Vehicle Detection and Orientation Estimation Using the Radon Transform

Rengarajan Pelapur[†], Filiz Bunyak[†], Kannappan Palaniappan[†],
Gunasekaran Seetharaman[‡]

[†] Dept. of Computer Science, University of Missouri, Columbia, Missouri
[‡] Air Force Research Laboratory, Rome, NY, USA

ABSTRACT

Determining the location and orientation of vehicles in satellite and airborne imagery is a challenging task given the density of cars and other vehicles as well as the complexity of the environment in urban regions around the world. We describe an accurate and flexible method for detecting vehicles using a template-based directional chamfer matching approach, and vehicle orientation estimation using a refined segmentation followed by a novel Radon transform based profile variance peak detection. The same algorithm was applied to both high resolution satellite imagery and wide area aerial imagery and initial results show robustness to varying illumination and geometric appearance distortions. Nearly 80% of the orientation angle estimates for 1585 vehicles across both satellite and aerial imagery were accurate to within 15° of the ground truth. In the case of satellite imagery alone, nearly 90% of the objects had an estimated error within ±1.0° of the ground truth.

Keywords: wide area motion imagery, orientation estimation, vehicle detection, intelligent transportation systems, traffic analysis, airborne surveillance

1. INTRODUCTION

Recognizing the pose of an object provides critical information for semantic scene understanding and improves object recognition,[1] spatial localization and object tracking.[2–4] Detecting the position and orientation of vehicles in satellite and aerial imagery is an important step in a number of applications such as emergency response, law enforcement, forensics, traffic analysis, intelligent transportation systems, urban planning, surveillance, improving web-based map imagery and vehicle tracking. Various commercial earth observation remote sensing satellites such as Worldview, Quickbird, GeoEye and Ikonos now routinely provide access to high resolution imagery captured with a ground sampling distance (GSD) resolution at the one meter to sub-meter level typically 0.5 m or 0.6 m panchromatic. Web-based geospatial imagery services providing online interactive access to maps, such as GoogleEarth and OpenStreetMap, combine satellite imagery with aerial (orthophoto) imagery acquired at even higher spatial resolutions (*i.e.* 0.15 m GSD). It is thus, beneficial to analyze such imagery for developing automated systems which can recognize objects in the scene including vehicles. Detection of objects of interest in high resolution aerial and satellite imagery pose a challenging problem due to various factors such as complex backgrounds, dense distribution of objects like cars in a parking lot or public event, and clutter or background objects with similar appearance which results in a large number of false alarms.

Change detection has always been an important task in aerial imagery and usually a lot of dynamic objects such as vehicles need to be excluded to produce reliable scene change detection results.[5,6] Availability of higher resolution imagery enables the development of traffic monitoring systems as part of an intelligent transportation system with adaptive routing based on traffic patterns, incident event detection, delivery of emergency services and reduced response times for law enforcement. Incident monitoring systems use orientation and velocity estimates as the main parameters to detect events.[7] Road extraction and vehicle detection in aerial imagery is also a well established approach in computer vision and some techniques perform road extraction before vehicle detection[8] or use prior map information in street and road databases. Vehicle detection has been studied as a template matching problem, as a learning problem that uses a very high dimensional feature space,[9] or as an online learning problem.[10] Different approaches have been used to make object detection more robust to variations in pose, occlusion and illumination.[11,12] Several such methods stress the need for specialized techniques to tackle the problem of different appearances for the same objects.[13] Pose estimation is thus an

important task in object recognition and a technique for orientation estimation can help by providing additional information for more robust object detection in visual scene analysis.

Explicit orientation estimation is a useful capability to explore since the majority of feature descriptors and car detection schemes are designed to be rotationally invariant. Additionally, automatic orientation estimation can speed up manual ground-truth collection that is often necessary for validation and performance analysis as well as to improve the quality of manual tracks. The typical method of manual tracking of vehicles involves marking only centroids since this is the simplest, or drawing a precise polygon boundary around the vehicle in each frame which is very time consuming. A good compromise is to use an oriented bounding box which takes vehicle geometry and pose into account and moves an axis-aligned bounding box drawn around the object of interest between frames. Vehicle detection and orientation estimation can speed up this manual ground truthing process as well as provide pose information to improve automatic vehicle tracking algorithms.[2, 14–16]

In this paper we describe an initial implementation of an overall approach to detect the location and orientation of vehicles in satellite and aerial imagery. The system consists of several modules that can be broadly divided into car detection, feature extraction, segmentation and orientation estimation. Our contributions are mainly in the use of the Hessian matrix for detecting prominent edges while decreasing the influence of background and lighting changes that may cause false detections, and in adapting the Radon transform for accurately determining orientations using the variance of the profile distributions. In this article only the evaluation of the vehicle orientation estimation is presented as this is the focus; the accuracy of the vehicle detection step will be presented in detail elsewhere. The segmentation module is also important as the car detection (or a tracking algorithm in the case of moving vehicles) can usually provide only an approximate localization of the object that needs refinement. A more accurate extraction of the vehicle-blob region using a simple segmentation process contributes significantly to the accuracy of detection and hence in reducing the angular error in the orientation estimate. Our performance evaluation used two types of imagery including satellite imagery from GoogleEarth and multi-camera wide area aerial imagery from Persistent Surveillance Systems.[17]

2. VEHICLE DETECTION IN AERIAL IMAGERY

Vehicle detection has been extensively studied in computer vision.[8–10, 18, 19] Our approach in this paper is based on using a set of rotated templates with directional chamfer matching for detecting locations in the image with car-like objects that can be further verified and analyzed for additional orientation information. The proposed approach is suitable for finding vehicles in aerial and satellite imagery and can be adapted for use in tracking vehicles in wide area motion imagery (WAMI).

2.1 Directional Chamfer Matching (DCM)

We first want to identify candidate positions in the image which contain car-like objects. Since the gray level appearance intensity distribution of car objects can change significantly with viewpoint, we instead use the distribution of edges within the car object to guide the search. Vehicle detection based on chamfer matching uses shape information provided by a number of templates of the same object in different orientations to produce a set of best possible alignments. chamfer matching is a popular method to find the best alignment between two edge maps; the smaller the matching distance the better the alignment. Standard chamfer matching, however, has difficulty in dealing with objects of different sizes and varying appearance. Directional chamfer matching (DCM) overcomes some of these limitations in a computationally efficient manner.[20]

