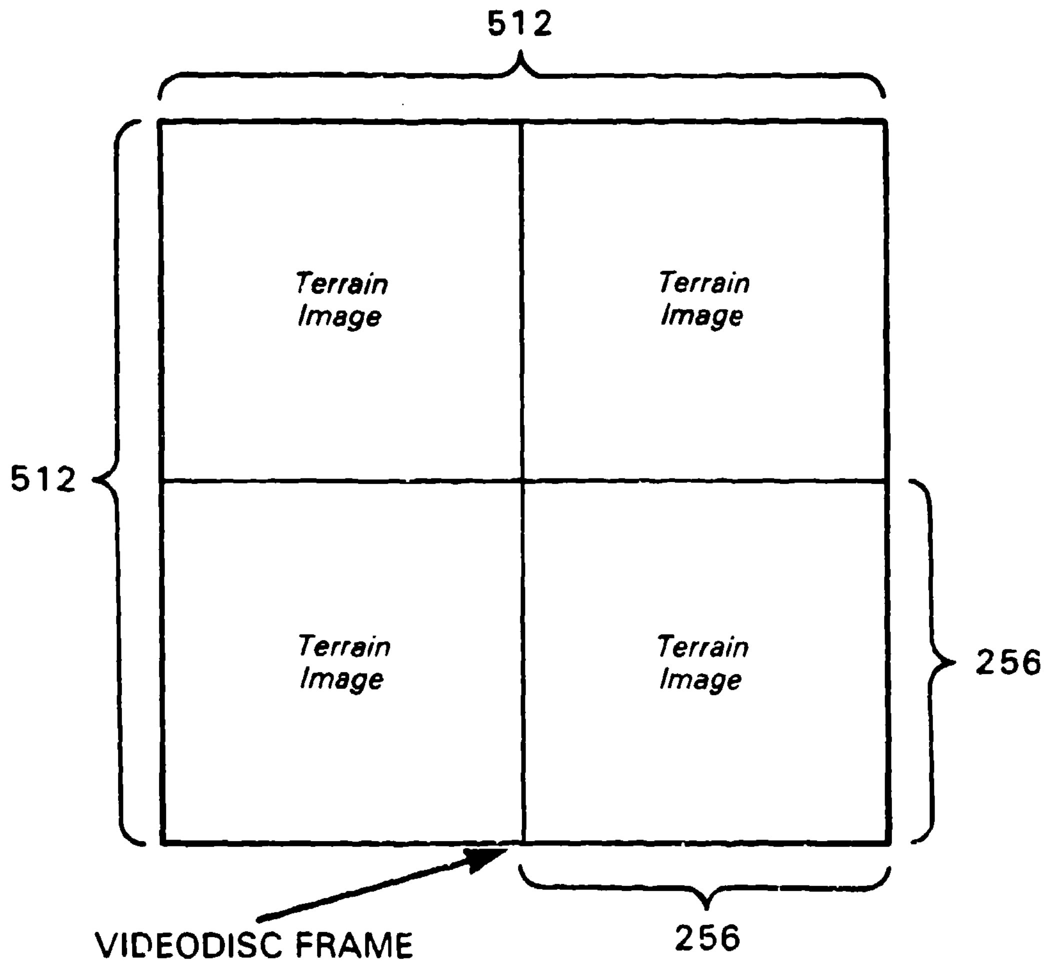

Figure 2-7
FOUR TERRAIN IMAGES PER VIDEODISC FRAME

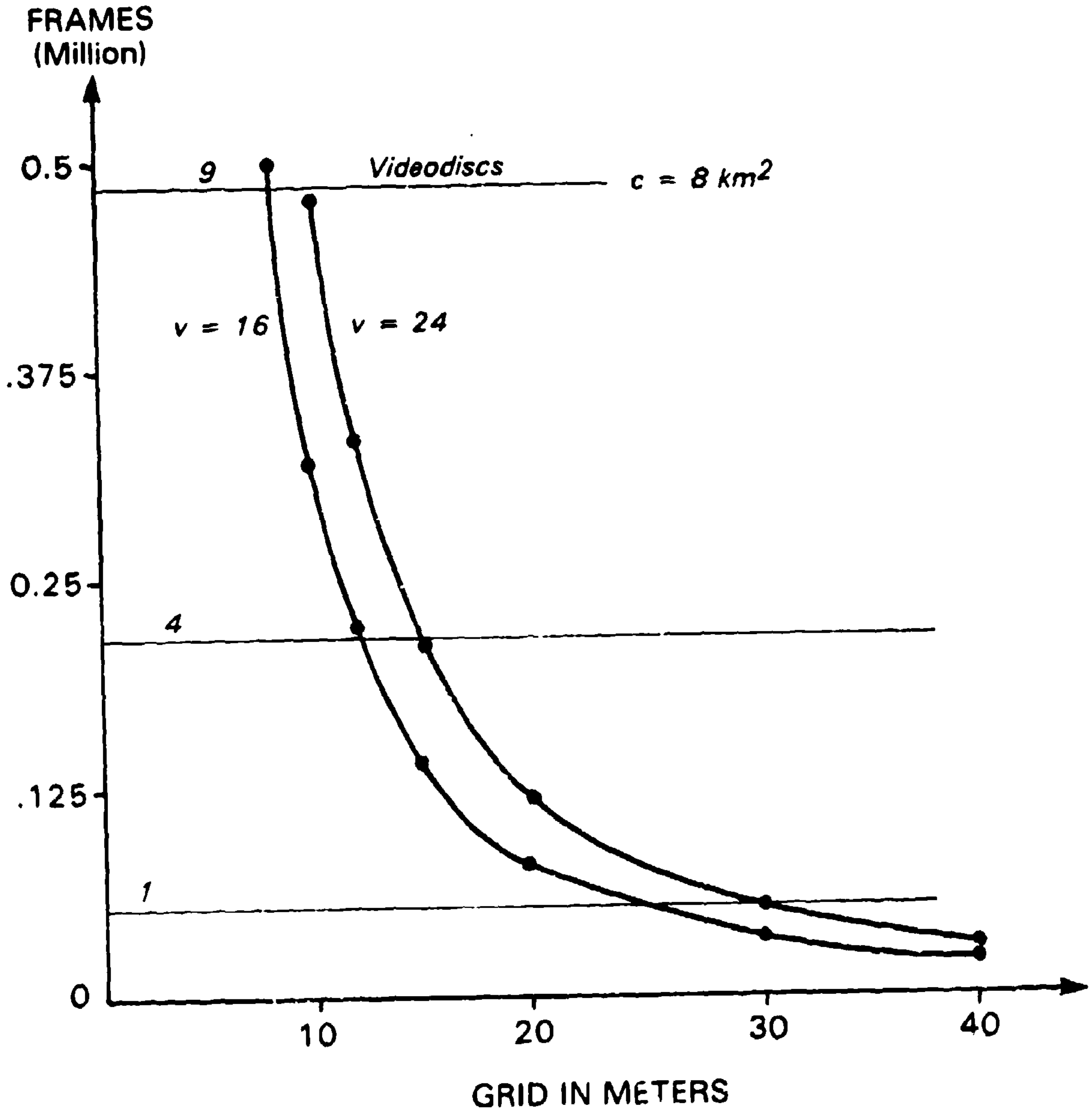

Figure 2-8

PLOT OF FRAMES VS GRID SIZE
(PACKING FOUR IMAGES PER FRAME)

number of frames per videodisc from 54,000 to 94,000. Unfortunately, Thomson CSF no longer produces transmissive videodisc players, nor does anyone else. Costwise, the Thomson transmissive videodisc player is three to four times as expensive as a non-transmissive (reflective) player.

2.5 Simulating Straight Travel, Turns, and Pivots

The user will have access to a controller, e.g., joystick, that will permit him to not only initiate movement forward (in the direction of view), but also to change the speed and direction of movement. For the purpose of this discussion, traveling across terrain will be divided into three components: straight travel, turns, and pivots. Straight travel and turns are associated with the movement (translation) of the user's vehicle across the terrain. Pivots are defined as rotations of the user (or more typically his head) for the purpose of scanning the terrain, and can occur while the user is stationary or moving. The following three subsections discuss each of these three components of traveling, in turn.

2.5.1 Straight travel - Traveling in a straight line will be simulated by the successive presentation of terrain images sampled from the grids lying along one of the discrete travel directions permitted by the system. Referring back to Figure 2-2, travel north from the lower left corner would be simulated by displaying terrain images lying along the vertical line for the fraction of a second that is between the Minimum and Maximum Presentation Times. The system will choose that travel direction which is closest to the user's indications, direct the appropriate videodisc players to begin searching for the needed images, and then sample terrain images from the videodiscs in sequence of movement. Once each videodisc player has been released from displaying its image, its player will begin searching for the next frame.

2.5.2 Turns - The user will indicate a turn in his movement through the control. Such a turn while moving is envisioned as being smooth and would be implemented by changing from one travel direction to the next (e.g., north to north northwest for d = 16) within one grid jump interval. This smooth change will be appropriate when the user wishes to make a slight adjustment to his travel direction because the discrete number of travel directions will not permit him to move in exactly the direction desired. The slow, smooth turns that result will also be appropriate in many other cases of traveling across open terrain. The user will be able to make one travel direction change for each grid jump that is made. It is only when sharp, quick direction changes are desired that the user must adopt a different procedure.

Since the user will be able to move in any direction that he is viewing when he is stopped, sharp direction changes can be effected by slowing to a stop, pivoting, and then resuming straight travel. This type of pivot is analogous to the manner in which a tank pivots, but can be permitted only when the user is stationary.

Turns will be implemented in a timely manner by having those videodiscs that are not actively searching to maintain straight travel, search for the grids which would allow one discrete change in travel direction. For the configuration of four videodisc players, the system can be prepared for a change in travel direction to either the left or right, but not both. Since travel would hesitate for a second to react to about half of the user's turns, this would be somewhat limiting. Greater flexibility would require more videodisc players (see Appendix A) or other image presentation alternatives.

