Multimodal Web-GIS:
Augmenting Map Navigation and Spatial Data Visualization with Voice Control

R. Daniel Jacobson and Ken Sam
Department of Geography
University of Calgary
2500 University Dr. N.W.
Calgary AB T2N 1N4

Telephone: +1 (403) 220 6192
Fax: +1 (403) 282 6561
Email dan.jacobson@ucalgary.ca | ksam@ucalgary.ca

Abstract

This paper describes the design and architecture of a prototype project that was implemented to augment the navigation and visualization of geospatial data for a web-mapping application. It leverages existing commercial off the shelf (COTS) web-mapping component and open specification Speech Application Language Tags (SALT) as building blocks for creating a multimodal web-mapping application. The voice-enabled multimodal web-mapping prototype application was developed using a combination of web, web-mapping, XML, and voice technologies. The application provides two types of interfaces. The first interface is the graphic user interface (GUI) that is exposed through the map viewer and the customized user controls for the application. The second interface is the voice user interface (VUI). The prototype application developed incorporated a set of voice-enabled commands that lets a user navigate the map without using keyboard or mouse. The prototype implements "crisp" voice commands – vocal equivalents of visual interface functionality and also "fuzzy" spatial language. Although imprecise in nature the spoken language feedback in the interface offers a clear economy of scale, providing an alternate level of precision through the exclusion of superfluous information. The prototype offers a platform for exploring the utility, tradeoffs, problems, and potential benefits when utilizing spatial language in a web based GIS.

1. Introduction

Geospatial data is primarily displayed visually to convey information to its end users. This has been the traditional paradigm of data visualization in the computer world. Efforts have been made in the software industry to better navigate and use software products. However, much more improvements are needed in the GIS software industry to help promote the visualization and interpretation of geospatial data through other cognitive means. To help with the visualization of geospatial data and navigation of a GIS for the visually challenged as well as for regular users, we present a prototype of multimodal web-mapping that augments map navigation and spatial data visualization with voice control and spoken language feedback.

The prototype application developed incorporated a set of voice-enabled commands that lets a user navigate the map without using keyboard or mouse. There are 115 voice-enabled commands that a user can issue while in voice-navigation mode. The voice commands are categorized into three main categories. They include the navigation related commands, location query related commands, and application settings related commands.

While voice is not a replacement for the mouse and keyboard interface, its usage enhances and improves the accessibility and usability of an application. Through the use of audio as a supplementary interface a multimodal GIS helps enhance a user's navigation of data and induces cognitive stimuli for the end-user to visualize geospatial data. It is hoped that the multimodal approach of combining voice with other user interfaces for navigation and data presentation is beneficial to the interpretation and visualization of geospatial data and makes GIS easier to use for its end-users.

The design and usability of GIS applications has traditionally been tailored to keyboard and mouse interaction in an office environment. To help with the visualization of geospatial data and navigation of a GIS application, this project presents the result of a prototype application that incorporates voice as another mode of interacting with a web-GIS application. While voice is not a replacement for the mouse and keyboard interface, it can act as an enhancement or augmentation to improve the accessibility and usability of an application. The multimodal approach of combining voice with other user interface for navigation and data presentation is beneficial to the interpretation and visualization of geospatial data and make GIS easier to use for all users.

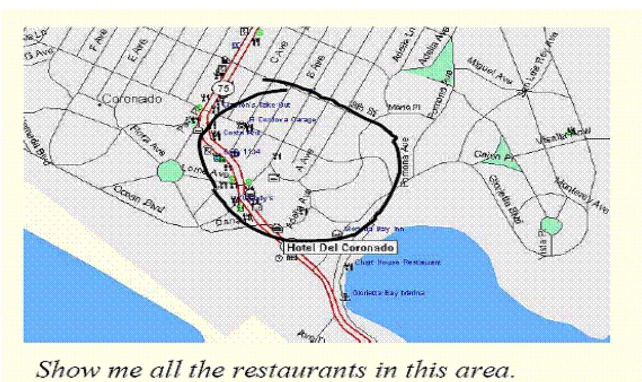
1.1 Background

Geovisualization has been increasing in human-computer interaction and flexibility, facilitating more complex exploratory data analyses and data mining. This evolution is contuing with further developments in multimodal interfaces (Golledge, Rice and Jacobson, 2005, 2006), geocollaboration (MacEachren Ref) and cybercartography (Taylor, 2003). There have been a number of research efforts in the past for data visualization on new ways of displaying data using (geo)graphic metaphors and interfaces. Internet technology has made information more accessible to many users, but most of the software applications created today do not take into consideration of the additional needs of users who are visually impaired or people who do not work in an office environment. The techniques used to interact with computers, and more generally, with digital content, have not changed much in the last 15 years. The WIMP (windows, icons, menus, pointer) paradigm using mouse and keyboard is still by far the most common human-computer interface (Rauschert et al., 2002) The focus on geospatial data visualization has been focused on the visual display. A visual display is synoptic like in nature, well suited to geospatial information is appealling and captivating; it draws the eye and engages the brain (Buckley, Gahegan, and Clarke, 2000). For many problems in Geographic Information Science, it is the optimal solution.

