Sound in Hypermedia

for

Geographical and Environmental Interpretation

and

Analysis

by

R. Ingvar Elle

A RESEARCH PAPER

submitted to

THE GEOSCIENCES DEPARTMENT

in partial fulfillment of the

requirements for the

degree of

MASTER OF SCIENCE

GEOGRAPHY PROGRAM

April 1994

Directed by

Dr. Jon Kimerling
# TABLE OF CONTENTS

INTRODUCTION ........................................................................................................ 1  
DEFINITIONS ........................................................................................................... 2  
HISTORY .................................................................................................................. 3  
   Multimedia........................................................................................................... 3  
   Hypertext............................................................................................................. 4  
   Object-oriented programming ........................................................................... 5  
HYPERMEDIA, MULTIMEDIA and the GEOGRAPHER ........................................ 8  
   Drilling Down .................................................................................................... 9  
   Alternative Information Access ...................................................................... 11  
   Multiple Views .................................................................................................. 13  
EXAMPLES ............................................................................................................. 14  
   Hypermedia and Multimedia in Collaborative Planning ............................. 14  
   Hypermedia and Biodiversity Visualization .............................................. 16  
   Multimedia GIS for Dynamic Spatial Analysis ......................................... 18  
DATA TYPES .......................................................................................................... 18  
   Video .................................................................................................................. 18  
   Animation .......................................................................................................... 19  
   Sound .................................................................................................................. 19  
      Nature of Sound ............................................................................................ 20  
      Sound Types ................................................................................................ 21  
      Human Voice ................................................................................................ 22  
      Environmental Sound .................................................................................. 22  
      Abstract Sound ............................................................................................. 23  
         Loudness ..................................................................................................... 24  
         Pitch ............................................................................................................ 24  
         Register ...................................................................................................... 24  
         Timbre ......................................................................................................... 24  
         Order ............................................................................................................ 25  
   Sound in Computers .......................................................................................... 25  
      Digital Sound Production and Storage ..................................................... 26  
      MIDI ............................................................................................................... 28  
      Data Driven Sound ......................................................................................... 29  
      Synthesized Voice ........................................................................................ 30  
DISCUSSION ............................................................................................................ 32  
   Legal Considerations ....................................................................................... 32  
   Educational Effectiveness .............................................................................. 33  
   Cost Effectiveness ............................................................................................. 33  
   Accuracy and the Friendly Front End ............................................................. 34  
CONCLUSION .......................................................................................................... 35
INTRODUCTION

This paper explores the use of hypermedia in geographical and environmental analysis, focusing on the use of sound. After providing essential definitions the paper presents a brief history of multimedia and hypertext (the combination of which lead to the term hypermedia) and object oriented programming. The concept of the hypermap and how spatially related data can be arranged in a hypermap for browsing of spatial features/information is considered next, along with the need for means of data retrieval other than the "drill down" technique. Several examples of hypermedia and multimedia applications that have been used in environmental interpretation and analysis are presented next. In the review of these examples the focus will be on the interface design and on multimedia data types that were used to convey important information.

After these examples, the paper will more deeply explore how different multimedia data types might be incorporated into hypermedia applications. Because multimedia can involve such a variety of media types, including high quality graphics, video, animation and sound, there is not enough room here to discuss them all in great detail. The reader will therefore find a short section on the use of video and animation in environmental and geographically related applications, but the focus here will be on sound.

The use of several sound types in hypermedia, including the human voice, environmental sound and abstract sound will be covered. Following this, examination of sound storage and different types of sound generation in computers. The paper ends with a look
at some of the potential problems involved with the use of hypermedia in geographic and environmental interpretation and analysis.

DEFINITIONS

We must first distinguish between the words multimedia and hypermedia. Multimedia allows the user to combine sounds, video, photos and text together in a computer.

Hypermedia is defined as a, "form of multimedia in which associations between information or ideas are identified by predefined, user-activated links." (Tilton and Andrews, 1993). As such, hypermedia offers the user the ability to "click" on words to get immediate definitions or to click on words or strings of words or pictures to get references to related ideas or to directly view related concepts textually, pictorially or acoustically. More relevant to the field of geography is that hypertext links can also be used to link hierarchically ordered levels of spatial information, and to link multiple views of physical features in applications referred to as hypermaps (to be discussed in detail later).

A brief discussion of the evolution of hypermedia components of multimedia, hypertext and object-oriented programming is in order.

HISTORY

Multimedia

Multimedia computing requires the ability to store large amounts of data that are generated by sound, high quality graphics
and video. This first became possible in the late 1970's with the development of the videodisc. Thirty minutes of motion video or 54,000 images could be stored in analog form on one side of these optical discs. Projects at MIT showed the power of integrating interactive systems with videodiscs, digitizer boards and graphics (Mackay and Davenport, 1989). Since this we have seen the development of the CD-ROM which, although it has been derided for being both slow and "klugey," (Hood, 1994) can hold as much as 700 megabytes of information and is the most cost effective digital storage medium (Fox, 1991). It is also being constantly upgraded to run faster (Hood, 1994).

In addition to high volume storage media, compression techniques have been and continue to be developed to store larger amounts of data. Compression techniques, although they can degrade image quality (Doyle, 1994a), have emerged and are standardized. These include JPEG and MPEG (Fox, 1991), (Le Gall, 1991) (Rosenthal, 1994).

