Retrieving Multimedia Travel Stories using Location Data and Spatial Queries

Gamhewage C. de Silva
Interfaculty Initiative for Information Studies
The University of Tokyo
chamds@hal.t.u-tokyo.ac.jp

Kiyoharu Aizawa
Dept. of Information & Communication Eng.
The University of Tokyo
aizawa@hal.t.u-tokyo.ac.jp

ABSTRACT

We propose a system for retrieving multimedia related to a person's travel, using location data captured with a GPS receiver, mobile phone or a camera. The user makes simple sketches on a map displayed on a computer screen, to submit spatial, temporal or spatio-temporal queries regarding his travel. The system segments the location data and images, analyzes sketches made by a user, identifies the query, and retrieves relevant results. These results, combined with online maps and *virtual tours* rendered using street view panoramas, form multimedia travel stories. We present the system's current status and conclude with future directions.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces

General Terms

Algorithms, Human Factors

Keywords

Multimedia retrieval, Sketch-based querying, Travel diary

1. INTRODUCTION

People create travel photo albums to preserve the good memories of their trips and share the experience with others. Digital cameras, wireless networks and online authoring software facilitate creating a photo album very quickly. Most of the existing authoring systems assume that a user either manually selects the content or specifies previously selected images. However, a photo album for a particular trip is not always created immediately. Further, most digital camera users neither tag nor organize their photos [7]. As a result, selection or retrieval of the right photos to create an album takes considerable time. Automated photo retrieval using cues related to travel can therefore greatly reduce the users' effort in creating travel photo albums.

MM'09, October 19–24, 2009, Beijing, China.

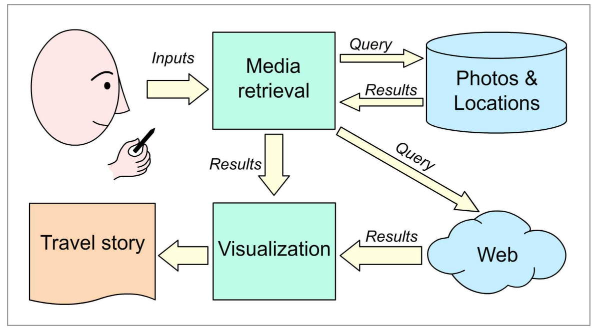

Figure 1: System overview.

Locations and movements are two important features related to travel. It is common to refer to a trip using locations and movements, as demonstrated by the following queries:

- "Photos taken around *German Corner*"
- "Photos that I took when I travelled from Frankfurt to Koblenz last year"

Ability to make *spatial queries* [2] such as the above makes photo retrieval faster and easier. Several new models of cameras are able to estimate the location and store it as metadata within the image or video file. Devices such as *iPhone* and *iPod Touch*, and cellular phone operators such as *NTT Docomo* of Japan provide support for continuous capture of location data. Online travel blog sites such as *EveryTrail* [3] facilitate creating "trips" combining photos and continuously recorded location data. Several researches have investigated the use of location data for multimedia retrieval [5][6]. However, intuitive and nonrestrictive input methods are necessary to facilitate entering such queries with desired precision (for instance, how close is "around" in the above query?).

This paper presents an interactive system that creates multimedia travel stories by combining photos and location data captured during travel (Figure 1). Unlike exiting systems that create static albums with manually selected content, this system automatically selects content for dynamic authoring. The user makes simple sketches on a map displayed on the screen, to specify the trip for which the story is to be created. The system interprets the sketches and retrieves the corresponding photos and location data. It also querries online multimedia collections for maps and panoramas. The data are combined using a visualization that cre-

