ZooMICSS: A Zoomable Map Image Collection Sensemaking System (The Katrina Rita Context)

Ross Graeber[1], Andruid Kerne[1], M. Kathryn Henderson[2]
Interface Ecology Lab
Computer Science Department[1], Sociology Department[2]
Texas A&M University, College Station, TX 77843, USA
rgraeber@tamu.edu, andruid@cs.tamu.edu, hendrsn@tamu.edu

ABSTRACT
Access to devices that integrate Global Positioning data with image and sound acquisition becomes more common, enabling people to build large collections of locative multimedia. As the size and number of these locative media collections grow, so too does the importance of systems that support collection sensemaking. Media semantics, which include automatically acquired location data, as well as user-supplied annotations, play a key role in these user-centered processes of collection utilization. This demo presents a Zoomable Map Image Collection Sensemaking System that enables the collection, organization, browsing, and annotation of locative images. The Zoomable Map Perspective is supplemented by event-based clustering. Dynamic views are generated automatically from captured media. The system is currently being used to document the location and condition of homes and neighborhoods in the aftermath of Hurricane Katrina.

Categories and Subject Descriptors
H5.2 Information Interfaces and Presentation (e.g., HCI): User Interfaces

General Terms Algorithms, Human Factors

Keywords digital collections, metadata, locative media

1. Introduction
Recent disasters in the Gulf of Mexico Region such as Hurricane Katrina and Hurricane Rita have displaced thousands of people. These people are separated from their homes, communities, and even their families. The recognition and rebuilding phase of hurricane recovery can be supported by collecting locative media and annotations to help evacuees tell their stories and re-form connections between one another and the homes they had to leave. This is a broad impact example of a type of phenomena that is increasingly common with the proliferation of digital cameras: photos are collected to represent significant events. Location is a key to understanding relationships between and meanings of the photos. The present research addresses these needs by developing a Zoomable Map Image Collection Sensemaking System (ZooMICSS).

ZooMICSS aids users in making sense of large collections of images with geospatial Global Positioning System (GPS) coordinates. Sensemaking is the process through which humans put together understanding of related information. Sensemaking has been said to involve changes in cognitive representations during a human information processing task [3]. *Collection sensemaking* involves understanding a collection of media entities, as a whole. One example of a sensemaking task is to compare the damage from Hurricane Katrina to homes, personal effects, and community buildings in different areas of New Orleans. Connected visual and semantic representations provide perspective to support users involved in collection sensemaking tasks. A zoomable map organizes images based on location at varying scales. Multiscale clusters based on zoom level organize images associated with events. The clusters afford contextualized thumbnail browsing and also maintain uniform information density on the map. Metadata enhances context and memory in the process of collection sensemaking.

MM'06, October 23-27, 2006, Santa Barbara, California, USA.
ACM 1-59593-447-2/06/0010.

2. System Description
The ZooMICSS system architecture is centered on the Locative Media Collection Repository (LMCR), which stores and organizes images and associated semantics (see figure 1). By "semantics", we mean relationships between images and their metadata, including cluster relationships. Three modules provide LMCR services: Image Acquisition and Storage, Semantics Processing, and Clustering. Image Acquisition and Storage (IAS) is responsible for the capture of images from the camera, and their binding with acquired metadata including GPS coordinates. Images are stored directly in the LMCR, while metadata and image references are passed to the Semantics Processing Module. The IAS module is written in Java and run from a mobile computer with attached GPS sensor and digital camera. Subsequent to acquisition, the Retrieval module extracts media and metadata from the repository for presentation and manipulation via the Sensemaking Interface. The Semantics Module processes metadata, binds it with images and clusters, and stores resulting data structures in the LMCR. The Clustering Module uses acquired metadata to create groups of images that are neighbors in space and time. These modules are run on a server

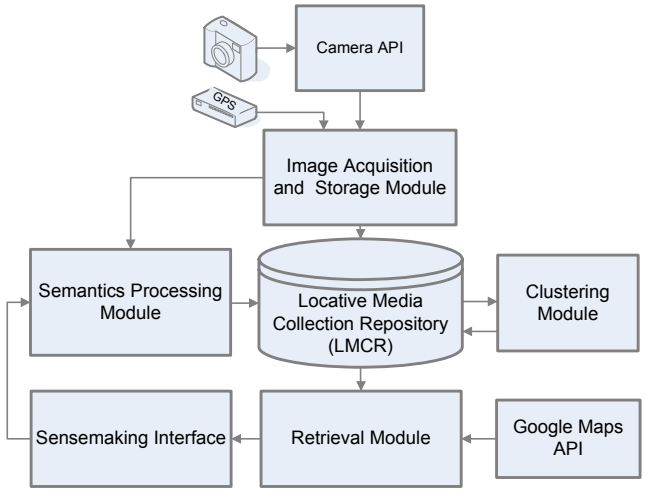

Figure 1: The ZooMICSS system architecture.

and coded in Java. The Zoomable Map Perspective, along with image and cluster browsing views, comprise the Sensemaking Interface, which allows participants to browse the collection, see relationships, and edit metadata associated with images and clusters. The interface is a web application coded in JavaScript utilizing the Google Maps API [2].

3. Application Design

When the user goes to take photos in the field, s/he brings a mobile device with the digital camera and GPS. With the Image Acquisition and Storage module running on the mobile device, the user simply clicks the camera's shutter. As the shutter is released, metadata from the camera and GPS are integrated and stored in relationship to the captured image.