In 1991 engineers from Apple Computer and SuperMac Technology developed primitive software compression/decompression (codec) algorithms that compressed digitized video. This resulted in a sophisticated file format called Quicktime. Prior to this time computer video technology was available but the addition of boards required to play the video was expensive. The Quicktime extension was simply dropped into the Macintosh and played. This lead to a more widespread use of video in the computer (Doyle, 1994b).
At first Quicktime movies could not play at many frames per second and were barely the size of a postage stamp, but the Quicktime architecture is open, which means that new compression schemes can be added as they are developed (Davis, 1993). This open architecture also means that Quicktime movies can be played on different platforms. New accelerator cards have also increased the quality of Quicktime movies so they can now be played at a 640-by-480 pixels with 24-bit color and 30 frames (with 60 fields) per second (Doyle, 1994b). Although Quicktime has been cited as the best means of playing video, several other formats exist that also offer cross-platform capability (Adams, 1992).

Support for sound in computers has varied between platforms. The first Macintoshes supported 8 bit sound. The development of 16 bit sound cards has increased the quality of computer sound. Mac's have offered 16 bit sound since the SE30, but cards must be added to PCs, which otherwise will not support sound at all (Burger, 1993). The quality of the majority of sound boards has been questioned (Howard, 1993) although with time and market pressure, this quality will undoubtedly improve. More will be said on sound in computers later in this paper.

Hypertext

The hypertext concept (i.e., the idea of providing easy access to related pieces of written information through user accessed links) was first published in 1945 by Vannevar Bush,* who conceived of a fairly elaborate device which would be built into a regular office

* The term "hypertext was actually not coined until 1965 by Ted Nelson.
desk. It would include a photographic plate to copy documents and two video display screens for viewing. The index of copied documents and the documents themselves would be stored on separate pieces of microfilm. As the user read one document he or she could link a part of it to another document on the other screen. The user could build a series of links and then store this series in a code for later use. Bush's machine, which he called the Memex, was never built. (Horn, 1989).

The first hypertext system was built by Doug Englebart for McDonnel Douglas beginning in 1962. Englebart is also credited with inventing the computer mouse, on-line integrated help systems, and several other firsts in the computer field (Van Dam, 1988).

Although educators and others have worked to develop usable hypertext systems since the late 60's, hypertext did not begin to catch on until the introduction of Hypercard, which was unveiled at the Mac-World Expo in San Francisco in 1987. Bill Atkinson, who developed Hypercard along with several others at Apple computer, insisted that the application be given away free with every Macintosh, which helped to greatly popularize the concept of hypertext and hypermedia (Horn, 1989).

The power of Hypercard is that it gives the "nonprogrammer" the ability to apply some basic aspects of object oriented programming to produce powerful applications.

Object-Oriented Programming
Most attribute the development of object oriented programming to Simula-67 and Smalltalk-80. However, as Thomas
(1989) points out, early exploration in object oriented programming was done by diverse groups within the computing community.

It is easiest to understand object-oriented programming by looking at its basic elements: objects, encapsulation, message passing and inheritance. Technically, an object is described as something that can not be broken into smaller parts. A message is any instruction that is passed from one object to another. One example illustrating these two concepts is to imagine two separate graphics stored in a computer, one of a house and another a room in the house. The pictures, which can also be called objects, can be "linked" together by a message that can be sent from the house picture to the room picture telling the room picture to display itself. In this case, the user might click on a window of the house that belongs to the room he wants to see. The message can be imbedded in the house picture by writing a simple script for the "button" that will be "hidden" behind the room window. Because the objects can only be operated on by messages, they are considered to be encapsulated.

Inheritance is another, slightly more difficult, concept of object oriented programming. Simple scripting languages like Hypercard do not contain this and so are not considered truly object-oriented by some (ibid). Inheritance can be a tremendous code saving device for the programmer, because it works on the notion that different objects can share many attributes. Therefore, when developing similar objects, the only extra code that is written is for those aspects of the object that are not shared.

Object-oriented programming can be used at various levels by programmers. Simpler languages, such as Hypertalk, used in
Hypercard, are referred to as authoring languages or fourth generation languages (4GLs). Grayson (1993) points out three advantages over more difficult third generation languages (3GLs) such as C++. First, a command in a 4GL does more work than a command in a 3GL, so there is less code to write. Second, 4GLs often come with interactive interfaces that make testing the code easier. Third, 4GLs are subjected to stricter tests at run time so that errors such as writing or reading from incorrect memory are decreased. In addition to these points, as Bottaci and Stewart (1992) point out, a major benefit of these languages is that they raise the level of abstraction for the programmer. In other words, the programmer is working in a language with syntax and modelling concepts that are closer to the English language and the real world, respectively.

Authoring languages therefore are becoming extremely popular, as they allow "non-programmers" to make powerful presentation and educational applications (Gray et al. 1993). Various multimedia authoring tools are constantly being developed and refined so that the description of any particular product may be soon outdated. A variety of authoring tools are available. Andrews and Tilton (1993), and Kindleberger (1993) describe some of these. The reader would also do well to look at one of the many multimedia and computer magazines such New Media, Multimedia, MacWorld and Windows. While authoring tools are constantly being updated and new tools are being produced, DiBiase (1994) and Andrews (1994) have stated that Macromedia Director is emerging as a de facto authoring and animation standard for the CD-ROM market.
More powerful object-oriented programming languages, such as C++ and LISP, have been used for a wider variety of applications, including the construction of object oriented databases for geographic information systems (Kainz, 1993). The comparison of object oriented databases, (used in GIS systems such as ONTOS and Zenith), with relational systems (such as ARC Info) are beyond the scope of this paper, although Egenhofer and Frank (1992), Macguire et al. (1990), Joseph et al. (1991), and Crosbie (1993) all help to shed light light on the topic. For this paper it is enough to note that both types of systems very recently or in the near future will support multimedia data types (Kemp, 1992) (Dangermond, 1993).

