TEC-0095

Site Model Based Image Registration and Change Detection

R. Chellappa P. Burlina
C.L. Lin X. Zhang
L.S. Davis A. Rosenfeld

University of Maryland
Center for Automation Research
Computer Vision Laboratory
College Park, MD 20742-3275

March 1998

Prepared for:
Defense Advanced Research Projects Agency
3701 North Fairfax Drive
Arlington, VA 22203-1714

Monitored by:
U.S. Army Corps of Engineers
Topographic Engineering Center
7701 Telegraph Road
Alexandria, VA 22315-3864

US Army Corps of Engineers
Topographic Engineering Center

T E C

Destroy this report when no longer needed.
Do not return it to the originator.

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

The citation in this report of trade names of commercially available products does not constitute official endorsement or approval of the use of such products.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 1998	Third Annual Sep. 1994 - Sep. 1995

4. TITLE AND SUBTITLE

Site Model Based Image Registration and Change Detection

5. FUNDING NUMBERS

DACA76-92-C-0024

6. AUTHOR(S)

R. Chellappa P. Burlina C.L. Lin X. Zhang
L.S. Davis A. Rosenfeld

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Maryland
Center for Automation Research
Computer Vision Laboratory
College Park, MD 20742-3275

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Defense Advanced Research Projects Agency
3701 North Fairfax Drive, Arlington, VA 22203-1714

U.S. Army Topographic Engineering Center
7701 Telegraph Road, Alexandria, VA 22315-3864

19. SPONSORING / MONITORING AGENCY REPORT NUMBER

TEC-0095

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The University of Maryland (UMD) is one of the BAA contractors performing research on aerial image understanding for the RADIUS project. UMD is contributing model-based and context-based change detection (CD) and monitoring algorithms. The principal accomplishments for this third annual report are the following: (1) improvements to detecting and counting of vehicles on parking lots as well as along roads; (2) monitoring vehicle formations, such as convoys and railroad cars, (3) integrating the vehicle detector within the RCDE platform. UMD used geometric and context information derived from site models to accomplish these tasks.

14. SUBJECT TERMS

Change Detection, Site Monitoring, Vehicle Counting, Image Positioning

15. NUMBER OF PAGES

29

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

Contents

List of Figures

List of Tables

PREFACE

This report was sponsored by the Defense Advanced Research Projects Agency (DARPA) and monitored by the U.S. Army Topographic Engineering Center (TEC) under contract DAC76-92-C-0024, titled, "Site Model Based Image Registration and Change Detection." The DARPA Program Manager was Dr. Thomas Strat, and the TEC Contracting Officer's Representative was Ms. Lauretta Williams.

1 Introduction

The University of Maryland (UMD) is contributing model-based and context-based change detection (CD) and monitoring algorithms to the RADIUS project. Change detection (CD) and monitoring are essential aerial image exploitation tasks [2]. Previous research in this area has relied on general-purpose methods that are unreliable because of changes in weather, illumination, and imaging conditions. The model-based detection schemes presented here incorporate IU techniques whose primitives can be used to direct the system to conduct spatially constrained analyses, whose outcomes may be indicative of occurrences of changes with intelligence significance. The aerial image exploitation system described here is site-model driven and emphasizes using geometric (i.e. object and spatial) primitives, which appear as site models implemented in the RADIUS RCDE environment [1]. The models encode the spatial relationships between fixed objects of interest in a site, such as buildings, roads, etc.

In this report, progress in context-based aerial image analysis is summarized. Model-based approaches are presented for the following: (a) detecting and counting vehicles on parking lots as well as a long road; (b) monitoring object formations, such as convoys on roads or railroad cars on tracks. In our system, specific approaches have been matched to the tasks at hand. A global method using spectral analysis is used for (b), and template matching is used for (a). Site model information is then incorporated in different ways and to various extents, according to the selected detection strategy.

Section 2 reports on detection of convoys and other configurations and Section 3 presents improvements to the vehicle detection and counting module, with special emphasis on integration into the RADIUS testbed. In each section, experimental results are given.

2 Monitoring Vehicle Formations

We considered specific site-model-supported monitoring tasks involving movable objects and vehicles. In the first task, vehicles in selected regions, such as in parking areas, roads, and training grounds, were detected and counted. In the second task, vehicle formations, such as convoys, were detected.

In this section, a detector is described that may be used as a tool for globally qualifying

the degree to which a parking lot is full or determining the presence of platoons or convoys along roads; this detector is primarily intended as an attentional mechanism. Our approach is based on detecting special vehicle configurations – more specifically, configurations that exhibit a periodic behavior. This module is of particular interest since it enables one to look globally for activity in an entire area, a process referred to as Global Activity Detection (GAD). It may also serve as a preprocessing step to the local vehicle detection process.

2.1 Principle of Operation

Let $\mathbf{r}$ denote the position of a point in the image, let $\mathcal{C} : g(\mathbf{r}) = 0$ be the representation of a road center, let s denote the curvilinear abscissa on $\mathcal{C}$, and let $\mathbf{r}_C(s)$ be the image point position as a function of s. Further denote by $k(s)$ the directional derivative along the tangent to $\mathcal{C}$ to the curve averaged over $L(s)$ in the normal direction at position $\mathbf{r}_C(s)$, i.e.

$$k(s) = D_t I(\mathbf{r}_C(s))$$

where w is the vehicle width and $L(s) = \{\mathbf{r} : \mathbf{r} = \lambda \nabla g + \mathbf{r}_C(s), \lambda \in (-\frac{w}{2}, \frac{w}{2})\}$. The presence of organized convoys can be detected from any regular or periodic behavior of directional edges. This can be achieved from spectral analysis of $k(s)$. Let $K(f) = F[k|f]$ denote the Fourier Transform (FT) of $k(s)$, a procedure essentially equivalent to using a linear boost on the FT, thereby eliminating any DC component. A convenient model for $k(s)$ for regularly placed objects, such as a convoy-like formation of vehicles on a road or vehicles in a sufficiently full parking area, is obtained by assuming a multiply-replicated function $c(.)$ corresponding to the edge profile (the directional derivative profile) of a vehicle, or corresponding to a truncated periodic version of this function with additive white noise $n(s)$ (to model clutter structure), i.e.

$$k(s) = \text{rect}(\frac{s}{n/f^*})([c(.) * |f^*| \,\text{comb}(f^* \times .)](s)) + n\ (s) \qquad (1)$$

with $\text{rect}(x) = 1$ if $|x| < 1/2$ and 0 outside, and $\text{comb}(x) = \sum_n \delta(x - n)$. The resulting $K(f)$ then equals

$$K(f) = [\frac{n}{f^*}\text{sinc}(. \times \frac{n}{f^*}) * (C(.)\,\text{comb}(./f^*))](f) + N(f)$$

Detecting convoys on roads or detecting significant parking area occupancy can be carried out by detecting impulsive spectral components and corresponding harmonics from the

PSD (Power Spectral Density) $K(f)^2$. Various strategies, some of which are described later, can be derived to measure this occurrence.

