Elevation-Based Parking Lot Analysis, Visualization, and Simulation

X. Wang & A.R. Hanson

CMPSCI TR 98-01

January 1998

Elevation-Based Parking Lot Analysis, Visualization, and Simulation *

Xiaoguang Wang and Allen R. Hanson

Department of Computer Science
University of Massachusetts
Amherst, MA 01003
Email: {xwang, hanson}@cs.umass.edu

Abstract

We propose an elevation-based approach to parking lot structure analysis from aerial imagery. In contrast to image-based methods, the new approach treats parked vehicles as 3-D microstructures and attempts to locate them in the elevation domain. The STME (Surface Texture and Microstructure Extraction) system is applied to the elevation map to extract the 3-D microstructures. A hybrid application of this system to both intensity and elevation maps results in a complete extraction of individual vehicles. Based on a comparison of texture exploitation techniques, a WCC (weighted combination criterion) algorithm is presented to generate a clean parking lot ground model, which facilitates visualization and simulation of parking lot activities with a high degree of visual realism.

Keywords: aerial imagery, parking lot analysis, 3-D microstructure, stereo vision, visualization and virtual reality

*This work was funded by the RADIUS project under ARPA/Army TEC contract DACA76-92-C-0041, NSF grant CDA-8922572, and ARO DURIP grant DAAG55-97-1-0026.

1 Introduction

Parking lots are an important object class in many military and civilian applications. Goals of parking lot surveillance include counting the number of parked motor vehicles, monitoring the changes of the parked vehicles over the time, identifying the location, the size, or even the type of each vehicle, simulating/visualizing the vehicle activities in the parking lot, and providing a context for monitoring human activity [8, 10].

The most notable success in parking lot analysis from aerial imagery is reported by Chellappa et al. [2]. In their work, a frequency domain computation is employed to detect the vehicle configurations that exhibit a periodic behavior, and a series of image processing algorithms are carried out to identify individual vehicles from optical intensity images. A problem with this approach is that it only uses information from a single image and is restricted to intensity domain analysis. Examining the parking lot in the Lockheed/Martin site (Fig. 1), we can see that the vehicle appearances are very different in brightness, texture, reflectance, etc. Shadows are not a reliable cue, either, since they change over time. Because of these inherent variations, an algorithm which uses only a single image can become complicated.

In this paper, we resort to a different domain – the elevation domain – to detect vehicles. In a parking lot, all vehicles have the common property that they are higher than the ground. Thus, if height information is available, we are able to abandon the intensity appearance of vehicles completely when locating the vehicles. This strategy is particularly useful when analyzing aerial imagery because multiple images are usually available and elevation information can be obtained fairly easily. Section 2 describes a stereo algorithm that generates an elevation map from a pair of images.

The problem with elevation data is that it is an unstructured approximation of the surface height relative to a reference plane. For the many goals of parking lot analysis,

structured geometric information must be extracted. While much work has been done on extracting man-made structures in urban sites from aerial images [4, 5, 6, 9, 11, 12], these algorithms are more focused on large-scale structures such as buildings rather than on small objects (which occupy fewer pixels and possess less structural cues). Wang and Hanson [18] proposed a system for extracting microstructures that are attached to the surfaces of large-scale structures in a regular pattern. An important property of their system is the simplicity of its microstructure extraction algorithm due to the powerful constraints provided by the large-scale structure models. The deficiency of the system is that it only works on 2-D microstructures (e.g. windows).

In this paper we treat the motor vehicles in parking lots as 3-D microstructures. A very simple extraction algorithm is applied to the elevation domain to extract the vehicles as cuboids (Section 3). The entire structure of the parking lot is then recovered by a hybrid use of the simple extraction algorithm in both the elevation and the intensity domains. Once the parking lot structure is recovered, we are able to visualize and simulate vehicle activities using visually realistic image textures. An algorithm for obtaining such textures from the image domain is presented in Section 4.

2 The Elevation Domain

To extract the height features of objects on the ground, a stereo terrain reconstruction algorithm [13, 14] is employed. Input to this algorithm are two aerial images, I^F and I^G, of the same region. With known camera parameters, the algorithm utilizes the epipolar geometry (Fig. 2) to compute an *elevation map*, which approximates the relative height above a reference plane at each pixel.

Every 2-D point P_F in I^F, is correlated with points within a prespecified window along its epipolar line in I^G, under the assumption that the terrain is a nearly Lambertian

Figure 1: Generating the elevation map from a pair of Lockheed/Martin images

(a) the aerial images

(b) the elevation map

surface. For a point P_G on the epipolar line in I^G, the correlation is computed as

$$\rho(P_F; P_G) = \frac{\text{Cov}[P_F, P_G]}{\text{Var}[P_F]\text{Var}[P_G]}, \tag{1}$$

in which $\text{Var}[x]$ is the intensity variance of the image patch around point x and $\text{Cov}[x, y]$ is the covariance of intensities in the image patches around x and y. The point $\tilde{P}_G$ with the highest correlation on the epipolar line in I^G tends to be the true correspondence of P_F; that is, P_F corresponds to $\tilde{P}_G$ if

$$\rho(P_F; \tilde{P}_G) = \max\{\rho(P_F; P_G)\}. \tag{2}$$

Once these correspondences are obtained, the elevation map can be computed by solving the stereo observation equations [15].

