COGNITIVE SPACE IN THE INTERACTIVE MOVIE MAP:

AN INVESTIGATION OF SPATIAL LEARNING IN VIRTUAL ENVIRONMENTS

by

ROBERT MOHL

B.S., Massachusetts Institute of Technology
1971

Submitted to the Department of
Architecture in Partial Fulfillment
of the requirements for the
Degree of

Doctor of Philosophy in
Education and Media Technology

at the
Massachusetts Institute of Technology
February 1982

Signature of Author . . . .
Department of Architecture
September 17, 1981

Certified by . . . .
Nicholas Negroponte
Thesis Supervisor

Accepted by . . . .
Stanford Anderson
Chairman, Departmental Ph.D. Committee

Cognitive Space in the Interactive Movie Map: An Investigation of Spatial Learning in Virtual Environments

by

Robert Mohl

Submitted to the Department of Architecture on September 17, 1981 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Education and Media Technology.

ABSTRACT

A working system for movement through the space of a virtual environment under full interactive control has been developed, implemented, and evaluated. This interactive movie map is an optical video-disc-based system under computer control. Users are able to learn about an unfamiliar urban space by 'traveling' around at will through sequences of photographic footage and 'helicoptering' above dynamic aerial photos and reference maps.

The interactivity of the human-machine interface is achieved by providing simple but powerful controls in a multi-channel user station that includes pictures, sound, text, and graphics on a touch-sensitive television screen.

The effectiveness of the interactive movie map as an instrument for conveying spatial knowledge is demonstrated in experiments analyzing nature of the cognitive maps which subjects construct. Movie map training leads to superior way-finding competence in the real setting. Significant differences, and certain advantages, result when the users have access to personalized reference maps in addition to street-level travel.

Key words: maps, navigational aids, optical video disc, interactive video, man-machine interface, computer graphics.

Thesis Supervisor: Nicholas Negroponte
Title: Professor of Architecture

The work reported herein has been supported by the Cybernetics Technology Division of the Defense Advanced Research Projects Agency, under Contract No. MDA-903-78-C-0039.

Project director: Andy Lippman
Principal investigator: Nicholas Negroponte

ACKNOWLEDGMENTS

A project of the magnitude of the interactive movie map was made possible through the contributions of a great many people.

First, I would like to express my appreciation of Nicholas Negroponte's singular vision in establishing the Architecture Machine Group and his leadership in launching a large variety of innovative projects, including the interactive movie map. I am also grateful for the financial and administrative support which he provided as chairman of my thesis committee.

I thank Andy Lippman for his role as director of the project. He coordinated the team and kept work moving forward by providing guidance, suggesting improvements and occasionally setting impossible deadlines. All were appreciated.

Serving as a constant source of intellectual inspiration were Judah Schwartz and Kristina Hooper. Judah, with the skill of a wise mentor, sprinkled our encounters with cogent thoughts about spatial knowledge that alternated between insight and challenge. Kristina approached problems with an enthusiasm that was contagious and an energy that was enviable. (I only wish I could think as fast as she can write.)

Gary Allen and Alex Siegel, with a vast experience in cognitive mapping research, helped guide me past a number of pitfalls in experimental design. Those which I did not see or still have not discovered are my responsibility alone. The courteous encouragement and pertinent references both provided rejuvenated my resolve several times along the way.

I would also like to thank Lynn Goldsmith for our fruitful discussions over lunch and Don Olivier and Bob Muller for their help in unraveling the mysteries of multidimensional scaling

If any readers are, and I encourage all to be, of the type who look at all the pictures but only glance at the text, then they will be able to appreciate the major contribution of Most Media Graphics. Christian Lischewski raised compulsive perfectionism to new heights in his detailed diagrams. Martha Leinroth matched those standards as the photographic print-maker. It is fortunate that I wasn't in their shoes, to listen to revisions I suggested after the fact. I could never have exhibited their patience and tolerance.

And for help in the inevitable eleventh-hour panic, I would like to acknowledge the assistance of Linda Terry and Marti Feerick in cutting, pasting, labeling, and numbering, and in prying my fingers from the key-board. Cynthia Ware provided indispensable service as the proofreader.

I am indebted to the band of students at the Architecture Machine Group who did much of the work of implementing the interactive movie map. They applied themselves with dedication, whether it meant programming through the night or abandoning classes for a week to film the streets of Aspen. First and foremost thanks go to Walter Bender, now a research associate. Also providing much appreciated help were Ken Carson, Mike Naimark, Scott Fisher, Dave Sheppard, Paul Heckbert, Steve Yelick, Howard Eglowstein, Bill Parker, Mark Stiffler, Paul Roush, and Peter Clay.

I am also grateful to staff and students at the Film Section and the Visible Language Workshop, especially Ron MacNeil and Rob Faught, for letting me use the x-y plotter as a filming platform.

And finally, of course, a special thanks must be extended to all the citizens of Aspen, who so graciously lent their town.

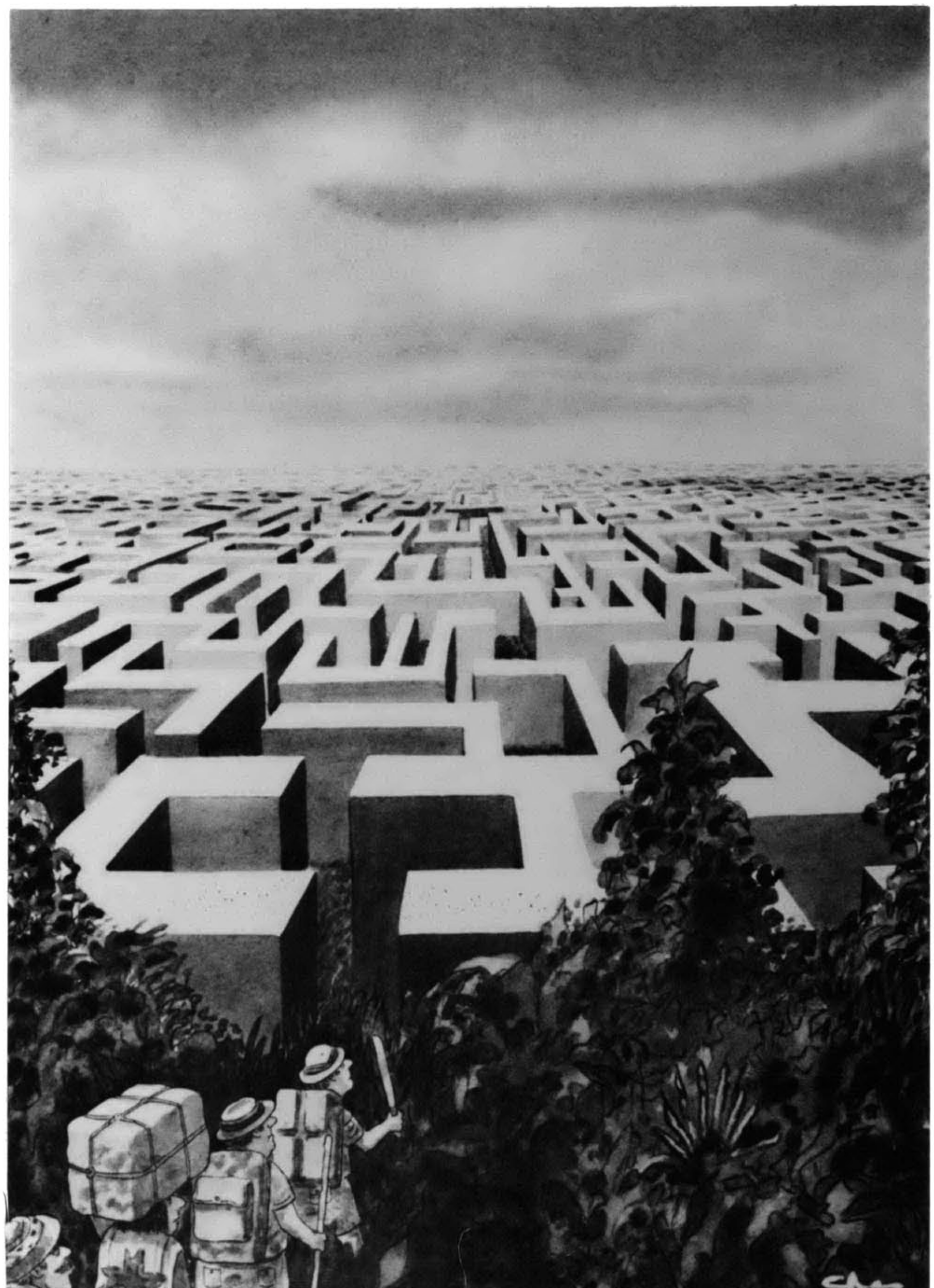
CONTENTS

CHAPTER 1

Introduction

The only way anyone ever manages to get from Logan Airport to M.I.T. on his own is to have made the trip before.

- Lippman, 1981



- Addams

For centuries newcomers to unfamiliar environments have had to rely almost exclusively on maps, diagrams, and verbal descriptions to find their way around. It is widely observed that people's competence in using such formats for representing spatial knowledge varies enormously. And yet many people who find it impossible to interpret the little red lines on their AAA maps the first time they go anywhere have little difficulty finding their way back a second time [Lippman, 1981]. In fact even proficient map users are likely to employ different spatial knowledge and strategies on a second visit. Part of the reason, presumably, is that first-hand experience in the veridical space provides the visitor with a different (and more directly accessible) internal model than is the case with abstract representations.

The premise of this document is that a novel form of spatial representation can be created which substitutes for the actual experience. This medium, called the interactive movie map, provides a mechanism for pre-experiencing an unfamiliar locale that allows the acquisition of spatial knowledge to take place in a meaningful, natural, and accurate way. Of course, it cannot measure up completely to the presence of being there. On the other hand, it provides some navigational aids which go beyond what is possible in the real environment.

1.1 Function

The interactive movie map is a computer-controlled, video-disc-based system which displays two representations of a virtual environment--as travel land and as map land. The user explores both worlds interactively.

In travel land the user "drives" around familiarizing him or herself* with the space by seeing real footage previously filmed on location. In this mode, called surrogate travel, he has complete control over speed, direction of travel, direction of view, and route selection, and even chooses a season of year in which to view the space. In addition the user can access pictures of individual building facades and see slide shows and facsimile data about each location.

In map land the user "helicopters" above aerial overviews of variable scale (and resolution) with customized navigational aids. Such aids include position pointers and route plotters.

With this dynamic overview the user is able to move his viewpoint above the surface (using track and zoom) and change the modality of representation by flipping between map and photo.

1.2 Objectives

This system addresses objectives in the following areas:

a) Mapping research.

The movie map is designed to contribute to research in the discipline of mapping by providing new ways of displaying spatial information and new means for acquiring spatial knowledge which

---

*Hereafter, use of third person singular pronouns is meant to refer to both genders. It is hoped that in exchange for less awkward syntax, the reader will refrain from reinforcing his sexual stereotypes.

are not possible with conventional alternatives. Inherent in the hard-copy medium of cartography, in street maps, for instance, are features which are abstract and static. While such standard formats serve well in some ways they clearly lack the ability to convey the dynamic, interactive, and pictorial nature of everyday experience in moving through real spaces. Furthermore, even when maps are used in conjunction with real-world experience, there is no precedent for making the content of the map automatically reflect the personal experience and needs of the user.

b) Spatial data access.

The interactive movie map provides a technique for storing and accessing a variety of data types within a three-dimensional space. The richness of the visual content of the system allows a vast amount of data to be distributed throughout the space and associated with appropriate locations. This capability is especially useful with data types which do not naturally lend themselves to alternative organization based on alphanumeric filing systems.

c) Prototype hardware configuration.

The interactive movie map demonstrates a complete system configuration as a general human-machine interface for any number of applications of optical video disc technology under interactive computer control with certain enhancements. The last include components such as multiple screen displays, video switching, video mixing, and sound channels and features such as run-time-generated graphics and multiple inputs (joystick, tablet, touch, voice).

d) Research instrument.

The interactive movie map is a robust research instrument for examining the processes by which the acqusition of spatial knowledge leads to the development of cognitive maps.

1.3 Overview

This thesis presents original work in three disciplines. First, it describes the systems development of a new computer-controlled, video-disc-based environment. Second, it establishes a general and unique experimental instrument for doing research in spatial cognition. Third, it formulates a series of original experiments to evaluate the interactive movie map with regard to the development of cognitive maps.

Chapter 2 traces the evolution and implementation of the interactive movie map. It discusses and evaluates the procedures employed in assembling the visual record of Aspen, Colorado--the site chosen. Material is classified under three categories: street travel, aerial overviews, and ancillary data. Next the hardware configuration and user operation of the system are discussed and evaluated under these three headings.

The remainder of the document addresses some of the cognitive implications of the interactive movie map. Chapter 3 develops a structural model of space as it is manifested in the operation of the system and hypothesizes about the implications for subjective or cognitive space. Time is similarly treated. A combination of the two is seen to form new space/time unit--events.

Chapter 4 surveys relevant theories of cognitive mapping. Discussion is divided between those that model the experiential component of way-finding and those that address the function of navigational aids like maps. The second half of the chapter examines related experimental research. The format of experimental design is addressed. Finally, the role of the technology of the research instrument is considered. A case is made for the unique contribution which the interactive movie map can make as a general research tool in this field.

Chapter 5 follows the theoretical discussion of the preceding chapter with empirical results. A series of experiments evaluating the system are presented. In the first subjects were interviewd, observed, and tested using surrogate travel to learn to find their way around an unfamiliar place. Evaluations were conducted first at the surrogate travel station and later in the real setting when subjects visited for the first time. To put these findings in a meaningful context, two other types of subjects were examined: residents of the town and first-time visitors. The investigation was carried out as pilot study. The experiment was informal and designed to expand the range of issues that warrant further study. A number of system revisions were suggested by this pilot study. The most obvious was to supply aerial overviews to complement the surrogate travel portion of the interactive movie map. With this additional mode of representation a second set of experiments was conducted specifically to study how each mode contributed to the construction of cognitive maps. The follow-up experiments consisted of two parts: one interactive and flexible, the other rigorously structured with carefully controlled differences which allowed separate examination of how the development of landmark, route, and configural knowledge was affected by the addition of navigational aids.

Chapter 6 summarizes the conclusions reached in this thesis. Further, it projects potential enhancements to the interactive movie map which can be implemented in the future to make it more effective. Finally, speculation is offered about the general implications the interactive movie map has for other disciplines. Some of the lessons learned in this research have particular relevance to educational purposes.

CHAPTER 2

Development and Implementation of the Movie Map

Rabbit puts the shift in neutral and pulls out the emergency brake and turns on the roof light and studies the map... He has forgotten the numbers of the routes he has taken and the names of the towns he has passed through. He remembers Frederick but can't find it and in due time realizes he is searching in a section due west of Washington where he has never been. There are so many red lines and blue lines, long names, little towns, squares and circles and stars. He moves his eyes north but the only line he recognizes is the straight dotted line of the Pennsylvania-Maryland border... He burns his attention through the film fogging his eyes down into the map again. At once "Frederick" pops into sight, but in trying to stead its position he loses it... The names melt away and he sees the map whole, a net, all those red lines and blue lines and stars, a net he is somewhere caught in.

- John Updike, Rabbit Run

A description of the evolution and operation of the interactive movie map follows. For background, the reader may consult the initial technical documentation of the system in Lippman [1980] and Mohl [1980], as well as additional published descriptions [New York Magazine, Filmmakers Monthly, Wall Street Journal]. The term "movie map" is used almost interchangeably with "surrogate travel." Both refer to the same system. The difference in connotation is that surrogate travel refers to the system as a model for experience and movie map refers to the system as a model for the spatial structure.

Starting with a quick sketch of the operation of the system may make the rationale for particular concerns and solutions clearer as the reader goes along. In a typical situation the user sits in front of a touch-sensitive television screen which shows a picture much like the one he would see through the front windshield of his car. Touching the speed control initiates movement down the street. By pushing another button the user changes course and turns down a side street. Now he selects a perpendicular view out of his side window. Next he calls up a map of the town which traces out the path which he has been following and permits him to decide on a subsequent route. Touching any point on the map causes it to zoom in to a detailed area. The user switches the mode of representation between street map, landmark map, and aerial photograph. Whether in map land or travel land, when he wants more information about a landmark of interest, the user touches building on the screen and is presented with a close-up of the facade. A "season knob" allows him to flip back and forth between winter and summer. Behind the facade is a slide show which lets the user stop in for a quick "visit."

2.1 Development

2.1.1 History

The first, embryonic version of the system was developed in early 1978. In a trial application of video disc technology, the corridor system of M.I.T. was filmed for surrogate travel. That version suffered from many limitations: the space was undifferentiated and without global landmarks, the film quality was poor, the hardware configuration was primitive. However, the experiment did suggest the potential of the movie map; and most important, it demonstrated the power of an interactive system. Simultaneously but independently, another movie map system had been developed [Ciccone, Lander, Weltman, 1978]. It used computer-generated animation (of a fictitious desert town, Dar-el-Mara) but was not interactive. These two precedents provided the impetus for developing the movie map as a comprehensive system.

A site was chosen to meet certain criteria suggested by experience: it had to be manageable in size, geographically bounded, and visually differentiated with distinctive landmarks. Aspen, Colorado was chosen for satisfying these conditions (Figure 1). The townsite, with an area of about one square mile, has only three routes of access. It is bounded on three sides by rivers and on the forth by a mountain. Mountain ranges visible in several directions provide unique exterior marks (Figure 2). Equally important are a number of additional properties: the street plan is a regular, rectangular grid oriented to the four cardinal directions. A unique geographical feature is that the major mountain landmarks also line up with the cardinal directions of the streets. (That this consistency

Figure 1. Map of Aspen

Figure 2. Oblique aerial view of Aspen

makes the problem of navigation easier was considered an advantage in this stage of implementing a movie map. Later versions could be designed for more problematical, unstructured environments.) Steep mountains so near provide convenient vantage points for filming oblique views of the townsite. Aerial photos are made practical by the presence of a local airport. Another advantage of this site are the seasonal changes, which drastically alter the perception of the space. Finally, because Aspen is a resort that attracts many first-time visitors, there is a large pool of potential subjects faced with the problem of finding their way around an unfamiliar town.

Filming of the town was divided into three categories: street travel, aerial overview, and cultural access. The first category included all ground-level movement through the street network of the town, the second all aerial movement above the town, recorded in aerial photos and overview maps, and the third spatial data on cultural seuqences associated with every landmark in town.

2.1.2 Street travel

The structural backbone of street travel consisted of:

a) straight footage, filmed driving up and down the center of every street, in both directions;

b) turn footage, making every turn from each street onto every intersecting street (Figures 3 and 4).

The cameras were oriented so that the field of view covered up to an entire 360-degree horizontal panorama. The town was shot in different seasons (fall and winter), different weather conditions (sunny, overcast, and

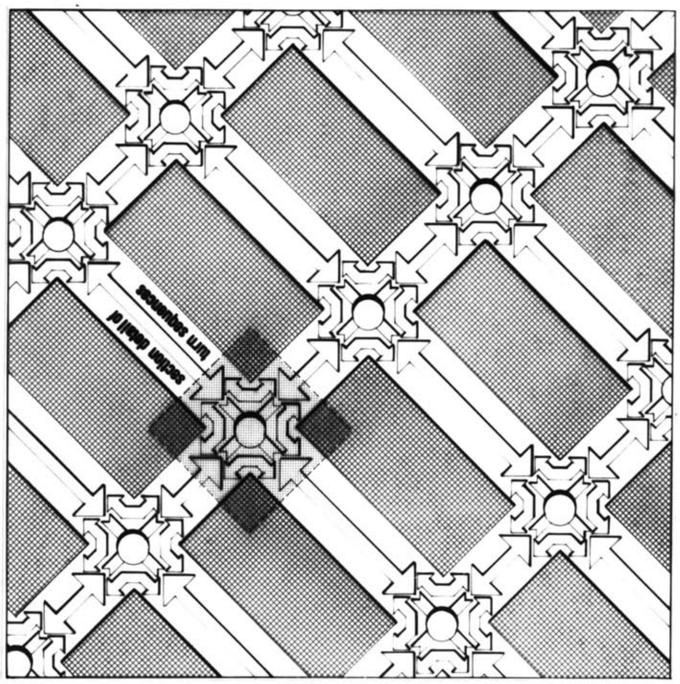


Figure 3. Straight and turn sequences

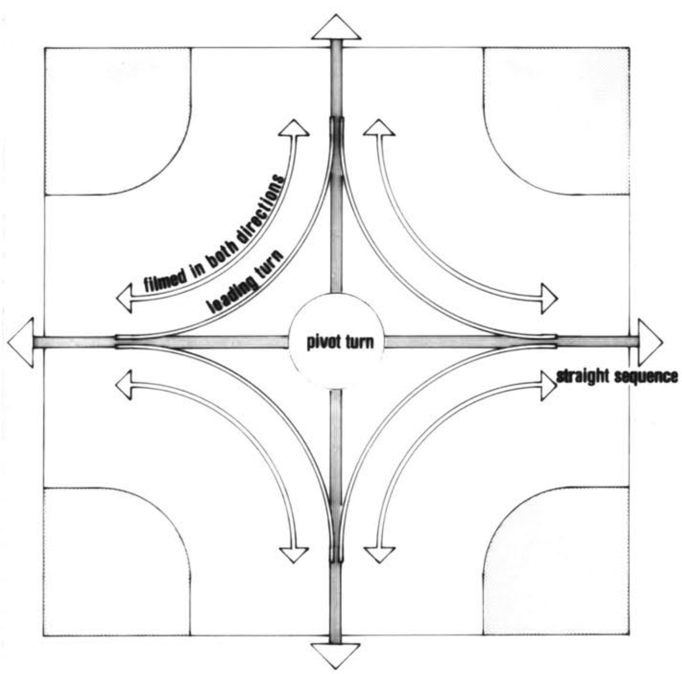


Figure 4. Detail of turn sequence

snowing) and at different hours (day and night). The entire town, consisting of 20 miles of streets, was filmed three separate times--fall, 1978, winter 1978-79, and fall 1979. Total footage in each case included 15,000 frames for straight sequences and 12,000 frames for turns.

The film was shot at a regular spatial (rather than temporal) interval, one frame every 10 feet (Figure 5). (It will be more appropriate to think in terms of a feet-per-frame interval than the conventional notation of frame-per-second rate.) The 10-foot interval was an arbitrarily determined distance which would allow all the footage to fit within the 50,000 frames of one side of a single video disc. A shorter inter-frame interval of 5 feet

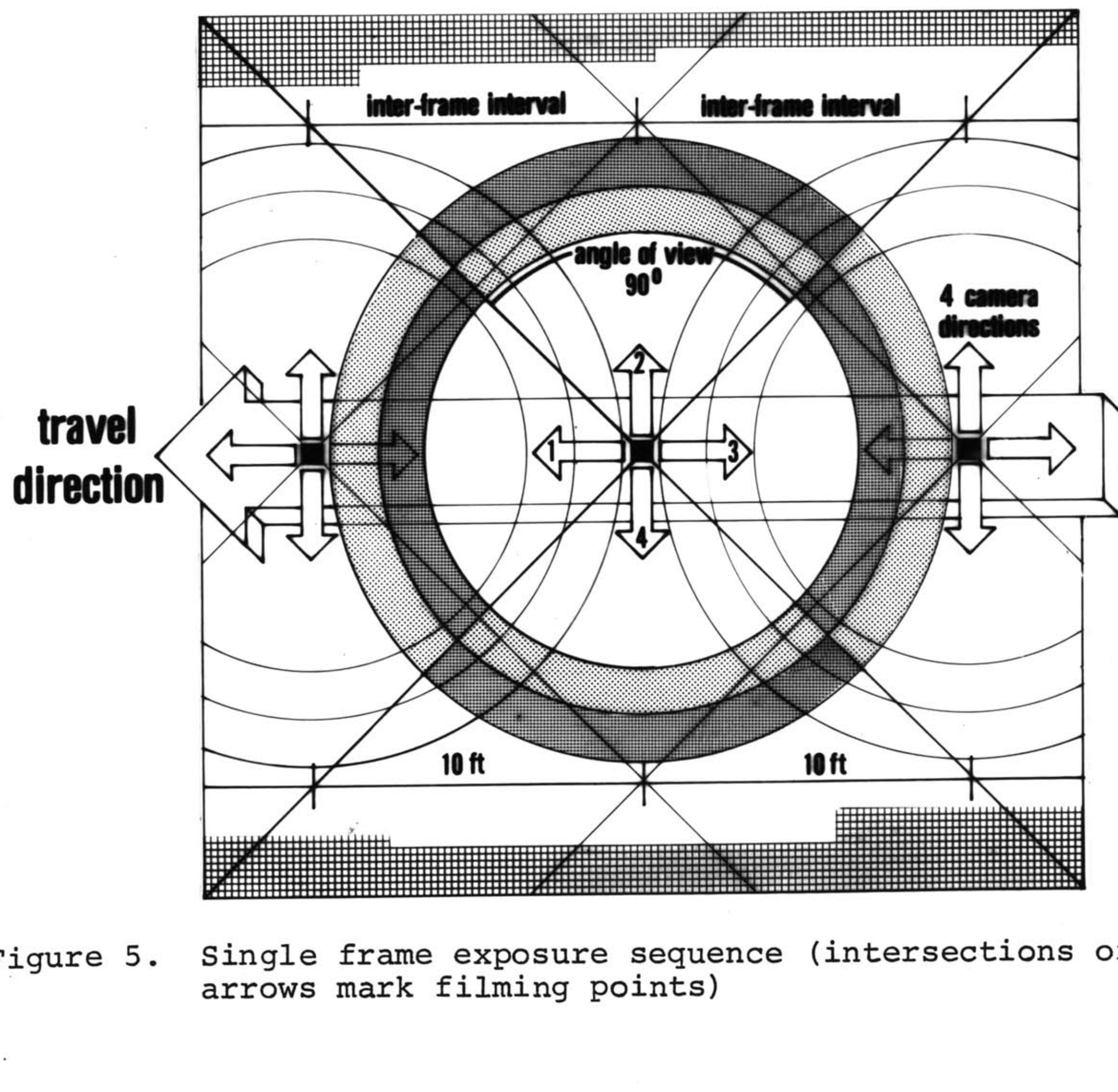


Figure 5. Single frame exposure sequence (intersections of arrows mark filming points)

was also tested. The mechanism used to trigger the cameras at regular spatial intervals was a bicycle wheel trailing the filming rig.

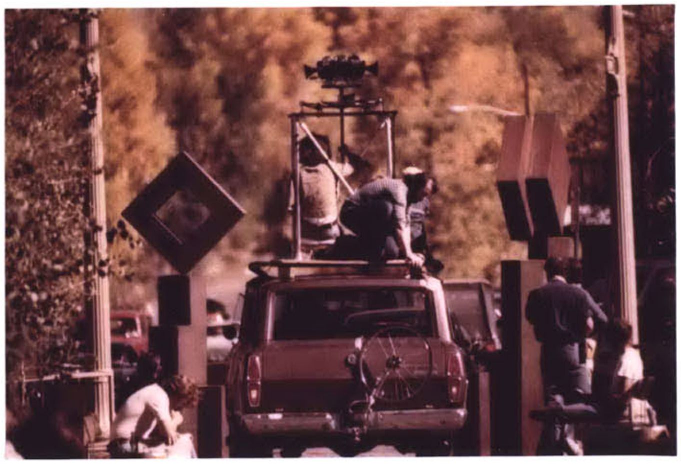


Figure 6. Filming rig

Figure 7. Camera platform

The basic filming rig was a van with a camera platform 11 feet above ground level (Figure 6). The rather high vantage point was chosen after test trials showed that lower points of view placed a distracting emphasis on incidental components of the scene, i.e., pavement and vehicles. The filming technique necessitated customized suspension of the camera platform to insure stability: because the time (and distance) between frames is subject to arbitrarily long intervals, any real-time camera unsteadiness is exaggerated, producing, in extreme cases, what filmmakers call a jump cut. Compounding the problem is the fact that during playback, the frame-per-second (fps) rate can be an order of magnitude faster, exaggerating camera jiggle even more and producing an effect like speeded-up Keystone Cop movies. In devising a solution to the problem, several different suspension systems were tested:

a) counterweighted platform with gimbaled support at the center of gravity and vertical shock absorbers;

b) the same "Steady-Cam"-type suspension as that just described with counterbalanced surface area to equalize wind resistance;

c) inertially stabilized rate gyro with servo-controlled hydraulics.

The first two versions were designed by John Borden, Peace River Films, Cambridge, Massachusetts. The third was provided by Tyler Camera Systems, Van Nuys, California.

The standard arrangement of film apparatus consisted of three 16mm Bolex EL cameras oriented at right angles aiming along and perpendicular to the direction of travel (Figure 7). A fourth, rear-facing camera used on the first shoot in 1978 was abandoned as redundant because each street was traversed in both directions. Experiments included varying the angle of the side cameras from 90 degrees off axis to 30 and 40 degrees both front and rear. An alternate film format also used for the entire town was the anamorphic Volpi lens mounted on a 35mm film camera pointing straight up (Figures 8 and 9). An example of the sequences shot with this lens is shown in Figure 10. This lens recorded a 360-degree panorama of the scene extending from 30 degrees above the horizon to 30 degrees below onto a ring (Figure 11). The field of view covered was significantly different from that of a conventional fisheye lens, which would have wasted much of the frame on sky unless the location were surrounded by tall buildings (Figure 12).

Many experimental variations in filming turns were conducted with the aim of improving the accuracy of interpretation by the viewer and/or decreasing the time and effort necessary for film turns. One of the first attempts, called the "pivot turn," consisted of rotating the camera 360 degrees while it was fixed at the center of each intersection. One frame was shot every 10 degrees, so a typical right-angle turn was 9 frames (Figure 13). Some tests were also made using a cherry picker to achieve a vantage point 50 feet above the subject. These aerial "pivot turns" used the same 360-degree rotation to provide a more comprehensive oblique perspective. The standard technique eventually adopted was the "truck turn." In this case the camera was fixed on axis with the direction of the filming vehicle as it executed a conventional turn (Figure 14). "Leading turns" were an integration of truck turns and pivot turns. They were constructed by pivoting the camera in the direction of the anticipated turn in advance of the intersection (Figure 15). As the path through the intersection was completed, the camera was rotated more and more slowly until it pointed once again on axis with the direction of travel. In fact, it was counter-rotated relative to the film rig axis. The angle of disparity

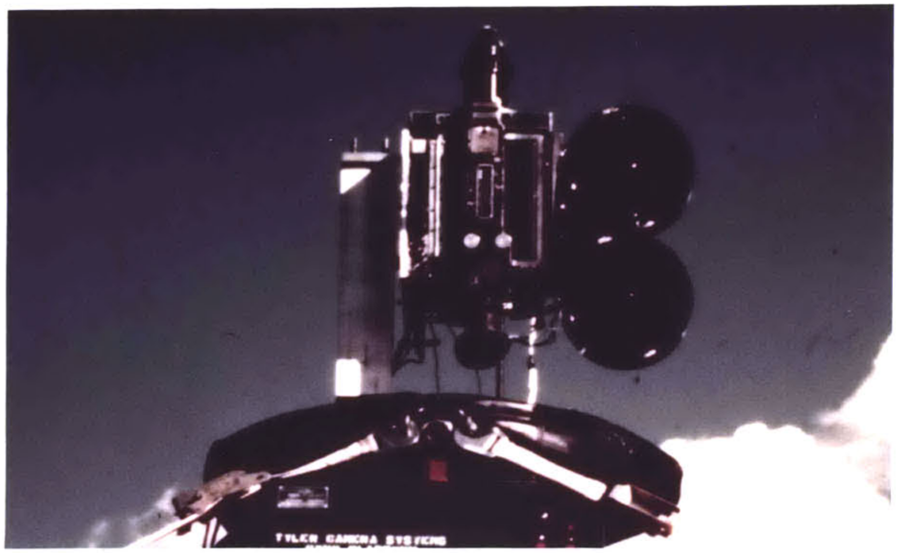


Figure 8. Camera with anamorphic lens

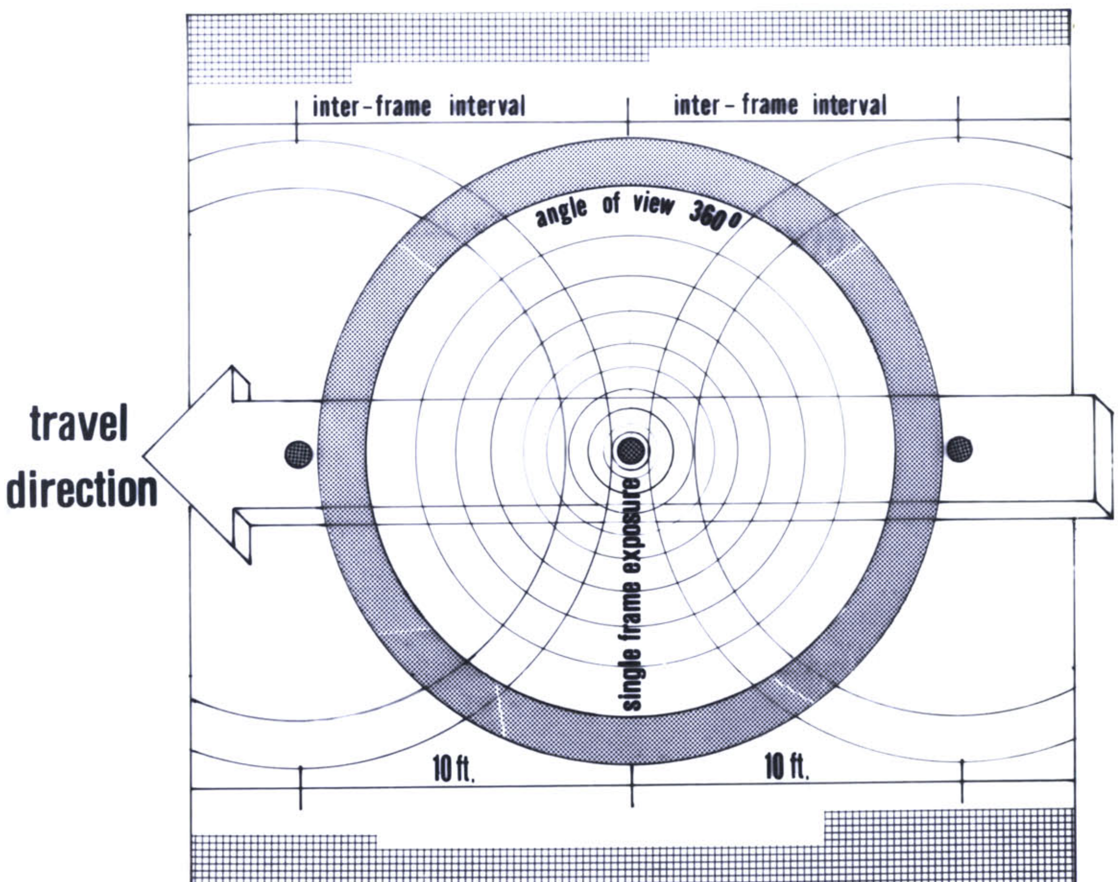


Figure 9. Single frame exposure sequence

Figure 10   Sequence of 360 degree anamorphic views

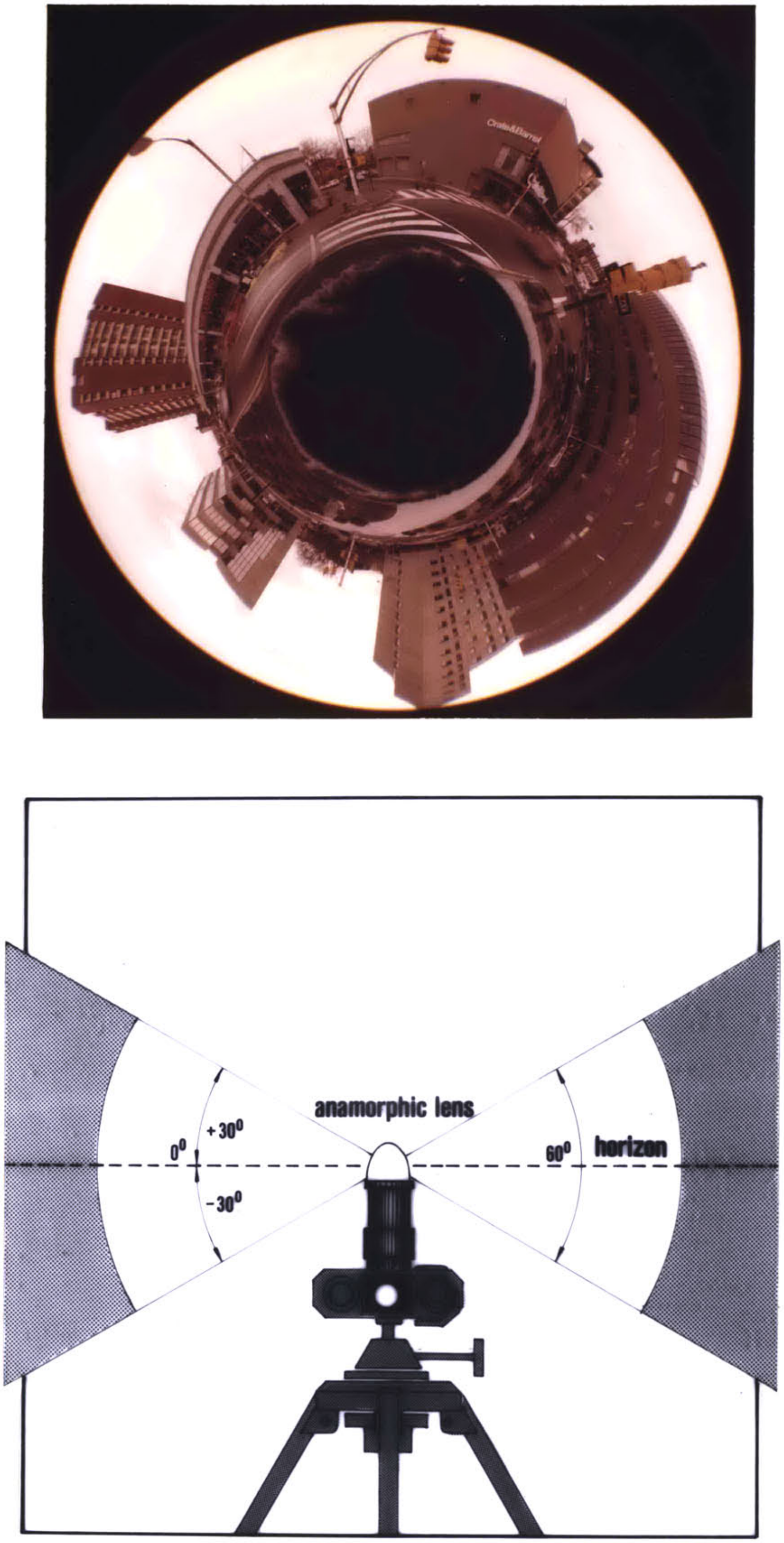


Figure 11. Anamorphic lens, field of view

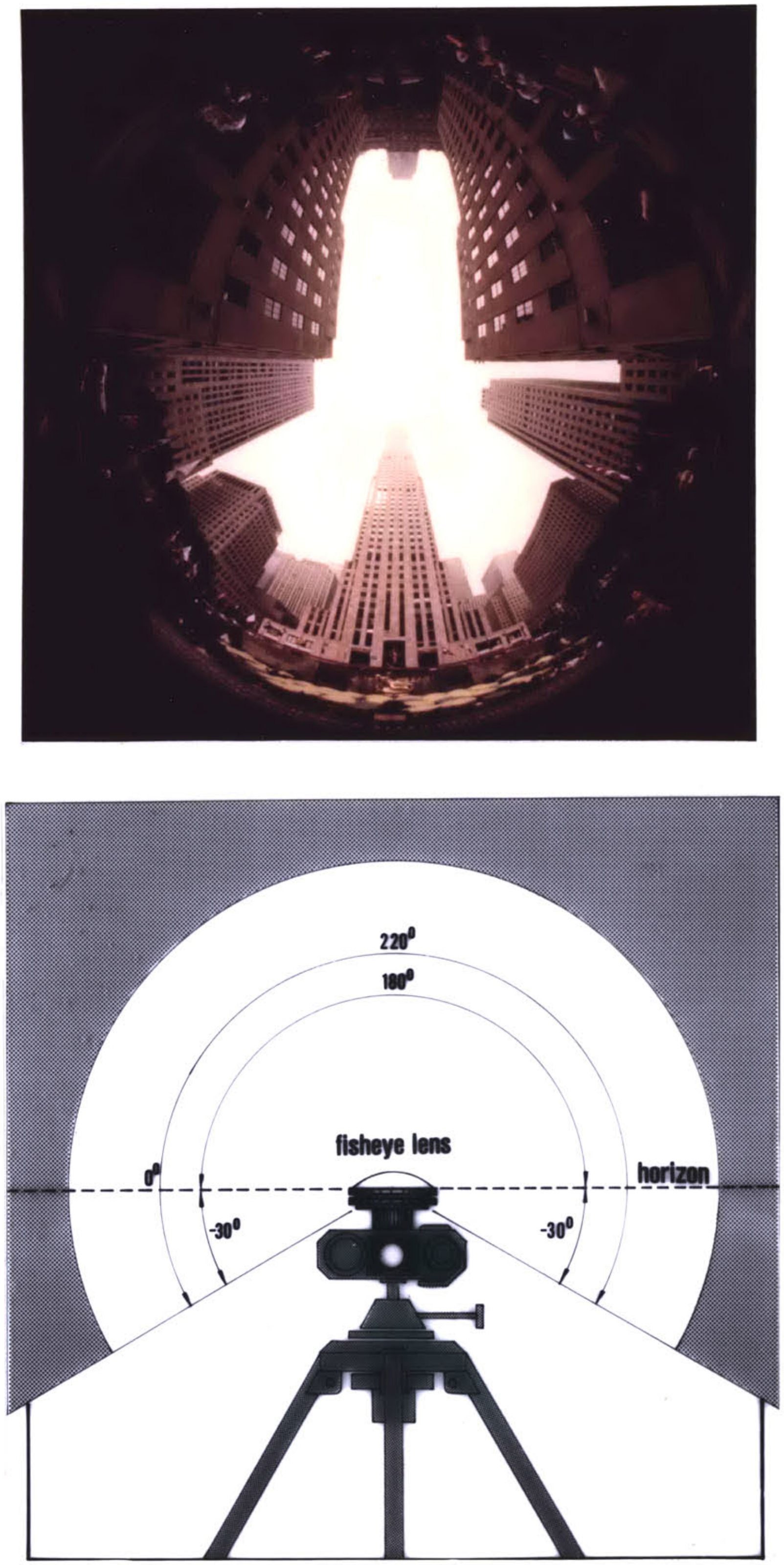


Figure 12. Fisheye lens, field of view

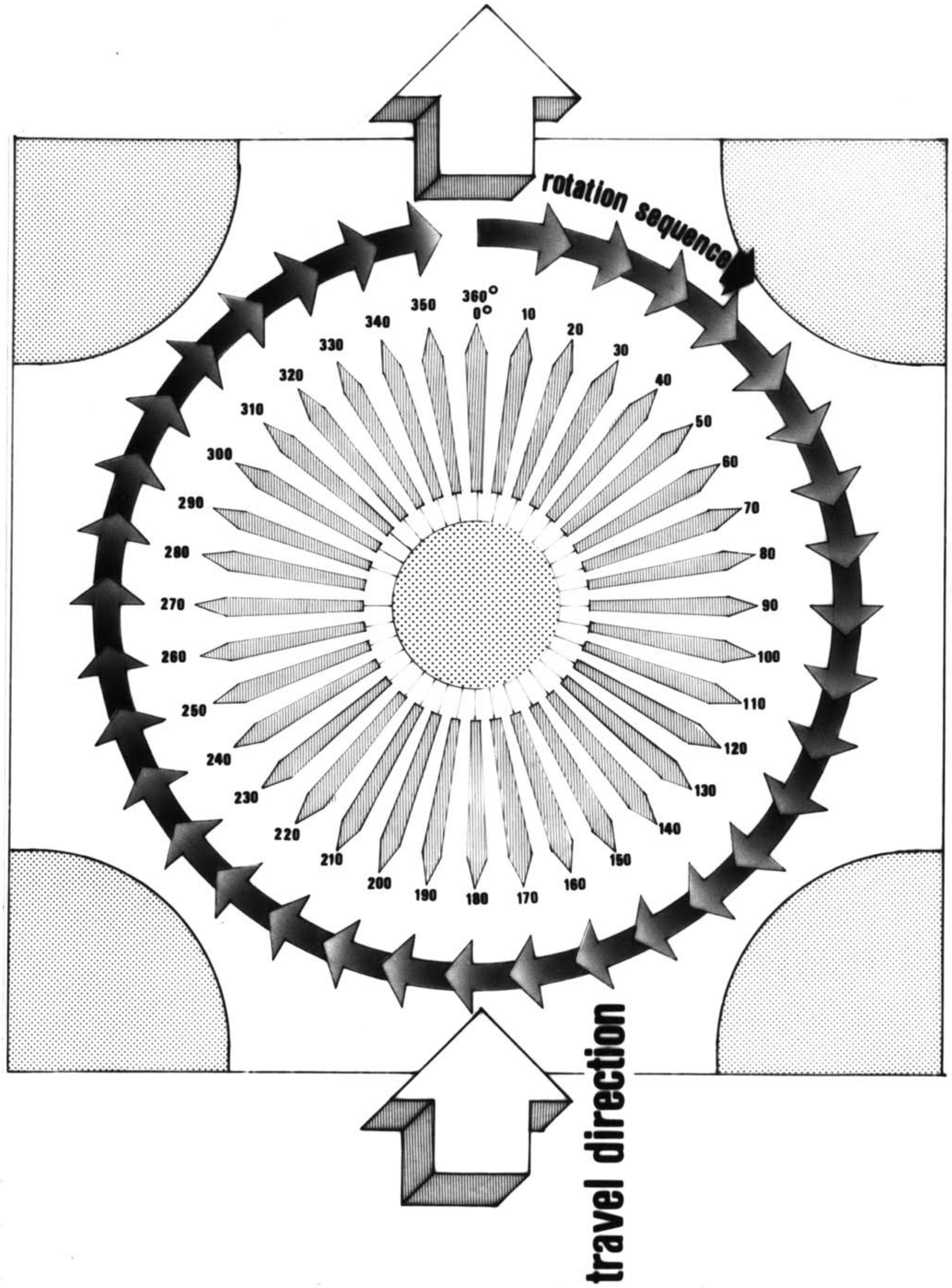


Figure 13. Pivot turn

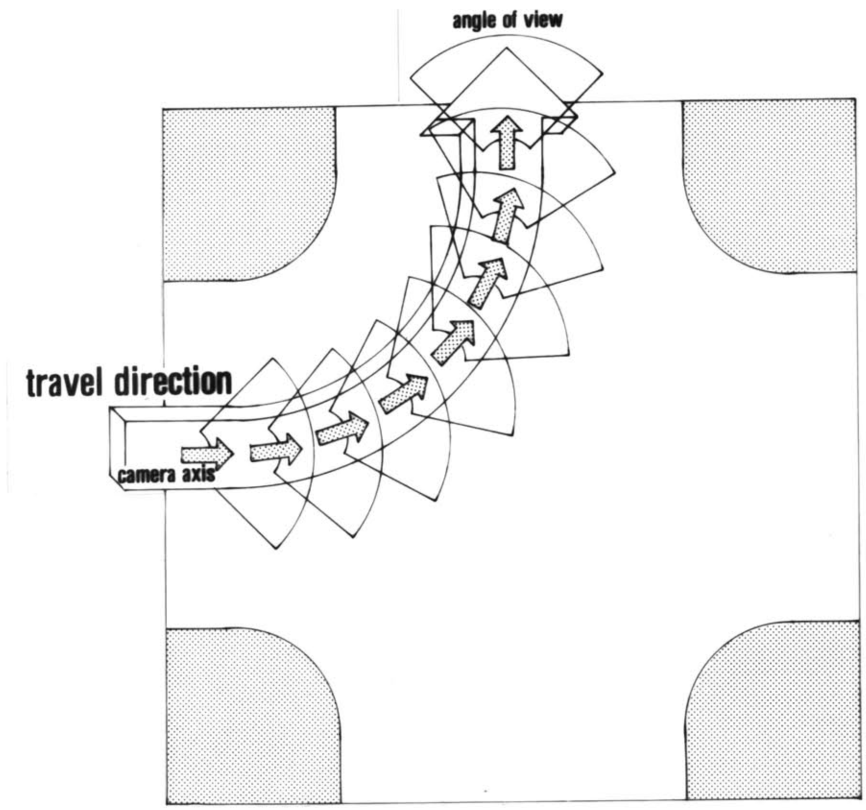


Figure 14. Truck turn

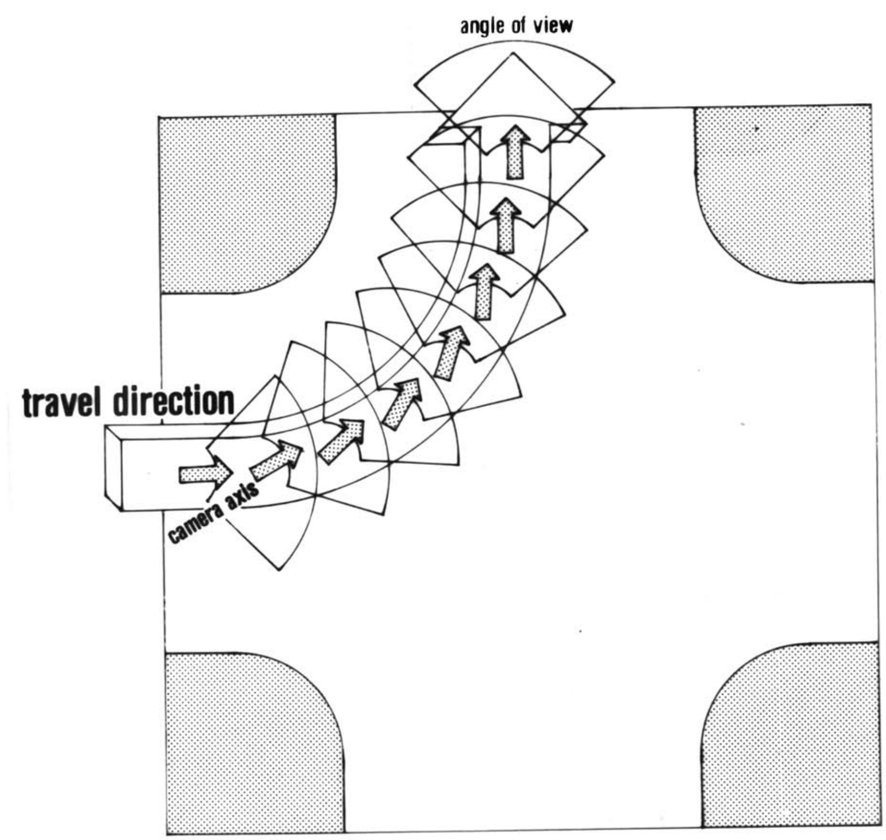


Figure 15. Leading turn

between the direction of view and the direction of motion increased from 0 degrees at the start of the camera turn to about 30 degrees at the start of the traveled curve and then decreased to 0 degrees at the conclusion of the turn. A variation of the pivot and leading turns produced the "tracking pivot turn" (Figure 16). In this case the pivoting camera was rotated while traveling in a straight line along the street axis up to the intersection. The camera began pointing in the direction of travel and ended pointing down the street exactly when it reached the center of the intersection.

