

GEODEC: ENABLING GEOSPATIAL DECISION MAKING

Cyrus Shahabi, Yao-Yi Chiang, Kelvin Chung, Kai-Chen Huang, Jeff Khoshgozaran-Haghighi, Craig Knoblock, Sung Chun Lee, Ulrich Neumann, Ram Nevatia, Arjun Rihan, Snehal Thakkar, Suyu You

Integrated Media Systems Center – Department of Computer Science,
University of Southern California, Los Angeles, CA 90089-2561

ABSTRACT

The rapid increase in the availability of geospatial data has motivated the effort to seamlessly integrate this information into an information-rich and realistic 3D environment. However, heterogeneous data sources with varying degrees of consistency and accuracy pose a challenge to such efforts. We describe the Geospatial Decision Making (GeoDec) system, which accurately integrates satellite imagery, three-dimensional models, textures and video streams, road data, maps, point data and temporal data. The system also includes a glove-based user interface.

1. INTRODUCTION

Recently, there has been a rapid increase in the availability of geospatial data, and this growth has motivated the construction of an information-rich and realistic three-dimensional (3D) visualization of a geographical location (e.g., a city).

The main challenge in such a task is how to accurately integrate and visualize all aspects of a geographic region, given the existence of heterogeneous data sources that have varying degrees of accuracy and consistency. Moreover, the user should be able to query the system and get information about the location that can then facilitate decision-making. Examples of information that could be used by decision-makers are: high-resolution imagery, maps, road networks, and 3D models of buildings. Such a system has applications in several domains, including urban planning, emergency response, online real-estate, simulation and training, computer games and army intelligence.

In this paper, we propose the Geospatial Decision Making (GeoDec) system, which is the result of applying relevant techniques developed independently in the fields of databases, artificial intelligence, computer graphics and computer vision to the problem described above. We present a general architecture for such a system and describe our current implementation. The current system not only allows navigation through a 3D model, but also issues queries and retrieves information as the user navigates through the area. In particular, the system seamlessly integrates satellite imagery, accurate 3D models, textures and video streams, road vector data, maps, point data, and temporal data for a specific geographic location. In addition, users can intuitively interact with the system using a glove-based interface. In the following sections, we discuss the specific application of this technology to the reconstruction of the University of Southern California campus located in Los Angeles, California.

2. RELATED WORK

Our proposed system has geospatial data integration, image/video pre-processing and visualization components. In the area of data integration, there has been some prior work on utilizing data integration systems for integrating geospatial sources [1][2][3]. While the other systems can retrieve different types of geospatial data from sources, they cannot integrate the data using different methods, such as integrating maps with imagery by conflating the maps to the imagery. Moreover, our integration system supports easy addition of new sources and more flexible integration.

In the area of image/video pre-processing in addition to providing a local map and possibly a satellite image of an area, the Google Earth project now offers 3D models of the area as well. However, unlike GeoDec, these models are typically based on LIDAR data. With LIDAR data, 3D building models are generated from high-resolution scanned range data [4][5]. The generated 3D building models are a mesh structure and do not make the structure explicit and need to be compensated for some holes in the generated models. The LIDAR data approach also requires accurate range data that may be difficult to obtain.

In the area of visualization, the main capability that differentiates our system from Google Earth is supporting fusion of video textures with the 3D model to create a compelling and realistic visualization of a geographic location.

3. GEODEC COMPONENTS

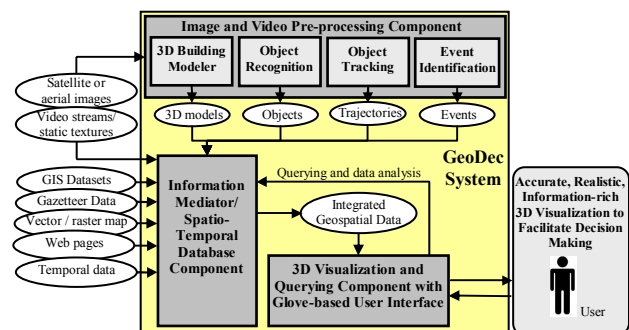


Figure 1: GeoDec System Architecture

As depicted in figure 1, GeoDec system is comprised of three main components: an information mediator with a spatio-temporal database, an image and video pre-processing component, as well as a 3D visualization component that supports queries through the

user interface. Our current system does not yet fully implement this architecture but nonetheless is an integrated end-to-end system comprising most parts of each major component.

The first step in our current system is to construct 3D models of buildings from (two) aerial photographs (not necessarily stereo images). In Section 3.1, we describe our system that facilitates the construction of the USC campus model rapidly and accurately with minimum human interaction.