Fig. 1 shows the elevation map generated from a pair of Lockheed/Martin images. Fig. 3(a) and (b) shows an orthographic version of the intensity and elevation maps. We can see that all the "bumpy" areas in the elevation map correspond quite well with the vehicles, while the ground areas appear to be flat and smooth. In this sense, we have

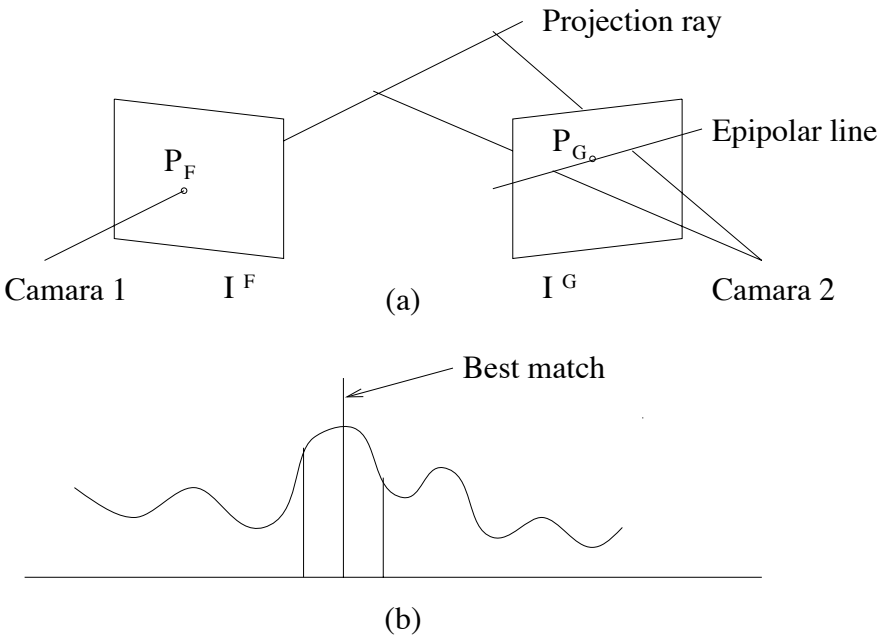


Figure 2: Concepts of epipolar geometry and correlation function

(a) epipolar geometry of an image pair

(b) correlation function of a 2-D point in the left image

recovered an important feature – the difference in height – that distinguishes the vehicles from the ground. Clearly, for the sake of correct elevation data, it is required that the image pair share the same vehicle configuration. Although this is not always true (cars may be in motion), statistically most part of a parking lot have a good chance to be static during the instant when two images are taken from a flying airplane.

3 3-D Microstructure Extraction

3.1 The STME system

The elevation map reveals important information about the vehicles in a parking lot; however, extracting the individual vehicles from the unstructured elevation data poses an

interesting problem. Fig. 3(c) shows a failed attempt to extract the vehicles by using a global thresholding algorithm applied to the raw elevation map. That the vehicles are higher than the ground is true only locally; in a global view, due to the natural incline of the ground or due to errors in the elevation estimates, global thresholding would clearly produce unacceptable results. Furthermore, when two (or more) vehicles are parked close to one another and the view angle does not allow the ground plane between the cars to be seen in both images, the elevation data for the two vehicles can be merged into one region (four cars in lower center of Fig. 3(a), for example).

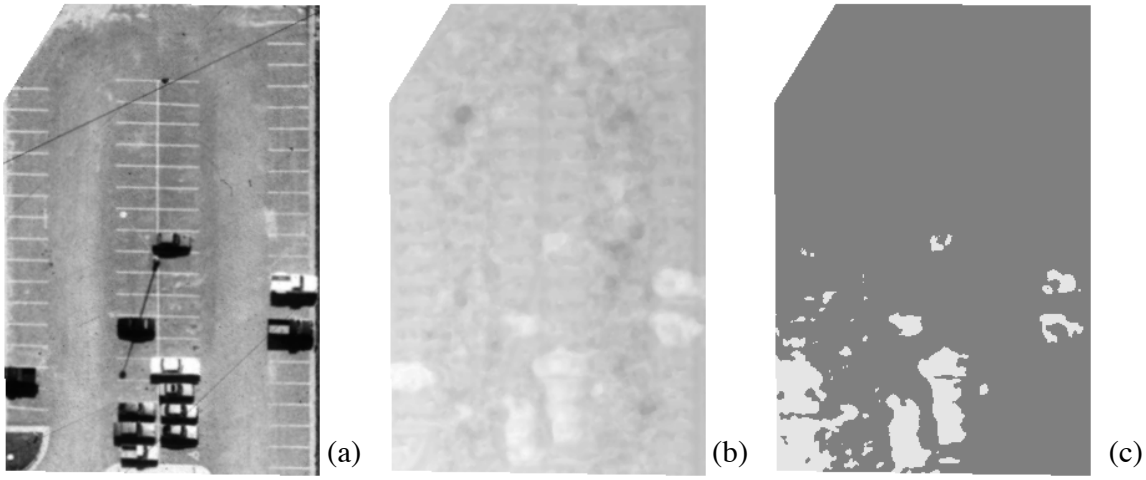


Figure 3: Orthographic versions of the intensity and elevation maps of the parking lot
(a) intensity image
(b) elevation map
(c) a failed attempt to extract vehicles by globally thresholding

