Forensic 3D scene reconstruction

Charles Q. Little*, Daniel E. Small, Ralph R. Peters, J. Brian Rigdon,
Sandia National Laboratories

ABSTRACT

Traditionally law enforcement agencies have relied on basic measurement and imaging tools, such as tape measures and cameras, in recording a crime scene. A disadvantage of these methods is that they are slow and cumbersome. The development of a portable system that can rapidly record a crime scene with current camera imaging, 3D geometric surface maps, and contribute quantitative measurements such as accurate relative positioning of crime scene objects, would be an asset to law enforcement agents in collecting and recording significant forensic data.

The purpose of this project is to develop a fieldable prototype of a fast, accurate, 3D measurement and imaging system that would support law enforcement agents to quickly document and accurately record a crime scene.

1: INTRODUCTION

The Department of Energy's robotics program has been interested in rapid world modelling technology to perform large-scale three-dimensional (3D) mappings of unstructured environments [1][2]. The goal, in this case, is to enable deployment of autonomous robots. The model is used to help navigate the robot in a previously unknown area. As this study developed, the world model concept was amplified to handle additional data. In the case of waste cleanup, a number of other sensors can be deployed. This data can be displayed on the same 3D map to show sensor results at the position where data was collected. Video was the first sensor to be mapped onto the geometry. The map would form the basis of a virtual worldview of the workspace, with accurate geometry, video texture mapping, and other color mappings for additional sensors (chemical, thermal, radiological, etc.).

It became apparent that this technology was useful in other areas outside of robotics for general 3D scene mapping. In particular, enough interest was generated to fund this project for 3D mapping to aid in crime scene documentation.

This paper is a report of the work being performed at Sandia National Labs to develop a general 3D scene mapping system. It covers the requirements, technical challenges, and a discussion of the Sandia system.

2. REQUIREMENTS

To address the needs of a generalized system for 3D scene reconstruction, the following goals were established:

- Generic range data input
- Room sized crime scene imaging
- Ease of use
- Multi-view registration
- Image texture mapping
- 3D model creation tools
- 3D graphics display
- Analysis tools
- Fieldable (man-portable) hardware

The need for a generic input format was to enable a variety of sensors to be used. There are a number of 3D sensors on the market now, and more coming shortly, in addition to the ones built in-house at Sandia National Labs.

Full 3D models will need data gathered from multiple viewpoints. This is analogous to camera shots from various angles to gather more complete information. The major issue with this is registration, which is knowing where one data set is in relation to the next. There are various ways to register data that will be discussed later (see **Technical Challenges - Registration**).

There are two major aspects to 3D scene reconstruction. The first is to acquire sensors that are small enough and capable enough to gather all the data that's needed in the field by an agent. This project does not address this issue. The second, which we are addressing, is to make the software system intuitive enough to allow an agent to collect, display, analyze, and present the data quickly and accurately enough to be useful.

* Email: cqlittl@sandia.gov, desmall@sandia.gov, rrpeter@sandia.gov, jbrigdo@sandia.gov

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000.

DISCLAIMER

DISCLAIMER

3. PERCEIVED USES

The implementation of a fieldable system revolves around the perceived use. First, the system must be capable of collecting both image data and 3D geometry data. This data will likely be captured by commercial sensors, so the issue is to have a method to input the data into a common data format. Data collection is by definition a field operation. Operations after this point happen on a computer, and therefore can be done on site or back in a lab.

After the data is collected, it will be combined to form the 3D model of the scene. This will be a human directed computer task, due to the desire to allow the user to select the most appropriate data to use (selecting images for example), and the options of editing the workspace.

Once the model is built, the use of the system will move to presentation and analysis. The model will allow users to measure distances between objects. Additional objects, such as human models, will likely be inserted to probe placement and proximity questions to determine how and why. The scene will likely be used as the basis of a virtual reality viewer to present the data in real-time to an audience. The model can be very useful for presenting the probable scenario of the crime from the agent to the department, then to the district attorney staff, who will then present the case to a judge and jury. Showing the scenario can be much more effective than telling it.