2.5.3 Pivots - Pivots will be used for two purposes: (1) a simulation of the user scanning his environment while stationary and (2) the presentation of terrain off-center to the user's direction of travel while moving. (Coming to a stop and pivoting to make a sharp, quick turn falls within the first category.) Since scanning one's environment can often take place rapidly, especially while stationary, and since the videodisc players are going to be sequenced for searching out grids for future jumps, it would be most efficient to use the videodisc player's "play" mode or "step" mode to simulate scanning. "Play" mode accesses adjacent frames at a rate of thirty per second, which for scanning 16 or 24 discrete views within a 360° revolution is much too fast. The "step" mode also accesses adjacent frames, and its rate can be controlled by the system. The maximum rate for the "step" mode is ten frames per second, which is reasonable for a quick 360° turn. However, since people are used to turning their heads and not rotating their vehicle when scanning, a 180° scan in one direction is the most that is needed.

We have been assuming that the number of directions in which the user can look is equal to the number of directions of view from a given grid location. However, a problem occurs when the step mode is used for scanning (turning the head to look). Since step mode can access only adjacent frames, a viewer facing north (N) at x grid can look 360° to the right but cannot look to the left because N is not adjacent to NW. Figure 2-9 illustrates this problem. However, a 180° view in either direction is possible for a viewer who begins his scan from a south (S) or SE point.

To achieve a 90° turn in the "step" mode from any possible direction of view, four additional (50% more) frames

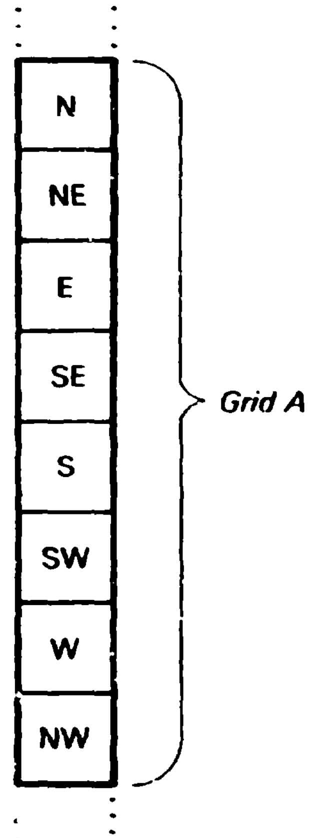

Figure 2-9a

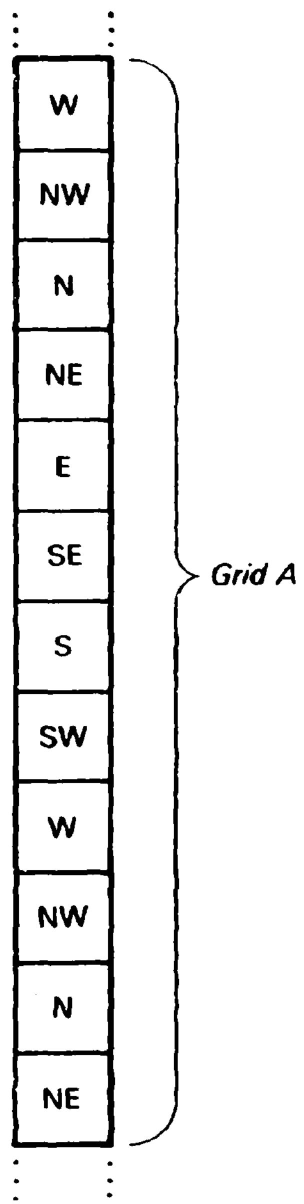

Figure 2-9b

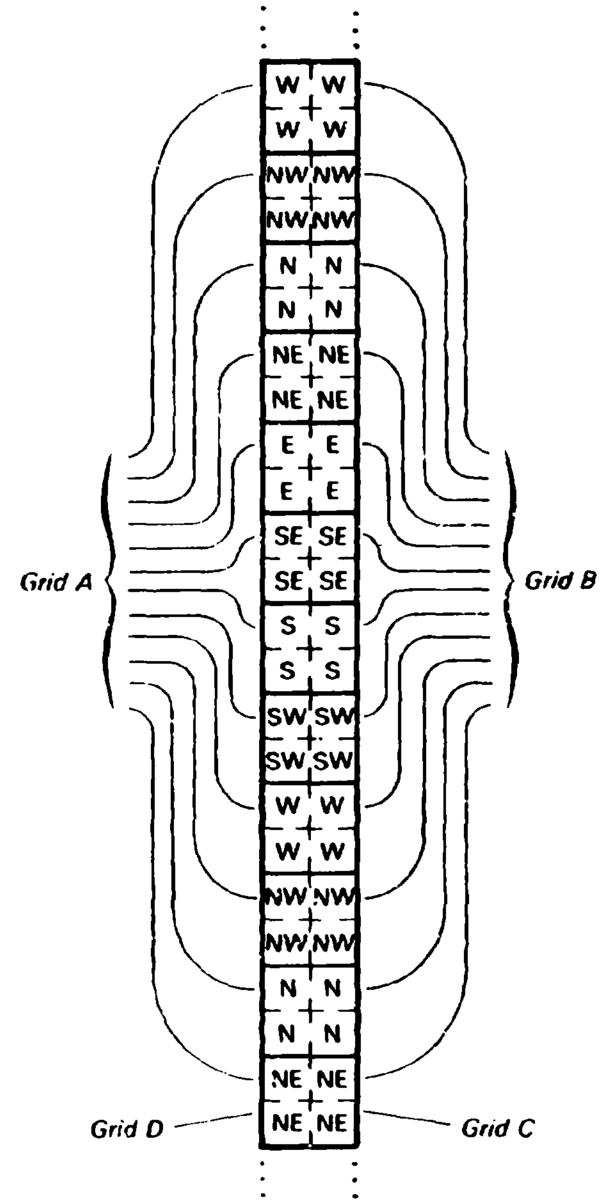

Figure 2-9c

VIDEODISC SEQUENCES FOR VIEW DIRECTIONS

must be added to the videodisc (Figure 2-9). This would result in a 33% decrease in Coverage for any of the illustrations discussed earlier in Section 2.0. In order to provide a scanning capability of 180° in each direction from any possible view, a 720° rotation must be stored on adjacent frames which is a 100% increase in frames (50% decrease in Coverage). Based upon the Coverage limitations discussed earlier, this latter option is probably excessive. However, the 90° turn in any direction is necessary and can be combined with the 256 x 256 resolution option to achieve a net gain in coverage with respect to the discussion in Section 2.2. Figure 2-9 shows the frames in sequence for scanning in "step" mode with the four grids (A, B, C, and D) packed into each frame. This design requires either a frame buffer or special effects generator, both of which will slow down the stepping time on the order of 1/30th of a second.

Pivots that are made while moving will have to be sequenced into the search process for the next grid jump using a procedure that is analogous to that used for turns. That is, if the user is traveling and looking north and wants to turn his head to the east, each successive grid jump will incorporate a new viewing angle that is one Direction of View more to the east than the last. So, if there are eight Directions of View, it will take two grid jumps to turn from north to east; four if there are sixteen Directions of View.

2.6 Hardware Configurations

The previous three subsections have discussed ATR system parameters and protocols with the intention of illuminating the non-hardware characteristics of the system design. Throughout those discussions, however, the hardware implications have been only briefly considered (e.g., capabilities requiring frame buffers). The following two subsections provide descriptions of two potential hardware configurations for an ATR system.

2.6.1 Low cost system configuration - The lowest cost ATR system would include user controls, a micro-computer, several videodisc players with controllers, a switch, and a color monitor. These items are pictured schematically in Figure 2-10. The major difference between this configuration and others of higher cost, is that this configuration has no frame buffer.

The micro-computer receives inputs concerning direction and speed of travel and view direction from the user through the controls. The videodisc players are sequenced to find the correct videodisc frames in a timely manner by the microcomputer. The display of the video signals from the videodisc players on the color monitor is controlled by the micro-computer via an electronic switch. Table 2-6 presents approximate costs for this system.

2.6.2 System configurations with a frame buffer - The term frame buffer is the generic name given to an image processing device which must receive the image in digital form, can refresh an image for continuous presentation on a CRT, and can process some image modifications such as zooming. The addition of a frame buffer to the ATR system makes the system more powerful, as was discussed in relation to frame packing in Section 2.4.2, and extended image presentation time in Section 2.1. Though there are a number of frame buffer packages on the market, they vary widely in their levels of capability. However, all frame buffers do require the transformation of the analog signal coming from the videodisc into a digital form. While some frame buffers have an internal analog to digital (A/D) converter, others do not. Since some means of