Many software applications are designed under the assumption that their users are using the software in an office environment sitting behind a desk and using keyboard and a mouse to interact with the application. Such assumptions are becoming less valid with increasing amounts of work being undertaken in the field or while in transit, using a variety of smaller personal computing devices. Recent legislation in the US has mandated that information and IT applications be more accessible to all users. In the US government agencies, section 508 of the Americans' with Disabilities Act (ADA) requires that Federal agencies' electronic and information technology is accessible to people with disabilities. It is believed that multimodal interface would facilitate human computer interaction for people who do not have full use of vision and people who do not use software applications in a traditional keyboard and mouse in an office environment. Multimodal interfaces are generally defined as an interface that employs two or more modes of interfacing with a computer system (Oviatt, 2003). It provides multiple modes of interfacing with the application and its data. It allows the user to interact with the application by a variety of interfaces, for example, keyboard, pointing device, touch screen, gesture, and voice. Through multimodal interface design, it is believed such human computer interface would allow users to better visualize and interpret geospatial data.

W3Cs Multimodal Interaction Working Group is currently in the process of defining the standards for multimodal interaction (W3C, 2006) . A multimodal interface may contain usage of GUI, Ink, Voice, Pointing device, and gesture. The diagram below shows a use case of using voice, ink and GUI to find locations on a map.

Figure 1. Multimodal GIS Use Case (source: W3C Multimodal Interface Working Group)



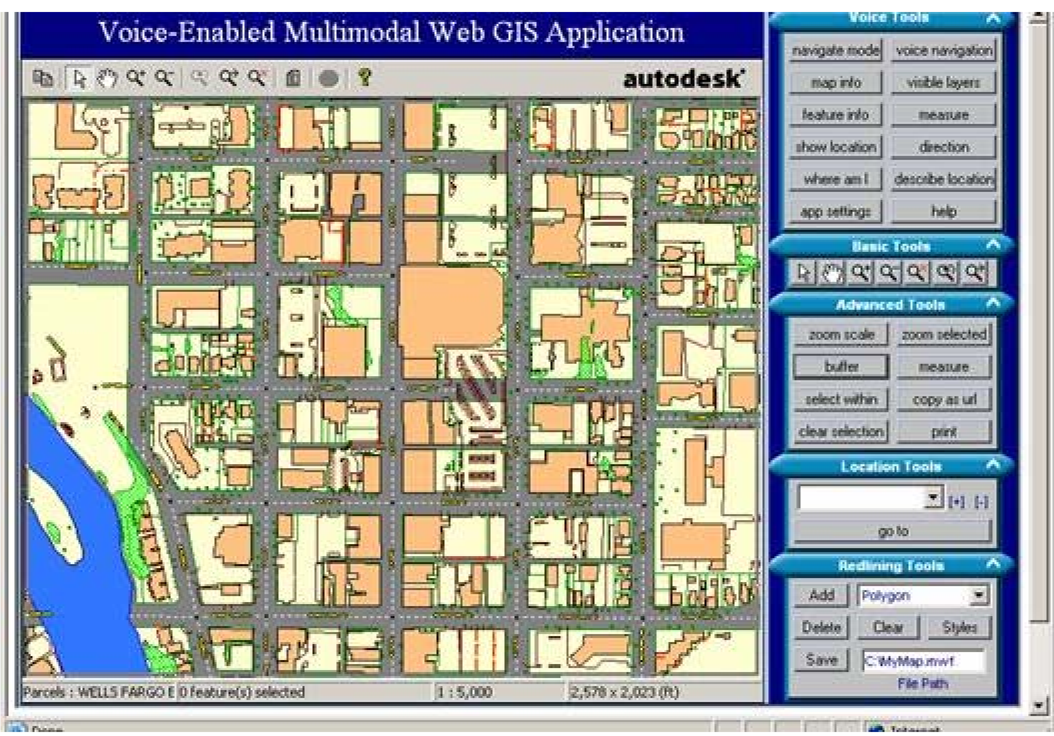
A multimodal interface needs to satisfy certain requirements (Dahl, 2002). For example, it needs to have the ability to let user select which mode to use and let the system constrain available mode. The system needs to use flexible mode of output, i.e. use output media in complementary fashion, for redundancy, and using media in a sequential fashion. The system needs to have flexible use of input modes. It allows user to choose mode(s) of preference and should have the ability to resolve inconsistent inputs. A multimodal interface should also support a wide range of devices and allow for multiple user/devices/servers to communicate. The system allows integrated and distributed architecture.