These problems have been solved successfully by the Surface Texture and Microstructure Extraction (STME) system. We have distinguished *microstructures* from *large-scale structures (LSS)* in model reconstruction problems [17, 18, 19]. LSS are structures such as buildings that show sufficient structural cues (corners, edges, etc.) in aerial images

for their unambiguous recognition. Microstructures are small structures (windows, roof vents, etc.) sized near the limit of resolution in the images. Because of the deficiency of supporting data in the images, these small structures cannot be extracted using the structural cues that many LSS extraction systems (e.g. [4, 5, 6, 9, 11, 12]) rely on.

The STME system was originally designed for symbolic extraction of 2-D surface microstructures. Its design philosophy is based on the fact that many man-made microstructures appear in rectilinear, repetitive patterns attached to regular, planar surfaces (e.g. building roofs and walls). With these constraints, symbolic extraction of the microstrcutures can be done in a very efficient way from noisy images, even if the small objects lack sufficient supporting pixels. Fig. 4 shows an application of the STME system to a window extraction problem in an extremely noisy environment. The presence of a window is at best a local intensity "dip" with a couple of supporting pixels. Fig. 4(b) shows the result of extracting windows using the algorithm described in [19]: windows are approximated at local intensity dips as rectangles that form a rectilinear lattice globally. Missing windows were filled in using the knowledge that the window lattice is regular (Fig. 4(c)).

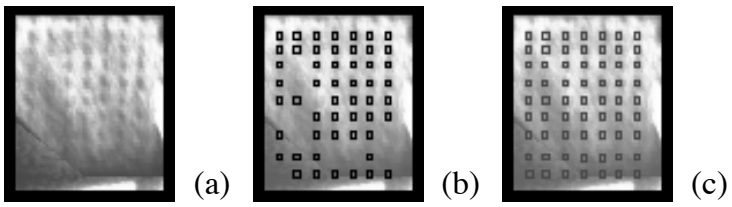


Figure 4: Window extraction using the STME system

(a) wall image in the facet coordinate

(b) extraction of local intensity dips

(c) filling in the missing windows

3.2 Vehicle extraction

Although vehicles in parking lots significantly differ from windows on building walls, they share the same basic characteristics of man-made structures: a repetitive rectilinear alignment of objects. Once the parking lot surface is mapped to the elevation domain, vehicles resemble windows: vehicles appear as "bumps" in an elevation map while windows are "dips" in an intensity map. Applying the STME system to the elevation domain enables the system to deal with 3-D microstructures.

Fig. 5(e) shows the extraction of the dips from the reversed version of the elevation map in (d) using the same approach used to extract windows in Fig. 4. The rectangles that signify possible vehicles are generated by an *oriented region growing (ORG)* algorithm. The algorithm starts at a local minimum on the reversed elevation map and grows the region from a small rectangle to a large one. The region is forced to grow as a rectangle, and only along two orthogonal directions. In each direction, the elevations on the frontier of the region are averaged, and the changes of the averaged elevation are inspected as the frontier is grown outward. The frontier ceases to grow at the position where the averaged elevation on the frontier reaches its maximal first order derivative. The rectangle thus obtained is taken as a hypothesis of a vehicle in the parking lot. A detailed implementation of the ORG algorithm, as well as other issues in the STME system such as finding the initial seeds of the growing regions and adjusting the final rectangles using knowledge of repetitive patterns, can be found in [17].

It is worth noting that the ORG algorithm is featured by its extreme simplicity of computation. This is partly because the image of a wall or a parking lot is stored in a *facet coordinate system*; that is, the edges of the rectangular objects are either perpendicular or parallel to the x and y axes. This involves finding the orientation of the parking lot. Chellappa et al.'s frequency domain analysis [1, 2] is one way to do it; in our work,

the facet coordinate is obtained by an LSS analysis of the parking lot as a precursor to the STME analysis, because the boundary of the parking lot (Fig. 1) provides enough structural cues to establish its orientation.

3.3 Individual vehicle identification

All the bumpy areas (dips in the reversed elevation map) have been extracted successfully in Fig. 5(e). However, not all the rectangles are correct hypotheses of individual vehicles. Typically, a vehicle hypothesis (e.g. the large rectangles in the middle columns) may correspond to more than one vehicle parked in adjacent spots for several reasons. One of these is noise in the elevation estimates and the second is a failure to see the ground between two cars because of sensor viewpoint.

This problem is again solved by the STME system. The general idea is to utilize the fact that normally a vehicle can occupy no more than one parking spot. If we can determine the position and the width of each spot, then we can reason how many cars there are in the rectangles. Recall that the STME system is capable of extracting microstructures in repetitive rectilinear patterns. The parking spot markers (white lines on the ground that separate the spots) can be considered to be such a 2-D microstructure pattern: they are small, but aligned regularly and repetitively. For the STME system that searches for intensity dips, the reversed intensity image serves as the input image from which the white markers are extracted (Fig. 5(b)). Among the 77 markers in the parking lot, 58 of them are correctly identified as long, narrow rectangular boxes. Due to the existence of image noise and occlusions caused by parked vehicles, some markers are missing and some false alarms occur (see Table 1).

