2010

A multi-modal interface for road planning tasks using vision, haptics and sound

Matthew Charles Newcomb
Iowa State University

A multi-modal interface for road planning tasks using vision, haptics and sound

by

Matthew Charles Newcomb

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Human Computer Interaction

Program of Study Committee:
Chris Harding (Major Professor)
Derrick Parkhurst
Eliot Winer

Iowa State University

Ames, Iowa

2010

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I'd like to thank my family for their support throughout my time at school, especially my late grandfather and father as well as my mother and brother.

I'd also like to thank the members of my committee, including Dr. Harding, who has always gone above and beyond what he needed to do to help me on my thesis. In addition, I'd like to thank Dr. Parkhurst and Dr. Winer for their helpful input at the earlier stages of this project.

I'd like to thank Dr. Cruz-Neira for introducing me to all the opportunities in the Virtual Reality Applications Center and the HCI program and also for introducing me to Dr. Harding.

I'm very grateful for the friends I've made here at VRAC, including Tyler Streeter, who has worked with me on so many class projects and always brings up interesting ideas; Dan Shipton and Tylor Husske who started our lunch ritual; and all the other students that have made working here more enjoyable.

Many thanks go to the great staff members who have been so helpful and to the great faculty here at the lab.

I'd also like to thank Beth Dunham for her editing expertise.

ABSTRACT

The planning of transportation infrastructure requires analyzing many different types of geo-spatial information in the form of maps. Displaying too many of these maps at the same time can lead to visual clutter or information overload, which results in sub-optimal effectiveness. Multimodal interfaces (MMIs) try to address this visual overload and improve the user's interaction with large amounts of data by combining several sensory modalities. Previous research into MMIs seems to indicate that using multiple sensory modalities leads to more efficient human-computer interactions when used properly.

The motivation from this previous work has lead to the creation of this thesis, which describes a novel GIS system for road planning using vision, haptics and sound. The implementation of this virtual environment is discussed, including some of the design decisions used when trying to ascertain how we map visual data to our other senses. A user study was performed to see how this type of system could be utilized, and the results of the study are presented.

CHAPTER 1 - BACKGROUND

Processing power and data storage are both growing at exponential rates, nearly doubling each year. This makes the task of viewing data and results much more difficult. Because there is a fixed limit to the processing capabilities of our visual senses, it has become increasingly important to find new ways of displaying and interacting with data via other sensory modalities.

When new pieces of infrastructure are planned, such as highways, railroads, pipelines or power lines, a Geographical Information System (GIS) is commonly used to work with the relevant geospatial data (Spear, 1998). In a GIS, this spatial data is typically viewed as overlapping layers. In the context of road planning that may mean looking at existing roads, developed lands such as cities, terrain elevation, and environmentally sensitive areas or landmarks.

1.1 Introduction to multimodal interfaces

Multimodal interfaces send and receive information through multiple sensory modalities such as haptics (touch) and sound. The very first MMI, Bolt's "Put That There" (1980), was a very simple speech and gesture recognition system. Modern examples of MMIs include cell phones, which can be set to different output modes like ring, which is a sound cue, or vibration, a tactile cue. Cell phones can also accept number pad input as well as speech input for programmed numbers. This introduction to MMI provides a basic understanding of some of the underlying theories that support performance improvement with the use of MMIs.

1.1.1 Theoretical background

The following sections describe some of the theories about how our minds process information, of which MMIs can be designed to take advantage. These theories suggest that the brain works as a parallel processor, which can be optimized by dividing up information into different sensory modalities.

1.1.1.1 Miller's magic number 7

According to Miller (1956), there is only a finite amount of discrete information "chunks" that we can keep in our immediate memory (i.e. the type of memory dealing with information stored for less than a second). He postulated that it could be possible to increase that amount by organizing data into different "dimensions":

> "The span of absolute judgment and the span of immediate memory impose severe limitations on the amount of information that we are able to receive, process, and remember. By organizing the stimulus input simultaneously into several dimensions and successively into a sequence or chunks, we manage to break (or at least stretch) this informational bottleneck." (Miller, 1956)

Note that Miller originally did not relate what he called different "dimensions" to different senses. However, others since then have successfully applied the concept of Miller's "dimensions" to different multimodal input channels. The results of a study conducted by Samman (2005) show that working memory capacity in multi-modal tasks surpassed the one-dimensional working memory limits by a factor of up to three. Subjects in the Samman study were shown stimuli and asked to respond in one of nine different ways. Some methods required verbal responses, while others required physical movement or spatially organizing some objects. There was an increase in the working memory capacity subjects when under most multimodal conditions. Although there appeared to be some interference between some of the nodes, overall there was an average threefold increase in working memory capacity under the multimodal conditions.

1.1.1.2 Multiple resource theory

The multiple resource theory (Wickens, 1984) suggests that individuals have several pools of resources that can be used. Previous theories described a single fixed resource pool that was used for working memory. The single resource theories failed to explain why task performance remained constant in a primary task when a secondary task's difficulty was increased at the same time. Multiple resource theory explains this because tasks can be split into multiple resource pools that act independently of each other instead of competing for resources from one common pool. Multiple resource theory also states that tasks which share sensory modalities should have a greater interference with each other. The classic example given is that it is very difficult to read directions while one is driving a car, but it is easy to listen to directions while driving (Wickens, 2002) because one task is visual and the other is aural, thus using different resource pools. Other results have shown that MMIs can help improve task performance by reducing cognitive load (Oviatt, 2004). MMIs can also help improve memory (Stefanucci, 2005). When subjects in the Stefanucci study were given a list of words to recall that were learned in different environments, the subjects were able to successfully recall more words that were learned in the environments that used more modes of input.

1.2 Haptics background

Haptics is the science of adding touch to human computer interaction. It allows users to feel, touch, and interact with virtual objects. The field of haptics is broken up into two major areas: tactile displays, which affect the sensory receptors on the skin and provide sensations such as vibrations, temperature variation, and pressure, and kinesthetic displays, which sense the body's joint positions, angles and muscle movement. Kinesthetic displays may also utilize input from the user, while tactile displays provide output only.

One of the most popular haptic devices commercially available, the SensAble PHANTOM (http://www.sensable.com), was the device used in this study. The Omni model can be seen in Figure 1.

Figure 1: SensAble's PHANTOM Omni Device.

The PHANTOM consists of a grounded robotic arm with a stylus on the end of the arm, giving the user a single point of interaction with which to touch and manipulate objects. This single point of interaction is visually represented by the proxy position in virtual space, which can be different from the true physical position of the stylus in 3D space. When the user interacts with virtual objects by making the physical position of the stylus penetrate an object, the visual (proxy) stylus point stays on the outside of the object and a spring force is applied between the proxy and the actual position that pulls the device outside the object again.

To obtain interactive frame rates in computer graphics the image must be refreshed anywhere from 30 to 60 frames per second. In comparison, the update rate for haptics requires an update speed of 1000 times a second. To achieve this update rate there is usually a dedicated haptic real-time loop to calculate the force and position of the haptic device and a second loop to perform graphics and other calculations that don't need to be performed in the real-time physics loop.

1.3 Sonification background

Sonification refers to the use of non-speech audio to display data. In order to represent changes in data, different sound properties such as pitch, volume, length of sound, and timbre, can be altered. These can be changed interactively, such as in this project, or can be rendered offline. Described below are some of the most common sonification techniques.

Parameter mapping is the most widely used sonification technique and is often the easiest to create. Numerical data within a certain range is directly mapped to a sound parameter, such as pitch. Different sound parameters can be mapped to different dimensions in data to provide richer information.

Earcons, as described by Blattner (1989), are sounds or tones, which have a meaning that can be combined with other earcons to have a more complex meaning. Verb and noun earcons can create of variety of different actions by sounding in various combinations. The problem with some earcons is that the sounds are often abstract, which requires learning by memorization.

Another way to use sound to represent data is by using auditory icons (Gaver, 1986). These are sounds that have a clear relationship to the type of data being presented. By using natural sounds, the meaning can be deciphered much faster and without the memorization required for earcons. These sonification techniques can be combined with graphics and haptics to provide more ways of representing data. Some systems that integrate haptics and audio with visual displays are discussed below.

1.4 Combining visual systems with haptic and/or audio cues

There seems to be a strong correlation between some audio and haptic characteristics. A study by Peeva (2004) found that the level of roughness correlated with the loudness of a sound when subjects were asked to match a given roughness to a volume and vice versa. There appears to be a similar, although weaker, link between roughness and pitch. A study by Emery (2003) on user performance 2D mouse-based drag-and-drop tasks showed that auditory cues improved performance when present, and haptic cues showed the best levels of improvement, especially when combined with audio. However, some tasks, like multi-modal texture recognition (Lederman, 2003), don't seem to show any clear benefits over unimodal cases. McGee (2000) tested the perception of surface roughness when using combined haptic and audio textures. Depending on the degrees of roughness and the sense (audio or haptic) with which it was presented, some combinations of multi-modal textures were helpful while other combinations produced interference. It is problematic to present haptic textures via a local, point-haptic device, such as the Phantom, rather than through a skin-pressure based, tactile device (Wall, 2003a). However, there seems to be some merit in using different friction values to distinguish between different values when using a PHANTOM (Wall 2003b). Crossan (2004) used haptic granular synthesis as an equivalent to auditory granular synthesis to combine visual graphs with a haptic expression of statistical uncertainty.

1.5 Traditional geographic information systems

Geographic Information Systems (GIS) are the tools used for storing, analyzing, visualizing, and manipulating different forms of geospatial data. This data can take the form of satellite images, zoning maps, digital elavation models, road data and more. Each of these data sets form a distinct layer that can be used to convey information about the geographical area they describe.