Multimodal interface provides a number of advantages to its users (Wilson, 2004). It allows parallel input - for example, the ability to use both key in commands and voice commands. It allows users to more quickly access and respond to information. In fact, multimodal systems don't just enable faster interactions, they also add value to the overall experience of interaction. Multimodal interfaces allow more room for user preference (giving users a choice of how they interact with the system) and reduce the overexertion that can result from single-modality interaction. Being able to switch between modes of interaction can lead to a lower incidence of error (because users can choose the mode most suited to different activities), as well as easier error recovery. And, finally, multimodal interfaces have the capacity to accommodate a wider range of tasks and environments. Recent research has indicated a variety of performance advantages associated with multimodal user interface. It's been observed that the advantages of multimodal interface include briefer task completion time, reduced errors, and a strong user preference to interact multimodally (Oviatt, 2003).

2. The Application

The voice-enabled web-GIS application focused on adding voice to the web-GIS application. Users can interact with the web-GIS by using either a graphic user interface with input from keyboard and a pointing device like a mouse, or they can use voice as an alternative means of interacting with the application. The voice-enabled multimodal web-GIS prototype application was developed using a combination of web, web-GIS, and voice technologies. The application provides two types of interfaces. The first interface is the graphic user interface (GUI) that's exposed through the map viewer and the customized user controls on the right hand side of the application. The commands available through the GUI are classified under their menu heading. They include the Basic Tools, Advanced Tools, Location Tools, Redlininig Tools and the Voice Tools. The development of the GUI tools mainly involved using Javascript and MapGuide Viewer APIs. Most of the GUI tools are integrated with voice feedback.

Figure 2. Screen Capture of Voice-Enabled Multimodal Web GIS Application



The second interface is the voice user interface (VUI). Most of the development efforts for this application were concentrated on the voice navigation mode section. The goal was to develop a set of voice-enabled commands that will let a user navigate the map without using keyboard or mouse. In this prototype application, there are 115 voice-enabled commands that a user can issue while in voice-navigation mode. The voice commands can be categorized into three main categories. They include the navigation related commands, location query related commands, and application/general settings related commands.

2.1 Data Set

In order to test the voice commands developed in the application, a map with data layers from a US city was used. The application is designed to work with any city level map that may contain street, building and parcel information. One of the aims of the prototype application is to develop a voice-enabled multimodal application that's data independent and the application can be used with any data set without much modification to the application, thus providing extensibility and an interoperable platform from which to base further developments.

2.2 Design and Robustness of Application

Voice interface is susceptible to background noise and incorrect user voice commands. In a multi-user environment, there is a certain degree of background noise that may introduce error to the system. The sensitivity of the microphone needs to be tuned to adjust to the users environment. This application includes grammar rules that are implemented for the recognition of voice commands so that only certain utterances issued by the user are recognized as valid voice commands. Depending on the setting of the user profile, the application would provide feedback to the user to indicate if the voice command is recognized as a valid command or not. If the error rate for the voice command recognition is high, then it would discourage users from using the application. One way to improve the robustness of the voice-recognition is by providing user the ability to use synonym voice commands.

2.3 Improving Responsiveness through Synonym Voice Commands

To help improve the responsiveness of the voice component of the application, it is important for the application to respond the voice commands that have similar meanings. In the researches that were done by the Geo-Collaboration Crisis Management team at the GeoVISTA Research Center at Pennsylvania Statue University, one of the areas identified for improvement of its Dialogue Assisted Visual Environment for Geoinformation (DAVE_G) is to improve the ability to process commands with a synonymous meaning (Rauschert et al., 2002). For example, when a user issued a voice command of give me information about this railroad, the system should also be able to handle a similar voice command like show me information about this railroad. The prototype web-GIS application has taken such recommendation in the design of the application. For example, when issuing a voice command to move the map to the right, the system can take any one of the six commands of move right, move east, go right, go east, pan right or pan east. In this application, there are a number of other voice commands that have synonym voice commands.

2.4 Improving Usability through User Profile Setting

One way to improve the usability of the application is through the use of user profile settings for the different type of users. Depending on the type of user, the application can apply a setting that best accommodate the preferred navigation and usage configuration according to the needs of the user. The prototype application includes three types of user profile settings. A user can say user profile one, user profile two, or user profile three to apply a set of pre-determined configurations for the application. Some of the settings for a user profile include the voice feedback, map feedback and size of the symbol representing the users location. Through the usage of user profile voice command, the application helps improve the usability of the application. It lets a user quickly adjust to a preferred setting for interacting with the multimodal application.