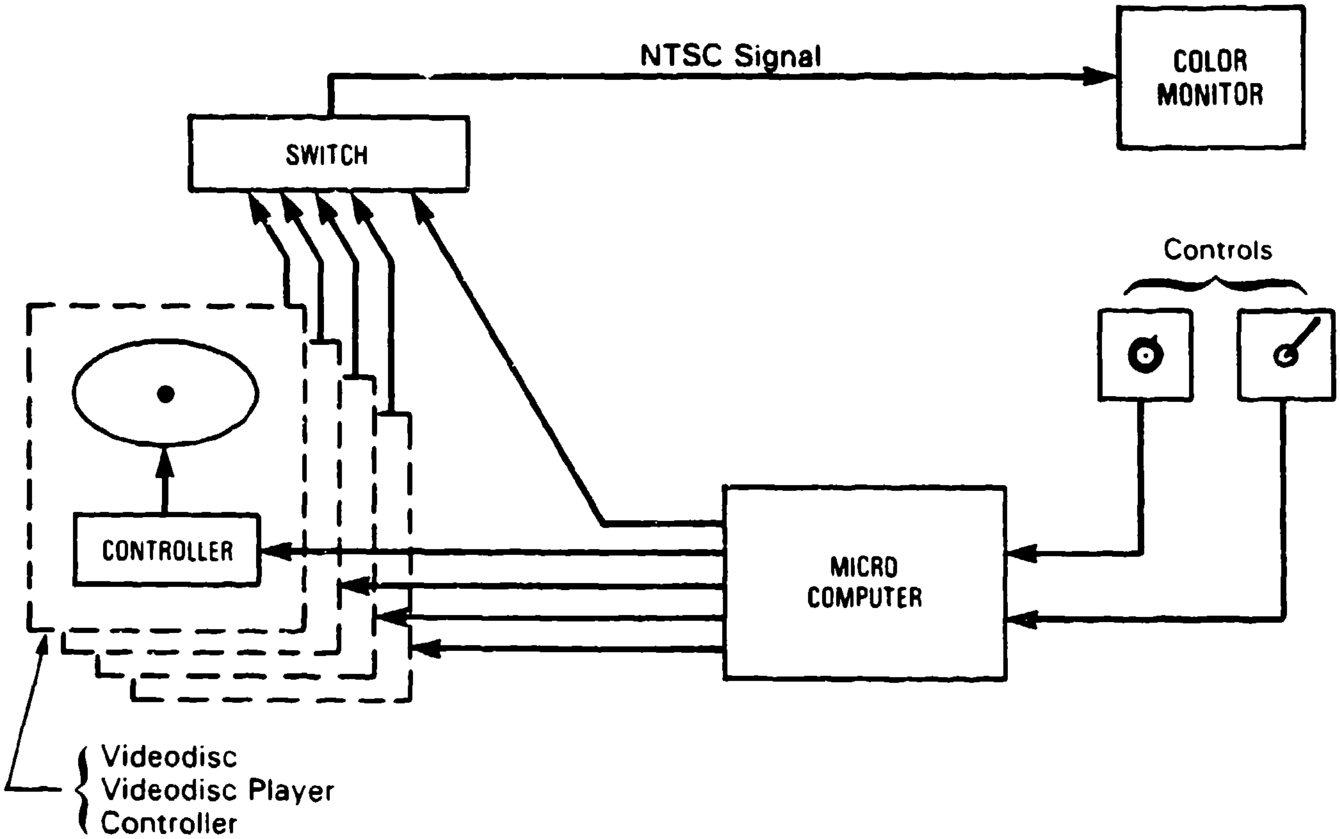

Figure 2-10
LOW COST SYSTEM CONFIGURATION

COMPONENT		$ COST
Controls		$ 1,000
Microcomputer		$ 5,000-6,000
Four Videodisc/Players		$10,000
Switch		$ 5,000-6,000
Color Monitor		$ 1,000
	TOTAL	$22,000-24,000

Table 2-6

ESTIMATES FOR LOW COST SYSTEM CONFIGURATION signal conversion must be available if the use of a frame buffer is planned, an external converter is shown in Figure 2-11.

Figure 2-11 shows a schematic for two possible system configurations using a frame buffer. The controls, microcomputer, videodisc players, and switch are central to all options, including the low cost configuration in Figure 2-10. The less expensive of the two configurations, shown in Figure 2-11, uses a black and white TV monitor and is shown with the dotted line on the top; the color configuration is shown below.

The frame buffers for both the black and white, and the color systems would provide continuous zoom capability. The differences in frame buffer costs between the two systems (see Appendix B) is accounted for by the use of color in one of the systems; the color system requires a more sophisticated frame buffer to mix red, green, and blue signals. Both of the frame buffers listed would have zoom increments on the order of 0.06, as opposed to a 0.002 increment provided by a top-of-the-line frame buffer.

With a frame buffer in a nine videodisc system, the desired Coverage and Minimum Speed tradeoff would be less difficult than for the low cost configuration shown in Figure 2-11. The ability to pack videodisc frames, which is inherent in the use of a frame buffer, eases this difficult tradeoff even more.

Finally, neither of the frame buffer configurations, as costed in Appendix B, have the capability required to overlay symbology on the terrain with proper intervisibility adjustments. This capability and the frame buffer requirements to support it are discussed in the next section.

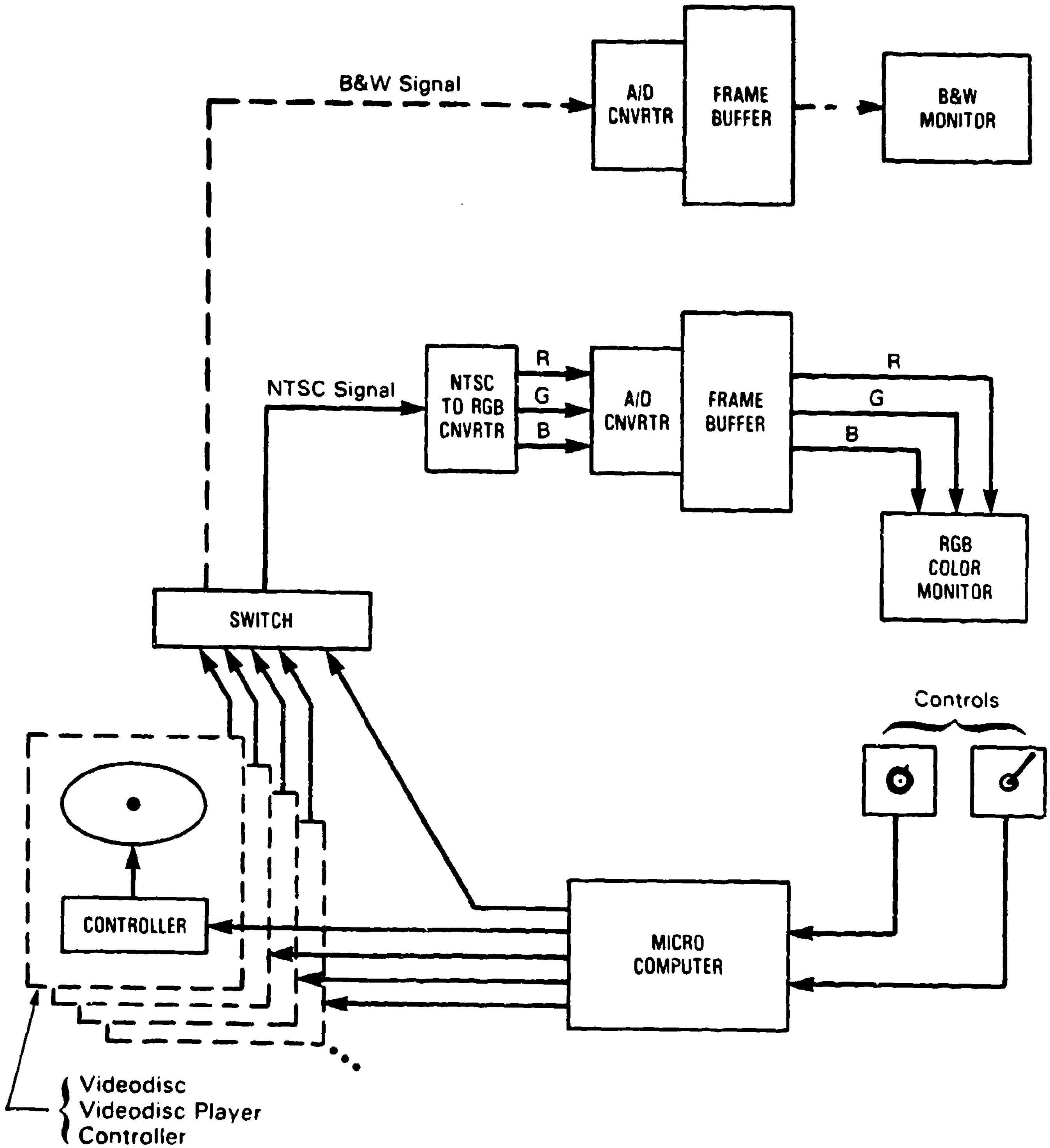

Figure 2-11

SYSTEM CONFIGURATION WITH FRAME BUFFER

3.0 TECHNOLOGICAL ISSUES AND SIMULATING INTERVISIBILITY

During the first year of ATR research the two major technological issues that developed were:

(1) Capturing the range data needed to calculate the occlusion of symbols to be overlayed on the terrain image.

(2) Displaying a 360° view with four or fewer images.

The first technological issue spawned two lines of investigation. The first line of investigation was oriented towards generating a range image using the real world as the information source. The second line of investigation was focused on the availability of non-real-world information sources (e.g. terrain boards and computer-generated imagery) since the range data may be more readily available from these two sources.

The second technological issue produced two concepts: panoramic photography and image rectification.

Section 3.1 provides a summary of our findings relative to each of these topics: range images, terrain boards, computer-generated imagery (CGI), panoramic images, and image rectification. Section 3.2 summarizes our concept for simulating the intervisibility that is required for the overlay of dynamic imagery.

3.1 Technological Research Issues

This section reports the results of DDI's consultation with a number of people at several organizations concerning the research issues identified in the ATR Task 2 Research Plan (Patterson, Kraft, and Buede, 1982). The section is organized into five subsections: range images, terrain boards, computer generated imagery (CGI), panoramic images, and image rectification.

3.1.1 Range images - A range image is a data source available to the computer containing information about the visual stimulus, or picture, seen on the CRT. A range image must contain distance information about every object in the field of view, where distance is measured from the observer to each element. Without this information, an overlayed tank might appear to be completely disoriented, hovering above the ground for example. The need for and means of acquiring a range image are documented in Sections 3.1.1.1 and 3.1.1.2.

3.1.1.1 Intervisibility calculations - The problem of overlaying dynamic imagery, e.g., enemy or friendly vehicles, can be divided into two components: (1) determination of the size, position, and orientation of the overlayed object in relation to the base image; and (2) determination of occlusion, or what is in front of or follows what. Determination of size, position, and orientation is relatively straightforward, see Figure 3-1. Given information concerning the actual size, position, and orientation of object "A" within the simulated terrain, and given information about the camera elevation and orientation for the background or image, projective geometry provides the size, position, and orientation of the image of "A" within the base image. The base or terrain image is the source material of the scene displayed on the CRT, i.e., photographs, CGI, and so forth.