Let $I(p)$ be the query image and $T(q)$ the template, with $E_I(p)$ and $E_T(q)$ being the respective edge maps, where p and q are the set of pixels in the image in lexicographic order. The chamfer distance between the two sets of edge maps, $E_I(p)$ and $E_T(q)$ is given by,[20]

$$d_C(E_I(p), E_T(q)) = \frac{1}{n} \sum_{\mathbf{e}_p \in E_I} \min_{\mathbf{e}_q \in E_T} |\mathbf{e}_p - \mathbf{e}_q| \tag{1}$$

where $\mathbf{e}_p$ and $\mathbf{e}_q$ belong to the two edge map sets E_I and E_T respectively. Directional chamfer matching (DCM) is a more reliable way to match two sets of edge maps in the presence of clutter and geometric distortions. DCM

jointly minimizes location and orientation errors between two edge sets. The DCM score is defined as,[20]

$$d_C(E_I(p), E_T(q)) = \frac{1}{n} \sum_{\mathbf{e}_p \in E_I} \min_{\mathbf{e}_q \in E_T} |\mathbf{e}_p - \mathbf{e}_q| + \lambda \min_{\mathbf{e}_q \in E_T} (|\phi(\mathbf{e}_p) - \phi(\mathbf{e}_q)|, 180 - |\phi(\mathbf{e}_p) - \phi(\mathbf{e}_q)|) \quad (2)$$

where ϕ is the directional term computed modulo π and λ is a weighting factor that determines the influence of the location and orientation error terms. The orientation error is computed as the minimum circular difference. Minimizing the directional chamfer distance in Eq. 2 gives the best alignment between the two edge sets. We use the fast sublinear time implementation for search optimization to find the best DCM match scores.[20]

The DCM approach uses a collection of template edge maps (at different orientations), for a selected set of model car templates, and allows us to find potential vehicle locations in the image. Directional chamfer matching has the highest detection rates compared to standard and oriented chamfer matching[21] on the ETHZ shape class dataset[22] and used five object classes with about 255 images in the evaluation reported by Liu *et al.*[20] In our tests we use several large high resolution images with about 1585 vehicle objects; we selected just 5 templates as training data in which each template is oriented from 0 to 179 degrees with a step size of 5 degrees or 36 orientations for a single template. This creates a collection of 180 car templates used in the DCM step for vehicle detection.

DCM-based vehicle detection tends to have high precision but low to moderate recall as shown in Figure 1. The result of DCM using Eq. 2 is a collection of potential vehicle centroids/locations which we denote as, (c_x, c_y). At each of these locations and a region of interest (ROI) around it, ridge information is extracted using the Hessian operator as described in the next section in order to do further filtering on detections and estimate vehicle orientation.

2.2 Multiscale Hessian-based Line Segment Feature Extraction

Color, intensity or texture are often not distinctive enough features in aerial and satellite images of cars and vehicles, hence we rely more on shape features for more accurately localizing car-like objects in imagery and determining the oriented bounding box around these objects. The distribution of vehicle line segment contours is used for both refining the object localization (produced by DCM) and orientation estimation. Both of these processes are sensitive to reliable detection of linear features so we use a multi-scale linear feature extraction approach. A linear feature (vessel) enhancement filter based on an eigenvalue analysis of the Hessian matrix across scale space has been used in a range of biological applications including vessel and cell segmentation[23–27] and can be adapted for vehicle segmentation. The Hessian matrix is defined using second-order image derivatives,

$$H(I) = \sigma_s^2 \begin{bmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{bmatrix} \quad (3)$$

where σ_s^2 is used to achieve invariance under image rescaling.[28, 29] Let λ_1, λ_2 be the eigenvalues and $\mathbf{e}_1$, $\mathbf{e}_2$ the corresponding eigenvectors of the Hessian matrix with $|\lambda_1| \geq |\lambda_2|$. The Hessian is computed after convolving the image with a 2D isotropic Gaussian smoothing filter,

$$G(x, y, \sigma_s) = (1/2\pi\sigma_s^2)exp(-(x^2 + y^2)/(2\sigma_s^2)) \quad (4)$$

where σ_s is the standard deviation of the Gaussian distribution for scale s. We use a suitable range of scales based on expected target size and define a mixed Frobenius (custom) norm of the Hessian matrices across scale with the maximum response as a measure of the second order structureness,

$$I_H(x, y) = ||H||_F = \left((\max_s \lambda_1(x, y, \sigma_s))^2 + (\max_s \lambda_2(x, y, \sigma_s))^2 \right)^{\frac{1}{2}} \quad (5)$$

Given the centroid and a region around it, based on the DCM step, and the second order structureness using Eq. 5, we extract the largest local connected component in the image by running a standard two pass connected component labeling algorithm.[30] We then find the best fitting ellipse, denoted as Ω_E given the centroid of the point set extracted in I_H in order to estimate an accurate bounding box. An iterative ellipse fitting approach is
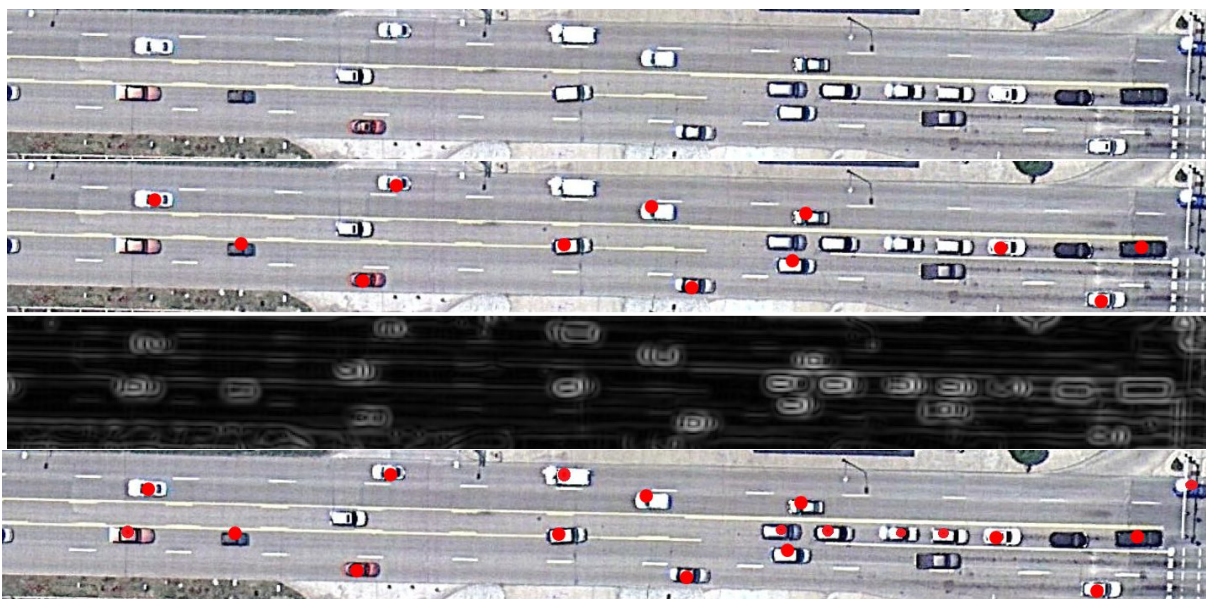
Figure 1: The top figure shows a small ROI from a GoogleEarth satellite image and the second figure shows the vehicle detection results using DCM based on Canny edge maps.[20] The DCM step produces accurate detections (red dots) with a low number of false alarms, that is high precision; however, the recall rate is only 57% (12 out of 21 cars) in this small test example. On the other hand we see that the multiscale Hessian, shown in the third figure, produces good (blob and ridge) responses for almost all of the vehicles and suggests an approach for improving the performance of the DCM. This is confirmed in the fourth (bottom) figure which shows DCM-based vehicle detections, but using the multiscale Hessian as input instead of Canny edge maps, with produces a higher recall rate of 90%. Note that in the experimental results we used Canny edge maps as input to DCM.

used with $x_E = c_x + r_1 \cos \alpha$ and $y_E = c_y + r_2 \sin \alpha$, r_1 and r_2 are incremented until the radii reach the edge of the largest connected component for a fixed set of angles α. The parameters covering the largest number of multiscale Hessian responses is used. Note that there may be multiple car-like objects within the ROI, due to closely spaced vehicles, and we identify the largest object not already extracted within this ROI.