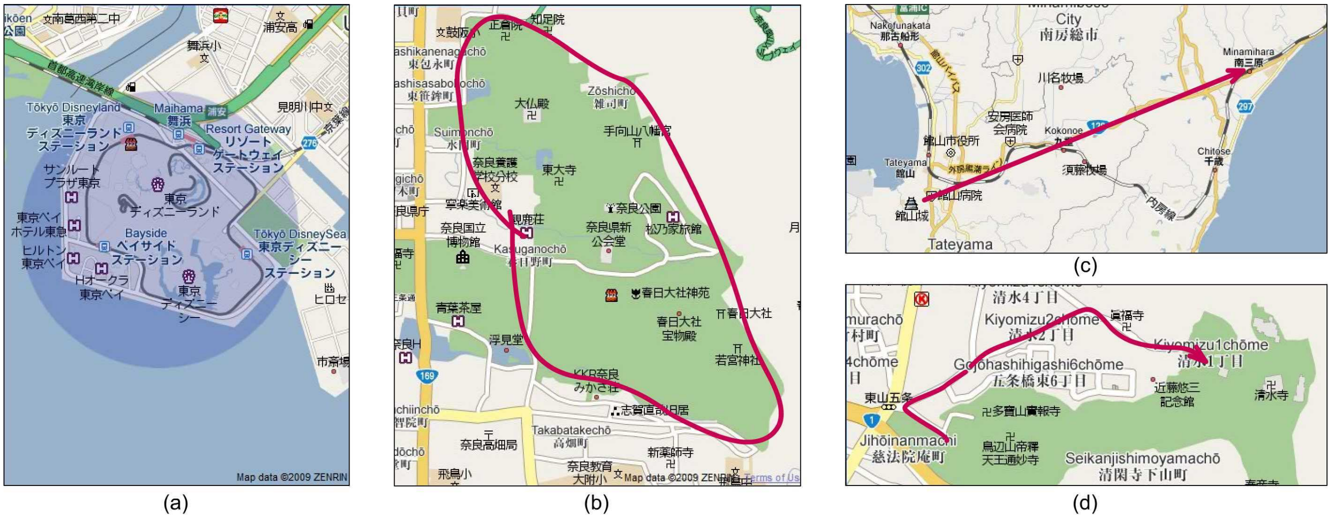

Figure 2: Different types of spatial queries.

ates a multimedia travel story that is easy to view. Two scenarios of location data acquisition are considered; continuous capture of data during the entire trip, and recording the location as image metadata when a photo is taken.

2. USER INTERACTION FOR QUERYING

We design the proposed system to support two types of queries. Spatial queries enable a user to specify information related to locations and paths. Temporal queries allow a user to specify the time interval for retrieval. It is also possible to submit spatio-temporal queries, combining both types.

2.1 Spatial Querying

The user submits spatial queries by making a sketch, using a mouse or a stylus, on a map displayed on the user interface. For this interaction strategy to be effective, the sketches should be intuitive, simple and unambiguous. We define the following three types of spatial queries to represent travel patterns:

- **Type 1: staying or traveling within a region**:

  We provide two methods to specify a region. In the first method, the user specifies the region by positioning a semi-transparent circle over it. The circle can be moved by clicking the left mouse button on it and dragging it. The size of the circle can be changed by rotating the mouse wheel. Figure 2a corresponds to the query "in Tokyo Disney Resort." For a more accurate specification of a region, the user can sketch a closed curve around it. Figure 2b corresponds to the query "in Nara Park."

- **Type 2: moving from one location to another, irrespective of the actual path taken**:

  The user draws an arrow from the originating location to the destination, by clicking the two locations in the correct order. Figure 2c corresponds to the query "From Tateyama castle to Minamihara."

- **Type 3: specific path**:

  The user draws the path that he/she wishes to retrieve, on the map. The path can be drawn either by making a freehand sketch, or clicking on points along the desired path to join them with straight line segments. It has to be drawn as an open curve, to prevent misdetection as a region. Figure 2d corresponds to a sightseeing walk in Kyoto, Japan.

The queries are sketched on images created by extracting map segments from the *Google Maps* Database [4].

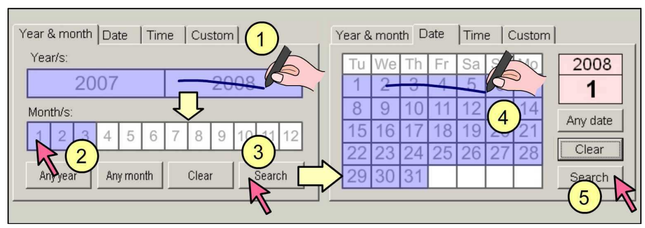

Figure 3: Sketching temporal queries.