The more difficult problem is to determine whether an overlayed object will be obscured or occluded, visually, by either the terrain or another object within the

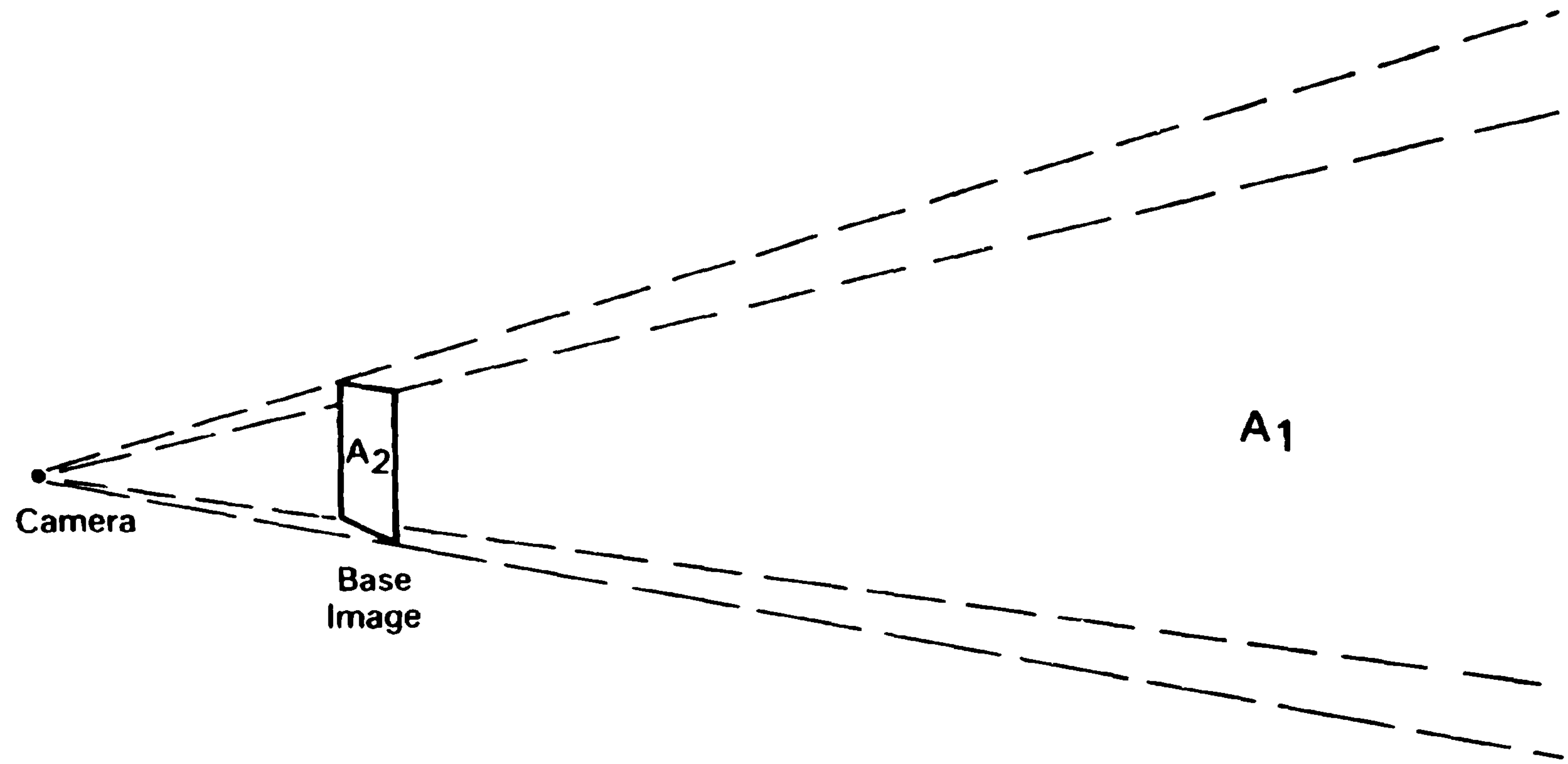

Figure 3-1

EXAMPLE OF OVERLAY SIZE AND POSITION DETERMINATION

(A_1 = OBJECT TO BE OVERLAYED UPON BASE IMAGE, A_2 = IMAGE OF OBJECT A_1 WITHIN BACKGROUND OR BASE IMAGE)

terrain image. Occlusion is a function of the portion of the base image that will be overlayed, and the position of the objects in the base image that are nearer or farther than the overlay image. The objects or pieces of objects that are not occluded should be fully displayed.

The base image does not, in itself, contain readily accessible information concerning the distance to objects. It is simply a planar projection of the three dimensional world, and as such, offers no obvious solution to the problem of occlusion.

Instead, the answer lies in constructing a second image to accompany the base image. This is called a range image. Point for point it must align with the base image; but rather than containing lighting information, it contains information about the distance to the objects that reflected the light.

Figure 3-2 characterizes the distinction between these two images. Here, both images are represented in the way a computer or a video signal represents the image. For each picture element (pixel) a magnitude is stored characterizing either the lighting intensity (base image) or the distance to the reflecting object (range image).

Although the present example envisions only three levels of magnitude for either image, a video signal will permit more than one hundred levels per pixel. This would provide a basis for discriminating every 30m of range, if the limit of visibility were 3km. Moreover, since close objects are likely to require finer range discrimination, a nonlinear scaling of ranges would improve the capability.

2	2	2	2	1	1	2	2	2	2
2	2	2	2	1	1	2	2	2	2
2	2	2	2	1	1	2	2	2	2
2	2	2	2	1	1	2	2	2	2
3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3
2	2	2	2	1	1	2	2	2	2
2	2	2	2	1	1	2	2	2	2
2	2	2	2	1	1	2	2	2	2
2	2	2	2	1	1	2	2	2	2

1 = Dark
3 = Bright

BASE IMAGE

3	3	3	3	2	2	3	3	3	3
3	3	3	3	2	2	3	3	3	3
3	3	3	3	2	2	3	3	3	3
3	3	3	3	2	2	3	3	3	3
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
3	3	3	3	2	2	3	3	3	3
3	3	3	3	2	2	3	3	3	3
3	3	3	3	2	2	3	3	3	3
3	3	3	3	2	2	3	3	3	3

1 = Near
3 = Far

RANGE IMAGE

Figure 3-2

EXAMPLE OF THE DISTINCTION BETWEEN THE BASE IMAGE AND THE RANGE IMAGE (PICTURE = HORIZONTAL WHITE BAR IN FRONT OF VERTICAL BLACK BAR ON GREY FIELD)

Figure 3-3 illustrates how a range image can be used to compute occlusion. First a threshhold is applied to the range image at the distance of the object that is to be displayed. This identifies regions of the base image that are nearer and farther than the overlayed object. Next, those portions of the overlay image that overlap a "near" region of the base image are deleted. Finally, the occluded object image is overlayed on top of the base image.

3.1.1.2 <u>Alternatives for achieving range images</u> - The range image technique works and is used by CGI houses to build their imagery. The following four alternatives were outlined:

(1) Laser range imaging;
(2) Stereo computation;
(3) Solid photography; and
(4) Computer generation.

In the Research Plan, we took the position that stereo computation (Hannah, 1974 and Gennery, 1980) and solid photography are time consuming, expensive, and not likely to be accurate enough for an ATR application. No new information has arrived to change this position.

Laser range images have been described by a number of authors (Lamberts, 1975, Milgram and Bjorkland, 1980; and Lynd). Laser range finders are readily available products for use in detecting the range from point A (location of the range finder) to point B in the real world. However systems for developing a range image from point A for an entire scene are few in number, highly proprietary, and not generally available as a product. Such laser range image systems use the concept of a laser range finder, integrated

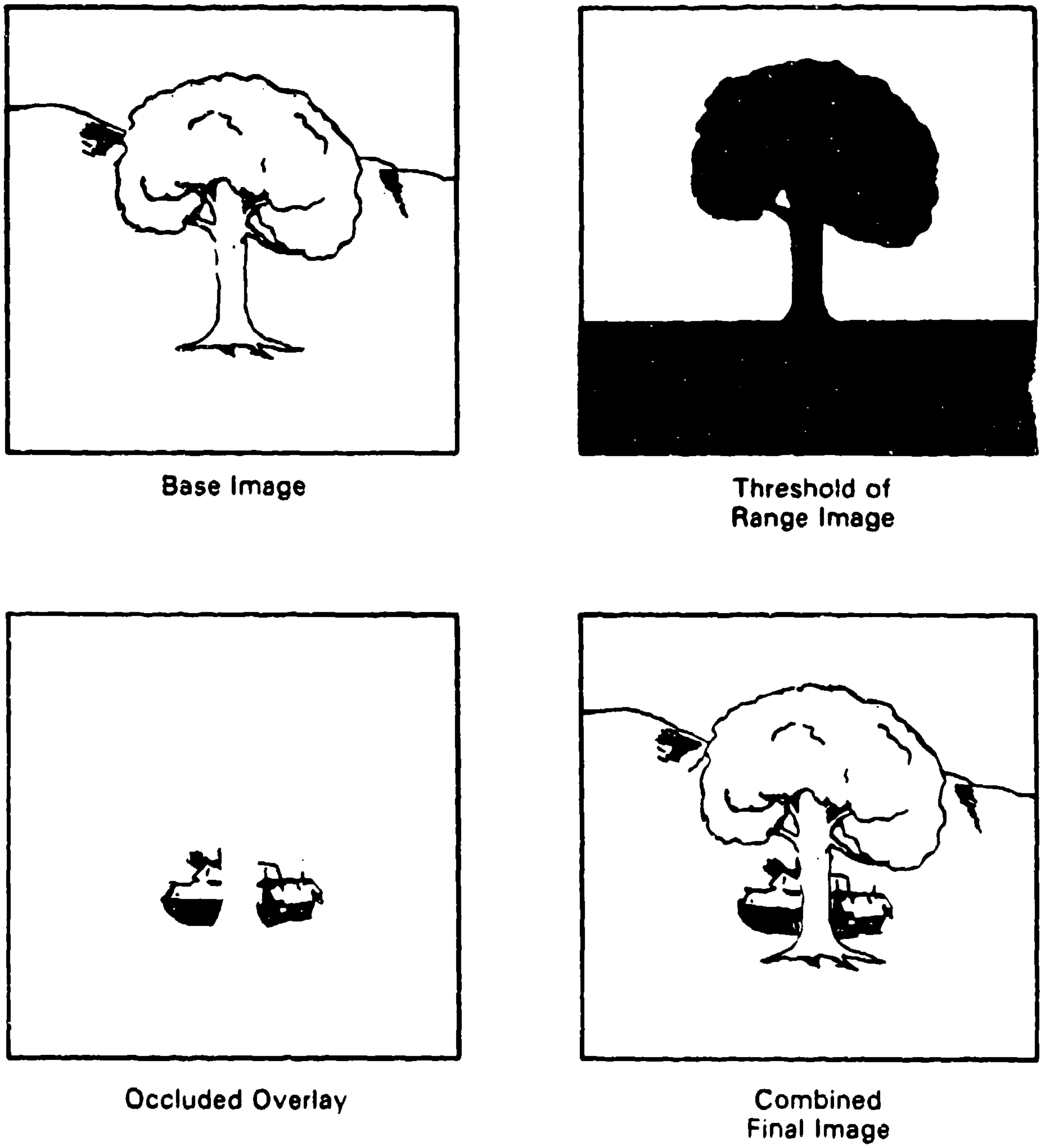

Figure 3-3
ILLUSTRATION OF THE COMPUTATION OF OCCLUSION

with an automated scanning capability, and augmented by algorithms for processing geometrical properties of the image for data enhancement. Proprietary systems exist or are under development at companies such as Hughes Research Laboratory, Ford Aerospace, Raytheon, General Dynamics, and the Environmental Research Institute of Michigan (ERIM).

