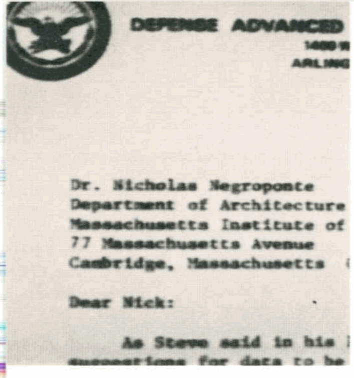
Spatial Data-Management



DEFENSE ADVANCED
1400 W
ARLING

Dr. Nicholas Negroponte
Department of Architecture
Massachusetts Institute of
77 Massachusetts Avenue
Cambridge, Massachusetts

Dear Nick:

As Steve said in his

Dr. Richard A. Bolt
Principal Research Scientist
Massachusetts Institute of Technology

Acknowledgments

Realization of our spatial data-management project derives from many minds and hands.

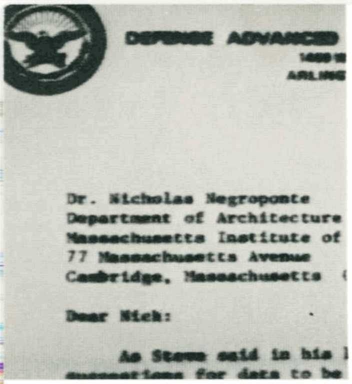
The Architecture Machine Group at the Massachusetts Institute of Technology, both as a collection of individuals and as a creative ambiance, generated the essential environment in which such an investigation could thrive. There does indeed operate the "spirit of a place," difficult to define, undeniable when present.

Systems implementation of spatial data-management system concepts was carried out under the leadership of Mr. William C. Donelson, who has performed throughout with exceptional initiative and ingenuity.

The energy and vision of Dr. Craig Fields of the Defense Advanced Research Projects Agency's Office of Cybernetics Technology provided extraordinarily vigorous and profitable sponsor interaction. We are all appreciative of his perceptive support.

My personal gratitude extends especially to Professor Nicholas Negroponte, Principal Investigator, whose élan and imagination has touched every aspect of this enterprise, indeed made it possible.

Richard A. Bolt


DEFENSE ADVANCED
ARLING

Dr. Nicholas Negroponte
Department of Architecture
Massachusetts Institute of
77 Massachusetts Avenue
Cambridge, Massachusetts

Dear Nick:

As Steve said in his
suggestions for data to be

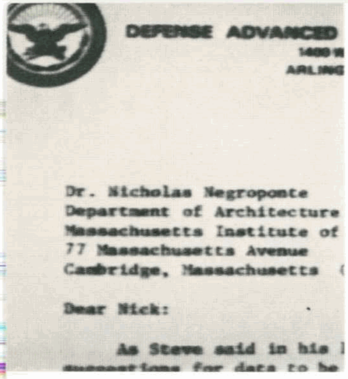
Contents

Prologue

Such startling advances and cost reductions are occurring in microelectronics that we believe future systems will not be characterized by their memory size or processing speed. Instead, the human interface will become the major measure, calibrated in very subjective units, so sensory and personalized that it will be evaluated by feelings and perceptions. Is it easy to use? Does it feel good? Is it pleasurable?

It is this interface that will bring computers directly to generals, presidents of companies, and six-year-old children. Thus, we ask the reader of this report to think about its contents in two ways. One is to judge the specific value of spatial constructs for nonspatial data. The other is to appreciate the general intention of making sound, visual, and tactile interfaces to serve conjointly as the modes and media of interaction with a computing resource, at once transparent and ubiquitous. Such quality is no longer a luxury but a requirement.

Nicholas Negroponte


DEFENSE ADVANCED
1400 W
ARLING

Dr. Nicholas Negroponte
Department of Architecture
Massachusetts Institute of
77 Massachusetts Avenue
Cambridge, Massachusetts

Dear Nick:

As Steve said in his

The Concept

The basis of spatial data-management is accessing a data item by going to where it is rather than referencing it by name.

Consider what occurs when you retrieve a book in your home library. You look at the left-hand bookcase, second shelf from the top, say, and scan about two-thirds down the shelf for a certain book spine. Thus, the book is located not by its title but by its customary site in a space defined by other books and the bookcase itself.

Consider how people ordinarily find items on their desk tops: the appointment book is up and to the right; the telephone is in the lower right corner; high priority memos are kept in an "in" box immediately to the left of the desk blotter; less urgent items are in the file folder, top middle of desk; and so forth.

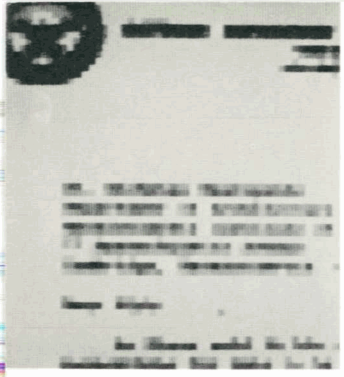
The person who uses this desk has organized the layout of items in a more or less systematic way. He or she refers to them constantly throughout the working day: reaching in that direction, that far, up, down, to the right, to the left. Through this activity, a mental image of the layout of the desk is elaborated in the "mind's eye." Additionally, through constant tactile interaction with the items, reaching for and touching them, a "motor memory" of where things are arises as well. A script for the act of retrieval becomes encoded into the musculature, as it were, according to where the item is located.

The practical importance of this spatial principle of organization is illustrated when someone happens to disturb our familiar arrangement of things. Perhaps some well-meaning soul tries to tidy up our seemingly chaotic desk top. Rather than being a benefit, the now "well-organized" desk is for us an organizational disaster: the former "messy" arrangement was a familiar and well traveled ground whereupon we could go directly to what we wanted. Such a loss of a familiar spatial layout can impress upon us the extent to which we take this subtle but powerful principle of organization for granted in our daily lives.

An even more striking instance of the role of spatiality in common experience is the coherence and persistence of "imagery." Two individuals, arguing a topic in front of a blackboard, will refer each other to diagrams, equations, and terms on the basis of where they had been written, even long after they have been erased!

It is surprising how pervasive the underlying notion of spatiality is, even in symbolic modes of thought. Consider a file of three-by-five notecards, organized alphabetically, but without letter tabs showing. "Filed under the letter *R*" translates into a tactile estimate of how far down the row of cards we must reach. The same holds true for a dictionary without thumb indexes. When we want to look up a word beginning with the letter *R*, we "guesstimate" where we should open the pages of the book to begin our search. Thus, *R* is somehow a relative distance, as well as a letter. There is more spatiality implicit in what is ordinarily thought of as symbolic retrieval than we may realize.

In summary, it is precisely the insight about how we tend to retrieve items from desk tops, from files and bookshelves, even from erased blackboards, that lies at the heart of the spatial data management concept: we find items on the basis of a more or less definite sense of their location in a familiar space, which space may be actually present or remembered.



This well-evolved human ability to organize information spatially remains essentially untapped in the realm of computer-based information handling. Typically, in such systems, retrieval on a symbolic or name basis is the norm, and must be the norm because the conventional keyboard interface is too limited a channel, the wrong mode and medium, to begin to offer the user a direct, palpable sense of spatiality.

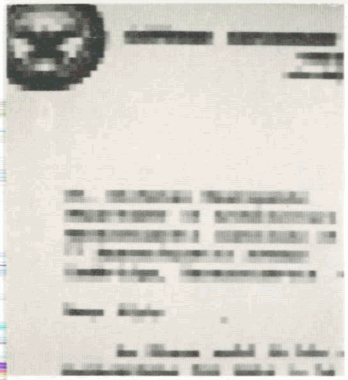
In the spring of 1976, the Architecture Machine Group at MIT proposed to the Cybernetics Technology Office of the United States Defense Advanced Research Projects Agency (DARPA) a program of research organized under the title "Augmentation of Human Resources in Command and Control through Multiple Media Man-Machine Interaction." Intrinsic to the ensemble of studies outlined in the proposal was a study recalling the ancient principle of using spatial cueing as an aid to performance and memory: the "Simonides Effect."

Simonides was a poet of ancient Greece famous for his ability to give long recitations entirely from memory. His secret, which as a teacher of rhetoric he shared with his students, was to tie each successive part of a to-be-remembered poem or speech to a specific locale within the mental floor plan of either an actual or imagined temple. Around the floor of the temple were statues, serving to augment the imagery of specific niches or corners.

To commit to memory a long poem or speech, the orator would pause in his imagination before the threshold, and with this mental image in mind, commit the introductory remarks to memory. Then, for each successive subsection of the talk to be given, the orator would mentally walk from place to place within the temple, rehearsing the appropriate material before some specific piece of statuary. The resultant path about the temple would then serve as a mental *schema* to organize the speech upon its eventual recitation, the speech being retrieved and reconstructed during oration by way of an imaginary tour from statue to statue around the temple floor.

Another study outlined in that same proposal looked to the exploitation of *virtual spatiality* as a matrix for organizing information: the "virtual bulletin board." This term refers to the formulation of an implicit space which occupies no "real" space but can nonetheless contain a great deal, as well as offer the same spatial organizational cueing as to the whereabouts of material that real work surfaces do. The tacit size of a "field" of information to be portrayed in computer graphics need not be limited by the size of the physical display. Since it is virtual, its apparent size and graphic organization are open to user definition in the unconstrained sense of Ivan Sutherland's characterization of a computer display as "a window on Alice's Wonderland" ("Computer Displays," *Scientific American,* June 1970).

These two spatial themes specifically resonated with Dr. Craig Fields, Senior Program Manager in the Cybernetics Technology Office. As a result we started to explore the potentialities of a spatially oriented data management system, "a multidimensional window into data." The outcome has been that over the past two years, under the aegis of DARPA, the Architecture Machine Group has been experimenting with the creation of an information management system whose distinguishing characteristic is that it exploits the user's sense of spatiality for purposes of organizing and retrieving data: a spatial data-management system, or SDMS.



The Setting

This page: The monitor to the user's right as seen over the arm of the user's chair. Fingertips of user's hand are on the small joystick. On the chair's arm, immediately in front of user's hand is a small touch-sensitive pad ready for capture of fingertip gesture input. The other arm of the chair is similarly instrumented.

Opposite page: The absence of any desk-like appurtenances situated in front of the user lends a tone of immediacy and directness to interaction with data. At the upper right corner of the large screen, one of the sound system's speakers is clearly visible.

The lapboard and stylus are not featured in these illustrations. For a view of them, see illustration at top of page 48.

The setting of our experimental version of SDMS is a multiple media information place. It can be construed as an image of an office of the future.

The physical aspects of our "media room" are easily described. The media room is about the size of a personal office. The focal point of the room is a wall-sized display screen. Directly in front of this large screen is an instrumented Eames chair. On either arm of the chair is a small, pressure-sensitive "joystick." Adjacent to the joystick on each arm is a two-inch-square touch-sensitive pad. Also associated with the chair as complementary instrumentation is a ten-inch-square data tablet, with stylus; the tablet is upholstered to serve as a lap pad. This tablet bears added instrumentation: a small microphone mounted at its top edge to allow voice input.

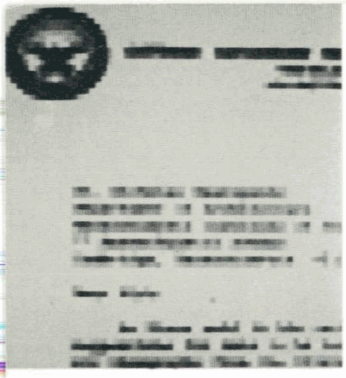
Two small television monitors are situated one on either side of the chair. Each of these monitors is touch-sensitive so that the user can point to data or can input gestures indicating actions to be taken.

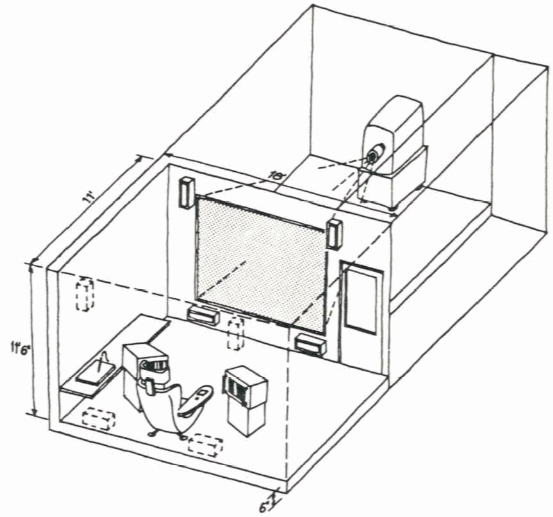
Finally, eight loudspeakers provide sound output for the media room. Four of them are placed around the perimeter of the large-format display to the user's front. The remaining four are correspondingly arrayed to the user's rear. Octaphonic sounds can be presented to the user as coming from various apparent directions in the space of the room, in coordination with actions on the visual display. The rate of playback is variable so that dynamic sound/motion effects, such as Doppler shifts, may be achieved. We call these "sound-synch graphics."

Notice that there is a conspicuous, but not mandatory, absence of a keyboard.

The configuration described is one which evolved through previous work for the Office of Naval Research (Contract number 0014-75-C-0460) and over the course of our experimentation, but this precise ensemble is certainly not the only one that might have been adopted. The kernel notion of managing data spatially is not necessarily tied to a room-sized terminal into which a user goes versus a desk-top arrangement in front of which he or she sits.


AMERICA





A television projector in the room adjacent to the media room back-projects upon the large screen to the user's front.

The relaxed ambiance conveyed by the presence of a chair of the Eames genre in lieu of something more utilitarian is intentional, however. It reflects convictions and positions about the nature and tone of human-computer interaction that we have attempted to actualize in the media room setting. Just as the hands-on immediacy of touch-sensitive pads suggests a literal impatience with intangibles about data, so the decor as epitomized in the selection of the style of chair rebuts the premise that system users must live in severe, ascetic settings. (See the Prologue to this report.)