4. TECHNICAL CHALLENGES

There are a number of technical challenges to consider when putting this system together, aside from the issues of what sensors to use.

4.1. Valid Data

The crime scene, practically speaking, only exists for a short period. After that, things are moved, cleaned, etc. The agents often only have one chance to gather data. Therefore, there is a need to verify that the data is good. Visually examining camera images at the scene is an example; this examination can happen with digital cameras or other types of frame grabbers. The use of 3D range sensors makes this need more pressing, due to the nature of range sensors. These systems are inherently sensitive to the scenes they scan, and therefore are adjusted on-site to account for varying conditions. Systems that rely on projecting and receiving a projected light, such as a laser based scanner, are affected by surface conditions. Ambient light can be a problem, especially direct sunlight. The color of objects can cause problems, since darker objects can absorb the light expected to return to the sensor. Surface reflectivity can be a big problem, both by causing the beam to never return to the sensor (no data) and by causing the beam to return via a reflection to another object (bad data).

The system must allow the user to preview the data at data collection to insure validity.

4.2. Adequate Coverage

The goal is to capture the entire scene. In a 3D sense this means getting the front, back, and sides of everything. A 3D range sensor does return 3D data, but it corresponds to more of a 2½ D scan, such as a relief map. To capture full 3D of a chair for example, a scan is needed from the front and back, and probably the sides. The best data from range sensors are from surfaces more perpendicular to it. In other words, from a single scan of an object, the front will be imaged well, the sides will be only partially captured, and of course the back is not captured at all. It will be necessary to move the sensor around the scene, as shown in figure 1. This represents a motel room layout. Here shots a and b may be enough to capture the nightstand on the left. Three shots, c, d, and e would be better to capture the three sides of the TV on the dresser.

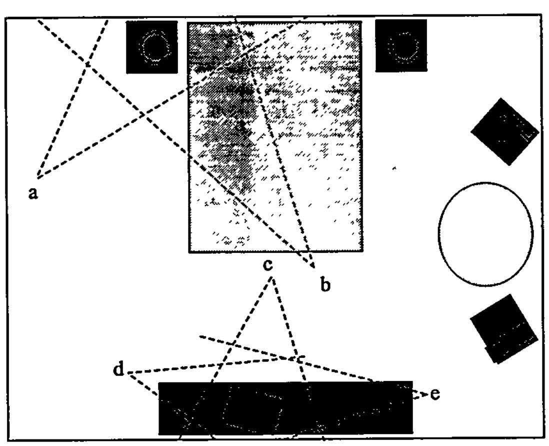

Figure 1. Multiple scans are needed to gather adequate data coverage

The easiest method to ensure adequate coverage is to put all the data together in a single view and let the user determine where the holes are. For simple scans this will be obvious as seen in figure 2. An automated system could also be implemented to help guide the user in placing the sensor.

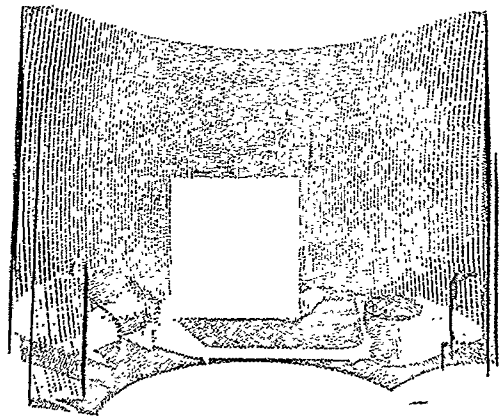

Figure 2. Hole shows incomplete range data of wall

It will not always be possible or practical to scan all surfaces, due to the inability or inconvenience in placing the sensor where it may need to go. These 'holes' may need to be filled manually.