HYPERMEDIA, MULTIMEDIA AND THE GEOGRAPHER

Hypermedia and multimedia are becoming popular with geographers as a means of integrating the large variety of spatial information that can be obtained from satellite photos, video photography, GIS, GPS and statistical analysis (Armenakis, 1993) (Kindleberger, 1989) (Lewis, 1989). This integration can take various forms in geography. Tilton and Andrews (1993), for example, have used hyper- and multimedia to produce a CD-ROM with a collection of famous and historical maps. Raper and Green (1989) have used Hypermedia to develop a GIS tutorial, and Linsey and Raper (1993) have used it to develop a friendly front end to ARC-Info. This paper will first focus the discussion of hypermedia and geography on the realm of the hypermap and move to the broader use of multimedia for geographic and environmental interpretation and analysis.
One of the most basic concepts in hypermedia applications is the use of links to allow the user to navigate between hierarchical levels of spatially related data, a process that Dangermond (1993) has referred to as “drilling down.”

Wallin (1990) describes the power of drilling down. He carries to hypermedia the concept of the territorial concern, which is not two-dimensional and so must be viewed from more than one vantage point for the viewer to gain an adequate understanding, and in the process produce a hypermap. Therefore, the cartographer must utilize the power of hypertext links to connect different spatial scale representations in the system. At lower levels, the degree of generalization is great so the only information conveyed is that of geometric shape. The lowest, termed level 0 by Wallin, may be represented by a small scale map of the entire region, which typically may be the highly generalized national or state map. As one moves up the levels the data representation becomes less abstract and more concrete. At level 7 one data structure may be a close-up photograph of the inside of a museum or even an electron micrograph of a micorrhizal fungi. By adhering to this highly intuitive method of organizing spatial features by levels of hierarchical abstraction, Wallin believes that the widely noted problem of “getting lost in hyperspace” (Conklin, 1987) can be avoided. It is important to note that the above mentioned number of levels was specific to a GIS development project that Wallin was involved in at the time of writing. The actual number will vary...
according to the purpose of the particular hypertext applications. It should also be noted that the base map of a hypermap does not have to be a map. For example Forer (1993) used a satellite photo as the base document and Woosley (1991) used a photograph.

Lindholm and Sarjakoski (1993) argue that for the hypertext concept to be most powerful the links must be between "features and not map sheets." Instead of providing a series of maps they emphasize the structural context of how objects relate to each other. They define context as a "set of links which impose a certain structure on a set of nodes." They distinguish between the vertical context, which can be divided into the functional and classificatory context, and the horizontal context. The functional context is formed out of inherently hierarchical units, such as political boundaries, which would progress from state to county to city. The classificatory context relates to zones of otherwise homogeneous units that are broken down into subunits based on progressively finer distinctions. An example is soil zones which are divided into subzones. The horizontal context refers to relations between features of equal importance, such as two counties. The user should be able to move freely between the vertical and the horizontal context of objects.

Laurini and Milleret-Raffort (1990) have devised a system for coding information between these levels of abstraction. Pyramid hierarchies are used to connect different levels of spatial abstraction and r-trees are used to deal with horizontally overlapping areas of concern. These r-trees are also used to connect the spatial areas to written documents that may be present.
Both Wallin's and Lindholm and Sarjakosky's methods of designing information in a hierarchical structure have been used extensively in atlases. Lindholm and Sarjakosky (1993) and Huffman (1993) discuss examples.

Alternative Information Access

In and of itself, the drill down concept is useful for linking hierarchical levels of spatial information and possibly, as Laurini and Milleret-Raffort (1990) propose, for providing access to written information relevant to spatial features. But with hypermaps the issue of alternative types of data access quickly arises, especially as the user becomes more accustomed to the system and wishes to more quickly access information.

Andrews and Tilton (1993) have noted that in an electronic environment users should not be forced to go to the information, as is the case in hypertext browsing, but should also have the option of having the information brought to them. In other words, a variety of query opportunities should be offered. Halasz (1988), in reflecting on the deficiencies of hypertext systems, notes that "navigational access by itself is not sufficient."

Forer (1992) illustrates several alternatives to the simple drill down for information access. He has developed a system for local area analysis which he refers to as a "type three system." The type three system can be distinguished from less sophisticated interactive cartographic display systems (types 1 and 2) in that it offers the opportunity for in depth analysis, and includes the use of GIS for overlay analysis. Figure 1 illustrates the components of the type 11...
three system. The reader should note that hypertext links are only one of the tools that Forer uses to organize and retrieve data.

Figure 1: Components of a Type 3 System for Multimedia Mapping

Export ← Mapper ← Analysis

User Workbench

Viewing Tools

Organisers

Data Objects

Maps, Photographs, Animations, Video, Data Matrices, Sound files, Text files, DTP documents, DTM etc

Tools for Selective Capture

Cameras, data loggers, spectral scanners, surveying, prose, drawing, tape recording, questionnaires etc

The Environment

(from Forer, 1993)
Multiple Views

So far, the fast linking or "hyper" function in hypermedia has been discussed. The other half of the term, of course, refers to the multimedia component. Multimedia is particularly useful to geographers in that it gives them the option of providing multiple views of geographic features using different media. For example, Forer (ibid.), Figure 2, has constructed a graph showing the relative advantages of using different types of media to relate aspects of a woodland.