We have attempted to create an interface which is not a tiny, narrow-band "port-hole" into an information bank, that bank itself an abstractly addressed set of intangibles. Rather, we have attempted radically to recast the setting as an "informational surround" wherein the user is directly engaged with data bodied forth in vision, sound, and touch, data inhabiting a spatially definite "virtual" world that can be interactively explored and navigated.

Finding Data

Dataland as it appears on the "world view" monitor to the user's right. Notice highlighted "you-are-here" rectangle situated over a map of northeastern United States located in upper-left region of Dataland. "You-are-here" marker is translucent, under joystick control, and indicates the region of Dataland visible on the large screen to user's front.

Spatial data-management involves the creation of a plausible and commodious "virtual" space at the computer interface, together with a way of getting around in that space quickly and easily.

The spatial world of SDMS consists of a single *plane.* This surface, called "Dataland," is presented to the user in two aspects:

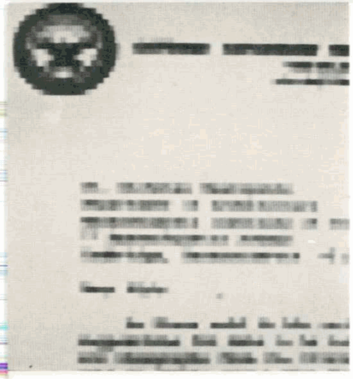
1
it is continuously visible to the user in its entirety in an "aerial," top-on view displayed on one of the monitors, arbitrarily situated to the user's right, called the "world view" monitor; and

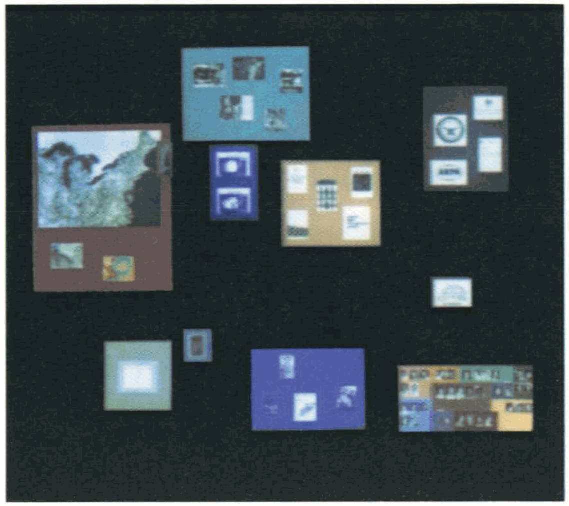
2
simultaneously, a small subsector of the Dataland surface is displayed on the ten-foot diagonal screen to the user's front, vastly enlarged and with appreciable gain in detail.

The logical relationship between these two views is akin to a mapping key, which places a particular image within a larger locale. The large screen effectively functions as a "window" and, as we shall see later, a "magnifying glass" onto Dataland.

The "world view" monitor serves specifically as a navigational aid to the user in getting around Dataland. The large display of whatever portion of Dataland is so "close up" that the user would get lost easily if there were not always on view a map of the entire Dataland world. A small, highlighted, "you-are-here" rectangle on that display shows the user at all times the position of the large-screen window on Dataland.

To illustrate how an actual Dataland might appear to a user seated in the media room's control chair, let us assume a database already created. In practice, individuals would design and structure their own Datalands according to personal preferences, tasks, and needs, just as the contents and organization of anyone's desk top come to reflect some combination of job requirements and individual taste.





Illustrations on this page and the next depict relationship betweeen "world view" monitor and the large screen. On Dataland view, this page, the translucent "you-are-here" marker resides over lower-left-most item. The item, a calendar, appears on the large screen, much magnified and with finer detail (opposite page).

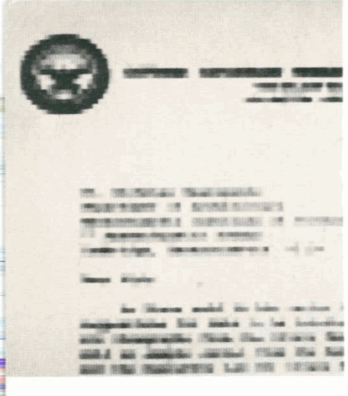
Now, what is visible to the user on the surface of this example of Dataland?

Items of various sorts may be seen. Some are small photographs: pictures of animals, of people, even a miniature LANDSAT satellite photo of part of New England. Other items look like small maps. We see what appear to be book covers, television sets, a business letter, even a small image of an electronic hand-held calculator. Yet other items resemble emblems or "glyphs," not dissimilar from business logos.

How does one get around in such an informational landscape?

User control of navigation around Dataland is provided by "joystick" action. The small joystick in the right arm of the chair permits a helicopter-like flight over the surface. The joystick itself does not deflect, but responds to pressure. It is "self-centering" in that it registers zero deflection when not touched. Push the stick "away" from you and the helicoptering motion over Dataland is north, or toward the top of the screen. Pull the joystick toward you and motion is south. Pressure to the right and left will direct you east or west.


24
24



The apparent rate of travel is proportional to the pressure the user exerts on the joystick. At maximum pressure, it takes about twenty seconds to travel across Dataland, as illustrated in this example database, from its eastern to western borders.

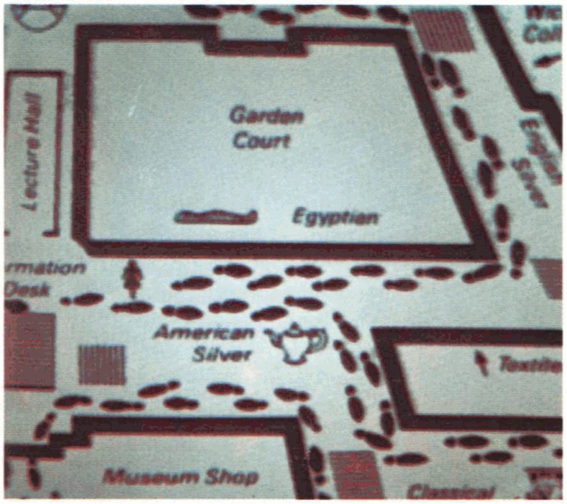
During flight, the user is given visual feedback of two kinds:

1
movement of the material across the "window" of the large screen;

2
tracking of the path of the window by the highlighted rectangle on the auxiliary monitor.

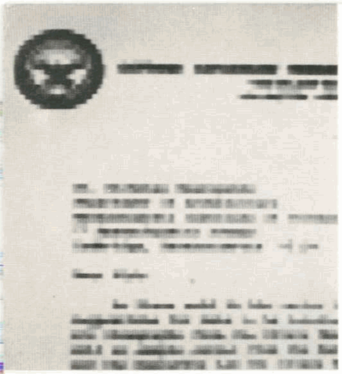
Let us say that you push the joystick to the right and away from you. This action will set you helicoptering across Dataland to the "northeast." On the big screen you will see material scudding by in a *southwesterly* direction; that is, the intuitive sense of movement will be that of a window looking out over a world passing by below. Simultaneously, the highlighted rectangle on the auxiliary view at your right hand will track your northeastward progress.

Complementing joystick control of flight over Dataland is "teleporation" by direct touch. Traversing relatively large distances with the joystick can be frustrating; for the veteran user, this often represents "dead time." Accordingly, the user is given the option of indicating where he wants to go directly on the touch-sensitive "world view." The highlighted "you-are-here" rectangle will disappear from its current location, to reappear centered at the spot indicated by the finger touch. On the large screen, the "magnified" view of the old location disappears, to be replaced by the just-indicated locale.



"Teleportation" about Dataland by direct touch: At top, user touch indicates item within a group with orange background as what he wants to go to on Dataland. This item, a "book cover", fills screen to user's front as user next indicates a location to the right and somewhat lower (middle). "You-are-here" marker immediately moves to this new locale, and large screen view now bears magnified view of the newly selected item, a schematic floor plan of a museum.

While its specific format is much a personal matter, the graphic layout of Dataland can aid navigation significantly. First, the directness and simplicity of Dataland itself provides a unitary conceptual substrate: a large plane conjointly given via two correlated views on the small and large screens. Second, the absolute and relative location of objects and items is evident within the frame of the Dataland world. Absolute locational cues are provided by the main "frame" of Dataland: the fact of edges, corners, up-down and left-right directionality. Any item is perceived and remembered in the context of and juxtaposed with certain other items. Third, the use of local background colors affords a further visual sense of neighborhood, grouping certain items together. Color backgrounding can, of course, be "nested," color background upon color background, as it is in our "rogues' gallery" of Architecture Machine Group personnel in the southeast corner of Dataland. Within the overall rectangular area there is further group intraorganization indicated by background colors.



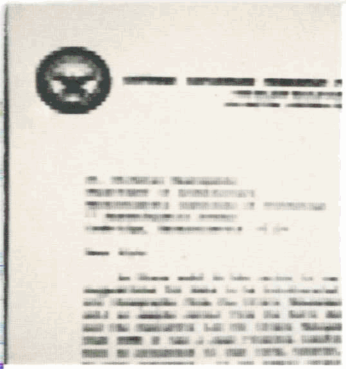
Faces which are unrecognizable and tiny on Dataland as seen on the "world view" monitor become detailed and clear when zoomed-in upon on the large screen.

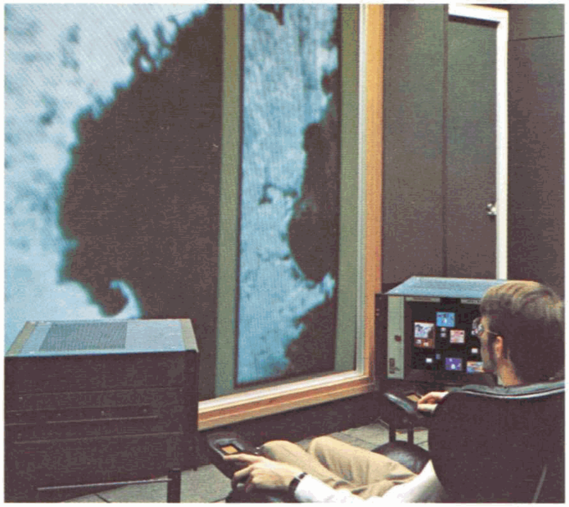
The level of visual detail in items provided on the auxiliary world-map view deserves comment. In everyday life, we need some minimal amount of visual detail to identify an item that we have never seen before. If an item is small in size, or small because distant, it can lack sufficient shape or figural detail to enable us to discern what it is. However, recognition of already familiar items requires less figural detail. We readily recognize friends way down a street, at distances where visual acuity is insufficient to provide facial details. We recognize them on the basis of gait, relative size, general shape or "visual envelope," as well as by *context,* that is, by expecting a figure with such-and-such-color clothes to appear in such-and-such a place.

Or, take the example of a Coca-Cola sign down the street. At a distance at which we cannot distinguish the individual letters of the sign, we yet recognize it as being a "Coke" sign. The fact that it appears over a drugstore or variety store supports contextually this act of recognition.

A similar principle holds in the area of *audition.* On a certain popular television game show ("Name That Tune"), persons who are particularly well informed about popular music as well as alert can sometimes identify ("recognize") a song on the basis of no more than its first two or three notes. Consider, for example, the first two notes of the melody "Auld Lang Syne." The clue that the song about to be heard is associated with a holiday, the distinctive tonal interval of the first two notes of the song, can be sufficient to define the intersection of context (New Year's Eve) and tonalities, as, indeed, "Auld Lang Syne."

The *recognition* situation, as opposed to the *identification* situation, is what we anticipate to be the operating condition for the typical user of SDMS. That is, the array of highly familiar items in Dataland is "searched," not by a deep visual analysis of every item before one's view, but by relatively "shallow" visual processing, quite adequate for the recognition of familiar items in familiar places.





If a user has been working with a certain personal Dataland configuration for some time, it is reasonable that all that should be needed on the auxiliary view of Dataland is sufficient detail to recognize items readily on that surface. With constant usage certain items could undergo further reduction in visual detail, even to the extent of becoming simple colored squares, circles, or triangles. This extreme represents a "cueing" situation, rather than a recognition situation.

Qualitatively, the figural resolution of items in Dataland as they appear on the world-map monitor fall in the intermediate range on the spectrum: cueing, recognition, identification. That is, they serve readily for *recognition* by an accustomed user, but not necessarily for *identification* upon a first-time look by a new observer.

Perceptibility indeed differs somewhat from item to item. For example, the LANDSAT satellite image does look like some bit of geography; but it is not clear whether, from a farthest-out view, it is a LANDSAT view, a map, or a drawing. The user who knows what it actually is of course readily identifies it and refers to it as "the LANDSAT photo." Or, consider the set of photographs in the southeast corner of Dataland. They are spontaneously classifiable as photos, but even for those who are personal acquaintances of the people pictured, the photos are ever so slightly too small and undetailed to permit positive identification without zooming in. Nonetheless, it is clear from "afar" to any observer that the items are in fact photos of faces.

Sound, too, can be an invaluable navigational aid. In SDMS, space as cued by sound can have an implicit extent much greater than that bounded by the visual "window" onto Dataland given by the large-screen view. Off-screen voices, becoming gradually audible as one approaches some item or area, can serve to orient and guide.

Two examples of sounds that we have employed include:

1
an off-screen voice, in stereo, reciting "This is MIT...This is MIT...," the voice leading the traveler in Dataland first to the vicinity of, and then to, a certain "glyph" bearing the initials MIT, which was the source of the sound;

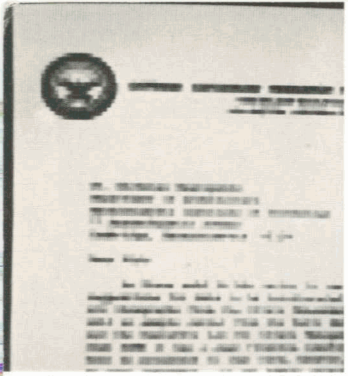
2
a gray telephone with white touch-tone buttons (*not* activated for "dialing," but potentially so), which rings louder and louder upon approach.