The two-dimensional case is more interesting since it allows for many possible object configurations. This case is also relevant for monitoring parking area activity or regular vehicle formations in open areas. Let $k(\mathbf{r}) = |\nabla I(\mathbf{r})|$, the gradient magnitude at $\mathbf{r}$, and consider the 2-D FT,

$$K(\mathbf{f}) = F[k|\mathbf{f}] = \int\int_{-\infty}^{+\infty} d\mathbf{r}\ k(\mathbf{r})\ \exp -j2\pi\mathbf{r}^T\mathbf{f}$$

with $\mathbf{f} = [f_0, f_1]^T$. For regular 2-D vehicle formations, a totally general configuration model is obtained by assuming arbitrary replications of an arbitrarily rotated object (see Figure 1). Its corresponding gradient magnitude can be described as

$$k(\mathbf{r}) = \sum_{\mathbf{n}} c(\mathbf{R}_{\omega_n}\mathbf{r} - \mathbf{v}_n) + n(\mathbf{r}) \tag{2}$$

with $\mathbf{v}_n$ and $\mathbf{R}_{\omega_n}$ denoting an arbitrarily chosen 2-D translation vector and rotation matrix. Since

$$F[x(\mathbf{Ar})|\mathbf{f}] = |\det(A)|^{-1} F[x(\mathbf{r})|\mathbf{Af}]$$

then with $k(\mathbf{r})$ as in Equation (2) we have

$$K(\mathbf{f}) = \sum_{\mathbf{n}} \exp\left(-j2\pi(\mathbf{R}_{\omega_n}\mathbf{f})^T\mathbf{v}_n\right) F[c(.)|\mathbf{R}_{\omega_n}\mathbf{f}] + N(\mathbf{f})$$

This structure, however, is too loose to yield any useful spectral signature. Further restricting $k(\mathbf{r})$, another convenient and still general configuration model is obtained by letting $\mathbf{R}_{\omega_n} = \mathbf{R}_\omega$ in Equation (2) (same object orientations) and by assuming periodic replication of the object within a support area A. Directions of replication are defined by the vectors $\mathbf{v}_0$ and $\mathbf{v}_1$, or equivalently by the 2-D periodicity matrix $\mathbf{V}$ whose columns are the $\mathbf{v}_i$'s, $i = 0, 1$ (see Figure 1), with $\det(\mathbf{v}_0, \mathbf{v}_1) = \det(\mathbf{V}) \neq 0$. This situation is general enough to describe, by simple selection of the shape of the support region A, a wide array of configurations of special interest such as circular, semi-circular, wedge, pyramid, block or linear formations, a few examples of which are shown in Figure 2. Then, letting $\mathcal{I}_A(\mathbf{r})$ denote the indicator function over A, the resulting spatial representation is given by

$$k(\mathbf{r}) = \mathcal{I}_A(\mathbf{r})([c(\mathbf{R}_\omega .) * *\mathrm{comb2}_{(\mathbf{v}_0,\mathbf{v}_1)}(.)]\,(\mathbf{r})) + n\ (\mathbf{r})$$

with the separable comb2($\mathbf{r}$) defined as $\text{comb2}_{(\mathbf{v}_0,\mathbf{v}_1)}(\mathbf{r}) = \sum_{\mathbf{n}} \delta(\mathbf{r} + \mathbf{V}\mathbf{n})$, with $\mathbf{n} = [i, j]^T$. Its resulting spectral representation is

$$K(\mathbf{f}) = \left[I_A(.) * *(C(\mathbf{R}_\omega .)\, \text{Comb2}_{(\mathbf{v}_0,\mathbf{v}_1)}(.))\right](\mathbf{f}) + N(\mathbf{f})$$

where $I_A(.) = F[\mathcal{I}_A|\mathbf{f}]$, and $\text{Comb2}_{(\mathbf{v}_0,\mathbf{v}_1)}(\mathbf{f})$ satisfies

$$\text{Comb2}_{(\mathbf{v}_0,\mathbf{v}_1)}(\mathbf{f}) = \frac{1}{\det(\mathbf{V})} \sum_{\mathbf{n}} \delta(\mathbf{f} - (\mathbf{V}^{-1})^T \mathbf{n})$$

As in the 1-D case, a particular configuration can be detected by searching for its corresponding 2-D base and harmonic impulsive spectral domain components.

Constraints, which are provided by a 3-D geometrical model of the vehicle, yield *spectral compliance windows* where one can search for impulsive components characteristic of certain types of regular configurations. Let f^* denote the base component arising as a result of a regular vehicle formation on a road, and let N_0 be the sampling rate along the curve $\mathcal{C}$. The average image vehicle length l taken on the image region covered by the road is derived from the known vehicle model and the camera orientation parameters. For safety reasons, vehicles follow each other at distances greater than ϵl, for some $\epsilon > 0$; an upper bound for f^* is then given by $N_0/((1+\epsilon)l)$. Similarly, vehicles following each other at great distances are not considered organized convoys. Alternatively, if the area under study is a parking area, it is considered as a significantly occupied parking area. The maximum distance can be expressed in terms of vehicle lengths l as μl with $\mu > \epsilon$. This yields a lower bound $N_0/((1+\mu)l)$ for the dominant spectral component. In sum, f^* must lie within a *spectral compliance window* W_c, defined as

$$f^* \in W_c \text{ with } W_c = \left\{ f : \frac{N_0}{(1+\mu)l} < f < \frac{N_0}{(1+\epsilon)l} \right\}$$

A convoy is hypothesized if an impulsive component is found at frequency f^*, also referred to here as the base or dominant spectral component. Then f^* must satisfy

$$f^* = \text{argmax}(|K(f)|), \quad \text{where } f \in W_c$$

When the dominant impulsive component is found in the spectral compliance window, corroborating evidence is subsequently checked at harmonic values of f^*. Detection is thereafter declared if a set of conditions are met for the corresponding power at f^* and inside the spectral compliance window (as well as on the corresponding harmonics). These

conditions are reported later in this section. Similar compliance windows are derived in the two-dimensional case. Consider the periodicity matrix decomposition as

$$\mathbf{V} = \hat{\mathbf{V}}\,\mathbf{diag}(\tau_0, \tau_1),\ \hat{\mathbf{V}} = [\hat{\mathbf{v}}_0|\hat{\mathbf{v}}_1]\,,\ |\hat{\mathbf{v}}_0| = |\hat{\mathbf{v}}_1| = 1$$

The spectral periodicity matrix $\mathbf{U} = \mathbf{V}^{-T}$, so that $\mathbf{U} = \hat{\mathbf{V}}^{-T}\mathbf{diag}(\tau_0^{-1}, \tau_1^{-1})$. Also, $\mathbf{u}_i.\mathbf{v}_j = \delta_{i,j},\ i,j = 0,1$, since $\mathbf{U}^T\mathbf{V} = (\mathbf{V}^{-T})^T\mathbf{V} = \mathbf{I}$. Therefore, the spectral periodicity vectors $\mathbf{u}_i$ are orthogonal to their *alternate* spatial counterparts, i.e $\mathbf{u}_i.\mathbf{v}_j = 0$ if $i \neq j$, Also, from $\mathbf{u}_i.\mathbf{v}_i = 1$, and $\mathbf{v}_i = \tau_i\hat{\mathbf{v}}_i$, we have $|\mathbf{u}_i| = |\cos(\mathbf{u}_i, \mathbf{v}_i)|^{-1}\tau_i^{-1} = |\sin(\mathbf{v}_j, \mathbf{v}_i)|^{-1}\tau_i^{-1}$, with $j = \bar{i}$, and we have the decomposition

$$\mathbf{u}_i = \hat{\mathbf{u}}_i\xi_i, \text{with}\, \xi_i = \tau_i^{-1}\,|\sin(\mathbf{v}_0, \mathbf{v}_1)|^{-1} \tag{3}$$

where it is assumed that det $(\mathbf{v}_0, \mathbf{v}_1) \neq 0$. Otherwise, in the degenerate case, one may write the gradient magnitude map as

$$k(\mathbf{r}) = \sum_n c(.) * \delta(\mathbf{r} - n\mathbf{v})$$

for which

$$K(\mathbf{f}) = C(.)\sum_l \delta(\mathbf{v}^T\mathbf{f} - l)$$

which gives rise to impulsive lines orthogonal to the direction of $\mathbf{v}$. In sum, the spectral signature offers clues to the spatial geometry of the detected configuration, and we may therefore determine spatial orientations of the object configurations. Alternatively, if site model information is present, compliance windows can be derived to guide the detection process.