2.3 Orientation Estimation

After determining probable locations of vehicles several filters are applied to reduce false positives by validating local information using a segmentation process and fitting a bounding box. The orientation of the bounding box that locates the vehicle is estimated using a Radon transform based approach that is outlined in this section. In Kembhavi et al.[9] since the orientation is unknown, they propose a strategy where the templates are coarsely rotated in increments of 30 degrees and then finely refined towards a more accurate orientation estimate by using smaller angular steps. Our proposed method for orientation estimation could be incorporate to speed up such an algorithm. Orientation information can also help in traffic analysis and a number of other applications such as occupancy and direction of travel in a given area such as a shopping mall parking lot.

The Radon transform is a set of line integrals that computes the projection sum of image intensities along a set of parallel lines oriented at different angles. The object's principal direction can be defined as the direction along which there are more straight lines thus the cumulative sum along the direction which is perpendicular to the principal direction will have the maximum variance. Finding the variance thus, will give us the principle direction. Several methods such as taking a second order derivative as in Khouzani et al.[31] has been known to disambiguate between multiple local maximas in case of textures. Similar methods can be applied to detect orientation of objects by applying the Radon transform of the image intensities, $f(x, y)$, within the region of interest. Defining a set of parallel image rays at angle θ, as

$$\rho = x \cos \theta + y \sin \theta \tag{6}$$

then the Radon transform of the thresholded feature map is defined as,

$$R(\rho, \theta) = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} f(x,y)\delta(\rho - x\cos\theta - y\sin\theta)dx \ dy. \tag{7}$$

where $\delta(\cdot)$ is the Dirac delta sampling function. Given the projections we can now calculate the variance of each profile in the Radon transform of the binarized Hessian second structureness feature map as,

$$V(\theta) = Var(R(\rho, \theta)) \tag{8}$$

where $f(x,y)$ is the image that is derived from the improved segmentation and refined estimate of the probable location of the car. The global peak in the variance profile graph, $V(\theta)$, then gives an estimate of the orientation of the vehicle. A number of parameters affects the overall performance of the Radon transform based orientation estimate including the accuracy of the initial centroid estimate (i.e. ROI), type of vehicle features extracted, the binarization threshold, the sampling of the Radon transform, the peak detection method and the density of neighboring foreground objects within the transform domain.

3. DATASETS AND EXPERIMENTAL RESULTS

For benchmarking the orientation estimation we have used several 1500×1500 pixel images from GoogleEarth as shown in Figure 4 and wide area aerial imagery of Philadelphia, PA from PSS.[17] All the images have manually marked vehicle aligned polygons as ground truth. We ran the tests on a set of 1585 objects which includes cars from the GoogleEarth imagery and the more challenging Philadelphia dataset. The challenges in the wide-area dataset include (non-rigid) appearance distortions and image stitching errors. We obtain a rough localization of the objects given by the DCM module as described in Section 2.1, then we process these probable locations with car-like objects through a segmentation module to more accurately localize the object using connected component labeling and ellipse fitting, and then these refined centroids are supplied to the orientation estimation module using the Radon transform approach.

The evaluation methodology consists of comparing each object's estimated bounding box (orientation estimate) with the manually produced ground truth where each vehicle has been marked by a bounding rectangle. This introduces a 180° ambiguity, since the car may be rotated 180° within the rectangle, which we ignore, since our goal is to primarily ensure that the car is contained within the rectangle (aligned with the bounding box). In this case we consider 0° and 180° as equivalent in terms of alignment orientation, so that the error for the case where ground truth is pointing up but the estimated orientation is pointing down is still zero degrees. This is a circular orientation error, and we are measuring the minimum angular difference between two (oriented) rectangles; that is the minimum rotation needed to align two arbitrary rectangles (of approximately the same scale). This is equivalent to measuring the acute angle between two intersecting line segments. However, we need to account for the change in coordinate systems; that is the ground truth database measures angles between $-90°$ and $+90°$ with the horizontal or x-axis being zero degrees, whereas, the Radon transform computation returns an orientation angle for the bounding box or medial line segment with the vertical or y-axis being zero degrees. Algorithm 1 describes the details of the angular error computation. The maximum angular error will be less than 90° for two line segments or rectangles.

Algorithm 1 Angular Error Between Two Rectangles (Acute Angle Between Intersecting Lines)

Input: Ground truth angle in the DB, $\theta_{DB} \in [-90, +90]$ //Zero deg is horizontal
         Estimated angle from the image, $\theta_{EST} \in [0, 180]$ //Zero deg is vertical
Output: Absolute angular error in degrees, $\theta_{Error} \in [0, 90]$ //Maximum error will be 90°
    $\theta_{GT} = \theta_{DB} + 90$ //Zero deg is vertical
    $\theta_{Error} = \min\{|\theta_{EST} - \theta_{GT}|, 180 - |\theta_{EST} - \theta_{GT}|\}$

The peak in the profile variance plot gives an estimate of the orientation of the vehicle as shown in twelve examples in Figure 2. The angles in the profile variance plot are to be interpreted as being measured with