3. Application Components

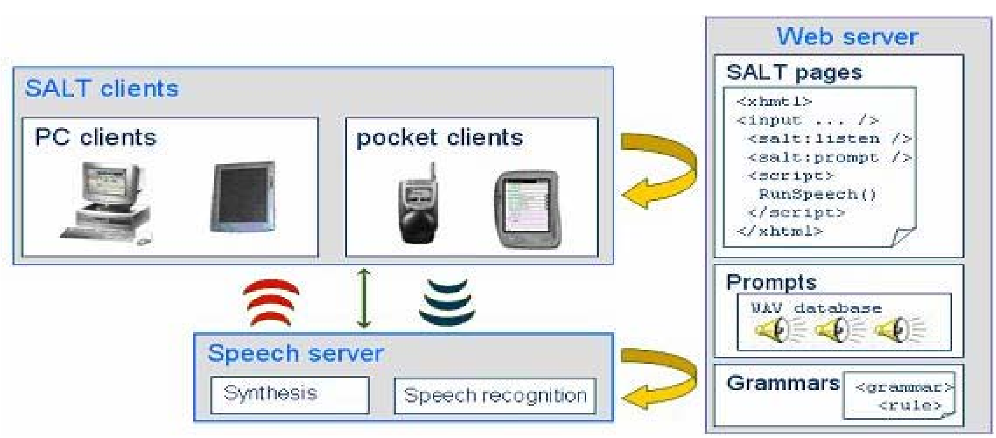
The Multimodal web-GIS application is mainly consisted of the three components:

1. Voice user interface using SALT.
2. GIS component using MapGuide.
3. Application program logics using client side JavaScript.

3.1 Voice User Interface using SALT and Text To Speech Engine

Speech Application Language Tags (SALT) is an extension of HTML and other markup languages. As defined by SALT Forum, SALT is "a small set of XML elements, with associated attributes and DOM object properties, events, and methods, which apply a speech interface to Web pages. SALT can be used with HTML, XHTML, and other standards to write speech interfaces for both voice-only (e.g. telephony) and multimodal applications." (Salt Forum, 2002). SALT is an open standard that's driven by a number of industry leaders like Microsoft, Intel, Cisco, SpeechWorks, Philips and Comverse. SALT is currently on version 1.0. It was officially released to the public on July 15th, 2002. Microsoft Speech Server SDK 1.0 is used for the development of this application. The SALT add-in is included with the installation of the Microsoft Speech Server SDK. For users who want to use the application, the SALT add-in available as a part of the SDK, or on its own, can be obtained from Microsoft's web site (Microsoft, 2002). A multimodal application using SALT is illustrated in Figure 3.

Figure 3. SALT Multimodal Environment (source: www.saltforum.org)



SALT mainly consists of four top level elements. Each top level element contains sub-elements that can be used to specify attribute values. The top level elements include: <prompt > for speech synthesis configuration and prompt playing ; <listen > for speech recognizer configuration, recognition execution and post-processing, and recording; <smex > for general purpose communication with platform components. Voice command recognition is done through SALTs listen element. Grammar or rules can be applied to the voice input to help with the recognition or determination of the correct voice input. The Text-To-Speech (TTS) component employs Microsoft's Text-To-Speech engine. Voice prompts and voice feedbacks are accomplished using the TTS engine.

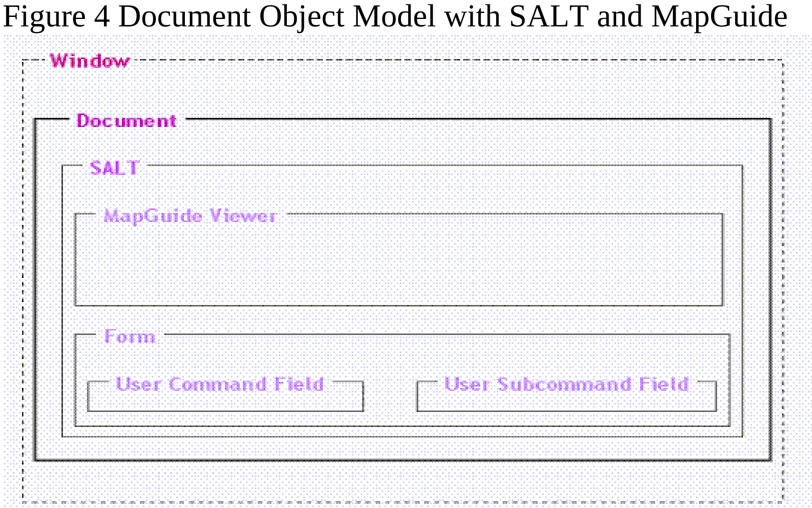
3.2 GIS Component using Autodesk MapGuide

This application uses Autodesk MapGuide as the web-GIS component for the prototype application. MapGuide has a number of components. For one to create a MapGuide web application, the following three components are needed: 1. Autodesk MapGuide Server for serving geospatial data. 2. MapGuide Author for creating or authoring the map. 3. MapGuide Viewer to be installed on the client's browser for viewing spatial data from the Server. The ActiveX Control version of the Autodesk MapGuide viewer is the viewer one would need to install on a system that uses the Internet Explorer browser. There are two main reasons for using MapGuide as the viewer. First, MapGuide Viewer provides a rich set of viewer APIs that offers client side customization through scripting language. Individual elements on a web page can be referred to using the browser's document object model (DOM). Secondly, MapGuide Viewer's APIs can be linked with SALT to offer a rich set of customized interactive GIS functions that are voice-enabled.