A scheme similar to the one used to fill in the missing windows in Fig. 4(c) is employed to fix these errors. First, a *parking spot hypothesis* is defined by two extracted

Table 1: Statistics of the marker extraction

	actual markers	extracted boxes
total	77	59
correctly identified	58	58
false neg. (missing markers)	19	
false pos. (incorrect boxes)		1

neighboring markers, and the width of the spot is calculated by the distance between the two centerlines of the two boxes delineating the parking spot. A clustering algorithm is used to cluster the spot hypotheses according to their widths, resulting in the four classes shown in Table 2. Because most markers have been extracted correctly, the largest class correctly reflects the majority of the parking spots and the average width of this class best reflects the actual spot width in this parking lot. For hypotheses in any of the other classes, they either represent unusual parking spots (with a width considerably different from the average) or areas where lane markers are missing or incorrectly identified. We use a merge/split scheme to deal with these spot hypotheses. A large spot hypothesis is split into m spots where m minimizes

$$r(m) = \left| \frac{w}{m} - w_a \right|, \tag{3}$$

in which w is the width of the spot hypothesis and w_a is the actual spot width. For example, a spot hypothesis of width 50.0 (in Class 3) would be split into 3 spots, each with width 16.7. Similarly, the spot hypotheses in Class 4 and 5 are split into 4 and 6 spots, respectively. The two spot hypotheses of average width 9.0 (in Class 1) are merged into one spot of width 18.0.

Fig. 5(c) shows the final parking spots after splitting and merging incorrect spot

Table 2: Using width to cluster the spot hypotheses

	total number of spot hypotheses	average width (pixels)
Class 1	2	9.0
Class 2	49	17.2
Class 3	1	50.0
Class 4	2	69.5
Class 5	2	102.5

hypotheses. A total of 73 spots have been obtained. It can be seen that the spots correctly reflect the 2-D structure of the parking lot; in particular, individual cars fall correctly into separate spots although some lane markers are absent due to occlusions.

Once the 2-D structure of the parking lot is known, separation of the connected vehicles becomes a simple reasoning process. Rectangular areas in Fig. 5(e) that cover more than one parking spot are segmented into separate vehicles. Of the 12 parked vehicles in the parking lot, 9 rectangular areas are extracted by the bump extraction algorithm from the elevation map, including 3 areas that contain more than one vehicle. With the information of the parking spots, the separation algorithm reorganizes these 3 areas, separating them into small areas in accordance with the spots. The 12 vehicles are finally identified in Fig. 5(f). A visualization of the parking lot with a bounding box for each vehicle is shown in Fig. 6. The height of each bounding box is the average of the elevation data within the box.

In summary, we provide a hybrid method for extracting the repetitive rectilinear structures in the parking lot using the same STME system: 2-D parking spot structures from the intensity map and 3-D vehicle structures from the elevation map. The combination of the results enables us to find the individual vehicle placement in the parking lot.

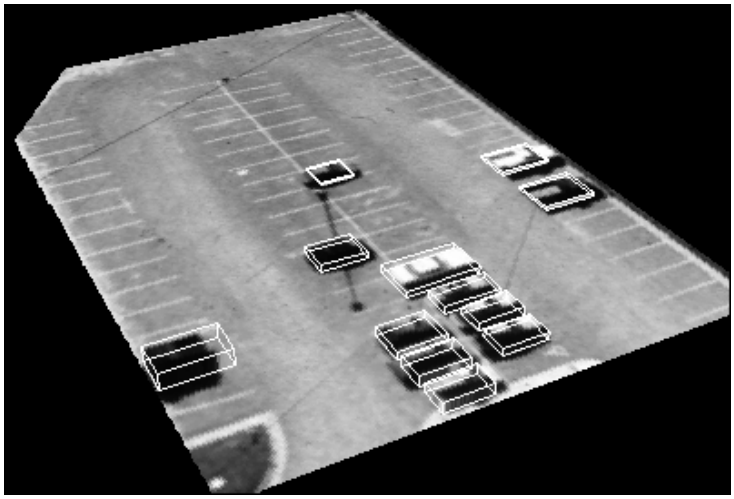


Figure 6: Visualization of the parking lot with individual vehicles extracted

4 Parking Lot Activity Simulation

4.1 Visualization with real textures

Simulation of parking lot activities includes placing new vehicles into the parking lot, animating them and visualizing their activities. These functions are useful in studies of parking lot management and traffic control.

In order to achieve visual realism in these functions, textures from real images are necessary components of the visualization subsystem. Coorg and Teller [7] proposed a technique based on median statistics to extract real textures of object surfaces from multiple images. The median texture algorithm works best in the circumstances that a large number of images are available. Since occlusions are not modeled in this technique, the removel of occlusions is not reliable when only a small number of images are provided. Wang and Hanson [18] proposed another architecture, called Orthographic Facet Image Library (OFIL). The OFIL makes use of multiple images to extract a combined texture map that is the composite of the "best" components of all images. In the process, the system handles occlusions (including self-occlusions) caused by modeled objects (such as

modeled vehicles) in the scene. This technique, however, still requires that every piece of the texture must have been seen from at least one view in the multiple images. The OFIL system cannot be used in the case of images in Fig. 1, because the two images were acquired simultaneously, and hence the parking lot textures occluded by the vehicles are not available.

