Tracking and Visualization

1. Research Team

Project Leader:	Prof. Ulrich Neumann, *IMSC and Computer Science*
Other Faculty:	Prof. Suresh Lodha, *UC Santa Cruz* Prof. William Ribarsky, *Georgia Institute of Technology* Prof. Promod Varshney, *Syracuse University* Prof. Suya You, *Computer Science, USC* Prof. Avideh Zakhor, *UC Berkeley*
Graduate Students:	Jinhui Hu, Bolan Jiang
Undergraduate Students:	Pamela Fox
Industrial Partner(s):	Airborne1 Inc.

2. Statement of Project Goals

This project addresses several research topics related to virtual reality, augment reality, geometry modeling, motion tracking, and visualization. The primary objective of this research effort is to devise new and innovative approaches and technologies that would improve the capabilities of information extraction, fusion, interpretation, and visualization from multiple sensor sources and datasets, suitable for time-critical applications.

Urban site modeling: While current sensing and modeling technologies offer many methods that are suitable for modeling a single or a small number of objects, an accurate large-scale urban model still remains costly and difficult to produce, requiring enormous effort, skill, and time, which results in the painfully slow evolution of such visual databases. This problem is the main impetus for our work. The objective of this research is the rapid and reliable creation of 3D models of large-scale urban environments such as city models. This research will benefit many applications including urban planning, geo-information systems, virtual reality, military training, and battlefield operations.

Hybrid motion tracking: This research front is for the development of innovated multi-sensor fusion technology that does camera motion tracking outdoors. The aim is to produce technology that is general-purpose, operates in unprepared-environments, and is feasible with current or near term technology. Particularly, we develop techniques towards the system requirements we consider essential in a hybrid tracking system. The effort of this research is complementary to current trends in NRL's BARS system. Other areas of potential use include virtual reality, HCI, and robot navigation.

Dynamic scene visualization: This research front focuses on the problem of dynamic visualization of multiple image/video/data streams from different sources and platforms. Image sensor can be on the ground as head-worn cameras, robotic panoramic cameras, aerial camera platforms, etc. They can be fixed or moving and their data can arrive in real-time or come from an archive. The problem we emphasize is providing some way for a person or group of people to digest and understand all this information easily with some sense of the spatial, temporal, and truthful nature of the information. The efforts of this research will benefit many applications in commercial, law enforcement, security surveillance, environment monitoring, traffic measurement, and military applications.

We developed an integrated prototype system by combing the research to illustrate the utility and benefits of our technologies.

3. Project Role in Support of IMSC Strategic Plan

The main goal and mission of IMSC is to create Immersipresence, or a state of highly efficient interactions between humans and machines. This project explores new methods for creating a highly comprehensible presentation of multiple channels of visual information. The techniques we develop may be extendible to GIS, security, and educational applications. Any time there are multiple views of a spatial setting, this approach has application. In the communication vision scenario, multiple views of each participant are treated as a case of multiple views of a single virtual space in order to perceptually-fuse the participants into one spatial setting.

4. Discussion of Methodology Used

We address a complete modeling system that can model complex building structures with irregular shapes and surfaces. Our approach is based on the use of airborne LiDAR which offers a fast and effective way to acquire models for a large urban environment. To verify and refine the reconstructed ragged model, we propose a primitive-based model refinement approach that requires minor user assistance. Given limited user input, the system automatically segments the building boundary, does the model refinement, and assembles the complete building model. By adapting a set of appropriate geometric primitives and fitting strategies, the system can model a range of complex buildings with irregular shapes.

Our tracking efforts focused on a hybrid-sensor tracking framework that integrates model-based tracking that combines extracted scene knowledge (such as artificial building models, global models such as LiDAR, DEM data, etc), natural occurring features (points and lines), auto-calibrated scene knowledge (online auto-calibration of point and line features constrained to the model database), and other sensor modules (such as INS and GPS). In this framework, the central unit will be the gyro sensor that continually reports real-time data streams in raw data form. These streams report the derivative of 3-axis orientation. The data are fed into a fusion filter combined with constrains provided by the model and tracked natural features for attitude computation and pose update. In our approach, the key feature is the compensation and correction of accumulated errors integrated into the attitude computation loop. The model constrains provide the necessary pose reference in the 3D world, and the natural features constrain the solution convergence to an optimum that achieves a stable and accurate pose estimate.