The two major problems that must be overcome before this technology will be useful to an ATR project are:

(1) The range image resolution is poor. Currently, 8 bits are available for the range spectrum (0-3km for ATR). The notion of "Ambiguous Range" is currently being used to enhance this resolution by recycling the 8 bits within specified range intervals. Computations within this ambiguous range data base are then used to determine to which range cycle each point belongs.

(2) The systems have not been ruggedized. At this point in time it is unlikely that the systems can be moved from point A to point B in the ATR terrain grid and function reliably.

However, researchers in this field foresee dramatic improvements for these systems in the next two to five years, so we recommend that ATR researchers stay abreast of this field.

As was mentioned at the outset of this section, range images exist for CGI terrain representations. Our conclusion is that the computer generation of range images for source material from either the real world or terrain boards is likely to be expensive and somewhat inaccurate, but is the only feasible alternative if other than CGI terrain imagery is chosen.

3.1.2 Terrain boards - The authors conducted a search to identify existing U.S. Army terrain boards located in the United States. A variety of terrain boards were identified, varying in sophistication from game boards used in the school at Fort Leavenworth for Dunn Kempf play, to the large, highly-detailed, wall-mounted architectural models at Fort Rucker which were designed to provide the visual displays for helicopter simulators. However, due to the highly exacting demands of the ATR project on visual representation of terrain, it was not necessary to investigate, in depth, all the various terrain boards identified in the search. But rather, it was decided that only the most highly detailed and textured boards located at Fort Rucker should be examined. The function of this investigation, then, was not to conduct an exhaustive analysis of terrain boards in the U.S. Army, but rather to assess the feasibility of using terrain boards as a basis for representing terrain from a ground perspective.

3.1.2.1 Terrain boards at Fort Rucker - The major purpose of observing the facilities at Fort Rucker was two-fold: (1) to observe terrain board representations of visual travel over terrain, and (2) to make a nonempirical comparison of visual travel over terrain represented by CGI versus terrain represented by specially constructed terrain boards.

Terrain board characteristics - Three major terrain boards were observed at Fort Rucker. Two were constructed on a scale of 1,500:1 and one on a scale of 1,000:1. All the boards represented the terrain of south central Alabama--lightly wooded rolling hills and small plains. In addition, a small portion of the 1,000:1 board contained a segment of Middle Eastern terrain represented on a scale of 500:1.

The 1,000:1 board contained individually mounted trees, fences, buildings, and so on and approached the scale needed for ATR. However, it is unlikely that this board would provide the type of imagery demanded by ATR. A terrain board representing a 500:1 scale might work, however, the most appropriate scale for ATR imagery would be 300:1.

In addition to problems of scale with the terrain boards at Fort Rucker, depth of field also appears to be a problem when representing ground-level views on film. A ground perspective leaves a large strip of the display out of focus. The Scheimpflug correction serves to attenuate this problem, although even after the correction is applied portions of the image remain noticably unfocused. Resolution of this issue, however, appears to be feasible. First, in the simulators at Fort Rucker terrain board imagery was created in real time. Because ATR does not demand real-time imagery, it may be possible to fine-tune the Scheimpflug correction to generate a wider band of focus. Secondly, the Link Flight Division of the Singer Company has developed a laser image generator that can be interfaced with their model boards, allowing for markedly improved depth of field, especially for 300:1 boards.

Finally, it should be noted that it is not within the scope of this project to construct a terrain board. Nor is it feasible to construct the stimulus capturing system for an already existing terrain board. If the ATR project is to use imagery from a terrain board, it must be collected from an existing system, similar in sophistication to the system at Fort Rucker.

A comparison of CGI and terrain board imagery - The visual displays examined in this analysis were part of Fort Rucker's sophisticated flight trainers used to simulate the U.S. Army's Blackhawk helicopter (UH6OFS). The two trainers--one with a computer generated visual display and one with an actual terrain board display--provided identical information to the user, with the exception of the visual information. Both provide the same actual movement, auditory information, and instrumentation. The following comparative observations were made concerning the simulation experiences of the two flight trainers:

- The most critical deficiency in the CGI display relative to the terrain board display concerns the lack of ground texture. In the CGI display, the ground appears as a smooth green or brown field, providing little information about velocity, self-location, distances to depicted objects, or steepness of hills. In fact, the psychological literature on perception of visual displays strongly supports the notion that texture information is critical in accurately perceiving <u>velocity</u> (e.g., Gibson, 1966; 1979), <u>egoloco motion</u> (e.g., Gibson, Olum, and Rosenblatt, 1955, Warren, 1976), <u>distances</u> (e.g., Gibson and Bergman, 1954; Gibson, Bergman, and Purdy, 1955; Purdy and Gibson, 1955; Wohlwill, 1962; Newman, 1971; Gibson, 1979), and <u>slant</u> (e.g., Gibson, 1950; 1966; 1979; Flock and Moscatelli, 1964; Eriksson, 1964; Kraft and Winnick, 1967; Newman, Whinham, and MacRae, 1973).

- Although the trees and other ground clutter are more detailed and realistic in terrain board imagery than in CGI, the prototype, diamond-shaped trees, and the abstract box-like buildings do not significantly interfere with visually guided behavior. That is, the cartoon-like quality of the objects may affect the face validity of the system (Nunnally, 1967), but not the critical behavior of the trainee.

- o Depth of field is narrower in terrain board imagery than in CGI, where it can be uniformly generated.

- o In both the CGI and the terrain board simulators, the presence of vibration, realistic auditory information, and side views added immeasurably to the experience of movement over the depicted terrain.

3.1.2.2 Terrain boards outside the United States - The Link Flight Simulation Division of the Singer Company has constructed several terrain board simulation facilities overseas. Terrain boards representing a scale of 300:1 are located in England, Holland, and the Middle East. Of particular interest is the terrain board facility in Bovingdon, England. The board is properly scaled, sufficiently detailed, and extensive enough to provide the type of imagery and coverage demanded by the ATR project. If access could be gained, and if there is a sufficient object model of this board, it could represent a feasible option for capturing the data necessary for ATR. However, it is being remodeled, and will not be completed until 1983. A similar terrain board is being constructed in Lansing, England and is scheduled to be available in January 1983. Finally, the terrain board in Holland is actively used for training.

3.1.3 Computer-generated imagery - Flight simulators have been using CGI as a source of imagery for some time. However, the CGI used in flight simulators generally does not have the detailed texture that is needed to provide the source information for an ATR system. Flight simulators have required real-time CGI in the past, and it is this real-time requirement that constrains the image complexity to an unacceptably poor level of texture. There are numerous non-real-time CGI systems in development and use for many purposes that have higher levels of quality, including ground texture.

As part of this research, we contacted the people at the organizations listed in Table 3-1 to ascertain the availability of high quality, non-real-time CGI. While most CGI has been generated via "edge models" in the past, current high quality CGI is produced via parabolic arcs, combinatorial shapes (e.g., spheres, cones, and boxes), and fractals (Mandelbrot, 1977). The range image information described in Section 3.1.1 is available for all CGI since it is computed in the process of setting up the terrain data base.

Current systems provide ground texture and shadows on rolling terrain, planted with vegetation, and marked by roads and paths.

The major problem in utilizing this high quality CGI is the cost. Cost can be broken into two components: the cost of defining the terrain (data base), and the cost of filming the terrain. The rough estimates we received from several of the contacts listed in Table 3-1 ranged from $350,000 to $1,000,000 for a ¼ x ½ mile region with a 50 foot grid and 16 view directions (25,088 pictures). This is well beyond the scope of the current [illegible] project.

3.1.4 <u>Panoramic images</u> - Panoramic imagery is a rather attractive alternative to discrete images. Rather than recording 360° of view in multiple planar images, a panoramic image records 360° in one continuous strip. This introduces some distortion into the image, but such distortion can be rectified by displaying the image onto a curved surface. The advantage of panoramic imagery lies in the ability to provide

Dr. Martin Cohen
Mathematical Applications Group, Inc.
3 Westchester Plaza
Elmsford, New York 10523

Mr. Gary Demos
Digital Productions
3416 South La Cienega Blvd.
Los Angeles, California 90016

Mr. Geoffrey Y. Gardner
Grumman Aerospace Corporation
111 Stewart Avenue
Bethpage, New York 11714
M/S A08-035

Dr. Kenneth C. Knowlton
SRI, International
Room EL393
333 Ravenwood Avenue
Manlo Park, California 94025

Mr. Bruce Laskin
Computer Grapnics, Inc.
P.O. Box 170
Wheatley Road
Old Westbury, New York 11568

Ms. Winnie Lee
Boeing Aerospace Corporation
P.O. Box 3707
Seattle, Washington 98124
M/S 9F-40

Mr. Alfred Senaughti
Information International, Inc.
5415 Jandy Place
Los Angeles, California 90066

Mr. Aldy Ray Smith
Lucas Films
P.O. Box 2009
San Rafael, California 94912

Table 3-1

CONTACTS FOR CGI QUALITY/AVAILABILITY

very smooth pivots, while reducing the number of videodisc frames required for each grid location.

