Three-dimensional visualization environment for multisensor data analysis, interpretation, and model-based object recognition

Michael E. Goss
goss@cs.colostate.edu

J. Ross Beveridge
ross@cs.colostate.edu

Mark R. Stevens
stevensm@cs.colostate.edu

Aaron D. Fuegi
fuegi@cs.colostate.edu

Computer Science Department
Colorado State University
Fort Collins, CO 80523

ABSTRACT

Model-based object recognition must solve three-dimensional geometric problems involving the registration of multiple sensors and the spatial relationship of a three-dimensional model to the sensors. Observation and verification of the registration and recognition processes requires display of these geometric relationships.

We have developed a prototype software system which allows a user to interact with the sensor data and model matching system in a three-dimensional environment. This visualization environment combines range imagery, color imagery, thermal (infrared) imagery, and CAD models of objects to be recognized. We are currently using imagery of vehicles travelling off-road (a challenging environment for the object recognizer). Range imagery is used to create a partial three-dimensional representation of a scene. Optical imagery is mapped onto this partial 3D representation. Visualization allows monitoring of the recognizer as it solves for the type and position of the object. The object is rendered from its associated CAD model. In addition to its usefulness in development of the object recognizer, we foresee eventual use of this technology in a fielded system for operator verification of automatic target recognition results.

Keywords: computer graphics, visualization, range data, ladar, multisensor fusion, automatic target recognition

1 INTRODUCTION

Model-based object recognition techniques use sensor data to find the three-dimensional position of an object in a scene which matches one of a set of object models known to the system. Monitoring progress and verifying correctness while the recognizer is finding objects requires a visualization tool capable of displaying the geometric plausibility of the resulting object pose and sensor registration. To support monitoring and verification, our visualization system allows a user to interact with sensor data and object models in a three-dimensional environment.

Our visualization environment, *Rangeview*, combines range imagery, color imagery, thermal (infrared) imagery, and CAD models of objects to be recognized. We are currently using imagery of vehicles in natural terrain. Range imagery is used to create a partial three-dimensional representation of a scene. Optical imagery is mapped onto this partial three-dimensional representation. Output from the vision system is registered with the scene.

The need for a verification tool for the multisensor object recognition algorithm was the primary motivating factor for development of Rangeview. Verification of features and objects recognized from multiple sensors is a non-trivial task.[1] Two important three-dimensional geometric relationships are computed by the recognizer:

1. Registration of the sensors (which are not bore-sight aligned), and
2. Alignment of the model to the sensor data.

Previous systems[2,3] locate a 3D target in a ladar image, and then render an image of the model in a 2D scene with the data. This 2D image does not allow a complete understanding of how well the model has been located. With our verification system the 3D model and sensor data can interactively be examined to determine how successful the match actually was. We have found the ability to arbitrarily change viewing parameters invaluable in the development of our coregistration and object recognition algorithms.

2 SENSORS

The Rangeview system integrates two basic sensor data types, range and optical. Rangeview provides the capability for the user to manually register data from multiple sensors, as well as displaying registration computed by the object recognizer.

Range images are generated by a ladar (laser-ranging) system. The ladar scans a scene in a series of parallel vertical strips, generating a rectangular array of range values with 12 bit resolution. The field of view of the current ladar system is approximately 15° horizontally and 3° vertically.

Two types of optical images are currently used by the object recognizer and Rangeview:

1. Color CCD: 24-bit color images captured by a color CCD digital camera. For testing purposes, images digitized from color film are also used.
2. FLIR: Forward-Looking Infrared thermal data.

The resolution and field of view of the optical imagery are typically both much greater than that of the range imagery.

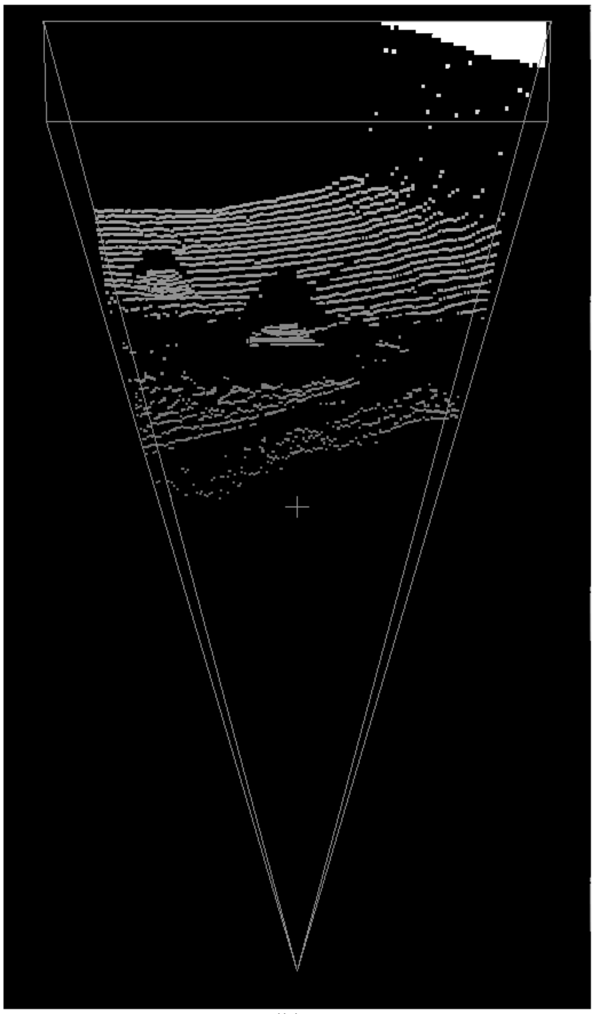
The sensors are configured such that, although not bore-sight aligned, the separation between the sensors is small compared to the distance to objects in the scene. As part of the object recognition process, the recognizer performs an eight degree of freedom coregistration process to register the ladar to each optical sensor.[4] Six degrees of freedom relate the object to the optical camera, and two degrees of freedom relate the range sensor to the optical sensor.

3 RANGE IMAGES IN THREE DIMENSIONS

Range images are conventionally displayed as gray scale (or pseudo-color) images, as shown in Figure 1(a), or as a 2D overhead "scatter plot" or "count image."[3] For a gray scale or pseudo-color representation, the gray level or color of each pixel corresponds to the distance of the range sample from the sensor. Some work has also been done in the display of range images as 3D surfaces with illumination and shading, but this requires the prior extraction of surfaces from the raw range data.[5,6]

We have developed a new method for display of raw range data which provides a three-dimensional view of the data, and allows the viewer to interactively roam through the data (see color Plate 1). The range data is displayed as a partial 3D

(a)



(b)

Figure 1: (a) Conventional view of range data as gray scale, (b) Three-dimensional view of range data

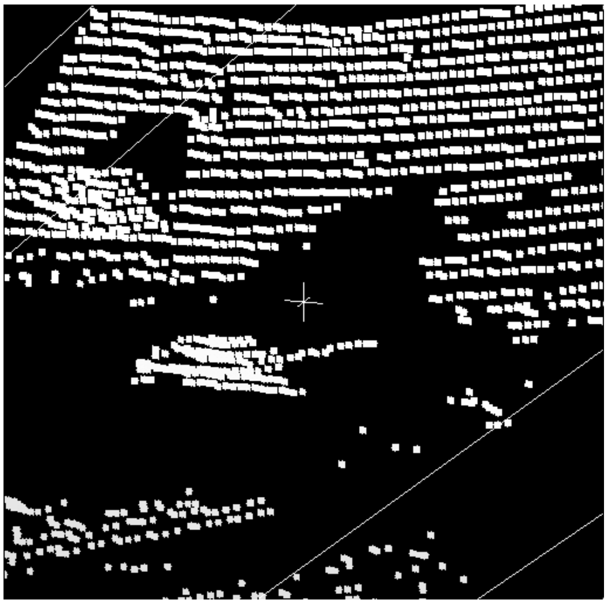


Figure 2: Oblique view of three-dimensional range data

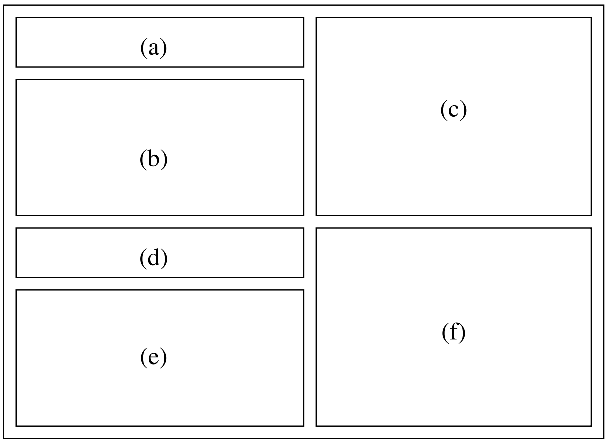


Figure 3: Key to color Plate 2: (a) FLIR image in pseudo-color, (b) FLIR image merged with range data, (c) Close-up of FLIR and range data with model, (d) Full resolution color image, (e) Color image merged with range data, (f) Pseudo-color range data with model

model of the scene in a coordinate system relative to the sensor. Each range data sample is displayed as a rectangle at the measured depth, occupying the angular space covered by the ladar pixel. Figure 1(b) shows a static view of this interactive system. The four-sided pyramid represents the view volume of the ladar sensor. In this image, the gray levels assigned to the pixels are the same as in (a). Figure 2 shows the same data from a different viewpoint. A pseudo-color option based on range is available to the user, as well as several color schemes described below which map optical sensor data onto the range data (see color Plate 2 and Figure 3).

Display of the range data in three-dimensional space gives the user a more complete representation of the spatial configuration of the data. Interactive control of viewing parameters allows the user to examine the range data from many different viewpoints at any magnification level. The user may select a point of interest in the data using the cursor in the viewer window, or by specifying a 3D location directly using sliders in the control panel. The control panel also provides sliders for the user to change the view azimuth, elevation, and distance relative to the point of interest. User controlled de-cluttering of the image is accomplished by specifying the range of depth values displayed relative to the viewer, the sensor, or both, using sliders on the control panel. Gray scale or pseudo-color values are rescaled to represent the selected depth range. Several built-in color palettes are provided, as well as the ability to load a color palette from a user-specified file. A workstation with hardware graphics acceleration is used to perform display update at interactive rates. Current implementations are for Hewlett-Packard 400/TurboVRX and 700/CRX24Z systems using the Starbase graphics library, and Sun Sparc10/24ZX systems using the PEX graphics library.

4 MERGING OPTICAL AND RANGE IMAGERY

The three-dimensional display of range data provides an intuitive method for the user to view the ladar output. Gray scale (or color) encoding of range is no longer needed, since range can be observed directly in the partial 3D scene representation. This allows representation of other sensor information using gray scale or color, allowing other sensors to be viewed simultaneously with the range data.

Figure 4(a) shows a color CCD image (shown here in gray scale) collected at the same time as the ladar range image shown in Figure 1. The color image was resampled to the resolution of the range data and then registered to the range data using the Rangeview interactive registration tool. This tool allows a user to mark corresponding points in a range image and an optical image. Rangeview uses these registration points to compute a 2D affine transformation which registers the two images.

The user can request that the color image be overlaid on the range data; the resulting image is shown in Figure 4(c). To maintain interactive display update speeds on the available hardware, the CCD image is normally resampled to the lower resolution of the range data (as shown in Figure 4), but it may also be displayed at full resolution (see color Plate 2(d) and (e), and Figure 3). The lower resolution capability is also available for thermal (FLIR) images, shown in Figures 4(b) and (d), and color Plate 2(a) through (c). Thermal images can be displayed in gray scale or pseudo-color.

Another experimental viewing mode allows the user to simultaneously view all sensors in a single three-dimensional image. Colors for the range data values are computed using the hue and intensity of the color CCD image; the saturation for each range value is computed from a function of the saturation of the source color image and the intensity of the thermal image. This results in an image which has vivid, highly saturated colors in areas of interest due to high thermal intensity (such as the tank) and has dull, unsaturated (gray) colors in areas of low thermal intensity.

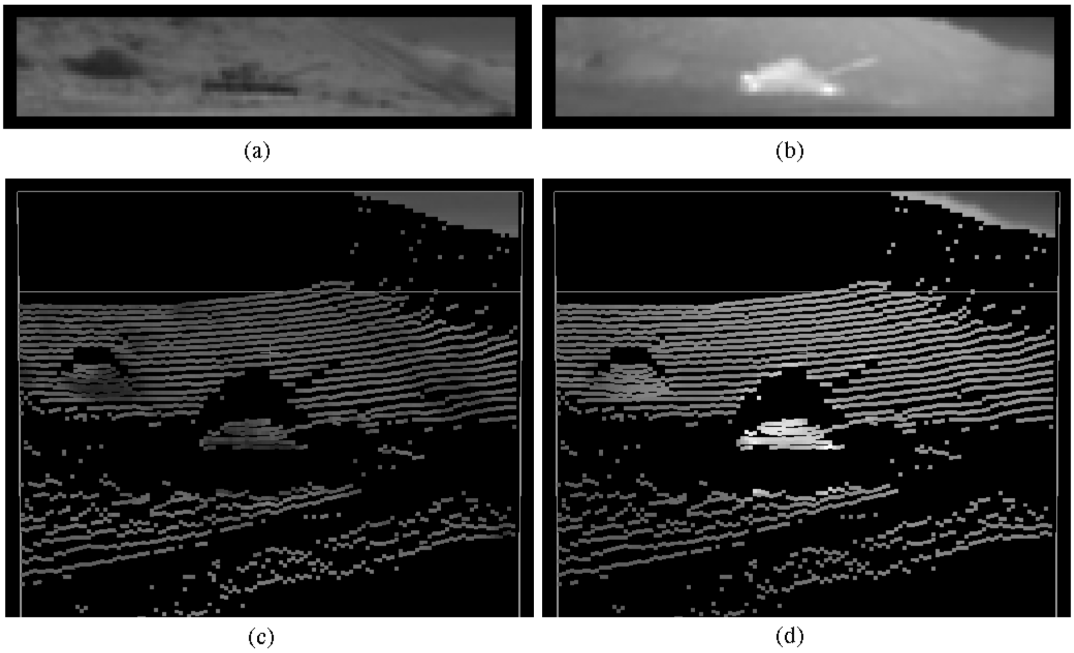


Figure 4: (a) CCD image, (b) FLIR image, (c) CCD image superimposed on three-dimensional range data, (d) FLIR image superimposed on three-dimensional range data

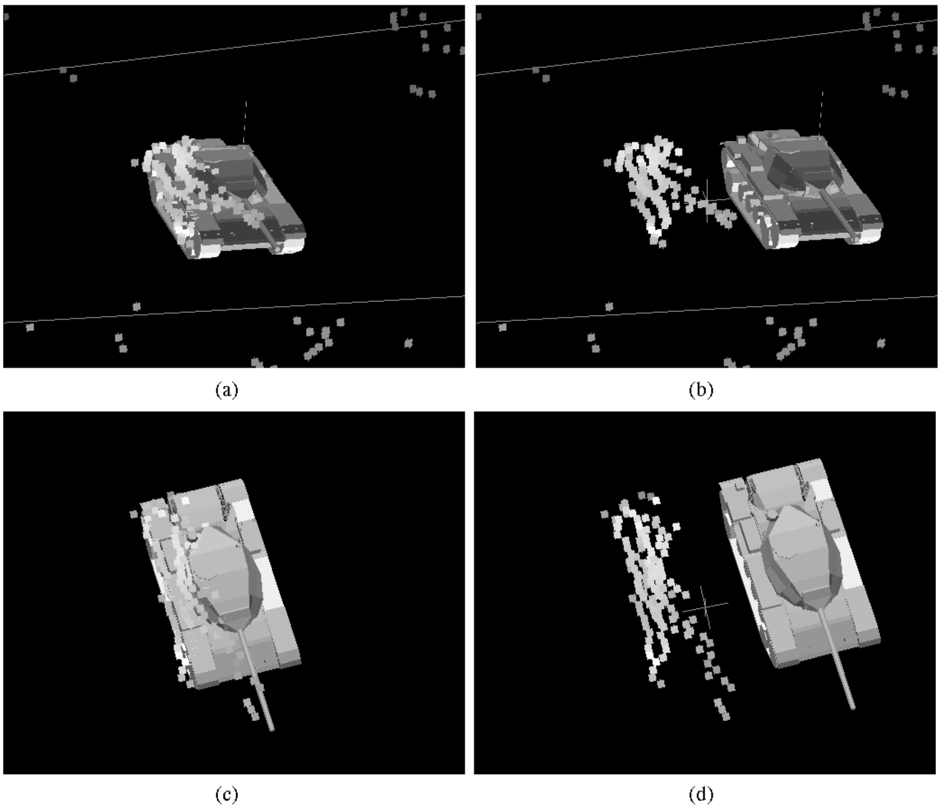


Figure 5: Model of vehicle in relation to three-dimensional range data

5 SENSOR REGISTRATION

Registration of optical sensor data with the range data can be performed interactively by the user, or automatically by the object recognizer. For interactive registration, the user is presented with a 2D viewer for the optical image to be registered, in addition to the 3D range image viewer. Color Plate 1 shows the FLIR viewer at top center, and the color image viewer at lower right. The quadrilateral superimposed on each shows the area of the optical image registered to the range data. The ladar viewer (top right) can display either the color or infrared image overlaid on the range data using the current registration transformation (the color CCD image is shown). All the viewers allow the user to magnify the images for accurate selection of points. The various viewer and control panel windows can be opened, closed, and repositioned at the discretion of the user. To perform registration, the user picks corresponding points in the range image and the corresponding optical image. Rangeview computes the registration transformation and immediately updates the merged image display. The user can interactively edit the selected points to modify an unsatisfactory registration.

When registration is performed automatically by the coregistration portion of the object recognizer, progress in the multi-step coregistration process is displayed in the same manner as for interactive registration. The user can monitor the progress at each step. All viewer controls are active while the recognizer is running, allowing the user to change viewing or other display parameters. Rangeview can also display the raw transformation values computed by the recognizer, a useful capability for development of the coregistration algorithm.

6 MODEL DISPLAY

In addition to display of sensor data, Rangeview can also display 3D polygonal models within the 3D viewer. The vehicle models used (such as the one in Figure 5 and Plate 2) have been derived from detailed BRL-CAD CSG (Constructive Solid Geometry) models.[7] The models were manually reduced in detail and then converted to polygons. The position of the model relative to the range sensor is specified by a 3D affine transform. This transform is computed iteratively by the object recognizer while matching the model to the sensor data. Rangeview also allows a user to specify model position interactively, using a popup control panel.

7 OBJECT RECOGNITION

Rangeview has been developed in parallel with the object recognition system,[4] which uses the range, thermal, and color images. To date, the coregistration portion of the recognizer has already been implemented. The recognition system now being implemented will automatically hypothesize matches between the scene and object models, and iteratively refine the three-dimensional geometric relationship between the model and the multisensor imagery. As the recognizer develops spatial relationships among the sensors and the model, it updates the three-dimensional display. The hypothesized position is shown by the display of a wireframe or solid object in the same three-dimensional space occupied by the range data. It is possible to view an animated trajectory of the object in real time (or user-controlled playback) as the coregistration module refines the model position. Figure 5 shows a manually-positioned model displayed in relation to range data. The range data can be displayed in any of the modes described above. In Figure 5(a) and (c), the model is closely registered to the range values; in (b) and (d) the model is moved slightly away for illustration purposes. Color Plate 2(c) (top right) shows the model positioned relative to range data, with infrared imagery superimposed using a pseudo-color palette. Plate 2(f) (bottom right) shows range data using pseudo-color determined by the range values, with the range compressed to include only the area immediately around the model.

8 SUMMARY

When development began on this system, the goal was to develop a useful tool for working with an object recognition system in a three-dimensional environment. The resulting tool, the Rangeview program, while satisfying our original goals, is also developing into a general purpose tool for working with multisensor data. By using the three-dimensional environment of Rangeview as a base, we have been able to integrate other useful tools (such as sensor registration) into a single program.

Early in the project, we realized the need to interactively register data from multiple sensors, and also the need to interactively position models for comparison with recognizer results. Since we had already developed a three-dimensional interactive environment, it made sense to integrate the registration and positioning tasks into this environment. This made it possible to display the source images and models within Rangeview, select registration points, and immediately view the results.

Rangeview is currently being used with ladar, FLIR and color CCD data collected at Fort Carson in November of 1993.[8] This data includes multiple frame ladar sequences taken at evenly spaced time intervals. Rangeview provides the capability to display these frames as an animated loop. The animation capability will be extended in the future to multisensor time sequences if data becomes available.

The Rangeview three-dimensional visualization environment is proving to be extremely useful in the development of multisensor object recognition and Automatic Target Recognition (ATR) algorithms. We foresee eventual use of this technology in fielded systems for operator verification of ATR results. Display of a target model registered to a partial three-dimensional scene could allow an operator to verify a target with greater confidence than is possible using current types of displays. In time-critical applications, multiple simultaneous views could be scanned rapidly by the operator.

9 ACKNOWLEDGEMENTS

We thank Hewlett-Packard Co. for the generous donation of graphics workstation equipment, the availability of which was invaluable to this project. Participation of J. Ross Beveridge was supported in part by Advanced Research Projects Agency (ARPA) grant DAAH04-93-G-422, administered by the U. S. Army Research Office.

10 REFERENCES

1. J. K. Aggarwal and N. Nandhakumar. Multisensor fusion for automatic scene interpretation. In Ramesh C. Jain and Anil K. Jain, editors, *Analysis and Interpretation of Range Images*. Springer-Verlag, 1990.
2. Steven K. Rodgers, Carl W. Tong, Matthew Kabrisky, and James P. Mills. Multisensor fusion of ladar and passive infared imagery for target segmentation. *Optical Engineering*, 28(8):881–886, August 1989.
3. Jacques G. Verly, Dan E. Dudgeon, and Richard T. Lacoss. Model-based automatic target recognition system for the UGV/RSTA ladar. In *Proceedings: Image Understanding Workshop*, pages 559–583. ARPA, November 1994.
4. Anthony N. A. Schwickerath and J. Ross Beveridge. Object to multisensor coregistration with eight degrees of freedom. In *Proceedings: Image Understanding Workshop*, pages 481–490. ARPA, November 1994.
5. R. Nevatia, K. Price, and G. Medioni. USC image understanding research: 1993-1994. In *Proceedings: Image Understanding Workshop*, pages 69–80. ARPA, November 1994.
6. Hillel Rom and Gérard Medioni. Part decomposition and description of 3D shapes. In *Proceedings: Image Understanding Workshop*, pages 1505–1522. ARPA, November 1994.

7. U. S. Army Ballistic Research Laboratory. *BRL-CAD User's Manual*, release 4.0 edition, December 1991.

8. J. Ross Beveridge, Durga P. Panda, and Theodore Yachik. November 1993 Fort Carson RSTA data collection final report. Technical Report CS-94-118, Computer Science Dept., Colorado State University, January 1994.