The image in Figure 2 shows how these layers of information can be stacked. When planning some new type of infrastructure like roads, pipelines, or power lines, many of these layers are used to determine the most suitable placement. There have been several systems that experiment with adding audio and/or haptics to a GIS. Most of them use force-feedback

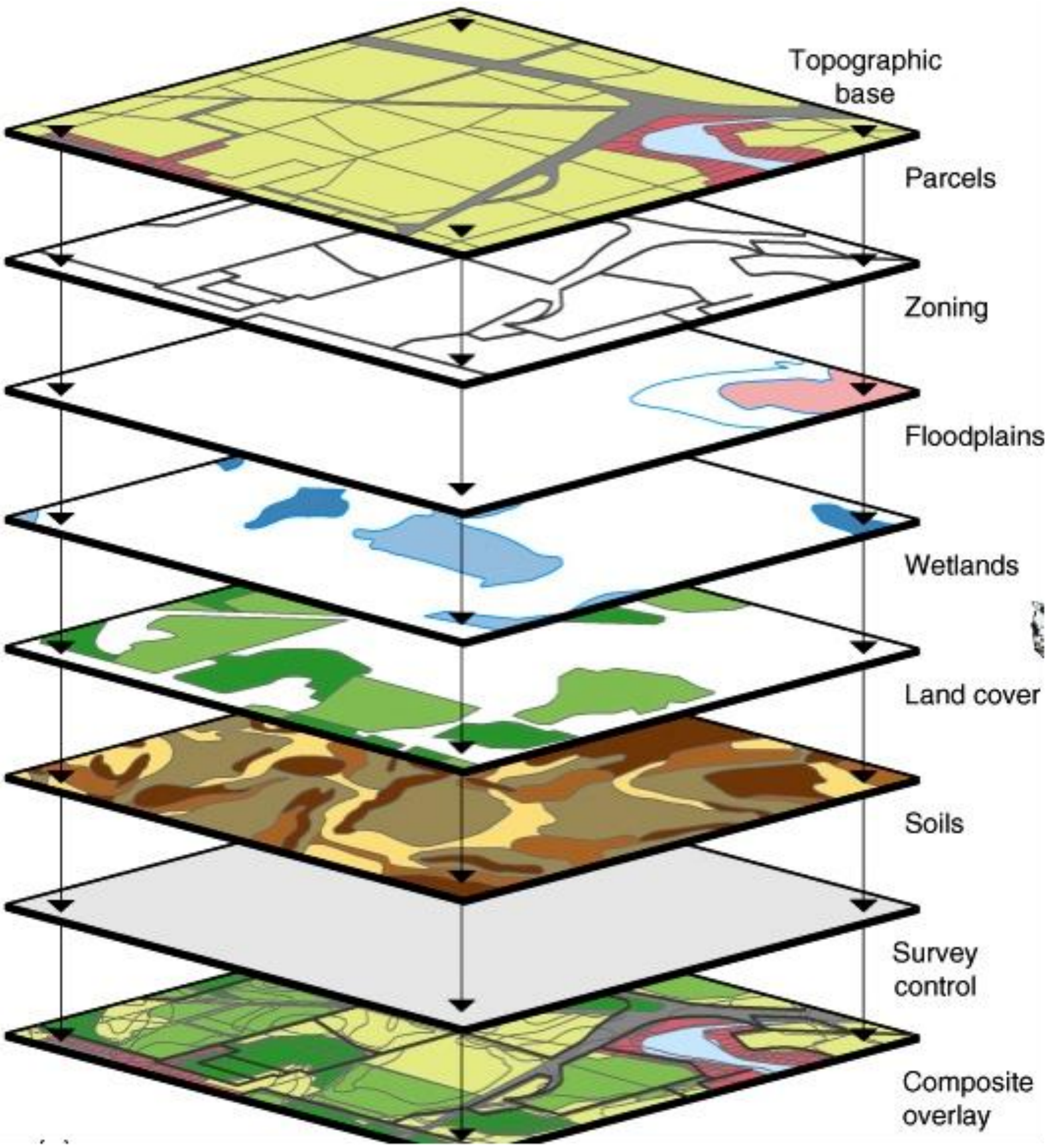

Figure 2: Overlapping GIS data layers.

mice or force-feedback joysticks as haptic displays, not the PHANTOM point-haptic device used in this study. Several systems were created to allow the blind and visually impaired access to geo-spatial information.

To help blind children gather information about a map, Treviranus (2000) used sound and haptic cues. Parente (2003) reported on the BATS system, which used a combination of text-to-speech synthesis, spatialized sound (auditory icons) and tactile feedback (vibrations and textures) to help students with visual impairments understand maps. Work by Griffin (2002) and Jacobson (2002) indicates that haptic and/or audio cues are effective at representing spatial information. Subjects in Jacobson's study used scanning and probing to explore a surface; scanning proved useful for finding overall trends in data, while probing was effective at finding out values at a specific location of interest.

Krygier (1994) presented a system of representing spatial data via realistic or abstract sounds that used analogies to visualization. Fisher (1993) used the duration of sound to communicate the reliability of pixels on a remotely sensed image. Zhao (2004) reported on the sonification of spatial information (choropleth maps) and suggested that the Shneiderman visualization mantra, "Overview first, zoom and filter, then details on demand" (Shneiderman, 1998), should also be considered when exploring spatial information via sound.

Jeong (2003) presented results of a feasibility study testing various identification tasks on a GIS with added haptic and auditory displays; haptic displays produced faster and more accurate performance than auditory displays and combined displays for more complex tasks. A predecessor to the current system discussed in this thesis was used for hapto-visual suitability analysis (Harding & Newcomb, 2004), a simple, generic process for dealing with the local suitability rated on a scale of 1 to 10 for various geospatial planning scenarios. In this system the PHANTOM was used to generate haptic gravity effects that guide the user in the direction of higher suitability by making higher suitability areas have an attractive force and areas of lower suitability a force that pushes the user away.

1.6 Introduction to road planning

As stated earlier, Geographical Information Systems (GIS) are used in the process of planning the placement of infrastructure like roads. A typical road planning project uses several overlapping layers of geospatial data which are interpreted by an expert to find the best path for the new road.

1.6.1 Challenges with existing planning systems

Planning roads in rural areas is a complex process with many stages, starting with the consideration of all of the contributing factors (spatial and non-spatial), then gradually narrowing the scope down to the process of determining a "strip" of land for the general route, and finally concluding with the precise placement of the actual road (US Dept. of Transportation, 2004). The more information available at the beginning of the process the more complex the solution becomes.

A detailed description for the creation of the "optimal" road is beyond this project's scope, but a few examples of some of the more intuitive (direct) rules for road placement include the facts that a shorter road requiring a tunnel through a mountain may be more expensive than a longer road around the mountain and that laying a power line through farmland is less expensive and more desirable than installing it within dense forest. Although the total direct building cost – based on factors such as land value, engineering expenses and predicted annual road maintenance – is certainly very important for the placement of the road, it is not the only factor.

Among the higher level (indirect) considerations that need to factor into the planning are the impact on environmentally sensitive areas, potential expandability in the future and even political issues such as social justice. However, as the "cost" of such indirect concepts cannot be quantified as easily as direct monetary costs, it is difficult to efficiently deal with them in the planning process.

In summary, the process of planning the placement of a new road is by no means straightforward; even sub-tasks, such as placing the initial rough “draft” of the road onto the landscape, are complex, iterative undertakings. Planning involves multiple overlapping sets of spatial data, and all of the rules cannot be effectively contained in formal guidelines. Only well-defined engineering tasks can be solved by automated (algorithmic), computer-optimized solutions; most other decisions still must be made by a human using his or her expert knowledge and experience. This project aimed to provide a novel multi-modal framework for the road placement phase, in which practitioners are able to transform their expert knowledge and past experiences into an initial road layout and do so more efficiently than with a traditional GIS.

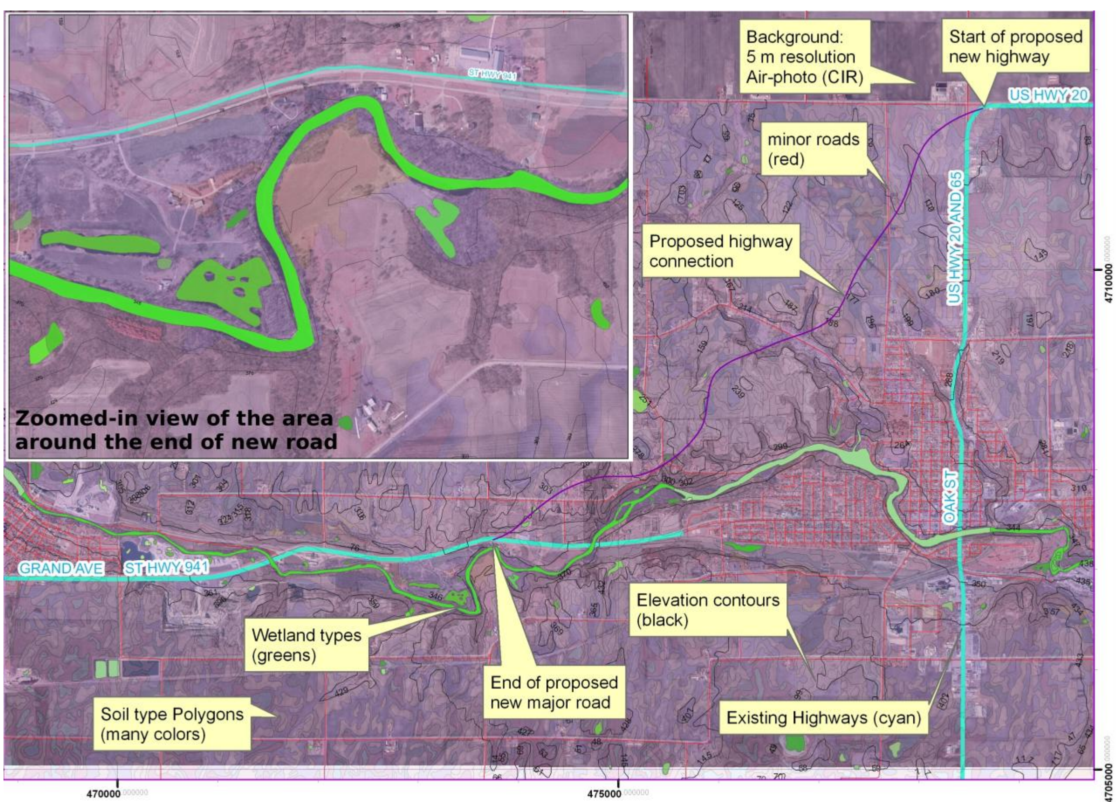

Figure 3: Example of combined geospatial layers used in traditional GIS. The exact placement of the new road, labeled "proposed highway connection", is based on information in these layers.

CHAPTER 2 - DESIGNING THE SYSTEM

Planning roads is traditionally performed with a GIS system, such as ArcGIS, which allows the user to choose which layers of data to display and provides varied options for displaying data. Planners are able to make certain layers semi-transparent and can also change the way that data is colored as well in order to help display more layers at once and differentiate between them. This project took a similar approach by creating different maps through touch and sound that can be overlapped with visual maps but designed so that only one layer can be displayed for each sensory modality. At most, the user can have one layer set for vision, one for touch and one for sound.

2.1 ReachIn system

The ReachIn system, seen in Figure 4, is an immersive visual and haptic display that uses a mirror to reflect 3D graphics so that they appear in the same location as the user's hand. The haptic device that users interact with is SenseAble's PHANTOM Desktop device. This co-location of the computer graphics and haptics is one of the main benefits of using the ReachIn system.

ReachIn (http://www.ReachIn.se) also provides a high level API to program the system that consists of three layers: VRML, Python, and C++. Each level allows for different complexity and functionality. The implementation of the system consisted of a VRML file that contained the geometry of the surfaces and the buttons of the interface. The functionality of the interface, which controlled the switching of data layers and the digitizing of the lines the users drew, was written in Python scripts that were linked to the main VRML file. The functionality of each representation of the data layers was programmed in C++. These classes calculated the pitch according to the sound layer data map in the graphics loop and calculated the force to apply to the stylus for certain haptic representations.