4.3. Registration

A critical aspect of this project is the need for registration. All data, both image and range, are by nature relative to the sensor. If one knows the position of the sensor, the data can be placed in a common coordinate frame. Only by doing so does it make sense to combine more than one set of data. In the case of 3D range sensors, the data must be registered in 6 degrees of freedom (DOF): 3 positional (x, y, z) and 3 orientational (roll, pitch, yaw). Camera images are a 2D projection of the 3D world. In addition to the 6 DOF pose of the camera the internal camera parameters are needed to back project the 2D image into 3D space. Gathering this data is referred to as calibration, and can be done by a knowledgeable user at their site.

Registration can be done in several ways. One method, which is the one we prefer, is to register the sensor before the data is collected by tracking the sensor in 6 DOF at the time of data capture. Attaching the camera and/or range sensor to a physically tracked platform, or tracking the unit through the air can do this. We have experimented with a magnetic field system from Polhemus, and a lighted diode tracking from Northern Digital. Other systems can be found in the area of motion tracking.

A second method of data registration is to use information in the data itself to correlate with a second data set. This can only be done if the information exists in both data sets. For example, if two scans overlap sufficiently to contain three or more non-collinear static points, and these points can be found in both images, a transform can be derived to register one data set with the other. There are automated ways of doing this, provided again there is common coverage between the two scenes. One technique involves a technique called Iterative Closest Point[3], and another uses a simulated annealing approach[4]. The goal of both is to iteratively transform one data set and calculate a closeness factor, where the closest fit is the best. Figure 3 shows an example of registered versus non-registered data. There are two data sets in these views that overlap in the center. Although the fit may be close, it must be better for latter use.

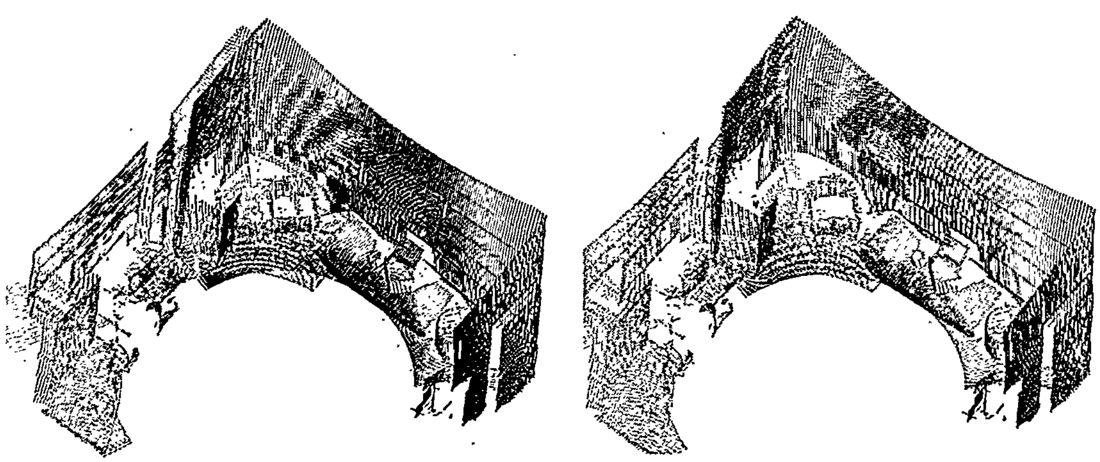

Figure 3. a) Non-registered versus b) registered data

4.4. Creating Surfaces

The output of range sensors are points in space. To be more useful for this project, this data is connected together to form a triangle mesh or surface. Generally, connecting nearest neighbors together can do this, but care must be taken to avoid problems. For large range gaps, such as moving from the back of a chair to the wall, a hole is generally left. This prevents the user thinking data was captured where it was not. Also, range data often contains error points that must be filtered out to avoid large surface errors from cluttering the scene. Usually it is best to create surfaces on a per scan basis to avoid difficulties in surfacing when multiple scans are combined.

4.5. Combining Data

As registered range data is combined, some editing techniques become desirable. For example, overlapping or redundant data builds up. This can be removed by filtering techniques, or stitching, which refers to seaming overlapping surfaces into a single surface. Often, however, it is desirable to divide a surface into multiple surfaces, or objects, that can be used later in analysis and in adding object specific reference information to the scene. This is called segmentation.