Using geometrical considerations similar to the ones presented in the one-dimensional case, the following compliance (longitudinal and lateral) constraints may be inferred:

$$\begin{aligned} \frac{N_0}{(1+\mu_w)w} < \mathbf{f}^T\hat{\mathbf{u}}_0\,|\sin((\mathbf{v}_0, \mathbf{v}_1)| < \frac{N_0}{(1+\epsilon_w)w} \\ \frac{N_0}{(1+\mu_l)l} < \mathbf{f}^T\hat{\mathbf{u}}_1\,|\sin((\mathbf{v}_0, \mathbf{v}_1)| < \frac{N_0}{(1+\epsilon_l)l} \end{aligned} \tag{4}$$

The search strategy to detect a compliant impulsive component follows various schemes depending on the available model information. If only the direction of the vehicle is

available, then the constraints in Equation (4) translate into radial search spaces on a half-plane along directions $\hat{\mathbf{u}}_0$ and $\hat{\mathbf{u}}_1$, or alternatively into a search sector if an uncertainty on these directions is specified. If only the object dimensions are known, the search is conducted along annular disks, generated by a full 2π revolution of the sets in Equation (4). Finally, if both dimension and directions can be inferred from the site model or from the context, the search space reduces to annular sectors.

We turn next to the detection strategy. A significant component due to clutter noise is present across the spectral compliance window, and detection is often complicated by poor imaging and illumination conditions, which lead to weak gradient magnitude. Inaccuracies in site and road boundary models, as well as occlusions and shadows, may lead to false alarms. With these complications in mind, the following detection rule is chosen for an impulsive component at a base and corresponding harmonic frequencies:
Detection is declared if conditions (a) and (b) are met simultaneously:

1. *Power Ratio Test*: $|K_r(f^*)| > \kappa$,

2. *Absolute Power Test*: $|K(f^*)| > \gamma$

where $K_r(f) = K(f)/K_{\text{med}}$, and K_{med} denotes the "median" power inside the spectral compliance window. If these conditions are met, then the corresponding harmonics of f^* are checked for the presence of an impulsive component using the above decision rules, i.e. $|K_r(f_i^*)| > \kappa_i$ and $|K(f_i^*)| > \gamma_i$, where the subscript i indicates the ith harmonic of the base frequency f^*, $f_i* = if^*$. In practice, it is often difficult to verify more than the first harmonic. Furthermore, in order to simplify tuning and keep thresholds constant across images and viewing directions, the spectral estimates are normalized by the aspect ratio. Call the resulting normalized power the absolute power K_a. Further tests can involve comparisons with the spectral components of $k(.)$ for other images where the given road is known to be empty.

Varying illumination conditions lead to variations in power associated with a given directional derivative map. To allow for these variations, the above detection rule on the absolute power K_a and power ratio K_r is replaced by an acceptance region defined on the space (K_a, K_r). This acceptance region is based upon the premise that peaks with lower absolute power should be tolerated for larger power ratios. We define a boundary region in the space (K_a, K_r), $K_r = b(K_a)$. The detection rule then becomes

$$d(K_a, K_r) = \mathcal{I}_{R_b}(K_a, K_r)$$

where $\mathcal{I}_R$ is the indicator function over region R, and R_b is defined as

$$R_b = \{(K_a, K_r), \text{s.t. } K_r \geq b(K_a)\}$$

Tuning the function $b(.)$ is neither an acceptable nor a practical option due to the large and diverse amount of remotely sensed imagery typically available for exploitation. The optimal acceptance region is instead learned from a representative control set of test images. Let $\mathcal{S}$ denote this control set. The acceptance region boundary is parameterized by vector $\mathcal{V}$, i.e. $b_{\mathcal{V}}(.)$ (an example of a possible parametric function is a decaying exponential). For the previous detection rule d and control set $\mathcal{S}$, we are interested in optimizing the choice of $\mathcal{V}$ for an objective function $O(d)$, which includes the probability of false alarm $P_f(d)$ and probability of detection $P_d(d)$,

$$O(d) = C_{nd}(1 - P_d(d)) + C_{fa}P_f(d)$$

where the coefficients C_{fa} and C_{nd}, respectively, denote the selected costs associated with occurrence of a false alarm and with non-detection. C_{nd} and C_{fa} reflect the priorities or the final function of the GAD (preprocessor, stand-alone). In sum, $\mathcal{V}^*$, the optimal choice of the parameter $\mathcal{V}$, is one that minimizes the selected cost function computed over the control set $\mathcal{S}$ of images, i.e.

$$\mathcal{V}^* = \text{Argmin}(O(d)), \text{ with } d(K_a, K_r) = \mathcal{I}_{R_{b(v)}}(K_a, K_r) \tag{5}$$

where $R_{b(\mathcal{V})} = \{(K_a, K_r), \text{s.t. } K_r \geq b(K_a, \mathcal{V})\}$. If the GAD is used as preprocessor to local object detection, then higher false alarm rates can be tolerated.

Global activity detection can naturally serve as a preprocessing step to local object detection and vehicle detection. We can infer the spatial positions of edges, giving rise to the periodic structure determined through frequency domain analysis, referred to here as *periodic loci*, and thereafter apply local vehicle detection schemes to corroborate the presence of a vehicle at these positions. In the 1-D case, $k(s)$ is band-pass-filtered around the base frequency f^* and harmonic frequencies (with impulse response denoted by $h(s)$). The locations of edges satisfying the periodic structure are then determined by the set of locations s satisfying the condition

$$|k(.) * h(.)\ (s)| > \nu k_m \tag{6}$$

for some $\nu < 1$, with k_m denoting the median of the absolute value of $k(s)$. Let $\mathcal{L}$ denote the set of *periodic loci* satisfying Eq. (6). A local vehicle detector can then be applied to a reduced spatial window centered at these locations.

2.2 Experiments

The global activity detector was run on the complete set of 40 RADIUS model board MB2 images. Detection and false alarm rates are reported in Table 1 for hand-tuned and learned-decision rules. Results for tuned road activity are given in Table 1. These detection results are very satisfactory, given that the decision rule and acceptance region remain the same across images. Detection results are shown in the following figures for various images. Figure 3(a) shows a newly acquired image; GAD is applied to operator-selected areas (shown in yellow). Figures 3(b) and (c) show detected active parking areas. The resulting detected active roads are highlighted in Figure 3(d) and Figure 4(a). Figure 4 (b, c and d) shows the results when GAD is applied to all the monitored areas shown in Figure 3(a). Active areas are shown in red, and inactive roads are shown in green. Inactive parking lots are not highlighted. The resulting set of periodic loci $\mathcal{L}$ is shown in Figure 5(a) and again in Figure 5(b) (enlarged). After applying the vehicle detector described next, the resulting detected vehicles are shown in Figures 5(c) and (d).