These and other sounds did much to create a sense of "region" and "neighborhood," a feeling of proximity, much as the bells, hammerings, traffic noise, and playground chatter of the real world fill things out. At moments in navigating across Dataland, depending upon the route taken, sound can be the *sole* directional cue, should one momentarily not be consulting the "world view" auxiliary monitor but watching the large screen.

We note in passing that the user is given interactive control of the loudness level of sound through gesturing upon either of the small, touch-sensitive pads on the arms of the chair. Circling with the fingertip in a counterclockwise direction attenuates the volume of sound, and a circling gesture clockwise causes volume to increase.

With respect to navigation around Dataland, we have so far talked only about movement over the surface of the plane of data at a constant "altitude." How then does the user manage to get at the data detected at a distance?

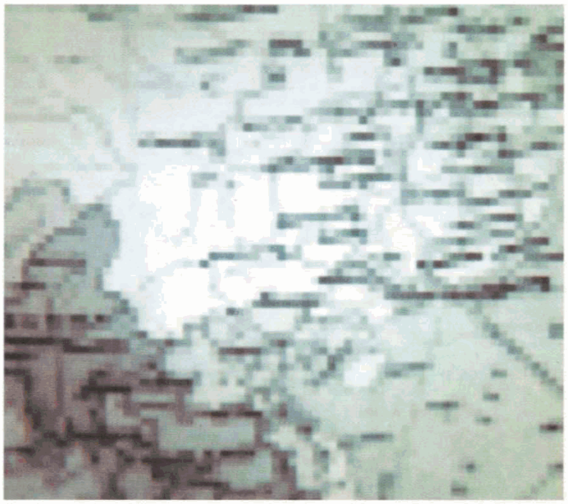
The Northeastern Coast of the United States seen from afar. A satellite image of the New England coastline is inset at the edge.

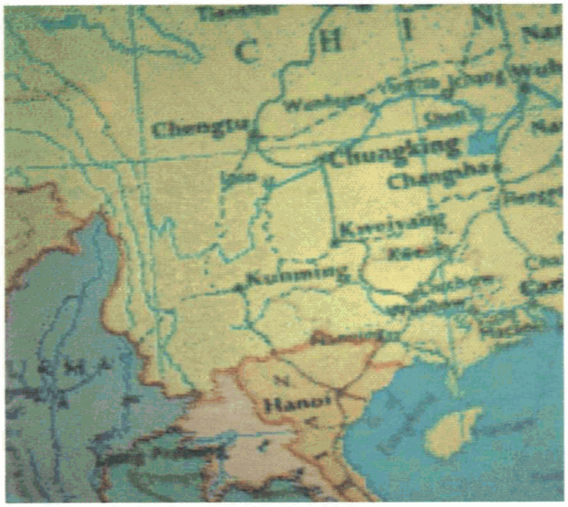


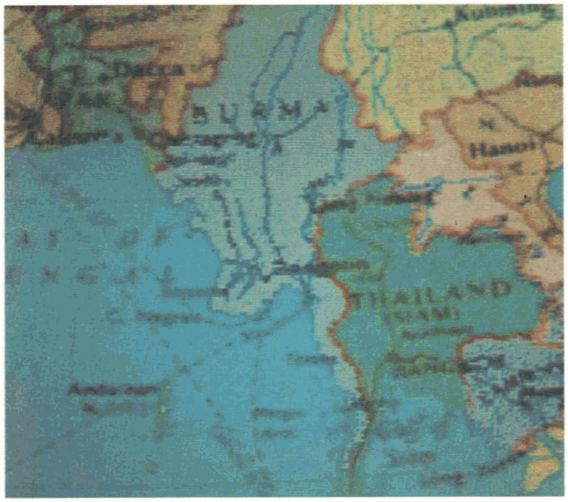
Zooming-in upon Dataland's map of Southeast Asia: successive images on this page depict a monochromatic map becoming larger and "blockier" in texture. The image then gains both clarity and color (opposite, top) for scrolling about (opposite, middle, and bottom).

While the right-hand joystick guides lateral and up-down movements, the joystick mounted in the left arm of the media room's control chair permits the user to "zoom" in upon Dataland for a closer look. Pushing forward on the left-hand joystick will cause zooming *in;* pulling back will cause zooming *out.* Upon zooming in, the image of the item on the large screen gains in size dramatically. The relative size of the "you-are-here" highlighted rectangle on the "world view" auxiliary monitor grows correspondingly smaller, yet remains on view so that the user will always know his location in the spatial context of Dataland.

For example, on the "world view" auxiliary map, we can recognize a small, monochromatic map of Southeast Asia located in the western part of Dataland, at about mid-latitude. Coming in closer upon the map by pressing forward on the left-hand joystick makes detail more discernible, as well as causing the cartography to gain color.

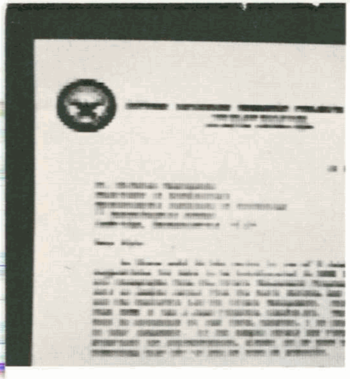






Notice that the enlargement of the map image takes place in "stages." That is, the image upon zooming in becomes larger in discrete increments, simultaneously more "blocky" in appearance. At a certain point in zooming in, the blocky image is replaced by a clear map image the same size and in the same register as the previous blocky one.

The increasing "blockiness" of the image accompanying the increase in size during zooming in, requiring eventual replacement with a clearer image, is a consequence of the *digital* nature of the television image and of the way the "zoom" effect is achieved.



In broadcast television, the "information" for the image, namely the brightness (luminance) and color (chrominance) information for each scan line composing it, comes over the airwaves as an analogue signal. With digital television, however, the information which modulates the image either is generated by computer computations or comes directly out of the computer's memory store. The color television images displayed in SDMS, with important exceptions to be noted later, actually reside as stored information in an area of computer memory termed a "frame buffer."

The digital color image in SDMS can be conceived of logically as composed of an array of tiny dots called "picture elements," or "pixels." Each pixel corresponds to certain numeric values located in addressable slots in the frame buffer. Given this mosaic, a ready way to blow up the image for a zoom-like effect is to repeat (in the sense of "displaying again") each picture element, in both the horizontal and vertical dimensions of the image, first once, then twice, then three times, and so forth. The repetition of pixels is patterned from center screen outward in emulation of a central-axis "optical" zoom. The result visually is a negatively accelerating rate of image expansion.

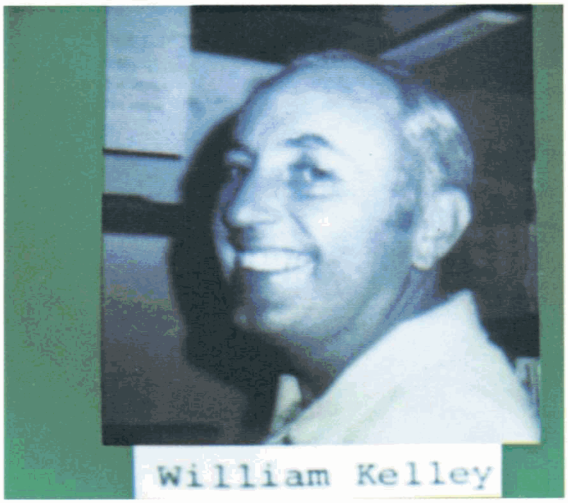
The basic Dataland image, as seen from the highest "ceiling" or altitude, actually incorporates one "repeat" to begin with, partly to avoid the rather large visual "lurch" inherent in the quadrupling of image size that occurs when each pixel is first repeated once vertically and horizontally. One initial "repeat" makes for a smoother total effect even though only half of the television screen's real resolution is being used.

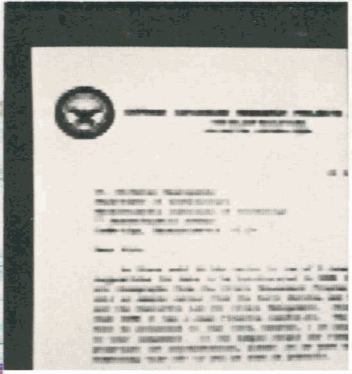
The increasing "blockiness" of the image upon zooming in derives precisely from the squarish *pattern* of repetitions of the individual pixels, the basic pixel being visually just a small dot of color. To counteract this side effect of the way we achieve zoom effects, the now larger and blockier image can be replaced by an image of the same size and in the same register, but with increased detail and clarity.

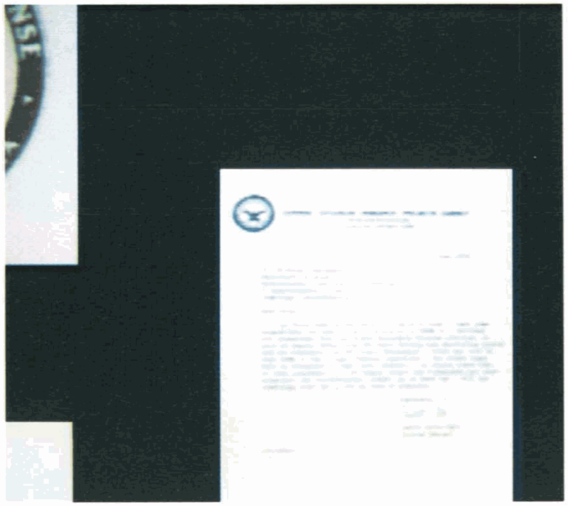
On the large screen to the user's front, image of William Kelley of our laboratory's staff, at first small (opposite page), becomes larger and "blockier" as the user zooms-in (top and middle, this page). At a certain point in approach, the blocky image is replaced (instantaneously) with a higher resolution image (bottom).

In practice, the zoom procedure we have adopted calls for the replacement of the scaled-up, "blocky" image at about a repeat count of seven, so that the loss in clarity with increasing blockiness is not too great. With successive rounds of repeat count incrementing followed by image replacements, a zooming in process could be as "magnifying" as one wished, the practical limitation being the resolution of the source image information.

Other techniques of zooming are under investigation, in particular interpolative zoom, which removes blockiness and jumpiness but manifests the same degradation in resolution.







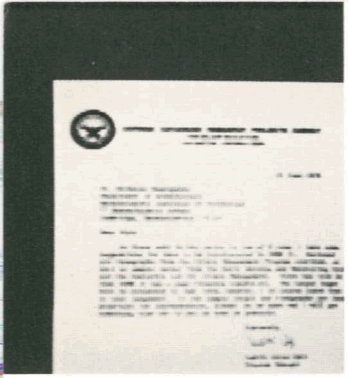
At a distance, and at low resolution, this image on Dataland is clearly a letter. Letterhead *logos,* despite the coarseness of scale, generally can serve as reliable cues for sender's identity.

While it might at first seem that such discontinuities in image sharpness during zoom can only be disadvantageous, there are circumstances in which a lack of low-order visual detail followed by the sudden introduction of finer detail makes good sense in terms of information presentation.

In the upper right, or northeast, corner of Dataland is an area allocated to Architecture Machine Group "correspondence" with our research sponsor, DARPA. In this area is a facsimile of a letter, complete with letterhead and signature. When approached, the letter indeed looks like a letter but is as yet unreadable. At a certain threshold nearness, the text suddenly gains in clarity and becomes perfectly legible.

To zoom in upon an item of interest is to "address" that item for purposes of inspection. For some items, simply to come up to them will cause them to gain not only in size but in information: increased detail, finer delineation, color, and, in some cases, sound and movement. To approach them is to "activate" them. With other items, some further user initiative such as "touch selection" may be involved as a part of a more interactive process of *perusal*.

Zooming-in upon the letter image results in legible text.



Perusing Data

Logically distinct from *finding* a certain item in Dataland is *perusing* that item in place. Once the user has found the item and zoomed in on it, the specific form that perusal takes is conditioned by the nature of the data.

In direct contrast to material that is "out there," visible, to be developed in full simply by approach, are materials which need to be "perused," where perusal means opening up a volume, starting a film clip, selecting part of a set of slides to be seen, and so on. The essential idea is that the user opens up and develops his interchange with the data more or less interactively, through direct touch in a "hands-on" manner. The dictionary sense of "perusal" is that of "survey or scrutiny, in some thoroughness or detail." It is this *engaged* quality of person-data interchange that we intend by this terminology.

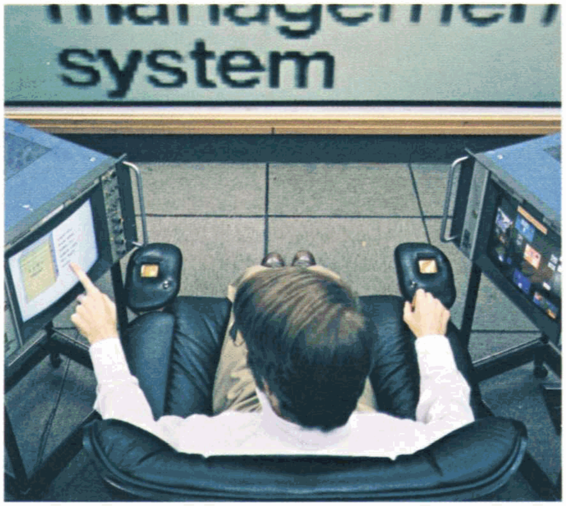
Our notion of perusal includes a class of "data types" which, when addressed, prove to be unusually rich in motion, color, and sound. One of these new data types is yet an old and familiar one: the "book."

The user zooms in on what appears from afar to be simply a rectangle. At close enough range, the color patch gains detail in the form of some horizontal black rows of "characters" implying a printed title. Upon even yet nearer approach, the black figures give way to readable text, giving the title of the item, which is now clearly seen as implying a "book-like" image.