Our goal is to provide a clean image of an empty parking lot to facilitate visualization and simulation tasks. The method is once again based on the fact that the parking spots present a repetitive pattern in the facet coordinate. It is also reasonable to assume that a parking spot has a similar intensity texture to its neighboring spots. Therefore, after extracting the parked vehicles we can replace the textures of the spots occupied by the vehicles by textures drawn from neighboring empty spots. In practice, a parked vehicle, together with its shadow, often corrupts its adjacent spots as well. Hence, the spots needing a texture replacement actually include both the occupied spots and their direct neighbors.

4.2 Texture replacement criteria

Given a spot S, which is a subimage containing the parking spot whose texture needs to be replaced, how to appropriately choose an uncorrupted, empty spot texture to replace S is an issue that determines the quality of the resulting visual realism. Let $\mathbf{Q}_S$ denote the set of all the spots qualified to replace S, i.e. the set of the spots that are neither occupied by a vehicle nor adjacent to an occupied spot. (For the experiments presented in this paper, $\mathbf{Q}_S$ is restricted to the qualified spots that are in the same column as S.) The problem is formulated as how to find out a spot texture, $f(S)$, to replace S using the information of $\mathbf{Q}_S$.

The simplest way is to use a *least distance criterion (LDC)*, i.e. to replace S by the

texture of the closest qualified spot. That is, $f(S) = Q_S$, in which $Q_S \in \mathbf{Q}_S$ and

$$\text{dist}(Q_S, S) = \min\{\text{dist}(Q, S) \,|\, Q \in \mathbf{Q}_S\}, \tag{4}$$

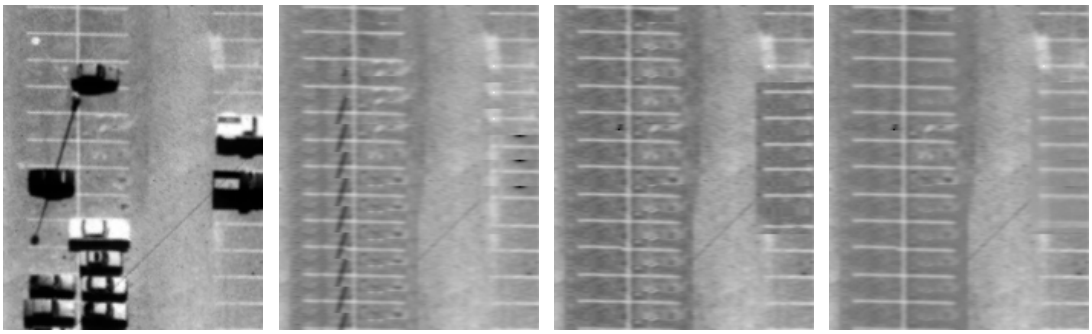
where $\text{dist}(Q, S)$ is the distance between two spots. Fig. 7(b) shows an application of LDC to a portion of the Lockheed/Martin parking lot, shown in (a). For example, all the parking spots on the lower part of the left column are corrupted spots, and are replaced by the fifth spot counting from the top. The advantage of LDC is that it retains intensity similarity, because neighboring spots tend to have similar intensities. However, the spot chosen by LDC might not have a good image quality. Using the example of the left column in Fig. 7(b), since the fifth spot was corrupted by a piece of a shadow of a pole, this corruption is inherited by all the spots replaced and this injects an unrealistic artifact into the image.

Because the texture inside a parking spot usually has a homogeneous intensity distribution, the intensity variance in the texture is a good measure to judge the quality of a spot. This leads to the *least intensity variance criterion (LVC)*, that is, $f(S) = Q_S$, in which $Q_S \in \mathbf{Q}_S$ and

$$\text{var}(Q_S) = \min\{\text{var}(Q) \,|\, Q \in \mathbf{Q}_S\}, \tag{5}$$

where $\text{var}(Q)$ is the intensity variance of the texture Q. Fig. 7(c) shows an application of LVC. While LVC provides much cleaner textures than LDC, it sometimes causes unsatisfactory results in that the selected texture Q_S might be so distant from S that their intensities differ significantly. On the right column in Fig. 7(c), the spots being replaced look very unrealistic for this reason. (In this example, the replacing texture Q_S is taken from a spot beyond Fig. 7(a); it is a spot on the top part of Fig. 3(a).)