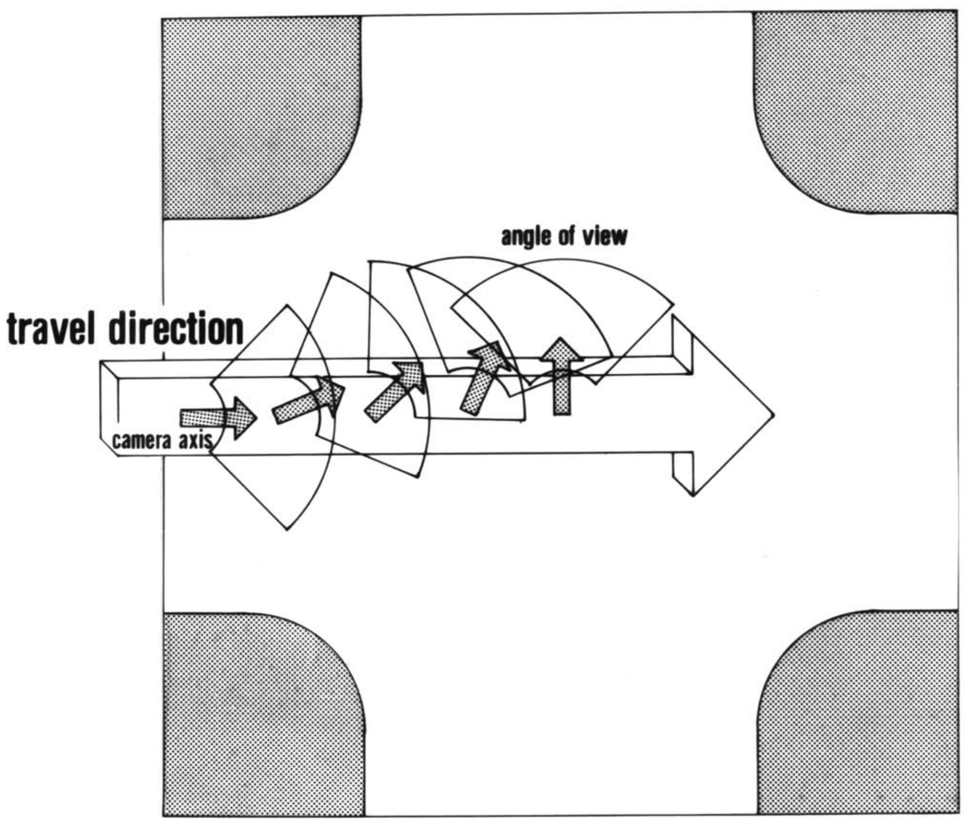


Figure 16. Tracking pivot turn

A technique for creating "synthesized turns" was also developed [see Yelick, 1980]. These turns were generated in-house by computer after production using pictures from on-site filming. Two perpendicular views from cameras pointing down the center of intersecting streets were digitized. The computer reconstructed the in-between views by interpolating between the two source images (Figure 17). One algorithm used was to simulate a pivot turn, applying perspective correction to these two perpendicular faces. Another experiment was simply to slide the abutting views across the frame without any perspective correction. Computer-synthesized turns were also created from the 360-degree anamorphic panorama views.

A separate and more comprehensive computer-generated graphics system was developed to show an abstracted representation of the town from any viewpoint. QADAS, or Quick and Dirty Animation System, as it was affectionately nicknamed, provided a three-dimensional model of the town [see Bender, 1980]. The data base was entered from topographical maps by tablet, digitizing the x,y coordinates of the base corners of each building in Aspen. Height was approximated according to number of stories, determined by visual inspection of previously shot travel footage. Time for data entry was about 16 man hours.

QADAS incorporated 7 degrees of freedom in determining the view: 3 for position (x, y, z), 3 for orientation (roll, tilt, yaw), and 1 for focal length of the viewing lens. The "painter's algorithm" was employed, sorting polygons by distance from the viewer and drawing from back to front. The computational time required to generate each frame was 30 seconds to 1 minute, depending on the complexity of the elements in the scene. As each frame was generated,it was photographed by a Matrix Color Graphic Camera onto 35mm film. Among the sequences filmed by computer animation was standard street-level travel, with inter-frame intervals and focal length to match the on-site street travel footage.

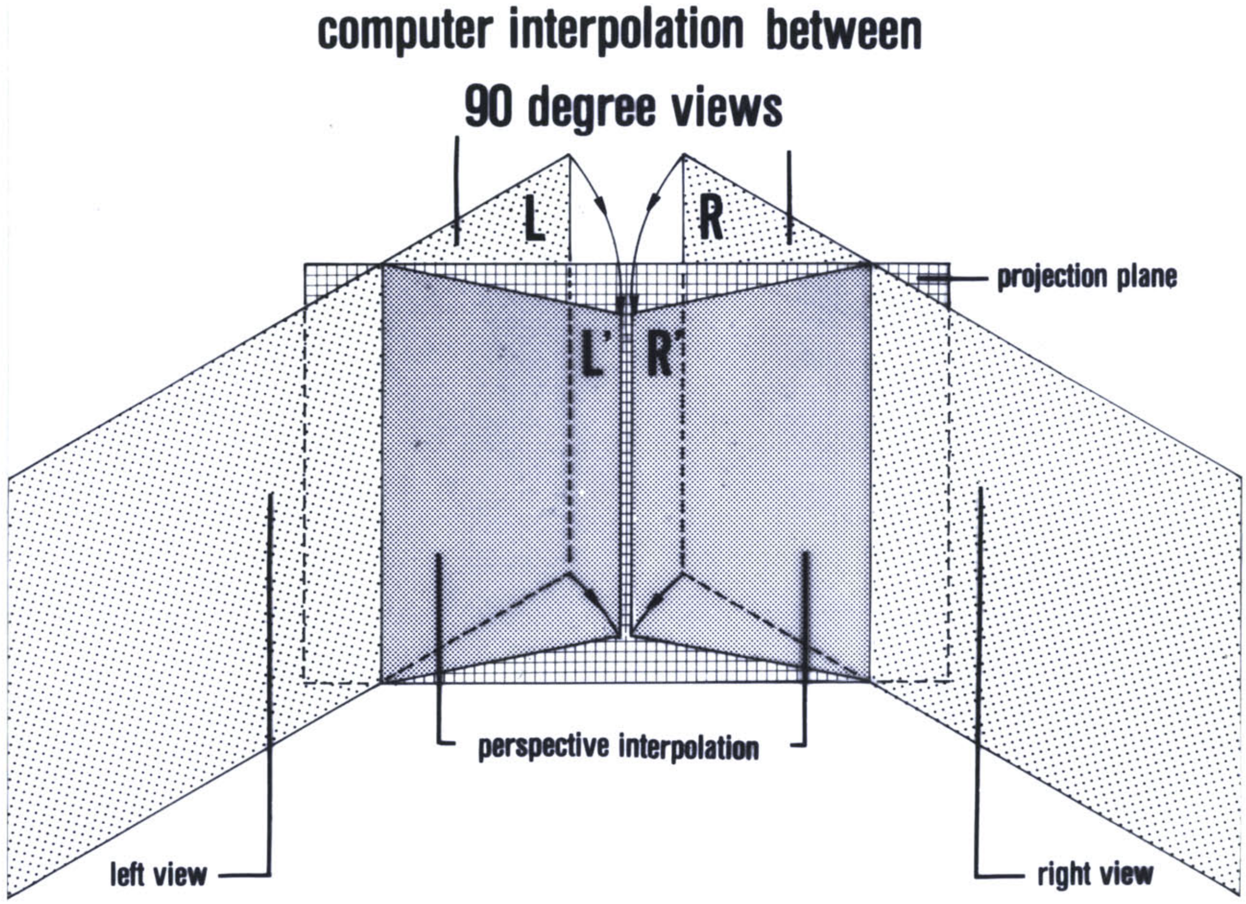


Figure 17. Synthesized turns

Certain sequences were filmed from viewpoints that could not be duplicated by real photography. Figure 18 shows examples of movement in z from ground level to 600 feet, diagonal movement in x and y at 80 feet in altitude, and yaw rotation. "Helicopter" sequences combined several of these degrees of movement to produce circumnavigation and helical landing approaches.

The procedure used in generating computer animation views is illustrated in Figure 19. The bottom picture illustrates three modes of representation: the underlying wire frame structure, displayed using hidden lines with smooth shading (dejaggied), solidly colored polygonal shapes, and polygons with photographic texture. Photographic texture was added to selected building facades to enhance the degree of visual realism. (Background mountains were represented by defining them as monster facades standing far in the distance.) The procedure was to digitize images of the more prominent landmarks and "clean them up" to remove deep shadows and obstructions like bushes or cars [see Bender, 1980]. The upper pictures in Figure 19 show examples of the original video disc image and the digitized one which was used for the computer animation. As QADAS drew one of these buildings, it mapped the stored photograph onto the perspective rendering of the facade. This "billboarding" process treated the facade as a two-dimensional surface, ignoring perspective transformation of any features with depth, like porches, awnings, and windowsills. Comprehensive techniques for image perspective transformations are discussed by Devich and Weinhaus [1980] and by Parsons and Henrikson. Comparisons between an original scene and the computer-generated version with perspective billboarding are shown in Figure 20.

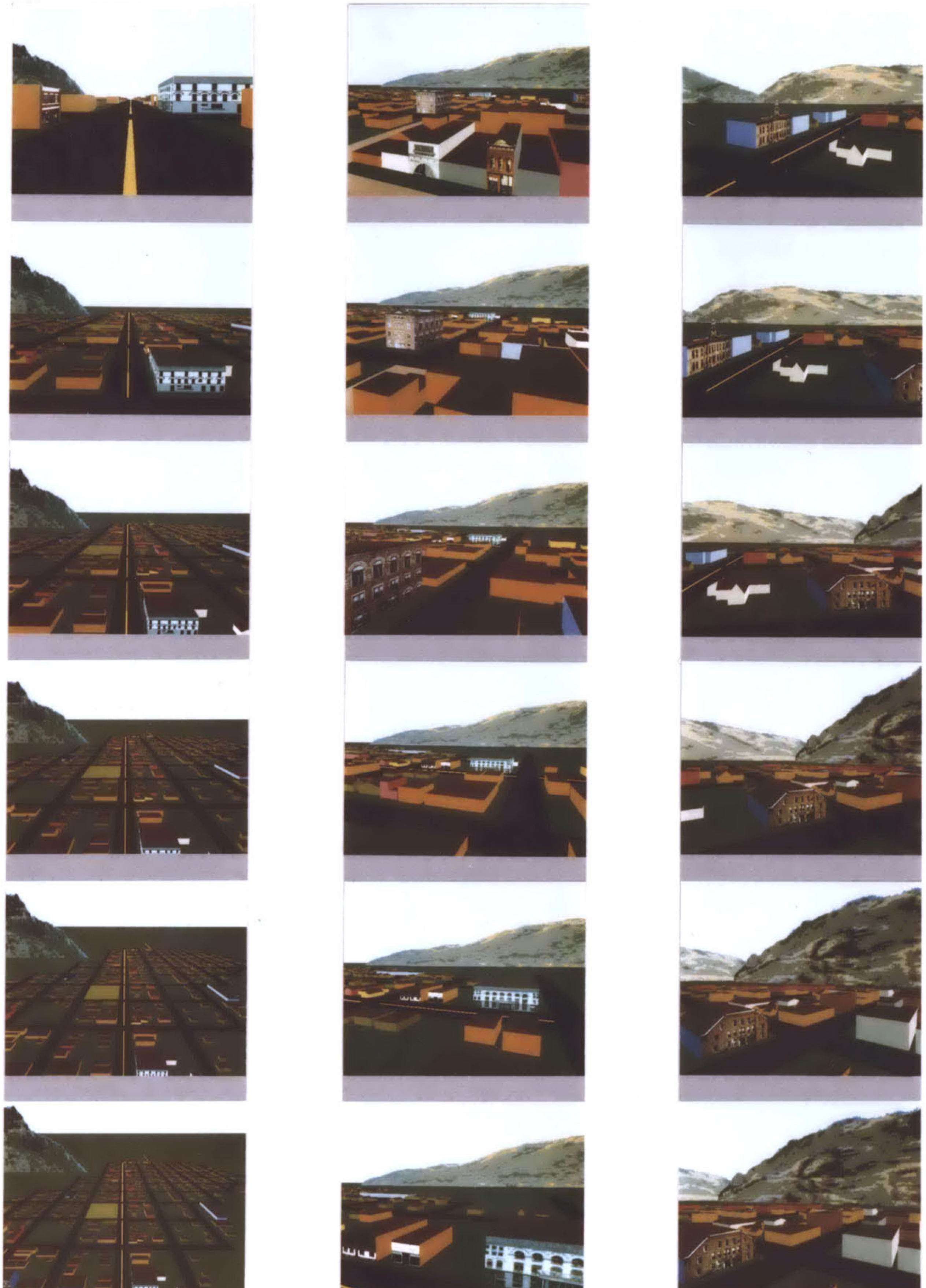


Figure 18 Computer generated views showing vertical,diagonal and rotational movement

Figure 19 Billboarding photographic detail onto computer generated views

Figure 20. Comparison between original view and computer-generated version (upper right, lower left)

2.1.3 Aerial overview

The aerial overview was conceived of as a dynamic window which allowed for a viewpoint that could "helicopter" above the town, moving laterally in x and y and zooming in z. Orientation of the viewpoint remained fixed perpendicular to the ground plane. Theoverviews were constructed from three modes of representation--photograph, street map, and landmark map.

The source for the first mode was a 4 ft. x 8 ft. black and white photograph collaged from smaller recitified aerial prints, which was used as copy for filming (Figures 21 and 22). The original map supplied by

Figure 21. Aerial photo

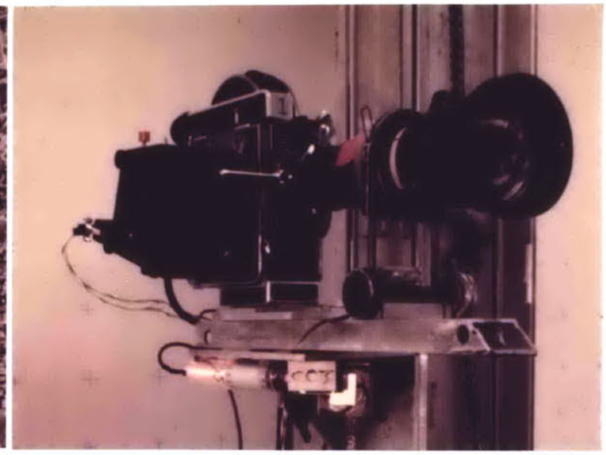


Figure 22. Copy camera

Cooper Aerial Survey, Tucson, Arizona. The aerial survey was flown in 1974 at an altitude of 3000 feet, shot in a standard film format 9 inches square, and printed at a scale of 1:50. The rectified prints were cut into irregular shapes, borders following contours of objects such as buildings, streets, and rivers to conceal boundary lines where prints were joined. The topography of the street map was derived from the aerial photograph. (Corresponding details are shown in Figure 23.) The map was custom-drafted in house, because standard Aspen street maps available had detail density and line width suitable for the medium of paper but not for video. Graphics were specially designed for the resolution and legibility of television. Sizes and positions of map labels were arranged so that, for each possible frame, the labels most appropriate for that level of scale would be present and readable. (For example, increasingly detailed information was displayed in smaller fonts, which would be conspicuous only when the viewpoint zoomed closeup.) While this map emphasized streets, the landmark map featured the outline of every building. The names and locations of particularly noteworthy places were accented with color coding and name identifiers.

Figure 23 Street map and aerial photo, matched

The structural format for filming the aerial overview consisted of a three-dimensional lattice of camera viewpoints above the ground plane. Two-dimensional positioning of viewpoint was achieved with a vertical 8 ft. x 12 ft. x-y plotter. Mounted on the moving arm of the plotter was a copy camera, an 18mm Bolex with a 12-240 zoom lens (Figure 22). Attached to the camera was a low-powered laser. Sighted on axis with the lens, it targeted an 1/8-inch spot of light on the flat work. Once calibrated with the viewing field of the camera, the laser enabled precise positioning of the camera without looking through the viewfinder. Before filming began, the coordinates of each intersection were digitized, using the laser for registration. Movement along the z-axis was simulated with the zoom lens.

Three different lattice networks were filmed: cubic, pyramidal, and lateral. The cubic lattice consisted of 7 viewpoint levels of scale, zooming perpendicularly into each of 130 intersections, as shown schematically in Figure 24. (For clarity, only alternate frames, labeled "F," are shown.) Once positioned above a given intersection the camera remained fixed while the lens was manually zoomed to preset focal lengths. These settings corresponded to magnification increases of a factor of 2 for each step. (This factor was chosen for implementation considerations which will be explained later.) The scale ranged from a 16-block-wide field at the top level to a 2-block-wide field at the bottom close-up. A fortuitous circumstance lent itself to the cropping proportions of the TV viewing field which determined the filming frame. The longer east-west axis of the block grid of the town closely matched the 4:3 horizontal-to-vertical aspect ratio of the viewing frame. As a result, framing symmetry with respect to x and y was maintained at all scales.

The pyramidal lattice also employed 7 levels of scale for each of the 130 intersections. In this case, however, the viewpoint of the camera was panned laterally between each level. The technique is diagrammed in Figure 25. (Again, for clarity, only alternate frames, labeled "F," are shown.) The top level always began centered over the town. As the field of view zoomed in, the x-y position of the camera, labeled "P," translated toward the target intersection by a distance proportional to the amount of the zoom. This algorithm had the effect of keeping the target intersection fixed in relation to the viewing frame throughout the zoom. An inspection of Figure 25 will demonstrate that for each step the coordinates of the target remain in constant proportion to the framing window.

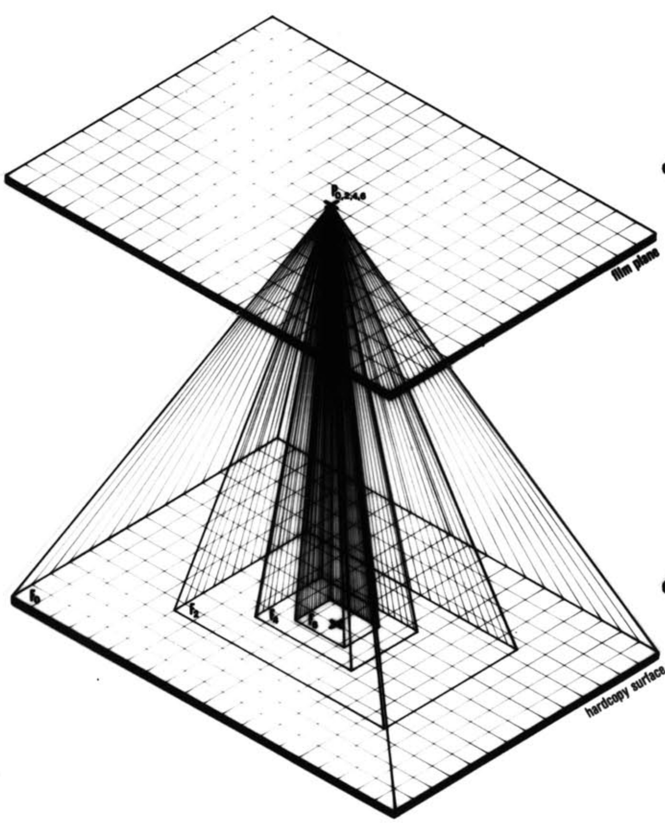


Figure 24. Cubic lattice

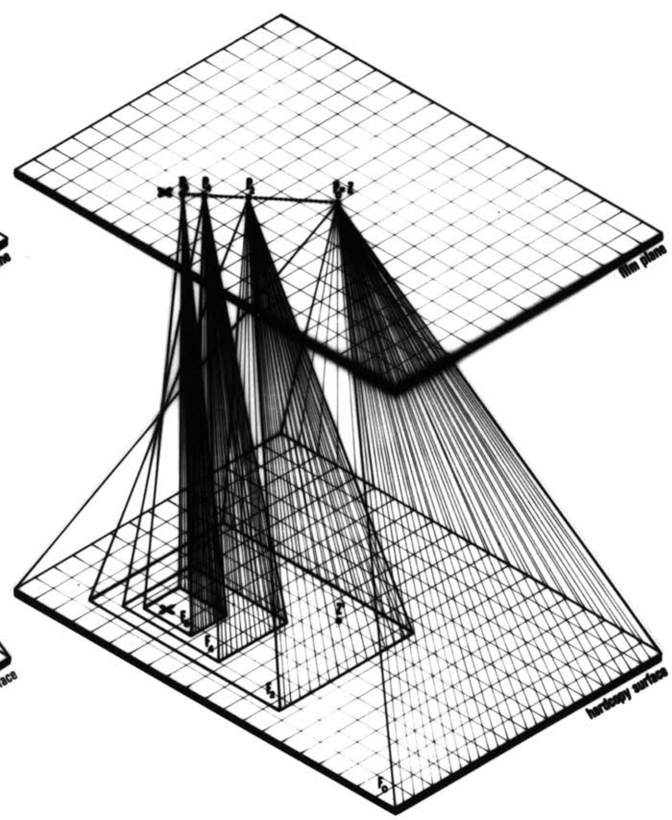


Figure 25. Pyramidal lattice

The pyramidal lattice kept the town centered at the top level and let the target intersection remain off-center as the camera zoomed in. The cubic lattice, on the other hand, maintained a field centered on an intersection, with the result that the town was off-center when the camera was zoomed out.

Lateral lattices were filmed at two different scales. At the top level (16 blocks wide) a series of five radial arms were filmed (Figure 26). The arms extended like spokes from a centering hub intersection of the town. The purpose of the arms was to serve as a transition from the centered view of the layout to the off-center view with which each intersection in the cubic lattice started at the top level. The number of frames required was much smaller than for the other lattices (160 compared with 910). It was not necessary for separate radial arms to be filmed for each intersection because (1) the arm to the outlying intersections would already include the correct frames for closer points along the same line and (2) at this large scale a given arm could serve a larger section of nearby intersections with negligible inaccuracy.

At an intermediate scale a rectangular lattice was filmed. The network corresponded to the street layout, tracking along every street (Figure 27). In this case three frames were shot per block--one centered on the intersection and the next two at one-third intervals along the block. The purpose was to create paths which would serve as transitions between intersections in the cubic lattice. This was the only lattice in which the laser calibration spot was left on during film exposure to serve as a meaningful position tracker.

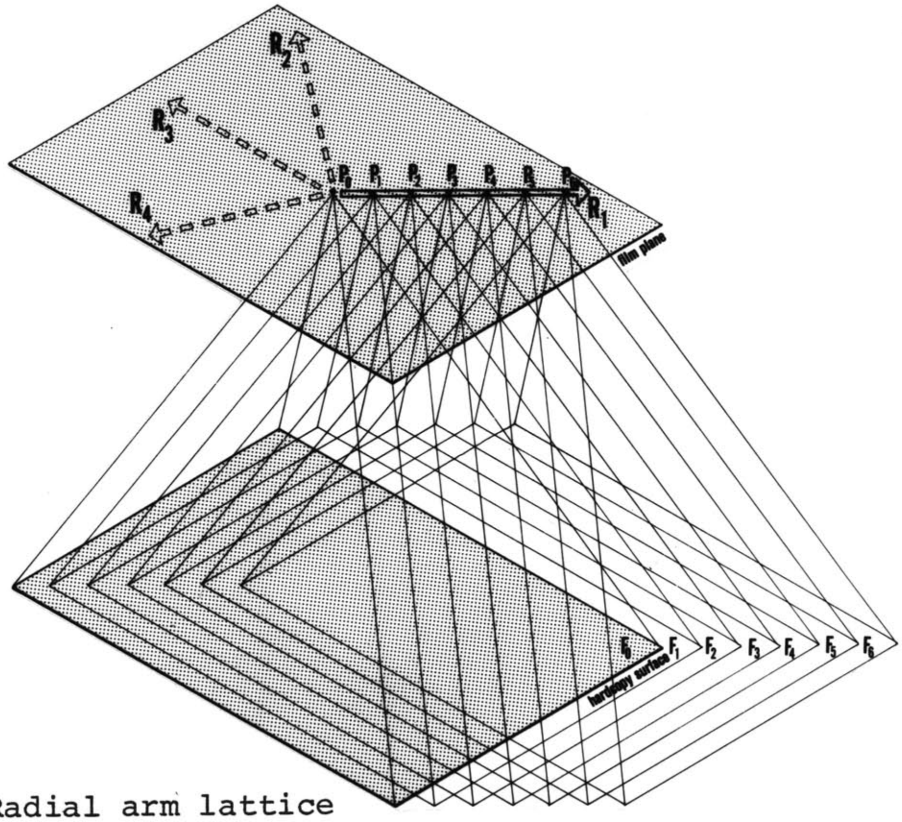


Figure 26. Radial arm lattice

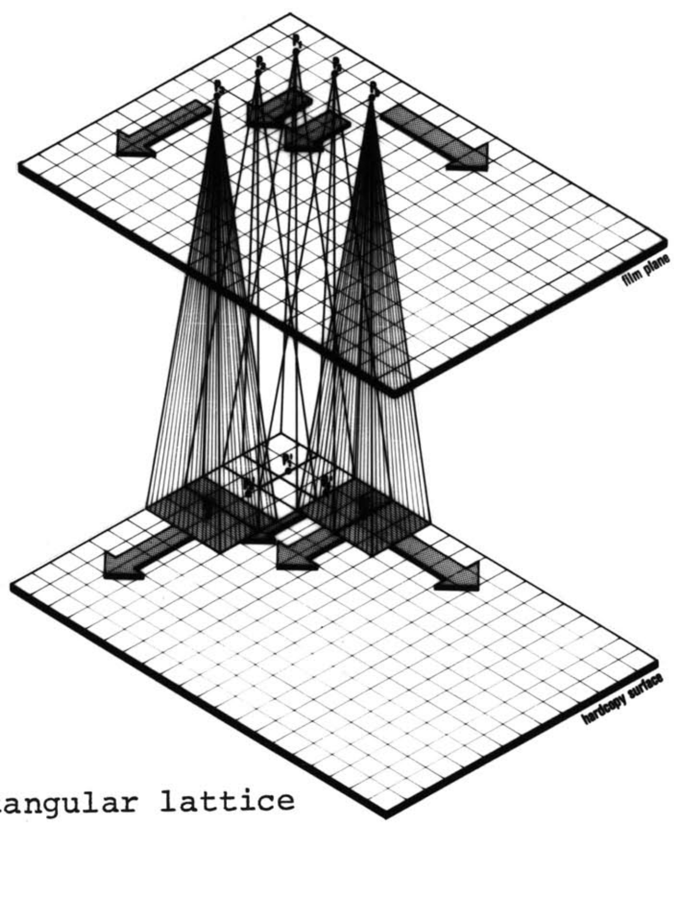
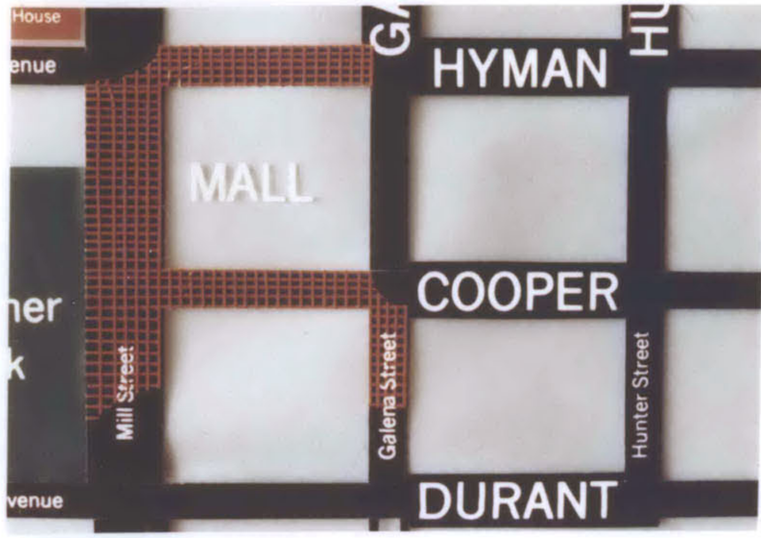


Figure 27. Rectangular lattice

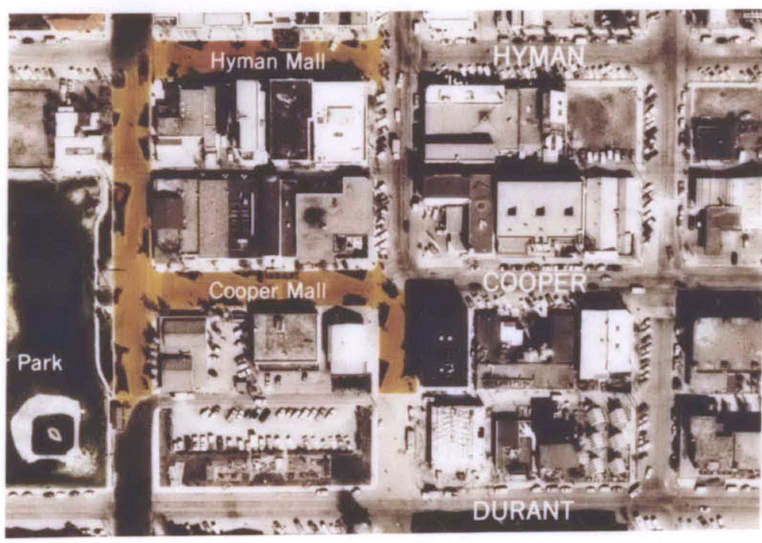
Within each lattice described above (including all intersections at all scales), the aerial photograph and the street map were identically filmed with matching registration. An isomorphic match in the topology was assured because the map measurements were derived from the photograph. An example of corresponding frames in the two modalities at an intermediate scale is shown in Figure 23. The landmark map was filmed only at the smallest scale (zoomed all the way in) and only in the centered cubic lattice. Figure 28 shows an example of three-way correspondence between modalities.

2.1.4 Cultural access

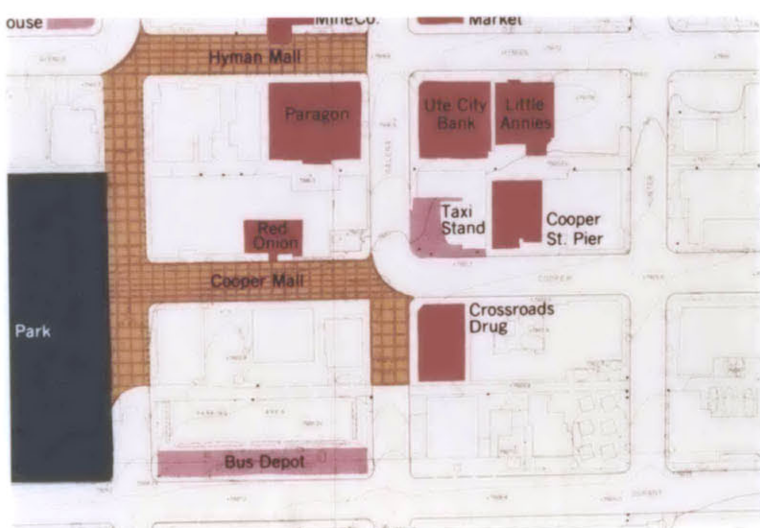
The material produced to provide access to examples of local culture included slide shows, short cinema verité films, and sound recordings. To provide a transition from surrogate travel to this ancillary data, separate pictures were shot of each building in Aspen. Using 35mm cameras, 2,000 facades (including parks and other undeveloped sites) were photographed in both fall and winter with matching registration (Figure 29). The initial registration technique was based on verbal descriptions spoken into a portable tape recorder each time a facade picture was shot. In the fall the photographer described exactly how the picture was being centered with reference to fiducial marks on the viewfinder focusing screen, where he was standing, and what focal length was being used. In the winter the photographer listened to the tape recording as he reshot the facades, carefully following the directions from the first outing. While the initial pairing was an approximate registration, a postproduction realignment and duplication of projected slides was necessary to achieve absolutely accurate alignment between fall and winter. In addition to seasonal



Street map



aerial photo



landmark map

with matching registration

Figure 28

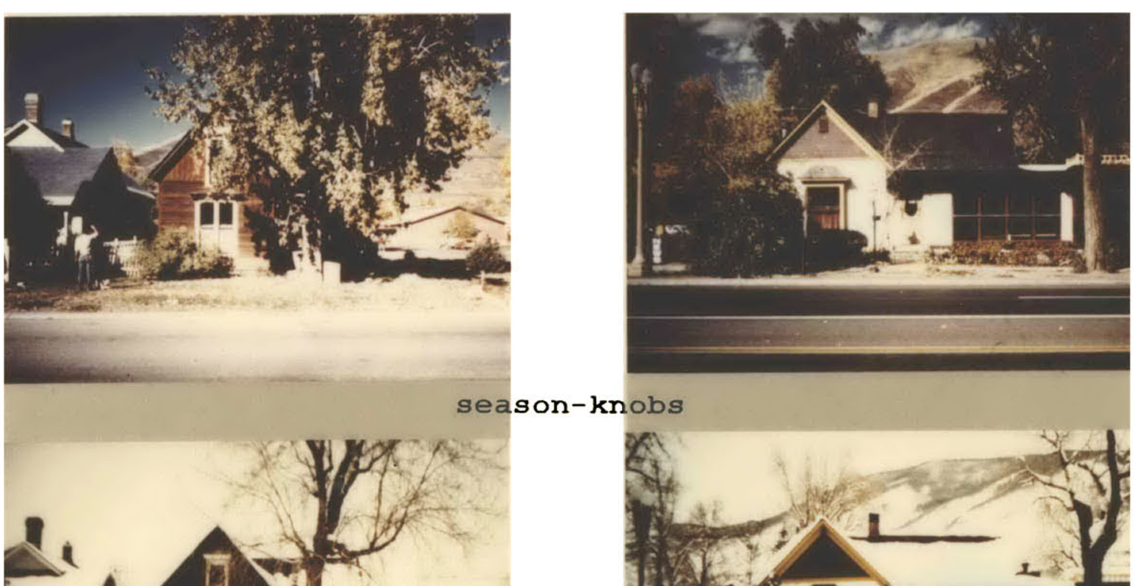


Figure 29

registration, historical registration was created where possible. For scenes of which historical photos existed from the turn of the century, modern-day facades were reshot to match the original views.

The interiors of hundreds of sites in Aspen were photographed separately with a 35mm camera to capture a sense of the cultural activity behind the facades of the more interesting locations in town. While some attempt was made to shoot a few frames showing the approach to the entrance way (serving as a transition from the facade shot to the cultural data), the concept of preserving spatial continuity was suspended once inside the building. Here the purpose was simply to record, in the form of a slide show, the "feel" and not necessarily the layout of the place. In addition to 6,000 shots taken on site, printed matter like menus, brochures, and newspaper clippings was collected and later copied as elaboration on the slide shows. Dozens of cinema verité "moviettes" (from 20 to 120 seconds long) were shot by professional filmmakers to document similar cultural slices. Both 16mm and 8mm film were used, the latter being blown up to 16mm before the transfer to video. In the case of both slide shows and film, sound was recorded in conjunction with the visuals. Special cases of audio included experiments with verbal guided tours by local residents as well as binaural sound to recreate the characteristic ambience of a space. A separate introductory slide show was also created to serve as an orienting overview for visitors unfamiliar with the town.

2.1.5 Mastering the optical video disc

Material from all of the categories described in this chapter was edited and transferred to optical video disc. The bulk consisted of 16mm film with either sync or wild sound on accompanying 1/4-inch half-track tape. Other formats included 35mm movie film and 2-inch quad video tape. In the pre-master transfer to 2-inch helical video, all film material was transferred at 30 frames per second with the exception of the cinema verité sequences which were transferred at the usual 24 fps with a 3:2 pulldown. In all, four discs were mastered. The first, with the full 54,000 frames, had the complete surrogate travel footage with examples of pivot turns, leading turns, tracking pivot turns, and cherry picker experiments. It also had facade pictures, slide shows, and extensive cinema verité footage. The second and third discs were largely experimental. The former included both fall and winter surrogate travel and thevarious tests of side cameras angles. The latter consisted of a catalog of ancillary cultural data plus examples of the first version of QADAS footage. The fourth disc was truly comprehensive, including street travel (different seasons), aerial overviews, facades (different seasons), cultural access (both slide shows and cinema verité), QADAS animation (version two), 360-degree anamorphic panoramas, computer-synthesized turns, and the visitor's introduction to the locale.

2.2 Production Evaluation

There are two stages at which it is appropriate to evaluate the success of the experiments and variations tested in developing the interactive movie map. The first is after the optical video disc has been mastered. The second, after the computer interface has been implemented, will be discussed later.

2.2.1 Street travel evaluation

The time required to film street travel forthe entire town dropped from approximately 200 man hours on the first trip, including time for debugging equipment and procedures, to a surprising 24 man hours on the final trial. In the latter case the three-man crew filmed all straight sequences during a four-hour span in a single day and all the turns the next day.

Examination of the first footage, taken in 1978, showed that the time-of-day latitude, ranging from 9:00 a.m. to 4:00 p.m., produced too much inconsistency in lighting, particularly with regard to shadows and contrast. Conditions. were optimized in the final filming, which was done under sunny skies, with all shooting taking place between 10:00 a.m. and 2:00 p.m. While overcast weather provided more uniform lighting, the advantage was nullified by a drop in con-rast and in the orientation cues provided by shadows.

The inertially stabilized platform with the Tyler mount provided the smoothest suspension for the filming rig, especially in keeping the horizon line level. It was themost successful in minimizing discontinuities between successive frames along the street and in matching turn footage with straight-ahead travel. The biggest problems with the terrain proved to be dips at intersections, steep streets, and snow-packed ruts in winter. It was established that while the 11-foot height of the filming cameras provided a vantage point foreign to everyday experience, the resulting perspective was more meaningful and more easily interpreted than the eye-level view, which was more occupied with cars and pavement. The 10-foot inter-frame interval was found to produce an unavoidable jerkiness. The

Tyler mount minimized the overall sense of jerkiness the most. Test shooting at 5-foot intervals was, of course, smoother than at 10, but the disadvantages were reduction in the area which could be covered on a single video disc and in the maximum speed of travel which could be replayed on video disc.

The perpendicular side camera views were a little bit more sensitive to the inter-frame interval. Unlike the forward view, an object in the side field of view maintained a relatively constant distance from the camera. The resultant speed of movement of objects across the frame was linear and inversely proportional to their setback from the center of the street. With a 90-degree angle of view and a setback distance equal to the inter-frame interval, the object would pass into and out of view in just three frames--too fast to convey a good sense of spatial relations. This proved to be a problem only in the case of encroaching objects like trees and street lamps on a mall. In the spirit of turning a bug into a feature, it was observed that the briefer life span of side view objects could serve to advantage: the film could be viewed backward with negligible sense of time reversal because the temporal cues like cars and pedestrians were too fleeting to give a sense of movement.

Because the filming of "truck turns," the easiest and quickest of all, still made up half of the total shooting time, particular attention was paid to the problem of turns. The most original solution was suggested by experience with synthesizing turns, namely, to film none and to reconstruct them later by interpolating between 90-degree views or rectifying 360-degree anamorphic panoramas. It must be noted, however, that while synthesized turns can cut on-site filming time in half, the postproduction time and expense is likely to surpass the savings.

The top and center sections of Figure 30 show examples of a synthesized pivot turn and a "slide turn." As a matter of fact, though not tested, this method could be used to simulate any of the other turn types filmed in Aspen. The slide turn created by sliding two abutting views across the frame was not nearly as convincing as the pivot turn with perspective correction (even though this figure may not be the best to illustrate the difference). An unfortunate but avoidable flaw in these examples is that the original angle of view was 86 degrees instead of 90, leaving a mullion running vertically between the two views. The more intrinsic problem with synthesized turns is that reprocessing involves some loss of resolution, an increase in contrast, and possibly color shift. If the procedure were carried out optically instead of computationally the loss in quality might be lessened.

The bottom section of Figure 30, showing an unwarped anamorphic view, is softer because of poor optics in the 35mm Volpi lens. The 2-1/4 format Volpi lens, tested with still pictures, produced far sharper images. The picture in Figure 31 was taken with the larger format. An interesting characteristic of these rectified anamorphic views is that resolution varies along the vertical dimension. The circumferences of the inner and outer rings are unequal, but both are mapped onto the same rectangular width. The result is noticeable in the lower part of the frames in Figure 30, bottom, where the texture is blockier.