The user, of course, knows that the rectangle is a "book," having put it there in the first place. However, if he is a subscriber to a computer-network-based "book of the month" service, and has perhaps requested or has a standing order for books, the book may be one he has yet to see, having just appeared on the Dataland surface in a certain area on the data plane serving as a "book drop" for incoming communications which are book-like.

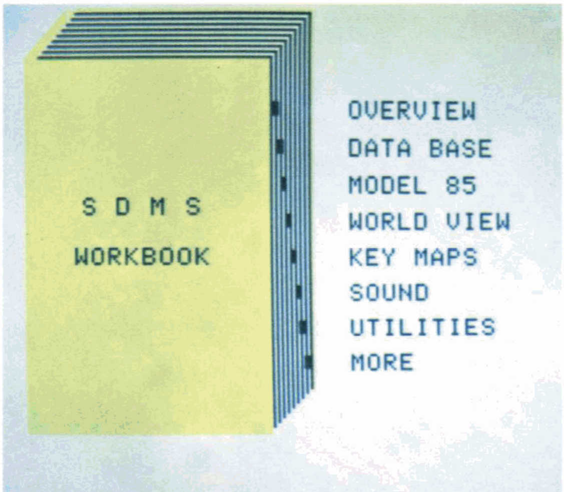
Upon nearest approach, when the image of the book is fully addressed in the visual sense of filling the large screen, a "table of contents" for the book appears on the second and so far unused color monitor to the user's left. This table-of-contents view is one of a class of interactive "key maps" made available to the user as adjuncts to certain data types, to permit direct user interaction with the data to be perused.

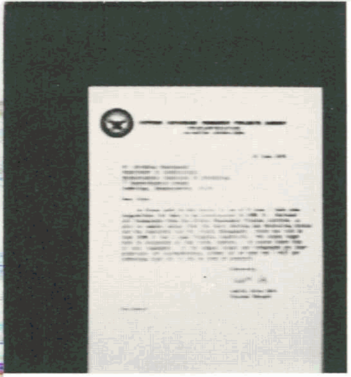
In this instance, the key map allows user access to the interior of the "book" on display, exactly as the table of contents of a "real" book lets the reader *(a)* know what the book offers by way of topics covered, and *(b)* serves as a "dispatching" table to send the reader directly to the subsection of the book that is of most immediate interest.



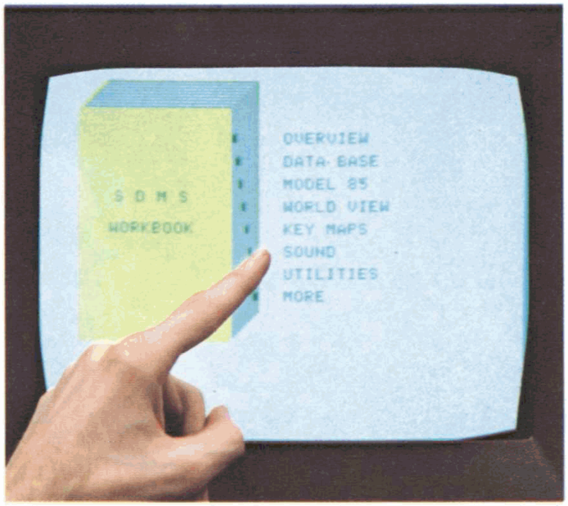
User interacts by touch with a "table of contents" key map for book-like data types.

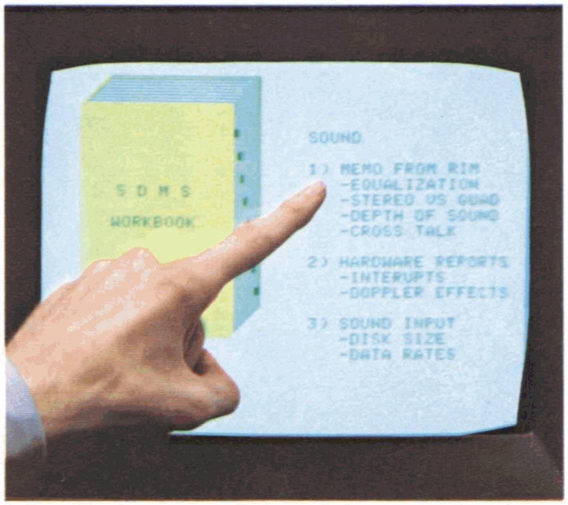
In close-up of this key map below, note touchable "tabs" aligned with heading taglines.





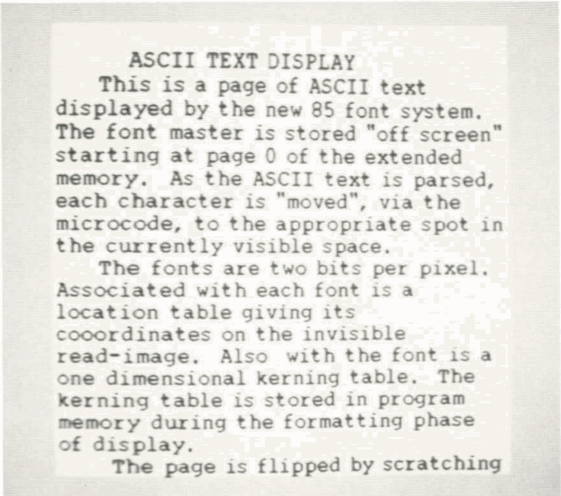
Upon touch indicating chapter heading of SOUND (top), the heading expands (bottom) to offer sub-headings under SOUND.

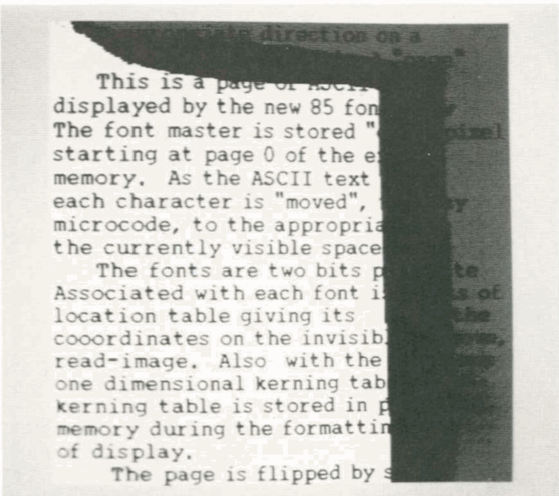


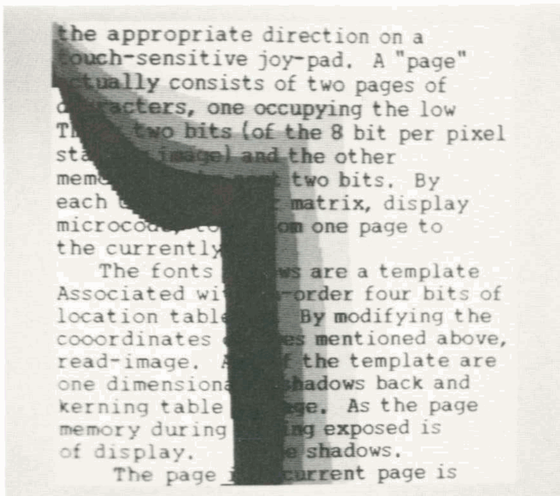


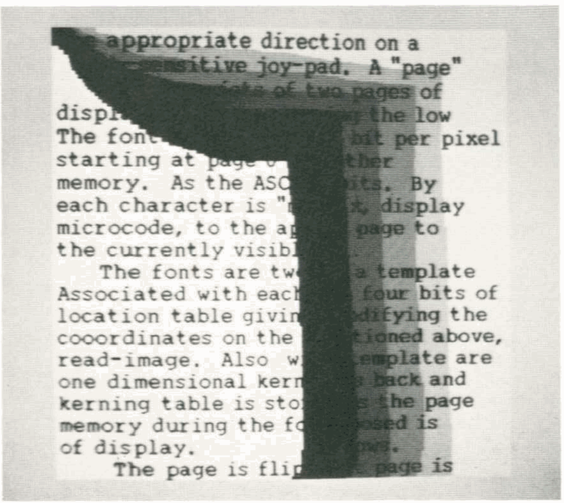
Notice that the table of contents can "expand" to show a finer breakdown of content area when the user selects a major heading of interest. In concert with the activity on the key map monitor screen, the main large-screen view will "go to" the section of the book selected by the user through the interactive key map so that the user may read the material.

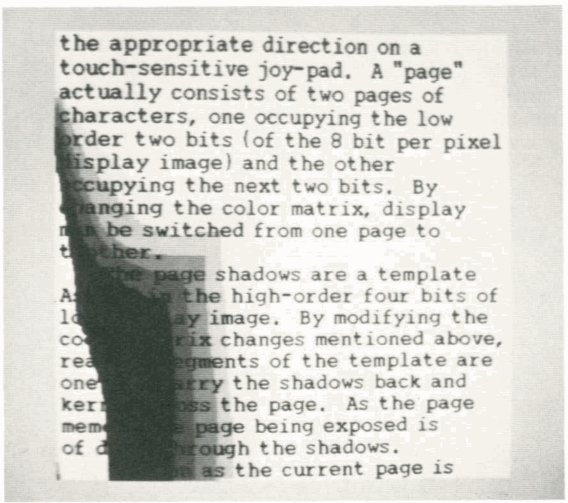
There is yet further interaction with the data type of book: that of turning its pages. This action is initiated by a specific user action: a page-turning *gesture* given as a right-to-left, top-to-bottom stroke on either of the touch-sensitive pads on the arms of the user chair. Any page can be turned back by a stroke across the pad in the opposite direction. The accompanying visual action is the display of an "animation" of a page actually turning. The upper right-hand corner of the page on view progressively "lifts" away and sweeps leftwards across the screen, immediately revealing the new page below.





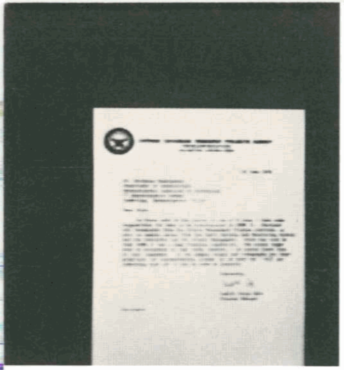






Initial page of text (left, top) begins to "turn" (left, middle) upon user indication (a stroke of the finger) on touch-sensitive pad on arm of user chair.

The next page of text is progressively revealed as the "page" edge sweeps leftwards, and is almost entirely on view near the end of the page animation sequence (right, bottom).



The image of the "television set" serves as locale marker for television data types on Dataland. On the tube face of the set can be a "program guide", or, as in the example on the following pages, a pictorial label which becomes dynamic upon approach.

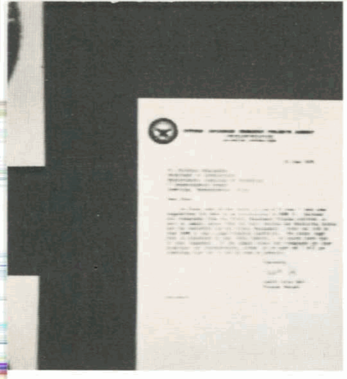
The intent is not to give a nostalgic impression of the way books used to look, like the electric fireplace with plastic back-lit logs. The purpose is quite functional. First, the page is a unit of apprehension of the book; it is a "bite-sized" chunk or packet of the material that one feels one can take in comfortably, in contrast to the stretching-ever-onward, rhythmless character of scrolled text. Second, the page serves as a progress marker through the text in a "milepost" sense. With continuous scrolling, such natural, discrete units of progress are not explicit in the medium, and, if felt necessary by the reader, must be effortfully abstracted, resulting in a more burdensome tone to the reading task.

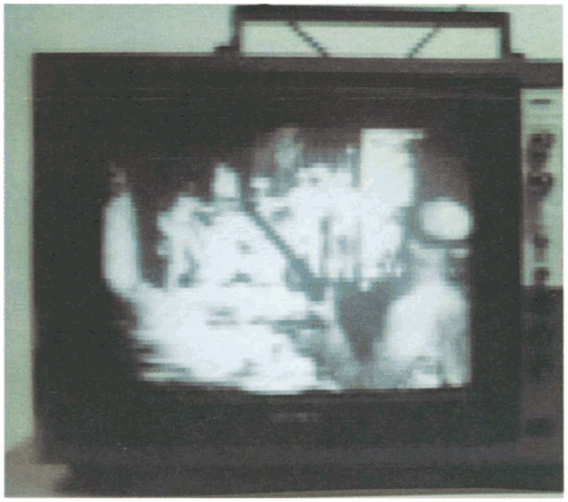
Thus, the "page-flipper" is not frivolous. It provides the user with feedback that a page has turned, the directionality of the animation corresponding to forward or backward. Specifically, it reinstates for the user, in the context of the reading of computer-television text, the functional equivalent of what the *fact* of the page gives to the reader of traditional paged printed media: the sense of knowing where you are in the material. More generally, it is an instance of what we choose to term a "media fiducial," a marking system, or built-in feature of any medium that serves the function of letting you know "where you are" based on locational cues arising out of the basic structure of the medium itself.

One radical departure from the ordinary concept of the book, except as adumbrated in such page-corner flip-animations as this report has in its upper left-hand page corners, is the fact that while some page illustrations will be still slides, other illustrations will be "movies," or animated diagrams, with or without sound accompaniments. The possibility of sound can, upon occasion, permit the book to comment upon its own text and illustrations, and even read itself to you.