The second step is to enhance the model by integrating different geospatial data from both online sources as well as private databases. This data includes road vector data and maps, gazetteer points, as well as temporal data, which are described in Sections 3.2.1, 3.2.2 and 3.2.3, respectively. In our proposed architecture, the system would query all the different geospatial data from a single mediator that integrates data from various sources. Our current system can obtain gazetteer point data as well as conflated vector and map data from the mediator

The third step is to complete the 3D visualization by mapping textures to the buildings, which includes fusing video streams (either live or stored) of the area with the models. Section 3.3.1 discusses the details of our approach. Moreover, since intuitive interaction is an integral part of decision making, we have also implemented a glove-based user interface described in Section 3.3.2 to enhance the user experience.

3.1. Rapid Construction of 3D Models

We have developed an interactive building modeling system that helps construct 3D building models from two or more aerial or satellite images. The system is designed to reduce user interactions and fatigue [6].

Even though the system uses two or more images, the images need not be true stereo images (i.e. they need not be taken at the same time) nor do they need to be rectified (i.e. have parallel epipolar lines); however, precise camera geometry is assumed to be given. The user interacts with only one image at a time and is not asked to fuse stereo images, reducing fatigue considerably.

The method consists of first approximating a complex building by a rectangular seed building. After each click, indicating the approximate position of a corner, the system finds nearby corners constructed from extracted line segments. This information is sufficient to construct a 3D model by combining information from the multiple views automatically. The user can then adjust the model height or sides using efficient editing tools.

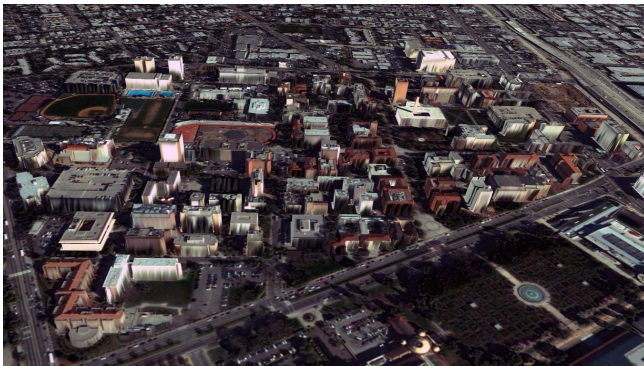


Figure 2: 3D Model Construction

Complex roof shapes can be modeled by adding or removing rectangular or triangular components. These modifications do not

typically require all three corners to be indicated; instead one side and one corner suffice. Multi-layered buildings are modeled from the bottom up so that new structures are extruded to the previous layer rather than the ground. By combining these operations, highly complex polygonal roof shaped buildings can be modeled.

Our modeling system is integrated in the ERDAS Imagine GIS software. Thus, it can handle a wide variety of image formats including commercial satellite images with accompanying RPC camera models. A modeling result with texture from our system for the USC campus is shown in Figure 2. In this case, 256 building components are generated in about two hours (user interaction time excluding image pre-processing time).

3.2. Geospatial Data Integration

3.2.1. Road network and map fusion

In order to generate a useful visualization of integrated geospatial data (i.e. vector data, satellite imagery, and raster maps), a simple superimposition of various geospatial sources is not sufficient to align the sources with each other for the following reasons:

- For many raster maps, the geocoordinates of the maps are unspecified
- For other geospatial data with specified geocoordinates, the projections and transformations used to produce the maps, orthorectified imagery or vector data are unknown

The current commercial tools to solve this problem require heavy user interventions. It is simply too slow and tedious to fully exploit the rich sources of information available in geospatial datasets. Towards this end, we have developed a set of techniques for automatically aligning maps and road vector data on orthorectified imagery. In order to allow integration of the aligned data with the 3D model described in Section 3.1, the maps and road vector data are aligned to one of the satellite images used to construct the 3D model.

Our vector-to-imagery conflation technique exploits a combination of the knowledge of the road network with image processing techniques. We first find road intersection points from the road vector dataset. For each intersection point, we then perform image processing in a localized area to find the corresponding intersection point in the satellite image. Finally, we compute the transformation metrics from the two sets of intersection points to align the vector data with the imagery. The running time is dramatically lower than traditional image processing techniques due to the limited image processing. Furthermore, exploiting the road direction information improves both the accuracy and efficiency of detecting edges in the image.