Figure 32 compares three of the different turn types filmed. For clarity some of the intermediate frames have been omitted. The pivot turn, shown

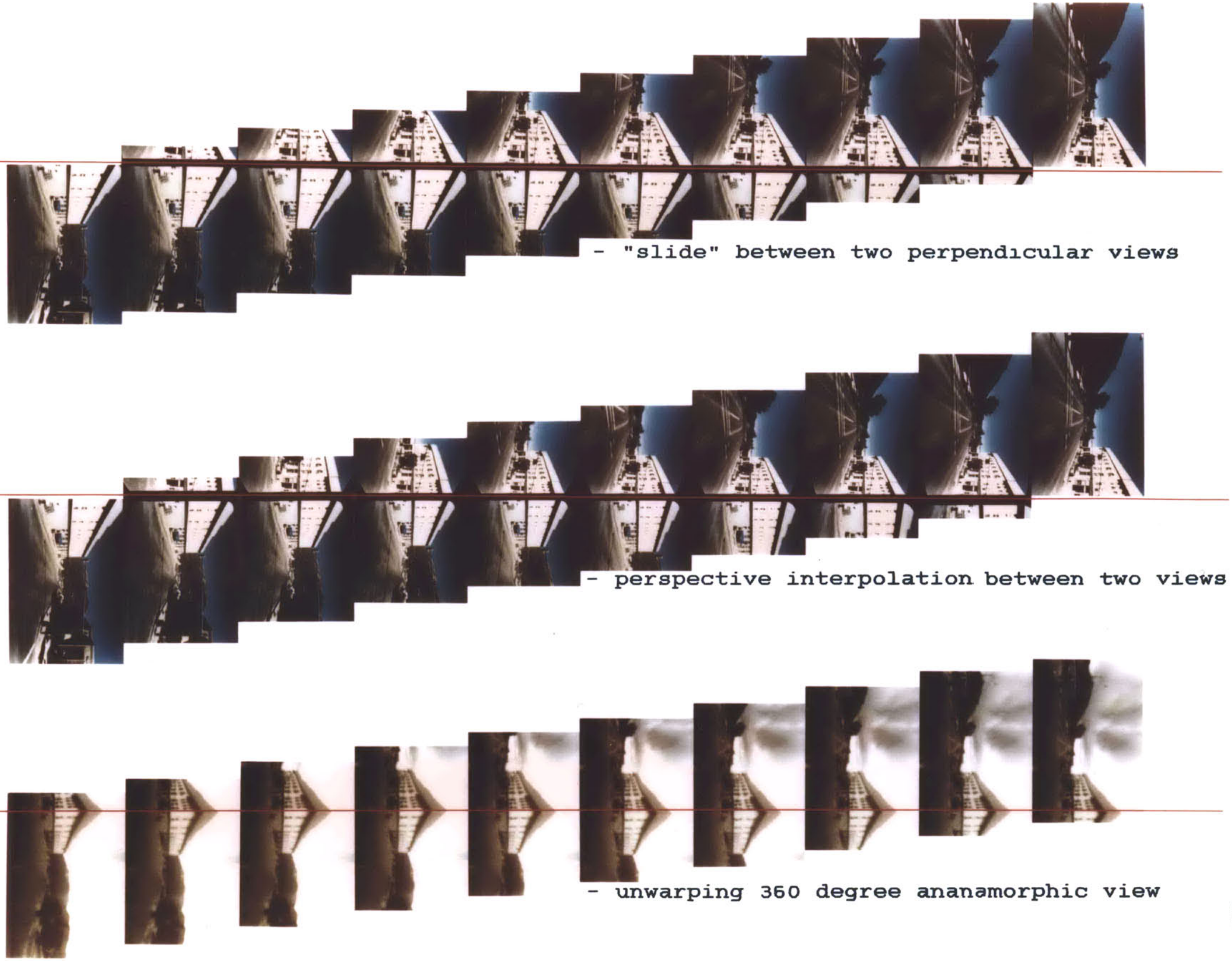


Figure 30 Comparison of synthesized turns

Figure 31 360 degree anamorphic panorama

pivot turn truck turn tracking pivot turn

Figure 32 Comparison of three turn techniques on the far left, raises an extreme case of the most basic problem in filming turns--how can the spatial context be preserved while the viewpoint undergoes a transition from one street to another? Informal tests suggested that viewers had particular difficulty interpeting the angular displacement between the two intersecting streets. A 90-degree turn could not reliably be distinguished from one of 60 or 120 degrees. A 180-degree rotation at one of the nondescript T-intersections on the rural outskirts was sometimes read as 90 degrees. In fact, it was possible to miss the fact of the turn completely, misinterpreting the rotating pan as a tracking shot moving past sharply angled streets. This phenomenon was first noticed in the Dar-el-Mara film, in which the salient parallax cues to distinguish between a pan and a track were particularly absent [Ciccone, Landee, Weltman, 1978].

There was a second problem with the pivot turn. Since the rig is already at the middle of the intersection and stopped before the camera is rotated in place to begin the turn, the view of the intersecting side streets has already passed out of the frame. This will be true for any lens short of a fisheye.

The truck turn, in the center column of Figure 32, which reproduces the exact path of the rig as it begins the turn slightly in advance of the intersection, is somewhat better; it keeps a glimpse of the side street in view as the turn is executed--but just barely.

The tracking pivot turn, on the far right in Figure 32, with camera rotation initiated independently from the rig and well in advance of the intersection, is perhaps visually the most persuasive. As has been pointed out, one of the reasons for its success is that it mimics what the automobile driver and passenger themselves do in cars [Miller, 1980]. Another feature is that the intended goal is maintained squarely in the frame during the entire turn.

The leading turn, not pictured, was also quite effective, as it combined elements of the truck turn and the tracking pivot turn. Although it lacked the full anticipatory preview of the latter, movement of the rig through the turn increased the credibility of the transition.

The practicality of filming each of these turn types is a separate matter. The pivot turn lends itself to the most efficient filming technique on two counts. It requires the rig to pass through each intersection only one time. In fact, no extra excursion is necessary if the pivots are filmed by stopping the rig at the center of each intersection during one of the four regular passes of the straight travel shooting. With the other turns a four-way intersection requires eight passes to film every permutation. Again, the pivot turn is more economical. The eight permutations are achieved with only four 90-degree rotations because each can be shown running forward or in reverse as two independent turns--right from street A onto street B or left from street B onto street A.

All of the other turns had to be filmed separately from the straight travel footage, requiring two additional passes through the entire town. Of these the easiest were the truck turns, which kept the camera fixed on axis with the film rig. Leading turns required special work by the cameraman to initiate camera rotation in advance of the intersection sighting on a point of reference along the inside radius of the turn. Introducing this kind of manual and subjective control into the operation raised the level of difficulty and yielded inconsistent results.

Tracking pivot runs were the most problematical. The cameraman had to decide when and how fast to start camera rotation so that the angle of view would become exactly perpendicular precisely when the rig reached the center of the intersection. Smoothly synchronizing these two independent motions proved to be virtually impossible with the technique which was employed. A possible solution which was never tested would be to film the turn in reverse. By this method, the cameraman would eliminate the problem of hitting a moving target by starting with the perpendicular view at the center of the crossing. Footage would then be shot, rotating the camera view backwards and keeping the side street in view, as the rig passes beyond the intersection. Showing this footage backwards should generate smoothly turned tracking pivot.

Cherry picker footage was, of course, prohibitively cumbersome to film and disappointingly unsuccessful in rendering a credible turn. On the other hand, the 50-foot height provided oblique views (intermediate between ground level travel and the aerial overview) that rendered the space in a most revealing way. The 360-degree panoramic vistas provided a wonderful way to "look around" and get a sense of the locale. And this led to the recognition of a new objective--how to provide the viewer with a more global context. Studies with subjects, which will be described later, led to the development of the comprehensive aerial overview system.

Interestingly, the pivot turn, which was marginally successful in serving as a change in path, was quite helpful in allowing a change of viewpoint. That is, it could serve as a mechanism for looking around at intersections to see what was down any of the side streets. In a related way it could also serve as a continuous transition between straight-ahead travel and the perpendicular side views.

2.2.2 Aerial overview evaluation

Preliminary tests showed that the great challenge with dynamic overviews is how to help the viewer keep his bearings. Because the visual patterns are so similar (especially with the regularity of the rectangular grid layout) the connection between two successive frames may be lost if the step between them is too great. The problem was first discovered in tracking across a rectangular lattice at a fixed scale. The lattice was shot modulo intersections, i.e., by successive frames centered on successive intersections. The result, painfully obvious in hindsight, was that no lateral movement along a street could be represented. The most salient graphic features, namely, the grid of the intersections, seemed to stay fixed like a standing wave. Only by noticing irregular features (like outlying rivers) could the viewer perceive motion. The perceptual phenomenon of integrating such successive views has been treated by Hochberg [1978].

From these findings it also seemed clear that shooting lateral movement at half-block intervals would not be effective either. It would differentiate between successive frames, certainly, but it would provide no cues to identify the direction of movement. Therefore, it was established that a third-block interval was the shortest inter-frame distance that would work. Results of filming with the lateral lattices bore this out.

A related problem was encountered with the introduction of change of scale. If the difference between steps was too great, even in the simple case of a perpendicular zoom in the cubic lattice, the relation between them could be baffling. Making the change of scale a factor of 2 insured that every element of the scene would move less than half of the field width between steps. The pyramidal lattice added another difficulty by combining lateral movement with zoom. The same principle applied of keeping element movement less than half a frame width by setting the lateral field movement proportional to the amount of zoom. Trial footage, in which this proportion was not exactly maintained, proved to be difficult to follow, it seemed because nothing in the field remained constant to serve as a frame of reference. The advantage of keeping lateral and perpendicular movement proportional was that the target intersection remained at a fixed location in the frame (Figure 25). This invariant was, of course, also a property of perpendicular zoom in the cubic lattice.*

As an extra precaution for maintaining continuity from one from to the next, nonhomogeneous elements were introduced to differentiate points of the grid network. The dozen most important landmarks of the town were

---

*The concept of the invariance of a visual element as a frame of reference is a recurring theme of the movie map. Its importance has already been seen in explaining the success of leading turns or tracking pivot turns and the failure of pivot turns. This principle will come up later in this section with regard to switching modality of representation and later in this paper with regard to seasonal registration, time registration, and transitions between front and side views.

singled out by overlaying them with colored icons. The increased legibility that resulted from this technique was demonstrated in the final product. A second circumstance that also served to enhance the ease of perception was intent--on the part of the user. The fact that the action on the screen was viewer-initiated made ensuing interpretation even more obvious.

As can be seen in the preceding discussion, highly accurate framing of the viewing field was critical to the success of filming the lattice overviews. Using a low-powered laser for registering camera position was relatively successful. Within a given lattice the consistency throughout all intersections and levels was excellent. However, imperfect registration between lattices was just noticeable in two cases. The view in the radial lattice was slightly rotated, about one degree, and the aerial photograph was offset vertically about a fifth of a block. The problem stemmed from the way in which the laser was calibrated to match the viewfinder of the camera. Alignment was done manually and visually with the lens zoomed all the way in on a "centering" test chart. However, the entire take was not shot on a single setup. Realignment introduced the slight error. A likely solution to this problem, which was not tested, would be to beam the laser back through the viewfinder and out the front lens so that the path of the beam would be locked into the camera optics. Calibration errors would be prevented. Another improvement would be to put the zoom control under computer control (as x and y positioning are).

One of the problems that has dogged cartographers most consistently throughout the centuries has been the problem of labeling. The overview map was no exception. The main problem was how to fit the identifying text into the graphics without obscuring thematic elements or interfering with other pieces of text. Compounding the problem was the convention we employed of making font size proportional to the importance or extent of the referent, typically resulting in nested text--"Chicago" appearing between the "S" and the "T" in "United States." In the case of the overview map, the solution was to position large global labels near the frame boundaries so that they would disappear upon zooming in to scales where they would no longer be relevant.

Also effective was the technique of varying font size with the distribution of labels so that in any given frame the boldest label would appear once and only once. Our one misguided method, which stemmed from a less than complete understanding of the medium, was employed in this area. The entire map was labeled before the zooms into each intersection were filmed. A more elegant procedure would have been to label only the intersection (and neighboring area) where the zoom was taking place. This would have excluded all extraneous clutter from the map by identifying only elements of current interest.

2.2.3 Cultural access evaluation

In terms of the amount of effort expended per frame in gathering, cultural documentation was by far the most time-consuming segment of the movie map (with the exception of 24 fps cinema verité filming). The reason, of course, is that everything had to be done manually; there was no equivalent of the fifth wheel on the back of the filming rig to trigger cameras. Shooting 2,000 facades took approximately 80 man hours and the 200 slide shows of interiors took approximately 100 man hours. Using 16mm film in single-frame mode instead of 35mm might have speeded up the initial shooting, but it would have greatly hampered editing. Cinema verité filming took about as long as a slide show for a given site, but, of course, the resulting footage comprised closed to 100 times as many frames. This ratio precluded all but 10 seconds of real-time footage from being allocated space on the final video disc.

In retrospect, photographing the facades manually on 35mm film was somewhat redundant. They had already been shot with the side cameras in street travel every 10 feet. Because of the arbitrary framing of the side cameras, seasonal post-registration would have been difficult if not impossible. On the other hand, a costly component of production could have been eliminated.

The importance of the "entering" transition from the facade to the interior was a lesson learned late. The jumps from exterior to interior views were often jarring. That spatial continuity was abandoned once inside was not nearly so objectionable.

Different sound types were transferred onto some of the early discs--sync, wild, and binaural. All were helpful in complementing the visuals. Binaural sound was the most interesting and probably the most effective because it conveyed a sense of ambience which directly helped to articulate the space. However, work with sound never progressed beyond the experimental stage. The final disc had no sound except for the 10 seconds of cinema verité. Part of the reason was that accommodation of sound would have increased the complexity of system implementation.

2.3 System Configuration

2.3.1 Hardware

The overall configuration of the interactive movie map had three components: computer control, audio/video processing, and user interface (Figure 33).

Computer

The host machine was a 32-bit Perkin-Elmer minicomputer, which was standard hardware in the Architecture Machine Group Lab. (Using this size amounted to real overkill, since most of the project could have been carried out on a small personal computer, as will be explained later.) The two functions fulfilled by the computer were data access and program execution. The data base consisted of three parts, reflecting the main components of the interactive movie map.

a) The basic topology of the town was modeled in terms of the connecting blocks and linking turns with the beginning and ending frame numbers on the video disc corresponding to each visual sequence.

b) The aerial overview data base was constructed, with frame numbers for each type of lattice for each intersection.

c) The name and/or address of every building in town was recorded in the data base along with its position and setback along the block. The frame number of each facade and the number of shots in the slide show behind it were also incorporated.

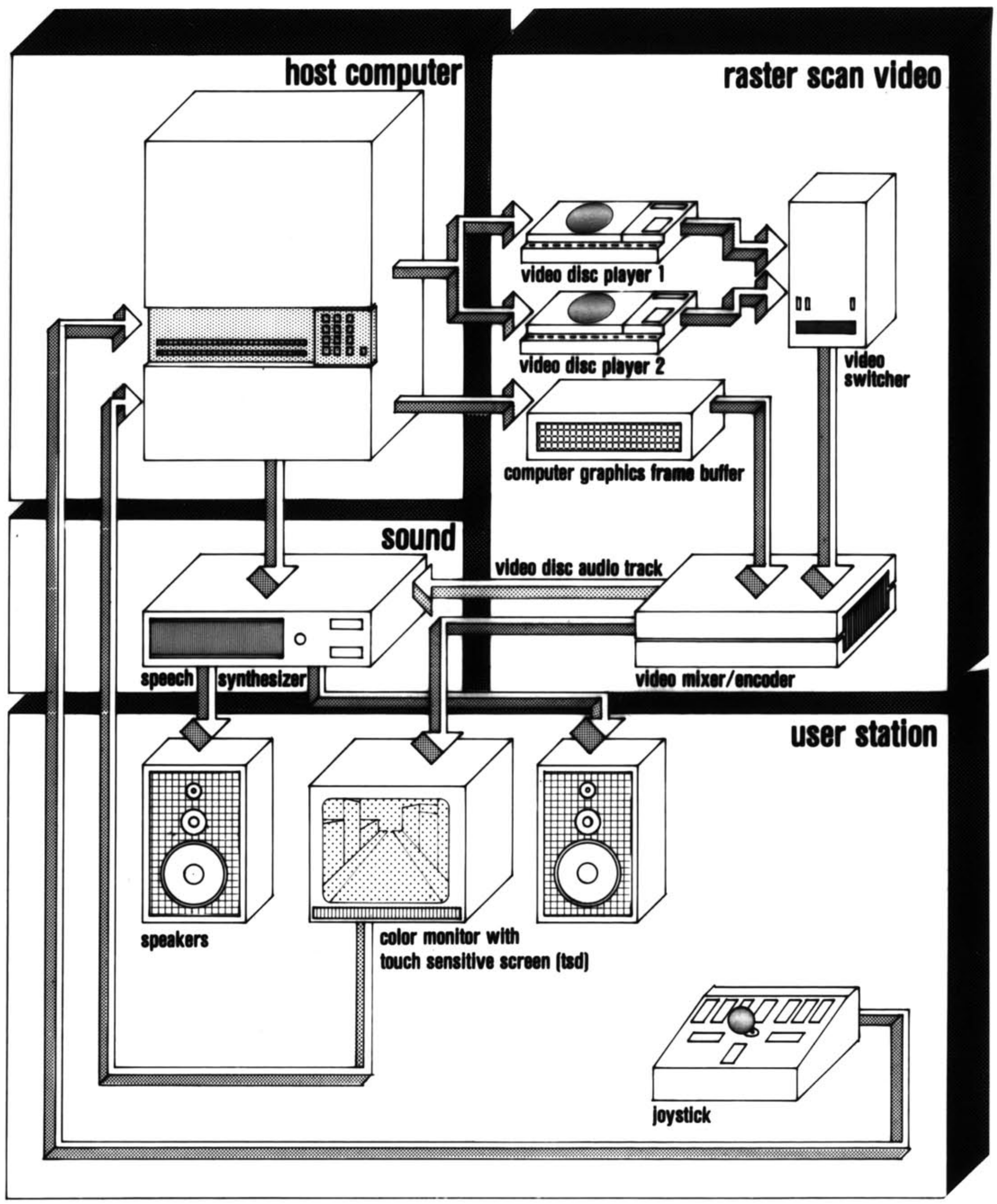


Figure 33. Movie map system configuration

The overall system control consisted of a corresponding set of three programs:

a) The basic surrogate travel routine, to handle street-level movement with control over direction of travel, direction of view, turns, speed, etc.

b) The aerial overview routine, to handle zooming modality flips between representations, route display, and route specification.

c) Cultural access routines, to compute which facade was being accessed from either of the first two routines, to initiate the appropirate audio identification, and to step through the seasonal and/or ancillary data behind the facade.

Video

Output from the computer was directed to a number of audio and video devices. The key component was the reflective optical video disc, with a storage capacity of up to 54,000 frames. The player, the MCA DiscoVision Model PR-7820, had a maximum seek time of 5 seconds. (A newer model has cut that time in half.) Commands were sent to two video disc players to control typical functions like searching starting frame number, step forward or reverse at specified rates, freezing on a single frame, etc. Any actions requiring searches (and therefore a temporary disruption of the video signal) were delegated to the video disc player not then being used by the viewer. Once the "invisible" player had accessed the target frame, the video path to the viewer's screen could be switched to that disc via a video switcher. The algorithm for swapping between the visible and invisible disc will be discussed shortly. The output of the switcher was time base corrected before proceeding in the video path.

A digital video signal was also generated by the computer and stored in a 9-bit Ramtek frame buffer. This form of graphics provided real-time feedback to the viewer in the form of menu choices, navigational aids, and ancillary data. The computer-generated graphics were encoded into NTSC and overlayed on the video disc signal with the keying function on a video mixer. The composite picture was sent to the viewer's monitor. All of the functions described were under computer controj except for the key level of the video mixer.

Control of sound also encompassed both analog and digital formats. The former consisted of the video disc audio track sent to a speaker at the user's station as part of the cinema verité sequence. The latter was a somewhat primitive Votrax speech synthesizer, chiefly used to identify street and building names.

User Station

The focal point of the user station was the color video monitor which displayed the movie map images (Figure 34). It had a touch-sensitive screen (consisting of a transparent plastic overlay by Elographics). This screen, recording the x-y position of a touching finger, was the user's interface for controlling his journey through the movie map. In an earlier version an x-y force joystick with button functions was the input device. The keyboard terminal was all but eliminated from the user station. Its only necessary function was to start the program running. All other functions were implemented through menu control on the touch-sensitive screen. Finally, of course, the station included audio speakers.

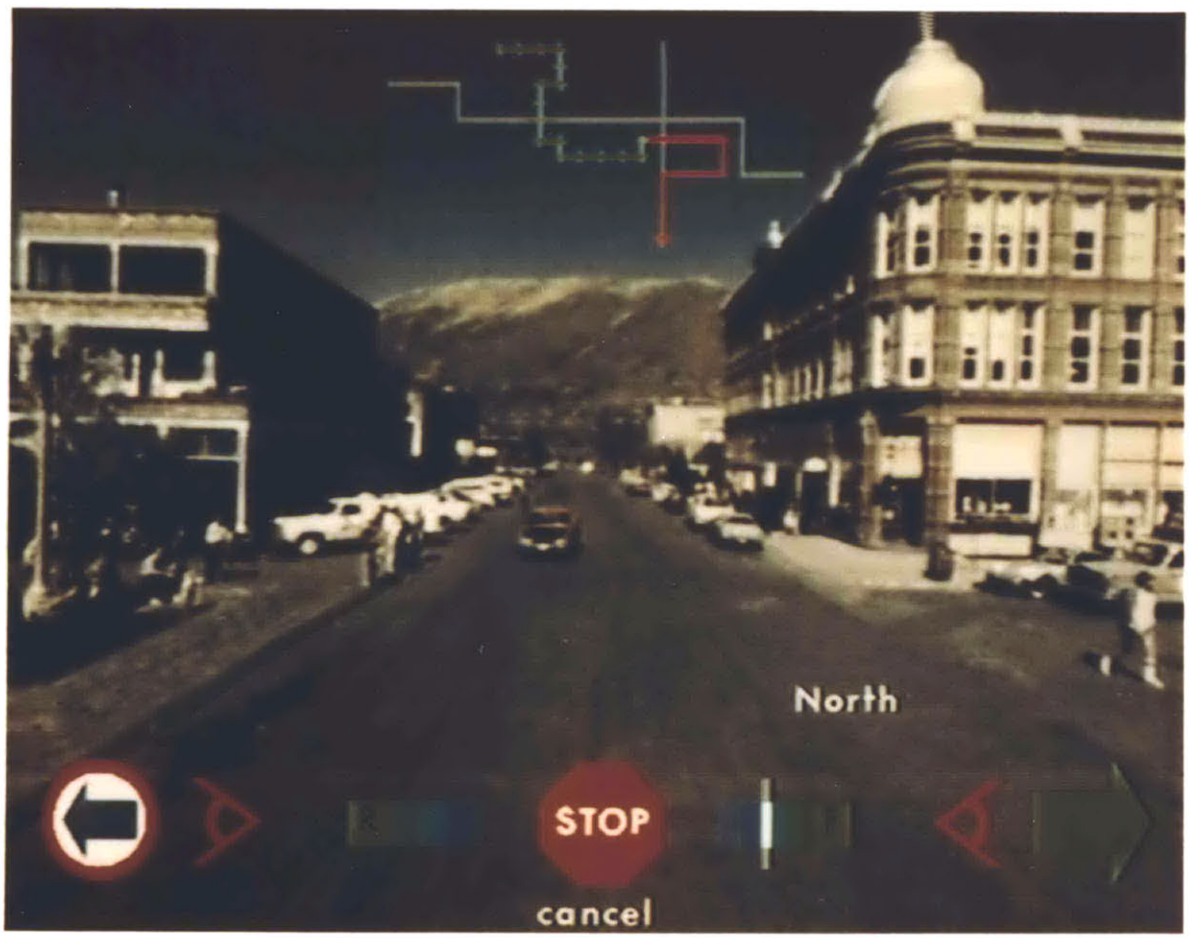


Figure 34   Surrogate travel - user station

2.3.2 Interaction

2.3.2.1 Surrogate travel

Speed and Direction

The virtual instrument control panel which the viewer used to engage in surrogate travel was a computer-generated overlay (Figure 35). Direction of movement was determined by the blue/green bars flanking the stop sign. "F," on the right, indicated for ward, "R" was reverse.* Speed of movement was a function of how far along the bar the user pressed, with a tick mark speedometer providing feedback. In response the visible video disc was step-framed at a variable rate up to 10 fps, or an apparent speed of 70 mph. Backing up was achieved by simply stepping the video disc in reverse.

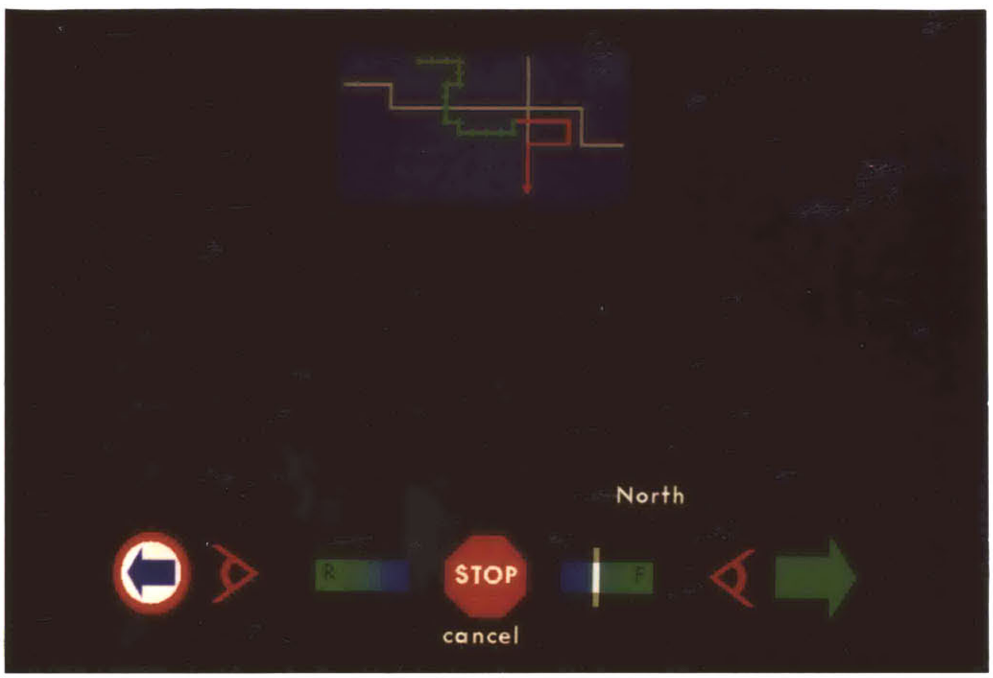


Figure 35. Computer overlay for surrogate travel

*Other forms of input control tested were a joystick and a position-seeking device by Polhemus called ROPAMS (Remote Position Attitude Measurement System), worn on the body.

An example of forward travel with the computer overlay is illustrated in Figure 36, reading top to bottom, left to right. A comparison of travel in animation, fall, and winter is shown in Figure 37.

Side Views

Side views were invoked by touching the "eye profile" icon beyond the speed controls. TKree methods of accessing perpendicular views were tried. The most immediate was simply cutting to the view "out the window" at any point along the street. The sequences in Figure 38 show the forward view in the center panel and corresponding perpendicular views to the left and right. The alternative to this abruptly discontinuous movement was to delay the shift to the side view until reaching the next intersection. At this point a smooth transition could be provided by a pivot turn, corresponding to a rotation of the surrogate driver's head. Or the surrogate vehicle could "crab" sideways by cutting to the side view along the cross street (which would match the front of the original street, since the grid layout was rectangular). In this last case, calling for a side view would also entail changing the path to an intersecting street.

When in a side view, control for direction of movement was screen-relative, i.e., the two blue/green rectangles were used for left and right. Instruction to exit back to a front view was indicated by touching the side view icon again. The transition was made in the reverse fashion--making a pivot rotation at the next intersection, proceeding forward at the next cross street, or cutting back directly. Notice that in the last case, two successive calls to the side view would effect a 180-degree U-turn.

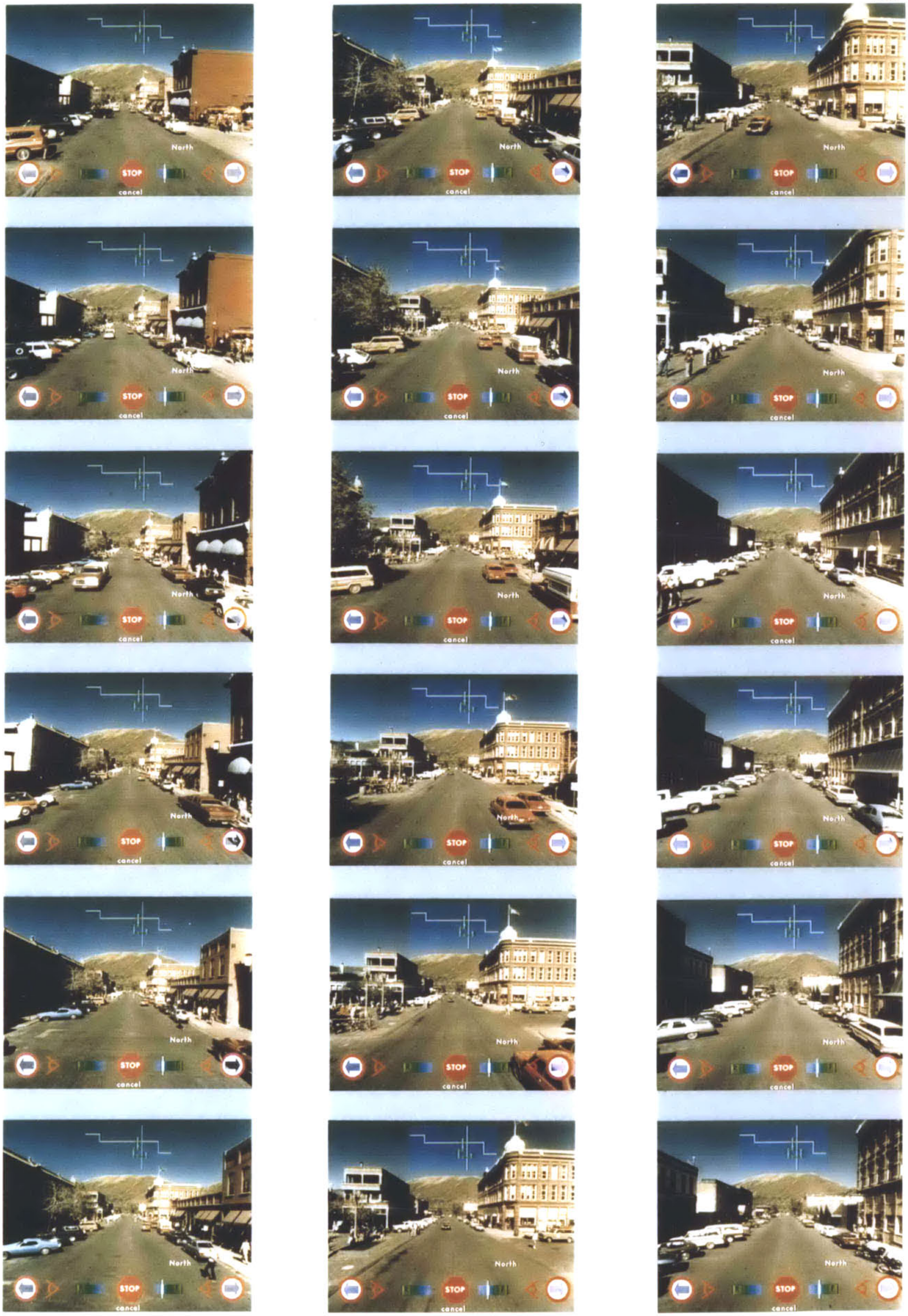


Figure 36 Surrogate travel - forward sequence

animation fall winter

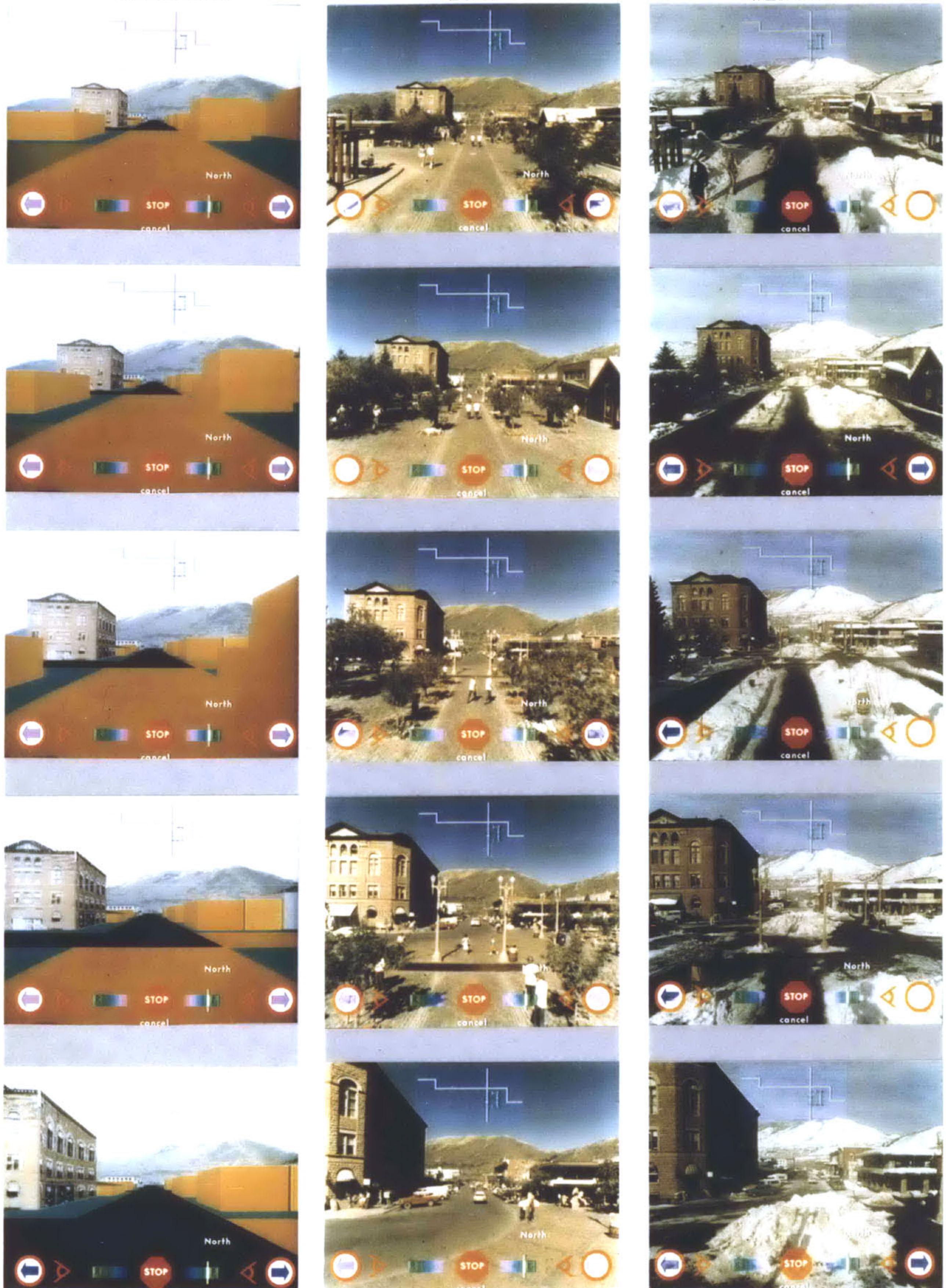


Figure 37 Surrogate travel in three matching "seasons"

Figure 38 Surrogate travel - front and side views

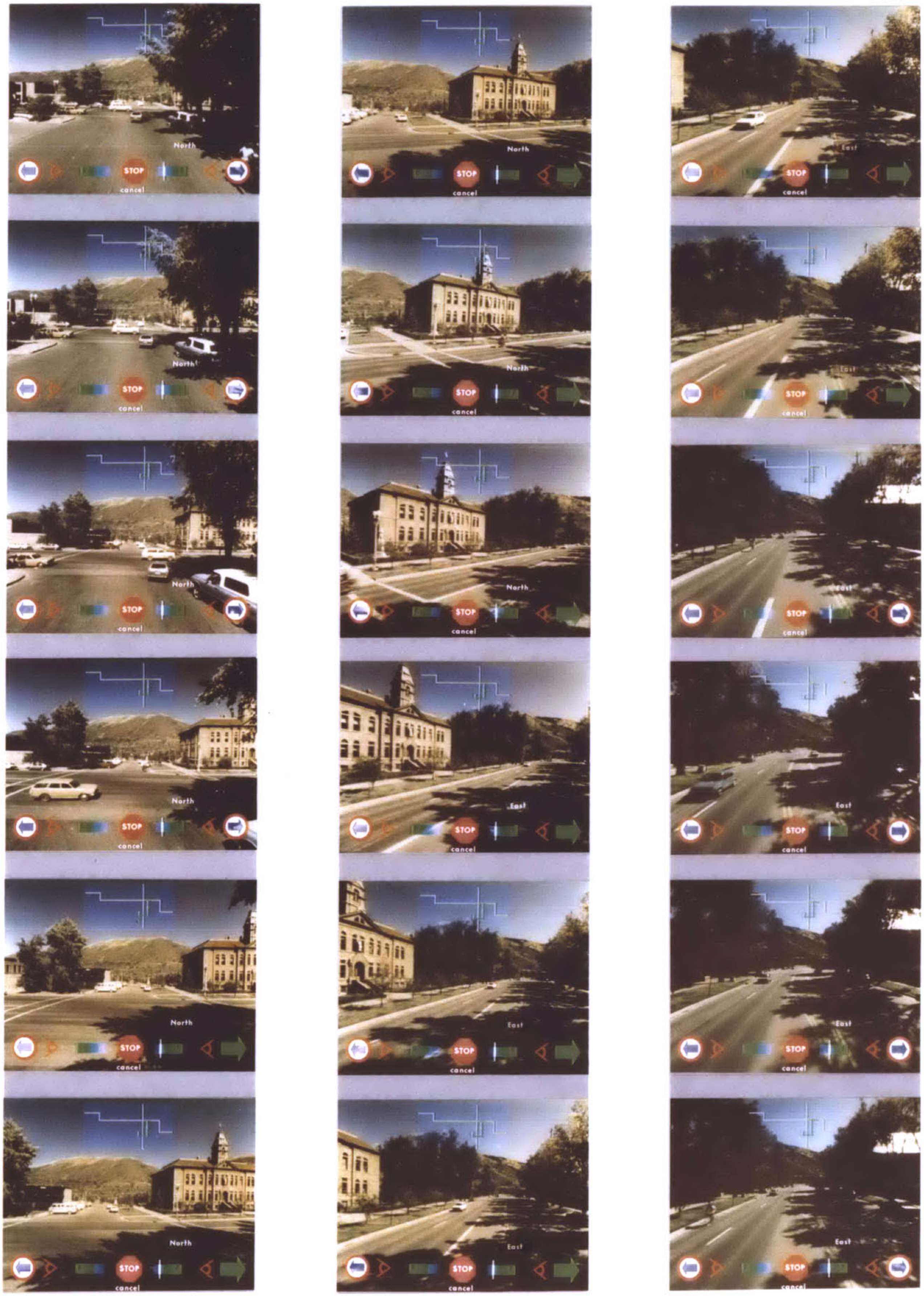


Figure 39 Surrogate travel - turn sequence

Turns

Turns were signaled by the arrows on the far left and right. As shown on the right in Figure 39, the arrow turned greenin feedback to the user to acknowledge an intention to turn at the next corner. Figure 40 illustrates the manner in which turns were edited onto straights to produce a continuous sequence. The shaded links indicate the pathy actually followed.

The second disc was, whenever possible, staged ahead at the starting frame of the next turn onto the cross street. As long as the user continued to travel straight, the second standby disc staged ahead, anticipating the next possible turn. If the user elected to turn, the switcher would cut to the second disc when the straight sequence reached a match frame with the turn sequence (as recorded in the data base). At that point the original disc became standby as it searched to the starting frame for straight travel down the side street. When the turn reached the match frame, the switcher but back to the original disc now on the new street. (In case the standby player was not yet cued up, the visible player paused on the match frame until the search was completed.) The action of the discs is diagrammed in Figure 41.

The unshaded segments refer to the standby (or "invisible") player. The play modes, which the user sees, are shaded. Left and right turns

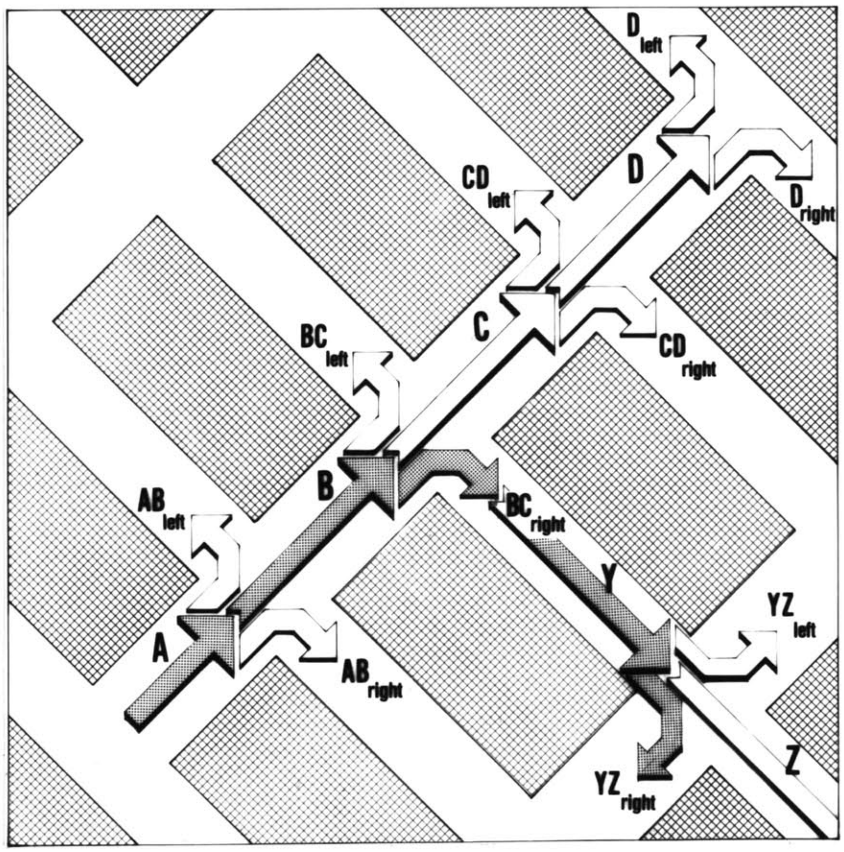


Figure 40. Sequencing straights and turns together

have been laid out "head-to-head" to give both options to the staging player on a single frame. In fact, this means four options are always available with two players, two for forward/backward and two for left/right (Figure 42).

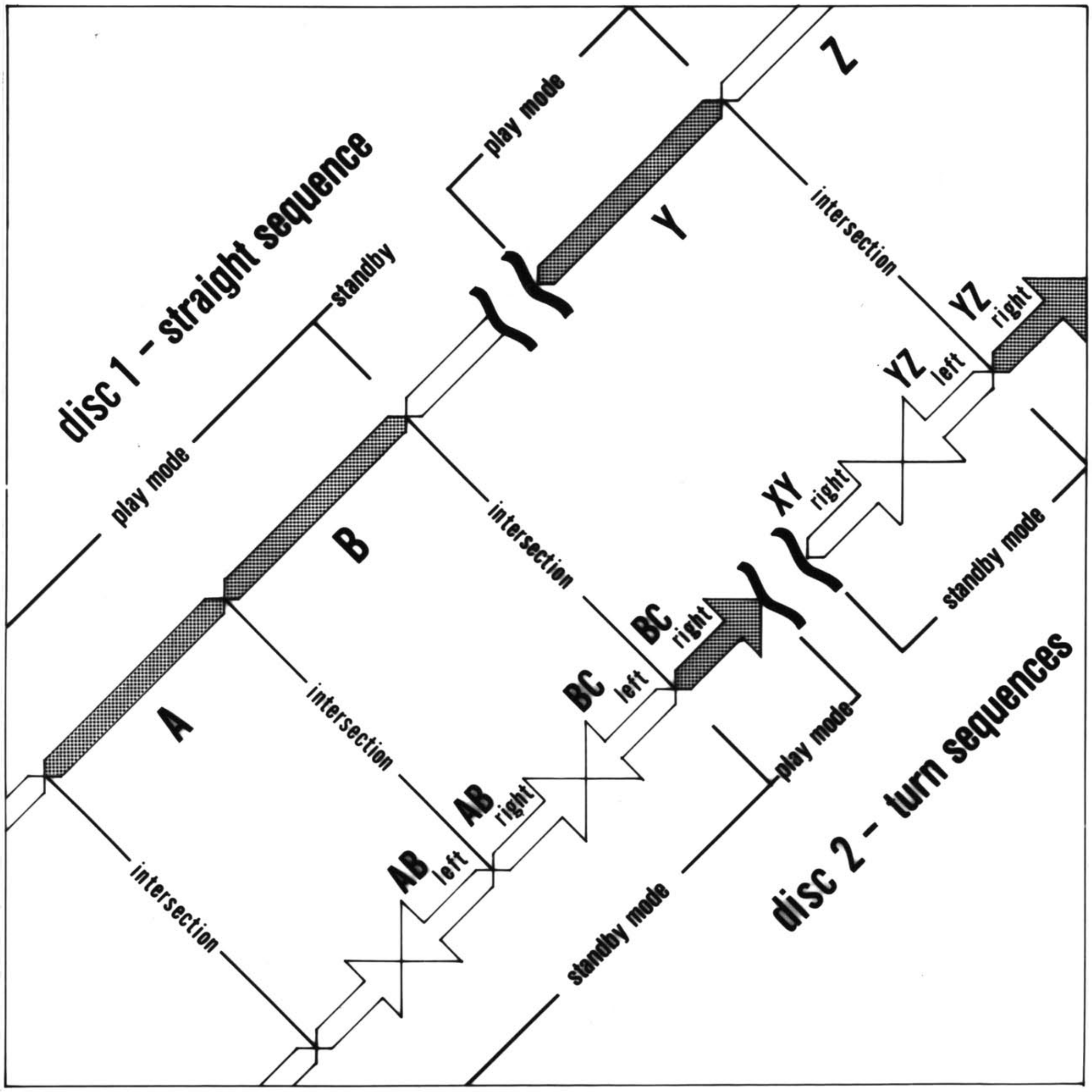


Figure 41. Standby and play mode with two discs

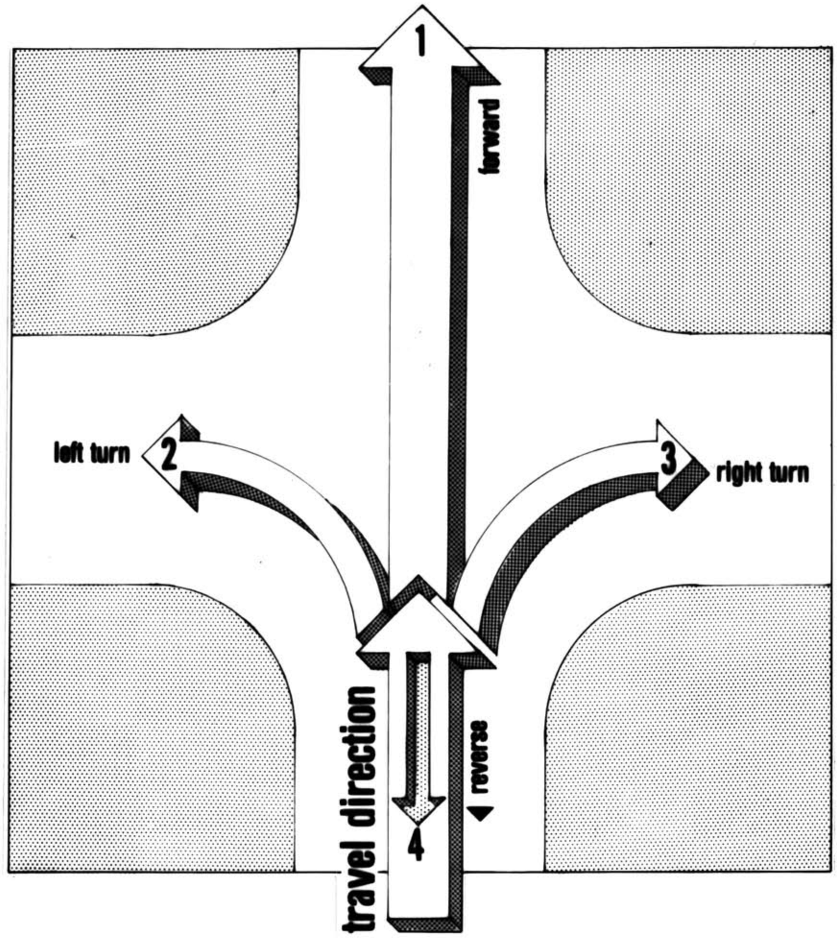


Figure 42. Four surrogate travel options

The Two-Disc Algorithm

Three simple rules account for the basic procedure employed in all phases of the interactive movie map for accessing material on the video disc:

a) The alternating two-disc format was the technique necessary to allow all cuts to be made without interruption of the picture to the viewer.

b) Staging the invisible ahead was a strategy designed to minimize, if not eliminate, any search-time delays.

c) Being able to play both discs either forward or in reverse made up to four different branching options instantly available to the user.