Figure 3-4 depicts the encoding scheme for placing panoramic images onto two videodisc frames. The original panoramic strip is divided into four overlapping sections each containing 180° of view. The overlap is defined in such a way that the left-hand portion of each section corresponds exactly to the right-hand portion of another. At the time of display, only one quarter of a frame is actually presented. Within a section, continuous scanning is accomplished by simply moving this quarter-frame window across the section. At the end of a section, scanning is continued by placing the window on the corresponding region of the neighboring section.

The panoramic approach to continuous scanning is not without its own difficulties. First, it sacrifices some image resolution to achieve the necessary encoding. Second, it can require special equipment to film and recode the panoramic strips. Finally, it will demand special display equipment (e.g. frame buffer or special effects generator) to allow the continuous scanning of the partial frames.

Our investigation into panoramic imagery was aimed primarily towards its application to terrain boards and CGI. We know its availability for the real world, but terrain boards and CGI seem to be better sources for the terrain imagery. Based upon conversations with personnel from Ft. Rucker and the Singer Company, we feel confident that a camera and processor that will produce panoramic imagery from terrain boards has not been developed. There is no significant reason why this cannot be done, other than time and expense. However, the ATR project does not warrant such a development at this time.

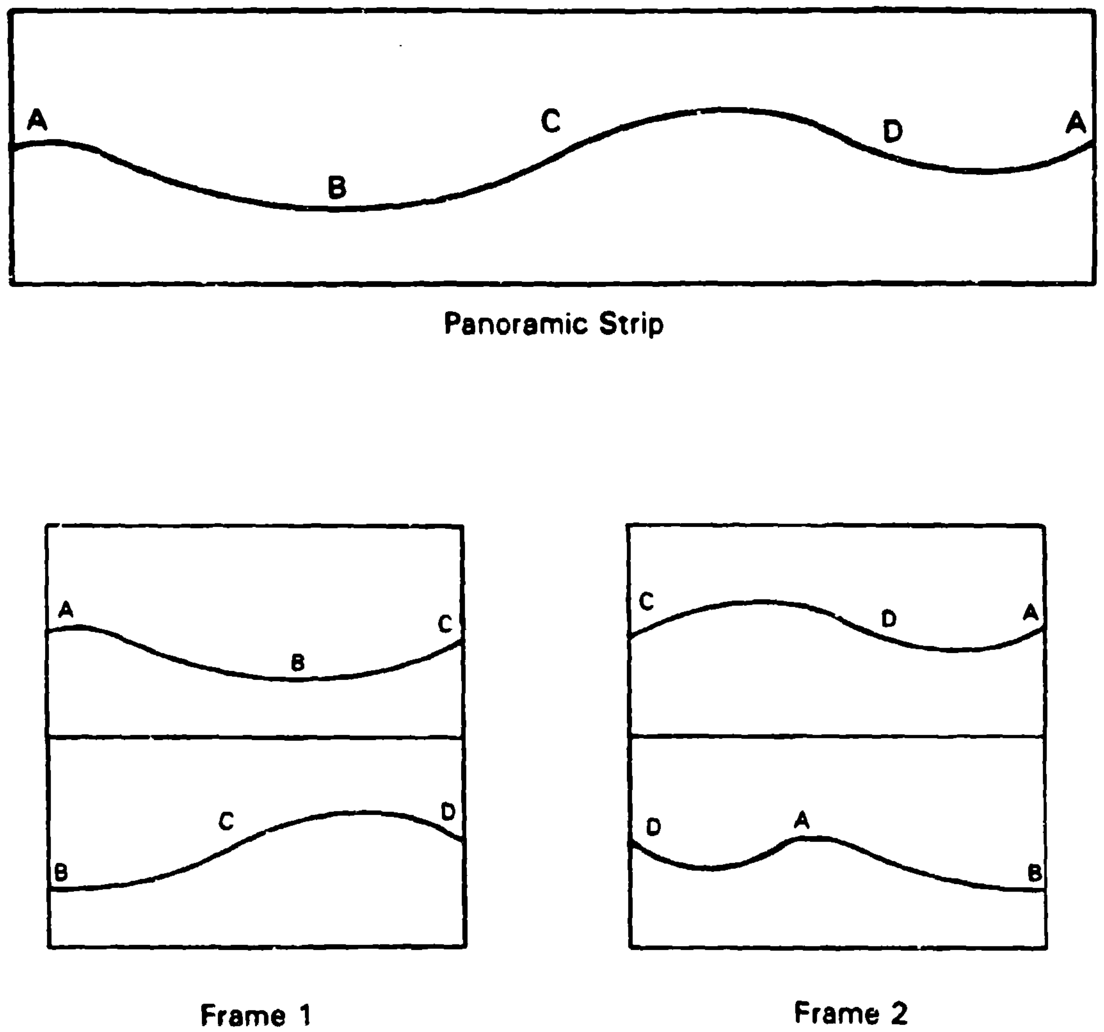

Figure 3-4
EXAMPLE OF
ENCODING SCHEME FOR PANORAMIC PHOTOGRAPHY

Finally, our search for a panoramic capability for CGI proved inconclusive. We were not able to find an existing application, but need to check additional sources.

3.1.5 <u>Image rectification</u> - As an alternative to panoramic imagery, image rectification can be used to produce a 360° view from four videodisc frames. The image rectification concept uses four discrete 90° Angle of View frames and a special display system for combining frames to provide interpolated images (Figure 3-5). In essence, a periscope is placed into the center of a box and discrete views are displayed on the walls of the box. Then, by pivoting the periscope, a 360° view can be reconstructed. Since this system compensates for the planar perspectives of the discrete images, it is called an image rectification device.

The image rectification device provides smooth turning with only a few discrete images, but it has certain disadvantages. First, it is mechanical and potentially more prone to error than an electronic technique. Second, it requires a hardware configuration that is expensive and needs further development. Finally, this device requires that the user view the displays through binoculars.

Additional research into the costs and requirements to implement the image rectification concept was not undertaken since this concept requires that views from the same grid be stored on different videodiscs. This storage design is contrary to that presented in Section 2.0. The design in Section 2.0 was chosen to enhance a wide range of speeds for travel and maintain high coverage.

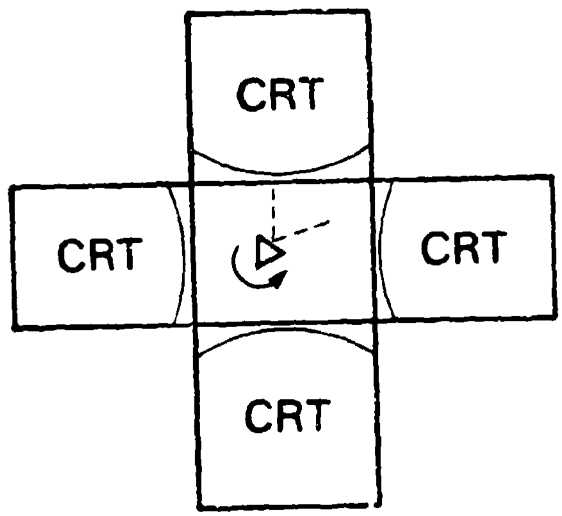

Figure 3-5
EXAMPLE OF IMAGE RECTIFICATION HARDWARE

3.2 Simulating Intervisibility

Section 3.1.1 (range images) provides a discussion of the intervisibility calculations needed to provide a realistic overlay of dynamic imagery for the user. This overlay of imagery is the most difficult technical problem to be solved in the development of an ATR system. Not only must a range image be produced that is an accurate representation of the terrain source environment, but good registration between the terrain image and the range image must be achieved so that points in the terrain line up with the correct pixels in the range image.

Summarizing our research on this topic over the past year:

(1) Range images for CGI source terrain are readily available, with perfect registration to the source image. (The major concerns about CGI are (1) the texture quality of the terrain image for conveying a sense of movement, distance, and terrain features to the user and (2) the cost of producing the imagery.)

(2) Presently, range images generated from terrain boards should be more accurate and have better registration than range images that would be obtained from the real world. This assertion is based on the likelihood that a terrain board is developed from a blueprint or an object model, or both. It is also true that the real world has a degree of randomness that could not be factored into a terrain board, making the board more regular and easily mapped. Furthermore, a range image of the real world is virtually impossible until a laser range imaging system is

available (at least two years away). Finally, there are factors working in the real world, such as weather and people in an environment that cannot be controlled; factors that would not occur when capturing base and terrain images from a terrain board.

The impacts on the ATR system configuration of having a range image for each terrain image are twofold: (1) the number of videodiscs/players doubles, and (2) a frame buffer with four frames is required. Given the grid allocation scheme discussed in Section 2.4, and the fact that the terrain and range images cannot be retrieved from the same videodisc and achieve the Minimum Presentation Times described in Section 2.0, there will have to be a duplicate set of videodiscs/players of the terrain for the range images.

In order to overlay dynamic symbols on the terrain image, four frames must be available within a frame buffer: one for the terrain image, the second for the symbol (e.g., tank), the third for the range image, and the fourth for the composite image of the first three. The DeAnza Image Array Processor is an example of hardware that possesses the capability to perform these operations (as well as many others) and costs about $60,000. This type of frame buffer can be substituted for that shown in Figure 2-14.

4.0 CONCLUSIONS AND RECOMMENDATIONS

On the basis of the research conducted over the past year, we conclude that the development of an ATR system for simulating travel over open terrain is feasible within the current state-of-the-art of computers and videodiscs. There are, however, several uncertainties concerning the quality and cost of such an ATR system still remaining.

The results of the empirical research (Kraft, Patterson and Mitchell, 1982) provided valuable information concerning the formatting of visual material. The technological research reported in sections 2.0 and 3.0 provided the necessary information to allow for the examination of tradeoffs between the ATR system parameters and to develop a definitive system design. The significant conclusions of this effort are as follows:

(1) The tradeoff between the system performance parameters of Coverage and Minimum Travel Speed is critical. Increases in Coverage increases require larger grids for travel jumps, while smaller grids are needed to decrease Minimum Travel Speed (assuming that a terrain image must be updated at least once every second to convey a sense of travel to the user).