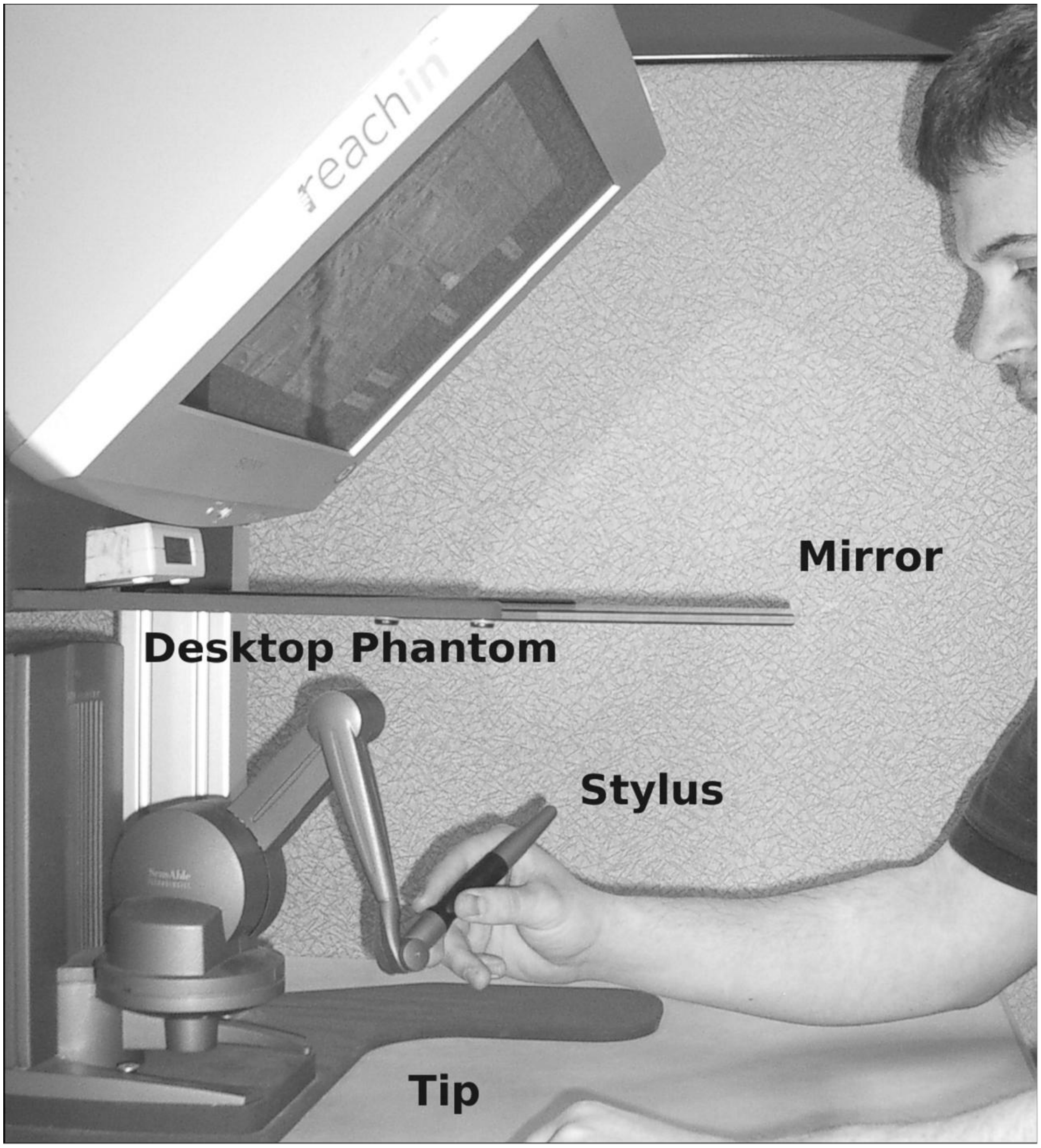

Figure 4: The Reachin system reflects the graphics from the monitor into the mirror so that the virtual stylus and tip match up with the actual physical location of the haptic device.

2.1.1 VRML

The most abstract layer of the ReachIn system, many popular 3D modeling programs such as Autodesk 3D Studio Max and Autodesk Maya can create this plain text XML-style file format. Because the files are in a readable format, the structure is very easy to visualize, and by adding in a few extra lines any basic VRML file can be rendered into touchable surface by adding properties to the 3D model data.

Graphical user interface objects, such as touchable menus, interactive sliders, and other widgets, are also created within the scene graph. This project used the ReachIn menu system for the interface to switch between the different representation layers.

2.1.2 Python

Python scripting can be inserted directly inside the VRML file itself or can be linked from external files. The scripting layer is suitable for any code that is not dependent on speed. In this project the scripts were mostly used for handling user interface tasks, like handling button presses. The scripting layer allows for rapid iteration without the need to compile any code.

2.1.3 C++

The previous two layers allow for the creation of a basic scene graph with some limited functionality, but to create new types of nodes that need to perform operations in the rendering or physics loops, then C++ is the only option. This low-level layer gives access to the rendering loop, called the renderer, and the high speed haptics loop, called the collider, which can be used to read values on a haptic surface, perform custom force feedback equations, and modify geometry at interactive speeds.

2.2 The road planning application

The users were presented with a simple interface that displayed a terrain surface on the left with a set of menu options on the right (Figure 5). The stylus could be used to select menu options in the interface as well as serve as a drawing device on the terrain surface for the purpose of planning the path for a new road.

Menu options were provided for the sound, visual and touch layers. The user could select which data set to map to each sense, or choose to turn the layer off completely. There were also a few line drawing menu options for saving one's progress, undoing the last portion of drawing, and earasing all the lines completely. A button on the stylus was used to draw on the map. The stylus also displayed the haptic layer of the map, and a pair of stereo speakers was used to display the sound layer.

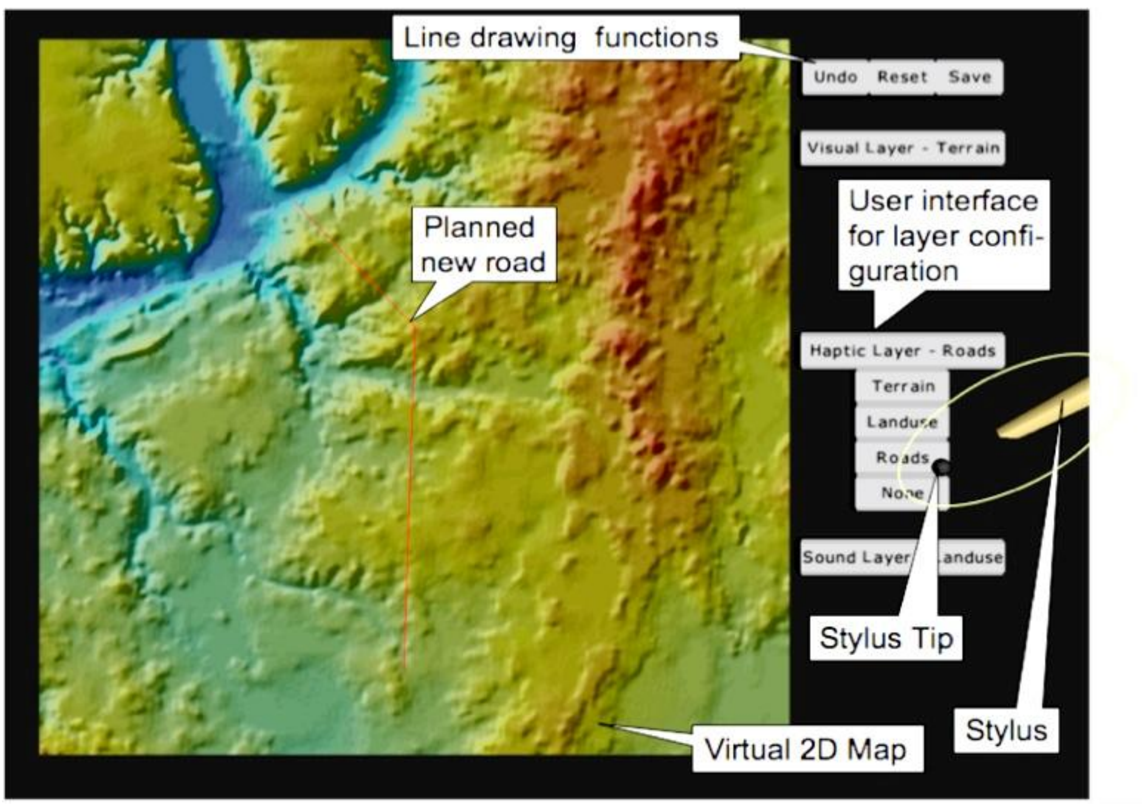

Figure 5: A screenshot of the system showing the terrain surface (left) and the haptic user interface (right). In this image the user is selecting the roads data to be displayed in the haptic layer. The tip of the stylus is used to press the interface buttons when selecting a data layer for a modality; it is also used to draw directly on the virtual map.

2.3 Layer representation

Three GIS data layers were taken from a section of northern Iowa and were used for the planning process: a map of the land use, a map of the terrain elevation and a map of the existing roads. The visual representation of these maps was very straightforward, but methods were needed to "view" these maps in the touch and sound layers. Each type of data required a slightly different approach.

2.3.1 Visual

Each map used a visual color scheme to display information. Terrain elevation, obtained as numerical data, was converted to color values; the elevation was represented by a spectrum of colors ranging from blue (low elevation) to red (high elevation). Hillshading was also applied to the map to simulate shadows. The land use map conveyed categorical data with different colors representing four types of land use: yellow for fields, green for forests, blue for water or wetlands, and red for populated areas. The third map of the existing roads used green coloring with a narrow buffer around the road to represent a dirt road, while highways were colored pink and had a larger buffer zone around them.

2.3.2 Touch

Different haptic effects were used to express each type of data. The three techniques used were haptic bump mapping, friction mapping, and gravity lines.

2.3.2.1 Haptic bump mapping

For the terrain elevation map to be effective as a touch layer, the user needed to be able to feel the difference in the elevation. The haptic bump mapping technique does this by taking a grayscale image of the elevation and displaces the height of the haptic pixel up or down depending on its elevation. Note that the terrain is still only rendered graphically in 2D.

This is very similar to displacement mapping in graphics, and a 3D model of the terrain could also be used if the data was available.