2.3.2.2 Aerial overview

Schematic Map

The blue rectangle at the top of the surrogate travel view (approximately where a rear-view mirror would be located) provided the user with a condensed diagram of the route just traveled (Figure 35). The two major cross streets in the town were drawn in to provide a frame of reference. Like the other graphic overlays, the schematic map served to provide more than just feedback to the viewer--it was also an active button which could be touched to enter "map land." Computer-generated graphics were again used as an overlay (Figure 43).

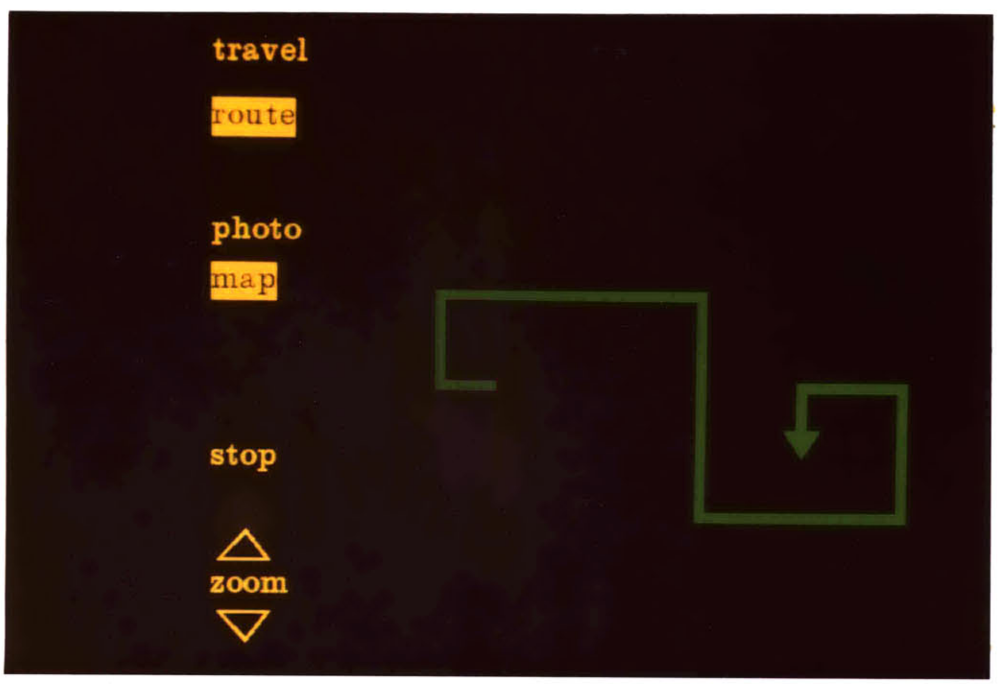


Figure 43. Computer-generated overlay for aerial overview

Route

The introduction to map land was an overview in the form of a conventional-looking map. However, it had some unconventional properties, the first two of which resemble those of the working memory in Kuipers' TOUR model [Kuipers, 1977]:

a) It displayed where the user had been.

b) It marked where he was currently.

c) It enabled him to specify where he wanted to go next.

The route which had been followed so far was traced out, with intensity a function of recency (Figure 44). The arrow at the end served as a "you-are-here" pointer. Selecting the route button from the menu on the left enabled the user to trace out a route he wanted to follow when he resumed surrogate travel. This autopilot mode invoked a chauffeur to guide the surrogate passenger around. On the schematic map the prescribed path, traced in red, turned green, block by block, as the trip unfolded.

Scale

Zooming in for a closer, more detailed view of any part of the map was initiated by touching an intersection in the desired area (Figure 45). If the pyramidal lattice mode was in use, the target intersection remained stationary under the user's finger as the scale changed. In the cubic lattice mode the target intersection was first centered at the top level by playing the appropriate radial arm sequence to "helicopter" over to it (Figure 46). Then the perpendicular zoom into that location was played. Using the same alternating disc technique described for surrogate travel,

Figure 44 Aerial overviews with route plotter

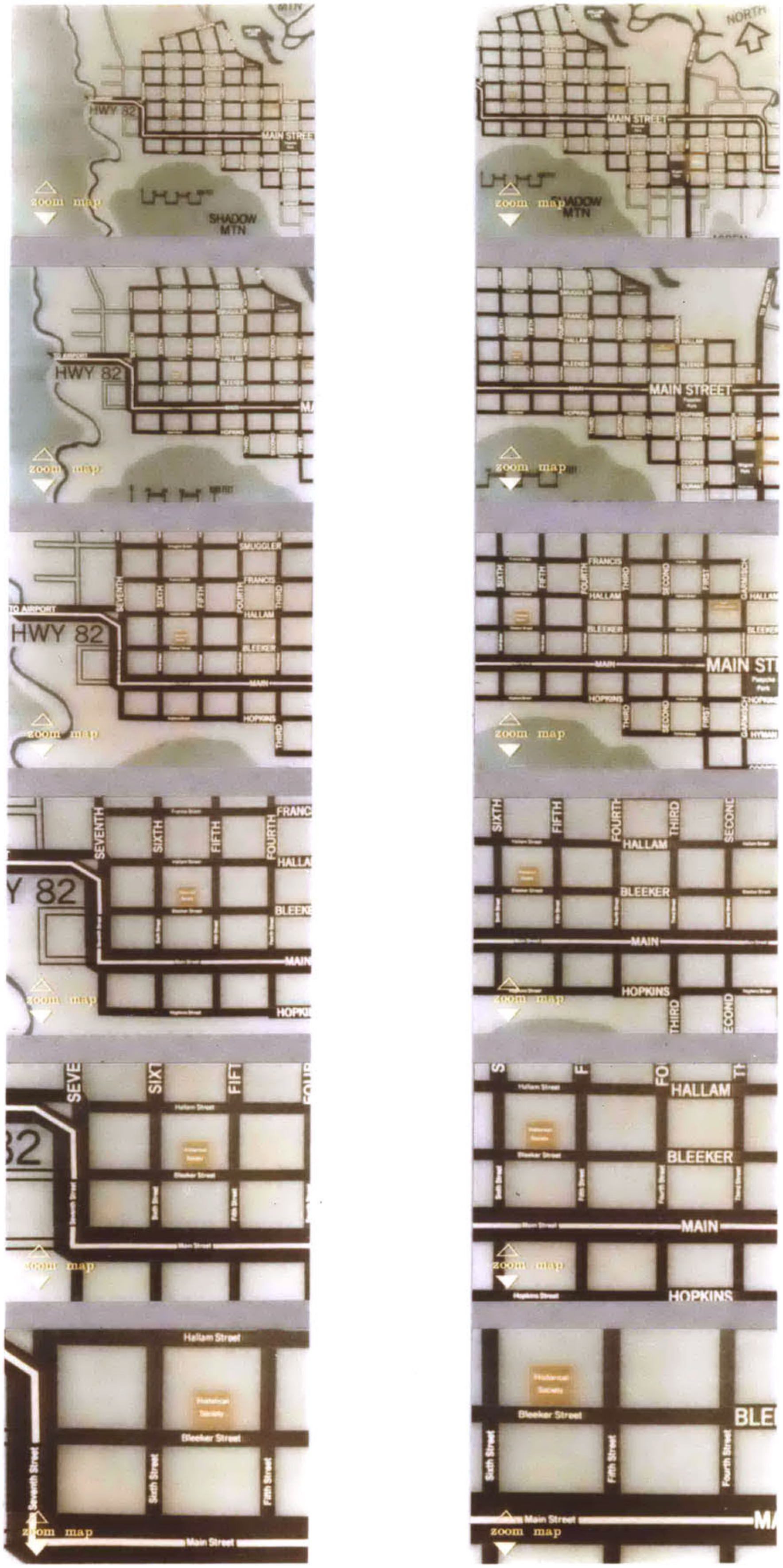


Figure 45 Cubic zoom and pyramidal zoom

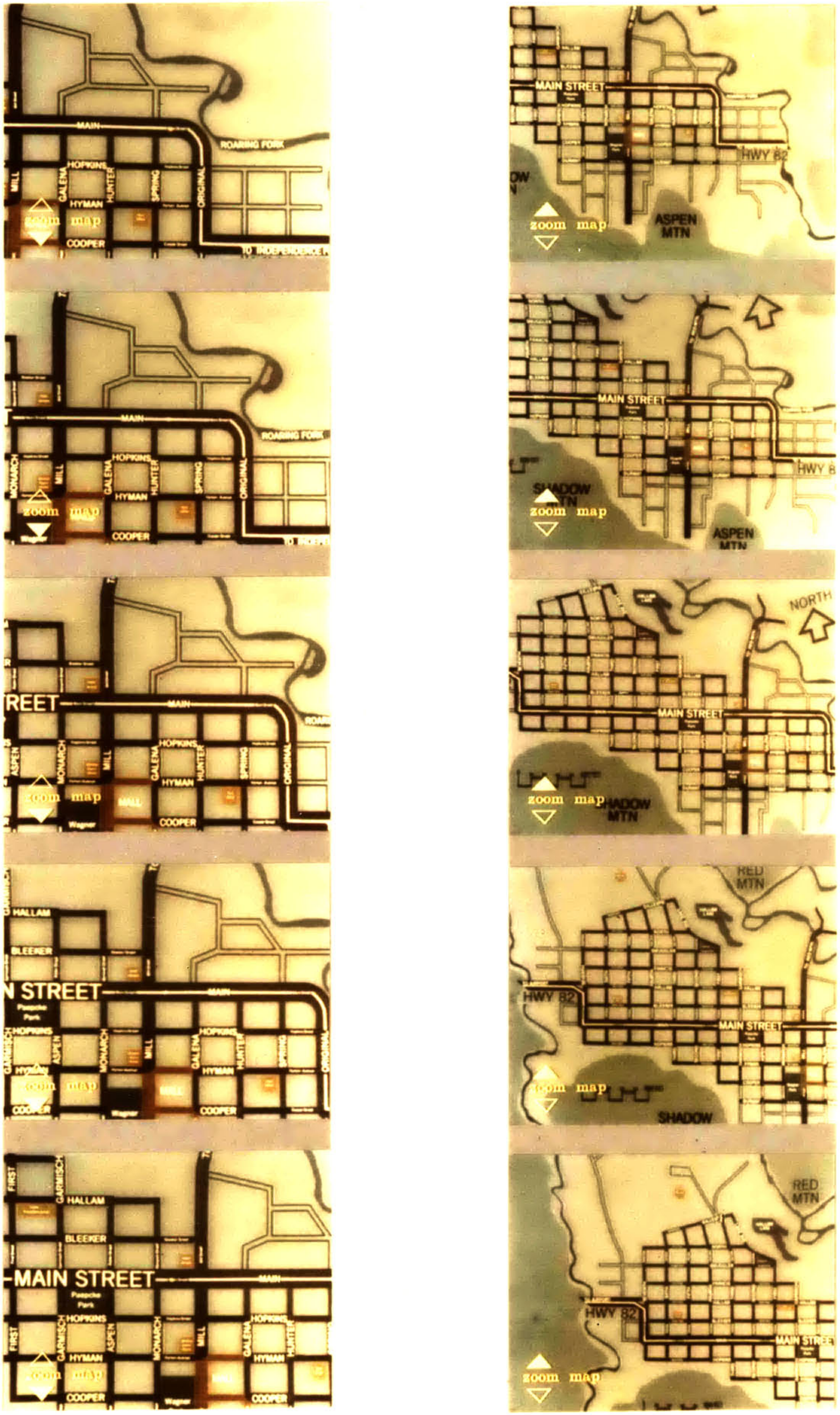


Figure 46   Lateral and diagonal movement across map

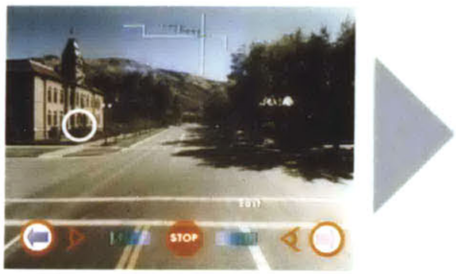
visual continuity was maintained by switching from one disc to the other on matching frames. The upper zoom button returned the viewer to the top level by reversing the video disc sequences. Zooms could be halted at intermediate levels of scale by using the stop control.

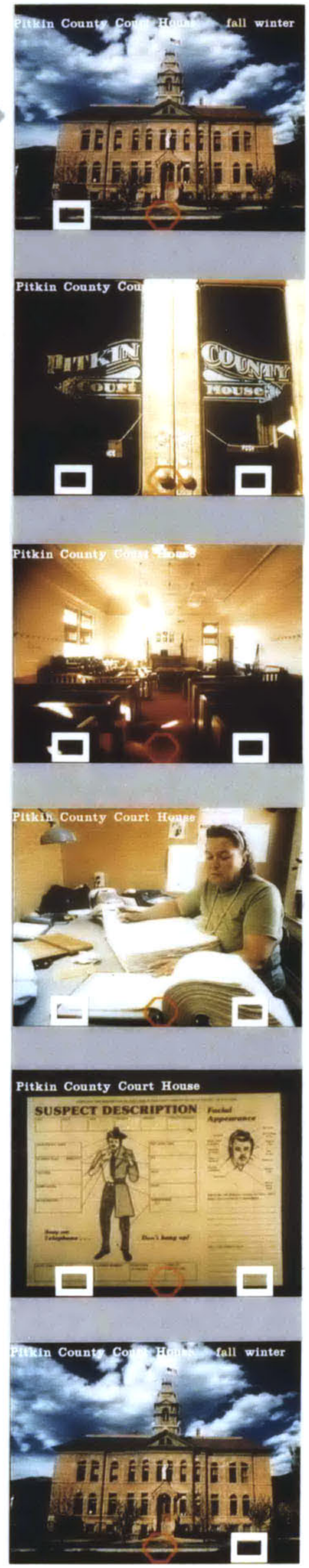
Changing Representation

At any scale in any lattice the mode of representation could be switched between map and aerial photograph through menu control. Flipping to the landmark map was possible only when zoomed all the way in. In this last case, three options were simultaneously available because the regular map and landmark map were on adjacent frames and the photograph was cued up on the second disc.

2.3.2.3 Cultural access

Ancillary data could be accessed for every location in both travel land and map land. In the former the user simply stopped and touched any building (or other built feature) in view. In map land, sites could be selected when zoomed all the way in. In either case the viewer was presented with the facade of the appropriate building and a verbal identification (Votrax) by the name and/or address. Depending on the site up to three options were then instantly available. Return to surrogate travel was always possible. In many cases a season knob, which was toggled by pushing the appropriate time of year, caused the scene to alternate between fall and winter in perfect registration. The leaves vanished, the snow appeared, but all permanent fixtures remained precisely in place. Finally the user could "stop in for a quick visit" by starting the slide show (Figure 47). The sequence could be stopped or reversed at any point.





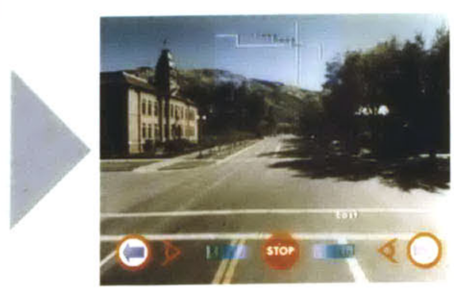


Figure 47

Cultural data access

2.4 Operation Evaluation

A human-factors-type analysis was part of a general study based upon the responses of subjects, which will be elaborated on in Chapter 5. Some of the more technical aspects of the operation will be addressed here.

2.4.1 Surrogate travel

Turns

The basic technique that made the entire concept of the interactive movie map possible was a convincing one. Cutting between different pieces of footage to edit in a turn on the fly produced a well-matched sequence in terms of lighting and registration. Three procedures for making the cut between straight-ahead movement and the turn footage were tried:

a) The shot used as the first frame of the turn was matched as closely as possible to thelast frame of the straight.

b) The shot used as the first frame of the turn was off-axis (i.e., starting the angular displacement) and approximately 10 feet farther along than the last frame of the straight.

c) The shot used as the first frame of the turn was approximately 10 feet farther along than the last frame of the straight, but its angle of view was still on-axis.

Procedure (a) was objectionable because it resulted in irregular speed of movement. The double frame at the cut point caused motion to be halted (or even slightly reversed) for a moment. However, in the case of a single-disc system (like the original M.I.T. corridor map) this method was preferable, because the fact that the blank screen interval was bracketed by matching pictures helped maintain a frame of reference. Procedure (b) maximized discontinuity by combining two changes in a single edit--shifting scenes by a jump cut and beginning angular displacement. Procedure (c) proved to be the most successful. It maximized continuity by separating the two changes to make for a longer and smoother transition.

The one problem that remained virtually inescapable was the discreteness of movement, owing to the 10-foot interval and the irregular pixillation of moving objects caused by the often inconstant speed of the filming vehicle. The result was a disconcerting separation of temporal and spatial linearity. More constant filming speed, achieved by the third trip, reduced the irregularities somewhat. Tests at 5-foot interframe intervals produced less pronounced effects. The "now-you-see-it-now-you-don't" phenomenon was exaggerated in cuts to and from turn footage, when transient objects appeared or vanished. Naturally, the problem was minimized in the less trafficked parts of town like the residential areas. Surprisingly, it was more noticeable in the final implementation of surrogate travel than the first. Apparently the more regular flow of straight travel achieved in the final shooting made the discontinuities in turns that much more conspicuous.

User Interface

Since one of the premises in designing the interactive movie map was that of a naive user should be able to use the system easily and comfortably with only a few seconds of direction, the responsiveness of the man-machine interaction was very important. In spite of several refinements to the interactive movie map two kinds of occasional hesitations remained in the system. One was due to video disc search-time delays; the other was a result of computational drag. The irregularity of such delays prevented the user from adapting quickly. Fortunately, the brevity of the interruption, maybe a second, and menu feedback minimized the problem.

The first input control device used with the system was a joystick with a set of buttons. Speed and direction were controlled by pushing forward and backward, turns by pushing left and right. Spatially arrayed buttons were used to toggle between perpendicular views. The joystick turned out to be a grossly inappropriate interface. Its degrees of freedom did not correspond to those available in the movie map. Continuous excursion in x and y produced the effect of a two-dimensional flight simulator with side-to-side motion available everywhere. But the movie map, of course, restricted movement off the street center axis to crossings. The inconsistency proved confusing for naive users, who tried to steer themselves around corners even after being told that the x-axis was simply a turn indicator. The second main shortcoming with the joystick was the lack of feedback. The only acknowledgement that the system had understood a command were red lights on the buttons, which divided the user's attention between the display and the joystick.

The touch-sensitive display with computer-generated graphic overlays was much more intuitive and direct. The viewer could perform all operations and experience all aspects of the movie map without taking his eyes off the screen. Immediate feedback was provided by the menu symbols. The small price paid was the extra clutter of the instrument control in the picture, which counteracted a sense of depth perspective. Using on/off menu buttons for binary functions like views and turns, and sliders for continuous functions like speed, achieved matching degrees of freedom between input and output. The fact that using the touch-sensitive display for such simple functions disguised the real nature of the medium was beside the point.

Operation tested the viability of other interface media. The ROMPAMS body tracking device worked but suffered from the same problem as the joystick--too many degrees of freedom. Voice input, on the other hand, not only permitted the right nuance of control but also avoided the limitations of other versions. It separated the driving controls from the visual display without requiring interference from the user. When the large, 13-foot diagonal rear projection video screen was used for surrogate travel, the reincarnation of Aspen achieved its greatest "presence." The size of the image relative to the viewer conveyed a feeling for the scale of the town that was missing on smaller displays. When the viewer was required to raise his eyes and tilt his head back to focus on the mountains, their height was more convincing.

2.4.2 Aerial overview

Schematic Map

The schematic map was useful in clarifying toplogical ambiguities. For example, it could show that the user had just circled a three- or four-sided block. It also displayed directional relations between different parts of a route. The coordinate axes provided by the two major cross streets helped set a global frame of reference. Missing were the boundary streets, which would have added valuable context by defining the outline of the town. The placement of graphics at the center top of the screen was not an intrusion in front views, as it obscured only sky. Unfortunately, it was not so unobtrusive in turns and side views, where buildings filled the same part of the display. The main advantage of the schematic map was that it was small enough to be included in the same picture as the travel footage. By the same token, its size precluded any significant amount of detail, necessitating full-size overviews.

Route

Early versions, which traced the traveled route quickly and in a uniform color, had no way of disambiguating the chronology of self-crossing paths. Tracing the route out more slowly in a human time frame and increasing the intensity of the color as a function of recency made even complicated routes easy to follow.

Scale

There were interesting tradeoffs between using one or the other of the two lattices for zooming in. The pyramidal lattice was effective for maintaining the frame of reference in two ways. Because it kept the target intersection fixed where touched, the user had little difficulty following the change of scale. The more interesting consequence of this property was that the intersection remained off-center in the frame in exact ratio to the amount it was off-center in the town. Thus the user had access to global information about the relative location of an intersection even when he zeroed in on a local scale. The disadvantage, on the other hand, was that the local area around an intersection was never displayed symmetrically. The viewer saw proportionally fewer outlying features toward the town boundaries than he did interior features, and no features past the boundaries.

Since the cubic lattice zoomed in perpendicularly, it succeeded in showing an omnidirectional area around each intersection. The advantage was that information about features exterior to the town boundaries was included (Figure 45). This lattice also insured that the object of interest was centered on the screen, whether it was an individual intersection or the whole town. However, the transition--the radial arm--that was required in this case introduced two drawbacks. The viewer had a harder time following the target intersection during the transition because it did not stay fixed on the screen. Secondly, this zooming technique was slower because it doubled the number of video disc searches required.

Changing Representations

Ability to change the modality of the medium at any scale between map and aerial photograph provided a powerful handle for the viewer. Duality of representation allowed the integration of different understandings of the same thing (one of the conceptual threads of the movie map). The map obviously presented a schematic diagram of the street organization and important landmarks while the photo provided much more detailed views of actual buildings and surfaces. Although rich in detail, the content of the photo was much less differentiated than the abstracted graphic of the map. Flipping between the two representations was most informative at intermediate levels of scale. At the top of the scale, viewing the aerial photo was not very profitable because features were too fine and too similar for good legibility in video. At the bottom of the scale, the features of the map were gross and provided little new information. Consequently, the most constructive use of the aerial overview consisted of a sequence that began with the map, zoomed in part way, switched to the photograph, and then zoomed in the rest of the way. The additional flip from aerial photo to landmark map available at the bottom level did for buildings what the flip to the full map did for streets--it provided identifying structure and labels. The fact that the photograph was black and white worked both pro and con. Lack of color resulted in more uniformity of field and fewer distinguishing features. On the other hand, adding color (in the form of computer-generated overlays or preselected landmarks) was more effective in spotlighting targets than it would have been had the entire background been in color.

2.4.3 Cultural access

Facades

Facades served as the gateway transition between travel land or map land and cultural access. At a pivot point between the spatial/visual continuity of the former and the smorgasbord sampling of the latter, facades had to work as a transition in both directions. Whey they were used as a lead-in to cultural access there was no problem because discontinuities were the rule. However, in cutting to and from surrogate travel, the facade view presented the same difficulty that obtained in cutting from a front view to a perpendicular side view: the jump could be disorienting to the viewer. The one factor working in the viewer's favor was his intent. He identified the specific point of reference which should be cut to. There was no mediating solution, like using a 90-degree turn to provide transition, because the new view couldn't wait for the user to reach the next intersection. A quasi-legitimate transition was provided by selecting the facade, after the user had started traveling, via a side view. In this case the side view was so similar to the facade that it raised the question of why a transition was even necessary. Discontinuity in switching from the aerial photo was equally problematical, with its own quasi-legitimate solution. In one test case the roof of the building was used instead of the facade as the transition to the slide show.

Season Knob

Flipping back and forth between seasons with the facade in perfect register was probably the most intriguing example of the synergy that can result from seeing the same thing in two different ways. The change of a scene from fall to winter provided some very specific information. It could show which trees were deciduous or could articulate the depth of overhangs through snow buildup. It also provided a general sense of the process of visual metamorphosis that a building goes through in a changing environment. The expectation was that the viewer would be better prepared as a result to correcly identify a building in novel conditions of weather, lighting, etc. No direct evidence on this point was gathered.

Ancillary Data

The only control given the viewer in running the slide shows was the choice of forward, reverse, or stop. There was no flexibility to allow branching or personally assembled sequences. However, the richness in the quality and variety of imagery introduced in the course of a slide show was usually quite high (and a direct result of the discontinuities which distinguished it from surrogate travel). The obvious function of the ancillary data was to provide the user with relevant information about a building of interest--the size of the interior space, type of goods or services provided, prices, hours of operation, etc. Two important consequences of this data in the acquisition of spatial knowledge were not fully appreciated in advance:

a) The variety of imagery which could be accessed almost anywhere the user drove kept his interest level much higher than if he had only seen the increasingly similar surrogate travel views of street scenes. A longer attention span and lower fatigure meant that the viewer could travel longer and learn more of the town.

b) More important, the unique experience associated with each slide show enabled the viewer to acquire a rich set of associations to help him remember each building visited and to distinguish it from all others. Since the capacity to learn about a landmark is a function of the number and variety of experiences with it, slide shows performed a significant function in this regard.

CHAPTER 3

The Nature of Space and Time

- Magritte

In Boston you can go around the block
and end up on a different street.

- *Anonymous*

A rose may be a rose may be a rose, but in the case of imagery, each medium for representing reality has its own laws of nature. Precedents exist for the interpretation of images from the retina, the film camera, the television screen, but surrogate travel is without precedent in common experience. In some ways it seems very natural; the user decides where to "drive," where to turn, when to speed up, slow down, stop. But in some ways it is counterintuitive; it is dynamic, but the environment is unevolving and scenes are repeated. It is interactive but not live. Space and time do not obey the familiar rules. This section will examine the properties and consequences of this unique model of reality.

In tracking down some of these subtleties, it is important not to lose sight of the general effectiveness of the surrogate travel system. It is remarkable that apparently any group of naive users may sit down in front of the television display and without prompting, interact with the system (and more importantly with each other) in completely conventional terms. A typical exchange between "front seat" and "back seat" drivers might go as follows:

> "Slow down" and
>
> "Take a right at the light. Oh, you went too far--go around the block."
>
> "How do I get to the downtown mall?"
>
> "I'm not sure exactly but downtown is that way" (pointing offscreen at the hanging plant).

It seems clear that whatever the differences, the techniques of surrogate travel are compatible with the experience of the real environment.

3.1 Space

It will prove useful to distinguish among four incarnations: real space, virtual space, playback space, and cognitive space.

a) Real space is, of course, the three-dimensional euclidean world.

b) Virtual space is space as it is reconstructed from a video disc during surrogate travel. It encompasses both how the space is filmed and how it is presented by the system.

c) Playback space has a much more limited meaning than virtual space. It refers only to the spatial arrangement of the user station.

d) Cognitive space is the internal model built up in the mind of the viewer.

3.1.1 Virtual space

Real space is generally experienced as continuous and uniform. People are accustomed to moving around in it with smooth and minute control. The virtual space, on the other hand, is not continuous or uniform and presents some interesting modular characteristics. It has the following properties:

a) In any linear dimension space is discrete.

b) The two-dimensional plane of the town space is a discrete network of linear paths, where the unit interval between paths is much larger than the unit interval along a path.

c) The volume above the town is made up of a three-dimensional lattice of points. The grid of the network matches the ground plane intersections.

Since surrogate travel presents a succession of frames take 10 feet apart, movement is in steps, and the apparent discontinuity is a function of plahing speed. At a maximum rate of 30 frames per second the discontinuity is minimized. But smoothness is then achieved at an apparent velocity of more than 200 mph.

Discreteness of space is also a function of mapping from the two-dimensional street network on the video disc to the linear path traveled by the user. Each change in direction, accomplished by editing in a turn, introduces two potential discontinuities, at the beginning and end of the insert. If the camera orientation before and after the cut do not match perfectly, there will be a slight jump in spatial registration. Besides the same angle of view, both cameras must have been in the same position to make a perfect cut. Since any given block may not be evenly sectioned into 10 foot intervals, the difference in position between the cameras could vary from 0 to 5 feet. The result is an occasional, perhaps imperceptible, irregularity in the modular unit of distance.

As informative as characterizing the space in terms of where the user can go is to examine where he can look. Wherever the user travels, a finite number of views are available. So virtual space is also directional. Mapped onto the space is a pattern of vectors representing the four directions of view associated with each point. Pivot turns, rotating in 10-degree intervals, increase the number of views to 36 at intersections.

The 360-degree panoramic views from the anamorphic lens yield a space in which the direction of view is continuously variable (even though the viewpoint position remains discrete).

The third dimension introduced by the aerial overview might be described as two-and-one-half-dimensional. Movement along the z axis does not result in corresponding perspective changes, only in scale changes. The town has been reduced to a flat surface. No parallax is seen in movement in x and y. As in the ground plane, the aerial space is discrete. The cubic lattice consists of linearly spaced points in a regular grid of parallel and perpendicular planes. Topology in the otherwise unstructured volume is specified by perpendicular projection of the ground plane. That is, the grid is modulo street intersection. Each vertical path is aligned with a unique crossing.

The pyramidal lattice also associates aerial paths one-to-one with intersections, but in other ways it creates a space with unusual properties. It is not linear and omnidirectional. Excursion in x and y is a function of z (because the pyramidal cross-section varies with height). Even more intriguing is the movement of points in the neighborhood of any end point. Movement along a path towards the ground plane creates a change of scale which is topologically equivalent to stretching the town on a rubber sheet. The area is stretched around the target intersection as if it were anchored by a nail. The amount of movement of neighboring points is a function of their distance from the anchor. An inspection of Figure 25 will show the proportionally different movements between the upper right corner (farther) and the lower left corner (closer) of the "F" frames.

3.1.2 Cognitive space

Now that a modestly formal description has been made of the properties of the virtual space the question must be addressed, What are the implications of these properties for the cognitive space of the viewer?

Moving in a discrete linear space is not an obvious handicap to the user. Sims [1974] has found that in many ways segmented media elicit a superior performance by the subject in the area of "cognitive response fidelity," i.e., the closeness of the match between the media simulation and the real environment in learned informational content. Sims compared a trip through the real environment with film, slides, drawings, and model photography. It cannot even be taken for granted that the real environment is the best medium for conveying the most important information. In fact, Sims found that the control subjects who traveled the real route were worse at estimating distance than five out of the six media groups tested. On the other hand, he concluded that recognition of directional change was best facilitated by certain continuous forms of movies, whose viewers again outperformed the real environment control group. Further, estimation of the magnitude of directional change was especially difficult in the case of discontinuous or segmented media.

Intersections seem to become emphasized in surrogate travel, both because of the option they present of changing paths and because of the increased number of views they offer. The comparison to real life is interesting. There are examples in the real world in which the opposite is true, where street intersections are only interruptions to the more interesting process of moving along each block. But these examples probably involve walking, window shopping, etc. Driving emphasizes the importance of intersections. They occur more frequently, require quick decisions, and are more likely to be pivotal points in an extended route of travel typical of driving. The one real-world element crucial to intersections not reflected in surrogate travel is danger. The user never has to worry about traffic and accidents, and driving down the middle of the street with seemingly reckless disregard for the pixillated crossing traffic may be exciting.

Virtual space around intersections is denser or richer than elsewhere. While the number of views is limited to 4 along straight segments--2 side views plus a front view in both directions of travel down the street--pivot turns treat the user to a visual "oasis" at the end of each block by increasing the number of views to 36. Depending on the actual combination of turns implemented in a particular version, the number of routes through an intersection can be as many as 24 compared with the 2 that are possible in opposite directions along a block. (Four arise from the 2 directions along the two intersecting paths, 8 from all possible leading or truck turns, 8 from tracking pivot turns, 4 from pivot turns. Not all are everywhere available; the average is about 12.) This means that surrogate travel is so configured as to provide the user potentially with the most information about the space in the vicinity of intersections.

Is the additional information an aid or a hindrance? It is quite reasonable to hypothesize that the amount of redundancy should vary inversely with the length of time the user has spent traveling. The interesting question is whether it is more important for the user to become familiar with the point of intersection than with points along the connecting segments. It is quite likely that landmarks near intersections play a more important role in the users' cognitive space than do those in mid-block. Corner landmarks are more uniquely identifiable, and the user wants to establish his location/orientation at intersections to facilitate decision making. There is evidence to suggest that subjective scaling in cognitive space along a route results in objects gravitating towards major landmarks, which act as anchors [Allen, Siegel, and Rosinski, 1978].

The closer the landmarks are to the beginning or end of a straight segment the more the richness of the intersection space is likely to be exploited. In fact the landmarks the viewers use may move closer to the corners as their familiarity increases. And in several instances landmarks exist in the intersections themselves--traffic lights, yellow lane dividers, etc.

A distinction can be made between two representations of an urban space: the ground-level, path-following experiential view and the aerial, map-like overview. The former is a network model of topological connectivity, the latter is a geographical model of geometrical coordinates [Kuipers, 1977]. It may be expected that straight segments are more crucial to the topological view and intersections more crucial to the geometrical view. That is, in learning which paths are topologically connected to which it may be sufficient to remember the straight segments and the sequences of turns to left and right. To learn more of the absolute geographical configuration of the space it may be important to attend closely at each intersection to the magnitude of each turn. The unique key offered by the interactive movie map is a mechanism for integrating these dual representations of the space. Experimental evaluation of the effect of this feature upon cognitive space will be more fully discussed in Chapter 5.

It is to be expected that access to aerial perspectives enhance the user's sense of geographical space, by providing a clear picture of the geometrical relations absent in the more topological ground-level representation. This effect is enhanced by the substitution of actual maps for the aerial photos. The lack of transition from ground-level view to the perpendicular perspective from the air presents a potential problem. The small schematic map provides a marginal connection. However, part of the hypothesis discussed in Chapter 5 is that maintaining the distinction between the two kinds of views can serve to advantage. There is no reason to believe that the space represented in cognitive maps exists in a unified internal form. The representations are likely to be as complicated and multi-modal as the experiences that produced them.

3.2 Time

As with space, there are four kinds of time to be distinguished: real time, virtual travel time, playback time, and cognitive time.

a) Real time is the constant and continuous flow measured by clocks in everyday life.

b) Virtual time is the constructed chronology that results when the video disc footage is edited.

c) Playback time is real time during playback.

d) Cognitive time is time as it is perceived by users during surrogate travel.

3.2.1 Virtual time

It is mathematically impossible for any mapping of a two-dimensional space (e.g., Aspen streets) into a one-dimensional record (film) to be continuous. There are discontinuities not only in space but in time. For the surrogate travelers, real film time jumps rather arbitrarily, almost every time they turn around (literally). As users drive down the street real time may be 11:00 a.m. Monday. Turnin right, the real time may switch to 2:00 p.m. Tuesday upon entering the turn, and then switch to 10:00 a.m. Monday on the new straightaway. Other discontinuities appear within the straight sequences. Frames are shot at regular spatial intervals, but not necessarily at regular temporal intervals. Whenever speed of the filming rig is not constant, especially when the rig is stopped, the flow of real film time is interrupted. Chronological order, of course, is preserved except when the film is reversed. This occurs, for example, when backing up. Time reversal also occurs when the side views are accessed while traveling in the opposite direction from the filming rig. Some of these artifactual properties of time are inherent in the medium of movie map, for example, in jumping between time zones; others are merely consequences of the particular way the space was filmed, for example, nonlinearity and time reversal.

3.1.2 Cognitive time

What are the cognitive consequences of such an idiosyncratic system of time? Clearly time is a secondary, almost subliminal dimension for the user, who perceives himself to be traveling in space, not in time. Discontinuities in time are only noticed when they mark events which exhibit discontinuity. What the user may notice is not that he is traveling down Main Street on Monday and turning onto Mill Street on Tuesday, but that a school bus has sudddenly materialized out of nowhere, or that the shadows are pointing the other way. When the objects and lighting from one shot correspond to those in the next shot the viewer is unlikely to experience any discontinuity in time, even though one exists in film time. (By the same token, if the camera exposure suddenly changes, implying a change in lighting, the viewer may interpret the discontinuity as a discontinuity in time.)

Generally, when first-time users sit down to the movie map system, the first thing they are struck by is the pixillated traffic. Cars zoom by in reckless burst of acceleration and deceleration. Typically users will make some comment indicating amusement at being in such a cartoon-like world (Roadrunner, etc.) where vehicles weave in and out at exaggerated speeds with absolute impunity. In this way users are immediately alerted to the fact that here is no ordinary flow of real time; and they are apparently able to accept this unusual convention of temporal flow quite readily. The introduction of this convention at the very beginning may set a precedent for the acceptance of the other unconventional time/space structures later.

Viewers notice certain peculiarities of traffic. But equally important is what viewers do not notice. Foot traffic draws little attention. So does oncoming traffic. In the former case relative speeds are low, in the latter they are high. It is only in the case of comparable relative speeds, like vehicles traveling in the same direction, that the world becomes curious. This is because the objects in question must remain on screen for a sufficient number of frames for the effect to be perceived. Pedestrians and oncoming cars disappear from view too quickly. (As a result, viewing time reversal does not present a perceptual problem when the moving objects are pedestrians or oncoming cars.)

The interesting thing implied by the relative speeds of moving objects is relative time. Events in the world of surrogate travel take place at a pace which users can never become a part of. On the average, surrogate travel time moves at about three times the observer's pace. This ratio arises from the fact that the flow of traffic always moves at approximately three times the speed of the surrogate traveler. So if playback time is to correspond to real time, surrogate travel speed must be around ten mph, one-third normal. In that case all events--pedestrians crossing, dogs running, leaves falling, cars stopping for lights--will take place at properly scaled intervals. Of course, the viewer's vehicular speed will be quite slow for a car. If playback speed is to simulate real driving speed in Aspen then all other time will speed up. Traveling down the street at 30 mph will speed up the rest of the cars to 90 mph. It turns out that this incompatibility of speeds is seen to be much more of a problem by those familiar with the design and production of the movie map system, who know that the footage was filmed as vehicular travel. Untrained users are not under the same bias to view their ego motion as vehicular and consequently not necessarily bothered by the relative time phenomenon. They may be interpreting their mode of travel as bicycle or are perhaps not forcing any particular interpretation. (After all, traveling down the center of the street eleven feet off the ground is a circumstance which is hard to relate to any experience.)

Because of the various ways in which traditional conventions of time are violated in movie maps, it has been suggested that here at least is an example of a movie that can portray physical properties without an inherent time referent. But in fact real playback time gives an accurate scaled model of real travel time. It takes users twice as long to go twice as far. Whether that relative indiex is useful in helping to judge distance is a question that will be addressed later.

3.3 Events

The most significant cognitive implications are consequences of the interplay of virtual time and virtual space. These will be addressed by looking at the world of events in surrogate travel. A viewer witnesses many events as he travels around--a pedestrian corssing at the corner, a red truck waiting for a left turn, a little boy flying a kite in his front yard, even something as subtle as sunlight glinting off a window. Each time a viewer repeats his exact path the same events will repeat exactly. (This is not true if he retraces his path going in the other direction.) As a result, surrogate travel is more than just a spatial record of geography; it is a temporal/spatial record of an events world. Events are fixed in this world almost as firmly as building are fixed in sapce. (The difference is that building generally remain constant over several passes through a given space, while constancy of events is not usually maintained beyond a single pass.) In surrogate travel, time is a function of space, not vice versa, as is the more typical experience (and as was the case during filming, when location was a function of time). The consequence is the rather interesting phenomenon that events are a function of space.

Events with the property of 100-percent repeatability have very interesting cognitive implications. Such events can be used by the viewer as orientation cues as dependably as can physical objects. For example, the viewer may verify that he is at the corner of Mill and Hyman, facing north, by recognizing the south facade of the Wheeler Opera House, or, with equal accuracy, by recognizing the horse-drawn buggy turning the corner. What seems to loom as a potentially hazardous artifact of surrogate travel is that viewers may sometimes rely on transitory cues to establish their orientation. Naturally, the horse-drawn buggy may not be at that corner when the viewers are actually traveling around Aspen. The Wheeler Opera House, one would assume, will still stand there. Actually a building can be considered an event; it takes place as a location in space for an interval in time. It happens that the duration is quite long in the scale of human events. The building, therefore, is given a high probability of being at that location at any time. Events of much more fleeting durations also have their own probabilities. That buggy, it turns out, does happen to spend a lot of time at that corner waiting to be hired by tourists; there is a high probability of encountering a bus along Ruby Park, which is the central bus stop; young people can almost always be seen lounging outside Burger Brothers on Durant Street. This means that when people travel around in the real world they must somehow be able to discriminate between those cues which are quite characteristic of a particular location and those which have only a low probability of being repeated. (The distinction is not simply between long and short durations, though there is undoubtedly some correlation.) How is such a discrimination made? Perhaps some kind of subconscious assessment of the frequency and periodicity of these cues is made. A cue or event with a long period or high frequency of occurrence has a greater probability of being repeated the next time. Periodicity is probably easier to guess. A stone building is much more likely to exist next month than a fruit stand, which is likely to last longer than a snowman. Frequency of occurrence is more difficult to judge, requiring exposure to multiple instances of an event or prior knowledge about the nature of such events. For example, pedestrians accumulate every minute or so at certain intersections with traffic lights; rush-hour traffic generally occurs twice a day, five days a week.

To what extent do surrogate travelers take advantage of the 100-percent repeatability of surrogate travel cues; and to what extent do they, out of habit, use their real-world procedures to decide which cues to depend on? That question may not have a definitive answer at this point, but some deductions follow. Surrogate travel is not really 100 percent repeatable; each distinct path through a given point gives rise to a different set of events. Since the number of possible paths varies from 4 along a straightaway to 24 at some intersections, any particular location has an associated pool of several possible events. This quite nicely, though artificially, establishes a probability for each event. The size of the events pool is greatest at intersections--just where viewers may be more likely to establish their orienting cues. It is not certain that viewer dependence on arbitrary transitional cues is detrimental to the acquisition of spatial knowledge, though it admittedly seems likely. One can imagine, and there is anecdotal evidence, that if movie-map-trained viewers were in the real environment, a transitory cue they had relied upon would "insert" itself into the real experience. An example of this phenomenon occurred with cultural slide shows. A viewer had two encounters with Cooper Street Pier, an outdoor cafe, first via the interactive movie map and then from actually being in Aspen. The mood of the lazy, sunlit afternoon, which was captured in the slide, so perfectly presaged the actual experience in this case that the person's later recollection of the scene much more strongly reflected the slide than the real visit. The image of the slide superseded the real view even though it was quite different in actual content. The on-site traveler may well "picture" the absent bus crossing his path just where it was actually seen in surrogate travel, or at least use the remembered event as an idiosyncratic label to identify that location.

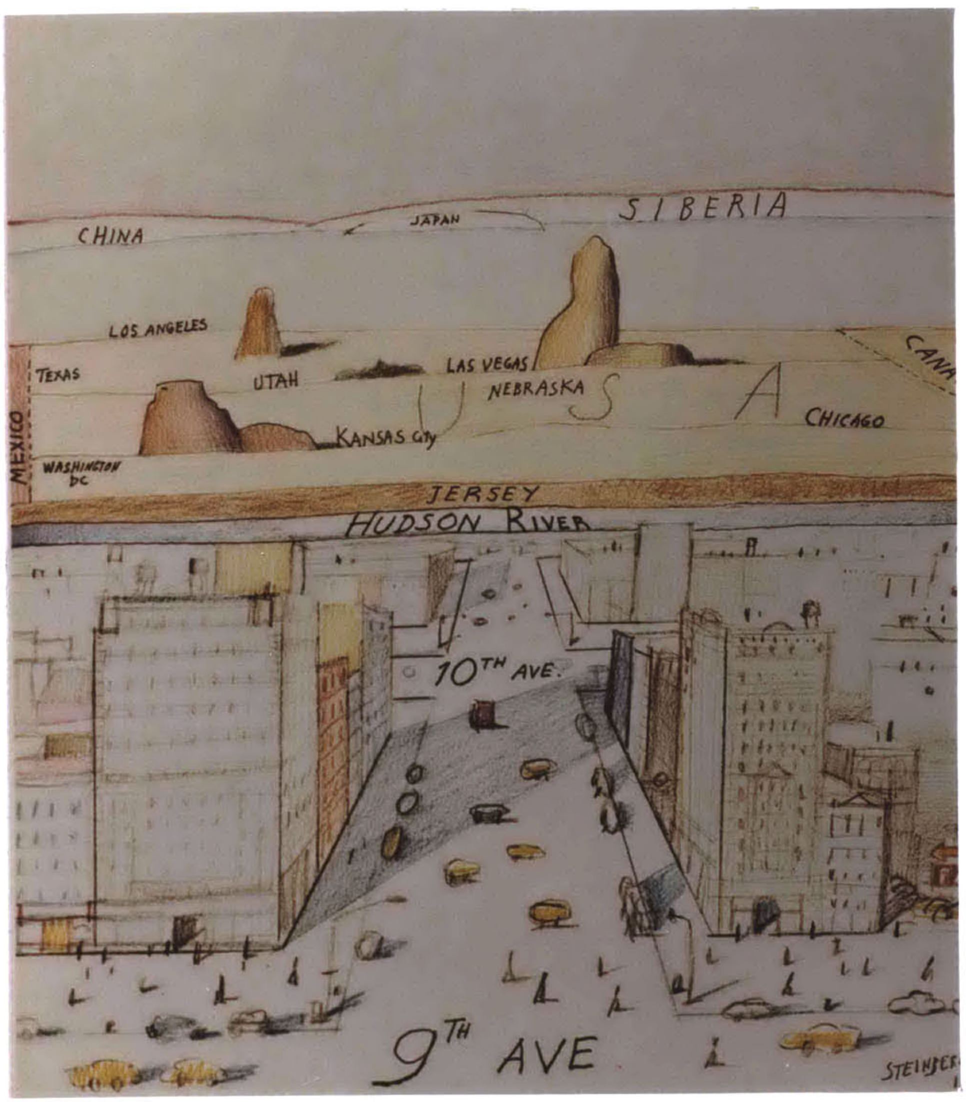
Despite this type of exception, one of the challenges which might be raised to question the success of surrogate travel is the tendency viewers may have to rely on the repeatability of cues that are normally transitory in real life. This problem will be discussed later in terms of actual cases.