To integrate maps with satellite imagery, we utilize common vector datasets as “glue” [7]. We first identify road intersections on imagery by the technique described above, and then we also detect the road intersections on maps [8]. Finally, we apply geospatial point pattern matching algorithm to find matches between the two point sets. Now that we have a set of matched control point pairs, we partition the map into small triangles, and utilize the local transformation matrix to transform the map piece by piece. By doing so, we find the location (i.e. the geocoordinate) of the map and align the map with satellite imagery.

Our proposed approach facilitates the close integration of vector datasets, imagery and maps, thus allowing the creation of intelligent images that combine the visual appeal and accuracy of imagery with the detailed attribution information often contained in diverse maps. Figure 3 illustrates an area in the campus map of University of Southern California where a campus tram route map



Figure 3: Conflated USC tram route map fused with 3D model is conflated with the aerial imagery and overlaid on the 3D model.

3.2.2. Data Integration and Efficient Geospatial Querying

In order to efficiently and accurately integrate a wide variety of information about a geographic location, GeoDec utilizes a geographic data integration system called Prometheus [9]. Prometheus organizes the available information sources in a domain hierarchy containing well-known domain concepts, such as, satellite image, map, or vector data. Moreover, Prometheus also models different types of integration operations and their effects. Examples of the integration operations include *Overlay* and *Align*. The *Overlay* operation may result in two information layers not aligning with each other while the *Align* operation described in Section 3.2.1 improves the alignment between two data layers.

When GeoDec receives a request to retrieve the data for a geographic location, it can send a request to the data integration system to obtain different types of information (i.e. vector data, raster maps, and satellite imagery). Prometheus determines the relevant sources for the requested data, retrieves relevant data from the sources, performs necessary alignment or integration operations, and returns the integration information to GeoDec which can then be displayed on the 3D model.

3.2.3. Temporal Data Fusion

One of the key aspects of GeoDec is to integrate temporal data with other available sources of information. The fusion of temporal data sources significantly enhances the ability of a user to visualize the dynamic nature of an environment over a given time period. In order to add this capability to the system, GeoDec tracks all the trams and buses that move in or around the USC campus in real-time.

The geocoordinates of each tram are transmitted to a server every 10 seconds via modems installed on the trams. The system can thus track their location in real-time and displays the trams on the 3D model. In addition, the integration of conflated raster map of bus routes, road vector data and real time tram locations adds to the decision making capability of GeoDec.

3.3. 3D Visualization

3.3.1. Texture mapping and video fusion

Textures enrich the visual complexity of a scene with elements that are not captured in the static 3D geometry of a model. Textures capture both the static and dynamic appearance of the scene surfaces to make the 3D world appear realistic. In particular, video

textures provide spatio-temporal interpretation of real world activity, thereby aiding human comprehension and decision making. Our method requires the building or use of existing models of scene structures. Imagery is then projected onto these models and aligned or calibrated based on 3D and 2D landmarks and correspondences.

Static textures are captured by still or video images. The acquisition method and texture mapping process employs a base texture buffer and image warping transformations. Dynamic textures are captured as live video streams. We dynamically fuse video textures from network cameras on a 3D model to reveal the spatial and temporal relationships between the video streams.



Figure 4: Video Stream Fusion with 3D Model

The example shown in Figure 4 depicts five real-time video images projected onto the building models of the USC campus. In this case, the video images are overlapped to seamlessly cover the entire street and provide a complete visualization of scene activity. The video fusion methods integrated in GeoDec are based on the Augmented Virtual Environment system [10]. The mappings between the model and the imagery are calibrated in advance, using point and line correspondences between multiple images and the 3D model to derive the unknown camera parameters. Projective projection techniques are based on graphics methods for shadows and lighting. Additional technical issues that must be resolved include video decoding, visibility and occlusion processing, multiple projection blending, and real-time rendering. Our solutions to these and other issues, as well as methods for leveraging the latest graphics hardware features for real-time visualization are described in [11].

3.3.2. Glove-based User Interface

Gloves can provide a natural interface for various applications [12], and the GeoDec system includes a glove-based user interface for decision-makers to navigate and query this immersive and information-rich 3D environment using gestures. The goal of providing such an interface is to allow the user to intuitively interact with the integrated geospatial data on a large display. Examples of gesture-based navigation commands supported include zoom, translation and rotation. The user can also issue commands using the gloves to display different geospatial information integrated with the model such as the conflated vector data or tram map described earlier. Our system interprets user commands based on hand gestures obtained using a pair of 5DT Data Gloves and an Ascension Technology Flock of Birds Extended Range tracking device.