3 Integration and Improvements of the Detection and Counting Module

In this section, we review briefly the basic principles of the vehicle detection scheme. Then, we report on the improvements made during the reporting period and on the RCDE integration of this module.

Principle of Operation. The vehicle detector can be triggered, or cued, by the previous GAD or can be used independently to monitor areas, such as parking lots, roads or training grounds. The general vehicle detection scheme relies on contour matching using information derived from the geometric model. As we are primarily concerned with high-altitude imagery in our implementation, vehicles are modeled as 3-D cuboids with given width, length and height specifications. The implementation of our vehicle detector consists of a prescreener, an extractor and a verifier. The first stage is replaced by the hypothesized vehicle locations at the periodic locus $\mathcal{L}$ if the GAD is used as a preprocessor, with the last two stages remaining the same. The prescreener is applied instead if the vehicle detector is used stand-alone. This prescreener relies on Hough Transform techniques to locate areas where the centers of vehicles are likely to lie. The extractor performs template matching and hypothesized vehicles are subsequently examined in more detail. The verifier checks

for shadows to guarantee the correctness of the results. Throughout the detection procedure, 3-D object models and site information (camera model, illuminant, etc.) are used. The above vehicle detection scheme has been fully integrated into the RCDE exploitation platform. Results of its application to both parking lots and roads are shown below on real images.

Assessing Operational Requirements for the Vehicle Detector. We have tested the vehicle detector on the Ft. Hood images so as to assess reasonable operational conditions. These conditions will be coded into Prolog rules in the SRI HUB [2].

Examples of results obtained with Ft. Hood images are shown in Figure 6 (a-d). Table 2 shows detection performance results on a set of Ft. Hood parking lots. Based on these experiments, we conclude that the current vehicle detector modules should be used with caution when any of the following conditions are present:

- Poor contrast.
- Very cluttered environment or very heterogeneous objects.
  (such as some Ft. Hood storage areas).
- Vehicles of greatly varying orientations in a parking lot when only one orientation was specified.
- GSD (Ground Surface Dimension) such that the image dimension of a vehicle is ≤ 6 or 7 pixels.
- Presence of water pools or oil marks (noted by J. Mundy).

The knowledge of these conditions is of great interest to us. The conditions will be coded into Prolog rules in a RCDE HUB subsystem, which will control the appropriate application of detection algorithms, according to the operational conditions.

An Automatic Ground Truth Verifier. The above statistics were obtained by inspection. We have now implemented an automatic ground truthing tool to allow for the automatic assessment of the statistical performance of our vehicle detector. GE has developed and integrated a very useful tool allowing systematic evaluation of results obtained with the vehicle detector: the ground truth verifier. This is primarily a developer's tool

allowing us to perform a fast evaluation of the results when a change in the vehicle detection scheme has been made. The current criterion for a correct detection is the presence of a detected vehicle within a pre-specified distance from the center of a specified ground truth vehicle. This distance could be replaced by a measure reflecting the GSD and the vehicle size.

We have modified the ground truth verifier to allow for this. As a first approximation, the original detection criterion was modified to declare a correct detection whenever the center of the detected vehicle is within the boundaries of a specified ground truth box. This rule will later be refined so as to dissociate the detection rate and the quality of the localization. We modified the module to allow for saving previous detection results and ground truth as site model objects. This is the preferred method for specifying the ground truth, since it can be reloaded and is more easily exchanged.

Results obtained with this ground truth verifier are reported for the Denver-TEC-2 images. Figures 7 (a-c) show an original image of the TEC-2 site with parking lots delineated, the ground truth vehicles, and the detection results. Table 3 shows the resulting statistics for the TEC-2 images.

Long-Road Detector. Most IU algorithms developed for aerial image exploitation fail to take into consideration situations where very large images are used. These images call for specific computational and sometimes algorithmic solutions. We have begun to address this problem in the context of vehicle detection on long roads. To this end, site model information in the form of the road model is fed back to the low-level processing stages. More specifically, the Canny edge detection module is modified so as to perform filtering, gradient computation, non-maxima suppression, and hysteresis thresholding only on masked regions. The block image structure of images will be exploited later, with processing performed only on the block traversed by a road, which further limits memory requirements. Lastly, template specifications in the case of a long road will consider the facts that vehicles are often isolated, while some edges in parking lots are occluded. Results of the application of the long-road monitor on Ft. Hood images are shown in Figure 8; statistical results are given in Table 4. Note that on the Ft. Hood images, the detection module is applied in situations where it is known to perform poorly, i.e. in cases where the apparent vehicle size is below 7 pixels.

Main modifications and Bug-Fix History .

The vehicle detector has been fully integrated into the RCDE platform. It has been implemented in C, and later reimplemented in C++. Its current integrated version, however, is still in C. The early timing of its integration and testing was as follows: in 1993 to 1994 the basic concept was tested and early integration occurred. In September 1994, the 2-D version was integrated into the RADIUS testbed and tested. In Spring 1995, the vehicle detector was integrated into the Lockheed-Martin Valley Forge (VF) Quick Look (QL) menus and a 3-D version was delivered to GE and VF. It was demonstrated at the San Diego ETS demo as well as the STISS demo.

The integrated vehicle detector has been tested by GE and Lockheed-Martin, with integration resulting from our collaboration with these institutions. VF has provided the needed RCDE upgrades, the testbed version, and disk-saves. It has provided us with feedback with bugs and fixes as well. GE has provided test images and a site model (Schenectady), and has provided assistance in getting RCDE upgrades and disk-saves. It has assisted us in the RCDE integration, particularly with the QL menus. Over multiple and regular integration trips made by R. Welty (GE), GE has given us feedback on implementation issues, bug detection and fixes. It has tested compatibility of our code with major RCDE upgrades. Some examples of the interactions are itemized below:

- **Formats:**
  - FBIP: The vehicle detector was found (VF) not to work on FBIP format. It was then modified (UMD) to work on FBIP images by removing all references to interior camera parameters (Spring 1995).
  - Lazy Images
    * The vehicle detector did not initially work on lazy images (sometimes viewless) (detected at VF, May 1995).
    * This was fixed by removing references to views (R. Welty, May 1995).
- **Interactive dimension specifications:**
  - These were inadequate in their initial implementation and were subsequently improved.
- **Orientation anomaly:**

- A mis-alignment between image coordinate systems and 2-D coordinate systems was detected (UMD, June 1995).
- This mis-alignment resulted in great deterioration of the detection performance.
- This was fixed by modifying the transformations used (UMD, June '95)

Current Features. This last item has greatly affected the detection performance. We expect that a version in which this problem has been fixed will be included in the next delivery in November 1995. The current vehicle detector features that the RADIUS testbed should have in the next delivery will include the following. In its current implementation the vehicle detector is accessible either stand-alone or under the QL menus. The detector allows for dual specification of the vehicle dimensions: 2-D or 3-D. It enables a two-way interactive specification of the dimensions, either graphically or by using sliding bars. In its current implementation, the user creates a sample image vehicle, modifies its dimensions, then uses it as a prototype. The user can modify its 3-D dimensions directly. The sample vehicle size is then modified accordingly. Alternatively, the user can modify the sample vehicle sizes and the 3-D dimensions are modified accordingly. The detector accepts FBIP format and lazy images. It is also equipped with a ground truth verifier.

4 Conclusion

We have presented the recent work performed under the RADIUS project for the development of model-supported algorithms for monitoring vehicle-related activities. Special emphasis has been placed on integrating the vehicle detector module into the RCDE environment.