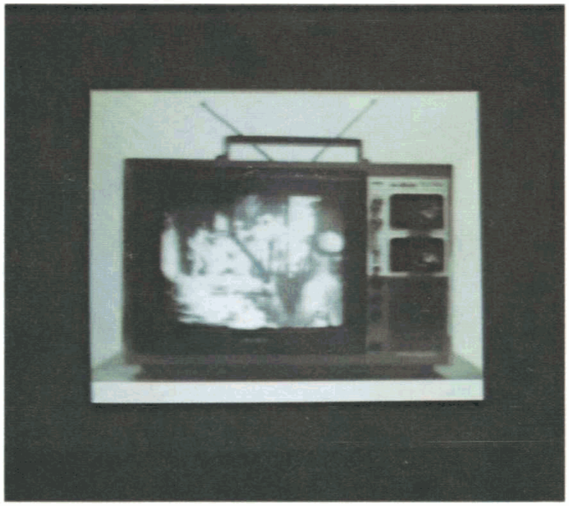
Another unusual data type residing in Dataland is that of television. An instance of this is the presence of a "glyph" on the Dataland surface which is the image of a Sony television. Zoomed in on, the television set becomes very large, its own screen filling the large screen directly before the user. What is at first a rather sketchy black-and-white image on the tube face of this "virtual television" is then replaced with live action: "Columbo," "The Sting," or a documentary on how to stop smoking.

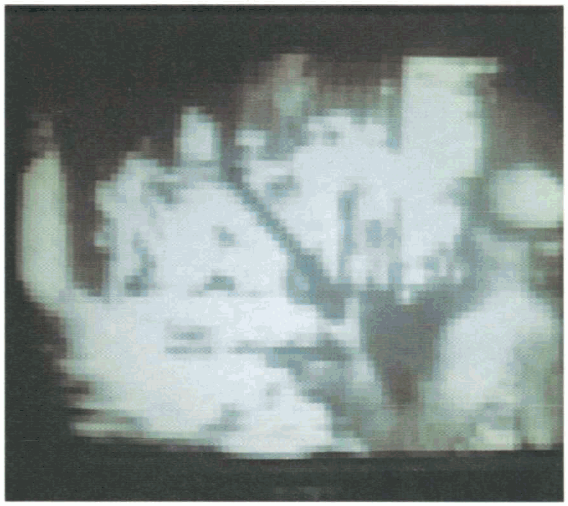
What may be seen on this television set in Dataland is anything that can be put on television, live or recorded, closed-circuit or broadcast. The user could, for example, indicate in the context of a session with SDMS that he next wants to view the twelve o'clock news, or see a rerun of some footage that was taped earlier.

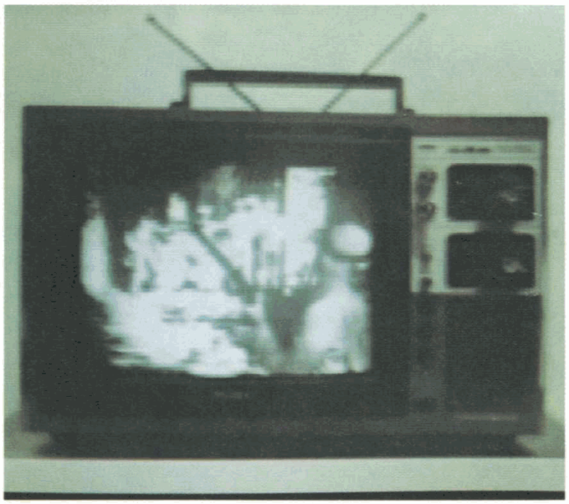


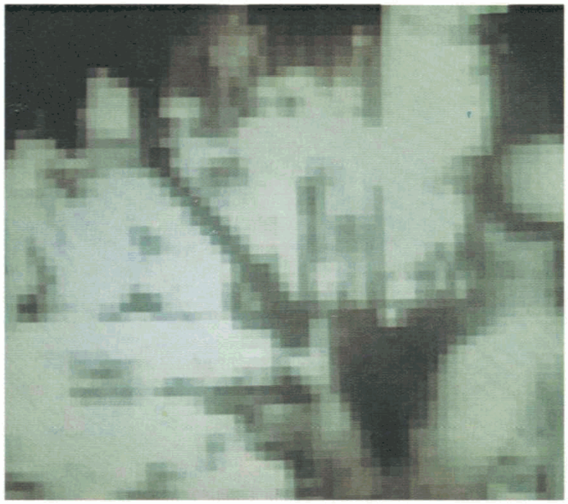


Zooming-in approach to "television set" on Dataland (left column), enlarges the television set's screen to become effectively as large as the screen to the user's front (right column). The "blocky" digital image becomes replaced by a dynamic image in color from videodisc, launching into continuous live action (opposite page).



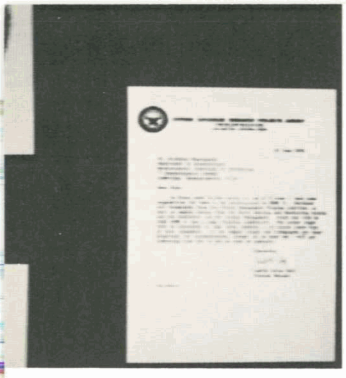


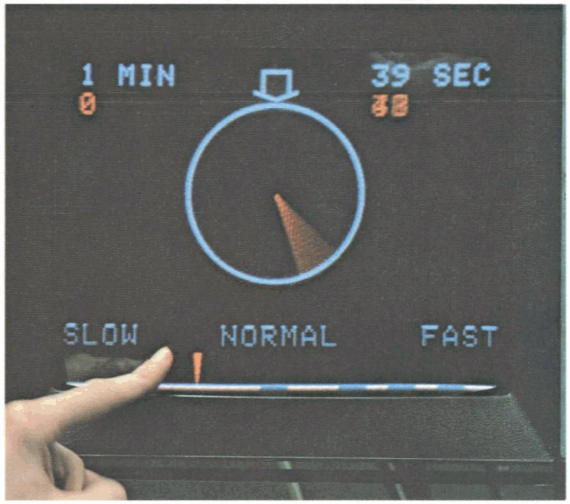








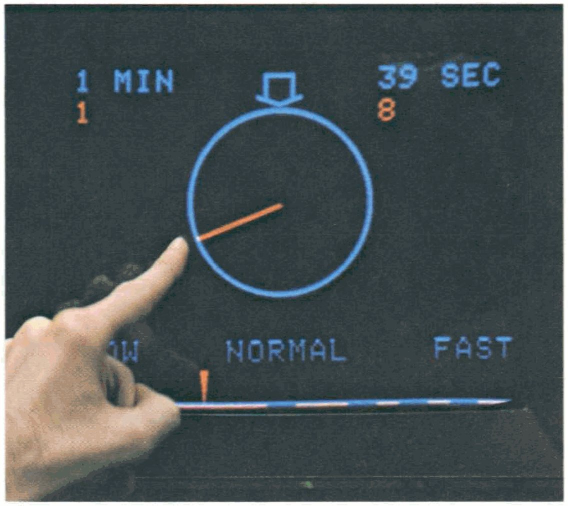





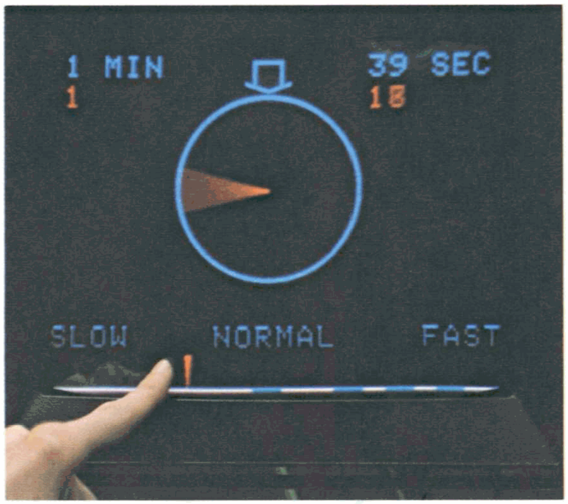
Since recorded sequences that might be viewed on the television are processes taking place over time, the key map for interaction with sequences on the virtual television set makes available to the user ways of controlling events that are oriented over a time span.

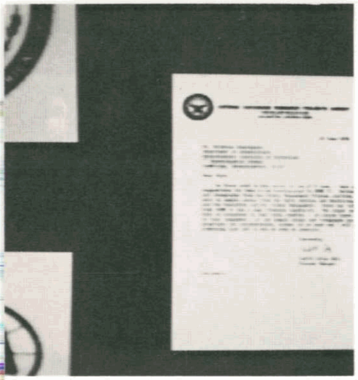
An example of such a key map is shown in the accompanying illustration. Central to the key map is a "dial," whose circumference indicates the time span to be covered by the *entire* movie sequence, with a pointer hand showing the current time in passage. One revolution of this *semantic* clock can be seconds, minutes, or hours. The total time-to-go is displayed by a digital counter on the key map; minutes and seconds elapsed are also displayed digitally.

Sliding-bar cursors toward the bottom of this key map allow the user to speed up or slow down the rate at which material is shown, as well as to have it shown in a forward or backward sequence. The material can be simply a "movie," in which case it probably makes little sense to show backward slow motion. But if the material happens to be an instructional movie, for example, about how to tie a sailing knot, or do a sleight-of-hand card trick, then this interactive ability to regulate time can be invaluable. The material to be shown in such a format can be, in general, any material that can be viewed visually and sequentially: movies, television, animations, slide sequences, as well as any electronic product of television mixing and special effects.



User interaction with "dial" key map: user sets pace to be somewhat slower than normal (top, left). Sequence is stopped (freeze frame) by putting finger on pointer of dial (middle), and resumes when finger is lifted (bottom, right).





MCA Discovision® Industrial Videodisc Player – Model P7820.

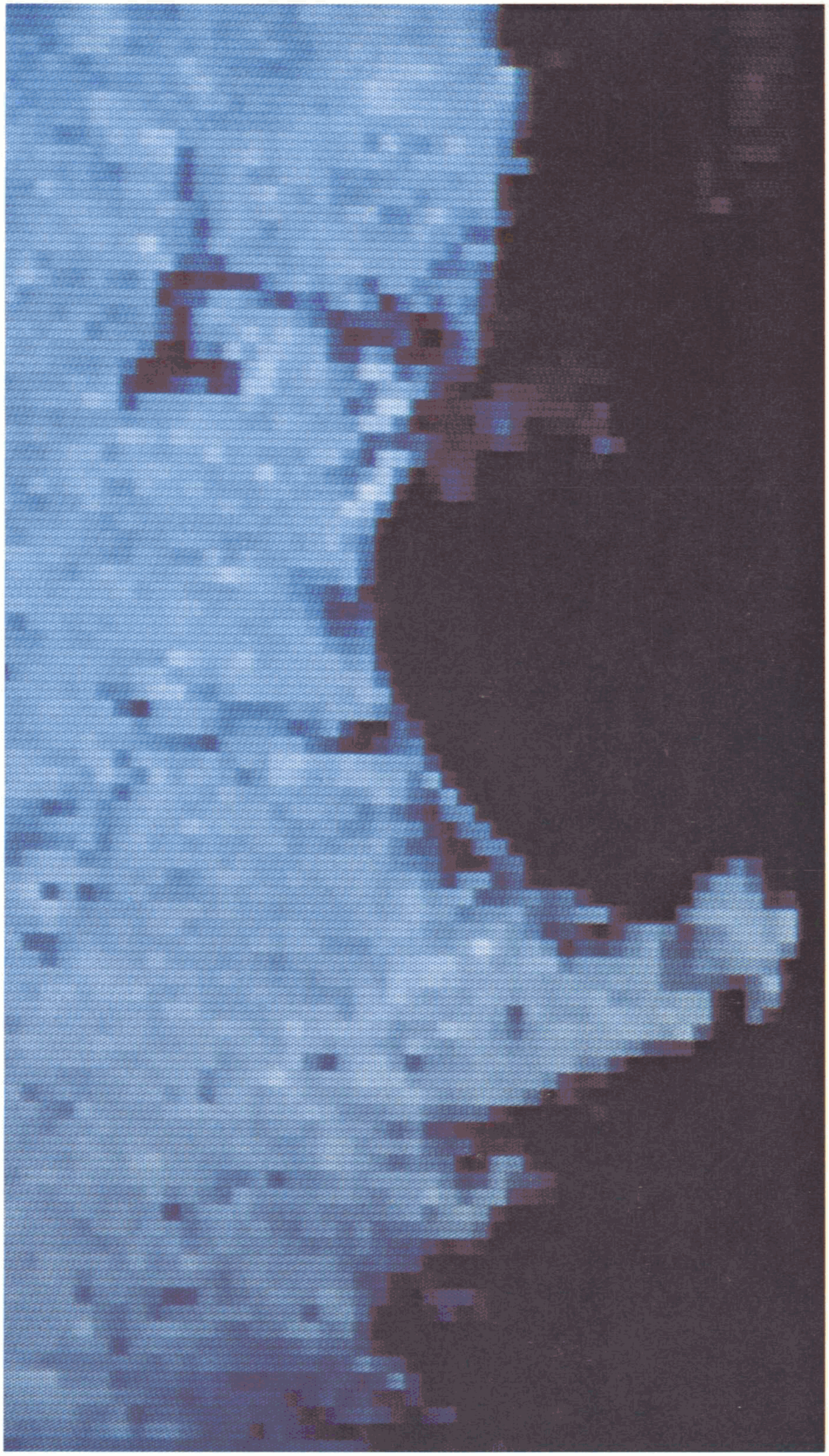
The visual repertoire of our virtual television screen, as well as other "slide collections" about Dataland, is markedly augmented by optical videodisc backup. Either side of this laser-scanned disk can hold up to 54,000 individual frames. Any frame can be individually addressed (freeze frame) or arbitrarily selected (random access), or sequences of frames can be exhibited forward or backward at varying rates. The optical videodisk as storage vehicle has a dramatic cost impact on SDMS, inasmuch as the player costs less than $2500.

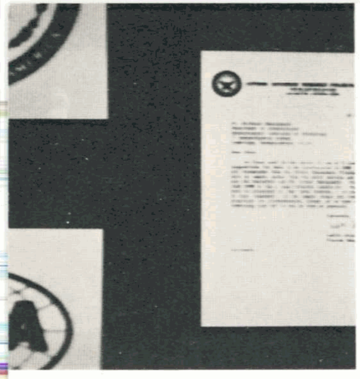
In a more general sense, the television "set" is an instance of any of a number of virtual "screens" which can reside in the geography of Dataland. The type of visual symbology or imagery surrounding any such screen indicates its usage: a "drive-in screen for movies," for example. What we are discussing here is a general capacity for "filmic" types of data, live and recorded, which are addressable in Dataland in user-designated locales.