Another form of combining data is putting images onto surfaces, commonly referred to as texture mapping. This is analogous to projecting an image onto a wall. It this case, the wall is the surface created from the range data. For an accurate model the image must match the surface. Therefore the image data must be properly calibrated and registered to the scene. If so, each pixel will project onto the proper corresponding surface patch, thus the need for registration mentioned earlier. Figure 4 shows a surface map of a wall with a single image texture mapped onto it.

Figure 4. Image texture mapped onto a surface

4.6. Maintaining Credibility

We suspect the very act of creating a model, especially editing the model, will generate concern that data was fabricated, and therefore not admissible or useful in the courts. To mitigate this, it will be important to be able to track the data back to the original data gathered, both range scans and images. An integral part of the model should be the capability to trace each piece of data to its origins.

4.7. Hardware / Software for Fieldability

A fieldable system must be deployable. Depending on the scene, this can mean fully self-contained and self-powered. It should be light enough to be handcarried. There should be little or no calibration needed in the field.

Cameras have fit this requirement for a long time. 3D range sensors are generally not there yet, but are progressing rapidly. Portable computers are now becoming powerful enough to perform the computation needed in the field. Registration with external tracking is still problematic, both with equipment size and weight, and power needs.

4.8. User Interface

The 3D graphics nature of the system defines the interface to be primarily graphical. Data should be viewed in the inherent form in which it is collected. 3D dots and surfaces require an interactive fly-through type display. On the other hand, image data will need an image display screen to aid the user in finding images. Both of these types of interfaces are ubiquitous. The challenge is to make the interface intuitive to the user.

The system needs editing capabilities in dots as well as surfaces. It must handle texture mapping. It must also be capable of handling very large data sets. Scans can easily return over 100,000 points each.

5. DISCUSSION OF SANDIA SYSTEM

Our crime scene reconstruction system consists of a 3D range sensor, an imaging camera, and a software system for creating and viewing a 3D model of the scene.

One of the primary goals of the system is to build something that can be replicated with minimum cost. This is driven by the limited resources available to most law enforcement entities.

A decision was made early on to use the fastest graphics display we have available to visualize the data, anticipating the need to view many thousands of polygons. We therefore use SGI's as a platform. We recognized, however, the need to move to a PC for reasons of cost, portability and support. Higher graphics speeds will undoubtedly be forthcoming. Our solution was to adapt a graphics toolkit that is written to support both platforms, as well as using portable coding.

5.1. Sensors

For this project, standard analog video cameras have been used to capture still frames for 2D imaging. The current camera in the system is a lipstick case size color video camera from Toshiba. A frame grabber board from Matrox accomplishes image capture. Digital cameras could be added easily. In addition, film based imaging can also be used if sufficient control can be maintained in digitizing the resultant image. Basically, the type of camera or type of lens is not important, so long as the image can be digitized and camera parameters can be gathered.

For a 3D range sensor, this project has used a laser based structured lighting system made in-house, called LAMA, consisting of a video camera (Cohu) and laser (Lasaris 30mW) mounted on a pan & tilt unit (Directed Perception). The laser output is spread into a line, which is projected onto the scene and imaged by the camera. The unit sweeps this line through the scene to gather data. This type of sensor is referred to as a structured light system. Figure 5 shows the setup of the 3D range sensor and the camera. One important note is a requirement we have imposed to maintain eye-safe operation. The laser we are using is class 3a in its fanned beam configuration. There are a number of range sensors now available, using a variety of technologies. The goal as stated is to be able to incorporate other range sensors, as they become available.