CHAPTER 4

Cognitive Mapping

My father, a man with a great sense of humor and no sense of direction, constantly led us on what he referred to as "scenic routes."

- Ellen Goodman, Boston Globe, 9/17/81



- Steinberg

This chapter will be divided into discussion of two components of research in cognitive maps--theoretical models and research methodology.

4.1 Theory

There are typically two ways in which knowledge about a space can be acquired: by actual movement through the space and by analysis of simplified, often symbolic, models of the space. The first part of this section will discuss the role of direct experience, the second the role of maps.

4.1.1 Experience

People's facility in acquiring spatial knowledge about their world is universal. Through repeated exposure they gain a familiarity which enables them to perform spatial operations. They recognize settings they have seen before, follow paths they have taken before. Clearly memory is an important component in the process by which they internalize their experiences of the environment.

However, the process is much more complex than simply recognizing features in context. People can recall spatial attributes of a place without being there. A common example is that most people can count the number of windows in their kitchen by reconstruction from an internal representation. Pylyshyn [1973] and Kosslyn and Pomerantz [1979] take opposite sides with regard to the existence of mental imagery. Beyond the two processes of recognition and recall just cited people can also infer novel spatial relations for which they have no direct experiential evidence. People can take short cuts in the woods and make detours around traffic jams.

The cognotive processes are rich and varied, as are the external manifestations. On an anthropological scale they range from native navigators of the Puluwat Islands, who pilot their canoes with unfailing accuracy over 100 miles of open sea [Gladwin, 1970] to cab drivers in Manhattan, who expertly find the shortest (or, perversely, the longest] route through the maze of one-way streets. The level of development ranges from that of three-year-olds, who tend to search for hidden toys with an egocentric organization of space [Acredolo, 1977] to that of engineers, architects,and other adults who can formalize their spatial concepts. A number of psychologists have suggested that microgenetic development parallels the process of ontogenetic development [Siegel and White, 1975].

The term "cognitive map," probably coined by Tolman [1948], is the most in vogue of a succession of labels for the internal models people develop in large-scale environments. As early as 1913 spatial images were called "imaginary maps" [Trowbridge]. Since then the list has included "mental map" [Shemyakin, 1962 and Gould and White, 1974], "environmental image" [Appleyard, 1969], "spatial image" [Boulding, 1956], and most recently the more generic "taxonomic category" offered by Siegel and White [1975]--"spatial representation." Avoiding any suggestion of reticence, the title of this document adds a term of its own--"cognitive space."

The first problem that needs to be addressed is whether each of these terms describes a single phenomenon and, if so, whether it is the same one in all cases. Generally the terms have been used to refer to both the internal representation and the external expression which a person formulates for a space. Clearly, the two imply different content and behavior. The first is a representation of experience; the second is a "'re-representation' and two levels removed from spatial activity" [Siegel, 1979]. Siegel also makes the case [1980] that cognitive mapping is fundamentally a kind of problem solving. To some extent the structure of an internal representation must be governed by the context and the purpose of the activity the individual is engaged in. A cat burglar presumably cases a new neighborhood with a different set of criteria than does a paper boy. Similarly the externalization of spatial knowledge is dependent upon the need it serves. Different elements of the space will be emphasized by a pedestrian planning a route than by the same person making the same trip by car.

For this discussion, "cognitive space" will be taken to mean an internalized model, representation, or strategy which is useful in accounting for spatial relationships. The term is chosen because, unlike "map," which suggests cartography, "image," which suggests pictures, or "representation," which suggests an internal state, "cognitive space" can include any of these notions until the evidence defines the constraints. The significant contribution of the term "cognitive space" is the explicit inclusion of process. A fundamental duality crops up again and again in the theories--process versus state, or, in information processing terms, the program versus the data. The state is an internal representation for structuring the spatial knowledge. The process is the cognitive strategy for internalizing and externalizing the state of spatial knowledge. One of the more persuasive demonstrations of process is the example of the commuter heading for Pittsburgh on Interstate 376 when on hearing that traffis is backed up to Philadelphia, he generates an alternate route which is completely new to him [Siegel, 1979]. A process of spatial rules operates on a body of spatial knowledge.

Theinformation processing model may imply a simpler reality than can be justified. Externalizations of cognitive space are typically fragmented, distorted, inconsistent. Both process and state are probably responsible.

A number of general models of cognitive space have been proposed. The development of children as they progress from preoperational to concrete operational stages has been characterized by Piaget, Inhelder, and Szeminska [1960] and by Piaget and Inhelder [1967]. Younger children's cognitive space is based on topological relations, while older children incorporate the euclidean concepts of metric dimensions and linear perspective. A major criticism of Piaget's position is that the research focuses on spatial knowledge in small-scale space only [Herman and Siegel, 1978]. Whether these conclusions can be generalized to large-scale environments, which are more representative of real-world experiences, is highly questionable. The two types of space are not just different in scale; they are different in kind, and there is no clear evidence that similar cognitive processes are operating in both.* Acredolo [1977] offers evidence to the contrary.

---

*The distinction in scale is not a function of size, but rather of visual properties. Large-scale environments are those in which there is no location from which all other points are perceivable [Siegel and White, 1975]. An integration of successive observations over time is, therefore, required [Kuipers, 1977]. Small scale denote a line-of-sight neighborhood in which all landmarks are simultaneously visible.

The large-scale space of the urban environment was the arena for what is perhaps the seminal work in modern cognitive mapping--The Image of the City [Lynch, 1960]. A wonderfully written book, full of rich anecdotal data, it popularized research in this field. Lynch, working mainly from image maps drawn by residents, concludes that internal models of the environment were constructed from five elements: landmarks, nodes, paths, edges, and districts. One of the important factors Lynch points out in his attempt to describe imageability is the role of functional interaction. People internalized and externalized their environment on the basis of what was meaningful and relevant, what was culturally and personally noteworthy. The history of experience with the corner drugstore can outweigh the significance of an irrelevant high-rise building. Largely missing from Lynch's account is a developmental theory relating his five elements.

Kuipers [1977, 1978] presents a comprehensive computer model, TOUR, to simulate cognitive processes in acquiring spatial knowledge and solving spatial problems. Closely echoing Lynch's categories, TOUR's "primitives" consist of places, paths, and regions. (Paths can act as edge boundaries and places can includes nodes.) Further, Kuipers [1977] draws a rough parallel between parts of his model and the Piagetian stages from topological to projective to euclidean space. Put briefly, the integration of learned route descriptions into path networks corresponds to the topological stage. Local and relative orientation appears to be similar to the projective stage. Global and absolute orientation matches the euclidean stage.

A comprehensive developmental model of spatial representation for large-scale space is proposed by Siegel and White [1975]. Again somewhat reminiscent of Lynch's work, their explanation proposes main components consist of landmarks, routes, and configurations. The authors hypothesize a developmental hierarchy which leads successively through these three elements. The model is more or less the same whether the developmental period is long-term, covering the growth of a child, or short-term, covering a visitor's weekend in an unfamiliar environment.

Landmarks are strategic foci which identify specific geographical locations. Interestingly, Siegel and White describe landmarks as unique perceptual events. The notion of the landmark as a general event (in a social context) is alluded to by Lynch and is documented in the next chapter. Routes, a natural integration of landmark information, are strongly characterized by sequence. Landmarks typically serve as beginning and end termini and punctuate the environment in between. Siegel and White emphasize the role sensorimotor behavior in routes (as opposed to landmarks). Configurations are overall gestalt knowledge of the global space, which can assume several forms--a boundary outline, a network skeleton, a figurative metaphor (e.g., the boot of Italy). Configurations function in both deductive and inductive fashions. They serve as aids in way-finding problems and as organizing structures in internalizing new experiences.

There is a certain elegance of geometry which links landmarks, routes, and configurations together. The underlying relation is that of point, line, plane. Siegel and White, in fact, allude to this interpretation of the first two stages, picturing them as visual pegs connected by sensorimotor lines.

It is interesting and presumably significant that the theories described share a common denominator, which is perhaps most succinctly stated in Siegel and White's model of three hierarchical levels. In addition, an underlying theme of polarity runs through most of the work in cognitive mapping. There are two distinct modes that appear to operate with a robustness that accounts for the various ways in which they show up in different research projects. The duality is fundamentally the distinction between large-scale and small-scale space that has already been mentioned. Kuipers [1977] offers the example of an entire city as small-scale when seen from an airplane. It will be important to keep in mind, with relation to the work presented in this document, that aerial photographs and maps are also examples of small-scale representation, in contrast with the large scale of the life-size world. The significance will be more fully explored later.

The duality of scale manifests itself in many contexts under various labels in the work of several researchers, but the underlying notion is always similar. In Piaget it is the difference between topological space and euclidean space. In Byrne [1979] it is the difference between a network map and a vector map. In the mathematical land of Abelson and DiSessa [1981] it is the difference between "turtle" geometry and cartesian coordinates. In Siegel and White it corresponds to the difference between routes and configurations. In Schon and Bamberger [1978] it is the difference between figural and formal. In the interactive movie map it is the difference between travel land and map land. The robustness of the duality is confirmed by the fact that it is carried over into a wide range of examples.

Interesting perceptual properties follow from the duality of large- and small-scale space when they share a common referent. The former can only be encountered sequentially; the latter can be apprehended simultaneously. The former is experienced locally; the latter is inspected globally. Crucial in the duality is the relation between the dimensionality of the space of the person and the dimensionality of the space of the environment. The bird's-eye view from the top of the World Trade Center [Siegel, 1980] and the fixed vantage point from the side of a room both are constrained to a two-dimensional plane outside of the physical space. In large-scale space the spatial dimensions of the vantage point cannot be separated from the dimensionality of the space. A necessary consequence is that, except in degenerate cases, views in large-scale space evolve with changing viewpoint. Dynamic occlusion and parallax are a function of locomotion. Small-scale space, on the other hand, remains largely unchanged by movement of the viewer.

4.1.2 Maps

The preceding discussion has set the stage for a discussion of the cognitive function of maps themselves. The vast majority of research studies and theoretical models have addressed spatial learning as it takes place in the absence of navigational aids like maps. And yet, the prominent role of maps in solving real-world navigational problems underscores the danger of ignoring their contribution to cognitive mapping. Even in the case of long-time residents intimately familiar with their home town, it would be a rarity for any of them not to have been exposed to a variety of maps of the locale. Experimental research, to be described in the next chapter, has established the importance of maps as a counterpoint to direct experience.

A longitudinal model of the mapping process has been developed by Hooper [1978, 1980]. Covering both design and use of maps, it identifies four mapping operators--data selection, map making, map reading, and problem sovling--which link five knowledge states: environment, data base, map, mental model, and plans for action. Hooper observes that the success of a map greatly depends upon the degree of compatibility between the objectives of the map designer and the goals of the user.

Because of the role of the map in the interactive movie map it may be worthwhile to address the question of map design and interpretation in some depth, especially with regard to scale and resolution of detail. Both are inherently linked to thematic content and therefore influence the type of cognitive tasks which can be performed by the user. Changes in scale arelinked to changes in concept. Robinson [1976] made the point clearly in observing that structure can be created through manipulation of scale. "Objects can be transformed from a state of 'separateness' to one of 'proximity' allowing the creation of an object from what may also be apprehended as a collection of discrete objects." Further, manipulation permits the domain to shift freely from alike/unalike to proximate-separate, which is a transformation from class associations to spatial associations. A geographic diagram of the same basic concept is used to illustrate Abler's [1971] "macro to micro viewpoint" (Figure 48). Another variation is Stafford Beer's "cones of resolution" model, which depicts level of detail varying with the level of scope (Figure 49). The capacity to effect a change of scale clearly increases the potential power of the

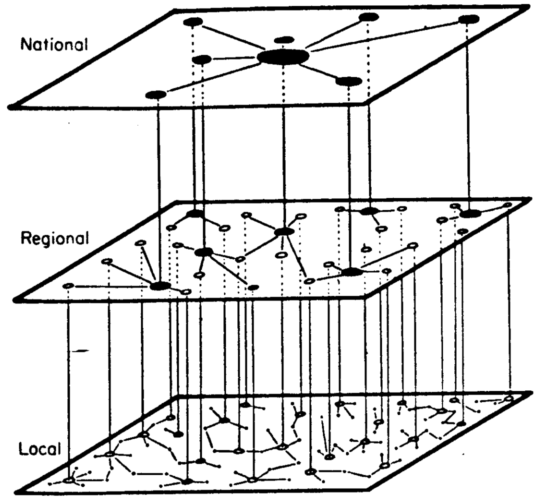


Figure 48. Alber's "macro-to-micro viewpoint"

representation. Changing scale is a transformation by which everything is changed and yet somehow remains the same. Thenotion is captured in a metaphor--"a building is a tomb or a coffin is a room" [Harrison, 1977]. The conceptual relationships change: the area of a city becomes the dot of a node. The spatial relationships are maintained: the Massachusetts Turnpike still runs west out of Boston.

Some attempts have been made to establish a theoretical explanation for the hierarchy. In a formulation reminiscent of the duality princple discussed in section 4.1.1, Chapman [1977] employes the concepts of relative and absolute space to analyze the process of reducing scale, generalizing a region from components. Such generalization, he maintains, cannot apply to objects defined in absolute space (since each position is unique). It can only occur in relative space. Experimental evidence is reported by Stevens and Coupe [Hooper, 1979]. Judgments about the spatial relations between cities were strongly affected by hierarchical relations

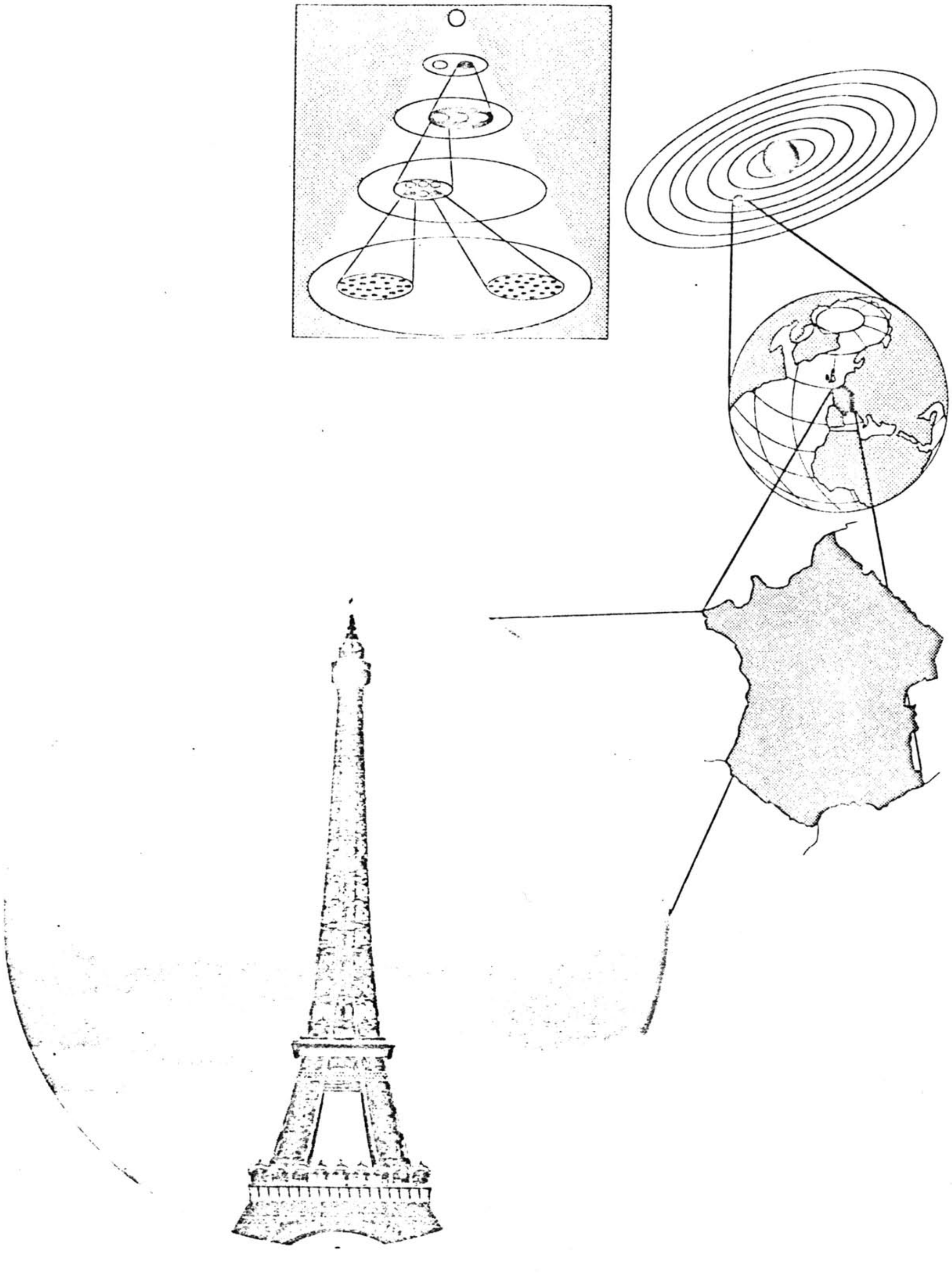


Figure 49. Beer's "Cones of Resolution"

between the superordinate categories, namely, the respective states.

A more comprehensive model by Feibleman [1954], called "a theory of integrative levels," addressed the concept of structure in general. Rather naive and prescriptive, it proposed:

a) Each level organizes the level below it plus one emergent quality. (As a corollary, it is impossible to reduce the higher level to a lower one because the quality and structure would be lost.)

b) For an organization at any level, its mechanism lies at the level below and its purpose at the level above. To analyze each requires moving up or down accordingly.

Another attempt by Philbrick [1957] presented nodes at one level alternating with homogeneous regions (parallel relationships) at the next in a functional rather than spatial hierarchy. For example, Figure 50 shows a homogeneous farm area in which single farms are nodal units. In an approach similar to Feibleman's, Philbrick proposed that "the structure of each level or scale of classification is definable in terms of the alterante type of relationship at the next larger scale." Aside from not demonstrating that homogeneous regions will always turn into nodes at smaller scales, Philbrick runs into trouble be endowing the user with unsubstantiated powers of visualization. By simply narrowing his attention to the appropriate scale of thinking and ignoring the details of scales below, he claims, the viewer has a device for examining the trees and the forest simultaneously.

The important principle which emerges is the need for the level of scale

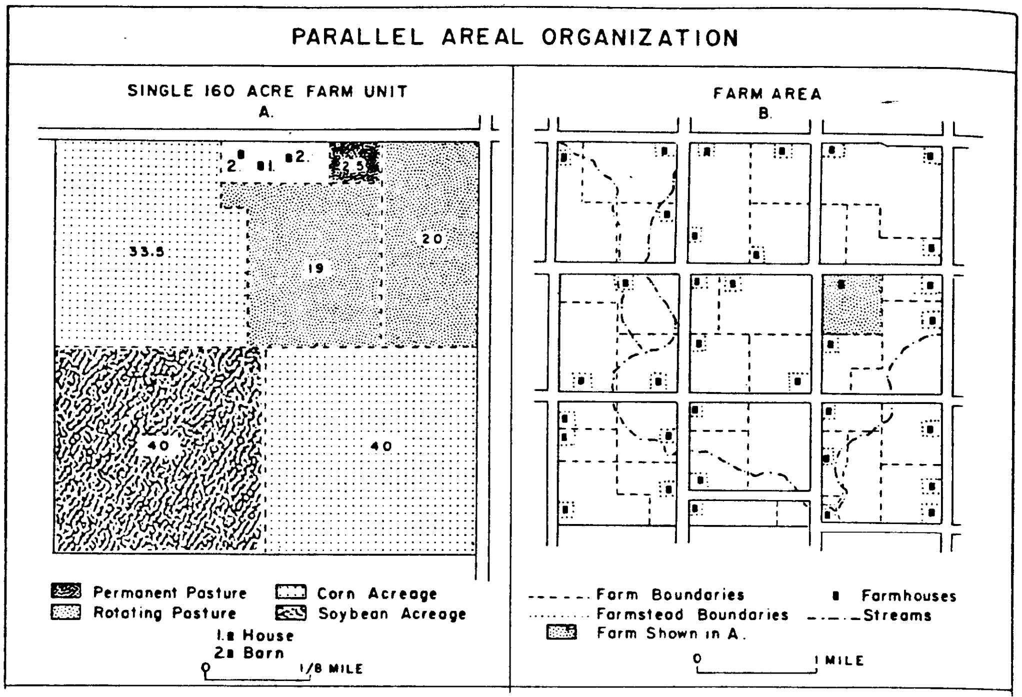


Figure 50. Philbrick's functional hierarchy

and consequent level of detail to be commensurate with the thematic units being addressed, more precisely, the need that the scope of the representation cover a field larger than the subject of inquiry, and that detail exist at a level finer than that of the subject of inquiry. For example, for the purpose of following street directions from a house in one neighborhood to another, it will be useful for the scope of the overview to include the larger context of the city to provide an appropriate and familiar frame to which to attach the new information being learned. It is pointless to include the whole state (unless that is truly the smallest context which theobserver is familiar with). And at the other end of the spectrum, it will be useful for the level of detail on the map to include elements which exist at a finer resolution than the street routes being followed. Path-marking landmarks would be a typical case in point.

The primary problem in the literature surveyed is the way in which the subject has been segmented. The main concern has been with conceptual units of representation. Some adopt the viewpoint of the map designer, some that of the observer. A whole other body of research addresses the development of cognitive space as it arises through perceptual processes in the real world of large-scale space. What is lacking is an integrated model which incorporates all of these approaches.

First of all, the physical environment of the map medium must be considered. Obviously, the constraints of resolution and legibility of the medium affect the scale of representation. But so can qualitative features. The map can be in the same medium as the referent (like a 3-D model) or in a different one (like verbal directions). The form of presentation can be static (like a book) or dynamic (like a movie). The map can be presented simultaneously with thelarger field of its reference or separately and sequentially. (The latter imposes different requirements on short-term memory than does the former.)

Even in the most limited case--say, that of a static paper map--the environment permits the representation of the map to be treated as a dynamic image. The viewer typically moves in and out, closer and farther from the surface, effectively changing the scale. Certainly the retinal projection changes in scale; and on a cognitive level, the viewer is attending to varying scopes, varying conceptual levels. The level of detail changes correspondingly. Close up to the surface the viewer is able to see details and read labels which he could not see father away, while global elements, like large, spread-out text, pass outside of the field of view and focus of attention. The extent to which that process can take place is dependent upon the environment--the size of the map, the range of viewing distances available, the resolution of the display medium.

When the size of the representation is very large (or the observer very small, as will be discussed shortly), a qualitative change can take place in the interaction. Blaut and Stea [1968] reported on a remarkable study in which the medium consisted of aerial photos so large in size and scale that children could crawl and walk over them on the floor. The result was a leap in the intensity of the children's intellectual involvement in the subject matter. Blaut and Stea theorized that, as the children became giants in walking over a real surface of the world, they themselves entered into the landscape. Their sensorimotor interaction was much closer to their normal real world experiences than would be the case in learning from small-scale representations. The authors' conclusion was that keeping the size of the representation as large as possible allows scale and changesin scale to be experienced in relation to the viewer's own physical size.

4.2 Research Methodology

There are two problems in research methodology which bear scrutiny, both of which are largely solved with the interactive movie map. The first is in design, the second in technology.

4.2.1 Design of the experiment

Not surprisingly, experimental research has commonly been directed at only one of the sources of spatial knowledge at a time--direct experience or map interpretation. In fact, there has been a dearth of experiments focusing on the latter at all.

Some of the best known research in the use of maps as navigational aids has been conducted by Thorndyke and Stasz [1980]. Their work addresses what they call the "map learning problem." Typical experiments included giving subjects hypothetical maps to be "learned" over a series of trials. The maps were conventional hard-copy representations with symbolic elements of both visual and written components. The tasks performed by the subjects after trials included map reproduction, problem solving, and verbal protocols. The experiments used only six subjects and concentrated on individual differences. The biggest criticism of the research is that it cannot claim any relevance to real-world experience. No effort has been made to design tasks which would require subjects to solve problems in the way they do in the everyday world. The results really have to do with "map memorization," not with "map learning." Further, the entire experiment stays within a single medium--the verbal symbolic format of maps.

Experiments which incorporate the dual components of spatial learning modes are even rarer. At the same time it is precisely those which may be the most relevant, because they can address the situation which is most commonly encountered in daily life. Most people most often are faced with situations in which they have incomplete map-derived data and incomplete exploratory experience. The one experiment which attempted to address the dual modalities of spatial representations was conducted by Ciccone, Landee, and Weltman [1978]. In fact, this work was based on a system, roughly similar in concept to the interactive movie map, except that it was noninteractive and entirely computer-generated imagery. Subjects were divided into three viewing categories: map only, map plus movie map, movie map only. Tests given required estimating distance and direction to certain landmarks, indicating a "you-are-here" pointer on a map, and matching written descriptions of locations with features. In spite of the apparent variety of tasks, the results were confounded. The media to be compared were separated in training but combined in testing. Subjects who had never seen a movie map image were shown a picture from the movie and asked to make distance/direction estimations. Surely, those data cannot be legitimately compared with the judgments of movie map subjects who did not have to transform their internal representation just to recognize the picture.

The richness and flexibility of the movie map system makes it possible to integrate dual sources of knowledge and to evaluate the resulting differences. But care must be taken to design the experiment to keep the distinctions being tested from being built into the test. There are two solutions. One is to keep each test in the same medium as the training. The challenge here, of course, becomes how to judge independent media equivalently. The other solution is to design the test in a third medium separate from the other two. Multidimensional scaling (MDS) methods [Kruskal and Wish, 1978] provide one mechanism for doing so. The procedure generates a two-dimensional spatial array from a series of one-dimensional pair-wise distance comparisons. The elegance of the technique is that a "map" is created without forcing the subject to draw an external representation of one. In this way one source of interference in the externalization of an internal representation may be eliminated [Kirasic, Allen, and Siegel, 1980].

4.2.2 The technology of the research instrument

Current research in cognitive space is vurtually defined, and certainly driven, by the medium of the experimental instrument; that is to say that the technology of the instrument often more than subtly shapes the design of the experiment. This is especially true of research which requires a visual display to present the space under study.

In the last decade the trend has been to bring experimental tasks out of the traditional laboratory into the rich domain of everyday experience. The result has been a surge in techniques for simulating large-scale spaces [Siegel, 1980] or actually using real-world environments. One of the problems with real-world settings as part of the experimental design is that they are very rich. It is difficult to know what cues the viewpoint of the subject or the content of the scenes as they unfold. Further, it is impossible to reproduce (or systematically vary) the setting exactly as the subject saw it for later spatial tasks (like landmark recognition or acquisition of route knowledge). The technical advantage of the interactive movie map as a research instrument is the facility it provides for controlling precisely what a subject sees and repeating or varying those images in subsequent tasks.

A number of types of research instruments have evolved in the simulation of large-scale spatial experience [Siegel, 1981]:

a) Life-size artificial environments. One of the more elaborate examples is one in which the subject is led through a space while particular areas light up to draw the subject's attention [Swedish reference unknown].

b) Slide show presentations. A virtual industry has grown up to represent paths through an environment by a sequence of overlapping views [Cohen, 1980; Siegel, 1980; Sims, 1974; and others].

c) Video tape presentations. Continuous motion sequences stored on film or video tape are a more expensive but more continuous way of presenting spatial sequences than slide shows. The material can be photographically derived or computer generated [Sims, 1974].

d) Simulation models. Small-scale models of the environment can be constructed as a course for photographic re-enactment. The Berkeley Simulator is a prominent example [Appleyard et al., 1973]. Even though the model is physically present, visual tours do not happen in real time. Flight simulators represent the other extreme; all images are computer generated but are displayed in real time by means of very expensive hardware.

Sims [1974] develops his own list of display techniques as part of his research into different media for simulating environmental experience.

His experiment compared the effectiveness of perspective drawings, photographic prints, projected slides, and film. He made a further division between photogaphing the real place and photographing a scale model. Perhaps his most interesting finding concerning the presentation format was that segmented media such as slides and drawings were not necessarily inferior to continuous media such as film.

In most of the categories discussed the issue of interactivity is not addressed (flight simulators being an exception). The reason is directly attributable to the technology of the research instrument. It is impractical if not impossible to allow the subject to make route-branching choices in the traditional display media. There have been a couple of bold attempts to force the technology to adapt to an interactive design. Sims [1974] cites two such experiments. The first, by Winkel and Sasanoff in 1966, used three-screen slide projection to simulate the interior of a museum gallery. By manually selecting the appropriate slides from a comprehensive record of the space, the operator could recreate virtual rotations and translations in space as instructed by subjects. In a follow-up experiment by Bonsteel and Sasanoff, a similar interactive technique was employed by using a closed-circuit television camera moving through a model of the museum. Like the Berkeley Simulator, this system worked in real time. The amount of brute force and human intervention which was necessary for limited success points even more convincingly to the impracticality of the traditional technology.

Each instrument has its own set of pros and cons. None is without shortcomings. It will be instructive to enumerate the probmes and at the same time keep the ineractive movie map in mind. In nearly all cases the video-disc-based system offers solutions to the problems without sacrificing the advantageous features.

a) Two major problems arise with model simulators. First, the environment is a stylized abstraction; fern sponges pass for trees and plastic figures suspended in time for people. The absence of human presence, motion of agents in the environment, events in general, removes this medium for the most part from the realm of real-world experience. Second, the cost of the hardware for such systems (in the hundreds of thousands of dollars) can be prohibitive. Further, the time and cost in building each new environment is similarly prohibitive.

b) Sims' [1974] findings about the effectiveness of segmented media cannot be generalized to all aspects of cognitive space. In the case of still-frame displays like drawings, photos, and slides, the temporal adn spatial gaps can be too large. "Life is not a series of snapshots" [Siegel, 1981]. Theproblem is simply one of technology. For example, no matter how much the operator seeks to reduce the gap, the cycle time of the slide projector remains the limiting factor.

c) The problem of gaps is solved by moving media like tape and film. Spatial continuity can be preserved. But then the advantageous feature of slide shows is lost, the ability to stop easily and fixate on a single scene. More generally, tape maintains spatial continuity only at a fixed speed of locomotion. The subjects do not have interactive control over their speed of travel.

d) None of the traditional media can easily be made interactive. The two attempts cited by Sims [1974] are awkward at best, and impossible as soon as the spatial network of branch choices becomes nontrivial. Computer control of video tape is a recently developed technique that allows complex branching. Still, the long search time delays keep video tape from being as effective as video discs.

e) Flight simulators (and other real-time computer-generated graphics systems) allow a high degree of interactivity. In fact, this is the one area in which they outperform the interactive movie map. They allow the user to assume any point of view, not just those which have been prerecorded. THey may well take the ideal research instrument beyond the level of the interactive movie map someday. In the meantime, the cost (in the millions of dollars) remains insurmountable except to institutions whose cost/benefit ratios put them into a different financial arena from that of prosaic psychological research. Moreover, flight simulators currently suffer from a lack of rich imagery. While provide much more realistic rendering of detail that a few years ago, the computer-generated graphics are still impoverished.

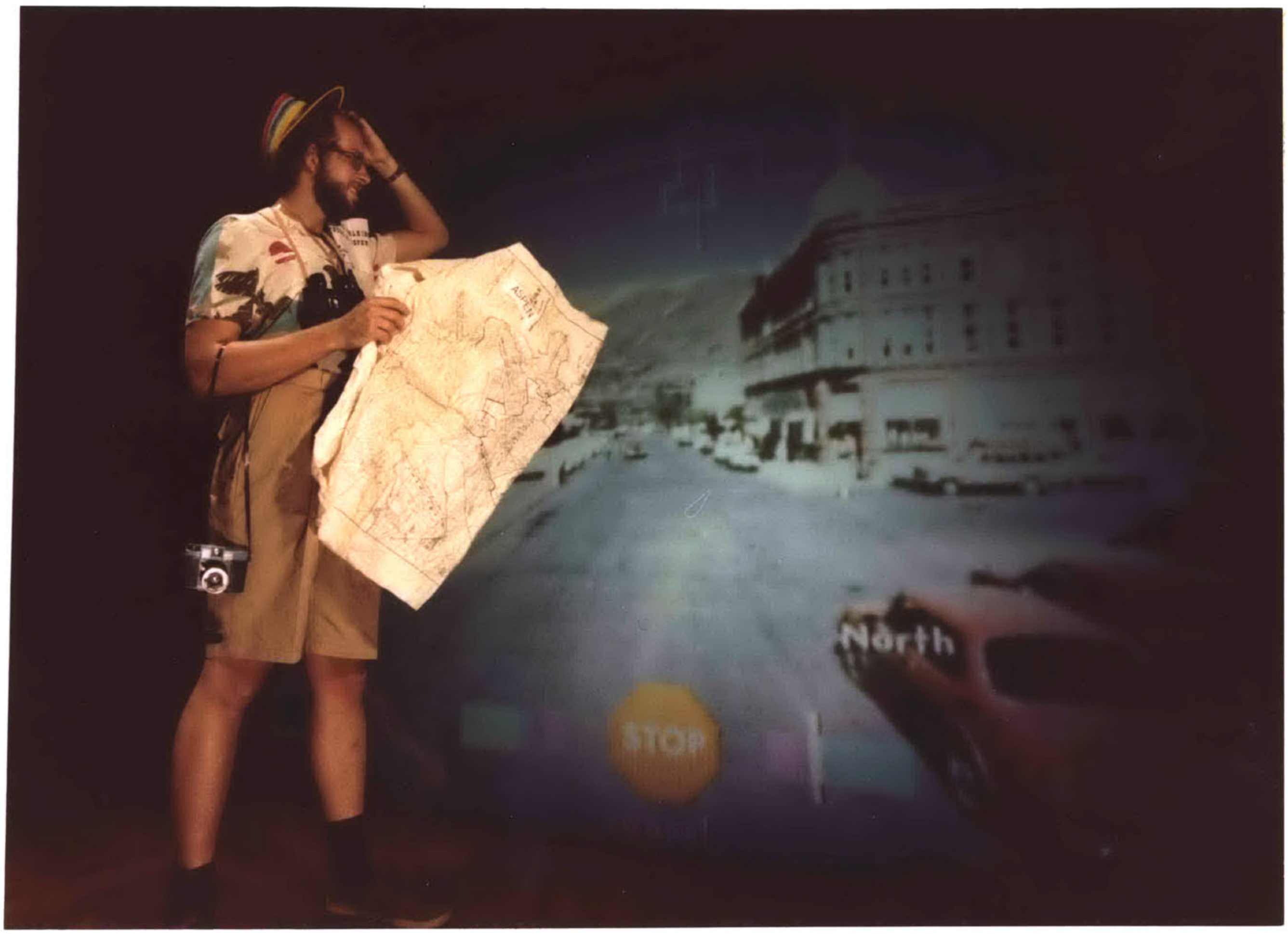
The interactive movie map is a versatile research instrument. It has the flexibility to present material in single frames at a time, at video frame-per-second rates, or at any speed in between. It can incorporate any of the modalities of representation in traditional media because these media are in fact source material for the video disc. That is, the format can be drawings, slides, film, etc. The one constraint is that the images must be presented as video, with its resolution, color, and aspect ratio. Thus the quality of a high-resolution medium like 35mm slides cannot currently be conveyed by television (and likely will not for a long time to come). The achievement heralded by video disc technology is its capacity to be dynamic, responsive, and highly interactive--and at the same time rich visually and aurally. It resolves what in other media has been a trade-off between competing poles. There can be little doubt that the development of the interactive movie map system has implications for wide-ranging areas of research in cognitive mappy.

CHAPTER 5

Experimental Studies in the Acquisition of Spatial Knowledge

...Spatial knowledge is represented in several different ways, and... important things happen when pieces of knowledge are transferred from one representation to another.

- Kuipers, 1977



The preceding two chapters have raised conceptual issues about the cognitive space of the interactive movie map in the absence of empirical evidence. This chapter documents some of the experimental investigations of the movie map system. Research was divided into two stages. The first was a wide-ranging pilot study conducted with the aim of establishing which avenues would hold promise for further and more specific study. The recommendations from that study led to a second, enhanced version of the system. The second part was a formal experiment using the movie map as a research instrument to learn about the role of dual representations in cognitive space. This chapter primarily addresses the first study. An effort is made to preserve the anecdotal nature of the protocols because of the richness of the evidence. This chapter should be interpreted in the same spirit and style in which it is presented. The findings are candidly speculative, not rigorously conclusive, as one might expect from the embryonic experiments designed to suggest future research.

One of the purposes of the movie map system is to help an individual acquire accurate spatial knowledge that will help him find his way around a real place. Any use to which the movie map is put in isolation from a real, verifiable environment falls short of this purpose. It follows that to understand the spatial knowledge of a virtual environment that is acquired by surrogate travelers, one must also be aware of the spatial knowledge of the real environment acquired firsthand. There are at least four groups of subjects worth distinguishing:

a) long-time residents of Aspen

b) first-time visitors to Aspen

c) movie map travelers who have never been to Aspen

d) movie map travelers who subsequently visit Aspen for the first time.

It is important to look at group (a) to establish a context of spatial knowledge about the real place. In group (b) the spatial model has been subjected to a time filter in an attempt to glimpse early stages in the process. Group (c) provides a comparison between spatial knowledge of the virtual environment and the real one. Group (d) yields the first real direct evidence about the success of movie maps.

The general format for conducting this pilot study was through interviews in which responses were written, tape recorded, or drawn (as by sketching maps). The questions used are listed in appendix 1. The option to probe for further detail suggested by any answer was liberally exercised. In the cases of Aspen residents and visitors the total interview time was about one hour. In the case of movie map users the total session lasted from one to three hours, with the time fairly evenly divided between surrogate travel and interview.

5.1 Aspen Residents

Interviews were conducted with 10 "locals" ranging in length of residency from one year to 24 years and in age from 10 years to 60 years. The interview was modeled after those conducted by Lynch [1960]. It is important to understand the conception of Aspen from the point of view of actual residents to provide a context in which the other cases can be considered. This does not necessarily imply that Aspen residents have a higher level of credibility in their view of the town, but that they can be expected to have a more comprehensive view. Any attempt to report an analysis of the distribution and frequency of response would be futile in a sample so small and diverse. Instead, the discussion will center on two kinds of responses--multiple occurrences which represent a consensus and comments by a single individual which provide a particularly illuminating or relevant view.

5.1.1 Map drawing

During the first half of the interview each resident sketched a rough map of the town covering all the main features. (See Figure 1 for the Chamber of Commerce map of Aspen.) The maps all consisted of a labeled street grid, a numberof identified landmarks, and wider regions. Landmarks were typically shops and public buildings; regions included the west end residential area, the downtown shopping malls, and the mountains. The scope of the maps was usually the city proper, bounded on the south by Aspen Mountain and on the other three sides by rivers (the same boundaries, incidentally, used in the movie map). Occasionally the scope was smaller, centering on the downtown area. (This was especially true of the two children's maps.) All maps were drawn with the paper oriented horizontally to accommodate the longer east-west axis of Aspen. This still permitted three variations: north at the top, south at the top, and no favored top (i.e., labels written in each of four possible orientations as the map was rotated). Roughly half of the maps had south at the top. Following is a summary of some of the most pertinent interpretations of the map data.

5.1.1.1 South-up orientation

This unconventional way of facing maps has already been noted by Wurman [1972], who observed that most published Aspen maps are oriented south-up but had no explanation for this habit. It seems that the prominent role of Aspen Mountain both physically and socially is probably the key factor (see Figure 2). It dominates the visual field, looming high, intruding closely; and it is the focus of the activity that most defines Aspen's identity--skiing. This hypothesis is reinforced by the testimony of local residents. One revealing instance was a resident who spontaneously observed that "Aspen Mountain isn't as important in summer as in winter. In winter I would likely have started with Aspen Mountain at the top of the map and worked from there." Other evidence provided by visitor response will be cited in the next section.

5.1.1.2 Sequence of features

The order in which individuals present features is a key to understanding how those features are organized in their minds. A good example of a sketch map by a resident is shown in Figure 51. Virtually all residents began by laying down the main route through town, Highway 82. This path was labeled as Main Street, which it is for all of its length except for the right/left jogs at each end of town. Main Street formed a spine that the rest of the grid grew out of. The most popular next addition was Mill Street, the most important major cross street, which becomes the only other exit out of town running north up Red Mountain. The streets most frequently chosen next were Hyman and Cooper, which constitute the pedestrian malls in the downtown shopping area. These two streets run east/west, leaving Mill Street as the only one of the first four running north/south. The most cited landmarks were noted in the following order: Wagner Park, Post Office, City Market, Hotel Jerome, Wheeler Opera House, pedestrian malls, Bank of Aspen, Aspen Mountain. It may help to clarify the roles of some of the special cases in this list. Wagner Park is the largest and most central athletic playing field, usually devoted to games of city teams in softball, soccer, and rugby. The Hotel Jerome, the largest and one of the oldest fixtures in town, has become one of the local "hangouts." Situated at the corner of Main and Mill Streets, it anchors the most important intersection in town. The Wheeler Opera House has a threefold identity (no longer including opera). It houses a movie theater, the Chamber of Commerce, and a locals' pub.

The significance of most of these data becomes more clear in comparison with similar data from visitors to Aspen, as will be seen in the next section. But one pehnomenon bears mentioning now. The most common technique in drawing a map was spatial, i.e., adding features in general spatial sequence according to proximity. Landmarks which are close to each other were drawn at about the same time. Remarkable exceptions to this rule were two cases in which features were added thematically, i.e., the sequence of features was based upon their role and meaning. If one park was drawn, for example, then a catalog of all parks scattered around the town would follow.

5.1.1.3 Density of detail

People generally became so enjoyably immersed in the map-making exercise that it seemed possible for the level of detail to increase indefinitely. It was necessary to encourage some of the more enthusiastic artists to declare arbitrarily their maps finished at some reasonable point. This was not true of the visitor subjects, who often felt that they had insufficient knowledge of detail even to begin drawing a map. Certain

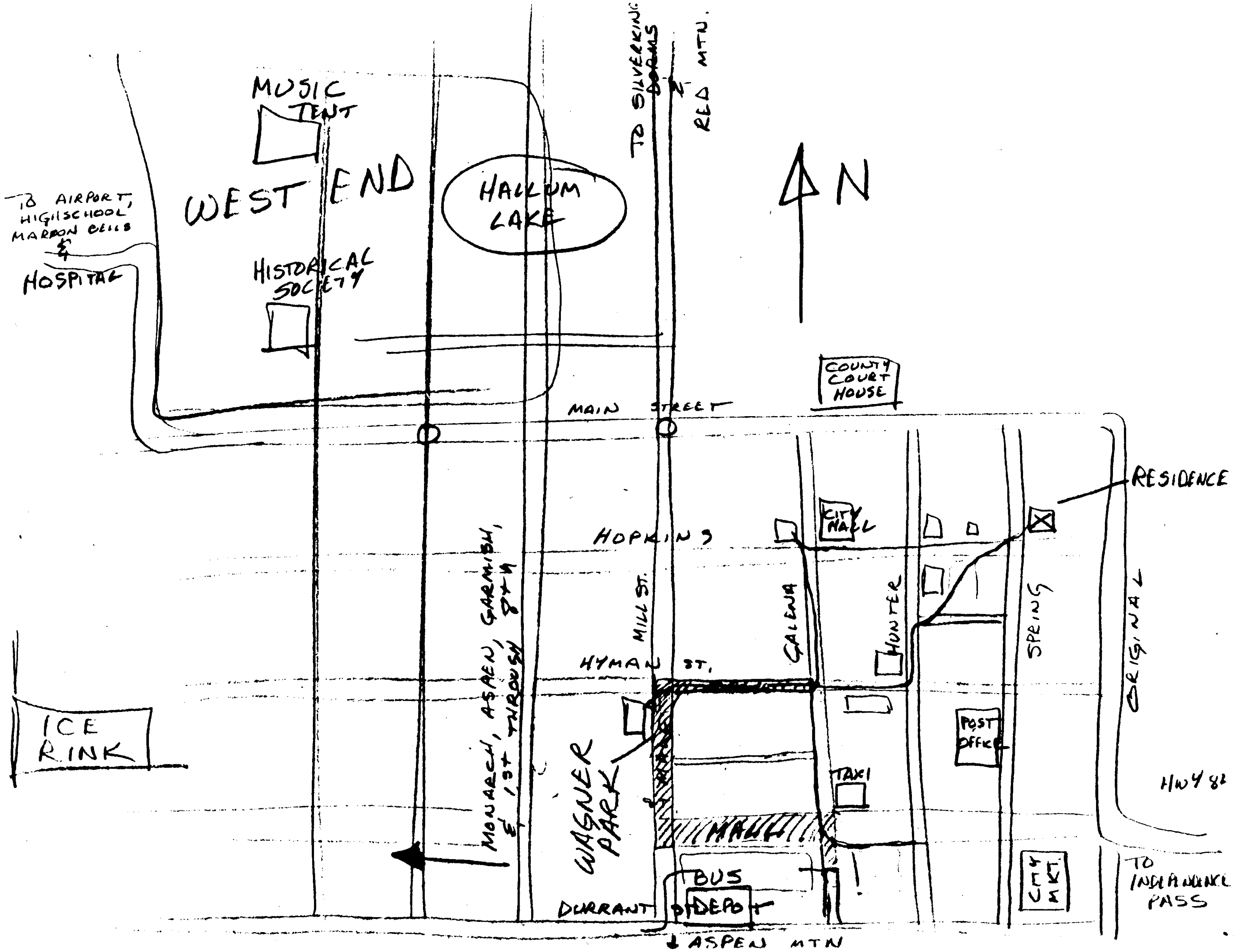


Figure 51. Sketch map by Aspen resident

consistencies were found among the residents' maps. The density of detail was much higher in the downtown area and along Main Street than in the residential area. The type of feature represented seemed to be subject to a proportional relation--the number of streets varied inversely to the number of landmarks. That is, on any particular map, if landmark density was high then street detail would be low and vice versa. Figure 52 and 53 show examples of each kind. Apparently some subjects organized the town more according to streets and others more according to landmarks; and these two modes were seemingly alternatives to each other.