We propose a *weighted combination criterion (WCC)* to take into account the factors of both the spot distance and the intensity variance. The texture $f(S)$ that replaces S is



(a) (b) (c) (d)

Figure 7: Repairing the textures of corrupted parking spots

(a) a portion of the orthographic intensity image of the parking lot

(b) repairing the occluded spots using the least distance criterion (LDC)

(c) repairing the occluded spots using the lease variance criterion (LVC)

(d) repairing the occluded spots using the weighted combination criterion (WCC)

a weighted linear combination of all the qualified textures in $\mathbf{Q}_S$:

$$f(S) = \sum_{Q \in \mathbf{Q}_S} G_Q Q, \tag{6}$$

in which

$$G_Q = \frac{G}{[\text{dist}(Q,S)]^\beta [\text{var}(Q)]^\gamma} \tag{7}$$

is a weight defined heuristically, and G is a constant satisfying

$$\sum_{Q \in \mathbf{Q}_S} G_Q = 1. \tag{8}$$

The value of G_Q is affected by both $\text{dist}(Q,S)$ and $\text{var}(Q)$. According to WCC, the weight of a qualified spot Q tends to be high when it is geographically close to S and when it has a low intensity variance. In this way, $f(S)$ takes advantage of both neighboring spots

and high quality spots. The experimental result (Fig. 7(d)) shows that WCC makes satisfactory replacements of corrupted parking spots. In (8) β and γ are constants that can be determined empirically to balance the effects of $\mathrm{dist}(Q, S)$ and $\mathrm{var}(Q)$. In the experiment, they have been set to $\beta = 1.0$ and $\gamma = 4.0$.

Having obtained a clean, empty parking lot, we can easily conduct visualization and simulation of parking lot activities. Fig. 8 is one such scene in which the user places the vehicles randomly onto the parking lot. The textures of the cuboid vehicles are taken from the real image using the OFIL system [18]. It is now possible to view the parking lot from any user selected position since the occlusions of the original vehicles have been removed.

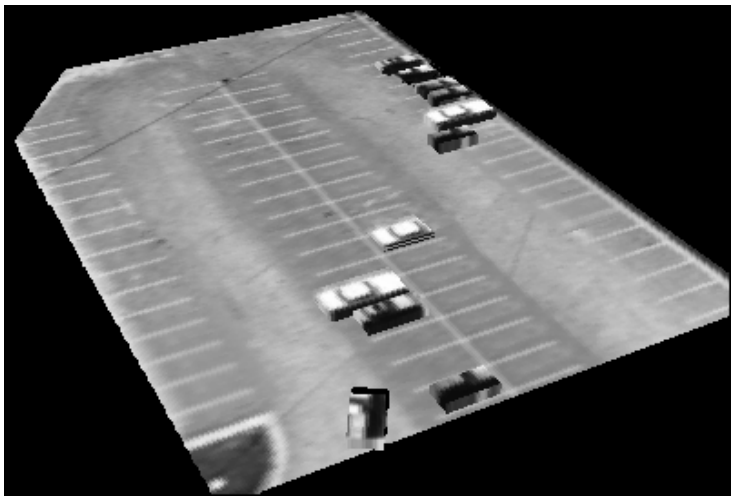


Figure 8: Simulation of parking lot activities using a cleaned ground surface and randomly placed, texture-mapped vehicles

5 Discussion

In this paper we have proposed some new approaches to parking lot analysis from aerial images: (1) the elevation domain provides features that distinguish the vehicles from

the ground; (2) parked vehicles are treated as microstructures rather than modeled as large-scale structures; and (3) textures of corrupted parking spots are repaired by using their repetitive appearance. A stereo algorithm is employed, bringing parking lot analysis into the elevation domain. An STME system is designed to extract a wide variety of microstructures, from 2-D window patterns and parking spot markers to 3-D motor vehicles. A hybrid application of the system to the elevation map and to the intensity map results in a complete extraction of individual vehicles. Finally, a new texture exploitation technique is proposed to generate a clean image of the parking lot without the vehicles. A combination of all the subsystems provides an ability to simulate/visualize parking lot activities with a high quality of visual realism.

The system is particularly useful in parking lot traffic studies. Cohen-Or et al. [3] have proposed a visualization method in which the objects on the ground are represented by voxels. In our system, because the vehicles are structured entities symbolically represented by cuboids instead of by unstructured voxels, the user has a higher degree of flexibility to handle the positions and movements of the vehicles.

The current system works most reliably on sparsely occupied parking lots for the following reasons. First, the stereo algorithm is good at detecting isolated vehicles. Due to the existence of noise and perspective distortion in the original image pair, adjacent cars may lead to a connected bump in the elevation map, which increases the possibility of mistakes in the vehicle extraction process. Second, the presence of more vehicles means a possibility of more occlusions on the ground and fewer visible parking spot markers. Generally, the extraction result is unreliable if too many spot markers are missing. Third, if the parking lot is nearly full then recovery of the ground texture is problematic.

This limitation can be reduced if the current system is combined with Chellappa et al.'s system [1, 2]. While the elevation based approach ignores features of vehicles in the

intensity domain, Chellappa et al.'s system relies on these features. For example, they use shadow as a cue to separate individual cars. In fact, their system would perform better in the case of densely parked area, for these intensity features are more significant in these areas. Therefore, a combination of the two systems would potentially improve the performance over either system in many situations.