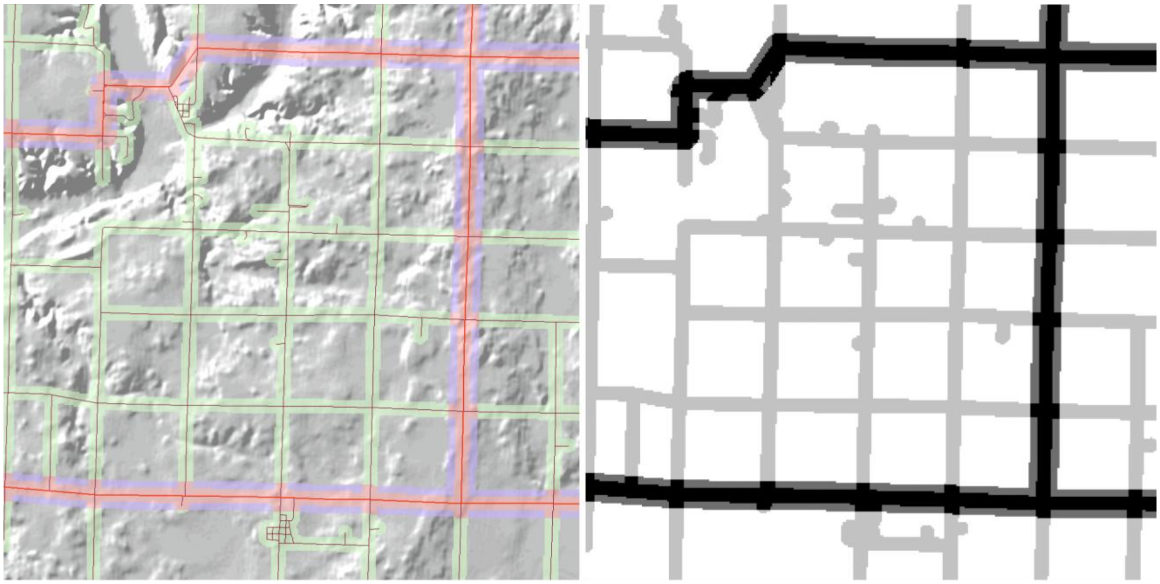

Figure 6: Visual layer of existing roads with buffer zones (left); audio/haptic representation of the road buffers (right).

2.3.2.2 Friction mapping

Categorical data, like land use, is more difficult to model haptically because there is no parallel to the visual way to distinguish categories, which is typically done by color. While there are ways to make different haptic textures, associating them with values would have to be arbitrary, since there isn't an immediately recognizable forest or field texture. Instead, each type of land use was represented with friction that correlated with the cost to build a road on each type of land. Fields, easier to build on and cheaper to buy, were assigned a very low friction value, while the slightly more expensive forests were assigned a slightly higher amount of friction. The most expensive locations – cities and wetlands – were designated with the highest amount of friction. The friction model has a dynamic and static friction component. The static friction component is larger than the dynamic friction component, which means it takes more force to move the stylus from rest than it does when the stylus is already in motion. The friction maps were also generated by grayscale images where white areas indicated the lowest friction and black areas indicated the highest friction.

2.3.2.3 Gravity lines

The haptic representation of a line needs to allow the user to follow a line that they cannot see. If a line branches off, the user needs to know the line hasn't simply ended and must also know which paths are available. To implement this, a structure called a gravity line, which attracts the user in the direction of the line, was created. When the user enters the boundary of the gravity line (the width of which can be set by the programmer), the stylus is snapped to the line and can follow the line with minimal effort. This was one of the custom nodes that were written for this application.

However, due to an API upgrade, this custom node was not available for the user study. Instead the haptic bump mapping effect was used to create troughs for each road that the user followed with the stylus. There is more discussion about this approach in the results section and why it is probably a less effective representation than the gravity line structure.

2.3.3 Sound

The sound layers used three different types of audio mappings to indicate changes in numerical data: natural tones to represent different types of categorical data, audio used to alert the user if they were within the proximity of an existing road, and parameter mapping which mapped a pitch to a corresponding elevation point.

2.3.3.1 Numerical representation

Numerical data sets, like terrain elevation, were represented by using tones with higher pitches for higher values and lower pitches for lower values. Pitch encoding can be used to represent changes in data relatively easily, even for untrained users (Flowers, 2005).

For this data, little or no change in pitch is desirable (it indicates flat terrain) and areas of high pitch changes indicate undesirable terrain (steep hills or valleys).

2.3.3.2 Categorical representation

Land use is categorical data by nature and requires a different form of representation. Sounds typical of each area's environment were used to indicate the type of land usage. Fields had chirping birds and insects, forests had the sounds of woodcutting and bears growling, water areas had the sound of running water and city zones had the sound of cars honking. This made it easy to identify the type of land on which the stylus was currently positioned without having to remember a difficult encoding. Each sound started with the signature sound such as a honk or a chirp and played the sound in a continuous loop until the user lifted the stylus off the map or moved into a new land use type.

2.3.3.3 Sonification of distance to roads

Direct sonification was used to indicate whether or not the stylus was located in a road. If the stylus was not on a road, no sound was played. If it was in the buffer zone for a dirt road, a low-pitched tone was played. If the stylus was over a highway, a higher-pitched sound was played with a slightly lower sound indicating the outer edge of the buffer. This sound helped the user know when he or she crossed a road, but did not provide information to tell the direction of the road.

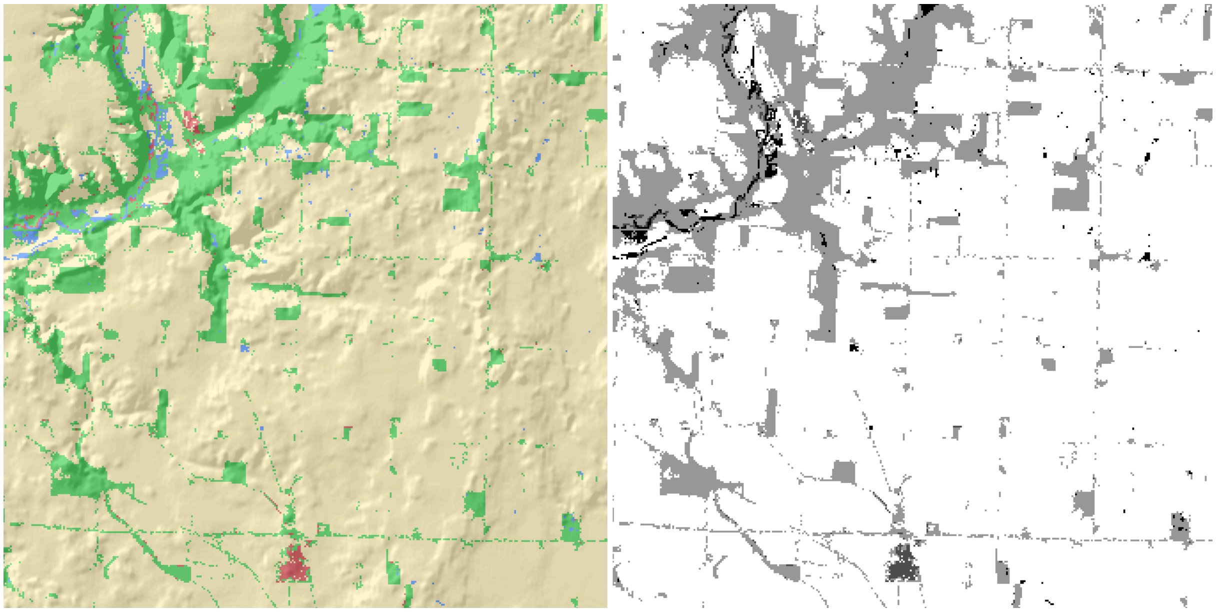

Figure 7: Visualization of land use data with four categories: fields, forest, rivers and cities (left) and their haptic/audio representation (right).

CHAPTER 3 - USER STUDY

A user study was conducted to determine how a person with basic planning experience would perform using the described MMI system to plan a road. After some initial interviews with a road planning expert a very simple multimodal system that used three different data layers was decided upon. Given the limited amount of available data on the use of MMIs in road planning, an ecological user study was implemented to provide some basic usage information so that a more formal quantitative study could be performed later to measure task performance, error rates, mental work load, etc.

3.1 Subjects

Subjects who had experience with road planning were desired, so volunteers from a highway design course were asked to participate in the study and were given 10 dollars and extra credit. Graduate students who had taken the class in the past were also recruited for this study. They were given $15 dollars to compensate for their inability to receive extra credit. There were a total of 12 subjects, 10 male and two female. The subjects' average age was 24 and they consisted of seven seniors, three master's students and one PhD student.

3.2 Procedures

Each user was provided with an informed consent document that outlined the tasks and procedures of the user study (see Appendix A.). They were then given an entrance survey with questions about their background and experience in planning as well as experience with computers and music. Without initially seeing the start and end points, the subjects were then told that their task was to design a road to go between two points on a map using the following criteria in order of importance:

1. The road should be as short as possible.
2. The road should avoid any steep increases or decreases in elevation.

3. Avoid creating unnecessary curves in the road.
4. The cheapest type of land to build on is fields. Forests are more expensive but are still acceptable to build on. Cities and wetlands should be avoided as they are the most expensive types of land use.
5. Try to reuse existing roads if possible, with a preference for major roads.
6. When crossing a major road, try to cross as closely to 90 degrees as possible.

The subjects were then given a chance to learn how to use the system during a short 5-minute training session. The subjects had the opportunity to experiment with each sensory layer (visual, touch, and sound) and look at each of the three maps provided in each sensory representation. They were also given several tasks that involved combining different data layers into different senses, such as "while looking at the terrain, try to find the southern city by listening for it in the land use map". After subjects became familiar with the data, they were given a chance to experiment with drawing a line; pressing the button on the stylus drew a yellow line while they were in contact with the map. Subjects were told to draw as many lines as they wanted to and that they would not be timed.

After asking the subject if they understood how to use the system and understood the task they were to perform, the videotaped recording was started. Users were asked to describe what layers they were changing and their reasoning behind changing it. They were also asked to "think aloud" or talk about their overall strategy for planning the section of road and ask questions if they ran into difficulty; if the user stopped talking, they were encouraged to start talking out loud again. The think aloud technique has been shown to be effective in discovering usability problems (Virzi et al., 1993). The list of line-placing criteria and a description of how each type of data was represented depending on the sense selected were displayed on a secondary screen for subjects to refer to as well. After subjects finished their first route, they were asked to check it against various maps and possibly decide on a secondary route. They were given a chance to make any minor changes if they wished. The videotape was stopped when the subjects said they felt satisfied with the roads that they had digitized.

After using the system, subjects were given an exit survey about their experience and the usefulness of the system's features. The exact forms given to the subjects can be found in Appendix A, while a visualization of data from the study can be found in Appendix B.

CHAPTER 4 - RESULTS

The results from the user study are presented in this chapter. Subjects were asked several questions about using the system and their responses are compared to how they actually used the system. Some usability concerns are also presented.