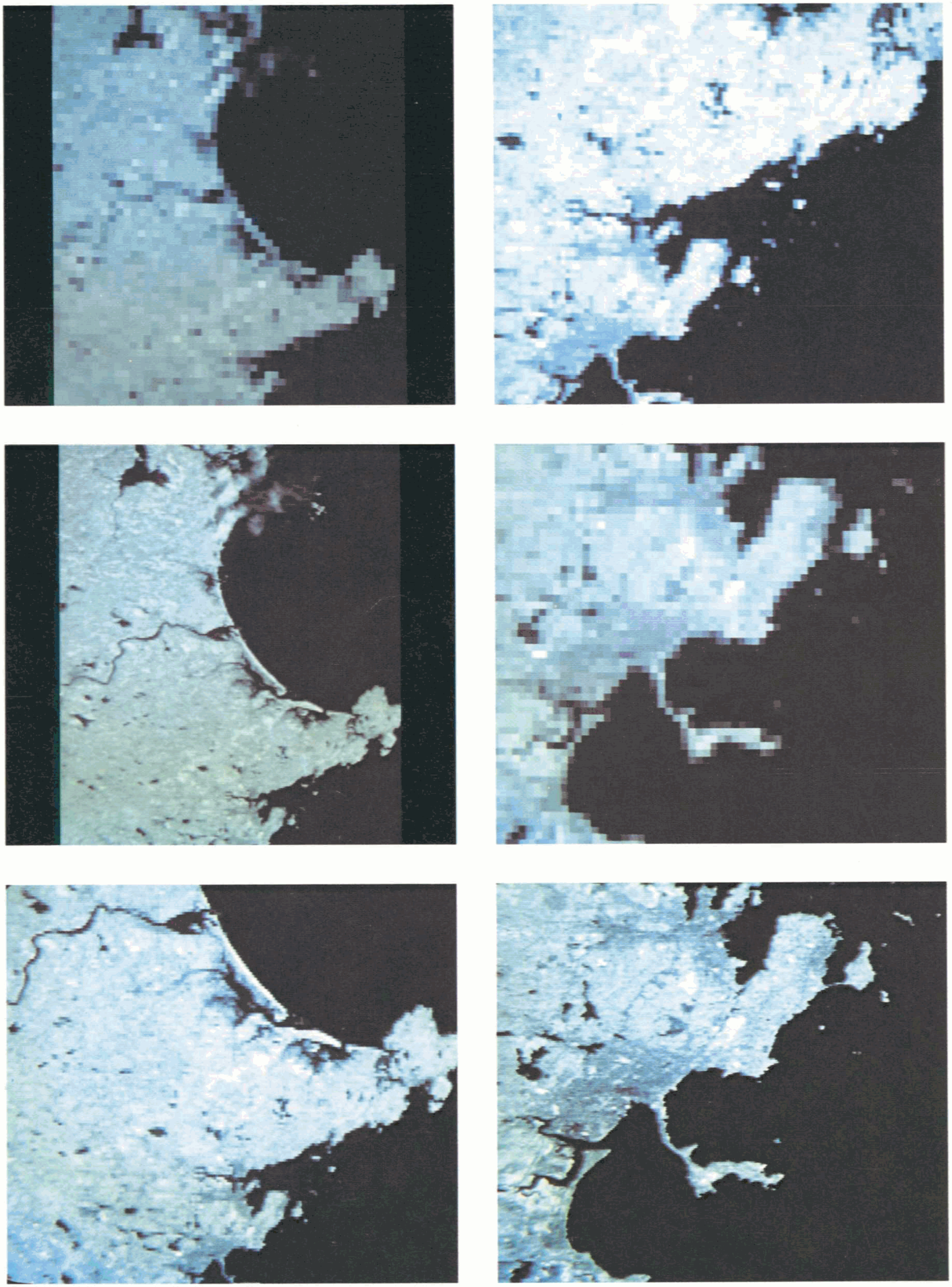
One of the most dramatic visual experiences on our experimental Dataland surface is that provided by LANDSAT photographs of about four thousand square miles of the New England coast, from the Boston area up into Maine.

In the northwest corner of our experimental Dataland, we have an image of the northeast sector of the United States, with an inset area of the New England coast. Upon traveling across Dataland to this patch of New England, it is possible to zoom in on any area.

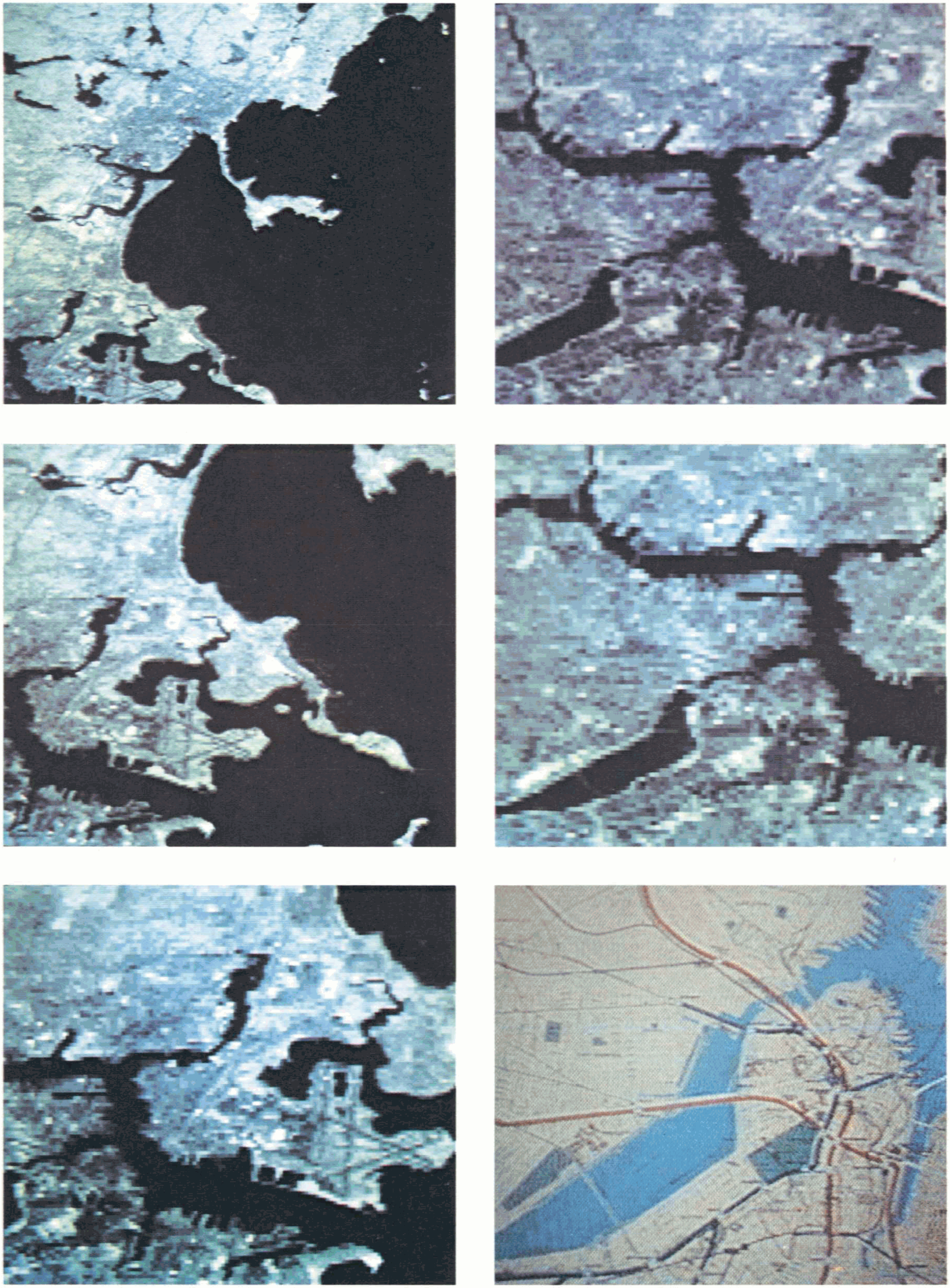
The perusal of the LANDSAT image has its own appropriate key map. It takes the form of a view on the key map monitor of, in this case, the entire sector of New England. The "key" consists of a highlighted rectangle moving around this key map image, tracking the course of the joystick-guided view visible on the large screen. Like the world map view, it offers direct teleportation about the LANDSAT image through touch.





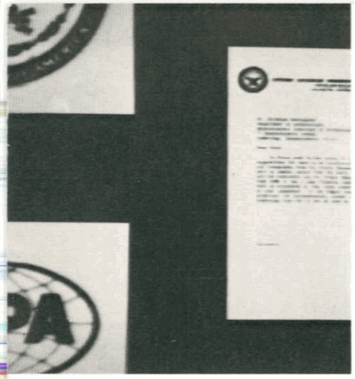


Zooming-in upon satellite image of Boston: In these views, approach is from the north over Cape Ann and Marblehead area. In first column, top-to-bottom, blocky image is replaced with finer resolution view (middle), while image expands and travel proceeds southward. In top and middle panels of the second column, approach continues over the peninsula of Nahant, with yet another image replacement (bottom).



Nearing Boston, the triangular runway patterns of Logan International Airport come into sight at lower left (left column). Finally, central Boston and Charles River basin appear (right column, top and middle).

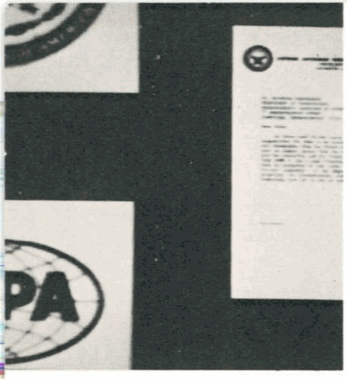
To this point, all images are retrieved from magnetic media. The final image, a map, replaces the satellite imagery; it comes off the optical videodisc.



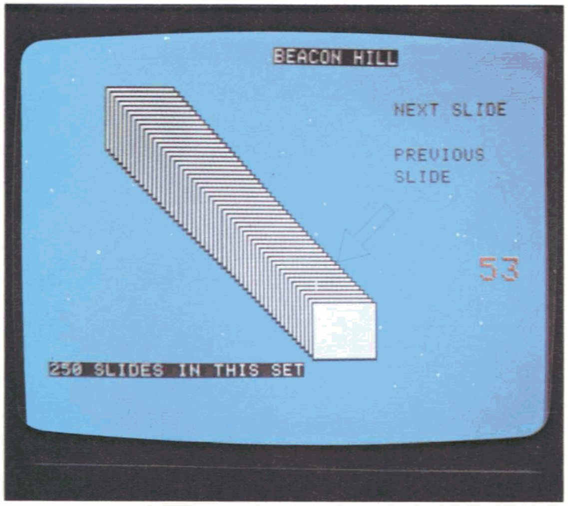
Zooming-in upon Boston.

As the user zooms in by a certain amount, the increasingly blocky material is replaced by a higher-resolution image in perfect registration. This replacement action is performed a number of times as one comes in on any portion of the New England area covered by the LANDSAT image and gives a dramatic impression of vast areas being contained within the confines of Dataland. In the sequence of pictures shown (coming over Marblehead toward Boston), the final limit of resolution is the resolution at which the original LANDSAT material was photographed and digitized. Each point of source data in the stored image represents about an acre in area.

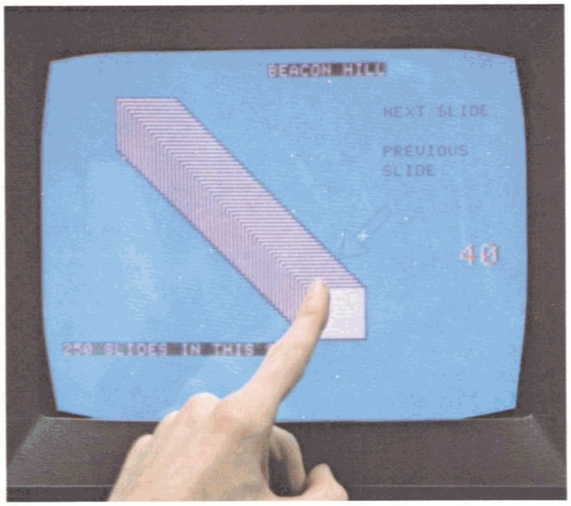
Should we zoom in on the city of Boston, we would find that the LANDSAT image of Boston is replaced at a certain level or altitude of approach by a *map* of Boston in the same size and register as the LANDSAT image it replaces. Then, as we zoom in on the Beacon Hill area of Boston, the map image suddenly gives way to a set of slide views of the central city area. The slide collection contains over 200 slides of Boston as seen from various vantage points. There is a key map which aids the user in perusing the collection of slides.



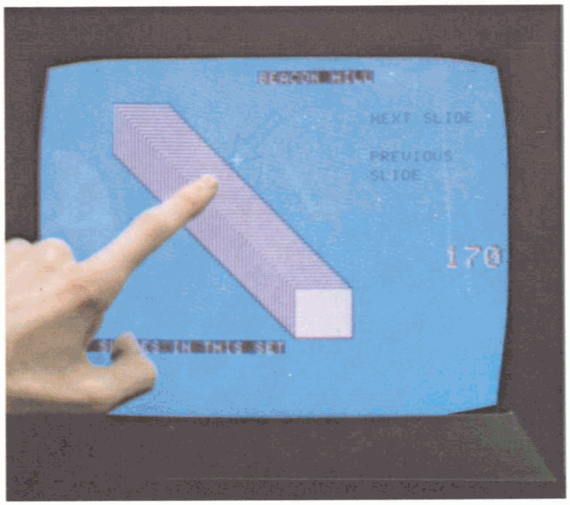
Close-up of slide collection key map used to access videodisc still frames. Region along spine of stacked "slides" permits sliding of cursor arrow by fingertip. Alternately, the captions "Next slide" and "Previous slide" can be tapped for one-by-one viewing, or continuously touched for rapid sequential showing.