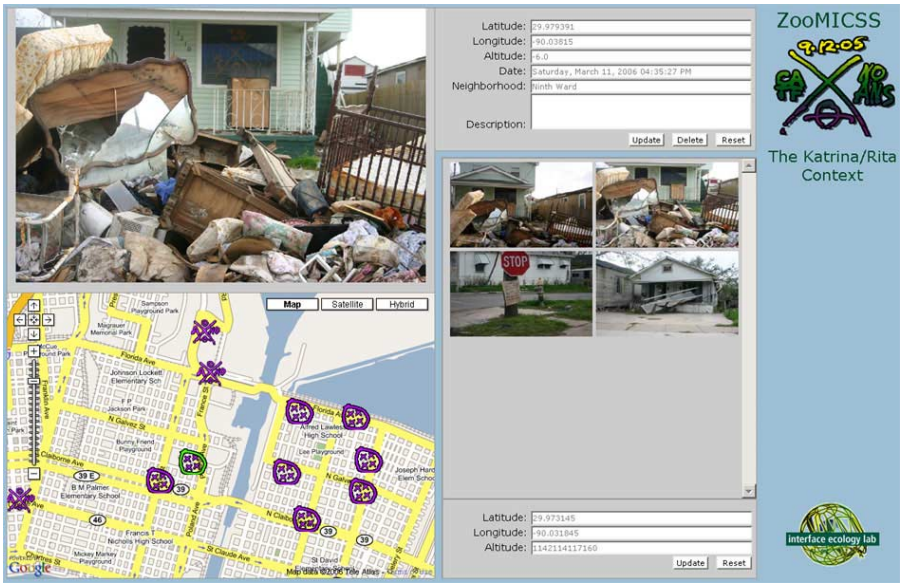

Figure 2: Locative Media Collection Sensemaking System: Personal possessions have been reduced to debris in the Upper Ninth Ward of New Orleans. The Zoomable Map Perspective in the lower left shows the Lower and Upper Ninth Ward. On the lower right, we see the cluster of which the selected image is a member, at the current zoom level.

Collection post processing is performed to create image clusters. Different cluster sets are formed for each discrete level of zoom provided by Google Maps, in order to maintain constant information density [5] and solve the occlusion problem. Without this multiscale clustering, as the user zooms in, image icons would pile up on the map, impeding the selection of individual icons for browsing. This shifting of representations to represent meaning appropriate to the selected zoom level is a form of semantic zooming [3].

Geospatial rules proscribe the performance of spatial joins from acquired GPS coordinates to GIS databases. Location is the key. This creates derived metadata, such as zip codes, and in turn, the labeling of each image with its neighborhood. We anticipate providing other derived metadata, such as flood water levels. Like the collaborative filtering employed by Davis, deriving metadata from GIS queries is a form of context-to-content inferencing [1].

The user accesses the Sensemaking Interface via a web browser. The interface consists of three panes: the Zoomable Map Perspective, the cluster view, and the image view (figure 2). The user starts browsing via the Zoomable Map Perspective, which affords panning and zooming. The map is populated with a set of icons. Each icon represents the images present in the currently visible region of the map at the selected zoom level. There are two icon types: image icons and cluster icons. Image icons are denoted by a single X-mark. Cluster icons are noted by the circular icon with multiple X's in them. When a user clicks on an icon, one of two things happens. Clicking on an image icon affords browsing the image and its metadata. With a click on a cluster icon, the cluster viewing pane on the right is populated with the thumbnails of the images that comprise the cluster. Meanwhile, on the map, the clicked icon is highlighted to make the focus visible. The map and cluster views provide two levels of context for collection sensemaking. The cluster thumbnails preview the images that are nearby. When an image is selected from the cluster, it is displayed in the upper left, with its metadata beside it. This provides additional context. Metadata may be entered to tell the stories that are associated with images. Clicking the image currently in the image module opens an additional browsing window with a full resolution version of the image, utilizing screen real estate available to users with multiple monitors. This view remains even when a new cluster and image are selected. The independence of the full resolution image view enables connection and comparison of information across contexts. An ethnographer using the system said, "This is an invaluable tool for connecting and analyzing images situated in context… For New Orleans residents and former residents it provides a grounding for telling their own stories in greater depth than might otherwise be possible."

4. The Katrina-Rita Context Collection

The Katrina-Rita Context Collection is currently composed of 252 images taken throughout the city of New Orleans approximately six months after the flooding caused by hurricanes Katrina and Rita. It includes photographs of the destruction caused by these disasters and of subsequent reconstruction. Some images show the sheer destructive force of the flooding, such as houses that were picked up by the waters and deposited in the middle of the street. Other images depict relief efforts. Through the Zoomable Map Perspective, the user can compare the progress of reconstruction in different parts of the city. Much of the Lower Ninth Ward remains a pile of rubble and devastation, while Lakeview residents are beginning to move back into repaired homes.

5. REFERENCES

[1] Davis, M., et al., From Context to Content: Leveraging Context to Infer Media Metadata. *Proc MM 04,* 188-195

[2] Google, Google Maps API, http://google.com/apis/maps/

[3] Perlin, K., Fox, D., Pad: an Alternative Approach to the Computer Interface. *Proc. CGIT 93,* 57-64

[4] Russel, D., Stefik, M., Pirolli, P., Card, S., The Cost Structure of Sensemaking, *Proc CHI 1993*.

[5] Woodruff, A., Landay, J., Stonebraker, M., Constant Information Density in Zoomable Interfaces. *Proc AVI*