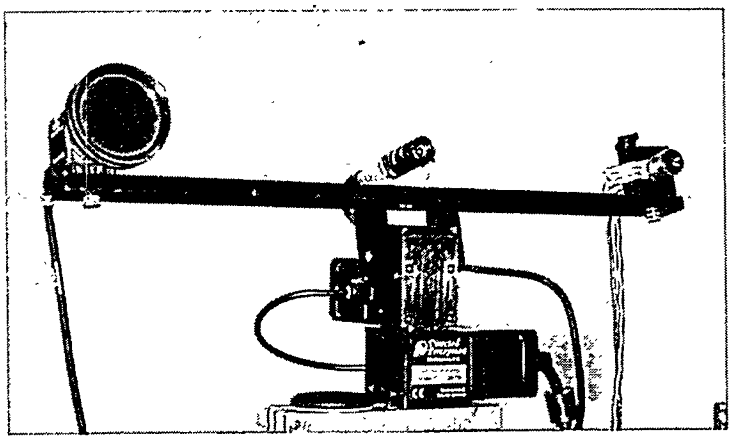

Figure 5. Sandia 3D range sensor setup

3D range data can also be determined using conventional manual correspondence techniques on images. This is possible where an object in the scene shows up in image views from different positions. Because the position information is recorded, triangulation can be used to locate the target object. This allows a user to go back and find geometric relationships between objects that may not have been thought important during the initial evidence gathering process, but were captured in the video survey.

5.2. Our Registration

The registration solution we are currently using involves two steps. The first is to setup a platform where all data is relative to a coordinate space on the platform. The second step is finding a transform between the platform and a common world coordinate space. We currently have mounted the camera with the laser mapper on a tripod stand. This provides partial registration. Also mounted to the stand is a portable coordinate measurement system (Faro) that consists of a 6 DOF arm. We mounted a single point laser range sensor (Disto) on the end of the arm. Figure 6 shows the Faro and Disto systems. The Disto and Faro units are calibrated, so for each press of the button, a range point is taken that is in the coordinates of the 6 DOF Faro arm. By viewing a number of laser tagged points from the Disto with the camera systems, we calculate the transform to put the camera and 3D range data in the coordinate frame of the Faro. This calibration process is only done once when the sensor system is first assembled, and corresponds to a factory calibration. When the unit is placed in a room, a common coordinate frame is identified by taking several points from each of three selected walls. This defines 3 planes, and their intersection point. The sensor and camera are then registered to this position in the room, which becomes a world coordinate system. 3D range data and camera images are then taken. The stand is then moved to a new position and the room registration process is repeated. Therefore all data is pre-registered to the same coordinate space.

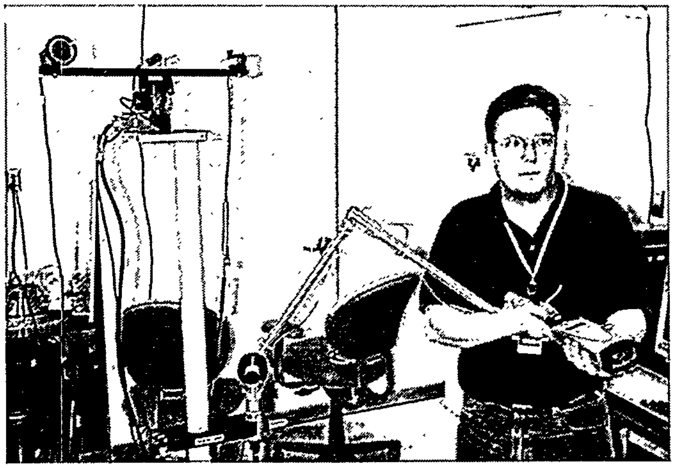

Figure 6. Sensor system with Faro arm and Disto mounted.

5.3. Software

The software is using a graphics toolkit for development that uses Performer on the SGI and Optimizer on the PC. The code is taken from simulation development software, which will allow us to add and delete modules at run time, meaning we can change filters, triangulation schemes, texture mapping schemes, etc. without restarting. This software will also enable us to add animated objects into the scene, such as human 'avitars', driven by independent programs.

The main software is written in C++. Multi-process interaction code is written in Tcl. We currently develop software on two platforms, an SGI running UNIX, and a PC running NT4.