(2) Grids must be allocated to multiple videodiscs in square block patterns in order to achieve a reasonable Minimum Presentation Time for each grid's terrain image and to anticipate changes in travel direction by the user. Our recommendation, at this time, is a nine videodisc player system, resulting

in a three by three block pattern that is repeated over the entire grid.

(3) Pivots and scanning could best be represented by using a panoramic (360°) image. However, the equipment needed to create panoramic photographs of a terrain board does not exist; and a range image of actual terrain is not practical for other reasons. Therefore, pivots and scans, at present, must be represented through the use of discrete images. If the user is to be allowed to scan 360° at will, from any point to any other point, 540° of sequential coverage on the videodisc is required because of the constraints implied in the use of the videodisc player "step" mode.

(4) The overlay of dynamic symbols (e.g., tanks) on the terrain image requires a range image in registration with the terrain image. Currently, the only means for creating such a range image is to build a CGI-like object model of the terrain. The ATR system requirement to achieve this symbol overlay is a frame buffer with four frames; one each for the terrain image, range image, symbol(s), and composite image.

The three major uncertainties which remain can only be resolved through research on a prototype system. Three areas of concern are:

(1) user acceptance which will be determined by the responsiveness of the system to the demands of open field travel;

(2) the perception of motion and the accuracy of distance estimation provided by CGI, should CGI be chosen for the terrain imagery; and

(3) the accuracy with which dynamic symbols can be overlayed on non-CGI terrain.

User acceptance could be determined by (1) creating a sample of each of the terrain images, i.e., photographs of actual terrain, of a terrain board, or CGI; (2) mastering a videodisc of these images; and (3) allowing tank commanders to interact with, and rate the system. If adequate terrain coverage can be generated, it would also be possible to provide a behavioral evaluation of the efficacy of an interactive system, and to acquire empirical data on the relative effectiveness of visual travel for the three image bases.

It seems likely that the implementation of CGI overlays must await the outcome of the research outlined above. Table 4-1 lays out the key attributes for choosing the three sources of terrain imagery. Dynamic overlays are a certainty if CGI is used; probable if terrain boards are the source of the base image; and highly unlikely if photographs of actual terrain are chosen. If an adequate sense of travel can be produced with CGI, dynamic overlays then become a viable option and should be considered for the final year of this project.

Source Material	Information Content	Quality/Control During Capture Process	Range Image Availability	Cost
Real World	High	Low	Low	High $
Terrain Boards	Medium	Medium	Low-Medium	Low $
CGI	Low	High	High	Medium-High $

Table 4-1

EVALUATION OF TERRAIN SOURCES

References

Eriksson, E. S. Monocular slant perception and the texture gradient concept. Scandinavian Journal of Psychology, 1964, 5, 124-128.

Flock, H. R., & Moscatelli, A. Variables of surface texture and accuracy of space perception. Perceptual & Motor Skills, 1964, 19, 327-334.

Gennery, Donald B. Modelling the environment of an exploring vehicle by means of stereo vision (AD A091081). Alexandria, Va.: Defense Technical Information Center, November 1980.

Gibson, E. J., & Bergman, R. The effect of training on absolute estimation of distance over the ground. Journal of Experimental Psychology, 1954, 48, 473-482.

Gibson, E. J., Bergman, R., & Purdy, J. The effect of prior training with a scale of distance on absolute and relative judgments of distance over ground. Journal of Experimental Psychology, 1955, 50, 97-105.

Gibson, J. J. The perception of the visual world. Boston: Houghton Mifflin, 1950.

Gibson, J. J. The senses considered as perceptual systems. Boston: Houghton Mifflin, 1966.

Gibson, J. J. The ecological approach to visual perception. Boston: Houghton Mifflin, 1979.

Gibson, J. J., Olum, P., & Rosenblatt, F. Parallax and perspective during aircraft landings. American Journal of Psychology, 1955, 68, 372-385.

Hannah, M. J. Computer matching of areas in stereo images (AD 786720). Alexandria, Va.: Defense Technical Information Center, July 1974.

Kraft, A. L., & Winnick, W. A. The effect of pattern and texture gradient on slant and shape judgments. Perception and Psychophysics, 1967, 2, 141-147.

Kraft, R. N., Patterson, J. F., & Mitchell, N. Psychological research on advanced terrain representation: Formatting the visual material (Interim Report DDI/PR 82-9-331). McLean, Va.: Decisions and Designs, Inc., October, 1982.

Lamberts, C. W. A long range scanned laser active imaging system (Report No. PHL 1975-52). The Hague, The Netherlands: National Defense Research Organization TNO, December 1975.

Lippman, A. Movie maps: An application of the optical videodisc to computer graphics. In SIGGRAPH 1980 Conference Proceedings. Seventh Annual Conference on Computer Graphics and Interactive Techniques. Also published in Computer Graphics, 1980, 14, 32-42.

Lynch, D. K. Range image enhancement via one-dimensional spatial filtering. Computer Graphics and Image Processing, 1981, 15, 194-200.

Mandelbrot, B. B. Fractals: Form, chance, and dimension. San Francisco: W. H. Freeman, 1977.

Milgram, D. L., & Bjorklund, C. M. Range image processing: Planar surface extraction. In Proceedings of the 5th International Conference on Pattern Recognition, pp. 912-919. New York: IEEE, 1980.

Newman, C. V. The influence of visual texture density gradients on relative distance judgments. Quarterly Journal of Experimental Psychology, 1971, 23, 225-233.

Newman, C. V., Whinham, E. A., & MacRae, A. W. The influence of texture on judgments of slant and relative distance in a picture with suggested depth. Perception and Psychophysics, 1973, 14(2), 280-284.

Nunnally, J. C. Psychometric theory. New York: McGraw-Hill, 1967.

Patterson, J. F., Kraft, R. N., & Buede, D. M. Issues to investigate for implementing an advanced terrain representation system (Research Plan RP 82-1-331). McLean, Va.: Decisions and Designs, Inc., February 1982.

Purdy, J., & Gibson, E. J. Distance judgment by the method of fractionation. Journal of Experimental Psychology, 1955, 50, 374-380.

Warren, R. The perception of ego motion. Journal of Experimental Psychology: Human Perception and Performance, 1976, 2, 448-456.

Wohlwill, J. F. The perspective illusion: Perceived size and distance in fields varying in suggested depth, in children and adults. Journal of Experimental Psychology, 1962, 64, 300-310.

APPENDIX A

Figure A-1 displays the grid allocation for the use of nine videodisc/players, allowing three videodisc players to search in parallel for every Travel Direction depicted in Figures 2-1a and 2-1b (d-8, 16, and 24). Unlike the layout using four videodiscs/players, this design permits travel on every other grid in the north/south, east/west, and 45° diagonal while maintaining three parallel videodisc players. For example, while traveling north from the lower-left corner of Figure A-2, the user can be presented images from grids 4,1,7, 4,1,7,3,1,7,3,1,... (every grid) or from grids 4,7,1,3,7,1,...- (every other grid). This flexibility for using either a sequence of adjacent grids or a sequence of odd-numbered grids in the north/south and east/west directions permits the Maximum Jump in those directions to be varied from g to 2g. Likewise, the Maximum Jump can be varied from 1.4g to 2.8g. These higher Maximum Jumps (2g to 2.8g) are more nearly equal to the highest Maximum Jump for 16 and 24 Travel Directions. This will be important in maintaining and even travel speed when changing directions of travel. Finally, with a videodisc random access search time equal to one second and no frame buffer, three parallel videodisc players will yield a Minimum Presentation of 0.5 seconds (two searching while one is displaying). Including the frame buffer for display drops t to 0.33 seconds.

Figures A-2 and A-3 show the grid allocations for the use of 16 and 25 videodiscs/players, respectively. The lines emanating from the lower left-hand corner of Figure A-2 show that four videodisc players can search in parallel for every Travel Direction discussed previously, cutting t to 0.33 without a frame buffer and to 0.25 with a frame buffer. (Note: the