4.1 Which data layer is best suited for each sense?

Participants were asked to indicate which sense (vision, touch or sound) they felt was best suited for each data layer (terrain elevation, land use or existing roads). These subjective results showed that each data layer had a dominant sense that users felt best represented that particular data. Most subjects preferred to use the terrain elevation data layer with the sense of touch, while sound was most favored for land use and the existing roads seemed to be best displayed in the visual sense.

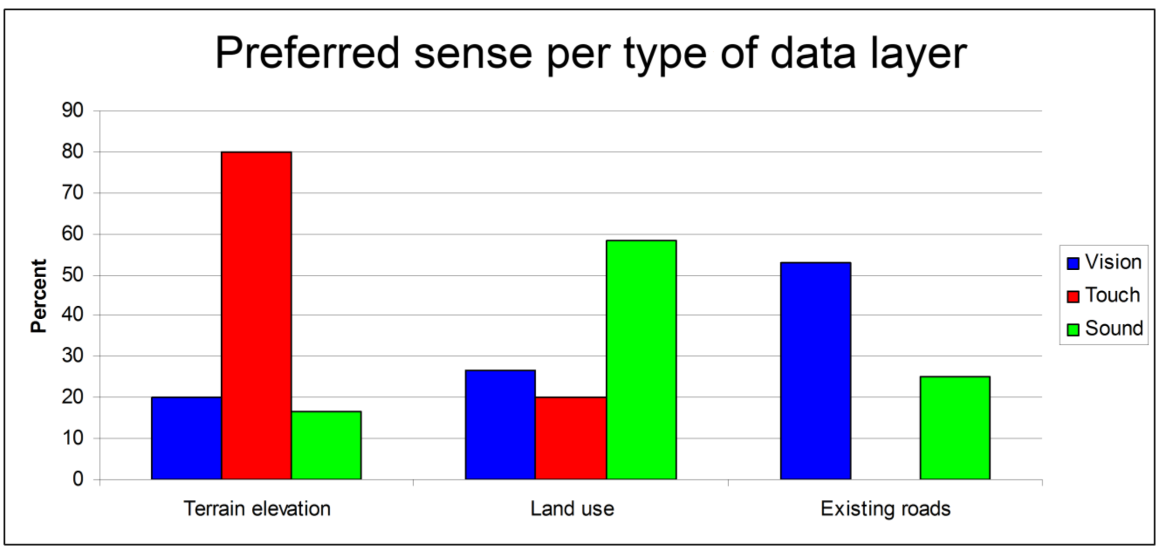

Figure 8: Subjectively preferred data layer for each sensory modality.

Based on user comments, the explanation for using touch with the terrain was that it gave the ability to ascertain the elevation quickly with a few quick movements. The land use layer was preferred with sound since the tones gave instant, common-sense indications of the land's suitability with each motion of the stylus. The complete network of roads in the test region could be viewed all at once in the visual layer, while audio and haptic output only provided information about the immediate area indicated by the stylus.

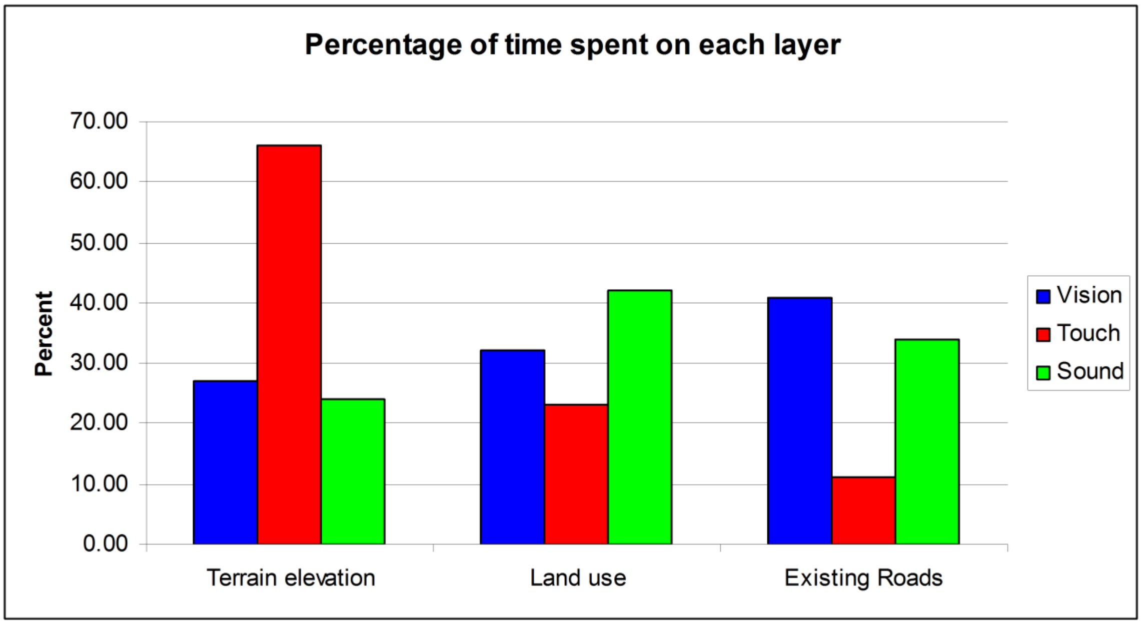

Figure 9: The objective percentage of time each data layer was active for each sense.

The subjects largely agreed on which layer they thought worked best with each sense. Objective data about which combinations were used the most when performing the task was also collected. To determine this, the time the subjects spent on each sense was totaled and a percentage was calculated by dividing that total by the total time spent on each data layer for all three senses. From this objective data in Figure 9, it can be concluded that the layers that subjects thought were best for each sense were indeed the layers that were active the most. The only exception to the alignment of subjects' preferences with the time spent in the preferred combinations is that the subjects used the sound output with the existing roads layer for a larger percentage of the time than indicated by subjective preference.

4.2 What is the importance of each sense?

Humans have one of the most acute visual senses on the planet, aside from birds of prey such as hawks, so it is not a surprise that a great deal of importance is placed on visual information. Humans can also utilize the senses of touch and sound to gain a meaningful amount of information for use during a task. Users were asked to rank the order of usefulness for each sense (see Figure 10), and vision, as expected, was picked by all but two subjects as the most important sense in performing the road planning task. Sound was picked by all but one subject as the least important sense, with touch ranking second. This data alone does not reveal how useful touch or sound is, but there are other indicators that can help explain why sound was clearly the least preferred, such as the total time each sense was disabled. Occasionally, users would turn off the sounds if they found them distracting or irritating, and some subjects turned on the higher-ranked senses first upon starting the study with all of the data layers turned off.

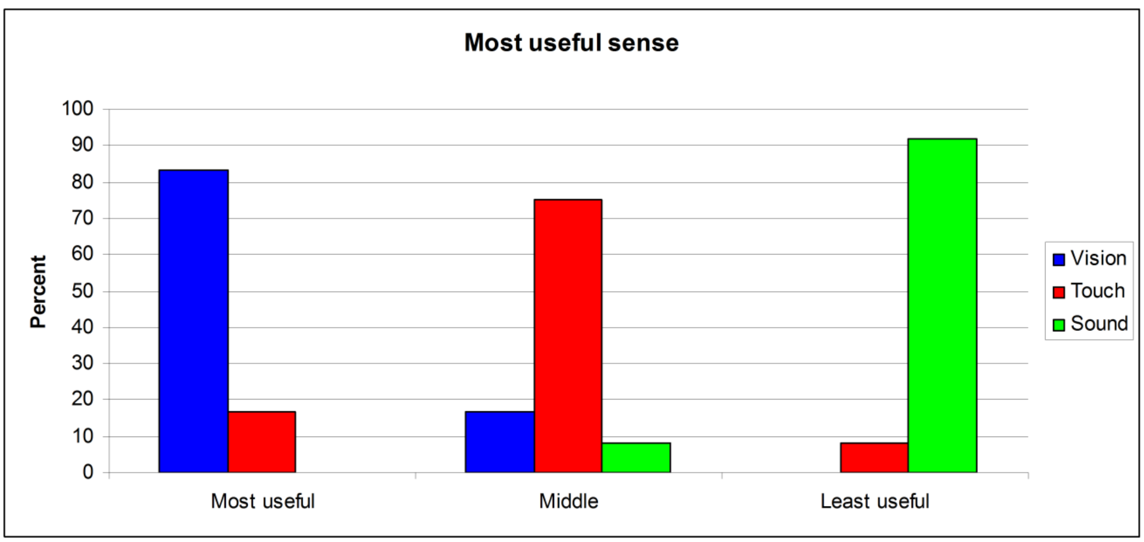

Figure 10: Subjective ranking of each sense by usefulness.

Looking at the time that each sensory layer was disabled (Figure 11, total time, in minutes, cumulatively spend by all users), vision was disabled the least with its total usage coming from the initial setup phase when no layers were active. Touch and sound, however, were sometimes turned off in the middle or end of the user study. The touch layer spent more time disabled than vision, but was disabled less often than sound. The main reasoning users cited for turning off the haptic layers was because they were trying to draw a segment of the road and the haptic layer made it difficult for them to accurately draw on the surface. Sound was inactive for the longest period of time; more than double that of touch. The reasons that subjects gave for disabling the audio were that the sound was not helpful or that it was too distracting. These reasons may explain why sound was rated the least useful sense used in this task.

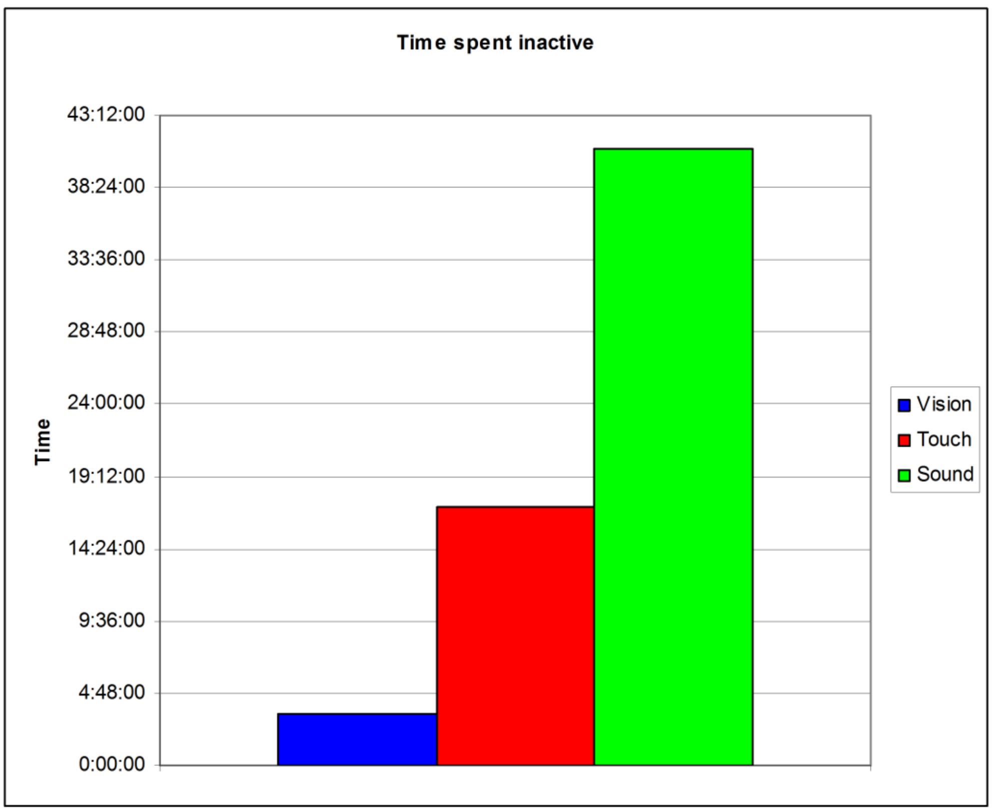

Figure 11: Total time (min:sec:msec) that each sensory layer had no active data

4.3 Typical Use

To visualize the way that users configured their environments and see how they changed over time, some usage timeline diagrams were created with time given in seconds.