3.3 Programming Logics using JavaScripts

The programming logics for the voice commands and functions in this application mainly relied on JavaScript for the customized web-GIS functions and voice navigation. JavaScript provides the ability to interact with the web-GIS application through the MapGuide Viewer APIs. All objects embedded in a HTML page are accessible by traversing the DOM using JavaScript. The diagram below illustrates how SALT, MapGuide Viewer and Form elements being embedded in the HTML as a part of the document object model.

Figure 4 Document Object Model with SALT and MapGuide

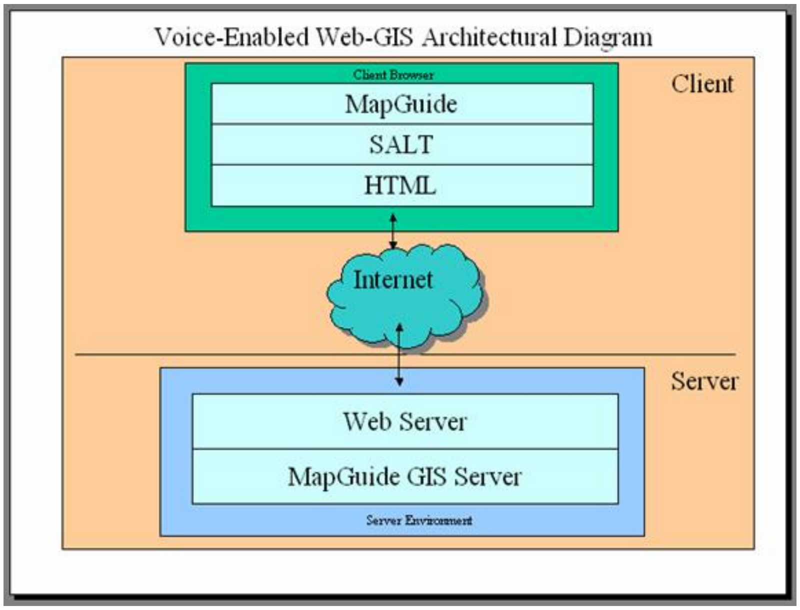


JavaScript also offers the ability to access SALT elements through the methods that are exposed by the SALT specification. This prototype application includes 95 level-one voice commands and 20 voice recognizable level-two sub-commands. The level-two sub-commands are used in conjunction with level-one commands to accomplish a task. For example, when a user issues a zoom scale voice command, the voice input is detected by a JavaScript function. The function calls the second component to let the user specify a parameter so that it can be used in conjunction with the zoom scale command. In the zoom scale example, one can specify a scale factor as the parameter. All the voice commands are bounded to textbox element of a HTML form. When a voice command is recognized, the value is entered into the textbox. The value from the textbox is subsequently accessed by JavaScript for additional processing or voice feedback.

4. Multimodal Web-GIS Architecture

This application mainly relied on MapGuide and SALT as the technologies for building the application. As illustrated by the architectural diagram below, the web-GIS application can be viewed as having two closely connected entities the Client component and the Server component. Spatial data for the application is served by a MapGuide Server through a web server by HTTP protocol. On the client side, GIS data is displayed by the MapGuide Viewer and the voice input and output for the application is serviced by the Microsoft SALT add-in. Both SALT add-in and MapGuide Viewer are wrapped by a web browser Microsoft Internet Explorer