The current system correctly identifies "spot-based" vehicles, i.e. cars, trucks, motorcycles, etc., that are parked in accordance with the parking lanes. This represents the majority situation in a parking lot. In real applications, improperly parked vehicles (e.g. a long truck occupying several spots) may take place in some occasions. The elevation of these vehicles can be obtained corretly in the system. However, some other information must be used to decide whether the bumpy area should be divided into individual vehicles. Again, the periodic information used in [2] may be useful in such situations.

Further studies include more aggressive use of multi-image resources. For the same parking lot, some spots might be occupied at one time, while being empty at another time. Consequently, the analysis of repetitive spot patterns could be an off-line process separated from the on-line elevation analysis on vehicles. Once the ground structure is completely known, it is an easy process to separate connected bumps into individual vehicles no matter how densely they are parked. The textures of parking spots can also be collected from different images wherever they are empty, instead of borrowing from neighbors in the same image. This could be done by combining our system with Wang and Hanson's OFIL system [16, 20], which collects textures from multiple images sources.

Acknowledgments

We would like to thank Howard Schultz who provided the elevation data for this paper, and Professor Edward Riseman for his discussions.

References

[1] R. Chellapa, P. Burlina, X. Zhang, Q. Zheng, C. Lin, V. Parameswaran, L. Davis, and A. Rosenfeld, "Site Model Mediated Detection of Movable Object Activities," *Image Understanding Workshop*, pp. 275-304, Palm Springs, CA, 1996.

[2] R. Chellapa, X. Zhang, P. Burlina, C. Lin, Q. Zheng, L. Davis, and A. Rosenfeld, "An Integrated System for Site Model Supported Monitoring of Transportation Activities in Aerial Images," *RADIUS: Image Understanding for Imagery Intelligence,* O. Firschein and T. Strat (Ed.), pp. 285-317, 1996.

[3] D. Cohen-Or, E. Rich, U. Lerner, and V. Shenkar, "A Real-Time Photo-Realistic Visual Flythrough," *IEEE Trans. on Visualization and Computer Graphics*, vol. 2, no. 3, pp. 255-265, 1996.

[4] R. Collins, Y. Cheng, C. Jaynes, F. Stolle, X. Wang, A. Hanson, and E. Riseman, "Site Model Acquisition and Extension from Aerial Images," *Fifth International Conference on Computer Vision*, pp. 888-893, Cambridge, MA, 1995.

[5] R. Collins, C. Jaynes, Y. Cheng, X. Wang, F. Stolle, H. Schultz, A. Hanson, and E. Riseman, "The UMass Ascender System for 3D Site Model Construction," in *RADIUS: Image Understanding for Imagery Intelligence,* Oscar Firschein (Ed.), pp. 209-222, 1996.

[6] R. Collins, C. Jaynes, Y. Cheng, X. Wang, F. Stolle, A. Hanson, and E. Riseman, "The ASCENDER System: Automated Site Modeling from Multiple Aerial Images," to appear in Special Issues of *Computer Vision and Image Understanding (CVIU)* on Building Detection and Reconstruction from Aerial Images, R. Nevatia, A. Gruen (Guest Ed.), 1998.

[7] S. Coorg and S. Teller, "Automatic Extraction of Textured Vertical Facades from Pose Imagery," MIT LCS TR-729, MIT Laboratory for Computer Science, January, 1998.

[8] L. Davis, R. Chellappa, Y. Yacoob, and Q. Zheng, "Visual Surveillance and Monitoring of Human and Vehicular Activitiy," *Image Understanding Workshop*, New Orleans, LA, pp. 19-23, 1997.

[9] Fua, P., "Model-Based Optimization: an Approach to Fast Accurate and Consistent Site Modeling from Imagery," in *RADIUS: Image Understanding for Imagery Intelligence,* O. Firschein and T. Strat (Ed.), pp. 129-152, 1996.

[10] T. Kanade, R. Collins, A. Lipton, P. Anandan, P. Burt, and L. Wixson, "Cooperative Multi-Sensor Video Surveillance," *Image Understanding Workshop*, New Orleans, LA, pp. 3-10, 1997.

[11] S. Noronha and R. Nevatia, "Detection and Description of Buildings from Multiple Aerial Images," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* Puerto Rico, pp. 588-594, June 1997.

[12] V. Ramesh, R. Haralick, A. Bedekar, X. Liu, D. Nadadur, K. Thornton, and X. Zhang, "Computer Vision Performance Characterization," in *RADIUS: Image Understanding for Imagery Intelligence,* O. Firschein and T. Strat (Ed.), pp. 241-282, 1996.

[13] H. Schultz, "Terrain reconstruction from widely separated images," *Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision II*, SPIE Proceedings Vol. 2486, Orlando, FL, pp. 113-123, 1995.

[14] H. Schultz, F. Stolle, X. Wang, E. Riseman, and A. Hanson, "Recent Advances in 3D Reconstruction techniques Using Aerial Images," *Image Understanding Workshop*, New Orleans, LA, pp. 977-982, 1997.

[15] C. Slama (Ed.), *Manual of Photogrammetry,* American Society of Photogrammetry, Falls Church, VA, 1980.

[16] X. Wang, R. Collins, and A. Hanson, "An Orthographic Facet Image Library for Supporting Site Model Refinement and Visualization," Technical Report #95-100, Dept. of Computer Science, Univ. of Massachusetts at Amherst, November 1995.