Abbreviations and Acronyms

The following abbreviations and acronyms were used in this report:

UMD:	University of Maryland.
VF:	Valley Forge (Lockheed-Martin).
GE:	General Electric.
RADIUS:	Research and Development in Image Understanding Systems.
RCDE:	RADIUS Common Development Environment.
FT:	Fourier Transform.
GAD:	Global Activity Detector.
FBIP:	Fast Block Interpolation Projection.
GSD:	Ground Surface Dimension.
QL :	Quick Look.

References

[1] SRI International and Martin Marietta, *RADIUS Common Development Environment:User's Manual*, July 1993.

[2] T. Strat and W. Climenson, "Site Model Content," in *Proc. ARPA Image Understanding Workshop* (Monterey, CA), pp. 277–285, 1994.

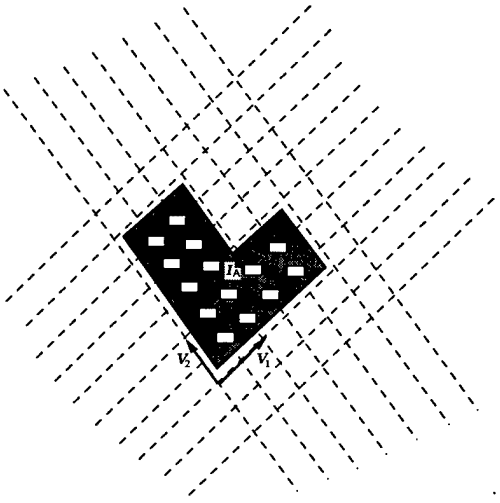


Figure 1: Canonical Configuration

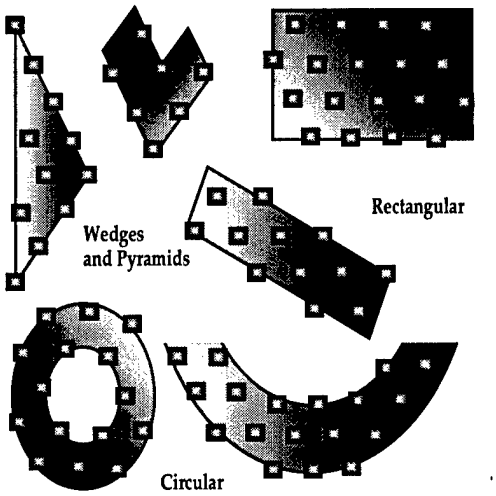
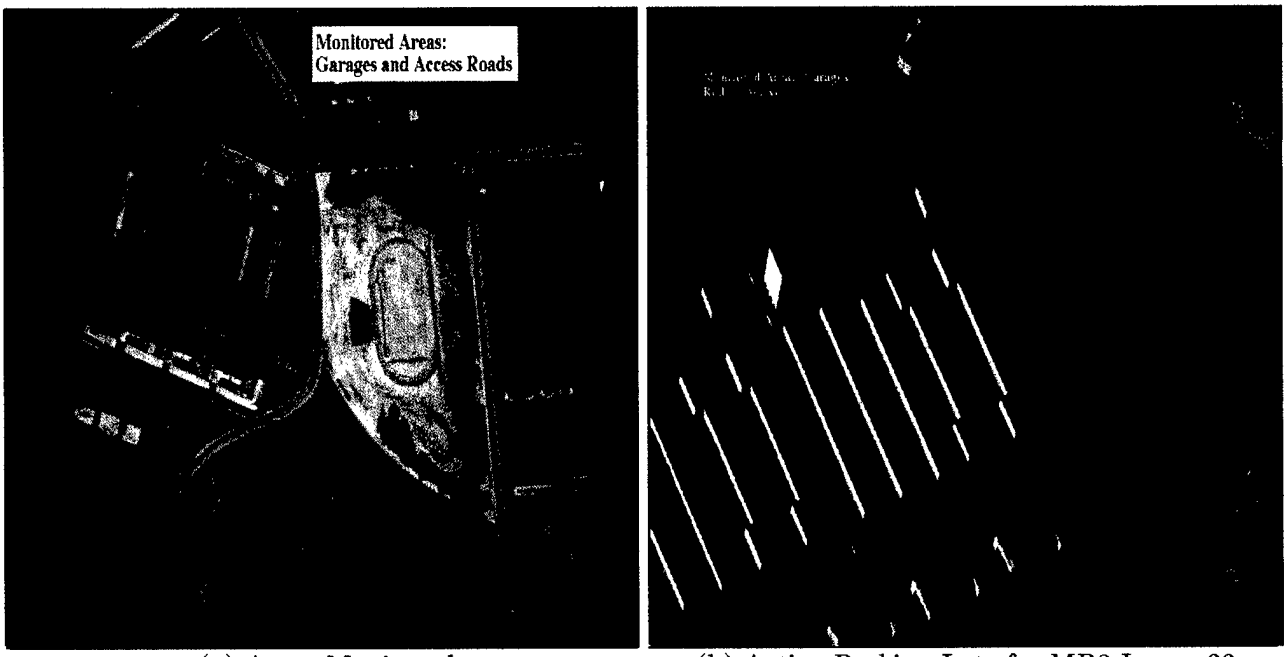