The system architecture is shown in figure 7. The data will be collected from various sensors through a common data format. The scenario generator will first serve as a pass-through to the visualizer to check for valid and adequate data. Then it will be used extensively to put the model together. The visualizer module will serve as the display for building, analysis, and later presentation.

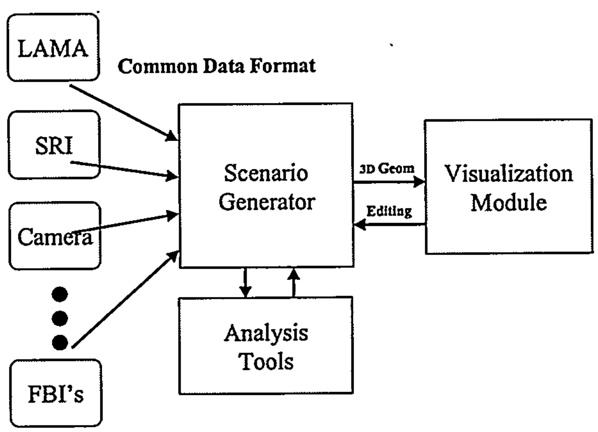

Figure 7. Software system architecture

5.4. Capabilities

The Sandia system can take 3D range and camera image data. The data is previewed for validity and stored. As additional scans become available, they are combined to check for adequate coverage. The data is converted into a common data format.

The presentation / analysis software is currently capable of data display, texture mapping of user selected camera images, and a few filters. There is a spatial filter for reducing dot densities and culling single dots (usually from sensor errors). There is also a decimation filter that removes data in flat areas, while maintaining data in high frequency areas such as along edges. There is also a smoothing filter to reduce the spatial texturing caused by sensor noise.

Lastly, we present several images showing the results of data collection. Figure 8 shows a photo and a texture mapped scene from the Sandia system. Figure 9 is a series of images showing various states of the model. These are taken from a single scan of the 3D range sensor. The images on the left show the full resolution, where the images on the right show a decimated, or data reduced version. The first row is a dot cloud; the second is a surfaced view, showing triangulation edges. The last row shows images texture mapped onto the surfaces. Figure 10 shows an overhead view of the scene with data from multiple views.

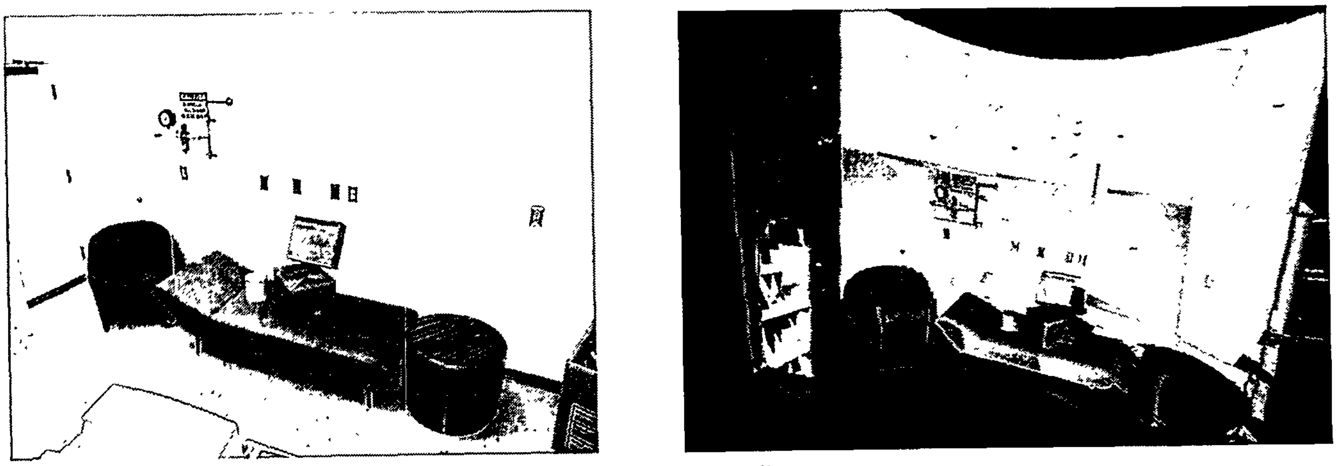

Figure 8. a) Photo of scene and b) view of 3D scene reconstruction from Sandia system