[17] X. Wang and A. Hanson, "A System for Surface Texture and Microstructure Extraction from Multiple Aerial Images," Technical Report #97-022, Dept. of Computer Science, Univ. of Massachusetts at Amherst, April 1997.

[18] X. Wang and A. Hanson, "Extracting Surface Textures and Microstructures from Multiple Aerial Images," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* pp. 301-306, Puerto Rico, June 1997.

[19] X. Wang, A. Hanson, R. Collins, and J. Dehart, "Surface Microstructure Extraction from Multiple Aerial Images," in *Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision III*, SPIE Proceedings Vol. 3072, Orlando, FL, 1997.

[20] X. Wang, J. Lim, R. Collins, and A. Hanson, "Automated Texture Extraction from Multiple Images to Support Site Model Refinement and Visualization," *The Fourth Int. Conf. in Central Europe on Computer Graphics and Visualization 96*, pp. 399-408, Plzen, Czech Republic, 1996.

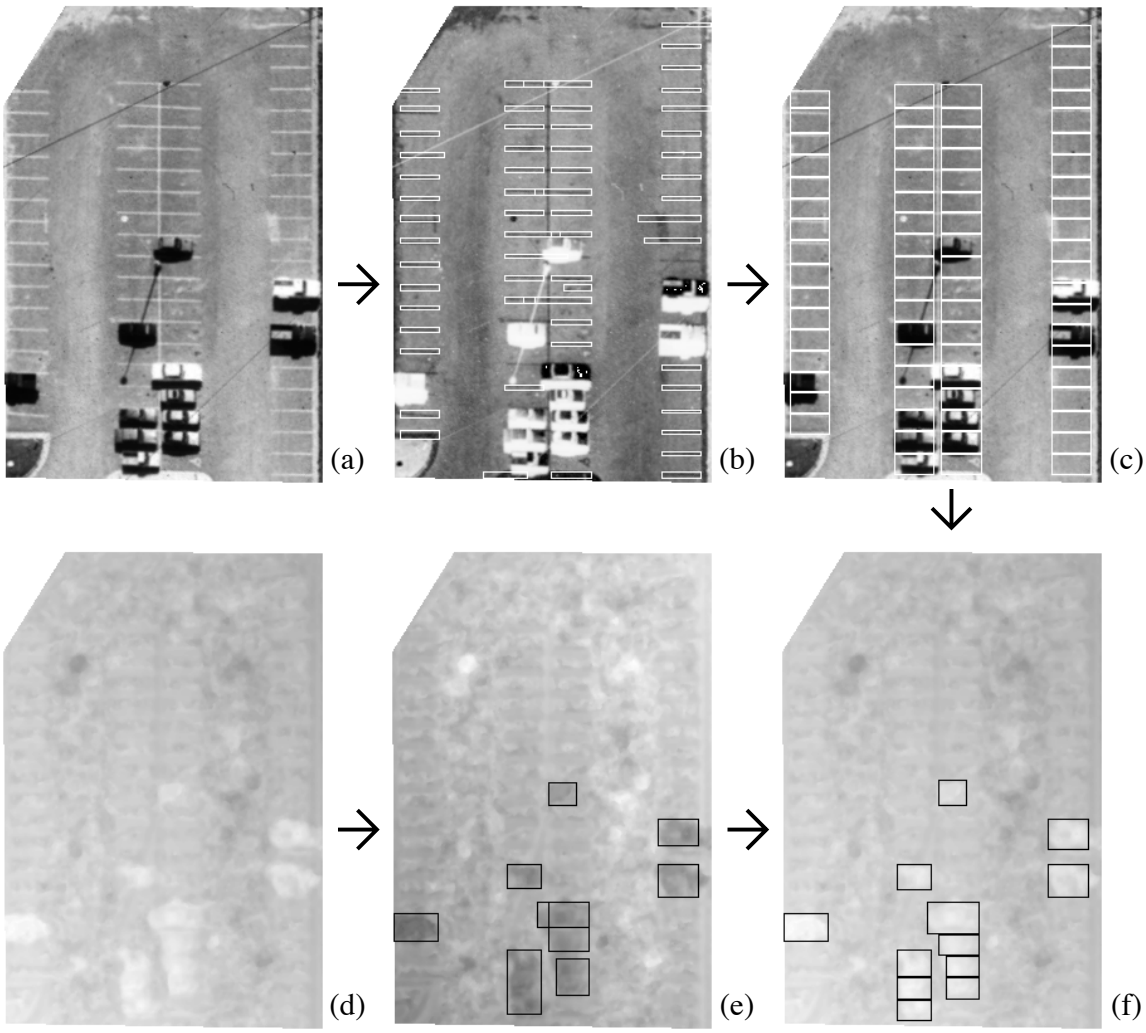


Figure 5: Parking lot analysis using the STME system

(a) orthographic intensity image of the parking lot

(b) extraction of spot markers as 2-D microstructures from the reversed intensity image

(c) the hypothesized parking lot structure based on extracted spot markers

(d) elevation map of the parking lot

(e) extraction of vehicles as 3-D microstructures from the reversed elevation map

(f) individual car identification by symbolic reasoning from (c) and (e)