2.2 Temporal Querying

The temporal queries also can be specified using sketches, making the interaction consistent with spatial queries. The users sketch on a calendar-like interface to select a duration to retrieve data from. Figure 3 shows how a user queries for the duration "from the 2^{nd} to the 5^{th} of January, 2008". Where only one item is selected, clicking can be used in place of sketching, enabling faster interaction.

Temporal querying is facilitated using two additional methods, to allow easier input and flexibility. The user can choose some frequently-used 'relative' time intervals directly from a combo box (Figure 4a). In addition to the above easy-to-use methods, controls for custom querying are provided in a separate tab (Figure 4b).

3. MULTIMEDIA RETRIEVAL

The first step of retrieval is to select images and location data for the time interval specified by the temporal query, if any. The resulting dataset is used as the domain for the spatial query. If continuous GPS data are available, the system groups the GPS data in to *locomotion segments* before

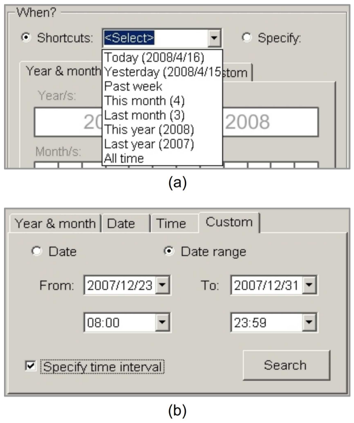

Figure 4: Other methods for temporal querying.

applying search algorithms. If location data are available only as image metadata, we cluster the locations using a self organizing map. The images in each cluster are ordered by time to form *sequences*.

3.1 Locomotion Segmentation

We combine the GPS data points to form a set of non-overlapping *locomotion segments* representing the nature of the person's movement. The segments belong to two classes: *Navigating* segments correspond to instances of locomotion where the person regularly changes his/her location with time; *Non-navigating* segments correspond to instances where the movement is zero or restricted to a small neighborhood. We employ a constrained hierarchical clustering algorithm [1] to cluster the data into locomotion segments.

The result of segmentation is a compact summary of movement that can be used as an index to the original data. Figure 5a shows a set of GPS data points recorded during a day, and Figure 5b shows the corresponding segments. The non-navigating segments are shown as circles with the mean location as the center and scaled standard deviation of noise as radius. The variable radius visualizes the confidence of the location estimation, allowing the user to identify the exact location with his memory and knowledge. The navigating segments are visualized with arrows. The starting segment for the given duration is shown in red on the map.

3.2 Search algorithms

- **Query type 1:**

  If continuous GPS data are available, we retrieve all segments (both navigating and non-navigating) that are contained within the region specified by the sketch (circle or closed curve). Otherwise, we retrieve all photos taken at locations within the region specified by the sketch. The results are ordered by the starting time of the segments.

- **Query type 2:** First, we define circular "search regions" around the starting and ending points of the sketch (in this case, an arrow). The center of each search region is the corresponding point on the sketch. The radius of the search region has been empirically set to the value of 20% of the geodesic distance between the two points. If continuous GPS data are available, we extract navigating segments that have starting points in the first region and end points in the second region. Otherwise, we search for sequences of images that start within the first search region and finish within the second region. The results are ordered by the starting time of the segments.

- **Query type 3:** We define a *search area* by expanding the dimensions of the bounding box of the sketched path by 10% in each direction. If continuous GPS data are available, we use a directional matching algorithm [1] to retrieve navigating segments that lie within this area and are similar to the sketched path, while preserving direction. Otherwise, the same algorithm is applied to find image sequences that are closest to the sketched path. The results are ordered by the similarity to the original sketch.

Figure 5: Results of locomotion segmentation.

4. VISUALIZATION

The retrieved images and location data are combined with online multimedia content to visualize a travel story that is more pleasing and easier to understand than a slideshow. Maps extracted from the *Google Maps* database are used as the base of the travel story.