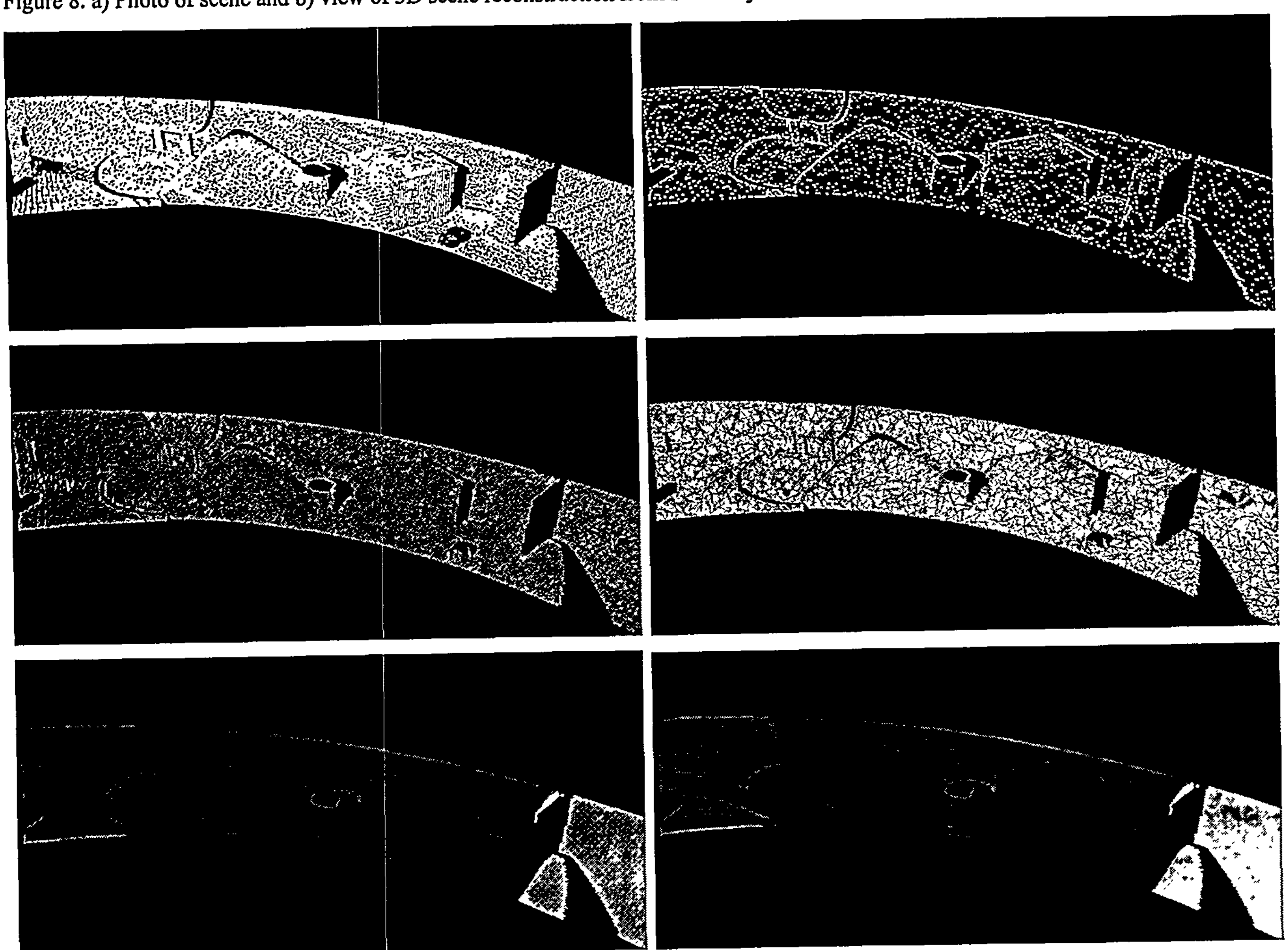

Figure 9. Data from a single scan of a 3D range sensor. Left side is full data set; right side is reduced version. Top is dot cloud, middle is surfaced data, and bottom is image texture mapped surface.

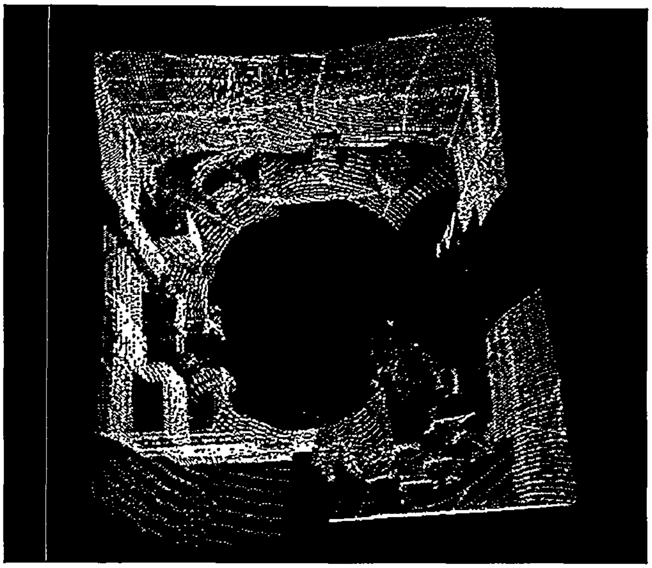

Figure 10. Overhead view of 3D scene reconstruction from Sandia system

6. FUTURE DIRECTION

We have a number of techniques already developed for model building and editing that have not been ported to this system. Including them will be the first future steps. These include surface stitching, surface fitting and replacement with simple quadric surfaces and planes. These steps will reduce redundancies in the model from multiply overlapped scan data, and have a tremendous capability to reduce the overall data size. Once the model building is forthright, we would like to include automated segmentation. This will allow separating the model into smaller pieces for ease in working with the model. At the same time we want to enhance the world model to capture other types of data, such as blood samples, fingerprints, and other chemical samplings. These, along with written or verbal notes, could be readily available in a fuller world model of the crime scene, allowing a greater virtual immersion into the scene. Temporal information would be next, to allow time based animated scenarios to unfold, including physics based modeling for gravity, inertia, and kinematics for thrown objects or spills.

7. SUMMARY

This purpose of this project is to develop a fieldable system for crime scene reconstruction. It takes input from camera and 3D range sensors, and converts this data into a 3D graphical world model. The resulting model will be a texture mapped replica, allowing the model to be used off-line as a record of the crime scene for discussion, analysis, and presentation.

REFERENCES

[1] Barry, R.E., Little, C.Q., and Burks, B.L., "Requirements and Design Concept for a Facility Mapping System," *Proc. ANS 6th Topical Meeting on Robotics and Remote Systems*, Monterey CA, pp. 775-783, February 5-10, 1995.

[2] Little, C.Q., and Wilson, C.W., "Rapid World Modeling for Robotics," *Proc. World Automation Conference '96*, May 1996.

[3] Besl, P.J., and McKay, N.D., "A Method for the Registration of 3-D Shapes," *IEEE Transactions of Pattern Analysis and Machine Intelligence*, pp. 239-56, 1992

[4] Luck, J. P., Little, C.Q., and Hoff, W., "Registration of Range Data Using a Hybrid Simulated Annealing and Iterative Closest Point Algorithm," submitted to *IEEE International Conference for Robotics*, April 2000.

[5] Luck, J. P., "Registration of Range Images Through a Hybrid Simulated Annealing and Iterative Closest Point Algorithm," *Master's Thesis*, Engineering Division, Colorado School of Mines, May 1999.