User 4, for example, configured their environment to show terrain elevation visually, display existing roads through touch and represent land use through sonification (Figure 12a). At the test's end they displayed existing roads visually, elevation haptically and land use aurally.

Most of the subjects in the study fell into two major groups: "lazy switchers" and "explorers" (Harding & Souleyrette, 2009). The former typically picked an initial configuration for the sound, haptic and vision layers and make very few switches from their original configuration. These users also tried to avoid mapping the same data layer to multiple sensory layers at once so that all layers were represented to them at all times. The approach used for this group was more linear, starting at one point and gradually working towards the other and trying different combinations when an obstacle such as an intersecting road or body of water was encountered. User 3 in Figure 12a is a good example of a lazy switcher. This individual changed their sound layer once and never touched their haptic layer; only the visual layer was changed multiple times.

The explorers switched their configurations at a much higher frequency. The visual layer still received the majority of switches, but the users in this group would try many different combinations throughout the study. They combined data layers with different senses more often, sometimes mapping one data layer to all three senses, and especially changed configurations when encountering an obstacle. It may not be apparent which mapping combination is the best for each problem encountered, so the subjects in this group seemed more willing to explore to find the combination that worked the best.

Across both groups, the visual layer saw the most changes by far. The average number of switches was 17.25 for the visual layer, while the averages for touch and sound were 4.91 and 3.75 switches respectively. This may be due to vision being the familiar way to process maps; it gives a large picture of the area at once instead of the limited information of the sound and touch layers.

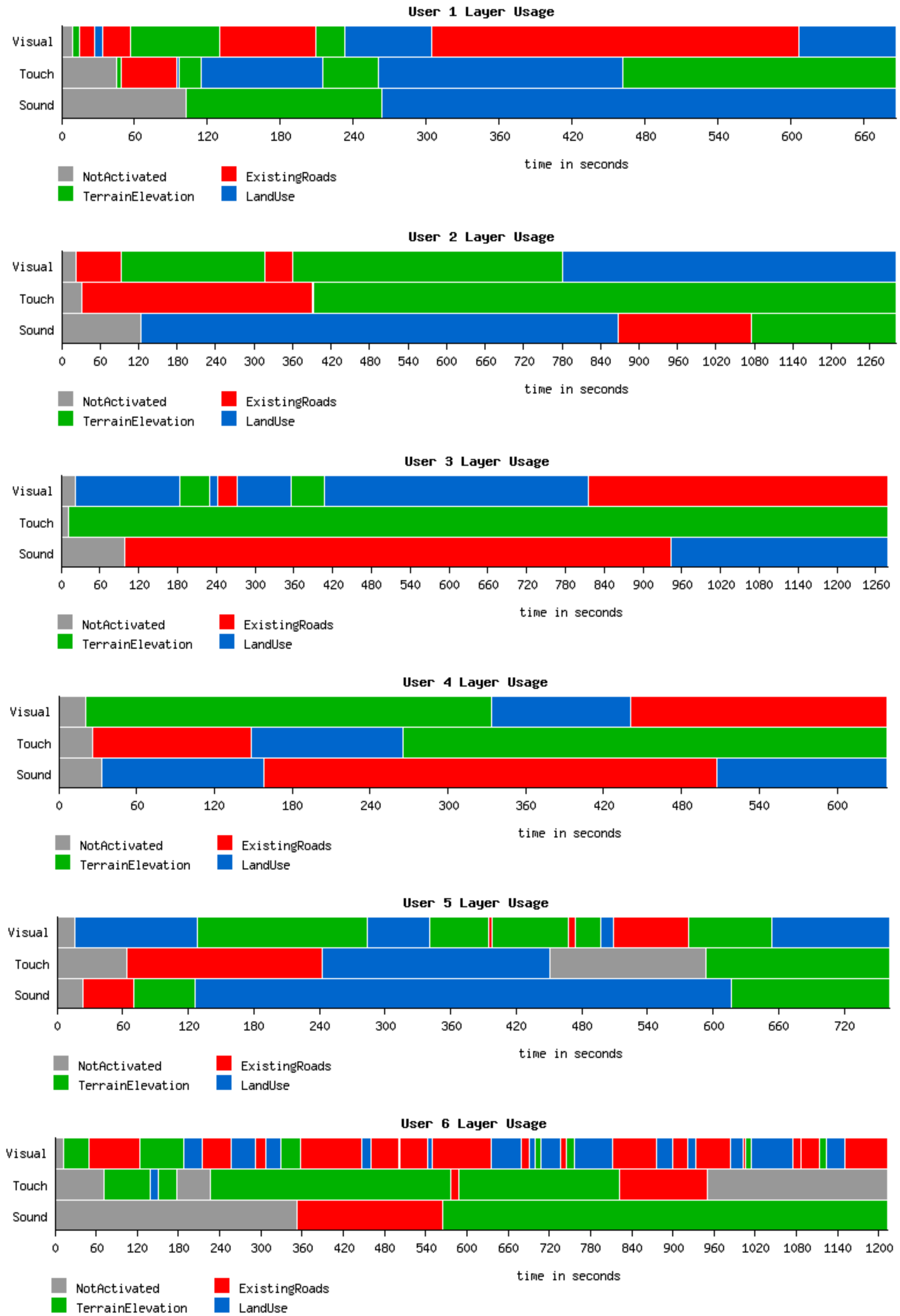

Figure 12a: Layer usage over time for Users 1 through 6.

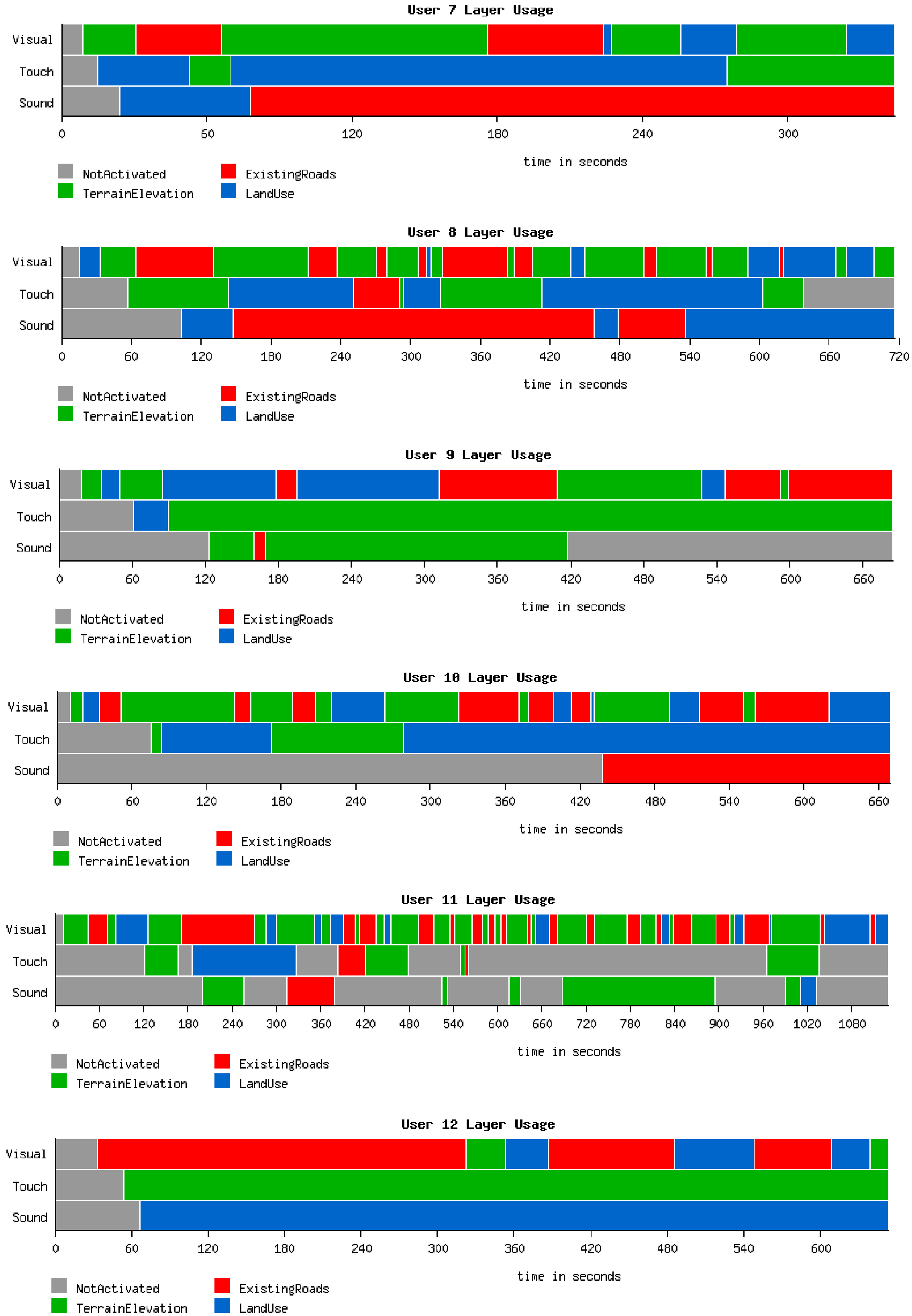

Figure 12b: Layer usage over time by Users 7 through 12.

4.4 Usability problems

There were several usability problems that came up while running the user study. The major problem was lack of feedback from the user interface. For the most part, the 3D menu system was easy to use, but once pressed the buttons gave no indication which layer was active for which modality. While it was obvious which visual layer was currently active, it was not so obvious for touch and sound, especially when some of the same sounds or haptic effects are used.