To cope with the aforementioned limitations of static models and visualizations of environments, we introduce the concept of an Augmented Virtual Environment (AVE) that has the capability to capture, represent, and visualize dynamic spatio-temporal events and changes within a real environment. The AVE is a comprehensive framework of data fusion, analysis, and visualization that incorporates and presents all the sensors, data, objects, and scenes within a common virtual world to produce a concise, coherent, and non-conflicting representation for time-space interpretation of real world scenery. In an AVE system, imagery sensors are modeled as "virtual projectors". The sensors can be static or dynamic and on different platforms, and their data streams can be live or recorded from archives of national or commercial sources. The geometry models of the environment are augmented with the projected images to create an augmented virtual environment that provides a visualization of both the geometry and dynamic imagery in a common 3D context. By projecting real-time video onto the 3D model substrate, the AVE not only adds the video information to the visualization, but adds it in the 3D context of the model, allowing dynamic images, events, and movements captured in video to be visualized in the same 3D context. The AVE is capable of presenting a single unified picture of real-time events from arbitrary viewpoints, offering a new framework for real-time data fusion and analysis.

5. Short Description of Achievements in Previous Years

We employed airborne LiDAR data to quickly collect 3D geometric samples of an environment. Through cooperation with Airborne1 Inc., an IMSC partner company, we acquired the LiDAR model of the entire USC campus and surrounding Coliseum area. This data is a cloud of 3D point samples registered to a world coordinate system. We implemented methods for LiDAR tessellation and assembly. Our model reconstruction approach can automatically processes the 3D point cloud and outputs the reconstructed 3D mesh model in VRML format.

We developed a texture projection method that enables dynamic projective texture mapping. Unlike traditional texture mapping where the mapping transformation is fixed during rendering, the texture projection approach paints the scene each frame with the updates of objects, camera pose, and viewpoints. The behavior of the dynamic mapping and update reflects the most recent information.

We constructed a basic testbed for testing and demonstrating our algorithm research efforts in a laboratory setting. The testbed comprises three parts that involve hardware and software components. The data acquisition system is a portable video and pose capture system. Two software systems include a geometry modeling module that is responsible for reconstructing 3D mesh models from LiDAR, and a visualization module that performs the visualization session including the dynamic texture projection.

5a. Detail of Accomplishments During the Past Year

Recent technical effort continues the algorithm and engineering research towards the system requirements we consider essential to an AVE. We built a prototype AVE system for the rapid creation of accurate wide-area virtual reality terrain and building models from multiple sensors. The models are used as a 3D substrate for the visualization of temporal image data, providing improved comprehension of scene activities. The core techniques we developed and integrated include model reconstruction, model refinement, building extraction, sensor tracking, real-time video/image acquisition, and dynamic texture projection for 3D visualization.

We developed a semi-automatic urban building extraction and refinement module. Given the limited user input, the approach automatically segments the building boundary, does the model refinement, and assembles the complete building model. By adapting a set of appropriate geometric primitives and fitting strategies, the system can model a range of complex buildings with irregular shapes. We demonstrated this system's ability to model a variety of complex buildings rapidly and accurately from LiDAR data of the entire USC campus.

We combined a differential GPS receiver with a commercial orientation sensor to produce a self-contained portable tracking system coupled to a video camera. The 6DOF pose of the camera is tracked by the sensors in real-time as the system moves in an open outdoor environment. The tracking data can be encoded and stored with the camera video stream in real time on the hard-drive of a laptop computer. The complete system is housed in a backpack for mobility.

Although the GPS-inertial tracking system provides an estimate of the camera pose that is adequate for some applications, its accuracy is inadequate for our AVE performance expectations. We overcome this problem by using a complementary vision tracker to stabilize the tracked camera pose. Vision tracking is also essential to overcome cases of GPS dropouts caused by satellite visibility occlusions. We developed a model-based tracking approach based on our previous auto-calibration work. The key feature of this method is that starting from a known initial estimate of camera pose (obtained from any method or sensors such as a GPS-inertial tracker), the camera pose is continually estimated using naturally occurring scene features (lines or edges) based on a prediction-correction strategy.

