A SCALABLE VIDEOGIS SYSTEM FOR GPS-GUIDED VEHICLES

Qiang Liu[1], KyoungHo Choi[2], JaeJun Yoo[2], and Jenq-Neng Hwang[1]

[1]Dept. of Electrical Engineering, Box #352500
University of Washington, Seattle, WA 98195
{liuq, hwang}@ee.washington.edu

[2]Electronics and Telecommunications Research Institute
161 Gajeong-dong, Yuseong-gu, Daejeon, Korea
{khchoietri, jjryu}@etri.re.kr

ABSTRACT

A VideoGIS system aims at combining geo-referenced video information with traditional geographic information in order to provide a more comprehensive understanding over a spatial location. Video data have been used with geographic information in some projects to facilitate a better understanding of the spatial objects of interest. In this paper, we present an on-going VideoGIS project, in which scalable geo-referenced video and geographic information (GI) are transmitted to GPS-guided vehicles. The hypermedia, which contains cross-referenced video and GI, are organized in scalable (layered) fashion. The remote user can request, through 3G mobile devices, the abundant information related to the objects of interest, while adapting to heterogeneous network condition and other factors such as display area size, CPU processing power, etc.

1. INTRODUCTION

The Geographic Information System (GIS) has been playing an important role in our daily life. It helps the system management of electricity, irrigation, water distribution, forest, route planning, transportation, etc. [1]. The map, either paper or digital, is a typical product of a GIS system and is widely used. A map often contains multiple layers of information, such as streets, landmarks, roads, rivers, etc.

As mobile computing environments are getting more popular due to the rapid advance of mobile communication technologies and widespread of mobile devices such as PDAs, new geographic information systems such as mobile GIS are being implemented. It is feasible today for a GPS-guided vehicle to receive geographic information (GI) through wireless communication. If the network bandwidth is allowed, additional layers of geographic information can be transmitted to the receiving device with a more detailed display.

When viewing a constantly updated map in a moving vehicle through a mobile device such as PDA, a user may get interested in viewing a related video corresponding to current spatial location or other spatial objects of interest. The user can switch to video mode after a target location is specified through the user interface. Also, the user can navigate from the video to other GI or video by clicking the hyperlink within the video frame.

By linking video data and geographic information, users can obtain more realistic information about geographic objects additional to some information that can be obtained using traditional GIS systems. It is very important to transmit video data and geographic information to heterogeneous users efficiently because most mobile devices in mobile environment have some limitations, such as varying network bandwidth, small display sizes, small memory size, etc. We are building a VideoGIS system, where we construct all the hypermedia data including geo-referenced video and GI and disseminate them to vehicles within a specific area. In this paper, we present the prototype of this system, including data preparation, organization, and adaptive transmission based on scalable layered coding and real-time end-to-end available bandwidth estimation.

The rest of this paper is organized as follows. We review some of the related work in Section 2. In Section 3, we present the framework of our VideoGIS system and how the geo-referenced video is constructed. Adaptive video transmission is discussed in Section 4. Descriptions of our primitive experiment results are presented in Section 5. Finally, Section 6 concludes the paper and discusses some of our future work.

2. RELATED WORK

Efforts to link video data and geographic information as a kind of hypermedia were started by Aspen Movie Map Project [2] to provide more visual and realistic data to users. The project used four cameras mounted on a truck and took an image every three meters on the streets of Aspen. The system consists of two videodisk players that allow users to “drive” through the city. In BBC Domesday Project [3], a map of Great Britain is presented and the users can watch (static) video clips from certain places. Shiffer presents a collaborative hypermedia tool for urban planning, where video clips showing the highway traffic can help planners to make decisions [4]. Christel and Olligschlaeger describe a project in the framework of the Information digital library, where geographical references are extracted from CNN news. The system allows to query for news from a specific place and to show a map of the corresponding location of the news story [5]. Navarrete et al. present a VideoGIS system that can segment and index video based on geographic information [6]. The system can provide videomaps and videoitineraries as an alternative interface to traditional maps for non-skilled users.

For the network transmission part, there has been a large body of work on congestion control for rate-based UDP applications. R. Rejaie, et. al. proposed RAP [7], where the sender controls the rate by using packet-loss and round-trip-time information. The rate increase/decrease is decided so as to be TCP-friendly. D. Sisalem and H. Schulzrinne proposed LDA, which is also a sender based adaptation scheme [8]. Packet loss and round-trip-time are feedbacked through RTP. Packet pair technique is used to estimate the bottleneck bandwidth. LDA applies dynamic determination of the Additive Increase Rate (AIR).