The second problem was the incomplete drawing functionality. Originally it was conceived that users would digitize points in a line segment by clicking a button on the stylus; this was to have an undo-last-point function and a reset button to help manage the line. Instead, subjects used something similar to drawing a line with a mouse. They could draw a line by holding down the button; this proved to have some benefit over the original idea, as many people drew small segments of line at each end of the area first, which would have required adding new segments and joining separate segments if the original point connection method had been implemented. However, there was no erase functionality, so subjects were told to cross out sections they didn't want to use and to feel free to draw multiple line segments.

CHAPTER 5 - FUTURE WORK AND CONCLUSIONS

Background on multimodal systems has been presented as well as some reasons why they may be well suited for GIS tasks, such as road planning, which uses multiple layers of information. A novel road planning interface was created using the ReachIn haptics system and a preliminary user study was conducted to determine how a multimodal GIS would be used for the specific application of road planning. A summary of the important results and areas of future work are presented below.

5.1 Future Work

The results from this study have shown that there is potential for sound and haptics in GIS systems. Haptics used to feel the terrain of a surface may be successful in helping generate a mental 3D model of terrain and helping with the specific placement of new road corridors. The use of sound needs to be investigated further. There did seem to be some usefulness of using natural sounds for land use, but there were a few user comments that suggested the sounds needed to be more distinct from each other.

A more formal experiment comparing this system against a traditional unimodal system is needed to obtain information about the differences in efficiency and workload strain. Testing for workload strain caused by stress could also help establish whether or not the task is more or less mentally taxing in a multimodal system. Comparing completion time and the correctness of the final solution may be used to determine whether or not users perform better with a multimodal system.

One future change that should to be implemented in the system to make it more powerful and user friendly would be to add pan and zoom functionality to the map surface so that larger areas can be used and specific areas can be concentrated on at will.

Another request that several subjects submitted was for the ability to overlap visual layers as other current systems do. Since it has been shown that there is a high amount of visual switching, adding transparency or another method of combining different visual layers at one time would be a way to help reduce the number of these visual switches. However, in this study users were only able to view one layer at a time.

5.2 Conclusion

Because this is a relatively unexplored area, the results from this ecological user study will be helpful in guiding the future design of multi-modal systems for geospatial data. This study has shown that for this specific task, users ranked vision as the most important sense, touch as secondary and sound as the least important. Users switched their visual layer much more often than their touch and sound layers, which they changed only once or twice on average.

The existing roads layer seemed to be best represented visually, because seeing the entire network of the roads at once is not possible with the single point of interaction that sound and touch use in the system. Terrain elevation was observed to be best represented with the sense of touch as users would be able to trace their path and see if they felt any significant changes in terrain where their path would be. The sense of sound was most associated with using natural sounds (or auditory icons) based on the categorical data of land use; this gave the user a realistic and effective audio cue for each type of land use. The preference for these combinations of roads can be found in the users' subjective reporting of how they rated the sensory-data layer combinations. The usage data shows that these combinations were also the combinations that were used most frequently throughout the experiment.

APPENDIX A - USER STUDY MATERIAL

Entrance questionnaire:

Pre Questionnaire for Subject # ___

You may skip any question that you do not wish to answer or that makes you feel uncomfortable

Age: __

Sex: M/F

Are you left or right handed? ___________

School Classification (Junior, Senior, Master, PhD, etc): ___________________________

Major or Area of Concentration: ____

Road planning experience outside of CE 453: _________ years, where? ___________________, in what role? ___________________

Years of active experience with GIS: _______ what Software? ____________________

Computer experience: Computer gaming: _____ years,   what games? _____________

Experience with Virtual Reality (VR) systems: what system? ________________

Musical background: How many years have you played a musical instrument or were actively involved in singing? ______ Years

Do you have any known hearing impairments? Y/N

If yes, please describe:

____

Do you have any of the following visual impairments?

Contacts/glasses? Y/N

color blindness? Y/N

other? (please describe):____

Any other disabilities (physical or learning)?

____

Exit Questionnaire for Subject # ___

You may skip any question that you do not wish to answer or that makes you feel uncomfortable

Which **sense** was **most useful** in the overall planning process (circle one):

Sound

Touch

Vision

Which **sense** was **least useful** in the overall planning process? (circle one):

Sound

Touch

Vision

Which **data** was the **most important** in the overall planning process? (circle one):

Terrain elevation

Land use

Existing roads

Which **data** was **least important** in the overall planning process? (circle one):

Terrain elevation

Land use

Existing roads

What combination of data and senses did you think worked best for you while planning the road? Connect each layer to the sense that was used by drawing a line (for example: Roads – sound, terrain – touch, land use – vision):

Data layer	**Sense**
Existing roads	Touch
Terrain elevation	Sound
Land use	Vision

What were the most important factors (valleys, hills, forest, rivers, cities, etc.) for developing your initial higher level planning strategy?

How much do you agree with the following statements?

1 – Strongly Disagree

2 – Disagree

3 – Neutral

4 – Agree

5 – Strongly Agree

I relied mostly on vision to initially develop a strategy for placing the road (higher level)

1 2 3 4 5

The precise placement of the road was mostly influenced by touch (lower level)

1 2 3 4 5

The precise placement of the road was mostly influenced by sound (lower level)

1 2 3 4 5

The precise placement of the road was mostly influenced by vision (lower level)

1 2 3 4 5

What did you like about the system?

What improvements could be made to the system?

Task description

Task: Your task is to plan a new section of highway (road) from the red sphere to the green sphere. You have 3 layers – although you can look at only one of them (at a time) visually.

We are especially interested in how you would use audio/haptics layers to augment visual layers. We want you to think about scenarios where you look at one type of layer with a certain talk in mind and configure the audio layer and/or haptic layer to help you with this task. For example, when looking at the roads to follow them or cross them at 90 deg., you could use the terrain via sound to keep the road level and use the land use via touch to avoid running into a forest area.

When digitizing, you are free to "snoop ahead" before actually putting paint on the map – just don't press the stylus button until you actually want to draw a line.

High level priorities:

1) Make the road as short as possible.

2) Avoid going over hilly terrain and steep slopes – try and stay level as best you can

3) Avoid unnecessary curves – keep the road as straight as possible.

4) Its best to go over fields (yellow), it is OK but not best to go through forest (green)

5) Avoid going through water (rivers, lakes; blue) and towns (red) if you can

6) Try and re-use existing roads if you can

7) Try and cross existing main roads (highways) at 75 - 90 degree angles – dirt roads can be crossed at any angle

Terrain Map

- Visual – Blue = low elevation, red = high elevation
- Haptic – The pen should sink down when entering areas of low elevation
- Sound – Higher pitch = higher elevation, lower pitch = lower elevation

Land Use Map

- Visual – Red = urban, Yellow = fields, Green = forest, Blue = wetlands

- Haptic – Areas are coded with friction, so that more friction = more expensive to build through. So fields have no friction, forests have a little friction, and urban areas and wetlands have the most friction.

- Sound – cars honking = urban, birds chirping = fields, chopping down trees = forests, running water = wetlands

Roads Map

- Visual – Existing roads are drawn with buffers around them. Green buffers indicate smaller dirt roads. The pink buffers indicate state highways.

- Haptic – You will feel haptic "troughs" along the existing roads, so you'll sink in with the stylus when you are on a road.

- Sound – No sound indicates you are not on a road. Low sounds indicate that you are on a dirt road. High sounds indicate that you are near or on a major road or highway.

APPENDIX B - STATISTICAL RESULTS

Time spent in each configuration (min:sec)

	Terrain elevation	Land use	Existing Roads	Not Active
Vision	46:33	54:06	70:28	3:23
Touch	91:09	31:04	15:27	17:09
Sound	30:25	52:46	43:10	40:59

Time spent in each configuration (% of total time)

	Terrain elevation	Land use	Existing Roads
Vision	27.00	32.00	41.00
Touch	66.00	23.00	11.00
Sound	24.00	42.00	34.00

Responses for question: What combination of data and senses did you think worked best for you while planning the road? Connect each layer to the sense that was used by drawing a line (for example: Roads – sound, terrain – touch, land use – vision):

Best configuration (number of votes)

	Terrain elevation	Land use	Existing roads
Vision	3	4	8
Touch	8	2	0
Sound	2	7	3

List of times in minutes:seconds and modality/map combination the subject changed to:

The first letter represents the sense (visual, touch or sound) and the second letter represents the data layer (terrain elevation, land use, and roads). For example, VT denotes the subject switching to the visual perception of the terrain elevation map. The total time and total number of switches per subject are shown in bold. Final time and total number of state changes are at the end of each list

User 01		**User 02**		**User 03**		**User 04**	
Frame	**State**	**Frame**	**State**	**Frame**	**State**	**Frame**	**State**
0:09	VT	0:23	VR	0:11	TT	0:21	VT
0:15	VR	0:32	TR	0:22	VL	0:26	TR
0:27	VL	1:33	VT	1:38	SR	0:33	SL
0:34	VR	2:04	SL	3:04	VT	2:28	TL
0:45	TT	5:17	VR	3:50	VL	2:38	SR
0:49	TR	5:35	VT	4:03	VR	4:18	VL
0:57	VT	6:00	VR	4:33	VL	4:25	TT
1:35	TL	6:30	TL	5:55	VT	5:34	VT
1:37	TT	6:33	TT	6:48	VL	7:21	VR
1:42	ST	13:00	VL	13:36	VR	8:23	TT
1:55	TL	14:28	SR	15:44	SL	8:27	SL
2:10	VR	17:55	ST	**21:18**	**11**	**10:38**	**11**
3:10	ST	**21:40**	**12**				
3:29	VT						
3:35	TT						
3:53	VL						
4:21	TL						
4:23	SL						
5:04	VR						
7:38	TL						
7:42	TT						
7:45	SL						
10:06	VL						
11:26	**23**						

User 05		**User 06**		**User 07**		**User 08**	
Frame	**State**	**Frame**	**State**	**Frame**	**State**	**Frame**	**State**
0:16	VL	0:13	VT	0:09	VT	0:15	VL
0:24	SR	0:50	VR	0:15	TL	0:33	VT
1:04	TR	1:12	TT	0:24	SL	0:57	TT
1:10	ST	2:03	VT	0:31	VR	1:04	VR
2:06	SL	2:11	VL	0:53	TT	1:43	SL
2:08	VT	2:19	TL	1:06	VT	2:11	VT
4:03	TL	2:31	TT	1:10	TL	2:24	TL
4:44	VL	2:58	TO	1:18	SR	2:28	SR
5:41	VT	3:07	VR	2:56	VR	3:32	VR
6:35	VL	3:35	VL	3:44	VL	3:57	VT
6:38	VR	3:47	TT	3:47	VT	4:11	TR
7:31	TO	4:17	VR	4:16	VL	4:31	VR