A key strength of the AVE system is its capability of dynamic data fusion of multiple spatio-temporal information sources from geometric models, images, video, and other sensing information. We developed techniques necessary for an AVE system that includes projective texture mapping, visibility and occlusion processing, real-time video projection, and image fusion for visualizing multiple images simultaneously.

We integrated the techniques described above and constructed a room-size 3D visualization environment. The system is used for demonstrations of the AVE concept and as a testbed for algorithm developments. The display consists of an 8x10 foot screen, back-projected by a sequential-frame stereo video-projector. A 3rdTech ceiling tracker is used to couple the rendering viewpoint to user's head position. A tracker also facilitates mouse-like interactions. The overall system provides the user with a high performance AVE visualization environment.

6. Other Relevant Work Being Conducted and How this Project is Different

The techniques we developed are novel and unique in many aspects. Most existing modeling systems facilitate geometric modeling through the manipulation of either standard geometric primitives, libraries of pre-modeled objects, or manual digitizing of key points. Creating models of a large environment in this fashion takes enormous effort, skill, and time, resulting in painfully slow evolution of such visual databases. Commercially available models of a city block, for example, can take several weeks to create even with several people working on it. Our modeling and refinement approach can rapidly model complex scene with irregular shapes and surfaces, suitable for producing large-scale urban models at modest cost. The dynamic texture projection was developed at SGI a number of years ago, and it's often used for video screen display in graphics scenes. Its use with tracked cameras for the purpose we pursue is the novel element we contribute. Similarly, tracked cameras are common in the augmented reality community. There are ONR and Columbia University projects pursuing GPS/inertial sensor tracking and image annotation. However, neither pursues the pose stabilization with vision. Furthermore, we first propose the concept of AVE visualizations of fusion multiple video with the 3D context of models, which offers the capability to capture, represent, and visualize dynamic spatio-temporal events and changes within a real environment.

7. Plan for the Next Year

Some ongoing and future work are: full automatic 3D model reconstruction and refinement; real-time pose tracking system with vision stabilization; extended landmark tracking combining with the scene knowledge and constraints; and some open issues related to AVE projection such as old projection texture persistence, and multiple video bandwidth requirements. In addition, we are planning to explore several potential areas to apply our AVE technology in commercial, law enforcement, security surveillance, and environment monitoring applications.

8. Expected Milestones and Deliverables

The developed technologies and system can be delivered in several forms including source code or libraries. The deliverable elements are summarized below:

- Backpack system for video and camera pose acquisition
- Software for hybrid vision, GPS, and inertial sensors tracking
- Software for automatic model mesh production from LiDAR data
- Software for complex building extraction and refinement
- AVE visualization system for dynamic image/video/model fusion

9. Member Company Benefits

The developed technologies can benefit many applications. Large-scale scene modeling is required for the fields of engineering, mission planning, urban planning, training simulations, entertainment, GIS systems, and military applications. Our work on LiDAR modeling is interested to LiDAR service providers like Airborne1 since we are using this data in novel ways that may lead to new markets for their services. Tracking is critical component in many applications, where require accurate estimating position and orientation of a camera moving freely in a wide area outdoor environment. Dynamic visualization and AVE projection has wide potential for various time-critical applications in commercial, law enforcement, environment monitoring, traffic measurement, and security surveillance. The visualization platform is developed with the assistance of Panoram Technologies. Our stereo display software and LiDAR data demonstrations may also open new markets for their display services.

10. References

[1] U. Neumann, S. You, J. Hu, B. Jiang, and J. Lee, "Augmented Virtual Environments (AVE): for Visualization of Dynamic Imagery", *IEEE Virtual Reality'03*, Los Angeles, 2003

[2] U. Neumann, S. You, J. Hu, B. Jiang, and J. Lee, "Augmented Virtual Environments for Visualization of Dynamic Imagery", to be appeared in *IEEE Journal of CG&A*, 2003

[3] S. You, J. Hu, U. Neumann, and P. Fox, "Urban Site Modeling From LiDAR", *II International Workshop on Computer Graphics and Geometric Modeling*, Montreal, 2003