respect to the vertical axis (zero direction). The 180° ambiguity in terms of the forward direction of the car is not treated as an error as we are only interested in vehicle alignment and an accurate oriented bounding box. All the cars in the second row, for example, illustrate this ambiguity. The forward direction of the car could be estimated from other information such as the velocity vector during tracking. The results show that the Radon transform based approach can result in accurate orientation estimates for different vehicle appearances as well as partial occlusions (see Figure 2 third row, third column).
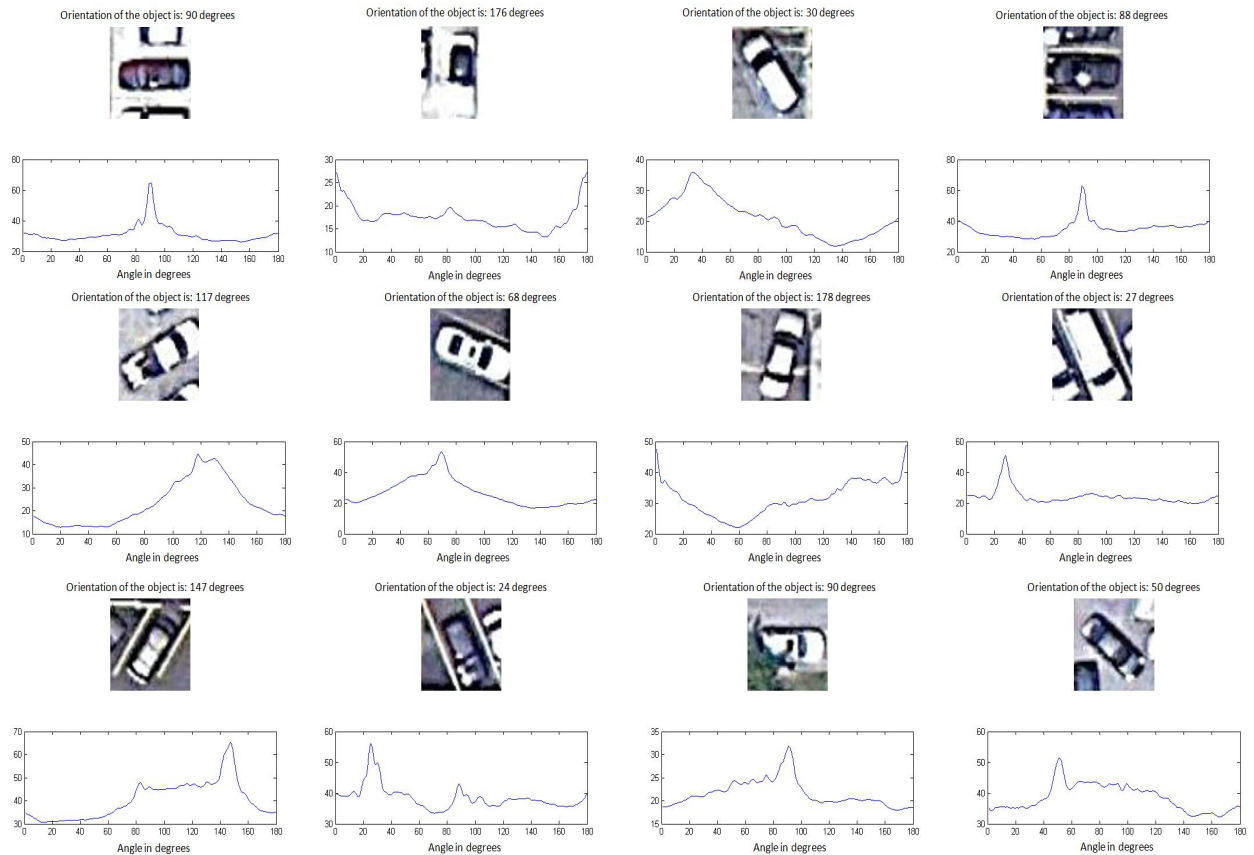

Figure 2: Results for a small set of randomly selected vehicles showing the object template above and the Radon transform profile variances plotted below the template. The peaks correspond to the estimated orientation of the car. The angles are to be interpreted as being measured with respect to the vertical axis (zero direction) and the 180° ambiguity does not contribute to the angular error. All the cars in the second row, for example, illustrate this ambiguity. The ground truth for the vehicle orientations are 90°, 176°, 30°, 88°, for the first row; 117°, 68°, 178°, 27° for the second row; and 147°, 24°, 90°, 50° for the last row.

Figure 3 shows the histogram plots. The first graph shows the distribution of vehicle orientations in the ground truth database and the estimated results using ten degree bins. The orientation estimates across the full range of vehicle orientations is quite accurate. This is quantified in the second graph in Figure 3 which shows the range of orientation angle errors. The orientation estimate for 650 vehicles out of the set of 1585 (about 41% of the dataset) is accurate to within ±5° and another 600 vehicles is accurate to within ±15° as seen in the plot in Figure 3. So nearly 80% of the vehicle orientations are estimated to within 15° of the ground truth which is quite promising for pose estimation and tracking vehicles through turns. When we looked at only the satellite imagery results almost 89% of the objects have an estimation error within ±1.0° of the ground truth (data not shown). Figures 5 to 8 show detailed results for four different cars at four different orientations. Each figure shows an image chip (zoomed up), the corresponding multiscale Hessian feature detector output, the associated

binarized Hessian, Radon transform profile graphs at selected angles to show the variance in the profiles at three selected angles including the maximum and minimum variance profiles, the 2D Radon transform, the variance of the Radon transform profiles and the second derivative of the variance which is sometimes used to detect the maximum in the variance profile.
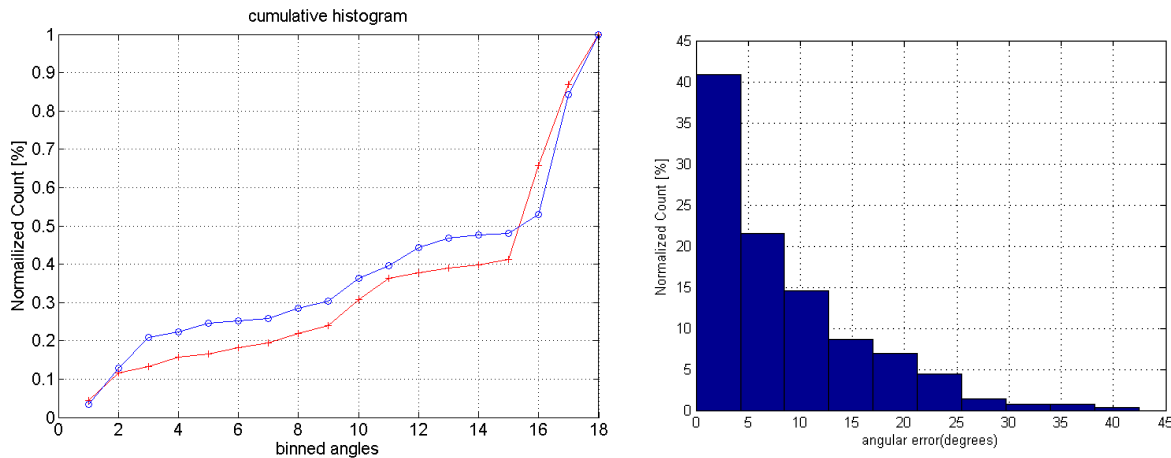

Figure 3: The first plot shows the cumulative histogram of the distribution of angles for 1585 vehicles with ground-truth orientations in blue and the estimated orientations in red using binned angles. The figure on the right is a histogram plot of the absolute angular error measured using Algorithm 1 for all vehicles (x-axis is absolute angular error in degrees and y-axis is the number of objects).

4. CONCLUSIONS

Pose estimation is an important task in many computer vision tasks and the proposed modular approach can be incorporated as part of large scene analysis systems. In this paper, we have shown that reliable vehicle detection and orientation estimation is possible, even when the scene is cluttered, the appearance of the vehicles highly variable, with partial occlusions or the scale of the objects is unknown. The same algorithm was applied to both high resolution satellite imagery and wide area aerial imagery to determine the pose of vehicles by using an approach based on directional chamfer matching and Radon transform profile variances that is robust to illumination changes and geometric appearance distortions. Nearly 80% of the orientation angles for 1585 vehicles were accurately estimated to within 15° of the ground truth. The tracking system LoFT[2, 14, 32, 33] includes the orientation estimation technique described in this paper and leads to improved tracking performance. The proposed approach can also be used within content-based image retrieval systems for geospatial applications.[34] Future improvements in the car detection module using better feature operators and using contextual cues and scene layout models are being explored.