7:48	VT	4:52	VL	4:35	TT	4:40	VT
7:54	VR	5:08	VT	4:39	VT	4:51	TT
8:17	VT	5:29	VR	5:24	VL	4:54	TL
8:29	VL	5:52	SR	**5:44**	**15**	5:07	VR
9:38	VT	5:57	VL			5:14	VL
9:54	TT	7:27	VR			5:18	VT
10:17	ST	7:40	VL			5:26	TT
10:54	VL	8:21	VR			5:28	VR
12:41	**20**	8:23	VL			6:23	VT
		9:03	VR			6:29	VR
		9:09	VL			6:45	VT
		9:24	ST			6:53	TL
		9:36	TR			7:18	VL
		9:48	TT			7:30	VT
		10:35	VR			7:38	SL
		11:19	VL			7:59	SR
		11:31	VT			8:21	VR
		11:40	VL			8:32	VT
		11:47	VR			8:47	SL
		12:16	VT			9:14	VR
		12:25	VL			9:19	VT
		12:37	VR			9:50	VL
		13:33	VL			10:03	TT
		13:43	TR			10:17	VR
		13:54	ST			10:21	VL
		14:37	VR			10:38	TO
		14:59	VL			11:06	VT
		15:22	VR			11:15	VL
		15:33	VL			11:39	VT
		15:51	TO			**11:56**	**40**
		16:24	VR				
		16:42	VT				
		16:46	VL				
		16:55	VR				
		17:54	VL				
		18:07	VR				
		18:34	VT				
		18:44	VL				
		19:11	VR				
		20:12	**50**				

User 09		**User 10**		**User 11**		**User 11**	**(cont.)**
Frame	**State**	**Frame**	**State**	**Frame**	**State**	**Frame**	**State**
0:19	VT	1.16.10	VT	0:12	VT	12:01	VR
0:35	VL	1.24.10	VL	0:44	VR	12:12	VT
0:50	VT	2.23.20	VR	1:11	VT	12:55	VR
1:01	TL	2.36.13	VT	1:23	VL	13:15	VT
1:25	VL	2.49.15	TT	2:02	TT	13:34	VR
1:30	TT	2.52.28	TL	2:07	VT	13:43	VL
2:03	ST	3.28.21	VR	2:47	TO	13:53	VT
2:40	SR	3.41.11	VT	2:51	VR	13:58	VR
2:50	ST	4.23.05	VR	3:06	TL	14:23	VT
2:58	VR	4.38.22	TT	3:20	ST	14:54	SO
3:15	VL	5.23.17	VT	4:17	SO	14:56	VR
5:12	VR	6.11.00	VL	4:31	VT	15:15	VT

6:49	VT
6:58	SO
8:48	VL
9:07	VR
9:52	VT
9:59	VR
11:24	**17**

6.19.07	VT
6.39.23	TL
6.53.29	VR
7.09.11	VT
7.12.03	VR
7.18.24	VL
8.12.22	VR
8.36.01	VL
9.11.19	VT
9.21.28	SR
10.20.03	VL
11:09:02	VR
11.27	VT
21.03	VR
34.26	VL
52.24	**25**

4:46	VL
5:00	VT
5:14	SR
5:27	TO
5:52	VL
6:02	VT
6:14	VL
6:19	SO
6:23	TR
6:32	VR
6:47	VT
6:53	VR
7:01	TT
7:15	VT
7:26	VL
7:35	VT
7:59	TO
8:13	VR
8:34	VT
8:44	ST
8:53	SO
8:55	VR
9:02	VT
9:09	TT
9:16	TR
9:20	TO
9:25	VR
9:39	VT
9:47	VR
9:56	VT
10:05	VR
10:13	VT
10:16	ST
10:31	SO
10:41	VR
10:45	VT
10:52	VL
11:11	VR
11:21	VT
11:27	ST

15:21	VL
15:33	VR
16:05	TT
16:09	VL
16:11	VT
16:31	ST
16:50	SL
17:13	SO
17:15	TO
17:18	VR
17:23	VL
18:24	VR
18:33	VL
18:49:32	**75**

User 12

Frame	**State**
0:33	VR
0:54	TT
1:07	SL
5:22	VT
5:53	VL
6:27	VR
8:06	VL
8:30	VR
8:46	VL
9:08	VR
10:09	VL
10:39	VT
10:53	**12**

Avg # of switches	**25.92**
Range	**11-75**
Mode	**18.5**

Entrance questionnaire – Summary

Average Age: 24

Sex: 2 Female, 10 Male

All 12 Civil Engineering majors

All 21 right handed

Classification: 1 PhD, 3 Masters, 8 Seniors

Average road planning experience: 0.44 years

Average GIS experience: 1.1 years

Average computer experience: 6.4 years

No one had previous VR experience

Average years of music experience: 4.1 years

7 wore contacts or glasses; 2 (both males) had colorblindness as well

Relied mostly on visual: 4.4 average

Placement depended on visual: 4.2 average

Placement depended on touch: 3.2 average

Placement depended on sound: 3.0 average

REFERENCES

Blattner, M., Greenberg, R., Sumikawa, D. (1989). Earcons and icons: Their structure and common design principles ; Human computer interaction, Vol. 4, 1989., No. 1, Pages 11-44.

Bolt, R., (1980). "Put-That-There: Voice and Gesture at the Graphics Interface," SIGGRAPH `80, 262-70.

Crossan, A., Williamson, J., Murray-Smith, R. (2004). Haptic Granular Synthesis: Targeting, Visualisation and Texturing. Proceedings of the Eighth International Conference on Information Visualisation.

Emery, V.K., Edwards, P.J., Jacko, J.A., Moloney, K.P., Barnard, L., Kongnakorn, T., Sainfort, F., Scott, I.U. (2003). Toward Achieving Universal Usability for Older Adults Through Multimodal Feedback. Proceedings of the 2003 Conference on Universal Usability.

Fisher, P.F. (1993). Visualizing uncertainty in soil maps by animation. Cartographica, 30(2&3), 20-27.

Flowers, J.H., (2005). Thirteen Years of reflection on auditory graphing: promises, pitfalls, and potential new directions. Proceedings of ICAD 05-Eleventh Meeting of the International Conference on Auditory Display. Limericak, Ireland. July 6-9, 2005.

Gaver, W. W. (1994). "*Using and Creating Auditory Icons*". Auditory Display:Sonification, Audification and Auditory Interfaces. G. Kramer, Addison-Wesley Publishing Company. XVIII: 417-446.

Griffin, A.L. (2002). 'Feeling it out: The use of haptic visualization for exploratory geographical visualization. *Cartographic Perspectives*. 39. 12 - 29.

Harding, C., Newcomb, M. (2004). Supporting Interactive Data Exploration for GIS Planning Tasks with a Multi-modal Virtual Environment. Proceedings of HAVE 2004.

Harding, C., Souleyrette, R. (2009). Investigating the use of 3D Graphics, Haptics, and Sound for Highway Location Planning. *Journal of Computer-Aided Civil and Infrastructure Engineering*. 24. 1-19

Jacobson, R.D. (2002). Representing Spatial Information through Multimodal Interfaces. Proceedings of the Sixth International Conference on Information Visualisation.

Jeong, W., Gluck, M. (2003). Multimodal geographic information systems: adding haptic and auditory display. J. Am. Soc. Inf. Sci. Technol. 54, 3 February, 2003.

Krygier, J.B. (1994). Sound and Geographic Visualization. In Visualization in Modern Cartography MacEachren, A. and Taylor D.R.F. (Eds.) New York: Pergamon. pp. 149-166.

Lederman, S.J., Martin, A., Tong, C., Klatzky, R.L. (2003). Relative Performance using Haptic and/or Touch-Produced Auditory Cues in a Remote Absolute Texture Identification Task. Proceedings of the 11th Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems.

McGee, M.R., Gray, P., Brewster, S. (2000). The Effective Combination of Haptic and Auditory Textural Information. Proceedings of the First International Workshop on Haptic Human-Computer Interaction.

Miller, G.A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. The Psychological Review, vol. 63, pp. 81-97, 1956.

Oviatt, S., Coulston, R., Lunsford, R. (2004). When Do We Interact Multimodally? Cognitive Load and Multimodal Communication Patterns. In Proceedings of the Sixth International Conference on Multimodal Interfaces (ICMI 2004), Pennsylvania, USA, October 14-15, 2004.

Parente, P. and Bishop, G. (2004). BATS: The Blind Audio Tactile Mapping System. In Proceedings of the ACM Southeast Regional Conference. March, 2003.

Peeva, D., Baird, B., Izmirli, O., Blevins, D. (2004). Haptic and Sound Correlations: Pitch, Loudness and Texture. Proceedings of the Eighth International Conference on Information Visualisation. 2004.

Samman S., Stanney, K., Sims, V. (2005). Think Multimodal to Maximize Information Management Capacity. HCI International 2005.

Shneiderman, B. (1998). Designing the User Interface: Strategies for Effective Human-Computer Interaction, 3rd Edition, Addison Wesley Longman Inc.

Spear, B. D., T. R. Lakshmanan (1998). The role of GIS in transportation planning and analysis. Geographical Systems. Vol. 5, p. 45-58.

Stefanucci, S., Proffitt, D. (2005). Multimodal Interfaces Improve Memory. HCI International 2005.

Treviranus, J. (2000). Adding Haptics and Sound to Spatial Curriculum. IEEE International Conference on Systems, Man, and Cybernetics.

US Dept. of Transportation (2004). (Fed. Highway administration): Planning for transportation in rural areas: Rural Transportation Planning, http://www.fhwa.dot.gov/planning/rural/planningfortrans/index.html. (accessed 10-14-04).

Virzi, R.A., Sorce, J.F., Herbert, L.B. (1993). "A Comparison of Three Usability Evaluation Methods: Heuristic, Think-Aloud, and Performance Testing," in *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, Human Factors and Ergonomics Society, Santa Monica, CA, pp. 309-313.

Wall, S., Brewster, S.A. (2003a). Assessing Haptic Properties for Data Representation. CHI 2003.

Wall, S., Brewster, S. (2003b). Scratching the Surface: Preliminary Investigations of Haptic Properties for Data Representation, Eurohaptics, Dublin, Ireland, pp. 330-342.

Wickens, C.D. (1984). Processing resources in attention. In R. Parasuraman & D.R. Davies (Eds.), *Varieties of attention*. (pp. 63-102). New York, NY: Academic Press.

Wickens, C.D. (2002). Multiple resources and performance prediction. Theoretical Issues in Ergonomics Science, 2002, vol. 3 no. 2, 159-177.

Zhao, H., Plaisant, C., Shneiderman, B., Duraiswani, R. (2003). Sonification of geo-referenced data for auditory information seeking: Design principle and pilot study, University of Maryland Tech. Report HCIL 2004-03.