Figure 2
Benefits of Alternative Representations of Woodlands.

<table>
<thead>
<tr>
<th>Shaded Area Map</th>
<th>Simple Symology</th>
<th>Satellite Images</th>
<th>Air Photos</th>
<th>Video</th>
<th>Sound</th>
<th>Text</th>
<th>DTM Drapes</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Shaded Area Map" /></td>
<td><img src="image2" alt="Simple Symology" /></td>
<td><img src="image3" alt="Satellite Images" /></td>
<td><img src="image4" alt="Air Photos" /></td>
<td><img src="image5" alt="Video" /></td>
<td><img src="image6" alt="Sound" /></td>
<td><img src="image7" alt="Text" /></td>
<td><img src="image8" alt="DTM Drapes" /></td>
</tr>
</tbody>
</table>

Key
- ● Particular strength exhibited in this area
- ○ Has capabilities in this area
- □ Limited strengths in this area
- ❌ Weak in this area
- ★ Denotes impact of DTM when combined with other data types.

(from Forer, 1993)
Hamilton and Flaxman (1991) have pointed out that multiple views are valuable for bringing together the wide variety of disciplines that may come into play in the environmental analysis process.

EXAMPLES

The following examples will help to illustrate how hypermedia, by utilizing its intuitive interface and multiple views, can be used in environmental interpretation and analysis.

Hyper- and Multimedia in Collaborative Planning

Shiffer (1993) describes a multimedia project developed by the Planning Support Systems Group at MIT's Department of Urban Studies and Planning. The project, called the Collaborative Planning System, examines various ways that planners might use multimedia to visualize different physical aspects of proposed planning projects. The system (Figure 3) includes a front end that uses Aldus Supercard for interactive manipulation and Apple's Quicktime for video. The system has been cleverly applied to land use analysis, automobile traffic analysis and visual environmental assessment using various combinations of sound, video, and numerical graphical representation.

For example, one module allows the user to point to a street on the map and have traffic levels displayed with different media types, including a video of the area, a bar graph showing the traffic level in cars per day, and the traffic sound, which is played at the true level of volume.
Figure 3:
Concept Diagram of Hardware Setup and Data Flow for a Collaborative Planning System (CPS)

(from Shiffer, 1993)
Hypermedia and Biodiversity Visualization

Hamilton and Flaxman have developed a hypermedia application for environmental interpretation and analysis (1991). In a project spanning six years, the authors assembled an extensive set of biodiversity data including information on genetic diversity and natural history, as well as spatial and temporal data relating to biological diversity.

Data types include a variety of sounds, images, video and surface representations. The highly flexible interface provides different ecological views, modes of data analysis and levels of user interaction, and will be described in detail here.

The interface, which was written using Hypercard, provides several spatial/ecological views: global, landscape, ecosystem/community, and species. The project area and features were put into their global view context using GPS derived coordinates that were attached to features on UTM coordinate system maps using custom external commands. At the landscape level the data are spatially referenced and graphically depicted using GIS. The ecosystem/community level view is shown using ground level images of the ecological communities. The species view is depicted by images of the specific species of interest.

Four "functional modes" were developed: Navigate, Interpret, Compare and Analyze. The user can use the Navigate mode to examine both large and small scale images associated with a certain point, as well as to enlarge or shrink graphics. In addition, because
many of the ground level images are comprised of collections of multiple overlapping frames, the user can "look" in any desired direction.

The Interpret mode can be used to select objects in an image or movie and then obtain text and sound files specific to the object.

Using the Compare mode the user can choose any two of the four ecological views described above and manipulate them at a higher resolution of information and graphics content.

The Analyze mode gives the user a set of software techniques called a "toolbox". One of these, the "link manager" tool, allows the user to follow pre-compiled links between objects and to add his or her own links. The "image processor" tool provides a library of object measurement and image manipulation functions. The "Spatial Analysis" tool allows the exploration of raster format map data for spatial processing. Finally, the user can export numerical data to spreadsheet and statistical analysis programs (Hamilton and Flaxman, 1991).

The system also provides four levels of user interaction which can be used in accordance with the user's level of familiarity with ecological processes. The Naturalist level is meant for school children and "non-science oriented" college students. Hence, the interface emphasizes acoustical and visual factors of biodiversity. The second level, Resource Manager, is designed to provide land managers, planners and "resource management focused users" with easy database access. The Ecologist level gives a wide range of analysis and representation tools, which include data entry forms. Finally, the Data Administrator user level allows addition of new features to the
interface and gives the user the opportunity to design personalized development tools.

Multimedia GIS for Dynamic Spatial Analysis

Fonseca et al. (1992) have developed a prototype multimedia geographic information system for dynamic spatial analysis. The authors use simple cellular automata rules adopted from Camara et al. (1990) to illustrate the movement of fire on the landscape. When the model is run, these behavior and interaction rules act to drive the movement of pictorial representations of the natural features. With the use of synthetic video, the simulations can be represented on aerial photographs and maps. Sound is used for background stereo and to help illustrate point scenes.

DATA TYPES

Having reviewed various hypermedia applications, I would like to now focus on some of specific data types used in these presentations. In my research for this paper it quickly became apparent that a full discussion of the various multimedia data types is beyond the scope of this paper. However, I will note what I consider valuable articles on the use of video and animation for hypermedia applications.