5.1.1.4 Ambiguities

Ambiguities, uncertainties, and errors are, of course, particularly useful sources of insight into cognitive models. As just mentioned, subjects were able to provide much less detail in the west end residential area. Through their maps and comments they revealed a vagueness about the actual structural elements in that region. This is quite likely a combination of two factors--the relatively undifferentiated environment and the limited amount of time spent there. Another uncertainty was the identity of Garmisch, Aspen, and Monarch Streets, which run north/south between the numbered streets of the west end and the well known cross streets of downtown. The sequence of the three streets was often in a wrong order or unknown. Perhaps the most interesting ambiguity was the distinction between Cooper Street and Highway 82, running out of the east end of town. East of Original Street these two paths are the same, most commonly referred to by the state highway label. West of Original the two paths are distinct; Highway 82 jogs north down Original Street and then runs west along Main Street, parallel to and three blocks over from Cooper

Figure 52. Resident sketch map--high street detail

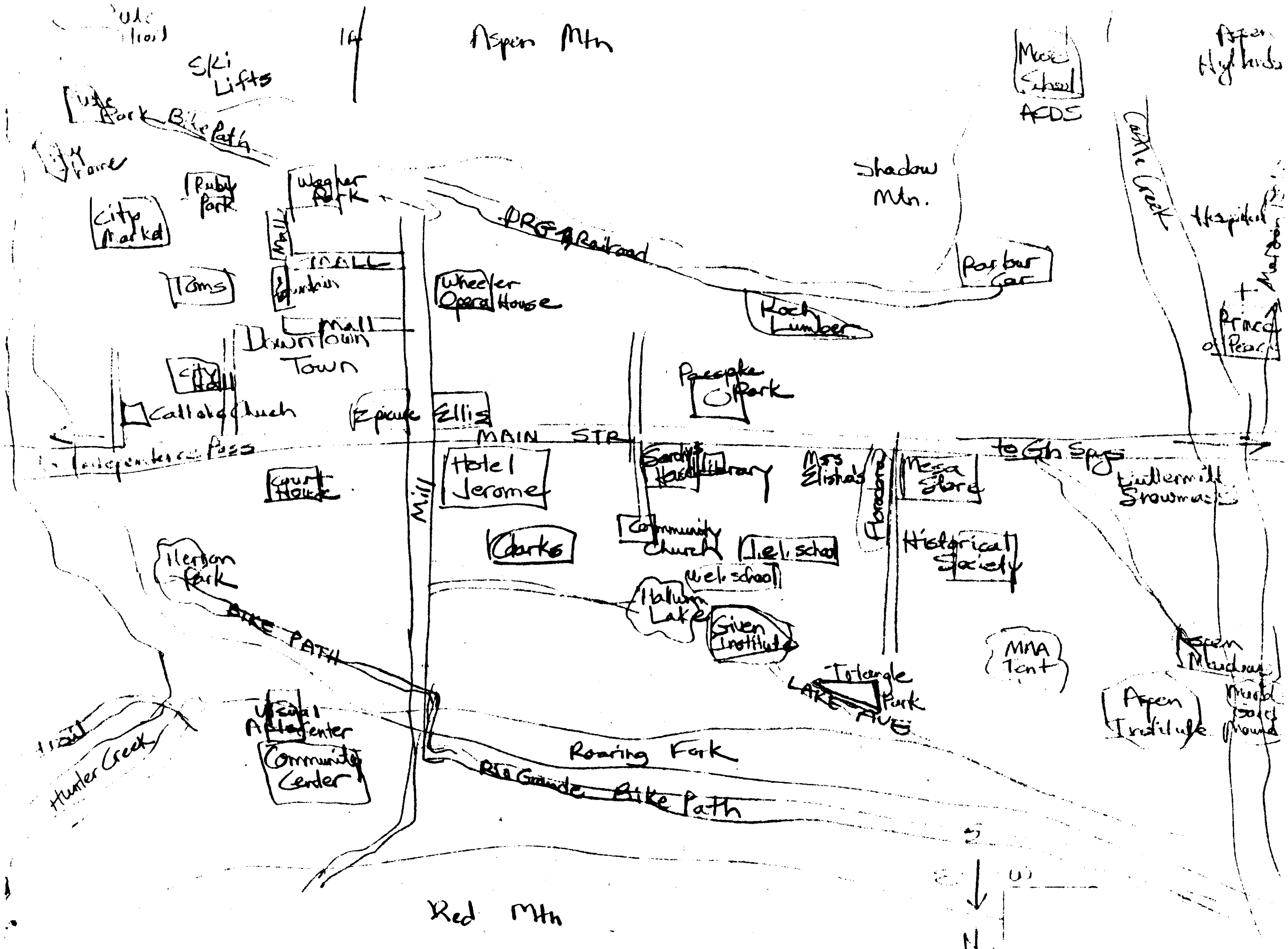


Figure 53. Resident sketch map--high landmark detail

Street. Frequently the connection between Cooper Street and east Highway 82 was left floating. As noted before, Highway 82 was typically drawn first as the spine of the town. Later, when the east end of Cooper Street was added, it often extended separately next to Highway 82, tenuously connected as an afterthought or left hanging unconnected. An example can be seen in Figure 54. The explanation is probably more subtle than the double identity of east Cooper Street. It also may stem from traffic patterns. When traffic exiting from Aspen on Highway 82 reaches the Cooper Street intersection its progress is halted both by a stop sign and no less significantly by the rising barrier of Aspen Mountain. The left turn onto Cooper becomes the obvious choice by default. However, the pattern is not symmetrical; traffic entering Aspen is not similarly constrained. Westbound traffic encounters no stop sign at the Original Street intersection. And it faces no visual barriers; Cooper extends right into the heart of Aspen. Consequently, a lot of traffic continues straight on Cooper. This pattern is the behavioral reflection of the dual identity of the eastern Aspen access, i.e, westbound traffic follows Cooper Street and eastbound traffic follows Highway 82.

In general ambiguities seemed to crop up most among residents with the shortest history in Aspen. The most recent arrivals (one to two years) in the sample were the least certain in distinguishing between Hyman and Cooper malls, for example. The one exception to this pattern was in knowing street names. The extent to which streets were correctly identified by name did not seem to vary much from new arrivals to old-timers; <u>neither</u> knew street names very well.

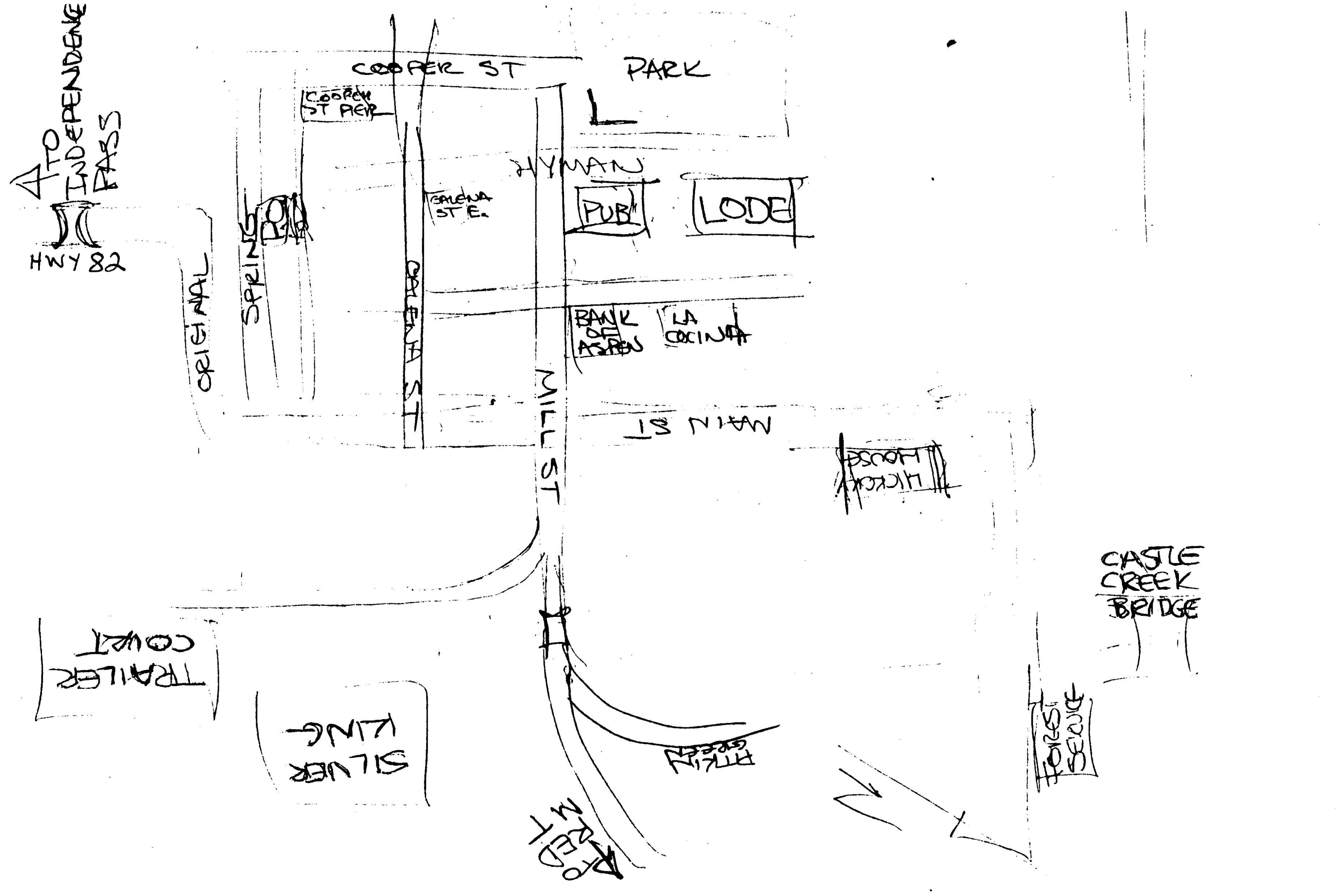


Figure 54. Ambiguity between Cooper Street and Highway 82

5.1.2 Verbal answers

The second half of the inerview consisted of a questionnaire (see appendix 2). Again, the most representative and the most interesting results will be cited.

A number of additional mabiguities were revealed. "Down valley" was a source of confusion to some residents, especially when they first arrived. "Down valley" refers to going down in altitude along the valley heading northwest. "But," as one subject pointed out, "you'd think 'down' should be south." Notice again how north and south orientations are reversed in Aspen. The crowning touch to these reversals is West End Street, a minor street on the <u>east</u> end of town. For some reason few people were aware of this contradiction and no one was bothered by it. Most residents said they were able to distinguish between Hyman Street Mall and Cooper Street Mall by particular stores in each one. The exceptions were those who were unable to verbalize how they made the distinction and one person observed that Hyman was the heavily shadowed mall, especially in winter (because its south side facades are much taller). One resident whose first exposure to Aspen had been in winter made a refreshing comment about Aspen Mountain. She said that every time spring came she was always surprised to be reminded that the ski slopes were really green meadows.

In all these cases, in both the interviews and the map-drawing exercises above, the ambiguity or confusion derives from an interplay of the physical environment, cultural conventions, and individual cognitive models. The underlying cause is the lack of complete coherence between two representations. Using the ambiguities of residents' perceptions as a reference will prove fruitful in later comparisons with the responses of visitors and movie map users.

Some of the responses were not surprising. The symbol most commonly used to identify Aspen was the aspen tree or aspen leaf. Other reactions were unanticipated. Almost all subjects included references to smells in their descriptions of trips through Aspen (bakery, Hickory House, sewer, traffic). As a matter of fact, smells were cited more often than sounds.

One of the most fascinating revelations was a unique path-choosing behavior which residents employed when they were in a hurry to get from one place to another. Aspen being a small, friendly town, everyone is on a first-name basis with a good percentage of the population. So a local person walking down the street will probably encounter at least one acquaintance in every block. The result is a rather leisurely progress sprinkled with conversations here and there, often over a cup of coffee. Naturally, hurrying over to the bank without being rude to friends presents a dilemma. The clever solution, pointed out to me by a subject, is to travel via alleys. The universality of this technique was spontaneously confirmed by another subject who referred to it by name--alley walking. Both subjects also commented that alley walking was becoming so popular that they were beginning to run into just as many acquaintances there as on the street--and that the topic of conversation was predetermined: "Why are you trying to avoid people?"

The last interesting finding came not in response to any particular questions, but upon introspection about the whole interview. All subjects enjoyed the interview and usually became quite engrossed in the exercise.

Some of them commented that they were surprised about the extent of knowledge and patterns of behavior which they became aware of for the first time through these interviews. For example, one subject was fascinated to discover that she did have favorite routes, that she made hundreds of subconscious path-choosing decisions during everyday outings, and that most of her choices were made for the simple joys--opportunities to see a pretty flowerbed along the way, to pass a friend's house, to smell the cottonwoods. Also revealed in these introspections was a particularly powerful insight which may prove to have important ramifications for the movie map system. At least two residents pointed out that events can assume the role of landmarks. For example, one subject commented that her daily route went by a cabin where her dog had once broken down the door. She reflected, "You know, every time I pass that cabin I still think about the incident with my dog, even though it happened over ten years ago."

5.2 Aspen Visitors

Interviews very similar to those for residents were conducted with 10 first-time visitors to Aspen, again of both sexes and a range of ages. Most visitors were eitherin Aspen on vacation or were taking part in professional activities at the Aspen Institute and the Music Festival. The average length of stay in Aspen at the time of the interview was about three-and-a-half days. When it was possible to contact these subjects upon arrival, they were encouraged to avoid looking at maps. In every case the amount of time spent looking at maps was not more than a few minutes. The most pertinent results will be summarized and compared with the residents' responses.

5.2.1 Map drawing

5.2.1.1 Orientation

A majority of maps were drawn with Aspen Mountain at the top, and of wrong guesses in assigning the cardinal directions, the most common placement of the arrow pointing north was in the direction of Aspen Mountain. This reinforcement of the orientation pattern exhbitied by residents will be further amplified by the verbal responses of visitors.

5.2.1.2 Density of detail

As one would expect maps of visitors were much sparser than those of residents (see Figure 55 for an example). And the scope was narrower; maps usually featured a particular subsection of town. Even when the extremities of Aspen were included, a particular subsection would usually be drawn larger, out of proportion. There was a greater variety among visitors' maps than among those of residents. The explanation may be that the familiar subsections which were favored varied among subjects as a function of where they stayed and visited in Aspen. There was no repetition of the inverse relation between the density of landmark and street detail which residents' maps had exhibited. Visitor maps were more or less balanced between the two.

5.2.1.3 Sequence of features

Visitors also began their maps with the Highway 82/Main Street spine, But they next generally added Cooper and Hyman Streets, sometimes simply labeled "mall." The next most frequent additiona was Durant Street. All

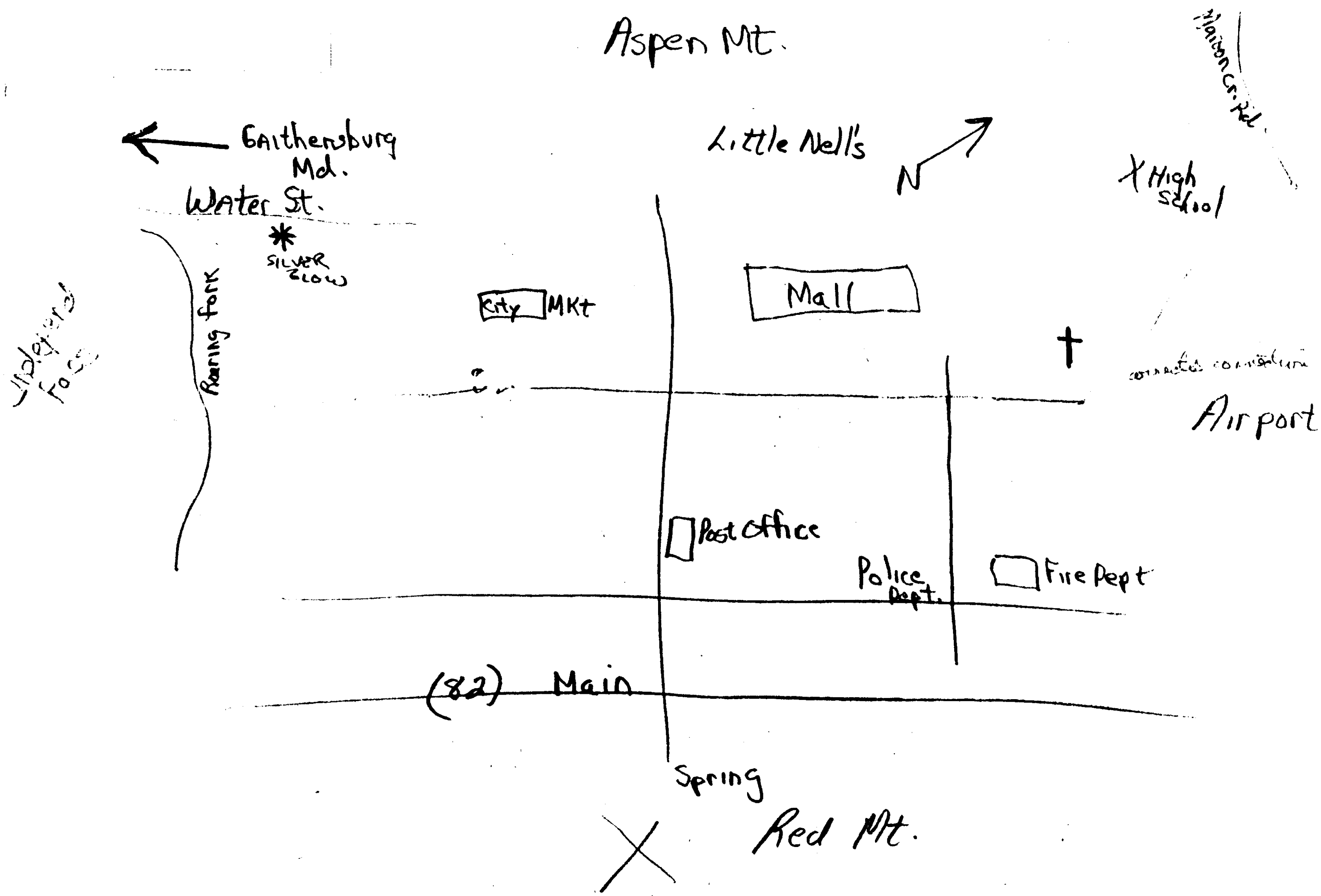


Figure 55. Visitor map of Aspen

the streets mentioned so far run east/west. When north/south streets were drawn next, they turned out to be Spring Street and Aspen Street, not Mill, which residents gave second place out of all the streets. One conjecture is thatthe importance of Mill Street as the northern access to homes on Red Mountain is not as likely to be felt by the visitor. But how does one account for Galena Street, which was also more emphasized by residents than the two favored by visitors? No compelling explanation comes to mind. Nevertheless, it is clear that visitors to Aspen fail to adopt that simple model of the residents--the Main Street spine, crossed by Mill Street, with the Hotel Jerome anchoring the intersection.

The landmarks most frequently cited also showed significant similarities with and differences from those chosen by residents. Wagner Park, City Market, and the Hotel Jerome were accorded the same weighting by both groups. Such a high showing by City Market seems surprising because it is not generally frequented by tourists and it is located quite out of the way. The other residents' landmarks which were not expected to figure prominently in the visitors' schedule were indeed missing, namely, the Post Office and the Bank of Aspen. Moving up on the list were two elements which quite naturally fit into the visitors' profile more readily than the residents': the pedestrian malls and the Music Tent. There was also a dramatic incrase in the emphasis upon Aspen Mountain. While this finding would fit very nicely with the pattern of other references to Aspen Mountain, this evidence cannot be given much credibility because of an ambiguity in instructions which may have encouraged subjects to draw in the mountain.

5.2.1.4 Ambiguities

Certainly the visitors' maps exhibited much greater ambiguities and errors than the previous group's. Significantly, most of these ambiguities were reminiscent of the residents' but written much larger. This correlation implies something which was not anticipated--that the two groups may share the same underlying cognitive models, differing only in stage of development. Of course the evidence is not formal enough to completely justify such a conclusion, but the examples are suggestive. The confusion about the duality of east Cooper Street and Highway 82 was even more pronounced on the visitors' maps (Figure 56). Most maps omitted the jog at the eastern exit from town and showed Highway 82 extending straight out from Main Street, parallel to and separate from Cooper Street. One subject illustrated the confusion perfectly. He drew Cooper Street correctly, running east out of town toward Independence Pass. Then he incorrectly extended Cooper Street toward the airport as the westward access and labeled the entire stretch as Highway 82. But he realized that something was wrong and moved the Highway 82 label to Main Street. He was able to picture Cooper Street as the main route through town and Main Street likewise, but what his schema lacked was the right/left jog to combine both streets into the same route.

The west end residential area, which was only sparsely detailed by locals, was either indicated as one big, undifferentiated blob or not acknowledged at all by visitors. This makes sense since visitors have even less reason to spend time in the residential area, except possibly passing through to the Music Tent. The malls were poorly distinguished as a rule. They were often represented as a single mall, lumped together

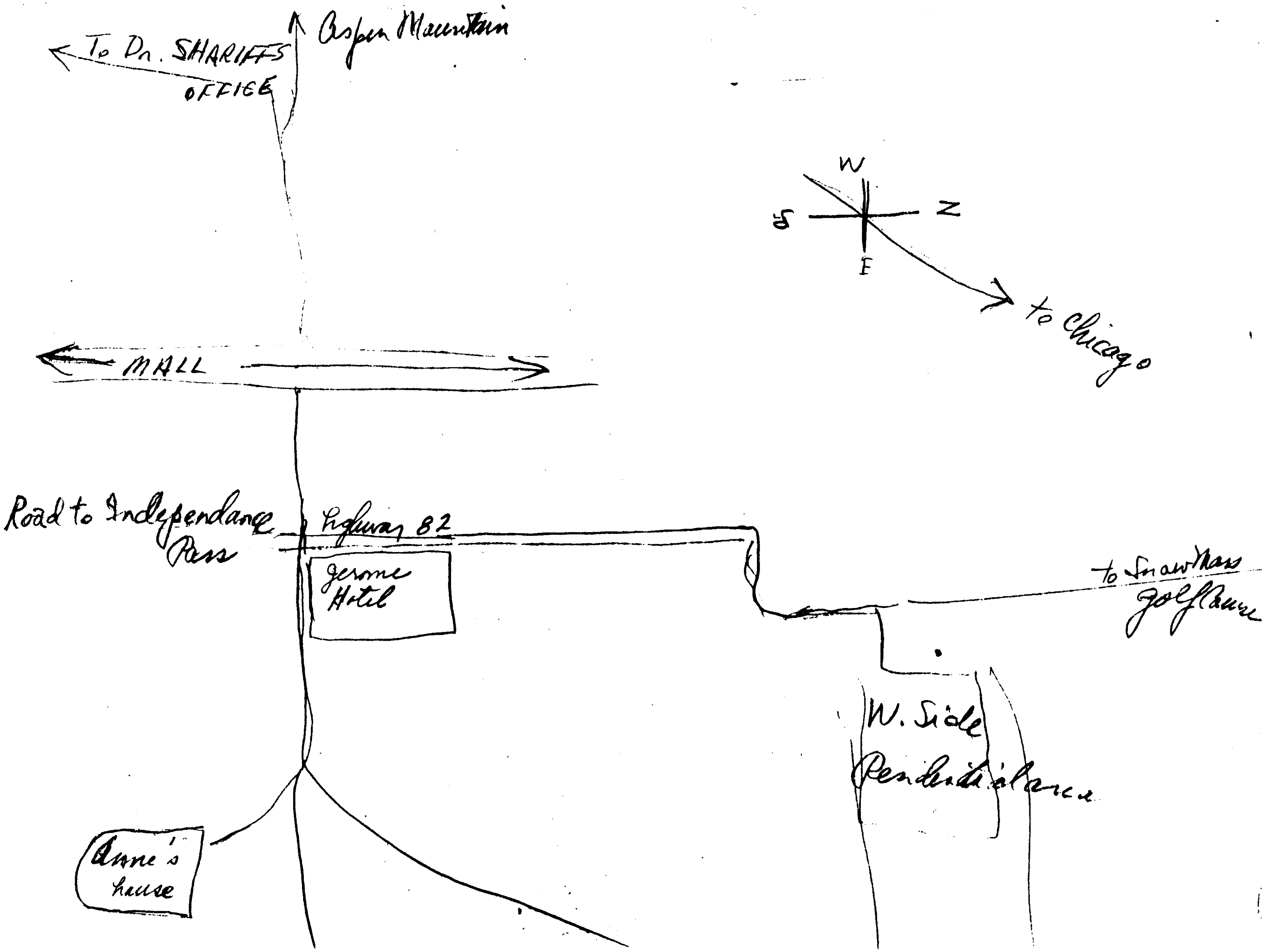


Figure 56. Visitor ambiguity about Cooper Street Mall and Highway 82

nto a single region, or separated into the two east/west malls on Hyman and Cooper. No one indicated the minor north/south mall sections on Mill and Galena. Residents were generally explicit about distinguishing th- two major malls, but only a few indicated the minor malls.

The phenomenon of "floating" areas was much more pronounced with the visitors. It has already been noted that residents' map had manifested one such example in the ambiguity of the Highway 82/Cooper Street intersection. There were other similar examples which showed that visitors often possessed accurate spatial knowledge about several subregions but not about the spatial relations between them. The process by which these submodels are integrated into more global structures with internal coherence is at the heart of spatial learning and is to some extent represented at two stages by visitors and locals. The suggestion is that spatial representations are manifested on global levels and on local levels. Naive subjects like visitors are in the process of creating both, trying to integrate the latter within the former. Native subjects are mainly adding detail at the local level and occasionally making refinements to the global structure. Part of the problem in interpreting subjects' representations is that global inaccuracies seem more flagrant than local ones, even though they may be errors of "equivalent magnitude" on a cognitive scale.

The final comparison of ambiguity concerns cardinal direction. All residents drew the north-pointing reference arrow on their maps correctly, though some only after much deliberation. Visitors, of course, did much poorly, with only 4 out of 10 marking north correctly. Half of the remainder showed north pointing toward Aspen Mountain, in the opposite direction. Several of those in error employed admirable strategies, like referring to the position of the sun, only to make a mistake in one of their premises.

5.2.2 Verbal answers

The verbal part of the residents' interview was slightly revised for visitors to include some questions about the techniques they used in learning they way around an unfamiliar environment.

When asked what symbolized Aspen for them, three-quarters of the visitors responded "skiing" or "mountains," where residents had singled out the aspen leaf. This pattern has been demonstrated time and again by visitors. They use Aspen Mountain to orient not only their maps and their compasses but even their image of the town.

The examples of ambiguities revealed in verbal answers helped clarify those already suggested in the maps. In terms of global structure, subjects, pressed to describe the overall regions of town, typically divided it into downtown, the mall, and everything else. The west end residential area was largely unknown and when it was referred to, the only real detail called to mind was the pretty Victorian architecture. One sentiment echoed several times was that all the buildings looked the same (reminiscent of racial stereotypes: all Orientals look alike to Westerners). This inability to distinguish extended to other major elements of the environment, namely, streets and mountains. The malls provided a source of extreme confusion. Visitors were uncertain about the size, layout, and number of malls. Many didn't realize there were more than one. Someone commented that it took a while to recognize them as two malls. Almost no one made the local distinction of the Hyman mall as the "main" mall, being the busier of the two with three times the number of stores. Visitors who did generally were able to differentiate between the two malls on the basis of a store or twho which they had visited.

Path-finding descriptions were good sources for understanding how visitors acted in an unfamiliar environment. In one case, a ten-year-old was able to provide more detailed descriptions of things seen in sequence along a path taken in Aspen than her father was. Her description even included detail at the level of describing a picture of a bearded man in a top hat that one would pass along the way. (The father, on the other hand, was able to adddress a much larger scope in his map-drawing.) This girl also verbalized her strategy for finding her way around in a new place. She would explore a small area around her home base and gradually venture farther and farther away, always staying within a psychological boundary where she was confident that she could find her way back. Being able to find her way back was clearly a source both of security and pride.

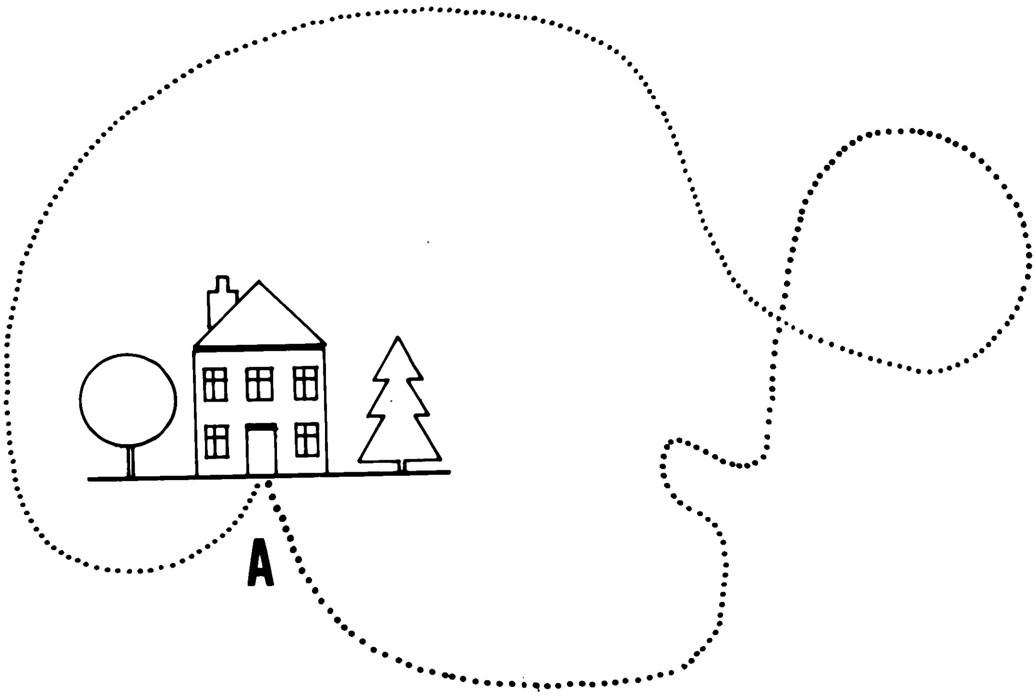
The procedure of another youngster, this time a fourteen-year-old boy, provided an alternative which is worth looking at in depth. The route he was most familiar with was one he took daily from his home on Waters Street at the southeast corner of town to the downtown malls. In his description of the path he recalled the first right/left mirrored as left/right. He later admitted that he sometimes said left when he really meant right and vice versa. But whether or not that factor was in play here, it was clearly more a symptom than a cause of the problem, because the same lapse showed up in two other nodes. In the first, map drawing, he introduced a similar error at the same corner, which had the effect of translating and rotating the spatial relations between some regions (and which should remind us that seemingly large mistakes can often be explained by very local errors). The second manifestation of the error actually happened in the field. After the first two parts of the interview we went for a ride with him to give directions for the path he had just described. I had expected that when faced with the real space he would correctly represent the unseen choices ahead. But he followed the established pattern and instructed me to make a left at the end of the street. As we approached the turn he began vacillating in his instruction and finally decided he would wait until we got to the intersection before he could make up his mind. Upon arriving at this point he was able to call for the correct right-hand turn without hesitation. (He also was immediately aware of his previous reversals in both the map-drawing and path-description exercises.)

This behavior raises an important issue. It suggests that the process of path finding and following may not be appropriately assessed by exercises which are inherently different from the initial activity. Verbal description involves recall; map generation involves both recall and abstraction (or transformation) of data. Battro [1979] has already pointed out that the problem of transformation and drawing skill can be largely overcome by replacing the mapping task with one of arranging a three-dimensional model. But mental recall is still the inherent mechanism in all these cases. The case of this boy suggests that recognition may also (or instead) be fundamental. In that case, the internal representation is dynamic over time, changing sequentially with feedback. This role of recognition would call for evaluations of a subject behavior to take place within the context of the real environment or a reasonable facsimile thereof. Just as the transformation from a sequential experiential path to a map is not obvious, neither is the converse--the assimilation of a map overview within path-finding experience. The transformation between these two kinds of knowledge and two kinds of experience remain one of the most fundamental issues which the movie map must ultimately address.

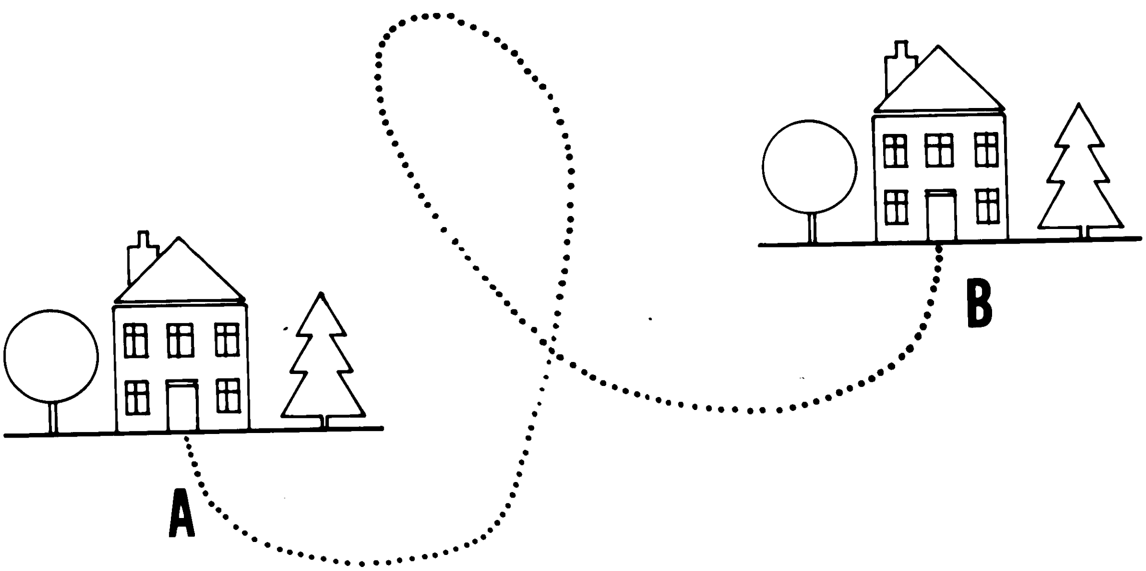
A few final glimpses into this case study raise related issues. Given the boy's record in recalling incorrect choices, how was he able to find his way back and forth to the mall with relative success? One technique which he employed was revealed in his discussion of the asymmetry between the back and forth paths. Going from his home to the malls was an easier task for a number of reasons: the destination target was bigger, most other pedestrian traffic converged on the malls, the malls were accessible from all directions. Going the other way was the opposite: the target was very small and no one else went that way. Also, there was only one successful access route: the only connection to his home on Waters Street was from West End Street, and the only way to get there on West End Street was from its intersection with Durant Street. He reported of the first time he left home, "I said to myself, 'Make sure you turn on West End or you're in trouble. Never turn on Original.'" Invoking this rule got him successfully through the one critical point in his route. The question is, How did he know he needed to establish that rule in the first place? Still, he had managed to get slightly lost a few times. In his terminology he ended up in "nowhereland" which he gave the charming definition, "when you don't KNOW WHERE you are."

An ambiguity cropped up in his recollection of the athletic field(s). He knew there was one large one downtown (Wagner Park), but also thought there was a second slightly smaller one someplace else where the girls played city league softball. In fact, they are the same. One essential construct of internalizing spatial knowledge is that thesame place, seen from two different views, must be identified as the same to preserve its location at a single position. While that may sound trivial, the consequence of failing to recognize that two experiences have the same referent causes a severe distortion of reality, in this case, space. Topologically the distortion is equivalent to turning a single point into two points, a closed curve into an open one (Figure 57).

Finally, this behavior raised the issue of "dynamic landmarks." These landmarks were identified not only with space but with time as well. They were events. The boy's description of his daily trip included comments about places along the way where "you probably see people on roller skates" or a street where "there's always lots of motorcycles." Other subjects also used events for landmarks, for example, "there's always a softball game going on (at Wagner Park." "You can hear the kids screaming two blocks away." Residents made many references to dynamic landmarks. It is interesting that visitors also made several such references. After all, recognizing such landmarks requires observation of multiple instances of an event to correlate frequency with location. It also requires an understanding of the causal relationships, which implies some knowledge of the cultural world of Aspen.



Thinking same house has been visited twice.



Thinking two different houses have been visited.

Figure 57. Two differing cognitive map topologies from the same trip

5.2.3 Summary

Before turning to the movie map subjects it may be helpful to review the most relevant findings of Aspen visitors and residents. Both groups possessed a lot of information (typically more than they realized) about the structure of Aspen. There was a difference in kind and detail of information between the residents and visitors. The experiences with, and meanings assinged to, the environment were key factors in the way the space they could structure cognitively. Many of the differences between visitors and residents seem to be due to differences in these relations with the environment. In other cases, visitors exhibited the same patterns as residents at a seemingly more primitive stage of development. In general subjects were able to externalize representations of the space in two distinct ways--by recalling the sequence along a path (one dimension of time, one dimension of space) and by generating maps (two dimensions of space). One characteristic of ground-level sequences along a path is that they are not necessarily reversible, i.e., the processes and strategies involved in making a trip one way may not apply in going the other way. Finally, subjects' cognitive structures included not just spatial referents but temporal ones as well; their models described an "event world."

5.3 Movie Map Users

Eight subjects who were unfamiliar with both movie maps and the town of Aspen were observed using the surrogate travel system and then interviewed in a format similar to that used for previous subjects. The subjects ranged in age from mid-twenties to mid-fifties, half men, half women.

(Investigations that include children are contemplated for the future.)

Different subjects were presented with different configurations of the surrogate travel system. There was considerable variety--regular-size television display versus an eight-foot-by-ten-foot screen, graphic overlays as navigational aids versus none, velocity control versus acceleration control versus fixed speed, and leading turns versus pivot turns. These trials were carried out while the system was still in the stage of development and debugging. At this stage access to aerial over-views had not been implemented.

Another eventual feature of the system which was not in place was a verbal "information service" which supplies, via sound and/or text, the names of streets, the direction of travel, and names of particular land-marks requested by the viewer. To simulate this feature, the observer provided this information for the subjects, generally at their request and refraining as much as possible from answering questions that the computer would not be likely to, such as, "Is this the rich, exclusive part of town?" The procedures for introducing users to the system were also being debugged; verbal instructions and explanations were modified as it became clear which were most useful.

Subjects were briefly told that the computer-controlled system allowed them to see film sequences taken in Aspen, to "drive" around, and to get to know the town. They were not told how the mechanics of the system worked (unless they wanted to know afterward). When they asked about the purpose of the system, they were told that it was an experiment to see how well someone could learn to find his way around a place where he had never been.

When users traveled around for the first time a certain amount of acclimation was necessary. This period usually extended from one to four minutes and was used to allow the subject to become familiar and comfortable with the joystick control, the responsiveness of the computer, and the feel of the motion with regard to seepd, smoothness, and size of image. During this warm-up subjects were told not to worry about where they were going, but just to see and feel how the system worked. Then they began exploring the town. If they became lost or for some reason wanted to start over they could begin again at the original starting point (or at any other requested point). This was also the procedure if the program halted or the computer crashed. The default starting place was on Main Street at the intersection with Mill Street heading west.

5.3.1 User response

5.3.1.1 Strategies

Viewers exhibited different strategies for exploration. The most common pattern was not anticipated. It consisted of striking out in an arbitrary direction and continuing more or less in that direction until it dead-ended at a movie map boundary at the edge of town. In most cases when the viewers then elected to start over, they set out from the original starting place in a brand new direction. They proceeded again until they could go no farther and started over again with yet a third heading. It is not as if they were assimilating global spatial relations. From their comments and questions it was clear they usually had not idea about the directional relationships among any of these radial trips. As a matter of fact, they had exposed themselves so much that they typically could not duplicate any of these forays. If they happened upon the same path again, comments would be of the sort, "Oh, am I at this place again?" It seemed their intent was not to structure the spatial relations of paths, but to sample the contents and find the edges. They wanted to find out how big the "fish tank" was and what sort of things it was stocked with. It seems likely that this "world-defining" pattern was promoted by two characteristics of the movie map:

a) No overview was provided (aerial photo, oblique view, map, verbal introduction).

b) There was almost no penalty for getting lost; it was trivial to start over.

This hypothesis was tested by giving later subjects a very quick (30-second) overview consisting of 4 oblique aerial slides from the edges of town and a few comments about the size, major regions, and main street. The two subjects given this overview spent much less time in boundary-seeking, world-defining behavior at the beginning. No conclusions can be drawn from such minimal data, but the importance of further examining the role of the overview and techniques for implementing it is suggested.

An opposite and more anticipated strategy was exhibited, particularly by one subject. He used what might be termed a "home base" method. In several instances, after he had ventured forth some distance and several turns, he would return by retracing his route to a familiar starting place. He seemed to be fixing one section at a time in his mind. Surprisingly, he rejected this suggestion, denying that he was aware of any strategy.

In a similar and perhaps later stage of this pattern another subject would crisscross through an increasingly familiar region, trying to anticipate and confirm where things were. She even said that she didn't want street names to be announced, but rather wanted confirmation after she announced her own prediction.

Finally, there was a strategy which was more unstructured even than the first one. This was exemplified by a sightseer, who never really cared where he was; he was content just to wander around looking at the sights. (Is it significant that of the five subjects who later went to Aspen, he is the only one who has not been heard from? Is he still wandering around, lost?)

Some features of surrogate travel will now be addressed independently. As stated early in this paper, the movie map provided an extremely viable format for naive subjects to learn about and experience a new space. All subjects were delighted and stimulated by their encounter with the system (and with Aspen). It is within the context of this overall success that the following examination of specifics should be considered.

5.3.1.2 Exterior landmarks

The mountains played an important orienting role for many but not all subjects. Subjects were most able to recognize Aspen Mountain and Red Mountain, which lie to the south and north respectively. The interpretation of these exterior landmarks was not always automatic. Sometimes it required deliberation, as demonstrated by the following excerpt: "Oh, is that always the same mountain? Every time I turn I'm headed for that mountain. . . But I was going east, now I'm going north. Either that's the same mountain or they all look the same. Aspen must be surrounded by mountains. It's in a valley, right?" Part of the reason the mountains were indistinguishable was suggested by the subject who cited a lack of prior experience of being in the mountains.

5.3.1.3 Turns

In several instances viewers commented that they lost their bearings during turns; they did not know how far they had just turned, sometimes mixing up 90 degrees with 180 degrees in pivot turns. In most of these trials the movie map used pivot turns much more predominantly than leading turns. (Pivot turns are subarcs of the 360-degree pans about a point; see Figure 13. Leading turns are those in which the camera movement anticipates the turn; see Figure 15.) Leading turns seemed to be easier to interpret in the few cases in which they were used. A more significant factor in the success of pivot turns may have been the success of other parts of the system. For example, when the joystick was still cumbersome and not fully debugged subjects said that they were distracted from concentrating on the screen; and similarly, without graphic overlays to provide feedback acknowledging that an instruction (like indicating a turn) had been received by the computer, user confidence was reduced. This mechanical uncertainty was perhaps translated into a perceptual uncertainty. When the joystick, overlays, and computer worked flawlessly, viewers seemed to follow the action more clearly. The hypothesis is that if viewers have faith that the system will execute the turn correctly, they are more willing to interpret the turn appropriately. Believing is seeing.

5.3.1.4 Discrete space and time

Subjects were typically surprised by the discreteness of movement along the street. Fewer number made comments about the way objects sometimes disappeared. In all cases except one, they quickly adapted to these artifacts of the system and gave them no further comment. The viewers became accustomed to the range of speeds available but retained a residual amusement at the 60 mph roadrunner pixillation. A few brought up the cartoon analogy themselves and one person compared it to a driving test where pedestrians jump out to test your reflexes. After observing the initial subjects' reactions it was decided to prepare subsequent viewers for most of these aspects in the verbal introduction. Subjects often expressed uncertainty about the distances they traveled between two points. They felt unsure of both the absolute and relative scales.

5.3.1.5 Editing

Jump cut edits (between two visual sequences) were generally interpreted smoothly. That is to say, shifts in registration, exposure, lighting, or color balance between edit frames were apparently seldom noticed, according to subjects' reports. Of these, registration was the most critical. Viewers could tolerate a larger variation in exposure if the important elements stayed lined up. Similarly viewers could tolerate a large variation in lighting if the overall exposure remained constant. That is, changes in shadow position were inconsequential. The only exception was frames that included large, opaquely shaded regions without detail. Interestingly, the smoothness of a given edit might be judged differently during separate observations if the viewer was looking at different parts of the frame which did not match up equally well in editing.

5.3.1.6 Viewpoint

Viewers seemed fairly oblivious to the height of their viewpoint above the ground. And after some initial reaction they seemed to be unconcerned that they were traveling down the middle of the street. Driving onto the pedestrian malls was not so casually dismissed. Most people thought that the malls would be off-limits to their vehicle, and were unnerved to find otherwise. Finally, several subjects commented on the narrowness of the field of view available and the lack of any global overview. The point was repeated even more strongly after they visited Aspen (and saw what they had been missing?).