ACKNOWLEDGMENTS

REFERENCES

[1] Felzenszwalb, P., Girshick, R., McAllester, D., and Ramanan, D., "Object detection with discriminatively trained part-based models," *IEEE Trans. Pattern Analysis and Machine Intelligence* **32**(9), 1627–1645 (2010).
[2] Pelapur, R., Palaniappan, K., and Seetharaman, G., "Robust orientation and appearance adaptation for wide-area large format video object tracking," in [*9th IEEE Int. Conf. Advanced Video and Signal-Based Surveillance*], (2012).

Figure 4: Areas of interest from GoogleEarth© imagery used for the experiments. The first row shows Columbia, MO and the second row shows a small region of New York City, NY. The last row shows wide area imagery from Philadelphia, PA acquired by PSS (Persistent Surveillance Systems).[17]

[3] Bunyak, F., Palaniappan, K., Nath, S. K., and Seetharaman, G., "Flux tensor constrained geodesic active contours with sensor fusion for persistent object tracking," *J. Multimedia* **2**, 20–33 (August 2007).

[4] Bunyak, F., Palaniappan, K., Nath, S. K., and Seetharaman, G., "Geodesic active contour based fusion of visible and infrared video for persistent object tracking," in [*8th IEEE Workshop Applications of Computer Vision (WACV 2007)*], Online (Feb. 2007).

[5] Blaschke, T., "Object based image analysis for remote sensing," *ISPRS Journal of Photogrammetry and Remote Sensing* **65**(1), 2–16 (2010).

[6] Crispell, D., Mundy, J., and Taubin, G., "A variable-resolution probabilistic three-dimensional model for change detection," *IEEE Trans. Geoscience and Remote Sensing* (99), 1–12 (2011).

[7] Kahaki, S., Nordin, M., and Ashtari, A., "Incident detection algorithm based on radon transform using high-resolution remote sensing imagery," in [*Int. Conf. Electrical Engineering and Informatics*], 1–5 (2011).

[8] Pacher, G., Kluckner, S., and Bischof, H., "An improved car detection using street layer extraction," in [*Proceedings Computer Vision Winter Workshop (SPRS)*], 1–8 (2008).

[9] Kembhavi, A., Harwood, D., and Davis, L., "Vehicle detection using partial least squares," *IEEE Trans. Pattern Analysis and Machine Intelligence* **33**, 1250–1265 (June 2011).

[10] Grabner, H., Nguyen, T. T., Gruber, B., and Bischof, H., "On-line boosting-based car detection from aerial images," *ISPRS Journal of Photogrammetry and Remote Sensing* **63**(3), 382 – 396 (2008).

[11] Huang, C., Ai, H., Li, Y., and Lao, S., "High-performance rotation invariant multiview face detection," *IEEE Trans. Pattern Analysis and Machine Intelligence* **29**(4), 671–686 (2007).

[12] Wu, B. and Nevatia, R., "Detection and tracking of multiple, partially occluded humans by bayesian combination of edgelet based part detectors," *International Journal of Computer Vision* **75**(2), 247–266 (2007).

[13] Liebelt, J. and Schmid, C., "Multi-view object class detection with a 3D geometric model," in [*Proc. IEEE Conf. Computer Vision and Pattern Recognition*], 1688–1695 (2010).

[14] Pelapur, R., Candemir, S., Bunyak, F., Poostchi, M., Seetharaman, G., and Palaniappan, K., "Persistent target tracking using likelihood fusion in wide-area and full motion video sequences," in [*15th Int. Conf. Information Fusion*], 2420–2427 (2012).

[15] Ersoy, I., Palaniappan, K., and Seetharaman, G., "Visual tracking with robust target localization," in [*IEEE Int. Conf. Image Processing*], (2012).

[16] Ersoy, I., Palaniappan, K., Seetharaman, G., and Rao, R., "Interactive tracking for persistent wide-area surveillance," in [*Proc. SPIE Conf. Geospatial InfoFusion II (Defense, Security and Sensing: Sensor Data and Information Exploitation)*], **8396** (2012).

[17] Palaniappan, K., Rao, R., and Seetharaman, G., "Wide-area persistent airborne video: Architecture and challenges," *Distributed Video Sensor Networks* , 349–371 (2011).

[18] Kluckner, S., Pacher, G., Grabner, H., Bischof, H., and Bauer, J., "A 3D teacher for car detection in aerial images," in [*Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*], 1–8, IEEE (2007).

[19] Zhao, T. and Nevatia, R., "Car detection in low resolution aerial images," *Image and Vision Computing* **21**(8), 693–703 (2003).

[20] Liu, M., Tuzel, O., Veeraraghavan, A., and Chellappa, R., "Fast directional chamfer matching," in [*IEEE Conf. Computer Vision and Pattern Recognition*], 1696–1703 (2010).

[21] Shotton, J., Blake, A., and Cipolla, R., "Contour-based learning for object detection," in [*IEEE Int. Conf. Computer Vision*], 503–510 (2005).

[22] Ferrari, V., Tuytelaars, T., and Van Gool, L., "Object detection by contour segment networks," in [*Proc. European Conference on Computer Vision*], 14–28 (2006).

[23] Frangi, A., Niessen, W., Vincken, K., and Viergever, M., "Multiscale vessel enhancement filtering," *Medical Image Computing and Computer-Assisted Intervention—MICCAI'98* , 130–137 (1998).

[24] Ersoy, I., Bunyak, F., Mackey, M., and Palaniappan, K., "Cell segmentation using Hessian-based detection and contour evolution with directional derivatives," in [*IEEE Int. Conf. Image Processing*], 1804–1807 (2008).

[25] Bunyak, F., Palaniappan, K., Glinskii, O., Glinskii, V., Glinsky, V., and Huxley, V., "Epifluorescence-based quantitative microvasculature remodeling using geodesic level-sets and shape-based evolution," in [*30th Int. IEEE Engineering in Medicine and Biology Society Conf. (EMBC)*], 3134–3137 (Aug. 2008).

[26] Bunyak, F., Palaniappan, K., Chagin, V., and Cardoso, C., "Cell segmentation in time-lapse fluorescence microscopy with temporally varying sub-cellular fusion protein patterns," in [*31st Int. IEEE Engineering in Medicine and Biology Society Conf. (EMBC)*], 1–5 (Sep. 2009).

[27] Prasath, V., Haddad, O., Bunyak, F., Glinskii, O., Glinsky, V., Huxley, V., and Palaniappan, K., "Robust filtering based segmentation and analysis of dura mater vessel laminae using epiflourescence microscopy," in [*35th Int. IEEE Engineering in Medicine and Biology Society Conf. (EMBC)*], (2013).