Jain and Dovrolis used *Pathload* [9] to estimate end-to-end available bandwidth. Their method is based on detecting the one-way delay increase trend, which is an indicator of congestion. The *Pathload* work is further extended by Liu and Hwang to allow the usage of packets with variable size, which can be the payload packets [10].

MPEG-4 fine-grain scalable (FGS) video codec [11], which is packet-loss resilient, has made it very simple and flexible for the video sender to adapt to the network dynamics.

3. VIDEOGIS SYSTEM FRAMEWORK

Our VideoGIS project is focusing on providing rich and cross-referenced video and geographic information to the mobile devices in GPS-guided vehicles. Currently the project has collected plenty of raw data from Dunsan city in Korea using a specially remodeled van. The video data are geo-referenced, i.e., they contain the spatial location information and also all the coordinates of the geographic object of interest in each frame. There were two cameras installed on the van to capture images of street blocks from two different angles. The van can receive GPS location information, which was recorded and incorporated into each captured image. The van was steadily driven through every major street in Dunsan city and two sets of pictures were taken each second (as shown in Fig. 1).

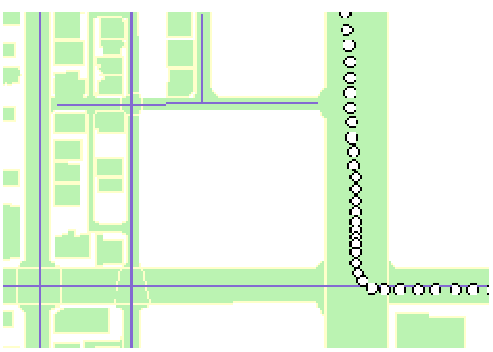

Fig. 1. The route of the van taking pictures. (Each circle in the figure corresponds to a picture capturing.)

The location information (obtained from GPS; in meter unit), camera parameters, and lens constants are recorded by the van for each video frame. Using these information, a post-processing program can extract the reference between the geographic objects of interest and the video data, such as the objects' ground coordinates (obtained from GPS; unit: meter), positions in the video frame (obtained by calculation; unit: pixel), and types. Fig. 2 shows the user interface of the post-processing program and Table 1 shows some samples of extracted reference data.

Table 1. Reference between GI and video data

Frame ID	ObjectID	Type	Position (in pixels)	Map Coordinate (in meters)
0	67108888	Signal Lamp	443, 164	235442.3, 317719.5
1	50331675	Street Lamp	394, 221	235474.8, 317816.0
2	33554460	Street Tree	249, 164	235502.9, 317858.7

Fig. 2. User interface of the post-processing program: map viewer (left-top window), video viewer (right window) and reference data extractor (center window).

Some other standalone videos, such as introduction to a building or the university campus, are prepared separately and can be referenced by the geo-referenced video. These video files are then compressed using scalable codec technique and can be transmitted adaptively based on the end-to-end available bandwidth to the receiver. We will discuss more about this part in the next section.

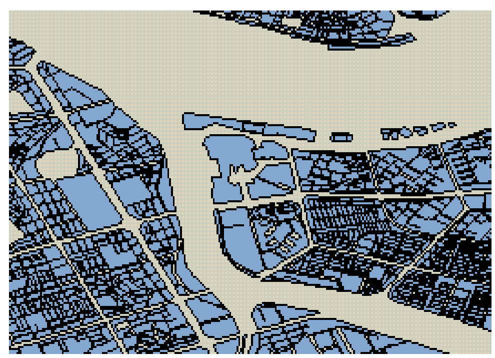

Fig. 3. Image showing the content of a shape file.

Some geographic information is saved in shape file format [12]. Shape file is a direct access, variable-record-length file in which each record describes a shape with a list of its vertices. Fig. 3 shows a typical image, which displays the content of a simple shape file. With more layers of GI, the user can have more detailed description regarding to a geographic object, such as a city block. Basically, the mobile device in a vehicle can receive map information regarding to its current spatial location or a specific location given by the user. The map is rendered to the display area of the mobile device and can follow the vehicle's movement. The level of the map details is determined by how many layers the device can receive. When the user finds an object of interest in the map, which has a hyperlink to a video file, the device can switch to video mode after user clicks the link. Similarly the user can also switch from video mode to GI mode by clicking hyperlink within the video display area. For our first prototype, we are going to use a single server to support these functionalities, even though the data source can be saved in a distributed way.