For situations where continuous location data are available, the visualization is based on locomotion segments. Figure 6a shows the results of a spatial query. The locomotoion segments retrieved by the query are shown as circles and arrows, in the same format as in Figure 5. The presence of the camera-like icon on a circle or an arrow indicates that photos have been taken during the corresponding segment. The user can select each segment to view more detailed results. Figure 6b shows the detailed visualization of a navigating segment. The location data plotted on the map change from blue to red with time, indicating the direction of movement. Figure 6e shows the detailed visualization for a non-navigating segment (corresponding to photographing the New Year's sunrise). The photos are arranged chronologically on a spiral starting from the mean location of the segment,to minimize overlap. The user clicks on a thumbnail to get an enlarged view of the corresponding image.

For situations where only the location data within the image are available, the visualization is simpler. A marker is positioned on the map showing the location of each image. The markers are joined by lines in chronological order, with the line color gradually changing from blue to red. Clicking on a marker shows the corresponding image. Figure 6c shows a spatial query for photos taken within "Ueno Zoological Gardens". The results are shown in Figure 6d. The dotted line indicates the actual path the user took during this trip.

Figure 6: Example queries and results.

For type 3 spatial queries, we also create a movie-like 'virtual tour' of the path followed by the user. This is performed by estimating a set of locations and viewing angles by matching the location data with streets on the map, and then rendering a sequence of panoramas from the *Google Street View* database. Figure 7 shows a path resulting from a type 3 query and the corresponding virtual tour. The stick figure on the map shows the current location of the displayed street view. The user can view either the entire sequence (which is similar to driving along the selected path), or view the direction changes. While this functionality is currently restricted to areas with Google Street View, it can easily be expanded to use other similar services such as *Microsoft Virtual Earth*.

The system was tested on two photo collections. The first consisted of images captured using two digital cameras during 11 months. The location data for this collection were continuously archived using a handheld GPS receiver. The second photo collection was captured using an *Apple iPhone* device, with location data embedded in the EXIF area of the images. Preliminary evaluations showed that it was possible to create travel stories quickly and easily using both temporal and spatial queries.

5. CONCLUSION AND FUTURE WORK

We have developed a system for creating multimedia travel stories using digital photos and maps, with the aid of continuously captured or in-image location data. The sketch-based interaction strategy provides an intuitive and fast way to submit queries and retrieve results. The virtual tours created using street view can also be used as a travel guide.

The travel stories can be improved further by allowing the user to enter titles and comments and add images from other sources such as *Flickr* and *Panoramio*. The search algorithms can be improved to achieve better results. We are designing user studies for evaluating both the interaction strategy and the quality of the created travel stories.

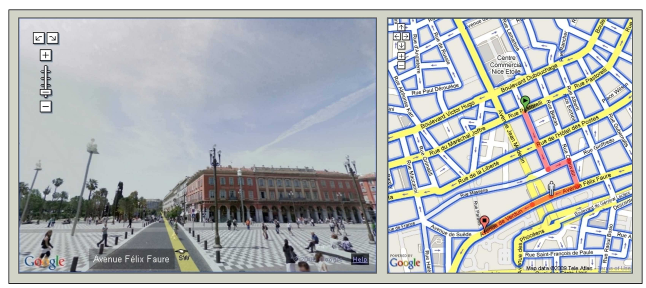

Figure 7: Virtual tour using street view panoramas.

6. REFERENCES

[1] G. C. de Silva, T. Yamasaki, and K. Aizawa. Sketch based spatial queries for retrieving human locomotion patterns from continuously archived GPS data. *To appear in IEEE Transactions on Multimedia*.

[2] M. Egenhofer. Spatial query by sketch. *J. Visual Lang. & Comp.*, 8(4):403–424, August 1997.

[3] EveryTrail. Everytrail - gps travel community. In *http://www.everytrail.com*, 2008.

[4] Google. Google maps api - google code. In *http://code.google.com/apis/maps/*, 2008.

[5] I. J. Kim, S. C. Ahn, H. Ko, and H. G. Kim. Persone : Personalized experience recording and searching on networked environment. In *Proceedings of ACM CARPE 2006*, October 2006.

[6] M. Naaman. *Leveraging Geo-Referenced Digital Photographs (PhD Thesis)*. Stanford University, 2005.

[7] K. Rodden. *Evaluating similarity Based Visualizations as Interfaces for Image Browsing*. PhD Thesis, University of Cambridge, September 2002.