Figure 5. Voice-Enabled Web-GIS Architectural Diagram using SALT and MapGuide



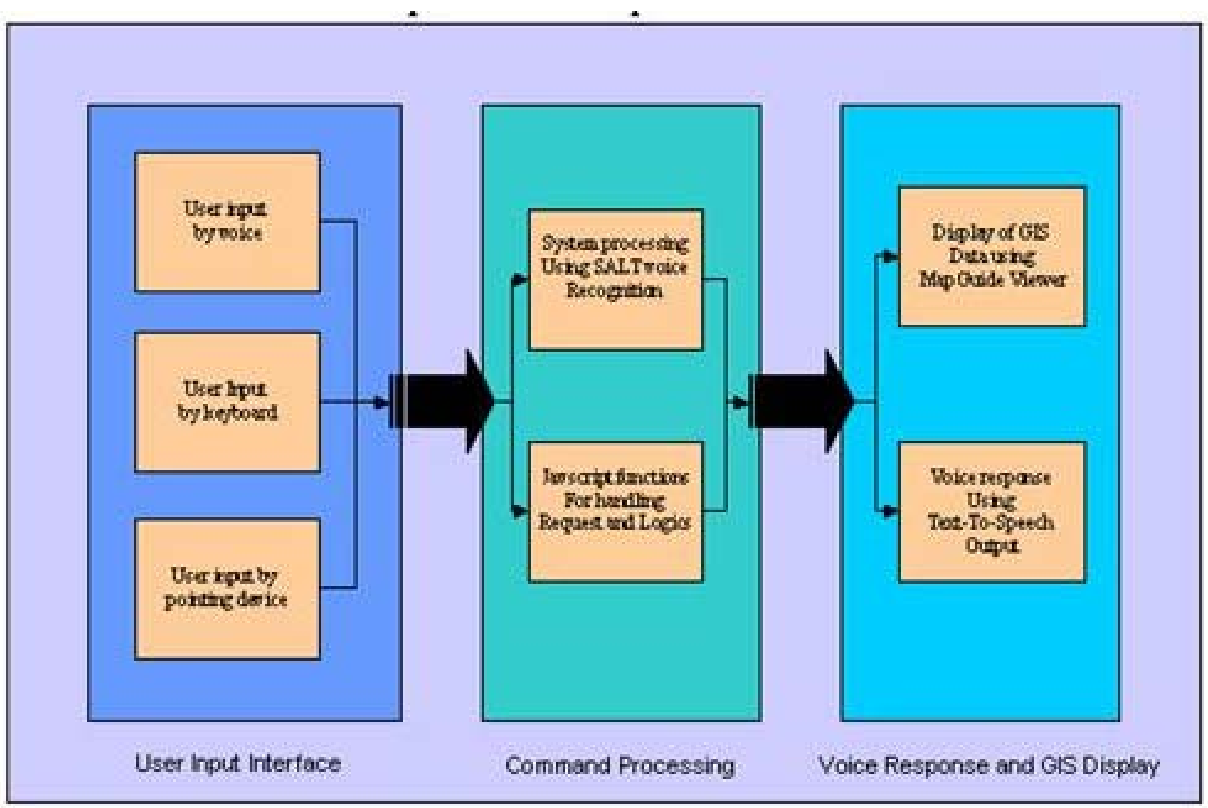
4.1 Voice-Enabling Application Using SALT

Upon loading of the application, the voice navigation mode is enabled by calling an initialize function using the onLoad event of the HTML's <body> tag. The application is launched into listen mode using SALT's listen element. This application mainly processes the user input through voice and graphic user interface. The user interface for input and output for this multimodal web-GIS application can be viewed as consisting of three components. They are grouped into logical grouping of User Input Interface component, Command Processing component, and Voice Response and GIS Display component. With this application, the command processing is handled by the processCommand() function. It acts as the receptor or handler for all the voice commands. On recognition of a voice command from the user, the application would call the processCommand() to hand off the request to other functions that services voice feedback or GIS related functionality. All the level-one commands are defined in the grammar rule of the getCommand listen tag. Grammar rules can be on the same page of the HTML or be pointed to from a source file. Besides the voice-enabling the application for all the voice commands, this application has taken into account of providing voice-feedback to the user while the application is not in voice-navigation mode. By augmenting voice as another feedback to the user while he/she is using a mouse and keyboard for navigation, there's a redundant level of information being feed to the user's cognitive system. It is believed that redundancy in communication can assist a user's reception and understanding of the information presented (Karpicke and Pisoni, 2000).

4.2 User Profiles and Use Case

The application was developed with three types of users in mind. The first group of users includes users who are sighted, but would benefit from a voice-enabled or voice-assisted interface to an application as warranted by use or domain specific tasks. Second type of users includes the users who are visually impaired and the third group are those who have no vision at all.

Figure 6. User Interface for input and Output of Multimodal Web-GIS



The multimodal web-GIS application developed for this project has taken into consideration of creating a user profile type for each type of user described above. A user profile setting helps the user adjust to the programs user interface by setting user configurable features for the application. User profile is especially important in an environment where heterogeneous devices are used to access all kinds off information and services. It allows both user and service provider the ability to customize ways of accessing services (Kotinurmi, 2001). Multimodal voice-enabled application can be beneficial to regular vision users as well as vision-impaired users. For example, a field technician who works for a power utility company may find voice navigation and voice feedback easier to use than the traditional keyboard and mouse interface. In many cases, a field technician may need to focus on task specific things such as connecting a power line rather than being tethered to a mouse and keyboard. In such situation, if voice commands are available to the technician as another mode of interfacing with a GIS, then the technician would not need to take his hands off the task at hand.