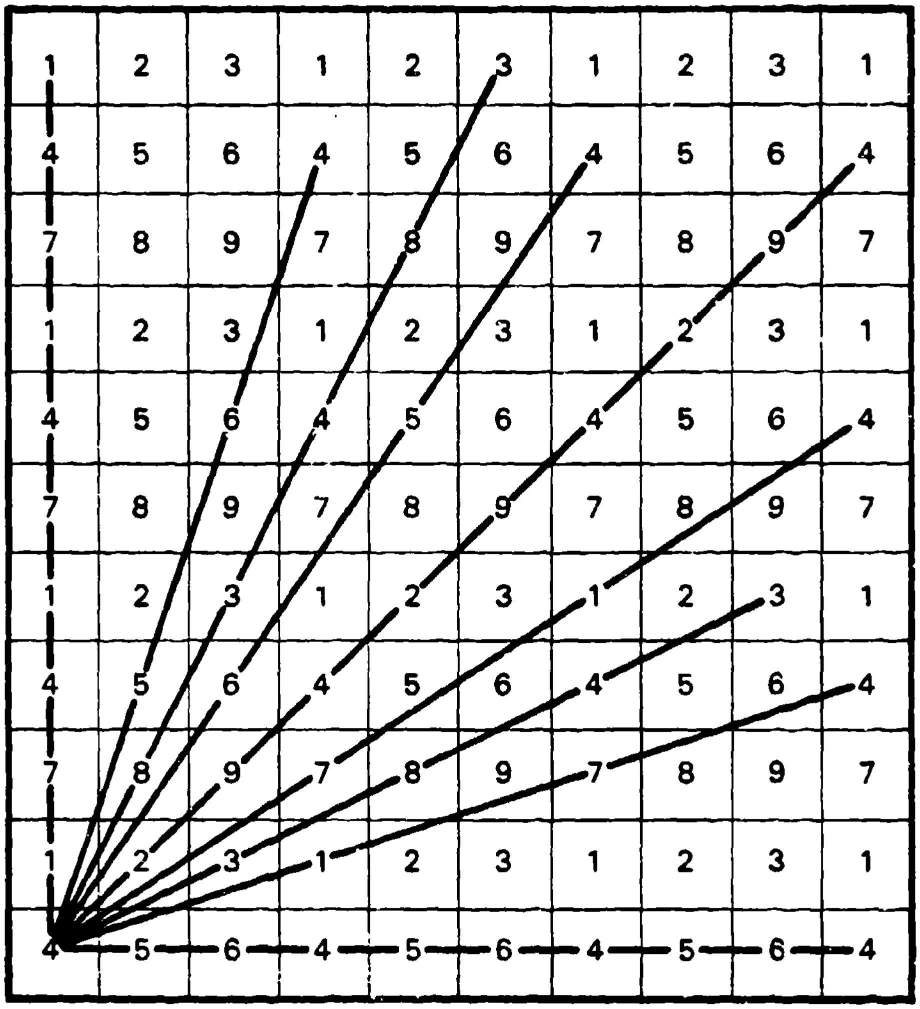

Figure A-1
GRID ALLOCATION FOR NINE VIDEODISCS/PLAYERS

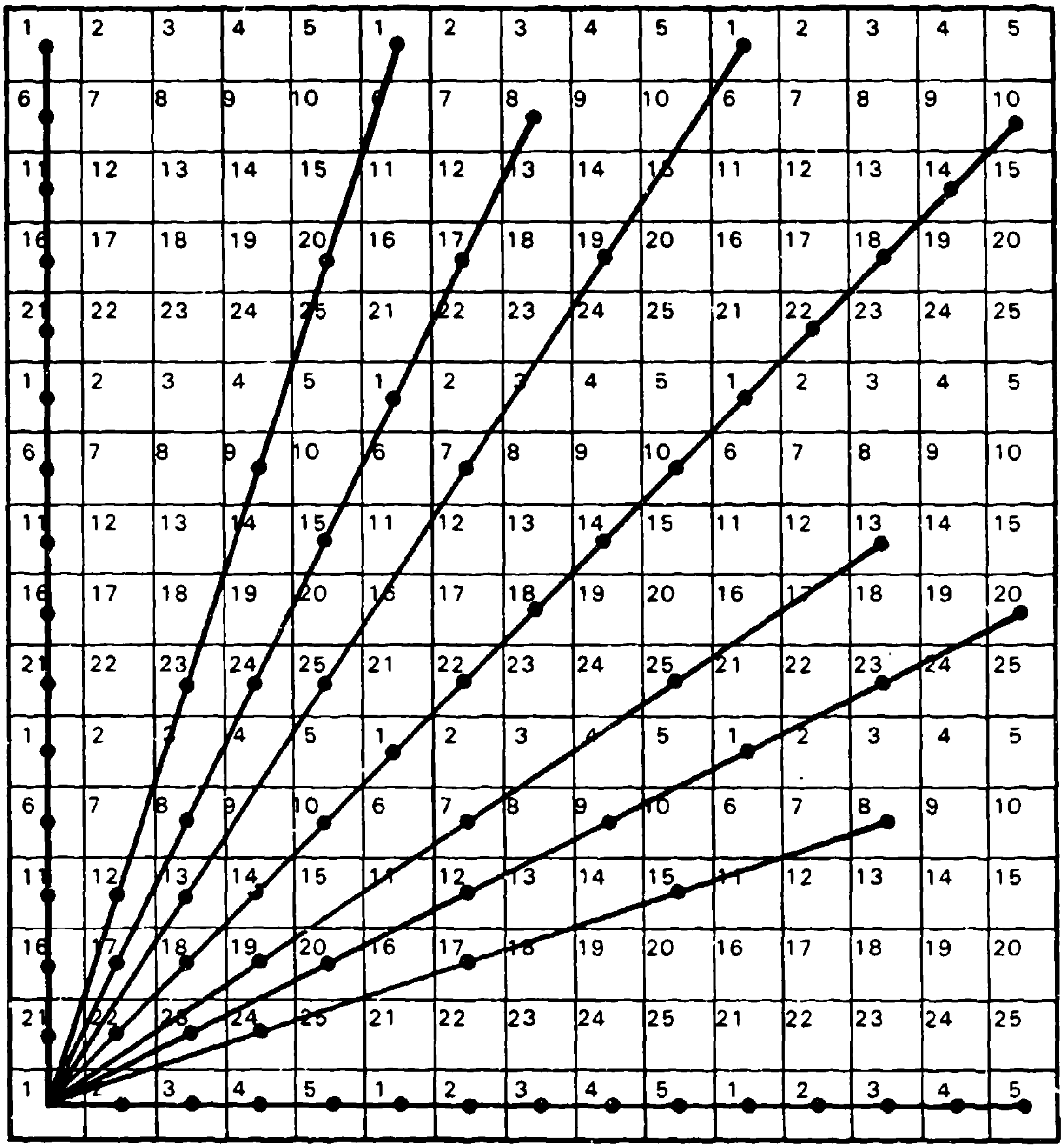

Figure A-2

GRID ALLOCATION FOR TWENTY FIVE VIDEODISCS/PLAYERS

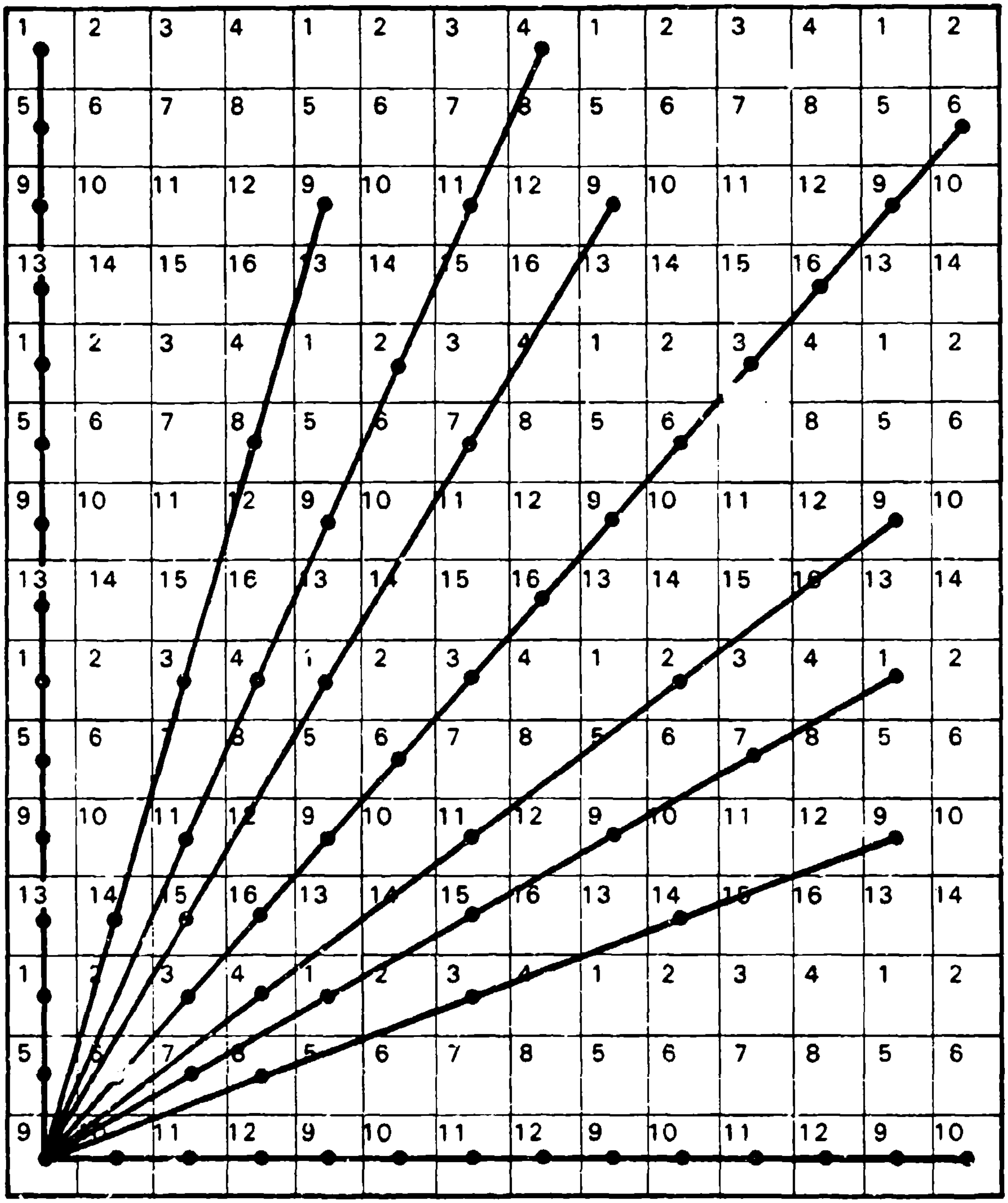

Figure A-3
GRID ALLOCATION FOR SIXTEEN VIDEODISCS/PLAYERS

north/south, east/west, and 45° diagonal travel directions require the display of adjacent grids to achieve the four parallel videodisc players.)

The use of twenty-five parallel videodisc players is depicted for all travel directions in Figure A-3. In this case, the flexibility exists for either using adjacent grids or skipping every other grid when traveling in the north/south, east/west, or 45° diagonal directions. This flexibility increases the Maximum Speed in these directions. The Minimum Presentation Time with a five parallel videodisc configurati. is 0.25 without a frame buffer and 0.20 with one.

APPENDIX B

BLACK AND WHITE CONFIGURATION WITH FRAME BUFFER

Components	$ Cost
Controls	$ 1,000
Microcomputer	$10,000
Nine videodisc players	$22,500
Switch	$ 5,000-6,000
Frame Buffer	$ 8,000-10,000
B&W Monitor	$ 500
TOTAL	$47,000-55,000

COLOR CONFIGURATION WITH FRAME BUFFER

Components	$ Cost
Controls	$ 1,000
Microcomputer	$10,000-15,000
Nine videodisc players	$22,500
Switch	$ 5,000-6,000
NTSC/RGB converter	$ 2,000
Frame Buffer	$20,000-25,000
RGB color monitor	$ 2,500
TOTAL	$63,000-74,000