[28] Lindeberg, T. and Fagerström, D., "Scale-space with casual time direction," in [*Proc. European Conf. Computer Vision*], 229–240 (1996).

[29] Pock, T., Beichel, R., and Bischof, H., "A novel robust tube detection filter for 3D centerline extraction," in [*Proc. 14th Scandinavian Conf. on Image Analysis (SCIA)*], 481–490 (2005).

[30] Dillencourt, M. B., Samet, H., and Tamminen, M., "A general approach to connected-component labeling for arbitrary image representations," *J. ACM* **39**, 253–280 (April 1992).

[31] Jafari-Khouzani, K. and Soltanian-Zadeh, H., "Radon transform orientation estimation for rotation invariant texture analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence* **27**(6), 1004–1008 (2005).

[32] Haridas, A., Pelapur, R., Fraser, J., Bunyak, F., and Palaniappan, K., "Visualization of automated and manual trajectories in wide-area motion imagery," in [*15th Int. Conf. Information Visualization*], 288–293 (2011).

[33] Palaniappan, K., Bunyak, F., Kumar, P., Ersoy, I., Jaeger, S., Ganguli, K., Haridas, A., Fraser, J., Rao, R., and Seetharaman, G., "Efficient feature extraction and likelihood fusion for vehicle tracking in low frame rate airborne video," in [*13th Int. Conf. Information Fusion*], (2010).

[34] Shyu, C. R., Scott, G., Klaric, M., Davis, C. H., and Palaniappan, K., "Automatic object extraction from full differential morphological profile in urban imagery for efficient object indexing and retrievals," in [*3rd Int. Symposium on Remote Sensing and Data Fusion Over Urban Areas (URBAN 2005)*], **36**(8) (2005).
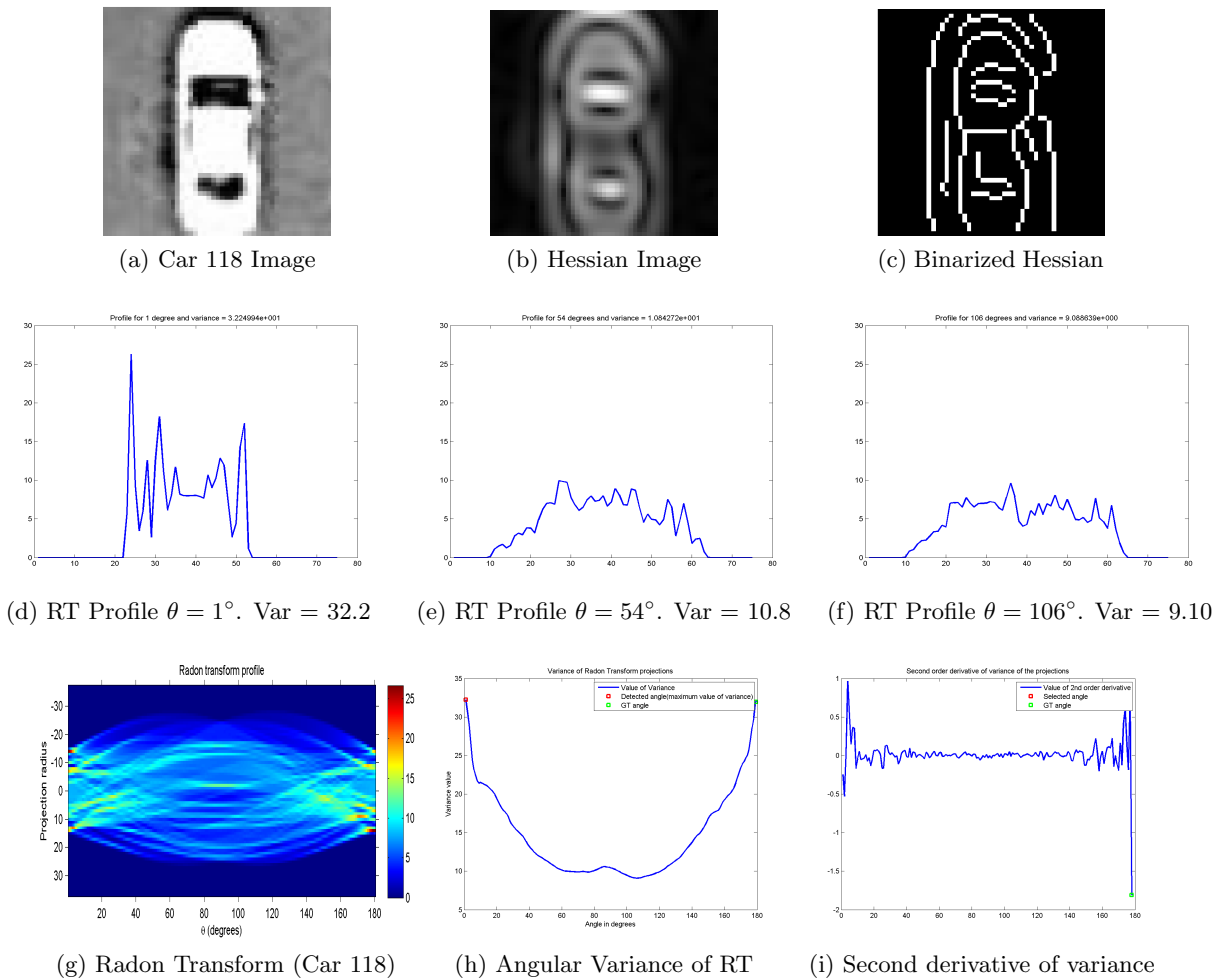
(a) Car 118 Image

(b) Hessian Image

(c) Binarized Hessian

(d) RT Profile $\theta = 1°$. Var = 32.2

(e) RT Profile $\theta = 54°$. Var = 10.8

(f) RT Profile $\theta = 106°$. Var = 9.10

(g) Radon Transform (Car 118)

(h) Angular Variance of RT

(i) Second derivative of variance

Figure 5: Results for white Car 118 at orientation angle $\theta_{DB} = 90°$, (equivalent to $\theta_{GT} = 180°$ or $0°$ wrt vertical). (a) Image chip showing Car 118. (b) Result of the multiscale Hessian feature detector. (c) Binarized detection image based on thresholding the Hessian. Graphs (d), (e) and (f) show the Radon transform profiles (vertical cross section through the Radon transform shown in (g)) at three representative projection directions including maximum and minimum variance angles. (g) Full Radon transform with blue for low values and red for high values. (h) Plot of the angular variance of the Radon transform for all the angles. (i) Second derivative of the angular variance curve.