5.3.1.7 Transitory cues

Subjects claimed that they filtered out irrelevant material in the movie map by using basic common sense. They commented that they ignored certain transitory cues, like fall leaves on the street, just as they would in real life. Viewers were much less cognizant of the 100-percent repeatability of traffic and other events on a given street than had been anticipated. The obvious reason is that they did not spend nearly as much time as designers of the movie map had traveling the same paths over and over again. The only section they were likley to repeat more than a couple of times was the starting place. The more subtle reasons is also perhaps more compelling: subjects were oblivious to the repetition of isolated events because they were ignoring them as irrelevant to begin with. No one had explained the mechanics of the system. Why should they expect everything to be redundant. Of course this interpretation raises an objection: if viewers are to induce event probabilities associated with location, they must distinguish between the exact repetition of film footage and the repetition of a similar event. An example would be the difference between seeing the same bus twice in the same place and seeing two different buses, implying a bus stop.

5.3.1.7 Video displays

The only surprising reaction from subjects was that the eight-by-ten-foot rear projection screen did not receive universal praise. The subject with the most exposure to it complained about the blurriness and jerkiness of the image. The limitations both of resolution and camera stability are, of course, accentuated on the large screen. The other side of increased visual involvement of the user is the danger of visual bombardment.

5.3.1.8 Navigational controls and aids

Graphic overlays confirming the user's instructions improved user comfort and confidence significantly. The joystick was more problematical. As already mentioned, in some ways it was too responsive, providing too many degrees of freedom for the actual driving options subject to user control. The touch-sensitive display had a more appropriate level of interaction. On/off controlls like turn signals were implemented with corresponding dual options. The second advantage of the touch-sensitive display was that the viewer was never distracted from the screen to attend to the control device. All action, control, and response happened on a single device. Verbal information was welcomed selectively. It obviously provided a needed context and reinforcement for the user, but the value of the information varied with circumstance. In the beginning, subjects found directional headings useful. Later the names of the mountains associated with the main directions were displayed (even though they were not necessarily requested). This procedure proved successful and quickly assimilated. After a few instances of each direction, requests for repetition of information diminished. This was true for most names supplied to the subject. After being sufficiently briefed, some subjects preferred having their own guesses and predictions confirmed rather than receiving the information passively, as reported above. Generally, they seemed to want to be informed of building names more than street names. Many subjects were not interested in street names (just as local residents had not been). Those who were modeled the street hierarchy in their requests. They wanted to know the name of the main street first and later wanted to hear the names of some of the lesser streets. Requests for repetitions of building names continued much longer. Lastly, some subjects said that they would have liked to have auxillary maps for occasional reference.

5.3.2 Map drawing

Naturally, there was a map-drawing exercise when subjects had finished traveling around. On a coarse level these maps were comparable to the maps that visitors to Aspen had drawn. The amount of detail was roughly equal, the number of errors and degree of uncertainty similar. But there was something different about the styles of these two groups that is hard to identify. The surrogate travel maps were perhaps more linear. They seemed composed of paths more than areas. Feature detail along the path spines was high, but dropped off rapidly with offset distance. The visitor maps gave a better sense of the overall grid and the boundaries of the town. Certainly part of the explanation may simply be exposure time. Visitors drew their maps after several days in Aspen and many hours wandering the streets. Movie map subjects spent an average of only one hour driving around. It was impressive that their maps demonstrated so much knowledge after such a short time. Another possible part of the explanation is too obvious to ignore. The more linear localized representation may be a consquence of the "blinders" in the system, i.e., lack of peripheral panoramic views and side-to-side head movement.

Other differences between the two groups of maps are straightforward. The surrogate travel group captured the honors for assigning the cardinal directions correctly. They fell down quite sharply in the spelling department, however. Both results can easily be attributed to differences in the kind of information each group tended to receive. Surrogate travel subjects exhibited slightly more uncertainty about distances in their maps. This was not just a question of parts being drawn out of scale, which was evident even in resident maps, but of subjects explicitly indicating that they just did not know how far it was between two points. The distance along Main Street was a common example. Surrogate travel has a very regular metric: frames are exactly then feet apart. However, it is so easy to change travel speed without accompanying cues like engine sounds and acceleration forces that frame rate does not automatically translate into a good subjective feel of distance.

One surrogate travel subjects was an exception to the general comparisons between groups above. Her map was probably more accurate and more representative than any of the visitor maps (Figure 58). It featured the key elements in the genre of resident maps. It showed the Main Street

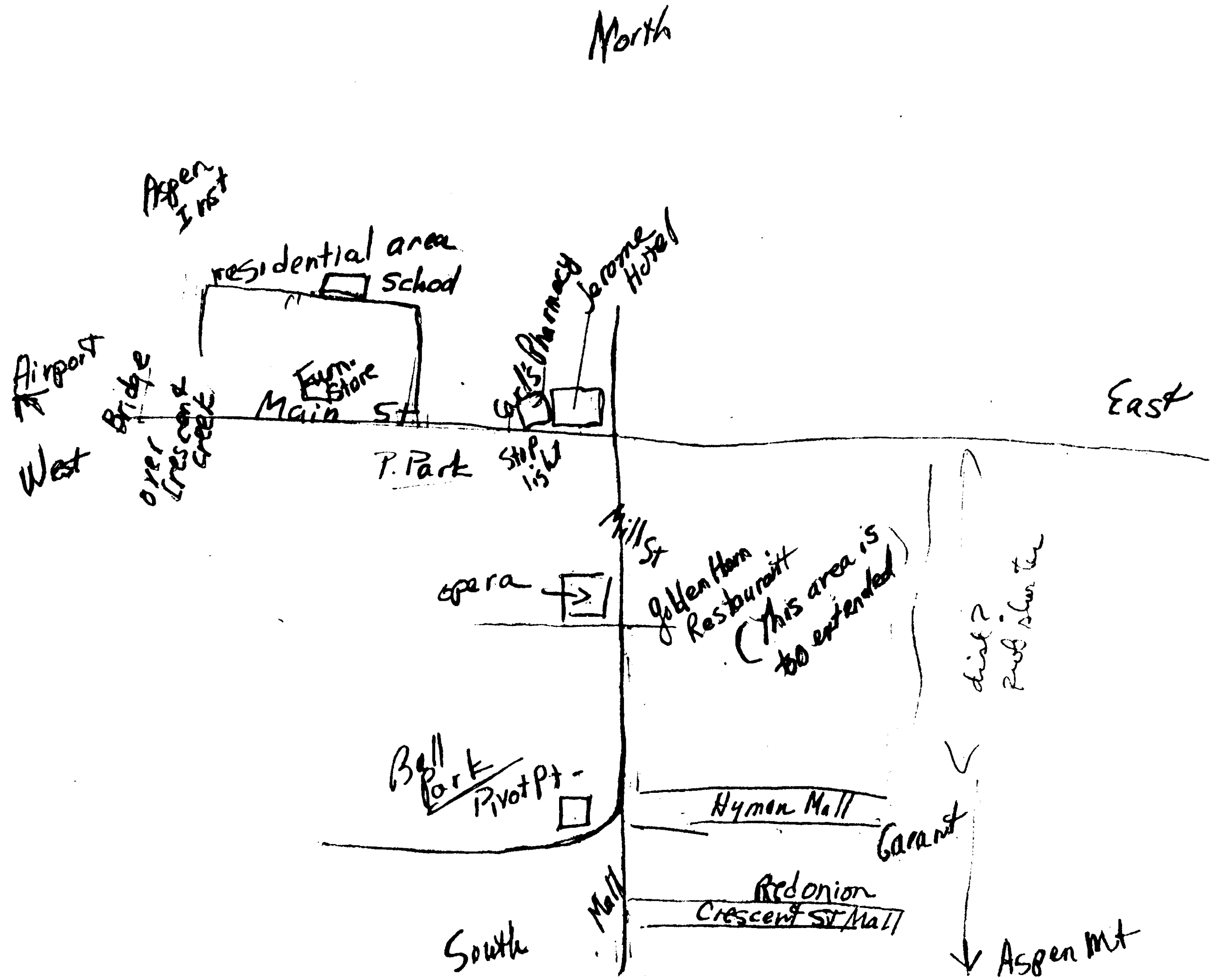


Figure 58. Sketch of movie map subject

spine, Mill as the major cross street, the Hotel Jerome anchoring the two, the two distinct malls in the downtown area, the west end residential area, and the major parks. It is tempting to try to isolate the exact configuration which led to her success. Her viewing station was the large-screen display; she was given a well formulated overview, including a couple of mnemonics (like the "figure eight" relationship between the downtown and residential areas); she was an experienced traveler with confidence in her sense of direction. Unfortunately, any conclusion that attempts to generalize from a single case is totally without foundation, except perhaps one: surrogate travel has the potential of working extremely well in some cases. Further evidence for judging its potential based on field results is presented in the next section.

5.4 Movie Map Users in Aspen

Responses were gathered from four of the subjects in the preceding group who went to Aspen. Their trips were all made independently, shortly after their surrogate travel experience. They were all first-time visitors to Aspen. They refrained from looking at maps or reading about Aspen before and during their visits. Subjects were each assigned a destination in Aspen that they were to try to find when they first arrived. This destination varied from subject to subject and in every instance was one which the subject had visited during the surrogate travel trial. In most cases, the destination was the place where the subject would actually be staying in Aspen. In all cases the subjects had attempted to preview in surrogate travel the trip they would be making to the destination upon entering town from the airport. Upon arrival they were usually picked up at the airport by friends. Once they entered the city limits the subjects made all decisions, directing the driver to the destination without any assistance or comments from other passengers. These instructions were tape-recorded. The subjects also recorded their reactions and their attempts to anticipate elements before they came into view. In addition they filled out a questionnaire after their first day in Aspen (Appendix 3). A follow-up interview was conducted upon their return from Aspen.

Three of the four subjects gave virtually flawless directions in guiding their drivers to the destination points. They did not get lost once. In some cases the route they chose was actually shorter than the one they had previewed in surrogate travel. (This would suggest that they were not merely following a rote path which they had memorized.) The tapes revealed some uncertainty and unfamiliarity during the approach to the town which, of course, had been outside the range of the movie map. But as soon as they crossed into town they began recognizing features and expressing confidence that they knew exactly where they were. Their own characterization of this unique experience was revealing and humorous. Comments ranged from "It almost felt like coming home; I had a strong sense of déja vu" to "I feel like Robin Williams on Mork from Ork." If anything, the passengers were as impressed by the subjects' facility as the subjects themselves. Looking at individual cases in more detail will give a richer sense of the nature of the subjects' experiences and the process of their interaction.

The first subject, who will be called Ben, was a man in his mid-twenties who was going to a friend's wedding in Aspen. Ben's whole family was going, but as they had visited before they did not take part in the surrogate travel exercise at M.I.T. When Ben arrived in Aspen he entered Aspen from the west end on Highway 82. He negotiated the right/left jog with confidence and proceeded east along Main Street. As requested, he made an effort to anticipate elements of the town before they came into view. He was able to verbalize half a dozen items along Main Street and to mention another six that he recognized after they had come into view. It is not known how many features Ben may have been familiar with but did not have time to record. Ben directed the driver to turn right off Main onto Aspen Street and after two blocks to turn right again on Cooper Street. Halfway down the block he pointed out his host's condominium on the left. Ben demonstrated an ability to extrapolate visually and spatially from his movie map experience. In his surrogate travel he had never turned onto Aspen Street; he did not even know its name. He decided to turn there because he remembered, in his prior experience at M.I.T., having gone two or three blocks out of his way. He evidently had enough confidence in his internalized spatial model to construct a short cut through unknown territory. It is clear that in several instances he relied on the visual cues of landmarks, not on street names or counting blocks. He said that the recognized the short-cut street he wanted to turn onto by the traffic light and the dress shop on the corner. He also confirmed the location of his host's condominium by recognizing the red ski lodge across the street. (The condominium had been obscured by heavy shadows on the movie map. It was similarly shaded upon his arrival in Aspen.)

The most spectacular testimony to the success of Ben's movie map introduction came over the next few days. His brothers, who had each been to Aspen several times before, kept getting lost. Ben, who had never been there before, knew his way around better than anyone else and was drafted as the official guide for the group. The only time he became lost (along with the rest) was in an area which he had not visited via surrogate travel.

Ben had several general reactions which were voiced by other subjects. He was quite surprised by the magnitude of the mountains and the steepness of streets at their base. He commented, "It's much more closed in than it seemed on film. In fact those mountains seem like they're right on top of you." He and other subjects shared an uncertainty about distances, along Main Street, for example.

Sally's trip to her destination was lacking in mistakes and surprises almost to the point of being uninteresting. Her tape-recorded description was so complete and so anticipatory (". . . coming to Carl's Pharmacy, on the right is . . . , just up ahead is. . . . ") that she sounded almost like an Aspen tour guide. One special circumstance of her trip was that she tnereed from the opposite end of town, arriving by car instead of airplane. This meant not just that everything was turned around, but that she came into an unfamiliar area, which she had not visited in surrogate travel. She felt unsure of herself until she came to the corner of Mill and Main and saw the Hotel Jerome. Then she felt that she was on "home ground. And away we go." Most subjects reported feeling a sharp transition when they passed from a familiar, surrogate travel area into a region they were visiting for the first time.

Sally also reported an incident which should testify to the subtle and pervasive power of the movie map. During surrogate travel, Sally had gone north on Mill Street a couple of times and encountered a dead end, which marked the edge of the movie map. At that time in system development there was a penalty in arriving at a dead end, because the user could not back up or turn around and so was forced to start over. In aspen Sally was invited to visit friends who lived up Red Mountain. The directions given her included turning onto Mill Street and heading north. Sally said that she listened to the instructions skeptically because in the back of her mind she remembered somehow that ". . . you can't really drive down that street, can you?"

The third subject, Henry, presented a challenge. He had rated his own sense of direction as extremely poor. Using the findings of Kozlowski and Bryant [1977] that people are accurate in this kind of self-evaluation, subjects were asked to rate themselves on a scale from 1 to 10. Henry gave himself a 2, which was the lowest chosen by any of the 35 subjects. He said he has always been uneasy in new places because he gets lost and disoriented so easily. At first his rating was thought to be a case of modesty because he seemed so successful in the surrogate travel and map-drawing exercises. But at the end of the M.I.T. interview Henry said goodbye and left, only to come back five minutes later and ask if he could be shown how to get out of the building. After he had thus bolstered the credibility of his low self-rating, it became that much more revealing to see his reactions in Aspen. His experience was much like the others'. He was a little nervous and uncertain until he actually cross the bridge into Aspen. After that he found his way with little difficulty, even though he had arrived at night. He said that it took longer to distinguish landmarks and recognize intersections because of the darkness. Henry reported that the surrogate travel experience helped him maintain his bearings in both familiar and unfamiliar areas. For example, when he parked his car on an unfamiliar street he still found it easier than was normal for him (and easier than his companions) to remember its location. It would seem that Henry had a better global framework within which to insert new local details. He was proud of being able to find the car when his friends were confused, and of assuming the role of tour director for the whole group, which was in Aspen for the first time. The real benefit seemed to be not merely cognitive but affective as well. Henry said, "It was fun and exhilarating. I could go about my business without worrying (about path finding)." He had a confidence and peace of mind that allowed him to enjoy the visit without the customary distraction of constantly worrying about where he was.

The careful reader will nave noticed that all of the preceding discussion has concerned subjects whose experiences were successful. The fourth case was equally spectacular in failure. The woman in this case, Kathy, had a difficult time both with the M.I.T. movie map and in Aspen. As a native of Boston she wondered if some of her difficulty in surrogate travel was due to an unfamiliarity with western environments. She had trouble distinguishing between different mountains. She was unaccustomed to relying on the fact that four right-angle turns around a block would return one to the start. (Indeed, it seldom works that way in Boston.) One of the most intriguing problems which cropped up in surrogate travel was her reptition of mistakes. In one case she was trying to get back to her goal, which was tantalizingly close by. She would turn one street too soon, proceed to the next intersection and then proclaim that she was lost. As she repeated this pattern, starting over each time, she reinforced the initial wrong turn. She would even say, "I know I turn right here," without being able to foresee that she would again come to her "I-am-lost" intersection. (This also illustrates the notion that cognitive space for some people is divided between "lost" and "found," and that making the transformation from the former to the latter is rare. Being lost is being in a space that is not related by euclidean rules to known space.) Kathy said that her sense of direction was poor except when she was using maps to find her way.

In Aspen she had great difficulty finding the way to her destination. Her initial reaction was how different everything looked--the mountains, the red dirt, all the houses. Upon crossing the west end bridge, she negotiated the right/left jog satisfactorily onto Main Street, proceeded down Main with some uncertainty, and made a left instead of a right. "Then I drew a blank. I panicked after I made the first boo boo." She recognized landmarks here and there that she had seen before but she couldn't place them in the context of any familiar points. As she colorfully put it, "It was like a jigsaw puzzle thrown into the air."

The axiom that failure may be more instructive than success is one of the underlying themes of the LOGO project [Papert, 1978]. Its system of "turtle geometry" (sequential and ego-relative) vividly demonstrates how small local errors in geometrical relations generate global consequences. The earlier example of the fourteen-year-old boy showed this. Some closer scrutiny is also warranted here. Turning the wrong way off Main (left instead of right) was undoubtedly induced by Kathy's movie map experience. As a result of a rather embarrassing oversight in structuring subjects' introduction to Aspen, they began traveling west on Main Street--in the exact <u>opposite</u> direction in which they would enter from the airport. So in Kathy's initial explorations to find a pathy to her destination, she made a left turn at that intersection over and over again. Later instances of surrogate travel along Main in the correct direction were less numerous and without the primacy of the first pattern. Kathy's response in Aspen was consistent with a verbal encoding of the path, but not with a visual/spatial one. She remembered the building which was a cue for her to turn. She made the turn on the basis of the remembered "left," not on the basis of her visual/spatial relation with that landmark, which would in this hypothesis have induced her to turn toward the building, not away from it.

After she realized how the directions had been reversed, Kathy still had great difficulty and changed her strategy. She remembered an important landmark which had been key in locating her destination--the park. Several times in surrogate travel she had "dead-ended" facing the west edge of Wagner Park at the T intersection of Monarch Street and Cooper Street. At those times I pointed out that her destination was 1-1/2 blocks directly behind her on the same street (Cooper). So she announced that she had to find the park; if she could just find the park she would know her way from there. In the end her passengers had to coax her to the destination, but the "park strategy" obsessed her for the four days she stayed in Aspen. She was able to find a T intersection between Cooper and Wagner Park, but her destination point was nowhere to be found--not 1-1/2 blocks back, not to the side, not on a parallel street. She said, "I walked those streets till I went crazy. No one could understand me. Until the last day, just as we were leaving, it never dawned on me that there were two sides to the park." It seemed reasonable to conclude that the problem was a manifestation of the initial reversal of the entering direction on Main Street. But upon further investigation it became clear that Kathy had acquired the misconception during surrogate travel. The map she had drawn at M.I.T. showed the crucial T intersection on the east side of the park, just where she looked for it in Aspen (Figure 59). (The source of the confusion may still be the same. The ill-fated design that first had subjects begin by heading west and later by heading east may have already confused the issue during surrogate travel.)

The affective component was equally severe. Her reaction was in sharp contrast to Henry's. "I was keyed up to do the test perfectly. I thought it was going to be a snap. But I got stymied. I started floundering . . . I was disgusted with myself. I couldn't believe I couldn't find anything."

The lesson from these few examples is clear. The interactive movie map medium is potentially powerful because it is so seductive. Its pervasiveness and its persuasiveness are potentially both a virtue and liability. The responsibility incumbent in the design of the movie map medium is to insure that the viewer is not led across that fine line. The implications of this paragraph will be addressed in the next chapter.

In addition to their spoken reactions, recorded as they first drove into Aspen and in post-Aspen interviews, all subjects filled out a questionnaire in Aspen (Appendix 3). An examination of their responses will conclude this section. Question 1-3 reinforced earlier statements about the extent to which the surrogate travel trial prepared subjects for Aspen.

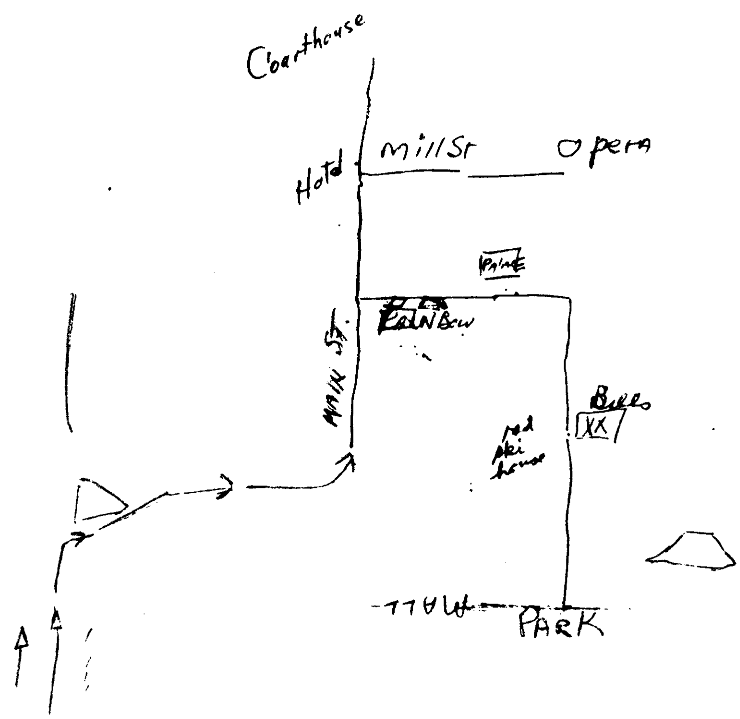


Figure 59. Park misplaced on Kathy's map

The consensus was, of course, that surrogate travel helped enormously. The dissenting opinion was that the preparation had done more damage than good. Generally subjects felt that the movie map was best at presenting specific views and spatial relations. There were easily remembered and recognized in Aspen. Subjects again expressed the sentiment that television display had presented a "tunnel vision" view.

Subjects felt that they were worst prepared for the dimensionality of things, for example, how far it was from Seventh Street to Mill Street. Because this finding was anticipated, question 4 was devoted to an expanded interrogation about many specific dimensions. If one can legitimately talk about patterns in a sample of four then the following reactions emerge as the most significant:

a) The mountains were larger and closer than subjects had anticipated.

b) Some streets were steeper than had been expected.

c) The number of buildings (residences, hotels, lodges, and condominiums) was larger than expected.

For what it is worth, the one subject who viewed the large, 8-foot-by-10-foot screen came the closest by far to having her anticipations realized. In a cumulative measure of deviation from expectation her score was 3 compared with 10, 11, and 14 for the other three.

Finally of interest were answers to the last three questions, which asked about whether and how a cognitive transformation took place while the surrogate experience was being replaced by the live, real-space experience. These questions were somewhat leading in nature and placed at the end to avoid biasing the other responses. Subjects hypothesized that their exposure to surrogate travel was equivalent to spending an hour driving around Aspen or being a visitor in town for a couple of days. In effect they had a head start on friends who arrived at the same time. With regard to the persistence of the surrogate travel experience, subjects reported that in the beginning the line of demarcation between familiar and unfamiliar spaces was quite distinct. They felt very certain of whether they had seen a particular path before during surrogate travel. One person commented that moving from an area previously visited in surrogate travel into a new one was just like the real-life experience. Subjects said that their conscious awareness of the movie map faded rapidly. One response was that by the second day or so, "the reality of Aspen became a fact," replacing the movie map version. Even so, some residual recollections from the move map cropped up much later. The hypothesis, of course, is that the cognitive spatial structures built during surrogate travel served as the foundation of further development on site, even after the origins of those structures were no longer available to active awareness.

5.5 Controlled Experiment

A subsequent experiment was conducted on the version of the movie map with aerial overviews to determine the cognitive significance of incorporating topological route representations with travel experience. Subjects were divided into two groups--those given routing diagrams overlayed on the aerial map prior to a twenty-block drive through the town (map-plus-travel) and those who were only given the street-level drive (street-travel-only). The groups followed identical routes at identical speeds. Twnety-four subjects were tested on landmark recognition, route knowledge and metric knowledge. This three-stage design was derived from the hierarchical model of Siegel and White [1975] described in chapter 4. Their developmental progress over six trials along the same route was examined in a series of tasks given after each trial.

The complete procedure for the experiement is documented in Appendix 4. The methodology consisted of presenting subjects with a set of pictures taken from the video disc. Among them were a subset of distractor images, not on the route followed but taken from an adjoining parallel route with highly similar characteristics. Three variations were included among the subset of correct pictures. Some were identical to views the subjects saw along the route, some were 90-degree rotations of views seen (i.e, a side window view), and some were 180-degree rotations (i.e., facing in the opposite direction).

The landmark recognition task required subjects to decide whether they recognized the location shown in each picture and if so, whether the view-point was the same or different than the one they saw on the route. A level of confidence rating was incorporated into their scores. The route know-ledge task tested the subjects' ability to sequence the pictures they recognized by both relative order and absolute position along the route. Metric knowledge was determined by a multidimensional scaling task that required subjects to estimate the real-space distance between each combina-tional pair of landmarks. The landmark recognition task was given after each trial, the route sequence task after trials 2 and 5, and the metric configuration after trials 3, 4, and 6.

Analysis of the results show two strong, albeit not surprising, trends. First, progressive improvement in landmark recognition over trials was clearly demonstrated (Figure). Accuracy increased from just above chance (60%) to 80% by trial five. Second, performance proved to be a function of the type of landmark picture being judged. Pictures identical in viewpoint to the ones encountered by the subjects were correctly identified with 88% accuracy over all trials. Those showing rotated views (either 90- or 180-degree transformations) were identified with 71%. Distractors were correctly rejected at only a chance level (53%). Presumably the reason for the poorest performance in the third category is that recognition requires evaluation of a single instance along the route while rejection requires an evaluation over the entire length of the route.

A less obvious result was that the map-plus-travel group exhibited some superiority in landmark recognition. After trial one they showed no advantage. But after trial five they outperformed the street-travel-only subjects on five out of the six landmarks for which the two groups differed. Further, the map-plus-travel subjects scored higher on every landmark for which analysis of variance could establish a significant difference at the .05 level. This group was also better in judging whether the test view presented was the same or different than the one encountered, especially with regard to 180-degree rotations.

When the distance estimations from the route ordering task were plotted on a metric scale, the map-plus-travel subjects showed more accuracy (averaging a remarkable 3% error or less from the true positions for eight of twelve landmarks!). However, street-travel-only subjects exhibited better sequential knowledge when the metric scale was omitted leaving just relative

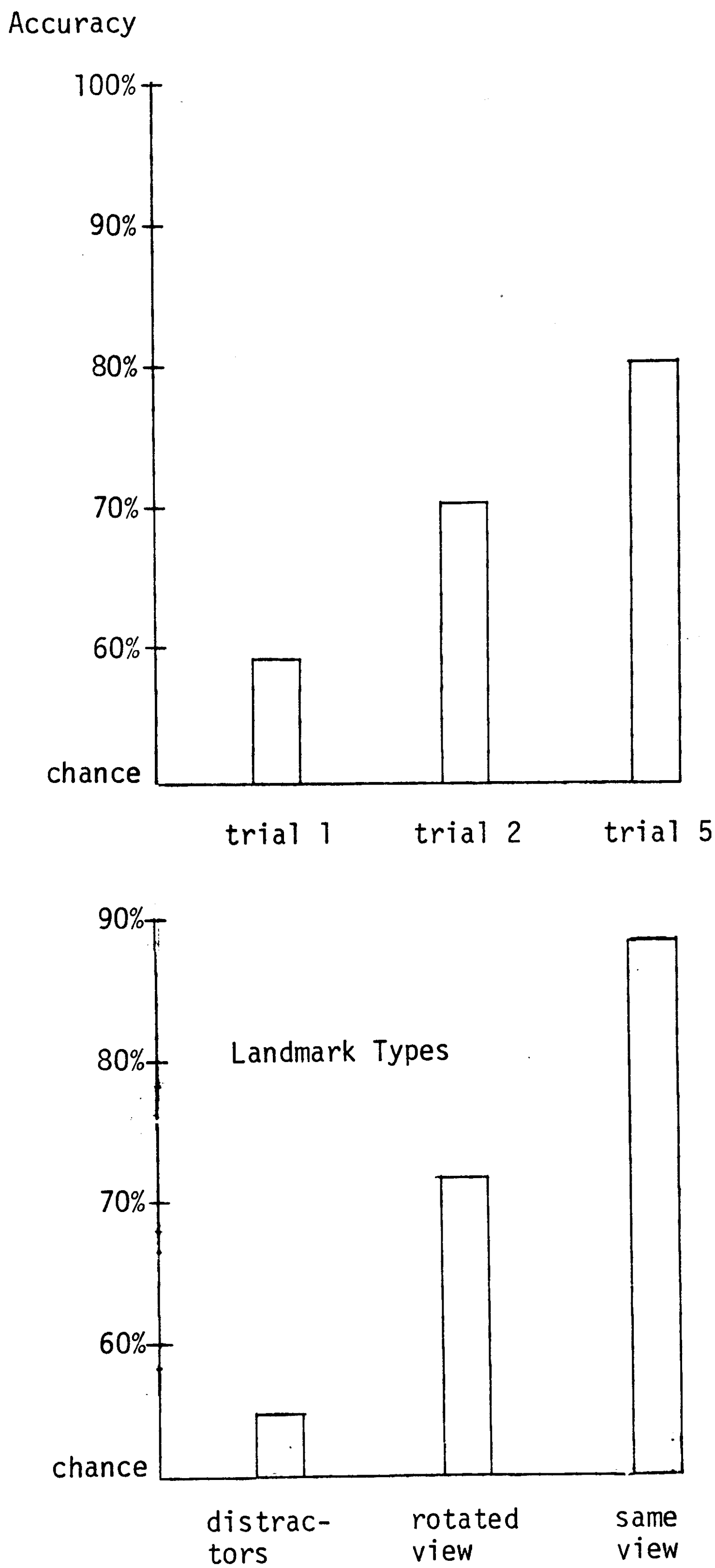


Figure 60. Landmark Recognition

position (Figure 61). The discrepancy is apparently accounted for by the fact that the latter subjects maintained accurate knowledge of order even though their relative scale of distance was distorted. Map-plus-travel subjects seemed to be able to assign the relative position of landmarks with respect to the entire length of the route but not with respect to each other.

Actual	Street Travel Only	Map Plus Travel
1	1	1
2	2	2
3	3	4
4	4	3
5	5	7
6	9	5
7	6	6
8	7	10
9	8	8
10	11	9
11	10	12
12	12	11

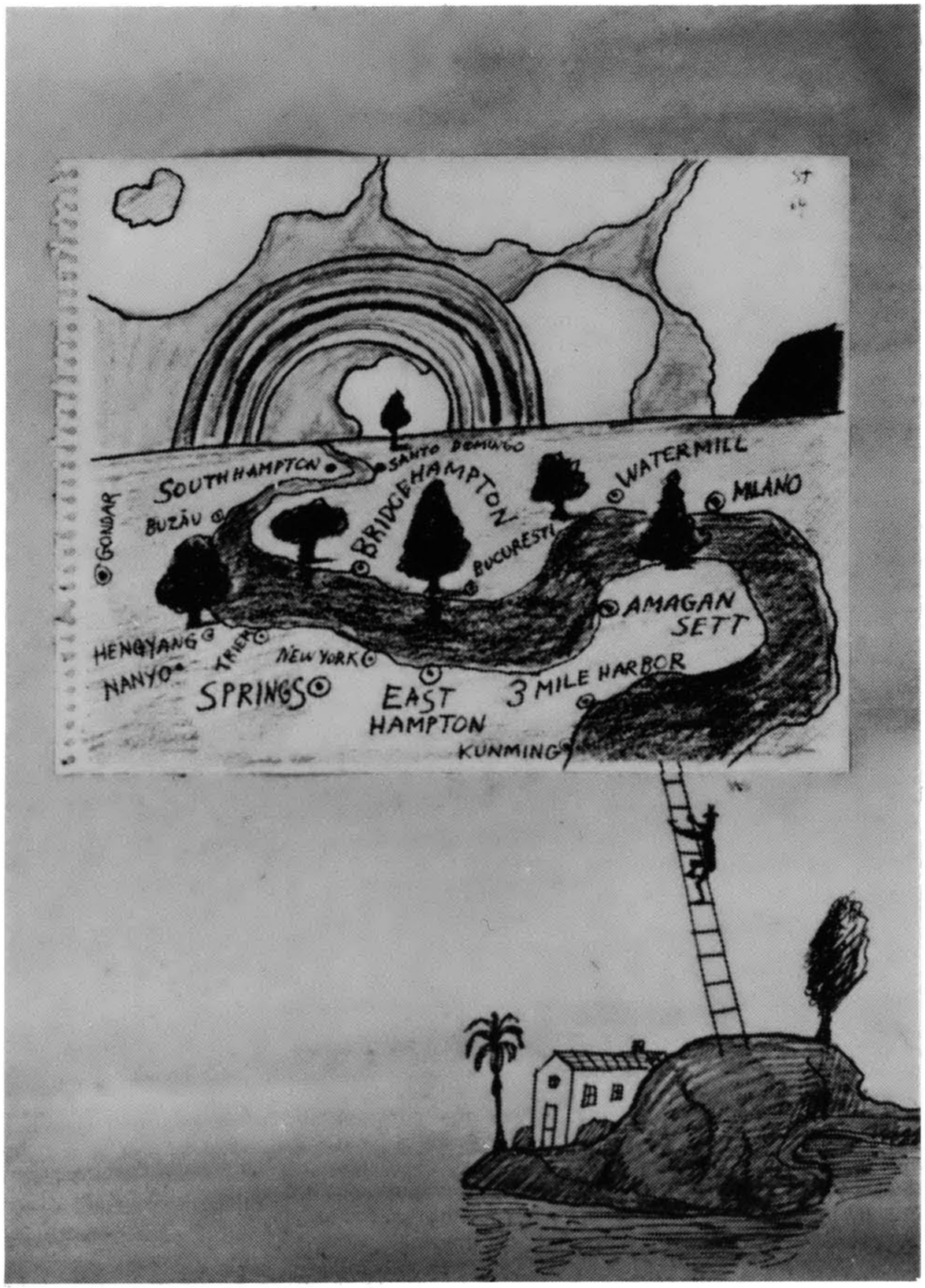
Figure 61
Sequencing Landmarks Along the Route

Preliminary analysis of multi-dimensional scaling of estimation of absolute distance estimations between landmarks (i.e., not necessarily along the route) showed, not surprisingly, that the map-plus-travel subjects performed more accurately. Complete results will be published separately.

Interpretation of these results began to cast doubt on whether the Siegel-White model of hierarchical development could be strictly applied to the movie map. The fact that topological routing diagrams could aid subjects in a landmark recognition task which has no direct spatial component suggested that the development did not proceed linearly from landmark to route to configurational knowledge. Instead it appeared that the relationship was circular. Each could mediate the others; in particular, configural knowledge could affect landmark recognition. The mechanism may be that configural knowledge provided an organizing scheme for internalizing non-configural features. While further research is required (and recommended), the experiment concluded thatintegration of multiple representations of space was not only advantageous in spatially dependent tasks but enhanced performance on non-spatial tasks as well.

CHAPTER 6

Summary, Conclusions, and the Future



- Steinberg

It has been clearly demonstrated that the interactive movie map is a viable concept which can be practically implemented. It is effective an an external spatial representation of an environment and as a process by which that model can be internalized as cognitive space. The nature of the learning which takes place with this interactive system has far-reaching implications for education in general.

6.1 Configuration of the Existing System

A number of techniques were discovered which have become fundamental to the concept of the interactive movie map. The two most general were:

a) Maintaining invariance of visual elements. The technique of keeping everything in the scene constant except for the property of interest was exploited in several different contexts. In the case of turn sequences, it accounted for the effectiveness of leading turns and tracking pivot turns by keeping a fixed point of reference in view. A similar example was evidenced in the transition from front to side view. In the case of showing scenes in different seasons and time periods, the invariance etched the constancy with clarity, identifying the elements and processes of change explicitly. In the aerial overview the invariance of between street map, landmark map, and photograph revealed the underlying topology upon which different thematic representations could be articulated. It was exploited to help the user keep his orientation during zooms by keeping the target intersection in a fixed location on the screen.

b) The two-disc algorithm. The technique of alternating between two discs was necessary to allow all editing to be done without visual discontinuity. The additional benefit achieved was that the invisible disc could be staged ahead to minimize, even eliminate, the search-time dealy which would otherwise have occurred at edit points. Optimizing disc layout to take the most advantage of this technique led to a format in which material was arranged to provide alternate options when played forward or backward. The result was that up to four different branching choices could be immediately available (without searching). The technique was available by default in forward travel (which could switch instantly to reverse). A more elegant example was the arrangement of related left and right turns starting on the same frame--one sequenced forward on the disc, the other backward. The aerial overview used the same technique to enable the user to switch between three modes of representation as well as zoom out. Cultural data access provided three options: viewing seasonal data or a slide show, or returning to surrogate travel.

6.2 Future Enchancements

The interactive movie map has evolved through a number of versions. Its refinement has stopped as a result of having achieved its goals within the scope of the original project. But that does not mean that speculation about future enhanced configurations of the system has to stop. A number of ideas have remained unrealized or only partially realized during the course of this project's development simply because of the allocation of resources. (An additional set of suggested improvements have arisen out of the experiments and will be discussed in the next section.)

The display configuration of the user's station was generally limited to a single monitor. One enhancement contemplated is to add a second monitor which would always give the user access to the aerial overview (Figure 62). Using computer-generated graphic overlays on both screens requires a second frame buffer. It is speculated that allowing the switch between the dual modes of representation with the ease and speed of eye movement instead of disc searching would improve effectiveness. In such a case a decision must be made concerning the positioning of the monitors. For reasons of symmetry and reinforcement of patial correspondence, the best solution may be to orient the overview monitor horizontally rather than side-by-side in relation to the travel monitor (Figure 63). Other examples of double monitors use are shown in Figures 64 and 65.

This change in hardware lends itself to a richer interaction which the user could employ in explicitly fashioning the nature of his cognitive map. Personalized landmarks could be added as icons on the aerial overview. They could be added automatically as a function of importance or as a function of points of interest revealed by the user in the process of cultural data access. More interestingly, the user could specify those landmarks which he finds to be the most effective as orientation cues. He would make a quick "rotoscoped" sketch of those features pictured on the travel screen which are the most important subjective identifiers (Figure 66). A small

Figure 62. Simultaneous access to dual representations

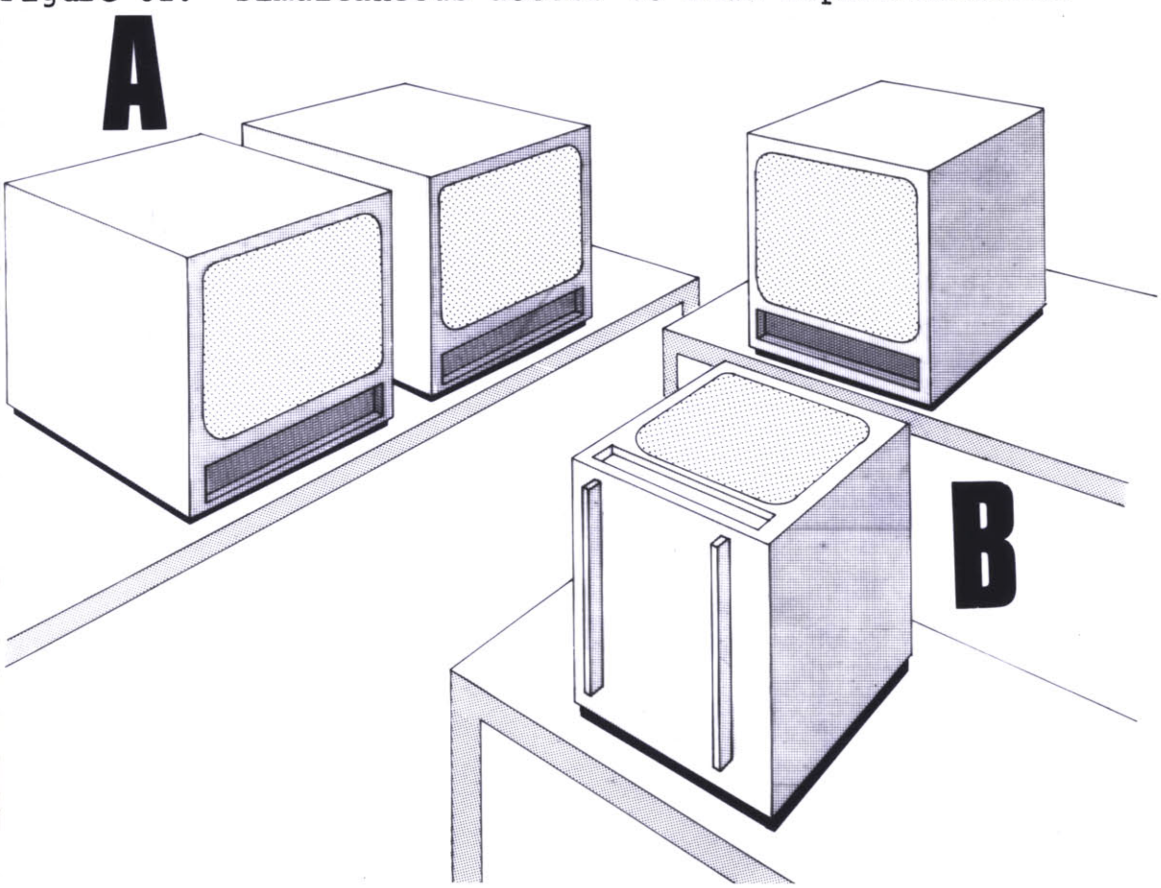


Figure 63. Possible monitor configurations

Figure 64. Ground view and aerial view

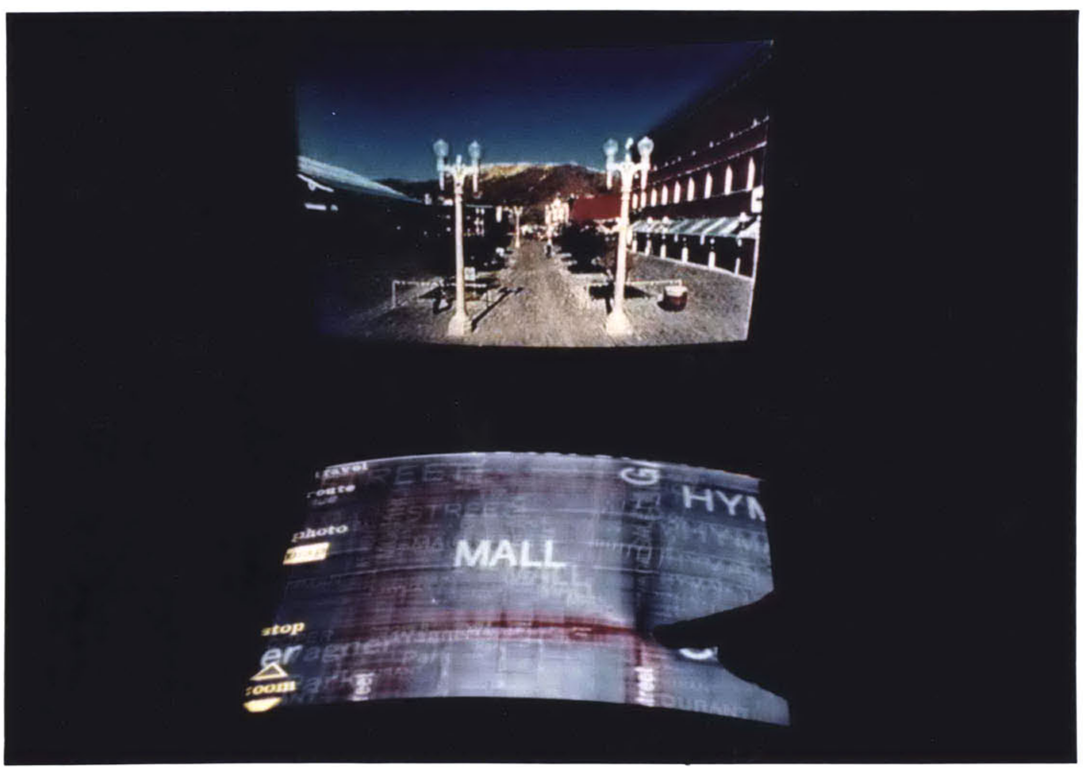


Figure 65. Map access during travel

icon of this sketch would then be inserted at the appropriate place in the overview (Figure 65).

Another variation tested in some temporary mock-ups included large-screen display and multiscreen formats. Preliminary evaluation confirmed the effectiveness of both enhancements, primarily because of the increased spatial correspondence which was achieved. The effect was first demonstrated showing the movie map in a media room with a 13-foot-diagonal, rear-projection television screen (Figure 68). Various combinations of these two elements are depicted as hypothetical configurations in Figure 69 (front and side views) and Figure 70 (travel land and map land). The full realization of both elements is shown in Figure 71.

The 360-degree panoramas were hardly exploited in existing implementations of the movie map. There are two easily envisioned display systems which could utilize these pictures. A prototype version of one solution which was briefly tested consisted of a conical mirror to rectify the anamorphic image (Figures 72 and 73). The anamorphic transformation using the mirror is shown in Figure 74. With the monitor lying horizontally on a "lazy susan," the viewer can rotate the point of view continuously, similar to the way in which he might move his head in the real space (Figure 75). With an eye-level viewing station the viewer could walk around the display (Figure 76). The major difference is that in the simulated case he is "on the outside looking in," exactly opposite from the real situation. A solution to this problem is to put the user inside the display surface. Figure 77 is a rendition of a cylindrical media room in which the same anamorphic lens used in taking the pictures is suspended over the viewer's head for 360-degree projection.