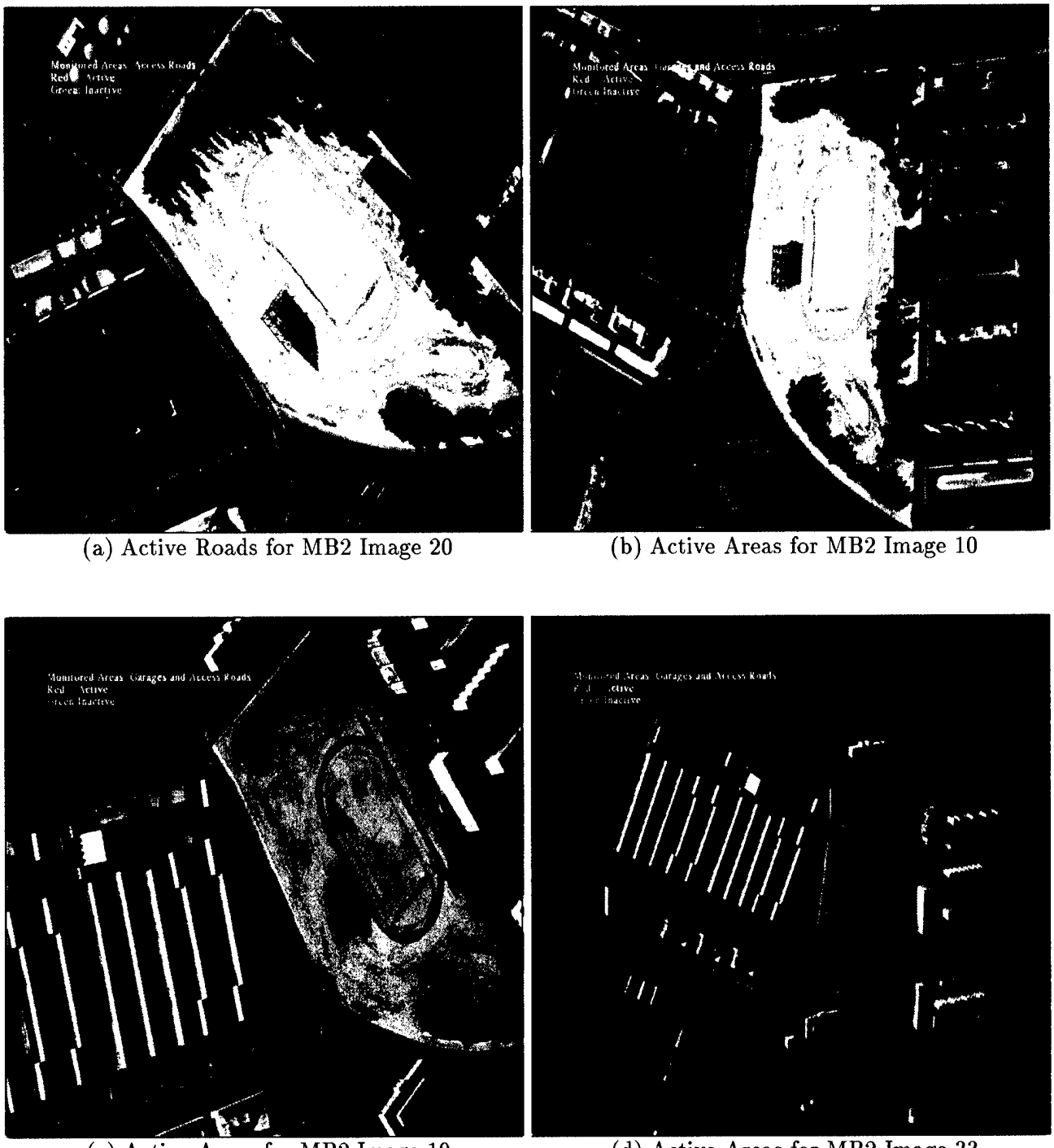
Figure 2: Typical 2D Configurations



(a) Areas Monitored (b) Active Parking Lots for MB2 Image 33

(c) Active Parking Lots for MB2 Image 23 (d) Active Roads for MB2 Image 18

Figure 3: Global Detector: Monitoring of Parking Lot Occupancy and Road Convoys.



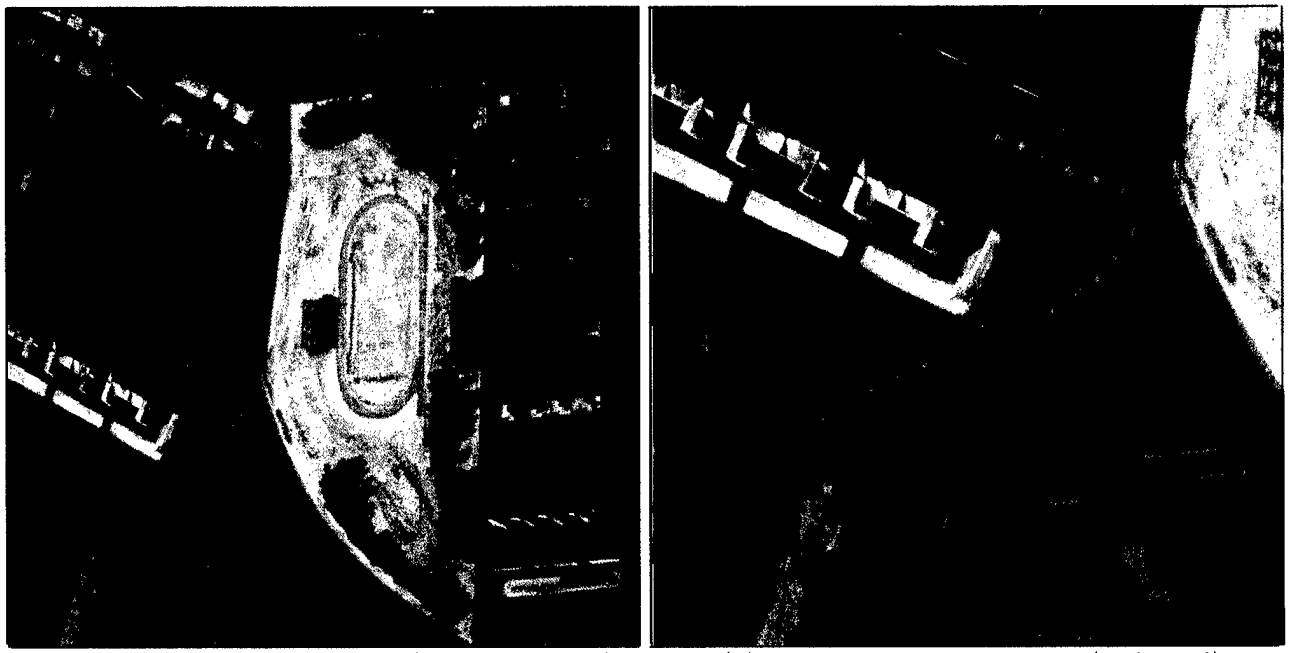
(a) Active Roads for MB2 Image 20

(b) Active Areas for MB2 Image 10

(c) Active Areas for MB2 Image 19

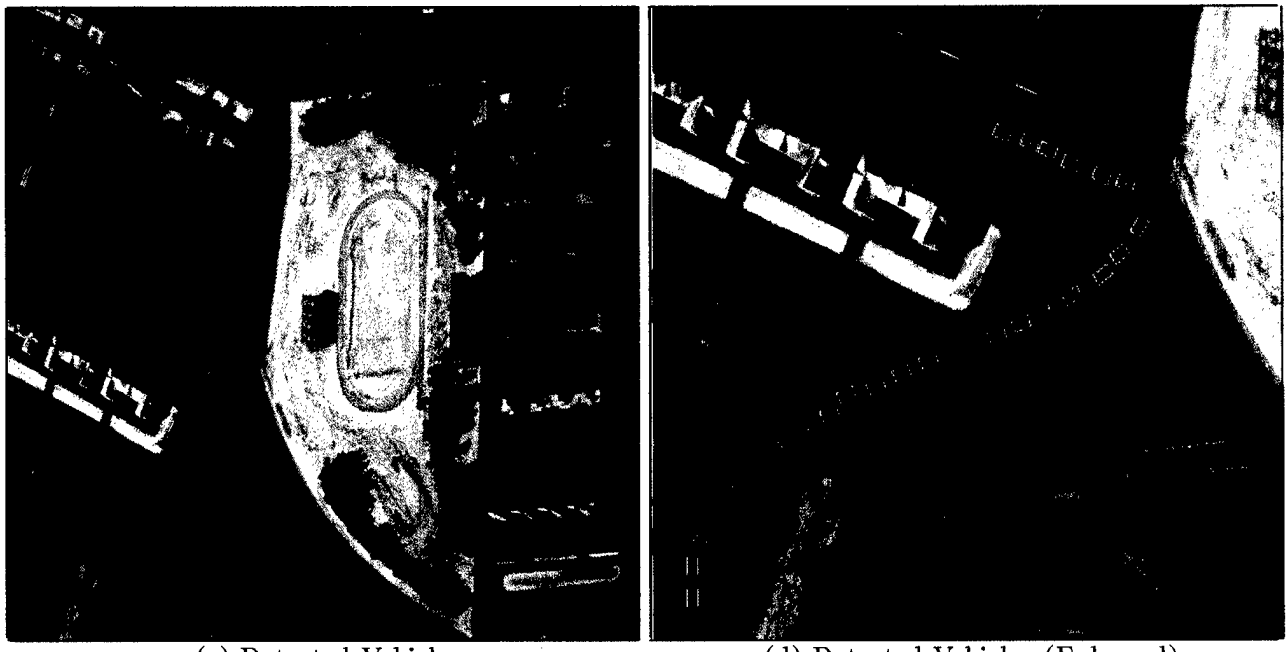
(d) Active Areas for MB2 Image 33

Figure 4: Global Detector: Monitoring of Parking Lot Occupancy and Road Convoys.