(a) Car 122 Image

(b) Hessian Image

(c) Binarized Hessian

(d) RT Profile $\theta = 129°$. Var = 11.5

(e) RT Profile $\theta = 156°$. Var = 15.7

(f) RT Profile $\theta = 177°$. Var = 29.3

(g) Radon Transform (Car 122)

(h) Angular Variance of RT

(i) Second derivative of variance

Figure 6: Results for occluded black Car 122 at orientation angle $\theta_{DB} = 90°$, (equivalent to $\theta_{GT} = 180°$ or $0°$ wrt vertical). (a) Image chip showing Car 122. (b) Result of the multiscale Hessian feature detector. (c) Binarized detection image based on thresholding the Hessian. Graphs (d), (e) and (f) show the Radon transform profiles (vertical cross section through the Radon transform shown in (g)) at three representative projection directions including maximum and minimum variance angles. (g) Full Radon transform with blue for low values and red for high values. (h) Plot of the angular variance of the Radon transform for all the angles. (i) Second derivative of the angular variance curve.
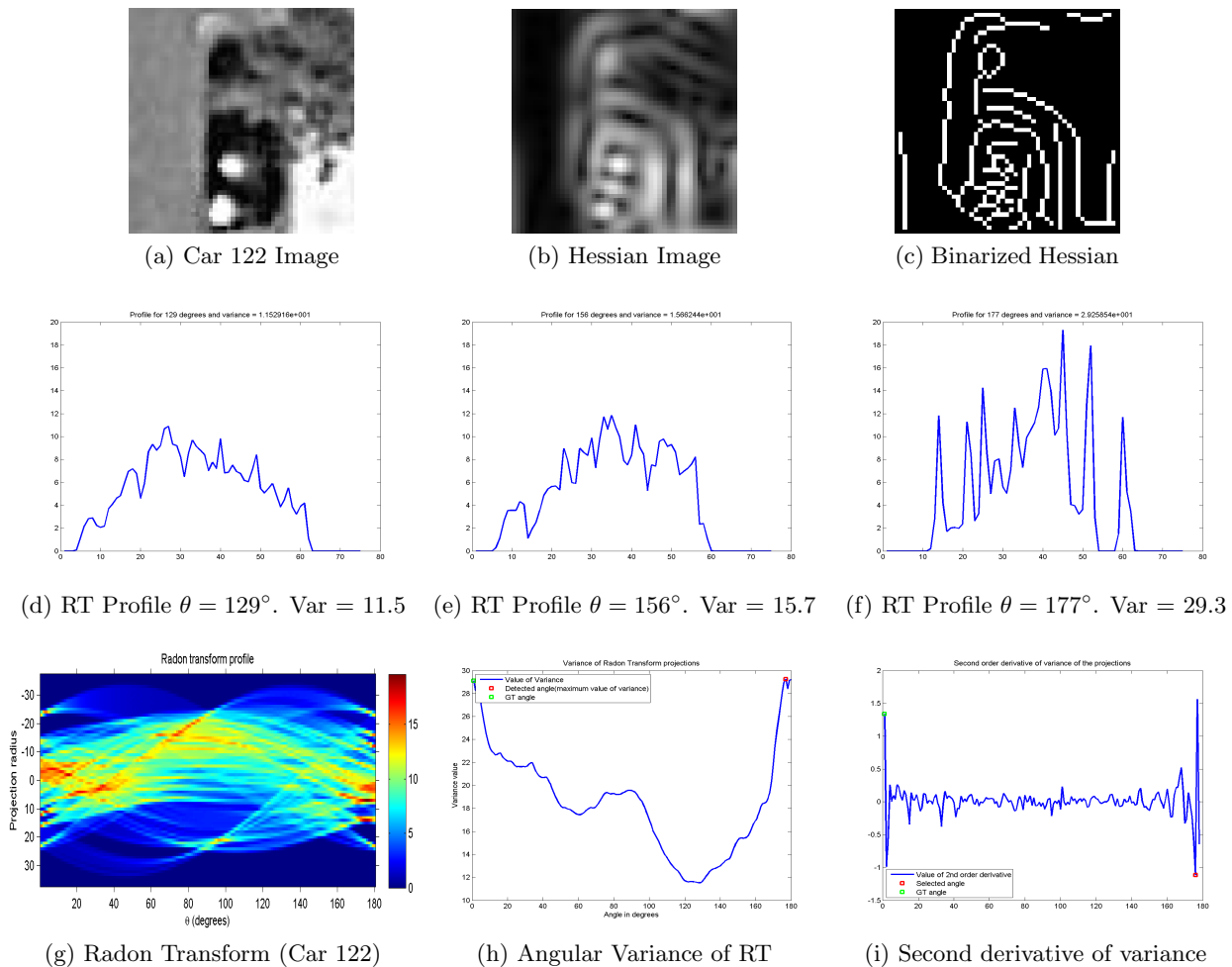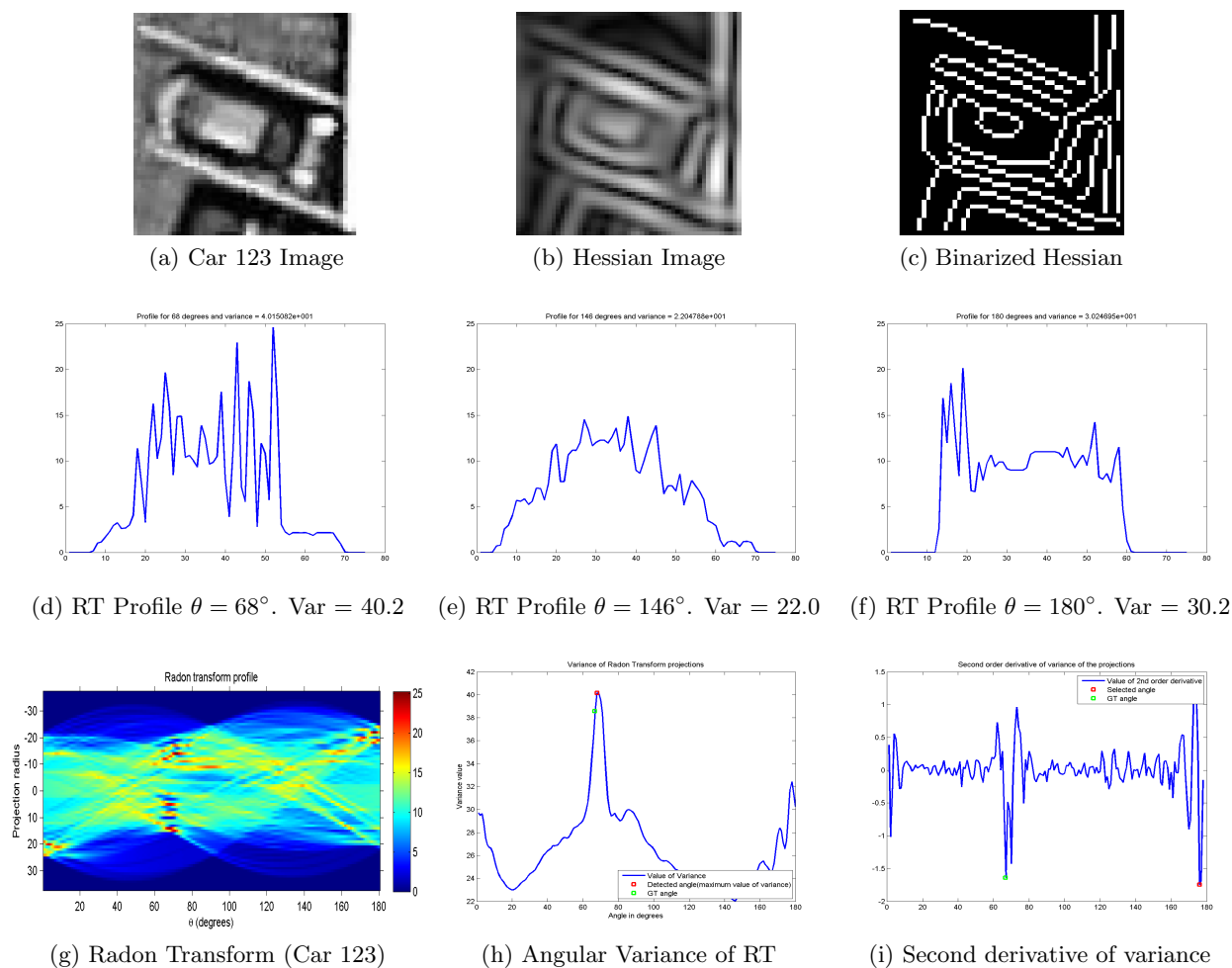
(a) Car 123 Image