4. CONCLUSION AND FUTURE WORK

In this paper, we described a system that accurately integrates

heterogeneous geospatial data sources to create a compelling, realistic and information-rich visualization of a geographic location. Specifically, our current system fuses satellite imagery, 3D models, textures and video streams, road data, raster maps, point data, as well as temporal data. To achieve this, the system utilizes various techniques and further details on each technique, including the accuracy and performance evaluations, are described in the respective references. Our system also includes a glove-based interface to facilitate intuitive user interaction and decision making.

In the future, we plan to continue to develop our system based on the three main architectural components described earlier. The integration of these components to fuse various modalities (e.g., geospatial data, images, video, 3D models) is what makes our system different from traditional information integration systems. However, it is part of our future work to perform a thorough evaluation of our end-to-end system and compare it with more traditional information integration systems. We plan to enhance the decision making aspects of our system through the integration of additional data sources as well as improved query capabilities using efficient spatio-temporal data structures. We are also working on a new middleware layer that offers a universal way of specifying the type of query as well as its parameters, and retrieves the result in a standard format. Therefore any visualization layer can sit on top of this universal query layer for its integrated query and access needs. In addition, we are also working on the formation of more sophisticated queries and visualization of the results in the user interface. Other potential enhancements include the modeling of more sophisticated building structures (e.g. sloping roofs), extraction of geospatial data from text documents, building image schematic representations of the dynamics and actions of objects within the GeoDec environment (e.g. cars in an intersection) and subsequently learning the object behavior, as well as exploring multimodal user interfaces such as those described in [13].

5. ACKNOWLEDGMENTS

This research has been funded in part by NSF grants EEC-9529152 (IMSC ERC), IIS-0238560 (PECASE), IIS-0324955 (ITR), and unrestricted cash gifts from Google and Microsoft. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

6. REFERENCES

- [1] A. Gupta, R. Marciano, I. Zaslavsky, and C. Baru, "Integrating gis and imagery through xml-based information mediation," *In Proceedings of NSF International Workshop on Integrated Spatial Databases: Digital Images and GIS*, 1999.
- [2] S. Adali and R. Emery, "A uniform framework for integrating knowledge in heterogeneous knowledge systems," *In Proceedings of the Eleventh IEEE International Conference of Data Engineering*, 1995.
- [3] M. Essid, O. Boucelma, F.-M. Colonna, and Y. Lassoued, "Query processing in a geographic mediation system," *In 12th ACM International Workshop on Geographic Information Systems (ACM-GIS 2004)*, pp. 101–108, 2004.
- [4] C. Vestri and F. Devernay, "Improving Correlation-based DEMs by Image Warping and Facade Correlation," *In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 438-443, 2000.
- [5] C. Fruch and A. Zakhor, "Constructing 3D City Models by Merging Ground-Based and Airborne Views," *Computer Graphics and Applications*, pp. 52-61, 2003.
- [6] S. C. Lee, A. Huertas, and R. Nevatia, "Modeling 3-D Complex Buildings With User Assistance," *IEEE Workshop on Application of Computer Vision*, pp. 170-177, 2000.
- [7] C-C. Chen, C. A. Knoblock, C. Shahabi, Y-Y. Chiang, S Thakkar, "Automatically and Accurately Conflating Orthoimagery and Street Maps," *The 12th ACM International Symposium on Advances in Geographic Information Systems (ACM-GIS'04)*, November 2004.
- [8] Y-Y. Chiang, C. A. Knoblock, C-C. Chen, "Automatic Extraction of Road Intersections From Raster Maps," *The 13th ACM International Symposium on Advances in Geographic Information Systems (ACM-GIS'05)*, November 2005.
- [9] S. Thakkar, J-L. Ambite and C. A. Knoblock, "A Data Integration Approach to Automatically Composing and Optimizing Web Services," *In Proceedings of 2004 ICAPS Workshop on Planning and Scheduling for Web and Grid Services*, June 2004.
- [10] U. Neumann, S. You, J. Hu, B. Jiang, and I. O. Sebe, "Visualizing Reality in an Augmented Virtual Environment," *Presence: Teleoperators and Virtual Environments Journal*, Vol. 13, No. 2, pp. 222-233, April 2004.
- [11] S. You, and U. Neumann, "V-Sentinel: a novel framework for situational awareness and surveillance," *International Conference on Sensors, Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense III, SPIE Defense and Security Symposium*, 2005.
- [12] D. J. Sturman and D. Zeltzer, "A survey of glove-based input," *IEEE Computer Graphics & Applications*, pp. 30-39, January 1994.
- [13] N. Krahnstoever, S. Kettebekov, M. Yeasin and R. Sharma, "A Real-Time Framework for Natural Multimodal Interaction with Large Screen Displays," *In 4th IEEE International Conference on Multimodal Interfaces (ICMI 2002)*, 2002.