Video

use of video for technical presentation of scientific data. Sheppard (1989) offers an extensive discussion on visual simulation for planning. Finally, Burger (1993) offers tips for video production and provides a number of references regarding video for multimedia production.

Cartographic Animation

The amount of literature on cartographic animation has grown sharply in the last few years. The following articles discuss cartographic animation in a broad context. The reader will find sufficient references to more specific animated cartography applications in these articles. Campbell and Egbert (1990) provide a discussion of the past thirty years of animation in cartography. Dorling (1992) suggests ways to combine animation and cartography. Gersmehl (1990) discusses various metaphors for cartographic animation. Karl (1993) outlines a comprehensive approach to cartographic animation. Finally, MacDougall (1992) discusses how dynamic graphics, data analysis, and statistical visualization can be applied to georeferenced data.

Sound

In the above examples of hypermedia systems, sound has been used for various purposes. I will now examine the physical nature of sound, point to different sound types, and show how they might be used in hypermedia presentations, and how they can be stored and generated in a computer.
Nature of Sound

Sound can be described as, "oscillations of air pressure that stimulate the eardrum and, by extension, the auditory nerves and the brain." (Burger, 1992). One of the most basic properties of sound is its frequency. Frequency is measured in cycles per second, or hertz. The average person has a maximum hearing range of 20 to 17 KHz. As we age this range may drop to around 15 KHz. The perception of frequency is referred to as pitch.

The largest musical measurement of pitch is the octave. Increasing a pitch one octave will produce twice the frequency of the original pitch. Conversely, taking a pitch one octave lower will produce a frequency that is one half the original pitch. The octave in Western music is subdivided into 12 half-steps. These half-steps correspond to the notes found on pitched instruments. Half-steps can be divided further into 100 smaller increments called cents.

Register refers to the relatively high or low frequency of a noise or a discrete sound.

Pitch is, of course, relative. Therefore a method of standardizing pitch has been developed. This standard, which is recognized internationally, is that the A above middle C on the piano is tuned to 440 Hz. This pitch is known as concert pitch and the note is called A-440.

Amplitude is a function of sound pressure and manifests itself as loudness. It is measured in decibels. One decibel is considered the smallest change in intensity that a trained ear can distinguish at a frequency of 1000 Hz. Zero decibels is considered the softest sound.
that a person can hear and is therefore established as the threshold of human hearing (ibid).

Related to amplitude are primary temporal elements of a sound. Attack is the amount of time that a sound takes to reach its highest loudness. Sustain level is the amplitude that maintains a sound at its constant rate. Release is the amount of time a sound takes to fall from sustain level to zero.

Timbre can be described as the tone color or "quality" of a sound. In the discussion of timbre it is necessary to distinguish between discrete sound and noise. The former represents a single fundamental frequency whereas the later is a combination of such a high number of fundamental frequencies as that any particular fundamental can not be distinguished. The timbre of sound, either a discrete sound or noise, can be altered by changing the wave form of a sound and or by adding harmonics to the fundamental frequency (Blattner et al., 1990).

Sound Types

Various types of sound can be used in hypermedia. These include the human voice, environmental sounds and abstract sound.

The Human Voice

In hypermedia we can take advantage of the psychological familiarity with human voice in several ways. Recalling that hypertext links can be used to obtain definitions for unfamiliar
words or terms, it is possible to provide these definitions in spoken as well as written form.

Voice can also be used, and is often preferred, in helping to acquaint the user with the interface (Walker, 1990) and to explain complicated changes taking place in the data (Monmonier 1993). The addition of voice as an aid to interface familiarity will vary according to the user's personality type (Cory, 1994) and his or her level of familiarity with the system (Marchionini, 1988). Therefore voice, as all sound in a computer, should be "turnoffable" (Ambron, 1990).

Narration can be used in both live and animated movies. Mckibben (1993), in describing a project in which GIS layers were animated to show predicted growth in the sewer system in Los Angeles, stresses the need for a professional narrator. Scalleti (1991) capitalized on our familiarity with the human "voice" by mapping the sounds of coughs to illustrate the increase in the amount of ozone in the air over time. Davenport et al. (1991) point out the need to log the specific attributes of a film sequence in order to provide potential users with ability to use the sequence in a variety of contexts. In the previously described project by Hamilton and Flaxman, voice annotation was used to document the existence of ecological information.

Environmental Sounds

Sound in the environment has been the focus of several geographers. Schafer (1977) provides a profoundly philosophical historical description of how sound in the environment has been changed by population growth and technological development.
Similarly, Pocock (1989) provides a deep discussion of the different ways in which sound is perceived in the environment. Ohison (1976), using a more technical approach, examined the range of sound in the “soundscape.” Environmental sound can be an important part of a hypermedia application. Hamilton and Flaxman collected over 2 gigabytes of digital bioacoustical recordings. The sounds were used for bioacoustical documentation, behavioral records, and to help “spatialize” the observation. Suchan (1994) has noted that traffic noise, which is often negatively perceived and is the most ubiquitous noise in the soundscape (Porteus and Mastin, 1985), can be a major concern in the planning process.

Abstract Sound As Hamilton and Flaxman have shown, hypermedia and multimedia cartographic display can be used for different knowledge levels. The inexperienced user may settle for pictures with accompanying natural sounds or stereo music, whereas the more expert subject may require more advanced methods of examining data. For the latter, sound can and has been used to help more closely visualize complicated geographic phenomena. For a detailed discussion of the use of sound in geographic visualization see Krygier (1994). This is borrowed from his section on abstract variables in listing their possible uses for visualization. The reader is advised to read his article for a more expanded treatment of the topic of sound in geographic visualization.
Loudness

Because loudness is naturally ordered from low to high or vice versa, it lends itself to representing ordinal data. Abrupt increases in loudness that punctuate a relatively low noise level can serve as notification of a change in state.