In a multi-user environment like an Emergency Operation Center (EOC), rapid access to geospatial information is crucial to decision-making (Rauschert et al., 2002). A multimodal interface not only allows users to work collaboratively, it also frees its users from the traditional unimodal interface of working with GIS. For example, if one were able to issue commands using voice and do not need to be constantly attached to a keyboard and sit in front of a monitor screen, then such user interface would enhance the productivity of the user.

5. Future Work

This project was a prototype project adding voice as another mode of interface for a web-GIS application. It is not to be viewed as a fully developed voice-enabled application that's ready for commercial use. This application prototype was developed in about two weeks. We believe that the application would benefit from additional refinement and usability testing. Furthermore, the application may be enhanced by future work in some of the areas identified below.

5.1 Integrating Voice with GPS Navigation

There are a number of GIS specific areas where voice augmentation would enhance the usability and accessibility of a GIS application. The auto industry has been leading the research in terms of in-car GPS navigation. There is ongoing research to help design better interface for in-car navigation system. Adding voice to GPS navigation is one area where a voice-assisted user interface would help with the navigation.

5.2 Server Side Processing

This application has mainly relied on the programming interface available on the client side for the interactivities of the application. While SALT and MapGuide Viewer technologies both are client side technologies, this application would also benefit by have server-side processing capabilities. For example, through server-side processing, it may be possible to create grammar rules that are generated dynamically for voice recognition. This would offer the multimodal application more robustness and interactivities. Also, having server-side processing would allow the application to connect to a database to provide additional attribute information to the user.

5.3 Integration of Gesture, Hand, Eye, and Head Movements

This application has mainly focused on adding voice as another mode of interacting with a GIS application. For an application to be truly multi-model, it should have the ability to respond to other modes of user input other than voice and keyboard/pointing device interface. Haptic, and gesture are other modes of user interface which would enhance the multimodality of the application. Simple and intuitive user interfaces are needed to give inexperienced users access to GISs (Schlaisichand Egenhofer, 2001). Additionally users' error rate can be lowered substantially when a multimodal interface is used for interacting with geographical information systems (Oviatt, 1996)

5.4 Technology Standards and Specifications

There are three main standards/specifications for voice-enabling applications today. They include SALT, VoiceXML and XHTML+VoiceXML (X+V) (Gannon, 2005). SALT is the specification that is used in this application. VoiceXML and X+V are two other technologies that are similar to SALT providing developers the ability to add voice to applications. VoiceXML has been adopted as a standard by the W3C. SALT and X+V and open specifications being proposed by industry leaders and they are both being proposed to W3C for standardization at this point. VoiceXML is the more mature technology available to developers today. However, VoiceXML is designed specifically for speech-based dialogues in telephony-based applications (Biber et al., 2005). SALT and X+V are specifications aimed at both phone applications and multimodal web applications. X+V is a specification being proposed by IBM, Motorola and Opera (Axelsson, 2001). It leverages XHTML and a subset of VoiceXML (McGlashan, 2004) as its specification for voice-enabling applications. This application has mainly focused on using SALT for the implementation of the project. Additional researches using X+V and other available technology specifications for multimodal interface will enhance our overall understanding of the topic and contribute to a more refined assessment of today's technology for multimodal web-GIS applications.

6. Conclusion

In this prototype project, Augmenting Map Navigation and Spatial Data Visualization With Voice, we have created a voice-enabled multimodal web-GIS application that uses both voice interface and graphic user interface for user input and output. Through the prototype application, we have shown how one can leverage existing web-GIS component and open specification to build a sophisticated multimodal interface for interacting with a GIS. We addressed the methodology and the technologies employed in the design and implementation of the project. The goal of the project was to demonstrate the ability to use voice as another mode of interacting, navigating and visualizing spatial data. The application built from this prototype project has demonstrated that GIS practitioners can use existing voice-enabling technologies to create multimodal applications today. It demonstrated that multimodal GIS application offers many potential benefits to a user. It not only helps visually impaired people to have a better understanding of spatial data displayed on a computer screen, it is also beneficial regular vision users by providing additional mode of interface to a GIS application. Multimodal application provides redundancy in both application input and application information feedback. Users who work in an emergency environment, or users who are out in the field that used to use hands in typing a message and use a pen, mouse and keyboard to interact with an application would find voice interface a valuable alternative to navigate and interact with GIS and mapping information.

7. References

Axelsson, J. (2001) XHTML+Voice Profile 1.0, Retrieved from http://www.w3.org/TR/xhtml+voice [last accessed April 20, 2006.]

Biber, C. Koziminski, A. and Langenberg, S. (2005) Six Steps for Creating a Speech Recognition Application or Speech-Enabling Your DTMF IVR. Speech Application Development, http://www.microsoft.com/speech/evaluation/newsletter/articles/0405article.htm [last accessed April 20, 2006.]