4. ADAPTIVE VIDEO TRANSMSSION

To support adaptive transmission, it is required to apply scalable coding for the video data. We use MPEG-4 Fine Grain Scalable (FGS) [11] coding scheme to encode the video data into base layer and enhancement layer. Only SNR scalability is used at this stage. The base layer is encoded using DivX encoder engine and the enhancement layer (residue) is encoded using DCT-based bitplane embedded coding. The bitrate of base layer is thus fixed and the enhancement layer is saved in macroblock (MB) units.

During the transmission, the base layer is sent out based on the bitrate defined in the encoding phase. The enhancement layer is sent in discrete levels adapting to the available bandwidth. The base layer video frame may be packetized into multiple packets if its size is larger than the target packet size, which can be configured by the user. The bitplane data are inserted into packets MB by MB. Once a packet exceeds the target packet size, a new packet is generated at the byte boundary of bitplane data. When the packet number reach the limit according to the corresponding target bandwidth (or if the bitplane data is exhausted), the sending of current video frame stops.

The receiver starts the transmission by only requesting the base layer data. Then the statistics of existing data being transmitted can be used to probe the end-to-end available bandwidth and more data can be requested if the bandwidth is sufficient [10]. During the probe, fixed numbers of packets are sent out at a specific rate (which is the rate to test). The one-way delay of these probe packets are recorded at the receiver and the receiver can test whether there is an increase trend among these one-way delays. As shown in [10], if there is an increase trend, it implies that the probe rate has exceeded the available bandwidth; on the other hand, if there is no increase tread, the available bandwidth is sufficient to support at the probe rate. The receiver can thus feedback this trend information to the sender, which is the server in our case, and let the sender to adjust its sending rate accordingly. The advantage of this algorithm for adapting sending rate is that it will not increase the sending rate continuously until packet loss occurs. As a result, the receiver can have a more stable video quality. The sender conducts the probe periodically and the probe period can be adaptive to the network condition and current sending rate.

When packet loss occurs, the receiver need to feedback the loss information immediately to the sender. The sender should drop its sending rate if the loss rate exceeds a predefined threshold. The sender can record the current sending rate when the loss occurs so that it can use a longer probe period when the sending rate approaches this rate again later. Besides the indication from the network congestion, the limitation of CPU processing power and display area size can also trigger a rate decrease.

Our GI data is organized in a similar way. Multiple layers of information can be fed to the receiver. Each layer has a priority according to its importance of geographic information. The layer with the highest priority is transmitted and displayed first. Users can adjust the layer priorities. The receiver can have a more detailed display with more layers received.

5. PRIMITIVE EXPERIMENT RESULTS

We have conducted some primitive experiments using current video and GI data available. For the proposed adaptive video transmission, we implemented and tested the transport protocol through various network conditions including LAN, Between ISPs, University to ISP, Across-Pacific, etc. The original video files are down-sampled to CIF and QCIF format to fit the limited display area of the typical mobile devices. The sender runs as a server and waits for receiver's connection. After being connected to the server, the receiver can choose a prerecorded video file, and define various protocol parameters such as target packet size, probe size (how many packets in a probe), default probe period, etc. During the session, the prerecorded video file is transmitted repeatedly from start to end. The client can view the receiving rate, loss rate, and frame rate throughout the session.

In most sessions, the receiver can find its optimal receiving rate quickly and keep at that level with some oscillations depending on dynamic end-to-end available bandwidth. Fig. 4 shows one QCIF video frame from a receiver at about 150Kbps. Fig. 5 and Fig. 6 show the short-term average of receiving bit rate and receiving frame rate for about 2 minutes.

Fig. 4. A QCIF video frame at about 150Kbps.

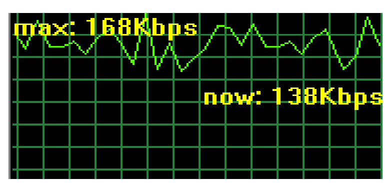

Fig. 5. Receiving bit rate.