Pitch

Pitch can be used to represent ordinal information. Scalletti (1991) observed that the most effective mapping of a variable where small changes needed representation was done with frequency. She also observed that high frequency sounds were more irritating than low frequency sounds. In addition, it was noted that subjects tended to associate low frequencies with larger and older things.

Register

As Krygier (1993) suggests, register can be used as a "broader quantitative distinction" than pitch. He provides an example in which he has constructed an interactive display to illustrate four different variables. A choropleth map is used to show different percentages of the population, by county, in the labor force. A graduated circle map is added to display median income. To prevent visual clutter he uses a single pitch in three separate octaves to represent a "drive to work index." He then adds a fourth variable, representing percent poor, by using a range of pitches within each register.

Timbre

Probably more than any other sound variable timbre has a distinctive personality and can therefore be effectively used as a nominal variable. For example, a low rumble may signify an ambling
bear, whereas a crisp tone may connote a small animal darting about. The potential problem with this type of data mapping is that it can appear very value-laden, which can be both an asset and a deficit.

Order

Krygier mentions order as a means of representing the passing of time. He provides an example (Krygier, 1993) of where low and high pitches are used to connote early and late time periods, respectively. This display, however, is hard to identify with and could possibly be improved by using exact spoken dates instead of different levels of pitch.

Combination of Variables

Fisher (1994) has developed an application for helping to visualize error in remotely sensed data (Figure 4). The variables he uses include pitch, duration, silence, and modulation. While each variable can be used separately, he has pointed out that it is possible to use sound variables in combination to get a single measure of reliability. For example, he associated long silence, long duration, and high pitch with good reliability. Conversely, short silence and duration and low pitch were associated with poor reliability short silence and duration.

Sound in Computers

Having covered various types of sound that may be used in hypermedia applications, it is now necessary to discuss various means of generating sound in computers.
Digital Sound Production and Storage

Sampling is the term that refers to the process of capturing digital representations of sound. The optimal sampling rate is dictated by the Nyquist equation, named after Harry Nyquist, an engineer who in the late 1920's theorized that a signal should be sampled at least twice as often as the highest frequency in that signal. Because the average human hears up to 20KHz, the standard for high quality CD sound has been set at 44.1 KHz (Nillson, 1993).

Note that 44.1 KHz is more than double the 20KHz limit.

If the frequency of a recorded sound is more than half the rate of sampling, a process called aliasing kicks in. The sound gets converted to a lower frequency, according to its harmonic values.
The result is poor sound. Therefore, low pass filters may be used to cut off any frequencies higher than half the sampling rate. Low pass filters require their own "headroom", which is why CD sound is a little more than twice the 20KHz limit (ibid).

Analog sound is, of course, continuous. The process of dividing the audible spectrum of analog sound into digital values is referred to as quantization. A computer stores sound values in binary form. In the case of CD quality sound 16 bits, or two bytes, are used, to give a range of 65,536 different values. Because two bytes are needed for each value and because as many as 44,100 values are recorded for each second of high quality monophonic sound, 176 kilobytes of storage is required for each second of high quality stereo sound (ibid).

There are several ways to reduce the amount of space needed. These include 1) using mono instead of stereo sound 2) taking fewer samples, which reduces the range of frequencies covered 3) using fewer bits per sample, which reduces the quality of the waveform reconstruction (Fox, 1991). These techniques are often used for sampling human speech, since high quality reproduction of the human voice is often not needed. Therefore recordings for the purpose of annotation are often sampled at 8 bits and 11KHz (Ettore, 1990). Sound compression is part of the MPEG format, which provides a compression ratio of 6 to 1 (Fox, 1991).

Many different formats for digital sound have been developed, so when using prerecorded sound it may be necessary to first convert the sound to a specific format. Inexpensive sound conversion software is available in the sound directory in the
Michigan archives of Macintosh freeware and shareware. This can be accessed through Internet Gopher servers at info.hed.apple.com. The reader will find software that can convert to and from formats used in a variety of different computer platforms including Unix, Amiga, DOS and Mac. In addition to sound conversion software you will also want a sound editor for recording and sound manipulation. Sound Sampler is a freeware sound editor (found at info.hed.apple.com) and includes cut, copy, paste, fade in and fade out editing functions. One of the more powerful and popular commercial sound editors is Sound Edit Professional by Macromedia. It contains a powerful suite of editing functions including Fourier frequency transforms (Yavelow, 1994), which can be used to eliminate background noise (Flaxman, 1994). Hied (1993) offers tips for recording using Sound Edit Professional. Most authoring systems allow the user to script recorded sounds to play at the desired time. If sound is to be played synchronously with a video or animation, a system, such as Quicktime, for handling time-based data types, should be used. Music, while not discussed earlier as a sound type, can be used as background in multimedia applications. In 1983 a group of electronic music manufacturers united to develop MIDI (Musical Instrument Digital Interface). A microprocessor was installed in synthesizers, allowing the use of serial communications protocol designed specifically for these...
of digital sampling, does not rely on the storage of enormous amounts of digital data, but instead utilizes sequencers containing instructions that control when and how digital synthesizers produce sound. MIDI interfaces are available for virtually every personal computer.