Buckley, A., Gahegan, M., and Clarke, K. (2000) Geographic Visualisation. UCGIS white paper. http://www.ucgis.org/priorities/research/research_white/ 2000%20Papers/emerging/Geographicvisualization-edit.pdf [last accessed April 20, 2006.]

Dahl, D.A., (2002) Standards for Multimodal interaction. Activities in the W3C Multimodal Interaction Working Group. http://www.conversational-technologies.com/images/tvuimmi.pdf [Last accessed April 20, 2006]

Duke, D. (2001) Cognitive and Computational Processes in Visualization, Paper presenated at Perceptually Adaptive Graphics (ACM SIGGRAPH / EUROGRAPHICS) Salt lake City Utah 2001. Retrieved from http://isg.cs.tcd.ie/campfire/davidduke.html. [Last accessed April 20, 2006]

Gannon, E. (2005) Multimodality: Simple Technologies Drive a New Breed of Complex Application Input, Devx.com magaxine, April 8, 2005, http://www.devx.com/wireless/Article/27878, [Last accessed April 20, 2006]

Golledge, R.G., Rice, M. and Jacobson, R.D. (2005) A Commentary on the Use of Touch for Accessing On-Screen Spatial Representations: The Process of Experiencing Haptic Maps and Graphics, *The Professional Geographer,* 57 (3). 339-349.

Golledge, R.G., Rice, M. and Jacobson, R.D. (2005) Multimodal Interfaces for Representing and Accessing Geospatial Information. In Rana, S. and Sharma, J. (eds.) *Frontiers of Geographic Information Technology.* Springer-Verlag.: Berlin & New York, 181-208.

Karpicke, J.and Pisoni, D. (2000) Memory Span and Sequence Learning Using Multimodal Stimulus Patterns: Preliminary Findings in Normal-Hearing Adults, Progress Report No.24 Speech Research Laboratory, Bloomington, IN.

Kotinurmi, P. (2001) User Profiles and Their Management, HUT, Software Business and Engineering institute, Retrieved from http://www.tml.hut.fi/Studies/Tik-111.590/2001s/papers/paavo_kotinurmi.pdf. [last accessed April 20, 2006.]

MacEachren, A. and Brewer, I. (2003) Developing a conceptual framework for visually-enabled geocollaboration, *International Journal of Geographic Information Science,* 17, 27-46.

McGlashan, S. (2004) Voice Extensible Markup Language (VoiceXML) Version 2.0, Retrieved from http://www.w3.org/TR/voicexml20 [last accessed April 20, 2006.]

Microsoft Corporation (2004) Microsoft .NET Speech SDK Version 1.0: http://www.microsoft.com/speech. [last accessed April 20, 2006]

Oviatt, S.L. (1996) Multimodal interfaces for dynamic interactive maps, in *Proceedings of Conference on Human Factors in Computing Systems: CHI '96,* New York, ACM Press,1996, 95-102.

Oviatt, S.L. *Multimodal interfaces.* In The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, J. Jacko And A. Sears,(Eds.) Lawrence Erlbaum Assoc., Mahwah, NJ. 286-304.

Oviatt, S.L., Cohen, P.R., Wu, L., Vergo, J., Duncan, L., Suhm, B., Bers, J., Holzman, T., Winograd, T., Landay, J., Larson, J. & Ferro, D. Designing the user interface for multimodal speech and gesture applications: State-of-the-art systems and research directions, *Human Computer Interaction*, 2000, vol. 15, no. 4, 263-322

Rauschert, I., Sharma, R., Fuhrmann, S., Brewer, S. and MacEachren, A. (2002) *Approaching a New Multimodal GIS Interface*, Paper presented at GIScience, September 25-28, 2002, Boulder, CO

Salt Forum (2006) http://www.saltforum.org/ [last accessed April 20, 2006.]

Schlaisich, I. and Egenhofer, M. (2001) Multimodal Spatial Querying: What People Sketch and Talk About**,** *1st International Conference on Universal Access in Human-Computer Interaction*, New Orleans, LA, C. Stephanidis (ed.), pp. 732-736, August 2001.

Section 508 of Americans with Disabilities Act: http://www.section508.gov/ [last accessed April 20, 2006]

Taylor, F. (2003) The Concept of Cybercartography, in Peterson, M. (ed) *Maps and the Internet*, Amsterdam: Elsevier, pp 405-420.

W3C Multimodal Interaction Working Group: http://www.w3.org/2002/mmi/ [last accessed April 20, 2006.]

Wilson, L. (2004) Look Ma Bell, No Hands! - VoiceXML, X+V, and the Mobile Device, XML Journal, August 3, 2004, http://sys-con.com/story/?storyid=45792 [last accessed April 20, 2006]