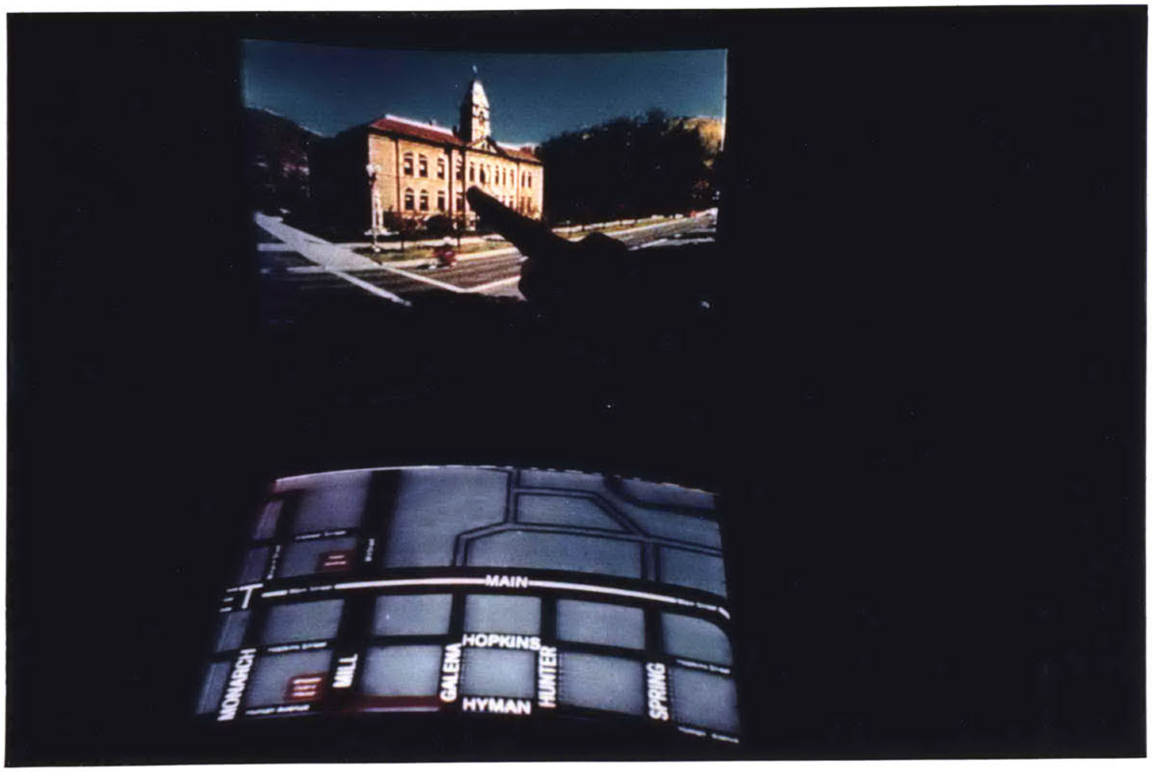


Figure 66. Selecting a personal landmark

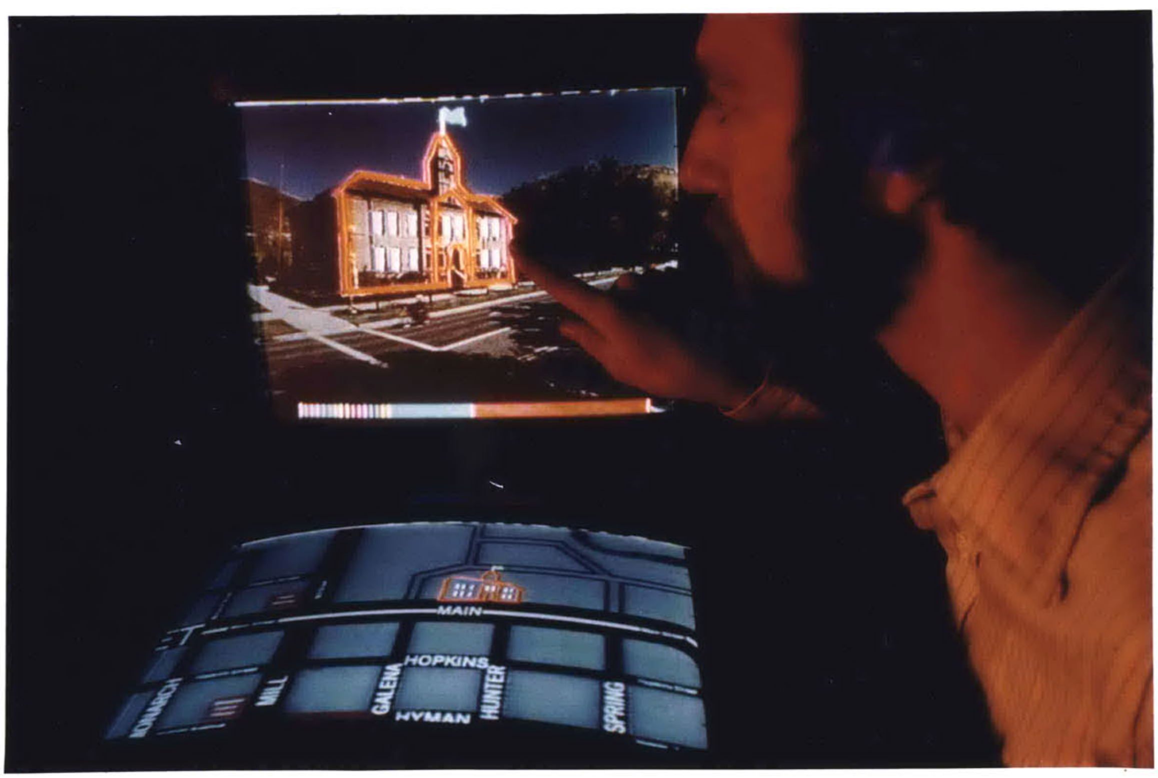


Figure 67. Symbol hand-drawn onto map

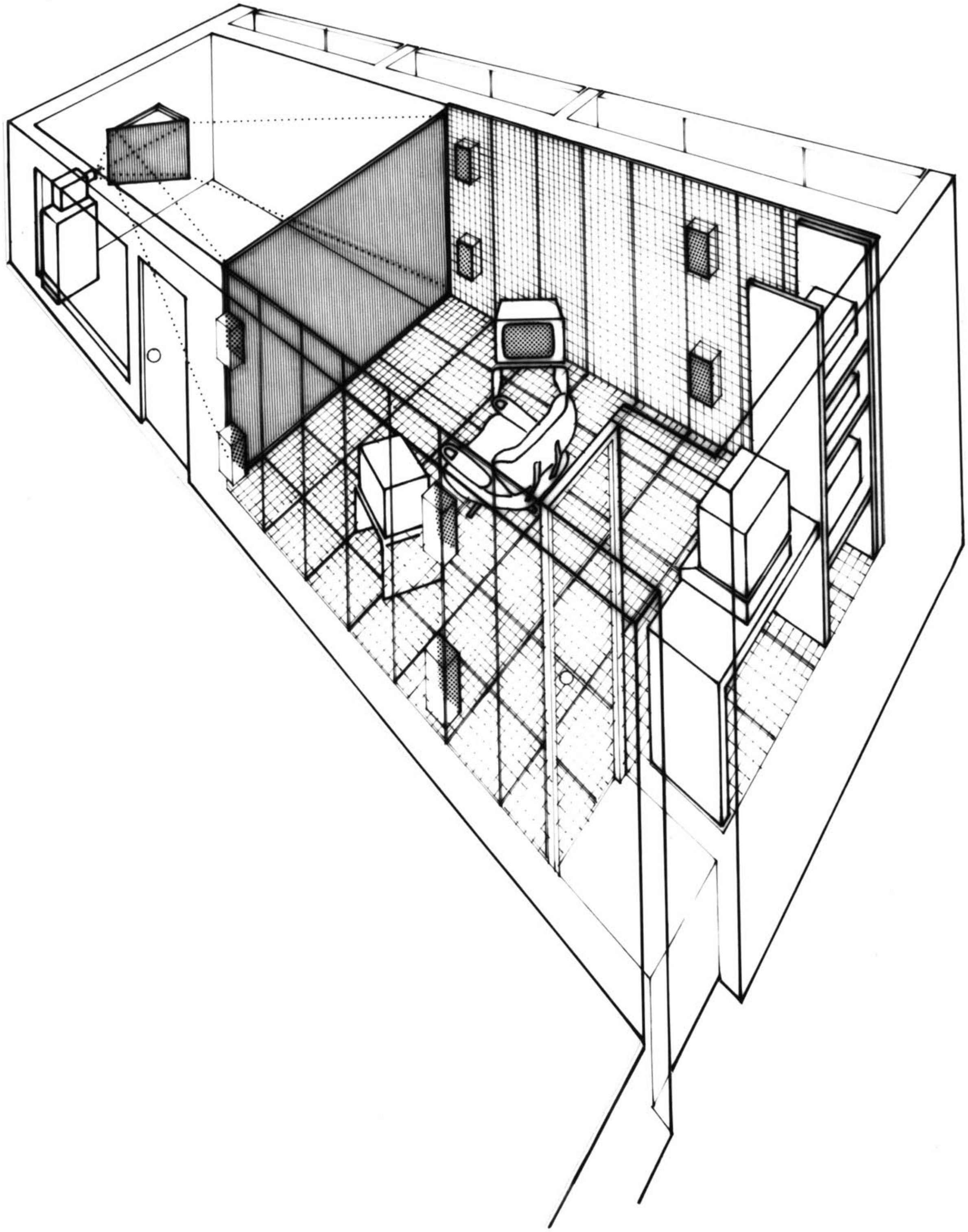


Figure 68. Existing media room

Figure 69. Multiscreen front and side views

Figure 70. Multiscreen travel land and map land

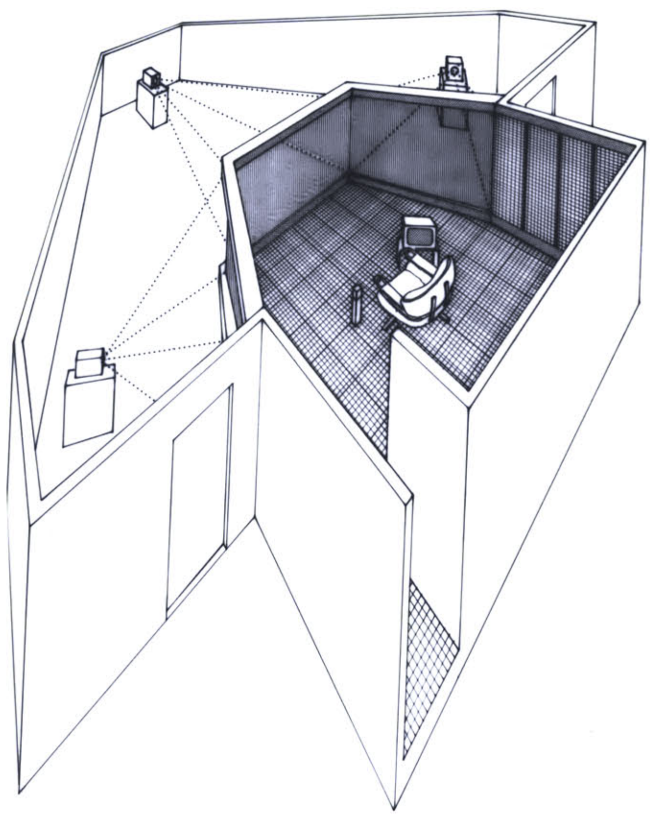
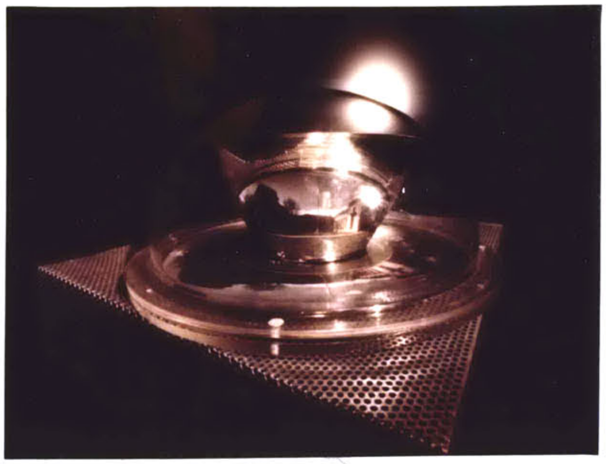
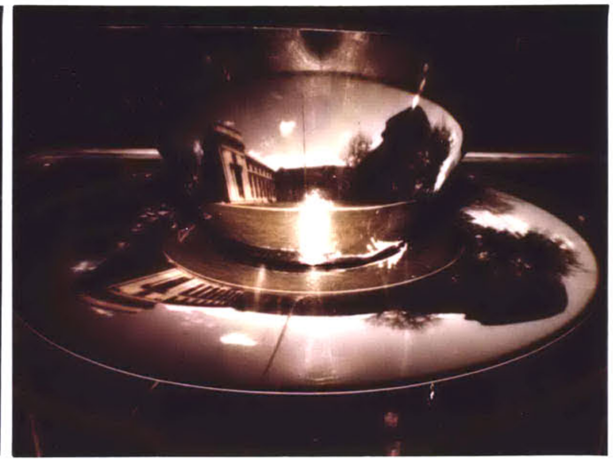


Figure 71. Large multiscreen media room, artist's conception





Figures 72, 73. Viewing dome for anamorphic photograph

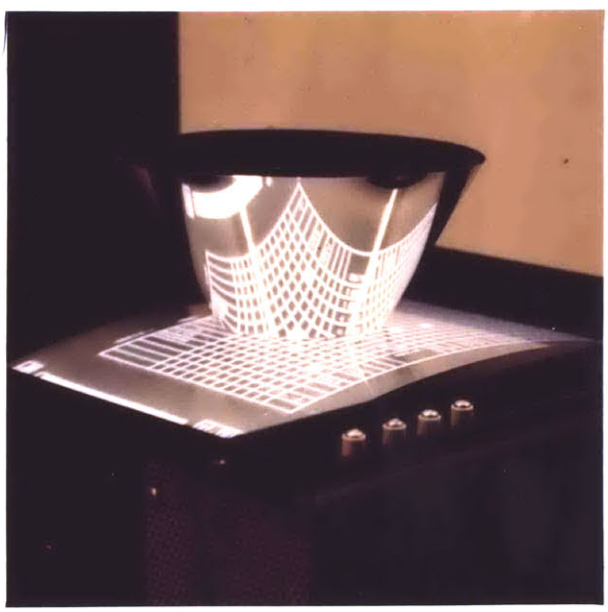


Figure 74. Warping by conical mirror

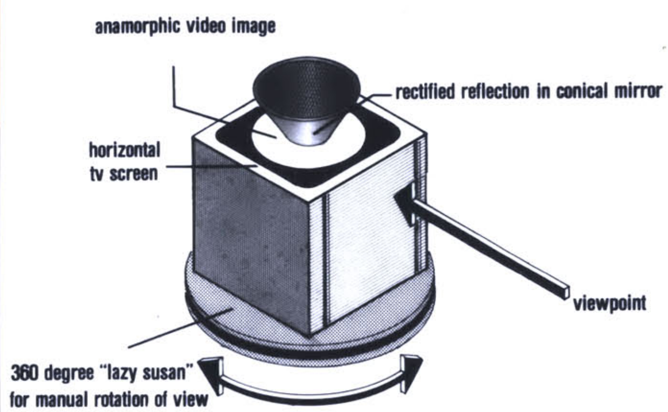


Figure 75. User's station for anamorphic images

Figure 76. Eye-level viewing station for anamorphic images

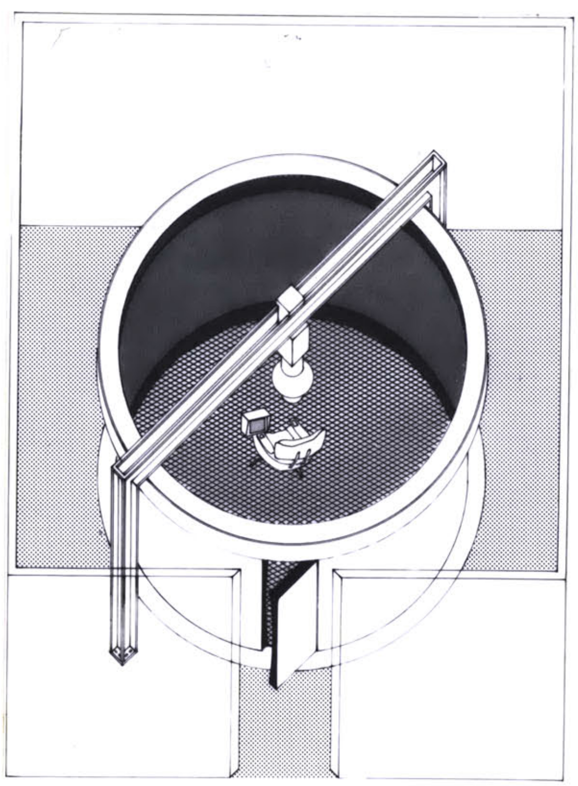


Figure 77. Cylindrical media room with 360-degree front projection, artist's conception

6.3 Experimental Results

One of the significant findings was the importance that spatial context played in the viewer's experience. It was manifested in three areas: global overview, peripheral vision, and spatial correspondence in head and eye movement. Preliminary evidence shows that a global overview offers more than navigational aids to help the viewer keep track of where he is. It actually affects the way in which his experience of the space is internalized--even to the extent of improving performance in nonspatial perception, like landmark recognition. A full summary of these results will be published separately.

With regard to the perceptual aspects of spatial context the following recommendations for a display space are made:

a) The playback space should incorporate a large field of view.

b) The peripheral field should be realized in three-dimensional space (or, more precisely, the two-and-one-half-dimensional space of curved surfaces).

c) Spatial correspondence should exist both in the displayed image and in the user's head and eye movements in viewing it.

The effectiveness of the interactive movie map as a medium for engaging the naive user in a convincing experience of surrogate travel has been demonstrated. Users find it easy and natural to drive around in an unfamiliar environment making real-time path-choosing decisions. The consequence of real significance is that the experience allows them to construct their own internal representations or cognitive maps of the space. Moreover, these cognitive maps are in many ways similar to those developed by travelers learning in the real space.

There is, of course, no question that the movie map and the subsequent cognitive maps ultimately fall short of the real-world experience. No one is fooled into thinking the video display is a window onto a live reality. The scale of the space is not always interpreted accurately (heights of mountains and linear distances, for example). It is anticipated that some of these shortcomings will be overcome in future versions of the system--by refining the user control interface, by increasing smoothness and camera stability, and by expanding the field of view and spatial correspondence. Still, one would expect that no matter how convincing the enhancements become, the gap to reality will remain obvious.

On the other hand, there is no reason to assume that the acquisition of spatial knowledge during the movie map experience will necessarily remain less successful than learning in the field. For example, the integration of ground level and aerial views, not possible in the real world, has demonstrated potential advantages. The dual representation seems to promote a synergistic cognitive mechanism for combining spatial and non-spatial features of an environment.

6.4 Educational Implications

The last section of this document takes the opportunity to reach beyond the immediate concerns discussed thus far. (In doing this, the author acknowledges some partiality.) A host of applications to interactive video-disc-aided instruction are on the horizon. Computer-aided instruction (CAI) has long suffered from the tedious medium in which it is presented, which is typically text-intensive. It is no wonder, therefore, that the advent of the optical video disc is being seen as a potential panacea for some of the severest shortcomings. As a visually rich and relatively inexpensive medium it is apparently a natural complement to the computer-aided format. Unfortunately (in the opinon of this author), one of the likely prospects is that the video disc will lead to the entrenchment of some bankrupt approaches.

The fallacies of traditional CAI include thedilemma that authors must establish an epistemiology of the subject matter domain as well as a generalized model of cognitive development to justify the prescribed learning sequences they adopt in programmed instruction [Schwartz and Olds, 1980]. And yet, these are both unresolved issues within the psychological and philosophical communities. The author sets forth a plan of tutorial instruction which includes a number of alternative branching paths depending on the responses of the student to predefined questions. But what if the student's difficulty hasn't been anticipated by the author and therefore does not fit into a ready-made diagnosis? What if his cognitive style is not consistent with any of the limited number of paths available?

With a small amount of extrapolation, the interactive movie map could turn CAI upside down. The fundamental tenet of the system is to put the user in complete control. He decides what he wants to know and how to go about learning it. He experiences continuous interaction, in contrast to the alternative, which is akin to a ping pong match [Negropointe, 1980]. It's either the student's turn or the computer's turn. One is always waiting for the other (and by implication not necessarily maintaining attention). True interactivity requires either agent to be able to intervene at any time deemed opportune. This is the kind of interaction achieved with the movie map. The user is the driver, not the passenger; program control is self-directed, not prescribed. In CAI the student is always called upon to supply the answer. In the movie map the user is never required to give answers, but instead is encouraged to ask questions. This manner of learning has achieved some success; it accounts for vast amounts of knowledge about the world which children acquire before they ever go to school.

How is such an alternative approach possible? Is the movie map really doing anything different and can it be generalized? The answer lies in the nature of the subject matter which is modeled in the system. In the case of spatial representation, the domain is described by a very simple but complete network of possible paths. The architects of the system did not have to anticipate which particular paths might be useful to the user. All possibilities were included (with the explicit and natural constraint: the user had to stay within the street system). Any subject matter with an inherent network structure is a likely candidate. In such cases, the problem of understanding the epistemiology of the knowledge domain is avoided. The system merely models the real world. The problem of incorporating a particular theory of cognitive development is also largely solved. Because of the complete and continuous interaction offered, the user himself adopts the individual cognitive style that best suits him.

The application of these ideas to other domains can be illustrated through a succession of minor modifications of the original system. For example, a movie map of seventeenth-century Florence (constructed partly through computer animation) would recreate a world and a culture with a richness and immediacy which is available only to a few scholars. Removing the constraint of geography, one can imagine a movie map of the human body. It might model the circulatory system, allowing a "fantastic voyage" through the veins and arteries to examine biological landmarks: heart, liver, lungs. Moving beyond the realm of physical networks, one might envision a movie map of a darwinian tree of evolution to explore, perhaps test, various theories of genetic development.

The particular examples are not important. Presumably hundreds of better ones could be suggested. The point is that they all belong to a pedagogy in which the aim is to model a specific world as an easily accessible environment in which the student conducts his own research.

BIBLIOGRAPHY

Abelson, H. and DiSessa A. 1981. Turtle geometry. Cambridge, MA: M.I.T. Press.

Abler, R. 1971. Spatial organization--the geographer's view of the world. Englewood Cliffs, NJ: Prentice-Hall.

Acredolo, L.P. 1977. Developmental changes in the ability to coordinate perspectives of a large-scale space. Developmental Psychology. 13:1-8.

Allen, G.L., Siegel, A.W., and Rosinski, R.R. 1978. The role of perceptual content in structuring spatial knowledge. J. of Exper. Psych. 4:617-630.

Appleyard, D. 1969. Why buildings are known. Envir. and Behavior. 1:131-156.

Appleyard, D., Craik, K.H., Klapp, M., and Kreimer, A. 1973. The Berkeley Environmental Simulation Laboratory: its use in environmental impact assessment. Institute of Urban and Regional Development. Berkeley: University of California, Berkeley. Working paper 206, February, 1973.

Baird, J.C. 1979. Studies of the cognitive representation of spatial relations: I. Overview. J. of Exper. Psych., General. 108:90-91.

Baird, J.C., Merrill, A.A., and Tannenbaum, J. 1979. Studies of the cognitive representation of spatial relations: II. A general environment. J. of Exper. Psych., General. 108:92-98.

Baird, J.C. and Merrill, A.A. 1979. Studies of the cognitive representation of spatial relations: III. A hypothetical environment. J. of Exper. Psych., General. 108:99-106.

Battro, A.M., Ellis, E.J., Flament, C., and Matalon, B. 1976. Urban space: experiments on the cognition of the environment. Estudos Cognitivos. 1:-49-70.

Bender, W. 1980. Computer animation via optical video disc. M.S. thesis, M.I.T.

Blaut, J.M. and Stea, D. 1968. Place learning. Place Perception Research Report Number 4, Clark University.

Boulding, K.E. 1956. The image. Ann Arbor: University of Michigan Press.

Byrne, R.W. 1979. Memory for urban geography. Quart. J. of Exper. Psych. 31:147-154.

Calhoun, J.B. 1970. Space and scale: insights from animal studies relating to the architecture and content of housing for the aged. In Pastalan and Carson (Eds.), Spatial behavior in older people. Ann Arbor: Wayne State University Press.

Chapman, G.P. 1977. Human and environmental systems--a geographer's appraisal. London: Academic Press.

Ciccone, D.S., Landee, B., and Weltman, G. 1978. Use of computer-generated movie maps to improve tactical map performance. Woodland Hills, CA: Perceptronics, Technical Report PTR-1033s-78-4.

Clarke, A. 1953. The city and the stars. New York: Signet.

Cohen, R. 1980. The representation of landmarks and routes. Child Devel. 51:1065-1071.

Collings, A. 1976. The effectiveness of an interactive map display in tutoring geography. Cambridge, MA: BBN Report 3346, August 1976.

Devich, F. and Weinhaus, F. Image transformation study. Sunnyvale, CA: ESL Inc., Semi-Annual Technical Report PR339, May 1980.

Downs, R. and Stea, D. 1977. Maps in mind. New York: Harper and Row.

Gladwin, T. 1970. East is a big bird. Cambridge, MA: Harvard University Press.

Gould, P. and White, R. 1974. Mental maps. New York: Penguin Books.

Harrison, R. 1977. Eccentric spaces. New York: Knopf.

Held, R., Dichsans, J., and Bauer, J. 1975. Characteristics of moving visual scenes influencing spatial orientation. Vision Res. Vol. 15. New York: Pergamon Press.

Herman, J. and Siegel, A. 1978. The development of cognitive mapping of the large-scale environment. J. of Exper. Child Psych. 26:389-406.

Hochberg, J. and Brooks, V. 1978. The perception of motion pictures. Handbook of perception, Vol. X, pp. 259-303. New York: Academic Press.

Hooper, K. 1979. The mapping process. In Fields, C. and Weltman, G., Maps: a guide to innovative design.

Hooper, K. 1980. Mapping summary. Santa Cruz, CA: Final Technical Report, Environmental Cognition Center, University of California at Santa Cruz.

Hooper, K. 1980. The use of computer controlled video discs in the study of spatial learning. Presentation at On-Line Computers in Psychology conference, St. Louis, November 1980.

Hooper, K. in press. Experiential mapping--the perceptual representation of environments. Defense Advanced Research Projects Agency, Final Technical Report, July 1977-November 1980.

Ittelson, W. . Perception of the large-scale environment. Trans. of the N.Y. Acad. Sci., 32: 807-815.

Kirasic, K.C., Allen, G.L., and Siegel, A.W. 1980. Methodological and conceptual issues in the externalization of cognitive maps. Presentation at meetings of Midwestern Psychological Assoc., St. Louis, May, 1980.

Kosslyn, S.M. and Pomerantz, J.R. 1979. Imagery, propositions, and the form of internal representations. Cog. Psych. :52-76.

Kozlowski, L.T. and Bryant, K.J. 1977. Sense of direction, spatial orientation, and cognitive maps. J. of Exper. Psych: Human Perception and Performance. 3:590-598.

Kruskal, J.B. and Wish, M. 1978. Multidimensional scaling. London: Sage.

Kuipers, B. 1977. Representing knowledge of large-scale space. Cambridge, MA: M.I.T. Artificial Intelligence Laboratory, Technical Report 418, July 1977.

Kuipers, B. 1978. Modeling spatial knowledge. Cog. Science. 2:129-153.

Kuipers, B. 1979. Refuting the 'map in the head' theory. Medford, MA: Tufts University Cognitive Science Working Papers 2, April 1979.

Lippman, A. 1980. Movie-maps: an application of the optical videodisc to computer graphics. Proc. SIGGRAPH '80, Seattle.

Lippman, A. 1981. Conversations, M.I.T.

Lynch, K. 1960. The image of the city. Cambridge, MA: M.I.T. Press.

Lynch, K. 1972. What time is this place? Cambridge, MA: M.I.T. Press.

Meyer, J. Map skills instruction and the child's developing cognitive abilities. J. of Geography. 72:27-35. September 1973.

Milgram, S. . A psychological map of New York City. Amer. Scientist. 60:194-200.

Miller, O.M. and Voskuil, R.J. 1964. Thematic map generalization. Geog. Rev. 54:13-19.

Miller, R. 1980. Notes on Aspen. Unpublished manuscript.

Mohl, R. 1980. The interactive movie map: surrogate travel with the aid of dynamic aerial overviews. Applications of interactive television and the video disc. Proc. Midcon.

Mohl, R. 1971. The experience of photographic sequences. Cambridge, MA: Unpublished undergraduate thesis, M.I.T.

Naimark, M. 1979. Spatial correspondence--a study in environmental media. Cambridge, MA: M.I.T. Master's Thesis.

Negroponte, N. 1970. The architecture machine. Cambridge, MA: M.I.T. Press.

Negroponte, N. 1975. Soft architecture machines. Cambridge, MA: M.I.T. Press.

Negroponte, N. 1979. Impact of optical video discs on film-making. Cambridge, MA: M.I.T. Architecture Machine Group, unpublished paper.

Negroponte, N. 1980. Various lectures.

Papert, S. 1978. Interim report of the LOGO project in the Brookline public schools. Cambridge, MA: M.I.T. A.I. Memo No. 484.

Parsons, J.R. and Henrikson, P.J. Interactive spatial warp and image registration. El Segundo, CA: Aerospace Corp., image processing applications note.

Philbrick, A.K. 1957. Principles of areal functional organization in human geography. Econ. Geography. 33:299-336.

Piaget, J. and Inhelder, B. 1967. The child's conception of space. New York: Norton.

Piaget, J., Inhelder, B., and Szeminska, A. 1960. The child's conception of geometry. New York: Basic Books.

Pylyshyn, Z.W. 1973. What the mind's eye tells the mind's brain: a critique of mental imagery. Psychological Bull. 80:1-24.

Robinson, A. and Petchenik, B. 1976. The nature of maps. Chicago: University of Chicago Press.

Schon, D. and Bamberger, J. 1978. The figural formal transaction. Cambridge, MA: M.I.T. Division for Study and Research in Education, unpublished manuscript.

Schwartz, J. and Olds, H. 1980. People and computers: who teaches whom.

Shemyakin, F.N. 1962. Orientation in space. In B.G. Anan'yev et al. (Eds.), Psychological science in the USSR, Vol. 1, Part 1 (Rep. No. 11466). Washington, D.C.: U.S. Office of Technical Reports. 186-255.

Siegel, A.W. 1979. The externalization of cognitive maps by children and adults: in search of ways to ask better questions. Presented at Conference on Spatial Representation and Behavior Across the Life Span, University Park, PA, May 3-5, 1979.

Siegel, A.W. 1980. Methodological and metatheoretical issues in the development of cognitive mapping of large-scale environments. In Lutkus, A.D. and Wall, D. (Eds.), Mind, child and architecture. New Jersey: Lawrence Erlbaum Assoc.

Siegel, A.W. 1981. Conversations, Harvard University.

Siegel, A.W., Herman, J.F., Allen, G.L., and Kirasic, K.C. 1979. The development of cognitive maps of large- and small-scale spaces. Child Devel. 50:582-585.

Siegel, A.W. and White, S.H. 1975. The development of spatial representations of large-scale environments. In Reese, H.W. (Ed.), Advances in child development and

Sims, W. 1974. Iconic simulations: an evaluation of their effectiveness as techniques for simulating environmental experience along cognitive, affective, and behavioral dimensions.

Thorndyke, P. and Stasz, C. 1979. Individual differences in knowledge acquisition from maps.

Tolman, E.C. 1948. Cognitive maps in rats and men. Psych. Review. 55: 189-208.

Trowbridge, C.C. 1913. Fundamental methods of orientation and imagery maps. Science. 38:888-897.

Watson, W.C. 1970. A study of the generalization of a small-scale map series. Internat. Year of Cartography. 24-32.

Wurman, R.S. 1972. Aspen visible: International Design Conference on Aspen. Cambridge, MA: M.I.T. Press.

Yelick, S. 1980. Anamorphic image processing. Cambridge, MA: B.S. thesis, M.I.T.

APPENDIX 1

Mapping Questionnaire for Aspen Residents

MAPPING QUESTIONAIRE FOR ASPEN RESIDENTS

1. I would like you to make a quick map of Aspen. Make it just as if you were making a rapid description of the town to a newly arrived friend covering all the main features. I don't expect an accurate drawing - just a rough sketch.
2. What first comes to your mind, what symbolizes the word "Aspen" for you? How would you briefly describe Aspen in a physical sense?

3a. Pick two places on your map that you are familiar travelling between. Picture yourself actually making that trip, and describe the things you would see, hear, or smell along the way, including the pathmarkers that have become important to you, and the clues that a stranger would need to make the same decisions that you have to make. I am interested in the physical pictures of things. It's not important if you can't remember the names of streets and places. (Probe where needed for detail).

b. Do you have any particular emotional feelings about various parts of your trip? How long would it take you? Are there parts of the trip where you feel uncertain of your location?

4. Now, I would like to know what elements of Aspen you think are the most distinctive. They may be large or small, but mention those that for you are the easiest to identify and remember.

5a. Would you describe _______ ? (from 4) If you were taken there blindfolded, when the blindfold was taken off what clues would you use to positively identify where you were?

b. Are there any emotional feelings that you have with regard to _________?

c. Would you show me on your map where _______ is? (and, if appropriate) Where are its boundaries?

6. Would you please show me on your map the direction of North?

7. How hard is it to learn your way around Aspen?

8. Thinking back to moving to Aspen, can you remember the processes by which you got to know the town?

9. On a scale of 1 to 10 (poor to excellent) how would you generally rate your own sense of direction?

10. Were you ever confused about the similarity between two places and confuse one with the other?

APPENDIX 2

Mapping Questionnaire for New Visitors to Aspen

MAPPING QUESTIONAIRE FOR NEW VISITORS

If you would be willing to answer some questions about how you learn to get around in Aspen (for a research project in mapping), please fill in the information below.

Name ____________________

Address in Aspen ______________________ Room no. _____

Phone 925-______

Have you ever visited Aspen before? ____
If so, when and for how long? ________________________

Date of arrival ______________

Date of departure ______________

I will contact you in about three days to ask you about your experience in getting to know Aspen. In the meantime please avoid the use of maps to find your way!

Thank you for your cooperation in this project.
I think you will find it fun.

Robert Mohl 925-3320

APPENDIX L

MAPPING QUESTIONAIRE FOR NEW VISITORS TO ASPEN

1. What first comes to your mind, what symbolizes the word "Aspen" for you? How would you briefly describe Aspen in a physical sense?
2. We would like you to make a quick map of Aspen. Make it just as if you were making a rapid description of the town to a newly arrived friend, covering all the features you have come to know. (for locals - all the main features) We don't expect an accurate drawing - just a rough sketch.

3.a Pick two places on your map that you have travelled between more than once. Picture yourself actually making that trip, and describe the sequence of things you would see, hear, or smell along the way, including the path-markers that have become important to you, and the clues that a stranger would need to make the same decisions that you have to make. We are interested in the physical pictures of things. It's not important if you can't remember the names of streets and places. (Probe where needed for detail).

b. Do you have any particular emotional feelings about various parts of your trip? How long would it take you? Are there parts of the trip where you feel uncertain of your location?

4. Now, we would like to know what elements of Aspen you think are most distinctive. They may be large or small, but tell us those that for you are the easiest to identify and remember.

5.a Would yo describe ________ to me? If you were taken there blindfolded, when the blindfold was taken off what clues would you use to positively identify where you were?

b Are there any emotional feelings that you have with regard to ______?

c Would you show me on your map where _____ is? (and, if appropriate:) Where are the boundaries of it?

6. Would you please show me on your map the direction of North? the airport? Ajax the ski mountain, your home town?

7. How hard is it to learn your way around Aspen?

8. Thinking back on the last few days can you remember the processes by which you got to know Aspen.?

9. On a scale 1 to 10 poor to excellent how would you generally rate your own sense of direction?

10. Were you ever confused about the similarity between two places and confuse one with the other?

11. Since arriving in Aspen how much time have you spent
    looking at maps?
    walking around?
    driving around?
    other exposure?

12. Did you have any preconceptions about Aspen? Have they changes?

3.c Sketch the path you would take coming from the airport going through Aspen and out the other side of town.

APPENDIX 3

Questionnaire for Movie Map Visitors in Aspen

Please answer the questionaire as much as you can without interfering with your activities with friends and relatives in Aspen. Whatever you can do will help me a lot.

When you begin the ride from the airport into town, please inform the driver that you will give all the directions to get to . Ask the others not to give you any hints or make comments when they think you are not taking the most direct route. If possible please tape record the adventure. As much as possible try and anticipate what you will see up ahead and record your comments. (ex. "A little ways past the bridge is ..." or "I think down the next side street on the right is a school." etc.)

If you have time and want to take a second mystery trip, please give the small envelope to a friend who can drive you to your starting point. (This is like a treasure hunt isn't it?) You must keep your eyes closed or blindfolded until you have been driven to the starting point. Now try and guide the driver Again record your comments if possible.

If you become hopelessly lost during either trip, you may ask for help or even give up.

Please turn to the questionaire after your first day in Aspen.

Thanks for your cooperation. I hope you have fun.

MOVIE MAP QUESTIONAIRE FOR NEW VISITORS TO ASPEN

Instructions: Please answer the following questions after your first day in Aspen. Use back of sheet for more room where necessary.

1. What did it first feel like to come into Aspen? How familiar did the town seem?

2. How easy is it to find your way around Aspen?

   Has the MIT movie map helped? If so, how?

   Has the MIT movie map hindered? If so, how?

3. Did you have any preconceptions about Aspen? How have they changed?

4. Think about how Aspen has matched your expectations in the following categories. Check the closest choice. If you had no prior expectation leave it blank. Please add comments where they seem helpful.

	Much more than expected	Somewhat more than expected	About the same as expected	Somewhat less than expected	Much less than expected
size of Aspen	☐	☐	☐	☐	☐
size of the mountains	☐	☐	☐	☐	☐
size of residential area	☐	☐	☐	☐	☐
size of downtown area	☐	☐	☐	☐	☐
height of buildings	☐	☐	☐	☐	☐
height of trees	☐	☐	☐	☐	☐
width of streets	☐	☐	☐	☐	☐
width of sidewalks	☐	☐	☐	☐	☐
steepness of streets	☐	☐	☐	☐	☐
amount of open space	☐	☐	☐	☐	☐
distance from airport	☐	☐	☐	☐	☐
distance from mountains	☐	☐	☐	☐	☐
number of streets	☐	☐	☐	☐	☐
number of residences	☐	☐	☐	☐	☐
no. of hotels/lodges/condos	☐	☐	☐	☐	☐
number of public bldgs.	☐	☐	☐	☐	☐
number of parks	☐	☐	☐	☐	☐

5. If you have not already, please comment on any of the following that were different than expected:

colors

sounds

climate

architecture

amount of activity in Aspen

amount of "history"

amount of tourism

how picturesque

how much fun

6. How difficult was the trip from the airport to your destination? Describe it. Were you disoriented or lost at times? Please draw a quick map on the back showing the route you actually took.

7. If you had time to make a second 'mystery' trip, where you were taken blidfolded to a starting point, please answer as in the previous question. How difficult was it? Describe it. Were you lost or disoriented at times? Please draw a map of your route on the back.

Name

Address

Date

How many hours have you been in Aspen?

What time of day did you arrive?

What was the weather like?

How much longer will you stay?

About how many times a year do you travel to a new place?

What is your average length of stay on these visits?

<u>Please do not turn to the questions on the last page until leaving Aspen</u>.

8. Would you say that your exposure to the MIT movie map is equivalent to spending a certain amount of time in Aspen? If so, how much time?

9. Had the effect of the movie map upon your experience in Aspen changed between the beginning of your visit and the end? If so, how had it changed?

10. If you answered in the affirmative to number 9, during what time period did this change occur? Was there a point where the experience of being in Aspen began to replace the experience of the movie map?

MYSTERY TRIP

Please take your blindfolded passenger to
the corner of Spring St. and Hopkins St. facing
north toward Main st.
Drive them there by a somewhat round-about route
so they can't memorize the directions or route you
are taking.
Your passenger will now direct you back to Billy's.
(Without the blindfold, OF COURSE!

APPENDIX 4

Procedure for Final Experiment

SUBJECT NAME ____________________

____

QUESTIONAIRE BEFORE SESSION

Personal data:

1. Name Sex Age
2. Address
3. Phone
4. Occupation

Familiarity:

1. What was your hometown(s) while growing up?

   Length of time lived there?

   Approximate population?

3. Have you ever visited Aspen or seen movies of the town?

Visiting new places:

1. How many times a year do you travel to new places?
2. What is your average length of stay on these visits?
3. Do you enjoy learning your way around a new place?

Spatial abilities:

1. Rate your own sense of direction according to the following scale:

poor		fair		average		good		excellent
1	2	3	4	5	6	7	8	9

2. Rate your own ability to read, interpret and use maps:

poor		fair		average		good		excellent
1	2	3	4	5	6	7	8	9

3. How much do you enjoy using maps and giving directions:

hate		dislike		neutral		like		love
1	2	3	4	5	6	7	8	9

SUBJECT NAME ____________________

____

QUESTIONAIRE AFTER SESSION

I am trying to evaluate the conditions under which subjects can most effectively learn a path through an unfamiliar environment. Please make any comments which will help in understanding and interpreting your experiences in this experiment.

1. Were any of the directions or tasks to be performed difficult to understand?

3. Were you bothered by the jerkiness of your motion traveling down the street? If so, did this artifact become less objectional as you became more accustomed to the system?

4. Would you have preferred to have interactive control to determine where to go, when to stop, etc.?

5a. Please describe the process by which you became increasingly familiar with the route as you saw it in successive trials. Did you develop any particular strategy? Was your learning more or less continuous or were there certain pivotal stages?

5b. To what extent did you make distance judgments about the picture pairs based on inferences about objects in the pictures instead of remembering and reconstructing the spatial relations between them?

5c. To what extent did you rely on transitory objects like cars and pedestrians for landmarks?

6. Would you have preferred to draw a little map as you were moving along the route? (Did you perhaps construct such a diagram in your head?)

7. If you were one of the subjects who saw the street map, were you able to integrate that representation with your pictorial travel experience? Or did the two representations always remain distinct?

SUBJECT NAME ____________________

TRIAL NUMBER _______

Picture Id	Likelihood this location was on the route (0-100)	Same view - S Different - D	Which percentile of the route? (0,10,20,etc.)
A			
B			
C			
D			
E			
F			
G			
H			
I			
J			
K			
L			
M			
N			
O			
P			
Q			
R			

SUBJECT NAME ____________________

TRIAL NUMBER _______

	A	B	E	G	H	K	M	N	P
RECOGNIZABILITY OF THIS LOCATION									
SAME (S) VIEWPOINT OR DIFFERENT (D) VIEWPOINT [IF TOP > 50]									

	A	B	E	G	H	K	M	N	P
A	0								
B		0							
E			0						
G				0					
H					0				
K						0			
M							0		
N								0	
P									0

OVERALL CONFIDENCE IN THESE
DISTANCE ESTIMATIONS (0-100):

INTRODUCTION TO SPATIAL LEARNING EXPERIMENT

[to be read to subjects]

We have developed a new system for learning your way around an unfamiliar space, in our case Aspen, Co. It is called a Movie Map and it shows you pictures similar to the ones you would see out of the window of your car driving around. The pictures are shown on a television screen under computer control allowing complete interactivity. The driver can choose any route, deciding how fast to go, when to stop, where to turn. (If you are interested in the technology of the actual system you can inquire at the end of the experiment.)

In this experiment you will shown a prerecorded route which someone drove earlier. After watching you will be shown a set of pictures. Some of these will be of places you passed, others will not. Among pictures of places you have passed some will show the same view you saw. Some will show the the same place but with the opposite view (driving down the street in the reverse direction). Still other will show the same place but with the point of view perpendicular to the view which you had ("side window" instead of "front window"). Because this sounds a bit complicated you will get a preview of the experiment on a much shorter route.

You may notice that the movement is a little discontinuous because each frame was taken ten feet apart. Sometimes cars or pedestrians will pop in or out of view. Sometimes there are hesitations in your motion at intersections. These discontinuities Try to get used to these artifacts of the filming technique as much as possible.

[Run preview tape]

You should ignore the control panel at the bottom of the screen with the exception of the turn indicator arrows at the right and left which turn green when the driver indicates an intention to turn, and the direction label in the lower right of the screen. Please be aware that the direction is not updated until after the turn has been FULLY completed. Let's look at the route again and try to pay attention to where it is going.

[Rewind, Run preview tape again]

At this point in the experiment you will be given a set of pictures. Your first task will be to identify which you recognize as places which you encountered on the route.

[Separate the proper ones]

These were not on the route even though the setting may look similar. These were on the route. Notice that while some show a different vantage point that the one you had, they are still of the same place.

For the second task you will be given a set of pictures and asked to estimate the distance between every pair. For example, these two pictures were taken 1 block apart. These two are different views of the same corner building so they are maybe 0. 1 blocks apart. In the case of these two pictures taken at the begining and ending of the route, the distance is about 3 blocks even though the diagonal distance "as the crow flies" is shorter. The town is layed out on a regular grid of perpendicular streets, uniformly spaced. All distance estimations which you make should be in terms of walking distance. In some cases you will be asked to judge distances between places in the order they were encountered along the route. In other cases you will be asked to make judgements about the shortest walking distance between two places independently of the actual route which you travelled. For example, consider going around 3 of the 4 sides of a square block. The route distance is 3 blocks; the absolute distance is 1 block.

Are there any questions?

O.K. we are now ready to begin viewing the real route. It is five minutes long. Do not be discouraged if you have difficulty remembering everything you see. You will see the same route many times to increase your familiarity.

[If subjects are viewing route with map overview]:

Before the travel sequence begins you will see an overview map which traces out the route you will be following.

Trial 1:
O.K. Now you are ready for the first task. Put your name on the paper and "trial 1". For each picture write down the id, the likelihood you passed that location on the route. (definite yes = 100, definite no = 0, uncertain = 50). For confidence levels above 50 decide whether the viewpoint in the picture is the same (S) or different (D) than the one your saw. Ignore the last column.

Trial 2:
O.K. Now we will view the same video tape again. Remember though, when making connections between the snapshots and the video tape, cars and pedestrians may not be the same. You may want to pay attention to buildings, mountains and other permanent landmarks.

Task 2 is the same as task 1 except this time fill in the last column. Indicate the nearest percentile (ie. 0,10,20 etc.).

Trial 3:

O.K. We will view the video tape over again. This time the task will involve distance estimations between locations in the town. Go through the pictures and for each give a confidence level and indicate the viewpoint as same (S) or different (D). You will make a series of pairwise comparisons between the pictures. Start with the first picture in column A and compare it with each of the rest. Then turn picture A over and repeat the process with the second picture as reference. Proceed until all pictures have served as the reference. If you are unfamiliar with a particular picture, make the best guess you can. You should be making your judments as much as possible from your direct knowledge and memory of your experience watching the video tape - not from trying to infer clues from the pictures themselves. Make an overall judgment about how accurate you feel you estimations are.

Trial 4:
O.K. We will do the same as for trial 3.

[This may be a good point to take a 5 minute break, possibly have some coffee]

Trial 5:
This time we will do the same as for trial 2, recognition of the original set of 18 pictures and position along the route.

Trial 6:
Last time. Same as trial 3 and 4.

Please fill out the post questionaire.