The advantages to MIDI are 1) the orchestration of different musical instrument sounds and each performance of the sound can be easily changed while still under MIDI control 2) since only the notes for the performance are recorded, there is no loss of musical quality as occurs with many recording processes 3) perhaps most importantly the amount of RAM needed to store the MIDI instructions is miniscule compared to that for digital audio. For example a typical four minute song may take only 50 KB of MIDI data as opposed to the 40 MB that would be required for a digital recording (Burger, 1992).

Data Driven Sound

While MIDI is excellent for the production of music, it may not be ideal for abstract sounds used in scientific visualizations.

Scalletti (1994) points out that:

"The timbres on MIDI synthesizers tend to have strong associations in most people's minds. It is something like doing a visualization of a density distribution, but instead of using dots or spheres, you decide to use images of human heads. The emotional impact of the image is so strong that it might overwhelm the abstract density information that you are trying to convey."
Scalletti advocates a means of tying continuously controlled, data driven sound to sonic visualization. Scalletti and Craig (1991) describe various tools for data driven sound. While this software is constantly being updated, these tools, described below, provide a firm base of components to be possibly added to in the future.

**Mapper** - This tool maps numerical data to sonic parameters. The stream of data can be mapped to frequency, changes in frequency, amplitude, stereo position and others.

**Comparator** - This simply feeds a different variable into each speaker.

**Sonic Histogram** - Like a graphic histogram the sonic histogram represents the relative magnitudes of different elements. Each element is given its own frequency, and the "time-varying magnitude or multiplicity of that element is mapped to the amplitude of the corresponding frequency."

**Marker** - A marker provides the ability to play a sound once a certain threshold is reached in another sound.

**Synthesized Voice**

Synthesized voice offers an option to the authentic human voice, and offers an advantage in that it does not take up as much space as digitally stored human voice. The only space required is that needed to store the text of words that will be "spoken". A certain amount of RAM is also needed (see below). An explanation of how voice is synthesized is beyond the scope of this paper, and for those interested a good starting point is "Principles of Computer Speech" by Ian Witten (1982).
In 1984 Macintosh made available a sound synthesizer called Macintalk, that "spoke" words which had been converted to their phonemic representations. However, the voice had a, "Drunken Scandinavian in an drum" quality that made it only appropriate for a limited number of applications, such as children's games (Andrews, 1993).

Apple has more recently replaced the Macintalk synthesizer with the Speech Manager, which is available for anonymous ftp at ftp.apple.com and is located in the /dts/mac/sys.soft/speech directory. The speech manager comes with several different voices, some of higher quality than others. The highest quality voice takes up 2.7 MB of RAM. The Speech Manager itself requires 150 KB of RAM while the Plaintalk extension (which must be used in all other MACs except the AV MACs) requires 70 KB. In addition to playing higher quality voices, the Speech Manager does not take control of the CPU as did the Macintalk synthesizer, therefore allowing other applications to run while the voice is being played (Weidl, 1993).

Those wishing to demo the Speech Manager can obtain an application named So-To-Speak, which is specifically designed for that purpose. It is very easy to use and is available via Gopher servers at info.hed.apple.com. The Speech Manager voices can also be played on Hypercard using extensions available at the same address.

In addition to text-to-speech, computers are now being developed that can recognize voice. Although this is presently still not perfected it will undoubtedly lead to more flexible methods of interaction and data retrieval (Lu, 1994).
DISCUSSION

A wide variety of technical knowledge is required to develop multimedia applications. Some of the fundamentals, such as compression techniques, are discussed in computer cartography courses. Others are not, and hopefully, those gaps have been decreased in this paper. In addition, much of the fine arts knowledge needed for good multimedia design (Liebhold, 1990), including familiarity with color, map drawing skills, and perspective drawing are also taught in cartography classes. However, there a few potential barriers to the production and diffusion of hypermedia applications reviewed here.

Legal Considerations

The main legal considerations for the development of multimedia applications concerns two major areas. The first concerns the use of copyrighted media (Martin, 1994). This is a particularly treacherous area considering that many "sources" of media, such as newsgroups on the Internet, may purport to be "giving away" material that is already under copyright protection.

The second legal consideration concerns a patent filed by Compton Multimedia Corp that seeks to obtain 1 percent of all gross earnings on every multimedia application that is sold. The patent covers "search and retrieval of text, picture, audio and animated data, especially when delivered on CD-ROM" (Vaughn, 1994). While it has been speculated that Compton may not bother to enforce the patent (DiBiase, 1994), the implications for state governments
seeking to distribute multimedia data for sale are potentially enormous.

Educational Effectiveness

A primary justification for the use of multimedia data in hypermedia applications has been that it is easier to compete with media such as television. This "media-is-the-message" argument may be undermined by the fact that systems for geographic education may simply never be as exciting as video games. Kocho (1994) notes that educational CD-ROMs that are not highly exciting do not hold the attention spans of children accustomed to the high action of Nintendo games. Bove (1994) suggests that educational multimedia should indeed incorporate many of the adventurous aspects of video games to be most effective. Ormerling (1993) suggests that hypermedia geographical atlases should incorporate the kinds of interactivity found in games such as SIMCity and Carmen San Diego.