(a) Backprojected Locations (MB2 Image 10)

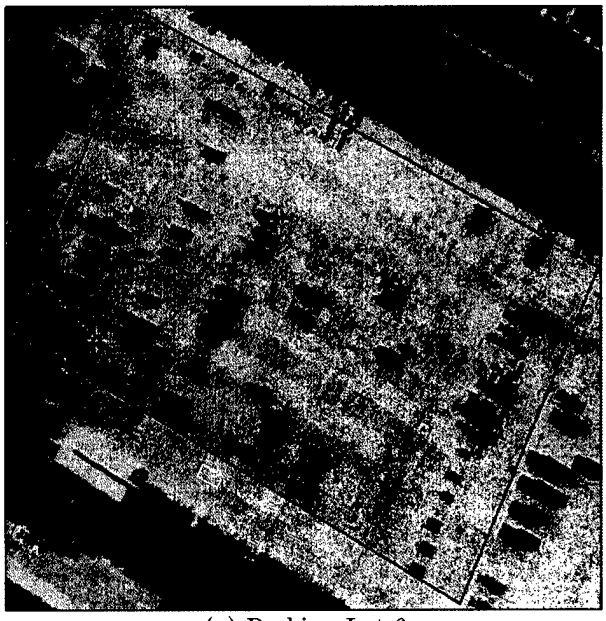
(b) Backprojected Locations (Enlarged)



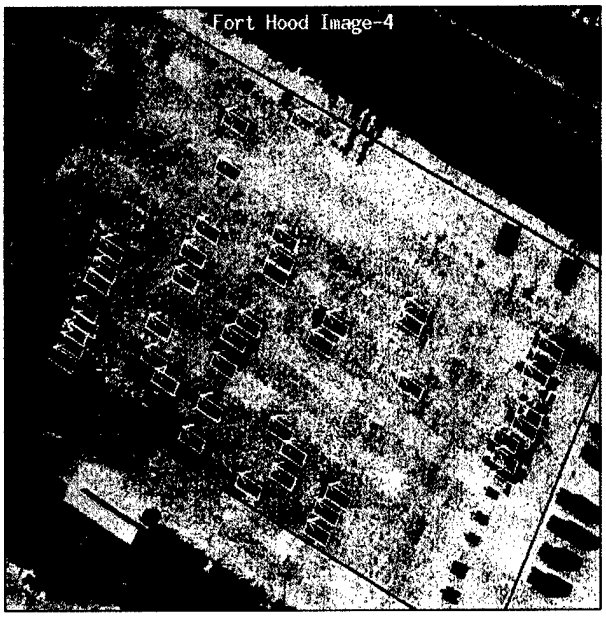
(c) Detected Vehicles

(d) Detected Vehicles (Enlarged)

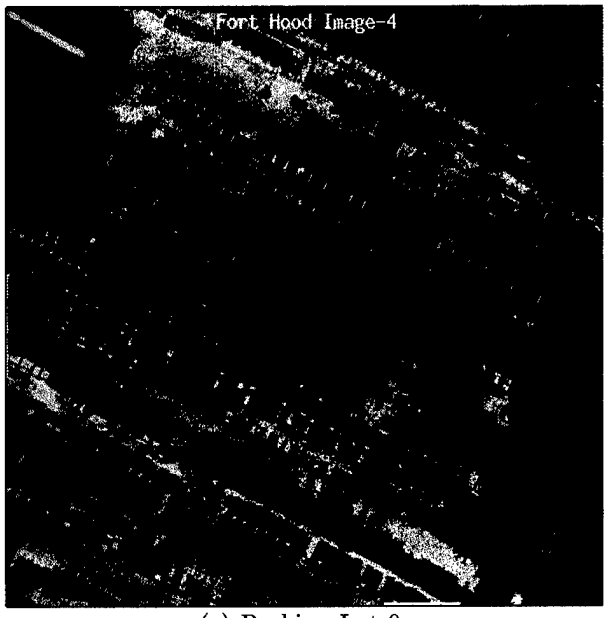
Figure 5: Backprojection and Local Vehicle Detection



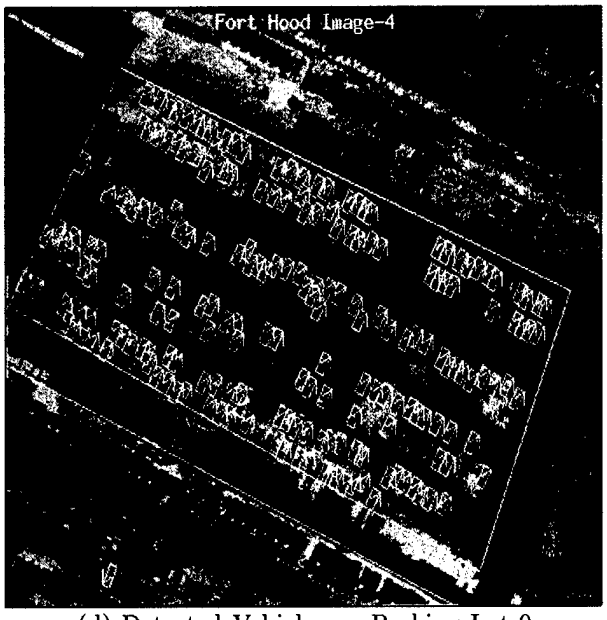
(a) Parking Lot 6.



(b) Detected Vehicles on Parking Lot 6.

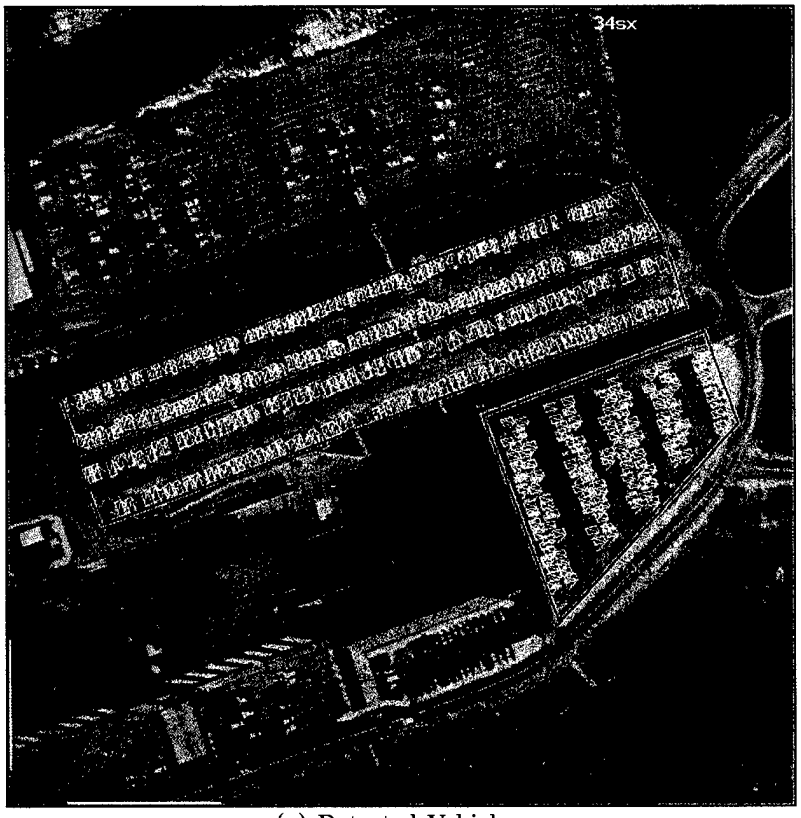


(c) Parking Lot 9.