(b) Hessian Image

(c) Binarized Hessian

(d) RT Profile $\theta = 68°$. Var = 40.2

(e) RT Profile $\theta = 146°$. Var = 22.0

(f) RT Profile $\theta = 180°$. Var = 30.2

(g) Radon Transform (Car 123)

(h) Angular Variance of RT

(i) Second derivative of variance

Figure 7: Results for gray Car 123 at orientation angle $\theta_{DB} = -22°$, (equivalent to $\theta_{GT} = 68°$ wrt vertical). (a) Image chip showing Car 123. (b) Result of the multiscale Hessian feature detector. (c) Binarized detection image based on thresholding the Hessian. Graphs (d), (e) and (f) show the Radon transform profiles (vertical cross section through the Radon transform shown in (g)) at three representative projection directions including maximum and minimum variance angles. (g) Full Radon transform with blue for low values and red for high values. (h) Plot of the angular variance of the Radon transform for all the angles. (i) Second derivative of the angular variance curve.
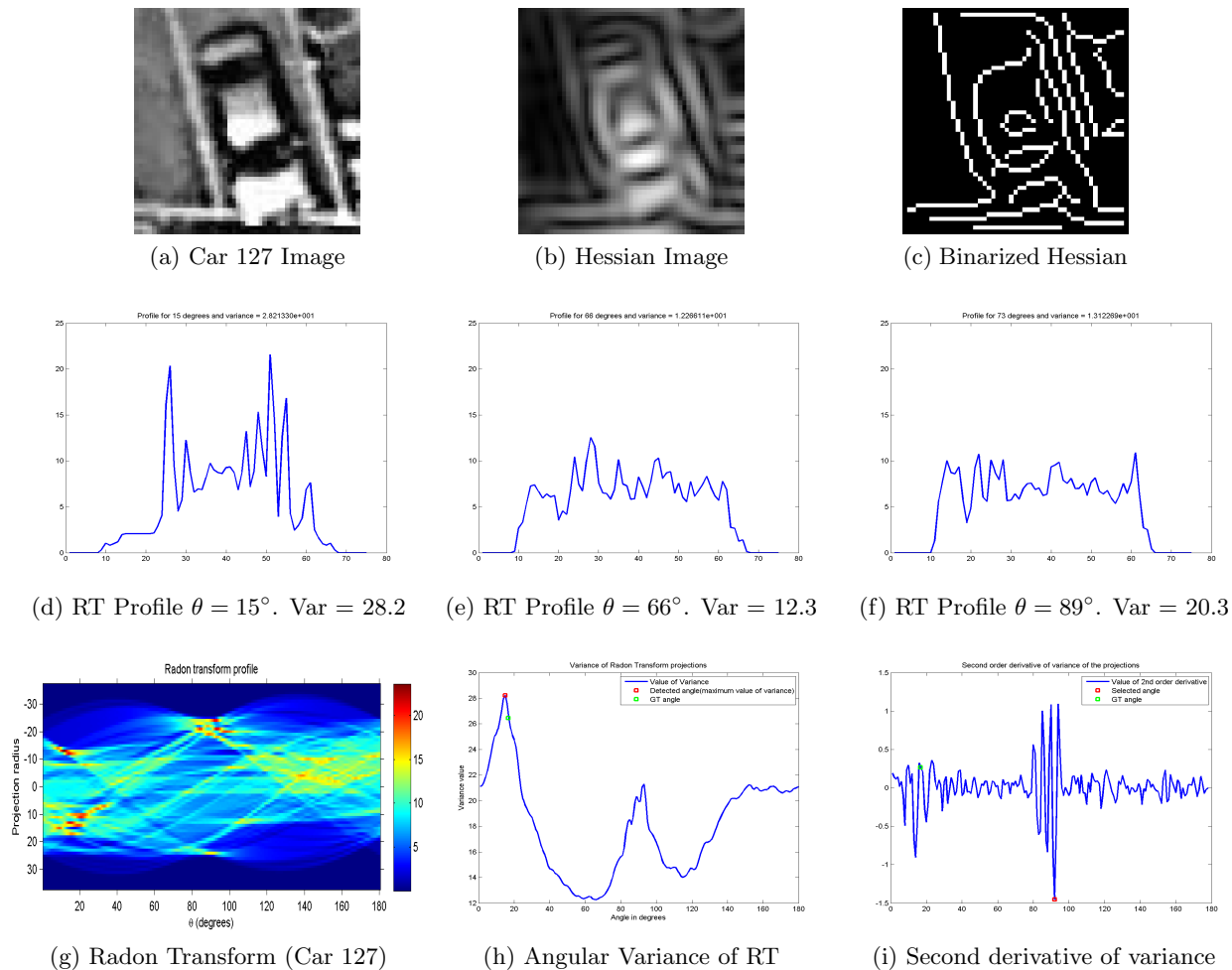
(a) Car 127 Image

(b) Hessian Image

(c) Binarized Hessian

(d) RT Profile $\theta = 15°$. Var = 28.2

(e) RT Profile $\theta = 66°$. Var = 12.3

(f) RT Profile $\theta = 89°$. Var = 20.3

(g) Radon Transform (Car 127)

(h) Angular Variance of RT

(i) Second derivative of variance

Figure 8: Results for gray Car 127 at orientation angle $\theta_{DB} = -75°$ (equivalent to $\theta_{GT} = 15°$ wrt vertical). (a) Image chip showing Car 127. (b) Result of the multiscale Hessian feature detector. (c) Binarized detection image based on thresholding the Hessian. Graphs (d), (e) and (f) show the Radon transform profiles (vertical cross section through the Radon transform shown in (g)) at three representative projection directions including maximum and minimum variance angles. (g) Full Radon transform with blue for low values and red for high values. (h) Plot of the angular variance of the Radon transform for all the angles. (i) Second derivative of the angular variance curve.