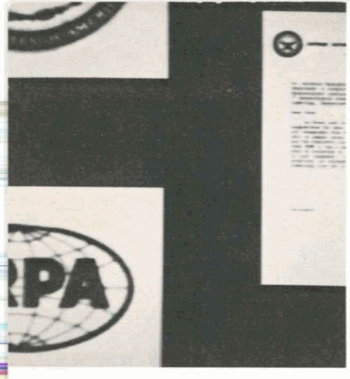


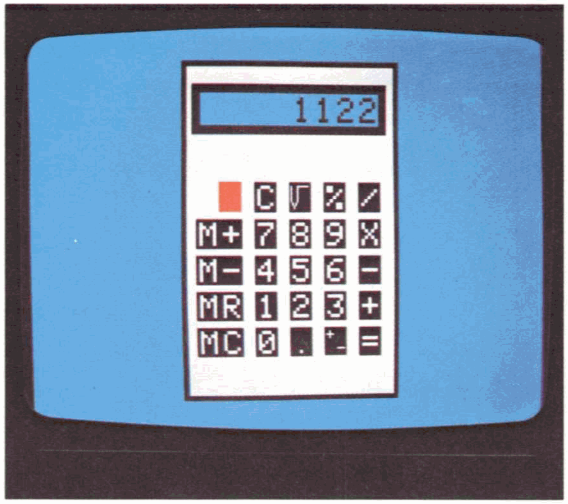
The interactive aspect of this key map is that the user can slide a colored-arrow "cursor" along the top edge of the diagonal row of "boxes" representing slides about Boston available to be seen. The analogue for more conventional media is that of remembering a slide tray. Moving the cursor up and down causes slides to be flashed in succession on the large screen as indicated by the finger touching and moving the display cursor.



Perusing slides of Beacon Hill area of Boston: slides 40 and 170 are shown being selected out for viewing.







The "active process," in this case a calculator, is illustrated by the photographic image on the large screen. The user can avail himself of a closer, touch-sensitive version, which can be invoked as a virtual machine. Other examples might include: a telephone, a dictaphone, even a television.

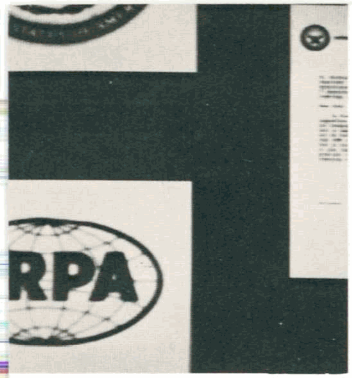
To recapitulate, the first key map was the "table of contents" of the "book," allowing the user access to selected sections of the text. The second type of key map, exemplified by the time-counter dial associated with the "television set," enables user interaction with what are in essence processes over time. The third type, the LANDSAT image key map, functions with respect to the LANDSAT image as the "world view" auxiliary monitor functions with respect to Dataland at large: a "you-are-here" marker, responding to user joystick navigation about a space. The key map associated with the "slide collection," residing at the map of Boston within the New England LANDSAT image, exemplifies the fourth type, used to manage data types which are essentially characterizable by "number": some ordered series of items, any one of which the user may dwell on for some unspecified period of time.

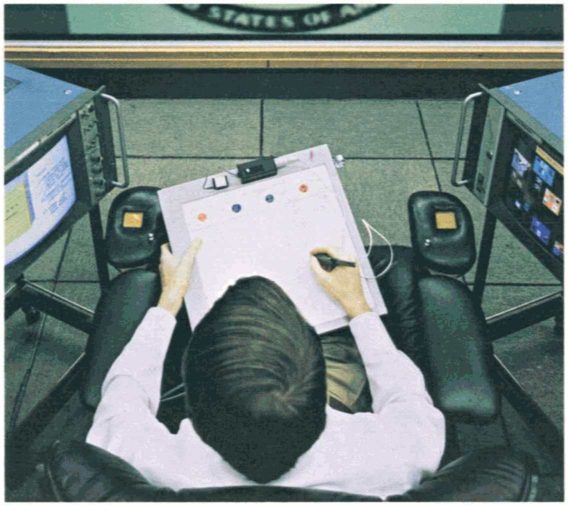
We expect the repertoire of key map types to grow along with enlargements of the repertoire of data types, as new data types require ordering by some principle not accounted for by any previously defined key map.

Further among data types in Dataland that depart from the usual are those that are *processes*. That is, SDMS's Dataland has nodes on its surface, indicated by appropriate glyphs, which make certain functions available to the user.

A specious example is the existence in Dataland of the image of a hand-held calculator. The user can zoom directly up to it. There then appears on the small key map monitor a copy of this calculator image. The user can directly touch the image surface on the monitor to input arithmetic problems, just as with an actual calculator. This particular example of a process, perhaps trivial or "cute," points the way to what will eventually be an interesting and powerful class of data types.


21.4476
C %
M+ 7 8 9 X
M- 4 5 6 -
MR 1 2 3 +
MC 0





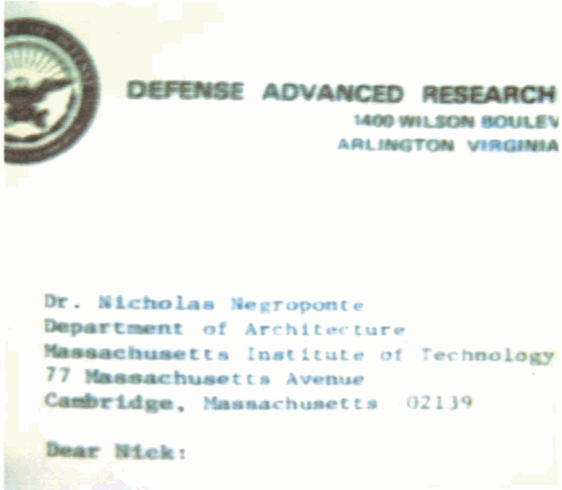
When making graphical annotations user enters commentary and jottings with lapboard and stylus.

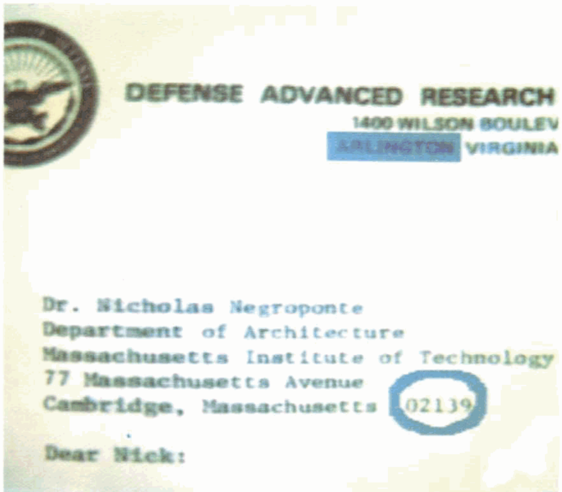
Noted was the idea that "perusing" implies the scrutiny in detail of items. This scrutiny in detail as an interactive process surely will trigger in the user the impulse to annotate, to append jottings, to comment "marginally." The traditional properties of computer media typically frustrate this impulse, one so readily indulged in on paper.

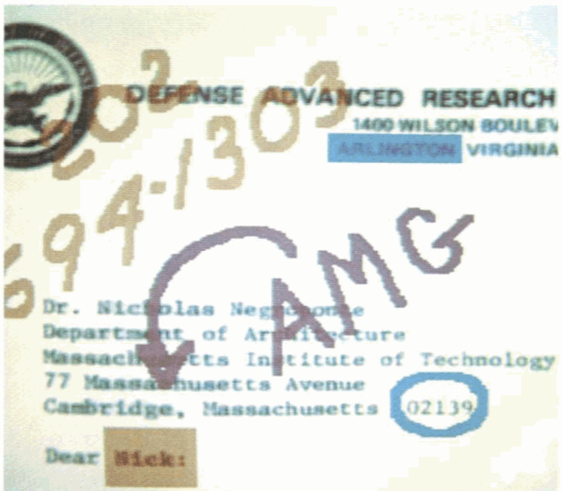
As an antidote to this shortcoming, we have introduced the capacity for making written annotations in "transparent ink," of varying line widths and colors. The ink, because transparent, does not obliterate material overwritten, and the notations made can be transient or permanent as the user wishes.

Notes and commentary can be not only in written form, but auditory as well. The user can employ the stylus and tablet to indicate an "anchorpoint" on the Dataland surface for to-be-recorded voice notes. The notes are played back subsequently upon coming within some critical range of the site of the recording.

Sound can be itself a data type as well as a vehicle for annotations or navigational aid. An instance of sound as data in our initial SDMS prototype was a digitally stored speech of Jerome Wiesner, President of MIT. A portion of this speech was activated and played back upon approach to an image of Dr. Wiesner, the volume increasing as one drew nearer by zooming in, or fading to the left or right as one withdrew from face-center toward the right or left, respectively. The speech sound was, in effect, synchronized with the location of the speech.



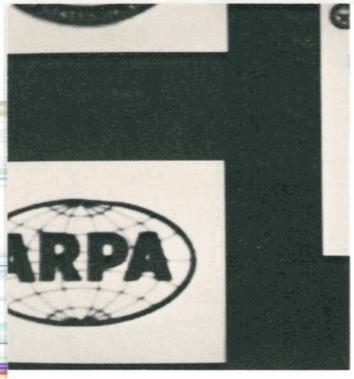




An octaphonic sound capability permits the creation in the media room of a veritable "cocktail party" of sounds and voices, any sound so re-created as to give it a spatially definite apparent source in the room. Endowing sound sources with specific locales effectively creates a set of "channels," aiding separation and identification of a number of simultaneously heard voices. The possibility of "browsing" through sets of auditory messages, redundantly cued in visual graphics, can make sound as a data type or means of annotation a rich offering to the user.

The ensemble of data types and associated interactions enumerated here form an initial set and by no means exhaust the long-term potentialities of user-system interaction. Before we turn to a consideration of what the future may hold, however, let us set down some of our major findings thus far about using space to organize data.

"Ink" is transparent, and does not obliterate material beneath. Note ink color options, and color-strip "highlighting" over address in letterhead and in salutation. Ink can be temporary, permanent, or placed in logically separable layers.



What We Learned

Diagram of spatial structure of Dataland in SDMS I, our initial SDMS prototype.

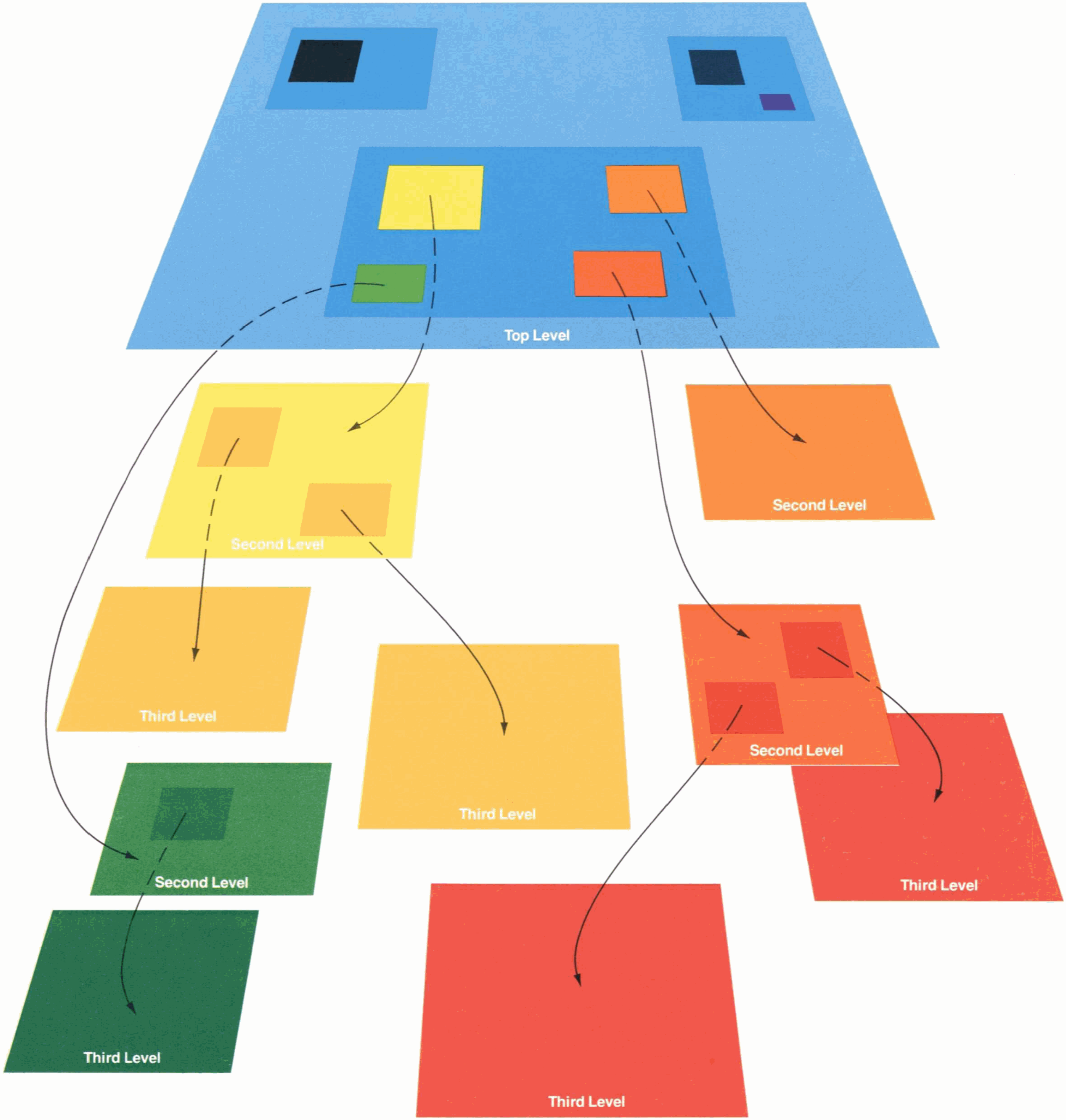
A major lesson of our experience in attempting to use space to articulate information was the virtue of straightforward simplicity in the presentational aspects of the user's "virtual" spatial world.

The system as outlined above did not spring into existence all at once but resulted from many years of Office of Naval Research research in television-based computer graphics and an initial version of a spatial data-management system, SDMS I, developed over the project's first year of DARPA support.

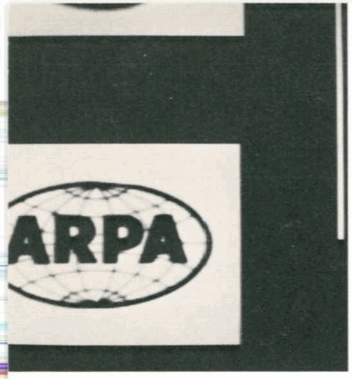
SDMS I was a "2-1/2-D" spatial world: a set of 2-D data planes, so linked as to define a hierarchical, "laminar" space. The top level of this structure was Dataland itself: a large data plane with "items" of various sorts arranged on its surface, some grouped on common background colors. As described earlier, the large screen in front of the user functioned as a "window" onto Dataland, which appeared in its entirety on the small monitor.

Importantly, the scheme of the data space in SDMS I included a system of "ports" on any data plane or layer, which, when zoomed in on, would give way to a next data layer, seemingly in back of or "beneath" it. In the accompanying illustration, we can see "ports" leading from the top Dataland layer to different, second-level data planes. One of the second-level data planes in the illustration is a *terminal* node, but the other second-level data planes themselves have ports, each leading down to third-level data planes. The information contained on the lower planes was logically subordinate to the topic at the port through which the user dove.

Travel around this system was up and down, chaining through the hierarchical structure of the nodes (ports) on the various data planes. One could travel deep down into a certain data structure and then "chain back" through the structure to get back to the point of entry.



Top Level
Second Level
Second Level
Third Level
Third Level
Second Level
Second Level
Third Level
Third Level
Third Level



Our SDMS I system was programmed to include up to twenty ports in any plane, and the "depth" of the structure in levels was in principle indefinite. As a practical matter, in our demonstration version the deepest "depth" in chaining down a hierarchy was about seven levels.

The spatial structure of SDMS I, however, generated a frustrating design paradox for any comprehensive data-world image for display on the "world view" monitor.

Although a number of candidate schemes for mapping views were advanced, including quite ingenious neural net-like schemes, some serious drawback seemed always to attend them. Those that were visually simple and easy to interpret were so at the expense of information content. Those that were comprehensive became "baroque": the more accurately they portrayed the rather complex paths that some user might take through a data space consisting of nested planes, the more the mapping view violated the principle that a display system be visually straightforward and easily interpreted.

The resolution of the dilemma was the insight that both the finding of items and the navigation in data space could best be served by a simple, single-plane Dataland. This sort of space seemed to offer all the primary benefits of locatability to be derived from the SDMS principle without having all the disadvantages of complexity that a laminar space might entail.

Thus, in our second prototype, SDMS II, the nested laminar space of SDMS I was replaced by a single, very large data plane. When approaching an item through the joystick "zooming" action, the user does not "pop through a port" down to another level of data space to obtain further information. Rather, once the viewer has come up close upon the item, the action from his or her point of view is then to "peruse" the item *in place,* as it were, with respect to the greater space of Dataland. That is, the perusal of any item occurs without any sense of further "travel" in data space of the sort implicit in going down through ports to deeper layers of data.

As a result, in SDMS II, *locating* an item, and *perusal* of that item once located, are, logically, two independent activities. In SDMS I, these activities were confounded each with the other, in that as you looked at more information, you also popped through down and deeper into "location space" as well.

What about an SDMS in 3-D, with 3-D "flight"?

At our project's inception, the possibility of creating a spatial data-management system in interactive 3-D space in the manner of a flight simulator was broached. The idea was an intriguing one: flying in and around a space of data items under the guidance of a variety of aviation-inspired navigational aids. For practical reasons, the idea was not pursued: acquisition of real-time flight simulation hardware, with visible surface portrayal of a 3-D world in color, would have far exceeded cost parameters for our project. But, apart from costs, there are factors intrinsic to dynamic 3-D flight to be considered in relation to the ends to be achieved in an SDMS.

While flight in simulated 3-D space can be dramatic and compelling, to the extent that 3-D flight is convincingly simulated, then to that extent the possibilities are expanded for becoming confused in that space. Spatiality, which in the SDMS context is supposed to be in the service of managing data, can then itself become so rich in directionalities that it generates sets of navigational issues which in their turn demand resolution.

Beyond the issue of user orientation in real-time 3-D flight is the problem of the recognition of objects in a dynamic 3-D world. The ability to approach any item in a 3-D space, from *any* angle, can seriously compromise the ready indentification of that item. Things simply do not look the same from different angles. The perception that a certain item is the "same" when seen from varying aspects (perceptual "shape constancy") is robust, to be sure, but it can and will break down.

Not only does the internal structure of a figure affect its recognition, but the "top," "bottom," and "sides" as determined by the viewer's relative orientation tend to condition how it is seen, and whether it is recognized or not. For example, even a 2-D silhouette map of Africa turned 90 degrees is much less readily identifiable as when seen in its usual north-south orientation (Rock, I., "The Perception of Disoriented Figures," *Scientific American,* January 1974).

To have all items always "face front" independent of the user's relative path of approach in flight is not necessarily a cure for the object recognition problem. The appearance of items taken *individually* is only a part of perceiving where you are and what you are looking at. Inter-item relational patterns, a significant contextual basis both for recognizing things and for navigating in "neighborhoods" of things, become effectively *different* patterns when a grouping of items is approached from different directions.

To summarize this point: the special virtue of a spatial data management system lies not in a sense of space and flight for their own sakes but in the invocation of a *sufficient* sense of space to activate the native, highly evolved capacity of humans to organize their perceptual experience spatially. The optimal space for an SDMS, then, is one that prompts spatial "schematizing" in the user, while minimizing the potential for disorientation.

Experience indicates that the majority of guest users trying out their hand at "flying" across Dataland for the first time invariably depend on the auxiliary view to navigate, using the large screen for "confirmatory" glances that they have arrived where they set out to go. The apparent aspect of Dataland is so large that this auxiliary view is vital for successful navigation, even for the constant, well-practiced user who may have, over time, gained a very detailed mental picture of the "lay of the land."

It turns out to be the case that navigation over Dataland with the joystick is so eminently and inherently "response compatible" that the novice user gains proficiency in flight in well under a minute. One of the prime attractions of such a system is the readiness with which the new user becomes a competent driver.

Accompanying the rapid acquisition of "piloting" proficiency is rapid assimilation by the novice of the graphic imagery of the Dataland surface in a "spatial" way. Almost immediately, the new user begins to refer to items by describing them as "the letter to the right and north of the green square," or "the yellow rectangles in the southeast corner." Although, of course, this way of regarding data is intended in the system, the beginning of discussing data in this mode by novice users is gratifyingly spontaneous.

Beyond sound as a data type and as an annotating modality, our work with sound as a "navigational aid" has built an appreciation of its utility for guidance and orientation, as well as for its enriching function as a redundant cue for what is going on visually. Those occasions when our own experimental sound system has been temporarily down have convinced us that we would be loath to return to a totally silent data world. Conclusion: sound, well-orchestrated, is crucial. A spatial data-management system must have it.

While we learned much concerning specific design features of our SDMS effort, realizations also emerged about further potentialities of the spatial theme. Some of these indications for future directions are set forth in the following section.

Future Directions

Progress in SDMS during our first two years of pursuing spatiality as a *motif* has encouraged us to chart a prospectus of next steps that is advisedly extreme in terms of positions taken.

Future directions elucidated include those of: topical spatiality, gaze-contingent displaying, eye-gesture coordination, voice-driving with eye contact, exploiting the user's near-field, user position sensing through a space-sensored "lab jacket."

The idea underlying "topical spatiality" is simple: the topography of Dataland is organized on the fly, as a function of user selected axes. These axes are chosen by the user on the basis of what data dimensions are judged appropriate for some specific task. In other words, unlike the current arrangement, pictograms are not stored and left in one place, but form part of a changing 2-D space.

As an example, consider data typical of a company's personnel file. Pictures of employees would reside on secondary storage media, and Dataland would have an indicative but blank array of rectangles. The user might then specify his horizontal axis to be a measure of age and the vertical axis as something like years of service, salary level, or managerial responsibility. Once selected, the pictures would be ordered on the fly in this 2-D space, with overlap and clustering, most probably. The data would be searched with the knowledge that Mr. Smith (the company's founding father) is somewhere at upper right and the delivery boy is in the southwest.

We propose to implement a topical SDMS, within the existing system, using building types. The selection of data comes from our 54,000-different-frame videodisc that was made in collaboration with the MIT architecture slide library and contains about 20,000 slides of American Architecture. At this writing, those "frames" are additionally described in a data structure which includes state, county, city, architects' names, dates of construction, and twenty building types. The data lend themselves to traditional spatiality (i.e., they can be viewed on a map of the US), but they also fall onto nonspatial axes like date of construction, physical scale, degree of public access, and the like. In the case of date and size (*X* and *Y*, respectively), if we limit our selection to post-Civil War, we can look for the World Trade Center in the northeast and Grant's tomb at the lower left.

This theme will include unclustering procedures that provide the proper changing, nonlinear axes while the user zooms, in order to "spread out" overlapping data. What we should learn are lessons concerning people's ability to create mental spaces and then to search them.

A critical focus of further research involves a set of themes having to do with the active engagement of the user in the environmental space of the interface, expanding the spatial domain to encompass the real-space characteristics of user actions. The overall intention is to render the system cognizant of how the user is operating in his personal space, as this space intersects with the shared space of the interface: directionalities in looking, speaking, touching, gesturing, together with a system awareness of user position and proximities.

Among these themes is the awareness of where the user is looking. Eye contact is an important adjunct in human interpersonal exchange and, in general, a prime indicator of the focus of active awareness.

The opportunity here is for, among other things, an eye-controlled "virtual camera," that is, a display that points, pans, zooms, travels, as a function of where the user is looking. In the context of the imagery in SDMS so far discussed, possibilities might include even the trivial example of a spatially dispersed set of items, perhaps faces, each of which "talks" when eye-addressed by the user, the directed gaze used by the system as an implicit "command" or act of selection. Another example would be an image, say a map, which locally expands and gains detail as a function of the point and duration of "looks".

The sense of space in and at the interface includes system awareness of the user near-field. The tone of appointments in the media room, especially the openness of the chair, invites the user to get up out of the chair the more immediately to engage himself with the data. Gesturing, arm waving, pointing, especially in conjunction with line-of-sight indications captured through unobtrusive eye tracking, would endow the immediate space between user and system with communicative meaning.

Space-ranging techniques that permit the system to know where the user *is* in the space of the interface, plus the user's body attitude, have intrinsic uses. A system of body-borne space sensors, unobtrusively disposed about a sort of "lab jacket" as buttons, cuff links, or epaulettes, would establish a computer-interpretable dot figure demarking the momentary position and attitude of the user. Among other functions, such knowledge can help to re-establish tracking contact of an errant oculometer, making sensing where one is looking a more realistic proposition in a dynamic environment.

With body sensing in the user near-field, pushing items of virtual data about in a now-tangible 3-D space arises as a possibility. Or, traffic-patrol-like direction of items streaming by in patterns for on-the-fly perusal. The dynamics of orchestration, literally like those of conducting a symphony, arise as realities, especially pertinent in an environment of spatially distinct sounds and voices in an octaphonic surround. Pointing to virtual talkers with indications to speak, to hurry up, to pause, becomes repertorial for the user, now "fleshed out" at least minimally in dot form in the augmented space of the interface.

Voice direction with eye contact, a natural combination in the extra system real world, persist in the realm of user-system interchange as a viable modality. The ability to voice-activate takes on more precise meaning and possibilities in conjunction with the ability to address by voice and eye simultaneously in a coordinated fashion. To be able to point by gesture as well defines yet an additional dimension of voice-and-eye dynamics.

These themes, somewhat synoptically expressed, implicate actions in an explicit space, coextensive with the real room space at the user-system interface, a shared space which can be artfully integrated with the virtual spaces which have been the focus of this, our work, so far.

At stake here is a user-system interface, a data space, so enriched and quickened that the human becomes more than a dimensionless point, the space at the interface something other than empty volumes. The long-term goal is the utmost in usability, where the user's natural modalities and capacities for communicative expression meet with perceptive, complementary system response.

Printed in the United States of America

Library of Congress Number 78-78256

Work Sponsored by
Defense Advanced Research Projects Agency
Office of Cybernetics Technology
Command Systems Cybernetics Program
Contract number: MDA903-77-C-0037
Contract period: 1 October 1976 through 30 September 1978

Work Performed by
Architecture Machine Group
Massachusetts Institute of Technology
Principal Investigator:
Professor Nicholas Negroponte

Distribution:
Unlimited

Key Words:
Management Information Systems
Computer Graphics
Man-Machine Studies
Communications Media

Printing:
This volume has been printed in an edition of 4000 copies. Supplemental funds for printing have been obtained from:
Office of Naval Research
IBM Corporation
Tektronix Corporation
Bell Northern Research
MIT Industrial Liaison Program

Photography by Christian Lischewski
Designed by Betsy Hacker,
MIT Design Services
Composed in Helvetica Light and Medium by Typographic House, Inc.
Printed by Rapid Service Press
Bound by Star Bindery