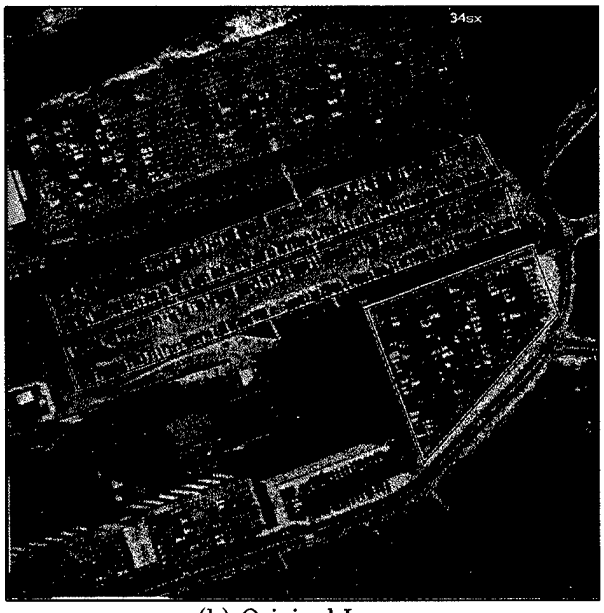


(d) Detected Vehicles on Parking Lot 9.

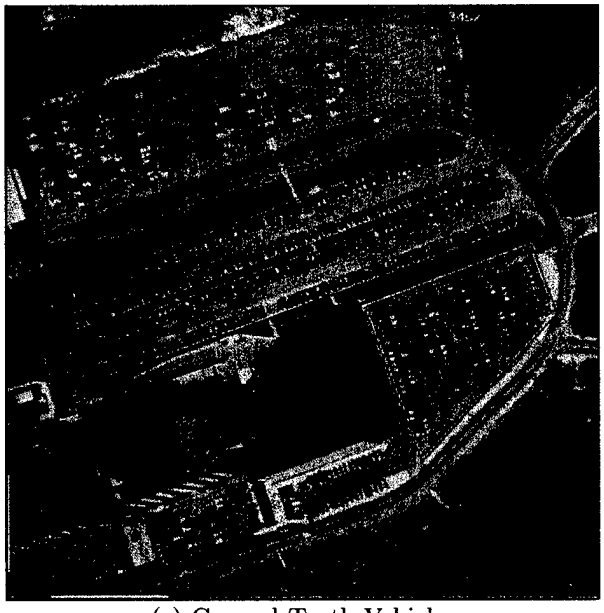
Figure 6: Vehicle Detection Results on Ft. Hood Image 4.



(a) Detected Vehicles.

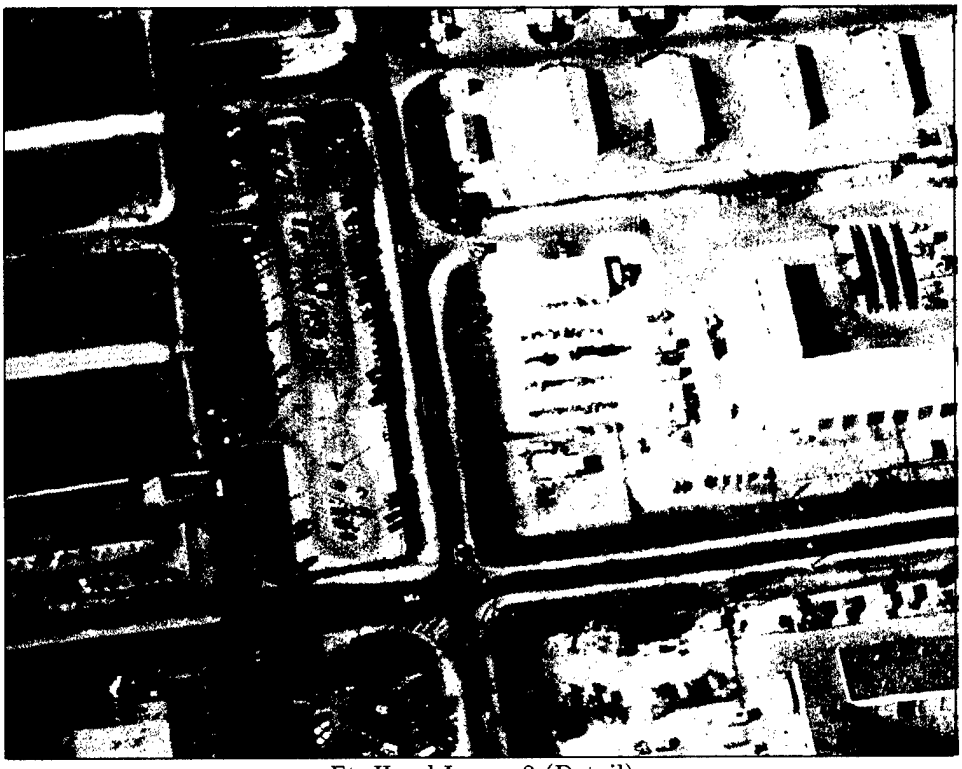


(b) Original Image.



(c) Ground Truth Vehicles.

Figure 7: Vehicle Detection Results on Denver Image.



Ft. Hood Image 2 (Detail).

Ft. Hood Image 2 (Detail).

Figure 8: Long Road Monitoring on Ft. Hood Image 2.

Table 1: Detection Results on Roads and Parking Areas

Areas	Active	Inactive	$P_f(d)$	$P_d(d)$
Parking areas (learned/optimal)	69	86	0.1123	0.9714
Parking areas (hand-tuned)	69	86	0.1046	0.9421
Road segments (hand-tuned)	24	235	0.1702	0.8340

Table 2: Detection Statistics for Ft. Hood

image-parking lot	true veh.	correct det.	false det.	missed det.	det. rate	false alarms
4-6	47	41	3	6	87.23 %	6.81 %
4-8	166	150	13	16	90.36 %	7.97 %
4-9	218	201	24	17	92.20 %	10.66 %
4-11	224	187	14	37	83.48 %	6.96 %
4-a1	33	22	1	11	66.66 %	4.34 %
4-a2	40	31	1	9	77.50 %	3.12 %
4-a3	74	58	5	16	78.37 %	7.93 %
4-a4	91	69	1	22	75.82 %	1.42 %

Table 3: Detection Statistics for TEC-2

image-parking lot	true veh.	correct det.	false det.	missed det.	det. rate	false alarms
34-1	79	67	15	12	84.81 %	18.29 %
34-2	53	36	4	17	67.92 %	10.00 %
34-3	51	45	9	6	88.23 %	16.66 %
34-4	149	118	12	31	79.19 %	9.23 %
34-5	165	122	10	43	73.93 %	7.57 %
34-6	165	150	26	15	90.90 %	14.77 %

Table 4: Detection Statistics for Ft. Hood Long Road

image	true veh.	correct det.	false det.	missed det.	det. rate	false alarms
Ft. Hood image 2	14	8	0	6	57.14%	0.00%
Ft. Hood image 3	34	23	0	11	67.64%	0.00%
MB2	50	44	4	6	88.00%	8.33 %