Cost Effectiveness

Cost effectiveness will be a big concern regarding the use of hypermedia systems for government agencies. Shiffer (1993) points out that to maximize the value of multimedia efforts, the data should not be designed for and stored in a single presentation, but should instead be "generalizable" and easily accessible for use in a variety of presentations. Nonetheless, many agencies, possibly perturbed over
the rapid and costly evolution of information system software and hardware (Van Demark, 1991) owing to new innovations and the increasing short lives of software and hardware standards (Nilison, 1993), may be hesitant to allocate scarce funds for hypermedia education and analysis.

Accuracy and the Friendly Front End

Members of the GIS community have expressed concerns regarding the integrity of multimedia data types (Hulce, 1993). The argument is that the "friendly front end" will in some cases be used to show flashy pictures and sounds that may act to sublimate concerns regarding data accuracy. There is no doubt room for concern. Much literature, for example, exists attesting to the potential inaccuracies in GIS data (Parent, 1989) (Craig, 1992) (Goodchild and Gopal, 1989). It is conceivable that these data, when turned into elegant animations or flyovers, would awe the viewer into neglecting data accuracy. However, it can also be argued that hypermedia can be used to help avoid errors. For example, systems for error detection using hypermedia have been developed (Buttenfield, 1992). In addition, hypermedia links can be used to offer fast access to metadata, for quick referencing. Finally, as Monmonier has pointed out, the potential for multiple levels and views offered by hypermedia can lead to a more complete representation than the traditional one map approach (Monmonier, 1991).

More insidious however, may be the urge to trade flash for substance. This temptation will undoubtedly increase with the
development of virtual reality software, which as an immersion experience in and of itself is very appealing. Care will have to be taken to make sure that data are being illustrated because they are meaningful and not simply because it is sexy or "so cool". Projects such as by Fonseca and Camara (previously described) in which dynamic spatial models are mixed with multimedia can perhaps be considered an appropriate use of virtual reality, as will cases in which dynamic flyovers are truly necessary. But such immersion will not always be necessary or appropriate (Lindholm and Sarjakosky, 1993), so the geographer should not forget the value of the simple regression line and bar chart that can be used very easily and effectively linked in hypermedia applications on most platforms (Nance, 1992).

CONCLUSION

In this paper I have looked at how hypermedia might help geographers to better analyze and interpret environmental phenomena. I began with a discussion of the development of multimedia, hypertext and object-oriented programming, the combination of which has lead to the advent of easy-to-use multimedia authoring tools. The fact that these tools are not standardized will mean that applications will not be viewable on all platforms. Just as the fax machine and VCR did not gain widespread use until standardized, hypermedia development will be limited until universal standards are reached and followed. Nonetheless, inexpensive authoring systems can be used to make electronic atlases and other hypermap-related presentations.
The use of freeware and inexpensive shareware such as that mentioned in this paper will help to minimize development costs. However, in the cases where images must first undergo a high level of processing, such as polygon overlay, expensive GIS software, such as Arc-Info, and more powerful hardware will be required.

The major benefits of hypermedia have been shown to be two-fold. The first is that the hypermap can act as a root document from which to drill down to more descriptive levels. But the simple drill down is often inadequate, especially for more advanced users. Therefore different methods of accessing information may be needed. These methods will in some cases be part of the software itself, such as the keyword search provision. In other cases they will need to be designed into the particular application. So as cartographers enter into the hypermedia production process, they will need to become skilled in developing interfaces that offer multiple and intuitive means of access to the information and that take into account the differing levels of the user's familiarity with computers in general and the content of the application domain in particular.

The multiple view benefit will enhance the way we visualize environments, and will help to bring together a variety of disciplines in the decision making process. In addition, the availability of different media types may help to reduce the cost of data acquisition and maintenance by offering media choices that involve less expensive capture and processing techniques.

Sound can be used in these applications in a variety of ways. The human voice, whether provided in digitized or synthesized form,
can be used to help familiarize the user with the interface and to put the material into context. The ability of the user to interact with the system, for example to have complicated concepts repeated, will offer a distinct advantage over traditional, noninteractive video.

In the future, with the perfection of voice recognition systems, speech may play a role in the retrieval of information. For example, the user may be able to ask for all of the images in the databases that correspond to a symbol on the hypermap. This will undoubtedly save the user from having to type the request, but it is important to note that the addition of the conversation metaphor will not add any intelligence to the system. In other words, the information that the application provides will still be limited by the robustness of the database and the complexity of the application algorithms. In addition, the environmental models, although they may be activated by human voice requests, will still suffer from the limitations posed by the inadequacies of the model itself.

The use of environmental sound will add a more holistic quality to the presentation, and will, in some cases, be valuable for visualizing sound related environmental impacts, such as the increase in traffic noise.

Abstract sound will pose a particular challenge to cartographers as it is for many an entirely new medium of communication. In addition, the choice of whether or not to use it will be guided not only by its relevance to the application but also by a mix of budget and time constraints. Nonetheless, it at least offers promise for advanced methods of visualizing data as well as errors in the data.
So we see that while sound may not add to the power of data retrieval process, and so can only help superficially with respect to information access, it may in fact be a major component of the multiple view side of hypermedia.

The production of hypermedia will be impacted by social and institutional barriers. Copyright restrictions will have to be heeded. These rules will restrict what can be added to hypertext nodes, in that they mean the difference between accessing a reference versus accessing the entire article.

The temptation to focus more on the medium itself may become stronger with the development of virtual reality software. Virtual reality will indeed have its place, but in other cases it will be inappropriate, in that so much of science is the process of abstracting information out of the landscape and much of this information can only be illustrated through statistical displays such as charts and regression lines. The fact that these statistical displays can be illustrated in hypermedia applications speaks well of the simple link.
References