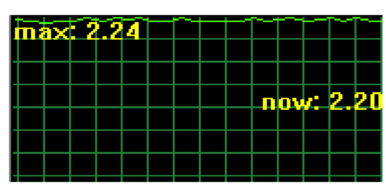

Fig. 6. Receiving frame rate.

In our experiments, the user can set the receiver to receive at a specific rate temporarily. By this way, one can monitor how the algorithm probes for spare bandwidth and backs off for congestion. The receiver can grab a reasonable share of bandwidth when there is competing TCP traffic (such as web page visit, file transfer, etc.) along the same path. However the fairness issue among various protocols is dynamically depending on several factors, such as round-trip-time, start sequence, various parameters of both protocols, etc.

For the geographic information, we have implemented a software package, MapView, for PDA. The user can choose to receive a specific layer of GI data, which is a shape file (see Fig. 7). The user can do feature identification, navigation (zoom in, zoom out, pan, etc.), scale-dependent display (such as Fig. 8), hyperlink to another file (document, GI, or video), distance measurement, labeling, etc.

Fig. 7. A screen snapshot for GI data rendering.

Fig. 8. Two city blocks at different scales.

6. CONCLUSIONS AND FUTURE WORK

We have presented a prototype of VideoGIS system for GPS-guided vehicles. Huge volume of geo-referenced video data and geographic information data have been collected for Dunsan city, Korea. The system can provide detailed and realistic information regarding to a specific location within Dunsan city. The mobile devices in the vehicles can real-time adaptively receive the video data depending on its available bandwidth with the server. The system supports cross-reference between the video part and GI part and the user can easily navigate between the two different modes. MPEG-4 FGS video scheme is used to prepare the prerecorded video files for adaptive transmission of video. Extensive experiments are conducted to test the performance. Also, a software package for receiving and rendering the GI data is implemented and tested.

We plan to integrate all the available parts into a working system in the near future. The transmission of video and GI data will be tested in a 3G mobile environment and the navigation between video and GI part is also to be realized as planned.

7. REFERENCES

[1] GIS URL, http://www.gis.com.

[2] Lippman, A., "Movie maps: An application of the optical videodisc to computer graphics", *SIGGRAPH'80*, 14(3):32-43, July 1980.

[3] S. Openshaw, C. Wymer, M. Charlton, "A geographical information and mapping system for the BBC Domesday optical disks", *Transactions of the Institutie of British Geographers*, 11, pp. 296-304, 1986.

[4] M.J. Shiffer, "Interactive Multimedia Planning Support : Moving from Stand-Alone Systems to the World Wide Web", *Environment and Planning B, Planning and Design*, vol. 22, pp. 649-664, 1995.

[5] M.G. Christel, A.M. Olligschlaeger, "Interactive Maps for Digital Video Library", Proceedings of *IEEE International Conference on Multimedia Computing and Systems*, pp 381-387, Florencia, Italia, July 7-11, 1999.

[6] T. Navarrete, J. Blat, "VideoGIS: Combining Video and Geographical Information", *5th AGILE conference on Geographic Information Science*, Palma, Spain, April 25-17, 2002.

[7] R. Rejaie, M. Handley, D. Estrin, "RAP: An End-toend Rate-based Congestion Control Mechanism for Realtime Streams in the" in Proc. *IEEE INFOCOM 99*, vol. 3, pp. 1337-1345, 1999.

[8] D. Sisalem and H., "The Loss-delay Based Adjustment Algorithm: A TCP-friendly adaptation" *Workshop on Net-work and Operating System Support for Digital Audio and Video*, July 1998.

[9] M. Jain and C. Dovrolis, "End-to-End Available Bandwidth: Measurement Methodology, Dynamics, and Relation with TCP Throughput," *IEEE/ACM Transactions on Networking*, Vol. 11, No. 4, pp. 537-549, August 2003.

[10] Q. Liu and J.N. Hwang, "End-to-end Available Bandwidth Estimation and Time Measurement Adjustment for Multimedia QoS", presented at *ICME 2003*, Baltimore, MD, USA, July 2003.

[11] H. M. Radha, M. van der Schaar, Y. Chen, "The MPEG-4 Fine-Grained Scalable Video Coding Method for Multimedia Streaming Over IP", *IEEE Transactions on Multimedia*, Vol. 3, No. 1, March 2001.

[12] ESRI URL, ESRI Shapefile Technical Description, http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf.