Mosaic Generation for Under Vehicle Inspection

Paul Dickson, James Li, Zhigang Zhu, Allen R. Hanson, Edward M. Riseman, Howard Sabrin[1],
Howard Schultz, Gary Whitten
Computer Vision Lab, Department of Computer Science
University of Massachusetts at Amherst, MA 01003, U.S.A.
[1]Atlantic Coast Technologies, Inc
Silver Spring, MD 20904
{pauld, jamesli, zhu, hanson, riseman}@cs.umass.edu, sabrinh@acoast.com, {hschultz,
whitteng}@cs.umass.edu

Abstract

As the threat of terrorism becomes more prevalent, the need for technological innovations for tighter security rises. Some of the areas where security needs to be increased are at border customs stations, airports, federal buildings, embassies, sporting events, and the like where checkpoints are being set up to inspect vehicles coming onto the premises. The proposed vehicle inspection system uses an array of cameras and mosaicing techniques to generate different perspective views that provide a 3D representation of a vehicle's undercarriage. These mosaics allow a complete inspection of the undercarriage using a graphical user interface without requiring the use of inspection mirrors or the vehicle to drive over a manned inspection bay. The system is designed for simplicity so that operators will not require long training periods.

1. Introduction

1.1. Motivation

Inspection stations are necessary at borders and secure locations to check vehicles for contraband. Among the places that must be examined are vehicle undercarriages. Two techniques are commonly used for this inspection. In the first, an inspector visually scans the undercarriage by running a convex mirror under the vehicle. This technique, while effective, does not give the inspector a complete view of the undercarriage and tends to leave areas unviewed because of occlusion caused by wheels, the viewing angle, and physical constraints on mirror position. In the second technique, the vehicle is driven onto an inspection bay built into the ground. An inspector visually scans the undercarriage. This technique allows a complete scan of the vehicle undercarriage but requires the construction of an inspection bay and exposes the inspector to possible danger.

1.2. Overview

The goal of this project was to create an under vehicle inspection system (UVIS) that would allow a complete undercarriage inspection of a vehicle without requiring significant construction while reducing danger to the inspector. The system is designed to complete all computation necessary for the inspection within 30 seconds and to display results in a form that enables the inspector to understand the data without extensive training. The system would also be robust enough to handle driver inconsistencies in both speed and steering. As this system is meant to be used outdoors, it will have to work under all weather conditions.

The basic design and operation of UVIS is shown in Figure 1. A linear array of cameras is placed in a low ramp housing that a vehicle drives over. As the vehicle passes, the cameras take continuous simultaneous

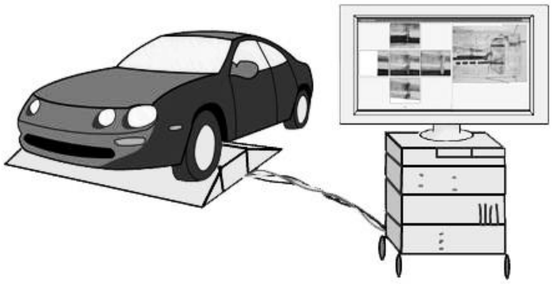

Figure 1. Basic UVIS design.

overlapping images of the vehicle undercarriage. The overlap of these images is then matched, and this information is used to generate mosaic images of the vehicle undercarriage from five different angular perspectives. Each contains a complete view of the undercarriage (see 1.4 for the theory behind their generation, 2.1 for the specifics of their generation) and the ensemble contains enough information to view the undercarriage in three dimensions from multiple viewpoints.. The images are then displayed in various ways through a flexible graphical user interface (GUI; described in 2.2).

1.3. Commercially available inspection systems

Almost all currently available inspection systems are based on an inspector manipulating a convex mirror under a vehicle [1,2]. Some of these systems have become more advanced and allow the inspector to perform this same search with a camera [3].

The one other type of system on the market has cameras that a vehicle drives over, much like our UVIS pictured in Figure 1 [4]. These systems use video cameras under the vehicle to send images to a monitor on the inspector's desk and sometimes to a VCR that records a video of the vehicle. These systems are essentially a more complete version of the handheld mirror.

What makes the system described in this paper unique is that it is the first to give a complete undistorted static view of the undercarriage being inspected. It also is the first system that allows the inspector to look around occluding portions of the undercarriage by using different perspective (stereo) views. Unlike the current camera-based systems, UVIS allows the inspector to focus on specific regions of the undercarriage (much as the inspector could with a mirror) while maintaining the complete view. The various images can be saved for later review.

1.4. Prior work

The mosaic generation techniques in UVIS were first developed for environmental monitoring, where the mosaics provide a compact representation of large amounts of aerial video data of forest canopies[5]. The techniques were further refined and generalized for use in stereo mosaic construction from a predominantly translating camera [6]. Flying a camera over a forest canopy is analogous to the situation of a vehicle moving over stationary cameras in UVIS. Mosaic generation techniques from our previous work [5,6] are especially useful as they are robust with respect to erratic shifts in speed and horizontal motion often encountered by small airplanes because of turbulence and shifting wind conditions. In UVIS this robustness allows the system to compensate for a driver who cannot manage a steady speed and direction. The primary focus of these methods was to build stereo mosaics from the data to generate 3D information from a given scene. These stereo techniques were used in the UVIS environment to generate the perspective views that handle issues of occlusion (see 2.1 for more detail).

Stereo panorama techniques have been developed for rotational cameras [7,8,9]. Huang and Hung [7] proposed methods for stereo generation from rotating pairs of cameras, and both Peleg and Ben-Ezra [8] and Shum and Szeliski [9] improved techniques for generating stereo mosaics from a single off-center rotating camera. As we [5,6] observed, the techniques based solely on rotation are not well suited to an environment where the motion of the camera is primarily translational. However, these rotational techniques shed light on the usefulness of epipolar geometry in the rendering of images. One important issue in creating stereo mosaics is motion parallax. Typical 2D mosaicing techniques, such as manifold projection by Peleg and Herman [10], cannot create seamless images in the UVIS domain because of the large depth differences in the undercarriage structure relative to the distance between the undercarriage and the cameras. This has led to the inclusion of local alignment techniques proposed by Shum and Szeliski [11]. Zhu et al. [6] proposed the use of an efficient ray interpolation technique based on local alignments. These local alignment techniques allow the generation of seamless mosaics, which are a more cohesive image of the vehicle undercarriage. Geo-registration techniques [12] that have often been used to aid the generation of 2D mosaics cannot be used in our case as we do not have a 3D model of the undercarriage of the vehicle.

1.5. Design considerations

The work that follows is primarily a tailoring of existing mosaic generation techniques to the specific problem domain of UVIS. Most of the techniques discussed show how domain constraints can be used to simplify the problem.

The amount of camera separation was one of the first considerations for the UVIS design. This issue was complicated by the limited distance between the cameras and the vehicle undercarriage. Through experimentation we discovered that a separation of 3 inches resulted in acceptable mosaic results. Applying the 3 inch camera separation to an equivalent separation in the direction of motion results in a vehicle velocity of approximately 5 mph, which is a realistic speed for a vehicle to traverse the camera platform. The mosaicing system is robust enough to generate acceptable results even at higher speeds. This ability to function with a greater separation is

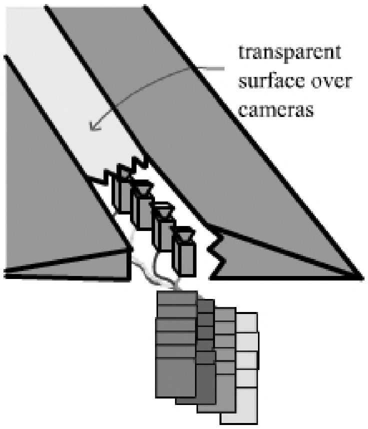

Figure 2. Camera platform of the UVIS system.

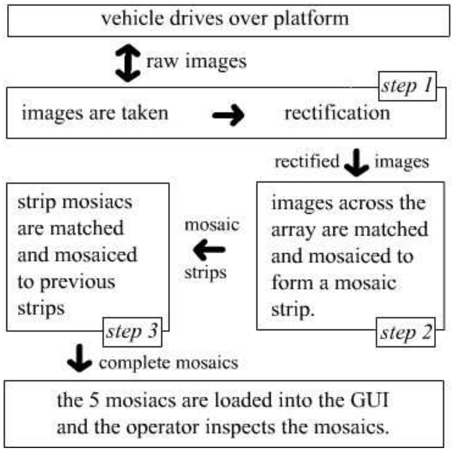

Figure 3. State diagram of UVIS operation.

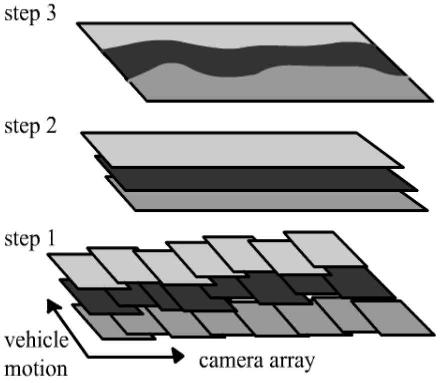

Figure 4. Mosaic building steps.

useful for handling data sets for drivers who cannot maintain 5 mph.

In order to reduce the effect of perspective distortion on the matching process, the distance between the cameras and the undercarriage was increased to three feet. This was accomplished by folding the optical path between the cameras and undercarriage so that the cameras could be placed horizontally under the inspection platform.

Another design consideration was the cameras' field of view. Initially, cameras with a 90° field of view were used to enable greater camera separation. We discovered that the occlusions caused by the combination of extreme perspective and relative closeness of cameras to undercarriage made matching impossible. The next step was to simulate lenses with a 45° field of view by keeping only the middle 50% in each axis of the 90° field of view image. We discovered that matching was possible with this simulated 45° field of view and plan to use 45° lenses in the prototype.

A different issue considered was lighting for the system. We discovered through experimentation that the mosaicing process would not function without fairly even lighting over the area of the images. Adding evenly spaced white LEDs around the cameras provided enough lighting for robust mosaicing.

One final issue was the use of epipolar geometry for increased matching speed. The epipolar geometry was not used in the current system because greater robustness was needed and the exact camera positions are not known. We assume that the epipolar geometry will function where cavities are present, but this will need to be proven through experimentation.

2. System description

The overall design of UVIS can be seen in Figure 1. The platform that the vehicle drives over appears in Figure 2. The cameras have a separation of 3 inches and are protected by a glass or Lexan window (material still to be determined). They capture synchronized images continuously as a vehicle drives over. From this data, five complete mosaics of the undercarriage are created. The complete system overview appears in Figure 3, which makes reference to Figure 4, which gives the order in which the mosaics are stitched together.

Each mosaic is generated from a different perspective corresponding to the portions of each image extracted to create it. Figure 5 shows how strips taken from different parts of the images can be combined into separate mosaics. Figure 5 is equally applicable to camera orientation across the vehicle or along the length of the vehicle. This give rise to six mosaics, the central ones in each case being the same. Thus, five unique mosaics are generated; left, right, center, forward, and backward (see 2.1 for specifics of mosaic generation). These mosaics are then displayed using a GUI that allows the inspector to

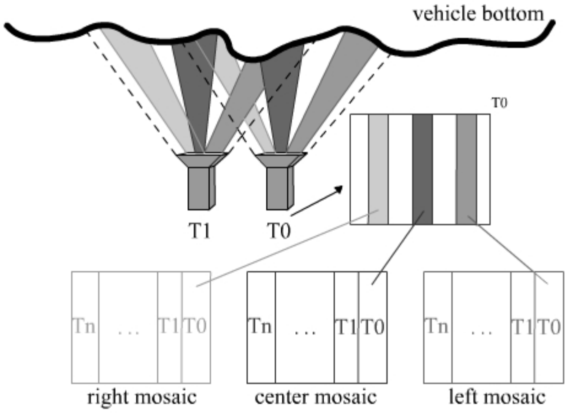

Figure 5. Example of how slices are taken from the original images to create perspective mosaics.

view all five perspective views of a given area of the undercarriage simultaneously.

2.1. Mosaic generation

The first step in the mosaicing process is to remove radial distortion from the camera images. This step is necessary as significant radial distortion warps images enough around the edges so that no match is possible between images. Examples of this can be found at the project website [13]. Unwarping involves two stages, the first of which, done at system setup, is to calibrate the cameras. During operation, every time an image is taken by one of the cameras, its calibration data are used to unwarp the image. Up to this point the polynomial unwarping package built into the commercial vision software ENVI has been used [14]. New unwarping techniques that should improve our mosaicing results are currently being implemented but have not been completed.

The next step in the mosaicing process is generation of match data between images to determine the rotation and translations necessary for stitching the images together. The matching process is currently performed using a correlation-based pyramid matcher. A least mean squares method is used to perform the motion parameter estimation between images. A new matching technique specifically designed for the constraints of UVIS is discussed in section 4.

We will show the geometry of mosaic generation from the point of view of a single camera image sequence as the principle is similar for cross camera mosaicing. Since the motion parameters between image pairs are necessary for mosaicing, the match data are used to generate a rotation matrix $\boldsymbol{R}$ and a translation matrix $\boldsymbol{T} = (T_x, T_y, T_z)^t$ for the kth frame for which the y-axis (assumed direction of motion) is the dominant direction. We can define the relationship among a world coordinate of $\boldsymbol{X}_k=(X_k, Y_k, Z_k)^T$ and reference coordinate of $\boldsymbol{X}=(X,Y,Z)^T$ as

$$\mathbf{X} = \mathbf{R}\mathbf{X}_k + \mathbf{T} \tag{1}$$

As the distance between the vehicle undercarriage and the cameras changes little as the vehicle passes over, we can simplify Equation (1).

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{Q}\begin{pmatrix} X_k \\ Y_k \\ Z_k \end{pmatrix} + \begin{pmatrix} T_x \\ T_y \\ 0 \end{pmatrix}, \quad \mathbf{Q} = \mathbf{R} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & T_Z / H \end{pmatrix} \tag{2}$$

where H is the average height of the undercarriage over the cameras. From this equation we get the projective transformation

$$\mathbf{u}_k^{\mathbf{p}} \cong \mathbf{A}\mathbf{u}_k, \quad \mathbf{A} = \mathbf{F}\mathbf{Q}\mathbf{F}^{-1}, \quad \mathbf{F} = \begin{pmatrix} F & 0 & 0 \\ 0 & F & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3}$$

where $\boldsymbol{u}_k$ and $\boldsymbol{u}^p{}_k$ are the original image point in the kth frame and its reprojected point after image rectification respectively, and F is the focal length of the camera. In general we need to estimate the 3D motion parameters $\boldsymbol{R}$ and $\boldsymbol{T}$ between each pair, and the intrinsic parameters of the camera (e.g., F) in order to generate mosaiced images. In UVIS, however, this is not necessary as the motion between two images is small, allowing the match data between successive image pairs to be used to generate the matrix $\boldsymbol{A}$ and the scaled translation parameter in Equation (5).

Let us first discuss the process of taking slices of the images and pasting them together into a single mosaic as shown in Figure 5. The mosaicing process begins with slices in the x-axis being taken from each image in sequence to contribute to the mosaic. In the cases where three perspective mosaics are being generated, three slices are taken from each image: call them left, center, and right. Example equations for selecting the portion of the image for the left perspective appear below. The center of the slice from image k is always at $(u,v)=(0,d_y/2)$, where d_y is the distance between the left and right slices. The widths of the slices from the front and rear of this center point are

$$w_l^k = F\frac{\left|T_y^k - T_y^{k-1}\right|}{2H}, w_r^k = F\frac{\left|T_y^{k+1} - T_y^k\right|}{2H} \tag{4}$$

The T_y components from equation 4 are the Y translation components of the translations between the images k-1, k, and k+1. The slice is then put into the right perspective mosaic centered at position

$$\left(t_x^k, t_y^k, \frac{d_y}{2}\right) = \left(F\frac{T_x^k}{H}, F\frac{T_y^k}{H} + \frac{d_y}{2}\right) \tag{5}$$

This process appears in greater detail in Zhu et al. [5, 6].

As the occlusions in the UVIS domain are large, a direct pasting together of these slits would result in prominent seams between all of the slices. To reduce the seams a parametric model of the motion between the images is created by local registration. This model is then used to interpolate between the features of the images [5,6].

2.2. GUI

The GUI is designed to simplify accessing data and ease the inspection process. Six small images appear on the left and one full image on the right, as in Figure 6.

The operator is able to inspect the complete image on the right and obtain enlarged perspective views of any suspicious area on the left all in real time. Of the six images on the left side of the GUI, the five in the cross display the portions of the perspective mosaics that correspond to the selected portion of the complete image

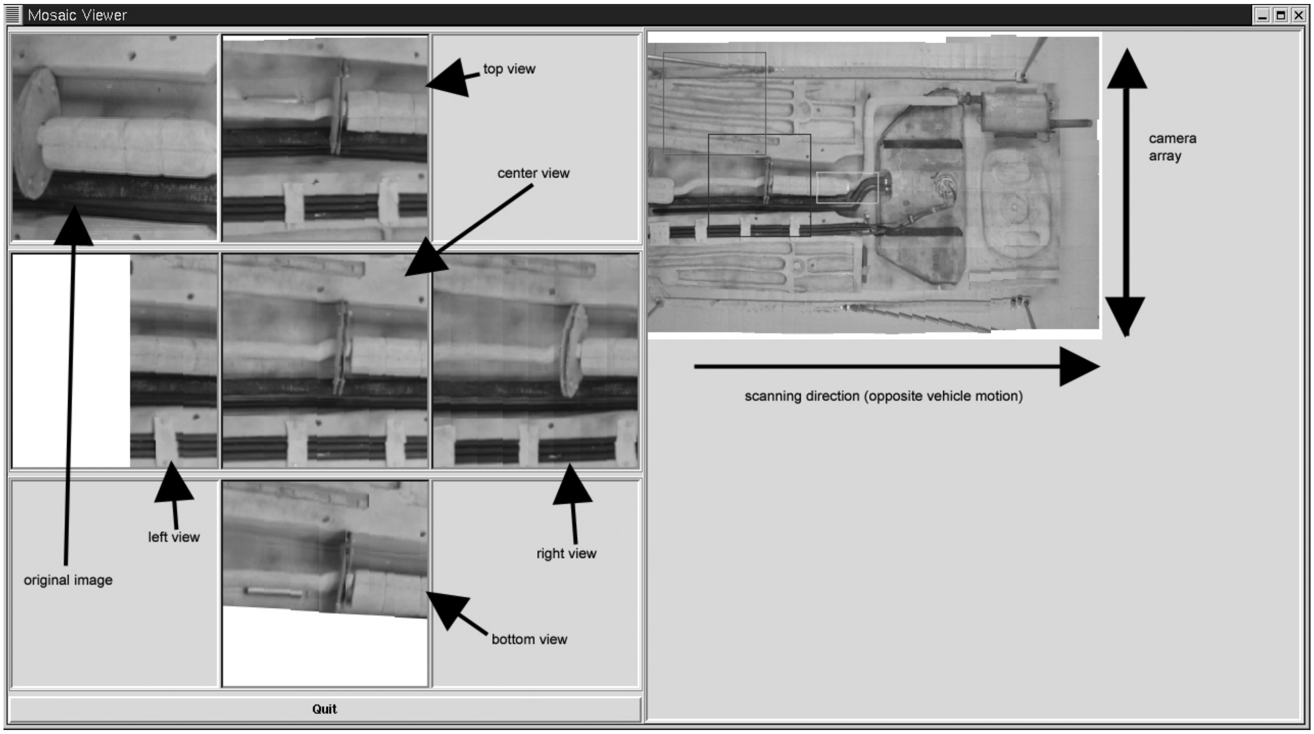

Figure 6. The GUI used to display generated data.

on the right. In effect this allows the operator a limited 3D view of the undercarriage. Items that are located behind occluding objects, such as the exhaust pipe in Figure 6, can be discovered by using the perspective views. The sixth image window on the left allows the operator to view any of the original individual frames used to create the mosaics. Viewing the individual frames allows the operator to inspect an area from a more acute angle than is present in the perspective mosaics. The GUI can be tested online or downloaded from our web page [13].

3. Experimental results

3.1. Experimental setup

As it was not feasible to build a full scale working prototype model of this system within our university laboratory, we built a scale model for performing experiments (Figure 7). The model was hand carved from Styrofoam based on images taken of a vehicle's undercarriage and painted to look like a used vehicle. To improve the realism of the model, an actual gas tank and muffler were built into the model (see Figure 7a). Four cameras are attached to a metal frame that surrounds the model. The cameras are WATEC America Corporation LCL-902C's which are ½" CCD black and white cameras.

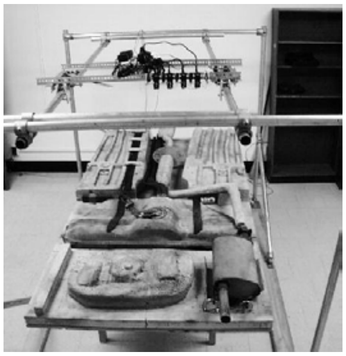

Figure 7. a. Full model and camera frame setup.

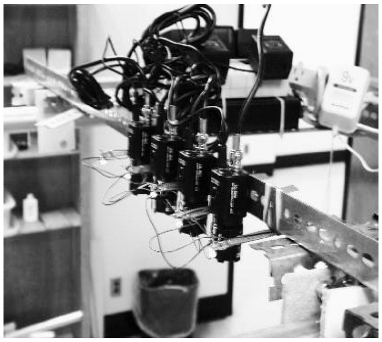

Figure 7. b. Cameras used in current model.

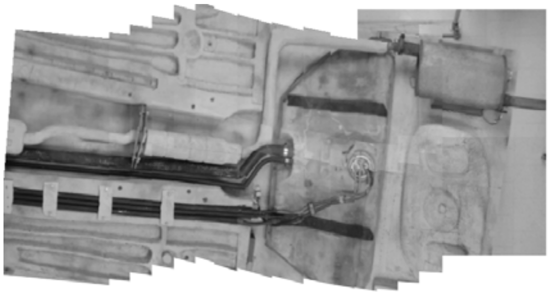

Figure 8. Sample of data from a simulated bad driver run.

They each have Computar H3616 FIC-3 lenses that are 3.6mm (90° FOV) at F1.6. The frame allows the cameras to be moved in the *x*, *y*, and *z* directions above the model. The separation between cameras can be varied. White LEDs are mounted around the cameras to even out the lighting of the undercarriage to improve mosaicing. The model was built on a 4’ by 8’ piece of plywood that is set on a table, which allows the model to be rotated while data are collected, simulating a vehicle turning while passing over the cameras.

3.2. Sample results

Complete mosaics of the modeled vehicle undercarriage were generated under a variety of driving scenarios. Part of a mosaic appears in the GUI in Figure 6 (shown in full at the project web site [13]). Sample data sets have been taken with erratic motions of the vehicle, including stopping, turning, and reversing while driving over the camera platform. Successful mosaics are generated even under these adverse conditions; a sample is shown in Figure 8. The mosaic displayed in the GUI (Figure 6) is more complete than that in Figure 8 because in Figure 6 the pure translational motion of the vehicle allowed the four cameras to be shifted to simulate 13 cameras.

4. Conclusions

A system has been designed that allows better inspection of the undercarriage of a vehicle through a flexible GUI, without fear of personal injury for the inspector or construction of an inspection bay. The system displays the vehicle undercarriage in a user-friendly manner that allows a quick visual inspection of the vehicle. The issues of processing time and all-weather operation have not yet been addressed. Further work is needed to build a working prototype that can be tested with real vehicles. Current results can be found at the project web site [13].

While the feasibility of the approach has been demonstrated, many issues have yet to be resolved. The matching process can be improved by exploiting the epipolar geometry between the cameras. No use is currently made of the fact that the locations of all of the cameras are known and hence that all epipolar geometry is known for the system. A computational system model must also be developed to handle the massive data flows, camera coordination, and multiprocessor coordination. The code will also need to be rewritten to take advantage of the eventual multiprocessor system for close to real time results.

A longer-term goal is to implement an autonomous system to locate contraband. The first step will be to develop algorithms to highlight likely points of interest for the operator. Part of this automatic inspection process will detect anomalies by comparing undercarriage images to previous images of the same vehicle type. The eventual goal will be to have the system automatically detect anomalies without human intervention.

5. Acknowledgments

This work was partially supported by Atlantic Coast Technologies, Inc., National Science Foundation grant 0105272, and Army Research Office grants DAAD19-99-1-0016 and DAAD19-02-1-0133.

6. References

[1] “Security mirrors \ observation systems”, available at http://www.insight-security.com/pf16.htm#Vehicle%20inspection%20mirrors, last checked 6/26/02.

[2] “Inspection Mirrors”, available at http://www.acrilconvex.com/inspection/default.htm, last checked 6/26/02.

[3] “Search Systems Incorporated: Technical Search & Survellance Equipment”, available at http://www.searchsystems.com/contraband.html, last checked 6/26/02.

[4] “UVI – Under Vehicle Inspection System”, available at http://www.uvisystems.com/ last checked 6/26/02

[5] Z. Zhu, A. R. Hanson, H. Schultz, F. Stolle, E. M. Riseman, “Stereo Mosaics from a Moving Video Camera for Environmental Monitoring”, *First International Workshop on Digital and Computational Video*, December 10, 1999, Tampa, Florida, USA, pp. 45-54

[6] Z. Zhu, E. M. Riseman, A. R. Hanson, “Parallel-Perspective Stereo Mosaics”, The Eighth IEEE International Conference on Computer Vision, Vancouver, Canada, July 4001, vol I, 345-354.

[7] H.-C. Huang, Y.-P. Hung, “Panoramic stereo imaging system with automatic disparity warping and seaming”, *Graphical Models and Image Process*., 60(3): 196-408, 1998.

[8] S. Peleg, M. Ben-Ezra, “Stereo panorama with a single camera”, *CVPR'99*: 395-401

[9] H. -Y. Shum, R. Szeliski, “Stereo reconstruction from multiperspective panoramas”, *ICCV99*, 14-41, 1999.

[10] S. Peleg, J. Herman, “Panoramic mosaics by manifold projection”, *CVPR'97*: 338-343.

[11] H. -Y. Shum, R. Szeliski, “Construction and refinement of panoramic mosaics with global and local alignment”, *ICCV'98*: 953-958.

[12] R. Kumar, H. Sawhney, J. Asmuth, J. Pope, S. Hsu, “Registration of video to geo-referenced imagery”, *ICPR98*,vol.4: 1393-1400.

[13] “Vision Lab at UMass - UVIS”, available at http://vis-www.cs.umass.edu/projects/uvis/index.html, last checked 6/26/02.

[14] “ENVI – RSI’s Remote Sensing Software”, available at http://www.rsinc.com/envi/index.cfm, last checked 6/26/02.